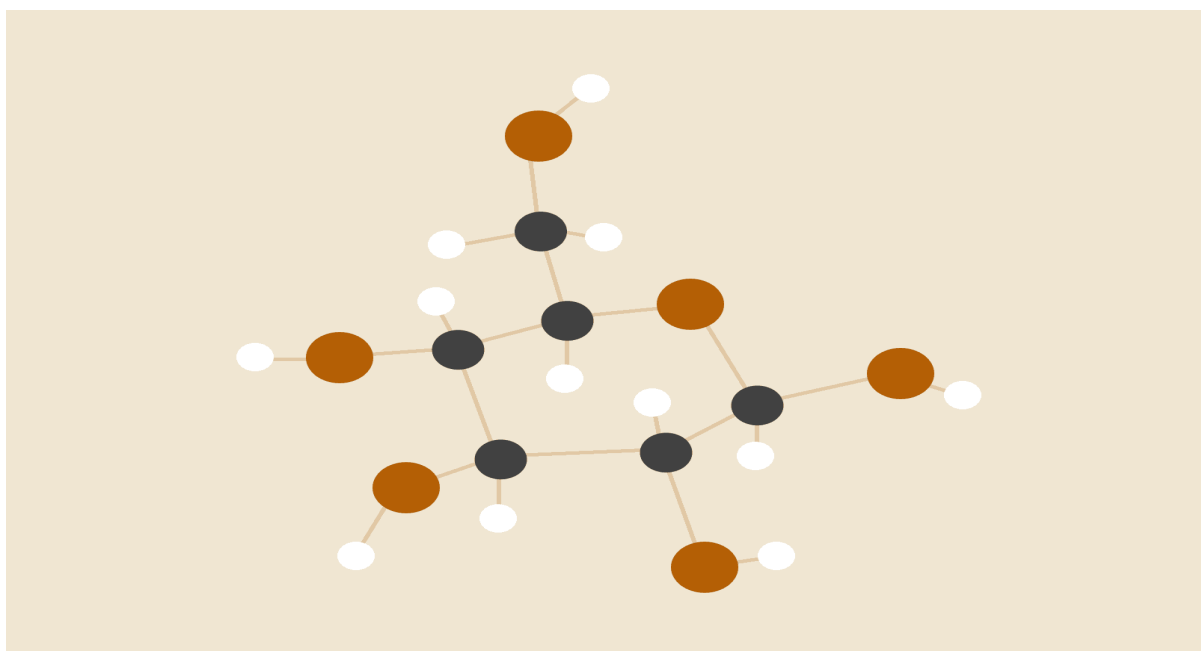


Evaluación de los efectos de los polimorfismos de nucleótidos únicos(SNPs) y sus mecanismos moleculares asociados a hipotensión ortostática neurogénica sobre la severidad y el pronóstico de la enfermedad de Parkinson.

Tesis presentada para optar por el grado de:

Doctor en Ciencias Médicas.



Por:

Doctorando : Guenson Chevalier, MD.

Director: Santiago Lloret Perez, MD, PhD.

Co-director: Francisco Capani, MD, PhD

UNIVERSIDAD ABIERTA INTERAMERICANA, UAI

Facultad de Medicina y Ciencias de la Salud

Buenos Aires, Argentina. Febrero del 2023.

Capítulo I- Enfermedad de Parkinson	23
1.1- Consideraciones generales sobre la Enfermedad de Parkinson	23
1.2- Fisiopatología y etiopatogenia de la enfermedad de Parkinson	23
1.3- Alteraciones genéticas y Enfermedad de Parkinson	25
1.4- Ganglios de la Base	31
1.5- Tratamientos actuales de la Enfermedad de Parkinson	33
Capítulo II- Hipotensión ortostática neurogénica	34
2.1- Sistema autónomo	34
2.2- Consideraciones generales de la Hipotensión Ortostática neurogénica	40
2.3- Fisiopatología de la hipotensión ortostática neurogénica y diagnóstico .	41
2.4- Hipotensión ortostática neurogénica y enfermedad de Parkinson	44
2.5- Hipotensión ortostática neurogénica y polimorfismos genéticos	48
2.6- Complicaciones de la Hipotensión Ortostática neurogénica en EP	49
2.7- Tratamientos actuales disponibles contra la NOH en EP	50
Capítulo III- Genética humana y Polimorfismo de nucleótido único.	51
3.1- Consideraciones sobre el genotipo y fenotipo	51
3.1.1- Transcripción y expresión genética	54
3.2- Polimorfismo de nucleótido único	57
3.4- Identificación de SNPs por secuenciación de última generación	61
3.4.1- Secuenciación de próxima generación	62
3.4.2- Análisis Bioinformáticos:	64
3.4.3- Análisis de genoma entero(GWAS)	65
3.4.4- Recombinación genética y desequilibrio de ligamiento	67
3.5- Fundamentos del análisis de expresión de genes basado en desequilibrio de ligamentos y eQTL	68
Capítulo IV- Métodos de machine learning para la predicción de enfermedades.	69
4.1- Generalidades sobre Machine Learning(Aprendizaje automática)	69
4.2- Neural Network(Redes Neuronales)	70
4.2.1- Métodos	70
4.2.2- Arquitectura	70
4.2.3- Aprendizaje	72
4.2.4- Unidades y pesos	72
4.2.5- Aprendizaje supervisado en redes neuronales	73
4.2.6- Retropropagación	74
4.3- Gradient Boosting Trees	76
4.4- Random Forest	76
Capítulo V- Hipótesis y objetivos	77
5.1- Justificación e impacto	77
5.2- Hipótesis de trabajo	78
5.3- Objetivo general	78
5.4- Objetivos específicos	78

Capítulo VI- Métodos y materiales	79
6.1- Base de datos del PPMI	79
6.2- Aprobaciones de protocolos estándar y consentimientos de los pacientes	80
6.3- Participantes	81
6.4- Medida de la presión arterial y clasificación de los pacientes en NOH o no NOH	81
6.5- Procesamiento de los datos genómicos y análisis genética	81
6.7- Preparación, obtención y mantenimiento de células dopaminérgicas.	83
6.8- Análisis de expresión de genes basado en desequilibrio de ligamiento	84
6.9- Análisis de enriquecimiento de SNPs asociados a NOH	84
6.11- Análisis de expresión de los genes según los alelos de los SNPs asociados a NOH(eQTL).	84
6.12- Identificación de módulos y genes Hub de los eQTL en análisis de network.	85
6.13-Análisis de enriquecimiento de los eQTLs y genes HUB	86
6.15- Análisis de la asociación entre los SNPs asociados a NOH y los datos clínicos.	86
Capítulo VII- Análisis de asociación entre SNPs y NOH	87
7.1- Asociación entre NOH y SNP en datos Baseline	87
7.1.1- Análisis de asociación por regresión logística simple, prueba de Chi cuadrado y/o prueba exacta de Fisher.	87
7.1.2- Análisis de asociación por regresión logística múltiple en modelo aditivo, dominante y recesivo.	90
7.1.3- Análisis epistático de pares de los SNPs	94
Capítulo VIII- Análisis de enriquecimiento de los SNPs asociados a NOH	97
Capítulo IX- Análisis de expresión de genes basado en LD de los SNPs asociados a NOH	99
Capítulo X- Efectos de los SNPs asociados a NOH sobre la severidad y progresión del EP	101
10.1- Análisis de asociación por regresión lineal y ANOVA entre UPDRS I, II, III, IV y los SNPs en los datos Baseline de los pacientes.	101
10.1.1- UPDRS I	102
10.1.2- MDS-UPDRS II	105
10.1.3- MDS-UPDRS III	106
10.1.4- UPDRS IV	110
Capítulo XI- Análisis de eQTL asociados a los SNP asociados a NOH.	113
11.1- El SNP rs2280104(BIN3) asociado a 75 genes en análisis de eQTLs	113
11.2- Identificación de módulos y genes hub asociados con la NOH, riesgo disminuido por rs2280104-BIN3/PPP1R14C	115
11.3- Procesos moleculares y celulares en los cuales participan BIN3-PPP1R14C	120
11.4- Propuesta de vías de señalización por las cuales el BIN3 disminuye el riesgo de NOH.	121
11.5- Los SNP rs55785911 (DDR1) está asociado a múltiples genes en análisis de eQTLs.	124
11.6- El SNP rs13294100 (SH3GL2) está asociado a múltiples genes en análisis de eQTLs.	127
Capítulo XII- Análisis diferencial de genes entre NOH y no-NOH.	136
12.1- Pre-procesamiento de los datos	136

	4
12.2- Análisis de cluster y de componente principal	137
12.3- Análisis diferencial de genes y downstream analysis.	138
12.4- Down-regulación de las vías de señalización de transmisión sináptica	143
12.5- Down-regulación de señalización trans sináptica	143
Capítulo XIII- Análisis de predicción de NOH en pacientes con enfermedad de parkinson por algoritmos de aprendizaje automático.	145
13.1- Algoritmos utilizados y procedimientos	145
13.2- Resultados de las predicción basadas en algoritmos de aprendizaje automático	147
13.2.1- Predicción sobre los datos tanto genéticos, clínicos demográficos	147
13.2.2- Predicción sobre los datos genéticos solamente	150
Capítulo XIV- Discusiones.	153
14.1- SNP y vías de señalización asociados con la NOH en pacientes con EP	153
14.2- SNCA	154
14.3- IP6K2	155
14.4- SH3GL2	155
14.5- Los SNP asociados con NOH alteran la expresión de varios genes en un mecanismo dependiente del desequilibrio de ligamiento.	159
14.6- La variante rs34311866-TMEM175 disminuye la expresión de SNP DGKQ basada en LD y aumenta el riesgo de NOH.	160
14.7- TMEM 175	161
14.8- FAM47E.STBD 1	163
14.9- SCN3A.SCN2A	163
14.10- SH3GL2	164
14.11- SIPA1L2	166
14.12- ITGA8	166
14.13- MIR4697	166
14.14- ELOVL 7.NDU FAF2	166
14.15- ACMSD/TMEM163	167
14.16- DDRGK1	167
14.17- EGF	168
14.18- BIN3	168
Capítulo XV- Conclusión	168

Intentionally left blank page

Resumen

Introducción

La hipotensión ortostática neurogénica (NOH) es un trastorno frecuente y limitante que se presenta en más de un 50% de los pacientes con enfermedad de Parkinson (EP). Resulta en una súbita caída de la presión arterial al cambiar de decúbito supino a posición de pie, lo cual puede generar mareos, desmayos y otros síntomas que incapacitan a los pacientes. La NOH causa serios problemas en pacientes con Enfermedad de Parkinson, especialmente al incrementar el riesgo de caídas y llevar a una importante morbilidad cerebrovascular y cardiovascular. Ya que es un factor de riesgo conocido de mortalidad por cardiopatías y accidentes cerebrovasculares. Se ha observado que la respuesta adrenérgica al ortostatismo está reducida en pacientes con NOH, lo que resulta en niveles circulantes disminuidos de norepinefrina. Aunque la base genética de la NOH en la Enfermedad de Parkinson no ha sido completamente explorada, se ha demostrado que ciertas variantes de nucleótido único (SNV) relacionadas con la enfermedad de Parkinson afectan al sistema nervioso autónomo y se asocian con riesgo de desarrollar NOH o incluso se comportan como factores protectores del mismo.

Materiales y métodos

Los pacientes con EP diagnosticados por tomografía por emisión de positrones (*PET*, por su sigla en inglés), libres de condiciones patológicas asociadas con la hipotensión ortostática, y/o no toman medicamentos asociados a la hipotensión ortostática, fueron seleccionados de las primeras visitas de la base de datos del PPMI. La NOH fue definida como un cambio en la presión arterial sistólica (≥ 20 mmHg) o diastólica (≥ 10 mmHg) dentro de los 3-5 minutos después de ponerse de pie los pacientes seleccionados según los criterios de inclusión. Se utilizó la regresión logística simple y múltiple, la prueba Chi cuadrada, el test de Fisher para relacionar los SNP con la NOH usando modelos genéticos aditivo, dominante y recesivo. Un Análisis de expresión cuantitativa de rasgos de locus (eQTLs) fue llevado a cabo entre el SNP-BIN3 asociados a NOH y datos de ARNseq (conteos de genes) extraídos de células dopaminérgicas derivadas de iPSCs de pacientes con EP usando modelo de regresión y análisis de la varianza. Un análisis de expresión diferencial de genes entre el grupo de pacientes con NOH con respecto al grupo control (pacientes que fueron clasificados estar libres de NOH). Además, se estudió la asociación entre algunos SNPs relacionados a la NOH y la severidad del EP y sobre todo se analizó cómo los alelos de riesgo de algunos SNPs influyen en la severidad de la EP. Se realizó un análisis de predicción y de pronóstico

de la NOH en los pacientes con EP utilizando algoritmos de aprendizaje automático en la plataforma KNIME.

Resultados

Un total de 21 SNPs de los 72 utilizados fueron encontrados asociados a la NOH. Los valores del UPDRS III son significativamente más altos en el grupo con NOH que además presentan rs104893877-SNCA_A53T, con respecto a los que no lo presentan. Hemos encontrado que los genotipos TT del SNP rs2280104-BIN3 se asocian con una sobreexpresión del gen PPP1R14C, el cual a su vez está conectado con otros 325 genes que se involucran en procesos moleculares de crecimiento y proliferación de las neuronas y neurogénesis. Además, hemos identificado procesos en los cuales está involucrado el gen PPP1R14C como uniones a citoesqueleto, uniones a quinasas y a microtúbulos. Hemos identificado en el análisis de eQTL, que los genotipos homocigotas y heterocigotas para el alelo menor del SNP rs55785911-DDRGK1 alteran la expresión génica de 17 genes tanto en el análisis de regresión lineal como la ANOVA(p-valor y FDR<0.05). De los 17 genes, NOTCH1, KIF4A, TSPYL2 son los que más importancia funcional tendría según los resultados del análisis de priorización usando la aplicación Causal-Path de Cystoscope.

585 genes eQTL estadísticamente significativos(p-valor y FDR <0.05) fueron encontrados asociados al SNP SH3GL2_rs13294100 en el análisis de regresión lineal y la ANOVA. Los clusters formados por los 585 eQTL enriquecen significativamente para procesamiento de ARNm, spliceosoma, para biogénesis, transporte y localización del complejo ribonucleoproteico, regulación del transporte nuclear y nucleocitoplasmático, procesos de transporte transmembranal de protones contra gradiente electroquímico, transporte de electrones acoplado al transporte de protones y respiración celular, enriquece para procesos de biosíntesis y metabolismo de fosfolípidos, lípidos y compuestos organofosforados. Se encontraron 5 genes con mayor importancia usando CausalPath de Cytoscape: ATP13A2, LAS1L, KIAA0895, CXADR, IKBKB.

En el análisis de predicción de la NOH, hemos encontrado 3 modelos de aprendizaje automático que mejor se ajustan a los datos, Gradient Boosted Trees, Logistic regression y Neural Network cuando usamos los datos genéticos solo o en combinación con datos clínicos y demográficos. Mientras tanto, cuando usamos solo los datos clínicos y demográficos, los resultados son menos efectivos como predictores de la NOH. Esto quiere decir que los datos clínicos y demográficos solo no pueden predecir la NOH mientras que los datos genéticos ya sea solos o en combinación con los datos clínicos y demográficos predicen eficazmente la NOH.

Palabras Claves: polimorfismos de nucleótidos únicos(SNPs)-hipotensión ortostática neurogénica - Enfermedad de Parkinson.

Abstract

Introduction

Neurogenic orthostatic hypotension (NOH) is a common and disabling non-motor feature that affects more than 30% of patients with Parkinson's disease (EP), sometimes seen in newly diagnosed patients even before starting antiparkinsonian therapy. The NOH can be defined as a sustained reduction in systolic and/or diastolic blood pressure (BP) after orthostasis (abruptly moving from recumbent to standing position), or a Tilt test. In the worst cases, symptoms include general weakness, lightheadedness, dizziness, or even syncope.

NOH causes substantial morbidity and mortality, mainly from the sequelae of falls, and is a risk factor for morbidity and mortality from heart disease, stroke, and cardiovascular mortality. By compromising the autoregulatory mechanism of cerebral blood flow, episodes of hypoperfusion lead to white matter damage, cognitive impairment, and dementia.

The adrenergic response to orthostatism, i-e, increased release of norepinephrine from the sympathetic terminals of the blood vessels, is reduced, and circulating norepinephrine levels are decreased in these patients.

The genetic basis of NOH in EP has not been sufficiently explored. Specific EP-related single nucleotide polymorphisms (SNPs) have been shown to affect the autonomic nervous system and the risk of developing NOH.

Materials and methods

EP patients, free of HO-associated pathological conditions, and not taking HO-associated medications were selected from the Parkinson's Progression Markers Initiative (PPMI) database. . The NOH was defined as a change in systolic (≥ 20 mmHg) or diastolic (≥ 10 mmHg) blood pressure within 3 minutes of standing on the initial visit. Simple and multiple logistic regression were used to relate SNPs to NOH using additive, dominant, and recessive genetic models. The SNPs found associated with NOH were used to perform gene expression and enrichment analysis based on SNPs in linkage disequilibrium with the SNPs associated with NOH. Analysis of eQTLs was carried out between NOH-associated SNPs and genes using simple linear regression analysis and analysis of variance. In addition, the effect of some SNPs associated with NOH on the severity of EP was determined. A forecast and prognostic analysis of the NOH in patients with EP using machine learning algorithms.

Results

A total of 23 SNPs out of the 72 used were found to be associated with NOH. SNPs rs104893877-SNCA, rs353116-SCN2A, rs13294100-SH3GL2, rs3910105-SNCA,

rs34311866-TMEM175, rs329648-MIR4697 increases the risk of NOH in a dominant genetic transmission model. While the SNPs rs11060180-CCDC62, rs6812193(FAM47E-STD1) increase the risk of NOH in a recessive genetic model.

The SNPs rs10906923-FAM171A1.ITGA8, rs2694528-ELOVL7-NDUFAF2 decrease the risk of NOH in a recessive model of transmission of genetic information. The SNPs GCH1_rs11158026, rs104893877-SNCA_A53T are associated with UPDRS III. The UPDRS III values are significantly higher in the group with NOH that also present the SNPs GCH1_rs11158026 and rs104893877-SNCA_A53T, with respect to those who do not present them. Patients with NOH and also presenting these SNPs develop more severe forms and courses of Parkinson's disease according to the results of our study.

We have elucidated that the TT genotypes of the rs2280104-BIN3 SNP are associated with an overexpression of the PPP1R14C gene, which in turn is connected with many other genes that are involved in cellular and molecular processes of growth and proliferation of neurons. In addition, we have found processes involved such as process binding to the cytoskeleton, binding to kinases and microtubules.

We have found, in the eQTL analysis, that homozygous and heterozygous genotypes for the minor allele of SNP rs55785911-DDRGK1 altered the gene expression of 17 genes in both the linear regression analysis and the ANOVA (p-value and FDR<0.05). Of the 17 genes, NOTCH1, KIF4A, TSPYL2 are the ones that would have the most functional importance according to the results of the prioritization analysis using the Causal-Path application of Cystoscope.

585 statistically significant eQTL genes (p-value and FDR <0.05) were found associated with the SNP SH3GL2_rs13294100 in linear regression analysis and ANOVA. A network analysis of the 585 eQTLs allowed the construction of 4 clusters. The first cluster significantly enriched mRNA, spliceosome processing. The second cluster enriches for biogenesis, transport and localization of the ribonucleoprotein complex, regulation of nuclear and nucleocytoplasmic transport. The third cluster enriches for processes of transmembrane transport of protons against electrochemical gradient, electron transport coupled to proton transport, aerobic electron transport chain, ATP synthesis coupled to electron transport, ATP metabolism processes, complex assembly NADH dehydrogenase, mitochondrial respiratory chain complex I assembly. The fourth cluster where our gene of interest (SH3GL2) is located, enriches processes of biosynthesis and metabolism of phospholipids, lipids, and organophosphorus compounds. In addition, it was found that the SH3GL2 gene where the SNP is located is connected to 5 genes, which are more important because they allow connections between the different clusters and the SH3GL2 gene. These proteins were chosen based on their degrees, their centralities in the Network. This allows us to add that

these 5 genes are central points in the network through which the information flow has to pass to travel from any node to another: ATP13A2, LAS1L, KIAA0895, CXADR, IKBKB.

In the predictive analysis of the NOH, we have found 3 best machine learning models that best fit the data: Gradient Boosted Trees, Logistic regression y Neural Network when we use both genetic and clinical and demographic data (AUC:0.84, accuracy:0.98, specificity and precision:1 each, sensitivity: 0.83 and NPV(Negative predictive value): 0.975, In the logistic regression the AUC was found in 0.85, accuracy at 0.96, precision and sensitivity at 0.83 each, and specificity and NPV at 0.974 each. The other best performing algorithm is neural network with an AUC of 0.91, accuracy of 0.95, specificity and precision of 1 each, sensitivity of 0.83 and NPV of 0.975. Another parameter is the error rate found at 0.02 in the case of the Gradient Boosted Trees algorithm and Neural Network and 0.04 in the case of logistic regression. In the 3 algorithms a false negative rate of 0.16 has been found).

In the case of the polygenic score, where only genetic data was used, omitting demographic and clinical data, we have not found differences in terms of the previous results. Meanwhile, when we use only the clinical and demographic data, the results are unusable as predictors. This means that clinical and demographic data alone cannot predict the onset of NOH while the genetic data either alone or in combination with clinical and/or demographic data can be used as predictors of NOH.

Intentionally left blank page

DEDICACIÓN

En primer lugar, quiero tomar un espacio especial para dedicar esta tesis a tres personas mas importantes en mi vida. El amor que les tengo es infinito. Mi abuela HERMENCE SEVERE y mi tía PHILOMENE SEVERE, las dos personas que me inculcaron la perseverancia, el amor a la vida. Recuerdo cuando las dos discuten sobre quién querían que yo fuera cuando yo sea grande. Mi abuela quería que yo fuera un “Savant” y mi tía decía que quería que yo fuera un “doctor”(risa). Aunque no estemos juntos, pero quiero que sepan que les agradezco mucho por lo que han hecho por mi y que les quiero mucho.

Mi padre, profesor Joel Chevalier, quien me inculcó el sentido de siempre preguntar sobre los eventos, las situaciones pero además de siempre buscar respuestas. Esto para mi es lo mejor que pudiste haberme enseñado y te seré siempre agradecido papá.

AGRADECIMIENTOS

Yo agradezco en primer lugar a mi familia por soportarme en este largo camino. Sobre todo mi abuela Hermence Severe, mi tía Philomene Severe y mi padre Joel Chevalier. Agradezco a mi madre Carole Severe por mi existencia, es una mujer luchadora y fuerte. Creo que yo y ella tenemos mucho en común. Te quiero mami. A mis hermanos y hermanas (Cleeford, Stephanie, Sonsón, Medjine, Flore, Jefferson, Mestarine), a mis hijas Lova Isabel Chevalier y Bel-Lisa Chevalier.

Agradezco al Dr Francisco Capani por haberme abierto las puertas de su laboratorio para poder doctorarme. Quiero que sepas que te seré siempre agradecido por esto, porque me diste la mano en el momento que más lo necesitaba y para mi es algo grande.

Agradezco a mi director de tesis Dr. Santiago Perez Lloret por su gran ayuda y por meter su confianza en mí. Gracias a ti pude terminar esta tesis en tiempo y forma. Te estoy muy agradecido Dr Santiago.

Agradezco a todo el equipo del CAECIHS, Liliana, un amor de persona. A Carlos Kusner mi gran amigo con quien yo hablo de todo tipo de tema(Ciencia, política, cultura, filosofía). Un gran personaje. Camilia Guilamares, no te puedo olvidar otro personaje grande, gracias Cami por los momentos de alegría que me compartiste. Un agradecimiento especial a Nicolas Toro por su ayuda. Nicolas fue mi primer amigo del grupo y el primero en entenderme. ¡Te agradezco campeón! Agradecimiento especial a Lucas por su gran ayuda, quien me hace entender algunos aspectos buenos que tengo en cuanto a investigación y docencia. El siempre me dice Guenson tienes don para la docencia. Es lindo encontrar

gente que te ayude a reconocer las buenas cosas que tienes en ti. Lucas mi amigo te agradezco mucho. Gracias a ti esta tesis pudo llevarse a cabo también.

Agradezco a todos, Juan, Sofia, Ines, Graciela, Gabriel, Kollikers, las dos Camillas, y todos los que de una forma u otra directa o indirectamente ayudan a que esta tesis pueda llegar a término.

Quiero agradecer también a la institución CAECIHS por abrirme las puertas para realizar mi tesis doctoral. Son oportunidades que se me dieron que nunca olvidaré.

Por fin, quiero agradecer a la institución CONICET por la oportunidad que me dio para poder doctorarme dándome una beca de 5 años. Yo me considero como alguien muy afortunado por esta oportunidad por esto quiero dedicar un agradecimiento especial para CONICET.

Abreviaturas

AINES	Anti-inflamatorios no esteroideos
AUC	Area under the ROC curve
AVP	arginina vasopresina
DBH	deficiencia de Dopamina beta hidroxilasa
DBS	Deep Brain stimulation
DOPAC	Dihydroxyphenylacetic acid
DGE	Análisis de expresión diferencial de genes
DM	Diabetes Mellitus
EP	enfermedad de Parkinson
eQTL	expresión quantitative trait loci
FC	frecuencia cardiaca
FDR	False Discovery rate
GWAS	Genome wide association study
GPI	Globo pálido interno
GPe	Globo pálido externo
HPFS	Health Study and the Health Professionals Follow-up Study
iPSC	Induced pluripotent stem cells
LCR	Líquido cefalorraquídeo
LD	Desequilibrio de ligamiento
MAS	multiple atrophy system
MET	Transición epitelio mesenquimal
MOCA	Montreal cognitive assessment
MVLR	Médula ventrolateral rostral
NGS	Secuenciación de próxima generación(Next sequencing generation)
NOH	Neurogenic Orthostatic Hypotension

NPV	Negative predictive value
NTS	Núcleo del tracto solitario
OH	Hipotension ortostatica
PAF	Pure Autonomic failure
PCA	Análisis de componente principal
EP	Parkinson disease
PPMI	Parkinson's Progression Markers Initiative
PCR	Reacción en cadena de la polimerasa
PIP	Predicción de la interacción proteína-proteína
RE	Retículo Endoplásmico
RAAS	Activación del sistema renina angiotensina
rCBF	Flujo sanguíneo Cerebral regional
ROC	Receiver operating characteristic
SNP	single nucleotide polymorphism
SNc	Sustancia nigra pars compacta
SPECT	Tomografía computarizada por emisión de fotones
SBS	Secuenciación por síntesis
SNC	Sistema nervioso central
SNS	Sistema nervioso simpático
TCE	Traumatismo craneoencefálico
UPDRS	Unified Parkinson's Disease rating Scale

Listas de las tablas

Tabla 1- Contenido en lisina y arginina de las histonas del nucleosoma.

Tabla 2- tabla de contingencia

Tabla 3- SNPs asociados a NOH en una regresión logística simple y un modelo genético aditivo

Tabla 4- Análisis de asociación por regresión logística simple . Modelo aditivo, recesivo y dominante

Tabla 5- Análisis de asociación por regresión logística múltiple. Modelo aditivo

Tabla 6- Análisis de asociación por regresión logística múltiple. Modelo dominante.

Tabla 7- Análisis de asociación por regresión logística múltiple. Modelo recesivo.

Tabla 8 - Cuadro comparativo entre los diferentes modelos aditivo, dominante y recesivo

Tabla 9- Análisis epistático, modelo aditivo contra modelo aditivo

Tabla 10- Análisis epistático, modelo recesivo contra modelo aditivo.

Tabla 11- Análisis epistático, modelo dominante contra modelo aditivo

Table 12. Enrichment analysis of SNPs associated with NOH

Tabla 13- Análisis de expresión de genes basado en el desequilibrio de ligamiento entre los SNPs

Tabla 14- Análisis de asociación entre UPDRS I y variantes genéticas por regresión lineal simple.

Tabla 15- Análisis de ANOVA de una vía sobre SNPs asociados a UPDRS I según los genotipos.

Tabla 16- Análisis de asociación entre los valores del MDS-UPDRS II y los SNPs asociados a NOH.

Tabla 17- Distribución de los valores del MDS-UPDRS II según los genotipos de los SNPs encontrados asociados usando el análisis de la varianza de una vía(ANOVA).

Tabla 18- Análisis de asociación entre los valores del MDS-UPDRS III y los SNPs asociados a NOH.

Tabla 19- Distribución de los valores del UPDRS III según los genotipos de los SNPs encontrados asociados usando el análisis de la varianza de una vía(ANOVA)

Tabla 20- Análisis de asociación entre los valores del UPDRS IV y los SNPs asociados a NOH.

Tabla 21- Distribución de los valores del MDS-UPDRS IV según los genotipos de los SNPs asociados a MDS-UPDRS IV. (ANOVA de una vía).

Tabla 22- Análisis de asociación por regresión lineal de los SNPs y la escala NYH.

Tabla 23- Valores de NYH Scale según los genotipos de los SNPs encontrados asociados a la escala NYH.

Tabla 24- Expresión de los genes bajo la influencia del SNP rs55785911-DRGK1

Tabla 25- Análisis de regresión entre el gen SH3GL2 y 25 eQTLs elegidos

Tabla 26- Expresión de los genes bajo la influencia del SNP rs13294100-SH3GL2.

Tabla 27 - Confirmación de los SNPs asociados a NOH por expresión diferencial de genes.

Tabla- 28- Los parámetros de evaluación de los algoritmos de aprendizaje automático en el conjunto de los datos(demográficos, clínico, genético).

Tabla 29- Evaluación de los algoritmos de aprendizaje automáticos para los datos genéticos(Excluyendo datos clínicos y demográficos)

Tabla 30- Evaluación de algoritmos de aprendizaje automático para predicción de NOH en datos clínicos y demográficos(Excluyendo a datos genéticos)

Listas de figuras

Figura 1(A-C)- Factores de riesgos asociados a la enfermedad de Parkinson

Figura 2. Vías directa e indirecta de los ganglios basales

Figura 3- Funciones del sistema autónomo.

Figura 4- Sistema nervioso autónomo y los receptores involucrados

Figura 5- Inervación y función del sistema nervioso autónomo

Figura 6: Circuito neuronal del arco barorreflejo

Figura 7- Estructura secundaria del ADN

Figura 8- Proceso de corte y empalme de ARNm

Figura 9: Resumen del dogma central de la genética

Figura 10: representación esquemática de variantes genéticas SNPs y genotipos.

Figura 11(A-B): Tecnología de secuenciación del ADN de la plataforma Illumina.

Figura 12- Workflow para análisis bioinformáticos de los datos genéticos tipo fastq

Figura 13- Arquitectura del algoritmo Neural Network Multilayer Perceptron

Figura 14- Representación esquemática de los datos del PPMI

Figura 15- Distribución de los valores del UPDRS I según los genotipos de los SNPs rs34043159-MAP4K4 y rs11158026-GCH1.

Figure 16- Distribución de los valores del UPDRS III según los genotipos de los SNPs

Figura 17A,B- Procesos moleculares enriquecidos para los eQTLs asociados al SNP rs2280104-BIN3

Figura 18(A-F)- Análisis de eQTLs asociados a BIN3 como factor protector de la NOH

Figura 19(A-C)- Enriquecimiento de los 7 genes Hubs encontrados en el cluster PPP1R14C

Figura 20- Análisis de redes de las proteínas extraídas de la base de datos STRINGs

Figura 21. A: ScatterPlot de los coeficientes de la regresión lineal entre los 17 genes eQTL cuyas expresiones son influenciados por los genotipos homocigotas y heterocigotas asociados a NOH.

Figura 22 A-E- Análisis de network y enriquecimiento de los eQTLs asociados al SNP SH3GL2_rs13294100.

Figura 23- Clusters de eQTLs asociados al gen SH3GL2

Figura 24- ANOVA de los eQTLs HUB con mayor influencia en el network.

Figura 25 A/B- Heatmap de la expresión de genes en grupos NOH con respecto a los No-NOH

Figura 26(A-D)- Análisis diferencial de genes entre NOH y no-NOH

Figura 27- Enriquecimiento de los genes diferencialmente expresados en NOH con respecto al grupo No-NOH. (Procesos moleculares)

Figura 28- Enriquecimiento de los genes diferencialmente expresados en NOH con respecto al grupo No-NOH. (Vías de señalización y/o metabólicas).

Figura 29- Enriquecimiento de los genes diferencialmente expresados en NOH con respecto al grupo No-NOH. (Vías de señalización y/o metabólicas) by KEGG database.

Figura 30: Representación de los algoritmos de aprendizaje automático en el conjunto de los datos(Clínicos, demográficos, genéticos).

Figura 31- Parámetros de evaluación de los algoritmos de aprendizaje automático sobre el conjunto de los datos(clínicos, demográficos, genéticos).

Figura 32- Parámetros de evaluación de los algoritmos de aprendizaje automático en los datos genéticos(Excluyendo datos demográficos y clínicos).

Figura 33- Matrix de confusión de los algoritmos de aprendizaje automático en los datos clínicos y demográficos(Excluyendo a datos genéticos).

Figura 34- Figura de resumen 1: Resumen de los resultados de los tres primeros objetivos de la tesis.

Figura 35- Figura de resumen 2: Análisis diferencial de genes entre los pacientes con NOH y grupo control.

Thesis Outline

Esta tesis está dividida en 3 partes en las cuales se distribuyen en 15 capítulos. La primera parte, formada por cinco capítulos, contiene el marco teórico del tema que se desarrolla en el presente trabajo. **El primer capítulo** se centra sobre la Enfermedad de Parkinson, aspectos generales y clínicas de la enfermedad de Parkinson, fisiopatología y etiopatogenia además de aspectos relacionados con los genes y SNPs(*single nucleotide polymorphism* por su sigla en inglés) conocidos por estar asociados con la enfermedad de Parkinson. Además, la última sesión de este primer capítulo habla de forma resumida del tratamiento actual de la Enfermedad. **El capítulo dos** se centra sobre la hipotensión ortostática neurogénica en los pacientes con enfermedad de parkinson sobre todo, los aspectos vinculados con la clínica de la misma, su fisiopatología, complicaciones y tratamiento disponible. **En el capítulo tres**, expresa someramente sobre la dogma central de la genética humana, fundamentos de la expresión genética, fundamentos del análisis de mapeo eQTL, análisis de enriquecimiento, explicaciones resumidas sobre genotipos y fenotipos, los diferentes análisis estadísticas realizados en este estudio como regresión lineal, regresión logística, ANOVA de una vía, Chi cuadrado, Test de Fisher y sus usos en el estudio de asociación en enfermedades de causas multilocus(poligénica). Este capítulo termina con un resumen sobre la base de datos PPMI y los diferentes datasets usados para hacer esta tesis. **El capítulo cuatro** explica de forma detallada la teoría que está detrás de algunos de los algoritmos de aprendizaje automático que se usan en este trabajo para hacer la predicción de los SNPs en la NOH. Los algoritmos de aprendizaje automático como Neural network, Random Forest, Naives Bayes, XGBoost Trees y modelo lineal generalizado lineal. **El capítulo cinco** habla de los objetivos e hipótesis del trabajo, justificación e impacto del estudio. La segunda parte de este trabajo por su lado contiene dos capítulos enfocados sobre la descripción de las metodologías utilizadas para poder llegar a los resultados. La parte dos empieza con **el capítulo seis** en el cual se detallan los materiales y métodos utilizados en este estudio. **El capítulo siete** detalla los resultados del análisis de asociación entre SNPs y NOH utilizando diferentes test estadísticos como regresión logística y prueba de Chi Cuadrado. En **el capítulo ocho**, se presentan los resultados del análisis de enriquecimiento de los SNPs encontrados asociados al NOH y vías de señalización asociados al NOH. Los resultados del análisis de Trans eQTL in silico usando los datos de la base de datos GTEX se presentan en el **capítulo nueve**. En **el capítulo diez** se presentó los resultados de los efectos de los SNPs sobre los valores del score MDS-UPDRS, el cual nos permite evaluar como los SNPs se asocian con gravedad de la EP. En **el capítulo once**, se presentaron los resultados del análisis de eQTL utilizando células dopaminérgicas derivadas de iPSCs de pacientes con EP. En **el capítulo 12** se

realizó un análisis diferencial de genes en pacientes con NOH con respecto a los que no. Con este análisis, se pretende descubrir nuevos marcadores de diagnóstico de la NOH y validación de los SNPs asociados a NOH en enfermos de parkinson. En el **capítulo trece** se usaron 7 algoritmos de aprendizaje automático en el análisis de predicción de la condición de NOH en pacientes con enfermedad de Parkinson. Para terminar, las discusiones generales del estudio se presentan en el **capítulo catorce** y las conclusiones en el **capítulo quince**. La idea detrás de estos análisis es en primer lugar descubrir los SNPs asociados a la NOH y usar algunos de los mismos para estudiar los mecanismos moleculares y vías de señalización que intervienen en el proceso de la enfermedad y sobre todo descubrir proteínas que se podrá usar como blanco terapéutico. Para llevar a cabo este objetivo, se hizo análisis de regresión logística, prueba de Chi cuadrado y otros análisis para descubrir dichas asociaciones SNP-NOH. Una vez obtenidos los SNPs asociados a NOH, se decidió estudiar los mecanismos moleculares que intervienen en la patogenia de la NOH en los pacientes con EP según los SNPs asociados. Por eso hicimos analisis de expresión diferencial de genes, analisis eQTL cis trans in silico(Base de datos GTEx) y usando los datos de los pacientes, además de analisis de enriquecimiento de genes y analisis de interacción proteína-proteína. Una vez conocidos los mecanismos moleculares, nosotros estudiamos el efecto que tiene estos SNPs asociados a NOH sobre la severidad y el pronóstico usando los scores UPDRS I, II, III, IV. Después de esto, se hizo un análisis de predicción usando los datos de los SNPs asociados a NOH, los datos clínicos y 7 algoritmos de aprendizaje automático. Los resultados de los algoritmos fueron comparados para filtrar los primeros mejores resultados. En resumen, con esta tesis, se pudo determinar SNPs que son asociados a NOH en pacientes con EP y asociarlos, en caso de algunos SNPs, con los mecanismos moleculares en los cuales intervienen, además de determinar el efecto de los mismos sobre la severidad de la EP. Se pudo construir un workflow de predicción de la NOH a partir de los SNPs.

Los artículos derivados de esta tesis son los siguientes:

1- **Chevalier Guenson**, Udovin L, Otero-Losada M, Bordet S, Capani F, Luo S, Goetz CG, Perez-Lloret S. **Genetics of Neurogenic Orthostatic Hypotension in Parkinson's Disease, Results from a Cross-Sectional in Silico Study**. Brain Sci. 2023 Mar 17;13(3):506. doi: 10.3390/brainsci13030506. PMID: 36979316; PMCID: PMC10046202.

2- **Chevalier Guenson**, Udovin L, Otero-Losada M, Bordet S, Capani F, Luo S, Goetz CG, Perez-Lloret S, Francisco Capani. eQTLs revealed by linear regression and gene correlation-based network analysis identify BIN3-PPP1R14C as a key protector factor in the neurogenic orthostatic hypotension from Parkinson disease patients. **(En revision para publicar)**.

3- **Chevalier Guenson**, Udovin L, Otero-Losada M, Bordet S, Capani F, Luo S, Goetz CG, Perez-Lloret S, Francisco Capani. Polygenic risk score analysis associated with NOH in Parkinson disease patients using Machine Learning algorithms(**en escritura**).

4- **Chevalier Guenson**, Udovin L, Otero-Losada M, Bordet S, Capani F, Luo S, Goetz CG, Perez-Lloret S, Francisco Capani. MicroRNAs and rs1803274 SNP-based BUCHE downregulation is associated with metabolic syndrome through ghrelin hydrolysis and eQTLs expression regulation in PD patients (**En revisión para publicación**).

PARTE I

INTRODUCCIÓN

Capítulo I- Enfermedad de Parkinson

1.1- Consideraciones generales sobre la Enfermedad de Parkinson

Descrito por vez primera en el año 1887 por James Parkinson, la enfermedad de Parkinson (PD) es la segunda enfermedad neurodegenerativa conocida después de la enfermedad del Alzheimer(7). Desde su descripción en aquel momento, se han producido numerosos avances en relación a su fisiopatología, descripción de subtipos clínicos, presencia de síntomas no motores y su relación con la severidad de la EP, métodos diagnósticos(1), tratamiento de su comorbilidad asociada, tratamiento no farmacológico y pronóstico de la enfermedad(2). Sin embargo, los tratamientos antiparkinsonianos disponibles en la actualidad no impiden eficazmente el avance hacia las formas graves de la enfermedad e incluso la muerte(3). Algunos medicamentos utilizados en la EP producen complicaciones inesperadas como es el caso del trastorno impulsivo asociado a la toma de agonistas dopaminérgicos(4). Si bien se han descrito diversos mecanismos patogénicos involucrados (genética, ambiental tóxica, inflamación e inmunitaria, disfunción mitocondrial y el estrés oxidativo)(5), no se conoce fehacientemente el mecanismo que desencadena o perpetúa la inexorable muerte de las neuronas dopaminérgicas. Considerando que la EP es la segunda enfermedad neurodegenerativa más frecuente en pacientes gerontológicos y que su prevalencia aumenta con la edad y el envejecimiento poblacional(6), el desarrollo de un adecuado tratamiento contra la misma resulta de máxima prioridad. Está considerado como un problema de salud pública porque afecta a más del 1.5% de la población mayor de 60 años y a más del 5% de la población mayores de 85 años(8). La incidencia de la enfermedad de Parkinson va desde 16 hasta 19 por cada 100.000 habitantes por año(9). El principal evento celular que conlleva a la EP es una pérdida progresiva y mantenida de neuronas dopaminérgicas de la sustancia nigra y de la vía nigroestriatal(10), (11). Siendo una enfermedad crónica y progresiva, empiezan a aparecer las manifestaciones clínicas de la enfermedad cuando la pérdida de las neuronas llega a unos 70-90% de las poblaciones neuronales de la sustancia nigra pars compacta(SNc)(12). Fearnley JM y colaboradores

han demostrado que al inicio de los síntomas del PD se produce una pérdida de células del 68 % a nivel de la región ventrolateral y una pérdida del 48 % en región caudal de la sustancia nigra(12). EL evento molecular bien conocido en la EP es la formación y depósito de cuerpos de inclusión llamada cuerpo de Lewy en las células de la sustancia nigra, constituido principalmente por precipitados de proteínas alfa-sinucleína(13)(14), descubierto por vez primera por Frederick Lewy en 1912, por eso lleva su nombre. En las últimas décadas, en investigaciones sobre la EP, se han descrito otros procesos moleculares que conllevan a la muerte celular dopaminérgica como disfunción mitocondrial incluyendo alteraciones de la cadena de transporte mitocondrial y síntesis de ATP, alteración en el equilibrio fisión/fusión mitocondrial, fragmentación mitocondrial, movilidad alterada de las mitocondrias, afectación de la mitofagia, producción en exceso de radicales libres, neuroinflamación mediada por la microglia, trastornos en el tráfico de vesículas endolisosomales y sinápticas, alteración de la homeostasis del calcio, alteración en la neurotransmisión y estrés oxidativo entre otros. Estos procesos moleculares son responsables de las muertes celulares dopaminérgicas en el pars compacta de la sustancia nigra y por consiguiente la neurodegeneración y enfermedad de parkinson. Esta deficiencia de las neuronas nigroestriatales conduce a la aparición de los síntomas principales de la EP como la bradicinesia, los temblores, alteraciones posturales y rigidez muscular(15),(16). A medida que se va estableciendo la enfermedad, van apareciendo otros síntomas no-motores como la depresión, alteración cognitiva, afectación del humor(7). En el siguiente apartado, vamos a detallar los mecanismos fisiopatológicos y etiopatogénicos que conllevan al establecimiento de la EP en los pacientes.

1.2- Fisiopatología y etiopatogenia de la enfermedad de Parkinson

Hasta el momento, se han identificado varios procesos etiológicos y fisiopatológicos de la enfermedad de Parkinson, los cuales se basan en una variedad de factores, que abarcan desde factores genéticos, ambientales, tóxicos, además de factores inflamatorios, inmunitarios, disfunción mitocondrial y la generación de radicales libres(Simon, Tanner, and Brundin 2020b)(figura 1A-C).

Varios científicos que investigan la enfermedad de Parkinson están de acuerdo en que el 10% de la EP está causado por factores genéticos(H. Deng, Wang, and Jankovic 2018). Mutaciones en los genes SNCA, LRRK2, GBA, VPS35 causan la forma familiar autosómico dominante de la enfermedad(H. Deng, Wang, and Jankovic 2018; Funayama et al. 2023)(Salles, Tirapegui, and Chaná-Cuevas 2024). Por otro lado, mutaciones en los genes parkin, PINK1 y DJ-1 causan la forma autosómica recesiva de la enfermedad y aparece a

temprana edad(Funayama et al. 2023),(Salles, Tirapegui, and Chaná-Cuevas 2024)(Figura 1A-C).

Por otro lado, existen variantes genéticas tipo SNVs(*single nucleotide variants* por su definición en inglés) localizadas en los genes que causan las formas familiares de la EP o que se localizan en otros genes están asociadas con riesgo elevado de enfermedad de parkinson o incluso algunas variantes pueden comportarse como factores protectores(Day and Mullin 2021). Estas variantes no causan la EP familiar pero interactúan con otras variantes genéticas y/o con factores ambientales para generar el desarrollo de la enfermedad de parkinson(Day and Mullin 2021; Lohmueller et al. 2003)(Figura 1A-C).

Según varios estudios, el 90% restante de la enfermedad de parkinson es de origen multifactorial que se desarrolla a partir de la combinación e interacción entre estas variantes genéticas y factores ambientales.

La edad es uno de los factores de riesgo más importantes en el desarrollo de la EP y los hombres son los más afectados por esta patología(Joseph Jankovic and Tan 2020).

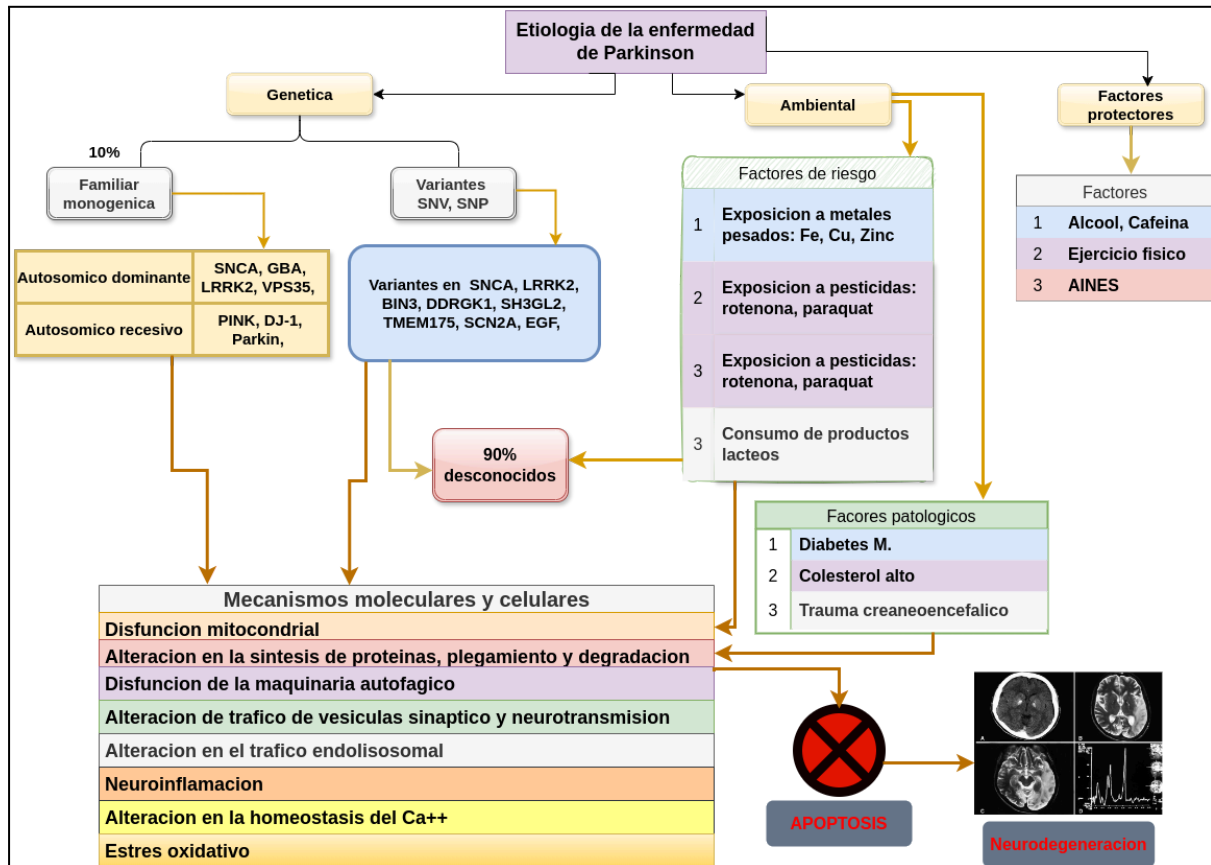
En cuanto a factores ambientales, varios estudios han demostrado asociaciones entre la enfermedad de Parkinson y eventos como exposiciones a metales pesados y pesticidas(Langston et al. 1983), y ocupacional(agricultura)(Petrovitch et al. 2002). Los metales pesados como el cobre se unen con la dopamina en el espacio sináptico y se forma el complejo dopamina-cu⁺⁺, el cual se introduce dentro de la célula utilizando el transportador de la dopamina. El complejo dopamina-Cu⁺⁺ se oxida dentro de la célula para dar formación a productos como el aminocromo y radicales libres los cuales son neurotóxicos y provocan muerte celular(Pyatha et al. 2022)(Figura 1A-C).

Antecedente de traumatismo craneoencefálico (Rugbjerg et al. 2008), consumo de productos lácteos en hombre pero no en mujeres(the Nurses' Health Study and the Health Professionals Follow-up Study (HPFS)(H. Chen et al. 2002), diabetes mellitus tipo II(Ascherio and Schwarzschild 2016) aumentan el riesgo de enfermedad de parkinson. Por otro lado, paradójicamente, sustancias conocidas como tóxicas para el cuerpo como el tabaco y el alcohol(cerveza principalmente) parece tener efecto protector en la EP(Thacker et al. 2007), (D. Zhang, Jiang, and Xie 2014). Otros factores protectores como la cafeína y el té, antiinflamatorios no esteroideos, actividad física y consumo de estatinas se han descrito en otros estudios(Ascherio and Schwarzschild 2016). Hasta la fecha muchos de estos estudios no podían ser reproducidos o confirmados en estudios posteriores. Por ejemplo, en el caso de los productos lácteos, en el mismo estudio en el cual lo han descubierto, se ha encontrado que los productos podrían estar contaminados con el heptacloro epóxido en los grupos que tenían EP en relación a los que no(Ascherio and Schwarzschild 2016). En el caso del alcohol, en un estudio en pacientes Swedish que incluye a más de 1000 casos de

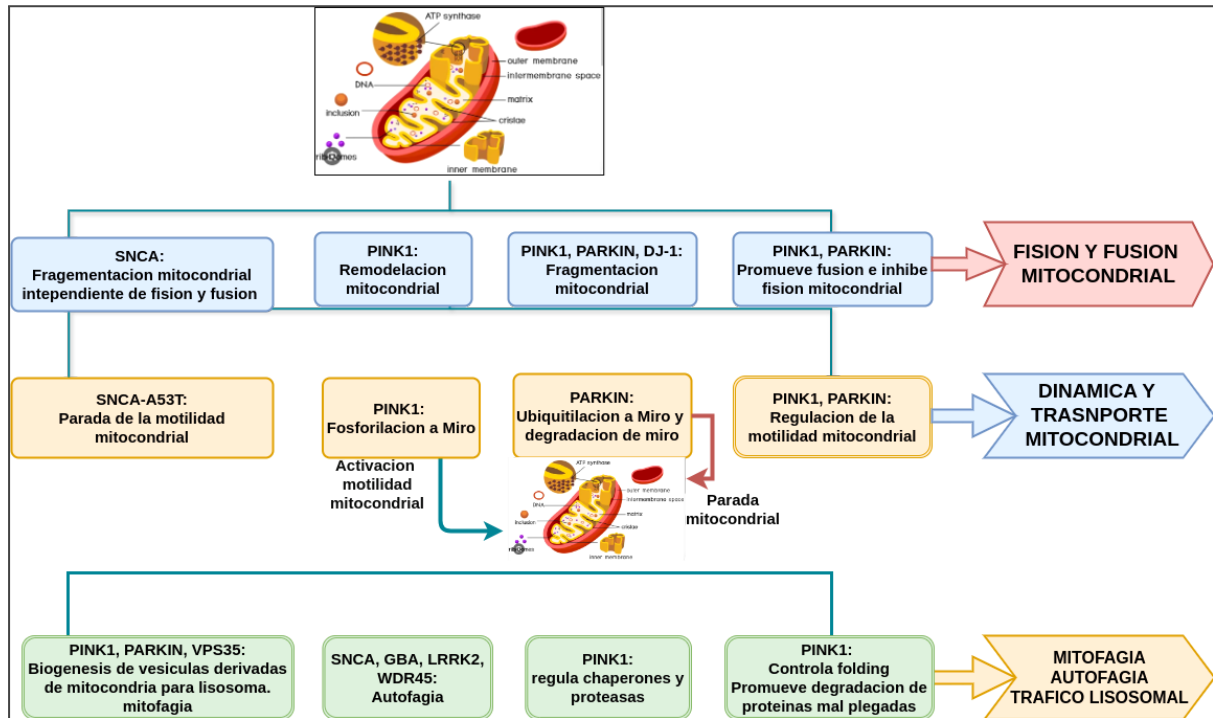
EP, se encontró al contrario un aumento del riesgo de EP en los individuos que tomaban alcohol (Eriksson et al. 2013). Lo que sí está comprobado es que los factores ambientales ya conocidos, tienen una participación importante en el desarrollo y la progresión de la enfermedad de parkinson.

Además de los factores ambientales, se han identificado numerosas variantes y mutaciones genéticas como factores de riesgo para la enfermedad de Parkinson. Estos hallazgos respaldan la influencia genética en el desarrollo de esta enfermedad neurodegenerativa (Day and Mullin 2021). Se calcula que la variación genética contribuye aproximadamente en un 25-36% al riesgo global de desarrollar EP del 90% (Keller et al. 2012), (Keller et al. 2012; Nalls et al. 2019). Todas son variantes de genes que participan en diferentes vías de señalización y funciones moleculares muy importantes en la célula como el transporte intracelular (van Veen et al. 2020), mecanismos de desintoxicación celular, identificación y degradación de proteínas por el sistema proteosomal (Y. Zhang et al. 2000), plegamiento de las proteínas (Day and Mullin 2021; Polymeropoulos et al. 1996), (Krüger et al. 1998), (Krüger et al. 1998; Benitez et al. 2016). A continuación, vamos a explicar algunos ejemplos de mecanismos moleculares ya conocidos por los cuales variaciones en algunos genes producen la enfermedad de Parkinson.

1-A



1-B



1-C

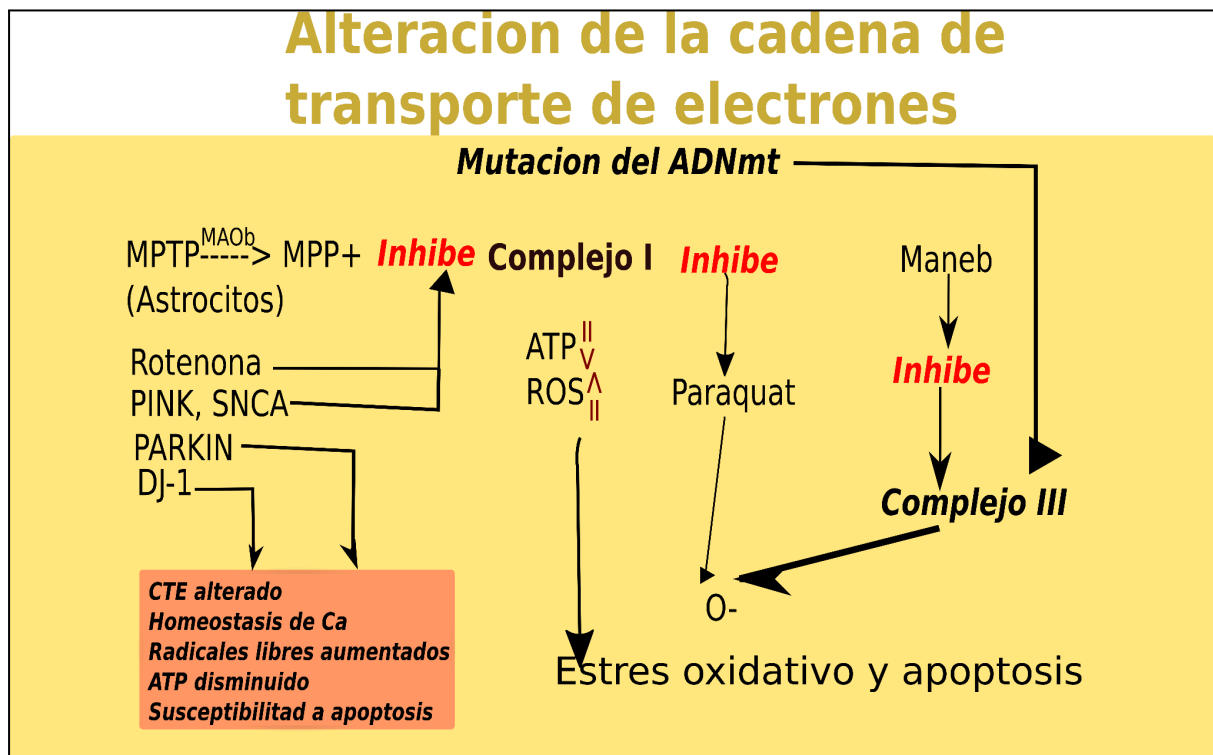


Figura 1 A-C: Factores de riesgos asociados a la enfermedad de Parkinson. Figura realizada en nuestro laboratorio. La figura hace referencia sobre los factores de riesgo asociados a la EP. Se describen tanto los factores genéticos, factores ambientales, patológicos asociados que aumentan el riesgo de la EP como factores que disminuyen el riesgo de EP como el

consumo de alcohol, cafeína, AINES y ejercicio físico. En la **figura 1A**, se describen los factores de riesgos o factores causales asociados a la enfermedad de Parkinson. En la **figura 1B** se describen las alteraciones mitocondriales, autofágicas, y tráfico endolisosomal y la participación de los genes SNCA, PARKIN, PINK1, DJ-1, VPS35 en las mismas. La figura 1C describe cómo los genes SNCA, PINK1, PARKIN, DJ-1 y los pesticidas como paraquat, rotenona y Maned alteran la cadena de transporte de electrones, la síntesis de ATP y la producción de radicales libres en las células.

1.3- Alteraciones genéticas y Enfermedad de Parkinson.

1.3.1- Polimorfismos de nucleótidos únicos y mutaciones

Los polimorfismos de nucleótidos únicos o SNPs (por su sigla en inglés, *single nucleotide polymorphism*), son marcadores genéticos que nos permiten entender la estructura de las poblaciones a nivel genético y su relación con ciertas enfermedades. Es un cambio puntual en el ADN donde una base es reemplazada por otra. Se considera SNP cambios en una base o algunos pocos pares de bases incluso hasta 50 pares de bases y algunos pequeños inserciones y deleciones (**Ver capítulo 3 para más detalles**). La frecuencia alélica del SNP debe ser igual o mayor a 1% para ser considerado como polimorfismo de nucleótido único. Muchos autores utilizan el término SNV (*Single nucleotide polymorphism*) para referirse a los SNPs obviando la frecuencia alélica (incluyendo a los SNPs y las mutaciones). Aunque por sí solos no nos dicen cuáles genes específicos están involucrados, sí nos indican una localización cromosómica que puede estar asociada con un fenotipo particular. Incluso si un SNP no está directamente relacionado con una enfermedad, puede servir como un marcador para identificarla. De ahí la diferencia de las mutaciones las cuales están directamente asociadas con un fenotipo generalmente patológico porque son cambios bruscos que se producen en el ADN y se presentan en una pequeña porción de la población con una frecuencia alélica menor a 1%.

1.3.2- Mutaciones, polimorfismos de nucleótidos únicos y Enfermedad de Parkinson

Recientemente, se han dedicado una gran cantidad de investigaciones para demostrar la etiología genética de la enfermedad de Parkinson (EP) y sus mecanismos moleculares. Aunque el 90% de esta enfermedad es multifactorial y está asociada a la combinación de componentes genéticos y ambientales, existen estudios que han descubierto la causa principalmente genética de algunas formas de EP hereditaria, conocida como EP familiar (Balestrino and Schapira 2020). Hasta la fecha se han descubierto varias mutaciones genéticas responsable de la enfermedad de parkinson análisis del genoma humano como la SNCA, LRRK2 (Lill 2016), (Olanow and Brundin 2013), (Nuytemans et al. 2010), MAPT, DJ1, PARK2, SCA1, SCA2 (Coppedè 2012). Ortega RA y colaboradores encontraron que, los pacientes con mutación en el gen GBA tiene una peor progresión de la enfermedad de parkinson en comparación con los con mutaciones en GBA/LRRK2 (Ortega et al. 2021).

SNCA

Se ha demostrado que las mutaciones en el gen SNCA aumentan la acumulación de la proteína alfa-sinucleína en forma de cuerpos de inclusión en células astrocitarias. Esto destaca la necesidad de investigar en profundidad el papel de esta proteína en enfermedades neurodegenerativas(Gu et al. 2010). Por ejemplo, la mutación A53T en la proteína alfa-sinucleína se encuentra asociada con una disminución en la entrada de glutamato dentro de las neuronas y la consiguiente muerte celular(Gu et al. 2010). Otros trabajos han identificado mutaciones en el gen SNCA en familias con enfermedad de Parkinson(Polymeropoulos et al. 1997). Hay que recalcar que hasta los pacientes con EP esporádico presentan la proteína alfa-sinucleína en el cuerpo de Lewy, una de las principales alteraciones neuropatológicas de la EP(Gómez-Benito et al. 2020). Otro caso relevante que evidencia aún más la asociación causal entre la SNCA y la EP es la presentada por Singleton y colaboradores en la cual encontró una triplicación de la SNCA en enfermos con la EP familiar de transmisión dominante autosómica(A. B. Singleton et al. 2003). Ahora queda explicar los mecanismos moleculares por los cuales la proteína alfa-sinucleína participa en el proceso patológico de la EP.

La alfa-Sinucleína es una proteína cuyo mecanismo molecular en la EP ha sido bien estudiado hasta el momento y representa un gran porcentaje en las causas de la EP(Du, Xie, and Liu 2020). La alfa-sinucleína tiene su mayor función en la formación de las vesículas intracelulares, además de la regulación de la concentración de las proteínas SNARE los cuales regulan la liberación de los neurotransmisores(Bendor, Logan, and Edwards 2013). La Enfermedad de Parkinson está relacionada con la proteína alfa-sinucleína por diferentes razones de las cuales se puede citar la acumulación o sobreexpresión de la SNCA. Por ejemplo, una triplicación del gen SNCA se encuentra asociado con la EP por aumentar la concentración de la proteína alfa-sinucleína dentro de las células(A. B. Singleton et al. 2003). Las mutaciones a nivel del gen SNCA de variantes A53T y A30P, los cuales inhiben la degradación de la alfa-sinucleína(A. B. Singleton et al. 2003; Cuervo et al. 2004), están asociados con la Enfermedad de Parkinson(Jo et al. 2002). Otro postulación tiene que ver con el hecho de que algunas mutaciones en la proteína alfa-sinucleína inducen un mal plegamiento de la proteína lo que engendra una disminución de la biodisponibilidad de las vesicular a nivel celular y consecuentemente una disminución en la liberación de dopamina y un aumento en la concentración intracelular de dopamina(Luth et al. 2014; Burré, Sharma, and Südhof 2015). Esta última es un neurotransmisor que tiene la propiedad de auto-oxidarse cuando se encuentra libre en el citoplasma dando como resultado producción de grandes cantidades de radicales libres como radical superóxido, peróxido de hidrógeno y aminocromo(Adams 2021). Esta

superproducción de radicales libres intracelulares provocan alteraciones de procesos biológicos y moleculares importantes por la inestabilidad genética que produce (mutaciones puntuales, de inserción, de delección y de translocación), peroxidación de los lípidos e inestabilidad de las membranas biológicas y mal plegamiento de las proteínas intracelulares (Berman and Hastings 1997), (Picca et al. 2021). Además se ha descrito también que alteraciones de la alfa-sinucleína provocan disfunciones mitocondriales, alteración en la producción energética y muerte celular por apoptosis con liberación de sustancias proinflamatorias en el medio extracelular (Tansey et al. 2022). Los productos liberados serán encargados de activar la respuesta inflamatoria por parte de la microglia que responde activando células inflamatorias como macrófagos y linfocitos para la reabsorción de las estructuras dañadas (Sanchez-Guajardo et al. 2013). La expresión aumentada de la alfa-sinucleína puede producir la secreción de la TNF en los pacientes con EP con relación a los controles además de incitar las células T-helper y citotóxicas a liberar citocinas como IFN γ , IL-2, IL-5 (Sanchez-Guajardo et al. 2013; Czirr and Wyss-Coray 2012). Pareciera que este proceso de respuesta inflamatoria por la microglia es responsable en parte de la neurodegeneración progresiva tanto por ser mantenida en el tiempo como por dañar a células íntegras en pacientes con EP (W.-W. Chen, Zhang, and Huang 2016), (Kwon and Koh 2020). La disfunción mitocondrial se produce dependiente de alteraciones a nivel del complejo mitocondrial I, lo cual afecta el potencial de membrana mitocondrial, alteración de la homeostasis del calcio y liberación del citocromo C (Luth et al. 2014). Otro mecanismo por el cual la alfa-sinucleína provoca alteración mitocondrial es a través de su unión al receptor TOM20, impidiendo de esta forma la unión de este último a su co-receptor TOM22 provocando alteración en el transporte mitocondrial de proteínas, lo que produce insuficiencia mitocondrial e incremento en la concentración de los radicales libres y alteración del potencial de membrana (Di Maio et al. 2016).

LRRK2

La quinasa 2 rica en leucina repetida (Leucine-rich repeat kinase 2), LRRK2 por su sigla en inglés también llamada PARK8, es una proteína enzimática de 280 KD de peso molecular y está formado por varios dominios de uniones que intervienen en la interacción proteína-proteína tal como dominio N-terminal de armadillo, dominio C-terminal WD40, anquirina y dominio rica en leucina repetida (Myasnikov et al. 2021). La LRRK2 interviene en diferentes procesos moleculares como la regulación de la autofagia, en el transporte de componentes citoplasmáticos hacia los lisosomas para ser degradados y reciclados (Bonet-Ponce and Cookson 2021). Además, se ha postulado que LRRK2 desempeña múltiples funciones, que van desde la remodelación del citoesqueleto hasta la expresión de proteínas, pasando por la transmisión sináptica y el tráfico de membranas,

mediante la fosforilación de diversos sustratos(Kiral et al. 2018), (Kiral et al. 2018; Yun et al. 2015), (Parisiadou et al. 2009). La LRRK2 fosforila un subgrupo de proteínas perteneciente a la familia RAB y regula sus capacidades para unirse con sus respectivas proteínas efectoras(Alessi and Sammler 2018). Las mutaciones en el gen que codifica la quinasa 2 rica en leucina (LRRK2) son factores de riesgo genético comunes para la enfermedad de Parkinson (EP) familiar y esporádica(Jeong and Lee 2020). Se sabe que casi todas las mutaciones a nivel del LRRK2 asociados a EP se encuentran en el dominio catalítico de la enzima. Por ejemplo, la mutación G2019S está asociada con un aumento en la actividad catalítica de la enzima(Jeong and Lee 2020) el cual se ha asociado con una forma de EP familiar de transmisión dominante y varias otras formas de EP esporádico(Zimprich et al. 2004). Se ha demostrado que la EP idiopática puede estar asociada con aumento de la actividad de la LRRK2 sin que haya necesariamente mutación en la proteína, lo que aumenta la fosforilación de la proteína Rab10 y las consecuencias derivadas incluyendo la disfunción mitocondrial y lisosomal(Di Maio et al. 2018). Varios papers describen como en cultivos neuronales primarios, la sobreexpresión de LRRK2 G2019S y otros variantes como I2020T, R1441C o Y1699C inducen toxicidad neuronal, evidenciada por el acortamiento de las neuritas, la muerte celular o el deterioro de las funciones de los orgánulos intracelulares(B. D. Lee et al. 2010), (B. D. Lee et al. 2010; West et al. 2007), (Smith et al. 2006). Se postula que el estrés del retículo endoplásmico es una alteración neuropatológica de la enfermedad de Parkinson(J. H. Lee et al. 2019) y se ha demostrado que la mutación G2019S del LRRK2 está asociado con aumento en las proteínas de estrés del retículo endoplásmico de los astrocitos y muerte celular(J. H. Lee et al. 2019). Además, Mamais y colaboradores demostraron que esta mutación G2019S de la LRRK2 produce una secuestación de la proteína Rab8 en los lisosomas y aumento de depósito de hierro en células microglia intrastratial(Mamais et al. 2021).

En resumen, la EP está relacionado con varias mutaciones y/o variantes genéticas a nivel de la proteína enzimática LRRK2, principalmente la variante G2019S, el cual aumenta la actividad catalítica de la enzima, aumentando de esta forma la fosforilación de la proteína Rab10 y la consecuente disfunción mitocondrial y muerte celular dopaminérgico en pacientes con EP.

GBA

La β -glucocerebrosidasa es una enzima con actividad de glucosilceramidasa que se necesita para escindir, por hidrólisis, el enlace beta-glucosídico del glucocerebrósido, un intermediario en el metabolismo de los glicolípidos que abunda en las membranas celulares(Vielhaber et al. 2001). Se localiza en el lisosoma, donde permanece asociado a la

membrana lisosomal(Vielhaber et al. 2001; Rijnboutt et al. 1991). Se ha demostrado la asociación entre la EP y mutaciones encontradas localizadas en la enzima GBA(G. Liu et al. 2016). Se ha estimado que las mutaciones en GBA confieren un riesgo de entre 20 y 30 veces mayor de desarrollar EP, y que al menos el 7-10% de los pacientes con EP tienen una mutación en GBA(Migdalska-Richards and Schapira 2016). Hasta la fecha, el mecanismo por el cual la GBA provoca la EP no está bien claro, pero varios autores postulan que las mutaciones en la GBA producen plegamiento defectuoso de la misma y afectación del buen funcionamiento y estrés del retículo endoplásmico(RE) por la acumulación de las proteínas mal plegada a nivel del RE(Migdalska-Richards and Schapira 2016). Esto a su vez produce la activación del proceso de respuesta a proteínas mal plegadas, que va desde el proceso de degradación por parte del sistema ubiquitina-proteasoma hasta la muerte celular por apoptosis. Li y colaboradores demostraron que la mutación L444P heterocigota del GBA produce disfuncionamiento mitocondrial e inhibición de la mitofagia, proceso por el cual se eliminan las mitocondrias disfuncionales(H. Li et al. 2019). Este cúmulo de mitocondrias disfuncionales dentro de las células provocan estrés oxidativo y muerte celular(H. Li et al. 2019).

Estrés oxidativo y enfermedad de Parkinson

Otros grupos de proteínas cuya mutación provoca enfermedad de Parkinson son las responsables en la detoxificación celular de los radicales libres producidos(Berman and Hastings 1997). Las especies reactivas de oxígeno son un grupo de moléculas incluyendo O_2^- , radicales hidroxilo $HO^\cdot$, y peróxido de hidrógeno. Las mitocondrias constituyen los componentes células que producen mayormente radicales libres dentro de la célula.

Las reacciones de oxido reducción son indispensables para la supervivencia celular por participar en procesos de señalización y procesos tisulares como el mantenimiento del tono muscular vascular, control de la ventilación pulmonar y la producción de eritropoyetina a través del monitoreo de la presión de oxígeno(Dröge 2002). Pero el equilibrio entre productos oxidados y reducidos y mantenimiento de las concentraciones de los radicales libres dentro de límites adecuados es de suma importancia para un buen funcionamiento de los procesos celulares. Para esto existen mecanismos establecidos dentro de la célula para mantener este equilibrio. Para entender bien los mecanismos celulares del mantenimiento del equilibrio oxido-reducción, vamos a presentar brevemente un modelo de reacción de producción y de eliminación de radicales libres a nivel de las mitocondrias. Todo puede explicarse empezando con la NADPH oxidasa que oxida una molécula de oxígeno en el anión superóxido en presencia de la xantina oxidasa(Valko et al. 2007), (Belarbi et al. 2017). El anión superóxido se convierte en peróxido de hidrógeno por la superóxido dismutasa(SOD)(Lubos, Loscalzo, and Handy 2011; Y. Wang et al. 2018), el cual a su vez

se reduce en agua por la Glutación peroxidasa(Lubos, Loscalzo, and Handy 2011). La glutatión peroxidasa en esta reacción oxida al glutatión el cual vuelve a reducirse por la glutatión reductasa en presencia de NADPH. A partir de la explicación anterior se podrá deducir que el sistema enzimático constituido por la glutatión peroxidasa y la glutatión reductasa participan en el mantenimiento del equilibrio de las reacciones de oxidorreducción dentro de la célula. También es de suma importancia recalcar que cualquier alteración en el sistema enzimático arriba mencionado llevará consecuentemente a un desequilibrio del sistema oxidación-reducción celular a favor del aumento de productos oxidativos llamados comúnmente radicales libres. En la enfermedad de Parkinson, se han encontrado alteración tanto cuantitativa como cualitativa en este sistema enzimático de protección celular contra los radicales libres. Diferentes estudios han demostrado una disminución considerable en la expresión de la enzima Glutathione peroxidase en modelos de enfermedades de Parkinson tanto murinos como en humanos(Trépanier et al. 1996),(Power and Blumbergs 2009) . La enzima glutatión transferasa, otra enzima que participa en el proceso de desintoxicación celular, se encontró relacionado con EP por una mutación tipo sustitución G/A producida en la posición 313 en presencia de una mutación A53T del PARK1(Golbe et al. 2007).

Hasta ahí, nosotros veníamos explicando brevemente los mecanismos moleculares tanto genético, inflamatorio, desequilibrio óxido-reducción como explicativos de la fisiopatología del EP, pero todavía nos falta por exponer los fenómenos cerebrales que siguen a estas alteraciones moleculares que también constituyen la base de la aparición de la sintomatología en el EP. Para entender estos fenómenos, sería de gran importancia presentar un recordatorio de los ganglios de la base en su estructura anatómica, su función y explicar cómo su alteración provoca la enfermedad de Parkinson.

1.4- Ganglios de la Base

Los ganglios de la base también llamados núcleos de la base son agrupaciones de cuerpos celulares en forma de núcleos situados en la profundidad de la sustancia blanca de los hemisferios cerebrales. Están constituidos por el núcleo caudado a ambos lados de los ventrículos laterales hacia dentro de la cápsula interna. Un poco fuera de la última se encuentra el putamen y el globo pálido con sus dos zonas importantes(GPi y GPe). Además forman parte del núcleo de la base, estructuras como la sustancia nigra y el núcleo subtalámico. El núcleo caudado y el putamen forman el striatum el cual se conecta con la sustancia negra por una red de fibras nerviosas dopaminérgicas.

Los núcleos basales son responsables de la coordinación de los movimientos y participan en el control cognitivo del movimiento. Permiten el control de la amplitud y de la expresión

de los movimientos. Para llevar a cabo estas funciones, los núcleos basales establecen una red de comunicación compleja con la corteza cerebral de la cual reciben impulsos excitatorios y al cual mandan a través del tálamo respuesta específica tanto excitatorio como inhibitorio según el caso después de integrar la información recibida. Todos los movimientos que se vayan a realizar, la corteza cerebral al mismo tiempo manda impulsos excitatorios hacia los ganglios de la base a través de fibras glutamatérgicas. Un grupo de fibras llamado cortico-caudal salen desde las áreas somatosensoriales, premotora y corteza suplementaria hacia el núcleo caudado. Desde allí se dirigen hacia el Globo pálido interno, el cual se encargará de mandar respuesta de regreso hacia el tálamo y de este último hacia la corteza prefrontal y a las áreas corticales somatosensoriales y suplementarias. Otros grupos de neuronas salen desde las mismas áreas corticales precitadas y se dirigen hacia el putamen. Desde el putamen se van produciendo dos vías de conducción una vía directa en la cual las fibras van desde el putamen hacia el globo pálido interno(GPi) y desde este último hacia el tálamo y hasta la corteza cerebral; y otra vía indirecta que antes de llegar al globo pálido interno(GPi) establece conexiones en el globo pálido externo(GPe), desde allí va al núcleo subtalámico y al final en el globo pálido interno(GPi) y desde este último se envían conexiones hacia el tálamo y a la corteza cerebral(Ver figura 2). Es de suma importancia recalcar, la vía directa activa los movimientos mientras que la vía indirecta inhibe los movimientos. A través de los receptores D1 del striatum, la sustancia negra a través de sus fibras dopaminérgicas activa la vía directa mientras que inhibe la vía indirecta por los receptores D2. De esta forma, la sustancia nigroestriatal participa en la ejecución de los movimientos, razones por las cuales en la enfermedad de Parkinson donde haya una deficiencia de las células dopaminérgicas nigroestriatales se producen alteraciones motoras de tipo bradicinesia, rigidez muscular y alteraciones posturales.

2A

2B

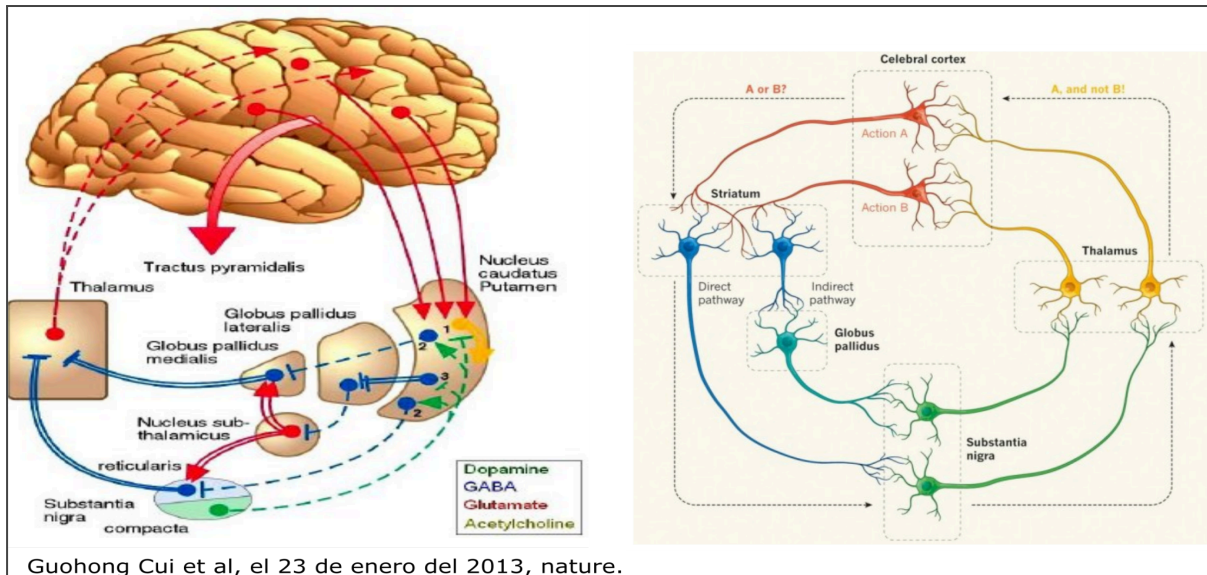


Figura 2. Vías directa e indirecta de los ganglios basales. **En la figura A** se presentan las conexiones nerviosas de los ganglios de la base. En esta imagen se describen las estructuras relevantes que conforman las vías directas e indirectas: sustancia nigra, globo pálido, putamen, núcleo subtalámico, núcleo caudado. Las fibras dopaminérgicas de la sustancia nigra pars compacta (SNpc) se proyectan hacia el cuerpo estriado. Las vías empiezan en esta estructura. **En la vía directa**, las fibras se proyectan desde el cuerpo estriado hacia el segmento interno del globo pálido. Desde ahí, salen proyecciones inhibitorias (GABAérgicas) hacia el tálamo y desde este último se proyectan hacia la corteza cerebral. En cuanto a la **vía indirecta**, las fibras se proyectan desde el cuerpo estriado hacia el segmento externo del globo pálido. Las fibras inhibitorias salen desde el globo pálido externo hacia el núcleo subtalámico y desde ahí regresa hacia el globo pálido interno.

Guohong Cui et al., el 23 de Enero del 2013, Nature.

En resumen, la enfermedad de Parkinson es de origen multifactorial donde participan diferentes factores tanto ambientales como genéticos. Factores como edad mayor a 60 años, la diabetes mellitus, exposición a metales pesados y pesticidas, traumatismos craneoencefálicos (TCE), productos lácteos incrementan el riesgo de padecer la EP mientras que el tabaquismo, el alcohol, ejercicio físico, cafeína, y Anti-inflamatorios no esteroideos (AINES) constituyen factores protectores contra la EP. En adición a estos, varios estudios demostraron la participación de una carga genética en el desarrollo de la EP, algunos que producen la EP familiar y esporádico y otros factores genéticos de tipo SNVs que en interacciones entre muchos de ellos (interacción gene-gene) o entre ellos y factores ambientales (interacción genes-ambiental). Los genes como GBA, LRRK2 y SNCA son los principales genes que se asocian con la EP esporádico y/o familiar como lo han demostrado diversos estudios. Pero otros genes como ATP13A2, VPS35, DJ-1 también tienen sus participaciones en el desarrollo de la enfermedad. Estas variantes genéticas provocan alteraciones en varios procesos celulares como disfunción mitocondrial, estrés del ER, estrés oxidativo y vías de señalización que en conjunto con los factores ambientales y de estilo de vida de los pacientes provocan muerte celular en regiones de ganglio basales.

específicamente de las células dopaminérgicas situadas en el pars compacta de la sustancia nigra.

1.5- Tratamientos actuales de la Enfermedad de Parkinson

En el momento actual, existen varias opciones de tratamientos con la EP, algunos más eficientes que otros o con más efectos adversos que otros. El tratamiento de la EP se divide en farmacológico y no farmacológico. Actualmente, las drogas dopaminérgicas son consideradas como las primeras líneas de tratamiento farmacológico contra la EP. Se ha propuesto una combinación entre las últimas y otras drogas como los inhibidores de la amino oxidasa tipo B, los inhibidores de catecol-o-metiltransferasa y anticolinérgicos para efectivizar el tratamiento de la EP sobre todo para controlar los síntomas motores y las complicaciones motoras inducidas por la Levodopa. También, los síntomas disautonómicos, insomnio, alteraciones psiquiátricas acompañados a la EP se están tomando en consideración para incrementar la calidad de vida de los pacientes. Otras medidas no farmacológicas son consideradas como efectivas contra la EP como la estimulación cerebral profunda (DBS por su sigla en inglés), terapia génica, terapia de reemplazo celular, terapia molecular dirigida entre otros.

Hasta la fecha, no existe un tratamiento específico efectivo que revierte la enfermedad de Parkinson. Los medicamentos que existen pueden en algunos casos disminuir los síntomas de la EP, pero muchos de los pacientes con estos tratamientos, a veces no responden a los mismos y desarrollan la forma severa de la EP o bien los mismos efectos adversos de los medicamentos contribuyen a una mala calidad de vida de los pacientes. De ahí, se hace indispensable seguir investigando sobre los mecanismos moleculares para poder encontrar blanco terapéutico y métodos de diagnóstico precoz de la EP de forma a prevenir la EP o tratar y prevenir las complicaciones de la misma. En este punto, es necesario recordar que la EP está considerada como un problema de salud pública que afecta a más de 10 millones de habitantes en el mundo. La EP afecta a más del 1% de la población mayor de 60 años y 5% a mayores de 80 años (Tysnes and Storstein 2017). La incidencia de la EP es de más de 19 por cada 100.000 habitantes por año (Twelves, Perkins, and Counsell 2003b). Estos datos refuerzan más la importancia de enfocarse en investigaciones cuyo objetivo es descubrir tratamientos efectivos contra la EP y las complicaciones de la misma.

Es importante enfocarse en la búsqueda de tratamiento y medios efectivos de diagnósticos precoz de la EP y sobre todo la predicción de las complicaciones de la misma, pero hay que recalcar también que existe subgrupos de pacientes caracterizados por la presencia de algún sintomatología que hace que la EP tiene un peor desenlace. Tal es el caso de los pacientes que presentan hipotensión ortostática (HO). Varios trabajos convergen en la idea

de que la OH está asociada con mayor mortalidad en los pacientes que tienen EP(Masaki et al. 1998). Un meta-análisis encontró que más del 30% de los pacientes con EP presentan NOH y disfunción autonómica(Velseboer et al. 2011).

En el siguiente capítulo nos detendremos a explicar detalladamente la hipotensión ortostática en los pacientes con enfermedad de Parkinson y su vínculo con los diferentes factores etiopatogénicos conocidos hasta la fecha como causantes de la EP. Sobre todo investigaremos sobre los antecedentes en la literatura que podrían vincular algunos de los factores etiológicos de la EP con el desarrollo de la NOH.

Capítulo II- Hipotensión ortostática neurogénica

2.1- Sistema autónomo

El fisiólogo y farmacólogo británico John Newport Langley fue el primero en usar el término “sistema nervioso autónomo” para designar a la parte del sistema nervioso central responsable de las funciones vegetativas, inconscientes e involuntarias de los órganos internos(Velseboer et al. 2011; Langley 1901). Después, se complementa esta idea con la propuesta de Walter B. Cannon y Henry Dale quienes dividieron este sistema en dos grandes componentes: el sistema nervioso simpático y parasimpático(Karemaker 2017).

El sistema nervioso autónomo, además últimamente se considera que tiene otro componente que es el entérico(Sternini 1997), (Lake and Heuckeroth 2013). Los dos mayores componentes del sistema nervioso vegetativo, a saber el simpático y el parasimpático, forman circuitos constituidos de fibras nerviosas tanto aferentes y eferentes que conectan el sistema nervioso central con los órganos internos y viscerales del cuerpo a través de impulsos nerviosos de tipo sensitivo y motor según si las fibras son aferentes o eferentes(Karemaker 2017). Además, este circuito se forma de tal manera que se conectan las fibras con cuerpos neuronales localizados en el SNC llamadas neuronas preganglionares y otras conectadas a cuerpos neuronales ganglionares periféricos llamadas neuronas postganglionares. El componente entérico por su parte está constituido por una extensa red de neuronas conectadas localmente entre sí y son responsables del funcionamiento y regulación de los procesos digestivos(Sternini 1997), (Lake and Heuckeroth 2013).

El SNS inerva a todos los órganos viscerales internos produciendo por ejemplo aumento de la frecuencia cardíaca y de la presión arterial, activa la glucogenolisis, disminuye el peristaltismo y procesos intestinales, broncodilatación, estimula la secreción de sudor, disminuye la excreción de orina(Kenney and Ganta 2014)(ver figura 3).

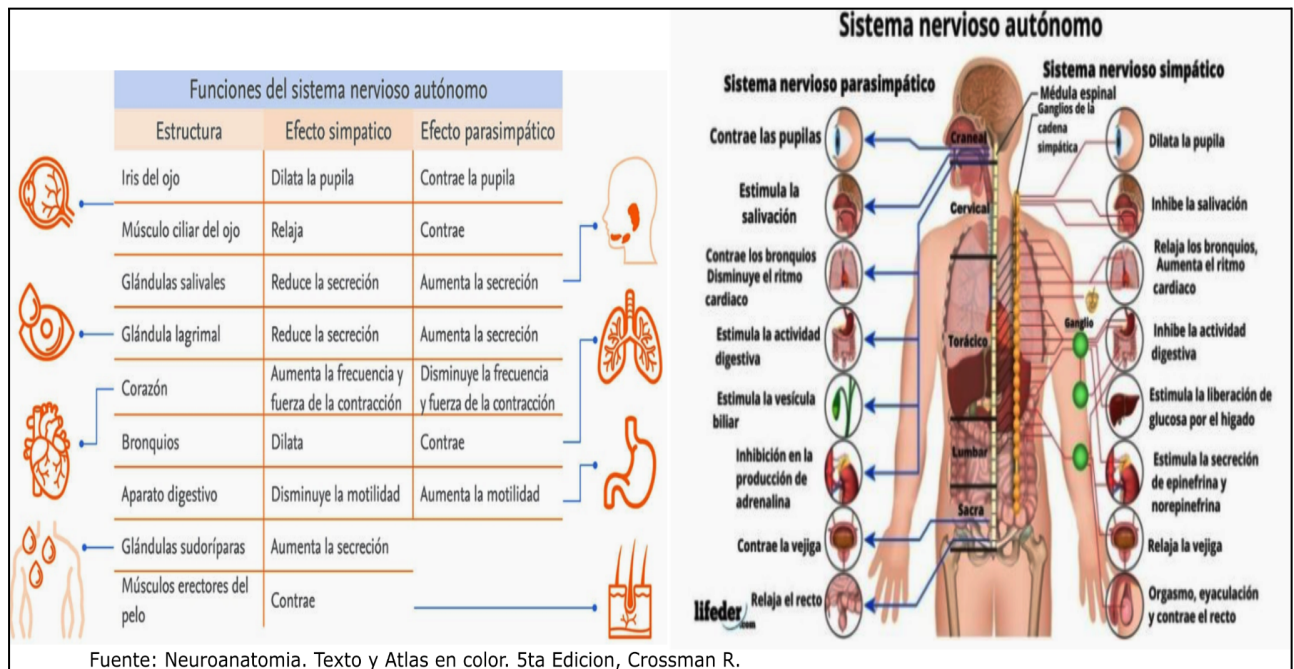


Figura 3- Funciones del sistema autónomo. Neuroanatomía. Texto y Atlas en color. 5ta Edición, Crossman R. En esta imagen se presentan las funciones del sistema nervioso autónomo. En tabla de izquierda, se presenta en una primera fila los órganos que corresponden a los efectos del sistema vegetativo simpático, en la segunda columna las estructuras específicas sobre las cuales actúan el sistema autónomo, en la tercera columna se describen los efectos que tiene el sistema autónomo simpático sobre las estructuras y en la cuarta columna los efectos del sistema vegetativo parasimpático sobre las estructuras de la columna 5. En la figura de la derecha se presentan las mismas informaciones antes descritos usando una imagen del ser humano señalando las partes y estructuras y los efectos de los sistemas simpáticos(a la izquierda de la imagen) y del sistema parasimpático(a la derecha de la imagen).

El sistema nervioso simpático es el encargado de preparar al cuerpo para responder a situaciones de estrés. Este mecanismo activa órganos y funciones corporales, como el aumento del ritmo cardíaco, la dilatación de los bronquios y la liberación de adrenalina, para permitir una respuesta rápida y efectiva ante situaciones de peligro o estrés. Este proceso es fundamental para la supervivencia y adaptación del organismo a diferentes desafíos. La fibra preganglionar del SNS se conecta con los receptores nicotínicos los cuales son activados por los neurotransmisores acetilcolina mientras que las fibras postganglionares activan o inhiben los receptores de los órganos periféricos a través de la adrenalina o la noradrenalina, las cuales se unen a receptores adrenérgicos alfa y/o beta. Se conocen 2 tipos de receptores adrenérgicos involucrados en los mecanismos de acción del sistema simpático: receptores alfa y beta. Son proteínas transmembranales asociadas a proteínas G los cuales activan o inhiben el segundo mensajero según el resultado buscado. Estos receptores se expresan tanto en el sistema nervioso central como en los órganos periféricos. Los receptores alfa a su vez se componen de alfa1 y alfa2. La alfa1 se acopla a las proteínas $G_{q\alpha}$, las cuales cuando se activan estimulan la fosfolipasa C. Este último es

una hidrolasa que hidroliza enlaces fosfodiéster del fosfatidilinositol bifosfato cuyos productos son segundos mensajeros (inositol trifosfato IP₃ y diacilglicerol DAG) que participan en procesos de señalización celulares como la neurotransmisión. El IP₃ como es soluble se difunde a través del citoplasma hacia el retículo endoplásmico donde promueve la liberación de calcio para la contracción muscular. Por esta razón los receptores alfa 1 se encuentran en tejidos donde se producen contracción musculares cuando el sistema simpático se activa como por ejemplo en vasos sanguíneos produciendo vasoconstricción. Por el otro lado, los receptores alfa 2 se acoplan a proteínas G_{iα} cuya activación está asociada con disminución en la concentración del AMPc intracelular, lo que conlleva a una inhibición de las enzimas PKA y la apertura de canales de K⁺, promoción del intercambio de Na⁺/K⁺ y estimulación de la fosfolipasa Cβ₂ que moviliza el Ácido Araquidónico y aumenta Ca⁺⁺ intracelular. Estos receptores producen vasoconstricción, disminución de la motilidad del músculo liso del tracto gastrointestinal. Los receptores beta por su parte se dividen en 3 subgrupos: beta 1, 2 y 3. La activación de los 3 tipos de receptores betas producen elevada concentración de AMPc intracelular. Los Beta 1 son más abundantes en el tejido cardíaco provocando efectos metabotrópicos y cronotrópicos en dicho tejido. Sus principales acciones se centran en el aumento de la frecuencia cardíaca y del gasto cardíaco y el subsiguiente aumento de la presión arterial por el incremento de la fracción de eyección producida por el gasto cardíaco aumentado. Mientras tanto los beta 2 están distribuidos mayoritariamente en músculos esqueléticos y lisos provocando dilatación de los mismos a través de la activación de la AMPc y la consiguiente activación de la PKA y captura de calcio en el retículo endoplásmico. Finalmente los subtipos beta 3 están más relacionados con procesos metabólicos como por ejemplo la activación de la lipólisis a nivel del tejido adiposo.

El sistema nervioso parasimpático, por su lado produce todo lo contrario al SNS en cuanto a los procesos biológicos que acabamos de mencionar como la disminución de la frecuencia cardíaca, vasodilatación. Las neuronas preganglionares del sistema nervioso parasimpático se conectan a los cuerpos neuronales del ganglio a través de receptores colinérgicos nicotínicos como lo hace también el SNS. Contrariamente a este último, el sistema nervioso parasimpático utiliza receptores colinérgicos muscarínicos ([ver figura 4](#)) para conectarse con los órganos periféricos. Los receptores muscarínicos están ampliamente localizados en el sistema nervioso central y se conoce que están asociados con procesos de aprendizaje y de memoria. También se encuentran distribuidos en los tejidos periféricos como el corazón, tejidos glandulares, fibras musculares lisas de vasos sanguíneos y digestiva (Botella Romero et al. 2009). Se conocen 5 subtipos: M₁, M₂, M₃, M₄, M₅. Los receptores M₁, M₃, M₅ se acoplan a proteínas G_{q/11} cuya activación provoca la estimulación de la PLC y la

subsiguiente activación del IP3 y DAG mientras que los M2, M4 se asocian a proteínas Gi/o que inhiben la enzima adenilciclase y activar canales de K⁺ e inhibir canales de calcio sensibles a voltajes. El subtipo M1 se encuentra especialmente distribuido en la corteza cerebral e hipocampo y cuerpo estriado. El subtipo M2 abunda más en el tejido cardíaco, músculo liso produciendo efectos inotrópicos, cronotrópicos y dromotrópicos negativos.

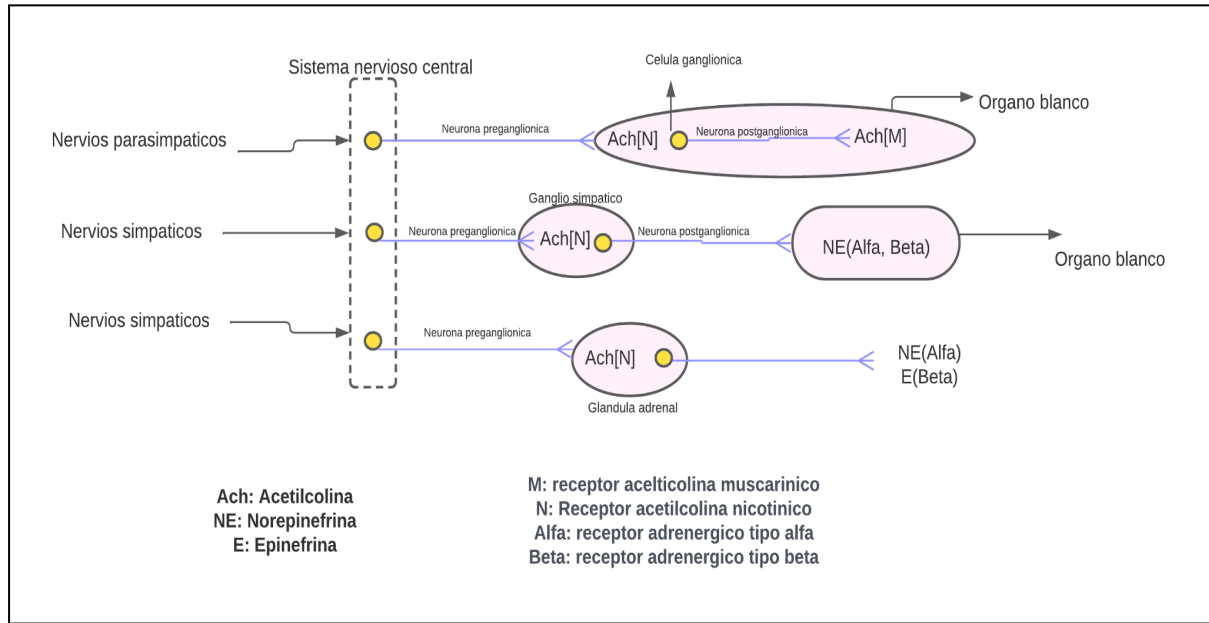


Figura 4- Sistema nervioso autónomo y los receptores involucrados. En el caso del parasimpático, el cuerpo neuronal de la fibra preganglionar está localizado en los núcleos ambiguos y motor dorsal del vago. De ahí salen los axones hacia los ganglios donde se libera acetilcolina responsable de la activación de los receptores nicotínicos de la neurona postganglionar. Las neuronas del ganglio mandan sus fibras hacia los órganos periféricos efectores para activar los receptores muscarínicos de los mismos. En el caso del sistema simpático, la primera neurona con su cuerpo localizado en la médula espinal manda sus fibras hacia el ganglio donde libera acetilcolina que activan los receptores nicotínicos de las neuronas postganglionares los cuales a su vez liberan noradrenalina y/o adrenalina hacia los órganos periféricos (por ejemplo el nodo sinusal y auriculoventricular del corazón para la activación de los receptores B1 y alfa de los vasos sanguíneos). Realizado en nuestro laboratorio basándose en la figura del sistema nervioso autonómico de la Enciclopedia Británica. Encyclopædia Britannica. 20 de abril de 2022. <https://www.britannica.com/science/autonomic-nervous-system#/media/1/45079/1154>

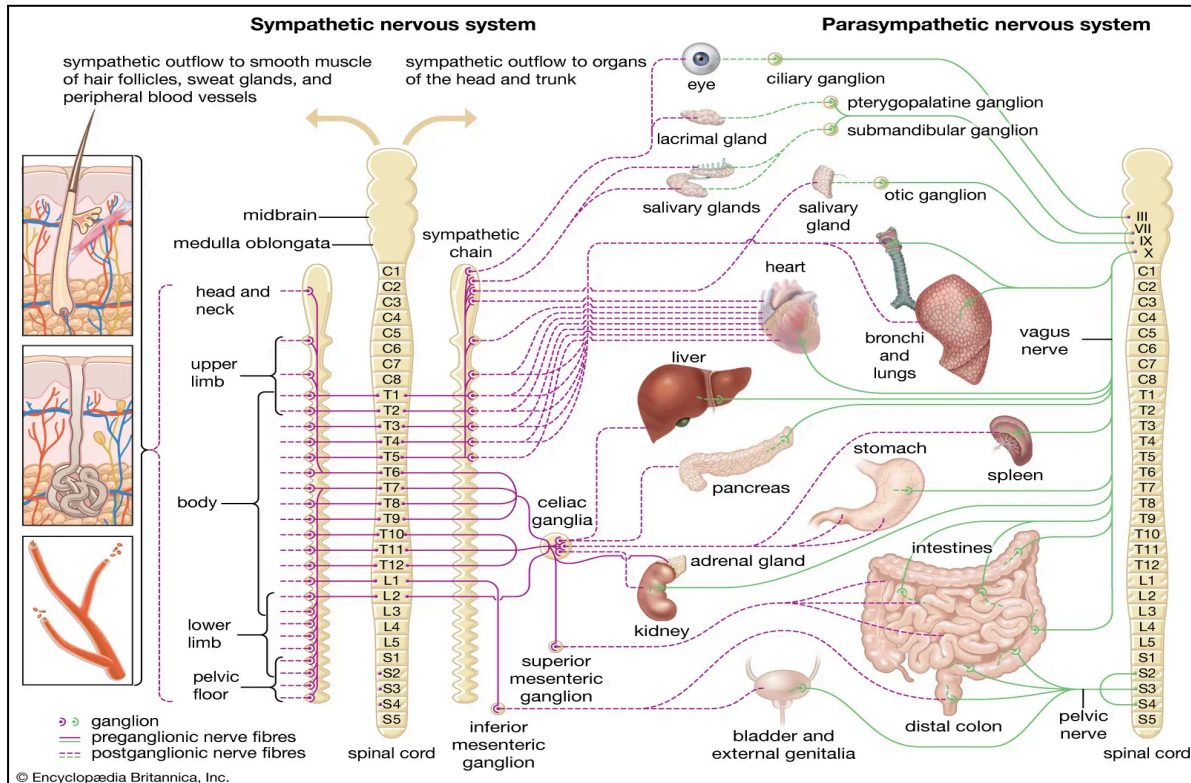


Figura 5- Inervación y función del sistema nervioso autónomo Encyclopædia Britannica. 20 de abril de 2022. <https://www.britannica.com/science/autonomic-nervous-system#/media/1/45079/1154>. En esta figura se presenta las inervaciones de los órganos por el sistema vegetativo. Se describen las diferentes regiones de la médula espinal de donde salen los nervios. Se dividen en 5 regiones: región cervical, región torácica, lumbar y sacra.

Hasta aquí hemos descrito someramente el sistema autónomo con sus respectivos componentes y sobre todo los receptores que permiten la comunicación o conexión entre el SNC autónomo y los órganos efectores. Ahora vamos a hablar un poco del sistema de regulación central del sistema vegetativo. Se sabe hace tiempos remotos que la vía aferente que toman estímulos desde los barorreceptores al cerebro, llega hasta una región dorsomedial de la médula oblongada llamada núcleo del tracto solitario (Seller and Illert 1969). Además, otros estudios han demostrado que las neuronas preganglionares se unen a cuerpos neuronales en la región ventrolateral rostral (Ross, Ruggiero, Joh, et al. 1984), (Ross, Ruggiero, Joh, et al. 1984; Ross, Ruggiero, Park, et al. 1984). Lo que hizo pensar que las vías simpáticas y parasimpáticas están integradas a estructuras más complejas en el SNC. Dichas estructuras fueron estudiadas en detalles últimamente a partir del desarrollo de la neuroimagen funcional. Involucran estructuras y regiones del cerebro como el tronco cerebral, el hipotálamo, el sistema límbico y la corteza cerebral, el cerebelo. Las regiones del cerebro anterior están involucradas directa o indirectamente en el control de las funciones autónomas, las cuales incluyen a la corteza insular, la corteza cingulada anterior, la amígdala y el hipotálamo. La corteza insular es una región de integración de sensaciones de dolor, funciones viscerales y temperatura (Craig 2003). Además se le atribuye funciones en el control de las actividades del sistema autónomo a través de su conexión con el

hipotálamo(Westerhaus and Loewy 2001). El área cingulada anterior está formada por amplias conexiones con la ínsula, la corteza prefrontal, la amígdala, el hipotálamo y el tronco cerebral, a través de las cuales controla las funciones simpáticas y parasimpáticas(Terreberry and Neafsey 1987).

Por ejemplo se pudo comprobar a través de MRI funcional que la parte dorsal se encuentra activada cuando los individuos realizan tareas que involucran la atención y el comportamiento y está asociado con la estimulación del sistema simpático(Critchley 2009).

Se ha encontrado en diversos estudios asociación entre el cíngulo anterior y cerebro anterior con la respuesta simpática mediada por barorreceptores(Kimmerly et al. 2005). La amígdala por su parte se involucra en la regulación del sistema autónomo a través de las conexiones establecidas por el núcleo central de la amígdala con el núcleo paraventricular y área lateral del hipotálamo, con regiones del mesencéfalo como la formación reticular medular ventromedial, rostral ventrolateral, núcleo periacueductal(Loewy 1982). El hipotálamo permite la integración entre el sistema endocrino y el autónomo en cuanto a la respuesta a estímulos para el mantenimiento de la homeostasis y la adaptación del individuo. La médula ventrolateral rostral (MVLr) contiene un grupo de neuronas adrenérgicas responsables de la presión arterial(Dampney et al. 2003),(Guyenet 2006) mientras que las neuronas glutamatérgicas del MVLr controlan el gasto cardiaco y el flujo sanguíneo hacia los órganos viscerales y músculos esqueléticos. Las neuronas simpáticas en su función excitatorias reciben estímulo de diferentes regiones como el tronco encefálico, y del cerebro anterior incluyendo a señales inhibitorias de las neuronas sensibles a barorreceptores del núcleo del tracto solitario(NTS)(Guyenet 2006). También el MVLr recibe señales desde el hipotálamo. Además, la MVLr recibe señales inhibitorias de neuronas GABAérgicas desde las neuronas barosensibles del núcleo del tracto solitario inhibiendo de esta forma el componente simpático del barorreflejo(Dampney et al. 2003).

En resumen, sabiendo que el sistema autónomo está formado por 3 componentes cada uno de los componentes conectados con el SNC a través de su primera fibra preganglionar, y con los órganos efectores con sus fibras postganglionares. Cada uno de estos componentes tienen funciones específicas en los órganos dependiendo del tejido y de los receptores que contienen. Además, estudiamos que las fibras nerviosas simpáticas y parasimpáticas forman redes o conectan con áreas más altas del sistema nervioso como el hipotálamo, el área prefrontal, el mesencéfalo etc. En el siguiente apartado nos dedicaremos a estudiar la hipotensión ortostática neurogénica.

2.2- Consideraciones generales de la Hipotensión Ortostática neurogénica

Se puede dar el caso que este sistema autónomo sufre alguna alteración en uno o algunos de sus componentes tanto los núcleos, los ganglios, las fibras nerviosas que lo constituyen a nivel periférico y/o central dando lugar a alteraciones en el mecanismo de regulación o disfunción del sistema nervioso autónomo. El proceso resultante se denomina disfunción autonómica o disautonomía. Esta última está constituida de un conjunto de síntomas y signos que lo caracterizan. Los síntomas van apareciendo según los nervios o regiones del sistema nervioso afectados. Se descartan síntomas como disfunción sexual; a nivel intestinal existe alteración de la motilidad del estómago y del intestino que se manifiestan por náuseas, vómitos, diarrea, molestias abdominales difusas y estreñimiento. La hipotensión ortostática es una de las principales manifestaciones clínicas de la disfunción autonómica. La hipotensión ortostática es una bajada brusca de la presión arterial 3-5 mn después de pasar del estado de decúbito a posición de pie y puede estar acompañado de la sensación de mareo, síncope, por disminución del flujo sanguíneo al encéfalo, todo de forma transitoria. Cuando la hipotensión ortostática está producida por una disminución en la liberación de norepinefrina a nivel de los nervios simpáticos postganglionares o alteración de estructuras vegetativas centrales, se le denomina hipotensión ortostática neurogénica (NOH)(Kalra, Raina, and Sohal 2020). Las causas etiopatogenia de la NOH se dividen en dos grandes grupos, los NOH de causas primarias y los secundarias. Los de causa primaria incluyen a un grupo de patologías neurodegenerativa caracterizadas por depósito de alfa-sinucleína en el sistema nervioso central y periférica incluyendo a la EP, PAF(Pure autonomic failure), MAS(Multiple atrophy system) y demencia con cuerpo de Lewy. Estas 4 condiciones se denominan sinucleinopatías y tienen en común la sintomatología NOH(Kalra, Raina, and Sohal 2020). Las NOH de causa secundaria son las provocadas por patologías que destruyen los nervios periféricos como la DM y la neuropatía diabética, neuropatía inducida por fármacos, neoplasias de la fosa posterior,iringomielia, amiloidosis, inflamatoria y autoinmunitaria(Kalra, Raina, and Sohal 2020). Los principales síntomas encontrados en los pacientes con NOH son el mareo y/o vértigo, cefalea, visión borrosa, síncope, alteración cognitiva y en raras ocasiones pueden presentar hasta convulsiones(Kalra, Raina, and Sohal 2020; Low et al. 1995). Pueden aparecer síntomas por hipoperfusión como cefalea suboccipital, dolor de hombro, claudicación de miembros, y dolor precordial por hipoperfusión del músculo cardiaco. Además, los pacientes con NOH también presentan otros síntomas como disnea ortostática por hipoperfusión del ápice pulmonar y hasta oliguria por hipoperfusión renal.

2.3- Fisiopatología de la hipotensión ortostática neurogénica y diagnóstico .

La hipotensión ortostática tiene una prevalencia de 6% aproximadamente en la población general, lo cual se aumenta exponencialmente con la edad de los individuos hasta estabilizarse entre 10 y 30 % incluso en pacientes que no tienen neurodegeneración (Joseph et al. 2017). Los factores de riesgo asociados a la hipotensión ortostática son los medicamentos como diuréticos, antidepresivos, hipotensores; polifarmacia, tabaquismo, diabetes mellitus y la amiloidosis (Poon and Braun 2005). Aunque paradójicamente, en otros trabajos, han encontrado que los pacientes con tratamiento con diuréticos, al añadir fluoxetina la bajada de la presión arterial está mejorada en pacientes con enfermedad de Parkinson y OH (Montastruc et al. 1998). La hipotensión ortostática en algunos pacientes constituyen en la mayoría de las veces el inicio de las alteraciones neurodegenerativas por sinucleinopatías en enfermos de Parkinson.

La hipotensión ortostática neurogénica es la hipotensión ortostática que se produce en los pacientes como consecuencia de la alteración de la vasoconstricción simpática debida a una liberación disminuida de norepinefrina por parte de las neuronas postganglionares (Idiaquez et al. 2021), (Idiaquez et al. 2021; Robertson 2008). Además, se debe diferenciar de los demás hipotensión ortostática que se produce por otra etiología como hemos explicado en párrafo anterior. La hipotensión ortostática se define como una reducción sostenida de la presión arterial sistólica de al menos 20 mm Hg o una reducción de la presión arterial diastólica de al menos 10 mm Hg, generalmente en los primeros 3-5 minutos de estar de pie o de inclinarse hacia arriba en una mesa basculante (Freeman et al. 2011).

La capacidad de los seres humanos para superar los efectos iniciales del estrés gravitacional sobre el sistema cardiovascular depende en gran medida de un sistema nervioso autónomo muy desarrollado que incluye un reflejo barorreceptor exquisitamente sensible para mantener niveles adecuados de presión sanguínea. Una clara comprensión de los mecanismos implicados en el mantenimiento de la presión arterial al ponerse de pie es clave para descifrar las causas y tratamiento de la hipotensión ortostática. La postura erguida provoca una rápida acumulación de sangre en las extremidades inferiores, con el consiguiente movimiento de agua plasmática hacia el intersticio y la microcirculación circundante (Jacob et al. 1998). Debido al efecto de la gravedad, en pocos segundos entre 500 y 700 mL de sangre se trasladan desde la parte superior del cuerpo a los vasos de capacitancia venosa de las extremidades inferiores y circulación esplácnica (Christopher J. Mathias 2002). De los cuales, 500 ml se acumulan en las extremidades y 200 aproximadamente en la región esplácnica y pélvica. Además, debido al aumento de la presión hidrostática, aproximadamente el 10% del plasma intravascular se desplaza hacia el

compartimento intersticial en un plazo de 10 a 30 minutos de la postura erguida (Christopher J. Mathias 2002; Hagan, Diaz, and Horvath 1978). Al acumularse la volemia en la parte inferior, se produce una disminución del retorno venoso, de la presión de llenado del ventrículo derecho y, en última instancia el gasto cardíaco en aproximadamente un 20%, de ahí la consiguiente caída de la presión sanguínea que puede comprometer la perfusión de los órganos por encima del corazón en ausencia de mecanismos de compensación. Los barorreceptores son los responsables de sensor el aumento o la bajada de presión de la volemia.

Los barorreceptores son terminaciones nerviosas mecanosensibles situadas en la pared de los vasos, que detectan pequeñas reducciones de la presión arterial o del volumen sanguíneo central, a través de deformaciones de la pared de los vasos. Los barorreceptores arteriales o de alta presión se localizan en los senos carotídeos y el arco aórtico y desempeñan el papel más importante en la adaptación aguda a la postura erguida (Jacobsen et al. 1993), (J. S. Sanders, Ferguson, and Mark 1988). Los barorreceptores de baja presión (o volorreceptores) se localizan principalmente en los principales vasos cardiopulmonares (J. S. Sanders, Ferguson, and Mark 1988) detectan una reducción del volumen sanguíneo central (J. S. Sanders, Ferguson, and Mark 1988). Estos mecanorreceptores están conectados con el sistema nervioso autónomo central a través de las vías aferentes y eferentes. Las vías aferentes en el seno carotídeo del corazón y los principales vasos cardiopulmonares transmiten información al cerebro a través de los nervios vagos y glosos-faríngeos. Existen numerosas conexiones cerebrales. Las vías eferentes comprenden la salida del impulso nervioso desde regiones simpáticas y parasimpáticas centrales hacia los vasos sanguíneos y el corazón. Estas vías ejercen un control sobre el latido cardíaco y de la presión arterial, siendo un componente vital la vasoconstricción inducida por el aumento de la actividad simpática (Ichinose and Nishiyasu 2012). La disminución de la presión venosa central al ponerse de pie, reduce el tráfico barorreflejo aferente al tronco cerebral al descargar los barorreceptores cardiopulmonares, lo que provoca un aumento del tono simpático (Persson 1996). Las Vías aferentes [nervios glosos-faríngeos (IX) y vagos (X) en el caso de los barorreceptores carotídeos y aórticos, respectivamente, y fibras nerviosas vagales no mielinizadas para los

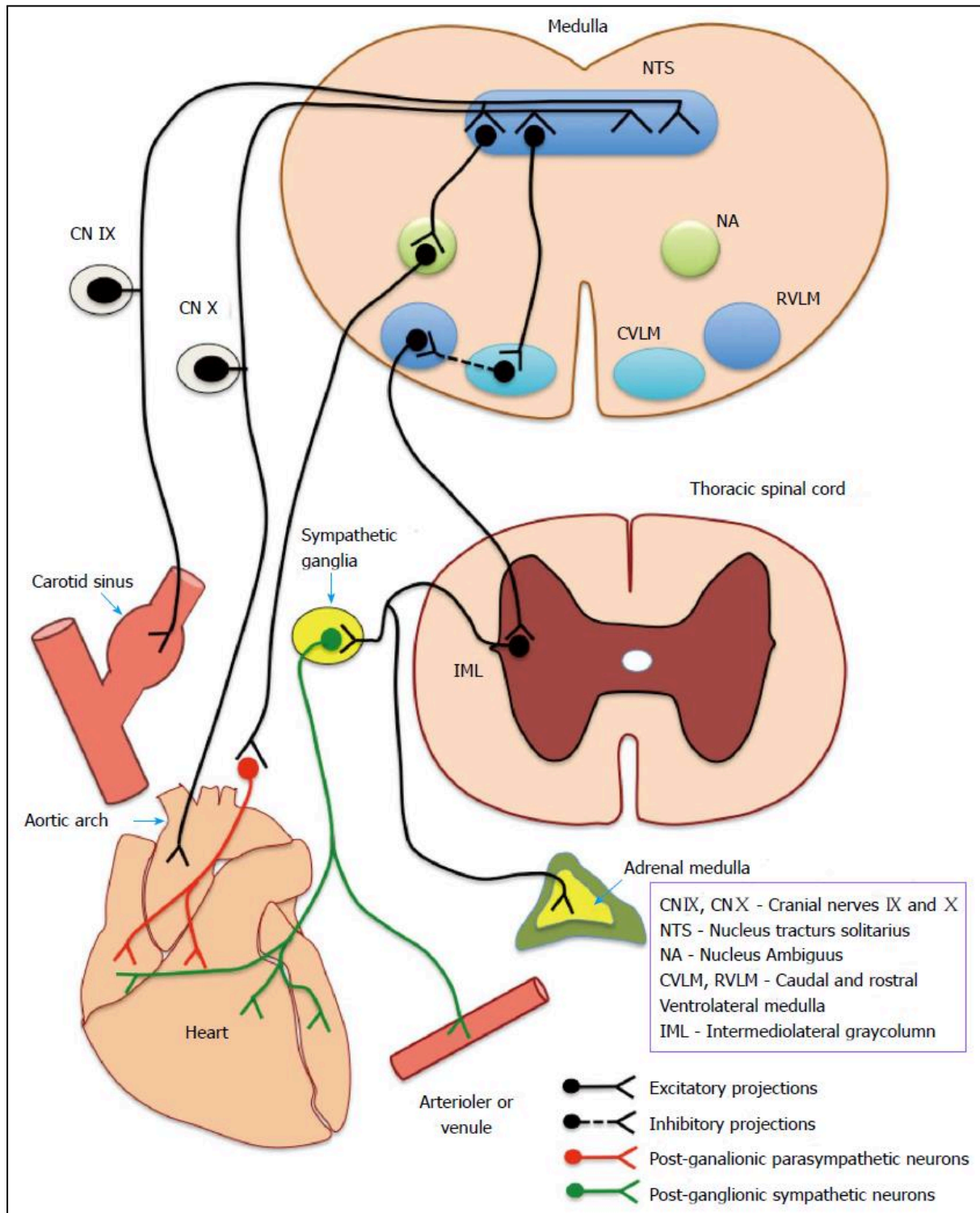


Figura 6: Circuito neuronal del arco barorreflejo. Representa el arco barorreflejo completo: partiendo de los barorreceptores (situados en el seno carotídeo y en el arco aórtico), los aferentes (IX y X nervio craneal) ascienden a los centros medulares y envían eferentes (simpáticos y parasimpáticos) a los órganos finales (corazón y vasculatura). CN IX y CN X: Nervio craneal IX y X; NTS: Núcleo del tracto solitario; NA: Núcleo ambiguo; CVLM: Médula ventrolateral caudal; RVLM: Médula ventrolateral rostral; IML: Columna gris intermedial (Kaur et al. 2016)

barorreceptores de baja presión, hacen sinapsis en el sistema nervioso central, especialmente en el núcleo del tracto solitario, el cual se activa por la liberación de

glutamato(Ohta, Li, and Talman 1996), (Ohta, Li, and Talman 1996; Andresen et al. 2001) . Los circuitos neuronales centrales como la médula ventrolateral caudal y la médula ventrolateral rostral(bulbo raquídeo), integran los impulsos nerviosos que controlan la salida simpática hacia las arterias y las venas mientras que el corazón a través de la columna de la médula espinal torácica, y por fin el núcleo ambiguo que controla la salida parasimpática al nodo sinoauricular del corazón(Andresen et al. 2001). Un descenso de la presión arterial provoca una disminución de la actividad de los nervios aferentes hacia el tracto solitario con la consiguiente inhibición de la región ventrolateral del bulbo raquídeo que libera poco señal inhibitorio(GABA) hacia la región ventrolateral rostral, el cual a su vez responde aumentando su señal activadora(glutamato) hacia los cuerpos neuronales preganglionares en las astas de la médula espinal lo que hará que se produzca mayor liberación de acetilcolina hacia los cuerpos neuronales de la cadena laterovertebral simpática y este último libera noradrenalina, lo que provoca un aumento del impulso simpático hacia los órganos efectores como vasos sanguíneos y corazón produciendo vasoconstricción y aumento de la frecuencia cardiaca, del volumen minuto y del gasto cardiaco y consecuentemente aumento de la presión arterial. Por otro lado, a nivel del sistema parasimpático, se produce una disminución del impulso nervioso hacia los cuerpos neuronales de los núcleos ambiguos y dorsal del vago lo que induce inhibición de la actividad parasimpática(Joseph et al. 2017), la activación del sistema renina-angiotensina-aldosterona (RAAS) (Persson 1996; Egan et al. 1983) y liberación de arginina vasopresina (AVP) (Taylor et al. 1995), (Jacob et al. 1998) y el consiguiente aumento de la presión arterial.

Cuando existe alteración a nivel del sistema nervioso autónomo, estos mecanismos son alterados y no se puede recuperar la presión arterial después de ponerse de pie.

2.4- Hipotensión ortostática neurogénica y enfermedad de Parkinson

La hipotensión ortostática neurogénica siendo una de las principales causas de discapacidad en los pacientes con EP, es una condición no motora acompañada de otros síntomas como mareos, aturdimiento, debilidad y fatiga. Los pacientes tienen tendencia a la caída y trastornos en las actividades diarias, llevándolos a tener una mala calidad de vida(Sin and Khemani 2020). La presencia de la hipotensión ortostática neurogénica en los pacientes con EP está asociado con una alta tasa de mortalidad que podría ser explicado por la consiguiente hipoperfusión cerebral que trae consigo. Masaki y colaboradores encontraron una asociación lineal significativa entre el cambio de la presión arterial sistólica de la posición de cubito a la posición de pie y alta tasa de mortalidad de los pacientes a los

4 años de seguimiento (Masaki et al. 1998). Incluso sin la presencia de la EP, la NOH puede constituir una fuente de causa de debilidad en pacientes mayores, lo que acarrea riesgo alto de caída y alta tasa de mortalidad a los 3 años de empezado la NOH si lleva un mal control de la presión arterial ortostática (Shaw et al. 2019), (Veronese et al. 2015). Gibbons y colaboradores encontraron en un estudio publicado en el año 2015 que el 31% de los pacientes con OH retardada desarrolla alfa sinucleinopatía y la tasa de mortalidad de los pacientes a los diez años de seguimiento fueron de 29% con respecto a 9% en los controles (Gibbons and Freeman 2015).

La hipotensión ortostática neurogénica se produce como resultado de una liberación defectuosa de norepinefrina desde los terminales simpáticos al ponerse de pie. Los estudios neuropatológicos también han demostrado la existencia de una patología de cuerpos de Lewy en el locus coeruleus, la principal fuente de noradrenalina en el cerebro. Los niveles de norepinefrina en el locus coeruleus se reducen en los pacientes EP con demencia en comparación con los pacientes EP sin ella (McDonald, Newton, and Burn 2016). Norcliffe y colaboradores encontraron asociación entre una alteración de la activación simpática y la presencia del NOH en pacientes con EP y demencia con cuerpo de Lewy. (Norcliffe-Kaufmann et al. 2018)

La hipotensión ortostática neurogénica es una de las condiciones más frecuentes que aumenta la severidad y el mal pronóstico de la Enfermedad de Parkinson como lo han demostrado muchos trabajos. Complicaciones asociadas a alteraciones cognitivas, motoras, y muerte son más frecuentes en los pacientes con EP que además tienen NOH.

Se sabe que la NOH constituye una causa del deterioro cognitivo en los pacientes con EP sobre todo cuando la NOH está asociada con la sinucleinopatía en dichos pacientes (McDonald, Newton, and Burn 2016), (McDonald, Newton, and Burn 2016; Bocti et al. 2017). No está del todo claro porque la hipotensión ortostática neurogénica provoca alteración cognitiva de los pacientes. Una de las hipótesis sobre la causa de la aparición del deterioro cognitivo en pacientes con NOH y EP, se relaciona con las alfa sinucleinopatías, en las cuales la alfa-sinucleína afectaría regiones cerebrales específicas relacionadas con la cognición y funciones ejecutivas como la corteza frontal (Kang et al. 2022), (McDonald, Newton, and Burn 2016). Algunas mutaciones en la proteína SNCA por ejemplo la E46K provocan depósitos de cuerpo de Lewy agresivamente en casi todo el sistema nervioso autónomo central y periférico produciendo condiciones no-motoras por la degeneración de fibras nerviosas (Zarranz et al. 2004). Zarranz y colaboradores reportaron un caso de un japonés que lleva la alteración en el gen SNCA p.G51D que desarrolló la EP tempranamente, acompañada de deterioro precoz de la cognición y trastornos psiquiátricos (Zarranz et al. 2004; Tokutake et al. 2014). Por el otro lado, España y

colaboradores refutaron la idea según la cual la sobreexpresión cortical de la alfa sinucleína solo provocaría disfunción cognitiva en animales de laboratorio. Encontraron que la sobreexpresión de la proteína inducida en la corteza frontal de ratones por vectores virales de expresión, necesita además inoculaciones de fibrillas preformadas de alfa sinucleína para producir deterioro cognitivo en los modelos de animales (Espa et al. 2019). Varias investigaciones anteriores demostraron que la sobreexpresión de la proteína SNCA es suficiente para producir EP porque esta proteína en concentraciones altas dentro de la célula es tóxica para la misma. Altas concentraciones de SNCA celular produce radicales libres intracelulares además de depósitos de inclusiones de proteínas intracelulares, lo que conlleva a la muerte celular por apoptosis. La sobreexpresión de la α -sinucleína aumenta la producción de ROS y la susceptibilidad de las células al estrés oxidativo (Winklhofer and Haass 2010), (Junn and Mouradian 2002). Las células dopaminérgicas derivadas de iPSCs de pacientes con triplicación del SNCA tienen altas concentraciones de productos de estrés oxidativo y una mayor susceptibilidad al H_2O_2 , lo que sugiere que una sobreexpresión de la SNCA está asociada con un aumento en la producción de ROS en dichas células. Las alfanucleinas acumuladas dentro de la membrana mitocondrial inhiben la actividad del complejo I lo que provoca una disfunción mitocondrial (Chinta et al. 2010) además de la capacidad de los oligoméricos de las alfasinucleinas de activar la microglia conllevando a una producción elevada de mediadores inflamatorios TNF-alfa, IL-1B e IL-6 (C. Kim et al. 2013). La transferencia de neurona a neurona y la propagación de la α -sinucleína desencadena más agregaciones y cascadas citotóxicas (Luk et al. 2012). Con estos datos, está claro que la sobreexpresión de la SNCA está asociada con toxicidad celular y apoptosis celular.

Pero, por otro lado, según Espa y sus colaboradores, pareciera que no es suficiente la sobreexpresión de la proteína SNCA cuando se trata de deterioro cognitivo. Lo que sí está ampliamente aceptado es que existen mutaciones específicas a lo largo del gen SNCA que producen afectaciones cognitivas como hemos mencionado anteriormente.

Otra hipótesis que parece ser también válida para explicar la aparición de alteraciones cognitivas en pacientes con NOH, se refiere a las consecuencias de los episodios de hipotensiones los cuales producen hipoperfusión cerebral y daños isquémicos y anóxicos, lo que provocará deterioro de la función cognitiva (McDonald, Newton, and Burn 2016). Modelos de Enfermedad de Parkinson en ratones presentan disminución del número de neuronas en el hipocampo y sustancia blanca evaluado por inmunohistoquímica además de activación de la microglia y concentraciones elevadas de IL-1B e IFN- γ en un estudio realizado por Gao y sus colaboradores (Y. Gao et al. 2019). Estos hallazgos pueden ser las causas de la aparición de deterioro cognitivo en los pacientes con hipoperfusión causado

por la NOH justamente por las lesiones a nivel del hipocampo y de la sustancia blanca. Abe y colaboradores demostraron que una hipoperfusión cerebral de la región occipital y parietal posterior de los pacientes con EP está asociado con una alteración de la cognición visual sin deterioro motor (Abe et al. 2003). Los pacientes con EP tienen un desempeño cognitivo (evaluado por pruebas neuropsicológicas) más pobre con respecto a los controles y con un bajo rCBF (flujo sanguíneo cerebral regional relativo por su sigla en inglés) (medido por tomografía computarizada por emisión de fotón único-SPECT) en los lóbulos temporales derecho e izquierdo (Jagust et al. 1992). En tales casos, no solamente las regiones del hipotálamo y de la sustancia blanca juegan un papel preponderante en el deterioro cognitivo de los pacientes con NOH pero también las regiones temporales del cerebro.

Si bien las alfa sinucleinopatías podrían estar involucradas en la aparición de la NOH como lo han demostrado varios trabajos (Mendoza-Velásquez et al. 2019), (Isonaka et al. 2019), (Isonaka et al. 2019; Palma 2019), (Goldstein, Li, and Kopin 2001) pero está más que claro que la hipoperfusión cerebral constituye uno de los mecanismos fisiopatológicos por los cuales la hipotensión ortostática neurogénica contribuye a las complicaciones de la enfermedad de Parkinson. Dependiendo de las zonas del cerebro donde se producen las hipoperfusiones y las alteraciones resultantes van apareciendo las complicaciones como deterioro de la cognición (Goldstein, Li, and Kopin 2001; Aarsland et al. 2021) o alteración motora. También, la hipoperfusión cerebral participa ampliamente en el proceso de la neurodegeneración (Raz, Knoefel, and Bhaskar 2016).

Además, otros autores vinculan la hipoperfusión como la causa o mecanismo por el cual la hipotensión ortostática aumentaría la tasa de mortalidad en pacientes con EP. Aunque para otros no está del todo claro de qué manera la NOH produce el aumento de riesgo de muerte en estos sujetos.

Múltiples evidencias científicas demuestran que también la NOH está asociada con alteraciones a nivel periférico sobre deterioro del sistema autónomo cardíaco (Dickson et al. 2008), (Isonaka et al. 2019).

Utilizando, análisis inmunohistoquímicos, Orimo y colaboradores encontraron acumulación de agregados de alfa sinucleína en los axones distales del sistema nervioso simpático cardíaco. Esta acumulación patológica de alfa sinucleína estaría precediendo a la encontrada en los cuerpos neuronales localizados en los ganglios simpáticos paravertebrales lo que informaría sobre el sentido centripeta de la neurodegeneración simpática en la EP (Orimo, Uchiyama, Nakamura, Mori, Kakita, et al. 2008). Otros estudios apuntan a que, la hipotensión ortostática en edades tempranas puede ser un signo precoz que informa sobre la futura aparición de la EP (Palma et al. 2013), (Palma et al. 2013; Palma and Kaufmann 2014). Esto podría explicarse por el hecho de que algunos pacientes sin diagnóstico de EP en su vida,

pero que en estudios postmortem presentan lesiones neuronales caracterizadas por depósitos de inclusiones de cuerpo de Lewy que contienen alfa sinucleína en fascículos nerviosos epicárdicos (Dickson et al. 2008), (Palma and Kaufmann 2014). Además, Se evidencia a través de neuroimagen in vivo, depósitos de alfa sinucleína en nervios simpáticos noradrenérgicos miocárdicos lo que podría estar conllevando a una neurodegeneración simpática miocárdica asociada a la NOH en pacientes con EP (Isonaka et al. 2019).

Los pacientes con EP con fallo neurocirculatorio simpático con baja radioactividad derivada de la 6-[18F] fluoro-dopamina no asociada con la gravedad, ni con la duración de la enfermedad, ni con L-DOPA, demuestran tener una fracción de eyección cardíaca disminuida de [3H]norepinefrina y prácticamente no había derrame de norepinefrina cardíaca o incrementos venoso-arteriales en los niveles plasmáticos de DHPG y L-dopa en el cateterismo cardíaco derecho (Goldstein et al. 2000). Di Francisco y colaboradores estudiaron la respuesta de la frecuencia cardíaca (FC), la presión arterial (PA) y la concentración de norepinefrina (NE) en pacientes con EP y NOH tanto en reposo como después de ejercicio. Encontraron que después del ejercicio la FC, PA y NE fueron significativamente más bajas que los de los controles. (DiFrancisco-Donoghue et al. 2009) Estos datos sugieren que la pérdida de inervación de catecolaminas en la enfermedad de Parkinson se produce no solamente en el sistema nigroestriado del cerebro pero también en el sistema nervioso simpático en el corazón.

La vasoconstricción periférica constituye el principal mecanismo de regulación que mantiene la presión de perfusión adecuada durante la hipotensión ortostática. Este mecanismo de control no se lleva a cabo en los pacientes con EP y NOH por esta misma denervación catecolaminérgica del sistema cardiovascular (Convertino 2014).

2.5- Hipotensión ortostática neurogénica y polimorfismos genéticos

Varios trabajos anteriores demostraron que la NOH en general tiene su componente genético al poder asociarla con diferentes polimorfismo de nucleótido único (SNPs). El principal evento clínico encontrado en los pacientes que tienen deficiencia de la enzima Dopamina beta hidroxilasa (DBH), una patología de transmisión autonómica recesiva, es la hipotensión ortostática neurogénica causada por disfunción de la vía simpática (Garland and Biaggioni 2003). Shibao y colaboradores encontraron una variante patogénica en el gen que codifica una subunidad del receptor de la acetilcolina que es importante para la transmisión sináptica rápida en el ganglio autonómico. Uno de los principales eventos clínicos de esta alteración es la NOH y niveles de norepinefrina bajas (Shibao et al. 2021). Hace falta resaltar

que estos pacientes no padecían de Enfermedad de Parkinson, y que la NOH se presentó como síntomas clínicos relacionados con las alteraciones genéticas. Estos dos casos de variantes genéticas asociadas a NOH entre muchas otras que no se van a presentar aquí, nos permiten entender la participación y la naturaleza genética de la etiopatogenia de la hipotensión ortostática neurogénica sin que haya presencia de EP en los pacientes. Además, varios trabajos han demostrado la existencia de variantes genéticas asociadas a la EP, las cuales podrían estar asociadas a su vez con disfunción autonómica y la NOH presentes en dichos pacientes. Disfunción autonómica, caracterizada por OH y denervación cardíaca simpática en paciente con EP, ha demostrado estar asociada con triplicaciones en el gen SNCA (Orimo, Uchihara, Nakamura, Mori, Ikeuchi, et al. 2008), (Orimo, Uchihara, Nakamura, Mori, Ikeuchi, et al. 2008; A. Singleton et al. 2004). Variantes patogénicas del GBA asociadas al EP como la GBA-N370S y GBA-L444P, fueron encontradas asociadas con varios eventos clínicos no motores incluyendo la disfunción autonómica y la NOH (Brockmann et al. 2011).

A partir de estos datos, hemos podido comprobar la participación de las variantes genéticas no solamente en las disfunciones autonómicas cuyo evento principal es la NOH sino además en el desarrollo de la NOH asociado a la Enfermedad de Parkinson. Aun así, son muy pocos los datos que son disponibles en cuanto a este tema por la importancia que tiene en la enfermedad de Parkinson. La NOH y las disfunciones autonómicas son complicaciones muy importantes de la EP porque aumentan la tasa de mortalidad en dichos pacientes cuando están presentes. Como habíamos presentado en apartados anteriores, la NOH puede provocar hipoperfusión cerebral e hipoxia a nivel del cerebro y por consiguiente muerte neuronal con las complicaciones que podrían encarar según las zonas afectadas. Además, otros trabajos demostraron asociaciones entre la NOH y denervación cardíaca periférica y algunas patologías cardíacas. Por esto, se hace imperiosamente importante enfocarse en el tratamiento de la NOH para mejorar la calidad de vida de los pacientes y disminuir la tasa de mortalidad en estos grupos de pacientes. En el apartado siguiente vamos a presentar las características clínicas y complicaciones de la NOH en los pacientes con EP.

2.6. Complicaciones de la Hipotensión Ortostática neurogénica en EP

Los síntomas más frecuentes cuando un paciente con NOH se levanta rápidamente son mareos, aturdimiento y visión borrosa o disminuida. Los movimientos del brazo al ponerse de pie pueden empeorar estos síntomas debido a un fenómeno de robo de la subclavia, que reduce el flujo sanguíneo del tronco cerebral. El presíncope y el síncope se observan a

menudo en pacientes con MAS(atrofia multisistémica) y PAF(fallo autonómico puro) durante las fases iniciales de la enfermedad. A medida que los pacientes aprenden a reconocer los síntomas del presíncope, son capaces de evitar los desmayos adoptando una posición sentada o recostada(C. J. Mathias, Mallipedi, and Bleasdale-Barr 1999). El comienzo del síncope puede ser más insidioso en los pacientes de edad avanzada con EP, y la presencia de NOH puede ignorarse inicialmente. Un síntoma menos frecuente asociado a la hipoperfusión cerebral es un déficit cognitivo transitorio al ponerse de pie(C. J. Mathias, Mallipedi, and Bleasdale-Barr 1999; Poda et al. 2012).

Hasta aquí hemos estudiado las complicaciones de la HO en general y la NOH en los pacientes con EP en particular. Hemos presentado complicaciones relacionadas con eventos coronarios agudos, isquemia cerebral y hasta muerte. Estos datos hablan de cuán importante es el tratamiento de la NOH en los pacientes con EP para disminuir las complicaciones y muerte en pacientes con EP además de mejorar la calidad de vida de los mismos. En el apartado siguiente hablaremos de los tratamientos disponibles en la actualidad con la NOH en los pacientes con Enfermedad de Parkinson.

2.7- Tratamientos actuales disponibles contra la NOH en EP

El tratamiento actual contra la NOH en pacientes con EP es de tipo sintomático con el objetivo de revertir la hipotensión arterial. Muchos de ellos están asociados con hipertensión arterial y patologías cardíacas. Los fármacos simpaticomiméticos son los más utilizados contra la NOH: la midodrina es una agonista alfa adrenérgica y la droxidopa es un precursor de la noradrenalina, que se convertirá en esta última por la enzima dopa-descarboxilasa. Estos medicamentos no se pueden administrar en pacientes con patologías cardíacas, insuficiencia renal y retención urinaria por las complicaciones que pueden producir en estos pacientes. Otra estrategia que se usa en el tratamiento de la NOH es aumentar la volemia sanguínea para de esta forma poder aumentar la presión arterial y revertir la NOH. Para este objetivo se usa la fludrocortisona que es un mineralocorticoide que aumenta la retención de sodio y agua a nivel de los riñones. Este medicamento tampoco se puede administrar en los pacientes con insuficiencia cardíaca y renal. La administración de eritropoyetina acompañada de suplemento de hierro puede colaborar en aumentar los glóbulos rojos y la hemoglobina en dichos pacientes pero habrá que tener en cuenta los posibles riesgos de policitemia y trastornos de coagulación por hipercoagulabilidad de la sangre. Se usan otros medicamentos en el tratamiento de la NOH como por ejemplo la Yohimbina, piridostigmina, ergotamina, dihidroergotamina, efedrina, desmopresina, indometacina y fluoxetina, pero la seguridad y la eficacia de sus usos hasta la fecha no han

sido comprobado en estos casos. Estas opciones farmacológicas pueden considerarse en casos de OH refractaria, en caso que otros agentes resultaron ineficaces contra OH o que causaron efectos secundarios.

En resumen, es de recalcar que no existe un tratamiento efectivo con la NOH en pacientes con EP que podrían disminuir las complicaciones provocadas por la misma o mejorar la calidad de vida de los pacientes. Como acabamos de presentar anteriormente, la NOH en estos pacientes están asociados con altas tasas de comorbilidad cardíaca como fibrilación auricular, hipertrofia ventricular izquierda, eventos coronarios agudos y cerebrales como isquemia cerebral. Además, varios estudios encontraron asociación entre la NOH y alto riesgo de mortalidad en dichos pacientes.

Por todo lo anterior, pensamos que una estrategia que puede ayudar sería estudiar los mecanismos genéticos y moleculares asociados a la NOH en pacientes con EP, sabiendo que otros trabajos han encontrado la participación genética en la etiopatogenia de la NOH en los pacientes con EP. En el siguiente capítulo presentaremos una revisión de la genética humana y las variantes genéticas como SNPs además los métodos usados para sus estudios.

Capítulo III- Genética humana y Polimorfismo de nucleótido único.

3.1- Consideraciones sobre el genotipo y fenotipo

James Watson y Francis Crick en 1953 fueron los primeros en presentar sus hipótesis sobre la naturaleza de doble hélice del ADN. Cuando las técnicas analíticas mejoraron se observó que los ácidos nucleicos, incluyendo el DNA, estaban formados por cuatro bloques moleculares similares llamados nucleótidos. Hacia 1910, Phoebus A. Levene propuso la hipótesis del tetranucleótido para explicar la disposición química de los nucleótidos en los ácidos nucleicos. No fue hasta la década de 1940 en que el trabajo de Erwin Chargaff hizo comprender que la hipótesis de Levene era incorrecta.

El DNA es un ácido nucleico, y los nucleótidos son las piezas que forman todas las moléculas del ácido nucleico. Estas unidades estructurales, que se denominan mononucleótidos, constan de tres componentes esenciales: una base nitrogenada, un azúcar pentóxido (un azúcar con cinco carbonos) y un grupo fosfato. Las bases nitrogenadas son de dos tipos. Las dos purinas son la adenina y la guanina, abreviadas como A y G, respectivamente. Las tres pirimidinas son la citosina, la timina y el uracilo, que se abrevian como C, T y U, respectivamente. La unión entre dos mononucleótidos consiste en un grupo fosfato unido a dos azúcares. Se forma entonces un enlace fosfodiéster, ya que

el ácido fosfórico se une a dos alcoholes (grupos hidroxilo de los dos azúcares) por una unión éster en ambos lados.

4.1.1- Estructura del ADN

Watson y Crick publicaron su análisis sobre la estructura del DNA en 1953 y construyeron modelos según el cual el ADN tiene la forma de doble hélice, y que tiene las características principales siguientes:

- A. El ADN está formado de dos largas cadenas polinucleotídicas enrolladas alrededor de un eje central, formando una doble hélice enrollada hacia la derecha (dextrógira).
- B. Las dos cadenas son antiparalelas; es decir, la orientación C-5-C-3 va en direcciones opuestas.
- C. Las bases de las dos cadenas yacen formando estructuras planas y perpendiculares al eje; están «apiladas» unas sobre otras, separadas 3,4 Å (0,34 nm), y se encuentran en el interior de la estructura.
- D. Las bases nitrogenadas de las cadenas opuestas están apareadas como resultado de la formación de puentes de hidrógeno (que se describen a continuación); en el DNA sólo se permiten los emparejamientos A-T y G-C.
- E. Cada vuelta completa de la hélice tiene una longitud de 34 Å (3,4 nm); de este modo, cada vuelta de la cadena contiene 10 bases.

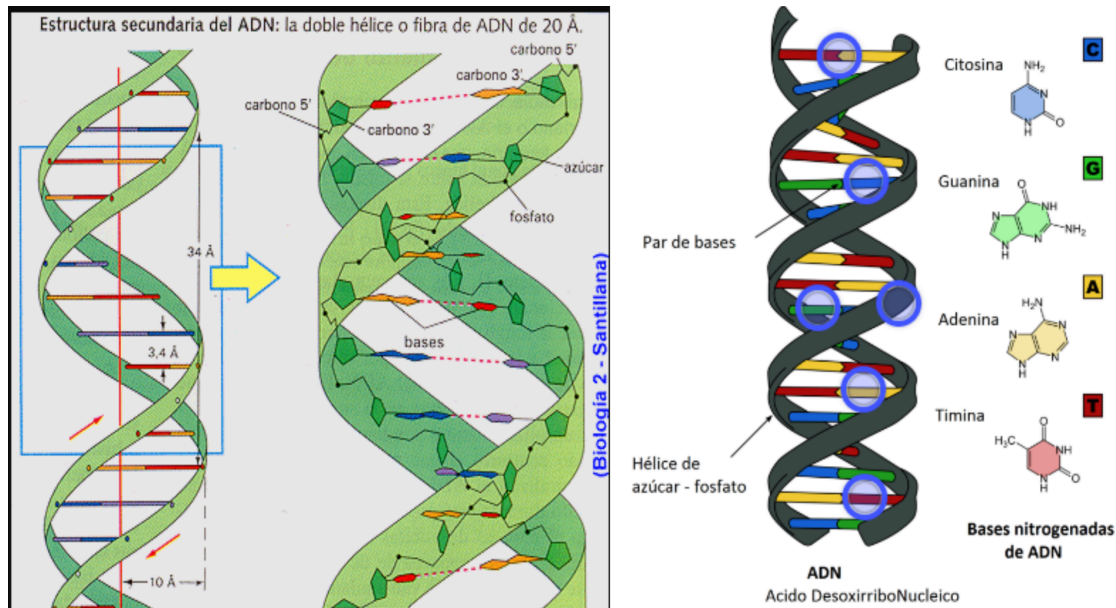


Figura 7- Estructura secundaria del ADN. En esta imagen se presenta la estructura secundaria del ADN. El ADN tiene una forma de hélice en el cual se apilan las bases en el eje central con las cadenas fosfato-azúcar hacia afuera de la estructura. Las bases se unen a través de puentes de hidrógeno (Dos en caso de A-T y 3 puentes de hidrógeno en caso de G-C). La estructura hélice del ADN hace que se forme el surco mayor y el surco menor.

La clave del modelo propuesto por Watson y Crick es la especificidad en el emparejamiento de bases. Los datos de Chargaff sugieren que las cantidades de A y T son iguales, y que las de G y C también. Watson y Crick se dieron cuenta de que si A se empareja con T y C lo hace con G, justificando estas proporciones, los componentes de estos pares de bases formarían puentes de hidrógeno, proporcionando la estabilidad química necesaria para mantener las dos cadenas juntas. Ordenadas de este modo, se hace aparente a lo largo del eje un surco mayor y uno menor. El emparejamiento específico entre las bases A-T y G-C es la base del concepto de complementariedad, que describe la afinidad química entre bases proporcionada por los puentes de hidrógeno. Como veremos, este concepto es muy importante en los procesos de replicación y de expresión génica del DNA. Antes de pasar al estudio de la expresión génica se hace importante explicar la importancia que tiene esta estructura del ADN el almacenamiento y la transmisión de la información genética. Las dos cadenas de ADN se enrollan de tal manera que las bases nitrogenadas que tienen las informaciones genéticas quedan dentro de la estructura mientras los grupos fosfatos quedan afuera interactúan con las moléculas de agua impidiendo de esta forma la conexión entre las bases nitrogenadas con el agua.

El ADN de la célula se organiza para dar formación a la cromatina. En la cromatina eucariótica, hay una cantidad substancial de proteína asociada al DNA cromosómico en todas las fases del ciclo celular. Las proteínas asociadas se clasifican en histonas cargadas positivamente (básicas), y en las no histonas, con una carga positiva menor. De las proteínas asociadas al DNA, las histonas desempeñan claramente la función estructural

más esencial. Las histonas contienen grandes cantidades de los aminoácidos lisina y arginina, que están cargados positivamente, lo que les posibilita unirse electrostáticamente a los grupos fosfato de los nucleótidos, cargados negativamente. Hay cinco tipos principales de histonas (Ver tabla 1). El modelo general de la estructura de la cromatina se basa en la suposición de que las fibras de cromatina, compuestas por DNA y proteína, experimentan un gran enrollamiento y plegamiento al condensarse dentro del núcleo celular. En este modelo, el núcleo de DNA de 147 pares de bases está enrollado alrededor de un octámero de histonas en una súper hélice levógira, completando 1,7 vueltas por nucleosoma.

Tipo de Histona	Contenido en Lisina y Arginina	Peso molecular en Da
H1	Rica en Lisina	23.000
H2A	Ligeramente rica en Lisina	14.000
H2B	Ligeramente rica en Lisina	13.000
H3	Rica en Arginina	15300
H4	Rica en Arginina	11300

Tabla 1- Contenido en lisina y arginina de las histonas del nucleosoma. En esta tabla se presentan una caracterización somera de las histonas, su clasificación según sus contenidos en lisina y arginina y además se presenta sus pesos moleculares.

3.1.1- Transcripción y expresión genética

Como vimos anteriormente, la estructura del DNA consiste en una secuencia lineal de desoxirribonucleótidos. En última instancia, esta secuencia dicta los componentes que forman las proteínas, el producto final de la mayor parte de los genes. Una cuestión central es cómo se puede decodificar en una proteína una información almacenada en forma de ácido nucleico. En el primer paso de la expresión génica, la información que se encuentra en una de las dos cadenas de DNA se transfiere a una secuencia complementaria de RNA mediante el proceso de transcripción. Una vez sintetizado, este RNA actúa de «molécula mensajera» que lleva la información codificada. De ahí su nombre, RNA mensajero (mRNA). Estos ARNm se asocian con los ribosomas, en los que se produce la decodificación en proteínas.

Es un proceso complejo realizado en múltiples pasos, y que empieza con la transferencia de la información genética almacenada en el ADN al RNA (Ver figura 9). El proceso por el que se sintetizan las moléculas de RNA sobre un molde de DNA se denomina transcripción. El resultado es una molécula de mRNA complementaria a la secuencia del gen de una de las dos cadenas de la doble hélice. La importancia del proceso de transcripción es enorme,

puesto que representa el paso inicial del proceso de flujo de información dentro de la célula..

La composición del ADN de los diferentes tipos de células de un organismo multicelular son esencialmente idénticas. La variación entre las células se debe a las diferencias en la expresión de los genes, principalmente a nivel de la transcripción: diferentes genes se transcriben en diferentes células según las necesidades de las mismas. Algunos genes conocidos como genes de mantenimiento, deben expresarse en casi todas las células, pero otros genes muestran una expresión génica específica para cada tejido o pueden expresarse en momentos específicos.

Normalmente, sólo una de las dos cadenas de ADN de un dúplex sirve de molde para la síntesis de ARN. Durante la transcripción, el ADN de doble cadena se une al ARN polimerasa. A continuación, el ADN se desenrolla, lo que permite que la cadena de ADN que actúa como molde para la síntesis de ARN (la cadena molde) formará un híbrido transitorio de ARN-ADN con la cadena de ARN en crecimiento. El ARN es complementario a la cadena molde de ADN con misma secuencia (Excepto que la U sustituye la T) y tiene la misma dirección 5'→3'. La hebra no molde suele llamarse hebra sentido, y la cadena molde suele llamarse cadena antisentido. Las ARN polimerasas sintetizan el ARN a partir de cuatro precursores de nucleótidos: ATP, CTP, GTP y UTP. La elongación implica la adición del residuo de monofosfato de ribonucleósido apropiado (AMP, CMP, GMP o UMP) al grupo hidroxilo 3' libre en el extremo 3' de la cadena de ARN en crecimiento. Estos nucleótidos se derivan de la división de un residuo de pirofosfato (PPi) de sus precursores ribonucleósidos trifosfato (rNTP). Sólo el nucleótido iniciador en el extremo 5' de un transcrito primario lleva un grupo trifosfato 5'.

El Splicing del ARN elimina las secuencias no deseadas del transcrito primario. Para la mayoría de los genes de los vertebrados -casi todos los genes que codifican proteínas- sólo una pequeña parte de la secuencia del gen se decodifica para dar el producto final. La transcripción de un gen produce inicialmente un ARN transcrito primario que es complementario a toda la longitud del gen, incluyendo tanto los exones como los intrones. Este transcrito primario se somete a continuación a un empalme (Splicing) de ARN, que es una serie de reacciones por las que los segmentos de ARN intrónico se eliminan y descartan, mientras que los segmentos de ARN exónico restantes se unen de extremo a extremo para dar lugar a un producto de ARN más corto (**Figura 8**). El empalme del ARN requiere el reconocimiento de las secuencias de nucleótidos en los límites de los exones e intrones transcritos (uniones de empalme). Los dinucleótidos en los extremos de los intrones están muy conservados: la gran mayoría de los intrones comienzan con una GT (que se convierte en GU en el ARN intrónico) y terminan con un AG (la regla GT-AG).

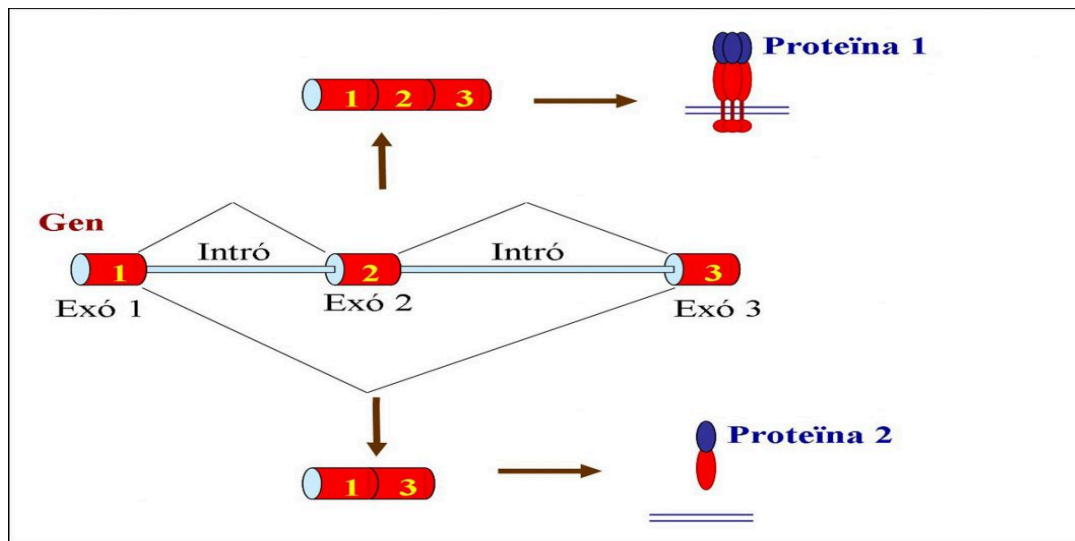


Figura 8- Proceso de corte y empalme de ARNm. En esta figura se presenta el proceso de splicing durante el proceso de maduración de los ARN mensajeros. En este proceso se eliminan los intrones para formar los ARNm maduros formados de exones (regiones codificantes del gen). Extraído de la página web: <https://ellipse.prbb.org/es/splicing-o-empalme>.

Aunque los dinucleótidos conservados GT y AG son cruciales para el empalme, no son suficientes para marcar los límites de un intrón. Las secuencias de nucleótidos que son inmediatamente adyacentes también están muy conservadas, constituyendo secuencias de consenso de empalme. Una tercera secuencia intrónica conservada que también es importante en el empalme se conoce como sitio de ramificación y suele estar situado a no más de 40 nucleótidos aguas arriba del AG terminal 5' del intrón. Otras secuencias exónicas e intrónicas pueden promover el splicing (secuencias potenciadoras del empalme) o inhibirlo (secuencias silenciadoras del empalme), y las mutaciones en estas secuencias pueden causar enfermedades.

Además del empalme del ARN, los extremos de los transcritos de la ARN sufren otras modificaciones: al extremo 5' se añade una variante de guanina mediante un enlace fosfodiéster inusual, y una larga secuencia de adeninas se añade al extremo 3'. Además de proteger los extremos de las exonucleasas celulares, estas modificaciones pueden contribuir al correcto funcionamiento de los transcritos de ARN.

El ARNm producido por los genes en el núcleo migra al citoplasma, donde se une a los ribosomas y otros componentes para iniciar la traducción y la síntesis de los polipéptidos. Sólo un segmento central de una molécula de ARNm eucariota se traduce para hacer un polipéptido. Las regiones no traducidas (la 5' UTR y la 3' UTR) se transcriben a partir de secuencias de exones presentes en los extremos 5' y 3' del gen. Estas regiones ayudan a unir y estabilizar el ARNm en los ribosomas y promueven una traducción eficaz.

El ensamblaje de un nuevo polipéptido a partir de los aminoácidos que lo componen se rige

por un código genético llamado codón. Dentro de un ARNm, la secuencia central de nucleótidos que se usa para fabricar el polipéptido se explora de 5' a 3' en el ribosoma en grupos de tres nucleótidos (codones). Cada codón especifica un aminoácido, y el proceso de decodificación utiliza una colección de diferentes moléculas de ARNt, cada una de las cuales se une a un tipo de aminoácido. Un complejo aminoácido-ARNt se conoce como aminoacil tRNA y se forma cuando una aminoacil tRNA sintetasa dedicada une covalentemente el aminoácido requerido a la adenosina terminal en el trinucleótido CCA conservado en el extremo 3' del ARNt.

Como se podría deducir, cualquier alteración tanto estructural como funcional en este proceso de transcripción y traducción en proteína de la información genética, puede acarrear a la aparición de enfermedades potencialmente nefastas para el individuo. Por esta razón, los estudios del genoma entero, del transcriptoma(conjunto de genes de las células) y de proteoma(conjunto de proteínas de las células) se utilizan para estudiar las causas genéticas y mecanismos moleculares asociados a las patologías humanas. En esta tesis, nosotros vamos a utilizar estudios del genoma para analizar asociación entre los genes particularmente los polimorfismos de nucleótido único(SNP) y la NOH, y también usaremos análisis de la transcriptoma para estudiar los mecanismos moleculares en los cuales involucran los genes(SNPs) asociados a NOH.

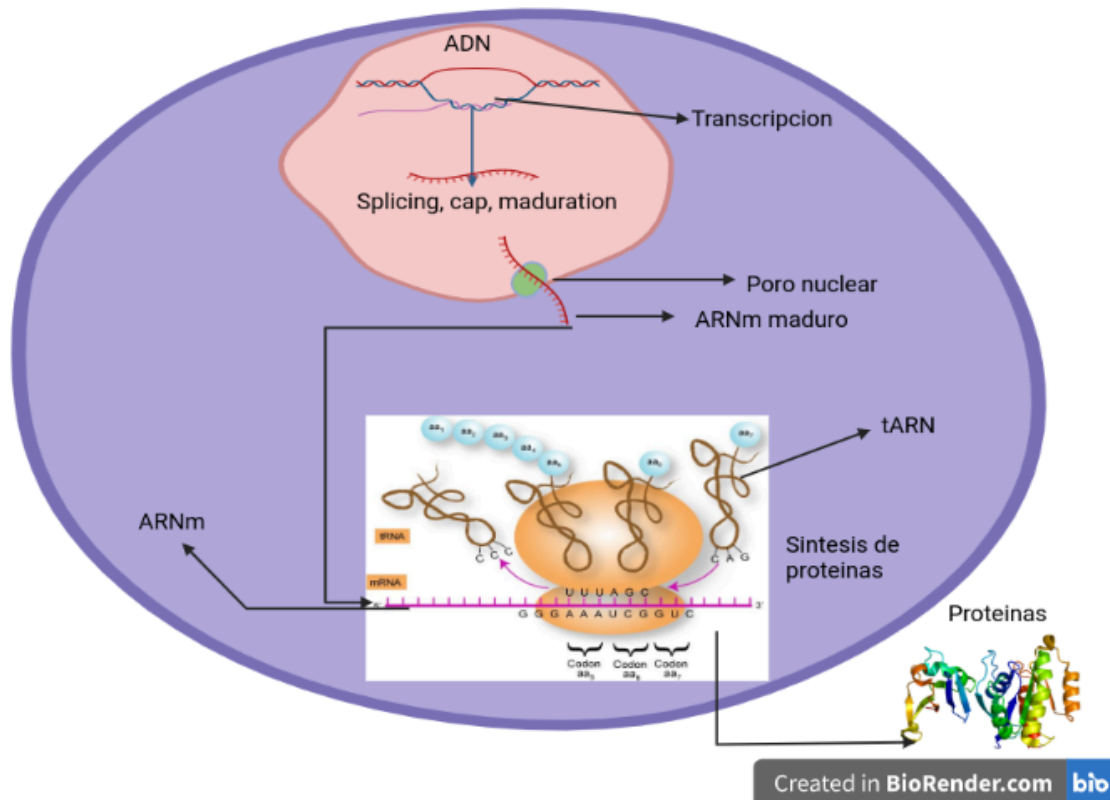


Figura 9: Resumen del dogma central de la genética. En esta figura presentamos un resumen de los procesos celulares de transcripción y traducción. En este caso obviamos la replicación. Realizado en nuestro laboratorio usando la aplicación Biorender.

3.2- Polimorfismo de nucleótido único(SNP)

En las dos últimas décadas, los avances en la secuenciación del genoma han proporcionado un acceso sin precedentes al genoma humano y han permitido documentar las variaciones de secuencia entre los individuos. Los seres humanos comparten un 99% de identidad a nivel de secuencia genómica, lo que implica que la diversidad fenotípica resultante se deriva del 1% restante de diferencia, así como de las modificaciones epigenéticas(Feuk, Carson, and Scherer 2006). Las diferencias de secuencia se deben a la presencia de repeticiones en tándem cortas y de número variable, a polimorfismos de inserción o delección y a polimorfismos de un solo nucleótido (SNP)(McCarroll et al. 2006). Entre los SNP, las transiciones ($A \leftrightarrow G$ o $C \leftrightarrow T$) son más frecuentes que las transiciones ($A \leftrightarrow C$ o T ; y $G \leftrightarrow C$ o T). Hay al menos 10 millones de SNP en el genoma, que se producen aproximadamente cada 100-300 pares de bases y con una frecuencia alélica superior al 1%, lo que hace que sean, con diferencia, el tipo de variante más común en el genoma humano. El tipo más común de variación genética humana es el polimorfismo de un solo nucleótido (SNP), en el que dos bases alternativas se dan con una frecuencia apreciable (>1%) en la población humana(D. G. Wang et al. 1998). Aunque en la actualidad se emplea

el término de SNV(single nucleotide variant) para referirse a variaciones genéticas sin tener en cuenta la frecuencia alélica. Esto implica que un SNV después de estudiarse puede ser un SNP o una mutación según si la frecuencia alélica(MAF) en la población es mayor o menor que 1%. En el caso que la MAF es menor que 1% de la población en estudio y que además está presente un fenotipo patológico asociado a esto se considera como una mutación.

Un polimorfismo puntual, también denominado de un solo nucleótido o SNP (Single Nucleotide Polymorphism, SNPs por su sigla en inglés), es una variación en la secuencia de ADN que afecta a una sola base (adenina (A), timina (T), citosina (C) o guanina (G)) de una secuencia del genoma(Sherry et al. 2001). Sin embargo, generalmente se considera que cambios de unos pocos nucleótidos, como también pequeñas inserciones y deleciones (indels) pueden ser consideradas como SNP(Sherry et al. 2001). Los SNP son regiones polimórficas de un solo par de bases del ADN que suelen variar de una persona a otra. Los SNP están presentes en todo nuestro genoma, con una media de un SNP por cada 72 pares de bases, y se han mapeado a lo largo de todos los cromosomas humanos(G. Li et al. 2014). Los SNPs por sí mismos no proporcionan información sobre genes específicos; simplemente indican una localización cromosómica que es probable que esté estrechamente asociada con un fenotipo dado(G. Li et al. 2014), por esto es recomendado completar los estudios genómicos con otros como la transcriptómica o la proteómica. Sin embargo, su estudio permite obtener más información sobre la estructura de las poblaciones a nivel genético y cómo estos se pueden correlacionar con ciertas enfermedades. Estas variaciones en la secuencia del ADN pueden afectar a la respuesta de los individuos a enfermedades, infecciones, fármacos y sustancias tóxicas.

Los SNVs tienen el potencial de alterar todos los pasos de la expresión génica dependiendo de su localización genómica. Cuando se produce a nivel de exones lo más probable será que se produzca una alteración estructural y funcional de la proteína, lo que conlleva a un mal funcionamiento de la misma y por consiguiente a enfermedades. Estos SNPs se llaman SNVs no-sinónimos. Si la variante provoca un codón de parada prematuro, puede dar lugar a la producción de un producto proteico truncado o a un fenotipo casi nulo debido al decaimiento mediado por el sinsentido(Mendell and Dietz 2001). Puede ser que se produzcan variantes de tipo silente causado por el hecho de que exista aminoácidos que son codificados por más de un codón y si se diera el caso que el cambio del nucleótido produjera un codón que codifica para el mismo aminoácido, entonces no habrá cambio en la secuencia del polipéptido. En tal caso, la proteína estará estructural y funcionalmente intacta. En caso que estos cambios surgen en las regiones intrónicas pueden tener un impacto negativo en el empalme del ARNm, la exportación núcleo-citoplásmica, la

estabilidad y la traducción. Muchos de estos SNVs no codificantes que residen en las secuencias no codificantes (por ejemplo, promotores, potenciadores y 5'/3'-UTR) que rodean a los genes codificantes de proteínas han demostrado tener profundos efectos en la expresión de los genes vecinos y pueden causar patologías humanas (Schaub et al. 2012), (Molineris et al. 2013) por lo que se denominan SNVs reguladores (rSNVs). En tal caso se dice que la variación es sinónima (es decir, no cambia la naturaleza del aminoácido), y las tasas de traducción o la vida media del ARNm pueden verse afectadas.

Con el auge del análisis del genoma entero llamado GWAS (*Genome wide association study* por su definición en inglés), se enfocan mucho en el descubrimiento de las variantes genéticas producidas a nivel de los intrones y regiones intergénicas (que no codifican para proteínas), que se asocian también a enfermedades. Los intrones intervienen en la formación de los ARN reguladores que intervienen en la regulación de otros procesos moleculares y celulares como la expresión de los genes. De esta forma, alteraciones de ADN intrones podrían interferir con las funciones normales de dichos procesos moleculares participando así en la ocurrencia de enfermedades. Hay bastante pruebas que sugieren que muchos SNPs no codificantes, especialmente los que se encuentran en las proximidades de los genes codificantes de proteínas, desempeñan un papel importante en la formación de la estructura de la cromatina y regulan la expresión de los genes, por lo que están implicados en una gran variedad de enfermedades (G. Li et al. 2014). Las variantes genéticas contribuyen a prácticamente todas las enfermedades humanas, confiriendo susceptibilidad o resistencia o influyendo en la interacción con los factores ambientales (Collins, Guyer, and Chakravarti 1997).

Aunque se ha prestado mucha atención a los SNVs en las secuencias codificantes en los estudios de asociación de todo el genoma (GWAS), el papel de los SNVs no codificantes, que cuentan más que los SNVs codificantes, está mucho menos estudiado. A diferencia de los SNVs codificantes, que causan un cambio en las secuencias de aminoácidos o no hacen nada, los rSNP pueden tener un efecto a nivel de transcripción de los genes vecinos. Se pueden atribuir múltiples mecanismos a dicho efecto, entre los que se incluyen la afinidad de unión del factor de transcripción de la proteína o la alteración de la metilación del promotor (Ziller et al. 2013). También es probable que los rSNV afecten la unión específica de la secuencia del factor transcripcional al promotor o enhancer. Se ha demostrado recientemente que los miRNAs y los RNAs no codificantes largos (ncRNAs) pueden regular la transcripción de los genes o la estructura de la cromatina de forma dependiente de la secuencia (Huang et al. 2012).

Las variaciones normales de la expresión de genes entre diferentes individuos suelen considerarse como normales por las variantes genéticas que cada individuo presenta en los

genes. Según recientes estudios estas variaciones están causadas por los SNVs encontrados en las regiones reguladoras no codificantes del ADN donde se unen factores de transcripción (Kasowski et al. 2010).

En cuanto a esta tesis nosotros nos enfocamos en dos métodos que se estudian en transcriptomics, los cuales se presentarán en apartados más adelante. Estudiaremos la expresión diferencial de genes de células dopaminérgicas de pacientes con NOH vs los que no presentan esta condición. Este análisis permitirá confirmar los genes que están asociados a NOH en los cuales se localizan los SNVs previamente encontrados asociados a NOH. Otro análisis que se hará en este presente estudio utilizando la transcriptómica será el análisis de eQTL (Expresión cuantitativa de rasgos de locus). Este análisis nos permitirá poder estudiar los efectos de los SNPs encontrados asociados a NOH sobre la expresión de los genes en células dopaminérgicas.

3.3- Genotipos y fenotipos

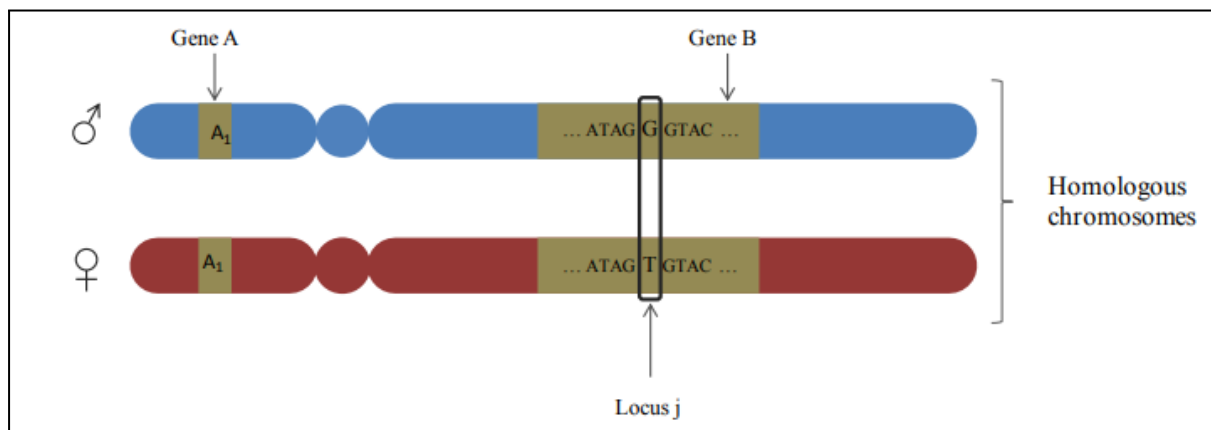
En genética, se define el fenotipo como la expresión del genotipo de un individuo dado incluyendo a rasgos tanto físicos, morfología, desarrollo, propiedades bioquímicas, fisiológicas, como conductuales, en función del medio ambiente con el cual él mismo interactúa. El genotipo y el fenotipo no están siempre correlacionados directamente. Algunos genes solo expresan un fenotipo dado bajo ciertas condiciones ambientales. Inversamente, algunos fenotipos pueden ser el resultado de la combinación de varios alelos localizados en locus diferentes, lo que se conoce como pleiotropismo: El modelo se denomina poligenico si un fenotipo está causado por múltiples variantes localizadas en locus diferentes, pleiotropía si una variante puede provocar varios fenotipos. El fenotipo está principalmente determinado por el genotipo del individuo caracterizado por los alelos que se encuentran en una o más posiciones del genoma del individuo. También su expresión está influenciada por el factor ambiente. La expresión del fenotipo como resultante de las interacciones entre ambiente y genotipo, se traduce en la ecuación siguiente:

$$\text{Fenotipo} = \text{Genotipo} + \text{Ambiente}$$

En este estudio, consideraremos solo los genotipos y sus asociaciones con el NOH omitiendo los efectos del medio ambiente sobre el fenotipo aunque en casos específicos consideramos variables como la edad, el sexo, la etnia de los individuos, tratamiento y algunas escalas clínicas. Por esta razón no entraremos en el estudio detallado de aquella interacción fenotipo-genotipo-ambiente.

Para el estudio de las asociaciones entre las variantes genéticas y fenotipos, se usan múltiples métodos o modelos según si las causas genéticas del fenotipo son múltiples o no.

En caso que una variante cause la enfermedad se le conoce como rasgos monogénicos. Generalmente, se trata de mutaciones encontradas a nivel de los exones que provocan alteraciones estructurales y funcionales de las proteínas codificadas por las mismas. El fenotipo resultante puede ser principalmente de tipo dominante, recesivo, codominante. Se habla de fenotipo dominante cuando se expresa el fenotipo con la presencia de al menos un alelo de riesgo y no es dosis dependiente, el fenotipo recesivo requiere la presencia de los dos alelos de riesgos (el individuo tiene que ser homocigota para el alelo de riesgo), y el fenotipo codominante cuando en los individuos heterocigotas los dos alelos se expresan de forma igual y no existe recesividad ni dominancia de un alelo sobre el otro. En nuestro estudio, en los análisis de asociación, consideraremos los modelos genéticos dominante, recesivo y aditivo (para ver cómo influyen la dosis de los alelos en los otros dos modelos).



Fuente 10: representación esquemática de variantes genéticas SNPs y genotipos. En el caso del gen A se aprecia que el individuo tiene el mismo alelo A₁ (Pudiera ser por ejemplo una adenina) en ambos cromosomas homólogos para este locus específico dentro del gen A. Se dice que este individuo es homocigota para el locus. Mientras tanto en el gen B del cromosoma heredado del padre, se ve que tiene el alelo G y en el otro cromosoma heredado de la madre el alelo T. En este caso se dice que este individuo es heterocigota para este locus. Esta imagen se obtuvo de la tesis doctoral de Helena Brunel Montaner. <https://www.tdx.cat/bitstream/handle/10803/134362/THBM1de1.EPf>.

3.4- Identificación de SNPs por secuenciación de última generación

Para identificar los SNPs en los pacientes, un largo proceso de laboratorio y análisis bioinformático usando múltiples workflows y aplicaciones informáticas se llevan a cabo. La primera fase del proceso es la secuenciación del ADN de los individuos usando diferentes muestras como la sangre, LCR, tejidos, u otros fluidos corporales. Frederick Sanger juega un rol importante en el desarrollo de los métodos de secuenciación del ADN que permite establecer las bases teórico-metodológicas de la misma. El invento la técnica llamada la secuenciación de Sanger (Pevsner 2009). La técnica de secuenciación de Sanger consiste en poner en 4 tubos el mismo fragmento de ADN (cadena simple) a secuenciar, un didesoxinucleótido (que no tiene el grupo hidroxilo en la posición 3' para que se introduzca un solo nucleótido a la vez y que no siga creciendo la cadena), el ADN polimerasa y los primers. Después se realiza un gel de agarosa para identificar cual de los ddNTPs se

introducen en el fragmento. Este método constituye la base de casi todos los métodos de secuenciación subsiguientes. Este método se automatizó usando la electroforesis capilar en donde se realiza PCR utilizando esta vez dNTPs marcados con fluoróforos. Los fragmentos se introducen dentro del capilar de electroforesis y las bases nitrogenadas son detectadas según la fluorescencia por el detector y analizados bioinformáticamente hasta obtener la posición de cada base en la secuencia.

Después de la terminación en 2003 del proyecto del genoma humano, que consiste en la secuenciación de todo el genoma humano (McElheny 2010), se han intensificado las investigaciones en el ámbito de la secuenciación del ADN que conlleva al desarrollo de nuevas técnicas de análisis de datos genómicos y nuevo campo de estudio como la genómica bioinformática y la genómica funcional. También, se multiplican la generación de los datos genómicos y se hace necesaria la construcción de grandes bases de datos adaptados a estos tipos de datos y además software que permiten acceder a los mismos y analizarlos. Unos 10 años después de la primera secuenciación del genoma entero, se completó otro proyecto llamado 1000 Genomes Project durante el cual se secuenció 1093 genomas humanos de diferentes etnias y edades (1000 Genomes Project Consortium et al. 2012). A partir de entonces, las bases de datos genómicos, los recursos bioinformáticos no cesan de incrementarse. En nuestro caso de la presente tesis por ejemplo, nos ocuparemos de la base de datos del PPMI que contienen datos fenotípicos y genéticos de pacientes que tienen enfermedad de parkinson. Últimamente son muchos los doctorandos que usan datos de bases de datos para llevar a cabo sus tesis doctorales.

3.4.1- Secuenciación de próxima generación

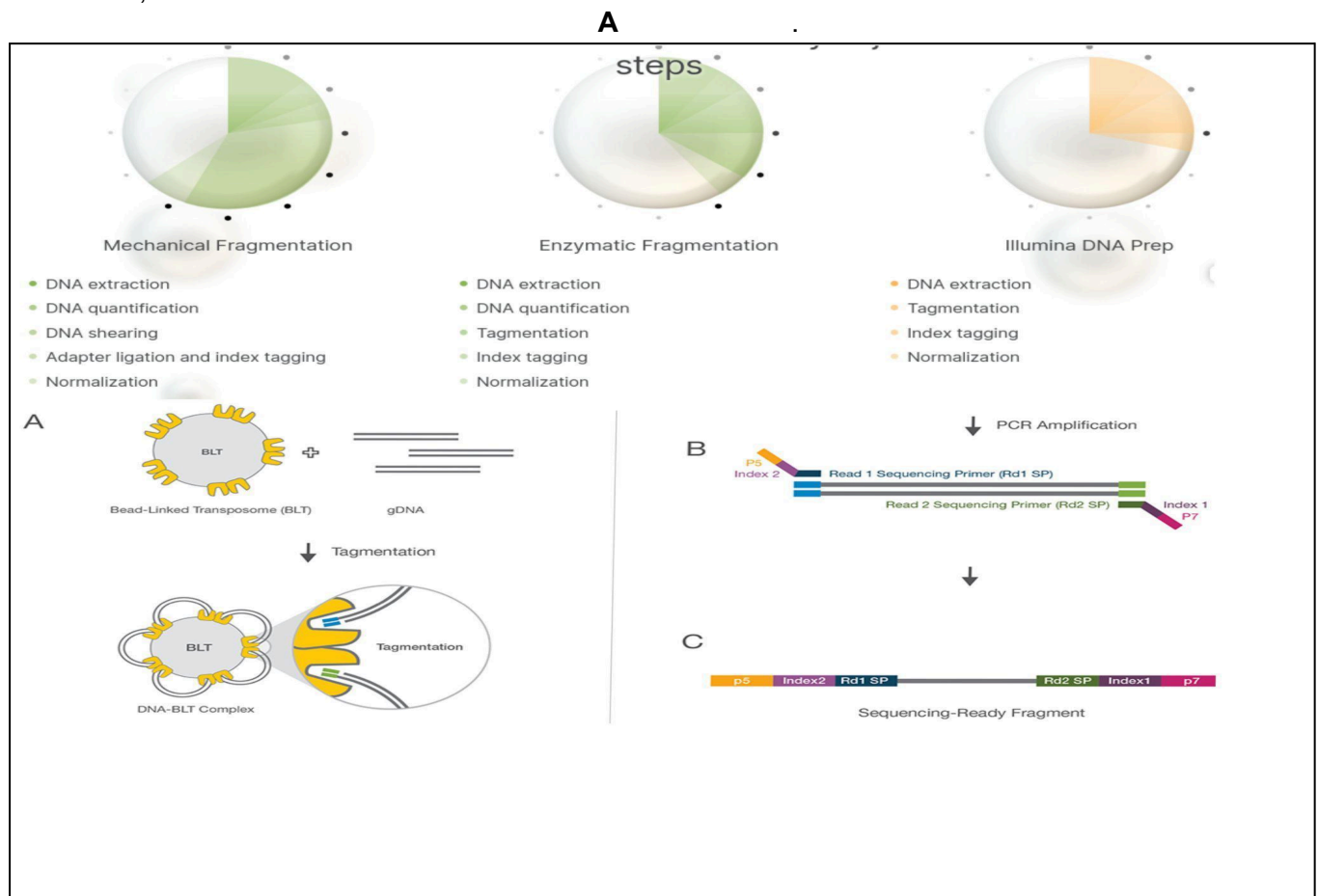
La secuenciación de nueva generación (NGS) es una tecnología de secuenciación masiva en paralelo que ofrece un rendimiento, una escalabilidad y una velocidad muy elevados. La tecnología se utiliza para determinar el orden de los nucleótidos en genomas enteros o regiones específicas de ADN o ARN. La NGS ha revolucionado las ciencias biológicas, permitiendo a los laboratorios realizar una gran variedad de aplicaciones y estudiar sistemas biológicos a un nivel nunca antes posible. La tecnología de secuenciación de próxima generación ha cambiado fundamentalmente el tipo de preguntas que los científicos pueden plantear y responder. Las innovadoras opciones de preparación de muestras y análisis de datos permiten una amplia gama de aplicaciones. Por ejemplo, la NGS permite a los laboratorios: Secuencia rápida de genomas completos, secuenciar en profundidad de regiones target, utilizar la secuenciación de ARN (RNA-Seq) para descubrir nuevas variantes de ARN y sitios de empalme, o cuantificar los ARNm para el análisis de la expresión génica, analizar factores epigenéticos como la metilación del ADN en todo el

genoma y las interacciones ADN-proteína, secuenciar muestras de cáncer para estudiar variantes somáticas raras, subclones tumorales, estudiar el microbioma humano, identificar nuevos patógenos, identificar nuevas mutaciones u otras variantes asociadas a enfermedades como la EP.

Existen actualmente en el mercado muchas tecnologías para la secuenciación del ADN: Illumina, Roche 454, Ion Torrent, Oxford Nanopore. Más del 90% de los datos genómicos disponibles en las bases de datos son secuenciados por la tecnología Illumina, por esta razón vamos a describir someramente esta tecnología, además que los datos que vamos utilizar en esta tesis fueron generados por esta misma tecnología de secuenciación.

Los procesos de laboratorio para la secuenciación empieza con la extracción del ADN y el control de calidad del mismo incluyendo a la determinación de la concentración de la misma. Después sigue el proceso de la construcción o preparación de la librería que consiste en varios pasos: fragmentación del ADN, ligación del ADN al adaptador, ligación del índice, y normalización. Este workflow es propuesto por la tecnología Illumina para la secuenciación del genoma entero(Figura 11 A).

La fragmentación consiste en fragmentar el DNA en secuencias cortas por métodos mecanicas,enzimáticas



B

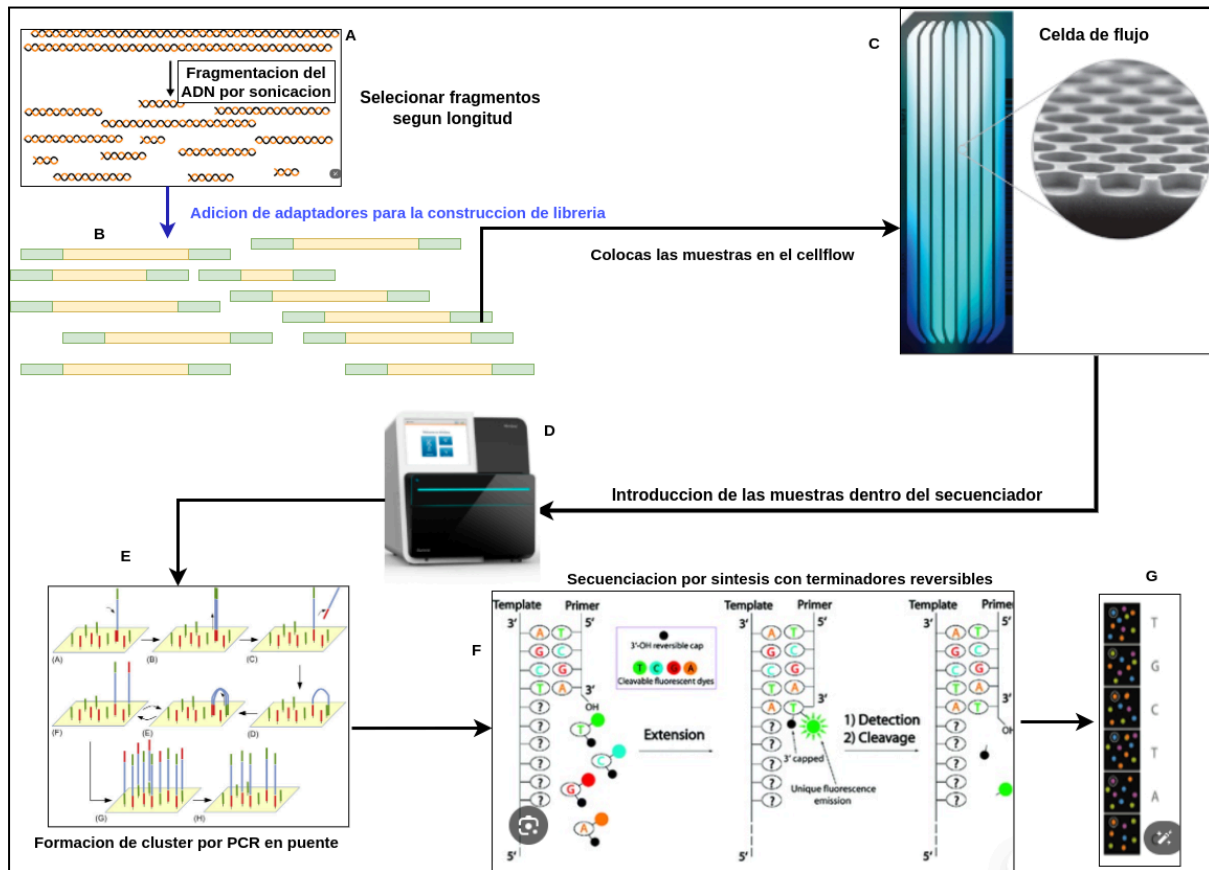


Figura 11(A-B): Tecnología de secuenciación de la plataforma Illumina. **Figura 11-A:** Preparación de librería de ADN por tecnología de secuenciación de Illumina. <https://www.illumina.com/>. En esta imagen se presentan los procesos de preparación de las librerías del ADN y los procesos de multiplexación. Existen maneras de hacer la fragmentación: fragmentación mecánica(sonicación), fragmentación enzimática y fragmentación. Se empieza con la extracción y purificación del ADN, la sonicación y fragmentación, selección de los fragmentos y multiplexación. La multiplexación consiste en la adición de fragmentos de oligonucleótidos que de un lado permite la identificación de las diferentes muestras, fragmentos que se usarán como primers y fragmento para la unión con la fase sólida del flow cell.

Figura 11-B: Workflow para la secuenciación de ADN en la plataforma Illumina. **A:** Fragmentación del ADN por sonicación, reparación de los extremos para que queden extremos romos, adición de una base adenina a cada extremo para complementarse con una timina del adaptador. **B:** multiplexación que consiste en la adición de los adaptadores a los extremos de los fragmentos. Los adaptadores tienen los oligonucleótidos siguientes: el código de barra también llamado index que permite identificar a las muestras(cada index corresponde a una muestra), primers para el PCR en puente, un oligonucleótido en los extremos para complementarse con otro fragmento del flowcell. Después de añadir los adaptadores se selecciona a través de gel de agarosa y purifica el ADN y se realiza PCR para enriquecer la muestra y aumentar la probabilidad de que los fragmentos puedan unirse con el flow cell(hay que tener en cuenta que este paso si bien es importante pero es mayor fuente de duplicación lo cual se tendrá que tratar durante los análisis bioinformático de los datos). **C:** En esta fase, la muestra se coloca en el flowcell. El flowcell es un chip de vidrio formado por canales en donde se colocan las muestras. Los canales están formados por más de 10 mil millones de nano wells donde se producen las reacciones de secuenciación. Los nanowells contienen incrustado covalentemente un oligonucleótido que se unirá con el adaptador cuando se añada la muestra. **D:** El flow cell con la muestra se introduce en el secuenciador para seguir con el proceso de secuenciación de la muestra. **E:** se realiza un PCR en puente para la formación de cluster constituido por clones de un mismo fragmento, lo que aumenta la señal de fluorescencia captada: el fragmento de cadena simple se añade al nanowell y se une por puente hidrógeno por complementariedad al oligonucleótido unido covalentemente a la superficie del nanowell. El fragmento del oligonucleótido unido covalentemente al flow cell crece por complementariedad al nuevo fragmento. Este último se eliminó como no se unió covalentemente. El fragmento de cadena única se encorva su extremidad libre para unirse por complementariedad con otro oligonucleótido unido al flow cell covalentemente. Este último oligonucleótido se crece sobre el fragmento por complementariedad resultando en un ADN doble cadena con los extremos unidos covalentemente al flow cell. Se desnaturaliza para liberar las cadenas simples que van a empezar de nuevo con el mismo proceso. Las dos cadenas cada una por su parte van a unirse con un nuevo oligonucleótido después de crecer para formar una nueva cadena complementaria. Se desnaturaliza para dar formación a dos cadenas nuevas. Se repite el proceso en varios ciclos para formar clusters de fragmentos idénticos. Después de desnaturalizar los fragmentos del cluster se eliminan los fragmentos para quedarse solo con los forward los cuales se usará en esta primera fase de secuenciación. **F:** Secuenciación por síntesis: la síntesis se realiza al añadir un nucleótido marcado con fluorescencia a la cadena creciente. El nucleótido se bloquea en el extremo 3'-OH para que se une un solo nucleótido a la vez por ciclo. El ciclo se termina al liberar el extremo 3'-OH, listo para recibir un nuevo nucleótido. Se excita el flow cell por láser y se detecta la longitud de onda emitida por la excitación de la fluorescencia. Se capta a través de una cámara. En cada ciclo se saca una imagen que contiene la fluorescencia captada en cada cluster por cada ciclo. Un cluster corresponde a un fragmento o una lectura. Illumina utiliza un algoritmo que permite hacer la llamada de

bases y asignar un score de calidad para cada base llamada según la intensidad captada. Al finalizar, el secuenciador devuelve un archivo formato fastq que contiene las lecturas(fragmentos), los nucleótidos que conforman cada fragmento y sus scores de calidad además de otras informaciones sobre la posición del fragmento en el flow cell.

Un avance importante es la tecnología de tagmentación del ADN, en la cual se utilizan perlas unidas a transposones para proporcionar una reacción de tagmentación más uniforme en comparación con las reacciones de tagmentación en solución. Después de los transposones unidos a perlas se saturan de ADN, y que no se puede producir una tagmentación adicional, lo que permite un proceso de normalización basado en la saturación altamente uniforme. La tagmentación incluye la fragmentación del ADN y la adición del primer de forma simultánea(Figura 11 A). Después, se realiza amplificación por PCR de los fragmentos listos obtenidos del paso anterior y se les adicionan los adaptadores y los índices, proceso llamado multiplexación. Con estos pasos estarán listos los fragmentos del ADN para la secuenciación.

Secuenciación

Durante el paso de secuenciación del flujo de trabajo de NGS, las librerías de ADN se cargan en una celda de flujo(Flow Cell) y se colocan en el secuenciador. Los grupos o clusters de fragmentos de ADN se amplificaron en un proceso denominado generación de clusters, lo que da lugar a millones de copias de ADN monocatenario. En la mayoría de los instrumentos de secuenciación de Illumina, la generación de clústeres se produce automáticamente. En un proceso denominado secuenciación por síntesis (SBS), los nucleótidos modificados químicamente se unen a la cadena molde de ADN mediante complementariedad natural. Cada nucleótido contiene una etiqueta fluorescente y un terminador reversible que bloquea la incorporación de la siguiente base. La señal fluorescente indica qué nucleótido se ha añadido, y el terminador se escinde para que la siguiente base pueda unirse. Tras la lectura de la cadena de ADN anterior, las lecturas se eliminan y el proceso se repite para la cadena inversa. Este método se denomina secuenciación por pares. Tras la secuenciación, el software del instrumento identifica los nucleótidos (un proceso denominado "llamada de bases") y la precisión prevista de las "llamadas de bases" (Ver figura 11 B).

3.4.2- Análisis Bioinformáticos:

Se obtienen los datos en formato fastq, que contiene las lecturas de los fragmentos de ADN y los scores de calidad y demás informaciones importantes para los procesos subsiguientes. El primer paso en el workflow del análisis bioinformático sería el control de calidad de los datos. En breve, este proceso de control de calidad consiste en demultiplexación de los datos, elección y filtro de las lecturas con scores de buena calidad, elección y filtro de las

lecturas según la cobertura etc. La demultiplexación consiste en eliminar los barcodes, los índices, los primers que se utilizaron tanto para la identificación de las muestras como para la realización de PCR. Se utiliza el programa fastQC para el análisis de los datos y la identificación de los elementos a eliminar según los criterios establecidos y otro programa como Trimomatic para el filtro y eliminación de los que no pasan los criterios. Los datos filtrados son mapeados contra un genoma de referencia. Después se realiza la llamada de variantes por cada muestra para poder descubrir INDELS, SNPs y después poolear todos en un solo archivo gVCF. Los SNPs e INDELS serán filtrados según criterios de calidad y son anotados (Ver figura 12).

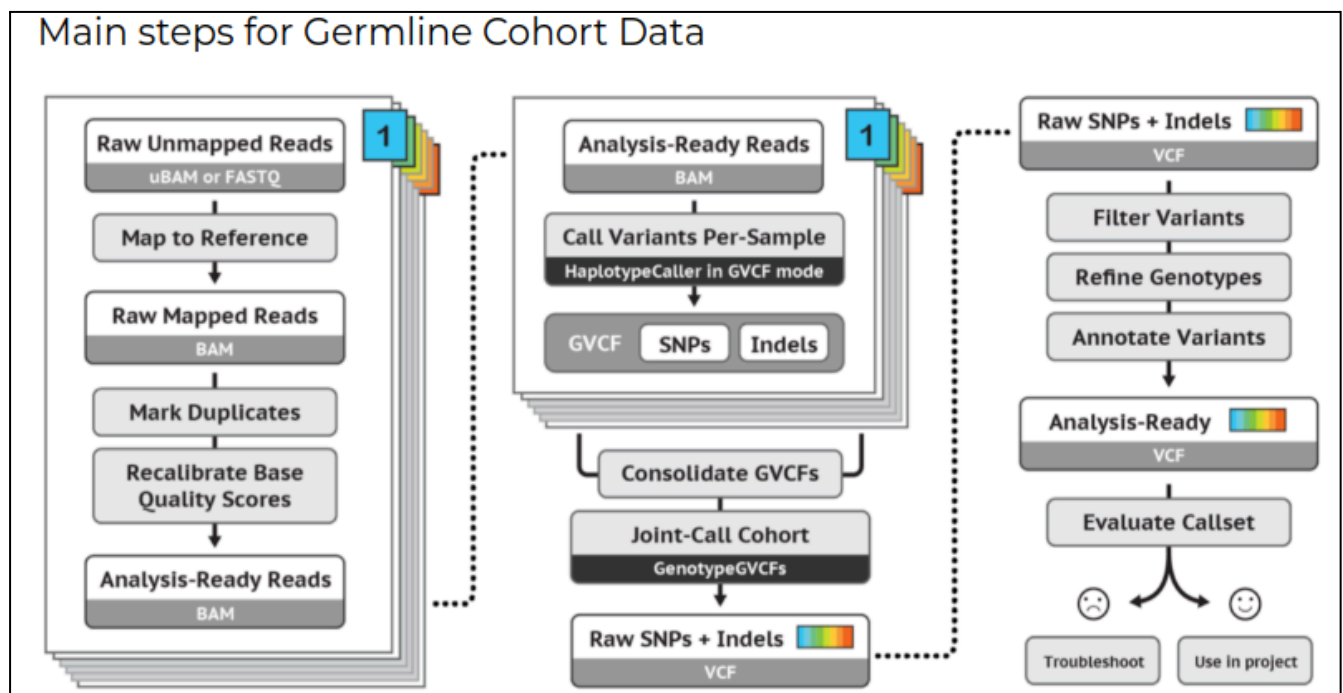


Figura 12- Workflow para análisis bioinformáticos de los datos genéticos tipo fastq. En esta imagen se presentan los pasos y workflows para los procesos informáticos de transformación de los datos de secuenciación hasta obtener los datos de SNPs. Del secuenciador se obtiene un archivo tipo fastq el cual se usa en este workflow. En primer lugar se hace un control de calidad de los datos durante el cual se filtran los datos con score de calidad Phred menor a 30 y se hace trimming para demultiplexar los datos (eliminar los barcodes, los índices, los primers etc.), y tratar los datos duplicados. Después se hace el mapeo de las lecturas contra un genoma de referencia previamente indexadas. Al final se hace una llamada de SNPs. <https://gatk.broadinstitute.org/hc/en-us/articles/360035535932-Germline-short-variant-discovery-SNPs-Indels->

3.4.3- Análisis de genoma entero (GWAS)

El análisis GWAS es un conjunto de análisis estadísticos realizados sobre los datos genotípicos con el objetivo de descubrir variantes asociadas a patologías. Se estudia una gran cantidad de individuos para aumentar el poder estadístico de descubrir variantes asociadas a la enfermedad. Generalmente, se realiza un estudio caso-control en donde los casos son los que tienen la enfermedad y los controles no la tienen y se comparan las frecuencias alélicas en ambos grupos. Son múltiples los análisis estadísticos para descubrir

las asociaciones existentes entre fenotipos y variantes genéticas. Un primer abordaje es el llamado asociación de un locus al fenotipo, que consiste en la determinación de la asociación entre cada variante y el fenotipo utilizando estadística como X^2 , odd ratio, regresión lineal o logística simple.

Sea la tabla de contingencia siguiente donde A1/A1 representa los homocigotas para el alelo mayor, el genotipo A1/a representa el heterocigota llevando el alelo a target y el genotipo a/a homocigota para el alelo menor. El chi cuadrado determina qué tan diferente están las frecuencias alélicas observadas y las esperadas y se puede definir por la ecuación siguiente:

	A1/A1=0	A1/a =1	a/a =2	Total
Casos	B1	B2	B3	B7
Controles	B4	B5	B6	B6
Total	B9	B10	B11	B12

Tabla 2- tabla de contingencia. En esta tabla se presenta una tabla de contingencia de 3 X 2 de las frecuencias genotípicas de los casos y los controles. Se realiza la suma de las frecuencias genotípicas homocigotas para el alelo menor, para el alelo mayor y para las heterocigotas tanto para los casos como para los controles. Después, usando la tabla se sacan OR, Chi cuadrado, Test de Fisher.

$$X^2 = \sum_{i=0}^6 (B_i - E_i) / E_i \quad (1.1)$$

en donde B representa las frecuencias alélicas observadas y E las frecuencias esperadas. La evaluación de los resultados se hace por el valor estadístico de X^2 obtenido y el p-valor, que determina la significancia de la asociación entre los genotipos y el fenotipo.

Por su lado, el odd ratio determina la fuerza de tamaño y fuerza de asociación entre dos variables binarias. Se define por la ecuación siguiente

$$OR = B_4(B_2 + B_3) / (B_5 + B_6)B_1 \quad (1.2)$$

La regresión es un enfoque para modelar la relación entre un resultado (el fenotipo Y) y una variable (un SNP S). En la regresión lineal los datos se modelan utilizando funciones lineales, y los parámetros desconocidos del modelo se estiman a partir de los datos utilizando la técnica de mínimos cuadrados o por algoritmo de optimización como el descenso del gradiente(gradiente descendente estocástico). Los modelos lineales intentan expresar la relación entre Y y S mediante una función afín como la descrita en la ecuación siguiente:

$$\gamma = \alpha + \beta S + \varepsilon$$

en donde β es el coeficiente de regresión calculado a partir de los datos y ε es una variable aleatoria que recoge todos aquellos factores de la realidad no controlables u observables y que por tanto se asocian con el azar. La determinación de una estadística basada en Test de student que evalúa la significancia de β , que representa la asociación entre Y y S, se lleva a cabo.

También se usa la regresión lineal múltiple que permite analizar la relación entre dos o más variables a través de ecuaciones. Constantemente en la práctica de la investigación estadística, se encuentran variables que de alguna manera están relacionadas entre sí, por lo que es posible que una de las variables puedan relacionarse matemáticamente en función de otra u otras variables. Este modelo cuenta con varias variables regresoras, por lo que cuenta con varios parámetros, para la regresión lineal múltiple, se asume que la variable de respuesta Y se relaciona con las variables regresoras $X_0, X_1, \dots, X_m$:

$$\begin{aligned} Y &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon \\ &= \beta_0 + \sum_{j=1}^m \beta_j X_j + \varepsilon \end{aligned}$$

donde los parámetros β_j , $\forall j \in \{0; 1; 2; \dots m\}$ Son llamados coeficientes del modelo de regresión múltiple.

Cuando los datos de respuesta son de tipo binario, se utiliza la regresión logística definida por la ecuación:

$$Y \approx \frac{1}{1 + e^{-z}}$$

$$\text{con } z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon.$$

3.4.4- Recombinación genética y desequilibrio de ligamiento

En genética se denomina desequilibrio del ligamiento o LD por sus siglas en inglés (Linkage Disequilibrium) a la propiedad de algunos genes de las poblaciones genéticas de segregar de forma dependiente, esto es, poseen una frecuencia de recombinación menor del 50 %. Esto suele deberse a que los dos loci implicados se encuentran en el mismo cromosoma, lo que imposibilita su transferencia a la progenie de manera aleatoria con la separación de los cromosomas en anafase (*Genética* - 7. Ed 2002). El componente básico de las métricas de LD es la diferencia entre las frecuencias observadas y las esperadas de un haplotipo

suponiendo que no hay asociación estadística. Si A y B son dos loci con dos alelos A1, A2 y B1, B2 respectivamente, el desequilibrio de ligamiento entre A y B es D se mide como se define en la ecuación siguiente:

$$D(A, B) = p_{A1B1} - p_{A1}p_{B1}$$

donde p_{A1} es la frecuencia alélica en locus A, p_{B1} es la frecuencia alélica en locus B y p_{A1B1} es la frecuencia alélica de la combinación A1B1.

3.5- Fundamentos del análisis de expresión de genes basado en desequilibrio de ligamentos y eQTL

El análisis de expresión de genes basado en desequilibrio de ligamiento llamado cis-eQTL es un análisis que se viene usando ultimamente en el descubrimiento de nuevos marcadores moleculares que sirven como blanco terapéutico y como marcador de diagnóstico molecular de las enfermedades (Mize and Evans 2022), (Duffy 2017; S. Lee et al. 2017). Este tipo de análisis permite determinar cómo los SNVs asociados a una patología, puede influenciar la expresión de los genes teniendo en cuenta la distancia entre los SNVs y los genes cuya expresión está afectada (Duffy 2017). Los mismos pueden verse sobreexpresados bajo las influencias de los SNVs o sub-expresados. Alteraciones en la expresión de los genes afectan el buen funcionamiento de los mismos y puede estar asociado con la ocurrencia de patologías humanas. Se sabe hasta este momento que muchas de las enfermedades humanas son de etiología multigénica quiere decir que en sus ocurrencias pueden participar muchos genes (J. Yang et al. 2014). El desequilibrio de ligamiento es el fenómeno por el cual los SNVs transmitidos a las generaciones tienden a estar relacionados los unos con los otros por el hecho que durante la recombinación genética, los SNPs tienden a segregarse en bloque y transmitirse a las generaciones (Lou et al. 2005). Los estudios que se basan sobre la determinación de los efectos de los SNVs en desequilibrio de ligamiento sobre la expresión de los genes cercanos donde se localizan los SNVs en DL aumentan o confirman aún más las relaciones no aleatorias entre los mismos y los genes cuyas expresiones se ven afectadas (Lou et al. 2005; Wu, Ma, and Casella 2002). Entonces, la idea detrás de estos análisis es no solamente demostrar estadísticamente significativo algún tipo de relación de desequilibrio de ligamiento entre al menos dos SNVs, sino también estudiar el efecto que tienen estos SNVs sobre la expresión del gen donde se localizan los mismos. Una vez conocido los SNVs que están asociados a una patología dado, con estos análisis, será posible mapear los genes que también contribuyen en el proceso patológico. Muchos métodos estadísticos fueron propuestos para estudiar la

expresión cuantitativa de rasgos de locus basado en desequilibrio de ligamentos. Uno de los métodos estadísticos que se utilizan en el análisis eQTL basado en LD es la regresión lineal simple (Palowitch et al. 2018), (Palowitch et al. 2018; Jia and Xu 2005) donde el objetivo es estudiar el efecto de alguna variable cuantitativa sobre la expresión de los genes, en este caso, los SNVs representados por los valores 0, 1, 2 según si se trata de genotipos homocigotas para el alelo mayor (no presenta el alelo de riesgo), heterocigotas y homocigotas para el alelo menor (doble dosis del alelo de riesgo) respectivamente. En nuestro estudio usaremos el análisis de regresión lineal y ANOVA de una vía usando workflow en KNIME construido por nosotros para poder estudiar los efectos de los SNVs sobre la expresión de los genes.

3.6- Fundamentos del análisis de expresión diferencial de genes.

El análisis de expresión diferencial de genes constituye una metodología que se viene utilizando para estudiar los patrones de expresión de los genes de forma diferencial de un grupo de individuos que tienen un fenotipo dado con respecto a otro grupo control. Conociendo los genes que son diferencialmente expresados, podemos determinar las vías metabólicas y/o de señalización en las cuales participan y cómo afectan dichas vías. Pero además, nos permite confirmar que un gen está asociado con alguna patología previamente estudiada por GWAS (Genome Wide Association Study). El estudio GWAS nos permite establecer asociación entre patologías y variantes genéticas como INDEL, SNVs. Esta variante encontrada asociada con la patología no quiere decir que sea la verdadera causante del problema. Podría ser que sea porque está en desequilibrio de ligamiento con el verdadero SNV responsable del problema. Uno de los análisis que se puede hacer para confirmar la asociación entre la variante y la patología es a través de un análisis diferencial de expresión de genes (DGE). La idea es determinar si el gen donde se encuentra localizado la variante está diferencialmente expresado en el grupo que tiene la patología con respecto al grupo control. Hay que recalcar también que un análisis DGE negativo para un gen (donde se localiza la variante encontrada asociada al fenotipo) tampoco implica que la variante no sea la causante del fenotipo.

Capítulo IV- Métodos de aprendizaje automático para la predicción de enfermedades.

4.1- Generalidades sobre aprendizaje automático.

El aprendizaje automático se define como una serie de técnicas estadísticas avanzadas que incluyen algoritmos de clasificación y de regresiones que se usan para poder hacer predicción sobre los datos. Se dividen en dos grupos: Aprendizaje automático de tipo

supervisado y no supervisado, aunque existen otros grupos que últimamente están en auge como el grupo semisupervizado y aprendizaje por reforzamiento. En este trabajo utilizaremos el aprendizaje supervisado para poder hacer análisis de predicción sobre los SNPs asociados a NOH. El ML se usa en investigaciones biológicas y médicas(Y. Deng et al. 2022; Akmal et al. 2022) y en diferentes métodos de diagnósticos, de predicción de enfermedades y otros métodos para búsqueda de medicamentos(Y. Deng et al. 2022). Específicamente, en lo que tiene que ver con nuestro estudio, muchos estudios anteriores han utilizado métodos de aprendizaje automáticos en la investigación genómica y transcriptómica. Dentro del grupo de los aprendizajes automáticos supervisados tenemos a la regresión logística, redes neuronales, Naives Bayes, Random Forest, XG Boost Trees y demás y del grupo de aprendizaje no supervisado tenemos a los métodos de conglomerados y agrupamientos.

4.2- Neural Network(Redes Neuronales)

4.2.1- Métodos

Las redes neuronales se concibieron originalmente como modelos computacionales del funcionamiento del cerebro humano. Al igual que el cerebro humano, están formadas por muchas unidades (análogas a las neuronas y a veces llamadas el mismo nombre) conectadas entre sí por enlaces de fuerza variable (análogos a los axones en el cerebro). Estos enlaces de fuerza variable son representaciones abstractas de la forma en que la mayoría de las neuronas se comunican entre sí en el cerebro a través de cambios en el ritmo o la frecuencia de los mensajes eléctricos o químicos. Esta técnica se ha inspirado en la forma en que los organismos biológicos (en particular los humanos) resuelven los problemas de computación en la naturaleza. Como modelos matemáticos, han encontrado un gran número de aplicaciones en la ciencia y tecnología, especialmente en el ámbito de las ciencias biológicas, en medicina y en Bioinformática. El atractivo de las redes neuronales es que pueden "aprender" relaciones entre conjuntos de variables tomadas de un sistema. Una vez entrenada, la red puede recibir nuevos ejemplos y se le pide que prediga el resultado de los nuevos datos basándose en los ejemplos anteriores que ha aprendido. Esta cualidad, conocida como generalización, es la capacidad de inferir las relaciones subyacentes en los datos y ser capaz de aplicarlas a nuevas situaciones y es la razón principal de su uso en una gran variedad de contextos. Esto puede sonar parecido al método por el que aprenden los humanos y, hasta cierto punto, es cierto.

4.2.2- Arquitectura

Una red neuronal está formada por unidades interconectadas, a menudo dispuestas en capas. La configuración de estas unidades se conoce como arquitectura, y puede variar mucho en función de la aplicación para la que se utilice. En las redes neuronales más sencillas sólo hay dos capas, una de "entrada" y otra de "salida" y se conocen como "perceptrones". Estas redes sólo son capaces de discriminar relaciones lineales entre variables porque sólo poseen una capa de pesos. El perceptrón "multicapa" más sofisticado añade varias capas de unidades "ocultas" y, por lo tanto, los conjuntos de pesos aumentan el poder de la red para inferir relaciones no lineales entre variables(ver figura 13). No existe un límite teórico para el número de capas que puede tener una red, aunque estas dos están entre las más populares.

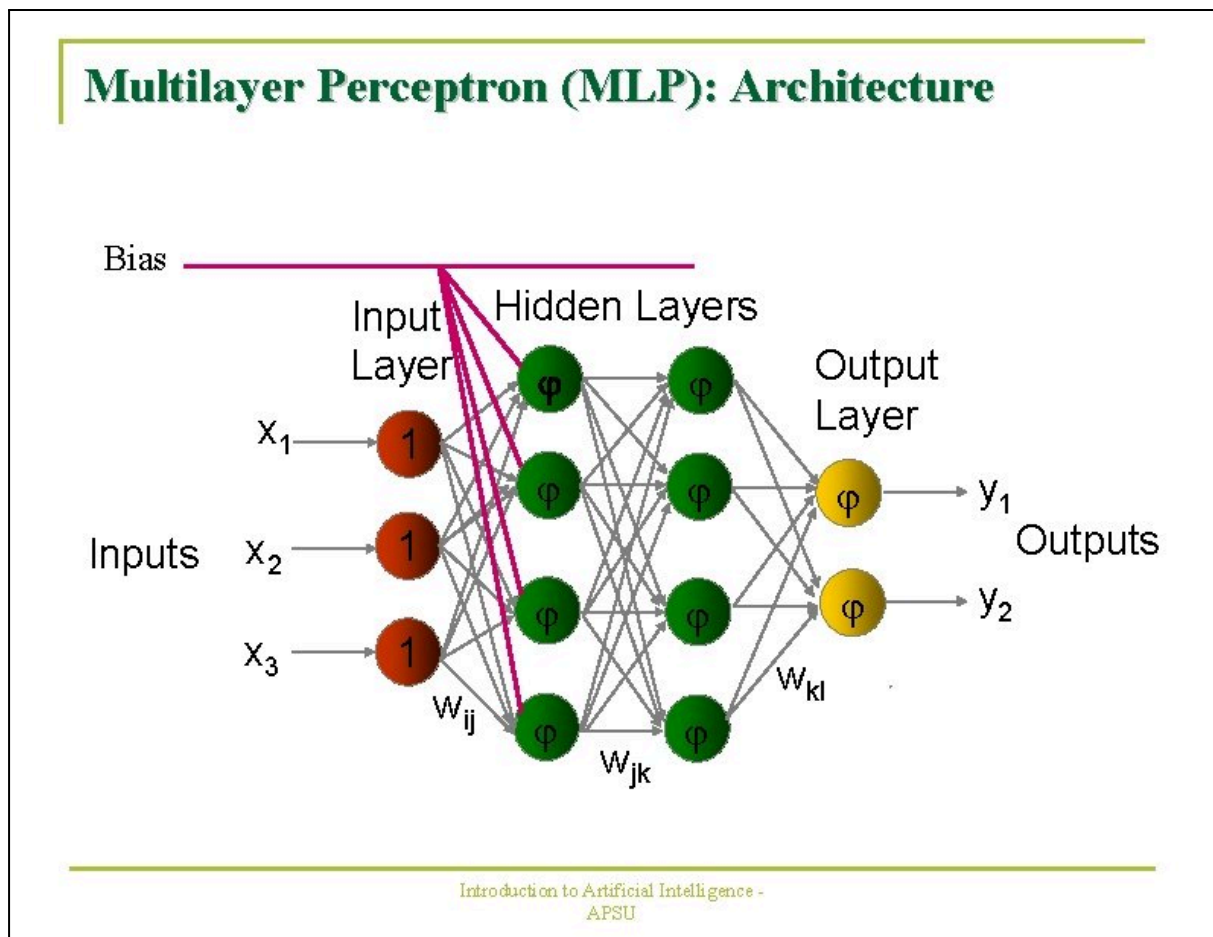


Figura 13- Arquitectura del algoritmo Neural Network Multilayer Perceptron. En esta imagen se explica la arquitectura del algoritmo de redes neuronales multicapa. Los datos a entrar ($x_1, x_2, x_3, x_4 \dots x_n$). La capa de entrada(globos de color rojo), dos capas profundas(globos en verde) y capa de salida (en amarillo).

La disposición de las unidades y los pesos en la red neuronal tiene un profundo efecto

en el rendimiento de la red e incluso en su propósito. Hay un gran número de arquitecturas que se han ideado para una variedad de propósitos, Aquí sólo se describen las arquitecturas más útiles en bioinformática. Las más utilizadas son las arquitecturas de retropropagación(Feed-Forward Backpropagation) hacia delante en las que las unidades se organizan en capas como se ha descrito anteriormente, y el algoritmo de aprendizaje se conoce como aprendizaje supervisado. El aspecto "feed-forward" del nombre de esta red se debe a la dirección del flujo de los datos, mientras que la retropropagación describe el hecho de que los errores en los que se incurre durante el aprendizaje se propagan de vuelta a través de la red.

4.2.3- Aprendizaje

En la mayoría de las aplicaciones de este tipo de técnica neuronal, la tarea de la red es relacionar las variables que recibe en la capa de entrada con algún comportamiento deseado en la capa de salida, presentando repetidamente los ejemplos a la red en un proceso conocido como entrenamiento. De forma análoga al aprendizaje en los bebés humanos, el entrenamiento de la red neuronal permite a la red determinar la respuesta correcta a los patrones de entrada que se le presentan. Una vez entrenada, la red neuronal debe ser capaz de predecir una salida a partir de un conjunto de entradas no vistas anteriormente. Este entrenamiento se conoce como aprendizaje supervisado porque la salida se conoce para los puntos de entrenamiento y, por lo tanto, se puede dar a la red la respuesta requerida durante el entrenamiento. Las redes neuronales supervisadas se diferencian del aprendizaje supervisado tradicional, como el utilizado por los árboles de identificación, en que el aprendizaje supervisado tradicional se ocupa sólo o principalmente de la clasificación y las redes neuronales supervisadas tienen una capacidad más general que ésta. Las redes neuronales pueden actuar también como transductores (convirtiendo una forma de entrada en otra forma de salida). Uno de los aspectos más importantes de la salida de una red neuronal es que puede tener un "valor real", mientras que los clasificadores tradicionales sólo pueden producir uno de los varios valores discretos que representan la clase a la que pertenece una muestra. Sin embargo, las redes neuronales también pueden utilizarse cuando no se conoce la respuesta requerida, por ejemplo, en tareas de agrupación de clústeres, con un enfoque no supervisado. Estas redes utilizan un entrenamiento basado únicamente en los datos de entrada, no tienen distinciones de salida ni capa oculta, y se utilizan con frecuencia en dominios en los que no se conoce la respuesta de salida. En las siguientes secciones se describen los componentes de la red neuronal, además de los regímenes de entrenamiento empleados en su uso.

4.2.4- Unidades y pesos

La unidad (también conocida como neurona o nodo) es el principal elemento de procesamiento de la red neuronal. Recibe un conjunto de señales de entrada y, en combinación con una función interna, convierte las señales de entrada en una señal de salida. La función interna puede ser tan simple como una función escalonada, o una transformación más compleja, como la función sigmoidea, comúnmente utilizada. Basándose en estas funciones, la función de activación toma una "decisión" sobre si disparar y propagar la señal más arriba en la red, basándose en el peso de las señales entrantes. La función escalonada garantiza una respuesta de "todo o nada", mientras que la función sigmoide produce una respuesta más graduada. La figura 7.1 muestra estas dos funciones en forma matemática y en forma de gráfico. La tarea de la función sigmoidea de la figura 7.1, por ejemplo, consiste en tomar los pesos (Weight) entrantes, comprobar si se supera un umbral interno y si lo es, calcular la salida como la función de 1 dividida por 1 más el exponente del valor que ha calculado. El efecto es producir una salida que varía entre 0 y 1, donde un gran número de valores negativos producen una salida que está cerca de 0 y grandes números de valores positivos entrantes producen una salida cercana a 1. Existen muchos otros tipos de funciones dentro de las neuronas para convertir los valores entrantes en un valor de salida, y estas funciones se denominan "funciones de transferencia" o "funciones de activación". Una de las principales características de las redes neuronales es que los enlaces entre los nodos están ponderados, por lo que la entrada a una neurona suele llamarse peso. Por tanto, aunque una neurona transmisora emita un "1", si el peso de un enlace que lleva ese valor a otra neurona es de 0,1 entonces la neurona receptora recibe 0,1. El principio importante es que, aunque las funciones en sí mismas son muy sencillas, cuando se ensamblan varias de ellas y se conectan con conexiones ponderadas, es posible realizar cálculos muy complejos. Las conexiones ponderadas propagan las señales de unidad a unidad modificando la fuerza de la señal según su peso. Los pesos se modifican en el proceso de entrenamiento y proporcionan gran parte de la capacidad de aprendizaje de la red.

4.2.5- Aprendizaje supervisado en redes neuronales

Como se ha descrito anteriormente, el aprendizaje supervisado es el proceso de aprender relaciones entre los datos de entrada y de salida. En este tipo de aprendizaje los datos de entrada se pasan a la capa de entrada, se propagan a través de las unidades y pesos de la red neuronal, y la respuesta de la red se compara con la respuesta requerida, que viene dictada por los datos de salida. La discrepancia entre ambas se calcula y la red realiza

entonces cambios en sus pesos internos para reducir el error la próxima vez que se presentan estos datos de entrada. Este proceso se repite para todos los datos de entrada y, una vez completado, constituye una "Epoch". La mayoría de las redes neuronales necesitan un número moderado (100-10 000+) de Epochs para aprender eficazmente las relaciones entre los datos de entrada y de salida. El aprendizaje supervisado puede aplicarse a muchos tipos diferentes de arquitectura, y permanece en gran medida sin cambios, independientemente del número de capas ocultas. Sin embargo, se utilizan dos algoritmos diferentes para los perceptrones y los perceptrones multicapa. El algoritmo se ocupa de calcular el error en los nodos de salida y de propagar este error por la red de capa en capa. El cálculo de este error es diferente para las capas de salida y las capas ocultas. La regla de aprendizaje más sencilla es la de la percepción de Rosenblatt y es la que se describe a continuación:

Regla de aprendizaje del perceptrón

1- Inicialice el entrenamiento con los pesos de la red y los umbrales de las unidades con valores aleatorios pequeños.

2- Presentar los datos de entrada ($i_1, i_2, i_3 \dots i_n$) y los datos de salida deseados (o) a la red para el aprendizaje.

3- Calcule la salida de la red mediante la expresión: $f(s) \left[\sum_{k=1}^n w_k(t) i_k(t) \right]$

4- Adapta los pesos:

(a) si es correcto $w_k(t + 1) = w_k(t)$,

(b) si la salida es 0 y debería ser 1 $w_k(t + 1) = w_k(t) + i_k(t)$,

(c) si la salida es 1 y debería ser 0 $w_k(t + 1) = w_k(t) - i_k(t)$,

donde $w_k(t)$ es el valor del peso en el momento t , $w_k(t + 1)$ es el valor del peso en tiempo $(t + 1)$ (es decir, después de la actualización) y $f(s)$ es la función utilizada para calcular la salida de la red (en este caso, la función escalón). Esta regla de aprendizaje garantiza que los pesos se modifiquen para que la próxima vez que se muestre este mismo patrón de entrada, los pesos modifiquen el comportamiento de la red para que se acerque más a la respuesta correcta. Esta es la regla de aprendizaje más sencilla y posteriormente se han añadido diversas modificaciones como la introducción de un término Δ para modificar los pesos los pesos más lentamente según la siguiente ecuación:

$$w_k(t + 1) = w_k(t) + \Delta i_k(t)$$

Esto es suficiente para permitir el aprendizaje en un perceptrón pero, como se ha comentado anteriormente es el más simple de las redes neuronales y sólo puede procesar interacciones lineales entre variables.

procesar interacciones lineales entre variables. La siguiente sección describe la regla necesaria para los perceptrones multicapa.

4.2.6- Retropropagación

Para arquitecturas más grandes con una "hidden layer(capa oculta)" se requiere una regla de aprendizaje más compleja. El perceptrón multicapa suele utilizar la función sigmoidea. Por lo tanto, cuando se calcula la diferencia entre los patrones de salida reales y los deseados los pesos deben cambiarse de acuerdo con la derivada de esta función, a diferencia del simple incremento/decremento cuando se considera la función escalonada(step function). Además, como se considera más de un peso(weight), la regla de aprendizaje debe tener en cuenta este hecho y contener algún método de retropropagación (backpropagation) del error a través de la red.

1- Inicializar los umbrales(thresholds) y los pesos(weight).

2 Presentar los datos de entrada ($i_1, i_2, i_3 \dots i_n$) y los datos de salida deseados a la red.

3 Calcule la salida de la red mediante la expresión $f\left[\sum_{k=1}^n w_{ki}i_k\right]$ para cada capa de la red. La salida de la última capa es el vector de los valores de salida.

4 Adapta los pesos, empezando por la capa de salida y moviéndose hacia atrás mediante la ecuación: $w_{ij}(t+1) = w_{ij}(t) + \delta p_j \cdot o_{pj}$ donde o_{pj} es la salida del nodo j para el patrón p y

δp_j se define como:

para la capa de salida: $\delta p_j = c o_{pj}(1 - o_{pj})(d_{pj} - o_{pj})$

para las capas ocultas: $\delta p_j = c o_{pj}(1 - o_{pj}) \sum_k \delta p_k \cdot w_{jk}$

donde d es la salida deseada, o es la salida real y c es una constante utilizada en la función sigmoidea. La suma mostrada para las capas ocultas es sobre los k nodos de la capa por encima de la capa actual para la que ya se habrá calculado el δp_j . Las ecuaciones anteriores, ilustradas en la figura incrementan (o disminuyen) los pesos de la red en función del error en la capa superior. En el caso de la capa de salida, el error se calcula directamente con la respuesta deseada, mientras que la capa oculta calcula su error basándose en el error ponderado propagado desde la capa de salida. Este proceso permite a la red adaptar sus pesos para corregir la diferencia entre su salida actual y la salida deseada. Cuando se aplica varias veces sobre todos los datos de entrada y salida, la red se reconfigura para ser más precisa y puede utilizarse para predecir el resultado de nuevos ejemplos. Este tipo de aprendizaje sólo es posible cuando se conoce la respuesta deseada.

respuesta deseada y, por lo tanto, el aprendizaje supervisado puede tener lugar. La respuesta deseada, sin embargo, no se conoce para algunos problemas y, por tanto, existen reglas de aprendizaje no supervisado para encontrar patrones interesantes en los datos y para dividirlos en clusters, lo que no se va a considerar aquí porque no los vamos a usar en este trabajo.

4.3- Gradient Boosting Trees

Gradient boosting o Potenciación del gradiente, es una técnica de aprendizaje automático utilizado para el análisis de la regresión y para problemas de clasificación estadística, el cual produce un modelo predictivo en forma de un conjunto de modelos de predicción débiles, típicamente árboles de decisión. Construye el modelo de forma escalonada como lo hacen otros métodos de boosting, y los generaliza permitiendo la optimización arbitraria de una función de pérdida diferenciable. La idea de la potenciación del gradiente fue originada en la observación realizada por Leo Breiman en donde el Boosting puede ser interpretado como un algoritmo de optimización en una función de coste adecuada. Posteriormente Jerome H. Friedman desarrolló algoritmos de aumento de gradiente de regresión explícita, simultáneamente con la perspectiva más general de potenciación del gradiente funcional de Llew Mason, Jonathan Baxter, Peter Bartlett y Marcus Frean. En sus últimos dos trabajos presentaron la visión abstracta de los algoritmos de potenciación como algoritmos iterativos de descenso de gradientes funcionales. Es decir, algoritmos que optimizan una *función de coste* sobre el espacio de función mediante la elección iterativa de una función (hipótesis débil) que apunta en la dirección del gradiente negativo. Esta visión de gradiente funcional de potenciación ha llevado al desarrollo de algoritmos de potenciación en muchas áreas del aprendizaje automático y estadísticas más allá de la regresión y la clasificación.

4.4- Random Forest

El Random Forest puede definirse como un método de aprendizaje automático versátil e inteligente que puede realizar tanto clasificaciones como tareas de regresión. También puede llevar a cabo métodos de reducción dimensional, valores atípicos, tratar valores perdidos y otros pasos de exploración de datos. Es una solución experta para la mayoría de los problemas. Se conoce como una forma de aprendizaje de conjunto, ya que un grupo de modelos débiles se combinan para formar este potente modelo.

Capítulo V- Hipótesis y objetivos

5.1- Justificación e impacto

La presencia de la hipotensión ortostática neurogénica en los pacientes con EP está asociado con una alta tasa de mortalidad que podría ser explicado por la consiguiente hipoperfusión cerebral que trae consigo. Masaki y colaboradores encontraron una asociación lineal significativa entre el cambio de la presión arterial sistólica de la posición de cubito a la posición de pie y alta tasa de mortalidad de los pacientes a los 4 años de seguimiento(Masaki et al. 1998). Incluso sin la presencia de la EP, la NOH puede constituir una fuente de causa de debilidad en pacientes mayores, lo que acarrea riesgo alto de caída y alta tasa de mortalidad a los 3 años de haber empezado la NOH si lleva un mal control de la presión arterial ortostática(Veronese et al. 2015), (Veronese et al. 2015; Shaw et al. 2019). Gibbons y colaboradores encontraron en un estudio publicado en el año 2015 que el 31% de los pacientes con OH retardada desarrolla alfasinucleinopatía y la tasa de mortalidad de los pacientes a los diez años de seguimiento fueron de 29% con respecto a 9% en los controles(Gibbons and Freeman 2005). La hipotensión ortostática neurogénica es una de las condiciones más frecuentes que aumenta la severidad y el mal pronóstico de la enfermedad de Parkinson como lo han demostrado muchos trabajos. Complicaciones asociadas a alteraciones cognitivas, motoras, y muerte son más frecuentes en los pacientes con EP que además tienen NOH. Hasta este momento, no existe un tratamiento muy efectivo que pueda controlar las complicaciones que provoca la NOH en pacientes con EP o que pudiera mejorar la calidad de vida de los pacientes. Además, son muy pocas las investigaciones en este campo que podrían demostrar las vías de señalización y/o las proteínas que participan en la ocurrencia de la NOH. Este estudio nos permitirá dilucidar el posible factor genético que explica la gravedad de la enfermedad de Parkinson en algunos pacientes con hipotensión ortostática neurogénica. Conociendo según estudios previos, las enfermedades cardíacas asociados a NOH constituyen de las principales causas de muerte en pacientes con enfermedad de Parkinson, este estudio nos permitirá comprender mucho mejor la fisiopatología, la relación genética y molecular entre la EP y la hipotensión ortostática pero también nos permitirá establecer y predecir los pacientes con EP con riesgo de desarrollar NOH. También se podrán descubrir nuevos marcadores y nuevas dianas terapéuticas contra la NOH en pacientes con EP.

5.2- Hipótesis de trabajo

Presumimos que existe algún factor genético y molecular que vincula la enfermedad de Parkinson (EP) y la aparición y progresión de la hipotensión ortostática neurogénica(NOH). También pensamos que una alteración genética alteraría rutas metabólicas y de señalización específicas, especialmente la expresión de los productos génicos que

intervienen en estas rutas. Por lo tanto, pensamos que un mapeo de eQTL (*Expression quantitative trait loci*) basado en la expresión de genes en pacientes con alteración genética asociada a NOH, nos permitiría estudiar las vías de señalización que intervienen y específicamente las proteínas implicadas en la enfermedad.

5.3- Objetivo general

Evaluar los efectos de los polimorfismos de nucleótidos únicos (SNPs) y sus mecanismos moleculares asociados a hipotensión ortostática neurogénica sobre la severidad y el pronóstico de la enfermedad de Parkinson.

5.4- Objetivos específicos

- Evaluar los pacientes con NOH según signos clínicos como la prueba de presión arterial ortostática (hipotensión ortostática) utilizando los datos disponibles en la base de datos PPMI.
- Evaluar la asociación entre variantes de nucleótidos únicos y la NOH en enfermos de Parkinson.
- Estudiar la asociación entre la severidad de la enfermedad de Parkinson y los SNPs asociados a la NOH.
- Estudiar los mecanismos moleculares usados por los SNPs relevantes del estudio para participar en el desarrollo de la NOH.
- Evaluar algoritmos de aprendizaje automático que mejor predicen la aparición de la NOH y pronóstico de los pacientes con Enfermedad de Parkinson.

PARTE II

Capítulo VI- Métodos y materiales

6.1- Base de datos del PPMI

La Fundación Michael J. Fox puso en marcha la Iniciativa de Marcadores de Progresión del Parkinson (PPMI, por sus siglas en inglés) en 2010 para comprender mejor el inicio y la

progresión de la enfermedad de Parkinson y acelerar así el desarrollo de nuevos tratamientos (Parkinson Progression Marker Initiative 2011). Desde entonces, se han mejorado los conocimientos que los científicos tienen acerca de la enfermedad y cómo intentar detenerla. La PPMI es un estudio longitudinal, observacional, multicéntrico de la historia natural de la enfermedad de parkinson. La PPMI evalúa la progresión de las características clínicas, los resultados de las imágenes, los marcadores biológicos y genéticos, y los resultados digitales de la enfermedad de Parkinson en todas las etapas de la EP, desde la EP prodrómica hasta la moderada y severa. El objetivo general del PPMI es identificar marcadores de progresión de la enfermedad para acelerar los ensayos terapéuticos de tal manera que los científicos pueden intervenir en la reducción de la progresión de la EP y evitar las complicaciones y discapacidad asociadas a la EP para de esta forma mejorar la calidad de vida de los pacientes e incluso eliminar definitivamente la EP si se diese el caso (Parkinson Progression Marker Initiative 2011).

El protocolo clínico PPMI está diseñado para adquirir datos longitudinales exhaustivos de aproximadamente 4000 participantes inscritos en unos 50 centros de todo el mundo. Se recuperan datos clínicos, demográficos, antecedentes patológicos personales y familiares, datos bioquímicos, genómicos, transcriptómicos, proteómicos y metabolómicos. Los pacientes se consultan de forma periódica (ver figura 14).

Study Synopsis - PPMI Clinical

Study Population* Target N = 4000+	Clinical Data Collection	Biologic Collection
1,000 Enrolled 2010-2018	• Motor Assessments • Neurobehavioral/Neuropsychiatric Testing • Autonomic, Olfaction, Sleep Testing • DaTSCAN, MRI, and other imaging sub-studies • Digital Data • Online patient reported outcomes (PROs)	• DNA, RNA • Whole Blood, Serum, Plasma, Urine • CSF • iPSC in Subset of 2010 - 2018 • Skin biopsy • Post-mortem tissue
500 PD		
100 HC		
400 Prodromal PD		
3,000 Enrollment Goal 2020-2023		
900 PD		
100 HC		
2000 Prodromal PD		

*Sample size numbers are approximate.

To view information on ClinicalTrials.gov, click [here for PPMI Clinical](#) and [here for PPMI Online](#).

Figura 14- Representación esquemática de los datos del PPMI. En esta figura se presenta un resumen de la base de datos del PPMI. En la primera columna a la izquierda se presenta la cantidad de pacientes incluidos en los cohortes desde 2010 hasta 2018 (500 PD: 500 pacientes con enfermedad de parkinson, 100HC: 100 controles sin enfermedad de parkinson y 400 pacientes en estado prodrómico). Los objetivos para 2023 para este entonces eran 900 pacientes con enfermedad de parkinson y 100 controles y 2000 en estado prodrómico. En la segunda columna de la tabla, se presentan los tipos de datos clínicos: datos sobre la evaluación motoras, prueba neuropsiquiátrica y neurocomportamental, prueba de sueño, olfacción,

autonómica y datos imagenológicos como DATScan MRI funcional y estructural. La imagen y las informaciones incluidas se obtuvieron de la base de datos del PPMI. WWW.PPMI.ORG.

6.2- Aprobaciones de protocolos estándar y consentimientos de los pacientes

Los sujetos participantes se derivan de la Iniciativa de marcadores de progresión de Parkinson (base de datos del PPMI)(Parkinson's Progression Marker Initiative 2011). Todos los participantes han acordado firmar el consentimiento informado que indica que están de acuerdo en utilizar sus datos genómicos para fines de investigación, incluido el intercambio de datos con otros centros de investigación. Todo esto tiene lugar antes del inicio de la inscripción del paciente en el estudio. Los datos obtenidos con la aprobación de las autoridades correspondientes, por supuesto, no contienen los datos de identificación de los participantes(Parkinson Progression Marker Initiative 2011).

https://www.ppmi-info.org/sites/default/files/docs/PA2_PPMI_Clinical%20Protocol_Final_01Feb2021.EPf.

6.3- Participantes

La Iniciativa de Marcadores de Progresión del Parkinson (PPMI, por sus siglas en inglés) es un estudio observacional multicéntrico en curso centrado en la identificación de biomarcadores de la enfermedad en pacientes con EP que asisten a centros clínicos desde América hasta Oceanía(Parkinson Progression Marker Initiative 2011). El protocolo es aprobado por la junta de revisión de cada centro, y todos los participantes firman un consentimiento informado por escrito. La información obtenida se comparte a ciegas con los investigadores involucrados y no involucrados. Sólo se extrajo información de la visita inicial de cada participante(datos baseline). Trescientos cuatro participantes con EP de 30 años o más fueron seleccionados de la base de datos del PPMI. Se excluyeron los pacientes con condiciones patológicas y/o medicamentos asociados con síndromes ortostáticos, incluyendo diabetes, alcoholismo, polineuropatía, amiloidosis, y usuarios de bloqueadores de los canales de calcio, diuréticos, inhibidores de la enzima convertidora de angiotensina, bloqueadores del receptor AT-1, nitratos, antagonistas de los α -adrenoreceptores, opioides o antidepresivos tricíclicos.

6.4- Medida de la presión arterial y clasificación de los pacientes en NOH o no NOH

La presión arterial se midió de 1 a 3 minutos después de descansar en posición supina, y de 1 a 3 minutos después de ponerse de pie(16). La NOH se definió como una caída de la PA sistólica de ≥ 20 mmHg (o ≥ 30 mm Hg en pacientes hipertensos) y/o una caída de la PA diastólica de ≥ 10 mmHg al cabo de 1-3 minutos de estar de pie(Lahrman et al. 2006; Freeman et al. 2011).

6.5- Procesamiento de los datos genómicos y análisis genética

La secuenciación del genoma completo se llevó a cabo mediante un secuenciador MacroGen Inc. (www.macrogen.com) de muestras de ADN extraídas de sangre completa. Un microgramo de cada muestra de ADN se fragmentó con el sistema Covaris y se preparó de acuerdo con la guía de preparación de muestras de ADN de Illumina TruSeq para obtener una biblioteca final con un tamaño medio de inserción de 300-400 pb. Las bibliotecas se multiplexan y secuenciaron en la plataforma Illumina HiSeq. Las secuencias de las lecturas pareadas se alinearon inicialmente con el genoma GRCh37-hs37d5 utilizando bwa-mem(v0.7.13). Se utilizó la herramienta Bamsortdup2 para filtrar los duplicados y ordenar los archivos bam alineados. Después de filtrar las secuencias de lectura duplicadas, las lecturas fueron alineadas y calibradas utilizando el pipeline de GATK(v3.5). Se utilizó el llamador de haplotipos en la línea de producción de GATK para llamar a las variantes, incluidas las variantes de un solo nucleótido (SNV) y las pequeñas InDels para generar los VCF del genoma. Utilizando los archivos VCF de la cohorte alineada hg38 de los datos de secuenciación del genoma completo, se extrajo la información del genotipo utilizando BCFtools y PLINK. Los alelos de las 72 variantes seleccionadas asociadas a la enfermedad de Parkinson para los sujetos de PPMI con datos de secuenciación del genoma completo disponibles fueron compilados en un archivo CSV para permitir el examen rápido de estas variantes en el conjunto de datos. Nos centramos en los SNP con una tasa de llamada mínima del 95%, una frecuencia alélica menor (MAF) $> 1\%$, y valores p de equilibrio de Hardy-Weinberg > 0.05 .

Usamos los modelos aditivos recesivos y dominantes para encontrar SNPs asociados a NOH. El modelo recesivo nos permitirá descubrir los SNPs que están asociados con la NOH solamente cuando están presentes los dos alelos menores en un locus en este caso en los pacientes homocigotos para este alelo(H.-M. Liu et al. 2021). Mientras tanto, el modelo dominante nos permite encontrar los SNPs asociados con NOH cuando está presente una sola copia del alelo de riesgo(H.-M. Liu et al. 2021). Entonces, los datos fueron codificado

de manera tal que en el modelo recesivo los genotipos homocigota del alelo que no se asocian a la enfermedad junto con el genotipo heterocigota se computo contra el genotipo homocigota del alelo que se asocia a la NOH. En el caso del modelo dominante, el genotipo homocigota del alelo que se asocia a la NOH junto con el genotipo heterocigota se computa contra el genotipo homocigota del alelo que no se asocia con la NOH(Bagos 2013). Usando cada uno de los modelos genéticos contruidos, hicimos análisis de regresión logística tanto simple para cada SNP, como múltiple donde se consideran los SNPs en conjunto(Bagos 2013; Hayes 2013), (M. H. Wang, Cordell, and Van Steen 2019). Además, chi cuadrado y test exacta de Fisher se computaron también para estudiar las asociaciones entre SNP y NOH. En la regresion logistica, el modelo se ajusta por el sexo, la edad y toma de medicamentos antiparkinsonianos(M. H. Wang, Cordell, and Van Steen 2019; Cao et al. 2020).

Además, hicimos análisis de epistasis para estudiar el efecto en conjunto de dos genes sobre la aparición de la NOH en pacientes con EP. Computamos en cada grupo de estudio(NOH y no-NOH) de forma individual la asociación de dos SNPs con NOH o con no-NOH y después se comparan los resultados entre ambos grupos. Se realizó prueba de chi cuadrado y análisis de regresión logística simple. Se consideró un p-valor <0.05 como significativo. Se aceptan los grupos de SNPs con efecto epistático cuando aparece significativo en el grupo NOH mientras tanto es no significativo en el grupo no-NOH. Se tuvo en cuenta los modelos genéticos aditivo, dominante y recesivo en el análisis de epistasis donde computamos modelo dominante contra aditivo, aditivo contra aditivo y aditivo contra recesivo."Knime Analytics platform" y R-4.2.2 fueron usados para realizar este análisis.

6.7- Preparación, obtención y mantenimiento de células dopaminérgicas.

Se llevaron a cabo cultivos de líneas iPSC en placas recubiertas de matrigel (Corning) utilizando medios Essential 8 Flex (E8, lotes 1, 2 y 3) o Essential 6 (E6, lotes 4 y 5) de Gibco durante aproximadamente un mes (5 pases). Se agregaron factores de crecimiento al medio Essential 6 para igualar su composición a la del Essential 8. Una vez que las células alcanzaron el 70-80% de confluencia, se disociaron en una suspensión unicelular con Accutase (Gibco) y se sembraron a una densidad de 200.000 células/cm² en placas de un pocillo recubiertas de matrigel (con código de barras, Greiner) adecuadas para el cultivo celular automatizado. Las células se mantuvieron en cultivo hasta que cubrieron la superficie de la placa, lo que normalmente ocurrió 24-48 horas después de la siembra unicelular. El tiempo necesario para alcanzar la confluencia fue variable y dependió de la tasa de crecimiento de cada línea iPSC. El protocolo de diferenciación DA se adaptó según el trabajo de Kriks y colaboradores(Kriks et al. 2011) con pequeñas modificaciones. Las

diferenciaciones se llevaron a cabo en un sistema de cultivo celular automatizado (Dhingra et al. 2020) con repeticiones manuales en los días 25 y 32 para la diferenciación final y la inmunocitoquímica (ICC), respectivamente (Dhingra et al. 2020). Las muestras para los ensayos se recogieron en diferentes días: 0 (iPSC), 25 (punto medio) y 65 (neuronas DA). Para los ensayos de ADN, las células se disocian con Accutase, se lavan con PBS y se centrifugan a 200 g. El sedimento celular se ultracongela o se procesa de acuerdo con los protocolos de ensayo. La mayoría de los ensayos de ARN requirieron células ultracongeladas recogidas raspando la superficie de la placa con PBS o tampón de lisis. Los ensayos de ARN-seq unicelular (sc) y scATAC-seq requirieron una suspensión unicelular preparada en albúmina sérica humana (HSA)/PBS al 0,04%. Todas las muestras se almacenaron a -80 °C hasta su posterior procesamiento. Para la criopreservación, los precursores de neuronas DA del día 25 se disocian con Accutase, se lavan con medio neurobasal (Gibco), se resuspenden en medio de criopreservación Synth-a-Freeze frío (Gibco) suplementado con 10 μ M Y-27632 y se alicuotan en crioviales con código de barras (NovaStora) a 10×10^6 células/ml/vial (en hielo). Los crioviales se colocan en contenedores de congelación celular CoolCell (Biocision), se mantienen toda la noche a -80 °C y se transfieren a nitrógeno líquido para su almacenamiento a largo plazo.

6.8- Análisis de expresión de genes basado en desequilibrio de ligamiento

Estudiamos los efectos de los SNPs asociados con NOH en la expresión de otros genes mediante un análisis cis-eQTL realizado utilizando la plataforma LDlink (Machiela and Chanock 2015), (Alexander and Machiela 2020), (Alexander and Machiela 2020; Myers, Chanock, and Machiela 2020). Los genes que mostraron un R-cuadrado mayor a 0.8 y un D' mayor a 0.9 fueron considerados estar en desequilibrio de ligamiento en este estudio (Bagos 2013; Hayes 2013). En este estudio se consideraron las poblaciones americana, europea y latina por ser las más representativas en la base de datos del PPMI. El análisis se hizo en datos RNA seq de 13 regiones del tejido cerebral y se consideraron los SNPs en una ventana de 500.000 pb al SNP query.

6.9- Análisis de enriquecimiento de SNPs asociados a NOH

Los SNPs relacionados con la EP asociados con NOH y aquellos en LD con ellos fueron el sustrato de un análisis de enriquecimiento. Este análisis nos permitió explorar la vía metabólica o de señalización en la que participaban los genes que contenían los SNPs asociados a NOH. Para este análisis utilizamos el software SNPnexus (Oscanova et al. 2020), (Chelala, Khan, and Lemoine 2009). Consideramos como enriquecidas estadísticamente las vías que tienen un valor $p < 0,05$.

6.10- Análisis estadísticas

Las variables numéricas se describieron como media $\pm$ desviación estándar y las categóricas con porcentajes. Las diferencias entre los pacientes de EP con NOH o no fueron analizadas por la prueba T o la prueba Chi-cuadrado. Usamos un modelo de regresión logística para encontrar los SNPs relacionados con la EP que estaban independientemente asociados con la NOH, ajustando por sexo, edad y consumo de medicamentos antiparkinsonianos. Este análisis fue realizado en la plataforma KNIME 4.1.0 (Knime AG, Zurich, Suiza). Todos los SNPs con valores $p < 0,05$ se seleccionaron para los análisis de anotación, LD, enriquecimiento y eQTL, como se ha descrito anteriormente.

6.11- Análisis de expresión de los genes según los alelos de los SNPs asociados a NOH(eQTL).

Los datos de ARN-seq normalizados fueron obtenidos de la base de datos del PPMI. Dichos datos fueron obtenidos a partir de células madres pluripotenciales inducidas (iPSCs) provenientes de células sanguíneas de algunos pacientes del programa. Las células madres pluripotenciales fueron inducidas para la obtención de células dopaminérgicas que se utilizan como modelos celulares para la extracción del ARN y estudios de ARNseq como lo explicaremos a continuación. El RNA-seq de todo el transcriptoma permite la detección de muchos tipos de RNA, incluyendo el lncARN y el circARN, además del ARN del exoma que codifican para proteínas. La preparación de la biblioteca se realizó con el kit SMARTer Stranded Total RNA Sample Prep Kit-HI Mammalian, que tiene la primera hebra como la hebra positiva. Bibliotecas de 100 pb x 100 pb fueron secuenciadas en un NovaSeq 6000. Los archivos fastqs obtenidos fueron demultiplexados y se utilizaron para cuantificar el ARN codificante, lncRNA y otros ARN no codificantes. Los datos de conteo de genes fueron obtenidos en formatos de texto de la base de datos del PPMI. Los últimos fueron transformados para obtener solamente los conteos del día 65 excluyéndose los días 0 y 25 porque, teóricamente las células a los 65 días ya se transforman en células dopaminérgicas. En primer lugar hicimos un análisis de reducción de dimensión eliminando los datos con varianza muy bajos utilizando KNIME v4.1.0 con los datos previamente normalizados con un sistema de normalización entre 0 y 1 que impide que se involucre la varianza y no se falsifique a la hora de excluir datos basando en la varianza. Después hicimos un análisis de componentes principales y extrajimos los componentes que mejor explican la varianza de los datos. La idea en este análisis de eQTL es mapear la expresión de los genes a los SNPs

encontrados asociados a NOH para ver si los alelos tienen algún efecto sobre la expresión de dichos genes. Para eso, realizamos un análisis de regresión lineal simple donde la expresión del gen es la variable dependiente y cada SNP la variable explicativa en un modelo aditivo. También, realizamos este mismo análisis utilizando modelos genéticos dominantes y recesivos.

6.12- Identificación de módulos y genes Hub de los eQTL en análisis de network.

Los eQTLs encontrados en el análisis anterior se usan para hacer un análisis de redes (*network analysis*) con el objetivo de encontrar genes Hub, vías metabólicas/señalización y procesos celulares claves en el desarrollo o la prevención de la NOH. En un análisis de redes, el *gen hub* es aquel que tiene altas conexiones en un módulo destacado y que puede tener importancia significativa en la enfermedad en estudio. Es decir mayor correlación un gen establece con otros genes, mayor se tiende a ser un *gen hub*. Se hizo un análisis de correlación de pearson con los datos previamente normalizados y reducidos por PCA. Los genes con p-valor menor a 0.05 y coeficiente de correlación menor a -0.5 y mayor a 0.5 fueron filtrados para posteriores análisis. El *network* fue construido en el software Cytoscape v3.9.1. Se hizo un análisis estadístico del *network* para determinar los nodos con mayor importancia. Además análisis de clustering se llevó a cabo para determinar módulos y cluster dentro del network. Además, los mismos genes del network fueron usados para extraer los datos de la base de datos STRING para hacer un análisis de interacción proteína-proteína. Se compararon ambos network para extraer los genes hub del PPI network. Al final los hub genes extraídos fueron validados haciendo una comparación en la expresión de los genes en los diferentes genotipos por un análisis de la varianza ANOVA(p valor<0.05).

6.13-Análisis de enriquecimiento de los eQTLs y genes HUB

Los eQTLs y los genes hub estadísticamente significativos fueron usados para hacer análisis de enriquecimiento para estudiar las vías metabólicas/señalización y procesos moleculares en los cuales participan los mismos. Para este análisis usamos la plataforma ShinyGo(Ge, Jung, and Yao 2020). Fueron considerados significativos enriquecidos los procesos moleculares y vías metabólicas/señalización con False Discovery rate(FDR)<0,05.

6.14- Análisis de regresión lineal entre los miembros de módulos

El módulo principal que hemos considerado estar asociado con la NOH se usa para hacer análisis de regresión lineal simple. Cada uno de los genes se regresan contra otro. Con este análisis, pretendemos analizar los genes que están relacionados linealmente dentro del módulo.

6.15- Análisis de la asociación entre los SNPs asociados a NOH y los datos clínicos.

Se realizaron análisis de regresión lineal entre los datos clínicos y los SNPs encontrados asociados a NOH, para estudiar los SNPs asociados a los mismos. Los datos como puntaje de MDS-UPDRS I, II, III, IV; MOCA, escala de Hoehn & Yahr, para estudiar los SNPs asociados a los mismos. Con estos análisis, pretendemos ver si algunos SNPs asociados a NOH están relacionados con la progresión de la enfermedad de Parkinson.

6.16- Análisis de predicción de NOH en enfermos de parkinson usando algoritmo de aprendizaje automático.

Datos demograficos como edad sexo, y clínicos como valores de MDS-UPDRS, Moca y escala de Hoehn & Yahr, edad de inicio, toma de medicamentos antiparkinsonianos junto con los SNPs encontrados asociados a NOH fueron utilizados para hacer análisis de predicción de la NOH en pacientes con enfermedad de Parkinson. Los algoritmos de aprendizaje automático support vector machine, redes neuronales y otros algoritmos de predicción. Después Hicimos una comparación entre los diferentes resultados obtenidos.

PARTE III

Resultados

Capítulo VII- Análisis de asociación entre SNPs y NOH

7.1- Asociación entre NOH y SNP en datos de primera visita.

7.1.1- Análisis de asociación por regresión logística simple, prueba de Chi cuadrado y/o prueba exacta de Fisher.

Para el análisis de asociación entre la NOH y los SNPs relacionados con EP, hemos realizado una regresión logística simple sobre un modelo aditivo de los datos genotípicos. Hemos encontrado que el alelo menor del SNPs rs34311866 localizado en el gen TMEM175 está asociado con NOH(OR: 3.4, p valor: 0.0347). El genotipo heterocigota del SNP rs3910105 localizado en el gen SNCA está asociado con NOH con un OR de 3.18 y p valor de 0.0248, mientras que el alelo menor con un OR de 3.07 y p valor de 0.0602; aunque en el análisis alélico, hemos encontrado una asociación con el alelo menor(p valor: 0.05) ([Ver tabla 3 y 4](#)). El genotipo heterocigota del SNP rs104893877 del gen SNCA-A53T está asociado con NOH con OR de 4.32 y un p-valor de 0.0227. El genotipo heterocigota del SNP rs329648 localizado en el gen MIR4697 está asociado con NOH con un OR de 2.41 p valor de 0.0396. El genotipo heterocigota del SNP rs12456492 está asociado con NOH con OR de 2.046 y p-valor de 0.0546. El alelo menor del SNP rs4818 localizado en el gen COMT con un OR de 2.7 y pv de 0.0554([Ver tabla 3 y 4](#)). En el análisis de regresión logística simple; de los 72 SNPs involucrados en el estudio, hemos encontrado que 9 SNPs aumentan el riesgo de NOH mientras que un solo SNP disminuye el riesgo de NOH en pacientes con Enfermedad de Parkinson([Ver tabla 3 y tabla 4](#)).

SNPs name	Alelo	Gene	Odd ratio	P-value
rs353116	menor	SCN3A.SCN2A	2,45	0,064*
rs34311866	menor	TMEM175	3,4	0,0347**
rs3910105	menor	SNCA	3.07	0,0602*

rs3910105	Hetero	SNCA	3,18	0,0248**
rs104893877	Hetero	SNCA A53T	4,32	0,0227**
rs2280104	menor	BIN3	0,27	0,0983*
rs329648	Hetero	MIR4697	2,41	0,0396**
rs11060180	menor	OGFOD2.CCDC62	2,48	0,0672*
rs12456492	Hetero	SYT4.RIT2	2,046	0,0546*
rs4818	menor	COMT	2,7	0,0554*

Tabla 3- SNPs asociados a NOH en una regresión logística simple y un modelo genético aditivo. En esta tabla se presentan los resultados de la regresión logística simple usando los datos codificados para un modelo genético aditivo. Primera columna: ID de los SNPs, segunda columna: genotipos y alelo asociados a la NOH. Tercera columna: nombre del gen donde se localiza el SNP. La cuarta columna: Odd ratio que informa sobre la fuerza de asociación a la NOH. La quinta columna: el P-valor nominal.

***Solo existen datos para alelo mayor y heterocigota.

** Existe asociación con P valor<0.05

* No existe asociación. P-valor con tendencia a la significancia, por lo que se considera para posteriores análisis.

Posteriormente, hemos realizado un análisis usando los diferentes modelos aditivo, recesivo y dominante. Computamos, el OR, el test de chi cuadrado, test de fisher en caso necesario para cada uno de los SNPs. Hemos encontrado para el SNP rs3910105, que en el modelo aditivo que existe asociación entre NOH y dicho SNP estadísticamente significativo (OR: 3.06, Chisq: 5.7, p-valor :0.05). También para el mismo SNP, el modelo dominante demuestra una asociación estadísticamente significativa (OR: 3.045, Chisq: 3.8, p-valor: 0.02); lo que significa que con solo tener una copia del alelo menor es suficiente para presentar el riesgo. Por eso, en el análisis aditivo, y dominante, hemos encontrado asociación significaba pero en el análisis recesivo no.

En el modelo tanto aditivo, dominante, el SNP **rs104893877** se encuentra asociado con NOH con OR de 4.42, 4.42, 4.2 respectivamente con pv < 0.05 mientras que en el modelo recesivo no existe asociación evidente. Entonces, también en tal caso con solo una copia del alelo menor es suficiente para la presencia del riesgo del NOH en dichos pacientes. El genotipo heterocigota del SNP **rs329648** localizado en el gen MIR4697 se encontró asociado con la NOH con OR de 3.06 y p-valor de 0.047 en el modelo aditivo. Mientras que el modelo recesivo ni dominante no existe asociación. Otro SNP encontrado asociado con NOH(OR: 2,136, pv:0,036) en el modelo aditivo es el **rs12456492** localizado en el gen SYT4.RIT2. Este caso es parecido al anterior donde no tiene asociación en los otros modelos(dominante ni recesivo). El SNP **rs34311866** se asocia con NOH en el modelo recesivo con OR de 3.2 y p-valor 0.04. Además se compuso asociación con este mismo SNP en el modelo alélico con

test de fisher de 0.06 y OR de 1.67. En el análisis de regresión logística se ve claramente que el alelo menor de este SNP está asociado con NOH(Ver [tabla 3](#)).

Estos análisis nos permitieron estudiar la asociación de cada SNP con el NOH por separado. Además hemos pensado en estudiar la asociación de los SNPs a la NOH en una regresión logística múltiple usando los mismos modelos.

SNP-name	Gene	Modelo	Alelos	Combinación Alélica	OR	RR	Xsq	P-value χ^2	Fisher
rs3910105	SNCA	Aditivo	T:C/C	CC vs CT vs TT	3.06	1.42	5.70	0.05	0.04
rs3910105	SNCA	Dominante	T:C/C	TC+CC vs TT	3.04	1.31	3.80	0.02	0.02
rs3910105	SNCA	Alélico	T:C/C	T vs C	1.67	1.28	3.55	0.05	0.05
rs104893877	SNCA-A5 3T	Aditivo	C:T/T	CC vs TT vs CT	4.42	3.98	4.10	0.043	0.03
rs104893877	SNCA-A5 3T	Dominante	C:T/T	TT+CT vs CC	4.42	3.98	4.10	0.043	0.03
rs104893877	SNCA-A5 3T	Alélico	C:T/T	C vs T	4.17	1.04	3.90	0.04	0.03
rs329648	MIR4697	Aditivo	T:C/T	TT vs CT vs CC	2.46	1.38	6.00	0.05	0.05
rs329648	MIR4697	Alélico	T:C/T	T vs C	1.21	1.08	0.37	0.54	0.50
rs12456492	SYT4.RIT 2	Aditivo	A:G/G	AA vs AG vs GG	2.14	1.38	6.76	0.03	0.03
rs353006	SCN3A.S CN2A	Recesivo	C:T/T	CC+CT vs TT	2.44	2.00	3.46	0.06	0.06
rs34311866	TMEM175	Recesivo	C:T-T	CC+CT vs TT	3.20	2.84	3.21	0.07	0.04

Tabla 4- Análisis de asociación por regresión logística simple . Modelo aditivo, recesivo y dominante. Primera columna: ID de los SNPs, segunda columna: nombre de los genes donde se localizan los SNPs, tercera columna: Modelo genético en el cual el SNP muestra asociación con NOH. Cuarta columna: los alelos del locus(A:B/C), el primero(A) es el alelo mayor(Mayor frecuencia alélica), el segundo es el alelo menor y el tercero es el alelo de riesgo(Alelo asociado con la NOH). Quinta columna:

codificación de los datos según los modelos genéticos. Sexta columna: Odd Ratio. Séptima columna: Riesgo relativo. Octava columna: Chi cuadrada. Novena columna: P-valor del chi cuadrado. Décima columna: P-valor para test de Fisher.

7.1.2- Análisis de asociación por regresión logística múltiple en modelo aditivo, dominante y recesivo.

Hicimos un análisis de regresión logística múltiple con un modelo aditivo. Hemos encontrado 15 SNPs asociados a NOH. De los cuales 5 coinciden con los resultados encontrados en los análisis simples del modelo aditivo (tanto regresión logística simple como chi cuadrada). Lo que nos lleva a pensar que algunos SNPs pueden estar asociados con la NOH en un formato poligénico. Si bien no tienen fuerza suficiente para la aparición del fenotipo cuando están solos pero cuando se asocian con otros SNPs juntos influyen la presencia del fenotipo. Por esa razón se nos ocurrió la idea de hacer un análisis de epistasis que consiste en la asociación de dos o más SNPs en la aparición de la NOH (Ver análisis epistasis en el apartado siguiente). En el análisis de regresión logística múltiple con los datos siguiendo un modelo dominante, encontramos que los SNPs rs34311866-TMEM175, rs104893877-SNCA, rs353116-SCN2A/SCN3A, rs319648-MIR4697, rs3910105-SNCA, rs13194100-SH3GL2 aumentan el riesgo de aparición de la NOH en pacientes con enfermedad de Parkinson mientras tanto los SNPs rs1694518-ELOVL7, rs34043159-IL1R1, rs10906913-FAM171A1/ITGA8 disminuye el riesgo de NOH (Ver tabla 6).

Por otro lado en un análisis de regresión logística múltiple donde los datos se codifican en un modelo recesivo, hemos encontrado que los SNPs rs34300866(TMEM175), rs34884107-TMEM175, rs11060180-OGFOD2, rs6801093-FAM47E, rs1180004-BIN3, rs4444903-EGF disminuye el riesgo de NOH (Ver tabla 7). Los SNPs rs319648-MIR4697, rs3431186-TMEM175, rs3910105-SNCA, rs353116-SCN2A, rs104893877-SNCA-A53T, rs13194100-SH3GL2, rs130188-GBA-E365K, rs11060180-CDC62 aumentan el riesgo de padecer la NOH en los pacientes mientras que los rs1694528-NDUFAF2, rs34043159-MAP4K4, rs10906913-ITGA8 disminuyen el riesgo en un modelo dominante.

SNP-name	Gene	Odd Ratio	p-value
rs34311866	TMEM175	9.97	1.30E-4

rs11060180	OGFOD2.CCDC62	4.01	4.32E-3
rs6812193	FAM47E.STBD1	4.53	6.32E-3
rs104893877	SNCA_A53T	73.69	9.44E-3
rs353116	SCN3A.SCN2A	3.52	1.30E-2
rs13294100	SH3GL2	3.97	1.33E-2
rs3910105	SNCA	3.49	1.67E-2
rs10797576	SIPA1L2	0.17	2.15E-2
rs10906923	FAM171A1.ITGA8	0.33	2.70E-2
rs329648	MIR4697	3.45	2.90E-2
rs34884217	TMEM175	4.18	5.60E-2
rs2694528	ELOVL7.NDUFAF2	0.15	0.07
rs6430538	ACMSD.TMEM163	0.39	0.07
rs34637584	LRRK2_G2019S	0.17	0.08
rs11158026	GCH1	0.41	0.08
rs55785911	DDRKG1	2.19	0.09

Tabla 5- Análisis de asociación por regresión logística múltiple. Modelo aditivo. Primera columna están los ID de los SNPs, la segunda columna los genes en los cuales o cerca de los cuales se localizan los SNPs. En la columna tres se encuentran los valores de los Odd Ratios y en la cuarta columna los P-valores. Existen algunos SNPs que no son significativos: tienen P-valores que van desde 0.07 hasta 0.09 pero tienen la tendencia hacia la significancia teniendo en cuenta la muestra relativamente pequeña que se usa para el estudio, los incluimos para comparar con los resultados de otros modelos genéticos (recesivo y dominante).

Nombre de SNPs	Gene	Odd Ratio	P value
rs319648	MIR4697	7.24	38E-3
rs34311866	TMEM175	4.62	1.17E-2
rs3910105	SNCA	6.95	1.45E-2
rs1694528	ELOVL7.NDUFAF2	0.14	2.62E-2
rs34043159	IL1R1.MAP4K4	0.29	4.70E-2
rs353116	SCN2A.SCN3A	3.63	4.80E-2

rs104893877	SNCA_A53T	13.46	0,05
rs13194100	SH3GL2	3.00	0,05
rs10906913	FAM171A1.ITGA8	0.32	0,05
rs1130188	GBA_E365K_	11.13	0,05
rs11060180	OGFOD1.CCDC61	3.12	0,08

Tabla 6- Análisis de asociación por regresión logística múltiple. Modelo dominante. En esta tabla se presentan los resultados del análisis de regresión logística múltiple usando un modelo dominante de los datos. En la primera columna encontramos los IDs de los SNPs, en la segunda los nombres de los genes en donde se localizan los SNPs, en la tercera columna los Odd Ratios y en la cuarta columna los p-valores. Existen algunos SNPs en este resultado que no tienen asociación con la NOH, pero tienen tendencia a la significancia estadística y decidimos incluirlos para posterior estudio y compararlos con los resultados de otros modelos(aditivo y recesivo).

Nombre de SNPs	Gene	Odd Ratio	p value
rs34300866	TMEM175	24.82	4.42 E-3
rs34884107	TMEM175	56.82	2.42 E-2
rs2280104	BIN3	0.10	2.72 E-2
rs00060080	OGFOD1.CCDC62	3.45	3.36 E-2
rs6801093	FAM47E.STBD1	4.48	3.56 E-2
rs4444903	EGF	0.13	4.92 E-2
rs55785911	DDRGK1	3.89	5.00 E-2
rs353006	SCN3A.SCN2A	3.09	8.425 E-2

Tabla 7- Análisis de asociación por regresión logística múltiple. Modelo recesivo. En esta tabla se presentan los resultados del análisis de regresión logística múltiple usando un modelo recesivo de los datos. En la primera columna encontramos los IDs de los SNPs, en la segunda los nombres de los genes en donde se localizan los SNPs, en la tercera columna los Odd Ratios y en la cuarta columna los p-valores. Existen algunos SNPs en este resultado que no tienen asociación con la NOH, pero tienen tendencia a la significancia estadística y decidimos incluirlos para posterior estudio y compararlos con los resultados de otros modelos(aditivo y dominante).

SNP name	Gene name	Aditivo	Dominante	Recesivo
rs34311866#	TMEM175#	AR***	AR*	AR**
rs11060180	OGFOD2.CCDC62	AR**	—	AR*
rs6812193	FAM47E.STBD1	AR**	—	AR*

rs104893877	SNCA_A53T	AR**	AR**	—
rs353116	SCN3A.SCN2A	AR*	AR*	
rs13294100	SH3GL2	AR*	AR*	
rs3910105	SNCA	AR*	AR*	
rs10797576	SIPA1L2	DR*	—	—
rs10906923	FAM171A1.ITGA8	DR*	DR*	
rs329648	MIR4697	AR*	AR**	
rs34884217	TMEM175#	AR*	AR*	AR*
rs2694528	ELOVL7.NDUFAF2	DR.	DR*	
rs6430538	ACMSD.TMEM163	DR.	—	—
rs34637584	LRRK2_G2019S	DR.	—	—
rs11158026	GCH1	DR.	—	—
rs55785911	DDRKG1	AR.	—	AR.
rs1180104	BIN3	—	—	DR*
rs4444903	EGF	—	—	DR*
rs1130188	GBA E365K	—	AR.	—
rs34043159	MAP4K4	—	AR.	—
rs12456492	SYT4	AR.	—	—
rs4818	COMT	—	—	AR.

Tabla 8 - Cuadro comparativo entre los diferentes modelos aditivo, dominante y recesivo. En esta tabla se presenta una comparación entre los diferentes modelos genéticos y los SNPs asociados con NOH en cada modelo. Tuvimos en cuenta también si los SNPs aumentan el riesgo de NOH o si lo disminuyen. Algunos SNPs se encuentran asociados a hipotensión ortostática neurogénica en más de un modelo genético. Al final se quedaron los SNPs con el mejor modelo (señal más alto y fuerza de asociación más alta).

AR: aumenta el riesgo de NOH DR: Disminución de riesgo

***: $p < 0.0001$ **: $p < 0.001$, *: $p < 0.01$, .: cercano a 0.05,

#Significativo para los modelos dominante y recesivo.

7.1.3- Análisis epistático de pares de los SNPs

Los resultados de los modelos que acabamos de presentar nos dejaron pensar que pudiera existir interacciones SNVs-SNVs en la aparición de la NOH. En tal caso puede haber SNVs no encontrados asociados a NOH pero que puede ser importante su asociación con algunos de los SNPs encontrados. Por esa razón decidimos hacer un análisis de epistasis de pares entre los 72 SNPs para encontrar algunas interacciones significativas. Este tipo de

análisis requiere poder de cómputo muy alto usando los paquetes disponibles, por eso nosotros usamos un modelo de regresión logística multinomial y/o binomial según la codificación de los datos donde cada SNP fue considerado como variable respuesta contra todos los demás SNPs como explicativas usando la plataforma KNIME. En este modelo tuvimos en cuenta tres modelos de los datos explicativos: modelo aditivo contra aditivo, aditivo contra recesivo, aditivo contra dominante y recesivo contra dominante. En el modelo aditivo contra aditivo, la variable respuesta(SNP) es cualitativa con tres niveles y las covariables son cuantitativas. En el caso del modelo recesivo-aditivo, la variable de respuesta se transforma considerándolo como el SNV recesivo(i-e se necesita dos copias del alelo menor para producir el fenotipo) y se computa como categórico(0, 1), las variables explicativas sigue siendo cuantitativas. Mientras tanto, en el modelo dominante-aditivo, la variable dependiente es categórica y caracteriza al SNV como dominante i-e que con solo una copia del alelo menor, tiene efecto sobre el fenotipo y las variables independientes son de tipo cuantitativo.

En el caso del modelo aditivo-aditivo, hemos encontrado asociación significativa entre los SNVs rs356181-SNCA y rs3910105-SNCA en los pacientes con NOH(OR:0.51, P-value: 0.007)([Ver tabla 9](#)). Esta misma asociación existe también en los pacientes sin NOH con un OR de 0.18 mucho menor que en el caso de los pacientes con NOH. Hicimos también un análisis de chi cuadrado para estudiar la asociación entre ambos SNPs, encontramos que existe asociación significativa con un valor de chi cuadrado de 19,48(gl:4) y p-value de 0.001 en los pacientes con NOH. En los pacientes sin NOH, existe dicha asociación con un chi cuadrado de 171,232(gl:4) y p-value de 5.68E-36. Esta asociación está presente independientemente que los pacientes padecen de NOH o no. En el análisis de asociación entre los SNVs usando el modelo recesivo, hemos encontrado resultados parecidos al caso aditivo([ver tabla 9](#)).

SNP1	SNP2	Gene 1	Gene 2	OR	P-value
rs356181(c)	rs3910105	SNCA	SNCA	0.51**	0.007**
rs3910105(c)	rs356181	SNCA	SNCA	0.49**/0.18*	0.007**/0.0*
rs3910105(b)#	rs356181	SNCA	SNCA	0.18*	1.44E-8*
rs3910105(b)#	rs11343	SNCA	COQ7/SYT17	0.71*	0.044*

rs5993883(c)	rs740603	COMT	COMT	2.32**/10.75*	0.002**/0.0*
--------------	----------	------	------	---------------	--------------

Tabla 9- Análisis epistático, modelo aditivo contra modelo aditivo. En esta tabla se presentan los resultados de un análisis de epistasis entre pares de SNPs en un modelo genético de combinación aditivo-aditivo. Se realizaron combinaciones de modelos genéticos (recesivo-recesivo, recesivo-aditivo, recesivo-dominante, aditivo-aditivo, aditivo-dominante, aditivo-recesivo etc.). Cuando la variable dependiente es el modelo aditivo (codificado con tres niveles), se utiliza un modelo de regresión logística multinomial y cuando la variable dependiente es el modelo recesivo o dominante (con dos niveles) se usa una regresión logística binomial. La primera columna representan los SNPs de variable dependiente y la segunda columna los SNPs de variable independiente. En la tercera columna se encuentran los genes donde se localizan los SNPs de variables dependientes, la cuarta columna los genes donde se localizan los SNPs de variables independientes. En la quinta y la sexta columna se encuentran los Odd ratios y los p-valores respectivamente que indican las fuerzas de asociación entre los SNPs y vienen de la siguiente forma: dos cifras separadas por una barra invertida. La primera cifra con dos asteriscos designan resultados de OR y p-valor para el grupo de pacientes con NOH y la segunda cifra con un solo asterisco detrás de la barra designa los OR y p-valor de los pacientes sin NOH.

*Grupo de pacientes sin NOH

** Grupo de pacientes con NOH

a: La asociación entre ambos SNPs son del genotipo homocigota del alelo mayor

b: La asociación entre ambos SNPs son del genotipo heterocigota.

c: La asociación entre ambos SNPs son del genotipo homocigota del alelo menor

#Los dos se asocian juntos al SNP (variable respuesta).

Se presentan los resultados de los análisis epistáticos realizados sobre los 72 SNPs conocidos asociados a la Enfermedad de Parkinson. Los SNP1 y SNP2 se localizan dentro o en la vecindad de los genes 1 y gene 2 respectivamente. Un solo asterisco significa los resultados para el grupo de pacientes sin NOH y dos para el grupo con NOH. El “/” se usa para separar los resultados de cada grupo en la tabla. En los casos en los que no aparecen resultados para un grupo dado es porque no existieron suficientes datos para hacer los análisis para este grupo en este caso particular. En tal caso no se tendrá en cuenta estos resultados porque la idea es comparar los resultados entre ambos grupos.

SNP 1(rec)	SNP 2(add)	Gene 1	Gene 2	OR	P-values
rs356181	rs3910105	SNCA	SNCA	0.55**/0.23*	0,021**/0.0*
rs3910105	rs356181	SNCA	SNCA	0.54**/0.26*	0.018**/3.44E-15*
rs5993883	rs740603	COMT	COMT	2.3**/5.5*	0.001**/5.218E-15*

Tabla 10- Análisis epistático, modelo recesivo contra modelo aditivo. En esta tabla se presentan los resultados de un análisis de epistasis entre pares de SNPs en un modelo genético de combinación recesivo-aditivo. Se utiliza un modelo de regresión logística binomial. La primera columna representan los SNPs de variable dependiente y la segunda columna los SNPs de variable independiente. En la tercera columna se encuentran los genes donde se localizan los SNPs de variables dependientes, la cuarta columna los genes donde se localizan los SNPs de variables independientes. En la quinta y la sexta columna se encuentran los Odd ratios y los p-valores respectivamente que indican las fuerzas de asociación entre los SNPs y vienen de la siguiente forma: dos cifras separadas por una barra invertida. La primera cifra con dos asteriscos designan resultados de OR y p-valor para el grupo de pacientes con NOH y la segunda cifra con un solo asterisco detrás de la barra designa los OR y p-valor de los pacientes sin NOH.

*Grupo de pacientes sin NOH ** Grupo de pacientes con NOH a: La asociación entre ambos SNPs son del genotipo homocigota del alelo mayor b: La asociación entre ambos SNPs son del genotipo heterocigota. c: La asociación entre ambos SNPs son del genotipo homocigota del alelo menor #Los dos se asocian juntos al SNP (variable respuesta).

rec: la columna con datos que siguen el modelo de recesivo **add:** columna con datos que siguen el modelo aditivo

NB: Se realizó un análisis de regresión logística en knime por ciclo reiterativo en donde en cada ciclo un SNP representa la variable respuesta contra otro SNP como variable explicativa hasta computar todos los SNPs en todos los ciclos. Cuando un SNPs es variable respuesta, se transforma en un modelo recesivo donde el alelo menor se modela contra los demás alelos como un solo grupo (Homocigota del alelo menor contra homocigota del alelo mayor+heterocigota). En cuanto a los resultados, vemos que los SNPs localizados en los genes SNCA y COMT se interactúan tanto en los grupos de pacientes con NOH o sin NOH. Si bien no sirve mucho para nuestro objetivo que era comparar los resultados de ambos grupos y seleccionar los SNPs que se interactúan en el grupo NOH, podríamos postular que estos SNPs participan en el desarrollo de la EP. Los SNPs localizados en la SNCA estarían interactuando para disminuir el riesgo de EP aunque paradójicamente se sabe que estos SNPs actuando solo aumentan el riesgo. Por otro lado, los SNPs localizados en el gen COMT se estarían interactuando para aumentar el riesgo de EP.

En el análisis de interacción entre genes en un modelo dominante, hemos encontrado que el SNVs rs199347 (GPNMP) actúa sobre el SNV rs8005172(GALC/GPR65)(tabla 10) con un p-valor de 0.05 en el caso del grupo de pacientes con NHO y 0.28 en el grupo sin NOH. Existe una diferencia entre ambos grupos en cuanto a la interacción entre ambos SNPs. Lo que nos llevó a pensar que estos dos SNPs, al interactuar contribuyen al riesgo de la NOH mientras que por separado no tienen efecto sobre la presencia o no del fenotipo en cuestión.

Además se dio a conocer otra interacción entre los SNVs rs823118(NUCKS1) y rs356181(SNCA) en el grupo NOH(OR:1.62, pvalor: 0.05) mientras que en el grupo control no existe tal asociación(OR:1.08, p-valor: 0.56). El efecto de un SNV sobre el otro es unilateral es decir que solo la acción del SNV rs823118 (NUCKS1)(Variable independiente) sobre el SNP rs356181(SNCA) dio significativo, no al revés(OR:1.49, p valor: 0.137)([Ver tabla 11](#)). Hemos hecho un análisis de asociación entre ambos SNPs a través de la prueba de chi cuadrada; hemos encontrado que existe asociación entre ambos SNPs(rs823118, rs356181) en el grupo NOH(Chi cuadrado:8, p valor:0.018) y tal no es el caso en el grupo control(Sin NOH)(Chi cuadrado:1.54, p valor:0.462)([Ver tabla 11](#)).

SNP 1(dom)	SNP 2(add)	Gene 1	Gene 2(add)	OR	P-values
rs356181	rs3910105	SNCA	SNCA	0.56**/0.3*	0.028**/0.0*
rs3910105	rs356181	SNCA	SNCA	0.56**/0.3*	0.028**/0.0*
rs199347	rs8005172	GPNMP	GALC/GPR65	0.61**/1.16*	0.05**/0.28*
rs8005172	rs199347	GALC/GPR65	GPNMP	0.62**/1.01*	0.065**/0.89*
rs823118	rs356181	NUCKS1	SNCA	1.62**/1.08*	0.05**/0.56*
rs34311866	rs12497850	TMEM175	NCKIPSD/CD C71/IP6K2	0.6**/1.35*	0.042**/0.025*

Tabla 11- Análisis epistático, modelo dominante contra modelo aditivo. En esta tabla se presentan los resultados de un análisis de epistasia entre pares de SNPs en un modelo genético de combinación dominante-aditivo. Se utiliza un modelo de regresión logística binomial. La primera columna representan los SNPs de variable dependiente y la segunda columna los SNPs de variable independiente. En la tercera columna se encuentran los genes donde se localizan los SNPs de variables dependientes, la cuarta columna los genes donde se localizan los SNPs de variables independientes. En la quinta y la sexta columna se encuentran los Odd ratios y los p-valores respectivamente que indican las fuerzas de asociación entre los SNPs y vienen de la siguiente forma: dos cifras separadas por una barra invertida. La primera cifra con dos asteriscos designan resultados de OR y p-valor para el grupo de pacientes con NOH y la segunda cifra con un solo asterisco detrás de la barra designa los OR y p-valor de los pacientes sin NOH según la columna. Los SNPs localizados en los genes GPNMP y GALC/GPR65 se interactúan para disminuir el riesgo de NOH. Los SNPs localizados en los genes NUCKS1 y SNCA se interactúan para aumentar el riesgo de NOH.

*Grupo de pacientes sin NOH

** Grupo de pacientes con NOH

a: La asociación entre ambos SNPs son del genotipo homocigota del alelo mayor

b:La asociación entre ambos SNPs son del genotipo heterocigota.

c: La asociación entre ambos SNPs son del genotipo homocigoto del alelo menor

#Los dos se asocian juntos al SNP(variable respuesta).

Dom: la columna con datos que siguen el modelo de recesivo

add: columna con datos que siguen el modelo aditivo

NB: Se realizó un análisis de regresión logística en knime por ciclo en donde en cada ciclo un SNP representa la variable respuesta contra todos los demás SNPs como variables explicativas. Cuando un SNPs es variable respuesta se le transforma en un modelo dominante donde el alelo menor junto con el heterocigota se modela como un grupo (Homocigota del alelo menor +heterocigota contra homocigota del alelo mayor).

Capítulo VIII- Análisis de enriquecimiento de los SNPs asociados a NOH

En el análisis de enriquecimiento de los SNPs asociados a NOH, hemos encontrado que el SNP rs13294100-SH3GL2 enriquece significativamente para vías de señalización como señalización retrógrada de neurotrofinas, regulación negativa de la actividad MET, regulación negativa de EGFR, desprendimiento de vesículas de la red trans-Golgi, señalización por EGFR, biogénesis de vesículas lisosomales, biogénesis de vesículas asociadas al Golgi, reconocimiento de Cargo para la endocitosis mediada por clatrina. Por otro lado, el SNV rs2694528-NDUFAF2/ELOVL7 enriquece para las vías de señalización de biogénesis del complejo I y del transporte respiratorio de electrones, síntesis de acil-CoA grasos de cadena muy larga, metabolismo de los ácidos grasos (Ver tabla). En la tabla aparecen otros genes que no se encontraron asociados a NOH en el análisis primario, pero después del análisis de desequilibrio de ligamiento, hemos encontrado SNVs en desequilibrio de ligamiento con los SNVs asociados a NOH, localizados en estos genes: QARS, PRKAR2A, SLC25A20 (Ver tabla 12).

Genes Involved	Metabolic Pathway	p-value
IP6K2	Synthesis of IPs in the nucleus	0.003
	Inositol phosphate metabolism	0.039
	Interferon alpha/beta signaling	0.038
SH3GL2	Retrograde neurotrophin signaling	0.012
	InIB-mediated entry of Listeria monocytogenes into a cell	0.013
	Listeria monocytogenes entry into host cells	0.017
	Negative regulation of MET activity	0.017
	EGFR downregulation	0.026
	trans-Golgi Network Vesicle Budding	0.039
	Signaling by EGFR	0.042
	Lysosome Vesicle Biogenesis	0.029
	Recycling pathway of L1	0.040
	Golgi Associated Vesicle Biogenesis	0.046
	Cargo recognition for clathrin-mediated endocytosis	0.05
SCN2A	Interaction between L1 and Ankyrins	0.026
	Phase 0 - rapid depolarization	0.035
SH3GL2-SCN2A	L1CAM interactions	0.004
	Axon guidance	0.034
ITGA8	Molecules associated with elastic fibers	0.032
	Elastic fiber formation	0.037
	Integrin cell surface interactions	0.046
NDUFAF2	Complex I biogenesis	0.041
	Respiratory electron transport	0.049
PRKAR2A	ROBO receptors bind AKAP5	0.013
	AC-mediated CREB1 phosphorylation	0.018
	PKA activation in glucagon signaling	0.025
	PKA activation	0.027
	PKA-mediated phosphorylation of CREB	0.029
	DARPP-32 events	0.035
	Glucagon signaling in metabolic regulation	0.048
SLC25A20	Carnitine metabolism	0.019
ELOVL7	Synthesis of very-long-chain fatty acyl-CoA	0.035
	Fatty acid metabolism	0.028
QARS	Mitochondrial tRNA aminoacylation	0.031
	Cytosolic tRNA aminoacylation	0.035

Tabla 12- Análisis de enriquecimiento de los SNPs asociados a NOH. En esta tabla se presentan los resultados de un análisis de enriquecimiento de los SNPs asociados con NOH. También se incluyeron en este análisis otros SNPs encontrados relacionados con los SNPs asociado a NOH en un análisis de desequilibrio de ligamentos. En la primera columna se encuentran los genes en los cuales se localizan los SNPs. En la segunda columna se encuentran las vías metabólicas y procesos moleculares en los cuales se involucran estos genes. En la columna tres se presentan los p-valores que informan sobre la significancia de los resultados. Este análisis se hizo utilizando la plataforma web [SNP Nexus](https://www.snp-nexus.org/)(<https://www.snp-nexus.org/>).

**Capítulo IX- Análisis de expresión de genes basado en LD de los SNPs
asociados a NOH**

SNP (Gene)	SNP en LD	Chr	Influencia sobre gene	Brain region	Effect size	p-values	R ²	D'
rs34311866 (TMEM175)	27	4	DGKQ	BHPC, BCRB, BFC	-0.47/ -0.19	5.90E-7	0.65/ 1.00	0.86 / 1.00
rs34311866 (TMEM175)	1	12	HIP1R	BCRB	0.22	3.90E-6	1.00	1.00
Rs3910105(SNCA)	61	4	SNCA-AS1	BCRB	-0.63 / -0.39	1.66E-12 / 2.11E-5	0.62 / 1.00	0.86 / 1.00
Rs353116 (SCN2A)	2	2	A27.3	BCRB	0.59	9.00E-6	0.17	0.88
Rs6812193 (FAM47E-STBD)	14	4	CCDC158	BC, BFC, BSC	0.32 / 0.48	1.20E-5	0.5 / 1.00	0.97 / 1.00
Rs6812193 (FAM47E-STBD)	99	4	FAM47E	BCB, BCRB, BC, BFC, BHPC, BNA, PB, BSC	0.26 / 0.45	7.59E-5/ 8.78E-9	0.91 / 1.00	0.99 / 1.00
Rs10906923 ITGA8	200	10	ITGA8	BCB,BC, BHY, BNA	-0.41 / -0.31	4.49E-5 / 7.67E-8	0.53 / 1.00	0.84 / 1.00

Rs12497850 (IP6K2)	1232	3	AMT	BCRB, BC, BCB, BNA, BFC, BHPC, BP	0.19/ 0.67	2.02E-4 / 5.28E-27	0.57/ 1.00	0.82 / 1.00
Rs12497850 (IP6K2)	351	3	GMPPB	BAMY, BAC, BCRB, BC, BFC, BHPC, BP, BSC, BSN	-0.51 / -0.51	1.48E-10 / 4.20E-4	0.57/ 0.88	0.82 / 0.95
Rs12497850 (IP6K2)	1602	3	WDR6	BAMY, BCB, BCRB, BCB, BC, BFC, BHPC, BNA, BP, BSC, BSN	-0.45 /0.40	2.03E-27 / 1.89E-4	0.50/ 1.00	0.88/1.00
Rs11158026(G CH1)	39	14	ATG14	BFC,BP, BCRB	-0.43 /0.35	1.62E-8/3. 20E-5	0.50/ 1.00	0.82 / 1.00

Tabla 13- Análisis de expresión de genes basado en el desequilibrio de ligamiento entre los SNPs. En esta tabla se presentan los resultados del análisis de expresión de genes basado en el desequilibrio de ligamiento(cis-eQTL). Los SNPs encontrados asociados a NOH fueron entrados en la aplicación web LDexpress(www.idlink.nih.gov/?tab=express). Se computaron los SNPs en desequilibrio de ligamiento con los SNPs targets(asociados a NOH), y se computaron la expresión de los genes donde se localizan aquellos SNPs. Los datos eQTL se obtuvieron de la plataforma GTEx. En cuanto a la tabla 13, en la primera columna se muestran los SNPs asociados a NOH(utilizados como entrada a la plataforma LDexpress) junto con los genes donde se localizan. En la columna dos se encuentran el número de SNPs que están en desequilibrio de ligamiento(LD) con el SNP target, la columna tres presenta los cromosomas donde se encuentran estos SNPs. La columna cuatro indica los genes donde se localizan los SNPs en LD con aquellos SNPs target(Asociados a NOH). En la columna cinco se encuentran las regiones del cerebro donde estos genes cuya expresión se ve afectada por los SNPs. En la columna seis muestra el tamaño del efecto(cuantifica como la expresión del gen cambia bajo el efecto del SNP). La columna siete, ocho y nueve presentan el p-valor, R², y D' respectivamente. Se consideraron R² >=0.5 y D' >= 0.8. Desde la columna seis hasta la columna nueve se presentan los resultados como intervalo que va desde el valor más bajo hasta el valor más alto para cada SNPs en cada región. El gen DGKQ(donde se localizan 27 SNPs en LD con el SNP rs34311866-TMEM175) está subexpresado en el hipocampo, cerebelo y en la corteza frontal. El gen HIP1R(donde se localiza un SNP en LD con el rs34311866-TMEM175) tiene su expresión alterada bajo el efecto de este SNP(rs43311866) en el cerebelo. El gen SNCA-AS1(donde se localiza 61 SNPs en LD con el SNP rs3910105-SNCA) está subexpresado en el cerebelo. El gen CCDC158(donde se localizan 14 SNPs en LD con el SNP rs6812193-FAM47E/STBD1) está sobreexpresado en la corteza cerebral, corteza frontal y médula espinal. El gen FAM47E(donde se localizan 99 SNPs en LD con el SNP rs6812193-FAM47E/STBD1) está sobreexpresado en la corteza cerebral y frontal, el cerebelo, el caudado, el hipocampo, el núcleo Accumbens, el putamen y la médula espinal. El gen ITGA8(donde están localizados los 100 SNPs en LD con el SNP rs10906923-ITGA8) está subexpresado en el caudado, el hipotálamo y el núcleo accumbens. El gen AMT(donde se localizan 1232 SNPs en LD con el SNP rs12497850-IP6K2) tiene su expresión alterada bajo el efecto de estos SNPs en el cerebelo, la corteza, el caudado, el núcleo accumbens, corteza frontal, el hipocampo y el putamen. El gen GMPPB(donde se localizan 351 SNPs en LD con el SNP rs12497850-IP6K2) está sobreexpresado en el cerebelo, la amígdala, la corteza singular anterior, corteza frontal, hipocampo, el putamen, médula espinal. El gen WDR6(donde se localizan 1602 SNPs en LD con el SNP rs12497850-IP6K2) tiene su expresión alterada en el cerebelo, amígdala, corteza singular anterior, corteza frontal, corteza cerebral, hipocampo, putamen, médula espinal, núcleo accumbens. El gen ATG14(donde se localizan 39 SNPs en LD con el SNP rs11158026-GCH1) tiene su expresión alterada en la corteza frontal, el putamen y el cerebelo.

Chr: chromosome in which the SNPs in LD could be found. BHPC: Brain - Hippocampus. BCRB : Brain - Cerebellum. BFC: Brain - Frontal Cortex. BC: Brain - Cortex. BSC: Brain - Spinal cord. BCB: Brain - Caudate (basal ganglia). BNA: Brain - Nucleus accumbens (basal ganglia). BP: Brain - Putamen (basal ganglia). BHY: Brain - Hypothalamus. BAMY: Brain - Amygdala. BAC: Brain - Anterior cingulate cortex. BSN: Brain - Anterior cingulate cortex.

El SNV rs34311866-TMEM175 está en desequilibrio de ligamiento con 26 otros SNVs(ver tabla 13) los cuales alteran la expresión de los genes donde se localizan. La expresión de los genes DGKQ y CPLX1 se encontraron disminuida por algunos SNVs en la región cerebelar(Ver tabla 13). El alelo del SNV que provoca la alteración en la expresión de los genes es el alelo menor en los 3 casos lo que concuerda con el análisis de asociación SNV-NOH donde el alelo menor del SNP rs34311866 está asociado con la NOH y además que altera la expresión de genes en LD con el mismo. Hemos encontrado que el SNP rs34311866 que provoca estos efectos sobre las proteínas se encuentra localizado en región 3'-río abajo del gen DGKQ y en la región codificante del gen TMEM175. Según el análisis de predicción de SIFT y polyphen este SNV no tiene un efecto deletéreo sobre la función de la proteína TMEM175 aunque se encuentra en la región codificante de la misma(Score de predicción:0.09). El SNV rs34311866 en LD con otro SNV está relacionado con aumento de la expresión del gen HIP1R en la región cerebelar. Otro SNP asociado a NOH rs3910105-SNCA en LD con 61 SNP, los cuales disminuyen la expresión del gen SNCA-AS1(Ver tabla 13). Los demás resultados son presentados en la tabla 13, referirse a ella para más detalles.

Capítulo X- Efectos de los SNPs asociados a NOH sobre la severidad y progresión del EP

10.1- Análisis de asociación por regresión lineal y ANOVA entre UPDRS I, II, III, IV y los SNVs en los datos de primera visita de los pacientes.

Hemos realizado un análisis de regresión lineal utilizando los parámetros MDS-UPDRS I, II, III, IV y los datos de las variantes para poder determinar si existe alguna asociación entre ambas variables. Los SNVs encontrados asociados con MDS-UPDRS se someten a otro análisis de ANOVA de una vía para ver el efecto de cada genotipo sobre el MDS-UPDRS.

10.1.1- UPDRS I

SNPs	Variables	Coef	P-valor	Año
SNCA_A53T_rs104893877	MDS-UPDRS I	6.18	8.71E-8	BL
IL1R2/MAP4K4_rs34043159	MDS-UPDRS I	0.76	0.034	BL
GCH1_rs11158026	MDS-UPDRS I	0.75	0.037	BL

SH3GL2_rs13294100	MDS-UPDRS I	-0.73	0.05	BL
SNCA rs3910105	MDS-UPDRS I	0.67	0.05	BL

Tabla 14- Análisis de asociación entre UPDRS I y variantes genéticas por regresión lineal simple. En esta tabla se presenta un análisis de asociación entre los 72 SNPs de la base de datos del PPMI y los Scores UPDRS I utilizando un modelo de regresión. La primera columna: SNPs ID y genes donde se localizan, columna dos: Escala UPDRS I. Columna tres: coeficiente beta. Columna cuatro: P-valor. Columna cinco: BL(baseline/primer visita).

SNPs	HDAM	HET	HDAME	P-valor
SNCA_rs3910105	4+-2.68* 5.09+-3.53**	6.26+-4.38* 5.68+-4.52**	8.3+-10.00* 6.87+-4.74**	0.39* 0.005**
SNCA_A53T_rs104893877	—	5.21+-3.74* 5.88+-4.26**	14.80+-11.90* 10.73+-5.38**	0.07* KW 2.8E-4**
IL1R2/MAP4K4_rs34043159	3.65+-3.06* 5.54+-3.93**	7.06+-4.52* 6.59+-4.73**	12.67+-10.97* 5.52+-4.05**	0.004* 0.09**
GCH1_rs11158026	5.33+-4.01* 5.77+-4.53**	5+-3.37* 6.26+-4.08**	15+-11.87* 6.54+-4.90**	0.001* 0.51**
SH3GL2_rs13294100	7.92+-8.98* 6.48+-4.55**	6+-4.07* 5.78+-4.29**	4.40+-4.34* 5.36+-3.97**	0.5* 0.25**

Tabla 15- Análisis de ANOVA de una vía sobre SNPs asociados a UPDRS I según los genotipos. En esta tabla presentamos los resultados de un análisis de ANOVA de los valores del UPDRS I con respecto a si el paciente es de homocigota para el alelo mayor, homocigota para el alelo menor o heterocigota. El análisis se hace en ambos grupos NOH y no-NOH por separado. Primera columna : SNPs ID y los genes donde se localizan. Columna dos: resultados para los homocigotas para el alelo mayor. Columna tres: resultados de los heterocigotas. Columna cuatro: resultados de los homocigotas para el alelo menor.

HDAM: Homocigota para el alelo menor HDAME: Homocigota para el alelo mayor HET: Heterocigota

*: Grupo con NOH **: Grupo sin NOH —: No hay suficiente datos para los cálculos por ANOVA

KW: Prueba no paramétrica de Kruskal Wallis(Se aplica cuando no hay homocedasticidad u homogeneidad de la varianza con Levene significativo, p-valor<0.05).

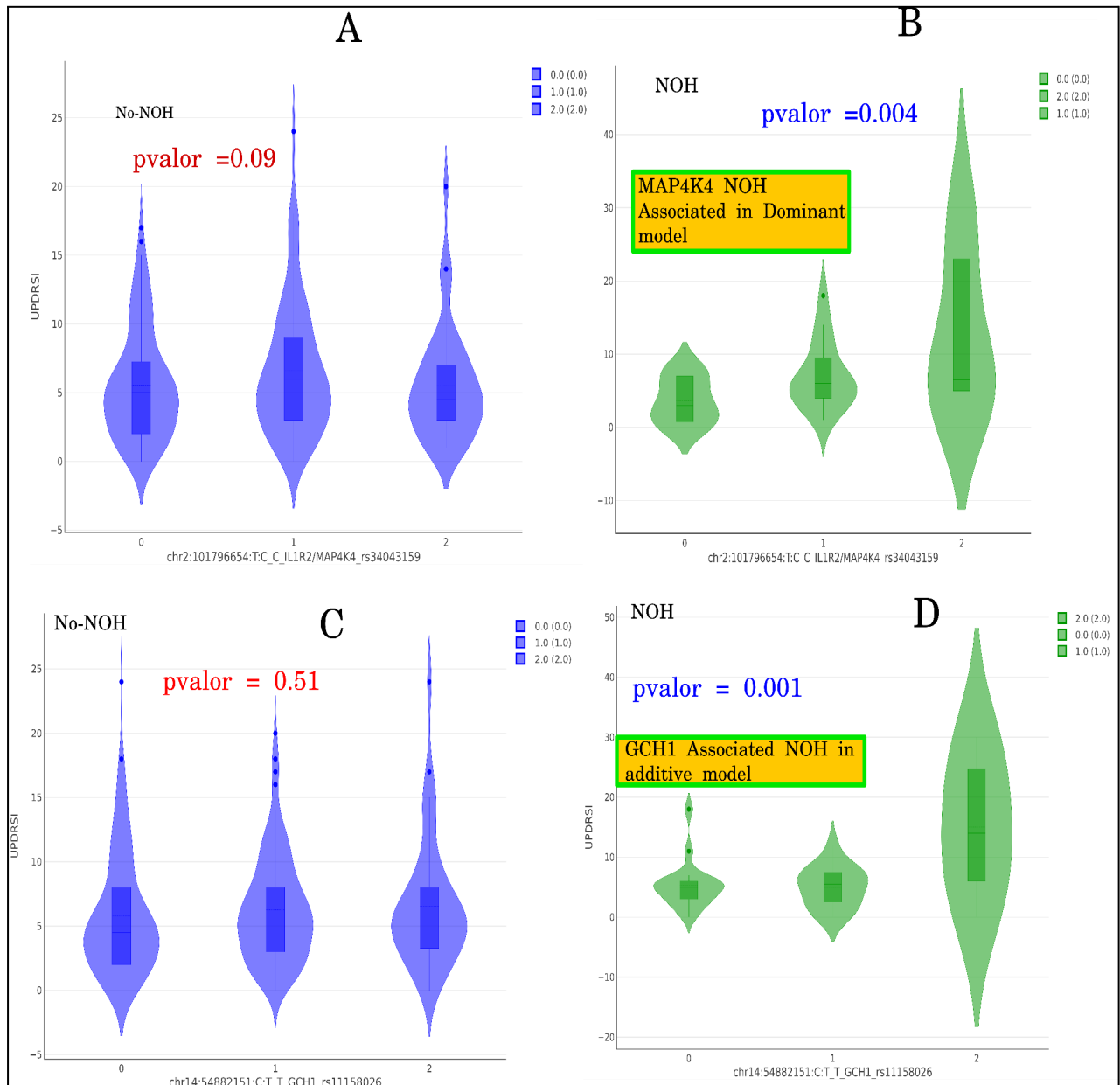


Figura 15- Distribución de los valores del UPDRS I según los genotipos de los SNPs rs34043159-MAP4K4 y rs11158026-GCH1. En esta figura se presentan para dos SNPs los valores de los UPDRS I según los diferentes genotipos. En el caso del SNP localizado en el gen MAP4K4 y en el grupo NOH, se nota que los pacientes homocigotas para el alelo menor tiene mayor cifras UPDRS I que los pares en el grupo control(No-NOH). Los mismos resultados son encontrados para el SNP localizado en el gen GCH1.

En el análisis de regresión lineal, hemos encontrado los SNPs rs104893877, rs3910105(SNCA), rs34043159(IL1R2/MAP4K4), rs11158026(GCH1), rs13294100(SH3GL2) se asocian con el MDS-UPDRS I(Ver tabla 14). El SNP rs104893877 localizado en el gen SNCA presenta una señal bastante alta para esta asociación con un p-valor de $8.71E-8$ y un coeficiente de 6.18. Por otro lado, en el análisis de ANOVA, hemos encontrado que esta señal que hemos visto se produce en el grupo sin NOH(NO, valor:0.07, No-NOH, valor: $2.8E-4$), lo cual no es de nuestro interés en este estudio(tabla 14). Pero los SNPs rs34043159(IL1R2/MAP4K4) y rs11158026(GCH1) presentan mayores cifras de

MDS-UPDRS I en los pacientes homocigotas para el alelo menor con NOH con respecto al grupo control(sinNOH)(ver tabla 15), (ver figura 15). También hay que recalcar que las cifras de MDS-UPDRS I de los pacientes rs34043159-heterocigotas del grupo NOH está ligeramente elevada con respecto a los homocigotas del alelo mayor, lo que coincide con nuestro hallazgo de dicho SNPs asociado a NOH en el modelo dominante. Quiere decir que con solo la presencia de una copia del alelo menor del SNP rs34043159(IL1R2/MAP4K4), se produce alteración en las cifras del MDS-UPDRS I. Mientras que en el caso del SNP rs11158026(GCH1), las cifras del MDS-UPDRS I de los pacientes heterocigotas y homocigotas del alelo mayor son parecidas con respecto a los homocigotas del alelo menor. Eso quiere decir que los cambios en las cifras del MDS-UPDRS I solo se producen cuando están presentes las dos copias del alelo menor(tabla 14 y 15). Este hallazgo es diferente de lo que hemos encontrado con este SNP en el análisis de asociación con NOH. Hemos encontrado que este SNP está asociado con NOH en un modelo aditivo y no en un modelo recesivo.

El MDS-UPDRS I es la primera parte de la escala unificada para la evaluación de la enfermedad de Parkinson. Esta escala permite evaluar el estado mental, el comportamiento y humor de los pacientes. MDS-UPDRS I estudia el deterioro intelectual, trastornos del pensamiento, depresión motivación e iniciativa. En la subescala motora se considera un cambio mínimo a 2.5 puntos, moderado a 5.2 e importante a 10.8. Para la MDS-UPDRS total, los valores que indican un cambio mínimo, moderado o importante son 4.3, 9.1 y 17.1 puntos respectivamente.

A partir de los hallazgos encontrados en este estudio, podemos resumir que los pacientes con NOH que presentan los SNP rs34043159(IL1R2/MAP4K4)(al menos una sola copia del alelo menor presente), rs11158026(GCH1)(las dos copias del alelo menor deben ser presentes), tienen mayor riesgo de tener alteración importante(promedio de MDS-UPDRS I de 12.67 y 15 respectivamente) de la esfera cognitiva, depresión y trastornos del pensamiento con respecto a los pacientes que no tienen NOH. También hay que recalcar que los pacientes con enfermedad de Parkinson en general tanto en el grupo con NOH o sin NOH además con las dos copias del alelo de riesgo del SNP SNCA_A53T_rs104893877 presentan mayor riesgo de tener alteración del MDS-UPDRS I .

10.1.2- MDS-UPDRS II

SNPs	Variables	Coef	P-valor	Año
------	-----------	------	---------	-----

SNCA_A53T_rs104893877	MDS-UPDRS II	6.02	4.16E-5	BL
ELOVL7/NDUFAF2_rs2694528	MDS-UPDRS II	-1.53	0.02	BL
GCH1_rs11158026	MDS-UPDRS II	1.072	0.012	BL
SH3GL2_rs13294100	MDS-UPDRS II	-0.85	0.056	BL
LRRK2_G2019S_rs34637584	MDS-UPDRS II	1.51	0.026	BL

Tabla 16- Análisis de asociación entre los valores del MDS-UPDRS II y los SNPs asociados a NOH. En esta tabla se presentan los resultados de una asociación entre los SNPs de la base de datos del PPMI y la escala UPDRS II. En la primera columna se encuentran los SNPs ID y los genes en donde se localizan. En la columna tres los coeficientes de regresión, en la columna cuatro los p-valores.

En el caso de la MDS-UPDRS II, hemos encontrado un total de 6 SNPs asociado con la MDS-UPDRS II (Ver tabla 16). Todos están asociados con NOH también. La señal del SNP SNCA_A53T_rs104893877 es la más fuerte de todos (p valor = 4.16E-5). Pareciera que los SNPs ELOVL7/NDUFAF2_rs2694528 y SH3GL2_rs13294100 tienen un efecto protector contra cifras altas de MDS-UPDRS II en el análisis de regresión lineal. Pero en la ANOVA ninguno de los SNPs encontrados asociados a MDS-UPDRS II tienen un efecto diferencial en cuanto a las cifras del promedio del MDS-UPDRS II (Ver tabla 16 y 17). Hay que descartar acá que los pacientes heterocigotas para el SNP SNCA A53T_rs104893877 tienen mayor riesgo de presentar alteración del MDS-UPDRS II tanto en pacientes con NOH como los sin NOH.

SNPs	HDAM	HET	HDAME	P-valor
SNCA_A53T_rs104893877	5.77+-4.82* 6.19+-4.86**	12.25+-7.76* 12.12+-5.93**	—	0.02* 8.4E-4**
ELOVL7/NDUFAF2_rs2694528	7.03+-5.76* 6.65+-5.32**	3.60+-2.51* 5.56+-3.36**	2.25+-0.96**	0.20* 0.08** KW
GCH1_rs11158026	6.44+-5.19* 5.76+-4.42**	5.21+-4.41* 6.57+-4.59**	11.5+-8.66* 8.03+-7.57**	0.13* 0.06**
SH3GL2_rs13294100	7+-7.01* 7+-5.70**	6.47+-5.11* 5.88+-4.18**	5.50+-3.41* 5.52+-4.43**	0.90* 0.14**
LRRK2_G2019S_rs34637584	6.76+-5.69* 6.07+-4.41**	4.75+-4.11* 6.96+-4.99**	20.5+-27.57**	0.50* 0.42** KW
Max*(EPMEDYN)	2.50+-2.52* 6.66+-5.13**	7.06+-5.61* 4.43+-3.39**	—	0.12* 0.01**

Tabla 17- Distribución de los valores del MDS-UPDRS II según los genotipos de los SNPs encontrados asociados usando el análisis de la varianza de una vía (ANOVA). En esta tabla se presentan los resultados de la ANOVA según los genotipos de los

pacientes. Se computan y comparan los valores de UPDRS II en los genotipos. Columna 1: SNPs ID y genes donde se localizan. Columna dos: homocigotas para el alelo mayor. Columna tres: heterocigotas. Columna cuatro: homocigotas para el alelo menor. Columna cinco: p-valor. No se encontró ningún resultado significativo.

*Grupo que tiene NOH

**Grupo sin NOH

10.1.3- MDS-UPDRS III

SNPs	Variables	Coef	P-valor	Año
SH3GL2_rs13294100	MDS-UPDRS III	-2.30	2.40E-13	BL
SNCA_A53T_rs104893877	MDS-UPDRS III	5.57	2.37E-11	BL
SIPA1L2_rs10797576	MDS-UPDRS III	2.42	4.22E-10	BL
FAM47E/STBD1_rs6812193	MDS-UPDRS III	1.78	1.04E-9	BL
Age	MDS-UPDRS III	0.12	5.80E-9	BL
GCH1_rs11158026	MDS-UPDRS III	1.65	1.29E-8	BL
NCKIPSD/CDC71/IP6K2_rs12497850	MDS-UPDRS III	1.33	3.04E-5	BL
OGFOD2/CCDC62_rs11060180	MDS-UPDRS III	1.18	3.52E-5	BL
GBA_E365K_rs2230288	MDS-UPDRS III	3.80	9.37E-5	BL
COQ7/SYT17_rs11343	MDS-UPDRS III	-1.00	3.50E-4	BL
GBA_N370S_rs76763715	MDS-UPDRS III	2.71	3.63E-4	BL
BIN3_rs2280104	MDS-UPDRS III	-0.98	7.76E-4	BL
IL1R2/MAP4K4_rs34043159	MDS-UPDRS III	0.95	9.57E-4	BL
LRRK2_rs76904798	MDS-UPDRS III	1.37	0.001	BL
ITPKB_rs4653767	MDS-UPDRS III	0.86	0.005	BL
ELOVL7/NDUF2F2_rs2694528	MDS-UPDRS III	-1.22	0.009	BL
CTSB_rs1293298	MDS-UPDRS III	0.73	0.020	BL
LRRK2_G2019S_rs34637584	MDS-UPDRS III	-0.95	0.026	BL

COMT_rs4818	MDS-UPDRS III	0.63	0.026	BL
VPS13C_rs2414739	MDS-UPDRS III	0.72	0.035	BL
DDRGK1_rs55785911	MDS-UPDRS III	-0.57	0.043	BL
MIR4697_rs329648	MDS-UPDRS III	0.61	0.043	BL

Tabla 18- Análisis de asociación entre los valores del MDS-UPDRS III y los SNPs. En esta tabla presentamos los resultados de asociación entre los 72 SNPs asociados a EP de la base de datos del PPMI utilizando un modelo de regresión lineal. Primera columna: SNPs ID y genes donde se localizan. Columna tres: coeficientes de regresión. Columna cuatro: p-valores. Todos los SNPs presentes en esta tabla están asociados con UPDRS III.

SNPs	HDAM	HET	HDAME	P-valor
SH3GL2_rs13294100	24.91+-12.63* 21.29+-11.90**	22.20+-14.09* 19.40+-8.24**	17.66+-4.72* 16.23+-7.36**	0.67* 0.14** KW
SNCA_A53T_rs104893877	19.93+-10.02* 19.79+-9.46**	34.00+-17.78* 23.875+-17.91**	—* —**	0.049* KW 0.89** KW
SIPA1L2_rs10797576	20.88+-12.94* 19.29+-9.49**	28.11+-12.14* 21.24+-11.58**	—* 25.14+-5.89**	0.15* 0.14**
FAM47E/STBD1_rs6812193	23.78+-12.85* 19.11+-9.17**	18.07+-10.68* 19.11+-9.17**	28.50+-15.29* 21.49+-13.32**	0.19* 0.36**
GCH1_rs11158026	15.33+-7.22* 19.46+-10.62**	24.16+-9.29* 20.48+-9.65**	46.00+-6.36* 20.71+-10.39**	5.02E-8* 0.68
NCKIPSD/CDC71/IP6K2_rs124978 50	21.23+-14.91* 18.73+-9.74**	24.16+-11.06* 20.98+-10.65**	—* —**	0.51* 0.14**
OGFOD2/CCDC62_rs11060180	18.25+-8.97* 18.90+-11.47**	25.83+-15.01* 20.39+-9.66**	20.55+-10.87* 21.40+-8.55**	0.33* 0.33**
GBA_E365K_rs2230288	23.18+-13.22* 19.88+-10.12**	15.50+-2.12* 24.40+-10.94**	—* —**	0.42* 0.16**
COQ7/SYT17_rs11343	28.40+-18.83* 21.33+-11.75**	23.06+-10.16* 19.25+-9.05**	15.88+-5.35* 20.23+-10.36**	0.20* 0.36**

GBA_N370S_rs76763715	22.68+-13.17* 19.79+-9.67**	23.33+-12.89* 22.71+-14.23**	—* —**	0.95* 0.64**
BIN3_rs2280104	23.43+-12.88* 20.33+-9.97**	22.23+-14.157* 20.13+-10.59**	21.50+-2.12* 19.03+-9.36**	0.95* 0.79**
IL1R2/MAP4K4_rs34043159	19.23+-12.84* 20.03+-11.48**	22.83+-11.77* 19.59+-8.85**	33.75+-15.67* 21.16+-10.33**	0.14* 0.68**
LRRK2_rs76904798	22.04+-9.48* 19.49+-9.805**	24.08+-18.35* 21.29+-11.10**	—* 24.5+-7.77**	0.66* 0.36**
ITPKB_rs4653767	22.000+-12.70* 19.81+-10.10**	24.25+-14.56* 20.00+-9.88**	21.50+-10.61* 21.26+-11.96**	0.88* 0.81**
ELOVL7/NDUFAF2_rs2694528	24.21+-13.67* 20.18+-10.28**	15.66+-4.93* 20.19+-9.50**	—* 8.00+-2.16**	0.14* 0.058**
LRRK2_G2019S_rs34637584	24.033+-12.92* 20.26+-10.53**	15.00+-11.46* 18.86+-8.28**	—* 31.66+-21.36**	0.15* 0.015** KW

Tabla 19- Distribución de los valores del UPDRS III según los genotipos de los SNPs encontrados asociados usando el análisis de la varianza(ANOVA). Primera columna: SNPs ID y los genes en donde se localizan. Columna dos: homocigotas para el alelo mayor. Columna tres: heterocigotas. Columna cuatro: homocigotas para el alelo menor. Columna cinco: Heterocigotas.

*Grupo NOH

**Grupo no NOH

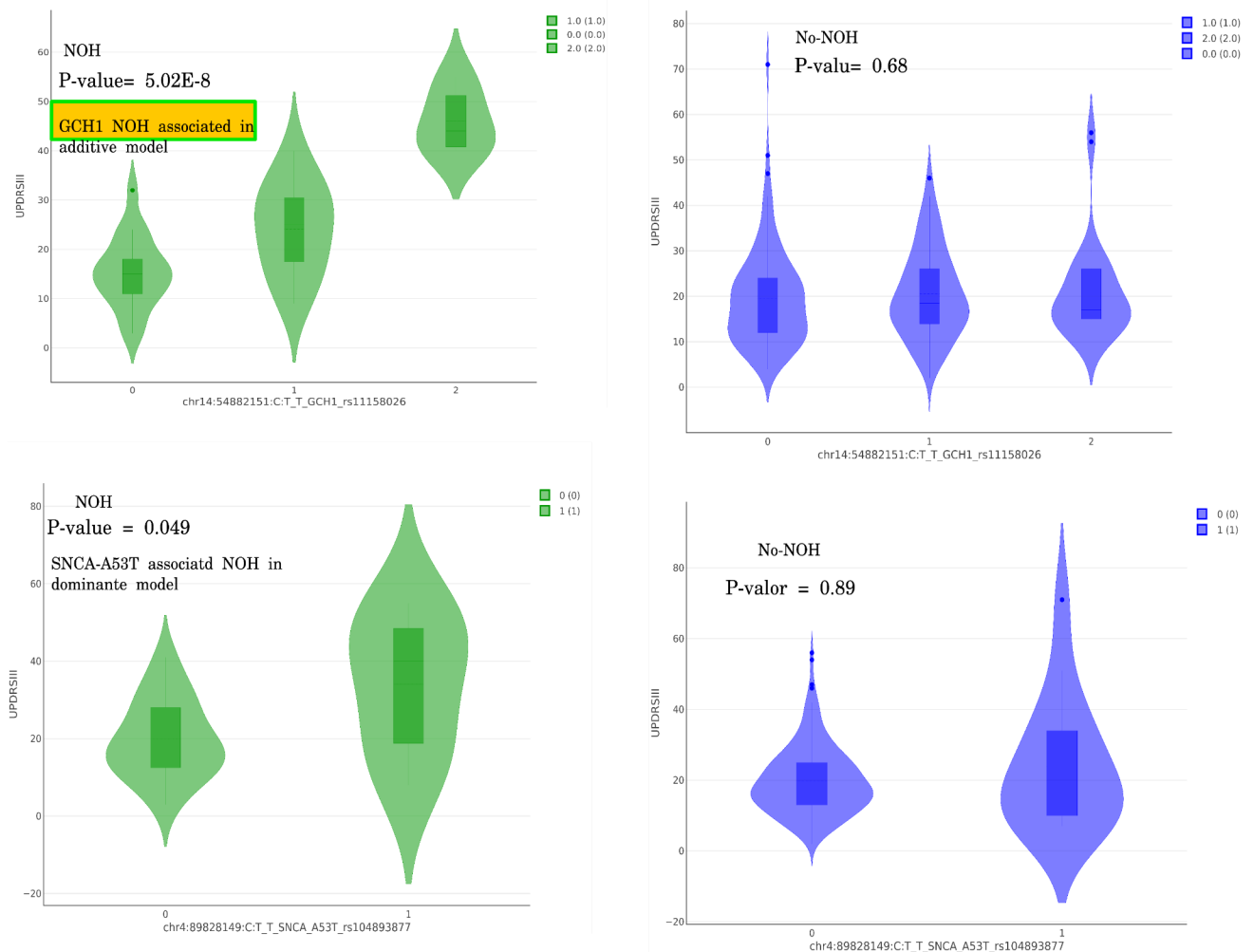


Figure 16- Distribución de los valores del UPDRS III según los genotipos de los SNPs. En esta tabla presentamos las distribución de los valores de UPDRS III para los genotipos de los SNPs localizados en el gen SNCA-A53-T y GCH1(solo los SNPs en este estudio, ver tablas 3-11). En el caso del SNP localizado en el gen GCH1, se identificó que los pacientes con NOH y homocigotas para el alelo menor y heterocigotas del SNP-GCH1 tienen las cifras de UPDRS III más altos que los pares del grupo control(no-NOH).

El MDS-UPDRS III está asociado con unos 21 SNP de los 72 dos analizados con p-valor menor a 0.05. Mientras que en el ANOVA encontramos diferencias de las cifras de MDS-UPDRS III en ambos grupos solamente en los SNP SNCA_A53T_rs104893877 y GCH1_rs11158026(Ver tabla 18 y tabla 19 y Figura 16).

10.1.4- UPDRS IV

SNPs	Variables	Coef	P-valor	Año
SNCA_A53T_rs104893877	MDS-UPDRS IV	3.14	0.0	BL
SCN3A/SCN2A_rs353116	MDS-UPDRS IV	-0.61	4.49E-12	BL
COMT_rs5993883	MDS-UPDRS IV	-0.58	1.349E-11	BL
LRRK2_G2019S_rs34637584	MDS-UPDRS IV	0.812	5.845E-10	BL
ELOVL7/NDUF2F2_rs2694528	MDS-UPDRS IV	-0.57	3.047E-5	BL
COMT_rs4818	MDS-UPDRS IV	0.35	5.93E-5	BL
CTSB_rs1293298	MDS-UPDRS IV	0.32	7.379E-4	BL
NCKIPSD/CDC71/IP6K2_rs12497850	MDS-UPDRS IV	-0.32	8.897E-4	BL
GBA_E365K_rs2230288	MDS-UPDRS IV	0.97	0.0011	BL
COQ7/SYT17_rs11343	MDS-UPDRS IV	0.27	0.0016	BL
BIN3_rs2280104	MDS-UPDRS IV	0.27	0.0023	BL
SNCA_rs3910105	MDS-UPDRS IV	0.25	0.0043	BL
SH3GL2_rs13294100	MDS-UPDRS IV	-0.26	0.007	BL
GBA_N370S_rs76763715	MDS-UPDRS IV	-0.588	0.012	BL
FAM171A1/ITGA8_rs10906923	MDS-UPDRS IV	-0.23	0.013	BL
TMEM175_rs34884217	MDS-UPDRS IV	-0.346	0.015	BL
GCH1_rs11158026	MDS-UPDRS IV	0.21	0.020	BL
SNCA_rs356181	MDS-UPDRS IV	0.182	0.035	BL

Tabla 20- Análisis de asociación entre los valores del UPDRS IV y los SNPs asociados a NOH. En esta tabla se presentan los resultados de un análisis de asociación utilizando modelo de regresión lineal entre los SNPs de la base de datos del PPMI y la escala de UPDRS IV. Primera columna: SNPs ID y los genes donde se localizan. Columna tres: Coeficiente beta de regresión. Columna cuatro: P-valores. Columna cinco: datos de primera visita (BL, baseline). Los SNPs presentes en esta tablas son todos asociados con UPDRS IV.

SNPs	HDAM	HET	HDAME	P-valor
SNCA_A53T_rs104893877	2.105+-2.45*	4.235+-3.44*	—*	0.013*
	3.36+-3.349*	5.83+-4.52*	—*	1.37E-4
CTSB_rs1293298	1.83+-2.21*	2.57+-2.67*	8.50+-1.29*	1.49E-5*
	3.38+-3.398**	3.40+-3.29**	4.21+-4.24**	0.23**

Tabla 21- Distribución de los valores del MDS-UPDRS IV según los genotipos de los SNPs asociados a MDS-UPDRS IV. (ANOVA). En esta tabla se presentan los resultados de análisis de la varianza de los scores UPDRS IV según los genotipos para los SNPs SNCA-A53-T y rs104893877-CTSB. Columna primera: SNPs ID y los genes donde se localizan. Columna dos: homocigotas para el alelo menor. Columna tres: heterocigotas. Columna cuatro: homocigotas para el alelo menor. Columna Cinco: P-valor de la ANOVA. Los pacientes que tienen el alelo de riesgo para el SNP SNCA-A53T presentan significativamente las cifras de UPDRS IV más altas tanto en el grupo NOH como el grupo control. Por otro lado, los pacientes homocigotas para el alelo menor y que presentan además NOH tienen cifras UPDRS IV significativamente más altas que los pares del grupo control.

* Grupo con NOH

** Grupo sin NOH

En el análisis de asociación entre los SNPs y los valores de UPDRS IV, 18 SNPs fueron encontrados asociados con el UPDRS IV estadísticamente significativo (tabla 20 y 21). Por otra parte, en el análisis de ANOVA solo dos SNPs fueron encontrados tener valores de MDS-UPDRS diferentes entre los genotipos. Pero estas diferencias se notan tanto en el grupo con NOH como en el grupo sin NOH para el SNP SNCA A53T_rs104893877. Por esto, este SNP no será considerado para posteriores análisis y/o conclusiones. El SNP CTSB_rs1293298 presenta valores de MDS-UPDRS IV más elevados en los pacientes homocigotas para el alelo menor (8.50+-1.29, pv: 1.49E-5) en el grupo con NOH con respecto al grupo que no tiene NOH (pv: 0.23) (Ver tablas 20, 21).

En resumen, los SNPs encontrado asociados con UPDRS I además tienen cifras de MDS-UPDRS-I diferentes entre los grupos estadísticamente significativas son IL1R2-MAP4K4_rs34043159 y GCH1_rs11158026. Mientras tanto para el MDS-UPDRS III encontramos SNCA_A53T_rs104893877 y GCH1_rs11158026 y en el caso de MDS-UPDRS IV el SNP rs1293298-CTSB. Pero este SNP no se encontró asociado a NOH en los análisis anteriores.

SNPs	Variables	Coef	P-valor	Año
Age	NYH Scale	0.012	2.46E-5	BL
SNCA_A53T_rs104893877	NYH Scale	0.43	1.54E-4	BL
LRRK2_G2019S_rs34637584	NYH Scale	0.21	7.42E-4	BL
LRRK2_rs76904798	NYH Scale	0.17	8.20E-3	BL
NCKIPSD/CDC71/IP6K2_rs12497850	NYH Scale	0.10	0.03	BL
CTSB_rs1293298	NYH Scale	0.10	0.03	BL

Tabla 22- Análisis de asociación por regresión lineal de los SNPs y la escala NYH. En esta tabla se presentan los resultados de asociación entre los valores NYH y los SNPs. Primera columna: SNPs ID y los genes donde se localizan. Columna tres los coeficientes de regresión y columna cuatro los p-valores. Columna cinco: BL(baseline, primera visita). NYH: Escala Hoehn and Yahr.

SNPs	HDAM	HET	HDAME	P-valor
SNCA_A53T_rs104893877	1.63+-0.49* 1.59+-0.53**	2.00+-0.82* 2.11+-0.60**		0.20* 0.0042**
LRRK2_G2019S_rs34637584	1.67+-0.55* 1.55+-0.53**	1.75+-0.50* 1.81+-0.51**	2.00+-0.00	0.77* 2.00E-3
LRRK2_rs76904798	1.60+-0.50* 1.57+-0.53**	1.89+-0.60* 1.71+-0.54**	2.00+-0.00	0.17* 0.08**
NCKIPSD/CDC71/IP6K2_rs12497850	1.73+-0.59* 1.59+-0.54**	1.63+-0.49* 1.55+-0.53**	1.96+-0.44**	0.59* 2.20E-3**

Tabla 23- Valores de NYH Scale según los genotipos de los SNPs encontrados asociados a la escala NYH. En esta tabla se presentó un análisis de ANOVA con la media de las cifras de la escala NYH según los genotipos. Primera columna: SNPs y los genes donde se localizan. Columna dos: homocigotas para el alelo mayor. Columna tres: heterocigotas. Columna cuatro: homocigotas para el alelo menor. Columna cinco: p-valores.

*Grupo NOH

**Grupo no NOH

Hemos encontrado cinco SNPs asociados a la escala NYH(ver tabla 22 y 23). Por el otro lado, no se encontró diferencias significativas en cuanto a los valores de la escala de NYH según los genotipos en el grupo de los NOH.

Capítulo XI- Análisis de eQTL asociados a los SNP asociados a NOH.

En este apartado, vamos a presentar el análisis de eQTL, que nos permite estudiar el efecto de las variantes genéticas sobre la expresión de los genes en el tejido cerebral. Hemos

usado datos de expresión génica provenientes de células dopaminérgicas derivados de células madre pluripotente inducidas y datos genómicos de los pacientes. La idea con este análisis es estudiar y descubrir los eQTLs (proteínas) asociados a cada SNPs que hemos encontrado asociados a NOH. Hemos obtenido de la base de datos del PPMI, un total de 45755 genes, los cuales fueron reducidos a 2828 después de realizar sobre ellos un análisis de componente principal(PCA) en KNIME para reducción de datos utilizando un workflow en Knime. Después se realizó un modelo de regresión lineal simple sobre cada par de genes-SNPs para estudiar los eQTLs asociados a los SNPs encontrados asociados a NOH en análisis anteriores ([ver tabla 6](#)). Los genes estadísticamente significativos (P -valor <0.05) son usados para hacer un análisis de protein-protein interaction incluyendo la proteína donde se localiza el SNP a estudiar en la aplicación STRING para estudiar alguna interacción funcional entre los genes y el SNP en estudio y también se realizaron análisis de enriquecimiento de dichos genes. Además, hemos realizado un análisis de Network para estudiar las interacciones entre los genes y extraer los genes claves. Para este análisis, en un primer instante usamos los 2838 genes para realizar una matriz de correlación por el análisis de correlación de Pearson. Después de eliminar los datos con P -valor mayor a 0.05 y valor de correlación mayor a 0.5 y menor a -0.5, decidimos cruzar los resultados con los eQTLs encontrados en el análisis de regresión lineal. Después, los resultados fueron usados para realizar el análisis de network en Cytoscape. Estos análisis fueron repetidos para los SNPs rs2280104(BIN3), rs55785911 (DDRGK1) y rs13294100 (SH3GL2) en esta tesis.

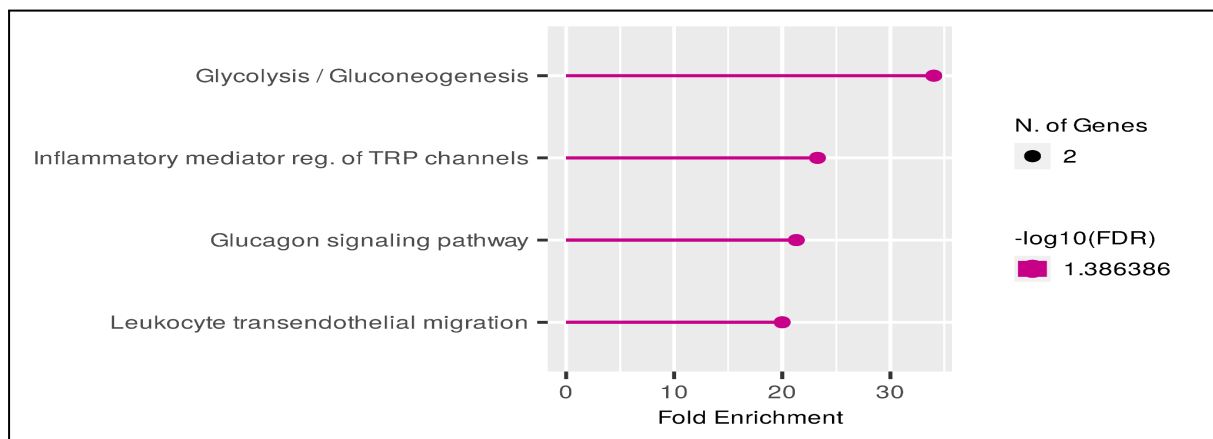
11.1- El SNP rs2280104(BIN3) asociado a 75 genes en analisis de eQTLs

Un total de 75 eQTLs fueron encontrados asociados al SNP rs2280104(BIN3)(p -valor <0.05). Este SNP tiene un efecto protector contra la NOH cuando están presentes las dos copias del alelo menor(transmisión recesiva) en los pacientes según análisis encontrados anteriormente(OR:0.1, p -valor:0.02)(tabla 5 y 6), mientras que en el análisis de dominancia genética no hubo significancia estadística. En tal caso, nosotros pensamos encontrar de los 75 genes cuales son aquellos cuya asociación con el SNP rs2280104-BIN3 involucran en proteger los pacientes contra la NOH y por qué mecanismos moleculares, vías de señalización y/o metabólicas. Para llevar a cabo esto, construimos un análisis de network basado en los puntajes de correlación de Pearson estadísticamente significativos (p -valor <0.05) (Ver resultado en apartado siguiente).

En el análisis de enriquecimiento de los 75 genes asociados a rs2280104-BIN3/NOH, hemos encontrado **enriquecimiento para vías metabólicas como vías de la glicolisis y**

gluconeogénesis, vías de señalización del glucagón, migración de leucocitos trans endoteliales. Si bien que las vías encontradas en el análisis son enriquecidas estadísticamente ($FDR < 0.05$), pero la cantidad de genes de nuestro set que participan en las mismas en relación al conjunto de genes que forman las vías es muy poco. Para esto se usa otra métrica que es el Fold enrichment que indica el porcentaje de genes del set de genes query perteneciente a una vía o pathway enriquecido. Como podemos ver en la imagen 1-A, los Fold-enrichments están por debajo de 50%. Este enriquecimiento encontrado no se va a considerar estadísticamente significativo. Además, en el análisis de interacción proteína-proteína, 17 proteínas fueron encontradas en la base de datos STRING de los 75 proteínas introducidos. De los 17 proteínas solo 4 fueron interconectados mientras que la aplicación nos cómputo unas interacciones esperadas a 2 con un p-valor de 0.13 (ver figura 17-A/B). Esto quiere decir que los resultados no fueron representativos con respecto a los datos introducidos.

A



B

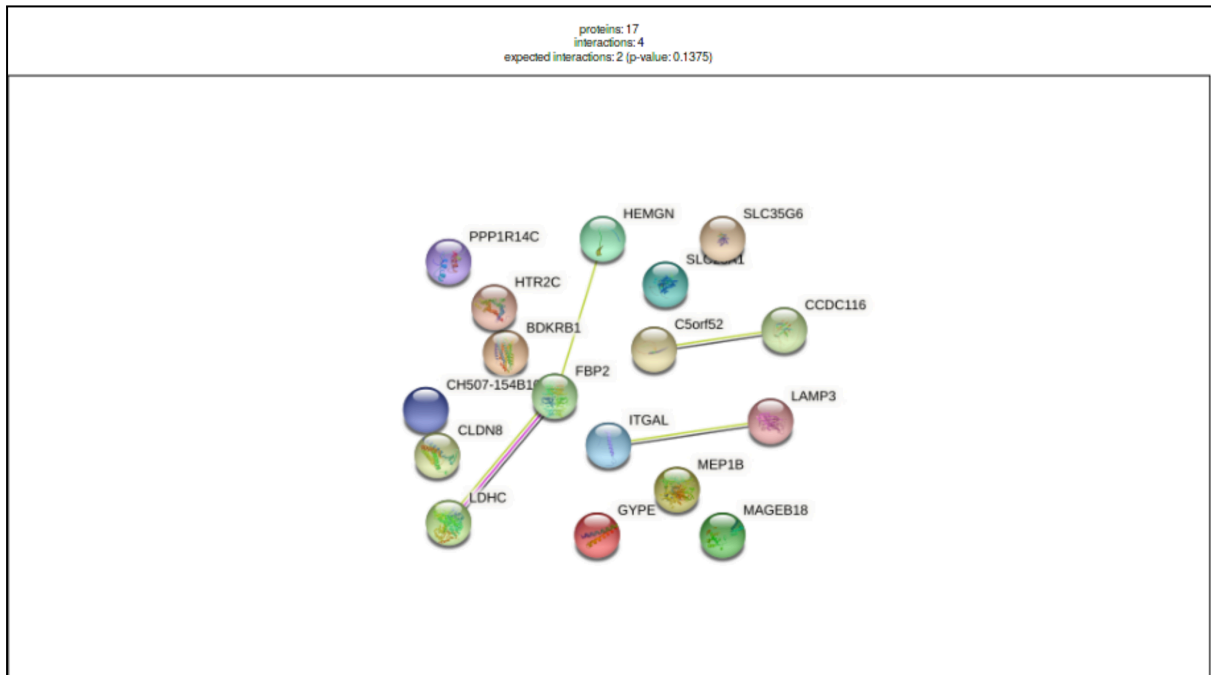


Figura 17A,B- Procesos moleculares enriquecidos para los eQTLs asociados al SNP rs2280104-BIN3. Análisis de enriquecimiento de los eQTLs asociados a NOH. Figura A: vías metabólicas encontradas a partir de los 75 eQTLs asociados a BIN3-NOH. Figura B: Protein-protein interaction in the STRING database.

11.2- Identificación de módulos y genes hub asociados con la NOH, riesgo disminuido por rs2280104-BIN3/PPP1R14C

Para este análisis, en un primer instante usamos los 2838 genes post-filtrados para realizar una matriz de correlación por el análisis de correlación de Pearson. Después de eliminar los datos con P-valor mayor a 0.05 y valor de correlación menor a 0.5 y mayor a -0.5, decidimos cruzar los resultados con los 75 eQTLs encontrados en el análisis de regresión lineal para extraer los genes que están correlacionados con los 75 genes(eQTL) encontrados asociados al SNP BIN3. Los datos resultantes se introdujeron en el software cytoscape para construir el network usando el paquete expresión correlation network de cytoscape. Para el network, los genes fueron representados por los nodos y los coeficientes de correlación por las conexiones(edges).

En el análisis de network explicado en el párrafo anterior, hemos encontrado 5 genes que se interrelacionan mayormente con los demás genes(ver figura 18 A-F). Se formaron 5 clusters a partir del análisis de cluster. El primer cluster está constituido de 325 nodos(cada nodo un gen) unidos al gen PPP1R14C(protein phosphatase 1 regulatory inhibitor subunit 14C) con mayor número de conexiones con los demás genes(unos 357 conexiones). En el

análisis de priorización, también los 5 genes mencionados aparecen como los más importantes en el network (**Figura 18B**). La aplicación ShinyGo(Ge, Jung, and Yao 2020) fue usada para hacer los análisis de enriquecimiento del cluster PPP1R14C. Hemos encontrado enriquecimiento estadísticamente significativo de genes para procesos como morfogénesis y desarrollo de dendritas, regulación de potenciales de membrana, morfogénesis de proyecciones neuronales, diferenciación neuronal, transmisión sináptica, exocitosis regulada por calcio, transmisión de neurotransmisores y muchos otros procesos involucrados en el crecimiento, desarrollo neuronal(Figura 18D).

Encontramos también que varios procesos moleculares participan tales como unión a microtúbulos, uniones a clatrina, uniones a enzimas quinasas, uniones a receptores GABA, uniones a proteínas del citoesqueleto(Ver figura 18C).

También habíamos realizado un análisis de interacción de proteína-proteína para ver las uniones funcionales que existen entre las proteínas, como se puede ver en la figura 18E; de los 325 proteínas de nuestro network, 309 se interconectan funcionalmente con 4691 conexiones observadas contra 2144 interacciones predichas(p-values: 0.0). Este análisis valida el análisis de network que hicimos basándonos en el análisis de correlación de Pearson; además nos indica que las conexiones entre las proteínas son altamente funcionales. Buscamos los genes hub en el network PPI. El network PPI generado en la base de datos STRING fue llevado a cytoscape para posterior análisis. Hemos encontrado 7 genes hubs los cuales enriquecen para procesos moleculares y celulares como el transporte de proteínas a lo largo de microtúbulos, regulación de ensamblaje postsináptico, morfogénesis celular envuelto en diferenciación celular, morfogénesis de proyección celular(FDR<0.05). Proteínas como NRXN(neurexina 1) que ayudan a conectar neuronas en la sinapsis, KIF5C(quinesina) que son una familia de proteínas motoras que median el transporte intracelular anterógrado sobre los microtúbulos que son componentes del citoesqueleto, SCN2A(subunidad alfa 2 del canal de sodio dependiente de voltaje), KIF1A(kinesina 1A) que también participan en el transporte a través de microtúbulos y que participan en procesos fisiológicos de supervivencia de las células y funciones cerebral altas, GAP43(proteína asociado a crecimiento) que participa en el desarrollo neuronal y regeneración axonal y por fin SNCB(sinucleína beta) que se sabe que es un inhibidor de la agregación de la alfa sinucleína y protege el cerebro contra la toxicidad de la alfa sinucleína en la EP(**figura 19A**).

Hemos encontrado que los 7 genes están enriquecido de forma significativa para vías metabólicas y/o de señalización de la secreción de la GnRH, sinapsis serotoninérgicas, sinapsis dopaminérgicas, vía de señalización del estrogen, vías de señalización de la oxitocina según la base de datos KEGG(**figura 19 A**). Por el otro lado, la base de datos

curado PANTHER enseño enriquecimiento para vías de señalización de la proteína G mediada por las subunidades alfa Gq y Go, vías de señalización del receptor muscarínico colinérgico 2 y 4(Ver figura 19B). Además se sabe que estas proteínas presentan funciones moleculares relacionadas con la presentación de dominio de unión a la familia de proteína neuroligina que son proteínas de adhesión celular encontradas en la membrana postsináptica y que se unen con la beta neurexina encontrados en la membrana presináptica para el mantenimiento de la sinapsis entre neuronas. También presentan dominio de unión a factores de crecimiento fibroblástico que son los cuales intervienen en el desarrollo y crecimiento neuronal. Otro proceso molecular importante que presentan estas proteínas es que presentan dominio de unión al ácido lisofosfatídico. El ácido lisofosfatídico(LPA, lysophosphatidic acid por su abreviación inglesa) es un fosfolípido involucrado en muchos procesos celulares a través de receptores acoplados a proteínas G específicas(Figura 19C). Diferentes estudios han demostrado la importancia de la LPA y los receptores a los cuales se une para llevar a cabo sus funciones moleculares en el cerebro por ejemplo proliferación y diferenciación celular, supervivencia celular, sinapsis y neurotransmisión en diferentes regiones del sistema nervioso central.

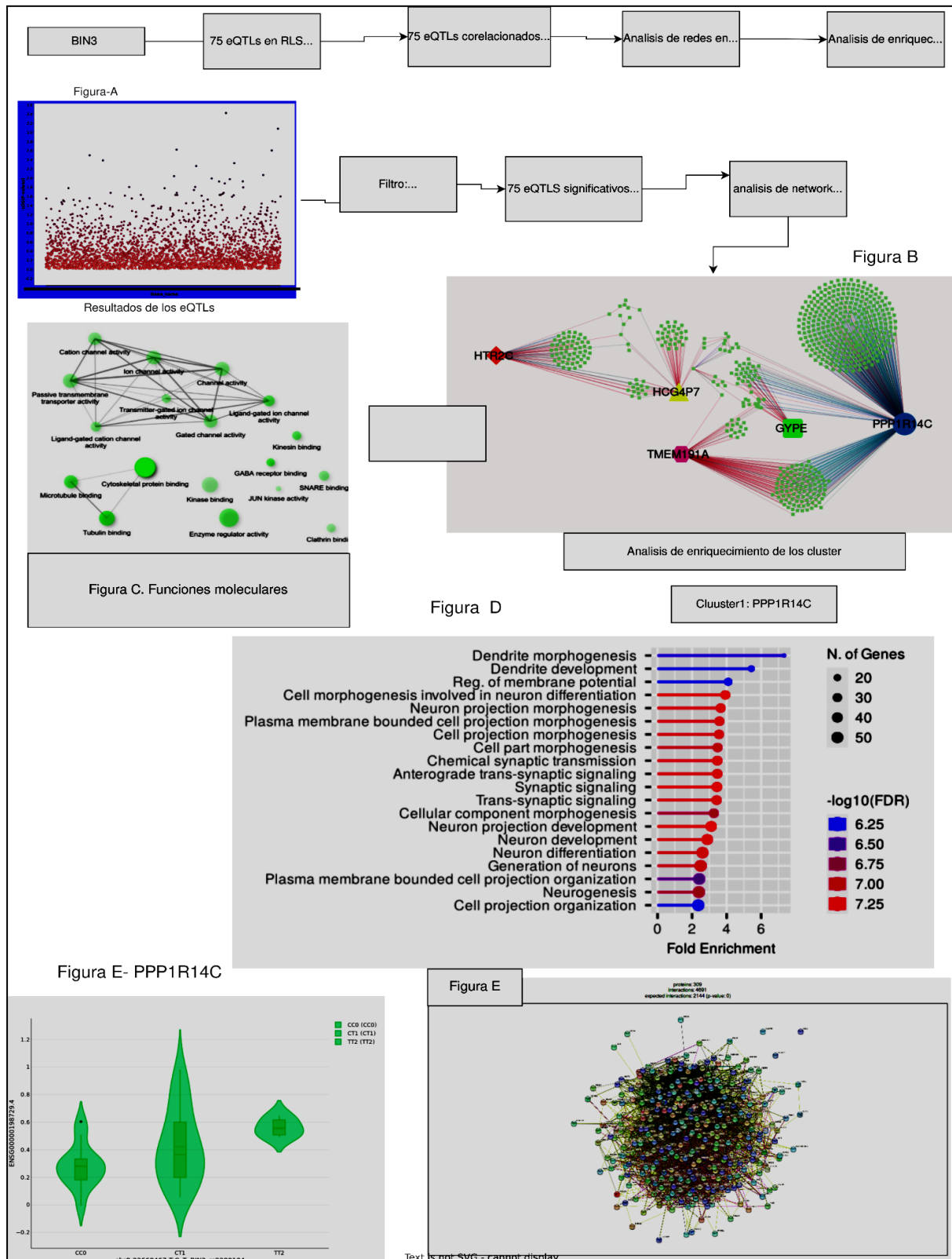
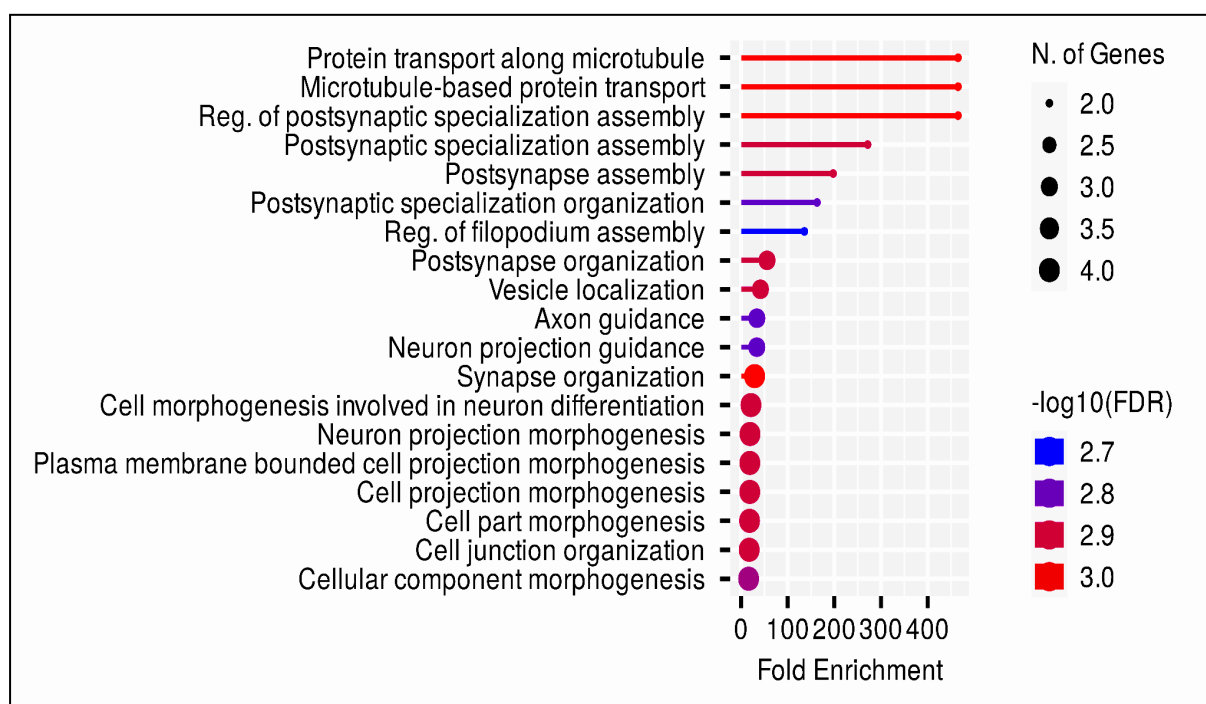


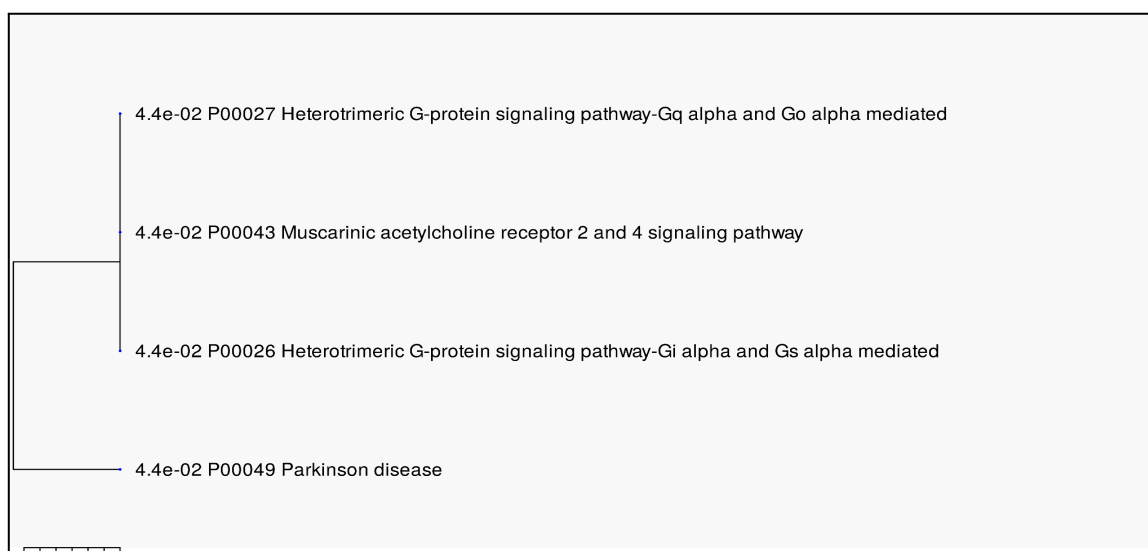
Figura 18(A-F)- Análisis de eQTLs asociados al SNP BIN3 como factor protector de la NOH. **Figura A:** Scatter plots de los eQTLs en función del $-\text{LOG}(p\text{-values})$ antes de filtrar (filtro: $p\text{-values} < 0.05$, Coeficiente de correlación > 0.5 y < -0.5). **Figura B:** Análisis de Network en cytoscape y priorización de genes. Los eQTLs encontrados con el SNP rs2280104-BIN3 fueron cruzados con los datos de correlación significativa entre los genes para captar conexiones entre los eQTLs y los genes correlacionados. Estos datos fueron utilizados para construir un network en Cytoscape. Se realizaron análisis de priorización y de cluster. En la **figura B** se presentan los 5 cluster encontrados. Uno de ellos, llamado cluster PPP1R14C tiene incluido un eQTL cuya expresión se ve influida por los genotipos del SNP-BIN3 y es el más importante dentro del cluster. Decidimos seguir los análisis con este cluster por esta razón. Además los demás clusters no tienen eQTL incluidos. El cluster está

constituido de 325 genes incluyendo al eQTL PPP1R14C. Este último está conectado con todos los 324 genes lo que le confiere una alta importancia y autoridad en el network. **Figura C:** Análisis de enriquecimiento en forma de network en cuanto a procesos moleculares: enriquecimiento significativo para procesos como uniones a microtúbulos, uniones a clatrina, uniones a enzimas quinasas, uniones a receptores GABA, uniones a proteínas del citoesqueleto. **Figura D:** Análisis de enriquecimiento en cuanto a procesos moleculares involucrados. En la figura d, se presentan los resultados de un análisis de enriquecimiento para procesos moleculares y el cluster PPP1R14C enriquece para morfogénesis y desarrollo de dendritas, regulación del potencial de membrana, morfogénesis de componentes y partes celulares, transmisión sináptica química, señalización anterógrada transináptica, desarrollo y generación de neuronas y proyecciones neuronales, diferenciación neuronal y neurogénesis. **Figura E:** Interacción Protein Protein en la base de datos STRING. Entramos en la base de datos STRING los 325 genes del cluster PPP1R14C, se mapeo 309 genes en la base de datos con alto número de conexiones **Figura F:** Análisis de expresión del gen target(PPP1R14C). Se presentó un análisis de ANOVA de la expresión del eQTL PPP1R14C según si el paciente es homocigoto para el alelo de riesgo o heterocigoto. Se nota que la expresión del eQTL PPP1R14C es significativamente más alta en los pacientes portadores del alelo de riesgo para el SNP-BIN3.

A



B



C

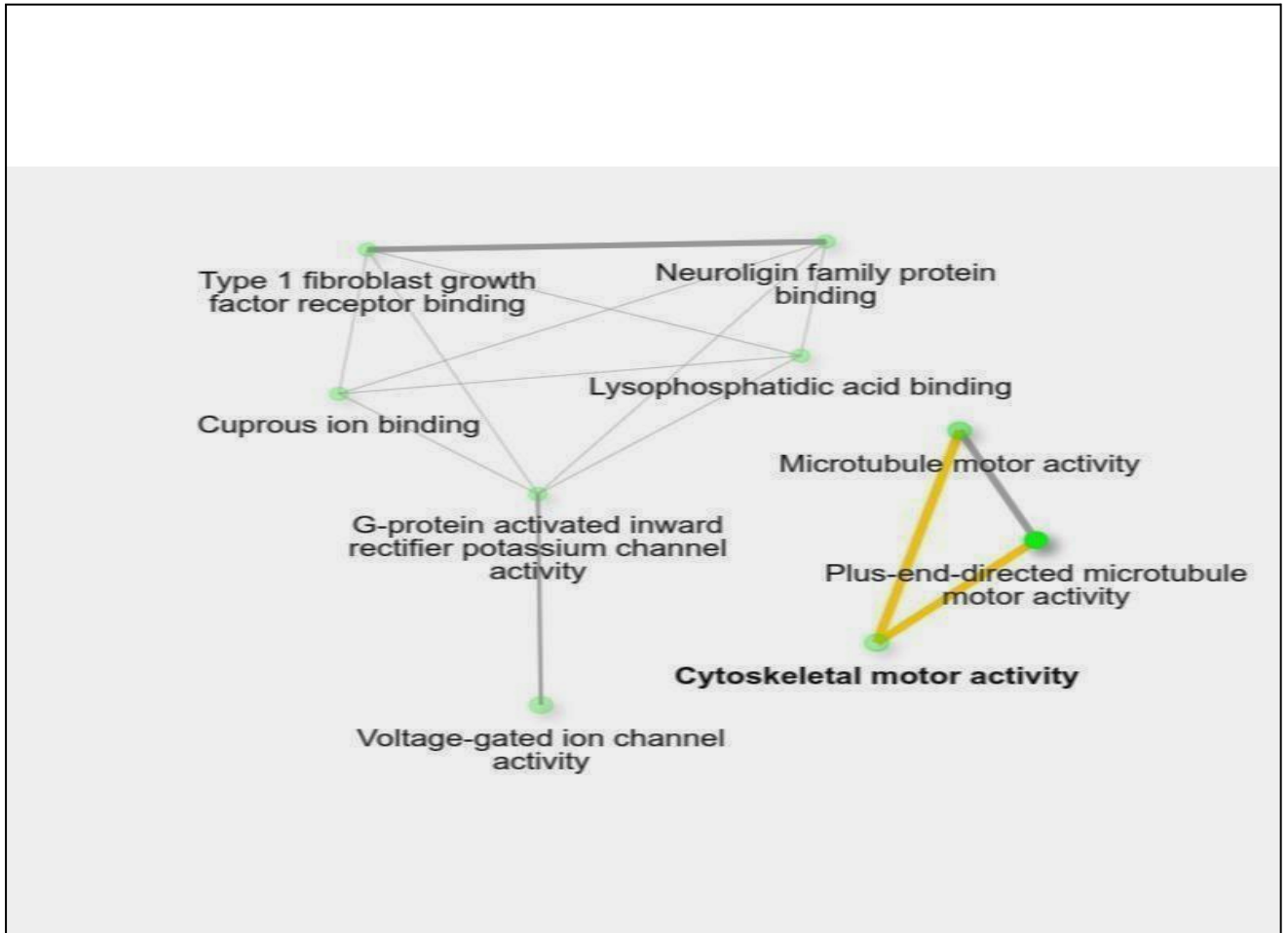


Figura 19(A-C)- Enriquecimiento de los 7 genes Hubs más importantes encontrados en el cluster PPP1R14C. **Figura A:** En esta figura se presenta el enriquecimiento de los 7 genes más importantes en el cluster PPP1R14C: transporte de proteínas a lo largo de microtúbulos, regulación del ensamblaje de especializaciones postsinápticas, localización de vesículas, morfogénesis celular involucrado en diferenciación neuronal, morfogénesis de componente y proyecciones celulares y neuronales. **En la figura B:** Enrique para vías de señalización como la vía de señalización de proteína G heterotrimérica mediada por Go y Gq alfa, vía de señalización de receptores de acetilcolina muscarínico 2 y 4, vía de señalización de proteína G heterotrimérica mediada por Gi y Gs alfa. **En la figura C,** presentamos la red formada por las vías de señalización y metabólicas del análisis de enriquecimiento de los 7 genes más importantes del cluster del eQTL PPP1R14C. Estos 7 genes enriquecen para procesos como actividad de canales iónicos gradiente dependiente, actividad motora de citoesqueleto, unión a receptores de factores de crecimiento fibroblástico tipo I, unión a proteínas familias de neuroliginas, unión a ácido lisofosfatídico, unión a iones de cobre. Los 7 genes hub son los siguientes: NRXN, KIF5C, SCN2A, KIF1A, GAP43, SNCB, PPP1R14C.

11.3- Procesos moleculares y celulares en los cuales participan BIN3-PPP1R14C

BIN3 (Bridging integrator 3), es un gen que pertenece a la familia de las proteínas con dominio BAR. Estos últimos son dominio de dimerización altamente conservada encontrado en proteínas involucradas en la dinámica de las membranas plasmáticas. Se localizan en zonas de mayor curvatura de las membranas para sensar las curvaturas de las mismas (Ren et al. 2006). El gen PPP1R14C es el eQTL encontrado asociado con el SNP

rs2280104-BIN3 con mayor importancia según el análisis de priorización. Los genotipos TT CT están asociados con una mayor expresión del gen PPP1R14C en relación al genotipo CC(p -valor <0.05)(figura 18E). Pareciera que el efecto del SNP rs2280104-BIN3 sobre este gen es de tipo dominante, una copia del alelo menor de este SNP es suficiente para aumentar proporcionalmente la expresión del mismo aunque se puede anotar que la expresión más alto en el caso del genotipo de riesgo TT(Figura 18E).

Hasta ahora hemos dilucidado que los genotipos TT del SNP rs2280104-BIN3 se asocian con una sobreexpresión del gen PPP1R14C, el cual a su vez está conectado con muchos otros genes que se involucran en procesos celulares y moleculares de crecimiento, proliferación de las neuronas. Además, hemos encontrado procesos involucrados como uniones a citoesqueleto, uniones a quinasas y a microtúbulos etc. El gen PPP1R14C es un inhibidor de la proteína fosfatasa 1(PP1). El sistema de fosforilación-desfosforilación juega un papel importante en la regulación del equilibrio que debe existir entre las proteínas fosforiladas y las no fosforiladas. Este sistema está regulado por las proteínas fosfatasas las cuales pueden desfosforilar las proteínas y cambiar las funciones de las mismas. Este mecanismo por cierto reversible regula una gran parte de procesos moleculares y funcionales de las células. Las proteínas fosfatasas 1(PP1) representante mayor del grupo de fosfatasas dependiente de treonina y serina, regula la síntesis proteica, división celular, transmisión nerviosa, transporte intracelular, contracción muscular y varios otros procesos celulares. Esta diversidad funcional es consecuencia de las uniones entre un número reducido de subunidades catalíticas con un número elevado de proteínas sobre las cuales la PP1 tiene efecto regulador.

11.4- Propuesta de vías de señalización por las cuales el BIN3 disminuye el riesgo de NOH.

En el modelo de regresión y el análisis ANOVA, se observó que la proteína BIN3 está vinculada a una mayor expresión del gen PPP1R14C (véase la Tabla 24). Este hallazgo plantea la cuestión fundamental del mecanismo subyacente a través del cual la proteína BIN3 influye en la expresión del gen PPP1R14C. Para abordar esta cuestión, se llevó a cabo una exploración de las posibles vías de señalización e interacciones proteicas.

Una búsqueda de BIN3 y PPP1R14C en la base de datos STRING no reveló una conexión directa entre ambas proteínas. Sin embargo, para investigar de forma exhaustiva las posibles vías, se introdujeron en la base de datos STRING 325 proteínas adicionales conectadas a la proteína PPP1R14C en el análisis de redes. Este conjunto de datos ampliado permitió un examen más exhaustivo de los posibles mecanismos a través de los cuales BIN3 puede influir en la expresión de PPP1R14C. La figura 20 presenta una vía de

señalización propuesta por la que BIN3 puede potenciar la expresión de la proteína PPP1R14C. El flujo de información dentro de esta vía implica a varias proteínas intermediarias, empezando por BIN3, pasando por ANK2 y NRXN1, y llegando finalmente a PPP1R14C. Esta vía propuesta proporciona un marco teórico para comprender cómo BIN3 puede influir en la expresión de PPP1R14C.

A partir de nuestro análisis se identificaron dos posibles rutas de señalización. La primera ruta propuesta implica a BIN3-RABGAP1L-SYT11-NRXN1-PPFIA2-PPP1R14C, mientras que la segunda ruta consiste en BIN3-RABGAP1L-ANK2-NRXN1-PPFIA2-PPP1R14C (Ver figura 20). Estas vías dilucidan los mecanismos potenciales a través de los cuales las señales de BIN3 pueden transmitirse para influir en la actividad y expresión de PPP1R14C.

Para investigar más a fondo la relación entre los genes identificados y la PPP1R14C, se realizó un análisis de regresión. Los resultados revelaron que todos los pares de genes estudiados estaban significativamente asociados entre sí, proporcionando un apoyo adicional a las vías de señalización propuestas (Ver tabla 24).

En resumen, este análisis ha revelado posibles vías de señalización a través de las cuales la proteína BIN3 puede influir en la expresión del gen PPP1R14C. Estos hallazgos proporcionan valiosos conocimientos sobre los mecanismos moleculares que subyacen a esta relación reguladora y ofrecen una base para una mayor validación y exploración experimental. La investigación continuada en este campo promete descubrir nuevas dianas terapéuticas y estrategias para modular la expresión génica en contextos biológicos relevantes.

Dependent Variable	Independent Variable	Coeficiente	p-Values	R Square	RMSE
TBC1D30 ENSG00000111490	PPP1R14C ENSG00000198729	0.70	2.08E-12	0.90	0.07
SYT11 ENSG00000132718	RABGP1L ENSG00000152061	0.874	0.0	0.94	0.05
SYT11 ENSG00000132718	NRXN1 ENSG00000179915	0.76	0.0	0.58	0.15
SNCB ENSG00000074317	PPP1R14C ENSG00000198729	0.70	7.34E-13	0.80	0.11
SCN2A ENSG00000136531	PPP1R14C ENSG00000198729	0.72	3.81E-13	0.91	0.07
RGS 17 ENSG00000091844	PPP1R14C ENSG00000198729	0.68	1.58E-12	0.90	0.07
RABGP1L ENSG00000152061	SYT11 ENSG00000132718	0.893	0.00	0.80	0.11
RABGP1L ENSG00000152061	ANK2 ENSG00000145362	0.857	0.0	0.71	0.13

PPP1R14C ENSG00000198729	NRXN1 ENSG00000179915	0.58	9.74E-11	0.82	0.10
PPP1R14C ENSG00000198729	PPFIA2 ENSG00000139220	0.68	4.44E-16	0.89	0.08
PPP1R14C ENSG00000198729	KIF5C ENSG00000168280	0.7	0.00	0.64	0.14
PPP1R14C ENSG00000198729	SCN2A ENSG00000136531	0.69	3.81E-13	0.83	0.1
PPP1R14C ENSG00000198729	KIF1A ENSG00000130294	0.65	5.45E-13	0.87	0.08
PPP1R14C ENSG00000198729	GAP43 ENSG00000172020	0.76	0.00	0.91	0.07
PPP1R14C ENSG00000198729	SNCB ENSG00000074317	0.62	7.34E-13	0.75	0.12
PPP1R14C ENSG00000198729	PCMT 1 ENSG00000120265	0.60	1.10E-10	0.75	0.12
PPP1R14C ENSG00000198729	RGS 17 ENSG00000091844	0.60	1.58E-12	0.84	0.09
PPP1R14C ENSG00000198729	TBC1D30 ENSG00000111490	0.60	2.08E-12	0.83	0.1
PPFIA2 ENSG00000139220	NRXN1 ENSG00000179915	0.90	0.0	0.99	0.02
PPFIA2 ENSG00000139220	PPP1R14C ENSG00000198729	0.74	4.44E-16	0.51	0.17
PCMT 1 ENSG00000120265	PPP1R14C ENSG00000198729	0.60	1.10E-10	0.75	0.12
NRXN1 ENSG00000179915	SYT11 ENSG00000132718	0.80	0.0	0.58	0.15
NRXN1 ENSG00000179915	ANK2 ENSG00000145362	0.87	0.0	0.97	0.04
NRXN1 ENSG00000179915	PPFIA2 ENSG00000139220	0.90	0.0	0.99	0.02
NRXN1 ENSG00000179915	PPP1R14C ENSG00000198729	0.62	9.74E-11	0.85	0.09
KIF5C ENSG00000168280	PPP1R14C ENSG00000198729	0.91	0.00	0.91	0.07
KIF1A ENSG00000130294	PPP1R14C ENSG00000198729	0.65	5.45E-13	0.88	0.08
GAP43 ENSG00000172020	PPP1R14C ENSG00000198729	0.80	0.00	0.93	0.06
BIN3 ENSG00000147439	RABGAP1L ENSG00000152061	0.66	8.88E-16	0.88	0.08
ANK2 ENSG00000145362	RABGP1L ENSG00000152061	0.832	0.0	0.74	0.12

ANK2 ENSG00000145362	NRXN1 ENSG00000179915	0.82	0.0	0.96	0.05
-------------------------	--------------------------	------	-----	------	------

Tabla 24- Análisis de regresión entre el gen PPP1R14C y los genes priorizados como más importantes del network. En esta tabla se presentan los resultados de un análisis de asociación entre los genes más importantes en el cluster PPP1R14 y el gen PPP1R14C. Primera columna: genes actuando como variable dependiente. Columna dos: genes actuando como variable independiente. Columna tres: coeficientes beta de regresión. Columna cuatro: p-valores. Columna cinco: R cuadrado. Columna seis: RMSE(raíz del error cuadrado medio). Partiendo de los datos obtenidos anteriormente donde el alelo de riesgo del SNP-BIN3 aumenta la expresión del gen PPP1R14C(análisis de eQTL), se observa en estos resultados que el gen PPP1R14C a su vez altera la expresión de los genes como KIF1A, NRXN1, SCN2A, SNCB, KIF5C, SYT11, RABGAP1L, ANK2, GAP43, PCMT1, PPFIA2, TBC1D30, RGS17.

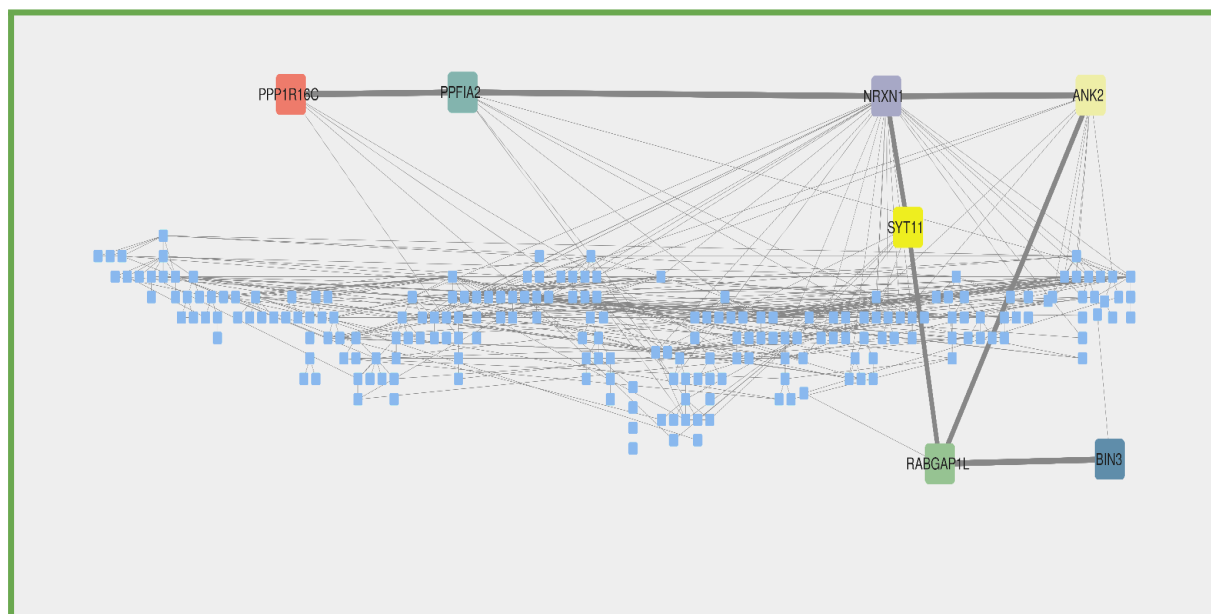


Figura 20- Análisis de redes de las proteínas extraídas de la base de datos STRINGs. Estas proteínas son los genes que hemos encontrado en análisis de redes de eQTLs. Fueron extendidos a la base de datos STRING. Usando Cytoscape se computo el camino más corto que lleva desde la proteína BIN3 hasta el inhibidor de la fosfatasa PPP1R14C. Se pudo identificar dos posibles vías teóricas:

BIN3-RABGAP1L-SYT11-NRXN1-PPFIA2-PPP1R14C / BIN3-RABGAP1L-ANK2-NRXN1-PPFIA2-PPP1R14C.

11.5- Los SNP rs55785911 (DDRKG1) está asociado a múltiples genes en análisis de eQTLs.

En el análisis de regresión lineal simple donde los genes fueron considerados como variables dependientes y el SNPs rs55785911-DDRGK1 como variable predictor para estudiar el efecto que tiene este último sobre la expresión de los genes. Se han encontrado un total de 469 genes eQTL significativamente asociados al SNP en estudio. Los genes fueron usados en un análisis de ANOVA para determinar cuáles de ellos los genotipos de riesgo verdaderamente influyen en su expresión.

De los 469 genes encontrados seleccionados, 17 fueron influenciados en su expresión por los genotipos homocigotas y heterocigotas para el alelo menor (p -valor <0.05). Estos 17 genes, además que fueron asociados al SNP rs55785911-DDRGK1 en cuestión en la regresión lineal, sino también en el análisis de ANOVA, se encontró que los genotipos influyen las expresiones de los mismos. También hicimos un análisis de correlación entre

los 17 genes y se encontraron que los mismos son significativamente correlacionados entre sí con un $R > 0.75$ (P -valor < 0.05). Además, se hizo otro análisis de regresión lineal entre cada par posible de genes para estudiar cómo se relacionan los mismos. Los datos obtenidos fueron usados para construir un network basado en relaciones lineales existentes entre cada par de los 17 genes. En el análisis de enriquecimiento de los 17 genes, no se encontró ningún proceso molecular en las bases de datos consultados (KEGG). Usamos la aplicación Causal-PATH en Cytoscape, la cual nos permite priorizar y seleccionar a 3 genes dentro del network con un potencial función en el proceso patológico molecular del SNPs rs55785911/DDRGK1/NOH. Los 3 genes son los que más conexión establecen dentro del network (Ver figura 21 B) con un network degree alto (11, 11, 10 respectivamente con respecto a otras conexiones establecidas por otros nodos del network). Se trata de los genes NOTCH1, KIF4A, TSPYL2. Las dos primeras se expresan significativamente más en los pacientes con genotipos homocigotas y heterocigotas para el alelo menor encontrado asociado con NOH en un modelo dominante en nuestro estudio (ver tabla 24, figura 21 C, D) (p -valor: 0.01 y 0.02 respectivamente). Por el otro lado, la expresión de TSPYL2 está subexpresada en los pacientes con genotipos homocigotas y heterocigotas para el alelo de riesgo (Ver figura 21 E, y tabla 24).

Genes	ANOVA-P Valor	Log2FC HeAm	Log2FC HAm	Efecto
NOTCH1	0.02	0.37	1.11	Expresión aumentada
KIF4A	0.01	0.88	1.36	Expresión aumentada
TSPYL2	0.02	-0.32	-0.96	Expresión disminuida
CDCA2	0.01	1.17	1.59	Expresión aumentada
ESCO2	0.02	0.79	1.65	Expresión aumentada
PBK	0.03	1.22	1.63	Expresión aumentada
CENPI	0.03	0.60	1.70	Expresión aumentada

Tabla 24- Expresión de los genes bajo la influencia del SNP rs55785911-DDRGK1.

A

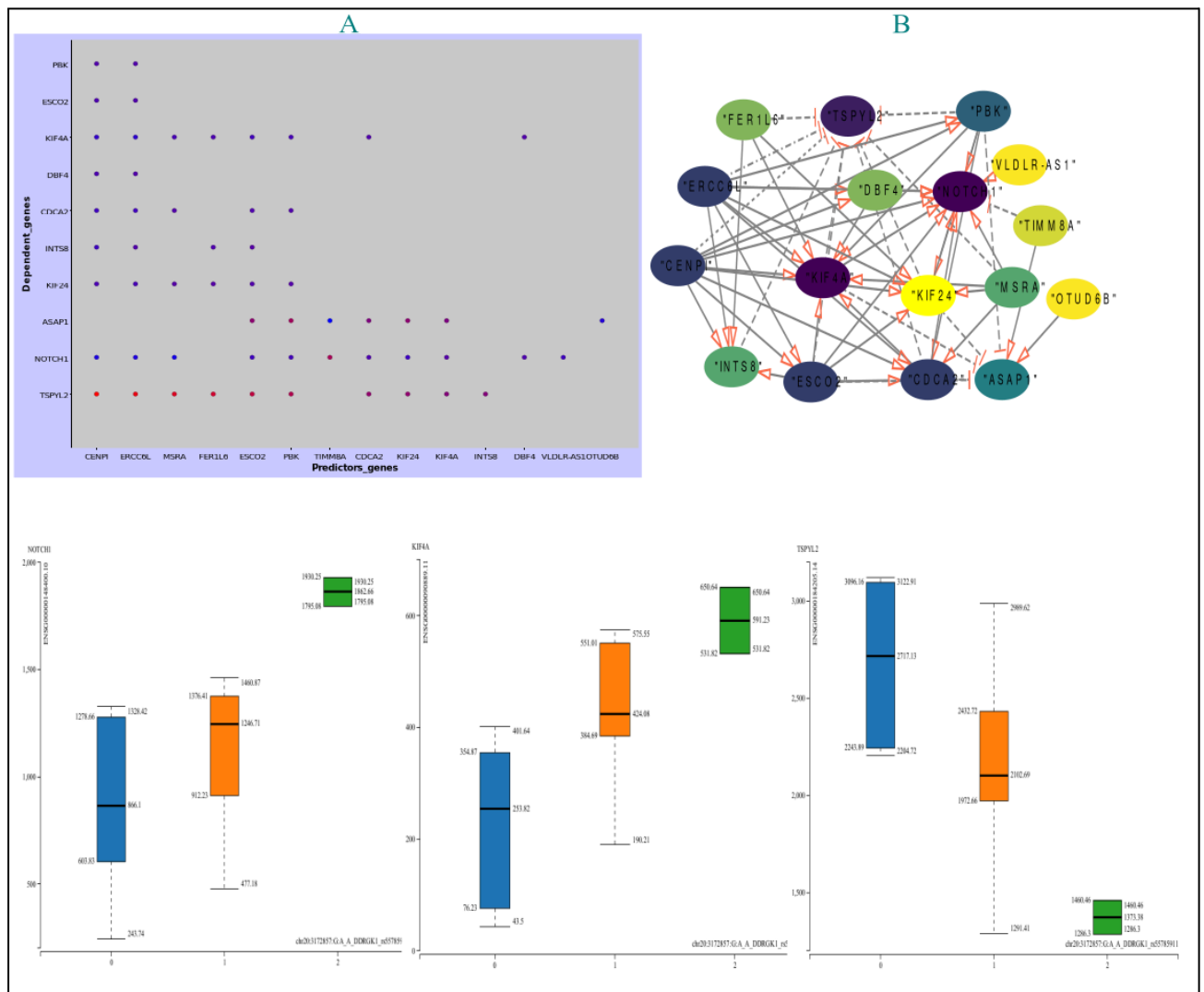


Figura 21. A: ScatterPlot de los coeficientes de la regresión lineal entre los 17 genes eQTL cuyas expresiones son influenciados por los genotipos homocigotas y heterocigotas asociados a NOH. Los puntos rojos y azules claros indican asociación negativa y positiva entre los genes. **B:** Network que determina la asociación entre los 17 genes usando el coeficiente de regresión lineal como parámetro de asociación. Cada esfera representa una proteína y las conexiones o edge representa el coeficiente de regresión lineal entre las proteínas. La flecha terminada en punto afilado indica que la proteína tiene un efecto positivo sobre la otra mientras que la flecha terminada en barra horizontal indica que el efecto es negativo. El color de las esferas indica el grado de conexión establece una proteína dentro del network, lo que vendría a decir su importancia dentro de la red. Las esferas de color púrpura tienen un network degree alto mientras que los amarillos tienen menos conexiones dentro de la red. Las azules y las verdes tienen conexiones de grado intermedio dentro del network. **C, D, E:** Boxplot que demuestran la expresión de los genes NOTCH1(C), KIF4A(D), TSPYL2(E) con respecto a los genotipos del SNPs rs55785911/DDRGL1. Las dos primeras boxplot para los genes NOTCH1 y KIF4A y el último boxplot para el gen TSPYL2. caja azul: homocigotas para el alelo mayor. Boxplot anaranjado: heterocigotas. Boxplot verde: homocigotas para el alelo menor(alelo de riesgo).

11.6- El SNP rs13294100 (SH3GL2) está asociado a múltiples genes en análisis de eQTLs.

Las mismas metodologías realizadas para el caso del gen DDRGL1 fueron utilizadas para analizar el efecto del SNP **SH3GL2_rs13294100** sobre la expresión de los genes utilizando

el análisis de regresión lineal entre SNP y genes (Ver apartado anterior). Fueron encontrados 1227 genes cuyas expresiones están asociados al SNP SH3GL2_rs13294100 (p-valor y FDR <0.05). Hemos utilizado los 1227 genes en un análisis de ANOVA para estudiar el efecto de los genotipos del SNP **SH3GL2_rs13294100** sobre la expresión de los mismos. Después del análisis de ANOVA, se recolectó un total de 585 genes eQTL estadísticamente significativos (p-valor y FDR <0.05). Los 585 genes encontrados en el análisis de ANOVA fueron considerados para realizar otro análisis de correlación lineal de a pares entre los 585 genes para poder construir después un Network usando estos resultados. 273 eQTL (genes asociados al SNP rs13294100_SH3GL2) fueron encontrados altamente correlacionados entre ellos mismos con coeficiente de correlación lineal mayor a 0.7, p-valor menor a 0.05 y FDR menor a 0.05. Se realizó un análisis de network usando la aplicación Cytoscape basándose en el coeficiente de correlación entre los genes. Se eligieron para la construcción del network solamente los genes que se correlacionan fuertemente (Coeficiente de correlación mayor o igual a 0.7, False discovery rate menor a 0.05). Sin uso de ninguna herramienta de agrupación podemos apreciar cómo de forma natural los nodos se agrupan en 4 clusters que parecen trabajar de forma coordinada. Por eso antes de cualquier otro análisis, decidimos estudiar cada cluster realizando sobre cada uno un análisis de enriquecimiento. El primer cluster al cual llamamos módulo en la figura 22, está significativamente enriquecido por los procesos biológicos que tienen que ver con el splicing alternativo y procesamiento de ARNm. Del otro lado, el segundo cluster enriquece para procesos de regulación del procesamiento y metabolismo del ARNm y de NcRNA, splicing alternativo del ARNm vía la reacción de la transesterificación y vía spliceosoma, biogénesis transporte y localización del complejo ribonucleoproteico, regulación del transporte nuclear y nucleocitoplasmático. El módulo 3 enriquece significativamente para procesos de transporte transmembrana de protones contra gradiente electroquímico, transporte de electrones acoplado al transporte de protones, cadena de transporte aeróbico de electrones, síntesis de ATP acoplado a transporte de electrones, procesos de metabolismo del ATP, el ensamblaje del complejo NADH deshidrogenasa, ensamblaje del complejo I de la cadena respiratoria mitocondrial (Figura 22A-E).



Figura 22 A-E- Análisis de network y enriquecimiento de los eQTLs asociados al SNP SH3GL2_rs13294100. En esta figura se presenta el enriquecimiento de los cuatro clusters encontrados en el análisis de network de los eQTLs asociados al SNP SH3GL2. **Fig A:** Los módulos computados son presentados en la figura 22A. **Fig B:** módulo dos: enriquece para transporte del complejo ribonucleoproteico, transporte nuclear, procesamiento de ARN, splicing por spliceosoma y por transesterificación, biogénesis del complejo ribonucleoproteico, regulación del transporte, metabolismo de ARNm. **Fig C:** módulo 1: enriquecimiento para procesamiento de ARN, splicing de ARNm, metabolismo de ARN. **Fig D:** módulo 3: enriquece para procesos como transporte membranar de protones contra gradiente de concentración, respuesta a hiperoxia, transporte electrones mitocondrial, síntesis de ATP mitocondrial acoplado a transporte de electrones, ensamblaje del complejo NADH deshidrogenasa, ensamblaje del complejo I de la cadena respiratoria mitocondrial, fosforilación oxidativa, respiración aeróbica, cadena de transporte de electrones, respiración celular, transporte transmembrana de protones, proceso metabólico de ATP, energía derivada de oxidación de compuestos orgánicos, generación de precursores metabólicos y energía, organización mitocondrial. **Fig E:** procesamiento biosintético de fosfolípidos, procesamiento metabólico de fosfolípidos y procesamiento biosintético de organofosforados.

respiración celular, fosforilación oxidativa y organización mitocondrial. El módulo 4 enriquece para procesos de biosíntesis y metabolismo de fosfolípidos, lípidos y compuestos organofosforados(figura 22 A-E). Se puede apreciar también que los 4 módulos están altamente conectados entre ellos, lo que nos lleva a pensar que los procesos celulares que se acaban de nombrar están relacionados de alguna forma. Los mismos genes fueron

usados para realizar un análisis de interacción proteína-proteína utilizando la base de datos STRING. Se introdujo en la base de datos STRING el gen SH3GL2 en conjunto con los demás genes de los 4 módulos para ver si existe alguna conexión a nivel de proteína entre el gen SH3GL2 (donde se localiza el SNP asociado a los eQTLs) y los genes eQTL de dichos módulos. Podemos apreciar que existe conexión directa entre el gen SH3GL2 y el gen SYNJ1 el cual a su vez se conecta con el gen MBOAT2. Otras 5 proteínas junto con las dos anteriores conforman una red de proteínas que se encuentra también en el network de eQTL que hemos construido (ATP13A2, HSDL1, SYNJ1, MBOAT2, PITPNM1, CDS1, RBM26). Hemos utilizado otra herramienta de Cytoscape llamado Causal-Path, el cual nos ayuda a determinar los nodos que más importancia tiene para cada módulo en nuestro network. Pudimos encontrar 4 nodos más relevantes por la aplicación causal Path desde Cytoscape, un nodo en cada módulo, PUS1, KATNA1, MTATP6, SYT13 en los módulos 1, 2, 3, 4 respectivamente. Además, elegimos otros nodos con mayor grado (network degree) en cada cluster. Elegimos 5 en el primer módulo ("PABPC1L", "CCER2", "MUSTN1-ITIH4", "NAP1L4P1", "NDST2"), cinco en el módulo 2 (SLTM, PNN, IWS1, CWC27, "PRPF38B"), dos en el módulo 3 (MT-ND1, MT-ND6) y 5 en el módulo 4 ("KIAA0895", "UBQLN1", "WDR43", "CXADR", "DGCR8"). Se buscaron los nodos que tienen conexiones con los nodos seleccionados (los más importantes por causal path y los con mejores grados) ([Ver figura 23](#)). Con este análisis pudimos seleccionar los nodos más importantes de cada módulo y que además que conectan los clusters entre sí.

Al final, seleccionamos 25 genes más importantes y que son responsables de las conexiones entre los diferentes módulos. Los cuales fueron usados para realizar un modelo de regresión lineal. Para este análisis, incluimos también al gen SH3GL2 con el objetivo de estudiar el efecto que tiene el gen SH3GL2 sobre la expresión de todos los 25 genes seleccionados.

Los nodos están altamente relacionados en la regresión lineal. El gen SH3GL2 está asociado significativamente ($FDR < 0.05$) con 20 de los 25 genes seleccionados. Se puede además apreciar en la tabla 25 los diferentes nodos (eQTLs) que conectan los diferentes módulos en este análisis. Por ejemplo, el eQTL LAS1L, DGCR8, IKBKB conecta el módulo dos con el uno, CXADR, SYNJ1, IKBKB, GNE conectan el módulo cuatro con el uno, ATP13A2 conecta el tres con el uno (Tabla 25). El gen SH3GL2 tiene un efecto negativo sobre la expresión de algunos eQTLs es decir que están asociados con disminución de la expresión de los mismos como por ejemplo ATP13A2, LAS1L, NCDN mientras tanto aumenta la expresión de otros eQTLs tal como CXADR, SYNJ1, IKBKB, KIAA0895, CDK20 (Ver tabla 25). Además, los genes eQTL están diferencialmente expresados según si

se tratan de homocigotas para alelo mayor o si los pacientes son heterocigotas y homocigotas para el alelo menor. Como habíamos hablado en este texto, este SNP está asociado al NOH en un modelo dominante. Como se esperaría, no existe diferencia significativa en la expresión de estos genes entre los heterocigotas y homocigotas para el alelo menor mientras si existe diferencia estadísticamente significativa con el genotipo homocigota para el alelo mayor(Ver [tabla 26](#)).

Gene_Var indep	Gene_Var dep	Coeficiente	FDR	Modulo de Connexion
SH3GL2	ATP13A2	-1.01	0.001	1-3, 2-3
SH3GL2	LAS1L	-0.42	4.204E-6	2-1
SH3GL2	NCDN	-0.19	0.004	2-4
SH3GL2	CISD2	0.03	3.991E-5	2-4
SH3GL2	DGCR8	0.13	5.069E-4	2-1, 2-4
SH3GL2	CXADR	0.24	3.35E-4	4-1
SH3GL2	SYNJ1	0.28	6.667E-5	4-1, 4-2
SH3GL2	SDCCAG8	0.29	0.003	2-1
SH3GL2	IKBKB	0.41	1.237E-10	1-4, 1-2
SH3GL2	MT-ND6	0.45	1.763E-6	3-1,3-2
SH3GL2	LRRC24	0.62	6.74E-10	2-4
SH3GL2	KIAA0895	0.68	9.317E-6	2-4
SH3GL2	UBQLN1	1.06	5.851E-8	2-4
SH3GL2	GNE	1.24	1.49E-8	4-2, 4-1
SH3GL2	LMTK2	1.26	1.689E-11	—(LMTK2-SYT13)
SH3GL2	HNRNPH1	2.26	6.113E-5	2-4
SH3GL2	RASA4CP	3.06	3.039E-5	—

SH3GL2	MTATP6P1	3.34	0.012	3-2
SH3GL2	GOT2	7.72	0.039	2-4
SH3GL2	CDK20	8.92	3.059E-7	---

Tabla 25- Análisis de regresión entre el gen SH3GL2 y 25 eQTLs hub más importantes de los clusters. En esta tabla se presentan los resultados de un análisis de asociación entre el gen SH3GL2 y los 25 genes más importantes de los 4 clusters(determinados por priorización). Primera columna: variable independiente(SH3GL2). Columna dos: variable dependiente. Columna tres: coeficiente de regresión. Columna cuatro: p-valores. Columna cinco: las conexiones entre los clusters. Las cifras de la columna 4 quieren decir: 1:modulo 1; 2:modulo 2; 3:modulo 3, 4:modulo 4. Ejemplo explicativo para la primera fila: el gen SH3GL2 está asociado significativamente con el gen ATP13A2 disminuyendo su expresión y permiten la conexión entre los módulos tres y dos y los módulos tres y uno.

Genes	ANOVA-P Valor	Log2FC HAm	Log2FC HeAm	Efecto
ATP13A2	0.035	-1.005	-0.672	Expresión disminuida
LAS1L	0.008	0.893	0.580	Expresión aumentada
NCDN	0.027	-0.845	-0.351	Expresión disminuida
CISD2	0.039	-0.562	-0.430	Expresión disminuida
DGCR8	0.027	0.597	0.362	Expresión aumentada
CXADR	0.016	-1.191	-0.888	Expresión disminuida
SYNJ1	0.022	-0.444	-0.537	Expresión disminuida
SDCCAG8	0.046	0.719	0.220	Expresión aumentada
IKBKB	0.011	1.133	0.421	Expresión aumentada
MT-ND6	0.002	2.150	0.470	Expresión aumentada
LRRC24	0.018	-1.103	-0.626	Expresión disminuida
KIAA0895	0.0201	-0.661	-0.682	Expresión disminuida
UBQLN1	0.048	-0.349	-0.247	Expresión disminuida
GNE	0.049	-0.320	-0.361	Expresión disminuida
LMTK2	0.048	-0.377	-0.368	Expresión disminuida
HNRNPH1	0.037	0.541	0.432	Expresión aumentada
RASA4CP	0.002	-0.729	-1.325	Expresión disminuida
MTATP6P1	0.006	1.484	-0.030	Expresión aumentada en HAm
GOT2	0.004	-0.628	-0.439	Expresión disminuida
CDK20	0.011	-0.711	-0.481	Expresión disminuida
CYB561	0.007	-0.444	-0.668	Expresión disminuida
FASN	0.029	-0.854	-0.409	Expresión disminuida
BAZ1A	0.032	0.909	0.889	Expresión aumentada

SYT13	0.040	-1.686	-0.976	Expresión disminuida
NOP2	0.010	1.469	0.474	Expresión aumentada

Tabla 26- Expresión de los genes bajo la influencia del SNP rs13294100-SH3GL2. En esta tabla se presenta un análisis de ANOVA para estudiar cómo los genotipos del SNP-SH3GL2 afecta la expresión de los 25 genes más importantes de los 4 módulos. Se calcula además los Log2 fold Change de las expresiones de los eQTLs para los genotipos heterocigotas y los homocigotas para el alelo menor con respecto a los genotipos homocigotas para el alelo mayor. Primera columna: los genes. Columna dos: p-valores de la ANOVA. Columna tres: Log2 FoldChange de la expresión de las proteínas para los pacientes homocigotas del alelo menor y columna cuatro para heterocigotas del SNP-SH3GL2 con respecto al homocigota del alelo mayor del mismo SNP. Los genes como SYT13, FASN, CYB561, CDK20, GOT2, RASA4CP, LMTK2, GNE, UBQLN1, KIAA0895, LRRC24, SYNJ1, CXADR, CISD2, NCDN, ATP13A2 están subexpresados en los pacientes homocigotas para el alelo menor y heterocigotas del SNP-SH3GL2. Los genes como LAS1L, NOP2, BAZ1A, MTATP6P1(aumentado en homocigotas del alelo menor y disminuido en heterocigotas), HNRNPH1, MTND6, DGCR8, SDCCAG8, IKBKB están sobreexpresión en los pacientes homocigotas para el alelo de riesgo y heterocigotas.

Para estudiar cuales de los eQTLs analizados son de mayor importancia en todo el network, hicimos otro análisis del network a partir de los genes con mejores valores de centralidad de intermediación(Betweenness Centrality) y centralidad de cercanía(closeness Centrality). Los mismos fueron incluidos en un nuevo network incluyendo al eQTL Synj1 por haberse encontrado conectado con la proteína SH3GL2 en el análisis de interacción proteína-proteína(PPI). La idea es inferir el posible Path por el cual el SNP **rs55785911**-SH3GL2 estaría provocando las funciones celulares encontradas en el análisis de enriquecimiento sobre los eQTLs asociados a dicho SNP. Entonces, pensamos que este SNP está asociado con NOH en los pacientes con EP porque altera la expresión de grupos de genes(llamados eQTLs) los cuales participan en vías de señalización y procesos celulares y moleculares importantes. Las alteraciones a nivel de la expresión de los mismos provocan alteración de las vías de señalización y procesos celulares en los cuales participan. Después de aplicar un filtro en el cual elegimos los nodos con un valor de centralidad de cercanía mayor de 0.014, obtuvimos un total de 94 eQTLs. Genes claves encontradas en el análisis anterior también se encuentran en este nuevo enfoque: LAS1L, CXADR, SYNJ1, IKBKB, MT-ND6, KIAA0895. Buscamos, dentro de este nuevo network de 94 genes, los nodos conectados directamente al gen SYNJ 1 porque este último es el gen conocido a nivel proteína por estar conectado con el gen SH3GL2. Encontramos un cluster de 7 proteínas unidos al gen SYNJ 1 que son UBQLN1, FASN, SYT13, KIAA0895, SYNJ1, SERINC1, CXADR, ATP13A2. Se eligieron los 7 genes junto con otros conectados directamente con los mismos en el network. De esta forma nos quedamos con 34 genes de los 94 seleccionados al inicio. Al final se seleccionaron los nodos con mejor centralidad de cercanía del network(formado por 43 genes) que son un total de 18 genes incluyendo al SYNJ 1. En el análisis de regresión lineal de a pares podemos apreciar el efecto que cada gen tiene sobre la expresión de otros. Al final encontramos 5 genes que conectan el gen SH3GL2 con los 4 módulos presentados anteriormente: ATP13A2, LAS1L, KIAA0895, CXADR, IKBKB. El ATP13A2 conecta los módulos tres con el uno y tres con el dos. eQTL LAS1L conecta el módulo 2 con el uno. KIAA0895 conecta el módulo 2 con el cuatro.

CXADR conecta el módulo cuatro con el uno. IKBKB conecta el módulo uno con el cuatro y uno con el dos.

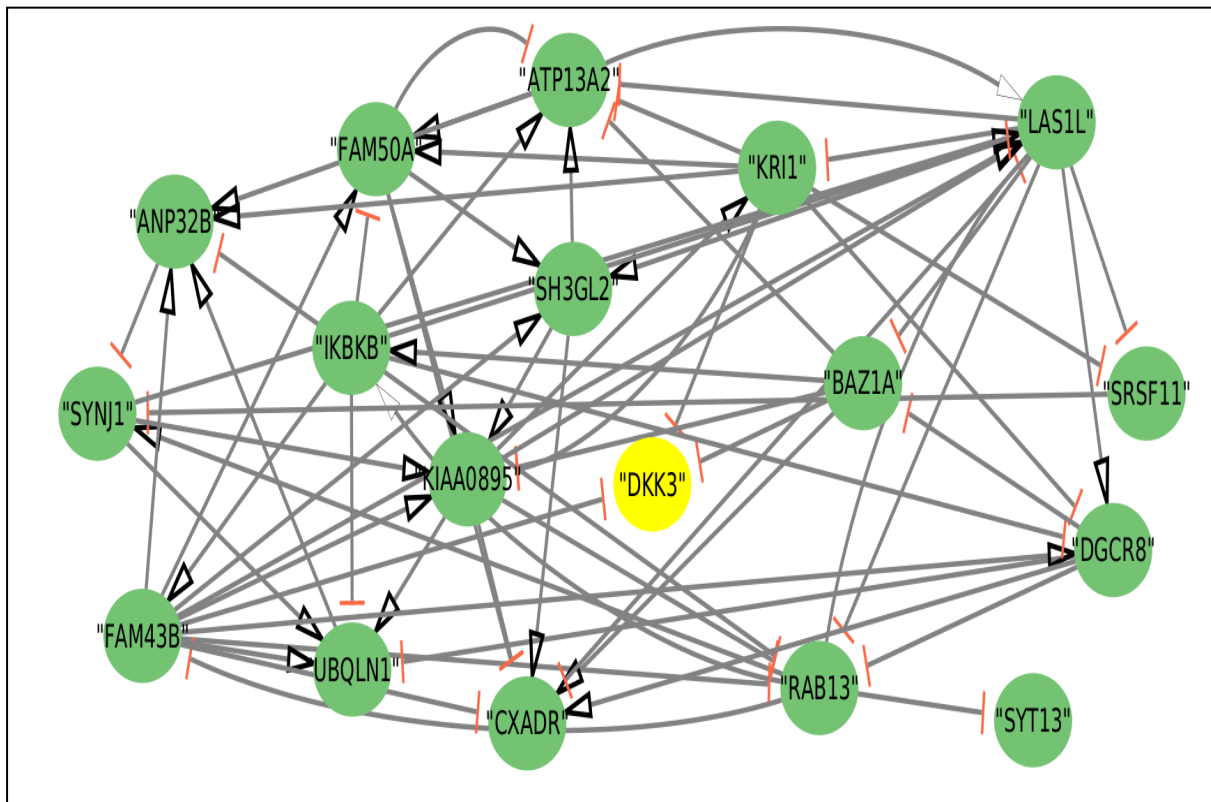


Figura 23- Clusters de eQTLs asociados al gen SH3GL2. En esta figura se encuentran los genes más importantes en cada módulo. Estos genes fueron obtenidos a partir de un análisis de eQTL entre el SNP-SH3GL2 y conteo normalizado y transformado de genes extraídos de células dopaminérgicas derivados de iPSC de pacientes con Enfermedad de Parkinson. La expresión de los eQTLs fueron encontrados influenciados según los genotipos del SNP-SH3GL2. Se seleccionaron los eQTLs en cada módulo con mayor importancia, mayor centralidad de cercanía, de intermediación y grado de conexión. Los mismos fueron usados para construir este network: los círculos son los genes, las flechas con punta quiere decir que esta proteína a la cual esta flecha apunta tiene aumentada su expresión bajo el efecto del gen de donde se origina el otro extremo de la flecha. Mientras que las flechas con barra roja quiere decir inhibición de la proteína hacia la cual apunta. En la tabla 25 se presentan los valores de la expresión de estos genes bajo la influencia del gen SH3GL2.

En resumen, los eQTL encontrados asociados con el SNP **rs55785911**-SH3GL2 forman 4 cluster los cuales enriquecen para procesos celulares bien definidos. Además, se encontró que el gen SH3GL2 donde se localiza el SNP está conectado con 5 genes los cuales tienen mayor importancia por permitir las conexiones entre los diferentes clusters y el gen SH3GL2. Estas proteínas se eligieron basándose en sus grados, sus centralidades en el Network. Esto nos permite añadir que estos 5 genes son puntos centrales en el network por los cuales el flujo de información tiene que pasar para viajar de cualquier nodo a otro. De ellos, los genes CXADR(4-1), SYNJ1(4-1, 4-2), IKBKB(1-4), KIAA0895(4-2) conectan el módulo 4 a los módulos uno y dos. Dicho de otra forma, estos 5 genes pueden constituir blanco terapéutico contra la NOH.

En la tablas 25 y 26, presentamos la expresión de los eQTL bajo la influencia del SNP **rs55785911**-SH3GL2. Se realizó un análisis de la varianza de una vía teniendo en cuenta los genotipos como variable dependiente y la expresión de los genes como independiente previa normalización de los datos. Se aprecian que los 5 eQTL (ATP13A2, LAS1L, KIAA0895, CXADR, IKBKB) se expresan de forma diferente según los genotipos(p -Valor <0.05). Se sacaron el Log2Change de la expresión en los genotipos heterocigotas y homocigotas para el alelo menor con respecto al homocigota del alelo mayor(tabla 26). Hay que recordar que el SNP **rs55785911**-SH3GL2 está asociado al NOH en un modelo dominante. Entonces, por eso buscamos cambios en la expresión de los genes de tanto heterocigotas y homocigotas para el alelo menor. Como cabría esperar, existen cambios significativos en la expresión de los 5 genes Hub en los genotipos asociados con la enfermedad con respecto al genotipo homocigota del alelo mayor. Además, se aprecia disminución de la expresión de los genes ATP13A2, CXADR, KIAA0895 mientras que existe sobreexpresión de los genes LAS1L, IKBKB(Ver tablas 25 y 26 y figura

24).

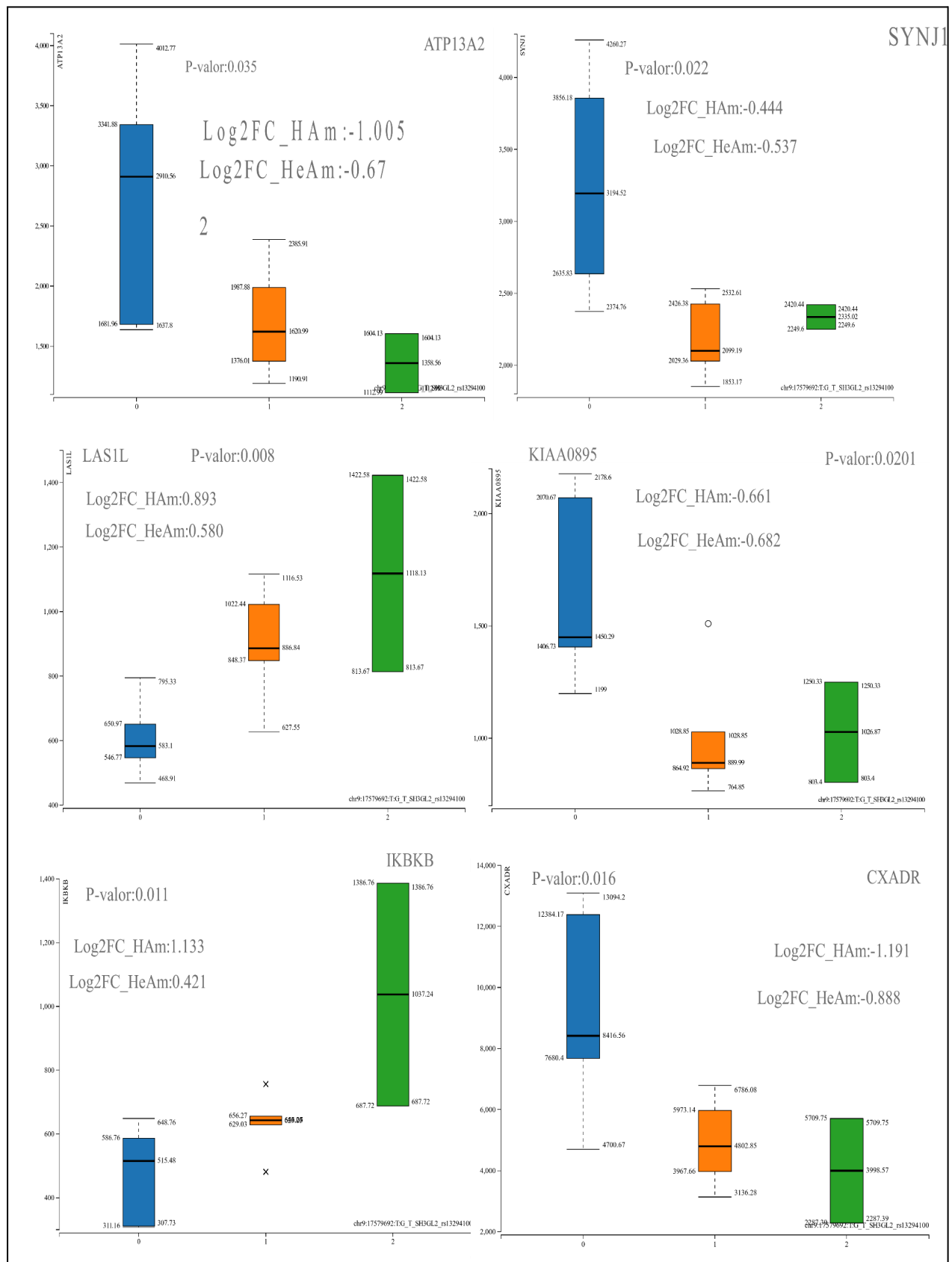


Figura 24- ANOVA de los eQTLs HUB con mayor influencia en el network. ATP13A2, SYNJ1, CXADR, KIAA0895 (expresión disminuida en heterocigotas y homocigotas para el alelo de riesgo del SNP-SH3GL2), LAS1L, IKBKB (expresión aumentada en pacientes homocigotas para el alelo mayor y heterocigotas para el alelo menor).

Capítulo XII- Análisis diferencial de genes entre NOH y no-NOH.

Para corroborar los resultados encontrados en la asociación entre SNP y NOH y además poder inferir sobre las vías de señalización involucradas en la aparición de la NOH en enfermos de Parkinson, decidimos realizar un análisis diferencial de genes. Usamos los datos procesados de conteo normalizado de genes de la base de datos del PPMI. Estos datos de ARN-seq provienen de células dopaminérgicas derivadas de células iPSC(induced pluripotent stem cells) que a su vez fueron obtenidas de células sanguíneas de los pacientes. Las células leucocitarias de la sangre extraída de los pacientes fueron inducidas a células pluripotenciales. Las últimas, después de soportar 65 días de tratamiento específico fueron transformadas a células dopaminérgicas. De las últimas fueron extraídos los ARNm, los cuales fueron secuenciados.

Los conteos normalizados de los datos de ARN-seq fueron obtenidos de la base de datos del PPMI. Hemos encontrado un total de 45574 genes disponibles en el dataset. Los pacientes fueron divididos en dos grupos, un grupo que tiene enfermedad de Parkinson con NOH y otro grupo que tiene enfermedad de Parkinson pero sin NOH. Cada uno de los grupos tiene 3 réplicas.

12.1- Pre-procesamiento de los datos

Hemos transformado los datos con el algoritmo “EdgeR: $\log_2(\text{CPM}+c)$ ”(Love, Huber, and Anders 2014) del paquete EdgerR embebido en la aplicación web iDEP(Ge, Son, and Yao 2018). Los ID de ensambl fueron anotados extrayendo los nombres de los genes y los datos fueron filtrados por el criterio default(0.5 count per millions). Obtuvimos un total de 17694 de los 45574 genes utilizados al inicio del análisis. Hemos observado una diferencia entre la librería NOH2 con respecto a las demás librerías específicamente con respecto a las del mismo grupo. Después de la transformación de los datos pudimos mejorar y obtener una varianza entre las librerías muy pequeña. Variaciones entre las réplicas fueron obtenidas muy pequeñas también($R=0.99$). Hay que recalcar aquí que se hizo un análisis de la varianza entre la media de los grupos y si el p-valor obtenido es menor a 0.01, se produce un mensaje automático que nos alerta que los datos tienen un factor confuso, tal no es nuestro caso. Se calcula también una razón entre los conteos totales máximo y mínimo de los diferentes grupos. Se cree que si la razón obtenida está dentro del valor 1.5, se puede aceptar los datos y seguir con el análisis de DGE(Ge, Son, and Yao 2018). Los datos

normalizados fueron utilizados para hacer otros análisis como Clustering y análisis de componente principal(Ver figura 25A y B).

12.2- Análisis de cluster y de componente principal

Se ha hecho un análisis de cluster jerarquizado para tener una idea somera de los datos usando los Top 1000 genes rankeados a partir de su normalización dividido por su desviación estándar. Este análisis nos permite especular que los grupos NOH y los no-NOH tienen bastante diferencia en la expresión de los genes seleccionados(Ver figura 25A). También se aprecia que las variaciones entre las réplicas no son mínimas. En el análisis de PCA, se puede apreciar en la figura 35B que el componente PC1 explica el 48% de las variaciones de los datos mientras que el segundo componente PC2 un 19% (Figura 25 B). El análisis de enriquecimiento de

Figura A

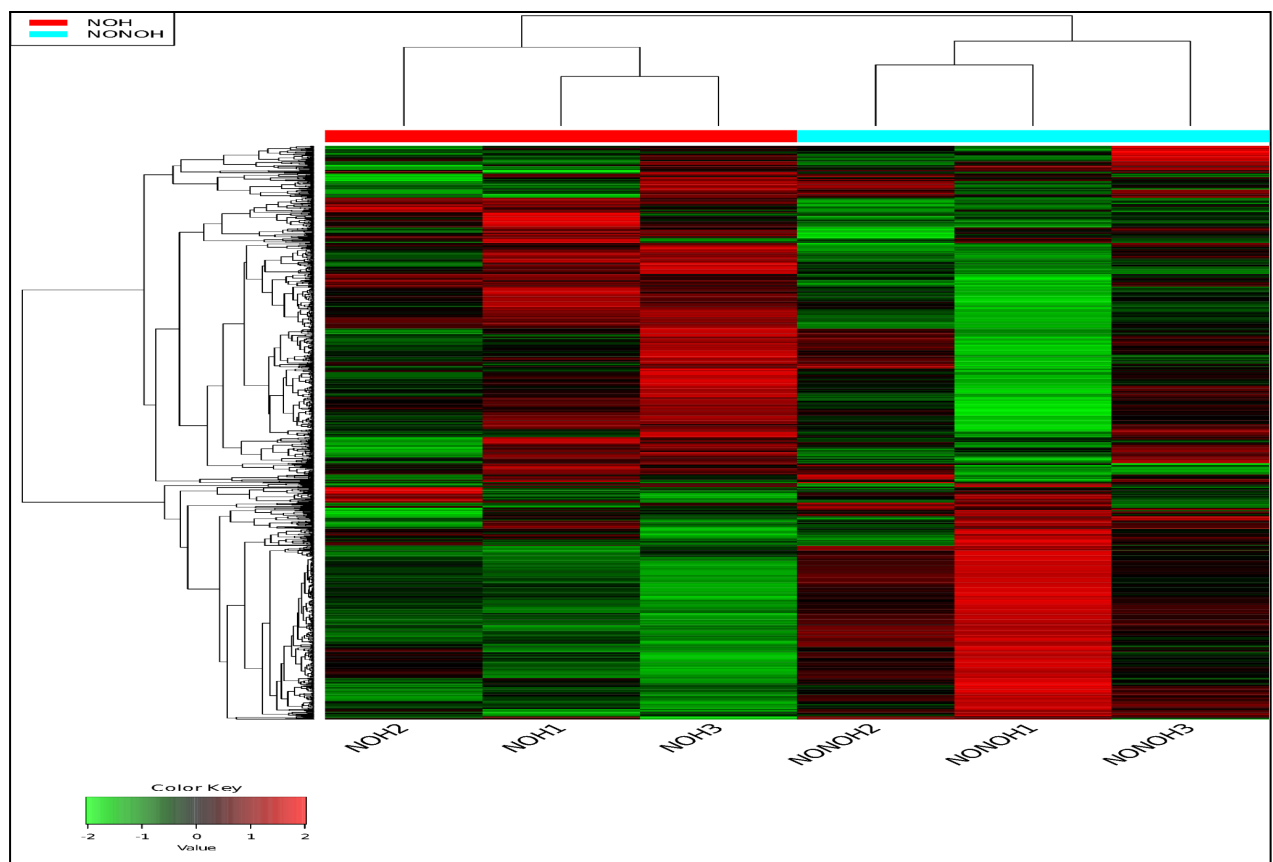


Figura B

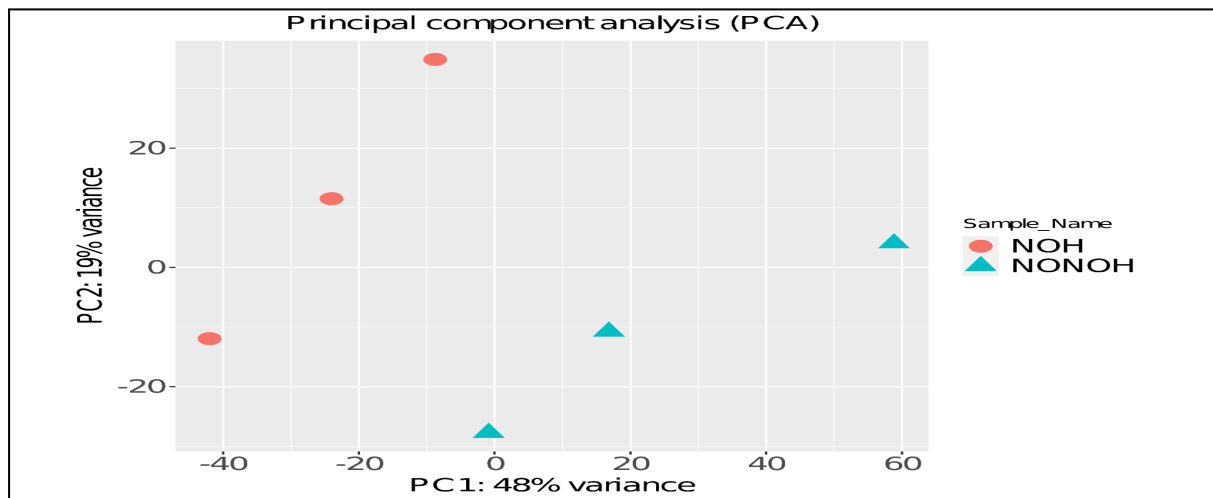


Figura 25- A: Heatmap de la expresión de genes en grupos NOH con respecto a los No-NOH(Cluster jerárquico). Se puede apreciar en la esquina abajo izquierda los 3 pacientes con NOH tienen un grupo de genes subexpresados(color verde) en relación al grupo control(No-NOH)(color rojo). B: Análisis de componente principal donde se aprecia cómo se separan los pacientes del grupo NOH del no-NOH. Los del grupo No-NOH se correlacionan con el componente principal PC1 y los del grupo NOH con el PC2.

los PC1 y PC2 informó un enriquecimiento para procesamiento de RNAm, transmisión química sináptica, reorganización cromosómica, señalización sináptica, secreción exportación desde la células mediado por vesículas, división celular, organización de microtúbulos, traducción de proteínas y biogénesis de ribosomas. Con esto podríamos inferir que los datos que más explican la variación de los datos tiene que ver con estos procesos celulares explicados anteriormente.

12.3- Análisis diferencial de genes y downstream analysis.

Análisis diferencial de genes entre los grupos NOH y NO-NOH se llevó a cabo usando el paquete Deseq2 de la plataforma iDEP. Hemos encontrado un total de 1438 genes diferencialmente expresados entre ambos grupos($FDR < 0.05$, $\text{Log}_2 \text{Fold Change} > 1$). De los cuales 592 genes sobreexpresados y 846 subexpresados(Ver figura 26A y B). Se han encontrado muchos genes diferencialmente expresados como lo demuestra las figuras 26A, B, C, D.

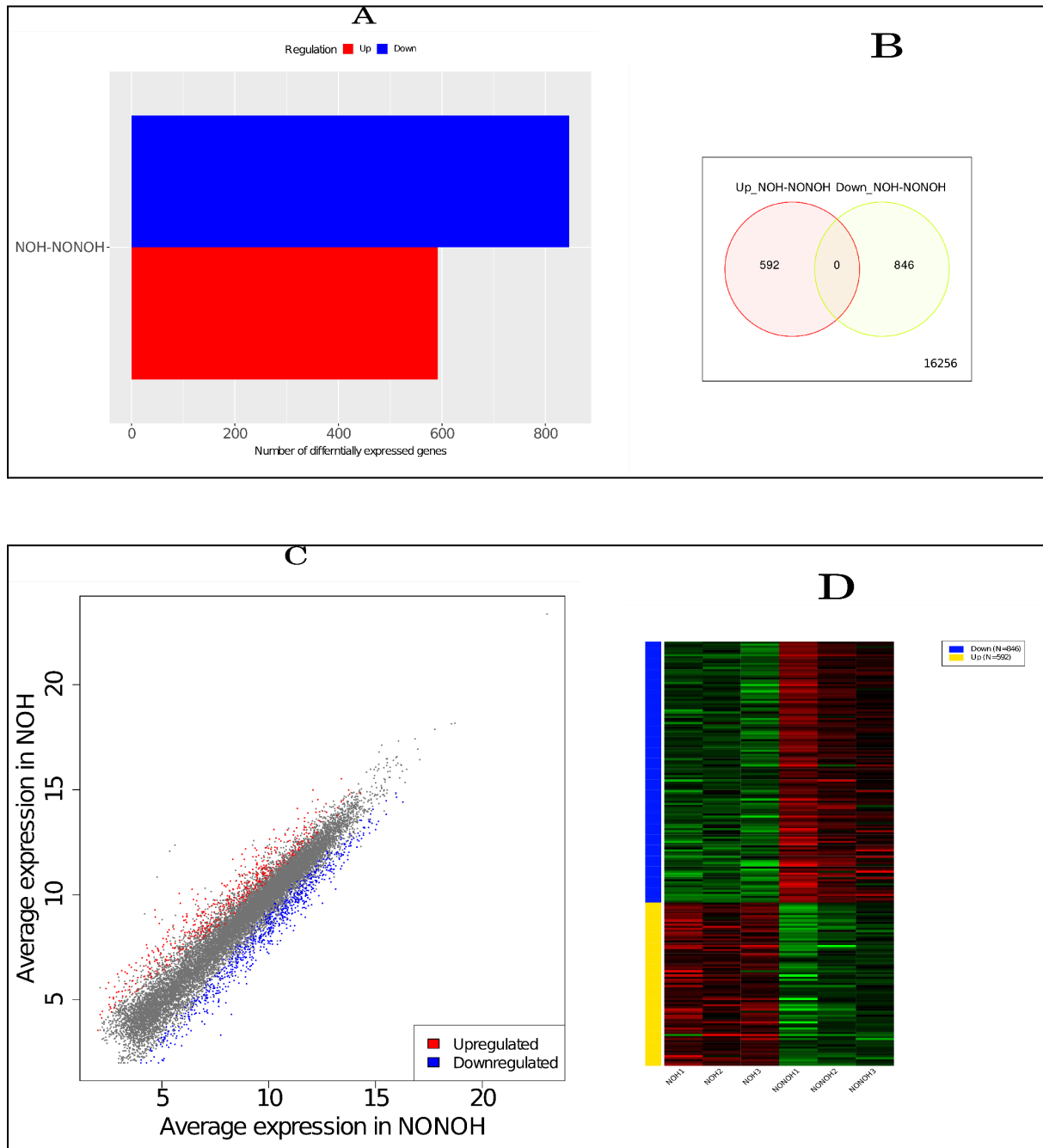


Figura 26 A-D: Análisis diferencial de genes entre NOH y No-NOH. Figura A: en azul están los genes que se subexpresan en el grupo con NOH y el rojo los que se sobreexpresan. B: se presenta un diagrama de venn que muestran 592 genes que se sobreexpresión y 846 genes que se subexpresan.

Los resultados de DEG fueron utilizados para ver los procesos moleculares por los cuales tienen enriquecimientos estadísticamente significativos. Encontramos que los genes bajamente expresados están enriquecidos para transmisión química sináptica, señalización sináptica anterógrada, regulación de exocitosis, liberación de señales desde las sinapsis, regulation de señal transinaptico, transporte sinaptico mediado por vesículas, secreción de neurotransmisores, regulación de la concentración de neurotransmisores, del otro lado los

genes sobreexpresados fueron encontrados enriquecidos para procesos de ensamblaje de axonema y proyecciones celulares, ensamblaje de organelas y cilium además de formación de microtúbulos, procesamiento de ARNm, Splicing alternativo vía spliceosoma y vía transesterificación, regulación del procesamiento de ARNm(Figura 27).



Figura 27- Enriquecimiento de los genes diferencialmente expresados en NOH con respecto al grupo No-NOH. (Procesos moleculares).

Por el otro lado, hemos encontrado enriquecimientos para vías de señalización y/o metabólicas como actividad de canales de transporte transmembranales, transporte pasivo, actividad de transporte de neurotransmisores en el caso de los genes bajamente expresados y enriquecimiento para los constituyentes estructurales de los ribosomas para los genes sobreexpresados(Figura 28).

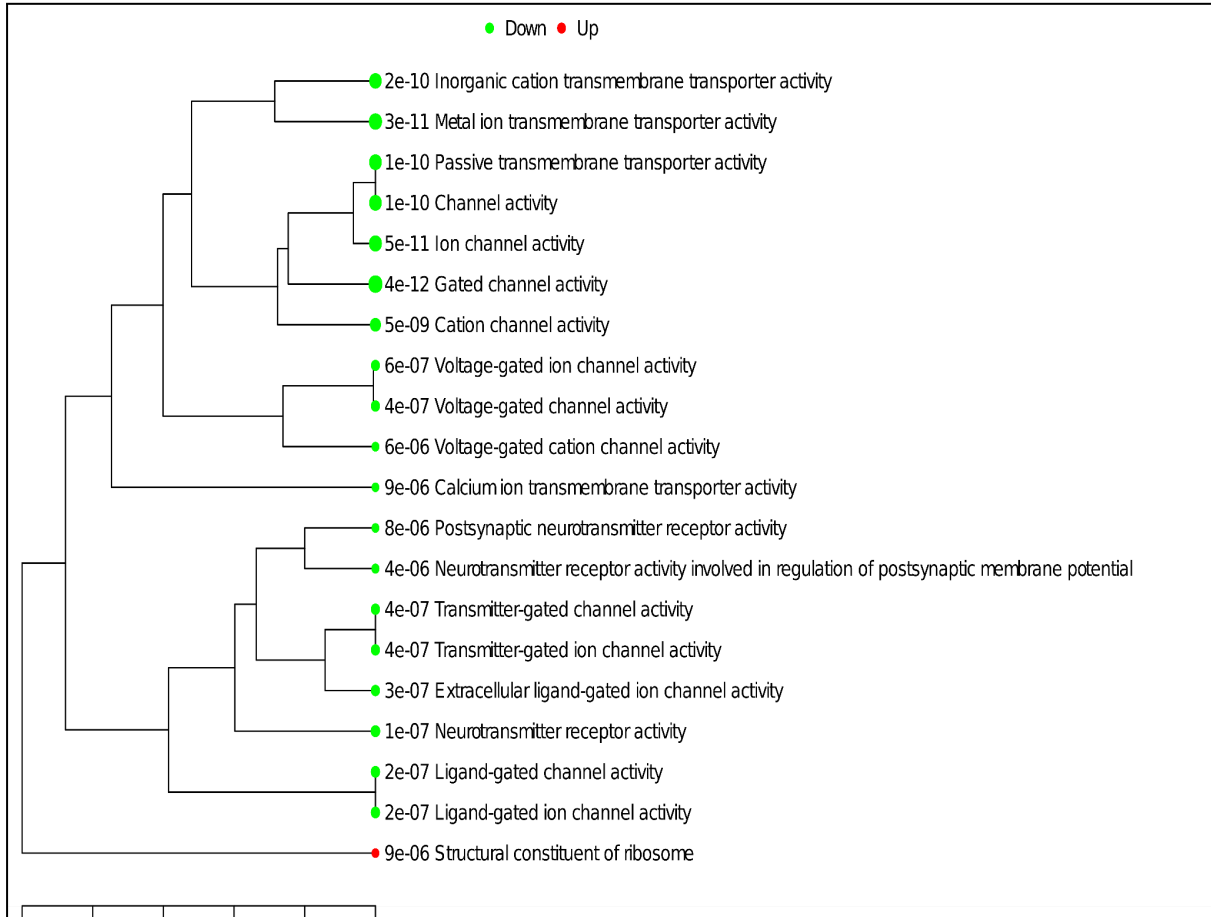


Figura 28- Enriquecimiento de los genes diferencialmente expresados en NOH con respecto al grupo No-NOH. (Vías de señalización y/o metabólicas KEGG).

También hemos estudiado las vías de señalización enriquecidas para los mismos genes diferencialmente expresados en pacientes con NOH con respecto a los controles(pacientes con Enfermedad de Parkinson sin NOH),

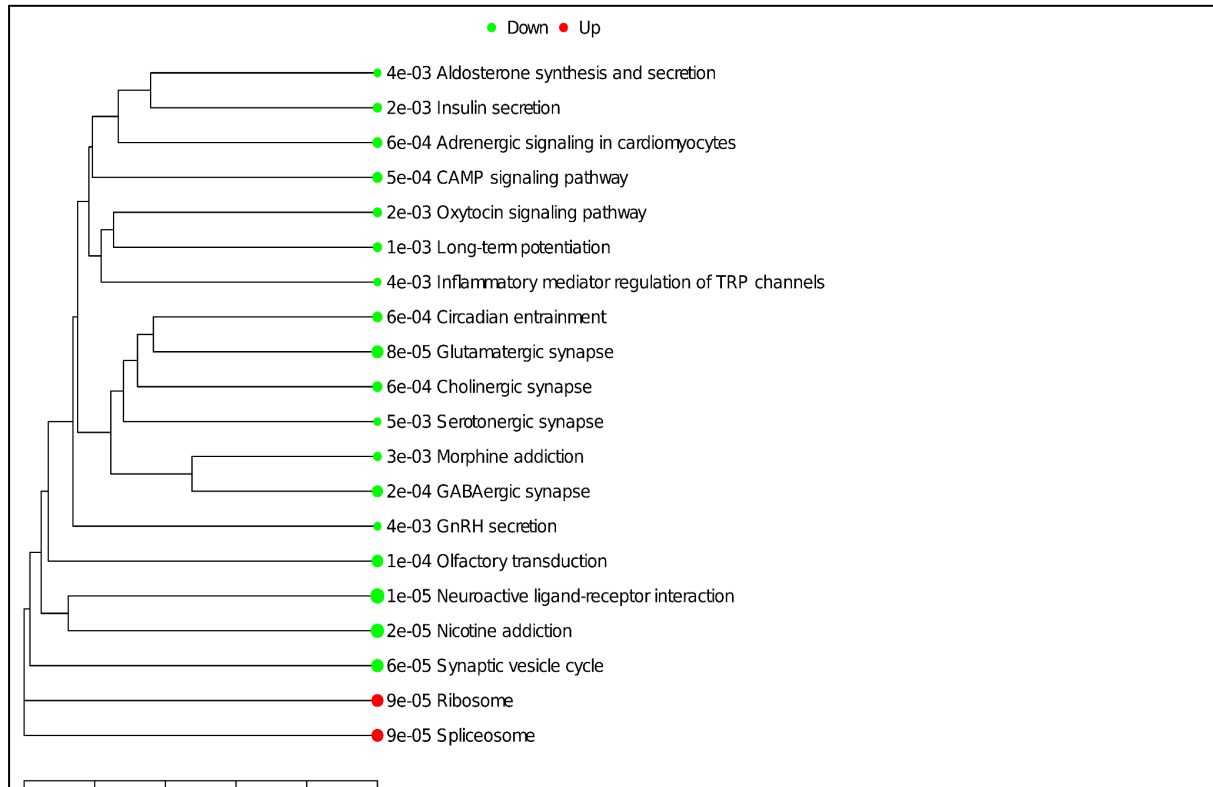


Figura 29- Enriquecimiento de los genes diferencialmente expresados en NOH con respecto al grupo No-NOH. (Vías de señalización y/o metabólicas) by KEGG database.

Hemos encontrado además que en los pacientes con NOH, los genes son bajamente expresados para los componentes celulares como la sinapsis, junction celular, proyección neuronal, membrana sináptica, axones, cuerpo celular, compartimiento somato dendrítica mientras que otros genes son sobreexpresados para estructuras como el nucleosoma, axonema, matrix extracelular, complejo proteína-DNA. Además de esto, usando STRING-DB hemos encontrado otros enriquecimientos no menores como generación y desarrollo de neuronas y neurogénesis.

Los genes fueron bajamente expresados para procesos moleculares y vías de señalización que tienen que ver con actividad de señales de receptor transmembranales, actividad de transporte de iones, transporte pasivo, actividad de receptor de neurotransmisores, voltage-gated channel activity mientras que los genes fueron sobreexpresados para pathway como actividad de inhibición de endopeptidasas, uniones a integrinas.

En la base de datos KEGG, encontramos que los genes bajamente expresados enriquecen para vías metabólicas interacción neuroactiva receptor ligando, sinapsis GABAérgico y Glutamatérgico, CAMP signaling pathway, sinapsis dopaminérgico y colinérgico, adicción a cocaína nicotina y morfina, ciclo vesícula sináptica, señalización adrenérgica. Mientras que por el otro lado se encuentra enriquecimiento para formación de trampas extracelulares de neutrófilos (Figura 29).

12.4- Down-regulación de las vías de señalización de transmisión sináptica

La vía de señalización de transmisión sináptica se encuentra sub-regulada en los pacientes con NOH. Este proceso celular incluye la liberación vesicular de neurotransmisores desde la membrana presináptica, la activación de los receptores de la membrana postsináptica, el efecto que tienen ambos procesos anteriores sobre el potencial de acción de la membrana postsináptica y en la composición iónica del citosol postsináptica. Los pacientes que tienen enfermedad de Parkinson y además NOH tienen mucho menos activados estos procesos de transmisión sináptica según nuestros datos. A continuación, presentamos los genes involucrados en este proceso de downregulation de la señalización de la transmisión sináptica:

SNCAIP, CREB1, GFAP, APOE, SYT4, AGT, DGKB, PPFIA2, SNCA, EGFR, STAT3, BDNF, LRRK2, SIPA1L2.

14 genes fueron encontrados estar involucrados en la disminución de la actividad de señalización de la transmisión sináptica. Nos llamó la atención 5 genes de los mismos fueron encontrados, en análisis anteriores, como asociados a NOH: SYT4, SNCA, LRRK2, SIPA1L2. Habíamos encontrado 5 variantes genéticas localizadas a nivel de estos 5 genes.

11.10- Down-regulación del potencial de membrana

La down regulation del potencial de membrana es otro proceso celular encontrado asociado a la NOH en este análisis. En este proceso participan un total de 381 genes, de los cuales presentamos los siguientes que son de interés en este estudio: SCN2A, SH3GL2, SNCA, GBA. Como podemos apreciar, habíamos encontrado variantes genéticas asociados a NOH, las mismas se localizan dentro o cerca de los genes SCN2A, SH3GL2, SNCA, GBA.

12.5- Down-regulación de señalización trans sináptica

El proceso celular de señalización trans sináptica fue encontrada down regulada en este estudio, en el cual intervienen también varios genes encontrados asociados a NOH en el estudio de asociación genética: SYT4, DGKB, SH3GL2, SNCA, LRRK2, SIPA1L2.

11.12- Down-regulación del transporte mediado por vesículas en sinapsis

El proceso de transporte celular mediado por vesículas sinápticas también fue encontrado asociado a NOH en este análisis. Este proceso fue encontrado down regulada, en el cual participan varias proteínas como por ejemplo SNCA, SH3GL2, PPFIA2, **SH3GL2**, SNCA, DGKQ, NRXN1, ACTG1, LRRK2. Hace falta recalcar además SH3GL2, SNCA, LRRK2 encontrado anteriormente estar asociado con NOH, encontramos también el gen DGKQ estar involucrado en este proceso. Habíamos encontrado que el gen DGKQ está asociado con NOH a través de la relación de desequilibrio de ligamiento, que algunos SNPs localizados en dicho gen establecen con el SNPs rs34311866-TMEM175(Encontrado asociado a NOH)(Ver tabla 6-10).

SNP name	Gene name	Log2FoldChange	P-values	P-val_Adjust
rs104893877 rs3910105	SNCA	-2.15	9.5E-5	4E-3
rs353116	SCN3A.SCN2A	-1.54	1.9E-3	2.8E-2
rs13294100	SH3GL2	-1.03		0.05
rs2694528	ELOVL7.NDUFAF2	-1.28	1.6E-3	2.2E-2
rs11158026	GCH1	-0.93		0.09
rs55785911	DDRKG1	1.31	8.3E-4	1.3E-2
rs12456492	SYT4.RIT2	-1.81	2.5E-5	1.3E-3

Tabla 27 - Confirmación de los SNPs asociados a NOH por expresión diferencial de genes. En esta tabla se presentan los resultados de algunas proteínas de interés que se encuentran diferencialmente expresados en el análisis diferencial de genes. Los genes como SYT14, GCH1, NDUFAF2, SH3GL2, SCN2A, SNCA están subexpresados en los pacientes con NOH. El gen DDRKG1 está sobreexpresado en el grupo NOH. Los SNPs encontrados asociados a NOH se encuentran localizados en estos genes.

Pudimos confirmar que 6 genes están asociados a la NOH en pacientes con enfermedad de Parkinson. Porque en análisis de regresión logística entre variantes y NOH, habíamos encontrado variantes localizadas en estos genes asociados a NOH. Además encontramos que estos genes se encuentran diferencialmente expresados en los pacientes con NOH con respecto a los sin NOH(Ver [tabla 6-10](#) y [tabla 27](#)). El gen GCH1 también tiene tendencia a estar diferencialmente expresados en ambos grupos(p valor ajustado: 0.09). Los genes como SYT4, GCH1, SCN2A, SNCA, NDUFAF2, SH3GL2 están con expresión baja en los pacientes con NOH según nuestro estudio. Por otro lado, el DDRKG1 está sobreexpresado en los pacientes con NOH según los resultados de nuestro estudio

Capítulo XIII- Análisis de predicción de NOH en pacientes con enfermedad de parkinson por algoritmos de aprendizaje automático.

13.1- Algoritmos utilizados y procedimientos

Hemos realizado análisis de predicción de la NOH en pacientes con EP usando el componente AutoML de la plataforma KNIME. AutoML es un módulo de aprendizaje automático de la plataforma de KNIME que se puede utilizar de forma libre y abierta por otros usuarios. Permite correr varios algoritmos de Machine Learning sobre los datos automáticamente y obtener resultados fiables. Los resultados obtenidos de los diferentes algoritmos elegidos por el usuario podrán ser comparados los unos con los otros determinando el algoritmo con el mejor performance para llevar a cabo la tarea en cuestión en este caso la predicción sobre los datos.

Los algoritmos de clasificación que se pueden utilizar en AutoML son la regresión logística, redes neuronales, XGBoost Trees, Random Forest, Gradient Boost Trees, modelo lineal generalizado, Naives Bayes. Para cada uno de los algoritmos, los datos fueron divididos en 2 juegos: un primer juego de datos para el entrenamiento el cual a su vez se dividió en 4 subjuegos de datos para la cross-validación, un segundo juego para la prueba. Hemos utilizado datos de 22 SNPs encontrados asociados con la NOH (Neurogenic Orthostatic Hypotension) en análisis anterior, edad, sexo, toma de medicamentos antiparkinsonianos, toma de medicamentos asociados a Hipotensión ortostática, scores de MDS-UPDRS I, II, III, IV. En primer lugar utilizamos todos los datos con todas las variables incluidas después dividimos las variables en clínicas que incluyen a todas las variables menos los 22 SNPs. También hicimos análisis de predicción del NOH usando separadamente los datos de MDS-UPDRS, SCOPA-AUT, y meramente clínicas sin los SNPs ni los Scores UPDRS ni SCOPA-AUT. Al final, utilizamos solamente las variables de los 22 SNPs para hacer un Score poligénico para la NOH. Este abordaje nos permite sopesar las funciones y el peso de cada grupo de variables en la predicción final del NOH por las variantes genéticas.

Para el entrenamiento, el 80% de los datos fueron utilizados y 20% para la prueba. Los datos de entrenamiento a su vez fueron divididos en subgrupos para la cross-validación donde se usaron un número de fold de 4. AutoML se trata de un módulo que actúa automáticamente hasta seleccionar y rankear los mejores modelos que se ajustan a los datos y que predicen bien la etiqueta elegida usando unas métricas preestablecidas. La métrica usada para rankear y seleccionar los mejores algoritmos fue el accuracy. Otras métricas de evaluación fueron utilizadas tal como curva de ROC(AUC), Error Rate, Balance

Accuracy, false negative rate, precisión, sensibilidad, especificidad, F, Youden's J Index y Negative predictive value(NPV).

Hemos utilizado estrategias de optimización e hiperparametros para poder encontrar de forma automática el mejor modelo que predicen la NOH en los datos.

En el caso del algoritmo de Gradient Boost Trees, el parámetro usado para la búsqueda es el número de Trees(Núm Trees) empezando con un valor inicial de 50 hasta un valor final de 100 en el cual se termine la búsqueda automáticamente, un Step Size de 20 y un número máximo de reiteración de 10. Se utilizó la optimización bayesiana como estrategia de búsqueda.

Para el algoritmo Naives Bayes, el def-prob fue usado como parámetro de búsqueda empezando con un valor de 0.004 hasta 0.01 con step size de 1.00E-4 y un número máximo de reiteración de 10. Se utilizó la optimización bayesiana como estrategia de búsqueda. En el caso de la regresión logística, el parámetro usado fue el Step-size que se inició con un valor de 0.01 hasta un valor final de 0.1, un step-size de 0.01 y un número máximo de reiteración de 10. También se utilizó la optimización bayesiana como estrategia de búsqueda. En el caso particular de Random Forest tres parámetros de búsqueda fueron usados: Tree-depth, node size, no-of-trees. El primero empezó con un valor de 4 hasta un valor de 8 y un step size de 4; el segundo empezó con un valor de 10 hasta 20 con un step size de 5 y el tercero con un valor de 100 hasta 200 y un step size de 100. Los tres métodos usaron la optimización bayesiana como estrategia de búsqueda y se configuró a un máximo de 10 repeticiones. Los parámetros usados en las redes neuronales fueron dos: Num-layers, Núm-neurons. El primero se inició con un valor de uno(1) hasta dos(2) con Step-size de uno y el segundo con un valor de 5 hasta 20, un step size de 10 y una reiteración máxima de 10 en ambos casos. Los dos usaron la optimización bayesiana como estrategia de búsqueda. También se usaron dos parámetros de búsqueda en el caso del algoritmo modelo lineal generalizado: Alpha y lambda. Ambos parámetros se inician con un valor de cero hasta un valor final de 1, un step size de 0.2 y un valor de reiteración máxima de 10. Ambos usan la optimización bayesiana como estrategia de búsqueda. El algoritmo XGBoost Trees usó dos parámetros el max-depth y el ETA. El primero empezó con un valor de 5 hasta 10, un step size de 5 y un valor de reiteración de 10 mientras que el ETA inició con un valor de 0.2 hasta 0.3, un step size de 0.1. En ambos casos se usa la optimización bayesiana como estrategia de búsqueda.

13.2- Resultados de las predicción basadas en algoritmos de aprendizaje automático

13.2.1- Predicción sobre los datos tanto genéticos, clínicos demográficos

En el caso donde se usaron todas las variables descritas anteriormente tanto clínicas, genéticas y demográficas, hemos encontrado 3 algoritmos con mejores performances, los cuales son Gradient Boosted Trees, Logistic regression y Neural Network(Ver tabla 27 y figura 28, 29). Se ha encontrado Gradient Boost Trees con un AUC de 0.84, accuracy en 0.98, especificidad y precisión en 1 cada uno, sensibilidad en 0.83 y NPV(Negative predictive value) en 0.975. Mientras que en la regresión logística el AUC fue encontrado en 0.85, accuracy en 0.96, precisión y sensibilidad en 0.83 cada uno, y especificidad y NPV en 0.974 cada uno. El otro algoritmo con mejor performance es neural network con un AUC de 0.91, accuracy de 0.95, especificidad y precisión de 1 en cada uno, sensibilidad de 0.83 y NPV de 0.975. Otro parámetro es el error rate encontrado en 0.02 en el caso de algoritmo Gradient Boosted Trees y Neural Network y 0.04 en el caso de Regresión logística. En los 3 algoritmos se ha encontrado un false negative rate de 0.16(Ver figura 30, 31, tabla 28).

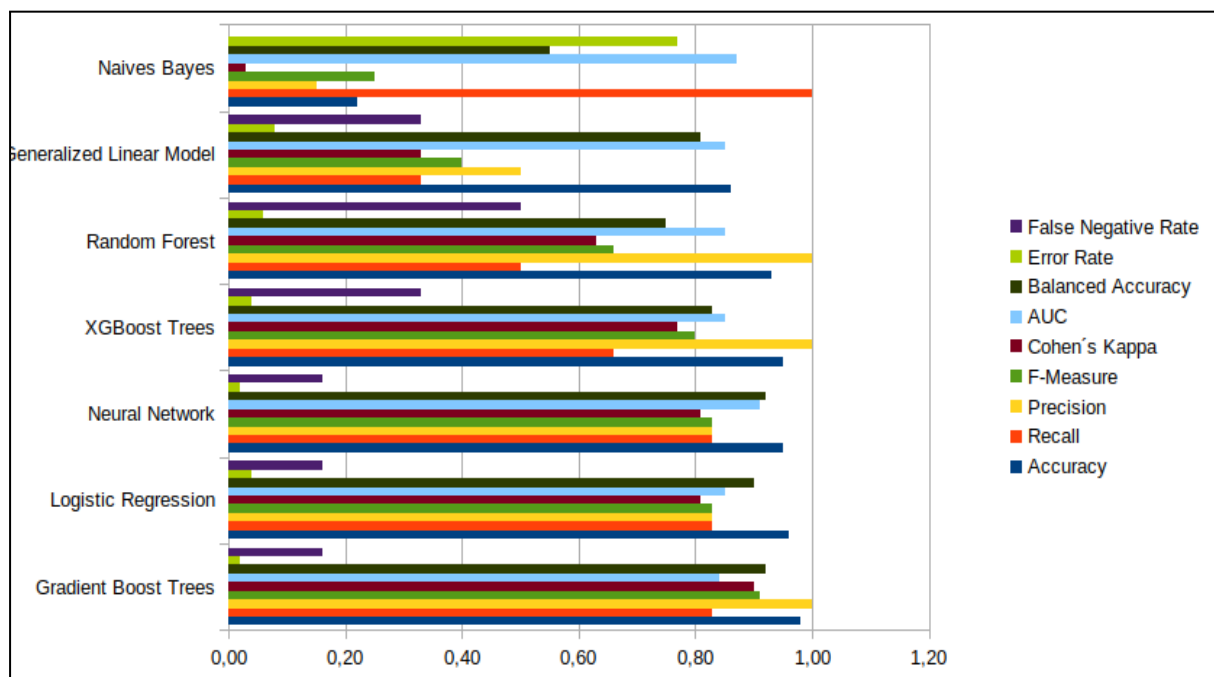


Figura 30: Representación de los algoritmos de aprendizaje automático en el conjunto de los datos(Clínicos, demográficos, genéticos).

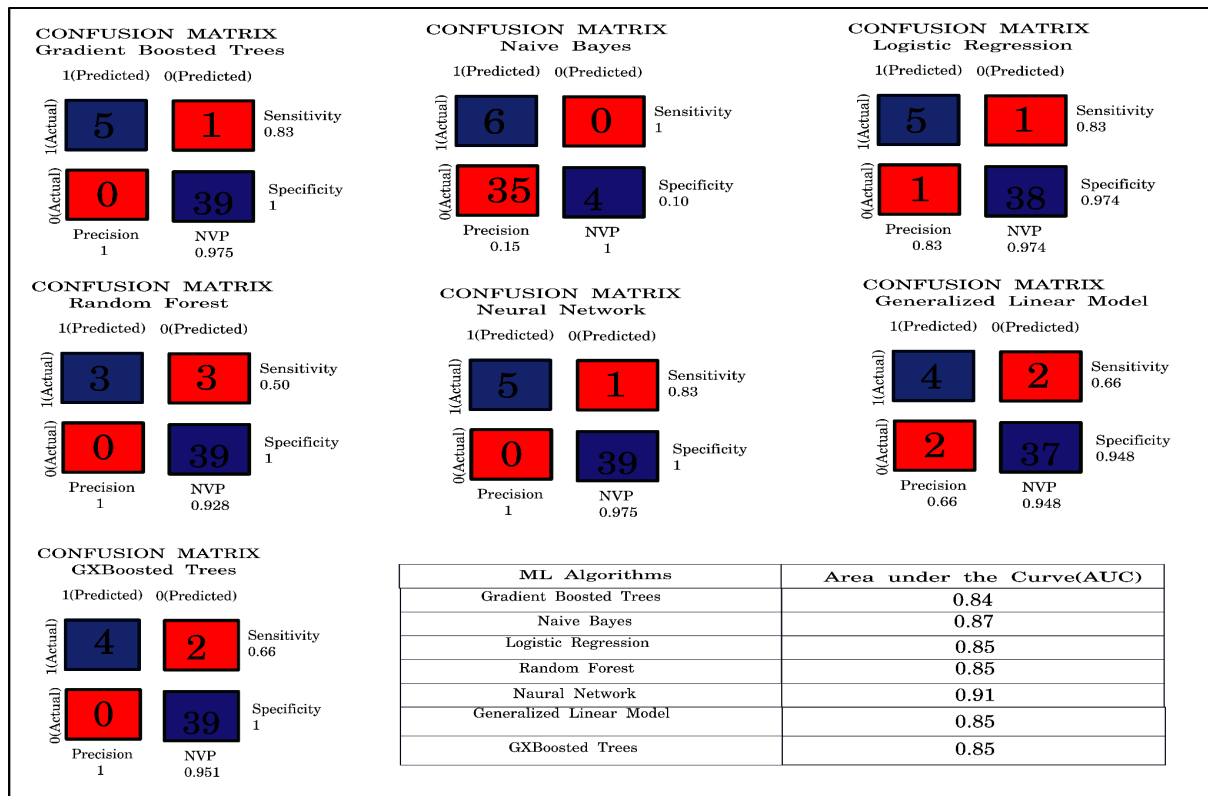


Figura 31- Parámetros de evaluación de los algoritmos de aprendizaje automático sobre el conjunto de los datos(clínicos, demográficos, genéticos).

	Accuracy	Recall	Precision	F-Measure	Cohen 's K.	AUC	Balanced Accuracy	Error Rate	False Negative Rate
Gradient Boosted Trees	0,98	0,83	1,00	0,91	0,90	0,84	0,92	0,02	0,16
Logistic Regression	0,96	0,83	0,83	0,83	0,81	0,85	0,9	0,04	0,16
Neural Network	0,95	0,83	0,83	0,83	0,81	0,91	0,92	0,02	0,16
XG Boost Trees	0,95	0,66	1,00	0,80	0,77	0,85	0,83	0,04	0,33
Random Forest	0,93	0,50	1,00	0,66	0,63	0,85	0,75	0,06	0,50
Generalized Linear Model	0,86	0,33	0,50	0,40	0,33	0,85	0,81	0,08	0,33
Naives Bayes	0,22	1,00	0,15	0,25	0,03	0,87	0,55	0,77	0,00

Tabla- 28- Los parámetros de evaluación de los algoritmos de aprendizaje automático en el conjunto de los datos (demográficos, clínico, genético).

13.2.2- Predicción sobre los datos genéticos solamente

Hicimos otros modelos de predicción usando solamente los datos genéticos sin los datos clínicos y demográficos para estudios de los scores poligénicos. Los mismos algoritmos y los mismos hiperparametros usados en los modelos anteriores fueron considerados para estos modelos. En el caso donde se usaron solamente las variables genéticas hemos encontrado 3 algoritmos con mejores performances, los cuales son Gradient Boosted Trees, Logistic regression y Neural Network(Ver tabla 29 y figura 32). Se ha encontrado Gradient Boost Trees con un AUC de 0.87, accuracy en 0.95, especificidad y precisión en 1 cada uno, sensibilidad en 0.83 y NPV(Negative predictive value) en 0.975. En el Neural Network, el AUC fue encontrado en 0.91, accuracy en 0.95, precisión y especificidad en 1 cada uno, y sensibilidad en 0.83 y NPV en 0.974 cada uno. El otro algoritmo con mejor performance es neural network con un AUC de 0.91, accuracy de 0.95, especificidad y precisión de 1 en cada uno, sensibilidad de 0.83 y NPV de 0.975. El algoritmo XG Boost Trees tiene un accuracy de 0.93 y AUC 0.85 precisión y especificidad en 1 cada uno, y sensibilidad en 0.83 y NPV en 0.974 cada uno Otro parámetro es el error rate encontrado en 0.02 en el caso de algoritmo Gradient Boosted Trees y Neural Network y 0.04 en el caso de GXBoost Trees.

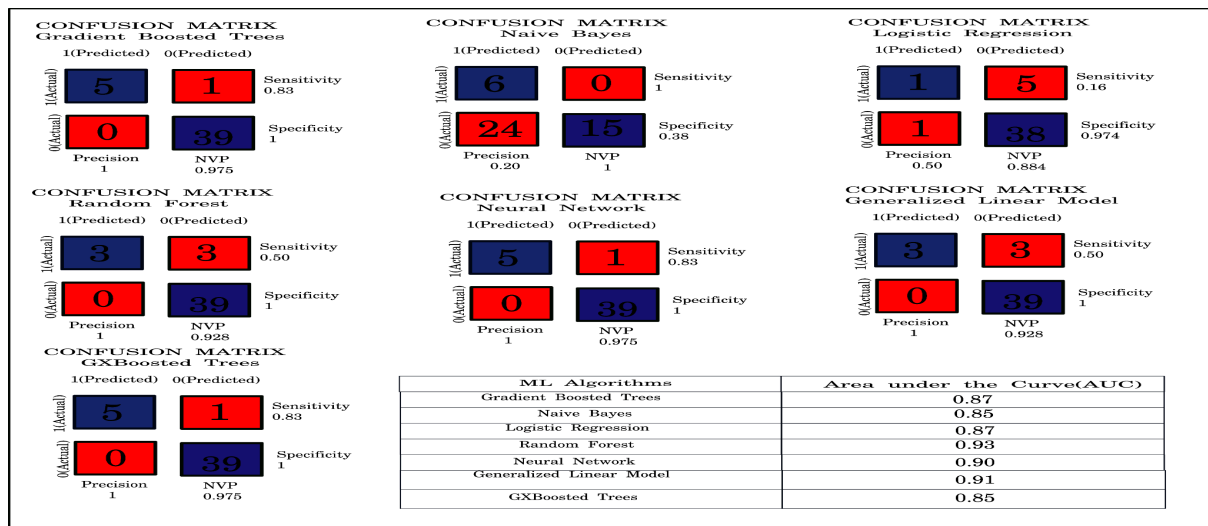


Figura 32- Parámetros de evaluación de los algoritmos de aprendizaje automático en los datos genéticos(Excluyendo datos demográficos y clínicos).

	Accuracy	Precision	F-Measure	AUC	Balanced Accuracy	Error Rate	False Negative Rate
Gradient Boosted Trees	0.95	1.00	0.90	0.84	0.92	0.02	0.16
Logistic Regression	0.93	0.83	0.83	0.85	0.90	0.04	0.16
Neural Network	0.95	1.00	0.91	0.91	0.92	0.02	0.16
XG Boost Trees	0.93	1.00	0.80	0.85	0.83	0.04	0.33
Random Forest	0.91	1.00	0.67	0.85	0.75	0.06	0.50
Generalized Linear Model	0.88	0.67	0.67	0.85	0.81	0.08	0.33
Naives Bayes	0.22	0.15	0.25	0.87	0.55	0.78	—

Tabla 29- Evaluación de los algoritmos de aprendizaje automáticos para los datos genéticos(Excluyendo datos clínicos y demográficos)

Modelo usando los datos Clínicos y demográficos:

Utilizando solo los datos demográficos y clínicos sin los datos genéticos, podemos apreciar que el performance de los datos para la predicción del NOH se ve grandemente afectados(Ver figura 33 y tabla 30). También aprovechamos acá para recalcar que, en el caso del usos de todos los datos tanto genéticos como clínicos demográficos para la predicción del NOH no se encuentran diferencias significativas cuando usamos solo los datos genéticos para la predicción. Con esto podemos concluir que con solo los datos genéticos pudimos obtener buenos performance en los algoritmos de predicción usados.

	Accuracy	Precision	F-Measure	AUC	Balanced Accuracy	Error Rate	False Negative Rate
Gradient Boosted Trees	0.93	1.00	0.80	0.83	0.83	0.04	0.33
Logistic Regression	0.84	NAN	—	0.65	0.50	0.13	1.00
Neural Network	0.91	1:00	0.67	0.82	0.75	0.06	0.50
XG Boost Trees	0.91	1:00	0.67	0.82	0.75	0.06	0.50
Random Forest	0.88	1:00	0.50	0.85	0.67	0.08	0.67
Generalized Linear Model	0.84	NaN	—	0.64	0.50	0.13	1:00
Naives Bayes	0.26	0.15	0.27	0.71	0.57	0.73	—

Tabla 30- Evaluación de algoritmos de aprendizaje automático para predicción de NOH en datos clínicos y demográficos(Excluyendo a datos genéticos).

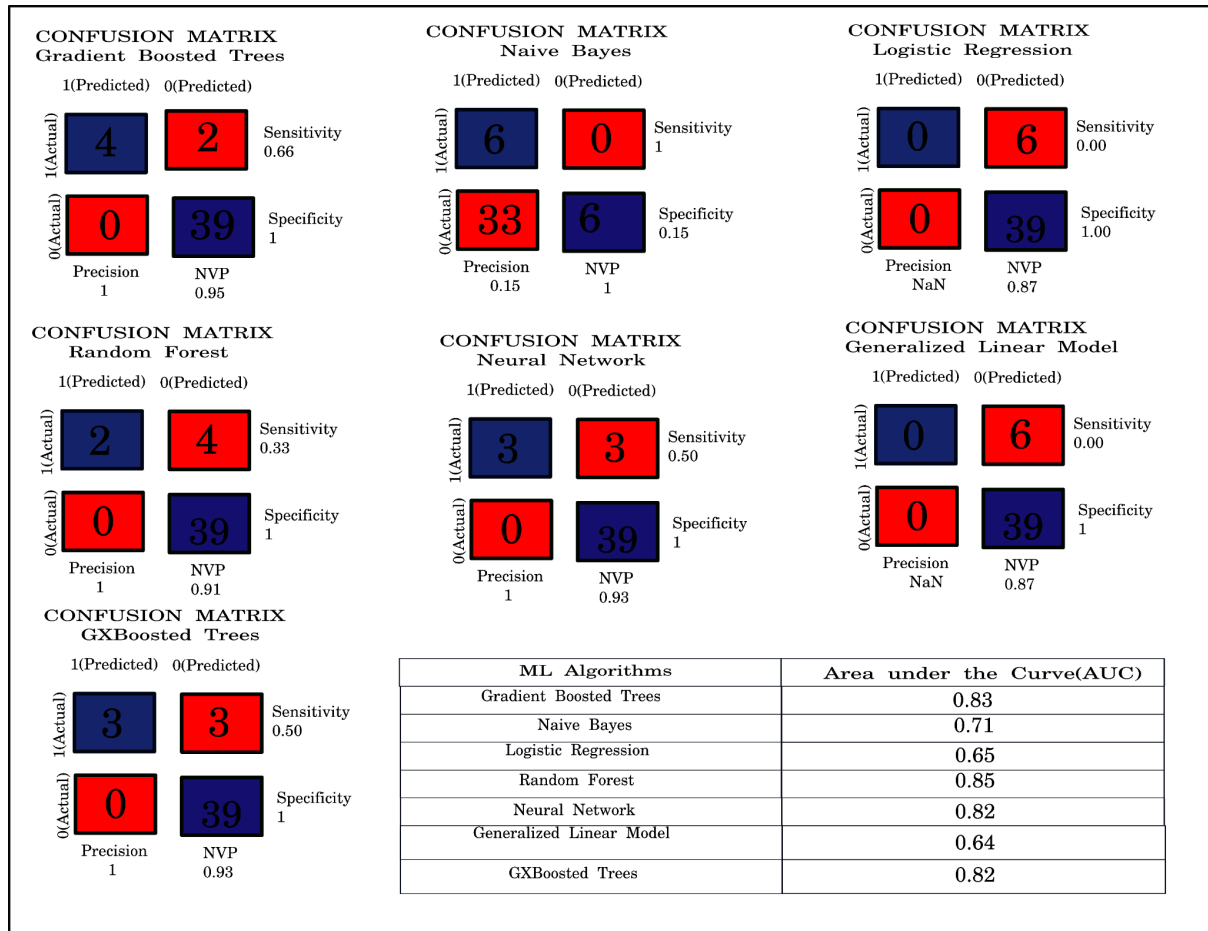


Figura 33- Matriz de confusión de los algoritmos de aprendizaje automático en los datos clínicos y demográficos(Excluyendo a datos genéticos).

Capítulo XIV- Discusiones.

14.1- SNP y vías de señalización asociados con la NOH en pacientes con EP

Como parte de la disfunción autonómica, la fisiopatología de la NOH se explica principalmente por la denervación simpática posganglionar (Rafanelli et al. 2019). Aunque los factores específicos relacionados con la EP que conducen a esta denervación simpática no se han estudiado satisfactoriamente. La base genética de la NOH en la EP no se ha estudiado de forma exhaustiva. En el presente estudio, se encontró que 21 de las 72 variantes examinadas están asociadas con NOH. Nuestro análisis estadístico se diseñó específicamente para descartar los efectos de confusión de variables como la edad, el sexo y los medicamentos para la EP. Los SNPs localizados en los genes SNCA-A53T, SCN2A, SH3GL2, SNCA-rs3910105, MIR4697 aumentan el riesgo de NOH en los modelos de riesgo aditivo y dominante. Esto quiere decir que con una sola copia del alelo de riesgo, el paciente presenta el riesgo de NOH y con dos copias el paciente presenta el doble de riesgo. Los SNPs localizados en los genes ITGA8, NDUFAF2 disminuyen el riesgo de NOH en los modelos de riesgo aditivo y dominante. Esto significa que con solo una copia del alelo de riesgo, el paciente será protegido de la NOH y con dos copias será doblemente protegido. Los SNPs localizados en los genes CCDC62, FAM47E/STBD1, DDRGK1 aumentan el riesgo de NOH en los modelos aditivos y recesivos. Esto quiere decir que para que el paciente presente el riesgo de NOH se necesita que sea homocigota para el alelo menor (alelo de riesgo). Los SNPs localizados en los genes SIPA1L2, TMEM163, LRRK2, GCH1 disminuyen el riesgo de NOH y SYT4 lo aumenta en un modelo aditivo, lo que quiere decir que el riesgo es dosis dependiente. La presencia de un alelo corresponde a una dosis de riesgo y dos alelos dos dosis de riesgo. Los SNPs localizados en los genes BIN3, EGF disminuyen el riesgo de NOH y COMT aumenta el riesgo de NOH en un modelo recesivo. Esto quiere decir que se necesitan dos copias para que se produzca el riesgo. Los SNPs localizados en los genes GBA, y MAP4K4 aumentan el riesgo de NOH en un modelo dominante solo quiere decir que con solo la presencia de una copia se produce el riesgo de NOH.

En un análisis de epistasis se encontró que los SNPs localizados en los genes GPNMP y GALC/GPR65 se interactúan para disminuir el riesgo de hipotensión ortostática neurogénica en enfermos de parkinson. En el análisis de interacción de proteína-proteína usando la base de datos STRING no se encontraron interacción a nivel de proteína entre los genes. Por otro lado, los SNPs localizados en los genes SNCA y NUCKS1 se interactúan para

aumentar el riesgo de NOH en enfermos de parkinson. Se encontró que a nivel de proteína que la proteína NUCKS1 y SNCA interactúan en la base de datos STRING a través del método de text mining.

En una análisis de enriquecimiento de los SNPs asociados a NOH usando la plataforma SNPnexus se identifica que los genes como IP6K2 enriquece para la síntesis y metabolismos de fosfoinositol, y señalización interferón alfa y beta. EL gen SH3GL2 enriquece para transporte retrógrada de neurotrofina, regulación a la baja del EGFR, red de vesículas trans-golgi, biogénesis de vesículas de lisosoma, biogénesis de vesículas asociadas a Golgi, reconocimiento de cargo para endocitosis mediada por clatrina. El gen SCN2A enriquece para la despolarización rápida e interacción entre Ankryns y L1.

14.2- SNCA

Estudios previos han informado que algunas variantes de SNP ubicadas en el gen SNCA (Chelban et al. 2018; Mendoza-Velásquez et al. 2019), aumentan el riesgo de deterioro autonómico y síntomas de NOH. De igual manera, en nuestro estudio, las variantes rs3910105 y rs104893877 localizadas en el gen SNCA aumentaron el riesgo de NOH (OR=3,05 y 55,29, respectivamente). Curiosamente, se han encontrado otras variantes SNPs-SNCA asociadas con falla autonómica en otras alfa-sinucleinopatías como la Atrofia Multisistémica en la que la NOH es un síntoma principal(Mendoza-Velásquez et al. 2019). Otro estudio encontró que la duplicación y triplicación del gen SNCA-MMRN1 en la población sueca conduce a una disfunción autonómica y parkinsonismo de inicio tardío(Fuchs et al. 2007).

Mapeamos la función de las vías metabólicas/de señalización, porque los datos recientes y nuestro resultado en el modelo de regresión logística sugieren que la EP probablemente abarca una variedad de entidades genético-moleculares muy importante(Lang and Espay 2018). Además, en este mismo estudio, pudimos confirmar a través de una análisis diferencial de expresión de genes que el gen SNCA está diferencialmente expresado en los pacientes con NOH con respecto al grupo control. La SNCA esta downregulada en los pacientes con NOH según los resultados de nuestro estudio. En otros estudios se ha demostrado que la disminución en la expresión de la SNCA está asociada con alteración de la exocitosis y la transmisión sináptica.

Los pacientes heterocigotas para el alelo menor del SNP SNCA_A53T_rs104893877 presentan mayor niveles de UPDRS III(34.00+-17.78, p-valor: 0.049 para los pacientes Heterocigotas) en el grupo NOH con respecto a los pacientes del grupo no-NOH (23.875+-17.9, p-valor: 0.89). El score MDS-UPDRS III estudia las actividades y

capacidades motoras de los pacientes con EP. Se puede utilizar para estratificar la severidad de la enfermedad de Parkinson (Lu et al. 2021). Dicho de otra manera, esta variante genética localizada en el gen SNCA está asociado con una mayor severidad y peor pronóstico de la enfermedad de Parkinson con respecto a los que no tienen esta variante.

14.3- IP6K2

Hemos encontrado que el SNP rs12497850 ubicado en el gen IP6K2 está asociado con NOH. Este gen codifica una proteína que forma parte de la familia de la proteína inositol fosfoquinasa (IPK). Se encuentra que este gen está significativamente enriquecido para las vías del metabolismo del fosfato de inositol, la síntesis de IP en el núcleo, la señalización de interferón alfa/beta. Múltiples estudios han demostrado que las altas concentraciones de difosfoinositol pentakisfosfato producido por la enzima IP6K2 y/o la sobreexpresión de IP6K2 están todas asociadas con la citotoxicidad y apoptosis de múltiples líneas celulares, mientras que la eliminación del gen IP6K2 disminuye la muerte celular (Nagata et al. 2005). (Nagpal et al. 2021) han demostrado que el cerebelo de ratones IP6K2^{-/-} Knockout, produce menos fosfocreatina y ATP por disminución de la expresión de CK-B, y genera altos niveles de especies reactivas de oxígeno y daño oxidativo a las proteínas, además a la disfunción mitocondrial asociada con la expresión alterada del complejo III de la subunidad c1 del citocromo de la cadena de transporte de electrones (Nagpal et al. 2021). Por otro lado, se ha demostrado que se requiere IP6K2 para mediar la detención del ciclo celular y la apoptosis dependientes de p53 en respuesta al estrés celular (Koldobskiy et al. 2010). Estos datos indican claramente que IP6K2 juega un papel importante en la apoptosis celular, y sus niveles de expresión o el deterioro de la concentración derivado de sus productos, producen trastornos en la función celular tales como disfunción mitocondrial, estrés oxidativo y muerte celular.

14.4- SH3GL2

El SNP rs13294100 ubicado en el gen SH3GL2 también se encuentra asociado con NOH. Además, el gen SH3GL2 está significativamente enriquecido para vías de señalización como el transporte retrógrada de la neurotrofina (Zweifel, Kuruvilla, and Ginty 2005), regulación negativa de EGFR, gemación de vesículas de red trans-Golgi, biogénesis de vesículas de lisosomas, biogénesis de vesículas asociadas a Golgi, reconocimiento de cargo para endocitosis mediada por clatrina. El transporte retrógrado de señalización de

neurotrofinas se caracteriza por el transporte retrógrado de factores de neurotrofinas desde los axones hasta el cuerpo celular para promover la supervivencia celular y prevenir la neurodegeneración (Ito and Enomoto 2016). Las alteraciones en este proceso conducen a la neurodegeneración (Ito and Enomoto 2016). Uno de los mecanismos descritos para el proceso de transporte retrógrado de neurotrofinas es a través del transporte vesicular de complejos neurotrofina-Trk que transmite una señal de supervivencia que involucra a PI3K y Erk5 (Ginty and Segal 2002). El gen SH3GL2 está enriquecido no solo para el transporte vesicular de la neurotrofina junto con sus receptores Trks, sino también para la biogénesis vesicular tanto en el aparato de Golgi como en el lisosoma. Estos datos están indicando que este SNP asociado a NOH ubicado en el gen SH3GL2 está afectando procesos celulares como el transporte vesicular de neurotrofina y la neurotransmisión. Además, este gen está enriquecido para la vía de señalización relacionada con la regulación a la baja de EGFR. Es bien sabido que la regulación a la baja de EGFR está asociada con la neurodegeneración en la corteza frontal de ratones knockout para EGFR-/- por aumento de la apoptosis mediada por un mecanismo dependiente de Akt-caspasa (Wagner et al. 2006). Con base en estos datos, podemos postular que los SNP ubicados en el gen SH3GL2 asociado con NHO alteran la transmisión sináptica de neurotransmisores y el transporte retrógrado de neurotrofinas a través del deterioro de la biogénesis vesicular tanto en los lisosomas como en el aparato de Golgi.

En nuestros datos, hemos identificado una disminución de la expresión del gen SH3GL2 en los pacientes con NOH con respecto al control (los pacientes que no tienen NOH) utilizando un modelo lineal generalizado con una función de enlace binomial negativa de la aplicación DESeq2. Este hallazgo permite formular la hipótesis de que los pacientes con NOH presentan disminución en la expresión del gen SH3GL2 y por consiguiente disminución en las funciones como transporte retrógrado de neurotrofinas, reconocimiento de cargo para la endocitosis mediada por clatrina, biogénesis de vesículas lisosomales y gemación de vesículas del complejo de Golgi. Todos estos procesos se encuentran disminuidos en los pacientes con NOH. John D Murdoch et al demostraron que una deficiencia del SH3GL2 (endofilina A1) produce alteración en los procesos normales de autofagia y en la degradación de proteína por el sistema ubiquitino-proteasomal y apoptosis de células neuronales (Murdoch et al. 2016). Pareciera que la deficiencia de SH3GL2 aumenta la expresión de ubiquitina ligasa E3 FBXO32 y su factor de transcripción FOXO3A, los cuales inducen apoptosis en células neuronales en cultivo (Murdoch et al. 2016). Por otro lado, la regulación a la baja de la expresión de la proteína endofilina A1 (codificado por el gen SH3GL2) es uno de los eventos encontrados asociados con la disminución funcional de la

maquinaria endocítica dependiendo o no de la clatrina (Buser et al. 2019). Estos datos van en el mismo sentido que los encontrados en nuestro estudio.

Se encontró que varios SNPs ubicados en el gen PRKAR2A están en desequilibrio de ligamiento con el SNP rs13194100-SH3GL2 asociado a NOH. El gen PRKAR2A está enriquecido para vías de señalización como los receptores ROBO que se unen a la vía AKAP5, la fosforilación de CREB mediada por PKA, la fosforilación de CREB1. La familia Roundabout (ROBO) codifica receptores transmembrana que regulan la orientación axonal y la migración celular. Las vías de señalización de la proteína quinasa A (PKA) y la proteína de unión al elemento de respuesta cAMP (CREB) mediante la neuroplasticidad y las respuestas de supervivencia neuronal por su capacidad para regular la expresión génica, como las neurotrofinas y otros genes de supervivencia (Clark, Monahan, and Drew 2018; Zanassi et al. 2001). Varios estudios han demostrado que un deterioro en la actividad o la expresión de las vías CREB está asociado con la neurodegeneración (Mantamadiotis et al. 2002), (Dragunow 2004). Esta alteración de la expresión del CREB podría ser un evento secundario a la alteración de la expresión del gen PRKAR2A por el SNP rs13194100 de SH3GL2.

En el análisis de eQTL (expresión cuantitativa de rasgos de locus) a partir de conteos de genes de células dopaminérgicas derivadas de iPSC, fueron identificados 273 genes cuyas expresiones dependen de los genotipos de los pacientes (eQTLs). Se computaron 4 módulos a partir del network contruidos con los eQTLs. Los genes que forman el primer y el segundo módulo intervienen en splicing de ARN, procesamiento y metabolismo de ARN, biogénesis y transporte del complejo ribonucleoproteico. Nuestros datos apuntan a que el gen SH3GL2 interviene en los procesos de metabolismos, procesamiento, splicing y transporte de ARNm y que la disminución en la expresión de SH3GL2 estaría induciendo alteración en estos procesos moleculares y celulares en pacientes con esta variante rs13194100 y NOH. En este estudio fue encontrado que los genes como HNRNPH1, DGCR8 del módulo dos, están asociados con el gen SH3GL2 en un modelo de regresión. Además, los pacientes homocigotas para el alelo de riesgo y heterocigota para el SNP rs13194100-SH3GL2 tienen expresión de estos dos genes y LAS1L aumentados. Mientras que en la regresión, la expresión del LAS1L está disminuida bajo el efecto del gen SH3GL2. LAS1L intervienen en la biogénesis de la subunidad ribosomal 60S y en la maduración del rRNA 28s (Castle et al. 2010) (Castle, Cassimere, and Denicourt 2012). HNRNPH1 son proteínas de unión a RNA y participan en el procesamiento, metabolismo y transporte de ARNm (Matunis, Xing, and Dreyfuss 1994). DGCR8 es una proteína que interviene en la biogénesis y procesamiento de microARN y en la reparación del ADN dañado por estrés oxidativo (Calses et al. 2017; Yeom et al. 2006). Por otro lado, en el módulo tres se

encuentran eQTLs que participan en el transporte membranar de protones contra gradiente de concentración, transporte de electrones mitocondriales, síntesis de ATP acoplado a transporte de electrones, ensamblaje del complejo NADH deshidrogenasa, ensamblaje del complejo I de la cadena respiratoria mitocondrial, fosforilación oxidativa y respiración celular. El SNPs rs13194100 altera la expresión de los genes que participan en estos procesos. En un análisis de asociación entre el gen SH3GL2 y otros genes encontrados en este módulo 3, encontramos que se disminuye la expresión del gen mitocondrial GOT2 en pacientes heterocigotas y homocigotas para el alelo de riesgo del SNP rs13194100-SH3GL2, se aumenta la expresión del MTATP6P1 en pacientes homocigotas para el alelo de riesgo, y se aumenta la expresión del gen MT-ND6 en pacientes homocigotas y heterocigotas para el alelo de riesgo del SNP rs13194100-SH3GL2. La proteína GOT2 mitocondrial participa en el malato-aspartato shuttle para proveer NADH a la cadena transportadora de electrones y la producción de ATP(H. Yang et al. 2015). MT-ATP6P1 es un pseudogen del ATP sintasa 6 mitocondrial. MT-ND6 es el gen que codifica la subunidad 6 de la NADH deshidrogenasa que participa en los procesos de la cadena de transporte de electrones. Una disminución en la expresión del gen SH3GL2 en pacientes con el alelo de riesgo del SNP rs13194100(tanto homocigotas para el alelo de riesgo como heterocigotas) estaría produciendo una disminución de la expresión del gen GOT2 y un aumento en la expresión de los otros dos genes que intervienen en la cadena de transporte de electrones y producción de energía. GOT2 es un gen nuclear mientras que MT-ND6 se localiza en el ADN mitocondrial. Pareciera que deficiencia en la expresión SH3GL2 estaría provocando disminución en la expresión del GOT2 en los pacientes con alelo de riesgo lo que afectaría la disponibilidad mitocondrial de NADH y producción de ATP a través del malato-aspartato shuttle y la disminución de la concentración mitocondrial de NADH induciría el aumento en la expresión de genes mitocondriales que intervienen en la producción de NADH y de ATP.

Es importante mencionar otros genes de los 25 que fueron elegidos como genes hub por su alto grado de conexión, centralidad de intermediación y de cercanía: ATP13A2, NCDN, CISD2, CXADR, SYNJ1, LRRC24, KIAA0895, UBQLN1, GNE, LMTK2, RASA4CP, CDK20, CYB561, FASN, SYT13. Todos estos genes están subexpresados(expresión disminuida) en los pacientes con el alelo de riesgo para el SNP rs13194100. El gen ATP13A2 es una enzima que interviene en la eliminación de la célula de ion metálico de transición. Protege la célula contra los efectos neurotóxicos del hierro, del manganesio y zinc(Tan et al. 2011; Rinaldi et al. 2015)(H. Gao et al. 2022). Marija et al demostraron que una deficiencia de la ATP13A2 produce disfunción en la maquinaria lisosomal y autofágica conduciendo a acumulacion de alfasinucleina y muerte neuronal en fibroblastos de pacientes con enfermedad de parkinson dependiente de ATP13A2(H. Gao et al. 2022; Usenovic et al.

2012). La NCDN es el gen que codifica la proteína neurocondrina la cual participa en el crecimiento de neuritas, desarrollo de proyecciones neuronales y la regulación de la plasticidad sináptica neuronal. La proteína codificada por el gen CSD2 es una proteína de dedo de zinc localizada dentro del retículo endoplásmico y participa en la homeostasis del calcio. SYNJ1 codifica la proteína synaptojanina-1 que es una fosfatasa de fosfoinositoles para regular los compuestos fosfatidilinositol-4,5-bisfosfato a nivel de las membranas celulares participando de esta forma en la transmisión sináptica y tráfico de membrana y endosomal(Choudhry, Aggarwal, and Pan 2021). Fasano et al encontró que la enfermedad de parkinson de aparición temprana, dependiendo de la disfunción del tráfico endolysosomal está asociada con mutación en el gen SYNJ-1(Fasano et al. 2018). UBQLN1 es una proteína llamada Ubiquilina con un dominio semejante a la ubiquitina en el extremo N-terminal y un dominio asociado a ubiquitina en el extremo C-terminal. Estos dominios se unen físicamente con el proteosoma y con la ubiquitina por lo que permite enlazar el sistema de ubiquitilación y el sistema proteasomal en la degradación de las proteínas. La Ubiquilina protege las células contra el estrés oxidativo(Fasano et al. 2018; Y. Liu et al. 2014). CYB561 es una proteína de la familia citocromo que regula la regeneración y homeostasis de ascorbato dentro de las vesículas secretoras. Mutación en este gen está asociado con hipotensión ortostática neurogénica con deficiencia de la actividad de la enzima dopamina beta hidroxilasa y disminución en la secreción de norepinefrina en tejido cerebral(van den Berg et al. 2018). SYT 13 codifica para la proteína sinaptotagmina-13 el cual interviene en procesos de tráfico vesicular, exocitosis y secreción de neurotransmisores. Además la sinaptotagmina 13 regula la exocitosis dependiente del calcio. La expresión de estos genes se encuentran disminuidas en los pacientes que tienen el alelo de riesgo para el SNP localizados en el gen SH3GL2.

14.5- Los SNP asociados con NOH alteran la expresión de varios genes en un mecanismo dependiente del desequilibrio de ligamiento.

Mapeamos el efecto de las variantes de SNP en la expresión del gen cis en una forma de desequilibrio de enlace de SNP asociado con NOH. Exploramos los efectos de las variantes encontradas asociadas con NOH en la expresión de otros genes, utilizando un análisis de cis-eQTL de ± 1 M de distancia(Tabla 3 y Material complementario).

Encontramos que algunos SNP ubicados en SCN3A aumentaron la expresión de SCN1A y SCN9A antisentido RNA1 (SCN1A-AS1) en el cerebelo. Otros estudios encontraron que las mutaciones en el gen SCN9A se asocian con frecuencia con disfunción autonómica en pacientes que portan esas mutaciones(Yuan et al. 2013). Encontramos estos genes

expresados en el hipocampo, la corteza frontal, la corteza cingulada anterior, la amígdala, el cerebelo, la médula espinal, los ganglios basales (núcleos caudado, accumbens y putamen) y el hipotálamo, todas las regiones del cerebro involucradas en la activación del sistema nervioso autónomo y regulación (Ruffle et al. 2021). Como se mencionó, la NOH en la EP se ha considerado tradicionalmente como resultado principalmente de la pérdida simpática posganglionar (Rafanelli et al. 2019). Nuestros resultados desafían esta noción, sugiriendo una disfunción mayor que involucra no solo al sistema nervioso periférico sino también al central.

14.6- La variante rs34311866-TMEM175 disminuye la expresión de SNP DGKQ basada en LD y aumenta el riesgo de NOH.

Hemos encontrado una expresión disminuida del gen DGKQ donde se ubican SNP que están en desequilibrio de ligamiento con los SNP rs34311866-TMEM175 asociados con NOH en nuestro estudio. La DGKQ (Diacylglicerol Quinasa Theta), enzima que contiene 941 aminoácidos, con una masa molecular aparente de 110 kDa (Houssa et al. 1997), con un dominio rico en cisteína, otro rico en prolina y un dominio pleckstrina, es responsable de la producción de ácido fosfatídico a partir de diacylglicerol (Houssa et al. 1997). Ácido fosfatídico y sus lípidos derivados aguas abajo, como el fosfatidilinositol-4,5-bisfosfato (PIP₂) (Falkenburger et al. 2010), lípidos de fosfatidilinositol fosforilados y otros fosfoinosítidos (Cremona et al. 1999) (Posor, Jang, and Haucke 2022), son lípidos involucrados en muchos procesos celulares, incluida la transducción de señales, la reorganización y regulación del citoesqueleto (Falkenburger et al. 2010), el tráfico celular, la formación de vesículas endocíticas y el reciclaje de vesículas (Cremona et al. 1999) crecimiento neural, crecimiento celular proliferación y supervivencia, y su homeostasis desordenada se asocia con muchas condiciones patológicas como la neurodegeneración (Tei and Baskin 2020). Goldschmidt y otros, usando el método de indicador óptico sensible al pH sinaptofisina-pHluorina para monitorear la dinámica de reciclaje de vesículas sinápticas, demostró que un agotamiento de la expresión del gen DGKQ (también llamado DGK θ o DGK theta) basado en la tecnología de silenciamiento de shRNA altera significativamente el reciclaje de vesículas sinápticas tanto in vitro como in vivo. en DGKQ^{-/-} (Knockout) modelos de ratones con neuronas. En este mismo estudio, se demostró que es probable que DGKQ sea importante para la transmisión sináptica cuando se requiere una actividad neuronal sostenida (Goldschmidt et al. 2016). Además, se encontró que el reciclaje vesicular en neuronas derivadas de ratones DGKQ-knockout, transfectados con el ^{myc}DGKQ que contiene el punto de mutación G648A que inactiva la

función catalítica de la enzima, no pudo ser rescatado a pesar de la expresión normal de ^{myc}DGKQ en esas neuronas (Goldschmidt et al. 2016). En base a estos datos, es lógico pensar que la depleción de DGKQ produciría una disminución de la concentración de ácido fosfatídico y sus derivados fosfoinosítidos lipídicos y los consiguientes trastornos en el proceso de reciclaje de vesículas cargadas de neurotransmisores a través de la membrana presináptica donde es necesaria la actividad catalítica de DGKQ. para el éxito de este proceso. Se demostró que la expresión de ARNm de DGKtheta mediante el método de hibridación in situ estaba fuertemente elevada en la corteza cerebelosa y el hipocampo del cerebro de ratones sanos(Houssa et al. 1997). En nuestro estudio, hemos encontrado, de manera dependiente del Desequilibrio de ligamiento del SNPs-rs34311866-TMEM175 asociado a NOH, una disminución de la expresión del gen DGKQ en la corteza frontal, el cerebelo y el hipocampo de NOH⁺ en comparación con los pacientes NOH. Muchos trabajos previos han estudiado el papel del cerebelo, el hipocampo y la corteza frontal en la aparición y gravedad de la NOH debido a la conectividad que fortalece con el tronco encefálico. Durante la prueba de Valsalva, los controles exhiben una conectividad significativamente mayor entre el tronco encefálico y el cerebelo (tanto el cerebelo anterior izquierdo como el cerebelo posterior bilateral, el vermis) y entre el tronco encefálico y la corteza (corteza prefrontal medial) (Baker and Kimpinski 2019)(Baker, Paturel, and Kimpinski 2019) y mayor activación en hipocampo derecho(Baker, Paturel, and Kimpinski 2018) que NOH⁺ pacientes en análisis de resonancia magnética funcional. Estos datos permiten postular que alteraciones en la homeostasis del ácido fosfatídico derivadas de una disminución en la expresión del gen DGKQ podrían alterar procesos celulares como la activación de la enzima PKC-ácido fosfatídico dependiente de la proliferación celular, la alteración en la transmisión sináptica y el reciclaje de neurotransmisores, y producir neurodegeneración. en el cerebelo, corteza frontal e hipocampo (regiones donde la expresión de DGKQ está disminuida), afectando el circuito de conexión de estas regiones con el tronco encefálico y la consiguiente aparición de disfunción autonómica y NOH. Sin embargo, se necesitan estudios más detallados para confirmar estos hallazgos. Aunque trabajos previos han demostrado la asociación entre la enfermedad de Parkinson y algunos SNPs localizados en el gen DGKQ(Y. P. Chen et al. 2013),(Simón-Sánchez et al. 2011), pero, hasta ahora, la asociación del gen DGKQ con NOH no se ha estudiado en pacientes con EP.

14.7- TMEM 175

TMEM175 es una proteína transmembranal de canales de potasio en los endolisosomas que intervienen en el transporte de ion de potasio en dichas estructuras celulares regulando el pH dentro de la luz endolysosomal participando de esta forma a las actividades autofágicas y lisosómicas (Cang et al. 2015). Este gen está asociado con enfermedades como Meier-Gorlin Syndrome 3 y enfermedad de Parkinson de inicio temprano (Cang et al. 2015; Qu et al. 2022). Según (Qu et al. 2022), pareciera que una inhibición de la actividad del TMEM175 por el Bcl-2 provocaría una alteración en la homeostasis mitocondrial de las células dopaminérgicas y producción de especies reactivas de oxígeno (ROS) provoca apoptosis independiente de la vía de las caspasas (Qu et al. 2022). Se produjeron alteraciones motoras y muerte celular dopaminérgica en modelos de animales parkinsoniano basado en MPTP (1-Metil-4-fenil-1, 2, 3, 6-tetrahidropiridina) con delección del gen TMEM175 (Qu et al. 2022).

Otros trabajos han encontrado que en modelos tanto in vitro como en animales Knockout para el gen TMEM175; tanto los heterocigotas como los homocigotas presentan acumulación intracelular de la proteína alfa-sinucleína fosforilada patogénica (en un 50%) en relación a los wild type, (Jinn et al. 2017). La proteína quinasa B (AKT) parece tener una función preponderante en mantener una conformación abierta del canal de potasio TMEM 175 (Wie et al. 2021) al unirse con el AKT independientemente de la actividad treonina-serina del AKT, para prevenir alteraciones celulares producidas por estrés.

Los SNP rs34311866 y rs34884217 localizados en el gen TMEM175, se encontraron asociados a la NOH. La primera variante rs34311866 (C/T) se encuentra en la región codificante 3'-UTR del gen TMEM 175 donde se cambió una metionina por una treonina en la posición M393T. Mientras que el segundo rs34884217 se localiza en la región no codificante intrónica y en un sitio de splicing alternativa donde el cambio es un 65 Q>P (Oscanoa et al. 2020). Los dos cambios fueron predichos como benignos según los softwares SIFT (Ng and Henikoff 2001) y Polyphen (Adzhubei et al. 2010) (Consultado el día 26/09/2022). El SNP rs34311866 tiene una Islot CpG de 638 pb representado por un 68.7% de GC y un ratio de 0.68 (% de GC Observado/GC esperado).

13.8- OGFOD2.CCDC62

El SNP rs11060180 localizado en la vecindad de los genes OGFOD2.CCDC62 se encuentra asociado a NOH en los modelos aditivos y recesivos. Este SNP se localiza en la región no codificante del gen CCDC62 (McLaren et al. 2016). Este SNP se localiza en regiones del

ADN donde se destaca una huella H3K4me1. Esta marca epigenética indica que haya una monometilación a nivel del cuarto aminoácido de la proteína histona 3 y generalmente está asociado con enhancers que controlan y regulan la expresión de genes. Otra marca epigenética encontrada en esta región es la H3K27me3 caracterizado por una trimetilación de la lisina 27 de la proteína histona 3 unida a la región del ADN donde se encuentra nuestro SNP. Se sabe que en células neuronales esta marca epigenética está asociada con disminución de la expresión genética. Ningún otro estudio ha demostrado anteriormente la asociación entre la NOH y esta variante. Otros estudios se necesitan para verificar los mecanismos por los cuales este SNP se asocia a la NOH.

14.8- FAM47E.STBD 1

El SNP rs6812193 localizado en las regiones intrónicas no codificantes de los genes STBD/FAM47E, se encontró asociado con NOH en este estudio en un modelo genético recesivo. Se predice estar localizado este SNP en una región donde se encuentra un enhancer activo en células neuronales (RegBuild) y una marca epigenética H3K27me3 (ENCODE) (ENCODE Project Consortium 2012) (Luo et al. 2020). Este SNP está fuertemente en desequilibrio de ligamiento ($R^2 : 0.98-1$, $D' : 1$) con otros SNPs que aumentan la expresión del FAM47E.

14.9- SCN3A.SCN2A

El SNP rs353116 localizado en la región intrónica no codificante del gen SCN2A está asociado con el NOH en el modelo genético dominante según nuestro estudio. No se han encontrado datos de predicción disponibles que nos indicarían si es patológico o no esta variante en los software polyphen (Adzhubei et al. 2010) (Ramensky, Bork, and Sunyaev 2002) ni SIFT. Dos marcas epigenéticas de tipo H3K4me1 (una metilación en la lisina 4 de la proteína Histona), enhancer que regula la expresión genética y H3K9me3 (trimetilación de la lisina 9 en la proteína histona) asociado con heterocromatina constitutiva (Volpe et al. 2002), están en la localización de este SNP (ENCODE) (ENCODE Project Consortium 2012), (Luo et al. 2020). SCN2A (Proteína del canal de sodio subunidad alfa 2), son glicoproteínas transmembranales constituidas de una subunidad alfa con 24 segmentos transmembranales y uno o más subunidades beta regulatorias que intervienen en la generación y transmisión del potencial de acción en las neuronas (Ahmed et al. 1992). Están heterogéneamente expresadas y distribuidas en todo el sistema nervioso y algunas mutaciones en su estructura están relacionadas con la epilepsia (Malacarne et al. 2001), (Malacarne et al.

2001; Heron et al. 2002) y autismo (S. J. Sanders et al. 2012). Hasta este momento, es el primer estudio en el cual se encontró asociación entre este gene particularmente esta variante y la NOH, aunque existe otro estudio que ha descubierto asociación entre la disautonomía y la SCN9A (canal de sodio activado por voltaje subunidad alfa 9) (Bartholomew et al. 2014), (M.-K. Kim et al. 2013), (Han et al. 2012), otros canales de sodio responsable de la generación y propagación del potencial de acción en las neuronas, los cuales son conocidos por su participación en el mecanismo del dolor (Cox et al. 2006), (Cox et al. 2006; Fertleman et al. 2006), (Eberhardt et al. 2014). En nuestros datos de análisis diferencial de genes, encontramos que el gen SCN2A está subregulado en los pacientes con NOH con respecto a los que no tienen NOH.

14.10- SH3GL2

En este estudio, el SNP rs13294100 encontrado asociado a NOH con un modelo dominante, se encuentra localizado en la región no codificante intrónica del gen SH3GL2 (chr9P22.2) (Cunningham et al. 2022), (Oscanova et al. 2020), (Dayem Ullah et al. 2018). No se encontraron informaciones en las bases de datos PolyPhen y SIFT en relación a la predicción de la capacidad patogénica de esta variante del gen SH3GL2. Este SNP se encuentra localizado en una región rica en islotos CpG de 1083 pb, constituido de 72% GC (ratio: 0.9 – % CG observado / % CG esperado). La base de datos RegBuid predice que esta región donde se localiza el SNP rs13294100 se encuentra un promotor activo y marcas epigenéticas tipo H3K27me3 (expresión génica a la baja), H3K4me3 en líneas celulares de neuronas bipolares (ENCODE Project Consortium 2012), (Luo et al. 2020) y otras neuronas (Roadmap, <http://www.roadmapepigenomics.org>). Estas marcas epigenéticas que comparten localización con el SNP rs13294100 nos estaría indicando que este último provoca disminución en la expresión del gen SH3GL2.

El gen se encontró significativamente enriquecido para vías metabólicas y de señalización como vía de señalización retrógrada de neurotrofinas (p-valor: 0.001299), Regulación a la baja del EGFR (p-valor: 0.002877), Biogénesis de las vesículas del lisosoma (p-valor: 0.003248), Biogénesis de las vesículas asociadas al Golgi (p-valor: 0.005197), Desprendimiento de vesículas de la red trans-Golgi (p-valor: 0.006682), Reconocimiento de carga en la endocitosis mediada por clatrina y Endocitosis mediada por clatrina (p-valor: 0.009745 y 0.013457). La vía de señalización retrógrada de neurotrofina interviene en procesos de supervivencia celular. Se sabe que actividad disminuida de esta vía de señalización está asociada con eventos de señalización apoptótica incluyendo la translocación de la proteína Bax a las mitocondrias, liberación de citocromo C, activación del apoptosoma y

muerte celular dependiente de la caspasa 1 en neuronas simpáticas(Deckwerth et al. 1998), (Zweifel, Kuruvilla, and Ginty 2005). Otra vía de señalización a destacar es la regulación a la baja de la expresión del EGFR, la cual se sabe estar asociado con muerte neuronal y neurodegeneración(Sibilia et al. 1998), mientras que sobreexpresión está asociada con aumento en el crecimiento neuronal y neurogénesis(Goldshmit, Greenhalgh, and Turnley 2004). Alteración de las vías de la biogénesis de vesículas asociados a lisosomas y golgi y el reconocimiento de segmento cargo asociado endocitosis mediada por clatrina nos estaría indicando que este SNP está asociado con alteración en los procesos de exocitosis y endocitosis de neurotransmisores y alteración de la neurotransmisión. No existen datos disponibles en la base de datos GTX sobre la expresión de este gen(SH3GL2) relacionado a este SNP. Está claro que las marcas epigenéticas(comentando anteriormente) cercanas al SNP rs13294100 estaría afectando la expresión del gen SH3GL2, lo que provocará alteración en la expresión del gen EGFR y en la disponibilidad de las neurotrofinas en las neuronas por la alteración en el proceso de transporte retrógrado de las mismas. En nuestros datos, hemos encontrado que el gen SH3GL2 está diferencialmente expresado en los pacientes con NOH con respecto a los que no tienen NOH. En estos pacientes la SH3GL2 se encuentra disminuida en su expresión según nuestro estudio.

En el análisis de eQTL para el SNPs localizado en el gen SH3GL2, 585 genes eQTL estadísticamente significativos(p-valor y FDR <0.05) fueron encontrados asociados al SNP SH3GL2_rs13294100 en el análisis de regresión lineal y la ANOVA. Un análisis de network de los 585 eQTLs permitió construir 4 clusters. El primer cluster enriquece significativamente para procesamiento de ARNm, spliceosoma. El segundo cluster para biogénesis, transporte y localización del complejo ribonucleoproteico, regulación del transporte nuclear y nucleocitoplasmático. El tercer cluster enriquece para procesos de transporte transmembranal de protones contra gradiente electroquímico, transporte de electrones acoplado al transporte de protones, cadena de transporte aeróbico de electrones, síntesis de ATP acoplado a transporte de electrones, procesos de metabolismo del ATP, el ensamblaje del complejo NADH deshidrogenasa, ensamblaje del complejo I de la cadena respiratoria mitocondrial. El cuarto cluster donde se localiza nuestro gen de interés(SH3GL2), enriquece para procesos de biosíntesis y metabolismo de fosfolípidos, lípidos y compuestos organofosforados. Además, se encontró que el gen SH3GL2 donde se localiza el SNP está conectado con 5 genes los cuales tienen mayor importancia por permitir las conexiones entre los diferentes clusters y el gen SH3GL2. Estas proteínas se eligieron basándose en sus grados, sus centralidades en el Network. Esto nos permite añadir que estos 5 genes son puntos centrales en el network por los cuales el flujo de

información tiene que pasar para viajar de cualquier nodo a otros: ATP13A2, LAS1L, KIAA0895, CXADR, IKBKB.

14.11- SIPA1L2

El SNP rs10797576 localizado en la región intrónica cercana al gen SIPA1L2(1q42.2), está asociado con la NOH en un modelo genético aditivo reduciendo el riesgo de padecer dicha condición en los pacientes con enfermedad de Parkinson. Los softwares SIFT y Polyphen no tienen datos de predicción de efecto patogénico para esta variante. Según la base de datos ENCODE, este SNP está cerca de marcas epigenéticas de tipo H3K4me1 y H3K36me3 (Roadmap) en neuronas. La proteína SIPA1L2 (Signal-induced proliferation-associated 1-like protein 2) activa la vía de señalización TrkB/BDNF al unir directamente la Trkb con la proteína dineína mediante su unión a la Snapina y el transporte retrógrado del complejo resultante(Trkb-dineína/Snapina/SIPA1L2/) hacia el cuerpo neuronal(Andres-Alonso et al. 2019).

14.12- ITGA8

El SNP rs10906923 localizado en la región intrónica del gen ITGA 8(10p13), está asociado con la NOH en un patrón de herencia de dominancia aumentando el riesgo del mismo en pacientes con enfermedad de Parkinson. No se ha encontrado ninguna información sobre marcas epigenéticas en las bases de datos consultados respecto a este SNP(Roadmap, ENCODE, miRNA, RegBuild). El SNP rs10906923 disminuye la expresión del gen ITGA8(effect size: -0.412, p-valor: 6×10^{-6}) en la corteza cerebral, el núcleo caudado y accumbens. Otros SNPs encontrado en desequilibrio de ligamiento con el SNP rs10906923 disminuyen la expresión del gen ITGA8(R: 0.5-1, D': 08-1).

14.13- MIR4697

El SNP rs329648 localizado en la región 3 downstream del gen MIR4697(11q25), está asociado con alto riesgo de NOH en un modelo dominante en pacientes con EP. El SNP se localiza en las cercanías de marcas epigenéticas de tipo H3K4me1, H3K27ac en neuronas bipolares(ENCODE), H3K4me1 en neuronas(Roadmap).

14.14- ELOVL7.NDU FAF2

El SNP rs2694528 encontró estar en asociación en un modelo dominante con la NOH en

pacientes con EP, se localiza en la región intrónica del gen **NDUFAF2**. No se encontraron información sobre la predicción patogénica de dichos SNP sobre el gen **NDUFAF2** en las bases de datos Polyphen y SIFT. Tampoco se encontraron datos sobre marcas epigenéticas regulatorios relacionados con el SNP rs2694528 y su localización en las bases de datos revisadas (RegBuild, ENCODE, Roadmap). Hemos encontrado que este SNP está enriquecido para las vías de señalización que tienen que ver con la vía de biogénesis del complejo I, cadena de transporte de electrones en la respiración celular, síntesis de ATP, la vía tricarboxílico (ciclo de Krebs). La proteína **NDUFAF2 (NADH:ubiquinone oxidoreductase complex assembly factor 2)**, es una proteína chaperona que interviene en el ensamblaje de la proteína NADH deshidrogenasa localizada en la región interna de la membrana mitocondrial. Se aprecia que este SNP está asociado con disminución de la expresión del gen **NDUFAF2** ($D': 1$, $R: 0.24$,) (effect size: -0.23 , p-valor: 4×10^{-5}) en corteza cerebral y putamen. Sería lógico pensar que una disminución de la expresión de este gen podría estar asociado con disminución de la actividad de la cadena de transporte de electrones y del ciclo de Krebs y por consiguiente defecto a nivel de la respiración celular y producción de ATP a nivel de las regiones cerebrales donde se ve afectado (Corteza cerebral y Putamen). Además, en nuestro estudio, hemos encontrado que este gen está downregulada en su expresión en los pacientes que tienen NOH con respecto a los que no tienen NOH.

14.15- ACMSD/TMEM163

El SNP rs6430538 se localiza en la región no codificante del gen **ACMSD**, está asociado con NOH en un modelo aditivo reduciendo el riesgo de padecer la NOH en pacientes con EP. La base de datos ENCODE predijo marcas epigenéticas de tipo H3K27me3 en la localización del SNP rs6430538 en neuronas bipolares. Hemos encontrado que este SNP está asociado con ligeros aumentos en la expresión de la ciclina T2 ($R^2: 1$, $D': 1$, Effect size: 0.171 , P-valor: 1×10^{-4}) en el cerebelo, y disminuye la expresión del gen **TMEM 163** ($R^2: 1$, $D': 1$, Effect size: -0.24 , P-valor: 4×10^{-6}) en la corteza cerebral y caudado. Se predice que la proteína **TMEM 163** permite la actividad de unión a iones de zinc, participa en la importación de iones de zinc a la vesícula sináptica. Se localiza en la membrana del endosoma y es activo en la vesícula intracelular y en la membrana plasmática y es un componente integral de la membrana de la vesícula sináptica (<https://www.genecards.org/cgi-bin/carddisp.pl?>).

14.16- DDRGK1

El SNP rs55785911 (20p13), localizado en región intrónica no codificante del gen **DDRGK1**, se encuentra con tendencia a la asociación a NOH (p-valor alrededor de 0.05) en nuestros

datos. En nuestros datos, hemos encontrado que el gen DDRGK1 está sobreexpresado en los pacientes con NOH con respecto a los sin NOH([ver tabla 27](#)).

Hemos encontrado, en el análisis de eQTL, que los genotipos homocigotas y heterocigotas para el alelo menor del SNP rs55785911-DDRGK1 alteran la expresión génica de 17 genes tanto en el análisis de regresión lineal como la ANOVA(p-valor y FDR<0.05). De los 17 genes, NOTCH1, KIF4A, TSPYL2 son los que más importancia funcional tendría según los resultados del análisis de priorización usando la aplicación Causal-Path de Cystoscope.

14.17- EGF

El SNP rs4444903(4q25) localizado en la región 5'-UTR del gen EGF, se encuentra asociado con la NOH disminuyendo el riesgo de padecer la última en pacientes con EP(Ensembl, Refseq). Ninguna predicción de la patogenicidad de este SNP se encontró en las bases de datos Polyphen ni SIFT.

14.18- BIN3

El SNP rs1180104-BIN3 está asociado en un modelo recesivo de transmisión genética, disminuyendo el riesgo de la aparición de la NOH en los pacientes con la EP. Un total de 75 eQTLs fueron encontrados asociados al SNP rs2280104(BIN3). Hasta ahora hemos dilucidado que los genotipos TT del SNP rs2280104-BIN3 se asocian con una sobreexpresión del gen PPP1R14C, el cual a su vez está conectado con muchos otros genes que se involucran en procesos celulares y moleculares de crecimiento, proliferación de las neuronas. Además, hemos encontrado procesos involucrados como uniones a citoesqueleto, uniones a quinasas y a microtúbulos etc.

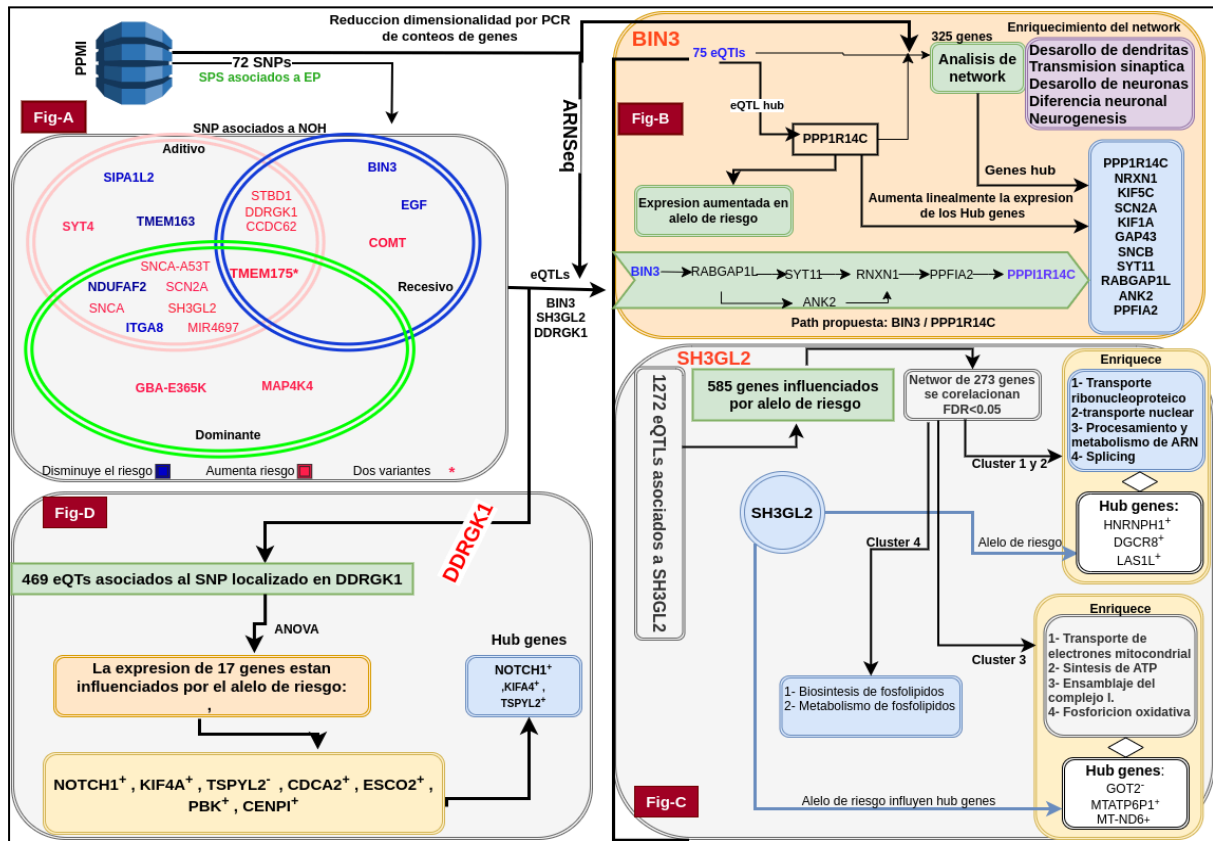


Figura 34. Figura de resumen 1: Resumen de los resultados de los tres primeros objetivos de la tesis. En la **fig-A** se describen los SNP (los genes donde se localizan para facilitar la presentación) y su asociación con la NOH según los modelos aditivo, recesivo y dominante considerado. Los de color azul disminuye el riesgo de NOH y los de color rojo aumentan el riesgo. El asterisco quiere decir que en este gen se encuentran más de un SNP asociados a NOH. Los SNPs localizados en los genes BIN3, EGF, COMT se asocian con NOH en modelo solo recesivo. Los SNPs localizados en los genes GBA, MAP4K4 se asocian con la NOH solo en modelo dominante. Los SNPs localizados en los genes SIPA1L2, TMEM163, SYT4 están asociados con NOH en modelo solo aditivo. Los SNPs localizados en los genes SNCA-A53T, SNCA, NDUFAF2, SCN2A, SH3GL2, ITGA8, MIR4697 se asocian con el NOH en modelos aditivo y dominante. Los SNPs localizados en los genes STBD1, CCDC62, DDRGK1 están asociados con NOH en los modelos aditivo y recesivo. Los dos SNPs localizados en el gen TMEM175 están asociados con la NOH en modelos tanto aditivo, dominante y recesivo. En la **fig-B**, se describe el análisis de eQTL realizado entre el SNP BIN3 asociado con la NOH y los genes de células dopaminérgicas derivadas de iPSCs de pacientes con EP. Se encontraron 75 eQTLs, los cuales se cruzan con el análisis de correlación entre los genes, previa reducción de dimensionalidad por PCA. Se encontraron que el eQTL cuya expresión está aumentada en pacientes con el alelo de riesgo forma un cluster con 324 genes los cuales enriquecen para crecimiento diferenciación neuronal y neurogenesis. Los genes más importante llamados hub genes que intervienen en estos procesos son: PPP1R14C, NRXN1, KIF5C, SCN2A, KIF1A, GAP43, SNCA, SYT11, RABGAP1L, ANK2, PPFIA2. El eQTL aumenta la expresión de los genes más importantes genes hub en una regresión lineal. En la **fig-C**, se describen los resultados del análisis de eQTL entre el SNP localizado en el gen SH3GL2 y los mismos genes utilizados en el análisis de la fig-A. De los 1272 eQTLs asociados con el SNP localizado en el gen SH3GL2, 585 tienen influenciados sus expresiones por el alelo de riesgo y de ellos 273 se correlacionan entre ellos mismos. Estos 273 forman 4 cluster en un análisis de network. Se presenta el enriquecimiento de cada cluster en la fig-C. El alelo de riesgo del SNP-SH3GL2 está asociado con aumento en la expresión de los genes hubs (más importantes en el network) HNRNPH1, DGCR8, LAS1L son cuales enriquecen para procesos de procesamiento, transporte de ARNm y splicing. Además el alelo de riesgo disminuye la expresión del gen GOT2 y aumenta la expresión de los genes MTND6 y MTATP6P1 que enriquecen para la cadena de transporte de electrones y síntesis de ATP. Además el alelo de riesgo está asociado con genes que enriquecen para la biosíntesis y metabolismo de fosfolípidos. En la **fig-C**, se describen los resultados de los eQTLs encontrados asociados con el alelo de riesgo del SNP localizado en el gen DDRGK1.

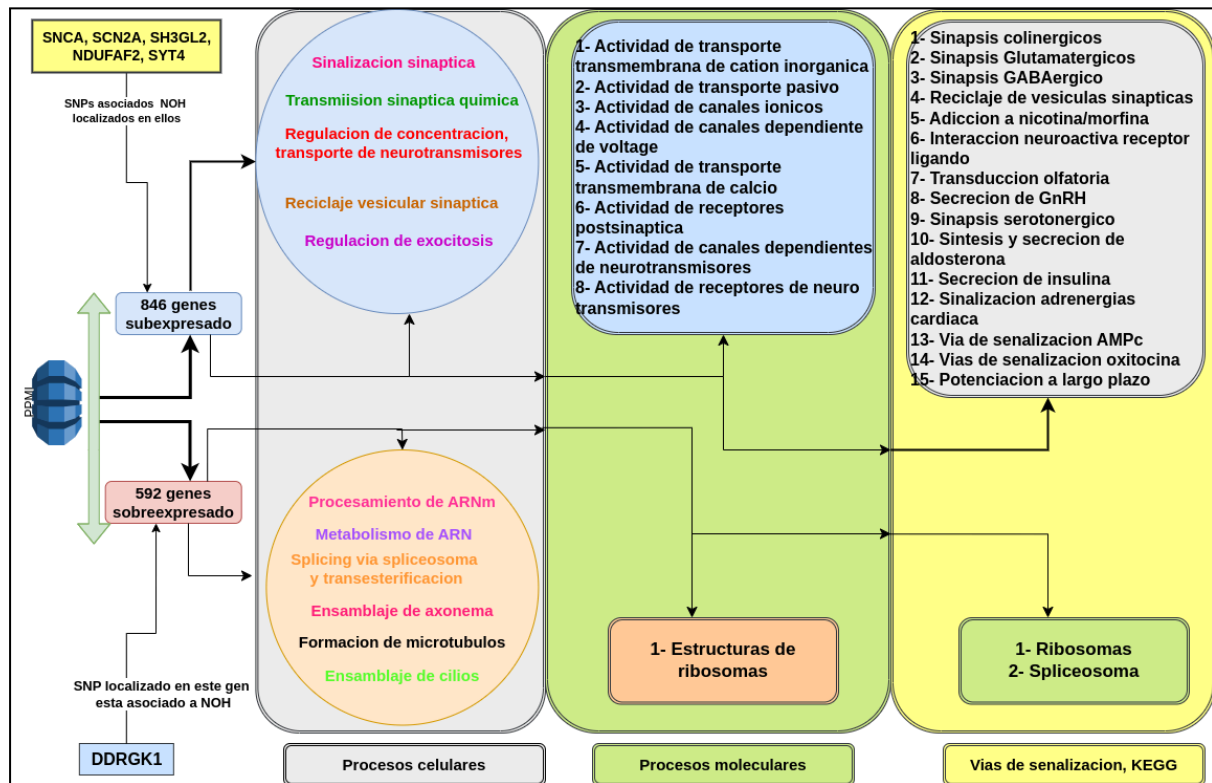


Figura 35. Figura de resumen 2: Análisis diferencial de genes entre los pacientes con NOH y grupo control. En esta figura se puede observar que los pacientes que tienen NOH presentan los procesos celulares como transmisión sináptica, neurotransmisión, regulación de la exocitosis y reciclaje de vesículas sinápticas disminuidos o inhibidos. Las actividades de canales iónicos, canales dependientes de voltaje, canales dependientes de neurotransmisores, transporte pasivo, transporte membrana de calcio, actividad de receptores postsinápticos están alteradas en los pacientes con NOH. Las funciones de las vías de señalización adrenérgicas cardíacas, síntesis y secreción de aldosterona, vía de señalización de AMPc, secreción de insulina, sinapsis glutamatérgicas, serotoninérgicas, colinérgicas, Gabaérgicas, secreción de GnRH, vía de señalización de oxitocina, transducción olfatoria, interacción neuroactiva receptora-ligando, reciclaje de vesículas sinápticas, potenciación a largo plazo, adicción a nicotina y morfina están disminuidas, alteradas y disfuncionales en los pacientes que tienen NOH. Por otro lado, los procesos como ensamblaje de axonemas y ciclos, formación de microtúbulos, procesamiento y metabolismo de ARN, formación de estructuras ribosomales, splicing por transesterificación y por spliceosoma están aumentados.

Capítulo XV- Conclusión

1- SNP asociados a NOH, epistasis entre SNPs y NOH, SNPs asociados a MDS-UPDRS

En conclusión, en esta tesis se encontraron 21 SNPs asociados a NOH en enfermos de parkinson. De los cuales , 7 aumentan y dos disminuyen el riesgo de padecer NOH en un modelo dominante mientras que en un modelo recesivo 5 aumenta y dos disminuyen el riesgo de NOH. Hemos encontrado que los SNPs rs199347(GPNMP) y rs8005172(GALC/GPR65) se interactúan en un modelo dominante para disminuir el riesgo de NOH mientras tanto los SNPs rs823118(NUCKS1) y rs356181(SNCA) se interactúan para aumentar el riesgo de NOH en pacientes con EP. El SNP encontrado asociado con UPDRS I además tiene cifras altas de MDS-UPDRS-I en grupo NOH con respecto al grupo control es el SNP IL1R2-MAP4K4_rs34043159. Mientras tanto para el MDS-UPDRS III encontramos SNCA_A53T_rs104893877 tiene valores de UPDRS III significativamente más elevados en pacientes con NOH con respecto a los que no tiene NOH. Además, se

identificó que el SNP localizado en el gen CTSB tiene cifras de UPDRS IV más altos en pacientes con NOH con respecto a los que no tienen NOH. Estos datos indican que los pacientes que presentan el alelo de riesgo para el SNP localizado en el gen MAP4K4 tienen alto riesgo de desarrollar alteraciones cognitivas, depresión, ansiedad asociados a la EP mientras que los pacientes presentando el alelo de riesgo para el SNP SNCA(A53T) tienen mayor riesgo de desarrollar complicaciones motoras de la EP. Los pacientes que presentan el alelo de riesgo para el SNP localizado en el gen SNCA(A53T) o que son homocigotas para el alelo menor del SNP localizados en el gen CTSB tienen alto riesgo de desarrollar la forma grave de la enfermedad de parkinson.

2- Algoritmos de aprendizaje automático para la predicción de NOH

De los 7 algoritmos de aprendizaje automático utilizado para la predicción de la NOH a través de los SNPs asociados a NOH, hemos encontrados 3 algoritmos que mejor predicen los datos tanto cuando se usan solo los datos genéticos(Score poligénico) o si se usan los datos genéticos junto con los clínico-demográficos. Ningún algoritmo pudo predecir la NOH al usar solo los datos clínicos o demográficos pero sí cuando se usan los datos genéticos.

En este trabajo de tesis, identificamos 21 variantes genéticas vinculadas a la hipotensión ortostática neurogénica (NOH) en pacientes con enfermedad de Parkinson (EP). Estos resultados asientan la base genética de la NOH en los pacientes con EP ya que ningún estudio anterior ha demostrado la amplia participación de variantes genéticas en esta condición tan importante en enfermos de parkinson. La NOH afecta al 50% de los pacientes con EP y se asocia con complicaciones y altas tasas de mortalidad en este grupo. La utilización de herramientas de bioinformática en entornos clínicos puede ayudar a predecir el riesgo de NOH en los pacientes, permitiendo así tomar decisiones más personalizadas en la prevención, tratamiento y diagnóstico temprano. En este trabajo, se desarrollaron métodos computacionales para calcular el riesgo poligénico y predecir la NOH en pacientes con EP, empleando algoritmos de aprendizaje automático y datos genotípicos y/o clínicos.

3- Análisis de eQTLs explican los mecanismos moleculares de la NOH

Nuestro análisis de eQTLs utilizando modelos de regresión y análisis de redes basados en la correlación de genes ha identificado a BIN3-PPP1R14C como un factor protector clave en el desarrollo de hipotensión ortostática neurogénica en pacientes con enfermedad de Parkinson. El alelo de riesgo del SNP localizado en el gen BIN3 aumenta la expresión del eQTL PPP1R14C asociado a BIN3. Se encontró que PPP1R14C formó un cluster con 325 proteínas que intervienen en el desarrollo de dendritas, morfogénesis celular, señalización sináptica, diferenciación neuronal y neurogénesis. Se identificaron seis proteínas como las más significativas en el análisis de priorización de redes (NRXN1, KIF5C, SCN2A, KIF1A, GAP43 y SNCB), y otras cinco están implicadas en la vía que conecta BIN3 con la proteína

PPP1R14C (NRXN1, PPFIA2, RABGAP1L, ANK2, SYT11). Este hallazgo sugiere que la interacción entre BIN3 y PPP1R14C puede desempeñar un papel crucial en la mitigación de los síntomas de hipotensión ortostática en individuos con enfermedad de Parkinson. La comprensión de los mecanismos genéticos y moleculares que subyacen a este efecto protector podría conducir potencialmente al desarrollo de terapias dirigidas para el manejo de la hipotensión ortostática neurogénica en estos pacientes. Por esta razón, el gen y producto génico del PPP1R14C es una propuesta en este estudio como blanco terapéutico contra la NOH en enfermos de parkinson. Las estrategias de aumento en la actividad y/o expresión de esta proteína serían adecuadas.

En esta tesis se aportaron nuevos datos que permiten dilucidar los mecanismos moleculares por los cuales el SNP localizado en el gen SH3GL2 aumenta el riesgo de NOH. En primer lugar, pareciera que el SNP se asocia con disminución en la expresión del gen SH3GL2. La base de dato RegBuild predice que esta región donde se localiza el SNP rs13294100 se encuentra un promotor activo y marcas epigenéticas tipo H3K27me3 (expresión génica a la baja), H3K4me3 en líneas celulares de neuronas. Estas marcas epigenéticas que comparten localización con el SNP rs13294100 nos estaría indicando que este último provoca disminución en la expresión del gen SH3GL2. Además, en nuestros datos, utilizando un análisis diferencial de genes, hemos identificado disminución en la expresión del gen SH3GL2 en los pacientes con NOH produciendo de esta forma disminución en la neurotransmisión, endocitosis, y transmisión sináptica además de disminución en el transporte retrógrada de neurotrofina y en la biogénesis de vesículas y tráfico endolisosomal. Además se encontró que los pacientes con el alelo de riesgo para el SNP localizado en el gen SH3GL2 presentan significativamente disminución en la expresión de genes claves que participan en la neurotransmisión como el SYT13, en la transmisión sináptica y tráfico endolisosomal como SYNJ-1, en el sistema de ubiquitilación y proteasomal como la ubiquilina (UBQLN1), en la homeostasis del calcio como C1SD2, en la eliminación en la protección del cerebro contra la neurotoxicidad de los iones metálicos como el ATP13A2. Se identificó en este estudio que los pacientes con el alelo de riesgo para el SNP localizados en el gen SH3GL2 tienen expresión disminuida del gen CYB561. Mutación en este gen se asocia con NOH y concentraciones de noradrenalina disminuida y actividad del gen dopamina beta hidroxilasa.

En un análisis de expresión diferencial génica entre el grupo NOH y el grupo control, se identificaron 6 genes que están regulados de forma distinta entre ambos grupos. Los genes SNCA, SCN2A, SH3GL2, NDUFAF2, SYT4 muestran una disminución en su expresión en pacientes con NOH, mientras que el gen DDRGK1 está aumentado en dichos pacientes. Estos hallazgos sugieren que estos genes podrían desempeñar roles clave como posibles

blancos terapéuticos y/o biomarcadores para el diagnóstico temprano de la NOH. Este descubrimiento podría tener implicaciones significativas en el desarrollo de estrategias para el tratamiento y diagnóstico de la NOH. Además este análisis permitió entender mejor los mecanismos moleculares del desarrollo de la NOH en los enfermos de Parkinson. Hemos podido dilucidar que los pacientes que tienen NOH presentan disminución o inhibición de los procesos moleculares de neurotransmisión, exocitosis, señalización sináptica y reciclaje de vesículas sinápticas en los cuales intervienen genes como SNCA, SCN2A, SH3GL2, SYT4, NDUFAF2. Estos últimos, como se explicó previamente, están subexpresados (expresión disminuida) en los pacientes con NOH. Seis de los SNPs encontrados asociados con NOH se localizan en estos genes. Los pacientes con NOH presentan las actividades de transporte de membrana, transporte pasiva, actividad de canales dependientes de voltaje, actividad de transporte transmembrana de calcio, actividad de receptores postsinápticos, actividad de canales dependientes de neurotransmisores, actividad de receptores de neurotransmisores disminuidos e inhibidos. También, están alteradas las vías de sinapsis colinérgicas, Gabaérgicas, glutamatérgicas, serotoninérgicas, interacción neuroactiva receptor-ligando, síntesis y secreción de aldosterona, señalización adrenérgicas cardíacas, vías de señalización AMPc, potenciación a largo plazo en los pacientes con hipotensión ortostática neurogénica.

En fin, en este estudio hemos logrado determinar los mecanismos moleculares que explican la reducción del riesgo de padecer la NOH en relación al SNP rs2280104-BIN3 en los enfermos de parkinson. Nuestros hallazgos señalan que el alelo de riesgo del SNP rs2280104-BIN3 aumenta la expresión del gen PPP1R14C, el cual actúa como un inhibidor de la fosfatasa y desempeña un papel crucial en otros procesos moleculares relevantes para la función neuronal, como la proliferación neuronal y la neurogénesis. Estos descubrimientos sugieren que PPP1R14C, la proteína inhibidora de la fosfatasa, puede representar un blanco terapéutico prometedor para el tratamiento de la NOH en general y la enfermedad de parkinson en particular. Además, estrategias que fomenten la expresión o activación de esta proteína podrían ser fundamentales para el desarrollo de tratamientos eficaces en pacientes con esta enfermedad.

Hemos identificado mecanismos moleculares que pueden explicar cómo los SNPs localizados en los genes DDRGK1 y SH3GL2 aumentan el riesgo de NOH en pacientes con EP. En el caso de DDRGK1, hemos identificado que el SNP localizado en este gen altera significativamente la expresión de 17 genes de los cuales NOTCH1, KIFA4, TSPYL2, CDCA2, PBK y CENPI. Además, la expresión de DDRGK1 está aumentada en pacientes con NOH como respuesta posiblemente al estrés del retículo endoplásmico. En el caso del SH3GL2, el SNP localizado en este gen disminuye la expresión de los genes

ATP13A2(Mutación en este gen está asociada con Parkinsonismo familiar), SYNJ1, UBQLN1, CDK20, SYT13, y aumenta la expresión de MT-ND6, MTATP6P1. Estos descubrimientos dan fundamentos para dilucidar los procesos moleculares de estos SNPs asociados a NOH.

Por fin, hemos identificado varios SNPs que están asociados con aumento de cifras de UPDRS I(MAP4K4), UPDRSIII(SNCA-A53T), y UPDRSIV(CTSB). En base a estos datos, encontramos que los pacientes que presentan el alelo de riesgo para este SNP localizado en el gen MAP4K4 tienen mayor riesgo de tener alteraciones cognitivas depresión y ansiedad. Mientras que los pacientes que presentan el alelo de riesgo para el SNP SNCA A53T presentan complicaciones motores y los pacientes con el alelo riesgo del SNP localizado en el gen CTSB tienen mayor riesgo de desarrollar la forma grave de la enfermedad.

Referencias

- 1000 Genomes Project Consortium, Goncalo R. Abecasis, Adam Auton, Lisa D. Brooks, Mark A. DePristo, Richard M. Durbin, Robert E. Handsaker, Hyun Min Kang, Gabor T. Marth, and Gil A. McVean. 2012. "An Integrated Map of Genetic Variation from 1,092 Human Genomes." *Nature* 491 (7422): 56–65.
- Aarsland, Dag, Lucia Batzu, Glenda M. Halliday, Gert J. Geurtsen, Clive Ballard, K. Ray Chaudhuri, and Daniel Weintraub. 2021. "Parkinson Disease-Associated Cognitive Impairment." *Nature Reviews. Disease Primers* 7 (1): 47.
- Abe, Y., T. Kachi, T. Kato, Y. Arahata, T. Yamada, Y. Washimi, K. Iwai, K. Ito, N. Yanagisawa, and G. Sobue. 2003. "Occipital Hypoperfusion in Parkinson's Disease without Dementia: Correlation to Impaired Cortical Visual Processing." *Journal of Neurology, Neurosurgery, and Psychiatry* 74 (4): 419–22.
- Adams, James D. 2021. "Possible Causes of Parkinson's Disease." *Frontiers in Bioscience* 26 (8): 387–94.
- Adzhubei, Ivan A., Steffen Schmidt, Leonid Peshkin, Vasily E. Ramensky, Anna Gerasimova, Peer Bork, Alexey S. Kondrashov, and Shamil R. Sunyaev. 2010. "A Method and Server for Predicting Damaging Missense Mutations." *Nature Methods* 7 (4): 248–49.
- Ahmed, C. M., D. H. Ware, S. C. Lee, C. D. Patten, A. V. Ferrer-Montiel, A. F. Schinder, J. D. McPherson, C. B. Wagner-McPherson, J. J. Wasmuth, and G. A. Evans. 1992. "Primary Structure, Chromosomal Localization, and Functional Expression of a Voltage-Gated Sodium Channel from Human Brain." *Proceedings of the National Academy of Sciences of the United States of America* 89 (17): 8220–24.
- Akmal, Muhammad Aizaz, Muhammad Awais Hassan, Shoaib Muhammad, Khaldoon S. Khurshid, and Abdullah Mohamed. 2022. "An Analytical Study on the Identification of N-Linked Glycosylation Sites Using Machine Learning Model." *PeerJ. Computer Science* 8 (e1069): e1069.
- Alessi, Dario R., and Esther Sammler. 2018. "LRRK2 Kinase in Parkinson's Disease."

Science 360 (6384): 36–37.

- Alexander, T. A., and M. J. Machiela. 2020. “LDpop: An Interactive Online Tool to Calculate and Visualize Geographic LD Patterns.” *BMC Bioinformatics* 21 (1): 14.
- Andres-Alonso, Maria, Mohamed Raafet Ammar, Ioana Butnaru, Guilherme M. Gomes, Gustavo Acuña Sanhueza, Rajeev Raman, Pingan Yuanxiang, et al. 2019. “SIPA1L2 Controls Trafficking and Local Signaling of TrkB-Containing Amphisomes at Presynaptic Terminals.” *Nature Communications* 10 (1): 5448.
- Andresen, M. C., M. W. Doyle, Y. H. Jin, and T. W. Bailey. 2001. “Cellular Mechanisms of Baroreceptor Integration at the Nucleus Tractus Solitarius.” *Annals of the New York Academy of Sciences* 940 (June): 132–41.
- Ascherio, Alberto, and Michael A. Schwarzschild. 2016. “The Epidemiology of Parkinson’s Disease: Risk Factors and Prevention.” *Lancet Neurology* 15 (12): 1257–72.
- Bagos, Pantelis G. 2013. “Genetic Model Selection in Genome-Wide Association Studies: Robust Methods and the Use of Meta-Analysis.” *Statistical Applications in Genetics and Molecular Biology* 12 (3): 285–308.
- Baker, Jacquie, and Kurt Kimpinski. 2019. “Reduced Brainstem Functional Connectivity in Patients with Peripheral Autonomic Failure.” *NeuroImage. Clinical* 23 (July): 101924.
- Baker, Jacquie, Justin R. Paturel, and Kurt Kimpinski. 2018. “Impaired Cortical Autonomic Responses during Sympathetic Activation in Neurogenic Orthostatic Hypotension Characterized by Postganglionic Autonomic Dysfunction.” *Journal of Applied Physiology* 125 (4): 1210–17.
- . 2019. “Cerebellar Impairment during an Orthostatic Challenge in Patients with Neurogenic Orthostatic Hypotension.” *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 130 (1): 189–95.
- Balestrino, R., and A. H. V. Schapira. 2020. “Parkinson Disease.” *European Journal of Neurology: The Official Journal of the European Federation of Neurological Societies* 27 (1): 27–42.
- Bartholomew, F., J. Lazar, A. Marqueling, C. Lee-Messer, S. Jaradeh, and J. M. C. Teng. 2014. “Channelopathy: A Novel Mutation in the SCN9A Gene Causes Insensitivity to Pain and Autonomic Dysregulation.” *The British Journal of Dermatology* 171 (5): 1268–70.
- Belarbi, Karim, Elodie Cuvelier, Alain Destée, Bernard Gressier, and Marie-Christine Chartier-Harlin. 2017. “NADPH Oxidases in Parkinson’s Disease: A Systematic Review.” *Molecular Neurodegeneration* 12 (1): 84.
- Bendor, Jacob T., Todd P. Logan, and Robert H. Edwards. 2013. “The Function of α -Synuclein.” *Neuron* 79 (6): 1044–66.
- Benitez, Bruno A., Albert A. Davis, Sheng Chih Jin, Laura Ibanez, Sara Ortega-Cubero, Pau Pastor, Jiyeon Choi, Breanna Cooper, Joel S. Perlmutter, and Carlos Cruchaga. 2016. “Resequencing Analysis of Five Mendelian Genes and the Top Genes from Genome-Wide Association Studies in Parkinson’s Disease.” *Molecular*

Neurodegeneration 11 (April): 29.

Berg, Maarten P. van den, Rowida Almomani, Italo Biaggioni, Martijn van Faassen, Pim van der Harst, Herman H. W. Silljé, Irene Mateo Leach, et al. 2018. "Mutations in Causing a Novel Orthostatic Hypotension Syndrome." *Circulation Research* 122 (6): 846–54.

Berman, S. B., and T. G. Hastings. 1997. "Inhibition of Glutamate Transport in Synaptosomes by Dopamine Oxidation and Reactive Oxygen Species." *Journal of Neurochemistry* 69 (3): 1185–95.

Bocti, Christian, François Pépin, Maxime Tétreault, Philippe Cossette, Francis Langlois, Hélène Imbeault, Nadia Duval, Guy Lacombe, and Tamas Fulop. 2017. "Orthostatic Hypotension Associated with Executive Dysfunction in Mild Cognitive Impairment." *Journal of the Neurological Sciences* 382 (November): 79–83.

Bonet-Ponce, Luis, and Mark R. Cookson. 2021. "LRRK2 Recruitment, Activity, and Function in Organelles." *The FEBS Journal*, July. <https://doi.org/10.1111/febs.16099>.

Botella Romero, Francisco, Elisa Martín Montañez, Eugenio Jiménez Gutiérrez, and José Pavía Molina. 2009. "Caracterización de receptores muscarínicos en células indiferenciadas tiroideas de ratas Fisher." *Endocrinología y nutrición: organo de la Sociedad Espanola de Endocrinología y Nutricion* 56 (3): 106–11.

Brockmann, K., K. Srulijes, A. K. Hauser, C. Schulte, I. Csoti, T. Gasser, and D. Berg. 2011. "GBA-Associated PD Presents with Nonmotor Characteristics." *Neurology* 77 (3): 276–80.

Burré, Jacqueline, Manu Sharma, and Thomas C. Südhof. 2015. "Definition of a Molecular Pathway Mediating α -Synuclein Neurotoxicity." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 35 (13): 5221–32.

Buser, Dominik P., Marie-Françoise Ritz, Suzette Moes, Cristobal Tostado, Stephan Frank, Martin Spiess, Luigi Mariani, Paul Jenö, Jean-Louis Boulay, and Gregor Hutter. 2019. "Quantitative Proteomics Reveals Reduction of Endocytic Machinery Components in Gliomas." *EBioMedicine* 46 (August): 32–41.

Calses, Philamer C., Kiranjit K. Dhillon, Nyka Tucker, Yong Chi, Jen-Wei Huang, Masaoki Kawasumi, Paul Nghiem, et al. 2017. "DGCR8 Mediates Repair of UV-Induced DNA Damage Independently of RNA Processing." *Cell Reports* 19 (1): 162–74.

Cang, Chunlei, Kimberly Aranda, Young-Jun Seo, Bruno Gasnier, and Dejian Ren. 2015. "TMEM175 Is an Organelle K(+) Channel Regulating Lysosomal Function." *Cell* 162 (5): 1101–12.

Cao, Xiaowen, Li Xing, Hua He, and Xuekui Zhang. 2020. "Views on GWAS Statistical Analysis." *Bioinformatics* 16 (5): 393–97.

Castle, Christopher D., Erica K. Cassimere, and Catherine Denicourt. 2012. "LAS1L Interacts with the Mammalian Rix1 Complex to Regulate Ribosome Biogenesis." *Molecular Biology of the Cell* 23 (4): 716–28.

Castle, Christopher D., Erica K. Cassimere, Jinho Lee, and Catherine Denicourt. 2010. "Las1L Is a Nucleolar Protein Required for Cell Proliferation and Ribosome Biogenesis."

Molecular and Cellular Biology 30 (18): 4404–14.

- Chelala, Claude, Arshad Khan, and Nicholas R. Lemoine. 2009. "SNPNexus: A Web Database for Functional Annotation of Newly Discovered and Public Domain Single Nucleotide Polymorphisms." *Bioinformatics* 25 (5): 655–61.
- Chelban, Viorica, Ekawat Vichayanrat, Lucia Schottlaende, Valeria Iodice, and Henry Houlden. 2018. "Autonomic Dysfunction in Genetic Forms of Synucleinopathies." *Movement Disorders: Official Journal of the Movement Disorder Society* 33 (3): 359–71.
- Chen, Honglei, Shumin M. Zhang, Miguel A. Hernán, Walter C. Willett, and Alberto Ascherio. 2002. "Diet and Parkinson's Disease: A Potential Role of Dairy Products in Men." *Annals of Neurology* 52 (6): 793–801.
- Chen, Wei-Wei, Xia Zhang, and Wen-Juan Huang. 2016. "Role of Neuroinflammation in Neurodegenerative Diseases (Review)." *Molecular Medicine Reports* 13 (4): 3391–96.
- Chen, Yong Ping, Wei Song, Rui Huang, Ke Chen, Bi Zhao, Jianpeng Li, Yuan Yang, and Hui-Fang Shang. 2013. "GAK rs1564282 and DGKQ rs11248060 Increase the Risk for Parkinson's Disease in a Chinese Population." *Journal of Clinical Neuroscience: Official Journal of the Neurosurgical Society of Australasia* 20 (6): 880–83.
- Chinta, Shankar J., Jyothi K. Mallajosyula, Anand Rane, and Julie K. Andersen. 2010. "Mitochondrial α -Synuclein Accumulation Impairs Complex I Function in Dopaminergic Neurons and Results in Increased Mitophagy in Vivo." *Neuroscience Letters* 486 (3): 235–39.
- Choudhry, Hassaam, Meha Aggarwal, and Ping-Yue Pan. 2021. "Mini-Review: Synaptojanin 1 and Its Implications in Membrane Trafficking." *Neuroscience Letters* 765 (November): 136288.
- Clark, Christine M., Kevin D. Monahan, and Rachel C. Drew. 2018. "Omega-3 Polyunsaturated Fatty Acid Supplementation Reduces Blood Pressure but Not Renal Vasoconstrictor Response to Orthostatic Stress in Healthy Older Adults." *Physiological Reports* 6 (8): e13674.
- Collins, F. S., M. S. Guyer, and A. Charkravarti. 1997. "Variations on a Theme: Cataloging Human DNA Sequence Variation." *Science* 278 (5343): 1580–81.
- Convertino, Victor A. 2014. "Neurohumoral Mechanisms Associated with Orthostasis: Reaffirmation of the Significant Contribution of the Heart Rate Response." *Frontiers in Physiology* 5 (June): 236.
- Coppedè, Fabio. 2012. "Genetics and Epigenetics of Parkinson's Disease." *TheScientificWorldJournal* 2012 (May): 489830.
- Cox, James J., Frank Reimann, Adeline K. Nicholas, Gemma Thornton, Emma Roberts, Kelly Springell, Gulshan Karbani, et al. 2006. "An SCN9A Channelopathy Causes Congenital Inability to Experience Pain." *Nature* 444 (7121): 894–98.
- Craig, A. D. 2003. "Interoception: The Sense of the Physiological Condition of the Body." *Current Opinion in Neurobiology* 13 (4): 500–505.

- Cremona, O., G. Di Paolo, M. R. Wenk, A. Lüthi, W. T. Kim, K. Takei, L. Daniell, et al. 1999. "Essential Role of Phosphoinositide Metabolism in Synaptic Vesicle Recycling." *Cell* 99 (2): 179–88.
- Critchley, Hugo D. 2009. "Psychophysiology of Neural, Cognitive and Affective Integration: fMRI and Autonomic Indicators." *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology* 73 (2): 88–94.
- Cuervo, Ana Maria, Leonidas Stefanis, Ross Fredenburg, Peter T. Lansbury, and David Sulzer. 2004. "Impaired Degradation of Mutant Alpha-Synuclein by Chaperone-Mediated Autophagy." *Science* 305 (5688): 1292–95.
- Cunningham, Fiona, James E. Allen, Jamie Allen, Jorge Alvarez-Jarreta, M. Ridwan Amode, Irina M. Armean, Olanrewaju Austine-Orimoloye, et al. 2022. "Ensembl 2022." *Nucleic Acids Research* 50 (D1): D988–95.
- Czirr, Eva, and Tony Wyss-Coray. 2012. "The Immunology of Neurodegeneration." *The Journal of Clinical Investigation* 122 (4): 1156–63.
- Dampney, R. A. L., J. Horiuchi, T. Tagawa, M. A. P. Fontes, P. D. Potts, and J. W. Polson. 2003. "Medullary and Supramedullary Mechanisms Regulating Sympathetic Vasomotor Tone." *Acta Physiologica Scandinavica* 177 (3): 209–18.
- Dayem Ullah, Abu Z., Jorge Oscanoa, Jun Wang, Ai Nagano, Nicholas R. Lemoine, and Claude Chelala. 2018. "SNPnexus: Assessing the Functional Relevance of Genetic Variation to Facilitate the Promise of Precision Medicine." *Nucleic Acids Research* 46 (W1): W109–13.
- Day, Jacob Oliver, and Stephen Mullin. 2021. "The Genetics of Parkinson's Disease and Implications for Clinical Practice." *Genes* 12 (7). <https://doi.org/10.3390/genes12071006>.
- Deckwerth, T. L., R. M. Easton, C. M. Knudson, S. J. Korsmeyer, and E. M. Johnson Jr. 1998. "Placement of the BCL2 Family Member BAX in the Death Pathway of Sympathetic Neurons Activated by Trophic Factor Deprivation." *Experimental Neurology* 152 (1): 150–62.
- Deng, Hao, Peng Wang, and Joseph Jankovic. 2018. "The Genetics of Parkinson Disease." *Ageing Research Reviews* 42 (March): 72–85.
- Deng, Yanyao, Yanjin Feng, Zhicheng Lv, Jinli He, Xun Chen, Chen Wang, Mingyang Yuan, et al. 2022. "Machine Learning Models Identify Ferroptosis-Related Genes as Potential Diagnostic Biomarkers for Alzheimer's Disease." *Frontiers in Aging Neuroscience* 14 (September). <https://doi.org/10.3389/fnagi.2022.994130>.
- Dhingra, Ashutosh, Joachim Täger, Elisangela Bressan, Salvador Rodriguez-Nieto, Manmeet-Sakshi Bedi, Stefanie Bröer, Eldem Sadikoglu, et al. 2020. "Automated Production of Human Induced Pluripotent Stem Cell-Derived Cortical and Dopaminergic Neurons with Integrated Live-Cell Monitoring." *Journal of Visualized Experiments: JoVE*, no. 162 (August). <https://doi.org/10.3791/61525>.
- Dickson, Dennis W., Hiroshige Fujishiro, Anthony DelleDonne, Joshua Menke, Zeshan Ahmed, Kevin J. Klos, Keith A. Josephs, et al. 2008. "Evidence That Incidental Lewy

- Body Disease Is Pre-Symptomatic Parkinson's Disease." *Acta Neuropathologica* 115 (4): 437–44.
- DiFrancisco-Donoghue, Joanne, Ahmed Elokda, Eric M. Lamberg, Nancy Bono, and William G. Werner. 2009. "Norepinephrine and Cardiovascular Responses to Maximal Exercise in Parkinson's Disease on and off Medication." *Movement Disorders: Official Journal of the Movement Disorder Society* 24 (12): 1773–78.
- Di Maio, Roberto, Paul J. Barrett, Eric K. Hoffman, Caitlyn W. Barrett, Alevtina Zharikov, Anupom Borah, Xiaoping Hu, et al. 2016. "α-Synuclein Binds to TOM20 and Inhibits Mitochondrial Protein Import in Parkinson's Disease." *Science Translational Medicine* 8 (342): 342ra78.
- Di Maio, Roberto, Eric K. Hoffman, Emily M. Rocha, Matthew T. Keeney, Laurie H. Sanders, Briana R. De Miranda, Alevtina Zharikov, et al. 2018. "LRRK2 Activation in Idiopathic Parkinson's Disease." *Science Translational Medicine* 10 (451).
<https://doi.org/10.1126/scitranslmed.aar5429>.
- Dragunow, Mike. 2004. "CREB and Neurodegeneration." *Frontiers in Bioscience: A Journal and Virtual Library* 9 (January): 100–103.
- Dröge, Wulf. 2002. "Free Radicals in the Physiological Control of Cell Function." *Physiological Reviews* 82 (1): 47–95.
- Duffy, David L. 2017. "Analysis of Quantitative Trait Loci." *Methods in Molecular Biology* 1526: 191–203.
- Du, Xiao-Yu, Xi-Xiu Xie, and Rui-Tian Liu. 2020. "The Role of α-Synuclein Oligomers in Parkinson's Disease." *International Journal of Molecular Sciences* 21 (22).
<https://doi.org/10.3390/ijms21228645>.
- Eberhardt, Mirjam, Julika Nakajima, Alexandra B. Klinger, Cristian Neacsu, Kathrin Hühne, Andrias O. O'Reilly, Andreas M. Kist, et al. 2014. "Inherited Pain: Sodium Channel Nav1.7 A1632T Mutation Causes Erythromelalgia due to a Shift of Fast Inactivation." *The Journal of Biological Chemistry* 289 (4): 1971–80.
- Egan, B. M., S. Julius, C. Cottier, K. J. Osterziel, and H. Ibsen. 1983. "Role of Cardiovascular Receptors on the Neural Regulation of Renin Release in Normal Men." *Hypertension* 5 (5): 779–86.
- ENCODE Project Consortium. 2012. "An Integrated Encyclopedia of DNA Elements in the Human Genome." *Nature* 489 (7414): 57–74.
- Eriksson, Anna-Karin, Sofia Löfving, Russell C. Callaghan, and Peter Allebeck. 2013. "Alcohol Use Disorders and Risk of Parkinson's Disease: Findings from a Swedish National Cohort Study 1972-2008." *BMC Neurology* 13 (December): 190.
- Espa, Elena, Erik K. H. Clemensson, Kelvin C. Luk, Andreas Heuer, Tomas Björklund, and M. Angela Cenci. 2019. "Seeding of Protein Aggregation Causes Cognitive Impairment in Rat Model of Cortical Synucleinopathy." *Movement Disorders: Official Journal of the Movement Disorder Society* 34 (11): 1699–1710.
- Falkenburger, Björn H., Jill B. Jensen, Eamonn J. Dickson, Byung-Chang Suh, and Bertil

- Hille. 2010. "Phosphoinositides: Lipid Regulators of Membrane Proteins." *The Journal of Physiology* 588 (Pt 17): 3179–85.
- Fasano, Dominga, Silvia Parisi, Giovanna Maria Pierantoni, Anna De Rosa, Marina Picillo, Giuseppina Amodio, Maria Teresa Pellecchia, et al. 2018. "Alteration of Endosomal Trafficking Is Associated with Early-Onset Parkinsonism Caused by SYNJ1 Mutations." *Cell Death & Disease* 9 (3): 385.
- Fearnley, J. M., and A. J. Lees. 1991. "Ageing and Parkinson's Disease: Substantia Nigra Regional Selectivity." *Brain: A Journal of Neurology* 114 (Pt 5) (October): 2283–2301.
- Fertleman, Caroline R., Mark D. Baker, Keith A. Parker, Sarah Moffatt, Frances V. Elmslie, Bjarke Abrahamsen, Johan Ostman, et al. 2006. "SCN9A Mutations in Paroxysmal Extreme Pain Disorder: Allelic Variants Underlie Distinct Channel Defects and Phenotypes." *Neuron* 52 (5): 767–74.
- Feuk, Lars, Andrew R. Carson, and Stephen W. Scherer. 2006. "Structural Variation in the Human Genome." *Nature Reviews. Genetics* 7 (2): 85–97.
- Freeman, Roy, Wouter Wieling, Felicia B. Axelrod, David G. Benditt, Eduardo Benarroch, Italo Biaggioni, William P. Cheshire, et al. 2011. "Consensus Statement on the Definition of Orthostatic Hypotension, Neurally Mediated Syncope and the Postural Tachycardia Syndrome." *Clinical Autonomic Research: Official Journal of the Clinical Autonomic Research Society* 21 (2): 69–72.
- Fuchs, J., C. Nilsson, J. Kachergus, M. Munz, E-M Larsson, B. Schüle, J. W. Langston, et al. 2007. "Phenotypic Variation in a Large Swedish Pedigree due to SNCA Duplication and Triplication." *Neurology* 68 (12): 916–22.
- Funayama, Manabu, Kenya Nishioka, Yuanzhe Li, and Nobutaka Hattori. 2023. "Molecular Genetics of Parkinson's Disease: Contributions and Global Trends." *Journal of Human Genetics* 68 (3): 125–30.
- Gao, Huiling, Hehong Sun, Nan Yan, Pu Zhao, He Xu, Wei Zheng, Xiaoyu Zhang, Tao Wang, Chuang Guo, and Manli Zhong. 2022. "ATP13A2 Declines Zinc-Induced Accumulation of α -Synuclein in a Parkinson's Disease Model." *International Journal of Molecular Sciences* 23 (14). <https://doi.org/10.3390/ijms23148035>.
- Gao, Yuyuan, Hongmei Tang, Kun Nie, Ruiming Zhu, Liang Gao, Shujun Feng, Limin Wang, et al. 2019. "Hippocampal Damage and White Matter Lesions Contribute to Cognitive Impairment in MPTP-Lesioned Mice with Chronic Cerebral Hypoperfusion." *Behavioural Brain Research* 368 (August): 111885.
- Garland, Emily M., and Italo Biaggioni. 2003. "Dopamine Beta-Hydroxylase Deficiency." In *GeneReviews®*, edited by Margaret P. Adam, David B. Everman, Ghayda M. Mirzaa, Roberta A. Pagon, Stephanie E. Wallace, Lora J. H. Bean, Karen W. Gripp, and Anne Amemiya. Seattle (WA): University of Washington, Seattle.
- Genetica - 7. Ed. 2002. McGraw-Hill.
- Ge, Steven Xijin, Dongmin Jung, and Runan Yao. 2020. "ShinyGO: A Graphical Gene-Set Enrichment Tool for Animals and Plants." *Bioinformatics* 36 (8): 2628–29.
- Ge, Steven Xijin, Eun Wo Son, and Runan Yao. 2018. "iDEP: An Integrated Web Application

for Differential Expression and Pathway Analysis of RNA-Seq Data.” *BMC Bioinformatics* 19 (1): 534.

Gibbons, Christopher H., and Roy Freeman. 2005. “Orthostatic Dyspnea: A Neglected Symptom of Orthostatic Hypotension.” *Clinical Autonomic Research: Official Journal of the Clinical Autonomic Research Society* 15 (1): 40–44.

———. 2015. “Clinical Implications of Delayed Orthostatic Hypotension: A 10-Year Follow-up Study.” *Neurology* 85 (16): 1362–67.

Ginty, David D., and Rosalind A. Segal. 2002. “Retrograde Neurotrophin Signaling: Trk-Ing along the Axon.” *Current Opinion in Neurobiology* 12 (3): 268–74.

Golbe, Lawrence I., Giuseppe Di Iorio, Katerina Markopoulou, Aglaia Athanassiadou, Spiridon Papapetropoulos, Ray L. Watts, Jeffery M. Vance, et al. 2007. “Glutathione S-Transferase Polymorphisms and Onset Age in Alpha-Synuclein A53T Mutant Parkinson’s Disease.” *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics* 144B (2): 254–58.

Goldschmidt, Hana L., Becky Tu-Sekine, Lenora Volk, Victor Anggono, Richard L. Huganir, and Daniel M. Raben. 2016. “DGK θ Catalytic Activity Is Required for Efficient Recycling of Presynaptic Vesicles at Excitatory Synapses.” *Cell Reports* 14 (2): 200–207.

Goldshmit, Yona, Christopher J. Greenhalgh, and Ann M. Turnley. 2004. “Suppressor of Cytokine Signalling-2 and Epidermal Growth Factor Regulate Neurite Outgrowth of Cortical Neurons.” *The European Journal of Neuroscience* 20 (9): 2260–66.

Goldstein, D. S., C. Holmes, S. T. Li, S. Bruce, L. V. Metman, and R. O. Cannon 3rd. 2000. “Cardiac Sympathetic Denervation in Parkinson Disease.” *Annals of Internal Medicine* 133 (5): 338–47.

Goldstein, D. S., S. T. Li, and I. J. Kopin. 2001. “Sympathetic Neurocirculatory Failure in Parkinson Disease: Evidence for an Etiologic Role of Alpha-Synuclein.” *Annals of Internal Medicine* 135 (11): 1010–11.

Gómez-Benito, Mónica, Noelia Granado, Patricia García-Sanz, Anne Michel, Mireille Dumoulin, and Rosario Moratalla. 2020. “Modeling Parkinson’s Disease With the Alpha-Synuclein Protein.” *Frontiers in Pharmacology* 11 (April): 356.

Gu, Xing-Long, Cai-Xia Long, Lixin Sun, Chengsong Xie, Xian Lin, and Huaibin Cai. 2010. “Astrocytic Expression of Parkinson’s Disease-Related A53T Alpha-Synuclein Causes Neurodegeneration in Mice.” *Molecular Brain* 3 (April): 12.

Guyenet, Patrice G. 2006. “The Sympathetic Control of Blood Pressure.” *Nature Reviews. Neuroscience* 7 (5): 335–46.

Hagan, R. D., F. J. Diaz, and S. M. Horvath. 1978. “Plasma Volume Changes with Movement to Supine and Standing Positions.” *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology* 45 (3): 414–17.

Han, Chongyang, Janneke G. J. Hoeijmakers, Shujun Liu, Monique M. Gerrits, Rene H. M. te Morsche, Giuseppe Lauria, Sulayman D. Dib-Hajj, et al. 2012. “Functional Profiles of

SCN9A Variants in Dorsal Root Ganglion Neurons and Superior Cervical Ganglion Neurons Correlate with Autonomic Symptoms in Small Fibre Neuropathy." *Brain: A Journal of Neurology* 135 (Pt 9): 2613–28.

Hayes, Ben. 2013. "Overview of Statistical Methods for Genome-Wide Association Studies (GWAS)." *Methods in Molecular Biology* 1019: 149–69.

Heron, Sarah E., Kathryn M. Crossland, Eva Andermann, Hilary A. Phillips, Allison J. Hall, Andrew Bleasel, Michael Shevell, et al. 2002. "Sodium-Channel Defects in Benign Familial Neonatal-Infantile Seizures." *The Lancet* 360 (9336): 851–52.

Houssa, B., D. Schaap, J. van der Wal, K. Goto, H. Kondo, A. Yamakawa, M. Shibata, T. Takenawa, and W. J. van Blitterswijk. 1997. "Cloning of a Novel Human Diacylglycerol Kinase (DGKtheta) Containing Three Cysteine-Rich Domains, a Proline-Rich Region, and a Pleckstrin Homology Domain with an Overlapping Ras-Associating Domain." *The Journal of Biological Chemistry* 272 (16): 10422–28.

Huang, Vera, Robert F. Place, Victoria Portnoy, Ji Wang, Zhongxia Qi, Zhejun Jia, Angela Yu, Marc Shuman, Jingwei Yu, and Long-Cheng Li. 2012. "Upregulation of Cyclin B1 by miRNA and Its Implications in Cancer." *Nucleic Acids Research* 40 (4): 1695–1707.

Ichinose, Masashi, and Takeshi Nishiyasu. 2012. "Arterial Baroreflex Control of Muscle Sympathetic Nerve Activity under Orthostatic Stress in Humans." *Frontiers in Physiology* 3 (August): 314.

Idiaquez, Juan Francisco, Juan Idiaquez, Juan Carlos Casar, and Italo Biaggioni. 2021. "Neurogenic Orthostatic Hypotension. Lessons From Synucleinopathies." *American Journal of Hypertension* 34 (2): 125–33.

Isonaka, Risa, Avi Z. Rosenberg, Patti Sullivan, Abraham Corrales, Courtney Holmes, Yehonatan Sharabi, and David S. Goldstein. 2019. "Alpha-Synuclein Deposition Within Sympathetic Noradrenergic Neurons Is Associated With Myocardial Noradrenergic Deficiency in Neurogenic Orthostatic Hypotension." *Hypertension* 73 (4): 910–18.

Ito, Keisuke, and Hideki Enomoto. 2016. "Retrograde Transport of Neurotrophic Factor Signaling: Implications in Neuronal Development and Pathogenesis." *Journal of Biochemistry* 160 (2): 77–85.

Jacob, G., A. C. Ertl, J. R. Shannon, R. Furlan, R. M. Robertson, and D. Robertson. 1998. "Effect of Standing on Neurohumoral Responses and Plasma Volume in Healthy Subjects." *Journal of Applied Physiology* 84 (3): 914–21.

Jacobsen, T. N., B. J. Morgan, U. Scherrer, S. F. Vissing, R. A. Lange, N. Johnson, W. S. Ring, P. S. Rahko, P. Hanson, and R. G. Victor. 1993. "Relative Contributions of Cardiopulmonary and Sinoaortic Baroreflexes in Causing Sympathetic Activation in the Human Skeletal Muscle Circulation during Orthostatic Stress." *Circulation Research* 73 (2): 367–78.

Jagust, W. J., B. R. Reed, E. M. Martin, J. L. Eberling, and R. A. Nelson-Abbott. 1992. "Cognitive Function and Regional Cerebral Blood Flow in Parkinson's Disease." *Brain: A Journal of Neurology* 115 (Pt 2) (April): 521–37.

Jankovic, J. 2008. "Parkinson's Disease: Clinical Features and Diagnosis." *Journal of*

Neurology, Neurosurgery, and Psychiatry 79 (4): 368–76.

Jankovic, Joseph, and Eng King Tan. 2020. "Parkinson's Disease: Etiopathogenesis and Treatment." *Journal of Neurology, Neurosurgery, and Psychiatry* 91 (8): 795–808.

Jeong, Ga Ram, and Byoung Dae Lee. 2020. "Pathological Functions of LRRK2 in Parkinson's Disease." *Cells* 9 (12). <https://doi.org/10.3390/cells9122565>.

Jia, Zhenyu, and Shizhong Xu. 2005. "Clustering Expressed Genes on the Basis of Their Association with a Quantitative Phenotype." *Genetical Research* 86 (3): 193–207.

Jinn, Sarah, Robert E. Drolet, Paige E. Cramer, Andus Hon-Kit Wong, Dawn M. Toolan, Cheryl A. Gretzula, Bhavya Voleti, et al. 2017. "TMEM175 Deficiency Impairs Lysosomal and Mitochondrial Function and Increases α -Synuclein Aggregation." *Proceedings of the National Academy of Sciences of the United States of America* 114 (9): 2389–94.

Jo, Euijung, Nola Fuller, R. Peter Rand, Peter St George-Hyslop, and Paul E. Fraser. 2002. "Defective Membrane Interactions of Familial Parkinson's Disease Mutant A30P Alpha-Synuclein." *Journal of Molecular Biology* 315 (4): 799–807.

Joseph, Adrien, Ruben Wanono, Martin Flamant, and Emmanuelle Vidal-Petiot. 2017. "Orthostatic Hypotension: A Review." *Nephrologie & Therapeutique* 13 Suppl 1 (April): S55–67.

Junn, Eunsung, and M. Maral Mouradian. 2002. "Human α -Synuclein over-Expression Increases Intracellular Reactive Oxygen Species Levels and Susceptibility to Dopamine." *Neuroscience Letters* 320 (3): 146–50.

Kalra, Dinesh K., Anvi Raina, and Sumit Sohal. 2020. "Neurogenic Orthostatic Hypotension: State of the Art and Therapeutic Strategies." *Clinical Medicine Insights. Cardiology* 14 (August): 1179546820953415.

Kang, Sung Hoon, Su Jin Chung, Jungyeun Lee, and Seong-Beom Koh. 2022. "Independent Effect of Neurogenic Orthostatic Hypotension on Mild Cognitive Impairment in Parkinson's Disease." *Clinical Autonomic Research: Official Journal of the Clinical Autonomic Research Society* 32 (1): 43–50.

Karemaker, John M. 2017. "An Introduction into Autonomic Nervous Function." *Physiological Measurement* 38 (5): R89–118.

Kasowski, Maya, Fabian Grubert, Christopher Heffelfinger, Manoj Hariharan, Akwasi Asabere, Sebastian M. Waszak, Lukas Habegger, et al. 2010. "Variation in Transcription Factor Binding among Humans." *Science* 328 (5975): 232–35.

Kaur, Manpreet, Dinu S. Chandran, Ashok Kumar Jaryal, Dipankar Bhowmik, Sanjay Kumar Agarwal, and Kishore Kumar Deepak. 2016. "Baroreflex Dysfunction in Chronic Kidney Disease." *World Journal of Nephrology* 5 (1): 53–65.

Keller, Margaux F., Mohamad Saad, Jose Bras, Francesco Bettella, Nayia Nicolaou, Javier Simón-Sánchez, Florian Mittag, et al. 2012. "Using Genome-Wide Complex Trait Analysis to Quantify 'Missing Heritability' in Parkinson's Disease." *Human Molecular Genetics* 21 (22): 4996–5009.

- Kenney, M. J., and C. K. Ganta. 2014. "Autonomic Nervous System and Immune System Interactions." *Comprehensive Physiology* 4 (3): 1177–1200.
- Khan, Asmat Ullah, Muhammad Akram, Muhammad Daniyal, and Rida Zainab. 2019. "Awareness and Current Knowledge of Parkinson's Disease: A Neurodegenerative Disorder." *The International Journal of Neuroscience* 129 (1): 55–93.
- Kim, Changyoun, Dong-Hwan Ho, Ji-Eun Suk, Sungyong You, Sarah Michael, Junghee Kang, Sung Joong Lee, et al. 2013. "Neuron-Released Oligomeric α -Synuclein Is an Endogenous Agonist of TLR2 for Paracrine Activation of Microglia." *Nature Communications* 4: 1562.
- Kimmerly, Derek S., Deborah D. O'Leary, Ravi S. Menon, Joseph S. Gati, and J. Kevin Shoemaker. 2005. "Cortical Regions Associated with Autonomic Cardiovascular Regulation during Lower Body Negative Pressure in Humans." *The Journal of Physiology* 569 (Pt 1): 331–45.
- Kim, Min-Kyeong, Ji-Won Yuk, Hyang-Sook Kim, Ki-Jong Park, and Dae-Seong Kim. 2013. "Autonomic Dysfunction in SCN9A-Associated Primary Erythromelalgia." *Clinical Autonomic Research: Official Journal of the Clinical Autonomic Research Society* 23 (2): 105–7.
- Kiral, Ferdi Ridvan, Friederike Elisabeth Kohrs, Eugene Jennifer Jin, and Peter Robin Hiesinger. 2018. "Rab GTPases and Membrane Trafficking in Neurodegeneration." *Current Biology: CB* 28 (8): R471–86.
- Koldobskiy, Michael A., Anutosh Chakraborty, J. Kent Werner Jr, Adele M. Snowman, Krishna R. Juluri, M. Scott Vandiver, Seyun Kim, Shira Heletz, and Solomon H. Snyder. 2010. "p53-Mediated Apoptosis Requires Inositol Hexakisphosphate Kinase-2." *Proceedings of the National Academy of Sciences of the United States of America* 107 (49): 20947–51.
- Kriks, Sonja, Jae-Won Shim, Jinghua Piao, Yosif M. Ganat, Dustin R. Wakeman, Zhong Xie, Luis Carrillo-Reid, et al. 2011. "Dopamine Neurons Derived from Human ES Cells Efficiently Engraft in Animal Models of Parkinson's Disease." *Nature* 480 (7378): 547–51.
- Krüger, R., W. Kuhn, T. Müller, D. Voitalla, M. Graeber, S. Kösel, H. Przuntek, J. T. Epplen, L. Schöls, and O. Riess. 1998. "Ala30Pro Mutation in the Gene Encoding Alpha-Synuclein in Parkinson's Disease." *Nature Genetics* 18 (2): 106–8.
- Kumar, Shivam, Lav Goyal, and Shamsher Singh. 2022a. "Tremor and Rigidity in Patients with Parkinson's Disease: Emphasis on Epidemiology, Pathophysiology and Contributing Factors." *CNS & Neurological Disorders Drug Targets* 21 (7): 596–609.
- . 2022b. "Tremor and Rigidity in Patients with Parkinson's Disease: Emphasis on Epidemiology, Pathophysiology and Contributing Factors." *CNS & Neurological Disorders Drug Targets* 21 (7): 596–609.
- Kwon, Hyuk Sung, and Seong-Ho Koh. 2020. "Neuroinflammation in Neurodegenerative Disorders: The Roles of Microglia and Astrocytes." *Translational Neurodegeneration* 9 (1): 42.

- Lahrmann, H., P. Cortelli, M. Hilz, C. J. Mathias, W. Struhal, and M. Tassinari. 2006. "EFNS Guidelines on the Diagnosis and Management of Orthostatic Hypotension." *European Journal of Neurology: The Official Journal of the European Federation of Neurological Societies* 13 (9): 930–36.
- Lake, Jonathan I., and Robert O. Heuckeroth. 2013. "Enteric Nervous System Development: Migration, Differentiation, and Disease." *American Journal of Physiology. Gastrointestinal and Liver Physiology* 305 (1): G1–24.
- Lang, Anthony E., and Alberto J. Espay. 2018. "Disease Modification in Parkinson's Disease: Current Approaches, Challenges, and Future Considerations." *Movement Disorders: Official Journal of the Movement Disorder Society* 33 (5): 660–77.
- Langley, J. N. 1901. "Observations on the Physiological Action of Extracts of the Supra-Renal Bodies." *The Journal of Physiology* 27 (3): 237–56.
- Langston, J. W., P. Ballard, J. W. Tetrad, and I. Irwin. 1983. "Chronic Parkinsonism in Humans due to a Product of Meperidine-Analog Synthesis." *Science* 219 (4587): 979–80.
- Lee, Byoung Dae, Joo-Ho Shin, Jackalina VanKampen, Leonard Petrucelli, Andrew B. West, Han Seok Ko, Yun-Il Lee, et al. 2010. "Inhibitors of Leucine-Rich Repeat Kinase-2 Protect against Models of Parkinson's Disease." *Nature Medicine* 16 (9): 998–1000.
- Lee, Jee Hoon, Ji-Hye Han, Hyunmi Kim, Sang Myun Park, Eun-Hye Joe, and Ilo Jou. 2019. "Parkinson's Disease-Associated LRRK2-G2019S Mutant Acts through Regulation of SERCA Activity to Control ER Stress in Astrocytes." *Acta Neuropathologica Communications* 7 (1): 68.
- Lee, Soyoun, Jie Yang, Jiayu Huang, Hao Chen, Wei Hou, and Song Wu. 2017. "Multi-Marker Linkage Disequilibrium Mapping of Quantitative Trait Loci." *Briefings in Bioinformatics* 18 (2): 195–204.
- Li, Gongcheng, Tiejun Pan, Dan Guo, and Long-Cheng Li. 2014. "Regulatory Variants and Disease: The E-Cadherin -160C/A SNP as an Example." *Molecular Biology International* 2014 (September): 967565.
- Li, Hongyu, Ahrom Ham, Thong Chi Ma, Sheng-Han Kuo, Ellen Kanter, Donghoon Kim, Han Seok Ko, et al. 2019. "Mitochondrial Dysfunction and Mitophagy Defect Triggered by Heterozygous GBA Mutations." *Autophagy* 15 (1): 113–30.
- Lill, Christina M. 2016. "Genetics of Parkinson's Disease." *Molecular and Cellular Probes* 30 (6): 386–96.
- Liu, Ganqiang, Brendon Boot, Joseph J. Locascio, Iris E. Jansen, Sophie Winder-Rhodes, Shirley Eberly, Alexis Elbaz, et al. 2016. "Specifically Neuropathic Gaucher's Mutations Accelerate Cognitive Decline in Parkinson's." *Annals of Neurology* 80 (5): 674–85.
- Liu, Han-Ming, Jin-Ping Zheng, Dan Yang, Zhao-Fa Liu, Zi Li, Zhen-Zhen Hu, and Ze-Nan Li. 2021. "Recessive/dominant Model: Alternative Choice in Case-Control-Based Genome-Wide Association Studies." *PLoS One* 16 (7): e0254947.
- Liu, Yanying, Lanhai Lü, Casey L. Hettinger, Gaofeng Dong, Dong Zhang, Khosrow Rezvani,

- Xuejun Wang, and Hongmin Wang. 2014. "Ubiquilin-1 Protects Cells from Oxidative Stress and Ischemic Stroke Caused Tissue Injury in Mice." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34 (8): 2813–21.
- Loewy, A. D. 1982. "Descending Pathways to the Sympathetic Preganglionic Neurons." *Progress in Brain Research* 57: 267–77.
- Lohmueller, Kirk E., Celeste L. Pearce, Malcolm Pike, Eric S. Lander, and Joel N. Hirschhorn. 2003. "Meta-Analysis of Genetic Association Studies Supports a Contribution of Common Variants to Susceptibility to Common Disease." *Nature Genetics* 33 (2): 177–82.
- Lou, Xiang-Yang, George Casella, Rory J. Todhunter, Mark C. K. Yang, and Rongling Wu. 2005. "A General Statistical Framework for Unifying Interval and Linkage Disequilibrium Mapping." *Journal of the American Statistical Association* 100 (469): 158–71.
- Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology* 15 (12): 550.
- Low, P. A., T. L. Opfer-Gehrking, B. R. McPhee, R. D. Fealey, E. E. Benarroch, C. L. Willner, G. A. Suarez, C. J. Proper, J. A. Felten, and C. A. Huck. 1995. "Prospective Evaluation of Clinical Characteristics of Orthostatic Hypotension." *Mayo Clinic Proceedings. Mayo Clinic* 70 (7): 617–22.
- Lubos, Edith, Joseph Loscalzo, and Diane E. Handy. 2011. "Glutathione Peroxidase-1 in Health and Disease: From Molecular Mechanisms to Therapeutic Opportunities." *Antioxidants & Redox Signaling* 15 (7): 1957–97.
- Luk, Kelvin C., Victoria Kehm, Jenna Carroll, Bin Zhang, Patrick O'Brien, John Q. Trojanowski, and Virginia M-Y Lee. 2012. "Pathological α -Synuclein Transmission Initiates Parkinson-like Neurodegeneration in Nontransgenic Mice." *Science* 338 (6109): 949–53.
- Lu, Mandy, Qingyu Zhao, Kathleen L. Poston, Edith V. Sullivan, Adolf Pfefferbaum, Marian Shahid, Maya Katz, et al. 2021. "Quantifying Parkinson's Disease Motor Severity under Uncertainty Using MDS-UPDRS Videos." *Medical Image Analysis* 73 (October): 102179.
- Luo, Yunhai, Benjamin C. Hitz, Idan Gabdank, Jason A. Hilton, Meenakshi S. Kagda, Bonita Lam, Zachary Myers, et al. 2020. "New Developments on the Encyclopedia of DNA Elements (ENCODE) Data Portal." *Nucleic Acids Research* 48 (D1): D882–89.
- Luth, Eric S., Irina G. Stavrovskaya, Tim Bartels, Bruce S. Kristal, and Dennis J. Selkoe. 2014. "Soluble, Prefibrillar α -Synuclein Oligomers Promote Complex I-Dependent, Ca^{2+} -Induced Mitochondrial Dysfunction." *The Journal of Biological Chemistry* 289 (31): 21490–507.
- Machiela, Mitchell J., and Stephen J. Chanock. 2015. "LDlink: A Web-Based Application for Exploring Population-Specific Haplotype Structure and Linking Correlated Alleles of Possible Functional Variants." *Bioinformatics* 31 (21): 3555–57.
- MacMahon Copas, Adina N., Sarah F. McComish, Jean M. Fletcher, and Maeve A. Caldwell. 2021. "The Pathogenesis of Parkinson's Disease: A Complex Interplay Between

- Astrocytes, Microglia, and T Lymphocytes?" *Frontiers in Neurology* 12 (May): 666737.
- Malacarne, M., E. Gennaro, F. Madia, S. Pozzi, D. Vacca, B. Barone, B. dalla Bernardina, et al. 2001. "Benign Familial Infantile Convulsions: Mapping of a Novel Locus on Chromosome 2q24 and Evidence for Genetic Heterogeneity." *American Journal of Human Genetics* 68 (6): 1521–26.
- Mamais, Adamantios, Jillian H. Kluss, Luis Bonet-Ponce, Natalie Landeck, Rebekah G. Langston, Nathan Smith, Alexandra Beilina, et al. 2021. "Mutations in LRRK2 Linked to Parkinson Disease Sequester Rab8a to Damaged Lysosomes and Regulate Transferrin-Mediated Iron Uptake in Microglia." *PLoS Biology* 19 (12): e3001480.
- Mantamadiotis, Theo, Thomas Lemberger, Susanne C. Bleckmann, Heidrun Kern, Oliver Kretz, Ana Martin Villalba, François Tronche, et al. 2002. "Disruption of CREB Function in Brain Leads to Neurodegeneration." *Nature Genetics* 31 (1): 47–54.
- Masaki, K. H., I. J. Schatz, C. M. Burchfiel, D. S. Sharp, D. Chiu, D. Foley, and J. D. Curb. 1998. "Orthostatic Hypotension Predicts Mortality in Elderly Men: The Honolulu Heart Program." *Circulation* 98 (21): 2290–95.
- Mathias, Christopher J. 2002. "To Stand on One's Own Legs." *Clinical Medicine* 2 (3): 237–45.
- Mathias, C. J., R. Mallipeddi, and K. Bleasdale-Barr. 1999. "Symptoms Associated with Orthostatic Hypotension in Pure Autonomic Failure and Multiple System Atrophy." *Journal of Neurology* 246 (10): 893–98.
- Matunis, M. J., J. Xing, and G. Dreyfuss. 1994. "The hnRNP F Protein: Unique Primary Structure, Nucleic Acid-Binding Properties, and Subcellular Localization." *Nucleic Acids Research* 22 (6): 1059–67.
- McCarroll, Steven A., Tracy N. Hadnott, George H. Perry, Pardis C. Sabeti, Michael C. Zody, Jeffrey C. Barrett, Stephanie Dallaire, et al. 2006. "Common Deletion Polymorphisms in the Human Genome." *Nature Genetics* 38 (1): 86–92.
- McDonald, Claire, Julia L. Newton, and David J. Burn. 2016. "Orthostatic Hypotension and Cognitive Impairment in Parkinson's Disease: Causation or Association?" *Movement Disorders: Official Journal of the Movement Disorder Society* 31 (7): 937–46.
- McElheny, Victor K. 2010. *Drawing the Map of Life: Inside the Human Genome Project*. Basic Books.
- McLaren, William, Laurent Gil, Sarah E. Hunt, Harpreet Singh Riat, Graham R. S. Ritchie, Anja Thormann, Paul Flicek, and Fiona Cunningham. 2016. "The Ensembl Variant Effect Predictor." *Genome Biology* 17 (1): 122.
- Mendell, Joshua T., and Harry C. Dietz. 2001. "When the Message Goes Awry." *Cell* 107 (4): 411–14.
- Mendoza-Velásquez, José Javier, Juan Francisco Flores-Vázquez, Evalinda Barrón-Velázquez, Ana Luisa Sosa-Ortiz, Ben-Min Woo Illigens, and Timo Siepmann. 2019. "Autonomic Dysfunction in α -Synucleinopathies." *Frontiers in Neurology* 10 (April): 363.

- Migdalska-Richards, Anna, and Anthony H. V. Schapira. 2016. "The Relationship between Glucocerebrosidase Mutations and Parkinson Disease." *Journal of Neurochemistry* 139 Suppl 1 (October): 77–90.
- Minchev, Danail, Maria Kazakova, and Victoria Sarafian. 2022. "Neuroinflammation and Autophagy in Parkinson's Disease—Novel Perspectives." *International Journal of Molecular Sciences* 23 (23). <https://doi.org/10.3390/ijms232314997>.
- Mize, Travis J., and Luke M. Evans. 2022. "Examination of a Novel Expression-Based Gene-SNP Annotation Strategy to Identify Tissue-Specific Contributions to Heritability in Multiple Traits." *European Journal of Human Genetics: EJHG*, November. <https://doi.org/10.1038/s41431-022-01244-1>.
- Molineris, Ivan, Davide Schiavone, Fabio Rosa, Giuseppe Matullo, Valeria Poli, and Paolo Provero. 2013. "Identification of Functional Cis-Regulatory Polymorphisms in the Human Genome." *Human Mutation* 34 (5): 735–42.
- Montastruc, J. L., M. Pelat, P. Verwaerde, C. Brefel-Courbon, M. A. Tran, O. Blin, O. Rascol, and J. M. Senard. 1998. "Fluoxetine in Orthostatic Hypotension of Parkinson's Disease: A Clinical and Experimental Pilot Study." *Fundamental & Clinical Pharmacology* 12 (4): 398–402.
- Murdoch, John D., Christine M. Rostovsky, Sindhuja Gowrisankaran, Amandeep S. Arora, Sandra-Fausia Soukup, Ramon Vidal, Vincenzo Capece, et al. 2016. "Endophilin-A Deficiency Induces the Foxo3a-Fbxo32 Network in the Brain and Causes Dysregulation of Autophagy and the Ubiquitin-Proteasome System." *Cell Reports* 17 (4): 1071–86.
- Myasnikov, Alexander, Hanwen Zhu, Patricia Hixson, Boer Xie, Kaiwen Yu, Aaron Pitre, Junmin Peng, and Ji Sun. 2021. "Structural Analysis of the Full-Length Human LRRK2." *Cell* 184 (13): 3519–27.e10.
- Myers, Timothy A., Stephen J. Chanock, and Mitchell J. Machiela. 2020. "r: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations." *Frontiers in Genetics* 11 (February): 157.
- Nagata, Eiichiro, Hongbo R. Luo, Adolfo Saiardi, Byoung-Il Bae, Norihiro Suzuki, and Solomon H. Snyder. 2005. "Inositol Hexakisphosphate Kinase-2, a Physiologic Mediator of Cell Death." *The Journal of Biological Chemistry* 280 (2): 1634–40.
- Nagpal, Latika, Michael D. Kornberg, Lauren K. Albacarys, and Solomon H. Snyder. 2021. "Inositol Hexakisphosphate Kinase-2 Determines Cellular Energy Dynamics by Regulating Creatine Kinase-B." *Proceedings of the National Academy of Sciences of the United States of America* 118 (6). <https://doi.org/10.1073/pnas.2020695118>.
- Nalls, Mike A., Cornelis Blauwendraat, Costanza L. Vallergera, Karl Heilbron, Sara Bandres-Ciga, Diana Chang, Manuela Tan, et al. 2019. "Identification of Novel Risk Loci, Causal Insights, and Heritable Risk for Parkinson's Disease: A Meta-Analysis of Genome-Wide Association Studies." *Lancet Neurology* 18 (12): 1091–1102.
- Ng, P. C., and S. Henikoff. 2001. "Predicting Deleterious Amino Acid Substitutions." *Genome Research* 11 (5): 863–74.

- Norcliffe-Kaufmann, Lucy, Horacio Kaufmann, Jose-Alberto Palma, Cyndya A. Shibao, Italo Biaggioni, Amanda C. Peltier, Wolfgang Singer, et al. 2018. "Orthostatic Heart Rate Changes in Patients with Autonomic Failure Caused by Neurodegenerative Synucleinopathies." *Annals of Neurology* 83 (3): 522–31.
- Nuytemans, Karen, Jessie Theuns, Marc Cruts, and Christine Van Broeckhoven. 2010. "Genetic Etiology of Parkinson Disease Associated with Mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 Genes: A Mutation Update." *Human Mutation* 31 (7): 763–80.
- Ohta, H., X. Li, and W. T. Talman. 1996. "Release of Glutamate in the Nucleus Tractus Solitarius in Response to Baroreflex Activation in Rats." *Neuroscience* 74 (1): 29–37.
- Olanow, C. Warren, and Patrik Brundin. 2013. "Parkinson's Disease and Alpha Synuclein: Is Parkinson's Disease a Prion-like Disorder?" *Movement Disorders: Official Journal of the Movement Disorder Society* 28 (1): 31–40.
- Olanow, C. Warren, Matthew B. Stern, and Kapil Sethi. 2009. "The Scientific and Clinical Basis for the Treatment of Parkinson Disease (2009)." *Neurology* 72 (21 Suppl 4): S1–136.
- Orimo, Satoshi, Toshiki Uchihara, Ayako Nakamura, Fumiaki Mori, Takeshi Ikeuchi, Osamu Onodera, Masatoyo Nishizawa, et al. 2008. "Cardiac Sympathetic Denervation in Parkinson's Disease Linked to SNCA Duplication." *Acta Neuropathologica* 116 (5): 575–77.
- Orimo, Satoshi, Toshiki Uchihara, Ayako Nakamura, Fumiaki Mori, Akiyoshi Kakita, Koichi Wakabayashi, and Hitoshi Takahashi. 2008. "Axonal Alpha-Synuclein Aggregates Herald Centripetal Degeneration of Cardiac Sympathetic Nerve in Parkinson's Disease." *Brain: A Journal of Neurology* 131 (Pt 3): 642–50.
- Ortega, Roberto A., Cuiling Wang, Deborah Raymond, Nicole Bryant, Clemens R. Scherzer, Avner Thaler, Roy N. Alcalay, et al. 2021. "Association of Dual LRRK2 G2019S and GBA Variations With Parkinson Disease Progression." *JAMA Network Open* 4 (4): e215845.
- Oscanoa, Jorge, Lavanya Sivapalan, Emanuela Gadaleta, Abu Z. Dayem Ullah, Nicholas R. Lemoine, and Claude Chelala. 2020. "SNPnexus: A Web Server for Functional Annotation of Human Genome Sequence Variation (2020 Update)." *Nucleic Acids Research* 48 (W1): W185–92.
- Palma, Jose-Alberto. 2019. "Dysautonomia in the Synucleinopathies: Not Just Orthostatic Hypotension." *Clinical Autonomic Research: Official Journal of the Clinical Autonomic Research Society* 29 (6): 547–48.
- Palma, Jose-Alberto, Maria-Mar Carmona-Abellan, Noelia Barriobero, Cristina Trevino-Peinado, Martin Garcia-Lopez, Elena Fernandez-Jarne, and Maria R. Luquin. 2013. "Is Cardiac Function Impaired in Premotor Parkinson's Disease? A Retrospective Cohort Study." *Movement Disorders: Official Journal of the Movement Disorder Society* 28 (5): 591–96.
- Palma, Jose-Alberto, and Horacio Kaufmann. 2014. "Autonomic Disorders Predicting Parkinson's Disease." *Parkinsonism & Related Disorders* 20 Suppl 1 (January): S94–98.

- Palowitch, John, Andrey Shabalin, Yi-Hui Zhou, Andrew B. Nobel, and Fred A. Wright. 2018. "Estimation of Cis-eQTL Effect Sizes Using a Log of Linear Model." *Biometrics* 74 (2): 616–25.
- Parisiadou, Loukia, Chengsong Xie, Hyun Jin Cho, Xian Lin, Xing-Long Gu, Cai-Xia Long, Evy Lobbestael, et al. 2009. "Phosphorylation of Ezrin/radixin/moesin Proteins by LRRK2 Promotes the Rearrangement of Actin Cytoskeleton in Neuronal Morphogenesis." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29 (44): 13971–80.
- Parkinson, James. 2002. "An Essay on the Shaking Palsy. 1817." *The Journal of Neuropsychiatry and Clinical Neurosciences* 14 (2): 223–36; discussion 222.
- Parkinson Progression Marker Initiative. 2011. "The Parkinson Progression Marker Initiative (PPMI)." *Progress in Neurobiology* 95 (4): 629–35.
- Persson, P. B. 1996. "Modulation of Cardiovascular Control Mechanisms and Their Interaction." *Physiological Reviews* 76 (1): 193–244.
- Petrovitch, Helen, G. Webster Ross, Robert D. Abbott, Wayne T. Sanderson, Dan S. Sharp, Caroline M. Tanner, Kamal H. Masaki, et al. 2002. "Plantation Work and Risk of Parkinson Disease in a Population-Based Longitudinal Study." *Archives of Neurology* 59 (11): 1787–92.
- Pevsner, Jonathan. 2009. *Bioinformatics and Functional Genomics*. Wiley.
- Picca, Anna, Flora Guerra, Riccardo Calvani, Roberta Romano, Hélio José Coelho-Júnior, Cecilia Bucci, and Emanuele Marzetti. 2021. "Mitochondrial Dysfunction, Protein Misfolding and Neuroinflammation in Parkinson's Disease: Roads to Biomarker Discovery." *Biomolecules* 11 (10). <https://doi.org/10.3390/biom11101508>.
- Poda, R., P. Guaraldi, L. Solieri, G. Calandra-Buonaura, G. Marano, R. Gallassi, and P. Cortelli. 2012. "Standing Worsens Cognitive Functions in Patients with Neurogenic Orthostatic Hypotension." *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology* 33 (2): 469–73.
- Polymeropoulos, M. H., J. J. Higgins, L. I. Golbe, W. G. Johnson, S. E. Ide, G. Di Iorio, G. Sanges, et al. 1996. "Mapping of a Gene for Parkinson's Disease to Chromosome 4q21-q23." *Science* 274 (5290): 1197–99.
- Polymeropoulos, M. H., C. Lavedan, E. Leroy, S. E. Ide, A. Dehejia, A. Dutra, B. Pike, et al. 1997. "Mutation in the Alpha-Synuclein Gene Identified in Families with Parkinson's Disease." *Science* 276 (5321): 2045–47.
- Poon, I. O., and U. Braun. 2005. "High Prevalence of Orthostatic Hypotension and Its Correlation with Potentially Causative Medications among Elderly Veterans." *Journal of Clinical Pharmacy and Therapeutics* 30 (2): 173–78.
- Posor, York, Wonyul Jang, and Volker Haucke. 2022. "Phosphoinositides as Membrane Organizers." *Nature Reviews. Molecular Cell Biology*, May. <https://doi.org/10.1038/s41580-022-00490-x>.

- Power, John H. T., and Peter C. Blumbergs. 2009. "Cellular Glutathione Peroxidase in Human Brain: Cellular Distribution, and Its Potential Role in the Degradation of Lewy Bodies in Parkinson's Disease and Dementia with Lewy Bodies." *Acta Neuropathologica* 117 (1): 63–73.
- Pyatha, Sarita, Haesoo Kim, Daeun Lee, and Kisok Kim. 2022. "Association between Heavy Metal Exposure and Parkinson's Disease: A Review of the Mechanisms Related to Oxidative Stress." *Antioxidants (Basel, Switzerland)* 11 (12). <https://doi.org/10.3390/antiox11122467>.
- Qu, Lili, Bingqian Lin, Wenping Zeng, Chunhong Fan, Haotian Wu, Yushu Ge, Qianqian Li, et al. 2022. "Lysosomal K Channel TMEM175 Promotes Apoptosis and Aggravates Symptoms of Parkinson's Disease." *EMBO Reports* 23 (9): e53234.
- Rafanelli, Martina, Kathleen Walsh, Mohamed H. Hamdan, and Laura Buyan-Dent. 2019. "Autonomic Dysfunction: Diagnosis and Management." *Handbook of Clinical Neurology* 167: 123–37.
- Ramensky, Vasily, Peer Bork, and Shamil Sunyaev. 2002. "Human Non-Synonymous SNPs: Server and Survey." *Nucleic Acids Research* 30 (17): 3894–3900.
- Raz, Limor, Janice Knoefel, and Kiran Bhaskar. 2016. "The Neuropathology and Cerebrovascular Mechanisms of Dementia." *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism* 36 (1): 172–86.
- Ren, Gang, Parimala Vajjhala, Janet S. Lee, Barbara Winsor, and Alan L. Munn. 2006. "The BAR Domain Proteins: Molding Membranes in Fission, Fusion, and Phagy." *Microbiology and Molecular Biology Reviews: MMBR* 70 (1): 37–120.
- Rijnboutt, S., H. M. Aerts, H. J. Geuze, J. M. Tager, and G. J. Strous. 1991. "Mannose 6-Phosphate-Independent Membrane Association of Cathepsin D, Glucocerebrosidase, and Sphingolipid-Activating Protein in HepG2 Cells." *The Journal of Biological Chemistry* 266 (8): 4862–68.
- Rinaldi, Débora E., Gerardo R. Corradi, Lucía Martínez Cuesta, Hugo P. Adamo, and Felicitas de Tezanos Pinto. 2015. "The Parkinson-Associated Human P5B-ATPase ATP13A2 Protects against the Iron-Induced Cytotoxicity." *Biochimica et Biophysica Acta* 1848 (8): 1646–55.
- Robertson, David. 2008. "The Pathophysiology and Diagnosis of Orthostatic Hypotension." *Clinical Autonomic Research: Official Journal of the Clinical Autonomic Research Society* 18 Suppl 1 (March): 2–7.
- Ross, C. A., D. A. Ruggiero, T. H. Joh, D. H. Park, and D. J. Reis. 1984. "Rostral Ventrolateral Medulla: Selective Projections to the Thoracic Autonomic Cell Column from the Region Containing C1 Adrenaline Neurons." *The Journal of Comparative Neurology* 228 (2): 168–85.
- Ross, C. A., D. A. Ruggiero, D. H. Park, T. H. Joh, A. F. Sved, J. Fernandez-Pardal, J. M. Saavedra, and D. J. Reis. 1984. "Tonic Vasomotor Control by the Rostral Ventrolateral Medulla: Effect of Electrical or Chemical Stimulation of the Area Containing C1

Adrenaline Neurons on Arterial Pressure, Heart Rate, and Plasma Catecholamines and Vasopressin." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 4 (2): 474–94.

Ruffle, James K., Harpreet Hyare, Matthew A. Howard, Adam D. Farmer, A. Vania Apkarian, Steven C. R. Williams, Qasim Aziz, and Parashkev Nachev. 2021. "The Autonomic Brain: Multi-Dimensional Generative Hierarchical Modelling of the Autonomic Connectome." *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 143 (October): 164–79.

Rugbjerg, Kathrine, Beate Ritz, Lise Korbo, Nick Martinussen, and Jørgen H. Olsen. 2008. "Risk of Parkinson's Disease after Hospital Contact for Head Injury: Population Based Case-Control Study." *BMJ* 337 (December): a2494.

Salles, P., J. M. Tirapegui, and P. Chaná-Cuevas. 2024. "Genetics of Parkinson's Disease: Dominant Forms and GBA." *Neurology Perspectives* 4 (3): 100153.

Sanchez-Guajardo, Vanesa, Christopher J. Barnum, Malú G. Tansey, and Marina Romero-Ramos. 2013. "Neuroimmunological Processes in Parkinson's Disease and Their Relation to α -Synuclein: Microglia as the Referee between Neuronal Processes and Peripheral Immunity." *ASN Neuro* 5 (2): 113–39.

Sanders, J. S., D. W. Ferguson, and A. L. Mark. 1988. "Arterial Baroreflex Control of Sympathetic Nerve Activity during Elevation of Blood Pressure in Normal Man: Dominance of Aortic Baroreflexes." *Circulation* 77 (2): 279–88.

Sanders, Stephan J., Michael T. Murtha, Abha R. Gupta, John D. Murdoch, Melanie J. Raubeson, A. Jeremy Willsey, A. Gulhan Ercan-Sencicek, et al. 2012. "De Novo Mutations Revealed by Whole-Exome Sequencing Are Strongly Associated with Autism." *Nature* 485 (7397): 237–41.

Schaub, Marc A., Alan P. Boyle, Anshul Kundaje, Serafim Batzoglou, and Michael Snyder. 2012. "Linking Disease Associations with Regulatory Information in the Human Genome." *Genome Research* 22 (9): 1748–59.

Seller, H., and M. Illert. 1969. "The Localization of the First Synapse in the Carotid Sinus Baroreceptor Reflex Pathway and Its Alteration of the Afferent Input." *Pflugers Archiv: European Journal of Physiology* 306 (1): 1–19.

Seppi, Klaus, K. Ray Chaudhuri, Miguel Coelho, Susan H. Fox, Regina Katzenschlager, Santiago Perez Lloret, Daniel Weintraub, Cristina Sampaio, and the collaborators of the Parkinson's Disease Update on Non-Motor Symptoms Study Group on behalf of the Movement Disorders Society Evidence-Based Medicine Committee. 2019. "Update on Treatments for Nonmotor Symptoms of Parkinson's Disease-an Evidence-Based Medicine Review." *Movement Disorders: Official Journal of the Movement Disorder Society* 34 (2): 180–98.

Shaw, Brett H., Dave Borrel, Kimiya Sabbaghan, Colton Kum, Yijian Yang, Stephen N. Robinovitch, and Victoria E. Claydon. 2019. "Relationships between Orthostatic Hypotension, Frailty, Falling and Mortality in Elderly Care Home Residents." *BMC Geriatrics* 19 (1): 80.

Sherry, S. T., M. H. Ward, M. Kholodov, J. Baker, L. Phan, E. M. Smigielski, and K. Sirotkin.

2001. "dbSNP: The NCBI Database of Genetic Variation." *Nucleic Acids Research* 29 (1): 308–11.
- Shibao, Cyndya A., Karen Joos, John A. Phillips 3rd, Joy Cogan, John H. Newman, Rizwan Hamid, Jens Meiler, et al. 2021. "Familial Autonomic Ganglionopathy Caused by Rare Genetic Variants." *Neurology* 97 (2): e145–55.
- Sibilia, M., J. P. Steinbach, L. Stingl, A. Aguzzi, and E. F. Wagner. 1998. "A Strain-Independent Postnatal Neurodegeneration in Mice Lacking the EGF Receptor." *The EMBO Journal* 17 (3): 719–31.
- Simon, David K., Caroline M. Tanner, and Patrik Brundin. 2020a. "Parkinson Disease Epidemiology, Pathology, Genetics, and Pathophysiology." *Clinics in Geriatric Medicine* 36 (1): 1–12.
- . 2020b. "Parkinson Disease Epidemiology, Pathology, Genetics, and Pathophysiology." *Clinics in Geriatric Medicine* 36 (1): 1–12.
- Simón-Sánchez, Javier, Jacobus J. van Hilten, Bart van de Warrenburg, Bart Post, Henk W. Berendse, Sampath Arepalli, Dena G. Hernandez, et al. 2011. "Genome-Wide Association Study Confirms Extant PD Risk Loci among the Dutch." *European Journal of Human Genetics: EJHG* 19 (6): 655–61.
- Singleton, A. B., M. Farrer, J. Johnson, A. Singleton, S. Hague, J. Kachergus, M. Hulihan, et al. 2003. "Alpha-Synuclein Locus Triplication Causes Parkinson's Disease." *Science* 302 (5646): 841.
- Singleton, Amanda, Katrina Gwinn-Hardy, Yehonotan Sharabi, Sheng-Ting Li, Courtney Holmes, Raghuvver Dendi, John Hardy, Andrew Singleton, Anthony Crawley, and David S. Goldstein. 2004. "Association between Cardiac Denervation and Parkinsonism Caused by Alpha-Synuclein Gene Triplication." *Brain: A Journal of Neurology* 127 (Pt 4): 768–72.
- Sin, Mo-Kyung, and Pravin Khemani. 2020. "Neurogenic Orthostatic Hypotension: An Underrecognized Complication of Parkinson Disease." *The Journal of Neuroscience Nursing: Journal of the American Association of Neuroscience Nurses* 52 (5): 230–33.
- Smith, Wanli W., Zhong Pei, Haibing Jiang, Valina L. Dawson, Ted M. Dawson, and Christopher A. Ross. 2006. "Kinase Activity of Mutant LRRK2 Mediates Neuronal Toxicity." *Nature Neuroscience* 9 (10): 1231–33.
- Sternini, C. 1997. "Organization of the Peripheral Nervous System: Autonomic and Sensory Ganglia." *The Journal of Investigative Dermatology. Symposium Proceedings / the Society for Investigative Dermatology, Inc. [and] European Society for Dermatological Research* 2 (1): 1–7.
- Tan, Jieqiong, Tongmei Zhang, Li Jiang, Jingwei Chi, Dongshen Hu, Qian Pan, Danling Wang, and Zhuohua Zhang. 2011. "Regulation of Intracellular Manganese Homeostasis by Kufor-Rakeb Syndrome-Associated ATP13A2 Protein." *The Journal of Biological Chemistry* 286 (34): 29654–62.
- Tanner, C. M., and S. M. Goldman. 1996. "Epidemiology of Parkinson's Disease." *Neurologic Clinics* 14 (2): 317–35.

- Tansey, Malú Gámez, Rebecca L. Wallings, Madelyn C. Houser, Mary K. Herrick, Cody E. Keating, and Valerie Joers. 2022. "Inflammation and Immune Dysfunction in Parkinson Disease." *Nature Reviews. Immunology* 22 (11): 657–73.
- Taylor, J. A., J. R. Halliwill, T. E. Brown, J. Hayano, and D. L. Eckberg. 1995. "'Non-Hypotensive' Hypovolaemia Reduces Ascending Aortic Dimensions in Humans." *The Journal of Physiology* 483 (Pt 1) (February): 289–98.
- Tei, Reika, and Jeremy M. Baskin. 2020. "Spatiotemporal Control of Phosphatidic Acid Signaling with Optogenetic, Engineered Phospholipase Ds." *The Journal of Cell Biology* 219 (3). <https://doi.org/10.1083/jcb.201907013>.
- Terreberry, R. R., and E. J. Neafsey. 1987. "The Rat Medial Frontal Cortex Projects Directly to Autonomic Regions of the Brainstem." *Brain Research Bulletin* 19 (6): 639–49.
- Thacker, E. L., E. J. O'Reilly, M. G. Weisskopf, H. Chen, M. A. Schwarzschild, M. L. McCullough, E. E. Calle, M. J. Thun, and A. Ascherio. 2007. "Temporal Relationship between Cigarette Smoking and Risk of Parkinson Disease." *Neurology* 68 (10): 764–68.
- Tokutake, Takayoshi, Atsuhiko Ishikawa, Nahoko Yoshimura, Akinori Miyashita, Ryozi Kuwano, Masatoyo Nishizawa, and Takeshi Ikeuchi. 2014. "Clinical and Neuroimaging Features of Patient with Early-Onset Parkinson's Disease with Dementia Carrying SNCA p.G51D Mutation." *Parkinsonism & Related Disorders* 20 (2): 262–64.
- Trépanier, G., D. Furling, J. Puymirat, and M-E Mirault. 1996. "Immunocytochemical Localization of Seleno-Glutathione Peroxidase in the Adult Mouse Brain." *Neuroscience* 75 (1): 231–43.
- Twelves, Dominique, Kate S. M. Perkins, and Carl Counsell. 2003a. "Systematic Review of Incidence Studies of Parkinson's Disease." *Movement Disorders: Official Journal of the Movement Disorder Society* 18 (1): 19–31.
- . 2003b. "Systematic Review of Incidence Studies of Parkinson's Disease." *Movement Disorders: Official Journal of the Movement Disorder Society* 18 (1): 19–31.
- Tysnes, Ole-Bjørn, and Anette Storstein. 2017. "Epidemiology of Parkinson's Disease." *Journal of Neural Transmission* 124 (8): 901–5.
- Usenovic, Marija, Emilie Tresse, Joseph R. Mazzulli, J. Paul Taylor, and Dimitri Krainc. 2012. "Deficiency of ATP13A2 Leads to Lysosomal Dysfunction, α -Synuclein Accumulation, and Neurotoxicity." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 32 (12): 4240–46.
- Valko, Marian, Dieter Leibfritz, Jan Moncol, Mark T. D. Cronin, Milan Mazur, and Joshua Telser. 2007. "Free Radicals and Antioxidants in Normal Physiological Functions and Human Disease." *The International Journal of Biochemistry & Cell Biology* 39 (1): 44–84.
- Veen, Sarah van, Shaun Martin, Chris Van den Haute, Veronick Benoy, Joseph Lyons, Roeland Vanhoutte, Jan Pascal Kahler, et al. 2020. "ATP13A2 Deficiency Disrupts Lysosomal Polyamine Export." *Nature* 578 (7795): 419–24.

- Velseboer, Daan C., Rob J. de Haan, Wouter Wieling, David S. Goldstein, and Rob M. A. de Bie. 2011. "Prevalence of Orthostatic Hypotension in Parkinson's Disease: A Systematic Review and Meta-Analysis." *Parkinsonism & Related Disorders* 17 (10): 724–29.
- Veronese, Nicola, Marina De Rui, Francesco Bolzetta, Sabina Zambon, Maria Chiara Corti, Giovannella Baggio, Elena Debora Toffanello, et al. 2015. "Orthostatic Changes in Blood Pressure and Mortality in the Elderly: The Pro.V.A Study." *American Journal of Hypertension* 28 (10): 1248–56.
- Vielhaber, G., S. Pfeiffer, L. Brade, B. Lindner, T. Goldmann, E. Vollmer, U. Hintze, K. P. Wittern, and R. Wepf. 2001. "Localization of Ceramide and Glucosylceramide in Human Epidermis by Immunogold Electron Microscopy." *The Journal of Investigative Dermatology* 117 (5): 1126–36.
- Volpe, Thomas A., Catherine Kidner, Ira M. Hall, Grace Teng, Shiv I. S. Grewal, and Robert A. Martienssen. 2002. "Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi." *Science* 297 (5588): 1833–37.
- Wagner, Bettina, Anuradha Natarajan, Sabine Grönaug, Renate Kroismayr, Erwin F. Wagner, and Maria Sibilia. 2006. "Neuronal Survival Depends on EGFR Signaling in Cortical but Not Midbrain Astrocytes." *The EMBO Journal* 25 (4): 752–62.
- Wang, D. G., J. B. Fan, C. J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, et al. 1998. "Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome." *Science* 280 (5366): 1077–82.
- Wang, Maggie Haitian, Heather J. Cordell, and Kristel Van Steen. 2019. "Statistical Methods for Genome-Wide Association Studies." *Seminars in Cancer Biology* 55 (April): 53–60.
- Wang, Ying, Robyn Branicky, Alycia Noë, and Siegfried Hekimi. 2018. "Superoxide Dismutases: Dual Roles in Controlling ROS Damage and Regulating ROS Signaling." *The Journal of Cell Biology* 217 (6): 1915–28.
- Weiss, Howard D., and Laura Marsh. 2012. "Impulse Control Disorders and Compulsive Behaviors Associated with Dopaminergic Therapies in Parkinson Disease." *Neurology. Clinical Practice* 2 (4): 267–74.
- West, Andrew B., Darren J. Moore, Catherine Choi, Shaida A. Andrabi, Xiaojie Li, Dustin Dikeman, Saskia Biskup, et al. 2007. "Parkinson's Disease-Associated Mutations in LRRK2 Link Enhanced GTP-Binding and Kinase Activities to Neuronal Toxicity." *Human Molecular Genetics* 16 (2): 223–32.
- Westerhaus, M. J., and A. D. Loewy. 2001. "Central Representation of the Sympathetic Nervous System in the Cerebral Cortex." *Brain Research* 903 (1-2): 117–27.
- Wie, Jinhong, Zhenjiang Liu, Haikun Song, Thomas F. Tropea, Lu Yang, Huanhuan Wang, Yuling Liang, et al. 2021. "A Growth-Factor-Activated Lysosomal K Channel Regulates Parkinson's Pathology." *Nature* 591 (7850): 431–37.
- Winklhofer, Konstanze F., and Christian Haass. 2010. "Mitochondrial Dysfunction in Parkinson's Disease." *Biochimica et Biophysica Acta* 1802 (1): 29–44.

- Wu, Rongling, Chang-Xing Ma, and George Casella. 2002. "Joint Linkage and Linkage Disequilibrium Mapping of Quantitative Trait Loci in Natural Populations." *Genetics* 160 (2): 779–92.
- Yang, Hui, Lisha Zhou, Qian Shi, Yuzheng Zhao, Huaipeng Lin, Mengli Zhang, Shimin Zhao, et al. 2015. "SIRT3-Dependent GOT2 Acetylation Status Affects the Malate-Aspartate NADH Shuttle Activity and Pancreatic Tumor Growth." *The EMBO Journal* 34 (8): 1110–25.
- Yang, Jie, Wei Zhu, Jiansong Chen, Qiao Zhang, and Song Wu. 2014. "Genome-Wide Two-Marker Linkage Disequilibrium Mapping of Quantitative Trait Loci." *BMC Genetics* 15 (February): 20.
- Yeom, Kyu-Hyeon, Yoontae Lee, Jinju Han, Mi Ra Suh, and V. Narry Kim. 2006. "Characterization of DGCR8/Pasha, the Essential Cofactor for Drosha in Primary miRNA Processing." *Nucleic Acids Research* 34 (16): 4622–29.
- Yuan, Junhui, Eiji Matsuura, Yujiro Higuchi, Akihiro Hashiguchi, Tomonori Nakamura, Satoshi Nozuma, Yusuke Sakiyama, Akiko Yoshimura, Shuji Izumo, and Hiroshi Takashima. 2013. "Hereditary Sensory and Autonomic Neuropathy Type IID Caused by an SCN9A Mutation." *Neurology* 80 (18): 1641–49.
- Yun, Hye Jin, Hyejung Kim, Inhwa Ga, Hakjin Oh, Dong Hwan Ho, Jiyoung Kim, Hyemyung Seo, Ilhong Son, and Wongi Seol. 2015. "An Early Endosome Regulator, Rab5b, Is an LRRK2 Kinase Substrate." *Journal of Biochemistry* 157 (6): 485–95.
- Zanassi, P., M. Paolillo, A. Feliciello, E. V. Avvedimento, V. Gallo, and S. Schinelli. 2001. "cAMP-Dependent Protein Kinase Induces cAMP-Response Element-Binding Protein Phosphorylation via an Intracellular Calcium release/ERK-Dependent Pathway in Striatal Neurons." *The Journal of Biological Chemistry* 276 (15): 11487–95.
- Zarranz, Juan J., Javier Alegre, Juan C. Gómez-Esteban, Elena Lezcano, Raquel Ros, Israel Ampuero, Lídice Vidal, et al. 2004. "The New Mutation, E46K, of Alpha-Synuclein Causes Parkinson and Lewy Body Dementia." *Annals of Neurology* 55 (2): 164–73.
- Zhang, Dongfeng, Hong Jiang, and Junxia Xie. 2014. "Alcohol Intake and Risk of Parkinson's Disease: A Meta-Analysis of Observational Studies." *Movement Disorders: Official Journal of the Movement Disorder Society* 29 (6): 819–22.
- Zhang, Y., J. Gao, K. K. Chung, H. Huang, V. L. Dawson, and T. M. Dawson. 2000. "Parkin Functions as an E2-Dependent Ubiquitin- Protein Ligase and Promotes the Degradation of the Synaptic Vesicle-Associated Protein, CDCrel-1." *Proceedings of the National Academy of Sciences of the United States of America* 97 (24): 13354–59.
- Ziller, Michael J., Hongcang Gu, Fabian Müller, Julie Donaghey, Linus T-Y Tsai, Oliver Kohlbacher, Philip L. De Jager, et al. 2013. "Charting a Dynamic DNA Methylation Landscape of the Human Genome." *Nature* 500 (7463): 477–81.
- Zimprich, Alexander, Saskia Biskup, Petra Leitner, Peter Lichtner, Matthew Farrer, Sarah Lincoln, Jennifer Kachergus, et al. 2004. "Mutations in LRRK2 Cause Autosomal-Dominant Parkinsonism with Pleomorphic Pathology." *Neuron* 44 (4): 601–7.

Zweifel, Larry S., Rejji Kuruvilla, and David D. Ginty. 2005. "Functions and Mechanisms of Retrograde Neurotrophin Signalling." *Nature Reviews. Neuroscience* 6 (8): 615–25.

