



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA
UNIDAD XOCHIMILCO

División de Ciencias Biológicas y de la Salud

Departamento de Sistemas Biológicos

Licenciatura: Química Farmacéutica Biológica

**Ejercicios bioinformáticos para el diagnóstico de la
enfermedad de Parkinson**

Proyecto genérico: Estudio farmacogenómico en población normal y en
pacientes con enfermedades neurológicas

Lugar de realización: Laboratorio de Genética Molecular (N-103),
Departamento Sistemas Biológicos, División Ciencias Biológicas y la
Salud, Universidad Autónoma Metropolitana Unidad Xochimilco.

Alumno: Castillo Refugio Alex

Matricula: 2163026259

Asesores internos:

Dra. Marisol López López. No. económico 7808

Dr. Alberto Ortega Vázquez. No. económico 35583

Abril de 2022

ÍNDICE

1. Marco teórico	1
1.1 Etiología	2
1.2 Bioinformática	4
2. Justificación	5
3. Objetivos	6
3.1 Objetivo General	6
3.2 Objetivos específicos	6
4. Metodología	6
5. Resultados	7
Ejercicio 1	15
Ejercicio 2	17
Ejercicio 3	19
Ejercicio 4	21
Ejercicio 5	24
Ejercicio 6	25
6. Discusión	27
7. Conclusiones	32
8. Bibliografía	32
9. Vo. Bo. del (la) o los (las) asesores (as) respecto a los contenidos académicos	39

1. Marco teórico

La enfermedad de Parkinson (EP) es un trastorno neurodegenerativo crónico progresivo que se caracteriza principalmente por el deterioro de las actividades motoras debido al deterioro del sistema nigroestriatal dopaminérgico. El inicio de la enfermedad ocurre décadas antes de que aparezcan los síntomas principales como temblor en reposo, bradicinesia, rigidez e inestabilidad postural, pero también se describen síntomas no motores que incluyen anosmia, trastornos del sueño, estreñimiento, dolor y fatiga, así como síntomas cognitivos y psiquiátricos (Ambrosio, *et al.*, 2016; Pajares, *et al.*, 2020). Patológicamente, la EP se define por la degeneración lenta y gradual de las neuronas dopaminérgicas en la sustancia negra compacta provocando niveles bajos del neurotransmisor dopamina en el estriado, los núcleos de cola y el putamen. Después de que el 50-60% de las neuronas dopaminérgicas mueren y los niveles de dopamina disminuyen alrededor del 80-85% aparecen los síntomas clínicos (Rabiei, *et al.*, 2019).

La EP es la segunda enfermedad neurodegenerativa más común, después de la enfermedad de Alzheimer, afectando al 1% de todas las personas mayores de 60-65 años; sin embargo, se espera que esta cifra se duplique para 2030 (Ambrosio, *et al.*, 2016; Rabiei, *et al.*, 2019), y en personas mayores de 75 años de edad aumenta al 3 % (Hernandez, *et al.*, 2016). Las tasas de prevalencia en México para las personas de 65 años o más son similares, con aproximadamente 3.14 individuos con EP por cada 1000 habitantes (Smith, *et al.*, 2020)

1.1 Etiología

La mayoría de los casos de EP son esporádicos, asociados al envejecimiento y probablemente tienen una etiología multifactorial como resultado de los efectos combinados de factores ambientales y genéticos; mientras que 5 al 10% de los casos son de origen genético (variantes genéticas de *PARK*). (Simon, *et al.*, 2020; Pajares, *et al.*, 2020).

En los últimos años, se han hecho progresos importantes en el entendimiento de la genética de la EP. Se han identificado mutaciones altamente penetrantes que producen formas raras y monogénicas de la enfermedad, así como variantes únicas con penetrancia incompleta que son fuertes factores de riesgo de EP (tabla 1) (Hernandez, *et al.*, 2016).

Tabla 1. Genes relacionados con la enfermedad de Parkinson.

Símbolo	Gen	Proteína	Localización cromosómica	Herencia	Función	Referencia
<i>PARK 1</i>	<i>SNCA</i>	α -sinucleína	4q21	AD	Involucrado en el tráfico vesicular	Hernandez, <i>et al.</i> , 2016; Jancovik & Tan, 2020; Simon <i>et al.</i> , 2020 y Tran, <i>et al.</i> , 2020.
<i>PARK 8</i>	<i>LRRK2</i>	Cinasa repetida rica en leucina 2	12q12	AD	Autofagia, tráfico vesicular, síntesis de proteínas y función citoesquelética	Hernandez, <i>et al.</i> , 2016; Jancovik & Tan, 2020; Simon <i>et al.</i> , 2020 y Tran, <i>et al.</i> , 2020.
<i>PARK 17</i>	<i>VPS35</i>	Clasificación de proteínas vacuolares 35	16q11.22	AD	Transporte retrógrado de proteínas en endosomas a la red trans-Golgi	Hernandez, <i>et al.</i> , 2016 y Tran, <i>et al.</i> , 2020.
<i>PARK 2</i>	<i>PRKN</i>	Parkina	6q26	AR	Control mitocondrial (mitofagia)	Hernandez, <i>et al.</i> , 2016; Jancovik & Tan, 2020;

						Simon et. al., 2020 y Tran, et. al., 2020.
<i>PARK 6</i>	<i>PINK1</i>	Cinasa putativa inducida por Pten 1	1p36.12	AR	Fosforila a la parkina para regular la mitofagia	Hernandez, et. al., 2016; Simon et. al., 2020 y Tran, et. al., 2020.
<i>PARK 7</i>	<i>DJ-1</i>	DJ-1	1p36	AR	Protege contra el estrés oxidativo	Hernandez, et. al., 2016; Simon et. al., 2020 y Tran, et. al., 2020.
<i>PARK 9</i>	<i>ATP13A2</i>	ATPasa tipo 13	1p36	AR	Autofagia-lisosoma, salud mitocondrial	Hernandez, et. al., 2016 y Tran, et. al., 2020.
<i>PARK 14</i>	<i>PLA2G6</i>	Fosfolipasa A2 grupo 6	22q13.1	AR	Hidroliza los glicerofosfolípidos	Hernandez, et. al., 2016 y Tran, et. al., 2020.
<i>PARK 15</i>	<i>FBXO7</i>	Proteína F-box 7	22q12-13	AR	Ayuda a la translocación de la parkina a la mitocondria	Hernandez, et. al., 2016 y Tran, et. al., 2020.
<i>PARK</i>	<i>GBA</i>	glucocerebrosidasa	1q22	Multifactorial	Descompone la glucocerebrosidasa en glucosa y ceramida	Hernandez, et. al., 2016; Jancovik & Tan, 2020; y Tran, et. al., 2020.
AD: autosómica dominante; AR: autosómica recesiva						

Se han identificado variantes en otros genes causantes de la EP monogénica dominante (*EIF4G1*, *DNAJC13* y *CHCHD2*) y autosómica recesiva (*GIGYF2*, *DNJAC6* *SYNJ1* y *VPS13C*) que son bastante raras y a menudo se manifiestan con características atípicas (Jancovik & Tan, 2020). Actualmente siguen surgiendo genes que son candidatos por causar y/o estar relacionados con la EP como *AIMP2*, *HMOX2*, *IMMT*, *MAN2C1*, *RHOT2*, *SLC25A39*, entre otros (Gialluisi, et al., 2021).

Los nuevos descubrimientos alientan a mejorar los estudios y a futuras investigaciones para confirmar el papel de los genes antes mencionados en el inicio o progresión de la EP (Gialluisi, et al., 2021).

1.2 Bioinformática

La bioinformática es un campo intrínsecamente interdisciplinario y de amplio alcance que utiliza aspectos de la informática, matemáticas y la estadística para almacenar, gestionar, analizar e interpretar datos biológicos. Con el advenimiento de la instrumentación de alto rendimiento se ha visto un crecimiento exponencial de los datos biológicos (Attwood, *et al.*, 2019). Implementar estos recursos en la atención médica contribuye significativamente para reducir la brecha entre el genotipo y el fenotipo, así como proporcionar información sobre biomarcadores y variantes genéticas. Esta nueva metodología permite a los profesionales de la salud un diagnóstico más preciso (Oulas, *et al.*, 2019).

En los trastornos neurodegenerativos, como en la EP, en el momento de la manifestación clínica (y ciertamente en el momento del diagnóstico), un número significativo de neuronas se ha perdido permanentemente. Por lo tanto, un diagnóstico temprano será fundamental para un tratamiento oportuno. (Schneider & Alcalá, 2020).

El sitio web del Centro Nacional para la Información Biotecnológica (NCBI, por sus siglas en inglés *National Center for Biotechnology Information*) es un centro dentro de la Biblioteca Nacional de Medicina de los Institutos Nacionales de Salud creado en 1988, que proporciona un gran conjunto de recursos en línea para información de biología molecular, incluida la base de datos de secuencias de ácidos nucleicos *GenBank* y la base de datos *PubMed* de citas y resúmenes publicados en revistas de ciencias de la vida (Sayers, *et al.*, 2021). La búsqueda de información genética es sencilla, permite a los usuarios navegar por una amplia gama de genomas, y se

puede ensamblar un paquete consistente en secuencias de genoma, transcripción y proteínas, así como anotaciones (Sayers, *et al.*, 2020; Sayers, *et al.*, 2021)

Ensembl es una plataforma de bioinformática para organizar la información biológica de grandes genomas, crea anotaciones detalladas y completas de estructuras genéticas, esto permite a mediante genética comparativa inferir la historia evolutiva de genes y genomas; además, trabaja de forma colaborativa con otras instituciones, incluidos el consorcio *GENCODE* y NCBI (Howe, *et al.*, 2021).

2. Justificación

La implementación de herramientas bioinformáticas y el avance de nuevas tecnologías aplicadas en el diagnóstico de enfermedades ha aumentado de forma exponencial por su gran potencial de mejorar la atención médica al ofrecer un diagnóstico temprano y preciso que permita una acción terapéutica oportuna e incluso su utilización encaminada hacia la atención personalizada.

La unidad de enseñanza-aprendizaje (UEA) de “Tecnologías moleculares para el diagnóstico y la terapéutica (UEA optativa en el 12 trimestre del programa de la licenciatura de Química Farmacéutica Biológica, Q.F.B.) aporta los conocimientos necesarios para que los alumnos en formación conozcan los fundamentos y aplicaciones de las nuevas tecnologías moleculares necesarias para la resolución de problemas del diagnóstico y la terapéutica de enfermedades humanas. Para ello, es necesario que los estudiantes conozcan la aplicación de las herramientas bioinformáticas. Para lograr este propósito es preciso diseñar una serie de ejercicios prácticos para que los estudiantes del área de la salud adquieran un mejor

conocimiento sobre estos temas y así contribuir al conocimiento y a la formación de profesionales de la salud con habilidades y experiencia en este ámbito.

3. Objetivos

3.1 Objetivo General

Conocer las aplicaciones de las herramientas de bioinformática aplicadas al diagnóstico de la enfermedad de Parkinson.

3.2 Objetivos específicos

- Describir el uso de las herramientas de bioinformática disponibles en la red para el diagnóstico de la enfermedad de Parkinson.
- Mostrar algunos ejercicios didácticos relacionados al diagnóstico de la enfermedad de Parkinson mediante el uso de herramientas de bioinformática.

4. Metodología

Se elaboró una serie de ejercicios relacionados al diagnóstico de la enfermedad de Parkinson utilizando las herramientas de bioinformática que proporcionan las páginas NCBI (<https://www.ncbi.nlm.nih.gov>) y *Ensembl*, (<https://www.ensembl.org/index.html>) para explicar el modo de uso de los recursos disponibles en las plataformas.

Se seleccionaron los genes *PRKN* y *PINK1* para ejemplificar los ejercicios que se elaboraron y la información necesaria fue recolectada principalmente de la base de datos *Pubmed* de acceso libre ofrecido también en la plataforma de NCBI.

1.- Se realizó una caracterización de los genes *PRNK* y *PINK1* de forma detallada que incluye localización, número de exones e intrones, tamaño en pares de bases, etc., utilizando el recurso de Genes & Expresión - gen disponible en NCBI.

2.- Se identificaron las variantes genéticas que se han reportado y que cuentan con evidencia disponible sobre su influencia en el desarrollo de la EP y se describieron brevemente los transcritos que presentan ambos genes y se mencionaron las diferencias entre los diferentes transcritos.

3.- A ambos genes se les realizó una alineación con genes ortólogos utilizando la herramienta *BLAST* para analizar las secuencias y realizar un árbol filogenético para poder observar y/o conocer su evolución con respecto a las especies seleccionadas.

4.- Se utilizó la herramienta *primer-BLAST* para diseñar oligonucleótidos específicos que se esperarían utilizar para analizar las secuencias de cada gen.

5.- Se elaboraron algunos ejercicios sobre la utilización de la bioinformática en casos de la EP hereditaria, donde se implementan los conocimientos adquiridos en los ejercicios anteriores para poder diagnosticar la enfermedad. Esto implica la identificación de la variante responsable de la EP, así como el diseño de oligonucleótidos para el análisis de la variante.

5. Resultados

El gen *PRKN* abarca una región de 1.38 Mb del cromosoma 6, tiene 12 exones que codifican para una proteína de 465 aminoácidos, y el gen *PINK1* abarca 18 kb, se encuentra en el cromosoma 1, y tiene 8 exones que codifican para una proteína de 581 aminoácidos (Hauser, *et al.*, 2017; Quinn, *et al.*, 2020). En la tabla 2 se muestran las características de ambos genes.

Tabla 2. Características del gen *PRKN* y *PINK1*

Gen	<i>PARK 2</i>	<i>PARK 6</i>
Símbolo oficial	<i>PRKN</i>	<i>PINK1</i>
Ubicación citogenética	6q26	1p36.12
Coordenadas genómicas	6:162727802-161342417	1:20633458-20651511
Hebra	Inversa	delantera
Número de nucleótidos	1380386	18054
Número de exones (número de nt)	1: 162727766-162727662 (105) 2: 162443473-162443310 (164) 3: 162262765-162262525 (241) 4: 162201252-162201131 (122) 5: 162054174-162054091 (84) 6: 161973417-161973302 (116) 7: 161785908-161785772 (137) 8:161569416-161569355 (62) 9: 161549003-161548854 (150) 10: 161386877-161386794 (84) 11: 161360205-161360088 (118) 12: 161350211-161347417 (2795)	1: 20633458-20633935 (478) 2: 20637842-20638129 (288) 3: 20639892-20639992 (101) 4: 20644490-20644672 (183) 5: 20645560-20645723 (164) 6: 20648505-20648632 (128) 7: 20648995-20649231 (237) 8: 20650434-20651511 (1078)
Número de intrones (número de nt)	1(284188), 2(180544), 3(61272), 4(146956), 5(80673), 6(187393), 7(216355), 8(20351), 9(161976), 10(26588), 11(9876)	1(3906), 2(1762), 3(4497), 4(887), 5(2781), 6(362), 7(1202)
ID NCBI	5071	65018
ID Ensembl	ENSG00000185345.24	ENSG00000158828.8

Las variantes homocigóticas y heterocigóticas compuestas de los genes *PINK1* y *PRKN* asociados con la EP recesiva de inicio temprano (generalmente menor de 40 años), se convirtieron en importantes líneas de investigación ya que la pérdida completa de cualquiera de las funciones resulta en la muerte selectiva de neuronas dopaminérgicas (Truban, *et al.*, 2017). Las variantes que afectan los productos génicos de *PINK1/PRKN* influyen en el proceso de homeostasis y de control de calidad de las mitocondrias y representan más del 50% de los casos de EP juvenil

autosómico recesivo (Quinn, *et al.*, 2020; Hou, *et al.*, 2020). En la tabla 3 se encuentran las variantes genéticas que están clasificadas como patogénicas en el repositorio de *ClinVar* (variación genómica y su relación con la salud humana www.ncbi.nlm.nih.gov/clinvar/).

Tabla 3. Variantes genéticas patogénicas de los genes *PINK1* y *PRKN*

<i>PINK1</i>		<i>PRKN</i>	
Variante genética	Cambio en la proteína	Variante genética	Cambio en la proteína
g.20633561C>T (c.13C>T)	p.Gln5Ter	g.161350176T>C c.1321T>C	p.Cys441Arg
g.20633798C>T (c.250C>T)	p.Gln84Ter	g.161350205G>T c.1292G>T	p.Cys431Phe
g.20633821Tdel (c.273Tdel)	p.Cys92Alafs*15	g.161350208G>A c.1289G>A	p.Gly430Asp
g.20638053delC (c.599delC)	p.Ala200Aspfs*21	g.161569357C>T c.931C>T	p.Gln311Ter
g.20638074delG (c.620delG)	p.Arg207Glnfs*14	g.161785793G>C c.850G>C	p.Gly284Arg
g.20638104C>A (c.650C>A)	p.Ala217Asp	g.161785820C>T c.823C>T	p.Arg275Trp
g.20639952C>T (c.736C>T)	p.Arg246Ter	g.161785839T>A c.804T>A	p.Cys268Ter
g.20639990C>A (c.774C>A)	p.Try258Ter	g.161785885G>A c.758G>A	p.Cys253Try
g.20644526C>A (c.813C>A)	p.His271Gln	g.161785900delT c.743delT	p.Val248Alafs*10
g.20644549G>A (c.836G>A)	p.Arg279His	g.161973317C>G c.719C>G	p.Thr240Arg
g.20644639G>A (c.926G>A)	p.Gly309Asp	g.161973401G>A c.635G>A	p.Cys212Try
g.20644651C>T (c.938C>T)	p.Thr313Met	g.161973403.A>T c.633A>T	P,Lys211Asn
g.20645640T>C (c.1040T>C)	p.Leu347Pro	g.162054149T>G c.560T>G	p.Leu187Ter
g.20646973_20651275del (c.807_5409del)		g.162201174delT c.491delT	p.Val164Aspfs*11
g.20649054G>A (c.1311G>A)	p.Trp437Ter	g.162201182A>T c.483A>T	p.Lys161Asn
g.20649109C>T (c.1366C>T)	p.Gln456Ter	g.162262600_162262561del c.337_376del	p.Pro113Thrfs*51
g.20649217C>T (c.1474C>T)	pArg492Ter	g.162262717_162262716dupTG c.220_221dupTG	p.Trp74Cysfs*8
g.20649232G>A (c.1488+1G>A)		g.162443326delA c.155delA	p.Asn52Metfs*29

g.20650515_20650518dup c.1570_1573TTAG	p.Asp525fs*	g.162443356G>C c.125G>C	p.Arg42Pro
g.20650542CAA[3] c.1597CAA[3]	p.Gln534dup	g.162443379_162443380delAG c.101_102delAG	p.Gln34Argfs*5
		g.162443380delA c.101delA	p.Gln34Argfs*10
		g.162443383G>A c.98G>A	p.Arg33Gln
		g.162443384C>T c.97C>T	p.Arg33Ter
		g.162727667T>C c.2T>C	pMet1Thr

Además de estas variantes del gen *PRKN*, hay más de 50 variaciones registradas en *ClinVar* de cambio en el número de copias que involucran a todos los exones, mayoritariamente son deleciones de exones completos (8-9, 8, 6-7, 3-7, 5, 3-4, 4, 2-4, 2-3, 2-5, 12, y 1) y duplicaciones o eliminaciones de los exones 3, 5-6, 7, 2, y 6.

Los genes *PRKN* y *PINK1* se expresan ampliamente en todo el cerebro y otros tejidos que requieren altas demandas de energía, la mayor abundancia de ARNm se encuentra en el cerebro, el corazón y el músculo esquelético (Barodia, et. al., 2017). El transcrito del gen *PINK1* (ARNm) tiene una longitud de 2657 pb (número de acceso: [NM_032409.3](#)) y el ARNm de *PRKN* tiene una longitud de 4178 pb (número de acceso: [NM_004562.3](#)), ambos transcritos codifican su respectiva proteína de longitud completa, consideradas como transcrito de referencia. Hay dos variantes del transcrito de *PRKN* causado por empalme alternativo, la variante 2 tiene una longitud de 4094 pb debido a que carece del exón 5, y la variante 3 con una longitud de 3731 pb ya que carece de los exones 3, 4 y 5; estas variantes

también se encuentran reportadas en la plataforma de *Ensembl* con diferencias en la longitud de los transcritos correspondientes.

Existen otras variantes, que no se han reportado en pacientes, generadas por computadora (predicciones), ya que gracias a diferentes herramientas (*in silico*) es posible descubrir variantes que sean potencialmente patogénicas y puedan aparecer en estudios futuros a estos genes (Hoppe, K., *et al.*, 2021).

Con la secuenciación continua de diversos organismos es posible reconstruir las historias evolutivas de miles de genes, las familias de genes evolucionan por los procesos de especiación, creando genes ortólogos o parálogos si divergieron a través de la duplicación (Gao & Miller, 2020; Glover, *et al.*, 2019). Además la conjetura ortóloga sugiere que estos genes tienden a conservar sus funciones ancestrales durante largos periodos de tiempo haciendo posible su rastreo hasta el último genoma ancestral común (Glover, *et al.*, 2019)

Ambas bases de datos cuentan con un apartado para genes ortólogos para el gen que se este investigando y *Ensembl* también proporciona una sección para genes parálogos. En la tabla 4 se muestran 15 especies diferentes con ortólogos a *PRKN* y a *PINK1*.

Tabla 4. Genes ortólogos y parálogos de *PINK1* y *PRKN*.

<i>PINK1</i>	<i>PRKN</i>
Ortólogos (especies)	Ortólogos (especies)

<i>Mus musculus</i> (ratón doméstico) <i>Rattus norvegicus</i> (rata noruega) <i>Danio rerio</i> (pez cebra) <i>Sus scrofa</i> (cerdo) <i>Macaca mulatta</i> (mono Rhesus) <i>Bos taurus</i> (vaca) <i>Pan troglodytes</i> (chimpancé) <i>Canis lupus familiaris</i> (perro) <i>Felis catus</i> (gato doméstico) <i>Gorilla gorilla</i> (gorila occidental) <i>Equus caballus</i> (caballo) <i>Chelonia mydas</i> (tortuga marina verde) <i>Oryctolagus cuniculus</i> (conejo) <i>Mesocricetus auratus</i> (hámster dorado) <i>Mustela putorius furo</i> (hurón doméstico)	<i>Mus musculus</i> (ratón doméstico) <i>Rattus norvegicus</i> (rata noruega) <i>Danio rerio</i> (pez cebra) <i>Sus scrofa</i> (cerdo) <i>Macaca mulatta</i> (mono Rhesus) <i>Bos taurus</i> (vaca) <i>Pan troglodytes</i> (chimpancé) <i>Canis lupus familiaris</i> (Perro) <i>Felis catus</i> (gato doméstico) <i>Gorilla gorilla</i> (gorila occidental) <i>Equus caballus</i> (caballo) <i>Chelonia mydas</i> (tortuga marina verde) <i>Oryctolagus cuniculus</i> (conejo) <i>Mesocricetus auratus</i> (hámster dorado) <i>Mustela putorius furo</i> (hurón doméstico)
Parálogos (Gen)	Parálogos (Gen)
PEAK3, PEAK1 y PRAG1	RNF144A, ARIH2, RNF144B, ARIH1, RNF14, RNF217, RNF19B, ANKIB1, RNF19A

Con el fin de poder apreciar la similitud de estos genes con respecto al humano (*Homo sapiens*), y conocer qué tan alejados o cercanos estamos evolutivamente de estas especies, se realizó la alineación de sus transcritos, utilizando la herramienta **BLAST** (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) obteniendo los resultados mostrados en las figuras 1 y 3, y el árbol filogenético de las secuencias comparadas en las figuras 2 y 4.

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len
Pan troglodytes PRKN	7285	7285	100%	0.0	98.57%	4215
Gorilla gorilla gorilla PRKN	7130	7130	99%	0.0	97.92%	4204
Macaca mulatta PRKN	6345	6345	99%	0.0	93.79%	4193
Equus caballus PRKN	1790	2682	88%	0.0	83.39%	3837
Canis lupus familiaris PRKN	1682	2176	61%	0.0	84.69%	3473
Sus scrofa PRKN	1590	1989	57%	0.0	82.23%	2690
Oryctolagus cuniculus PRKN	1563	1563	36%	0.0	83.09%	1753
Mesocricetus auratus PRKN	1562	1562	35%	0.0	83.49%	3159
Mus musculus parkin PRKN	1527	1671	43%	0.0	83.15%	3232
Bos taurus PRKN	1439	1439	30%	0.0	84.83%	1691
Felis catus PRKN	1012	2207	73%	0.0	83.26%	3881
Rattus norvegicus PRKN	1004	1580	35%	0.0	85.47%	3548
Mustela putorius furo PRKN	747	1346	30%	0.0	86.12%	1715
Chelonia mydas PRKN	612	860	24%	7e-177	75.81%	2712
Danio rerio PRKN	326	384	26%	1e-90	68.00%	1465

Figura 1. Resultado de la alineación del gen *PRKN* con 15 ortólogos en *BLAST*.

La secuencia de referencia de transcrito del gen *PRKN* (número de acceso: NM_004562.3) se alineó con los genes ortólogos (tabla 4), arrojando los resultados de la siguiente manera: puntaje máximo de la mejor secuencia alineada (*Max Score*), suma total de los puntajes de todas las secuencias alineadas (*Total Score*), porcentaje de la secuencia alineada (*Query Cover*), número de *hits* esperados (valor E); porcentaje de identidad (Per. Ident) y longitud de acceso (Acc. Len).

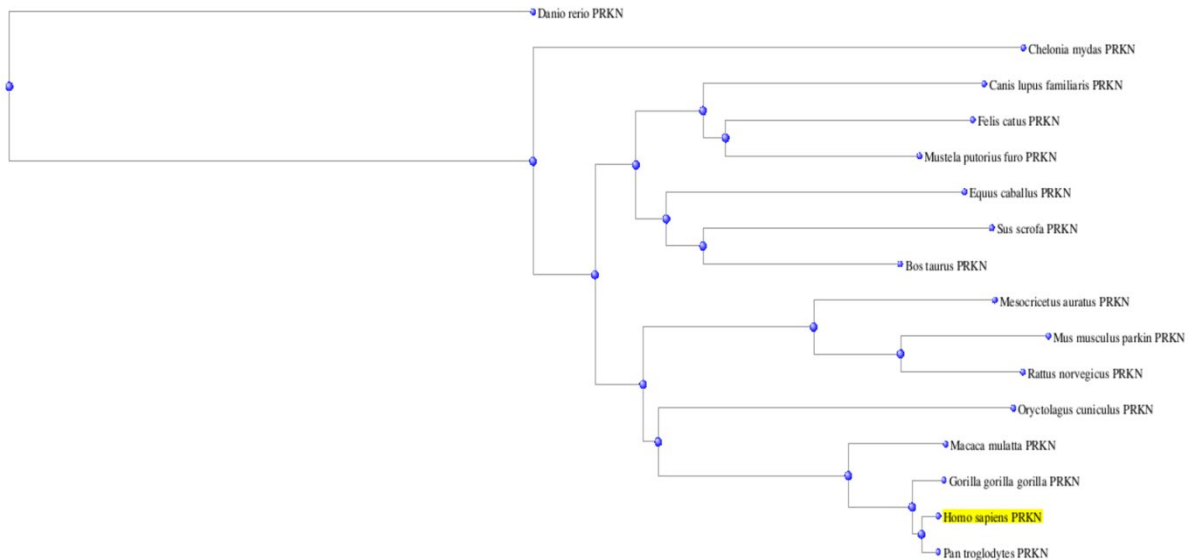


Figura 2. Árbol filogenético (*PRKN*) con base en los resultados de la alineación. Se aprecia la distancia evolutiva entre las diferentes especies alineadas con respecto al gen *PRKN* humano, generado por el programa *BLAST* (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len
Gorilla gorilla PINK1	4635	4635	99%	0.0	98.68%	2651
Pan troglodytes PINK1	4347	4347	92%	0.0	99.15%	2460
Macaca mulatta PINK1	3910	3910	92%	0.0	94.97%	2471
Equus caballus PINK1	2333	2333	70%	0.0	87.73%	2250
Felis catus PINK1	2214	2264	74%	0.0	85.27%	2931
Canis lupus familiaris PINK1	2145	2145	70%	0.0	85.54%	2857
Sus scrofa PINK1	2127	2127	72%	0.0	84.06%	3090
Bos taurus PINK1	2058	2058	68%	0.0	84.82%	2970
Mustela putorius furo PINK1	2030	2098	75%	0.0	83.61%	2945
Mesocricetus auratus PINK1	1814	1814	69%	0.0	81.88%	3025
Oryctolagus cuniculus PINK1	1701	1701	68%	0.0	80.93%	2380
Mus musculus PINK1	1697	1697	66%	0.0	80.80%	2429
Rattus norvegicus PINK1	721	721	27%	0.0	81.78%	1146
Chelonia mydas PINK1	566	782	54%	5e-163	72.27%	3067
Danio rerio PINK1	301	301	27%	3e-83	70.16%	2088

Figura 3. Resultado de la alineación de gen *PINK1* con 15 ortólogos en BLAST.

La secuencia de referencia de transcrito del gen *PINK1* (número de acceso: NM_032409.3) se alineó con los genes ortólogos (tabla 4), arrojando los resultados de la siguiente manera: puntaje máximo de la mejor secuencia alineada (*Max Score*), suma total de los puntajes de todas las secuencias alineadas (*Total Score*), porcentaje de la secuencia alineada (*Query Cover*), número de *hits* esperados (valor *E*), porcentaje de identidad (*Per. Ident*) y longitud de acceso (*Acc. Len*).

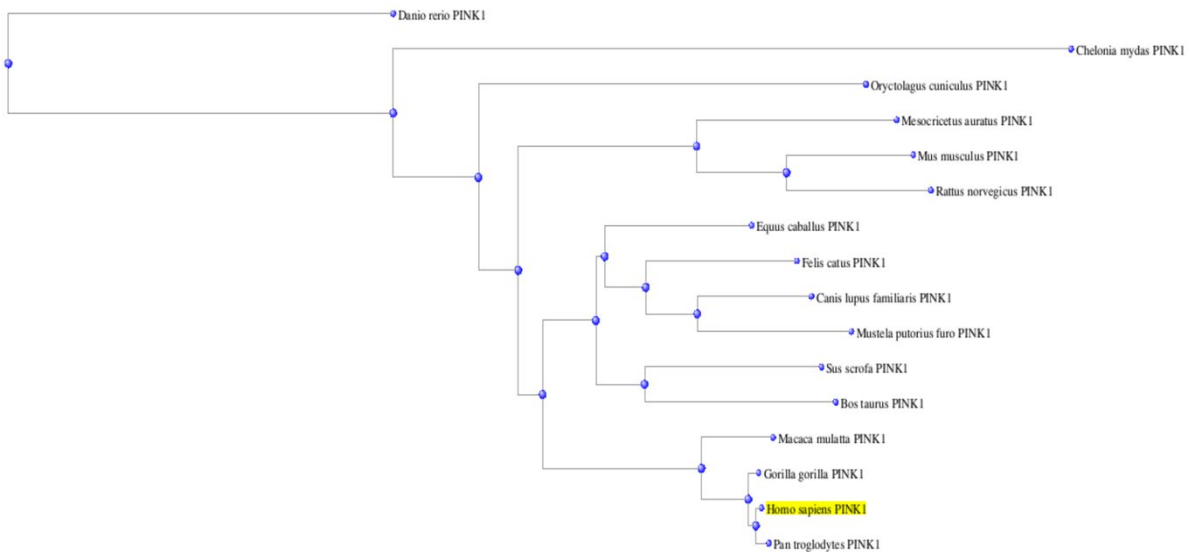


Figura 4. Árbol filogenético (*PINK1*) con base en los resultados de la alineación. Se aprecia la distancia evolutiva entre las diferentes especies alineadas con respecto al gen *PINK1* humano, generado por el programa *BLAST* (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

El diseño de oligonucleótidos es fundamental para examinar una secuencia específica de un gen, ya que estas permiten mediante una reacción de PCR amplificar la secuencia de nuestro interés delimitada por los cebadores que se diseñan para este fin. El programa *BLAST*, antes mencionado, tiene funciones adicionales una de las cuales es la herramienta para el diseño de oligonucleótidos específicos para la plantilla utilizada en la técnica de PCR.

Una vez dentro de esta herramienta bioinformática debes ingresar la secuencia de nucleótidos que deseas amplificar (plantilla) en formato FASTA, y el programa genera los oligonucleótidos específicos para dicha región (pueden variar con base en la secuencia que ingreses).

Ejercicio 1

Una persona sana con antecedentes heredofamiliares, abuelo paterno con EP, refiere que a su pariente se le realizó un análisis molecular mediante secuenciación resultando homocigoto para la variante c.719C>G del gen *PRKN*. La persona no presenta síntomas, sin embargo, quiere saber si es portador de dicha variante. Diga si la variante genética analizada es responsable de la EP. Diseñe los oligonucleótidos necesarios para amplificar la región que contiene la variante y diga los cambios que presenta la variante a nivel genómico y proteico.

Solución

La variante c.719C>G indica la ubicación en ADN codificante (ADNc), eso quiere decir que se encuentra en uno de los 12 exones del gen *PRKN*. En la tabla 3 podemos ver su ubicación genómica o se puede consultar en el repositorio de *ClinVar* y con esta información podemos ubicar la región que se analizará; además, en la tabla 2 tenemos las coordenadas de los 12 exones del gen, así que podemos identificar en qué exón se encuentra la variante.

El diseño de oligonucleótidos se realizó mediante una herramienta bioinformática para el diseño oligonucleótidos específicos llamada *Primer-BLAST* (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), donde se ingresó la secuencia de nucleótidos en la ventana “*PCR template*” (Plantilla de PCR), en este caso se usó una región del gen de 280 nt (g.161973163-161973442), en parámetros seleccionamos “*Database*” y cambiamos a “*Genomes for selected organisms*”, finalmente el programa genera los oligonucleótidos automáticamente (figura 5). Es importante mencionar que los resultados arrojados pueden ser de más de un par de oligonucleótidos y estos varían dependiendo la región y tamaño que seleccione como plantilla, pero lo importante es que abarque la región de interés.

Primer pair 1

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTTTAAATGTGGAGCACACCCC	Plus	22	34	55	59.77	50.00	6.00	2.00
Reverse primer	TGTGCACGTAATGCAAGTGATG	Minus	22	133	112	60.10	45.45	6.00	0.00
Product length	100								

Products on intended targets

>[NC_000006.12](#) Homo sapiens chromosome 6, GRCh38.p13 Primary Assembly

product length = 100

Features associated with this product:

[e3 ubiquitin-protein ligase parkin isoform 2](#)

[e3 ubiquitin-protein ligase parkin isoform x3](#)

Forward primer	1	CTTTAAATGTGGAGCACACCCC	22
Template	161973409	161973388
Reverse primer	1	TGTGCACGTAATGCAAGTGATG	22
Template	161973310	161973331

Primer pair 2

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	TTTAAATGTGGAGCACACCCC	Plus	21	35	55	58.76	47.62	6.00	2.00
Reverse primer	ACGTCTGTGCACGTAATGCAA	Minus	21	138	118	61.14	47.62	6.00	2.00
Product length	104								

Products on intended targets

>[NC_000006.12](#) Homo sapiens chromosome 6, GRCh38.p13 Primary Assembly

product length = 104

Features associated with this product:

[e3 ubiquitin-protein ligase parkin isoform 2](#)

[e3 ubiquitin-protein ligase parkin isoform x3](#)

Forward primer	1	TTTAAATGTGGAGCACACCCC	21
Template	161973408	161973388
Reverse primer	1	ACGTCTGTGCACGTAATGCAA	21
Template	161973305	161973325

Figura 5. Resultados arrojados por la herramienta *primer-BLAST*, nótese que en este caso el programa arroja dos pares de oligonucleótidos bastante similares, varían ligeramente su

posición con respecto a la plantilla y tamaño del producto de PCR. Ambos pueden ser usados como solución al problema.

Seleccionamos un par de oligonucleótidos para resolver el problema, en este caso se seleccionó el primero “*Primer pair 1*” (figura 5), aunque ambos son correctos. Los criterios de selección que se usaron para elegir este par fueron el tamaño del producto amplificado, el tamaño de los oligonucleótidos y el valor de la T_m (temperatura de alineamiento) la cual debe ser muy similar. En el caso que las T_m estén muy alejadas se deberá realizar un gradiente de temperatura para determinar la temperatura óptima en la que el templado haya amplificado mejor.

Oligonucleótido *forward*: CTTTAAATGTGGAGCACACCCC

Oligonucleótido *reverse*: TGTGCACGTAATGCAAGTGATG

Longitud del producto de PCR: 100 (g.161973409-161973310)

Esta variante se encuentra en el exón 6 (g.161973317C>G), esto ocasiona un cambio de aminoácido en la proteína p.Thr240Arg, reportada en *ClinVar* como patogénica.

Ejercicio 2

Un individuo se quiere realizar un diagnóstico genético presintomático ya que tiene ambos padres son heterocigotos para una variante de *PINK1*, la madre porta p.Thr313Met y el padre la variante p.Ala217Asp. Describa la ubicación exacta de estas variantes y diseñe los oligonucleótidos necesarios para amplificar dichas regiones.

Solución

El problema nos dice que la variante es un cambio de un solo nucleótido ya que causa un cambio de aminoácido en la proteína, entonces la variante está en ADNc, siguiendo el mismo razonamiento del problema anterior, y apoyándonos en la tabla 3 y 2, podemos contestar lo que se pide. La variante del padre se encuentra en el exón 4 del gen y el de la madre en el exón 2, por lo tanto los oligonucleótidos deben ser diseñados para amplificar la región de estos dos exones.

Exón 2

Oligonucleótido *forward*: TGCAGTGCTGCTGTGTATGA

Oligonucleótido *reverse*: CAGGCACCTTTCCTGTGGAT

Longitud del producto de PCR: 304 (g. 20637950-20638253)

La ubicación genómica es g.20638104C>A ó c.650C>A

Exón 4

Oligonucleótido *forward*: AACATCATCCGGGTTCTCCG

Oligonucleótido *reverse*: ATAACGAGGAACAGCGTCCG

Longitud del producto de PCR: 137 (g.20644530-20644666)

La ubicación genómica es g.20644651C>T ó c.938C>T

El diseño de los oligonucleótidos de este ejercicio se realizó siguiendo el mismo procedimiento como en el ejercicio anterior utilizando la herramienta *primer-BLAST*

de *BLAST* (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Por ejemplo, para el exón 4 se seleccionó como plantilla una región del gen de 490 nt (g.20644345-20644834), y en este caso sólo arrojó un par de oligonucleótidos (figura 6).

Primer pair 1

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	AACATCATCCGGTTCTCCG	Plus	20	186	205	59.82	55.00	4.00	2.00
Reverse primer	ATAACGAGGAACAGCGTCCG	Minus	20	322	303	60.18	55.00	3.00	2.00
Product length	137								

Products on intended targets
 >[NC_000001.11](#) Homo sapiens chromosome 1, GRCh38.p13 Primary Assembly

product length = 137
 Features associated with this product:
[serine/threonine-protein kinase pink1, mitochondrial prec...](#)

Forward primer	1	AACATCATCCGGTTCTCCG	20
Template	20644530	20644549
Reverse primer	1	ATAACGAGGAACAGCGTCCG	20
Template	20644666	20644647

Figura 6. Resultado arrojado por la herramienta *primer-BLAST*, el programa solo generó un par de oligonucleótidos específicos para la región del exón 4.

Ejercicio 3

En un estudio molecular a pacientes vietnamitas con EP de inicio temprano realizado por Ton *et al.*, (2020) se encontró una alta frecuencia de la variante c.1018G>A homocigota en *PINK1*.

Con esta información conteste lo siguiente:

- Identifique la posición a nivel genómico donde ocurre dicha variante.
- Mencione el cambio que produce a nivel de proteína y si es considerada patogénica.
- Diseñe los oligonucleótidos necesarios para examinar la variante y adjunte la región amplificada señalando el cambio.

Solución

A nivel genómico la variación se encuentra en g.20645618G>A en el exón 5 de *PINK1*, esto ocasiona un cambio de aminoácido en la proteína de alanina a treonina (p.A340T). Se ha demostrado con herramientas de predicción *in silico* que esta variante es “benigna” consistente con la clasificación de *ClinVar*, sin embargo, este cambio de nucleótido podría estar relacionada con la EP.

Oligonucleótido *forward*: GCCAGCTATCCCTGTACCCT

Oligonucleótido *reverse*: AGCAGGTTCTACCTGGGTC

Longitud del producto de PCR: 182 (g.20645555- 20645736)

```
GCCAGCTATCCCTGTACCCTGCGCCAGTACCTTTGTGTGAACACACCCAGCCCCCGCCTC
GCCAACCATGATGCTGCTGCAGCTGCTGGAAGGCGTGGACCATCTGGTTCAACAGGGCAT
CGCGCACAGAGACCTGAAATCCGACAACATCCTTGTGGAGCTGGACCCAGGTAGGAACC
TGCT
```

La secuencia resaltada en amarillo corresponde al exón 5 y el nucleótido resaltado en rojo es el sitio de la variación.

En el mismo estudio se estudiaron tres hermanos de una familia. El análisis de las secuencias codificantes mostró que son heterocigotos compuestos para el gen *PRKN*, en un alelo carecen de los exones 3-5 y en el otro se identificó una variante de un solo nucleótido. A continuación, se muestra la secuencia de la región donde se encuentra la variante presente en los tres hermanos.

```
TGATGTCTAGGCTAGCGTGCTGGTTTGGGAATGCGTGTTTTCCAGGTAACTTGCTGCGAAC
CCACCACACCTTTGTTTTCTGCCCCAACAGGAGGCTGCATGCACATGAAGTGTCCGCAG
CCCCAGCGCAGGCTCGAGTGGTGTGGAAACTGTGGCTGCGAGTGGAAACCGCGTCTGCA
TGGGGGACCACTGGTTCTGACGTGTAGCCAGGG
```

- Identifique la variante.

- Mencione el cambio a nivel genómico, ADNc y proteína.
- Diseñe los oligonucleótidos que se utilizarían para amplificar dicha región.

Solución

Los tres hermanos tienen un cambio de nucleótido en la posición g.161350176T>C, a nivel codificante en c.1321T>C, es parte del último exón y causa en la proteína un cambio de cisteína por arginina (p.C441R), dicha variante está reportada en *ClinVar* como patogénica.

Oligonucleótido *forward*: TGTTTTCTGCCCCCAACAGG

Oligonucleótido *reverse*: TGGCTACACGTCGAACCAAGT

Longitud del producto de PCR: 135 (g.161350230-161350096), es la parte subrayada en la secuencia.

TGATGTCTAGGCTAGCGTGCTGGTTTGGGAATGCGTGTTTTCCAGGTAAGTGTCCGCAG
 CCACCACACCTTTGTTTTCTGCCCCCAACAG **GAGGCTGCATGCACATGAAGTGTCCGCAG**
CCCCAGCGCAGGCTCGAGTGGTGTCTGGAAGTGTGGCTGCGAGTGAACCGCGTCTGCA
TGGGGGACCACTGGTTCGACGTGTAGCCAGGG

Ejercicio 4

Basado en un estudio realizado por Monroy-Jaramillo *et al.*, (2014), en mestizos mexicanos, resuelva el siguiente ejercicio. La secuencia muestra un fragmento del gen *PRKN* correspondiente a un paciente de origen mexicano con EP previamente diagnosticado por manifestaciones clínicas.

AATGTTGCATTATTTAATTAATAATTGAGCAAACCTATTTGACTTGATCTGAGTTTATATTACAAATAAT
 GTAAGATCAAAAGTGTAGGGCTGCCACAAAGAAGACCTTGAAAATGGTCTGATGGAGTTTGGAAATTTATT
 TGCAGGCTGGGGAAACCATTAAGGAAATTGAGCAAACCGGTACCAGGAAGAAATGACATCAGGGACACA

AGAAATCATGCTTTAAGAATAAGAATAATGATAAAGGGAAAATGTGATTGTTTAAACATTAGTCTTCTTCT
AGCTCATTGATATTGTGTGGGCTGCCAATATTATGAGGTGGAATGCTGGGCCGATTTTTCAAAAATAAATAT
AAAACCTTGTTGAATTTTAAATAGAGGCATGACAATATATTGCTTTATAAGATTGATAGCGATGTATTAA
AAGTGTTTCATCATTTAATACATAAAATATGTTCCCTCAAAAATGTATATGAAATAAGATTGTTAAAATACAGT
TGGTAGAAGCGATCTTATAAAAATACAAAACCTTTGGAAAATCTATTTTCTGAAGACAAAAAAGCAAAAAGT
AAACACACTTATATGTCTTGCCAGTTGAATTTGGTAGACTTTTACTGCATTGAACCATATTTCGTCTCTCC
AAATTTATGTATTTATTTTGGTGAAAACAGTTATATGAAGATTTTCTGCTTTGGTTGCTTATGGAAGTAT
TTACTTTAAGATGCTGTGAGCCACTAGGCCAGACTTTTAAACTTTTCTCCCAGCACTCAGCTGATGCAGA
GCAAAAATTTATTCTTTCTTCAGAGTTGCCTTTAAGTAGACAGGAGTGAAGGAGAAAAGAATTCTCCATCA
GTGACGCTTAGTGACATTGTTTGCCTGAGTATTTCAAGAAAATAAAATGAAGGTTAATGTGGAAAACCTTG
AAATGTTATTATGTCATAAAAATCATTTTAAGTAGAGATGAACAAAACATAATGGTGTCTCTTATGGAAGC
AGAATAAGAAATGTGGAAAATGTTTGTCTAAGAGGTCAAACCATGTGACCTGGCCGGCTTTTCAAATGA
ATCTATGGCTGCAAGCTGATGGCTCAAGGGGGGAAGAAGCATAAECTCAAACACAAAAAGGCTTACTAAAT
TATAGGTTAAAATGACAAATGGATAATAATCCTGATATTAATTTTCATGCAGCCTGGAAAAGTGAGAAAG
TGTATGTGCATATAAAAAGTCTATTGAATTTTGTGAGAAAACATTTCTTTTAAAATTGCAACTTCATGAGG
AGTTTTGTGCAAGAAGGAGCACAGATAGCAGAGTTGTATTAGCCAGTCTTGATTTGGGGCTGAGCACTGT
AAGTCATCTAGAAATAAATTAGCATTGGTTCCCTTAGTAAAGCTTCTGTATGTCATGCAGTGGGGCTC

Identifique la variante asociada a la EP utilizando sus habilidades en bioinformática.

Solución

La secuencia resaltada corresponde a la región genómica g.161549802-161549173 y la secuencia no resaltada a g.161548682-161547913, lo que nos indica una delección de 490 nt entre estas dos ubicaciones genómicas (g.161549172-161548683), esta delección abarca el exón 9 completo (g.161549003-161548854), esto significa que el paciente es homocigoto de esta delección.

AATGTTGCATTATTTAATTAATAAATTGAGCAAACCTATTTGACTTGATCTGAGTTTATATTACAAAATAAT
GTAAGATCAAAAAGTGTAGGGCTGCCACAAAAGAAGACCTTGAAAATGGTCTGATGGAGTTTGGAAATTTATT
TGCAGGCTGGGGAAACCATTAAGGAAAATTGAGCAAACCGGTACCAGGAAGAAAATGACATCAGGGACACA
AGAAATCATGCTTTAAGAATAAGAATAATGATAAAGGGAAAATGTGATTGTTTAAACATTAGTCTTCTTCT
AGCTCATTGATATTGTGTGGGCTGCCAATATTATGAGGTGGAATGCTGGGCCGATTTTTCAAAAATAAATAT
AAAACCTTGTTGAATTTTAAATAGAGGCATGACAATATATTGCTTTATAAGATTGATAGCGATGTATTAA
AAGTGTTTCATCATTTAATACATAAAATATGTTCCCTCAAAAATGTATATGAAATAAGATTGTTAAAATACAGT
TGGTAGAAGCGATCTTATAAAAATACAAAACCTTTGGAAAATCTATTTTCTGAAGACAAAAAAGCAAAAAGT
AAACACACTTATATGTCTTGCCAGTTGAATTTGGTAGACTTTTACTGCATTGAACCATATTTCGTCTCTCC
AAATTTATGTATTTATTTTGGTGAAAACAGTTATATGAAGATTTTCTGCTTTGGTTGCTTATGGAAGTAT
TTACTTTAAGATGCTGTGAGCCACTAGGCCAGACTTTTAAACTTTTCTCCCAGCACTCAGCTGATGCAGA
GCAAAAATTTATTCTTTCTTCAGAGTTGCCTTTAAGTAGACAGGAGTGAAGGAGAAAAGAATTCTCCATCA
GTGACGCTTAGTGACATTGTTTGCCTGAGTATTTCAAGAAAATAAAATGAAGGTTAATGTGGAAAACCTTG
AAATGTTATTATGTCATAAAAATCATTTTAAGTAGAGATGAACAAAACATAATGGTGTCTCTTATGGAAGC
AGAATAAGAAATGTGGAAAATGTTTGTCTAAGAGGTCAAACCATGTGACCTGGCCGGCTTTTCAAATGA
ATCTATGGCTGCAAGCTGATGGCTCAAGGGGGGAAGAAGCATAAECTCAAACACAAAAAGGCTTACTAAAT
TATAGGTTAAAATGACAAATGGATAATAATCCTGATATTAATTTTCATGCAGCCTGGAAAAGTGAGAAAG
TGTATGTGCATATAAAAAGTCTATTGAATTTTGTGAGAAAACATTTCTTTTAAAATTGCAACTTCATGAGG
AGTTTTGTGCAAGAAGGAGCACAGATAGCAGAGTTGTATTAGCCAGTCTTGATTTGGGGCTGAGCACTGT
AAGTCATCTAGAAATAAATTAGCATTGGTTCCCTTAGTAAAGCTTCTGTATGTCATGCAGTGGGGCTC

Para otro paciente en el mismo estudio con una variante homocigota en el gen *PINK1* se encuentra la siguiente secuencia, identifica y resalta la variación. Resalta la secuencia correspondiente al exón y menciona las coordenadas de la variante.

```
TGCTATCTTGGGACCTCAGCACACTCATTTCATAAAATTGTGTCTTTTGGGTGAGATTTGTCTTGGGGTCTT
CAAAACCCCTGAGACTGTGGACACATACAGCAGCCTAACGGCATTGAGGAGGGCAAGCATTCAACAGTTA
GGACAATGTGAACTGTAGCTCAGCTCTGCTAGGTACCTTCACAAGACCTCGAATGCTGCCCTTACTATG
CCTCGGTTTTCTTATCTATAAAAACGGCATTTCCTTATCTTGGTGGAGCTGTTAAATAAAATTTAAAGAC
GTAAAGGGTCTGGCACCATGGTTGGCAAAAAATATCAGTTTCCCTTCTCGACTTCTCGATTTTGCCAGG
ACCAGTGATGTTACATTCAGGACCTGCCTGAACCGGCAAGCCCTCCACGTGGGTCCAAAGTGCAAAGGG
AAAGTCACTGCTAGAGGCGCCAGTACCAGCATAGCGCCCCACGCGCCGAGTCGGGGAACTGCCGCGGGG
GCCGGCCCCGCCACCAGCGCTGCGCTGCGCAGAGGCACCGCCCCAAGTTTGTGTGACCGGCGGGGG
ACGCCGGTGGTGGCGGCAGCGGGCGGCTGCGGGGGCACCGGGCCGCGGCGCCACCATGGCGGTGCGACAGG
CGCTGGGCGCGGCCTGCAGCTGGGTGAGCGCTGCTGCTGCGCTTCACGGGCAAGCCCGGCCGGGCCTA
CGGCTTGGGGCGGCCGGGCCCGGGCGGGCTGTGTCCGCGGGGAGCGTCCAGGCTGGGCCGCGAGGACCG
GGCGCGGAGCCTCGCAGGGTTCGGGCTCGGGCTCCCTAACCGTCTCCGCTTCTTCCGCCAGTCGGTGGCCG
GGCTGGCGGGCGCGGTTGCAGCGGCAGTTTCGTGGTGCGGGCTGGGGTGCAGCGGGCCCTTGCAGCGGGCA
GTCTTTCTGGCCTTCGGGCTAGGGCTGGGCCTCATCGAGGAAAAACAGGCGGAGAGCCGGCGGGCGGTCT
CGGCCTGTCAGGAGATCCAGGTGAGCGGGGCCGGTCTTAAGCCGAGCGGAGGACGAGCTAAGCGCGGG
GGCGGGTCCCTCAGCTGGGTGGGGGCGGGGCTAGGTGTGGAGGCGGGGCTCTGAGCAGATCGAGGGCCGAG
GCGAGGGTCCCTAAAGCTCATCTATTTACCATTACTGATCGGCTGCTATAAAATAAAGCCAGCACCTCCC
ATTTGTTTTAATGTTTCCCTTCCCTCAAATGAAGACATGTTGCCGATTACAGCTCCTGTGCGAGCACAGCA
AAAGGCTTTGTGTAATTTTCTAAAATGTACGGACAATAAATCATAACATTCCTATCCCTTTGAGGTAG
TTGCCGTCCCTAATTTATGGAGAAGGAAAGTCTCAGGTGAAGGGACTTGCTCGAAGTCACACAGCTAA
```

Las coordenadas de la secuencia genómica anterior corresponden a g.20632935-20634334 y la variante es una delección del nucleótido timina (g.20633821Tdel), esto causa un corrimiento en el marco de lectura generando un codón de parada prematuro (p.Cys92Alafs*15). La parte resaltada en amarillo corresponde al exón 1 de *PINK1* (g.20633458-20633935), la resaltada en rojo es el codón de parada prematuro y la resaltada en verde es donde ocurrió la delección entre esos dos nucleótidos. Esta variante está registrada como patogénica.

```
TGCTATCTTGGGACCTCAGCACACTCATTTCATAAAATTGTGTCTTTTGGGTGAGATTTGTCTTGGGGTCTT
CAAAACCCCTGAGACTGTGGACACATACAGCAGCCTAACGGCATTGAGGAGGGCAAGCATTCAACAGTTA
GGACAATGTGAACTGTAGCTCAGCTCTGCTAGGTACCTTCACAAGACCTCGAATGCTGCCCTTACTATG
CCTCGGTTTTCTTATCTATAAAAACGGCATTTCCTTATCTTGGTGGAGCTGTTAAATAAAATTTAAAGAC
GTAAAGGGTCTGGCACCATGGTTGGCAAAAAATATCAGTTTCCCTTCTCGACTTCTCGATTTTGCCAGG
ACCAGTGATGTTACATTCAGGACCTGCCTGAACCGGCAAGCCCTCCACGTGGGTCCAAAGTGCAAAGGG
AAAGTCACTGCTAGAGGCGCCAGTACCAGCATAGCGCCCCACGCGCCGAGTCGGGGAACTGCCGCGGGG
GCCGGCCCCGCCACCAGCGCTGCGCTGCGCAGAGGCACCGCCCCAAGTTTGTGTGACCGGCGGGGG
```

ACGCCGGTGGTGGCGGCAGCGGGCGGCTGCGGGGGCACCGGGCCGCGGCGCCACCATGGCGGTGCGACAGG
 CGCTGGGCCCGCGGCCTGCAGCTGGGTTCGAGCGCTGCTGCTGCGCTTCACGGGCAAGCCCGGCCGGGCGCTA
 CGGCTTGGGGCGGCCGGGCCCCGGCGGGCGGGCTGTGTCCGCGGGGAGCGTCCAGGCTGGGCCCGCAGGACCG
 GGCGCGGAGCCTCGCAGGGTTCGGGCTCGGGCTCCCTAACCGTCTCCGCTTCTTCCGCCAGTCGGTGGCCG
 GGCTGGCGGGCGCGGTTCAGCGGCAGTTCGTGGTTCGGGCGCTGGGCTGCGCGGGCCCTTGGCGGCCGGGCA
 GTCTTTCTGGCCTTCGGGC TAGGGCTGGGCCTCATCGAGGAAAAACAGGCGGAGAGCCGGCGGGCGGTCT
 CGGCCTGTCAGGAGATCCAGGTGAGCGGGGCCGGGTCTAAGCCGAGCGGAGGACGGAGCTAAGCGCGGG
 GCGGGTCTCAGCTGGGTGGGGCGGGGCTAGGTGTGGAGCGGGGCTCTGAGCAGATCGAGGGCCGAG
 GCGAGGGTCTTAAAGCTCATCTATTTACCATTACTGATCGGCTGCTATAAATAAAGCCAGCACCTCCC
 ATTTGTTTTAATGTTTTCCCTTCTCAAATGAAGACATGTTGCCGATTACAGCTCCTGTCGCAGCACAGCA
 AAAGGCTTTGTGTAATTTTCTAAAATGTACGACAACATAAATCATAACATTCTATCCCTTTGAGGTAG
 TTGCCGTCCCTAATTTATGGAGAAGGAAAGTCCTCAGGTGAAGGGACTTGCTCGAAGTCACACAGCTAA

Ejercicio 5

Dentro de los siguientes ID de NCBI, NM_001106694.1, XM_045063875.1, NM_016694.4, XM_017345477.1, XM_023258242.2, XM_001153913.5, XM_021095478.1, XM_014737247.2, XM_038654595.1, NM_001017635.1, se encuentran cinco transcritos de diferentes especies que codifican para la proteína *PINK1* y cinco para la proteína *PRKN*. Elabore una tabla que indique la especie a la que pertenece, la longitud del transcrito y el gen al que pertenece.

Solución

ID del transcrito	Especie	Longitud del transcrito	Gen
NM_001106694.1	<i>Rattus norvegicus</i>	1146 pb	<i>PINK1</i>
XM_045063875.1	<i>Mustela putorius furo</i>	2945 pb	<i>PINK1</i>
NM_016694.4	<i>Mus musculus</i>	3232 pb	<i>PRKN</i>
XM_017345477.1	<i>Oryctolagus cuniculus</i>	1753 pb	<i>PRKN</i>
XM_023258242.2	<i>Felis catus</i>	2931 pb	<i>PINK1</i>
XM_001153913.5	<i>Pan troglodytes</i>	4215 pb	<i>PRKN</i>
XM_021095478.1	<i>Sus scrofa</i>	3090 pb	<i>PINK1</i>
XM_014737247.2	<i>Equus caballus</i>	2250 pb	<i>PINK1</i>
XM_038654595.1	<i>Canis lupus familiaris</i>	3473 pb	<i>PRKN</i>
NM_001017635.1	<i>Danio rerio</i>	1465 pb	<i>PRKN</i>

Ejercicio 6

Busca los ortólogos del gen *PRKN* y *PINK1* que se piden abajo y con sus secuencias (transcritos) de cada especie elabora un árbol filogenético con respecto al humano.

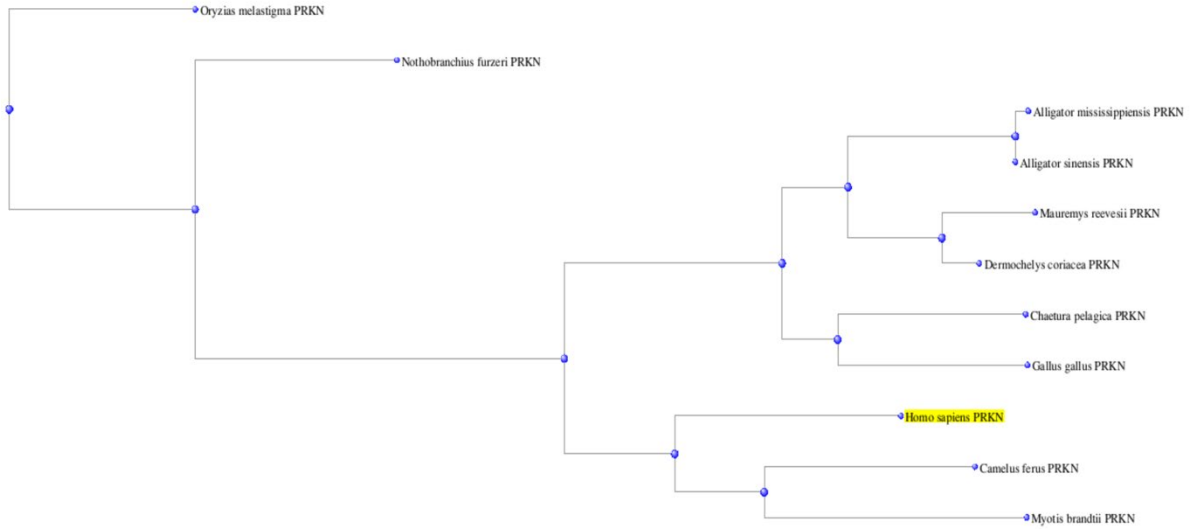
- *Chaetura pelagia*
- *Gallus gallus*
- *Mauremys reevesii*
- *Dermochelys coriacea*
- *Cemelus ferus*
- *Myotis brandtii*
- *Nothobranchius furzeri*
- *Oryzias melastigma*
- *Alligator sinensis*
- *Alligator mississippiensis*

Anexe el árbol filogenético de cada gen y conteste lo siguiente:

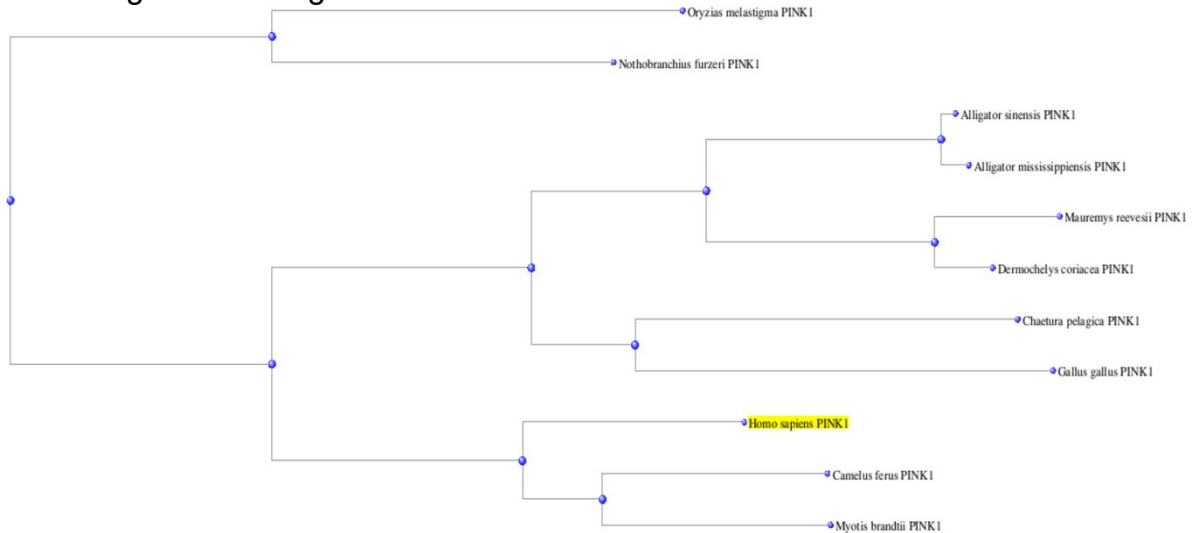
- ¿Qué especie está más lejos y cuál más cerca evolutivamente del humano?
- ¿Qué diferencia hay entre estos árboles filogenéticos?

Solución

Árbol filogenético del gen *PRKN*.



Árbol filogenético del gen *PINK1*



Ambos genes tienen una evolución bastante parecida, las especies más cercanas al humano son *Camelus ferus* y *Myotis brandtii*, mientras que las más alejadas son *Oryzias melastigma* y *Nothobranchius furzeri* para el gen *PINK1* y para *PRKN* la más alejada es *Oryzias melastigma*. También podemos observar que hay una diferencia entre ambos árboles ya que *Oryzias melastigma* y *Nothobranchius furzeri* están más relacionados en el árbol filogenético de *PRKN*; sin embargo, en el árbol filogenético de *PINK1* se encuentran menos relacionados.

6. Discusión

El uso de las bases de datos de NCBI y *Ensembl* son excelentes para llevar a cabo investigaciones científicas. Este caso la caracterización de los genes abordados es sencilla, ya que ambas plataformas ofrecen amplia información genética, además proporcionan enlaces directos a bases de datos colaboradoras y herramientas de análisis, su uso no requiere de conocimientos especializados, y tienen la ventaja de ofrecer un manual, así como videos para aprender a utilizar estos sitios web. Las características de cada gen están bien definidas en ambas bases de datos, prácticamente no difieren en la información que proporcionan. Como se puede observar en la información de la tabla 2 el gen *PRKN* tiene un tamaño bastante grande, la secuencia de sus 12 exones solo representa el 0.3% de su secuencia total, en cuanto al gen *PINK1* la secuencia de sus 8 exones representa el 14.71% de su secuencia total. A pesar de la gran diferencia de tamaño genético la proteína que codificada *PINK1* es 116 aminoácidos más grande que la proteína de *PRKN* y si comparamos ambos genes, el gen *PINK1* completo solo es el 1.3% del gen *PRKN*.

Los grandes avances en el análisis genético, particularmente por la utilización de la secuenciación de próxima generación, han acelerado considerablemente el descubrimiento e identificación de genes involucrados en muchas enfermedades hereditarias que han dado lugar a significativos avances en el ámbito de la investigación en genética humana, sin embargo, junto con este crecimiento exponencial de conocimiento nuevos desafíos se presentan, como lo es el interpretar si el número extraordinario de variantes genéticas actualmente identificadas son patogénicas o benignas (Nishio & Usami, 2017).

La creciente disponibilidad y el decreciente costo de realizar la secuenciación de ADN ha llevado al desarrollo de métodos de predicción *in silico* como una de las varias líneas de evidencia para la interpretación de variantes (Hoppe, K., et al., 2021). Sin embargo, a pesar de la valiosa información que ofrecen las herramientas bioinformáticas en variantes causantes de una enfermedad, actualmente no deben aplicarse para una clasificación definitiva, la determinación de una variante patogénica siempre debe incluir información adicional como análisis de segregación y análisis funcionales. (Ernst, C., et al., 2018; Hoppe, K., et al., 2021).

En el caso particular de los genes *PINK1* y *PRKN* cuentan con más de 200 y 500 variaciones respectivamente reportadas en *ClinVar*, la gran mayoría de estas variantes aún siguen siendo de interpretación incierta, ya que carecen de información o es insuficiente para poder determinar su influencia en la enfermedad. Además, como dice Hauser et al. 2017, la demostración de segregación puede ser particularmente difícil para los fenotipos recesivos si el tamaño de la familia es demasiado pequeño o los antecedentes familiares y los genotipos no están disponibles. Por otro lado, en un estudio realizado por Manroy-Jaramillo, et al., (2014) en 122 pacientes mestizo mexicanos mostraron una alta prevalencia de deleciones de exones en el gen *PRKN*, principalmente los exones 9 y/o 12, y para el gen *PINK1* se registró la deleción del exón 1 con mayor incidencia; hay que tomar en cuenta que México posee una gran diversidad poblacional, y los resultados de estudios futuros pueden variar.

En cuanto a las variantes de empalme de transcrito del gen *PRKN* ambas codifican una proteína más corta, ya que los exones faltantes no afectan el marco de lectura del ARNm, sin embargo hay que considerar que la variante 3 se clasifica como patogénica puesto que el producto proteico carecería de los aminoácidos 58 al 206 y la proteína no contaría con el residuo de serina en la posición 65, fundamental para su activación por fosforilación dependiente de *PINK1*, esto consiste con lo reportado en un estudio realizado por Ton, *et al.*, (2020), donde tres pacientes hermanos con la enfermedad portaban una variación heterocigota compuesta, una variante patogénica (p.C441R) y la otra carente de estos exones. Al igual que en el caso anterior la exclusión del exón 5 en la variante 2 no afecta la traducción, sin embargo, se ha reportado que la eliminación del exón 5 en pacientes con la enfermedad en estado homocigoto en ClinVar (VCV000663416.2), preserva el marco de lectura y está considerada como patogénica.

Al realizar una alineación de secuencias probablemente lo que más nos interesa es la relación de homología entre dos o más secuencias y con esta información poder desentrañar la historia evolutiva que nos relaciona a otros seres vivos. En los próximos años se prevé que miles de genomas sean secuenciados, haciendo el uso necesario de herramientas bioinformáticas que puedan manejar la gran cantidad de datos y nos permita consultar, analizar, comparar, manipular e interpretar dichos datos (Armstrong, *et al.*, 2019). Los ortólogos se pueden usar para la creación de filomas, que son colecciones completas de árboles filogenéticos para cada gen codificante en un genoma dado, siendo uno de sus principales objetivos encontrar genes que se han conservado evolutivamente (Glover, *et al.*, 2019).

Un buen ejemplo de esto es la alineación que se realizó con los genes ortólogos de ambos genes (*PINK1* y *PRKN*), podemos notar en la figura 1 y 3 que su porcentaje de identidad con respecto a los genes humanos están por encima del 80% a excepción de *Chelonia mydas* y *Danio rerio*, siendo los primates con el mayor porcentaje de identidad (>93%) en ambos, esto indica que los genes comparten gran proporción de nucleótidos. Por otro lado, el porcentaje de alineación cambia para primates con un porcentaje mayor al 90% mientras que las especies restantes por debajo del 76%, con estos resultados podemos decir que las especies más cercanas al humano son los pertenecientes a los primates. Hablando taxonómicamente, el humano pertenece a esta orden, lo cual se confirma en los árboles filogenéticos (figura 2 y 4) de cada gen, mientras que las especies más alejadas evolutivamente es el pez cebra y tortuga marina verde (*Danio rerio* y *Chelonia mydas*, respectivamente), ambos esquemas son parecidos, sin embargo, hay diferencias que se pueden resaltar como el hecho de que en base al gen *PRKN*, después de los primates, la especie más cercana al humano es el conejo (*Oryctolagus cuniculus*), contrariamente a lo que se observa en el gen *PINK1* donde el conejo es el más alejado después de *Danio rerio* y *Chelonia mydas*.

Los ejercicios planteados son ideales para que el profesional en formación en áreas de la salud se familiarice con este tipo de herramientas, ya que su resolución implica la utilización de herramientas como lo es *BLAST* para alineaciones de nucleótidos, *BLAST-primer* para el diseño de oligonucleótidos, la consulta de secuencias de nucleótidos de ambos genes y sus ortólogos, las variantes que se han reportado a la fecha, e incluso la posibilidad de visualizar la distancia evolutiva que nos separa

de otras especies y con este conocimiento adquirido explorar otras aplicaciones que son posibles gracias a estas herramientas e incluso mantenerse actualizados, ya que estas plataformas se actualizan constantemente.

A medida que la complejidad y el tamaño de los datos en ciencias de la vida aumenta, crece la necesidad de capacitar a los científicos para la utilización de herramientas bioinformáticas, haciendo de esta una disciplina esencial en la trayectoria de los profesionales de la salud, mejorando así la investigación biomédica y la atención médica, esto conlleva a un esfuerzo por desarrollar competencias básicas para la capacitación en bioinformática (Mulder, *et al.*, 2018). El diagnóstico molecular ofrece a las familias información sobre el pronóstico de la enfermedad y proporcionar algunos parámetros sobre qué esperar en el futuro e incluso existe la posibilidad de ofrecer evaluar a las familias del paciente (si así lo desearan), ya que, es posible que los familiares sean portadores de estas condiciones recesivas y así poder dar un seguimiento (Horton & Lucassen, 2019).

Una buena infraestructura informática es esencial para la ciencia moderna de datos biomédicos y requiere potencia informática, almacenamientos de datos, redes y administración de sistemas, mientras que algunos laboratorios, instituciones educativas, organismos de financiación y otros organismos ya han adoptado la biología computacional, otros que no corren el riesgo de quedar muy por detrás de la vanguardia (Way, *et al.*, 2021). A pesar de que existen cursos disponibles para los estudiantes que se especializan en ciencias de la vida, la mayoría tiende a centrarse en la teoría dejando de lado la parte práctica (Fan, *et al.*, 2021); con los ejercicios planteados se espera impulsar al profesional en formación a explorar las

bases de datos y/o herramientas de las que dispone en línea de forma gratuita y aprender a utilizarlas para un diagnóstico molecular.

7. Conclusiones

Existen varias herramientas bioinformáticas disponibles gratuitamente en la red, dentro de estas NCBI y *Ensembl* son una excelente opción para comenzar a descubrir las funciones disponibles, y además su utilización no requiere de conocimientos especializados.

La información analizada aporta conocimientos importantes en el uso de la bioinformática para los profesionales de la salud en formación y a personas interesadas en este tema.

Los ejercicios planteados contribuyen a los profesionales de la salud a una mejor comprensión y aplicación de las herramientas de bioinformática en el diagnóstico de la EP.

La información analizada servirá de guía para los estudiantes y profesores profesionales de la salud que requieren del uso de bioinformática para el diagnóstico molecular y la terapéutica.

8. Bibliografía

1. Ambrosio, L., Portillo, M. C., Rodríguez-Blázquez, C., Rodríguez-Violante, M., Castrillo, J., Arillo, V. C., Garretto, N. S., Arakaki, T., Dueñas, M. S., Álvarez, M., Ibáñez, I. P., Carvajal, A., & Martínez-Martín, P. (2016). Living with

- chronic illness scale: international validation of a new self-report measure in Parkinson's disease. *NPJ Parkinson's disease*, 2, 16022. <https://doi.org/10.1038/npjparkd.2016.22>
2. Armstrong, J., Fiddes, I. T., Diekhans, M., & Paten, B. (2019). Whole-Genome Alignment and Comparative Annotation. *Annual review of animal biosciences*, 7, 41-64. <https://doi.org/10.1146/annurev-animal-020518-115005>
 3. Attwood, T. K., Blackford, S., Brazas, M. D., Davies, A., & Schneider, M. V. (2019), A global perspective on evolving bioinformatics and data science training needs. *Briefings in bioinformatics*, 20(2), 398-404. <https://doi.org/10.1093/bib/bbx100>
 4. Barodia, S. K., Creed, R. B., & Goldberg, M. S. (2017). Parkin and PINK1 functions in oxidative stress and neurodegeneration. *Brain research bulletin*, 133, 51-59. <https://doi.org/10.1016/j.brainresbull.2016.12.004>
 5. Ernst, C., Hahnen, E., Engel, C., Nothnagel, M., Weber, J., Schmutzler, R. K., & Hauke, J. (2018). Performance of in silico prediction tools for the classification of rare BRCA1/2 missense variants in clinical diagnostics. *BMC medical genomics*, 11(1), 35. <https://doi.org/10.1186/s12920-018-0353-y>
 6. Fan, W., Mencius, J., Du, W., Fan, H., Zhu, H., Wei, D., Zhou, M., & Quan, S. (2021). Online bioinformatics teaching practice: Comparison of popular docking programs using SARS-CoV-2 spike RBD-ACE2 complex as a benchmark. *Biochemistry and molecular biology education : a bimonthly publication of the international Union of Biochemistry and Molecular Biology*, 49(6), 833-840. <https://doi.org/10.1002/bmb.21566>

7. Gao, K., & Miller, J. (2020). Primary orthologs from local sequence context. *BMC bioinformatics*, 21(1), 48. <https://doi.org/10.1186/s12859-020-3384-2>
8. Gialluisi, A., Reccia, M. G., Modugno, N., Nutile, T., Lombardi, A., Di Giovannantonio, L. G., Pietracupa, S., Ruggiero, D., Scala, S., Gambardella, S., International Parkinson's Disease Genomics Consortium (IPDGC), Iacoviello, L., Gianfrancesco, F., Acampora, D., D'Esposito, M., Simeone, A., Ciullo, M., & Esposito, T. (2021). Identification of sixteen novel candidate genes for late onset Parkinson's disease. *Molecular neurodegeneration*, 16(1), 35. <https://doi.org/10.1186/s13024-021-00455-2>
9. Glover, N., Dessimoz, C., Ebersberger, I., Forslund, S. K., Gabaldón, T., Huerta-Cepas, J., Martin, M. J., Muffato, M., Patricio, M., Pereira, C., da Silva, A. S., Wang, Y., Sonnhammer, E., & Thomas, P. D. (2019). Advances and Applications in the Quest for Orthologs. *Molecular biology and evolution*, 36(10), 2157-2164. <https://doi.org/10.1093/molbev/msz150>
10. Hauser, D. N., Primiani, C. T., & Cookson, M. R. (2017). The Effects of Variants in the Parkin, PINK1, and DJ-1 Genes along with Evidence for their Pathogenicity. *Current protein & peptide science*, 18(7), 702-714. <https://doi.org/10.2174/1389203717666160311121954>
11. Hernandez, D. G., Reed, X., & Singleton, A. B. (2016). Genetics in Parkinson disease: Mendelian versus non-Mendelian inheritance. *Journal of neurochemistry*, 139 Suppl 1(Suppl 1), 59-74. <https://doi.org/10.1111/jnc.13593>
12. Hoppe, K., Jurkat-Rott, K., Kranepuhl, S., Wearing, S., Heiderich, S., Merlak, S., & Klingler, W. (2021). Relevance of pathogenicity prediction tools in

- human RYR1 variants of unknown significance. *Scientific reports*, 11(1), 3445. <https://doi.org/10.1038/s41598-021-82024-7>
13. Horton, R. H., & Lucassen, A. M. (2019). Recent developments in genetic/genomic medicine. *Clinical science (London, England : 1979)*, 133(5), 697-708. <https://doi.org/10.1042/CS20180436>
14. Hou, X., Watzlawik, J. O., Fiesel, F. C., & Springer, W. (2020). Autophagy in Parkinson's Disease. *Journal of molecular biology*, 432(8), 2651-2672. <https://doi.org/10.1016/j.jmb.2020.01.037>
15. Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., El Houdaigui, B., Fatima, R., Gall, A., ... Flicek, P. (2021). Ensembl 2021. *Nucleic acids research*, 49(D1), D884-D891. <https://doi.org/10.1093/nar/gkaa942>
16. Jankovic, J., & Tan, E. K. (2020). Parkinson's disease: etiopathogenesis and treatment. *Journal of neurology, neurosurgery, and psychiatry*, 91(8), 795-808. <https://doi.org/10.1136/jnnp-2019-322338>
17. Monroy-Jaramillo, N., Guerrero-Camacho, J. L., Rodríguez-Violante, M., Boll-Woehrlen, M. C., Yescas-Gómez, P., Alonso-Vilatela, M. E., & López-López, M. (2014). Genetic mutations in early-onset Parkinson's disease Mexican patients: molecular testing implications. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*, 165B(3), 235-244. <https://doi.org/10.1002/ajmg.b.32228>

18. Mulder, N., Schwartz, R., Brazas, M. D., Brooksbank, C., Gaeta, B., Morgan, S. L., Pauley, M. A., Rosenwald, A., Rustici, G., Sierk, M., Warnow, T., & Welch, L. (2018). The development and application of bioinformatics core competencies to improve bioinformatics training and education. *PLoS computational biology*, *14*(2), e1005772. <https://doi.org/10.1371/journal.pcbi.1005772>
19. Nishio, S. Y., & Usami, S. I. (2017). The Clinical Next-Generation Sequencing Database: A Tool for the Unified Management of Clinical Information and Genetic Variants to Accelerate Variant Pathogenicity Classification. *Human mutation*, *38*(3), 252-259. <https://doi.org/10.1002/humu.23160>
20. Oulas, A., Minadakis, G., Zachariou, M., Sokratous, K., Bourdakou, M. M., & Spyrou, G. M. (2019). Systems Bioinformatics: increasing precision of computational diagnostics and therapeutics through network-based approaches. *Briefings in bioinformatics*, *20*(3), 806-824. <https://doi.org/10.1093/bib/bbx151>
21. Pajares, M., I Rojo, A., Manda, G., Boscá, L., & Cuadrado, A. (2020). Inflammation in Parkinson's Disease: Mechanisms and Therapeutic Implications. *Cells*, *9*(7), 1687. <https://doi.org/10.3390/cells9071687>
22. Quinn, P., Moreira, P. I., Ambrósio, A. F., & Alves, C. H. (2020). PINK1/PARKIN signalling in neurodegeneration and neuroinflammation. *Acta neuropathologica communications*, *8*(1), 189. <https://doi.org/10.1186/s40478-020-01062-w>

23. Rabiei, Z., Solati, K., & Amini-Khoei, H. (2019). Phytotherapy in treatment of Parkinson's disease: a review. *Pharmaceutical biology*, 57(1), 355-362. <https://doi.org/10.1080/13880209.2019.1618344>
24. Sayers, E. W., Beck, J., Brister, J. R., Bolton, E. E., Canese, K., Comeau, D. C., Funk, K., Ketter, A., Kim, S., Kimchi, A., Kitts, P. A., Kuznetsov, A., Lathrop, S., Lu, Z., McGarvey, K., Madden, T. L., Murphy, T. D., O'Leary, N., Phan, L., Schneider, V. A., ... Ostell, J. (2020). Database resources of the National Center for Biotechnology Information. *Nucleic acids research*, 48(D1), D9-D16. <https://doi.org/10.1093/nar/gkz899>
25. Sayers, E. W., Beck, J., Bolton, E. E., Bourexis, D., Brister, J. R., Canese, K., Comeau, D. C., Funk, K., Kim, S., Klimke, W., Marchler-Bauer, A., Landrum, M., Lathrop, S., Lu, Z., Madden, T. L., O'Leary, N., Phan, L., Rangwala, S. H., Schneider, V. A., Skripchenko, Y., ... Sherry, S. T. (2021). Database resources of the National Center for Biotechnology Information. *Nucleic acids research*, 49(D1), D10-D17. <https://doi.org/10.1093/nar/gkaa892>
26. Simon, D. K., Tanner, C. M., & Brundin, P. (2020). Parkinson Disease Epidemiology, Pathology, Genetics, and Pathophysiology. *Clinics in geriatric medicine*, 36(1), 1-12. <https://doi.org/10.1016/j.cger.2019.08.002>
27. Smith, E. R., Perrin, P. B., Tyler, C. M., Lageman, S. K., & Villaseñor, T. (2020). Cross-cultural differences in Parkinson's disease caregiving and burden between the United States and Mexico. *Brain and behavior*, 10(9), e01753. <https://doi.org/10.1002/brb3.1753>

28. Schneider, S. A., & Alcalay, R. N. (2020). Precision medicine in Parkinson's disease: emerging treatments for genetic Parkinson's disease. *Journal of neurology*, 267(3), 860-869. <https://doi.org/10.1007/s00415-020-09705-7>
29. Ton, N. D., Thuan, N. D., Thuong, M., Ngoc, T., Nhung, V. P., Hoa, N., Nam, N. H., Dung, H. T., Son, N. D., Ba, N. V., Bac, N. D., Tai, T. N., Dung, L., Hung, N. T., Duong, N. T., Ha, N. H., & Hai, N. V. (2020). Rare and novel variants of PRKN and PINK1 genes in Vietnamese patients with early-onset Parkinson's disease. *Molecular genetics & genomic medicine*, 8(10), e1463. <https://doi.org/10.1002/mgg3.1463>
30. Tran, J., Anastacio, H., & Bardy, C. (2020). Genetic predispositions of Parkinson's disease revealed in patient-derived brain cells. *NPJ Parkinson's disease*, 6, 8. <https://doi.org/10.1038/s41531-020-0110-8>
31. Truban, D., Hou, X., Caulfield, T. R., Fiesel, F. C., & Springer, W. (2017). PINK1, Parkin, and Mitochondrial Quality Control: What can we Learn about Parkinson's Disease Pathobiology?. *Journal of Parkinson's disease*, 7(1), 13-29. <https://doi.org/10.3233/JPD-160989>
32. Way, G. P., Greene, C. S., Carninci, P., Carvalho, B. S., de Hoon, M., Finley, S. D., Gosline, S., Lê Cao, K. A., Lee, J., Marchionni, L., Robine, N., Sindi, S. S., Theis, F. J., Yang, J., Carpenter, A. E., & Fertig, E. J. (2021). A field guide to cultivating computational biology. *PLoS biology*, 19(10), e2001419. <https://doi.org/10.1371/journal.pbio.3001419>

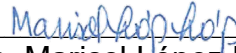
9. Vo. Bo. del (la) o los (las) asesores (as) respecto a los contenidos académicos



Dr. Alberto Ortega Vázquez

Profesor Asociado D, Tiempo completo

No. económico: 35583



Dra. Marisol López López

Profesor Titular C, Tiempo completo

No. económico: 7808