UNIVERSITA' VITA-SALUTE SAN RAFFAELE

CORSO DI DOTTORATTO DI RICERCA INTERNAZIONALE IN MEDICINA MOLECOLARE

Curriculum in Neuroscience and Experimental Neurology

Identification of genetic variants contributing to pain-related symptoms in familial and sporadic patients affected by peripheral neuropathy

DoS: Dr. Federica Esposito   *Federica Esposito*

Second Supervisor: Dr. Uberto Pozzoli

Tesi di DOTTORATO DI RICERCA DI Kaalindi Ashok Misra
Matr. 014326
Ciclo di Dottorato XXXIV
SSD MED/03

Anno Accademico 2020/2021

Release of PhD Thesis

Il/la sottoscritto/*I*	Kaalindi Ashok Misra
Matricola / *registration number*	014326
nat_ a/ *born at*	Mumbai, India
il/*on*	02/09/1987

autore della tesi di Dottorato di ricerca dal titolo / *author of the PhD Thesis entitled*

"Identification of genetic variants contributing to pain-related symptoms in familial and

sporadic patients affected by peripheral neuropathy."

☐ AUTORIZZA la Consultazione della tesi / AUTHORIZES *the Consultation of the thesis*

☒ NON AUTORIZZA la Consultazione della tesi per ..12.. mesi /NOT AUTHORIZES *the Consultation of the thesis for* ..12.. *months*

a partire dalla data di conseguimento del titolo e precisamente / *from the PhD thesis date, specifically*
Dal / *from* ..08../..03../..2022..   Al / *to* ..08../..03../..2023..

Poiché /*because*:

☐ l'intera ricerca o parti di essa sono potenzialmente soggette a brevettabilità/ *The whole project or part of it might be subject to patentability;*
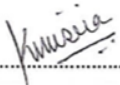
☒ ci sono parti di tesi che sono già state sottoposte a un editore o sono in attesa di pubblicazione/ *Parts of the thesis have already been submitted to a publisher or are in press;*

☐ la tesi è finanziata da enti esterni che vantano dei diritti su di esse e sulla loro pubblicazione/ *the thesis project is financed by external bodies that have rights over it and on its publication.*

E' fatto divieto di riprodurre, in tutto o in parte, quanto in essa contenuto / *It is not allowed to copy, in whole or in part, the data and the contents of the thesis*

Data /Date ....08/02/2022....   Firma /Signature ....*[signature]*....

Declaration

This thesis has been composed by myself and has not been used in any previous application for a degree. Throughout the text I use both 'I' and 'We' interchangeably.

All the results presented here were obtained by myself, except for:

1)   ***Sample recruitment:*** *performed by the Neuroalgology Unit and Skin Biopsy, Peripheral Neuropathy and Neuropathic Pain Center at the IRCCS Foundation "Carlo Besta" Neurological Institute (Milan, Italy) in agreement with the approval from the Ethics Committee of the same institute (Ethics Committee Number 65/11-04-2019).*

2)   ***DNA extraction*** *was performed by the researchers of Neuroalgology Unit and Skin Biopsy, Peripheral Neuropathy and Neuropathic Pain Center at the IRCCS Foundation "Carlo Besta" Neurological Institute (Milan, Italy).*

3)   ***Preparation of libraries*** *For PROPANE study, the libraries were processed in part by the Centre of Translation Genomics and Bioinformatics, IRCCS San Raffaele Scientific Institute (Milan, Italy). And, in part, by researchers of the Laboratory of Human Genetics of Neurological Disorders. For the PAIN-Net study, the libraries were prepared by the Centre for Advanced Studies, Research and Development in Sardinia, Italy.*

4)   ***Sample Sequencing*** *For the PAIN-Net study, all samples were sequenced by the Centre for Advanced Studies, Research and Development in Sardinia, Italy*

All sources of information are acknowledged by means of reference.

Acknowledgements

Abstract

Peripheral neuropathy (PN) is a debilitating disorder that reduces the patient's quality of life and increases healthcare costs (Cruccu & Truini, 2017). PN is caused by peripheral nerve injury; if the damage is limited to A and unmyelinated C fibres, it is referred to as Small Fiber Neuropathy (SFN) (Hovaguimian & Gibbons, 2011), and it is frequently accompanied by neuropathic pain (NP) (Van Hecke et al., 2014). In the general population, the prevalence of SFN is estimated to be around 53 per 100,000 people. However, patients are frequently underdiagnosed because regular neurologic examinations do not identify abnormalities in small fibres (Peters et al., 2013).

SFN is associated with several diseases, including diabetic neuropathy and sarcoidosis, making diagnosis more difficult. If the cause of SFN is uncertain, it is referred to as idiopathic SFN. Recent research suggests that mutations in voltage-gated sodium channels (VGSCs) genes are responsible for SFN due to their function in nociceptive transmission (Faber et al., 2012), which explains 3% of the disease pathogenesis (Devigili et al., 2008). The discovery of mutations in VGSC genes also highlighted the disease's genetic aetiology. Finding the genetic markers of PN patients based on their phenotypes could aid in resolving the pathophysiology of PN and classifying high-risk subjects.

The current study is being conducted within the international consortium PAIN-Net, which emerged from a previously European Union-funded project called the PROPANE study. Both initiatives, with slight changes, seek to better understand the genetic architecture of PN to identify high-risk painful PN patients. Consequently, the principal aim of this PhD thesis was to identify new genes and genetic variants associated with pain-related symptoms in patients affected by PN. For this purpose, whole exome sequencing (WES) was performed in twelve families with 34 affected and 19 unaffected members, as well as in 43 sporadic cases diagnosed with PN or SFN, with an early onset of clinical manifestations.

Our findings indicate that painful PN is genetically heterogeneous, with more than 40% of cases caused by genes other than VGSC genes. We detected 48 mutated genes in the familial cohort and 77 mutated genes in the sporadic cohort. Most of these genes are associated with neurotransmission rather than ion channel function. In addition to VGSC genes, the *ATP7B* gene, a copper ion transporter, had a high number of mutations in both familial and sporadic cohorts. *ATP7B* gene variants suggest that PN symptoms may be exacerbated by dysfunctional copper metabolism. Additionally, we discovered TRP gene variants, specifically *TRPM2*, that were causative for NP and conclusively segregated into two families. Pathway analysis of qualifying genes obtained from familial and sporadic cohorts revealed the nicotinic acetylcholine receptor signalling pathway as a process that appears to be altered in the disease, and that is shared between familial and sporadic subjects. This result strengthened the notion that synaptic transmission plays a role in the pathogenesis of PN. This genetic study revealed new pain-related genetic markers that may be associated with disease onset and pain modulation. For future research studies where family history or phenotypic information is suggestive of painful PN, we recommend a list of 113 genes that could be utilised to select and analyse data supplied by NGS applications. These mutated genes will be researched further to establish the pathophysiological causes of painful PN, facilitate the stratification of high-risk individuals, and assist in developing targeted therapies to alleviate pain.

Table of Contents

Acronyms and Abbreviations

PN: peripheral neuropathy

IPN: inherited peripheral neuropathy

CNS: central nervous system

DRG: Dorsal root ganglion

SFN: small fiber neuropathy

IENF: intraepidermal nerve fiber

VGSC: voltage gated sodium channel

NP: neuropathic pain

VUS: variant of uncertain significance

HMSN: hereditary motor and sensory neuropathy

CMT: Charcot-Marie-Tooth disease

HMN: hereditary motor neuropathy

HSN: hereditary sensory neuropathy

HSAN: hereditary sensory and autonomic neuropathy

NCV: nerve conduction velocity

EMG: electromyography

QST: quantitative sensory testing

QSART: Quantitative Sudomotor Axon Reflex Test

TRP: transient receptor potential family

IASP: International Association for the Study of Pain

PEPD: paroxysmal extreme pain disorder

CIP: congenital insensitivity to pain

WES: whole exome sequencing

GATK: Genome Analysis Toolkit

NGS: next generation sequencing

SNV: single nucleotide variations

Indels: insertion deletion variants

CNV: copy number variants

SV: structural variants

GnomAD: Genome Aggregation Database

WGS: whole genome sequencing

HPO: Human Phenotype Ontology database

OMIM: Online Mendelian Inheritance in Man

FDR: false discovery rate

nAChR: Nicotinic acetylcholine receptor

List of figures and tables

<u>Figures:</u>

Figure 1: Causes of SFN. This figure describes the causes of SFN into two main categories: Primary and Secondary causes of SFN. This figure has been adapted from the tables mentioned in the following articles: Themistocleous *et al*, 2014; Al-Shekhlee *et al*, 2002; Cazzato & Lauria, 2017; and Alsaloum *et al*, 2021

Figure 2: Digital confocal images of skin biopsy sections showing epidermal innervation in healthy control (A) and in a patient with diabetic SFN (B). In green: small fibre nerves marked with PGP9.5 antibody. In blue: endothelia and epidermis marked with ULEX europaeus. In red: basement membrane and blood vessels marked with collagen IV antibody. Scale bar: 100 µm. This image was retrieved from (Nolano et al, 2020).

Figure 3: The genetic markers mentioned above have been extracted from the two reviews as they were associated with NP. (Veluchamy et al, 2018; Calvo et al, 2019)

Figure 4: A pictorial representation of the transmission of a stimulus perceived at the periphery and the electric pulse is transported to CNS while interacting with different ion channels and biomolecules. It also shows the impact of nociceptive signals on different ion channels. This figure depicts three ion channels as voltage-gated sodium channels (Na$_v$s), voltage-gated calcium channels (Ca$_v$s), and hyperpolarization-activated cyclic nucleotide-gated channels (HCNs) are pro-excitatory while potassium channels (K$_v$s) and TRPM8 ion channels are not. At the same time, thrombospondins (TSP) are involved in synaptic transmission by operating through Ca$_v$ $α_2δ_1$ subunits. N-methyl-D-aspartate receptor (NMDA), α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor, and mGluR are glutamate receptors that aid in synaptic transmission at the spinal cord. DRG, dorsal root ganglion. This image is modified from Finnerup et al, 2021. (Raouf et al, 2010; Finnerup et al, 2021)

Figure 5: Schematic representation of VGSCs where on the left side, α subunit has four domains as D1-4 where each domain has six transmembrane segments (S1-6). IFM stands for isoleucine, phenylalanine, and methionine motif. The β-sub circle denotes the β-subunit interaction section. VSD is for the voltage-sensing domain. ID1-2, ID2-3, and ID3-4 are interdomain linkers. On the right side is the topology of the β-subunit (Detta *et al*, 2015)*.*

Figure 6: The stages of the mechanism of action of VGSCs adapted from (Xu *et al*, 2019)

Figure 7: Pictorial representation of the basic protocol of an NGS experiment. (Adapted from Shendure & Ji, 2008)

Figure 8: A pictorial representation of different genomic variations identified through NGS. Figure 6 (A) depicts short variations such as SNVs and INDELs. (B): It shows different types of structural variants determined by NGS technology. (C): It represents the copy number variants which is a typical structural variant and (D): It shows how many base pairs (bp) are constitute each type of genomic variation. Adapted from: 'Review: How Long-Read Sequencing Is Revealing Unseen Genomic Variation - PacBio'.

Figure 9a: Depicts first two stages of WES experiments. Adapted from (NGS Workflow Steps | Illumina sequencing workflow)

Figure 9b: Depicts last two stages of WES experiments. Adapted from (NGS Workflow Steps | Illumina sequencing workflow)

Figure 10: A schematic of GATK's best practices workflow where the joint genotyping aspect (highlighted in blue) is only used when more than 40 samples are analysed together. (Poplin et al, 2018)

Figure 11: A diagram representing the different sequence ontology terms for the transcript structure. Extracted from EMBL-EBI. (Calculated consequences)

Figure 12a: Average coverage depth per family from PROPANE and PAIN-Net study. Box plots show the minimum, 25% percentile, median, 75% percentile and maximum of average coverage depth covered

Figure 12b: Average coverage for each sample in sporadic cohort from PROPANE and PAIN-Net study. Box plots show the minimum, 25% percentile, median, 75% percentile and maximum of average coverage depth covered. Black dots denote an individual sample

Figure 12: Box plots showing the trend of target bases covered across familial (A) and sporadic (B) cohorts. Box plots show the minimum, 25% percentile, median, 75% percentile and maximum percentage of target bases covered.

Figure 13: A) Box plot showing Ti/Tv ratio per family with variant quality ranging from 2.35 to 2.55 and B) Box plot showing Ti/Tv ratio per sample in sporadic cohort by PROPANE and PAIN-Net studies with variant quality ranging from 2.37 to 2.53 where the black dot denotes each individual sample. Box plots show the minimum, 25% percentile, median, 75% percentile and maximum percentage of Ti/Tv ratio.

Figure 14: Family pedigree of PROP01 family. The proband (II.1) is denoted by a black arrow. Individuals marked with an asterisk were selected for WES experiments

Figure 15: Pedigree of the PROP02 family. The proband (III.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 16: Pedigree of the family PROP03. The proband (I.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 17: Pedigree of the family PROP04 family. The proband (II.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 19: Pedigree of the PROP05 family. The proband (II.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 20: Pedigree of the PROP06 family. The proband (II.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 21: Pedigree of the family PNET01 family. The proband (III.4) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 22: Pedigree of the PNET02 family. The proband (II.4) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 23: Pedigree of the PNET03 family. The proband (II.1) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 24: Pedigree of the PNET04 family. The proband (II.2) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 25: Pedigree of the PNET05 family. The proband (II.1) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 26: Pedigree of the PNET06 family. The proband (II.2) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

Figure 27: A pie chart of categorized 48 unique genes from familial cohort into four groups, namely, 'Ion channel', 'Neurotransmission', 'Metabolism' and 'Immune response' according to 'BRITE functional hierarchies'.

Figure 28: The count of qualifying variants at each step of variant prioritization and the count of variants for each category of classification are reported.

Figure 29: A pie chart of categorized 77 unique genes from sporadic cohort into four groups namely, 'Ion channel', 'Neurotransmission', 'Metabolism' and 'Immune response' according to 'BRITE functional hierarchies'.

Figure 30: A Venn diagram of the number of mutated genes shared amongst the familial and sporadic cohorts

Figure 31: The count of qualifying SVs at each step of variant filtration and prioritisation from sporadic cohort and the number of variants for each classification category.

Figure 32: SCN9A channel structure comprises four domains, each with six transmembrane segments. The N represents the amino side of the protein, and the C represents the carboxyl side of the protein. The fourth transmembrane segment is filled with a '+' sign in all domains to denote the positively-charged voltage-sensor of the channel. The 'P' in the loop denotes the phosphorylation site of Protein Kinase A. The location of genetic variants discovered in our study is displayed in yellow circles within the protein structure. The SCN9A mutations discovered in familial and sporadic groups are listed below the structure. Data such as ACMG classification status, family/sample identifier, and clinical classification of that family/subject were compiled. The samples from the sporadic cohort are denoted by the suffix 'PROPEO,' followed by the internal PROPANE study code and 'PNETEO' followed by the internal code for the PAIN-Net study. This picture has been modified from Swanwick et al., 2010.

Figure 33a: PROP04 and PNET04 family pedigrees, as well as a list of qualifying variants found in these families. The variant segregation within the individuals is shown in the tables below the pedigree for each family. The gene that may contribute to pain sensitivity is highlighted in bold letters. Individuals with a grey individual identifier are painless.

Figure 33b: A pictorial representation of the CACNA1E channel consisting of four domains (DI – DIV) with six segments, each numbered as S1 to S6 in domain one (DI). The domains are then linked with loops shown as I-II, II-III and III-IV. N stands for amino-terminus, and C stands for carboxyl-terminus. The location of the genetic variants from PROP04 and PNET04 is illustrated with yellow rectangles. This picture has been modified from Weiss & Zamponi, 2021

Figure 34: SCN10A channel structure comprises four domains each with six transmembrane segments. The N represents the amino-side of the protein, and the C represents the carboxyl side of the protein. The fourth transmembrane segment is filled with a '+' sign in all domains to denote the positively-charged voltage-sensor of the channel. The 'P' in the loop denotes the phosphorylation site of Protein Kinase A. The location of genetic variants discovered in our study is displayed with yellow circles within the channel structure. The SCN10A mutations discovered in familial and sporadic groups are listed below the channel structure. Data such as ACMG classification status, family/sample identifier, and clinical classification of that family/subject were compiled. The sporadic cohort samples are denoted by the suffix 'PROPEO' followed by the internal code for the PROPANE study and 'PNETEO' followed by the internal code for the PAIN-Net study. This picture has been modified from Swanwick et al., 2010.

Figure 35: Pedigree of the PROP02 family. Table with qualifying variants and their ACMG classification status displayed below the pedigree. Grey column headers denote individuals with TRPM8 mutations who have not been diagnosed with NP.

Figure 36a: Pedigrees of the PROP06 and PNET06 families. Tables with qualifying variants and their ACMG classification status are shown below each pedigree. Grey column headers denote individuals with TRPM8 mutations who have not been diagnosed with NP.

Figure 36b: The pictorial representation of the TRPA1 channel consisting of six transmembrane segments (S1-S6) where N stands for amino terminus, C stands for carboxyl terminus. This channel consists of ankyrin repeat domain and cysteine residues at the N-terminus. The location of the genetic variants observed in PROP06 and PNET06 families is displayed with yellow circles within the protein structure. This picture has been modified from Lapointe & Altier, 2011

Figure 37: A pictorial representation of the TRPM2 channel consisting of six transmembrane regions (S1-S6) where N-terminal stands for the amino end of the protein and C-terminal is the carboxyl-side of the protein. The squares on the N-terminal represent the conserved region. Circle with TRP stands for transient receptor potential motif at the C-terminal along with coiled-coil domain and Nudix hydrolase (NUDT9) substrate attached with Adenosine diphosphate ribose (ADPR) (Perraud et al., 2001). The location of the genetic variants observed in PNET01 and PNET06 families is displayed with yellow circles within the protein structure. This picture has been adapted from Nilius & Flockerzi, 2014b

Figure 38: A pictorial representation of ATP7B protein structure consisting of eight transmembrane segments (TM1-TM8) with NH2 denoting the amino-side of the protein and COOH the carboxyl side. Circle with MBD represents metal-binding domains (MBD1-MBD6) at the N-terminus. The squares represent other domains: the A-actuator domain, P-phosphorylation domain and N-nucleotide binding domain. The location of genetic variants observed in our cohort is displayed with yellow circles within the protein structure. The ATP7B mutations discovered in familial and sporadic groups are listed below the channel structure. Data such as ACMG classification status, family/sample identifier, and clinical classification of that family/subject were compiled. The sporadic cohort samples are denoted by the suffix 'PROPEO' followed by the internal code for the PROPANE study and 'PNETEO' followed by the internal code for the PAIN-Net study. The protein structure image has been modified from (Hasan et al., 2012)

Figure 39: Visual representation of the 'Nicotinic acetylcholine receptor signalling pathway'. It shows four compartments in which the pathway plays a role: 1. Nerve terminal, 2. Striated muscle, 3. Synaptic vesicle and 4. Sarcoplasmic reticulum. ACh stands for Acetylcholine, Acetyl-CoA stands for acetyl coenzyme A, nAChR stands for Nicotinic acetylcholine receptors, CAT stands for Choline acetyltransferase, VAChT stands for Vesicular acetylcholine transporters, AChE stands Acetylcholinesterase, CHT1 stands for The high-affinity choline transporter, and α, β, γ, δ represent different types of Acetylcholine receptor proteins. The reactive components are shaded orange or pink, and the channels are shaded in blue (Mi et al., 2021).

Figure 40: A schematic representation of the selected cohort of subjects for WES experiments from two different projects (PROPANE Study and PAIN-Net).

Figure 41: A schema for WES data workflow from raw reads to annotated file

Figure 42: A schematic representation of the variant interpretation workflow applied to the familial and sporadic cohorts

Tables:

Table 1: In-depth comparison of NGS DNA application technologies

Table 2: Variant classification terminology provided by ACMG guidelines and its definition

Table 3: Clinical characteristics of the familial cohort recruited in the study along with the phenotypes observed in each family. The presence of the phenotype in a family is represented with √ mark.

Table 4a: Table showing the causative/protective genetic segregation models adapted for PROP01 family with respect to SFN, NP and itch. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) after applying the filtration criteria mentioned in Methods section. Individuals highlighted in red are affected members, whereas unaffected members are in green. The Ref stands for Reference variant, Het stands for Heterozygous variant and Hom stands for Homozygous variant.

Table 4b: Table of prioritized qualifying variants for PROP01 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 5a: Table showing the causative/protective genetic segregation model in 'candidate gene' for PROP02 family with respect to PN and itch. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 5b: Table showing the causative/protective genetic segregation model at "open exome" level for PROP02 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 5c: Table showing the prioritised list of qualifying variants for the PROP02 family. The affected member of the family have been listed before and the unaffected member at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 6: Tabulated list of published SCN9A mutations seen in PROP03 known to be associated with SFN and NP (Faber *et al*, 2012a; Zeberg *et al*, 2020)*.*

Table 7a: Table showing the causative/protective genetic segregation model for the PROP03 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 7b: Table showing the causative/protective genetic segregation model for the PROP03 family with respect to SFN. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 7c: Table showing the prioritised list of qualifying variants for the PROP03 family. The qualifying variants belonging to the causative or protective model are reported in the last two columns. The SCN9A:p.Met932Leu + SCN9A:p.Val991Leu are in cis as they impact the gene together.

Table 8a: Table showing the causative/protective genetic segregation model for PROP04 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 8b: Table showing the causative/protective genetic segregation model for the PROP04 family with respect to SFN. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 8c: Table showing the prioritized list of qualifying variants for the PROP04 family. The qualifying variants belonging to the causative or protective model is reported in the last two columns. The SCN9A:p.Met932Leu + SCN9A:p.Val991Leu are in cis as they impact the gene together.

Table 9a: Table showing the causative genetic segregation model for the PROP05 family with respect to SFN and NP. After applying the filtration criteria mentioned in the Methods section, it includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 9b: *Table showing the prioritised list of qualifying variants for the PROP05 family. The qualifying variants belonging to the causative or protective model are reported in the last two columns.*

Table 10a: Table showing the causative/protective genetic segregation model for PROP06 family with respect to SFN and NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 10b: Table showing the causative/protective genetic segregation model for PROP06 family with respect to itch. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 10c: Table of the prioritised qualifying variants for the PROP06 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 11a: Table showing the causative/protective genetic segregation model for the first group of six individuals in the PNET01 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 11b: Table showing the causative/protective genetic segregation model for the second group of four individuals in the PNET01 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 11c: Table showing the causative/protective genetic segregation model for the third group of three individuals in the PNET01 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 12a: Tabulated representation of mutated genes shared amongst the affected members, which were absent in the unaffected members. This list of genes was obtained after applying the variant filtration criteria mentioned in the Methods section.

Table 12b: Table of prioritized qualifying variants for three individuals of the PNET01 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 13a: Table showing the causative/protective genetic segregation model for PNET02 family with respect to SFN, NP and itch at 'open exome' level. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 13b: Table showing the prioritised list of qualifying variants for the PNET02 family. The affected member of the family have been listed before and the unaffected member at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 14a: Table showing the causative/protective genetic segregation model for the PNET03 family with respect to SFN and itch at 'open exome' level. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 14b: Table of the prioritised qualifying variants for the PNET03 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 15a: Table showing the causative/protective genetic segregation model for the PNET04 family with respect to SFN and NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 15b: Table showing the prioritised list of qualifying variants for the PNET04 family. The affected member of the family have been listed before and the unaffected member at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 16a: Table showing the causative/protective genetic segregation model for the PNET05 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Table 16b: Table of prioritized qualifying variants for PNET05 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 17a: Table showing the causative/protective genetic segregation model for the PNET06 family with respect to PN and NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant
Table 17b: Table showing the prioritised list of qualifying variants for PNET06 family. The affected member of the family have been listed before and the unaffected member at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Table 18a: The count of each variant after each step of prioritization per family.
§ denotes one out of the total number of variants was classified as 'Likely benign'

Table 18b: The count of each variant by its classification status for each family

Table 18c: The number of classified qualifying variants seen more than once in families is summarized, the grouping of the variants according to the gene they map.

Table 18d: The 8 qualifying genes that overlapped amongst the families with at least one recurrence in two families are summarized

Table 19: Clinical aspects of the sporadic cohort recruited under PROPANE and PAIN-Net study

Table 20: The list of 11 genetic variants seen in 12 different subjects as the variant marked § is seen in two different individuals

Table 21: List of 17 mutated genes shared in more than one sample in the sporadic cohort. A sample and variant count has been recorded for each mutated gene.

Table 22: List of mutated genes in the entire cohort, considering both familial and sporadic samples, with their sample and variant counts

Table 23: List of 12 mutated genes in common between familial and sporadic cohorts along with their gene names

Table 24: The count of input genes annotated by WebGestalt tool to run pathway analysis.

Table 25a: Tabulated pathway analysis result from mapped input qualifying genes from the familial cohort. The table contains five columns: 1. Name of the pathway 2. The number of genes in that pathway 3. The number of genes overlapped between the pathway and the given input 4. Enrichment ratio 5. p-value of each hypergeometric test 6. FDR for each test.

Table 25b: Tabulated pathway analysis result from mapped input qualifying genes from the sporadic cohort. The table contains five columns: 1. Name of the database 2. Name of the pathway 3. The number of genes in that pathway 4. The number of genes overlapped between the pathway and the given input 5. p-value of each hypergeometric test 6. FDR for each test.

Table 26: The number of samples with variants in the five genes that were common between the familial and sporadic cohorts

Table 27: Tabulated results of SV analysis for the familial cohort. The first column gives the total SVs detected in the affected members of each family but absent in unaffected members of the family. The second column gives the count of variants that passed filtering criteria. The third column gives the count of variants after variant filtration. The fourth column shows the number of SVs classified as VUS or in higher classes and the number of SV regions that involved at least one candidate gene.

Table 28: Tabulated list of potential SV regions observed in 80 samples. The gene count column explains the count of genes involved in that SV region. DEL stands for Deletion SV, BND stands for Break-end SV, and DUP stands for Duplication SV.

Table 29: List of qualifying variants observed in the proband of PNET01 family and the segregation of these variants in other family members. The unaffected members are in the last two columns shaded in grey colour.

Table 30: Phenotypes and keywords selected as input to create the virtual gene panel

Tables in Appendices

Appendix A.1:

Table A.1: List of genes that have been reported to be involved in a type of IPN. This table has been adapted from (Scherer et al, 2020)

Appendix A.2:

Figure A.2: A schematic representation of the attributes given to a particular variant and its ranking along with its classification result

Appendix A.3:

Table A.3: List of candidate genes used in this project as 'candidate gene' filter that was curated by using multiple databases in a multi-step procedure.

Appendix A.4:

Table A.4: List of 113 prioritised genes observed in the entire cohort.

Introduction

Peripheral sensory system and Peripheral Neuropathy

An organism's ability to perceive its environment is fundamental for its survival. In humans and other mammals, the somatosensory system enables them to perceive temperature, touch sensations, pain, and its position in space. The somatosensory system is a complex system composed of several structures, including receptors, such as thermoreceptors, proprioceptors, nociceptors, mechanoreceptors, and nerve fibres (Hedayat & Lapraz, 2019). It is involved in the identification of stimuli and conveys the information through electrical signals to the central nervous system (CNS).

The peripheral sensory system includes three classes of peripheral nerve fibres: A, B, and C fibres. These peripheral nerve fibres are categorised by diameter, degree of myelination, and conduction speed. High conduction velocities are observed in nerve fibres with big diameter axons and increased myelin insulation; the conduction velocity is fastest in A-fibres and slowest in C-fibres.

The class A-fibres are responsible for passing on the signal from the periphery to the dorsal horn of the spinal cord. The signals are then transmitted to different parts of the brain designated for pain recognition through ascending pain pathways.

Furthermore, class A nerve fibres are divided into four sub-categories: Aα, Aβ, Aγ, and Aδ. These sub-categories were defined based on the degree of myelination and axon thickness of each nerve fibre, enabling the transmission of signals at different velocities. The Aα fibres are involved in proprioception, whereas the Aβ and Aδ fibres are responsible for providing sensory innervation at the periphery.

The class B and C-fibres are part of the autonomic sympathetic nervous system involved in understanding temperature and its position in space (Glatte *et al*, 2019).

Different nerve fibres are responsible for sensitizing different sensory receptors. In particular, proprioceptors are sensitized by type Aα and Aβ sensory fibres, mechanoreceptors by Aβ and Aδ sensory fibres, and nociceptors and thermoreceptors by type Aδ and C sensory fibres (Varga & Mravec, 2015).

Peripheral neuropathy (PN) is a common nerve disorder that affects the peripheral nerves. PN can be caused by damage to the axons or the myelin sheath, which covers the neurons. A combination of axonal injury and demyelination can also causes PN (Hughes, 2002; Torpy *et al*, 2010). Depending on the degree of damage to the axons and/or the myelin sheath and the kind of neuron damaged, PN can impact sensation, mobility, pain transmission, gland or organ function (Tesfaye & Selvarajah, 2012; Quasthoff & Hartung, 2002; Watson & Dyck, 2015b).

Epidemiology

The prevalence of PN ranges from 2.4 to 8 % in the general population; a higher frequency is seen in patients over 55 years old. PN is a common complication for diabetes mellitus, as 50 % of patients develop PN, also known as diabetic peripheral neuropathy (Watson & Dyck, 2015a; Tesfaye & Selvarajah, 2012; Daousi *et al*, 2004). Diabetes is one of the most common metabolic disorders in the global population, and its prevalence is increasing. This disorder has been estimated to inflict about 592 million people by 2035, and almost half of these people will have PN (Guariguata *et al*, 2014). Other than diabetes, PN is also a known complication of human immunodeficiency virus infection, alcohol use, injury, and it is also often observed in patients undergoing chemotherapy (Watson & Dyck, 2015a; Castelli *et al*, 2020).

Symptoms

The symptoms in PN can be highly variable and impact different disorders involving sensory, motor, and autonomic nerve fibre dysfunctions. Symptoms can range from altered sensations, muscle atrophy, autonomic problems, and Neuropathic Pain (NP). A patient could present with only one symptom or a combination of symptoms. Sensory dysfunctions include numbness, tingling, paraesthesia, hyperalgesia, and allodynia. Autonomic symptoms include orthostatic dizziness, abnormal sweating, urinary retention, constipation, and sexual dysfunction. NP is found in approximately 30% of people with PN. If the patient is diagnosed with PN, a complete clinical history and examination are required to determine the extent of the symptoms. (Watson & Dyck, 2015a).

Pathophysiology

The classification of PN is defined according to the number of nerves affected and the type of nerve fibre predominantly involved. PN can be mononeuropathy if a single nerve is damaged or polyneuropathy if multiple nerves are affected (Hughes, 2002). According to the type of damaged nerves, it can be a motor, a sensory, or an autonomic PN. The aetiology of PN progresses from damage to small- or large-diameter nerve fibres to components within the nerve fibres such as nerve cell bodies, myelin sheaths, or axons. The large-nerve fibres modulate motor, sensory, vibration, and proprioception functions, whereas small-fibre nerves affect temperature, pain, and autonomic functions. Because of this, patients are distinguished according to their symptoms and the type of nerves involved (Watson & Dyck, 2015a; Castelli *et al*, 2020). It is essential to get a proper clinical history and a physical examination to classify the neuropathy either by pattern distribution like length-dependent, length-independent, focal, or multifocal or by its symptoms' modality such as sensory, motor, autonomic, or mixed (Castelli *et al*, 2020).

Causes

PN can be caused by a number of factors, including cellular and molecular pathways.

PN can be inherited or acquired, but in many cases, no specific cause can be found; hence the condition is classified as idiopathic. Charcot-Marie-Tooth (CMT) disease is a well-known hereditary form of PN. Comorbidity, infections, immunological dysfunction, inflammation, external toxins, autoimmunity, and cancer are all examples of acquired PN, which are categorised according to their triggering factors (Dyck, 2010).

The inherited and acquired peripheral neuropathies are described below.

Inherited peripheral neuropathies

Inherited peripheral neuropathies (IPN) comprise a variety of disorders characterized by various clinical manifestations. IPN is classified into four categories:

Hereditary motor neuropathy (HMN)

HMN is also known as distal hereditary motor neuropathy as it affects the distal motor region of the peripheral nervous system (PNS) in a length-dependent manner. Because of its clinical and genetic heterogeneity, this type of neuropathy is divided into seven sub-categories. Also, this neuropathy has often been linked with amyotrophic lateral sclerosis (Liu et al, 2020). However, hardly any subject is seen with sensory deficits.

Hereditary sensory neuropathy (HSN):

In this type of neuropathy, sensory axonal damage results in limb ulcers, infections, and, at times, amputations. This neuropathy progresses slowly, but subjects sometimes also present with NP and autonomic dysfunctions. Owing to these symptoms, hereditary sensory neuropathy originates from another type of neuropathy known as hereditary sensory and autonomic neuropathy (Cintra et al, 2021).

Hereditary sensory and autonomic neuropathy (HSAN)

This type of neuropathy has evolved from hereditary sensory neuropathy as the symptoms include both sensory and autonomic dysfunction. Therefore, the impairment of small and unmyelinated nerve fibres is more suggestive than large myelinated fibres. This neuropathy has eight main subtypes, which involve both modes of inheritance (autosomal dominant and autosomal recessive) with variants known in 12 genes (Rotthier *et al*, 2012; Schwartzlow & Kazamel, 2019a).

Hereditary motor and sensory neuropathy (HMSN)

Hereditary motor and sensory neuropathy is another name for the CMT disease. There are four main groups under this neuropathy where peripheral nerves are impaired due to degeneration and demyelination. In this neuropathy, mostly large nerve fibres are affected, but there have been cases with sensory impairment. The most common form of this neuropathy is CMT-1A, which is caused by the duplication of the *PMP22* gene, where large fibres demyelination occurs at a slow pace (Scherer et al, 2020; Bienias, 2020).

One of the most common IPNs is CMT disease, which is axonal and demyelinating in nature (Scherer *et al*, 2020). CMT reported incidence is of 14.5 per 100,000 in the global population. IPN symptoms vary, but they are known to be subtle at first and gradually progress over time. It includes sensory symptoms such as burning pain, tingling or numbness in hands or feet, distal-motor symptoms like ataxia, gait disturbances, or autonomic difficulties like hyperhidrosis and orthostatic intolerance. Some patients might present with leg deformities such as thin calf muscles or hammertoes.

Owing to its genetic heterogeneity and varied Mendelian inheritances, about 102 genes have been associated with CMT (Theadom *et al*, 2019; Hereditary Motor Sensory Neuropathies: Charcot-Marie-Tooth). Most of the genes (*AARS*, *GARS*, *HARS*, *KARS*, *MARS, SBF1*, and *YARS*) involved in CMT disease encode cytosolic tRNA synthetases, which have a role in the metabolism of proteins or lipids. In addition, there are also genes like *PMP22* that are involved in axon growth and guidance.

In HMN, the proteins involved also include cytosolic tRNA synthetases, which play a role in membrane trafficking.

Other genes involved in HSN or HSAN include *NGF, NTRK1, SCN9A, SCN10A,* and *SCN11A,* which modulate sensory perception and axon guidance. A detailed list of genes that involve IPN is in Appendix A.1 (Scherer *et al*, 2020).

Thanks to recent advances in bio-computational technologies, genetic testing in suspected cases of IPN is practically standardized and could aid in the classification and assessment of IPN cases (Scherer *et al*, 2020).

Acquired peripheral neuropathies:

Neuropathies induced by metabolic disorders

These neuropathies progress slowly over time in a length-dependent manner, and their physiology shows an axonal loss. Diabetes, or glucose intolerance, is one of the most common causes of PN. The development of PN post-diabetes depends on the duration of diabetes, glycaemic index control, and the presence of retinopathy or nephropathy (Chung *et al*, 2014). The most common form of diabetic PN is distal symmetrical polyneuropathy. The complex pathogenesis of diabetic PN involves vascular and metabolic factors at all phases of diabetic PN (Tesfaye & Selvarajah, 2012).

Another example of neuropathy associated with metabolic disorders is caused by abnormal thyroid hormone production. Hypothyroidism-related neuropathy can occur due to less production of hormones, leading to swollen tissues affecting peripheral nerves (Yeasmin *et al*, 2007).

Vitamin deficiencies are another cause of acquired neuropathy. The most common vitamin deficiency is B_{12} which causes loss of proprioception and hyperreflexia. The other known causes of PN are B_6 and E deficiencies.

Patients with chronic renal dysfunction with a lower glomerular filtration rate lead to blood toxicity. This chronic renal dysfunction-related neuropathy leads to nerve tissue damage, causing PN. Lastly, the development of PN is also seen in patients with chronic liver disease (Chung *et al*, 2014).

Immune-mediated and inflammatory neuropathy

The immune-mediated neuropathies are a group of diseases in which immune dysregulation leads to inflammation of peripheral nerve fibres and their demyelination. The most common acute type of immune-mediated neuropathy is Guillain–Barré syndrome. The Guillain–Barré syndrome is a group of heterogeneous conditions with diverse pathology (Chandra *et al*, 2018). The immune-mediated neuropathies also include progressive or relapsing-remitting chronic inflammatory demyelinating polyneuropathy, or multifocal motor neuropathy (Chung *et al*, 2014).

Another type of immune-mediated neuropathy is vasculitic neuropathy, in which the pathologic process is initiated in the blood vessels, leading to nerve ischemia or infarction, causing multifocal motor and painful sensory neuropathy. This neuropathy is frequently seen in patients with systemic vasculitis or develops over time after receiving a diagnosis of rheumatoid arthritis or viral infections (Gwathmey *et al*, 2014).

Toxic neuropathy

In toxic neuropathies, specific neurotoxins affect the entire nerve, from the neuron cell body to the terminal axon at the neuromuscular junction. This neuropathy causes axon degeneration proximal to the injury, leading to disruption of the nerve cell body and cell death. This phenomenon is also called "dying-back" neuropathy, which does not allow for regeneration if the cell body dies (Chung *et al*, 2014). These neurotoxins can be prescribed medications, environmental and industrial toxins, or alcohol abuse. The most common neurotoxin that causes toxic neuropathies is chemotherapeutic drugs (Watson & Dyck, 2015a).

Neuropathies related to autoimmune diseases

In autoimmune-related neuropathy, autonomic dysfunction leads to the triggering of autoimmunity in the PNS. In some patients, the autoimmune dysfunction may occur as a systemic autoimmune disease like Sjögren syndrome or the antiphospholipid syndrome. The other diseases that can induce PN are systemic sclerosis, systemic lupus erythematosus, and sarcoidosis (El-Abassi *et al*, 2022).

Neuropathy related to cancer

In this neuropathy, the infiltration of tumour cells into cranial nerves and nerve roots can lead to PNS dysfunction, as it is commonly observed in leukaemia and lymphoma patients (Chung *et al*, 2014). Recently, a study demonstrated the involvement of C-type fibre in neurofibromatosis, causing painful sensory neuropathy (Farschtschi *et al*, 2020).

Neuropathies related to infection

In this type of neuropathy, infection by viruses or bacteria causes the pathogen agents to enter nerve fibres, causing severe nerve damage. For example, *Mycobacterium leprae* causes leprosy and leads to PN. The other example is the AIDS-causing human immunodeficiency virus (HIV) that has been associated with several neurological symptoms, including PN. Other examples of viral and bacterial infections causing PN are as follows: Lyme disease, the diphtheria-causing bacteria Corynebacterium diphtheriae, Epstein-Barr virus, Cytomegalovirus, and Varicella-zoster virus (Chung *et al*, 2014).

Diagnosis

Due to the phenotypes' complexity in PN, a detailed medical history and extensive examination are needed to precisely evaluate a patient and make a diagnosis. The diagnostic evaluations can include blood tests to detect glucose intolerance, vitamin deficiencies, liver or kidney dysfunction, hypothyroidism, and abnormal immune system activation. Also, to understand the involvement of immune-mediated neuropathies, cerebrospinal fluid can be tested to examine specific markers of autoimmunity.

To confirm IPN, clinicians thoroughly examine the subject's family medical history. If the clinician suspects IPN, the patient can be asked to undergo genetic testing. The investigation of the various types of nerves involved and their distribution pattern is critical in determining a patient's clinical pattern. Muscle strength assessment and evidence of cramps or fasciculation are investigated to confirm motor fibre deficits. The patient's ability to perceive vibration, light touch, pain, temperature, and body position is evaluated to confirm sensory nerve damage. This test also determines whether the nerve fibres are small or large (Lehmann *et al*, 2020).

Electrodiagnostic studies such as nerve conduction studies (NCS) and needle electromyography (EMG) are performed to: i) confirm the clinical manifestations of PN; ii) exclude mimickers of PN; iii) assess subclinical involvement of clinically unaffected nerve and fibre modalities; iv) evaluate the mechanism of nerve damage (axonal loss or demyelination); and v) assess disease severity (Lehmann *et al*, 2020).

In NCS, an electrical stimulation device with electrodes is kept along the skin surface underlying the nerve to measure the speed of impulse propagation by transmembrane ions. Demyelination is predicted if the transmission rates are slow and nerve conduction block is observed. In comparison, the axonal loss is indicated by reducing nerve action potentials and normal or marginally reduced nerve conduction velocity (Abrams, 2006; Lehmann *et al*, 2020).

The EMG uses an insertional electrode to record electrical muscle activity both at rest (spontaneous) and when it is contracted (voluntary motor unit). If abnormal activity is detected at rest, active denervation can be predicted, and the severity of the nerve injury can be estimated at the contracted stage (Chung *et al*, 2014).

Autonomic dysfunction or sensory nerve damage is investigated using pathological parameters such as nerve biopsy and skin biopsy. Autonomic testing examines a variety of functions, including cardiovagal function, adrenergic function, and postganglionic sudomotor function. Cardiovagal and adrenergic functions are assessed by measuring blood pressure after the Valsalva manoeuvre and recording heart rate responses to deep breathing. The quantitative sudomotor axon reflex test (QSART) is the most quantitative test, it measures action potential transmission along the postganglionic sympathetic sudomotor axon (Schwartzlow & Kazamel, 2019b; England *et al*, 2009).

The sural nerve biopsy is an invasive procedure that investigates the morphological changes between axon degeneration and demyelination. It is performed when the clinical manifestations are suggestive of inflammatory neuropathy.

Skin biopsy is a minimally invasive technique for assessing small, unmyelinated fibres. The density of intra-epidermal nerve fibres (IENF) can be identified and counted using a skin biopsy (Lauria & Devigili, 2007), as these fibres are invisible to routine neurophysiological tests. Skin biopsy can aid researchers in their study of myelinated fibres and their receptors, as well as provide pathophysiological insights through morphometry of skin myelinated fibres in patients with symptoms consistent with small fibre neuropathy (Nolano *et al*, 2020).

Treatment

The available treatment options for PN concentrate on addressing the underlying cause. Because there are numerous causes of PN, the clinician will recommend a variety of treatment options. Patients have the option of receiving either pharmacological or non-pharmacological treatment.

Non-pharmacological treatments entail proper foot hygiene, weight reduction, gait rehabilitation, and physical therapy to improve symptoms associated with lower limb PN.

NP is the most common phenomenon in PN patients (Cruccu & Truini, 2017). As a result, patients are prescribed antidepressants (e.g., venlafaxine), anticonvulsants (e.g., gabapentinoids), and opiate analgesics (e.g., oxycodone) to treat their NP symptoms (Trivedi *et al*, 2013; Castelli *et al*, 2020). The recommendation of topical medications is given mainly for localised NP.

For diabetic PN, treatment options include the optimization of the glycaemic index and medication such as tricyclic antidepressants (e.g. amitriptyline) or serotonin/norepinephrine reuptake inhibitors (e.g. duloxetine) (Farhat & Yezback, 2016). The administration of intravenous immunoglobulins or steroids can treat inflammatory and autoimmune-related neuropathies (Chung *et al*, 2014).

Regarding IPN, the treatment options are scarce, but certain gene replacement therapies approved by the Food and Drug Administration (FDA) are available. For example, adeno-associated virus-9 (AAV9) and adeno-associated virus-2 (AAV2) vectors for spinal muscular atrophy and *RPE65*-linked retinal dystrophy, respectively (Motley *et al*, 2022). Gene replacement therapy is the process of identifying a faulty gene and delivering a fragment of DNA in its proper form to the gene via a viral vector (known as the carrier molecule), thus replacing the detected faulty gene with the correct copy. (Pandey & Balekar, 2018). Small Fiber Neuropathy

Small Fiber Neuropathy

Small Fiber Neuropathy (SFN) is a form of PN characterised by damage to the thinly myelinated Aδ or unmyelinated C-fibres, also called intraepidermal nerve fibres (IENFs) that act as nociceptors and thermoreceptors in the skin.

Peripheral axons are an essential component of the autonomic nervous system because Aδ fibres connect the CNS to the ganglionic nerves of the PNS through the preganglionic fibres, whereas C-fibres connect the ganglion to the effector organs via the postganglionic fibres, such as sweat glands, the heart, and blood vessels (Themistocleous et al, 2014).

These small peripheral nerve fibres facilitate pain transmission, temperature perception, enteric and autonomic functions. Therefore, patients with SFN may experience autonomic dysfunction (Hovaguimian & Gibbons, 2011)

SFN can manifests with or without the involvement of large nerve fibres, or it can progress gradually to involve large fibres. Because of these phenotypic differences, SFN is difficult to diagnose.

Epidemiology

In the general population, the prevalence of SFN is estimated to be around 53 per 100,000 people. However, patients are frequently underdiagnosed during routine neurologic examinations because it is difficult to detect abnormalities at the small fibre level (Peters *et al*, 2013). Another study reported a steady increase in incidence and observed that SFN was more prevalent in obese people (Johnson *et al*, 2021). Furthermore, another study from Switzerland attempted to determine the prevalence of SFN in their population. They also observed a significant increase in SFN incidences, from 4.4 to 131.5 cases per 100,000 inhabitants over a 4.8-year period (Bitzi *et al*, 2021). After a thorough work-up of two-thirds of the patients, Bitzi *et al*. discovered that an etiological relationship could be attributed. The majority of patients in their study were unable to be etiologically identified, and they stressed the significance of more research into this area.

Symptoms

The clinical representation of SFN is heterogeneous. It is characterised by sensory disturbances and, at times, autonomic dysfunction. Many patients report a slow progression to distal symptoms that include unclear lower limb disturbances. For example, the patients feel like there are small stones or mud in their shoes (Hovaguimian & Gibbons, 2011).

Most of the patients present with NP, with or without abnormal thermal perception. For patients affected by SFN, the symptoms are distributed in the lower extremities in a length-dependent manner. They can, however, differ in how the disease manifests itself, with symptoms affecting either the upper or all four limbs and whether or not they have a length-dependent distribution. Patients may have significant somatosensory symptoms as well as a variety of unpleasant sensations (e.g., burning, electrical, or shooting pain). These painful sensations can vary in intensity throughout the day and night. Some individuals may experience dysesthesia (e.g. prickling sensations, tingling sensations, pruritus, numbness, cold, or stabbing pain) (Hovaguimian & Gibbons, 2011).

Allodynia and hyperalgesia are strong clinical indicators of SFN. Some individuals have also reported pain in the tongue, scalp, face, and torso (Sène, 2018). At times, the patients have reported symptoms similar to restless leg syndrome.

Also, autonomic dysfunction might manifest along with other symptoms. The commonly observed autonomic manifestations include abnormal sweating, episodic flushing, constipation, gastroparesis, micturition disorders, orthostatic hypotension, and cardiac arrhythmias (Sène, 2018).

Pathophysiology

The small sensory nerve fibres include thinly myelinated Aδ fibres responsible for mechanical allodynia to pinprick stimuli and unmyelinated C-fibres that transmit warmth sensation and pain sensitivity to thermal stimuli (Politei *et al*, 2019). The degeneration of peripheral nerve fibres involved in pain processing is indicated by the reduction in these small sensory nerve fibres. As a result, one of the mechanisms underlying SFN could be axonal degeneration in small dorsal root ganglion (DRG) neurons.

A study demonstrated that partially inactivating voltage-gated sodium channels leads to intracellular calcium toxicity (Estacion *et al*, 2015). Thus, one of the possible explanations for axonal degeneration could be calcium dyshomeostasis (Gross & Üçeyler, 2020).

The authors, Gross & Üçeyler, stated that in addition to calcium dyshomeostasis, impaired axonal transport, overexpression of nerve growth factor, and local inflammation could also lead to axonal damage in small fibres. The nociceptors, which are the terminal sensory organs in the skin, muscles, joints, and viscera, respond to potentially tissue-damaging stimuli (Gold & Gebhart, 2010), and their sensitization, the increment of excitability, could lead to increased pain sensitivity. In this way, nociceptor sensitization could be another possible explanation for SFN pathogenesis. (Gross & Üçeyler, 2020).

Causes

There are many potential causes of SFN, which makes it complex and variable, making it harder to diagnose. SFN may be defined as primary or secondary. In the case of primary forms of SFN, the discrimination between genetic and idiopathic forms is fundamental. Approximately 40% of the cases in a cohort of patients affected by SFN tend to be idiopathic (Devigili *et al*, 2008; Bitzi *et al*, 2021). Also, many of these cases might initially present as SFN but then develop into large fibres, making them cases of polyneuropathy. In the case of secondary forms of SFN, the specific cause has to be investigated. SFN, like PN, can have inherited, acquired, or idiopathic aetiology. The different aspects of inherited and acquired SFN have been elaborated more below, whereas Figure 1 summarises the primary and secondary causes of SFN.

Hereditary SFN

Several genetic studies were conducted to determine if there are any genetic factors associated with SFN and, as a result, to better understand the pathophysiology of SFN (C. *et al*, 2012). The subgroup of inherited SFN consists of patients carrying mutations in voltage-gated sodium channel (VGSC) genes, namely, *SCN9A*, *SCN10A*, *SCN11A*, and *SCN2B* (Faber *et al*, 2012a, 2012c; Huang *et al*, 2014; Alsaloum *et al*, 2021). Patients with sodium channel gene mutations can be inherited or sporadic in nature. The gain-of-function mutations in VGSCs have been associated with paroxysmal extreme pain disorder (Fertleman *et al*, 2006), erythromelalgia (Cummins *et al*, 2004), and familial episodic pain (Faber *et al*, 2012c). Because of these genetic studies on VGSCs, patients who were first classified as idiopathic SFN following genetic screening were reclassified as sodium channel-related SFN in about 25% of cases.

Tangier's disorder is an autosomal recessive disease characterised by high-density lipoprotein deficiency due to mutations in the *ABCA1* gene. In Fabry's disease, an X-linked lysosomal disease involving the *GLA* gene, SFN is the main manifestation, together with autonomic dysfunction. SFN has also been documented as a pre-symptomatic stage of familial amyloid polyneuropathy caused by *TTR* gene mutations. (Cazzato & Lauria, 2017; Sène, 2018).

Finally, when the nerves affected by HSAN concentrate exclusively in small sensory fibres, the condition is known as "Hereditary sensory and autonomic small fibre neuropathy." Appendix A.1 lists the genes that cause this hereditary SFN disorder. The clinical manifestations of this condition might range from painful to painless, including familial dysautonomia. (Terkelsen *et al*, 2017).

Acquired SFN

Metabolic causes

Diabetes mellitus appears to be the most common recognised metabolic cause of SFN, comparable to PN, involving approximately one-fourth of SFN cases. Furthermore, SFN is associated with impaired glucose intolerance and metabolic syndrome. Other known causes of SFN include chronic liver illness, hypothyroidism, vitamin B_{12}, and B_6 deficits (Cazzato & Lauria, 2017; Sène, 2018).

Inflammatory and autoimmune causes

Two conditions are potentially involved in leading to SFN, namely, Sjögren's syndrome and sarcoidosis. Sjögren's syndrome is an autoimmune disorder that generally affects the exocrine glands. In patients with Sjögren's syndrome, about 25%-45% end up developing SFN (Birnbaum *et al*, 2019).

Sarcoidosis is a multi-systemic granulomatous disease. A study demonstrated that 44% of patients with sarcoidosis had NP and sensory dysfunction (Hoitsma *et al*, 2002).

The other inflammatory and autoimmune diseases which have been reported with SFN are as follows: coeliac disease (de Greef *et al*, 2018), sensorimotor polyneuropathy (Oaklander & Nolano, 2019), systemic lupus erythematosus, vasculitis, inflammatory bowel syndrome, paraneoplastic syndrome, monoclonal gammopathy and rheumatoid arthritis (Cazzato & Lauria, 2017).

Toxic agents, including medication

SFN due to alcohol abuse is more commonly reported compared to other known toxic agents of SFN (Terkelsen *et al*, 2017). SFN due to alcohol abuse generally manifests with NP. Other toxic agents associated with SFN are as follows: statins, anti-retroviral drugs, chemotherapeutic agents, antibiotics, flecainide, tumour necrosis factor (TNF) inhibitors, and vitamin B_6 toxicity.

Infectious causes

A strong association has been observed between SFN and patients affected by HIV (Cazzato & Lauria, 2017). SFN was reported in patients with hepatitis C, Lyme disease, influenza, Chagas disease, leprosy (Cazzato & Lauria, 2017; Themistocleous et al, 2014; Al-Shekhlee et al, 2002; Sène, 2018).

Syndromic causes

In syndromic causes, patients have been diagnosed with other diseases, but their clinical examination either presents with loss in IENF density or dysfunctional small fibres. For instance, around 40% of patients with fibromyalgia have impaired small fibres according to different diagnostic methods used for SFN. The other conditions involved in syndromic causes are Ehlers-Danlos syndrome, Parkinson's disease, amyotrophic lateral sclerosis, and Pompe disease (Cazzato & Lauria, 2017).

PRIMARY

Idiopathic

→ Idiopathic small fiber neuropathy
→ Burning mouth syndrome

Hereditary/genetic

→ Sodium channel mutations (*SCN9A,SCN10A,SCN11A,SCN2B*)
→ Familial amyloid polyneuropathy
→ Fabry's disease
→ Tangier's disease
→ Hereditary sensory-autonomic neuropathy

SECONDARY

Metabolic

→ Impaired glucose tolerance
→ Diabetes mellitus
→ Hypothyroidism
→ Vitamin B_{12} deficiency
→ Dyslipidaemia
→ Chronic kidney disease
→ Treatment-induced in diabetes
→ Metabolic syndrome

Infections

→ HIV
→ Hepatitis C
→ Influenza
→ Chagas disease
→ Lyme disease
→ Leprosy

Toxin & Drugs

→ Anti-retrovirals
→ Antibiotics: metronidazole, nitrofurantoin, linezolid
→ Statins
→ Chemotherapy: bortezomib
→ Alcohol
→ Vitamin B_6 toxicity
→ Flecainide
→ TNF Inhibitor
→ Thallium

Immune-mediated

→ Coeliac disease
→ Sarcoidosis
→ Sjögren's syndrome
→ Rheumatoid arthritis
→ Systemic lupus erythematosus
→ Vasculitis
→ Inflammatory bowel disease
→ Paraneoplastic
→ Monoclonal gammopathy/amyloid

Syndromic

→ Fibromyalgia
→ Ehlers-danlos syndrome
→ Parkinson's disease
→ Amyotrophic lateral sclerosis
→ Pompe disease

Figure 1: Causes of SFN. This figure describes the causes of SFN into two main categories: Primary and Secondary causes of SFN. This figure has been adapted from the tables mentioned in the following articles: Themistocleous *et al*, 2014; Al-Shekhlee *et al*, 2002; Cazzato & Lauria, 2017; and Alsaloum *et al*, 2021

Diagnosis

The involvement of the IENFs, which regular neurologic examination and electrophysiological testing generally are not powerful enough to detect due to the small calibre of nerve fibres, makes SFN diagnosis challenging. In this setting, the density of IENFs and their decline over time must be assessed for diagnosis.

Skin biopsy

The skin biopsy is utilised as a diagnostic test and can consistently demonstrate the loss of IENFs in SFN, allowing the diagnosis to be validated. In this diagnostic method, IENF density is measured with a 3 mm to 5 mm punch skin biopsy taken from the distal leg after local anaesthesia for evaluation. The biopsy undergoes microdissection, after which the three randomly chosen 50 μm thick sections are then immunolabelled against pan-neuronal marker protein gene product (PGP9.5), expressed in the cytoplasm of the neurons, allowing the detection of small fibres in immunofluorescence microscopy, as shown in Figure 2.

An individual is diagnosed with SFN if the IENF density is reduced compared to the normative values of age-adjusted and sex-adjusted IENF density of the distal leg (Lauria *et al*, 2005).
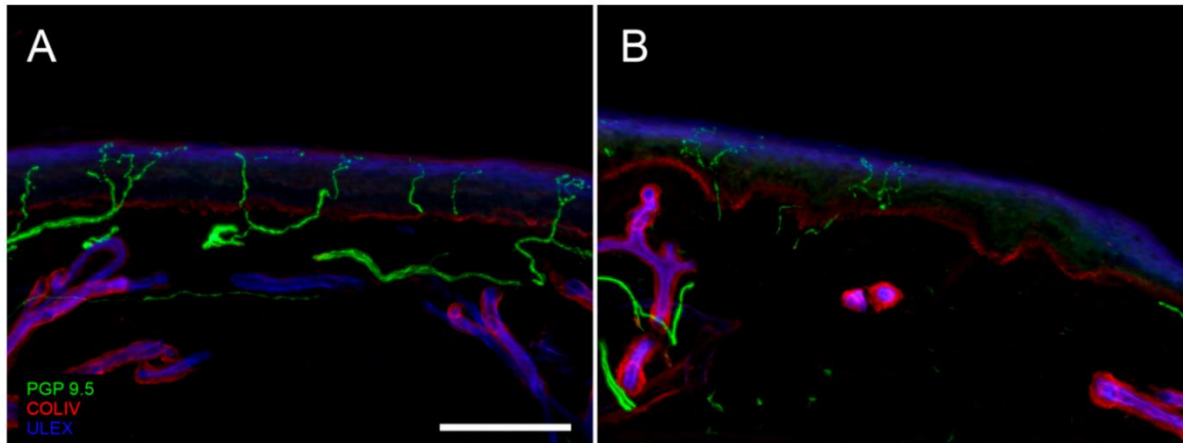
Figure 2: Digital confocal images of skin biopsy sections showing epidermal innervation in healthy control (A) and in a patient with diabetic SFN (B). In green: small fibre nerves marked with PGP9.5 antibody. In blue: endothelia and epidermis marked with ULEX europaeus. In red: basement membrane and blood vessels marked with collagen IV antibody. Scale bar: 100 μm. This image was retrieved from (Nolano et al, 2020).

According to the European guidelines, IENF density less than 7.6/mm at the distal leg has a diagnostic value of 88% to 90% to 95% specificity and 70% to 83% sensitivity for SFN (Lauria *et al*, 2005; Sène, 2018). Owing to its high diagnostic value, this minimally invasive tool is considered a gold standard for diagnosing SFN (Zhou, 2021).

Other diagnostic tools
The other diagnostic tools recommended together with skin biopsy are as follows:

–   The Quantitative Sudomotor Axon Reflex Test (QSART) investigates the postganglionic sympathetic C-fiber sudomotor nerve function, and it is recommended for patients with abnormal sweating and autonomic symptoms (Zhou, 2021). Together, QSART and skin biopsy can raise the diagnostic value of SFN (Thaisetthawatkul et al, 2013).

–   The Quantitative sensory test (QST) is used to study the perception in response to external stimuli of controlled intensity and to define the pain threshold (Cazzato & Lauria, 2017).

- Corneal confocal microscopy detects abnormalities in the C-fiber innervations of the cornea. It is a non-invasive and reproducible technique. This tool measures the corneal nerve fibre density through automated analysis of corneal nerve images. The decreased density of corneal nerve fibres was closely related to the results obtained from a skin biopsy. Moreover, multinational normative values for this tool have also been developed and can be added as a routine diagnostic tool for SFN (Cazzato & Lauria, 2017; Sène, 2018).

- Laser evoked potentials to assess Aδ fibres and heat-evoked potentials for small fibres can also be used as diagnostic tools. But, these neurophysiological tests have not been added to the clinical diagnostic routine (Oaklander & Nolano, 2019).

- Questionnaires: several questionnaires have been developed to detect symptoms related to SFN.
    - The Small Fiber Neuropathy and Symptoms Inventory Questionnaire (SFN-SIQ) gives information for 13 conditions: abnormal sweating pattern, diarrhoea, constipation, micturition dysfunction, dry eyes, dry mouth, dizziness on standing from sitting or supine positions, palpitations, hot flashes, sheet intolerance, burning feet, heat intolerance, and restless legs (Bakkers et al, 2014). Another SFN-related questionnaire is for patients with sarcoidosis-induced SFN, called the Small Fiber Neuropathy Screening List (SFN-SL). This questionnaire comprises 21 questions that focus on NP and dysautonomia (Terkelsen et al, 2017).

    - Another specific questionnaire for autonomic symptoms is called the Autonomic Symptom Profile and the Composite Autonomic Symptom Score-31 (COMPASS-31). This questionnaire comprises 31 questions that assess six domains of autonomic function. This test has been validated and has good reproducibility (Terkelsen et al, 2017).

- Lastly, neuropathic pain is a common symptom of SFN, and pain can be subjective. To that effect, pain questionnaires facilitate the measurement of pain intensity by attributing a quantitative value to it. The pain intensities are measured with different types of scales. One such scale is the Numerical Rating Score (NRS), where the levels of pain intensity are as follows: mild (1-4), moderate (5-6), and severe (7-10) (Caraceni & Shkodra, 2019).
The different types of available NP questionnaires are as follows: S-LANSS (The Self-Administered Leeds Assessment of Neuropathic Signs and Symptoms) (Birnbaum et al, 2019), DN4 (Douleur Neuropathique 4 questions), and painDETECT (Caraceni & Shkodra, 2019).

Treatment

There are two main treatment options for SFN: (i) treatment of symptoms and (ii) treatment of the cause. The treatment is determined by the clinical aspects of the patient, and one of the two approaches is recommended.

In the treatment of symptoms, NP is a common feature of SFN, and medication such as tricyclic antidepressants, serotonin-norepinephrine reuptake inhibitors, or gabapentinoids is recommended for alleviating NP (Sène, 2018). For localized pain, topical application of lidocaine or capsaicin is also suggested.

Regarding the treatment, which has an underlying cause, the drugs are recommended according to the acquired condition. For example, the clinicians might suggest intravenous immunoglobulin therapy for Sjögren's syndrome. Moreover, mexiletine, a class IB antiarrhythmic drug, is prescribed to patients with erythromelalgia with appreciable success. This drug is known to block sodium channel function (Cregg *et al*, 2014).

Genetics

As previously stated, mutations in VGSC genes are partly responsible for the pathophysiology of SFN. These genes are found in primary nociceptive neurons, which include DRG neurons (Laedermann *et al*, 2013).

Because the VGSCs on these neurons are in charge of generating and transmitting the electric pulse that transmits information from the periphery to the CNS, mutations in VGSCs can modulate pain perception. Electrophysiological studies on VGSCs revealed that mutations in specific functional regions of the VGSCs could affect the kinetics of channel opening, closure, and inactivation, resulting in hyper/hypoactivation in these neurons (Han *et al*, 2012b). These findings point to a molecular genetic basis for painful SFN.

Furthermore, in a recent review, the authors suggested that there are genes other than VGSCs that are potentially involved in pain modulation. In their systematic review, Veluchamy *et al*. collected and classified 28 genes associated with NP, involved in signal transduction, synaptic transmission, the immune system, iron metabolism, and drug catabolism. The authors then concluded that these genetic factors might be instrumental in the pathophysiology of pain and PN, encompassing SFN (Veluchamy *et al*, 2018).

Neuropathic pain and its genetics

NP is caused by a lesion or disease in the somatosensory nervous system. It's a common PN phenomenon, just like SFN (De Moraes Vieira *et al*, 2012; Attal *et al*, 2008).

There are two kinds of NP: spontaneous pain and provoked pain. Burning, tingling, numbness, sheet intolerance, pain caused by mild pressure, electric shock-like discomfort, pain from cold or heat, itching, and numbness are all symptoms of spontaneous pain. Provoked pain, on the other hand, is defined by external stimuli such as mechanical, thermal, or chemical, which can cause hyperalgesia or allodynia (Cruccu & Truini, 2017).

The heterogeneity in pain perception, as well as the distribution of symptoms, adds to the complexity of NP. NP is also subjective because it is heavily influenced by factors such as cognition, emotions, and an individual's education. Thus, to evaluate the severity of NP, several questionnaires have been employed to provide objectivity for measuring NP. There are different types of questionnaires available that concentrate on different NP descriptors to generate a score that determines if the pain is neuropathic or not, along with identifying specific characteristics of NP. It is recommended that a combination of screening tools be used to evaluate the degree of painful manifestations associated with PN and SFN.

The risk of inheriting pain-related phenotypes has been calculated to be between 25 and 50%, where the heritability of chronic pain is estimated to be around 50% (Nielsen *et al*, 2012).

Parallel to the problems in describing and reporting NP traits, several studies aimed at investigating the potential role of NP, focusing on its genetic aspects, and they could not satisfactorily associate genes with NP. This is mainly because most of these studies have drawbacks, such as small cohort size or uncertain phenotypic characterization of the samples (Calvo *et al*, 2019).

Veluchamy *et al*, highlighted a set of genes related to NP divided into four basic categories according to their biological pathways: ion channels, neurotransmission, metabolism, and immune response, as shown in Figure 3 (Veluchamy *et al*, 2018; Calvo *et al*, 2019). The authors divide the genes into various groups depending on their function or associated mechanisms, such as receptor signalling or iron metabolism or genes connected to pro-inflammatory cytokines.

Some examples are: the *SCN9A* gene that encodes for $Na_v1.7$, a VGSC, was sorted into the 'ion channel' group; *COMT*, which encodes for the catechol-O-methyltransferase, was classified into the 'neurotransmission' group. Similarly, the genes in the 'Metabolism' group are related to iron metabolism and were observed in individuals with distal peripheral neuropathic pain and low iron-binding.
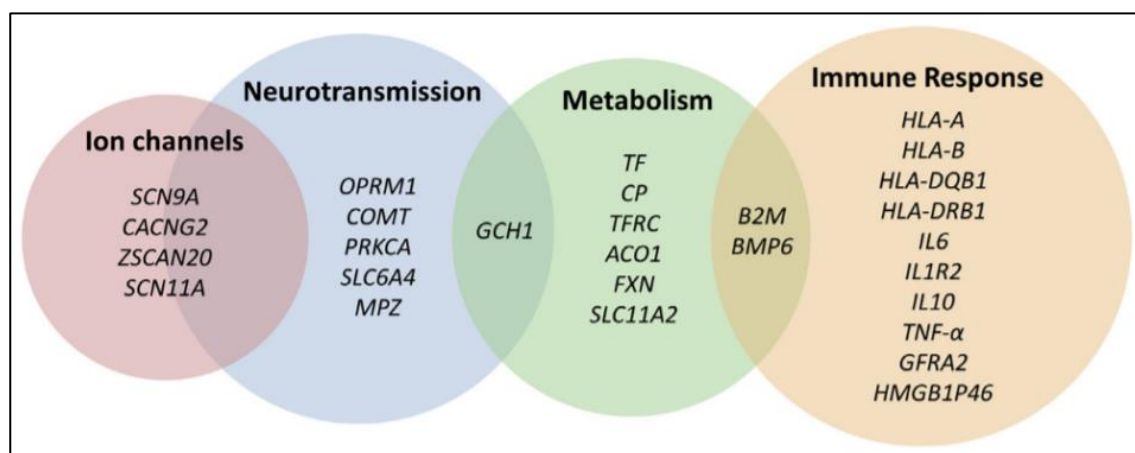
Figure 3: The genetic markers mentioned above have been extracted from the two reviews as they were associated with NP. (Veluchamy et al, 2018; Calvo et al, 2019)

The genetics of NP includes both rare and common variants, with the former occurring at a frequency of less than 1% in a population and the latter occurring at a noticeable frequency of more than 1% in a population. Recently, a database was released, listing different variants related to pain phenotypes; this database linked 125 distinct variants to NP and nociception (Meloto *et al*, 2018).

Despite all the research performed to date, projects such as genome-wide association studies or next-generation sequencing studies are still needed to broaden the genetic knowledge of NP and PN.

Nociception

Nociception is the process by which an animal's sensory system perceives and encodes noxious stimuli. The nociceptors, a set of specialised sensory receptors, are located in the PNS. These receptors identify and respond to heat, cold, chemical, and mechanical stimuli that can be potentially damaging. The nociceptors are present as free nerve endings at the periphery and as cell bodies in the DRG. (Cohen *et al*, 2018).

Within seconds of the first contact with noxious stimuli, peripheral sensitization occurs with free nerve endings (nociceptors) in the periphery. The conversion of the stimuli into electrical energy induces the activation of the VGSCs that generate action potentials and transmit the signal along the axon through the DRG.

Ion channels that convert stimuli to signals at the periphery include transient receptor potential (TRP) channels, acid-sensing ion channels (ASIC), and ATP-gated purinergic channels (P2X). This process is also known as signal transduction. The other important feature of nociception is the hyperpolarization of potassium channels that can inhibit the action potentials created by TRP channels at the periphery for regulating transmission. Meanwhile, in the spinal dorsal horn, nociceptors terminate and contact nociceptive-specific spinal projection neurons via N-methyl-Daspartate receptors (NDMAR), -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors (AMPA), metabotropic glutamate receptors (mGluR), and kainate receptors on the postsynaptic terminal. Voltage-gated calcium channels control neurotransmitter release with the help of the soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex that mediates synaptic vesicles in neurons (Finnerup *et al*, 2021). The signal then ascends, carrying sensory information from the body to the brain via the spinal cord, which is known as the ascending pathway. This process has been illustrated in Figure 4.

Figure 4: A pictorial representation of the transmission of a stimulus perceived at the periphery and the electric pulse is transported to CNS while interacting with different ion channels and biomolecules. It also shows the impact of nociceptive signals on different ion channels. This figure depicts three ion channels as voltage-gated sodium channels (Na$_v$s), voltage-gated calcium channels (Ca$_v$s), and hyperpolarization-activated cyclic nucleotide-gated channels (HCNs) are pro-excitatory while potassium channels (K$_v$s) and TRPM8 ion channels are not. At the same time, thrombospondins (TSP) are involved in synaptic transmission by operating through Ca$_v$ $a_2\delta_1$ subunits. N-methyl-D-aspartate receptor (NMDA), a-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor, and mGluR are glutamate receptors that aid in synaptic transmission at the spinal cord. DRG, dorsal root ganglion. This image is modified from Finnerup et al, 2021. (Raouf et al, 2010; Finnerup et al, 2021)

From the TRP channels, *TRPV1-4*, *TRPM2-3*, *TRPM8*, and *TRPA1* are known to be expressed in peripheral afferent neurons and associated with nociceptive pathways. Pharmacological studies concerning *TRPV1* and *TRPA1* have demonstrated that inhibition of the channels can ameliorate pain (Basso & Altier, 2017). The VGSCs participate in nociception by regulating the sensory nerve excitation in DRG neurons (Finnerup *et al*, 2021). The hyperpolarization-activated and cyclic nucleotide-gated (HCN) channels found in DRG are also associated with NP. (Jiang *et al*, 2008). Dysfunction of HCN channels is also implicated in the amplification of electrical signals, causing NP(Santello & Nevian, 2015).

Voltage-gated calcium channels are involved in nerve cell excitability and synaptic transmission in nociception. The $Ca_v3.2$ channel that encodes for the *CACNA1H* gene is expressed in DRG neurons and in the spinal cord. Li *et al* demonstrated that after administration of paclitaxel, the expression of the $Ca_v3.2$ channel increased, and it co-expressed with toll-like receptor 4 in DRG neurons and the spinal cord, causing nociceptive sensitization. Thus, the $Ca_v3.2$ channel blockers are anticipated as novel therapeutic agents.

The downregulation of potassium channels, $K_v1.1$ (encoded by *KCNA1* gene) and $K_v1.2$ (encoded by *KCNA2* gene) at the periphery has been linked with cold allodynia (González *et al*, 2017).

Based on all this evidence, ion channels at the periphery and the spinal cord dorsal horn play a significant role in painful PN.

Ion channel and their features

In this section, we discuss the different ion channels that have been associated with pain-related disorders and highlight their physiological features.

Sodium channels

VGSCs are large transmembrane proteins comprised of α-subunits that control the influx of sodium ions through a voltage gating system and β-subunits that mediate membrane transport, inactivation, activation of VGSCs, and ligand binding (Xu *et al*, 2019). The α-subunit consists of four transmembrane domains (D1-4) connected by long intracellular loops, where each domain is comprised of six transmembrane segments (S1-6). The first four transmembrane segments are voltage-sensing domains (S1-4), and the last two (S5-6) are pore domains linked with P-loops (pore loops) and a selectivity filter. The β-subunit includes a single transmembrane helix, an extracellular amino-terminal, and a small intracellular carboxy-terminal, as depicted in Figure 5 (Xu et al, 2019).

There are ten isoforms of α-subunit ($Na_v1.1$ to $Na_v1.9$ and Na_x) and five isoforms of β-subunit (β1, β1B, β2, β3, and β4). $Na_v1.1$–$Na_v1.3$ and $Na_v1.6$ are mainly expressed in the CNS, whereas $Na_v1.6$–$Na_v1.9$ are expressed in DRG neurons. Meanwhile, $Na_v1.4$ and $Na_v1.5$ are expressed in skeletal muscles and cardiac muscles, respectively (De Lera Ruiz & Kraus, 2015).

All the VGSCs partake in creating the sodium ion currents that initiate and transmit action potentials in nerves and muscle fibres (De Lera Ruiz & Kraus, 2015). The VGSCs are regulators of sensory nerve excitation, and so, understanding their physiological mechanisms has provided many insights into pain processing.
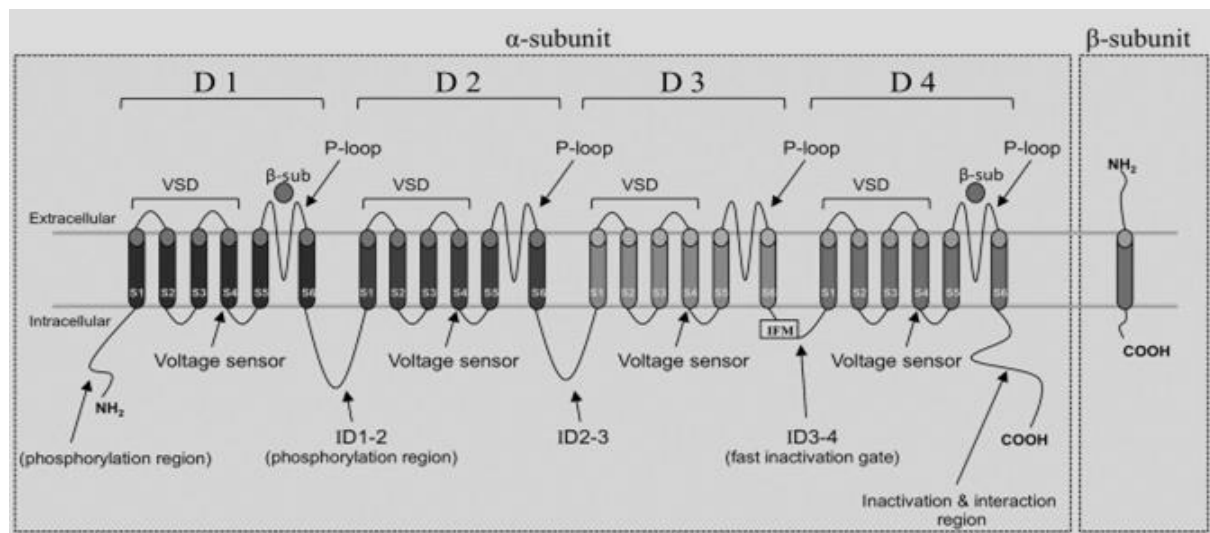
Figure 5: Schematic representation of VGSCs where on the left side, α subunit has four domains as D1-4 where each domain has six transmembrane segments (S1-6). IFM stands for isoleucine, phenylalanine, and methionine motif. The β-sub circle denotes the β-subunit interaction section. VSD is for the voltage-sensing domain. ID1-2, ID2-3, and ID3-4 are interdomain linkers. On the right side is the topology of the β-subunit (Detta et al, 2015).

The mechanism of action of the VGSCs is governed by three stages: i) resting or closed, ii) activated or open, and iii) inactivated or closed. When membrane depolarization occurs, the voltage-sensing domains in VSGCs undergo conformational change and modulate the opening of the gate to activate the channel. Upon activation, they initiate the influx of sodium ions and then further depolarize the membrane, initiating the action potential's rising stage. Within milliseconds, the inactivation of these channels is partly involved in the falling stage of the action potential, resulting in the termination of ion conduction. In the case of continued membrane depolarization or recurrent stimulations, the channel moves into a slow inactivation stage wherein the action potentials are amplified (Xu *et al*, 2019), as illustrated in Figure 6. The S4 transmembrane segments of α-subunits are responsible for the opening and closing of VGSCs (De Lera Ruiz & Kraus, 2015). The VGSCs are typically rapid transient channels that inactivate or activate within milliseconds (De Lera Ruiz & Kraus, 2015).
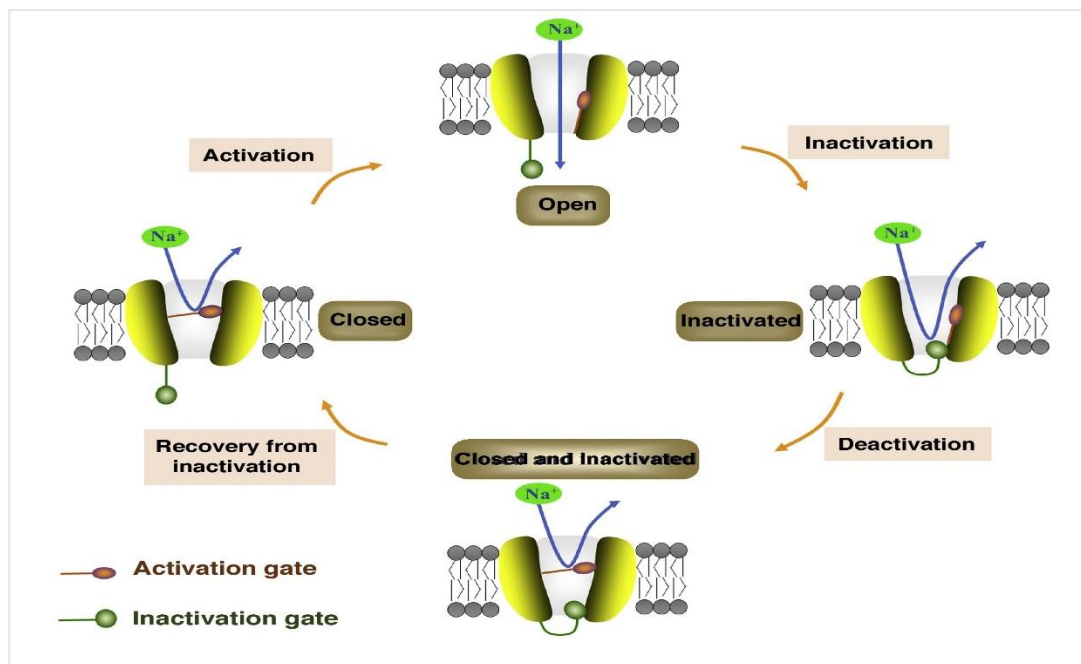
Figure 6: The stages of the mechanism of action of VGSCs adapted from (Xu *et al*, 2019)

The $Na_v1.1$–$Na_v1.9$ are encoded by *SCN1A*, *SCN2A*, *SCN3A, SCN4A*, *SCN5A*, *SCN8A, SCN9A*, *SCN10A*, and *SCN11A* genes, and Na_x is encoded by the *SCN7A* gene. The variants found in *SCN1A*, *SCN2A*, *SCN3A,* and *SCN8A* are responsible for the dysfunctional rapid inactivation of VGSCs, leading to recurrent sodium currents causing ataxia and epilepsy (Eijkelkamp *et al*, 2012). Several studies have associated mutations in *SCN9A* with pain-related conditions that occur due to i) structural change, ii) shift of voltage by dysfunctional voltage-sensing domains, iii) associating with proteins, and iv) generation of resurgent currents (production of a large current by sodium influx) (Lampert *et al*, 2014). The gain-of-function mutation, which modifies the gene product to enhance its function, in *SCN9A* is associated with painful disorders. In contrast, loss-of-function mutations, which modify the gene product to inactivate the function either partially or completely, are associated with pain insensitivity (Lampert *et al*, 2014).

The $Na_v1.8$ channels are resistant to the neurotoxin tetrodotoxin and are important for NP. Compared to the $Na_v1.7$ channel, $Na_v1.8$ gets activated after large depolarized potentials to modulate most of the sodium ion transport in neurons. Mutations in the $Na_v1.8$ channel gene cause hyperexcitability and abnormal firing of DRG neurons, associating them with painful SFN (Faber *et al*, 2012c).

Another channel resistant to tetrodotoxin is Na$_v$1.9 encoded by the *SCN11A* gene. This channel produces only tetrodotoxin-resistant sodium ion current in neurons with extremely slow activation and inactivation kinetics. They create a persistent sodium ion non-inactivating current with a negative activation threshold close to the resting potential. As a result, they aid in controlling the membrane potential and regulating nerve cell excitability (Baker & Nassar, 2020). Due to the nature of the Na$_v$1.9 channel, fast membrane depolarization causes hyperexcitability, whereas prolonged membrane depolarization leads to impaired signal transduction in neurons (Huang *et al*, 2017). Thus, mutations in the *SCN11A* gene have been observed both for pain sensitivity and insensitivity, and usually, the gain-of-function mutations in *SCN11A* are associated with painful SFN (Huang *et al*, 2014). However, Leipold *et al*. demonstrated that a gain-of-function mutation in *SCN11A* blocked further transduction of signals, causing pain insensitivity.

TRP channels

The transient receptor potential channel is a super-family of genes involved in nociception and pain perception (Nilius & Flockerzi, 2014b). All the genes in this super-family are non-selective cation-permeable channels known to conduct calcium ions (Gees *et al*, 2010). Pain-related conditions are linked to human gene mutations in *TRPA1* and *TRPV1* channels.

The name of the Transient Receptor Potential cation channel Ankyrin 1 (*TRPA1*) was given because of the ankyrin repeat domains found in its structure (Paulsen *et al*, 2015). The amino-terminal of the *TRPA1* structure consists of 14-18 ankyrin repeat domains with six transmembrane helical domains (TM1-6), the pore loop located between TM5 and TM6, and the C-terminal side of the protein has a TRP-like domain and a C-terminal coiled-coil (Paulsen *et al*, 2015; Jaquemar *et al*, 1999). *TRPA1* is a well-known chemosensor of nociception as it is activated by several chemical and environmental irritants (Andersson *et al*, 2008). All these irritants activate the channel by causing painful burning sensations in individuals (Kremeyer *et al*, 2010). The *TRPA1* gene is known to be co-expressed with the *TRPV1* gene, which modulates thermosensation, causing both pain and inflammation (Bandell *et al*, 2004).

TRPA1 is expressed in the brain, in skeletal muscles, and in epithelial cells. More importantly, it is expressed in nociceptive neurons such as DRG and trigeminal neurons (Kono *et al*, 2010; Takahashi *et al*, 2008). From functional studies, it is known that the Asn855Ser variant in the *TRPA1* gene is a gain-of-function mutation causing hyperalgesia from cold in patients with familial episodic pain syndrome. (FEPS) (Meents *et al*, 2019; Kremeyer *et al*, 2010).

Szolcsányi & Sándor defined the six transmembrane domains as well as the six ankyrin repeat segments. *TRPV1*, similar to *TRPA1*, has a wide expression pattern as it is expressed in somatic and visceral afferent neurons at both the spinal cord and peripheral junctions (Messeguer *et al*, 2005). *TRPV1* is a polymodal nociceptor that responds to vanilloids, protons, heat (42°C), lipids, voltage, and a variety of other stimuli (Messeguer *et al*, 2005; Julius, 2013; Pingle *et al*, 2007). *TRPV1* is a known potential molecular target as a treatment for various pain-related phenotypes such as hyperalgesia, inflammatory pain, hypersensitivity, and even acute burning pain (Kumar, 2020; Brandt *et al*, 2012). *TRPV1*-null animal models have significantly reduced thermal pain sensitivity in response to inflammation. Moreover, a recent study showed that knockout mice for models of three TRP genes (*TRPV1*, *TRPA1*, and *TRPM3*) experienced nearly no thermal sensation with normal nociception, indicating a combination of TRP channel genes engages in thermal hyperalgesia (Kaneko & Szallasi, 2014; Vandewauw *et al*, 2018). There have been many pharmacological and genetic studies targeting the *TRPV1* gene, but there have been no known rare functional genetic variants reported with a direct association to NP.

PROPANE and PAIN-NET studies

As already mentioned, neuropathic pain is a frequent feature of PN, adversely impacting patients' quality of life and increasing health care costs. Not all individuals with PN develop pain, and it is impossible to identify who is more or less vulnerable among those with similar risk exposure. In this field, a consortium of partners from different countries, including San Raffaele Hospital, coordinated by the IRCCS Foundation 'Carlo Besta' Neurological Institute (CBNI), was engaged in the project PROPANE STUDY ("Probing The Role Of Sodium Channels In Painful Neuropathies") from 2014 to 2017 and more recently in the PAIN-Net project ("Molecule-to-man pain network") from 2018 to 2021, both founded by the European Community. Both projects, with some differences, aim to assess the genetic architecture of PN, achieving a stratification of subjects at high risk of developing painful PN.

The current thesis is being completed as part of the PAIN-Net, a Marie Skłodowska-Curie funded project that evolved from the previous PROPANE STUDY. Many of the genes mentioned above have been tested by PROPANE STUDY consortium members using targeted next-generation sequencing focused on sodium channels and other candidate pain-related genes coding for ion channels and receptors expressed in dorsal root ganglion, in order to define the genetic components involved in the PN and NP. Here, within the PAIN-Net study, thanks to the unbiased approach of Whole Exome Sequencing (WES), the investigation has been enlarged to include all the contributing genes in order to identify new pain-related genes involved in PN and NP.

Next Generation DNA Sequencing

Deoxyribose Nucleic Acid (DNA) is a biomolecule that encompasses an individual's genetic code, which is inherited from parents to their offspring. During the lifetime of a human being, the DNA sequence undergoes many changes in genetic code, known as genetic variation. The genetic variant is referred to a precise location in the genome that differs from the reference genome. An allele is one of two or more gene variants. For each gene, an individual receives two alleles, one from each parent. As a reference, one of the two alleles is called a "reference allele" when it is equal to the one present in the reference genome, while the other is called an "alternate allele" when it is different from the one present in the reference genome.

A coding region of a gene contains all the information necessary to generate a protein, and a variant in a coding region could have different types of effects. In the case of one-base mutations (point mutations), three types of effects are observed as follows: synonymous, nonsense, and missense variants. In a synonymous variant, the point mutation will not alter the amino acid sequence. A nonsense mutation puts a premature stop signal and shortens the protein. Lastly, a missense mutation changes the amino acid sequence, potentially impacting the protein product. The change caused by a missense mutation could be neutral if the replaced amino acid has the same chemical property as the original amino acid or if it occurred in a protein region unimportant to the protein's function. On the other hand, a missense mutation could render a protein non-functional, leading to a disorder (Minde *et al*, 2011).

Some variants are longer than point mutations and are called insertions or deletions. If these variations fall into a coding region and the length of these variations is divisible by three (three base codes for an amino acid), they will cause the deletion or insertion of a complete amino acid in a protein. These types of variants are also called in-frame deletions or insertions. But, if the length is not divisible by three, then a shift in the amino acid sequence of the protein is seen to disrupt the structure of the protein. These types of variants are also called "frameshift variants."

These genetic differences from one individual to another could affect an individual's physical characteristics (or phenotype). By studying these genetic differences, the phenotype of an individual could be better explained and interpreted.

The research community needed a human genome blueprint as a reference to understand and identify genetic variations in a subject, so the researchers established the Human Genome Project. This project used a "Sanger sequencing" approach to sequence the human DNA. In this method, small fragments of DNA are sequenced one by one, making it time-consuming and expensive. As a result, the need to sequence DNA strands in parallel and achieve a high throughput arose, which revolutionised the art of DNA sequencing.(Shendure & Ji, 2008). This technique is known as Next Generation Sequencing (NGS).

Though NGS is cheaper and faster than Sanger sequencing, it lacks quality. In NGS, DNA fragmentation is necessary. These fragments are ligated with an index to match each fragment to the subject and with an adapter that allows the fragment to attach to the surface of the flow cell during sequencing. The output is a sequence that includes the DNA fragment, the index sequence (sample identifier), and the adapter sequences, termed libraries. These libraries are then immobilised by the adapter sequence onto a flow cell, creating clusters where these clusters are amplified simultaneously. Finally, a sequence is identified by its fluorescent labels, as depicted in Figure 7, as each fragment is sequenced multiple times to give many copies of each base in that fragment, defining the quality confidence of the sequencing. If a base in that fragment is sequenced 30 times, that means 30 reads cover a nucleotide, so that 30 is the "coverage" of that nucleotide. As a general rule of thumb, higher coverage of nucleotides results in identifying variants with greater accuracy.
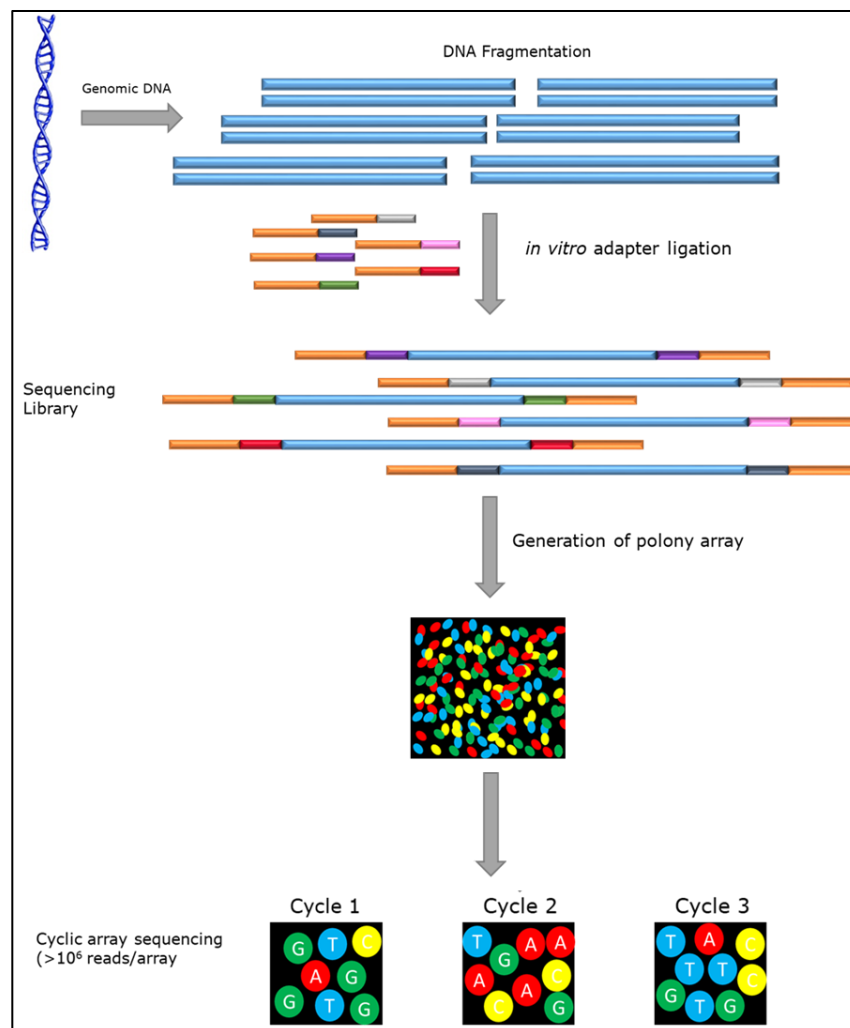
Figure 7: Pictorial representation of the basic protocol of an NGS experiment. (Adapted from Shendure & Ji, 2008)

NGS technologies have several uses, such as the identification of genetic markers involved in a disease, the stratification of subjects by their traits, the investigation of a missing or an over-response to a treatment in a subject, and other possible applications in personalised medicine.

NGS applications are widely used and they are classified into three categories: "Targeted sequencing," "Whole-exome sequencing" (WES), and "Whole-genome sequencing" (WGS). These methods are useful for identifying various types of genomic variation, such as single-nucleotide variations (SNVs) or point mutations, insertions and deletions (INDELs), structural variants (SVs), and copy number variations (CNVs), as depicted in Figure 8.
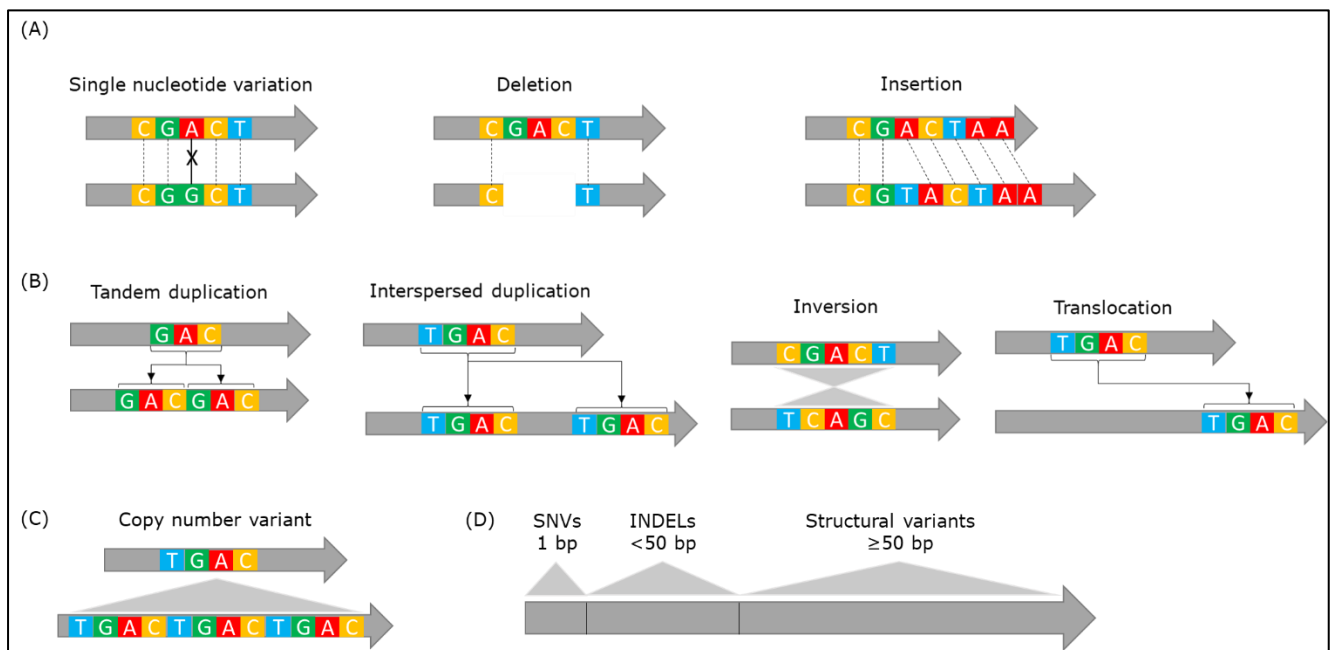
Figure 8: A pictorial representation of different genomic variations identified through NGS. Figure 6 (A) depicts short variations such as SNVs and INDELs. (B): It shows different types of structural variants determined by NGS technology. (C): It represents the copy number variants which is a typical structural variant and (D): It shows how many base pairs (bp) are constitute each type of genomic variation. Adapted from: 'Review: How Long-Read Sequencing Is Revealing Unseen Genomic Variation - PacBio'.

The targeted sequencing is based on a pre-conceived list of genes derived from previous research projects where a genotype was linked to a trait or phenotype for a specific disease. As the sequencing libraries are made to cover this list of genes, targeted sequencing is not recommended if the causative genes are unknown. For example, in PN, which is a multi-factorial and genetically heterogeneous disorder, WES or WGS would be useful in understanding its genetic mechanisms. Also, over the years, sequencing costs for WES and WGS have gone down comparatively, and it is estimated to go down further with more technological advances in this field.

In WGS, the whole three billion nucleotides of the DNA per sample are sequenced, the complete genome. Thus, the major drawback is that it spills out humongous amounts of information, making it hard to detect accurate genetic markers. The processing of large amounts of data also takes time computationally, affecting the turn-around time of a project. WGS is also recommended for diseases where the structural changes in the physical traits of an individual could have a genetic cause. For instance, Charcot-Marie-Tooth neuropathy type 1A (CMT1A) is caused by a CNV of 1.5 megabases in the *PMP22* gene. On the other hand, WES emits comparatively less information, and it is computationally less extensive and manageable. WES can identify such disruptive variations if the SV falls into a coding region. Thus, WES is a good choice among targeted sequencing and WGS (Sun *et al*, 2015).

Different aspects of the three applications have been illustrated in Table 1 (Seleman *et al*, 2017; Sun *et al*, 2015; Foo *et al*, 2012).

	Targeted sequencing	WES	WGS
Number of genes covered	50 – 500	About 2 % part of the genome (~ 30,000)	Complete genome
Gene coverage	100 % similar to Sanger sequencing	98% with 60X depth of coverage	90% with 60X depth of coverage
Number of variants	Depends on the gene panel size	~25,000	~4,000,000
Disease type	Best suited for Mendelian disease or disease for which a well-known genetic cause has been previously established	Disease with genetic heterogeneity or with a genetic cause is unknown or incomplete	Suitable for diseases that are linked with structural defects or complex diseases
Identification of novel genes	No	Yes	Yes
Incidental Findings	No	Yes	Yes
Intronic variant detection (>30bp from splice-sites)	No	No	Yes
CNV variant detection	It depends on the gene panel size and a control cohort of at least 200 samples for the same gene panel	Yes, CNVs > 3 exons are detected, also needs a control cohort from the same sequencing kit	Yes
Structural variant detection	Not recommended	Chances are low	Yes
Sequencing cost per sample	It depends on gene panel size, but for 300 genes, it is estimated to be around $300-500	$950	$1500-2000

Table 1: In-depth comparison of NGS DNA application technologies

Whole exome sequencing

The WES approach focuses on the exonic region of a genome, which aids in screening for rare variants having a strong effect on disease susceptibility (Frésard & Montgomery, 2018). WES is an extension of targeted gene panel sequencing, where all the coding regions are covered rather than just a few regions. One of the main advantages of WES is that, while performing variant analysis, we could initially focus on candidate genes and then proceed to assess the entire coding exons, searching for genotype-phenotype relationships not found in candidate genes. WES has some limitations, such as the sequencing coverage could vary across regions or samples, GC-rich regions are not captured qualitatively, and CNVs are not easy to detect because they are composed of intronic regions. Regardless, WES has been used successfully in determining the genetic signatures corresponding to different types of neurological disorders, especially inherited PN such as CMT disease (Foo *et al*, 2012).

The WES workflow consists of four stages, from sample selection to variant interpretation, as shown in Figures 9a and 9b below. The first two stages of WES consist of the laboratory procedure, also known as the "wet-part," whereas the latter two stages are performed using high- computing workstations as the "dry-part."
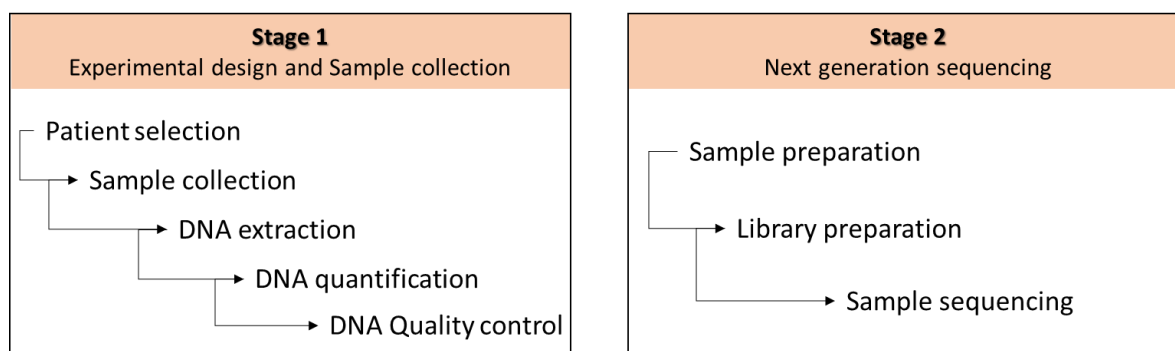
Figure 9a: Depicts first two stages of WES experiments. Adapted from (NGS Workflow Steps | Illumina sequencing workflow)

Stage 1: Experimental design and sample collection

The primary step is to understand the hypothesis and requirements for WES experiments. The clinician selects the patient for the study following specific inclusion/exclusion criteria and collects blood or saliva samples from the recruited subject. After sample collection, extraction of high-quality DNA is performed.

Stage 2: Next-generation sequencing

The exome library is prepared using the extracted DNA. At first, the DNA is fragmented either by a biological enzyme or by a mechanical process called sonication. These fragments are then ligated with adapters and index sequences. These adapters aid these fragments to bind to the flow cell, whereas the index sequences help identify the patient sample. These samples are prepared for amplification on a pre-determined probe set that would hybridise DNA fragments of the sample and amplify the exonic regions. After exon enrichment, the library sequences are amplified to generate millions of short reads. If forward and reverse strands of DNA are sequenced, the sequencing is called paired-end. Otherwise, if only one strand is sequenced, the sequencing is called single-end. The first type of sequencing attains more information and provides better alignment accuracy during data analysis (Seaby *et al*, 2016).

Because NGS technologies aid in the discovery of putative variants and the linking of genotypes to phenotypes, they have the potential to be used as a diagnostic tool in both research and clinical settings. From Figure 9b, data management and data analysis are completed with the help of bioinformatics, a crucial part of WES data analysis.
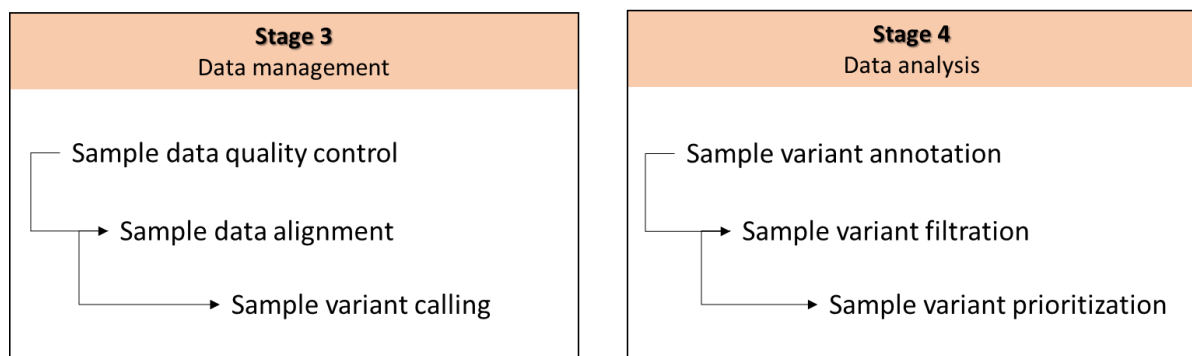
| Stage 3 Data management	Stage 4 Data analysis
Sample data quality control	Sample variant annotation
→ Sample data alignment	→ Sample variant filtration
→ Sample variant calling	→ Sample variant prioritization

Figure 9b: Depicts last two stages of WES experiments. Adapted from (NGS Workflow Steps | Illumina sequencing workflow)

Stage 3: Data Management

Quality control

Quality control is the first phase of the WES data analysis workflow, where the reads obtained per sample are evaluated qualitatively. The sequencing reads obtained from the sequencer are stored in base call files. These base call files are converted to files where each base in a read is given a score known as the PHRED score. The PHRED score, which assigns a probability of error to each called base in a read, supplements the Sanger-sequencing quality score (Ewing *et al*, 1998; Ewing & Green, 1998). The PHRED score is then used to weed out low-quality reads from further processing. The quality check of the reads is also performed after adapter removal to ensure the retention of the DNA insert and proper removal. This step is performed by the FastQC tool developed by Simon Andrews of Babraham Bioinformatics (Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data).

Read Alignment

The quality reads from the adapter trimming step are then aligned to the reference genome to localise the reads chromosome-wise in the genome. In the case of WES analysis, an aligner should be able to map at least 96% of the reads to the genome for the successful detection of variants in the exonic region. Several aligners can be used for this step, but the preferred one for the human reference genome is the Burrows-wheeler aligner (MEM algorithm) (Li, 2013). Usually, an aligner is chosen based on the output file that is emitted from the tool, as this file should be compatible for downstream analysis. The preferred output file format is either Sequence Alignment Map (SAM) format or Binary Alignment Map (BAM) format. Each sample data set is aligned to the human reference genome. The expected alignment score should be more than 99%. After the sample alignment step, the mean number of reads that cover the target regions (i.e., coverage) is calculated for each sample. The expected coverage per sample should be greater than 75x. The last step of data management is to perform variant calling, where the aligned reads are investigated and the difference between reference and sample reads is registered in a file.

Variant discovery

The alignment outputs are then processed by the Genome Analysis Tool Kit (GATK), which is an assemblage of several tools for recalibrating and calling variants for NGS technologies. The GATK's best practises workflow has helped standardise variant discovery in high-throughput sequencing analysis (Van der Auwera *et al*, 2013). Figure 10 describes the complex collective workflow that is implemented in WES analysis for variant discovery (Poplin *et al*, 2018).
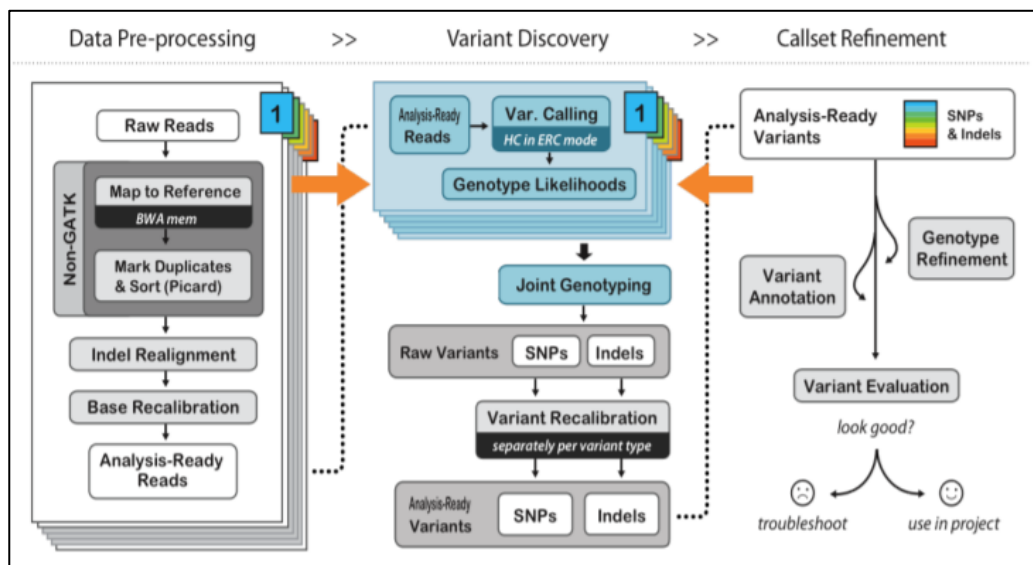
Figure 10: A schematic of GATK's best practices workflow where the joint genotyping aspect (highlighted in blue) is only used when more than 40 samples are analysed together. (Poplin et al, 2018)

For GATK, the resultant file has only small variants, less than 50 bp, but for SV discovery, another workflow must be adopted. Therefore, for SV discovery, it is recommended to use a combination of multiple variant callers and report the intersection of the variants seen amongst the chosen SV detection tools (Chapman, 2014; Pabinger *et al*, 2014). Also, these intersected SVs should be compared against a similar number of healthy control samples, and only the rare ones should be reported. Only CNVs from all structural variants can be confidently reported in a WES analysis; if other SVs are seen in WES analysis, they have to be further evaluated by functional studies.

After the sample variant discovery step, a quality check is performed where the ratio of transitions to transversions (Ti/Tv) is calculated. To elaborate, the structure of DNA is made up of two-ring purine-based nucleotides, Adenine (A) and Guanine (G), and one-ring pyrimidine-based nucleotides, Cytosine (C) and Thymine (T).

Transitions are mutations that occur between purine-derived nucleotides (A <-> G and C <-> T) or between pyrimidine-based nucleotides (C <-> T). Transversions are mutations that happen between purine-based nucleotides and pyrimidine-based nucleotides (A <-> C, A <-> T, C <-> G, G <-> T). In theory, if these mutations occur at random, the Ti/Tv ratio should be 0.5, meaning that in that position, any variant is present. The Ti/Tv ratio for an individual WES sample should range between 2.2 and 3.3.(Wang *et al*, 2017; Carson *et al*, 2014).

Stage 4: Data Analysis

In this step, the variant call format file for each sample is annotated by annotation tools. The variants are filtered according to different functional prediction tools. These filtered variants are then prioritised according to the individual's qualitative trait. While WES can be a robust tool, it also emits between 25,000 and 45,000 variants per sample (Foo *et al*, 2012). Therefore, proper bioinformatics support is needed for an apt variant interpretation for each sample.

Variant Annotation

In this step, each called variant is given a functional meaning by adding details about the location of the mutation in the gene, and the impact of that mutation on the gene's function is reported. Thanks to the sequence ontology database, it was possible to assign a feature to the variant even if one uses different annotation tools, as depicted in Figure 11 (Eilbeck *et al*, 2005).
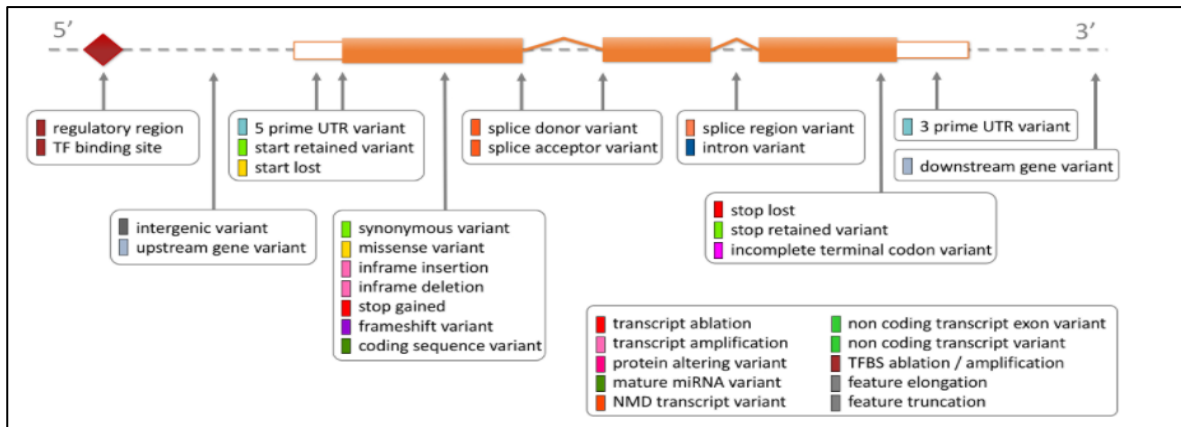
Figure 11: A diagram representing the different sequence ontology terms for the transcript structure. Extracted from EMBL-EBI. (Calculated consequences)

Also, other variant descriptors, such as providing the mutation's frequency in the global population and the conservation of that variant through different species are added to each variant in the Variant Call Format (VCF) file.

The two most common annotators used in WES are Variant Effect Predictor (VEP) by Ensembl and SnpEff annotator (McLaren *et al*, 2016; Cingolani *et al*, 2012). Following the annotator's assignment of a variety of information from various sources to each variant, the most important step is to select putative variants.

Variant prioritization

The variant discovery step outputs about 25,000 variants per sample. The variants are then prioritized in a multi-step process. The variants are filtered for a specific allele frequency in the global population and predicted variant impact. Henceforth, this subset of genetic variants will be referred to as qualifying variants (Petrovski *et al*, 2017). Richards *et al* acknowledged these conventional filtering criteria and included criteria for pathogenicity with biological support rather than relying solely on computational predictors. They advocated using five terms to classify qualifying variants, as shown in Table 2. Each variant must be manually examined for a specific individual while providing biological reasoning for its selection is a rate-limiting step. Therefore, certain strategies are applied with different modes of inheritance when performing familial segregation analysis or analysing sporadic cases.

In familial segregation analysis, the family members are tested to see if a variant segregates with the disease and to establish the gene-disease relationship. For example, in trio families (parents are unaffected and the offspring is affected), for a recessive disorder, both the parents are assumed to be heterozygous carriers, and the affected offspring are homozygous for the variant. Similarly, for a dominant disease model, the affected subject should be heterozygous for the variant, while the unaffected members should not share the heterozygous variant. These disease models inherently look for causal variants but not protective variants. An additional strategy can be applied to get information on protective variants. In the protective autosomal dominant model, the unaffected members of the family should be heterozygous for the variant, and the affected members should be homozygous for the variant.

For sporadic samples, due to a lack of information about the mode of inheritance, each sample is examined on a case-by-case basis. If more than one sporadic sample is available, then a "sharing" model is applied where the variant is prioritised if it is shared among more than one sample but less than 25% frequent in the analysed cohort. The qualifying variants are manually examined and ranked according to their traits, and then assigned the terminology as shown in Table 2.

The synonymous variants or variants that do not impact a protein's amino acid sequence are classified as "Benign." The missense variants with substitutions that have a neutral impact on the amino acid sequence or the altered amino acid that is located in a region that does not impact the protein's functionality are classified as "Likely benign." Suppose the variant (an SNV or an INDEL) is found in a gene that has been functionally associated with the phenotype and has been reported to segregate in a family conclusively with a phenotype similar to the analysed subject. Then, in that case, that variant is classified as "Likely pathogenic." When several studies report a variant and confirm its pathogenicity with a phenotype similar to the analysed subject, that variant is classified as "Pathogenic."

But, if the variant is associated with the phenotype without any functional study or familial segregation study, the variant will be classified as a "Variant of Uncertain Significance" (VUS). Moreover, the variants classified as "Pathogenic" or "Likely pathogenic" in the literature for a specific disease can differ from the analysed subject's trait.

Classification status	Definition
Benign/Likely benign	Variants that are not conserved or are too frequent in the population or the control samples Also, the protein change is tolerated in the gene.
Likely pathogenic	Variants with some clinical significance, at least one functional study
Pathogenic	Variants with strong clinical significance Variants recorded at multiple sources with functional studies
Variants of Uncertain significance (VUS)	Variants in a gene linked to the disorder or discovered in a related pathway, but no functional studies

Table 2: Variant classification terminology provided by ACMG guidelines and its definition

The prioritised variants are then further validated with Sanger sequencing. A functional study can be considered if the discovered variants have a significant impact or are shared by multiple individuals with the same phenotype.

Aim of the work

The aim of this thesis is to identify new genes and genetic variants possibly involved in peripheral neuropathy and neuropathic pain, in order to better elucidate the genetic factors involved in the disease and to investigate the genetic architecture that could explain the biological mechanisms associated with neuropathic pain.

This was accomplished by performing whole exome sequencing experiments in twelve families, comprised of 34 affected and 19 unaffected members, affected by peripheral neuropathy or small fiber neuropathy, and in 42 sporadic subjects affected by the same diseases and selected because of the early onset of the clinical manifestations.

To achieve the specific aim, a genotype to phenotype association at the genetic variant level was performed, evaluating the sharing of variants between familial and sporadic cases and also grouping the variants according to the family of genes known to be involved in pain. Moreover, a pathway analysis at the gene level was performed separately in familial and sporadic cases, to evaluate which are the processes that appear to be dysregulated and that could explain the onset of peripheral neuropathy and neuropathic pain

Results

In this study, subjects were recruited through two different projects as mentioned in the Introduction: PROPANE and PAIN-Net. The individuals were recruited in the PROPANE project between 2014 and 2017, whereas the participants were recruited in the PAIN-Net study between 2018 and 2020. All the families recruited under PROPANE project will have a prefix 'PROP' followed by a code whereas all the families from PAIN-Net project will have a prefix 'PNET' followed by a code. In case of sporadic cases, the samples will be denoted as 'PROPEO' followed by a code for PROPANE project and 'PNETEO' followed by a code for PAIN-Net.

Virtual gene panel creation

It is known that the number of raw exonic variants obtained using a WES approach could be very high, being around 25,000 variants per sample. Even after applying a quality check on variants and specific filter for coverage depth and frequency, about 2,000 to 5,000 variants remain for an individual sample. In WES data analysis for familial samples, where segregation analysis can be used, the number of qualifying variants is fewer than in sporadic cases. For example, in a familial segregation analysis, variant count ranges from 100 to 1000 variants while, for a sporadic case it is over 1000 variants per sample. The interpretation of each qualifying variant is a long and detail-oriented process because several aspects have to be considered, such as the association to the phenotype and the impact of the variant on the protein function. In order to increase the efficiency of variant detection and to focus the attention on genes with a greater likelihood of being associated with the phenotype, a phenotype-driven gene panel was created based on several criteria in a multi-step procedure. The creation of this gene panel would also enable us to choose variants in genes that are predisposed towards the disease.

An initial list of 762 genes was manually curated by selecting them according to specific keywords and phenotypes, as reported in Methods section, and according to the information retrieved from the following sources: i) 94 genes were selected from International Association for Study of Pain (Pain Gene Resource | Pain Research Forum); ii) 192 genes were taken from Human Pain Genetics database where the phenotype description was related to "Pain" and not to "Migraine", "Cancer pain" and "Temporomandibular disorders" (Meloto *et al*, 2018); iii) 476 genes associated with peripheral neuropathy, neuropathic pain and associated symptoms were extracted from Human Phenotype Ontology (HPO) database and Online Mendelian Inheritance in Man (OMIM). (Köhler *et al*, 2021; Amberger & Hamosh, 2017).

At this point, the list was further processed to retain the genes recognized by OMIM database. A final list of 592 genes was curated and used as the 'candidate gene' filter for variant analysis, meaning that variants located in those genes were selected for further evaluation. This panel of genes can be found in Appendix A.3.

In case the application of 'candidate gene' filter did not give new insights in the familial and sporadic subjects selected in this thesis, an 'open exome' approach was adopted considering the variants located in all the genes captured by the WES.

Assessment of WES workflow

In this project, the samples retrieved from the combination of the PROPANE and PAIN-Net studied were evaluated. The individuals belonging to the families were grouped together for assessment, whereas the individuals in sporadic cohort were grouped by the project in this section.

For precise variant selection, specific quality control assessments were done in the WES workflow. The workflow was evaluated by two main aspects: coverage of the target regions per sample and transition-transversion (Ti/Tv) ratio to determine the variant quality.

Since two separate target capture kits, Sure Select V5 and Sure Select V7, were utilized in our investigation, and while the two kits are equivalent in terms of covered exons, they have a minor discordant percentage of covered regions. As a result, a coverage per sample in relation to these kits was requested.

The average coverage of the families with Sure Select V5 as target region were as follows: for PROPANE study was 66.2X ± 17.5 (mean ± standard deviation) with range of 20X to 104X and for PAIN-Net study was 91.3 ± 19.4 ranging from 58.5X to 120X. Due to the commercial non-availability of Sure Select V5 probes, for the last two families recruited in the PAIN-Net study (PNET05 and PNET06) Sure Select V7 target capture kit was used. The coverage of these families was 154X ± 25.6 (mean ± standard deviation) with a range of 86X to 180X as shown in Figure 12a. These last two PAIN-Net families, sequenced with the newest target-enrichment kit by Agilent Technologies, achieved higher coverage than the other samples. Due to this difference, these samples were analysed separately to minimize the bias in the variant calling.
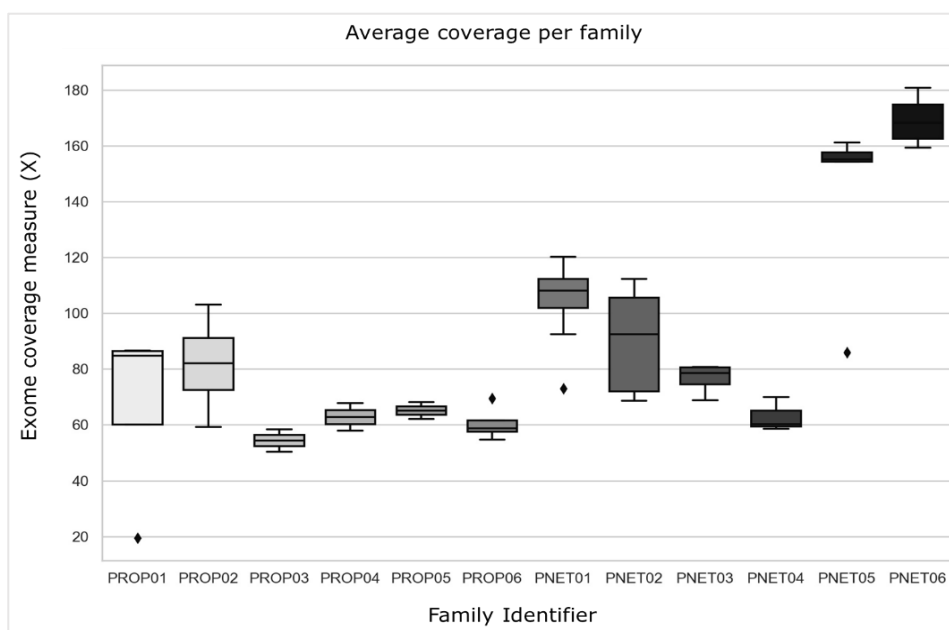
Figure 12a: Average coverage depth per family from PROPANE and PAIN-Net study. Box plots show the minimum, 25% percentile, median, 75% percentile and maximum of average coverage depth covered

Only one sample from a PROPANE study (family PROP01, subject I.2) had WES coverage less than 75X coverage depth (acceptable coverage for WES samples). For this sample, it was observed that not only coding but also non-coding regions were covered. Several hypotheses were discussed with the manufacturer of the target-enrichment kit but without any outcome. The Ti/Tv ratio for this sample was greater than 2.0, determining the variant quality was within acceptable range. The low coverage of this sample did not affect the downstream analysis. Therefore, the sample was retained for further process.

In comparison to families, the average coverage of the sporadic cohort of PROPANE and PAIN-Net studies were 58.1X ± 10.4 (mean ± standard deviation) ranging between 35X to 80.8X as shown in Figure 12b. Since the subjects with coverage less than 75X had Ti/Tv ratio greater than 2.0, they were retained for downstream analysis.
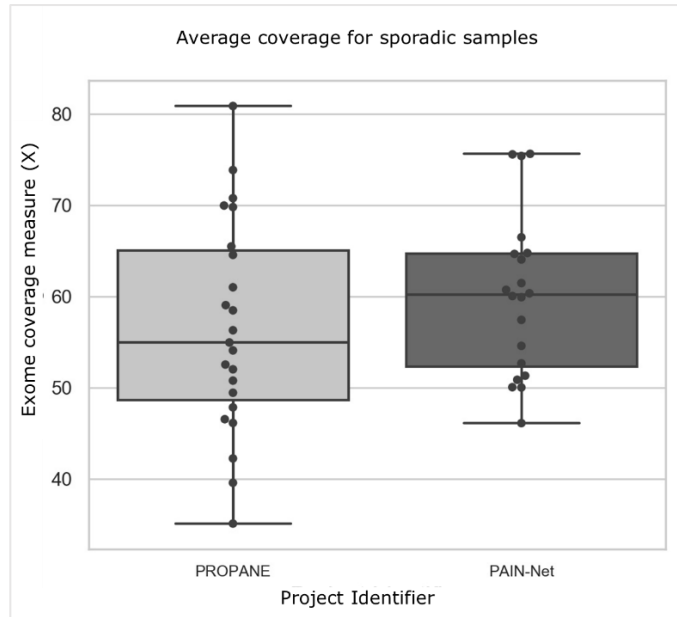
Figure 12b: Average coverage for each sample in sporadic cohort from PROPANE and PAIN-Net study. Box plots show the minimum, 25% percentile, median, 75% percentile and maximum of average coverage depth covered. Black dots denote an individual sample

The consistency of target region coverage in our samples was assessed Because the depth required for accurate variant selection should be larger than 25X (Lelieveld *et al*, 2015), the percentage of target bases covered at the 30X threshold was chosen, giving the requisite consistency of coverage across all samples in our investigation. The average percentage of target bases covered for familial cohort was 82.9 % at 30X whereas for sporadic cohort was 74.8 % at 30X, as shown in Figure 13.
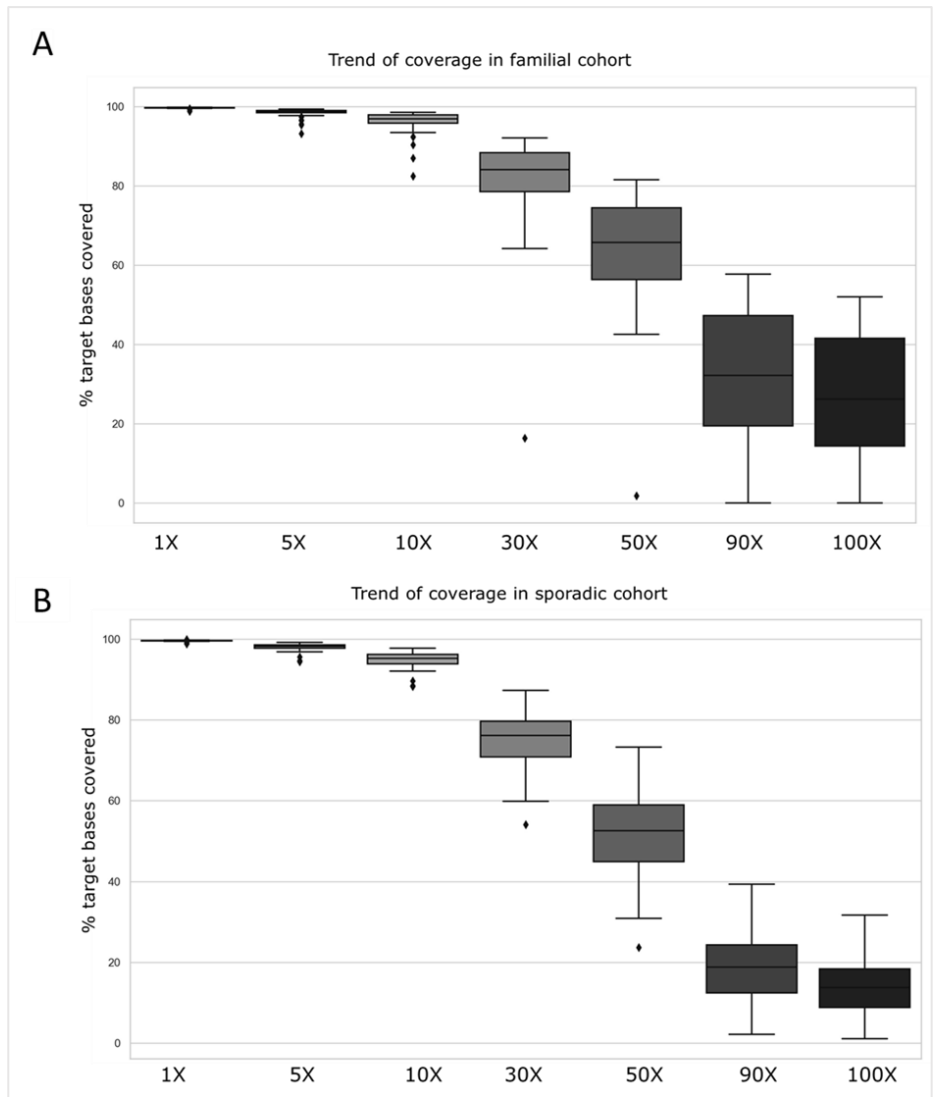
Figure 13: Box plots showing the trend of target bases covered across familial (A) and sporadic (B) cohorts. Box plots show the minimum, 25% percentile, median, 75% percentile and maximum percentage of target bases covered.

To complete the assessment step of WES data, the variant quality through Ti/Tv ratio was determined. For familial and sporadic cohorts the average Ti/Tv ratio was 2.41, as reported in Figure 14.
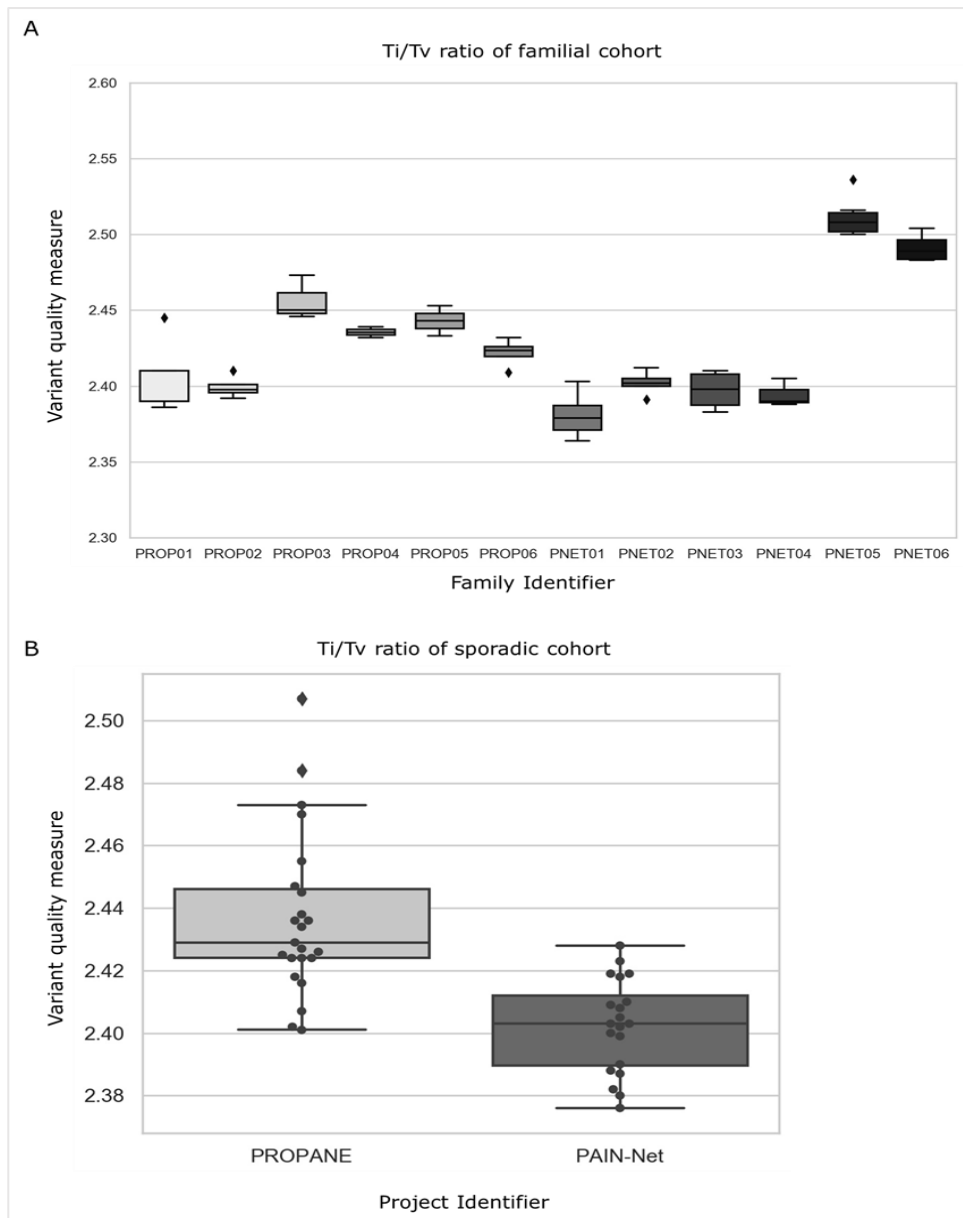
Figure 14: A) Box plot showing Ti/Tv ratio per family with variant quality ranging from 2.35 to 2.55 and B) Box plot showing Ti/Tv ratio per sample in sporadic cohort by PROPANE and PAIN-Net studies with variant quality ranging from 2.37 to 2.53 where the black dot denotes each individual sample. Box plots show the minimum, 25% percentile, median, 75% percentile and maximum percentage of Ti/Tv ratio.

Cohort description and variant prioritization

In this section, we will discuss different aspects of our cohort and explain how the variant was prioritized within a family and provide a biological reasoning for its selection in that family. For sporadic cohort, the prioritized variants along with their classification have been reported.

Familial Cohort

Six families from PROPANE project and six families recruited under PAIN-Net project were selected for the present study. The families that did not exhibit a type of pain typically seen in patients with PN/SFN were not included in the present study. The features of the familial cohort, the number of subjects, the gender and the number of affected and unaffected individuals in each family is reported in Table 3.

Family identifier	PN	NP	Itch	All individuals	Affected		Unaffected	
					Female	*Male*	*Female*	*Male*
PROP01	√	√	√	5	3	0	1	1
PROP02	√	√	√	4	1	2	1	0
PROP03	√	√		3	1	0	2	0
PROP04	√	√		2	1	0	1	0
PROP05	√	√		2	1	1	0	0
PROP06	√	√	√	4	1	1	1	1
PNET01	√	√		11	5	4	2	0
PNET02	√	√	√	5	2	1	1	1
PNET03		√	√	4	1	1	0	2
PNET04	√	√		3	1	0	1	1
PNET05	√	√		6	4	0	0	2
PNET06	√	√	√	4	3	0	0	1
			Total	53	24	10	10	9

Table 3: Clinical characteristics of the familial cohort recruited in the study along with the phenotypes observed in each family. The presence of the phenotype in a family is represented with √ mark.

This family was composed of two generations where three individuals were affected by paroxysmal itch followed by NP. There were three affected members (I.2, II.1 and II.2) and two unaffected members (I.1 and II.3) in this family, as shown in Figure 15. The affected family members were diagnosed to have SFN as their nerve biopsy showed considerable reduction in IENF density indicative of small fiber involvement.
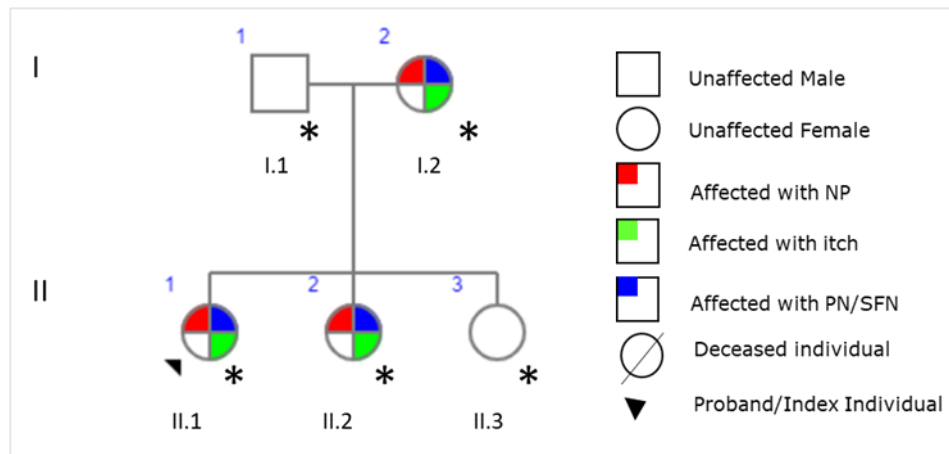
Figure 15: Family pedigree of PROP01 family. The proband (II.1) is denoted by a black arrow. Individuals marked with an asterisk were selected for WES experiments.

The 38 year old proband (II.1) from the age of 15 has been suffering from daily itch attacks at the trunk area and upper limbs. These itch attacks were stimulated by heat, spicy food and hot drinks. These attacks would last anywhere from 2 to 4 hours with a frequency of 3 to 4 times in a day along with episodic flushing of the upper body. But, at the age of 34, she started having burning pain after the itch attacks decreasing her quality of life severely. She also had other autonomic complaints such as sweating, hot flushing, and orthostatic dizziness (Devigili *et al*, 2014a). The 66 year old individual (I.2), mother of the proband, also complained of itch attacks by same triggers at the same region as the proband but her attacks started from the age of 54. These itch attacks went on for about 2 hours and their frequency was 5 times in a month. The mother also underwent cardiac ablation at the age of 55 years due to arrhythmia of unknown cause. Following similar pattern as the proband, after 5 years of itch attacks she started feeling burning pain after episodic flushing of upper body (Devigili *et al*, 2014a). The 43 year old proband's sister (II.2) followed similar pattern as their mother but from the age of 39 years. Her itch attacks also lasted for 2 hours but their frequency was 3 per month. Burning pain and episodic flushing appeared after 3 years, a little earlier compared to her mother.

Histamine mediated itch was ruled out as the test came negative for all three affected members. No other cause was identified for itch attacks and consecutive burning pain. All three affected members were treated with Pregabalin 300 mg daily, with a good response. This family has been studied intensively and has been previously published in the following articles: Devigili *et al*, 2014 and Martinelli-Boneschi *et al*, 2017.

The two genetic mutations that have been published for this family are: *SCN9A*:p.Ile739Val and *COL6A5*:p.Arg2162Ser. The mutation in the *SCN9A* gene was associated with painful SFN and the mutation in *COL6A5* to the neuropathic itch phenotype. A functional analysis was carried out on the *SCN9A* variant and it was confirmed that it aids in the hyperexcitability of DRG neurons, inducing NP (Han *et al*, 2012a). As regards to the *COL6A5* variant, the expression pattern and the RNA mediated-decay profile of the gene across patients and controls have been performed to correlate the variant with neuropathic itch. Thanks to the previous publications, the re-analysis of PROP01 for the present thesis is a proof of concept of the applied multi-step bioinformatics pipeline adopted for data analysis, which revealed to be reliable and reproducible in the detection of the functionally validated variants. Taking into account that the applied bioinformatics workflow has been updated according to the revised population database and clinical variant database, we wanted to reanalyse this family also to investigate if new findings were revealed.

In this family, we considered four different types of segregation models with respect to SFN, NP and itch, evaluating both the causative and protective segregation. These models have been further described in Table 4a, in which, along with the model used, the number of variants retrieved at three different Minor Allele Frequency (MAF), less than 10%, 5% and 1% obtained from the GnomAD database of the non-Finnish European population, was reported *(Karczewski* et al*, 2019)*.

Model type: SFN, NP, itch	I.1	I.2	II.1	II.2	II.3	10%	5%	1%
Causative	Ref	Het	Het	Het	Ref	4	4	4
Causative	Het	Hom	Hom	Hom	Het	0	0	0
Protective	Het	Ref	Ref	Ref	Het	2	2	2
Protective	Hom	Het	Het	Het	Hom	0	0	0

Table 4a: Table showing the causative/protective genetic segregation models adapted for PROP01 family with respect to SFN, NP and itch. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) after applying the filtration criteria mentioned in Methods section. Individuals highlighted in red are affected members, whereas unaffected members are in green. The Ref stands for Reference variant, Het stands for Heterozygous variant and Hom stands for Homozygous variant.

All the qualifying filtered variants were evaluated one by one and classified into a class of prediction status according to the criteria shown in the Method section in order to associate a possible effect on the protein function. The prioritized qualifying variants from both causative and protective models for this family have been listed in Table 5b. In conclusion, the published genetic variants were recaptured; however, due to the new guidelines given by ACMG, the *SCN9A* variant was not considered as a 'Pathogenic' variant but as a 'Variant of uncertain significance' (VUS) owing to the increase in the number of homozygotes for this allele in the global population. Also, a paper by Eijkenboom *et al* demonstrated that two families with this variant segregated inconclusively with respect to SFN; they concluded that *SCN9A*:p.Ile739Val variant is a VUS/risk factor for painful SFN.

Regarding the published variant in *COL6A5* genes, we recaptured it and classified it as VUS.

We also found a variant in the *BBS1* gene conclusively segregating in this family. This particular variant has been previously associated with 'Bardet-Biedl syndrome 1' that present in an autosomal recessive inheritance pattern (Mykytyn *et al*, 2002). This variant was selected because *BBS1*-null and *BBS4*-null mice are known to affect peripheral sensory function, and this variant, in its homozygous state, affects the ependymal cell cilia function (ventriculomegaly) (Tan *et al*, 2007; Davis *et al*, 2007).

Table 4b summarizes the selected variants for PROP01. The variants in *SCN9A, COL6A5, BBS1,* and *AGL* come from the causative model, while the variants in *MYT1L* and *HCRTR2* are from protective segregation.

Prioritized qualifying variants	Classification status	I. 2	II. 1	II. 2	I. 1	II. 3	Causative	Protective
SCN9A:p.Ile739Val	VUS (Risk factor)	√	√	√			√	
COL6A5:p.Arg2162Ser	VUS	√	√	√			√	
BBS1:p.Met390Arg	VUS	√	√	√			√	
AGL:p.Ala750Ser	VUS	√	√	√			√	
MYT1L:p.Pro496Ser	VUS				√	√		√
HCRTR2:p.Arg248Cys	VUS				√	√		√

Table 4b: Table of prioritized qualifying variants for PROP01 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

This family was comprised of three generations with three affected members (I.2, II.1 and III.1) and one unaffected member (I.3), as shown in Figure 16.
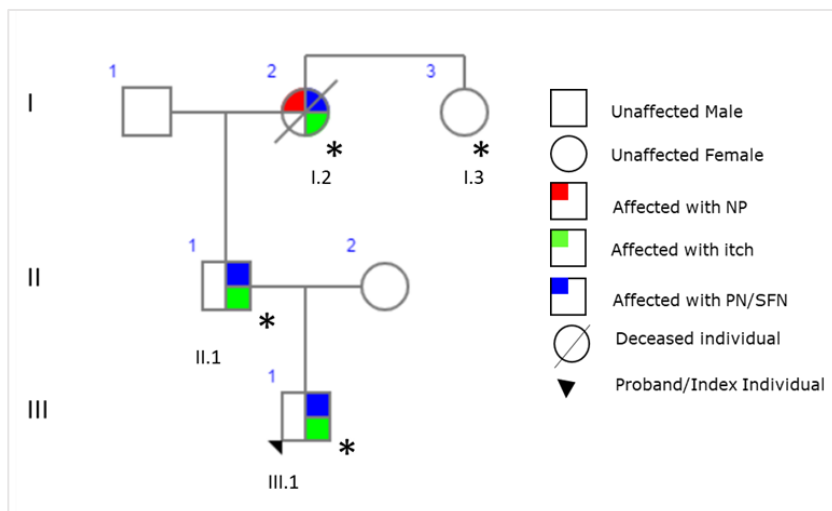
Figure 16: Pedigree of the PROP02 family. The proband (III.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

The proband (III.1) at 33 years of age complained of prolonged chronic itch for approximately 5 years. The itch attacks manifested daily in the upper limbs, shoulder and back and lasted for 30-60 minutes. These attacks were followed by episodic flushing for about an hour. The proband did not have any other sensory or autonomic complaints, and neither could he pinpoint the triggers for the itch attacks. Over the years, the proband had learned to exercise self-control to scratch his body during mild itch attacks. Similarly, his 65 year old father (II.1) had itch attacks in the same manner from the age of 30. Although, the itch attacks had intensified over the last few years. Even the father could not identify any triggers for the itch attacks. The paternal grandmother of the proband (I.2) also complained of itch attacks, and her age at onset was 40 years with subsequent pain. All three of them were diagnosed for PN and itch, with the addition of NP in proband's grandmother. Shortly after the diagnosis was made, the grandmother (I.2) passed away. The only unaffected member in this family is the sister (I.3) of the paternal grandmother.

Other than what is previously mentioned, the medical history of all the members was unremarkable. Nerve biopsies from the proband and father depicted loss in IENF density, whereas a normal value was observed in the grandmother. All the three affected members were treated with gabapentin 900 mg daily with relief. This family has also been previously published with two heterozygous genetic mutations in *COL6A5* corresponding to itch phenotype(Martinelli-Boneschi *et al*, 2017).

According to Martinelli-Boneschi *et al.*, *COL6A5*:p.Glu2272Ter was observed segregating conclusively with respect to itch in two families. In 2017, this variant was considered a rare variant, but now it has been reclassified as a polymorphism due to the high number of homozygotes seen in the global population. Therefore, this family was selected to see if the reanalysis would reveal any new genetic mutations pertaining to the PN and itch phenotypes.

In this family, segregation models were performed in two ways: i) genetic segregation with respect to PN and itch, and ii) genetic segregation with respect to NP. The application of the 'candidate gene' filter did not give new insights for this family. Therefore, the models were reapplied with an 'open exome' approach. The results of these transmission models with 'open exome' are reported in Table 5a and 5b.

Model type: PN, itch (open exome)	I.3	I.2	II.1	III.1	10%	5%	1%
Causative	Ref	Het	Het	Het	61	43	20
Causative	Het	Hom	Hom	Hom	0	0	0
Causative	Ref	Hom	Hom	Hom	0	0	0
Protective	Het	Ref	Ref	Ref	89	84	0
Protective	Hom	Het	Het	Het	0	0	0

Table 5a: Table showing the causative/protective genetic segregation model in 'candidate gene' for PROP02 family with respect to PN and itch. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Model type: PN (open exome)	I.3	I.2	II.1	III.1	10%	5%	1%
Causative	Ref	Het	Ref	Ref	134	98	42
Causative	Het	Hom	Het	Het	0	0	0
Causative	Ref	Hom	Ref	Ref	0	0	0
Protective	Het	Ref	Het	Het	5	0	0
Protective	Hom	Het	Hom	Hom	0	0	0

Table 5b: Table showing the causative/protective genetic segregation model at "open exome" level for PROP02 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

A list of qualifying variants was identified from causative/protective models for PROP02 and have been listed in Table 5c. The variants in *KRT16* and *KCNJ12* genes were causative for SFN and itch phenotypes. The variants in *TRPM8* and *CACNA1E* genes were protective for pain, whereas variants in *ATAD3A* and *PRDM16* were causative for pain.

Prioritized qualifying variants	Classification status	I. 2	II. 1	III. 1	I. 3	Causative	Protective
COL6A5:p.Glu2272Ter	Benign (COL6A5 Polymorphism)	√	√	√		NA	NA
KRT16:p.Thr215Ala	VUS	√	√	√		√	
KCNJ12:p.Glu380Lys	VUS	√	√	√		√	
TRPM8:p.Lys599Asn	VUS		√	√			√
TRPM8:p.Pro949Ser	VUS		√	√			√
CACNA1E:p.Val58Ile	VUS		√				√
ATAD3A:p.Leu77Val	VUS	√				√	
PRDM16:p.Arg823Cys	VUS	√				√	

Table 5c: Table showing the prioritised list of qualifying variants for the PROP02 family. The affected member of the family have been listed before and the unaffected member at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

PROP03

The PROP03 family is comprised of two generations with affected members (I.1, II.1 and II.2) for SFN, but only proband was diagnosed with NP, as shown in Figure 17.

Figure 17: Pedigree of the family PROP03. The proband (I.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

All of them were diagnosed with SFN, but their pain profiles were variable. The 70 year old proband experienced NP, but his children (II.1 and II.2, respectively 47 and 44 years old) did not experience pain. During the examination of variants, we noticed peculiar inconclusive segregation of *SCN9A* mutations. The *SCN9A* mutations were observed in the proband (I.1) and his son (II.2) with painless SFN. These mutations have been functionally proven to be responsible for painful SFN (Faber *et al*, 2012a; Zeberg *et al*, 2020), and their distribution in this family is shown in Table 6.

SCN9A variants	I.1	II.1	II.2
SCN9A:p.Met932Leu/ *SCN9A*:p.Val991Leu	√	√	-
SCN9A:p.Asp1908Gly	-	√	-

Table 6: Tabulated list of published SCN9A mutations seen in PROP03 known to be associated with SFN and NP (Faber et al, 2012a; Zeberg et al, 2020).

The first two *SCN9A* variants are common and are in *cis*; that is, the presence of these variants in combination impacts the gene, causing slow inactivation of the channel while increasing the pain stimuli (Faber *et al*, 2012). While the presence of all three *SCN9A* mutations increases the risk of NP, as demonstrated by Zeberg *et al* in their recent publication.

All three *SCN9A* variants are seen in the son (II.1), who does not feel NP but tested positive for SFN. This family was added to the current project to find out if there were other genetic variants along with *SCN9A* mutations that are causative for the proband but protective for the children.

The causative/protective genetic segregation models were assessed with respect to NP and SFN separately, as described in Table 7a and 7b.

Model type: NP	I.1	II.1	II.2	10%	5%	1%
Causative	Het	Ref	Ref	7	6	5
Causative	Hom	Het	Het	2	0	0
Protective	Ref	Het	Het	13	8	4
Protective	Het	Hom	Hom	0	0	0

Table 7a: Table showing the causative/protective genetic segregation model for the PROP03 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Model type: SFN	I.1	II.1	II.2	10%	5%	1%
Causative	Het	Het	Het	5	5	2
Causative	Hom	Hom	Hom	0	0	0

Table 7b: Table showing the causative/protective genetic segregation model for the PROP03 family with respect to SFN. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

The list of qualifying variants from both causative and protective models that were prioritized for PROP03 is reported in Table 7c. Here, the variant in the *WFS1* gene matched with the SFN phenotype. We found two heterozygous variants in the *TRPM4* gene (compound heterozygous) that have been previously reported for 'Progressive familial heart block, type IB' (Kruse *et al*, 2009). These variants were included because they were segregated as protective for NP. Also, the *SCN5A* gene, which is known for 'Progressive heart block, type 1A' (Probst *et al*, 1999), had a novel variant also seen only in children. Another variant in VGSC genes (*SCN8A*) was seen segregating inconclusively along with the other *SCN9A* mutations.

Prioritized qualifying variants	Classification status	I.1	II.1	II.2	Causative	Protective
WFS1:p.Ala738Asp	VUS	√	√	√	√	
TRPM4:p.Ala432Thr	VUS		√	√		√
TRPM4:p.Gly582Ser	VUS		√	√		√
SCN5A:p.Gln1873His	VUS		√	√		√
SCN8A:p.Ser476Arg	VUS	√	√		√	
SCN9A:p.Met932Leu+ *SCN9A*:p.Val991Leu	Pathogenic	√	√		√	
SCN9A:p.Asp1908Gly	VUS		√		√	

Table 7c: Table showing the prioritised list of qualifying variants for the PROP03 family. The qualifying variants belonging to the causative or protective model are reported in the last two columns. The SCN9A:p.Met932Leu + SCN9A:p.Val991Leu are in cis as they impact the gene together.

PROP04

In the PROP04 family, only two samples were sequenced from two generations where the father of the proband (I.1) had reduced in IENF density indicative of SFN but had no complaints related to NP. While the proband (II.1) was diagnosed with NP and SFN, as shown in Figure 18.
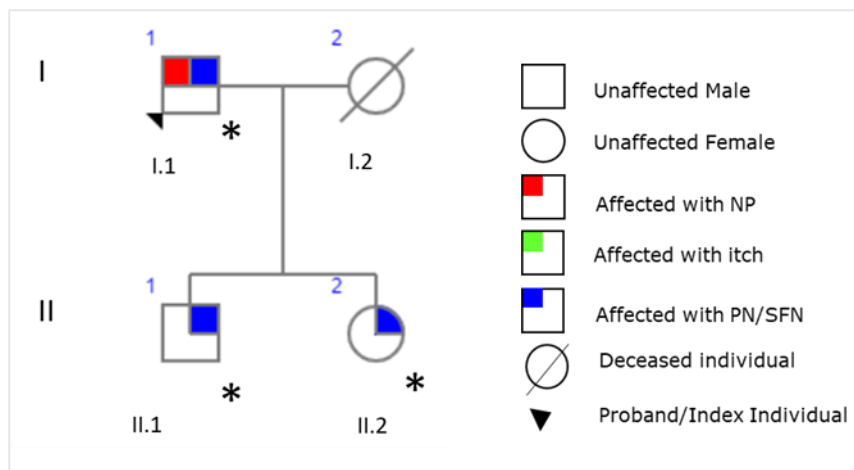

Figure 18: Pedigree of the family PROP04 family. The proband (II.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

Similar to PROP03, both the individuals (I.1 and II.1) shared three *SCN9A* variants, as depicted in Table 6. As mentioned in PROP03, the same *SCN9A* mutations were also observed in this family, and they segregated conclusively with respect to SFN but not for NP.

As a result, this family was included in the project to investigate if there are any other similarities between PROP03 and PROP04. In addition, new variants must be found to explain the peculiar genetic distribution that appears to be opposite across the two families. In fact, in the PROP03, the symptoms are transmitted from the father to the siblings, reducing their intensity, while in the PROP04, the father has mild symptoms compared to the daughter's more intense symptoms. For this family, causative/protective genetic model segregation was applied separately with respect to NP and SFN, as shown in Table 8a and 8b.

Model type: NP	I.1	II.1	10%	5%	1%
Causative	Ref	Het	13	9	3
Causative	Het	Hom	0	0	0
Protective	Het	Ref	21	19	1
Protective	Hom	Het	0	0	0

Table 8a: Table showing the causative/protective genetic segregation model for PROP04 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Model type: SFN	I.1	II.1	10%	5%	1%
Causative	Het	Het	15	13	1
Causative	Hom	Hom	1	0	0

Table 8b: Table showing the causative/protective genetic segregation model for the PROP04 family with respect to SFN. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

The record of all the prioritized qualifying variants for both causative and protective models have been illustrated in Table 8c. In this list, the mutations in VGSC genes, namely *SCN9A*, *SCN3A*, and *SCN11A*, were causative for SFN. The other variants in genes such as *NTRK1* and *CLTCL1* were also causative for SFN; these variants have been linked with insensitivity to pain (Axelrod *et al*, 1993; Nahorski *et al*, 2015). We also found two novel variants in genes *NTRK3* and *CHRNA2* from the same causative model. Comparatively, variants in *KCNK18* and *CACNA1E* genes were found to be protective for NP. We also observed a known polymorphism in the *TRPV1* gene in the homozygous state in the proband and in the heterozygous state in the father (I.1). This polymorphism in its homozygous state has been linked with functional dyspepsia but has also been implicated as a risk allele for dysfunctions in the somatosensory system (Tahara *et al*, 2010; Binder *et al*, 2011a).

Prioritized qualifying variants	Classification status	I.1	II.1	Causative	Protective
SCN9A:p.Met932Leu + *SCN9A*:p.Val991Leu	Pathogenic	√	√	√	
SCN9A:p.Asp1908Gly	VUS	√	√	√	
SCN11A:p.Gln520His	VUS	√	√	√	
SCN3A:p.Leu1861Phe	VUS	√	√	√	
NTRK1:p.Arg780Trp	VUS	√	√	√	
CLTCL1:p.Glu330Lys	VUS	√	√	√	
NTRK3:p.Arg645Cys	VUS		√	√	
CHRNA2:p.Met498Ile	VUS		√	√	
KCNK18:p.Ser252Leu	VUS	√			√
CACNA1E:p.Glu2148del	VUS	√			√
TRPV1:p.Met315Ile	Risk factor	√	√	√	

Table 8c: Table showing the prioritized list of qualifying variants for the PROP04 family. The qualifying variants belonging to the causative or protective model is reported in the last two columns. The SCN9A:p.Met932Leu + SCN9A:p.Val991Leu are in cis as they impact the gene together.

PROP05

In PROP05, the 67 year old mother of the proband (I.2) was diagnosed with distal myopathy and widespread pain. At the same time, the 44 year old proband (II.1) was diagnosed with Crohn's disease, neurosensorial hearing loss and distal pain. Both the individuals had low IENF density, indicating SFN and NP, as depicted in Figure 19.
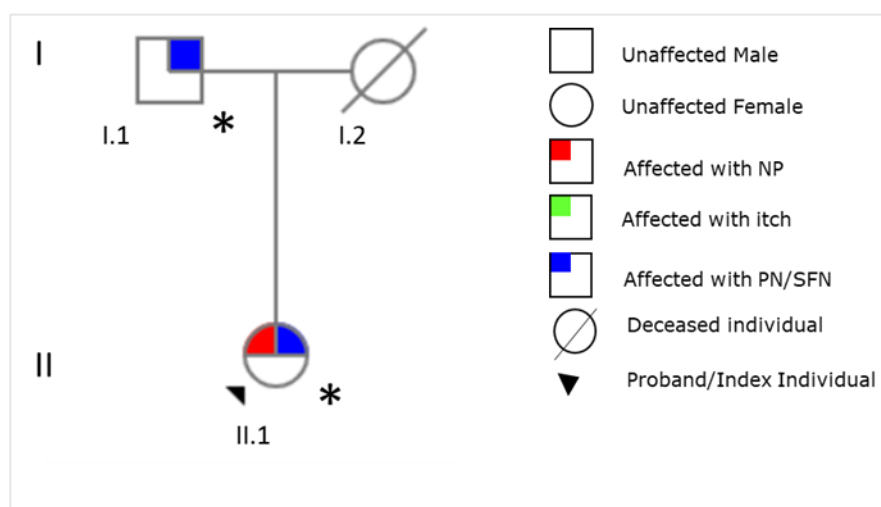

Figure 19: Pedigree of the PROP05 family. The proband (II.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

This family was selected to study painful SFN in autosomal dominant mode. As both affected members had the same phenotype, the genetic segregation models were performed with respect to SFN and pain, as depicted in Table 9a. Due to the absence of unaffected members in this family, we could assess only the causative model.

Model type: SFN, NP	I.1	II.1	10%	5%	1%
Causative	Het	Het	11	7	4
Causative	Hom	Hom	0	0	0

Table 9a: Table showing the causative genetic segregation model for the PROP05 family with respect to SFN and NP. After applying the filtration criteria mentioned in the Methods section, it includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

The qualifying variants selected for this family are shown in Table 9b.

The variant in *COL6A2* was selected as the mother (I.2) complained of distal myopathy, and variants in this gene have been previously associated with myopathy (Jobsis *et al*, 1996).

We found a variant in the *ATP7B* gene that has been associated with 'Wilson's disease' but in autosomal recessive mode (Bandmann *et al*, 2015). A functional study linked the *ATP7B* gene with MEDNIK syndrome, which stands for Mental retardation, Enteropathy, Deafness, peripheral Neuropathy, Ichthyosis, and Keratoderma (Jain *et al*, 2015a). As three phenotypes from MEDNIK syndrome matched with the proband's medical history, the *ATP7B* gene variant was selected and marked as 'Likely pathogenic'.

Other than that, a variant in the *TNFRSF1A* gene was also prioritized as an animal model study showed that it participated in skin inflammation (Kumari *et al*, 2013). This particular variant has been classified as 'Likely pathogenic' in the ClinVar database with confidence. But, as we could not find a specific related phenotype to either SFN or NP, we have classified it as 'VUS'.

Similar to the PROP04 family, the presence of a known risk factor for functional dyspepsia in *TRPV1*:p.Met315Ile in the homozygous state (Binder *et al*, 2011a; Tahara *et al*, 2010) was also observed in both the affected members of the PROP05 family.

Prioritized qualifying variants	Classification status	I.2	II.1	Causative	Protective
COL6A2:c.1572+1G>A	Likely pathogenic	√	√	√	NA
ATP7B:p.Met665Val	Likely pathogenic	√	√	√	NA
TNFRSF1A:p.Asp41Glu	VUS	√	√	√	NA
TRPV1:p.Met315Ile	Risk factor	√	√	√	NA

Table 9b: *Table showing the prioritised list of qualifying variants for the PROP05 family. The qualifying variants belonging to the causative or protective model are reported in the last two columns.*

PROP06

This family was comprised of two generations, each of which had an affected and unaffected member. The 37 year old proband (II.1) and his 64 year old mother (I.2) were both diagnosed with NP and SFN, as their nerve biopsies showed a reduction in IENF density. The mother (I.2) also complained of itch, as depicted in Figure 20. Based on this family, it appears that painful SFN is inherited in an autosomal dominant manner. Even though the itch phenotype is not inherited in this family, this aspect will shed light on genetic signatures that protect against itch.
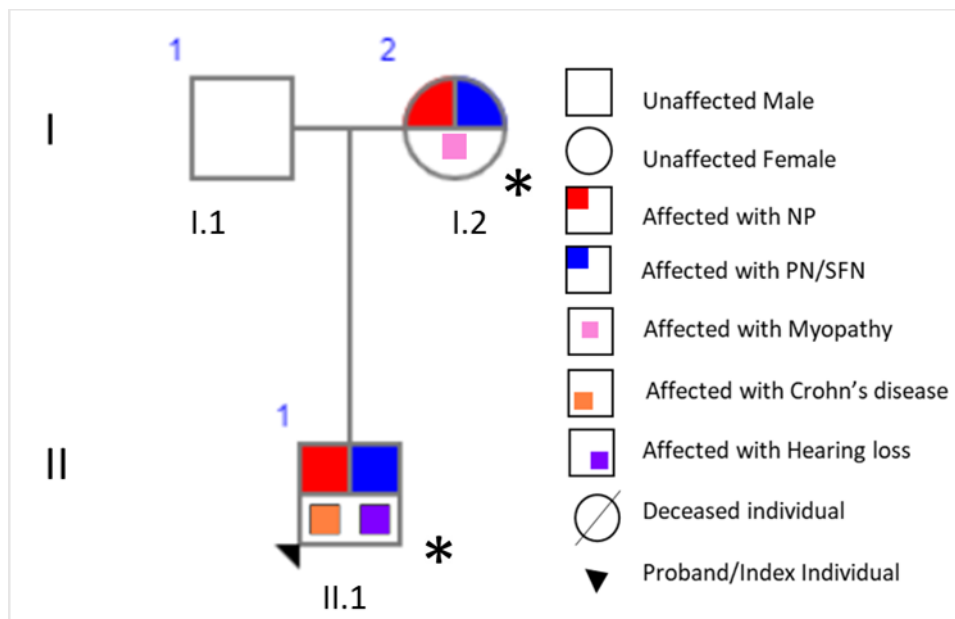

Figure 20: *Pedigree of the PROP06 family. The proband (II.1) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.*

In this family, the genetic segregation models were performed with respect to two phenotypes: i) SFN & NP and ii) itch separately, as elaborated in Table 10a and 10b.

Model type: SFN, NP	I.1	I.2	II.1	II.2	10%	5%	1%
Causative	Ref	Het	Het	Ref	6	6	2
Causative	Hom	Het	Het	Hom	0	0	0
Causative	Het	Hom	Hom	Het	0	0	0
Protective	Het	Ref	Ref	Het	2	2	0
Protective	Hom	Ref	Ref	Hom	0	0	0

Table 10a: Table showing the causative/protective genetic segregation model for PROP06 family with respect to SFN and NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Model type: itch	I.1	I.2	II.1	II.2	10%	5%	1%
Causative	Ref	Het	Ref	Ref	4	4	3
Causative	Hom	Het	Hom	Hom	0	0	0
Causative	Het	Hom	Het	Het	0	0	0
Protective	Het	Ref	Het	Het	15	12	2
Protective	Hom	Ref	Hom	Hom	0	0	0

Table 10b: Table showing the causative/protective genetic segregation model for PROP06 family with respect to itch. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

The prioritised list of qualifying variants obtained from the causative genetic segregation model has been illustrated in Table 10c. Only two variants were selected for this family, one in *SV2C* and the other in the *CAST* gene from the causative model. *SV2C* variant was selected as it is known to be crucial for neurotransmission (Löscher *et al*, 2016). *CAST* gene, which encodes for calpastatin protein, was selected as it takes part in synaptic plasticity (Nakajima *et al*, 2008).

A *TRPA1* variant was also found in this model but, as it had a high number of homozygotes in the global population, it was predicted to have a benign effect on the gene. So, it was ruled out as 'Likely benign'.

The genetic segregation models applied corresponding to itch did not yield any variant with sufficient biological support.

Prioritized qualifying variants	Classification status	I.2	II.1	I.1	II.2	Causative	Protective
SV2C:p.Ala676Val	VUS	√	√			√	NA
CAST:p.Ser113fs	VUS	√	√			√	NA
TRPA1:p.Asn109Lys	Likely benign	√	√			√	NA

Table 10c: Table of the prioritised qualifying variants for the PROP06 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

PNET01

This was the biggest family recruited in our cohort, where 11 individuals across three generations were available for WES experiments. This family had NP as its common phenotype. Two subjects, the proband at the age of 69 years (III.4) and her cousin at the age of 59 (III.11), complained of burning pain and underwent a comprehensive neurological examination, including tests through which they were diagnosed with painful SFN. The proband's younger brother (III.5) had diabetes, while the proband's cousin (III.9) complained of mild sensory disturbances, as depicted in Figure 21.
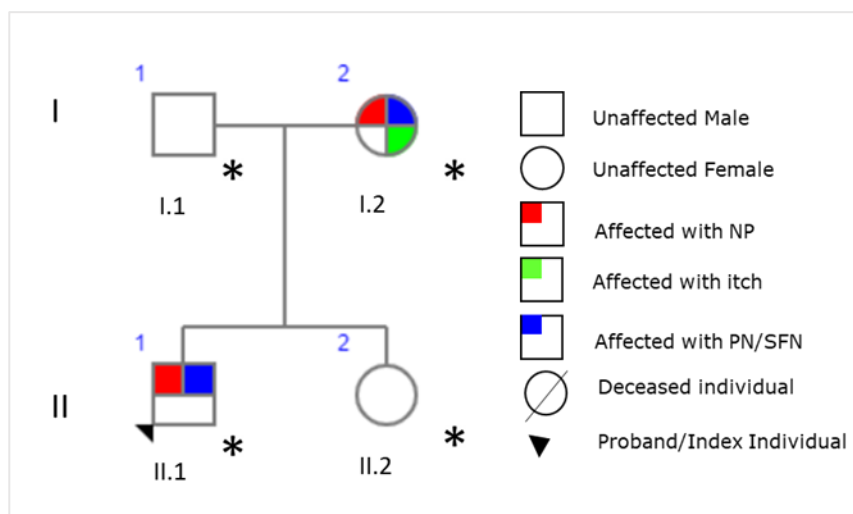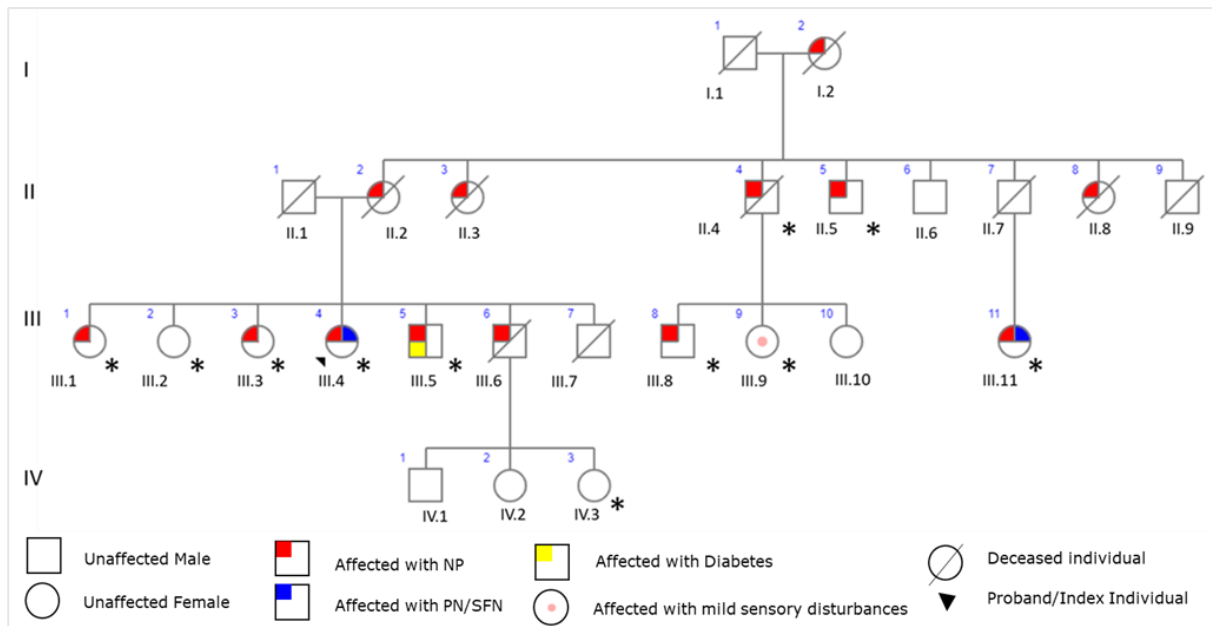
Figure 21: Pedigree of the family PNET01 family. The proband (III.4) is denoted by a black arrow. Individuals with an asterisk were selected for WES experiments.

This family was selected because of the large number of subjects involved, with NP phenotype across three generations. To evaluate this family, we adopted the following strategy: A) a genetic segregation model according to the phenotype, by dividing the family into three sections; and B) a genetic sharing model applied to the previous three groups, where cases were assessed against the unaffected members.

As regards to the three subdivisions of the family, we subsetted the following groups:

A.1) affected members: III.1, III.3, III.4, III.5; unaffected members: III.2 and IV.3. The different transmission models applied to this group have been described in Table 11a.

A.2) affected members: II.4, II.5, III.8 and III.9; unaffected member: III.2. The different transmission models applied to this group have been described in Table 11b.

A.3) affected members: II.5, III.11; unaffected member: III.2. The different transmission models applied to this group are shown in Table 11c.

Model type: NP	III.1	III.2	III.3	III.4	III.5	IV.3	10%	5%	1%
Causative	Het	Ref	Het	Het	Het	Ref	1	1	1
Causative	Het	Hom	Het	Het	Het	Hom	0	0	0
Causative	Hom	Het	Hom	Hom	Hom	Het	0	0	0
Protective	Ref	Het	Ref	Ref	Ref	Het	1	0	0
Protective	Ref	Hom	Ref	Ref	Ref	Hom	0	0	0

Table 11a: Table showing the causative/protective genetic segregation model for the first group of six individuals in the PNET01 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Model type: NP	II.4	III.8	III.9	III.2	10%	5%	1%
Causative	Het	Het	Het	Ref	8	7	6
Causative	Het	Het	Het	Hom	0	0	0
Protective	Ref	Ref	Ref	Het	18	13	5
Protective	Ref	Ref	Ref	Hom	0	0	0

Table 11b: Table showing the causative/protective genetic segregation model for the second group of four individuals in the PNET01 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

Model type: NP	II.5	III.11	III.2	10%	5%	1%
Causative	Het	Het	Ref	5	2	0
Causative	Het	Het	Hom	0	0	0
Protective	Ref	Ref	Het	16	11	6
Protective	Ref	Ref	Hom	2	0	0

Table 11c: Table showing the causative/protective genetic segregation model for the third group of three individuals in the PNET01 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

We then evaluated the sharing of genetic traits across each affected individual, comparing the presence of the variants in the affected members with their absence in the unaffected members. The sharing among the individuals was performed both at the gene level and at the variant level.

Regarding the results for sharing at the gene level, the data is reported in Table 12a.

Gene symbol	II.4	II.5	III.1	III.3	III.4	III.5	III.8	III.9	III.11	Sum
GABRR3	1		1	1		1		1		5
TRPM2		1	1	1	1					4
CACNA1A	1						1	1		3
LAMA2	1	1						1		3
KCNAB2				1	1					2
LZTR1		1				1				2
MFN2							1	1		2
NCOA7					1	1				2
OPRL1		1							1	2
SCN7A							1	1		2
SLC9A9			1			1				2
SV2B								1	1	2
COL11A2								1		1
KCNN3								1		1
SCN9A									1	1
TRPM7							1			1
ZFHX2		1								1

Table 12a: Tabulated representation of mutated genes shared amongst the affected members, which were absent in the unaffected members. This list of genes was obtained after applying the variant filtration criteria mentioned in the Methods section.

The prioritized list of qualifying variants obtained from all the groups with different transmission models is illustrated in Table 12b. For group A.1, the variants in gene *TRPM2* and *LZTR1* gene were prioritized as these genes have been associated with NP either by animal model study or indirectly linked with the phenotype (Matsumoto *et al*, 2016; Chung *et al*, 2015; Farschtschi *et al*, 2020). *MEN1* gene variant had inconclusive segregation but, as this gene has been previously linked with NP (Shen *et al*, 2014), it was added to the list. We noticed two variants in two potassium channel genes: i) the variants in the *KCNE2* gene was causative for NP; ii) the variant in the *KCNQ4* gene segregated inconclusively in the protective model. In the research by Strutz-Seebohm *et al* they demonstrated that *KCNE* beta-subunits modulate the function of the *KCNQ4* gene. As these two variants were found both in the proband, we have prioritized these variants to study them further.

For group A.2, the prioritized *CACNA1A* gene variant was observed in the second group comprised of the first maternal uncle of the proband (II.4) and his children (III.8 and III.9). This *CACNA1A* gene variant was selected as mutations in this gene were linked to sensory dysfunction along with episodic ataxia (Zafeiriou *et al*, 2009).

For the last group, A.3, similarly to the first group (A.1), the *KCNE2* gene is also mutated in the III.11, the individual diagnosed with painful SFN. We found a variant in the *OPRL1* gene that segregated within two generations; this gene encodes for nociceptin receptor, which is known to mediate nociception (Noda *et al*, 1998). Also, a *SCN9A* variant was observed in III.11 as causative for NP. This is a loss-of-function variant previously reported for congenital insensitivity to pain (Cox *et al*, 2010; Marchi *et al*, 2018).

Lastly, a *GABRR3* genetic variant was observed in five out of nine affected members, but it was not prioritized because there was not enough information for this variant or gene.

Prioritized qualifying variants	Classification status	II.4	II.5	III.1	III.3	III.4	III.5	III.8	III.9	III.11	III.2	IV.3	Causative	Protective
TRPM2: p.Pro112Ser	VUS		√	√	√	√							√	
LZTR1: p.Arg468His	VUS		√				√						√	
MEN1: p.Gly111Asp	VUS			√	√		√				√		√	
KCNE2: p.Ile57Thr	VUS			√	√	√				√			√	
KCNQ4: p.Glu551fs	VUS				√						√	√		√
CACNA1A: p.Arg481His	VUS	√						√	√				√	
OPRL1: p.Cys218Ter	VUS		√							√			√	
SCN9A: p.Arg896Gln	VUS									√			√	

Table 12b: Table of prioritized qualifying variants for three individuals of the PNET01 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

PNET02

This family is comprised of five individuals across two generations. The proband (II.4) and his step-sisters (II.1, II.2, and II.3) have the same father but different mothers, as shown in Figure 22. Five individuals were selected from this family for the study where individuals II.1, II.2, II.4 were the affected members of the family and II.3 and III.2 were selected as unaffected members of the family.
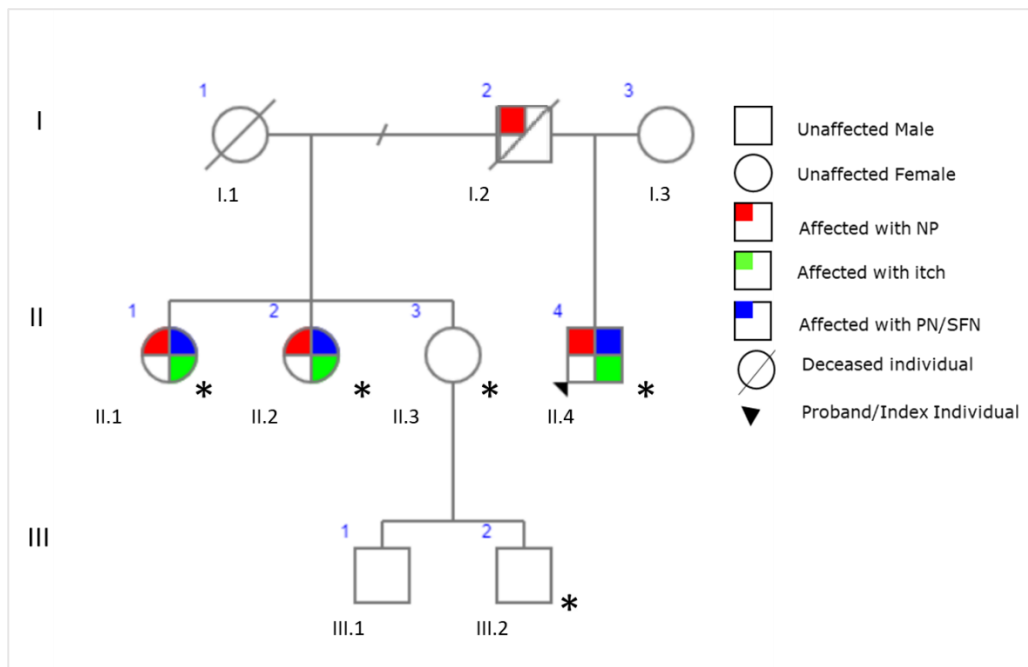
Figure 22: Pedigree of the PNET02 family. The proband (II.4) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

The proband's deceased father had symptoms similar to NP. The proband (II.4) had symptoms compatible with NP at the age of 52; these symptoms started after the surgical operation with meniscectomy. After a few months, he had itch attacks in his limbs and genitals. He also complained of visceral itching in the abdominal region along with joint pain. His skin biopsy evaluation showed a loss in IENF density, confirming SFN. The oldest step-sister (II.1), at 60 years of age, underwent surgery for anal fissures with spinal anaesthesia, and after a year, she complained of burning dysesthesia in her feet, which then extended to her hands. Similar to the proband, she started having itch attacks in her shoulders at night. The second oldest step-sister (II.2), at the age of 54, underwent hystero-appendectomy, and after that, she started to complain about itch attacks in her legs and genital region.

She also complained of burning pain in the feet along with vasomotor dysfunction. The youngest sister (II.3) and her son (III.2) did not have similar phenotypes and were characterized as unaffected members of the family. This family was selected for WES experiments given the peculiarity that the symptoms developed after surgical treatment. This family could shed light on the genes that are potentially involved in the triggering of NP after an external insult.

To evaluate the PNET02 family, the candidate gene filter was not applied as the literature search performed for creating the panel did not include genes involved in post-operative surgical pain. Other than that, the same criteria were used to assess the family, as depicted in Table 13a.

Model type: SFN, NP, itch (open exome)	II.1	II.2	II.3	II.4	III.2	10%	5%	1%
Causative	Het	Het	Ref	Het	Ref	21	14	5
Causative	Het	Het	Hom	Het	Hom	3	2	0
Causative	Hom	Hom	Het	Hom	Het	0	0	0
Protective	Ref	Ref	Het	Ref	Het	64	52	21
Protective	Ref	Ref	Hom	Ref	Hom	1	1	

Table 13a: Table showing the causative/protective genetic segregation model for PNET02 family with respect to SFN, NP and itch at 'open exome' level. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

The resultant qualifying variants from causative and protective models have been listed in Table 13b. The variants in genes *GABRD* and *ATAD3A* corresponded to SFN, NP and itch traits retrieved from the causative model, whereas the *MYLK2* genetic variant resulted from the protective model. The gene *GABRD* was associated with nucleoside reverse transcriptase inhibitors induced NP (Wu *et al*, 2021), and variants in the *ATAD3A* gene have been associated with axonal neuropathy (Harel *et al*, 2016). *MYLK2* gene has been associated with cardiomyopathy and motor neuropathy (Davis *et al*, 2001; Shtilbans *et al*, 2011). Although, its pathophysiology in motor neuropathy is still unknown.

Prioritized qualifying variants	Classification status	II.1	II.2	II.4	II.3	III.3	Causative	Protective
GABRD:p.Arg157*	VUS	√	√	√			√	NA
ATAD3A:p.Ala239Thr	VUS	√	√	√			√	NA
MYLK2:p.Arg553Cys	VUS				√	√		√

Table 13b: Table showing the prioritised list of qualifying variants for the PNET02 family. The affected member of the family have been listed before and the unaffected member at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

PNET03

The PNET03 family is comprised of four individuals in two generations where I.2 and II.1 were the affected members and II.2 and II.3 were selected as unaffected members of the family, as shown in Figure 23.

Figure 23: Pedigree of the PNET03 family. The proband (II.1) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

The proband (II.1) was 18 years old when he first experienced hyperhidrosis and intense fatigue. But, from the age of 40, he started having recurring itch attacks in the legs and back. He was not able to recognize any trigger for these attacks and did not have any other symptoms following them. Although, the proband reported a general discomfort with tachycardia in strong sunlight. His skin biopsy showed a marked decrease in IENF density indicative of SFN. He was then diagnosed as familial SFN with recurring itch. His mother (I.2) also had an itch phenotype, and a skin biopsy exhibited a decrease in IENF density, confirming SFN.

Proband's siblings (II.2 and II.3) had none of the symptoms and were then characterized as the unaffected members of the family. This family was selected to study the relationship between SFN and itch attacks. Initially, the 'candidate gene' filter was applied to the causative/protective genetic segregation analysis with respect to SFN and itch, but no valid results were obtained. Therefore, the same analysis was repeated at the 'open exome' level, and the assessment has been described in Table 14a.

Model type: SFN, itch (open exome)	I.2	II.1	II.2	II.3	10%	5%	1%
Causative	Het	Het	Ref	Ref	51	38	5
Causative	Het	Het	Hom	Hom	1	0	0
Protective	Ref	Ref	Het	Het	71	33	3
Protective	Ref	Ref	Hom	Hom	1	1	0

Table 14a: Table showing the causative/protective genetic segregation model for the PNET03 family with respect to SFN and itch at 'open exome' level. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

The final list of qualifying variants for PNET03 from the causative model has been illustrated in Table 14b. As mentioned before, for PROP05, also here we report *ATP7B* that has been linked with peripheral neuropathy by a functional study (Jain *et al*, 2015b). Variants in the *NEFH* gene have been associated with CMT (Rebelo *et al*, 2016). No variants were yielded from the protective model.

Prioritized qualifying variants	Classification status	I.2	II.1	II.2	II.3	Causative	Protective
ATP7B:p.Leu1188Pro	VUS	√	√			√	NA
NEFH:p.Ser124Cys	VUS	√	√			√	NA

Table 14b: Table of the prioritised qualifying variants for the PNET03 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

<u>PNET04</u>

The PNET04 is a trio family where the father (I.1) and mother (I.2) are unaffected, but their daughter (II.2), the proband, was diagnosed with SFN and NP at the age of 5 years old, as depicted in Figure 24.
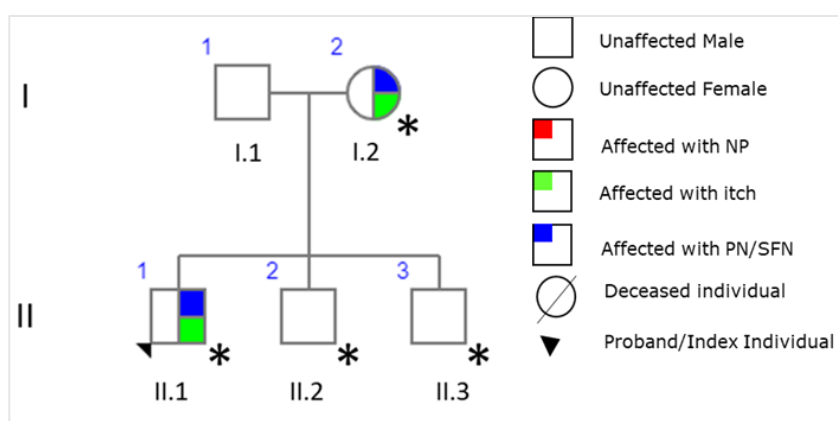

Figure 24: Pedigree of the PNET04 family. The proband (II.2) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

The proband had the onset of symptoms one month after getting chickenpox. She had burning pain of the feet and redness and heat during summer or after long walks.

We observed no variants after examination from the initially adopted transmission models. It was presumed that maybe the daughter had inherited variants from her parents. Thus, we applied an autosomal dominant inheritance model for further variant analysis. The revised transmission models have been illustrated in Table 15a.

Model type: SFN, NP	I.1	I.2	II.1	10%	5%	1%
Causative	Het	Het	Het	67	52	21
Causative	Hom	Hom	Hom	0	0	0
Protective	Het	Het	Ref	32	15	2
Protective	Hom	Hom	Ref	0	0	0

Table 15a: Table showing the causative/protective genetic segregation model for the PNET04 family with respect to SFN and NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

After variant evaluation, it was observed that three variants located in VGSCs genes and the *GBA* gene were inherited in the proband from the father. The *SCN9A*, *SCN11A,* and *GBA* genes have been previously linked with either NP or SFN (C. *et al*, 2012; Huang *et al*, 2014; Devigili *et al*, 2017). While the proband inherited variants in *TRPM8* and *NAGLU* genes from her mother, these variants have also been associated with either PN or NP (Binder *et al*, 2011b; Tétreault *et al*, 2015). It was concluded that the mode of inheritance could be 'Digenic autosomal dominant' where one variant inherited from the father and another from the mother could cause the phenotype in the child. Further examinations have to be done to confirm these results. In addition to this, a variant in the *CACNA1E* gene was found only in the father, which could be protective in nature. The final list of qualifying variants for PNET04 has been illustrated in Table 15b.

Prioritized qualifying variants	Classification status	II.2	I.1	I.2	Causative	Protective
SCN9A:p.Arg830Gln	VUS	√	√		√	
SCN11A:p.Arg1147Trp	VUS	√	√		√	
GBA:p.Thr222Pro	VUS	√	√		√	
TRPM8:p.Pro363Ser	VUS	√		√	√	
NAGLU:c.384-1C>G	VUS	√		√	√	
CACNA1E:p.Arg36fs	VUS		√			√

Table 15b: Table showing the prioritised list of qualifying variants for the PNET04 family. The affected member of the family have been listed before and the unaffected member at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

This family comprised of six individuals where the proband (II.1), her mother (II.2), and proband's maternal aunties (I.3 and I.4) were selected as affected members. Proband's father and brother were selected as unaffected members, as shown in Figure 25.
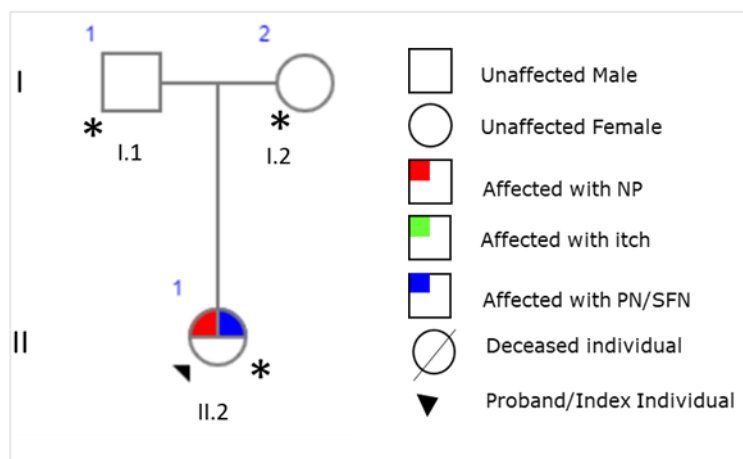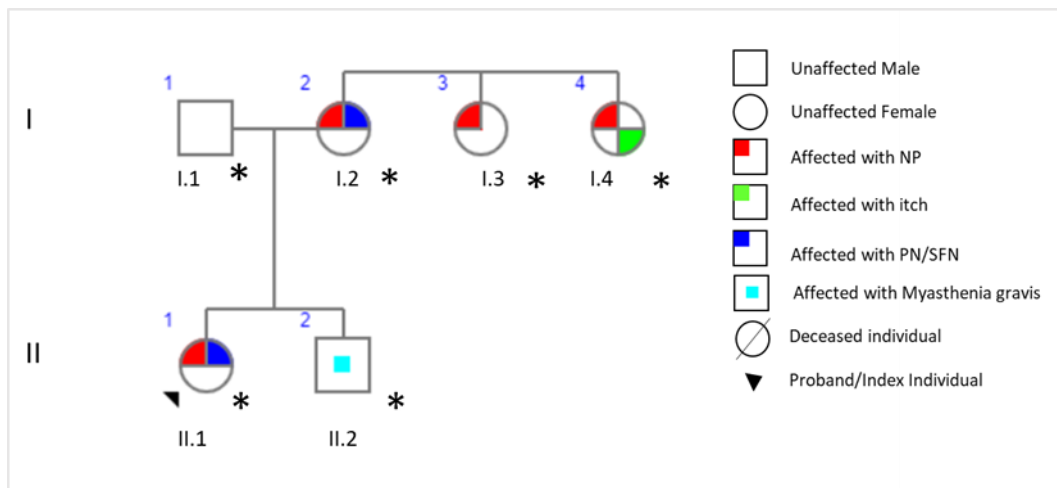

Figure 25: Pedigree of the PNET05 family. The proband (II.1) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

The proband had an onset of pain-related symptoms at the age of 17 and was diagnosed with painful SFN. The mother (I.2) was also diagnosed with NP. The mother (I.2) has sisters who were also diagnosed with NP. One of the mother's sisters (I.4) has talipes equinovarus since birth, along with itch in a non-length dependant manner. The other mother's sister (I.3) was diagnosed with NP as she experiences burning pain in her legs. The proband's brother (II.2) has Myasthenia Gravis with no other complaints. So, the brother (II.2) and father (I.1) with unremarkable medical history was selected as the unaffected members of this family. This family was selected to assess the transmission of NP from the maternal side of the family. To evaluate this family, the causative/protective genetic segregation models with respect to NP were adopted, as illustrated in Table 16a.

Model type: NP	I.1	I.2	I.3	I.4	II.1	II.2	10%	5%	1%
Causative	Ref	Het	Het	Het	Het	Ref	45	3	13
Causative	Hom	Het	Het	Het	Het	Hom	0	0	0
Causative	Het	Hom	Hom	Hom	Hom	Het	0	0	0
Protective	Het	Ref	Ref	Ref	Ref	Het	1	4	4
Protective	Hom	Ref	Ref	Ref	Ref	Hom	0	0	0

Table 16a: Table showing the causative/protective genetic segregation model for the PNET05 family with respect to NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

The prioritized list of variants from causative models has been illustrated in Table 16b. Only two genes segregated conclusively with complete penetrance in this family, whereas other variants had an incomplete penetrance. The *SPTBN2* variant was selected because it is known that the mutations in the *SPTBN2* gene affect neuronal morphology and have been associated with 'Spinocerebellar ataxia 5' (Clarkson *et al*, 2014). The genetic variant in *GLA* has been previously reported for SFN along with Fabry's disease, and it was classified as a risk factor as it is a common polymorphism (Lenders *et al*, 2013). Also, *GLA* genetic variant was heterozygous in the affected members but in homozygous state in the proband's brother (II.2), who has Myasthenia Gravis. A novel *SCN9A* mutation was found in this family. It was weakly conserved across species, and therefore it was classified as 'Likely benign'. The mutations in *COL1A2* have been associated with 'Ehlers-Danlos syndrome' (T *et al*, 1987), and it was selected as one of the affected members with this mutation had the itch. Lastly, the *PNOC* gene is involved in pain transmission (Okuda-Ashitaka *et al*, 1998), so the variant in this gene was prioritised.

Prioritized qualifying variants	Classification status	II. 1	I. 2	I. 3	I. 4	I. 1	II. 2	Causative	Protective
SPTBN2:p.Arg1125Gln	VUS	√	√	√	√			√	
GLA:p.Asp313Tyr	Risk factor	√	√	√	√		√	√	
SCN9A:p.Glu1129Ala	Likely benign	√	√		√			√	
COL1A2:p.Pro89Ser	VUS	√	√		√			√	
PNOC:p.Leu159Pro	VUS	√	√					√	

Table 16b: Table of prioritized qualifying variants for PNET05 family along with their interpreted classification status. The affected members of the family are listed first and the unaffected members at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

PNET06

The PNET06 is a family comprised of four individuals spread across two generations. In this family, the proband (II.2), proband's mother (I.2) and proband's sister (II.1) were selected as the affected members of the family. While the father (I.1) was selected as the unaffected member of the family, as shown in Figure 26.
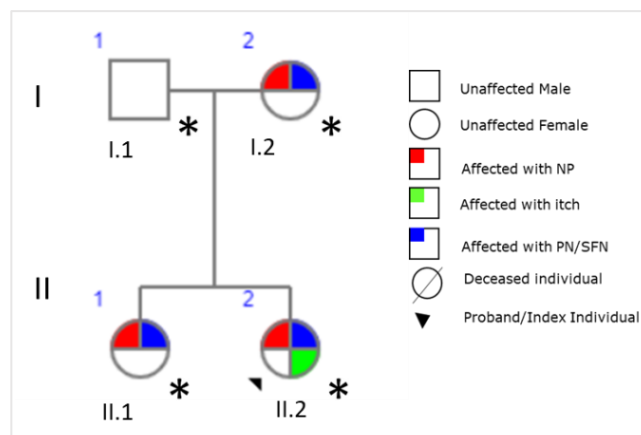
Figure 26: Pedigree of the PNET06 family. The proband (II.2) is denoted with a black arrow. Individuals with an asterisk were selected for WES experiments.

The mother (I.2) and her two daughters (II.1 and II.2) have been suffering from intense pain. The proband (II.2) is the youngest daughter, and she had NP phenotype along with itch. The proband started experiencing these symptoms from the age of 20. All three affected members of the family were diagnosed with painful PN. This family was selected to investigate the genetic components of PN inherited in an autosomal dominant way. The different types of transmission models applied to this family have been illustrated in Table 17a.

Model type: NP, PN	I.1	I.2	II.1	II.2	10%	5%	1%
Causative	Ref	Het	Het	Het	55	29	7
Causative	Hom	Het	Het	Het	0	0	0
Protective	Het	Ref	Ref	Ref	49	44	25
Protective	Hom	Ref	Ref	Ref	0	0	0

Table 17a: Table showing the causative/protective genetic segregation model for the PNET06 family with respect to PN and NP. It includes the number of variants retrieved at three different MAF (less than 10%, 5% and 1%) points after applying the filtration criteria mentioned in the Methods section. Individuals highlighted in red are affected members and unaffected members in green. The Ref stands for Reference variant, Het stands for Heterozygous variant, and Hom stands for Homozygous variant

The list of classified qualifying variants from causative models has been illustrated in Table 17b. A novel variant in *WARS* was found to segregate conclusively according to a causative model. It was selected because the variants in the *WARS* gene have been previously linked with distal hereditary motor neuropathy (Tsai *et al*, 2017; Wang *et al*, 2019). Another mutation was in the *TRPM2* gene that has been functionally linked with NP (Matsumoto *et al*, 2016; Chung *et al*, 2015), but no mutations have been reported yet. The *TRPA1* variant had an incomplete penetrance, but it was still selected as it has been previously linked with familial episodic pain and itch (Moore *et al*, 2018; Kremeyer *et al*, 2010). In addition to that, a variant in the sugar transporter *SLC45A1* gene variant was added to the list because it was classified as 'Pathogenic' by the OMIM database, but this gene is associated with intellectual disability and epilepsy in an autosomal recessive inheritance model (Srour *et al*, 2017). Because of the difference in phenotype and inheritance pattern, the *SLC45A1* genetic variant was classified as 'VUS' in this study.

Prioritized qualifying variants	Classification status	II.2	I.2	II.1	I.1	Causative	Protective
WARS:p.Ile435Leu	VUS	√	√	√		√	
TRPM2:p.Arg411Trp	VUS	√	√	√		√	
TRPA1:p.Thr311Asn	VUS	√	√			√	
SLC45A1:p.Ala210Val	VUS	√	√			√	

Table 17b: Table showing the prioritised list of qualifying variants for PNET06 family. The affected member of the family have been listed before and the unaffected member at the end. The qualifying variants belonging to the causative or protective model are reported in the last two columns.

Prioritisation of variants in the familial cohort

The results obtained from applying causative/protective genetic segregation models for each family were assessed in three steps as mentioned in the Methods section:

1. variant filtration
2. filtration by candidate genes
3. assignment of classes as defined in ACMG guidelines

For each step, the numbers of variants belonging to each family were counted and reported in Table 18a.

Family identifier	Number of variants after variant filtration	Number of qualifying variants selected after classification
PROP01	48	6
PROP02	89	7
PROP03	37	8
PROP04	45	12
PROP05	67	4
PROP06	38	3§
PNET01	131	9
PNET02	8	3
PNET03	117	2
PNET04	45	6
PNET05	37	5§
PNET06	76	4

Table 18a: The count of each variant after each step of prioritization per family.
§ denotes one out of the total number of variants was classified as 'Likely benign'

In more detail, for each of the filtered qualifying variants in Table 18b, it is possible to appreciate their distribution according to the ACGM classification status across the twelve families. For instance, in PROP01, the count of variants after filtration is 48, out of which 41 variants were classified as 'Benign/Likely benign'. As regards to the remaining 7 qualifying variants, 6 were classified as 'VUS' whereas 1 was classified as a 'Risk factor'.

Family identifier	Risk factor	Benign/Likely benign	VUS	Likely pathogenic	Pathogenic
PROP01	1	41	6		
PROP02		82	7		
PROP03		29	6		2
PROP04	1	33	9		2
PROP05	1	63	1	2	
PROP06		36	2		
PNET01	1	122	8		
PNET02		5	3		
PNET03		115	2		
PNET04		39	6		
PNET05	1	32	3		
PNET06		72	4		

Table 18b: The count of each variant by its classification status for each family

Furthermore, the prioritised qualifying variants seen more than once were tabulated according to their classification. The variants are grouped according to the gene they map to, as shown in Table 18c.

Gene symbol	Risk factor	Benign/ Likely benign	VUS	Likely pathogenic	Pathogenic	Total
SCN9A	1	1	3		2	7
CACNA1E			3			3
TRPM8			3			3
ATAD3A			2			2
ATP7B			1	1		2
COL6A5		1	1		1	2
SCN11A			2			2
TRPM2			2			2

Table 18c: The number of classified qualifying variants seen more than once in families is summarized, the grouping of the variants according to the gene they map.

It would be pertinent to note that the *SCN9A*:p.Met932Leu and *SCN9A*:p.Val991Leu variants classified as 'Pathogenic' were found in two different families (PROP03 and PROP04), matching the SFN phenotype but not the NP phenotype in our cohort.

There were 48 distinct genes among the variations reported in Table 19b identified as 'VUS,' 'Likely Pathogenic,' and 'Pathogenic.' 8 out of 48 genes overlapped in two or more families as depicted in Table 18d.

Gene symbol	PROPANE Study						PAIN-Net Study						Sum
	01	*02*	*03*	*04*	*05*	*06*	*01*	*02*	*03*	*04*	*05*	*06*	
SCN9A	1		1	1			1			1	1		6
CACNA1E		1		1						1			3
COL6A5	1	1											2
ATAD3A		1						1					2
ATP7B					1				1				2
SCN11A				1						1			2
TRPM2							1					1	2
TRPM8		1								1			2

Table 18d: The 8 qualifying genes that overlapped amongst the families with at least one recurrence in two families are summarised

To better describe the nature of the 48 genes, we clustered them using 'BRITE functional hierarchies' through the KEGG database (Kanehisa *et al*, 2012). These genes were classified by the tool into four categories as follows and as shown in Figure 27: 'Ion channel', 'Neurotransmission', 'Metabolism' and 'Immune response'. Most of the genes were found in the 'Neurotransmission' category, with the 'Ion channel' category being the second most common.
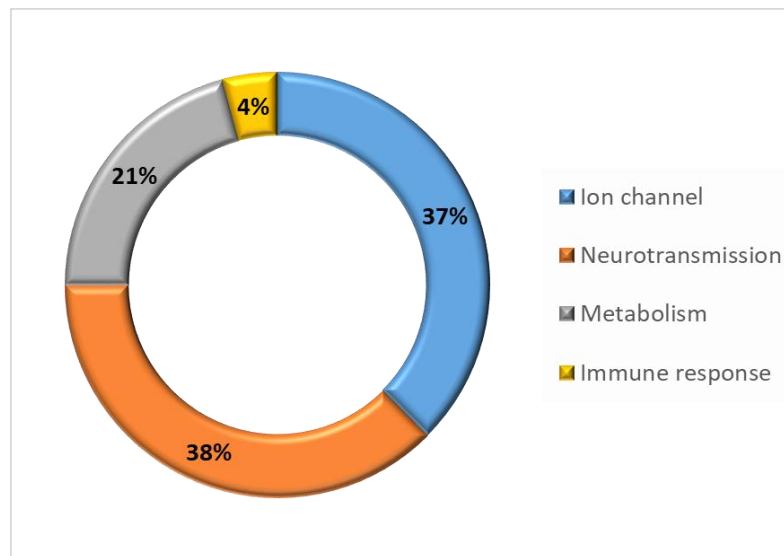
Figure 27: A pie chart of categorized 48 unique genes from familial cohort into four groups, namely, 'Ion channel', 'Neurotransmission', 'Metabolism' and 'Immune response' according to 'BRITE functional hierarchies'.

Sporadic Cohort

The sporadic cohort consists of 43 individuals: i) 23 subjects were recruited within the PROPANE study, and ii) 20 subjects were recruited under the PAIN-Net project.

The clinical aspects of the sporadic cohort have been illustrated in Table 19. The final analysed cohort had 25 females with a mean age of 36.4 years ± 6.54 (average ± standard deviation) and 18 males with a mean age of 35.7 years ±7.71 (average ± standard deviation).

Gender	Number of samples
Female	25
Male	18
Age range (years)	**Number of samples**
10 – 19	2
20 – 29	11
30 – 39	11
40 – 45	19
Clinical classification	**Number of samples**
Painful PN	38
Itch + painful PN	5

Table 19: Clinical aspects of the sporadic cohort recruited under PROPANE and PAIN-Net study

<u>Filtration and prioritisation of variants in the sporadic cohort</u>

The sporadic cohort patients went through the same variant filtering criteria as the families, with the exception of using the causative/protective segregation model, so only heterozygous and homozygous variants were considered for sporadic samples. Each subject in the sporadic cohort was individually assessed according to its phenotype, and each qualifying variant was stratified according to the ACMG guidelines along with their classification status. An additional filter was applied to each sample in order to avoid the selection of common variants in the studied population: the mutation should be less than 25% frequent in the sporadic cohort (so present in less than 11 sporadic subjects). Figure 28 depicts the number of variants produced by each step and their classification according to ACMG guidelines.

Number of variants after variant filtration
341

Number of variants selected after classification
130

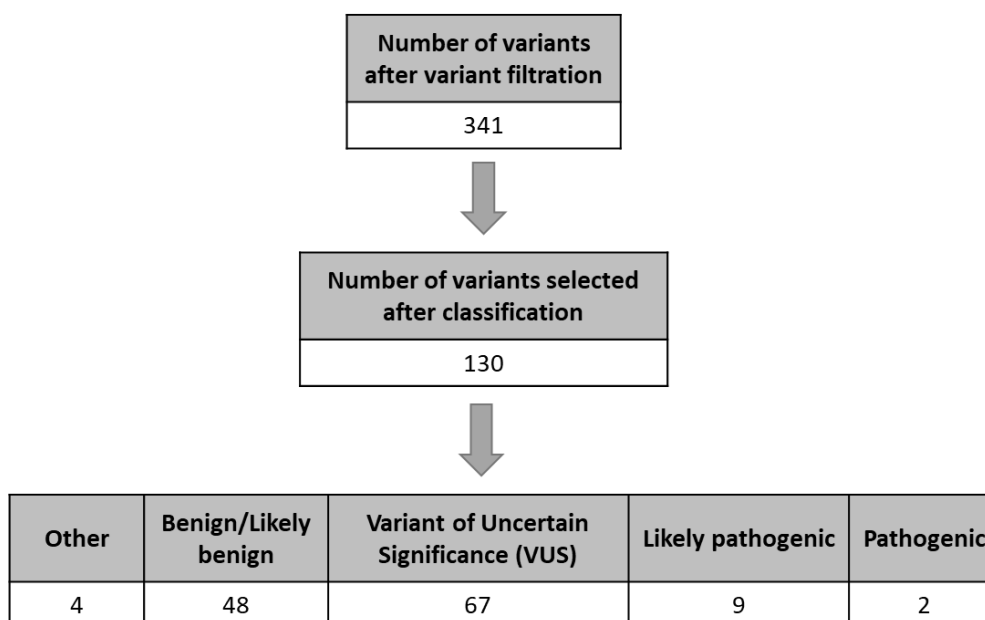| Other	Benign/Likely benign	Variant of Uncertain Significance (VUS)	Likely pathogenic	Pathogenic
4	48	67	9	2

Figure 28: The count of qualifying variants at each step of variant prioritization and the count of variants for each category of classification are reported.

The 11 variants that were classified as 'Pathogenic' or 'Likely pathogenic' were distributed across 12 different sporadic subjects. The 11 mutations have been listed in Table 20.

Prioritized qualifying variants	Classification status	Clinical indication
MEFV:p.Lys695Arg[§]	Pathogenic	Painful PN
WFS1:p.Val142fs	Pathogenic	Painful PN
ATL1:p.Arg416His	Likely pathogenic	Painful PN
ATP7B:p.Asp765Asn	Likely pathogenic	Painful PN
ATP7B:p.Tyr713Cys	Likely pathogenic	Itch + Painful PN
COL7A1:p.Val168fs	Likely pathogenic	Painful PN
FSHR:p.Arg283Trp	Likely pathogenic	Painful PN
LZTR1:p.Arg362Ter	Likely pathogenic	Painful PN
MYT1L:p.Arg558Cys	Likely pathogenic	Painful PN
PMP22:p.Val30Met	Likely pathogenic	Painful PN
SCN10A:p.Ile1225Thr	Likely pathogenic	Painful PN

Table 20: The list of 11 genetic variants seen in 12 different subjects as the variant marked § is seen in two different individuals

The MEFV:p.Lys695Arg, a 'Pathogenic' variant, was discovered in two people. This variant has been mentioned in several publications, and mutations in the MEFV gene have been associated with SFN (Amarilyo *et al*, 2020). Another variant is SCN10A:p.Arg14Leu, classified as 'VUS', was found in two sporadic individuals. This particular variant has been observed in 4 patients in a published genetic study (Eijkenboom *et al*, 2019). In addition, variants in the *SCN10A* gene had the highest variant and sample count: 4 variants were shared among 5 patients. There were 17 genes that had qualifying variants in at least two samples in the sporadic cohort. The list of these 17 mutated genes as well as related sample and variant counts have been depicted in Table 21.

Gene symbol	Sample count	Variant count
SCN10A	5	4
ATP7B	3	3
SCN9A	3	3
TRPA1	3	3
COL7A1	2	2
DNMT1	2	2
MEFV	2	1
MYT1L	2	2
NAGLU	2	2
NF1	2	2
NOTCH3	2	2
PIEZO2	2	2
RYR1	2	2
STOML3	2	2
TRPM5	2	2
TRPV1	2	2
WFS1	2	2

Table 21: List of 17 mutated genes shared in more than one sample in the sporadic cohort. A sample and variant count has been recorded for each mutated gene.

The qualifying variants that were classified as 'VUS', 'Likely pathogenic' and 'Pathogenic' amounted to 77 unique genes. To better classify these 77 genes, they were sorted into the aforementioned four categories according to their functional biological system obtained from 'BRITE functional hierarchies' through the KEGG database (Kanehisa *et al*, 2012), as shown in Figure 29. Compared to the familial cohort, the prioritized qualifying genes in the sporadic cohort were mostly seen in the 'Ion channel' category and then in the 'Neurotransmission' category.
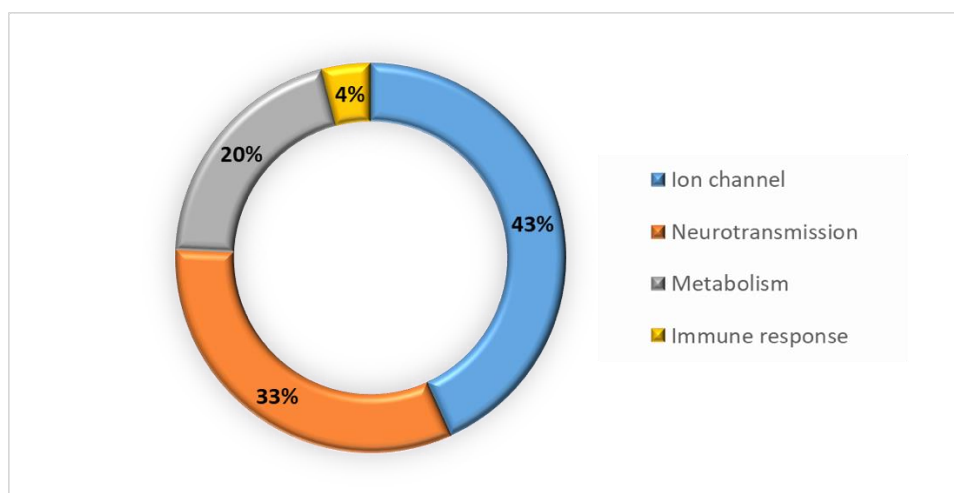
Figure 29: A pie chart of categorized 77 unique genes from sporadic cohort into four groups namely, 'Ion channel', 'Neurotransmission', 'Metabolism' and 'Immune response' according to 'BRITE functional hierarchies'.

Comparison between familial and sporadic cohort

Only one genetic variant classified as 'VUS' overlapped between the familial and sporadic cohort (*SCN9A*:p.Ile739Val). But, there was a higher overlap between the mutated genes both in familial and sporadic subjects. In particular, 12 genes overlapped between the 48 genes from the familial cohort and 77 genes from the sporadic cohort, as shown in Figure 30.
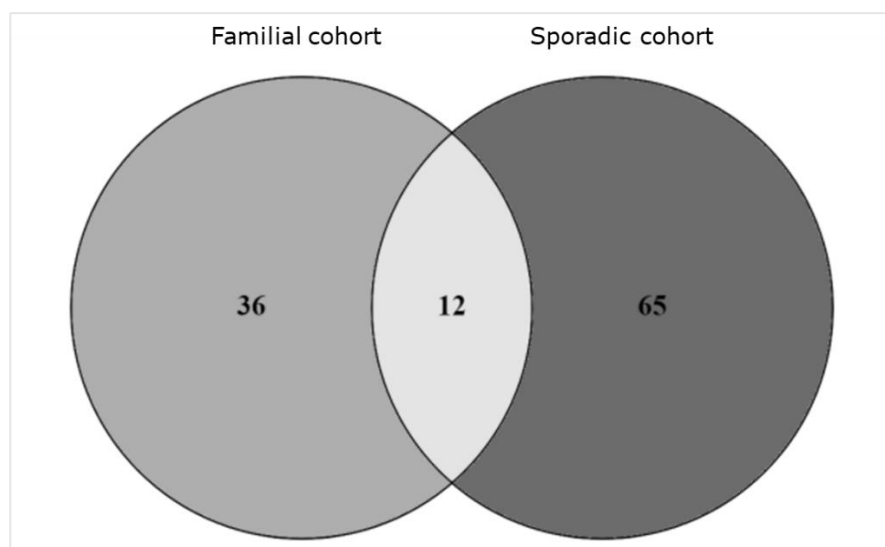
Figure 30: A Venn diagram of the number of mutated genes shared amongst the familial and sporadic cohorts

Collectively, among familial and sporadic subjects, the gene mutated in most of the samples was *SCN9A* followed by *ATP7B*, a copper transporter, as shown in Table 22. These results indicate that genes other than VGSCs might be playing a pivotal role in painful PN.

Gene symbol	Sample count	Variant count
SCN9A	17	8
ATP7B	7	5
SCN10A	6	5
TRPA1	5	4
TRPM8	5	4
CACNA1E	4	4
NAGLU	4	3
NEFH	3	2

Table 22: List of mutated genes in the entire cohort, considering both familial and sporadic samples, with their sample and variant counts

In summary, 113 genes were prioritised in the entire cohort and this list can be found in Appendix A.4. The 12 genes that were in common between the two cohorts are listed in Table 23.

Gene symbol	Gene name
ATP7B	ATPase, Cu^{2+} transporting, beta polypeptide
CACNA1E	Calcium channel, voltage-dependent, alpha 1E subunit
KCNK18	Potassium channel, subfamily K, member 18
LZTR1	Leucine-zipper-like regulator-1
MYT1L	Myelin transcription factor 1-like
NAGLU	N-acetylglucosaminidase, alpha-
NEFH	Neurofilament, heavy polypeptide
SCN9A	Sodium voltage-gated channel, alpha subunit 9

Table 23: List of 12 mutated genes in common between familial and sporadic cohorts along with their gene names

Pathway analysis results

The pathway analysis was performed to assess if the qualifying genes selected chosen from familial and sporadic cohorts using the variants filtering step were enriched in specific biological processes and for a better interpretation of the functional role of genes involved in these pathways in relation to the phenotype. This analysis was performed separately for familial and sporadic cohorts following the process reported in the Methods section. We used the manually curated PANTHER database for the pathway annotation through the WebGestalt tool.

To conduct the pathway analysis, we chose genes from each family based on variants that segregated for the PN/SFN & NP or PN/SFN & itch phenotype. Similarly, for the sporadic cohort, genes for pathway analysis were chosen because mutated genes shared by subjects with PN/SFN & NP or PN/SFN & itch phenotypes were observed in less than 25% of cases. The resultant number of qualifying genes from the familial cohort was 185 and for the sporadic cohort was 2897.

Taking into account the mapping performed by the WebGestalt tool to limit the study to unambiguously annotated genes, the number of genes that survived for pathway analysis using the PANTHER database has been described in Table 24 for the two analysed cohorts.

	Mapped input	
Pathway database	**Familial cohort**	**Sporadic cohort**
PANTHER database	31	348

Table 24: The count of input genes annotated by WebGestalt tool to run pathway analysis.

As a result, we obtained two tables per cohort from the WebGestalt tool comprised of 10 pathways. The pathways that matched the selection criteria as explained in the Methods section for each cohort were then prioritized. The resultant top five pathways that emerged from familial and sporadic cohorts have been described in Table 25a and 25b, respectively.

Pathway name	Gene set	Mapped input	Enrichment ratio	p-value	FDR
Nicotinic acetylcholine receptor signalling pathway	93	7	5.501908	1.84E-04	0.020738
Integrin signalling pathway	166	6	2.642052	0.022336	1
Hypoxia response via HIF activation	26	2	5.622829	0.048022	1
Inflammation mediated by chemokine and cytokine signalling pathway	200	6	2.192903	0.049818	1
FAS signalling pathway	31	2	4.715921	0.065883	1

Table 25a: Tabulated pathway analysis result from mapped input qualifying genes from the familial cohort. The table contains five columns: 1. Name of the pathway 2. The number of genes in that pathway 3. The number of genes overlapped between the pathway and the given input 4. Enrichment ratio 5. p-value of each hypergeometric test 6. FDR for each test.

Pathway name	Gene set	Mapped input	Enrichment ratio	p-value	FDR
Nicotinic acetylcholine receptor signalling pathway	93	33	2.3105	8.03E-07	9.08E-05
Integrin signalling pathway	166	39	1.5298	0.00275	0.1554
Insulin/IGF pathway-protein kinase B signalling cascade	35	12	2.2325	0.00417	0.15711
Endothelin signalling pathway	77	19	1.6067	0.02022	0.57115
Alzheimer disease-amyloid secretase pathway	63	15	1.5504	0.04927	0.92187

Table 25b: Tabulated pathway analysis result from mapped input qualifying genes from the sporadic cohort. The table contains five columns: 1. Name of the database 2. Name of the pathway 3. The number of genes in that pathway 4. The number of genes overlapped between the pathway and the given input 5. p-value of each hypergeometric test 6. FDR for each test.

Interestingly, the shared pathway between the familial and sporadic cohorts with FDR less than 0.10 was the 'Nicotinic acetylcholine receptor signalling pathway' retrieved from the PANTHER database. This pathway is comprised of 93 genes.

In the familial cohort, 7 of the 31 input genes mapped to the 93 genes of this pathway. The seven genes are the following: *STX16, CHRNA10, MYH7B, CHRNA2, MYO15A, PLEKHH3, and MYH13*. In the PROP05 family, two genes, *MYH7B* and *CHRNA10* were mutated. The other genes, PROP02, PROP03, PROP04, and PNET06, mutated only once in four families.

For the sporadic cohort, 33 out of 348 genes mapped to the 93 genes of the pathway. These genes are the following: *ACHE, BCHE, CACNA1D, CACNA1F, CACNB1, CACNB2, CHAT, CHRNA10, CHRNA3, CHRNA7, CHRNE, MYH1, MYH10, MYH11, MYH13, MYH2, MYH4, MYH6, MYH7, MYH7B, MYO15A, MYO18A, MYO18B, MYO19, MYO1A, MYO1D, MYO1E, MYO3B, MYO5B, MYO5C, MYO7A, MYO7B,* and *PLEKHH3*. The top 10 genes with a high variant count in the sporadic cohort are *MYH7B, MYO7A, MYH6, MYO15A, MYO5B, CHRNA10, MYH7 MYO7B, MYH13,* and *MYO19*.

Only five genes were shared in the two output gene lists: *CHRNA10, MYH7B, MYO15A, PLEKHH3,* and *MYH13*. The *MYH7B* gene was the most mutated of the five common genes in all samples, whereas the *PLEKHH3* gene was the least. The sample count for common genes between the two cohorts has been shown in Table 26.

	Sample count		
Gene name	**Familial cohort**	**Sporadic cohort**	**Total**
MYH7B	2	15	17
MYO15A	1	6	7
CHRNA10	2	4	6
MYH13	3	3	6
PLEKHH3	1	3	4

Table 26: The number of samples with variants in the five genes that were common between the familial and sporadic cohorts

Structural variant prioritization

WES concentrates mainly on the protein-coding region of the genome with the aim of identifying genetic variants potentially related to a specific phenotype. Along with this, WES is also useful to study variation in the genome at the structural level. Taking advantage of this, we performed structural variant (SV) analysis on our samples.

After applying all the criteria regarding coverage and numbers of samples reported in the Methods section, a total of 80 samples (42 from familial cohort and 38 from sporadic cohort) were selected for this analysis. The families PNET05 and PNET06 were not considered for this analysis as they originated from a different enrichment kit. In addition, one PROP01 (I.2) sample from the familial cohort and five samples from the sporadic cohort were excluded because their coverage depth was less than 50X.

To increase the power in SV detection, we took advantage of 40 unrelated in-house healthy controls processed by the Sure Select V5 kit. These samples were then used as reference samples for the SV callers. This process is done to normalise the data and reduce false positives in the output. The SV regions were then subsetted by filtering criteria mentioned in the Methods section. Genetic segregation analysis was not performed for the familial cohort because the selected SV callers did not provide genotype information.

We then adopted a case-control approach as described in the Methods section. For the familial cohort, the resultant output of SV regions was observed in the affected members but absent in the unaffected members and the control set. The resultant qualifying SVs for ten families have been illustrated in Table 27.

For sporadic cohort, the resultant SV regions per individual samples were unique to that individual but absent in control set, as elaborated in Figure 31.

Family identifier	Number of SVs observed	Number of SVs after variant filtration	Number of qualifying SVs after classification
PROP01	324	51	30
PROP02	334	40	7
PROP03	300	31	5
PROP04	289	33	9
PROP05	272	42	6
PROP06	350	46	11
PNET01	586	127	35
PNET02	486	88	7
PNET03	440	130	14
PNET04	558	122	8

Table 27: Tabulated results of SV analysis for the familial cohort. The first column gives the total SVs detected in the affected members of each family but absent in unaffected members of the family. The second column gives the count of variants that passed filtering criteria. The third column gives the count of variants after variant filtration. The fourth column shows the number of SVs classified as VUS or in higher classes and the number of SV regions that involved at least one candidate gene.

Number of SVs after variant filtration

4912

⬇

Number of SVs selected after classification

295

⬇

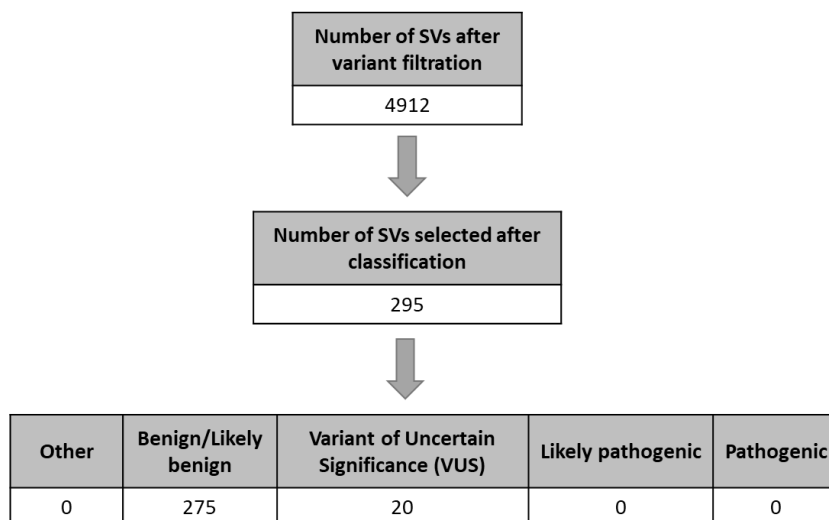| Other	Benign/Likely benign	Variant of Uncertain Significance (VUS)	Likely pathogenic	Pathogenic
0	275	20	0	0

Figure 31: The count of qualifying SVs at each step of variant filtration and prioritisation from sporadic cohort and the number of variants for each classification category.

We could not find any probable SV from both cohorts that would relate to the phenotype. However, SV regions were selected for further evaluation through SV visualization software because they consisted of one or more genes that are also found in candidate genes. These regions have been described in Table 28.

Chromosome	Region start	Region end	SV type	Gene count
16	89596790	89621227	DEL	1
14	78285140	80981295	DEL	6
16	89596929	89621087	DEL	1
6	1.52E+08	1.53E+08	DEL	4
9	1.17E+08	1.22E+08	DEL	20
12	1993497	1993498	BND	1
12	1993751	1993752	BND	1
8	24771776	24771777	BND	1
8	24772077	24772078	BND	1
20	62045407	62045408	BND	1
20	62045549	62045550	BND	1
17	29683598	29683599	BND	1
17	29684214	29684215	BND	1
2	2.38E+08	2.38E+08	DEL	1
2	2.38E+08	2.38E+08	DEL	1
X	69502565	99597249	DUP	131
X	1.35E+08	1.53E+08	DUP	177
16	16248286	16260017	DEL	1

Table 28: Tabulated list of potential SV regions observed in 80 samples. The gene count column explains the count of genes involved in that SV region. DEL stands for Deletion SV, BND stands for Break-end SV, and DUP stands for Duplication SV.

These shortlisted qualifying SV regions were observed in the following genes: *CACNA1B, KIF1A, SPG7, HCN2, TRPM2, NEFM, P2RX1, NF1, KCNQ2, CACNA2D4, ABCC6, ABCB7, ATP7A* and *NRXN3*. These SV regions were found mostly in ion channel genes and needed to be further investigated with biological inference.

Discussion

Painful peripheral neuropathy is a clinically heterogeneous and complex disease involving several characteristics. An in-depth understanding of its pathophysiology is needed to improve the quality of life of individuals affected with painful PN. Thus, the main contribution of this thesis is to identify new genes and genetic variants that can assist in the stratification of affected individuals and shed light on the biological mechanisms involved in painful PN. The application of WES in this study provided the ability to ascertain novel genetic components, which expanded our understanding of painful PN and could also contribute to the identification of potentially new targets for future therapeutics applications.

To better follow the discussion of this thesis and for an exhaustive comprehension of the genetic heterogeneity of PN and NP across familial and sporadic subjects, tables and figures which summarize the main findings will be presented from here on. Moreover, the main findings will be discussed by segmenting them according to the gene families involved in pain and that have been selected in our cohorts due to the presence of several variants within these genes.

SCN9A mutations in different phenotypes

SCN9A gene codes for the voltage-gated sodium channel, alpha subunit 9 (Na$_v$1.7). *SCN9A* is highly expressed in peripheral neurons and subcortical regions of the brain. The genetic variants found in this gene have been associated with primary Erythromelalgia, Congenital Insensitivity to Pain (CIP), Hereditary sensory and autonomic neuropathy, Paroxysmal Extreme Pain Disorder (PEPD) and SFN (McDermott *et al.*, 2019). The heterozygous gain-of-function mutations in this gene tend to hyperexcite the DRG neurons due to recurring and increasing currents, causing painful symptoms. The homozygous loss-of-function mutations have been attributed to CIP.

In the present study, eight *SCN9A* mutations were seen in six families and two sporadic patients. One *SCN9A* variant was shared among a family and a sporadic patient.

A summary of the eight *SCN9A* variants and their location in the channel is described in Figure 32. The summarised variants are then further discussed one by one.
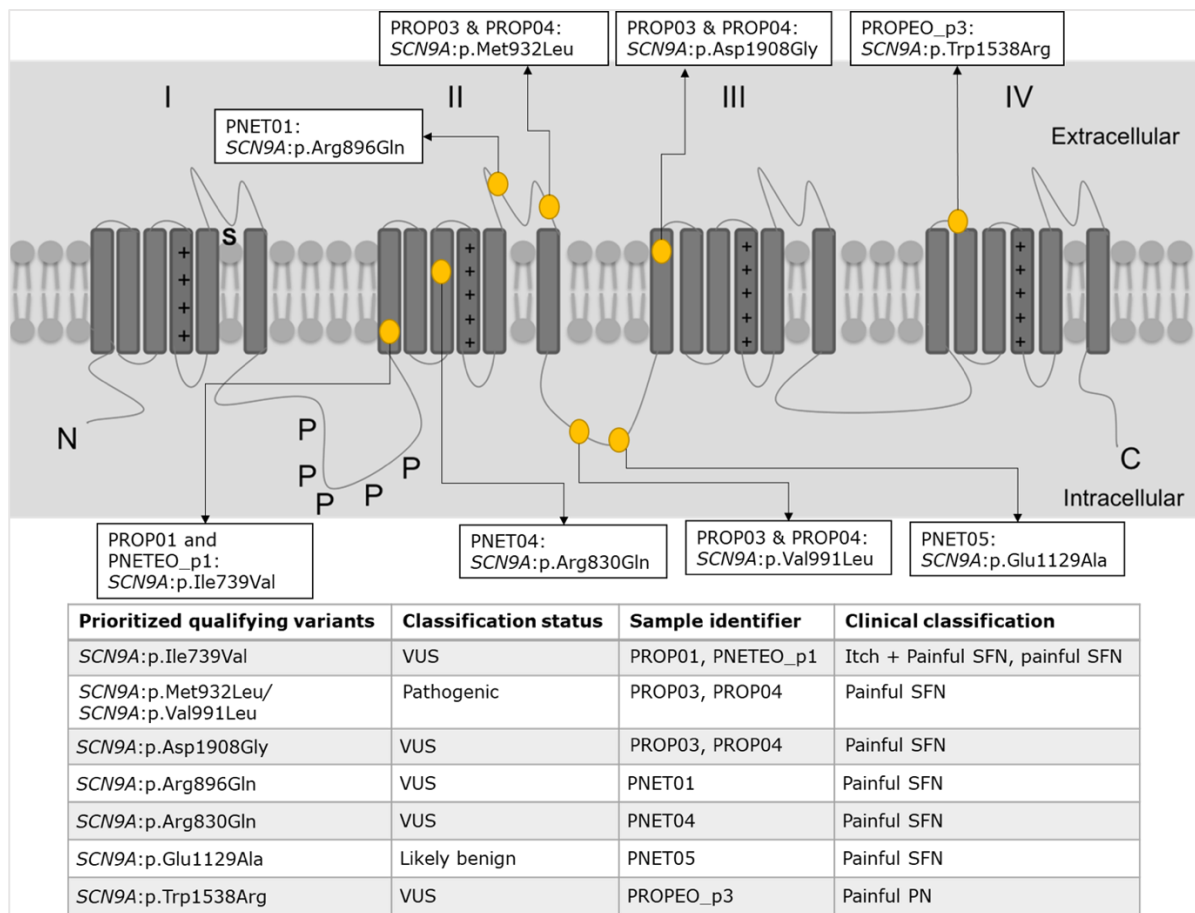
Figure 32: SCN9A channel structure comprises four domains, each with six transmembrane segments. The N represents the amino side of the protein, and the C represents the carboxyl side of the protein. The fourth transmembrane segment is filled with a '+' sign in all domains to denote the positively-charged voltage-sensor of the channel. The 'P' in the loop denotes the phosphorylation site of Protein Kinase A. The location of genetic variants discovered in our study is displayed in yellow circles within the protein structure. The SCN9A mutations discovered in familial and sporadic groups are listed below the structure. Data such as ACMG classification status, family/sample identifier, and clinical classification of that family/subject were compiled. The samples from the sporadic cohort are denoted by the suffix 'PROPEO,' followed by the internal PROPANE study code and 'PNETEO' followed by the internal code for the PAIN-Net study. This picture has been modified from Swanwick et al., 2010.

According to our findings, not all *SCN9A* mutations were found to be pathogenic. One out of eight mutations was classified as 'Likely benign' due to high allele frequency in the global population or because the gene predictably tolerates the protein change.

Two of the resulting variants segregated inconclusively because they were observed in unaffected family members of PROP03 and PROP04. In more detail, the variants *SCN9A*:p.Met932Leu and *SCN9A*:p.Val991Leu are in *cis*, meaning that the presence of these variants in combination affects the gene.

109

When the two variants are combined, the impact is defined as 'Pathogenic,' whereas when they are separated, the impact is classified as 'VUS'. The combination of these variants has been shown to functionally hyperexcite DRG neurons, resulting in increased pain sensitivity in patients (Li *et al.*, 2015).

Along with Met932Leu + Val991Leu, a third heterozygous *SCN9A* mutation, *SCN9A*:p.Asp1908Gly, was discovered in the PROP03 and PROP04 families. Two publications, Li *et al.* and Zeberg *et al.*, stated that these three mutations (Met932Leu + Val991Leu + Asp1908Gly) resulted in individuals with a predisposition to NP. These studies found that subjects with these mutations are more likely than controls to experience pain (odds ratio >1.2). Two out of 4 individuals with these mutations in our cohort were asymptomatic to pain. They were reanalysed without the candidate gene filter in order to find potentially protective variants. Despite the presence of gain-of-function mutations in these subjects, they had normal pain sensitivity. This disparity in genotype-phenotype in subjects could be explained by differences in variant frequency between the population studied in published studies and the population studied in the analysed population. (Li *et al*, 2015; Zeberg *et al*, 2020). At the same time, we cannot rule out the possibility that the pain assessment was made at the time of sampling, but that the phenotype related to pain could change over time. The published evidence that the presence of three variants (Met932Leu + Val991Leu + Asp1908Gly) leads to a predisposition to develop pain is reflected in our findings in the pain-affected subjects in our cohort.

We observed a loss-of-function mutation in a subject (PNET01-III.11) who was symptomatic for NP. This variant, *SCN9A*:p.Arg896Gln, was found in one out of nine affected members in the PNET01 family and it has been reported as a loss-of-function mutation in two studies. These studies showed that in conjunction with another variant (in compound heterozygous state) or a homozygous state, the variant causes CIP (Cox *et al.*, 2010; Marchi *et al.*, 2018). In the studied cohort the variant, in the heterozygous state, matches with painful phenotype, with no second mutation on the same gene. Marchi *et al.* also reported that the individuals with Arg896Gln mutation in the heterozygous state had a normal medical history. At the same time, Cox et al. conducted an electrophysiological study that demonstrated that this variant completely truncates *SCN9A* gene function rather than contributing to neuron hyperexcitability. Given the contradictions, we hypothesised that this variant was most likely not contributing to the phenotype.

The other gain-of-mutation observed in our cohort is in position 739, where the amino acid changes from Isoleucine to Valine in *SCN9A* protein, *SCN9A*:p.Ile739Val. For this particular variant, functional studies have shown hyperexcitability of DRG neurons associating the variant with painful SFN (Faber *et al.*, 2012a; Han *et al.*, 2012; Devigili *et al.*, 2014). This variant segregates conclusively in the PROP01 family with complete penetrance (all affected members have the variant). However, another study with European ancestry subjects affected with SFN reported it to be segregating inconclusively in two families (Eijkenboom *et al.*, 2019). This variant was found in subjects with a variety of phenotypes, including: i) affected members in PROP01 family with itch attacks and painful SFN, ii) PNETEO_p1, a sporadic case without itch and affected by painful SFN; in addition, a study by Faber et al. revealed that subjects with Ile739Val variants had several autonomic complaints. These observations increase the possibility that other genetic components along with *SCN9A* mutation might be contributing to the symptoms. These different outcomes could result from sub-population differences or heterogeneity in recruitment practices.

A novel *SCN9A*:p.Arg830Gln variant was observed in the trio family PNET04, with the affected daughter and the unaffected father both harbouring the same variant. The characteristics of this variant are not available in the literature, given that no functional studies were performed for it. However, at the same amino acid position in the protein a stop gain variant has been previously reported for CIP (Ramirez *et al.*, 2014). This mutation is associated with a premature translational stop signal in the *SCN9A* gene, causing insensitivity to pain. Thus, more investigation is needed about the nature of Arg830Gln variant before formulating a conclusion of its impact on the phenotype. It would be prudent to note that, though the father in the trio had no medical complaints, he did not undergo skin biopsy to rule out the presence of SFN completely. We concluded that we would select this variant for further functional studies.

The *SCN9A*:p.Glu1129Ala variant was classified as 'Likely benign'. This variant lies in a repeat region of the *SCN9A* channel, which is weakly conserved through species. Due to this poor conservation, it is less likely to have a deleterious effect on the protein. Furthermore, this variant had an incomplete penetrance because it was not present in all affected members. However, a functional study is required to confirm it as a benign variant.

Lastly, *SCN9A*:p.Trp1538Arg variant, found in a painful PN sporadic patient, has been previously studied and was seen in subjects with Erythromelalgia (Cregg *et al.*, 2013). An electrophysiological study for this variant depicted a gain-of-function mutation. Still, its hyperexcitability of DRG neurons was not as prominent due to lesser action potential firing per milliseconds compared to other gain-of-function *SCN9A* mutations (Cregg *et al.*, 2013). The subject, PROPEO_p3, with this variant had autonomic dysfunctions and painful PN. We theorise that this variant may be partially responsible for the symptoms observed in PROPEO_p3.

In our cohort, loss-of-function mutations were found in painful PN subjects, while subjects harbouring gain-of-function mutations did not have any neurological symptoms. Therefore, we speculate the co-presence of other mutations that, in combination with *SCN9A,* can fully explain the phenotypic variability in our cohort. In this regard, a recent publication described the presence of the *KCNQ3* gene variant and a *SCN9A* gain-of-function mutation in a subject with no pain, whereas the subject's daughter only had the *SCN9A* mutation suffered with severe pain (Yuan *et al.*, 2021). The authors of this study demonstrated that the presence of both the *KCNQ3* and *SCN9A* mutations contributed to pain sensitivity. Meanwhile, only the presence of the *SCN9A* mutation resulted in severe pain with SFN (Yuan *et al.*, 2021). This finding opened a new avenue to variant interpretation and could explain the phenotypic variability observed within *SCN9A* mutations

Co-presence of mutations in SCN9A and CACNA1E and KCNK18 channels in painless familial subjects

The *CACNA1E* gene encodes for Ca$_v$2.3 channel that is a voltage-dependent calcium channel. These channels are involved in releasing neurotransmitters during synaptic transmission in the CNS. The genetic variants in this gene are associated with developmental and epileptic encephalopathies (Helbig *et al.*, 2018). *CACNA1E* genetic variants were observed in PROP04 and PNET04 families, as depicted in Figure 33a. Figure 33a illustrates that in both the families, the fathers are asymptomatic for NP and harbour variants in VGSC genes along with variants in *CACNA1E* gene.
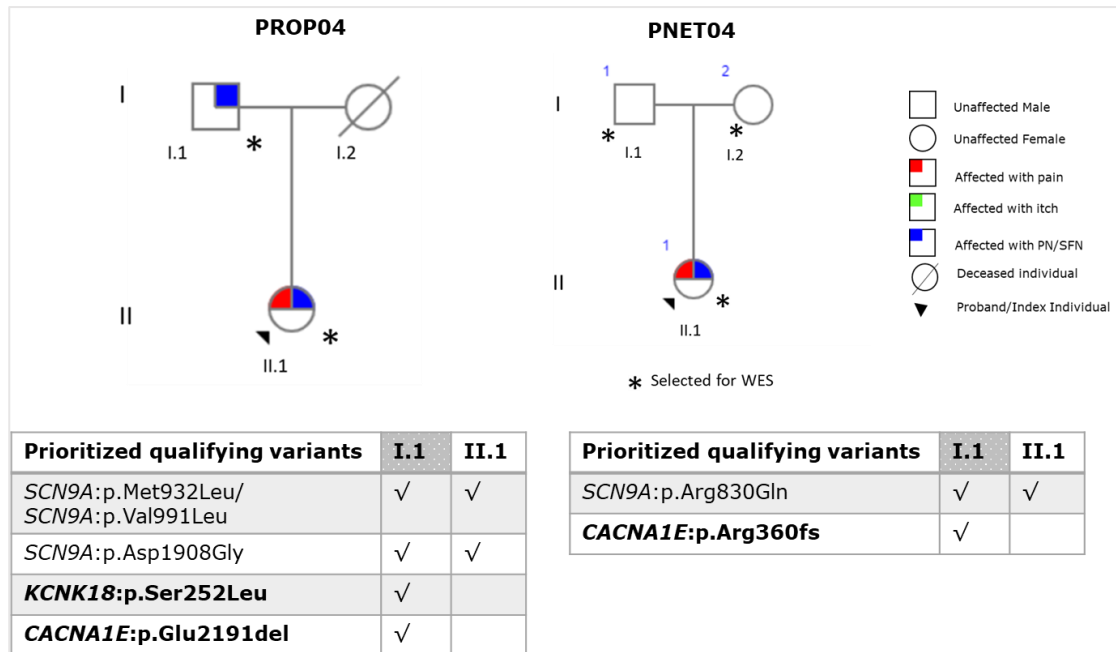
Figure 33a: PROP04 and PNET04 family pedigrees, as well as a list of qualifying variants found in these families. The variant segregation within the individuals is shown in the tables below the pedigree for each family. The gene that may contribute to pain sensitivity is highlighted in bold letters. Individuals with a grey individual identifier are painless.

CACNA1E was selected because it is known to correlate negatively with painful conditions, and it is associated with variable pain profiles in published studies. For instance, Polli et al. showed that the miRNAs related to the CACNA1E gene had a reduced expression in subjects with Fibromyalgia, explaining a reduced function of the channel. A study by Amano et al. demonstrated that the carriers of a single nucleotide polymorphism (rs3845446) in the CACNA1E gene had contradictory post-operative pain profiles. The carriers of the rs3845446 variant in 351 patients who underwent orthognathic surgery required less opioid medication, whereas the 112 individuals who underwent gastrointestinal surgery required more opioids due to severe pain, depicting the aforementioned contradiction in pain sensitivity. Moreover, a mouse model study on the CACNA1E gene displayed that complete ablation of CACNA1E resulted in normal responses for pain stimulus and lessened reaction to somatic inflammatory pain. This study also demonstrated that CACNA1E $^{+/-}$ heterozygous mice showed decreased reaction to visceral inflammatory pain (Saegusa et al., 2002). This mouse model study concluded that the loss-of-function mutations in this gene might be contributing to pain insensitivity.

Both the fathers (PROP04-I.1 and PNET04-I.1) harboured a variant in the CACNA1E gene where CACNA1E:p.Arg360fs is located in the linker region between domain I and domain II and CACNA1E:p.Glu2191del is located in the cytoplasmic area linked to domain IV at the carboxyl side, as shown in Figure 33b.
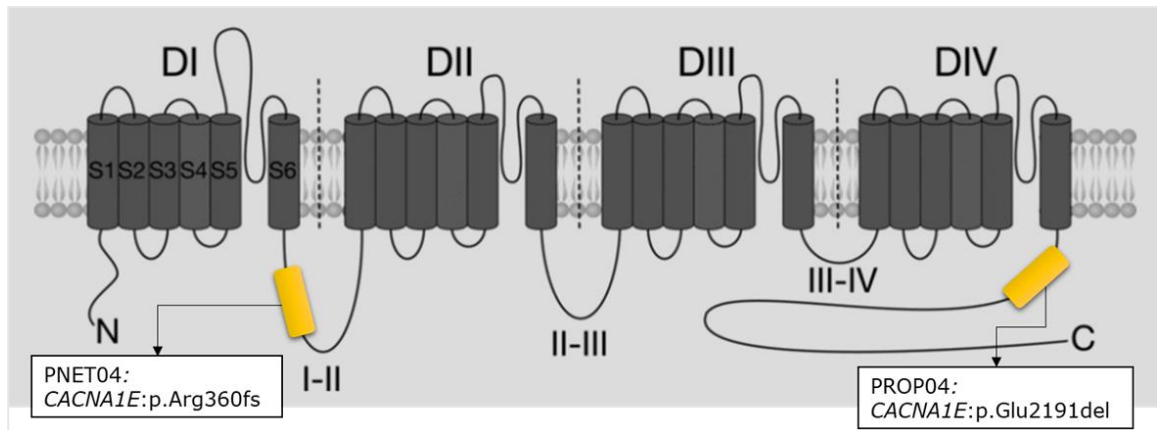
113

Figure 33b: A pictorial representation of the CACNA1E channel consisting of four domains (DI – DIV) with six segments, each numbered as S1 to S6 in domain one (DI). The domains are then linked with loops shown as I-II, II-III and III-IV. N stands for amino-terminus, and C stands for carboxyl-terminus. The location of the genetic variants from PROP04 and PNET04 is illustrated with yellow rectangles. This picture has been modified from Weiss & Zamponi, 2021

The *CACNA1E*:p.Glu2191del, seen in PROP04-I.1, is a variant located outside the hotspot region (the genic region where mutations tend to have a deleterious effect on the protein) and might have an insignificant impact. In addition to *CACNA1E*, PROP04-I.1 presents another variant in *KCNK18*:p.Ser252Leu in a potassium channel, which could contribute to pain insensitivity.

The *KCNK18* gene encodes for a two-pore domain potassium channel, and this gene is mainly expressed in the brain. Mutations in this gene have been associated with migraine with aura (Lafrenière *et al.*, 2010). This particular genetic variant, Ser252Leu, is located at this channel's phosphoryl regulatory site, which alters the channel activity (Pavinato *et al.*, 2021). As a result, we hypothesise that the presence of these two mutations in tandem may be protective for NP.

The other *CACNA1E*:p.Arg360fs was observed in PNET04-I.1. This variant is a loss-of-function mutation that induces a premature translational stop signal, which shortens the protein. This variant has a disruptive impact on the protein. As Saegusa *et al.* demonstrated, the effects of ablated *CACNA1E* gene led to a normal reaction to pain stimulus, and this null variant also contributed to pain resilience.

According to these hypotheses, the presence of *CACNA1E* gene mutations, either alone or in combination with other genes, has a negative impact on pain transmission. As a result, all three genetic variants (two from the *CACNA1E* gene and one from the *KCNK18* gene) will be validated further with Sanger sequencing and considered for the electrophysiological study.

Another, *CACNA1E*:p.Val58Ile variant was also observed in PROP02-II.2, associated with itching and PN phenotype, but not with pain. Whereas his mother, PROP02-I.2, exhibited pain. However, the familial segregation of this variant was inconclusive, so it was not chosen for further investigation.

KCNQ4 and KCNE2 genes

The largest family of the cohort, PNET01, had mutations in two potassium channel genes. One gene, *KCNE2*, corresponded to painful phenotype and the other *KCNQ4* to painless phenotype.

One of the mutations found was in the *KCNE2* gene, which encodes for 'Minimum potassium ion channel-related peptide 1'. This gene is expressed in heart muscle, thyroid and DRG neurons (Liu et al., 2014). The variants in this gene have been associated with Long QT syndrome, a cardiac condition that can cause fast, chaotic heartbeats (Liu *et al.*, 2014). The *KCNE2* genetic variant in our study has been associated with Long QT syndrome according to a digenic autosomal dominant inheritance and has been classified as a risk factor for sporadic atrial fibrillation (Heida *et al.*, 2019; Nielsen *et al.*, 2014). But, in the case of the PNET01 family, it was seen segregating with incomplete penetrance, meaning that it is not seen in all affected family members.

The other mutation was observed in the *KCNQ4* gene in the proband (III.3) and two unaffected members (III.2 and IV.3) of the family. The *KCNQ4* gene encodes for a voltage-gated potassium channel, subfamily Q, member 4 mediating the electrical signalling. This gene is expressed in ciliary and trigeminal ganglions and has been associated with autosomal progressive deafness (Kharkovets *et al.*, 2000).

These two variants were selected because they were observed in the proband (III.3). But, the *KCNE2* variant was causative for NP while the *KCNQ4* variant was protective for NP, as depicted in Table 29.

Prioritised qualifying variants	II.5	III.1	III.3	III.4	III.5	III.11	III.2	IV.3
KCNE2:p.Ile57Thr		√	√	√		√		
KCNQ4:p.Glu551fs			√				√	√

Table 29: List of qualifying variants observed in the proband of PNET01 family and the segregation of these variants in other family members. The unaffected members are in the last two columns shaded in grey colour.

The genetic variants in the *KCNE2* and *KCNQ4* genes have yet to be associated with painful PN. Moreover, Strutz-Seebohm *et al.* demonstrated that *KCNE* beta subunits are known to modulate the properties of *KCNQ4*.

We hypothesise that the impact of *KCNQ4* might be affected in the proband due to the presence of the *KCNE2* gene. Also, the *KCNE2* genetic variant is seen in III.11 (the cousin of the proband) diagnosed with painful SFN. This reinforces the notion that the *KCNE2* genetic variant might be causative for NP. Simultaneously, the presence of only a *KCNQ4* mutation in unaffected members may contribute to a painless phenotype. These variants will be further validated by Sanger sequencing.

SCN10A mutations in a sporadic cohort

SCN10A encodes for voltage-gated sodium channel, subunit alpha 10 ($Na_v1.8$). This channel is resistant to tetrodotoxin (TTX) and is widely expressed in peripheral sensory neurons (Swanwick *et al.*, 2010). The genetic variants in this gene have been previously linked with Familial Episodic Pain (Faber *et al.*, 2012b). Eijkenboom *et al.* reported that in their cohort of 1,139 patients affected by SFN, the prevalence of *SCN9A* variants was 5.1% and for *SCN10A* was 3.7%. Comparatively, in our sporadic cohort of 43 subjects, *SCN9A* variants' prevalence was 7.1 % and *SCN10A* variants' prevalence was 11%. The occurrence of *SCN10A* variants was relatively higher in our cohort than in the previously published study. This disparity could be attributed to sub-population differences (e.g., Dutch versus Italian) or cohort recruitment selection criteria.

The location of a variant in the *SCN10A* channel determines the regulation of this channel. These *SCN10A* qualifying variants, their location in the channel and the subjects mutated are depicted in Figure 34.
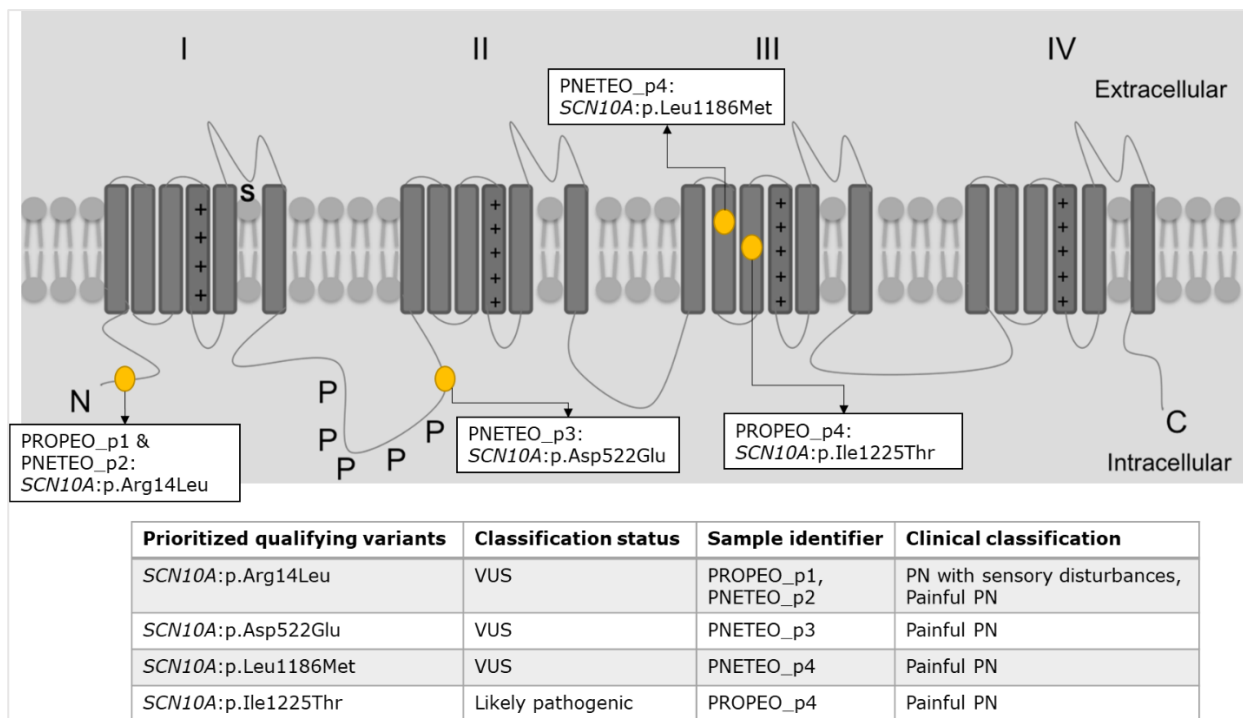
Prioritized qualifying variants	Classification status	Sample identifier	Clinical classification
SCN10A:p.Arg14Leu	VUS	PROPEO_p1, PNETEO_p2	PN with sensory disturbances, Painful PN
SCN10A:p.Asp522Glu	VUS	PNETEO_p3	Painful PN
SCN10A:p.Leu1186Met	VUS	PNETEO_p4	Painful PN
SCN10A:p.Ile1225Thr	Likely pathogenic	PROPEO_p4	Painful PN

Figure 34: SCN10A channel structure comprises four domains each with six transmembrane segments. The N represents the amino-side of the protein, and the C represents the carboxyl side of the protein. The fourth transmembrane segment is filled with a '+' sign in all domains to denote the positively-charged voltage-sensor of the channel. The 'P' in the loop denotes the phosphorylation site of Protein Kinase A. The location of genetic variants discovered in our study is displayed with yellow circles within the channel structure. The SCN10A mutations discovered in familial and sporadic groups are listed below the channel structure. Data such as ACMG classification status, family/sample identifier, and clinical classification of that family/subject were compiled. The sporadic cohort samples are denoted by the suffix 'PROPEO' followed by the internal code for the PROPANE study and 'PNETEO' followed by the internal code for the PAIN-Net study. This picture has been modified from Swanwick et al., 2010.

The variants located in the transmembrane segments tend to sensitise the channel and, in turn, invoke pain. There are two variants in our cohort, *SCN10A*:p.Leu1186Met in PNETEO_p4 and *SCN10A*:p.Ile1225Thr in PROPEO_p4, with painful PN phenotype, which are located in the transmembrane segments of domain III of this channel. Out of these two variants, the amino acid change Ile1225Thr was reported to be 'Likely pathogenic' because this variant segregated conclusively in a family affected by SFN (Eijkenboom *et al.*, 2019). The other two variants are *SCN10A*:p.Arg14Leu in PROPEO_p1 and PNETEO_p2, and *SCN10A*:p.Asp522Glu in PNETEO_p4 associated with painful neuropathy. These two variants were also found in the putative region of the channel and may be involved in channel sensitisation. However, the physiological mechanisms of these four putative variants have yet to be functionally investigated. As a result, these variants will be further validated through Sanger sequencing and an electrophysiological study. Furthermore, it would be interesting to learn about the differences in the prevalence of VGSC genes in the studied population against other European populations.

Transient Receptor Potential (TRP) channel mutations

TRP channel is a superfamily of genes involved in nociception and pain signalling (Nilius & Flockerzi, 2014b). In this study, we found that patients with painful PN have mutations in TRP channel genes, specifically *TRPM8*, *TRPA1*, and *TRPM2*. So, in combination with other genetic variants, they could play a role in the modulation of pain.

TRPM8 channel mutations

TRPM8 stands for Transient Receptor Potential Melastatin, member 8; it is known as cold-menthol receptor as it is activated by cold temperature and menthol (Farooqi *et al.*, 2011). *TRPM8* gene is expressed in small DRG neurons, and the genetic variants have been associated with cold-like pain (Soeda *et al.*, 2021). We observed two variants in *TRPM8* in two members of the PROP02 family affected with itch and PN but not NP. The prioritised variants are seen only in the painful members, as depicted in Figure 35.

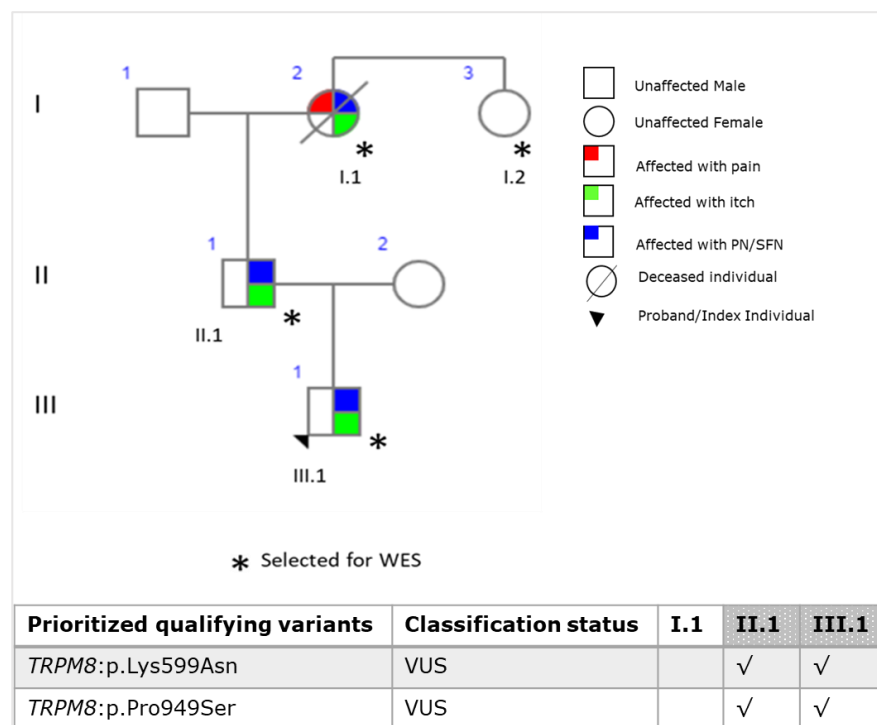| Prioritized qualifying variants	Classification status	I.1	II.1	III.1
TRPM8:p.Lys599Asn	VUS		√	√
TRPM8:p.Pro949Ser	VUS		√	√

Figure 35: Pedigree of the PROP02 family. Table with qualifying variants and their ACMG classification status displayed below the pedigree. Grey column headers denote individuals with TRPM8 mutations who have not been diagnosed with NP.

The two amino acid changes in the *TRPM8* gene, Lys599Asn and Pro949Ser, are in a region that may alter protein function. Lys599Asn is localised in a conserved region near the N-terminus of the channel. Pro949Ser is situated in the TRP domain near the C-terminus. Specifically, Pro949Ser falls inside the TRP domain, a site for phospholipid phosphoinositol 4,5-bisphosphate (PIP_2) binding. The PIP_2 molecule modulates the *TRPM8* channel activity upon cold stimulus or by *TRPM8* agonists while inducing pain (Latorre *et al.*, 2011; Daniels *et al.*, 2009). The presence of a variant Pro949Ser in this region could affect the activation of this channel (Moore *et al.*, 2018). At the same time, the presence of another variant Lys599Asn in the conserved region of the channel could compromise its function completely (Morgan *et al.*, 2015). As the individuals carrying these variants do not experience pain, we postulated that these *TRPM8* variants might downregulate pain perception together. De Caro *et al* demonstrated that *TRPM8* inhibition causes significant reduction in pain perception and has an analgesic effect.

Moreover, there has been evidence of menthol, a *TRPM8* agonist, effective against pruritus. We hypothesise that *TRPM8* participates in itch attacks (Kardon *et al.*, 2014). However, a clear mechanistic role of these variants in the *TRPM8* gene is yet to be examined functionally.

Finally, the coexistence of these two TRPM8 mutations may have analgesic effects on the affected members of this family. Thus, it would be useful to evaluate the involvement of these variations in the TRPM8 channel and, by extension, pain perception. These variants will be further validated with Sanger sequencing and considered for electrophysiological tests.

TRPA1 channel mutations

TRPA1 (Transient Receptor Potential cation channel, subfamily A, member 1) is expressed in primary afferent neurons. This channel has a long ankyrin repeat domain at the N-terminus. Mutations in the *TRPA1* gene localised in the repeat domain of the channel structure are known to sensitise the channel invoking pain (Meents *et al.*, 2017). The published genetic variants in *TRPA1* have been associated with familial episodic pain (Kremeyer *et al.*, 2010). We observed two different *TRPA1* variants located in the ankyrin-repeat domain of the *TRPA1* channel in members affected with painful SFN in PROP06 and PNET06 families. The family pedigrees and the respective variant segregation is shown in Figure 36a, whereas location of the variants in *TRPA1* channel is shown in 36b.

Prioritized qualifying variants	Classification status	I.1	I.2	II.1	II.2
TRPA1:p.Asn109Lys	Likely benign		√		√

Prioritized qualifying variants	Classification status	I.1	I.2	II.1	II.2
TRPA1:p.Thr311Asn	VUS		√		√

Figure 36a: Pedigrees of the PROP06 and PNET06 families. Tables with qualifying variants and their ACMG classification status are shown below each pedigree. Grey column headers denote individuals with TRPM8 mutations who have not been diagnosed with NP.

Figure 36b: The pictorial representation of the TRPA1 channel consisting of six transmembrane segments (S1-S6) where N stands for amino terminus, C stands for carboxyl terminus. This channel consists of ankyrin repeat domain and cysteine residues at the N-terminus. The location of the genetic variants observed in PROP06 and PNET06 families is displayed with yellow circles within the protein structure. This picture has been modified from Lapointe & Altier, 2011

The *TRPA1*:p.Asn109Lys variant seen in the PROP06 family was classified as 'Likely benign' because the *in silico* tools predicted that the gene tolerates the amino acid change in that position. We can speculate that the variant is partially responsible for the symptoms because it is located in a hotspot region, which is a region that may be involved in altering protein function. Furthermore, no functional studies have been conducted to support the classification of this variant as 'Benign'. We theorise that the variant may be a risk factor or an enhancer contributing to the painful stimuli in these patients.

The other *TRPA1*:p.Thr311Asn variant observed in the PNET06 family had an incomplete penetrance as two out of three affected members shared this variant. This *TRPA1* variant was predicted as deleterious by the *in silico* prediction algorithms. This variant was also seen in the ankyrin-repeat domain, a hotspot region of this channel. As explained before, we also hypothesise that this variant may contribute to a painful phenotype.

Investigating the two *TRPA1* variants discovered in our cohort could shed light on *TRPA1* channel sensitisation and aid in genetic screening for pain. Consequently, Sanger sequencing will be used to validate these variants further.

TRPM2 channel mutations

TRPM2 stands for Transient Receptor Potential Melastatin, member 2, and it is highly expressed in brain and peripheral blood cells such as neutrophils (Farooqi *et al.*, 2011; Nilius & Flockerzi, 2014b). No known genetic variants in the *TRPM2* gene have been associated with any human disease. This gene has been widely associated with NP through several animal models (Matsumoto *et al.*, 2016; Chung *et al.*, 2015). For example, Haraguchi *et al.* demonstrated that allodynia and hyperalgesia were diminished in *TRPM2* knock-out mice models.

In our cohort, the *TRPM2* genetic variants were observed in two families in PNET01 and PNET06 segregating according to pain phenotype. The *TRPM2*:p.Pro1102Ser had incomplete penetrance in PNET01 but the *TRPM2*:p.Arg411Trp variant in PNET06 segregated with complete penetrance. The variants were prioritised because of their location in the channel. One of the two genetic variants, Arg411Trp, is localised in the conserved region on the N-terminus side, and the other Pro1102Ser is near the pore region between segments five and six, as depicted in Figure 37.
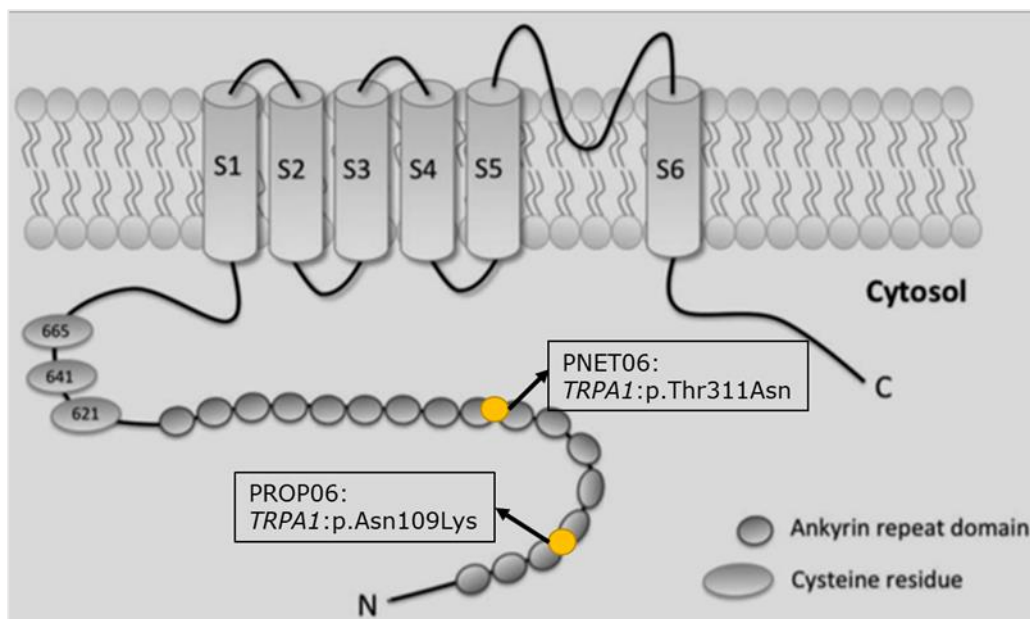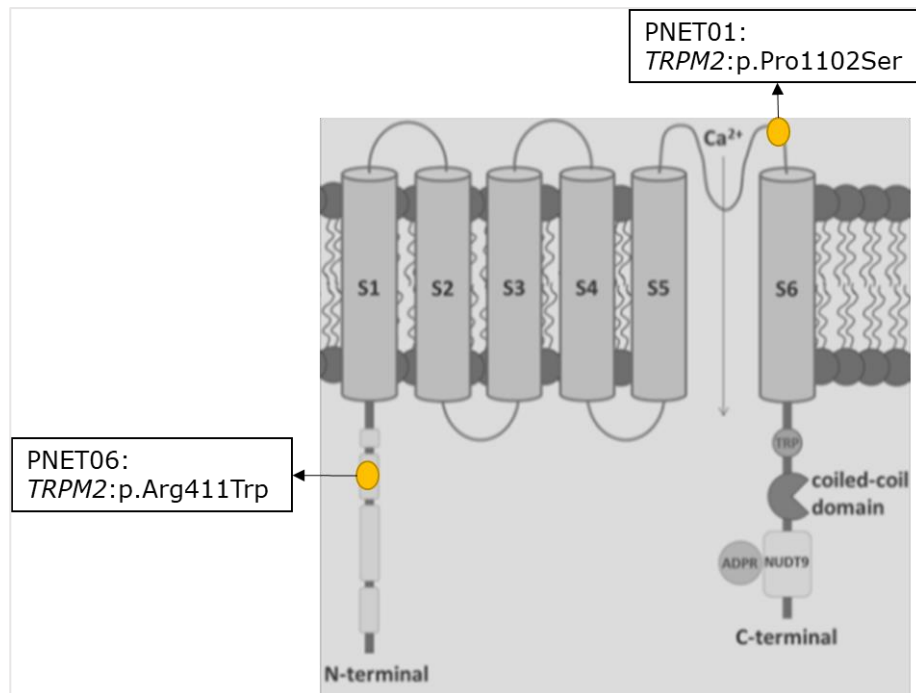
Figure 37: A pictorial representation of the TRPM2 channel consisting of six transmembrane regions (S1-S6) where N-terminal stands for the amino end of the protein and C-terminal is the carboxyl-side of the protein. The squares on the N-terminal represent the conserved region. Circle with TRP stands for transient receptor potential motif at the C-terminal along with coiled-coil domain and Nudix hydrolase (NUDT9) substrate attached with Adenosine diphosphate ribose (ADPR) (Perraud *et al.*, 2001)*. The location of the genetic variants observed in PNET01 and PNET06 families is displayed with yellow circles within the protein structure. This picture has been adapted from Nilius & Flockerzi, 2014b*

Inside the conserved region, a calmodulin-binding domain from 406-416 amino acid aids in the *TRPM2* channel activation (Tong *et al*, 2006). As the Arg411Trp variant falls into this region, we hypothesise that it possibly modulates the channel into inducing pain. The other Pro1102Ser variant lies near a pore region involved in calcium ion homoeostasis. *TRPM2* channel sensitisation depends on the calcium ion influx. It was then postulated that the Pro1102Ser variant might be involved in the activation of this channel.

In conclusion, the *TRPM2* gene has been frequently linked to NP, but no mutations in this gene have been associated with painful PN. *TRPM2* genetic variants were effectively segregated in two families, increasing the likelihood of *TRPM2* being one of the genetic causes of painful PN. Thus, a clear mechanistic role for these mutations must be investigated further with functional studies.

ATPase copper transporting beta (ATP7B) genetic variants

We observed an enrichment of heterozygous mutations in the ATPase copper transporting beta gene in familial and sporadic cohorts. This gene is expressed in the skin, skeletal muscle, liver, and brain.

This gene is associated with an autosomal recessive disease called Wilson's disease, characterised by high levels of copper accumulated in the liver, brain (CNS, hippocampus and cerebellum), and cornea of an affected individual. The symptoms manifest as liver dysfunction and neurological defects (Bandmann *et al.*, 2015). Additionally, cases with symptoms related to demyelinating polyneuropathy have been published in the presence of mild symptoms of Wilson's disease (Jung *et al.*, 2005). Moreover, a cross-sectional comparative study demonstrated a significant reduction in corneal fibres in patients affected by Wilson's disease using corneal confocal microscopy. They concluded that these alterations in corneal fibres indicate the development of small fibre peripheral neuropathy (Sturniolo *et al.*, 2015).

The *ATP7B* gene is also associated with MEDNIK syndrome. MEDNIK stands for Mental retardation, Enteropathy, Deafness, peripheral Neuropathy, Ichthyosis and Keratodermia (Vernon, 2020). This disease manifests itself due to abnormal copper metabolism and mutations in the Adaptor-Related Protein Complex 1, Sigma-1 subunit (*AP1S1*) gene (Montpetit *et al.*, 2008). This gene encodes for small subunit of the Adaptor-related Protein complex 1 (AP-1) (Incecik *et al*, 2018). In more detail, neurons are highly polarized cells with unique axonal and somatodendritic protein compositions. AP-1 inactivation prevents somatodendritic sorting of the relevant proteins, resulting in non-polarized distribution in rat hippocampus neurons (Jain *et al*, 2015a). Jain *et al*, demonstrated that the AP-1 protein is involved in the polarized sorting of the copper ion transporter *ATP7B* to rat hippocampus neurons' somatodendritic region. They also discovered that changes in copper ion levels in neurons can alter the polarity of *ATP7B*, resulting in aberrant copper metabolism. As a result, a link between the *ATP7B* gene and the MEDNIK syndrome was established.

Based on these studies, the variants in the *ATP7B* gene were prioritised in the PROP05 family and three sporadic subjects, all affected by painful SFN and in the PNET03 family affected by SNF and itch, as shown in Figure 38.

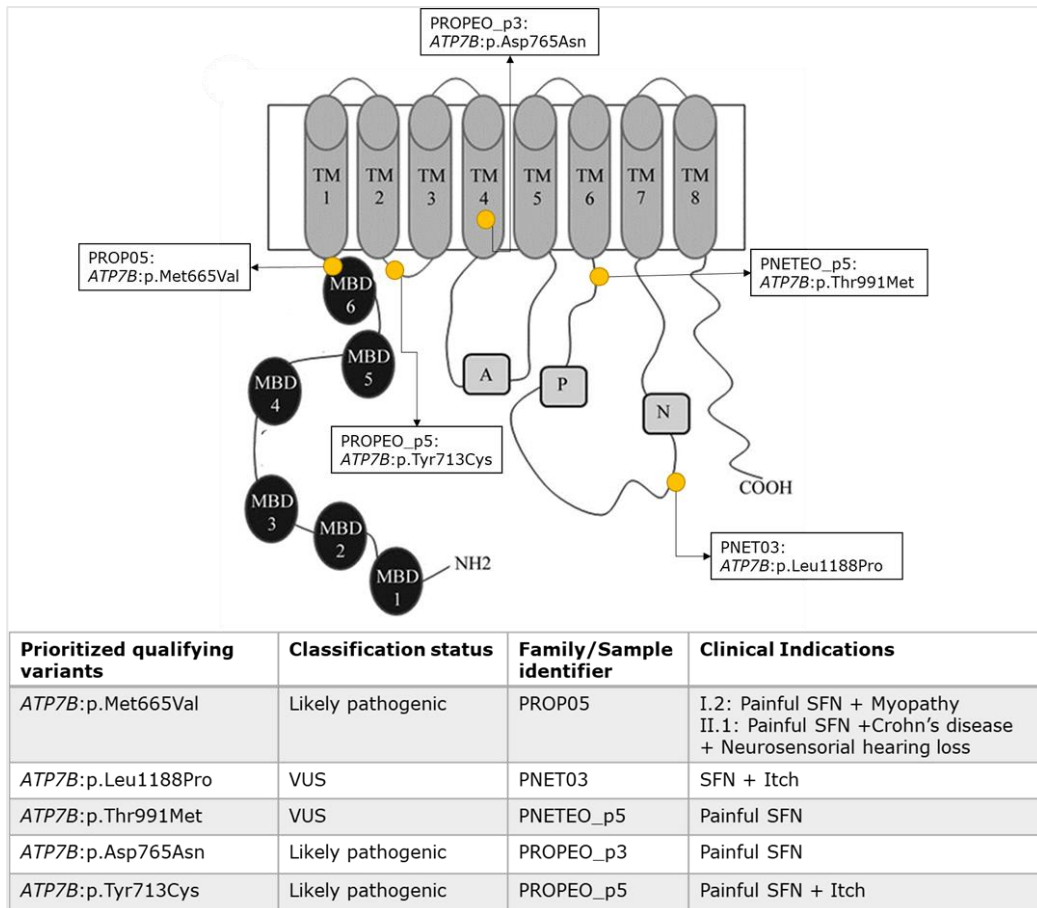| Prioritized qualifying variants	Classification status	Family/Sample identifier	Clinical Indications
ATP7B:p.Met665Val	Likely pathogenic	PROP05	I.2: Painful SFN + Myopathy II.1: Painful SFN +Crohn's disease + Neurosensorial hearing loss
ATP7B:p.Leu1188Pro	VUS	PNET03	SFN + Itch
ATP7B:p.Thr991Met	VUS	PNETEO_p5	Painful SFN
ATP7B:p.Asp765Asn	Likely pathogenic	PROPEO_p3	Painful SFN
ATP7B:p.Tyr713Cys	Likely pathogenic	PROPEO_p5	Painful SFN + Itch

Figure 38: A pictorial representation of ATP7B protein structure consisting of eight transmembrane segments (TM1-TM8) with NH2 denoting the amino-side of the protein and COOH the carboxyl side. Circle with MBD represents metal-binding domains (MBD1-MBD6) at the N-terminus. The squares represent other domains: the A-actuator domain, P-phosphorylation domain and N-nucleotide binding domain. The location of genetic variants observed in our cohort is displayed with yellow circles within the protein structure. The ATP7B mutations discovered in familial and sporadic groups are listed below the channel structure. Data such as ACMG classification status, family/sample identifier, and clinical classification of that family/subject were compiled. The sporadic cohort samples are denoted by the suffix 'PROPEO' followed by the internal code for the PROPANE study and 'PNETEO' followed by the internal code for the PAIN-Net study. The protein structure image has been modified from (Hasan et al., 2012)

The *ATP7B*:p.Met665Val variant found in the PROP05 family, segregating with pain and SFN, is localised between the sixth metal-binding domain and the first transmembrane segment. This sixth metal-binding domain is the receiver of copper ions. The metal-binding domain region of *ATP7B* regulates the resilience towards cisplatin; it has been demonstrated that a variant near these domains could result in failed protein function due to aberrant conformation of the protein (Safaei *et al.*, 2012). We theorised that this variant might be critical for copper metabolism. As Met665Val change has not been reported, we hypothesise it as possibly causative for the phenotype.

The *ATP7B*:p.Leu1188Pro was observed in the PNET03 family with the phenotype of SFN and itching. The variant *ATP7B*:p.Leu1188Pro is located at the ATP binding site, an important site for the correct protein function.

The genetic variants found in this site affect the protein's function (Kenney & Cox, 2007). According to Machado *et al.*, an *ATP7B*:p.His1069Gln variant located at the ATP loop, in heterozygous form, was observed in a patient with sensory disturbances. They also speculated that subjects with this variant had a late-onset of the disease and neurological defects. Moreover, this variant is observed in a family with itch and no pain. Studying this variant could provide insights into neuropathic itch. Due to its location in the gene and its conclusive segregation in the PNET03 family with respect to SFN and itch, this variant is likely to be pathogenic. It should be considered for further functional studies.

The *ATP7B*:p.Asp765Asn and *ATP7B*:p.Thr991Met variants were observed in sporadic cases with painful SFN. The location of Asp765Asn is inside the transmembrane domain number four, which is critical for copper transportation, while Thr991Met is near the transmembrane domain number six. According to the literature, Asp765Asn would lead to protein misfolding and cellular mislocalisation, whereas Thr991Met would compromise copper transport (Zhu *et al.*, 2013; Kenney & Cox, 2007). We theorise that the altered copper metabolism may partially contribute to PN, but a proper functional study is necessary to confirm it.

Lastly, the *ATP7B*:p.Tyr713Cys variant is observed in a sporadic case with widespread pain, itch, and SFN. This variant is localised in the linker loop between transmembrane domains numbers two and three. UniProt database reported this variant as 'Likely pathogenic' (Bateman *et al.*, 2021). Huster *et al.* stated that the loop between transmembrane two to three is a region that predominately affects copper transportation. This variant will also be included as a potential candidate for further study.

In summary, we have observed five mutations in the *ATP7B* gene that localise at crucial junctions of the gene. According to the literature, the localisation of these mutations can result in dysregulated copper metabolism. Thus, it would be pertinent to study the role of abnormal copper metabolism in PN.

Nicotinic acetylcholine receptor signalling pathway

The pathway analysis done for familial and sporadic cohorts revealed an intriguing shared pathway enriched between the cohorts based on PN & NP or PN & itch phenotype: the 'Nicotinic acetylcholine receptor signalling pathway' (Panther code: P00044). The pathway consists of 93 genes and 13 components. As shown in figure 39, this pathway is involved in four different compartments and 13 components.
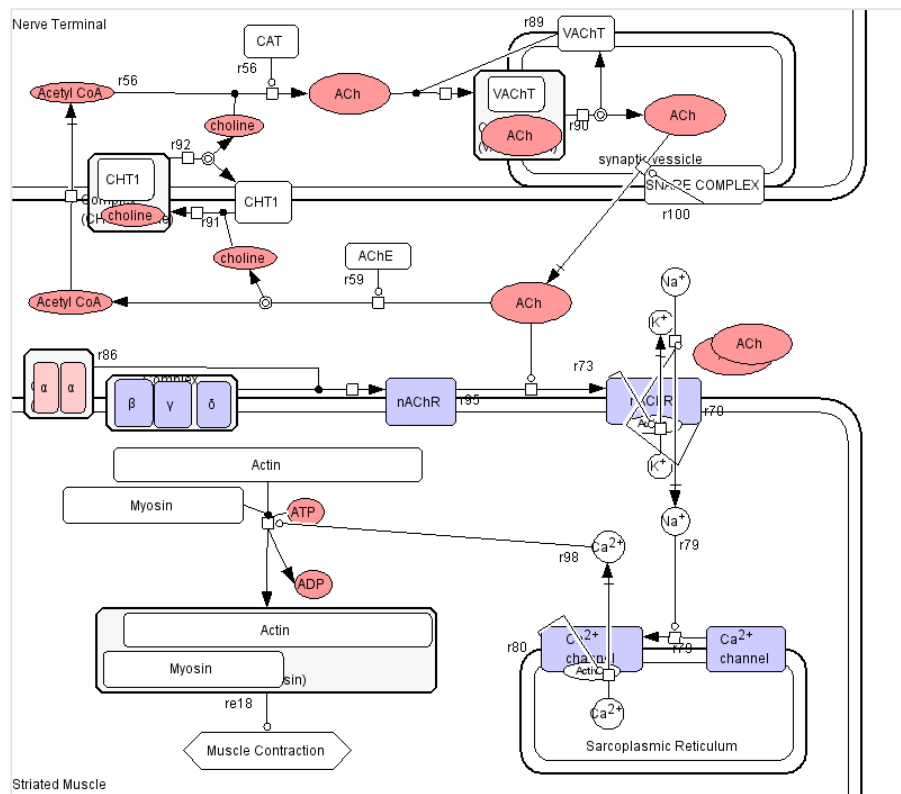

Figure 39: Visual representation of the 'Nicotinic acetylcholine receptor signalling pathway'. It shows four compartments in which the pathway plays a role: 1. Nerve terminal, 2. Striated muscle, 3. Synaptic vesicle and 4. Sarcoplasmic reticulum. ACh stands for Acetylcholine, Acetyl-CoA stands for acetyl coenzyme A, nAChR stands for Nicotinic acetylcholine receptors, CAT stands for Choline acetyltransferase, VAChT stands for Vesicular acetylcholine transporters, AChE stands Acetylcholinesterase, CHT1 stands for The high-affinity choline transporter, and α, β, γ, δ represent different types of Acetylcholine receptor proteins. The reactive components are shaded orange or pink, and the channels are shaded in blue (Mi *et al.*, 2021)*.*

This pathway describes the chemical synapse between motor neurons and skeletal muscle fibres at the neuromuscular junction, and it has been extensively studied in order to obtain insight into to aetiology of muscular disorders, by understanding synaptic plasticity and neuronal transmission at synapses.

In brief, the names of the components in this pathway are as follows: 1. Nicotinic acetylcholine receptors, 2. Choline O-acetyltransferase, 3. Vesicular acetylcholine transporter, 4. High-affinity choline transporter, 5. Actin, 6. Myosin, 7. Calcium ion channel, 8. N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE Complex), 9. Acetylcholinesterase, 10. Nicotinic acetylcholine receptor alpha, and 11-13. Acetylcholine receptor protein (beta, gamma and delta).

Figure 39 depicts the interplay of the 13 components at the neuromuscular junction as follows. Acetylcholine (ACh) is a neurotransmitter synthesized from choline by choline O-acetyltransferase (CAT) in cholinergic neurons (Picciotto *et al.*, 2012). The neurotransmitter ACh is laden onto a synaptic vesicle with the aid of a vesicular acetylcholine transporter (VAChT) encoded by the *SLC18A3* gene. ACh then activates the ionotropic nicotinic acetylcholine receptors (nAChR). The hydrolysis of ACh terminates the fast signal from the receptors into acetyl and choline by acetylcholinesterase (AChE). The released choline is taken up by a high-affinity choline transporter (CHT1) encoded by the *SLC5A7* gene. The activated nAChRs are depolarised at the neuromuscular junction, leading to a cation influx, which generates an action potential to trigger muscle contraction (Vautrin & Barker, 2003).

Nicotinic acetylcholine receptors in neuropathic pain
The nAChRs are a part of ligand-gated ion channels spread across the nervous system. These receptors arbitrate fast signal transmission at synapses (Orr-Urtreger et al., 1995). Functional nAChRs constitute of pentameric protein structure subunits. These subunits are categorized as alpha (α1- 7, α9-10) or non-alpha subunits (β1-4, δ, γ or ε) types. These subunits are encoded by the following genes: *CHRNA1-7*, *9-10*, *CHRNB1-4* and *CHRND*, *CHRNG* and *CHRNE*, respectively (Stokes *et al*, 2015). The different subunits of nAChRs form a subtype of receptors involved at the neuromuscular junction. For example, the most common nAChRs in the CNS are receptors that contain α4 and β2 subunits, which create α4β2* (the asterisk means that there are other subunits probably present in the native nAChRs). In the PNS, nAChRs consists of α3 and β2 subunits that form the α3β2* and α3β4* subtypes; they are highly expressed, in particular in ganglionic nerve cells (Hone & McIntosh, 2018).

Dorsal root ganglia (DRG) are nociceptive neurons whose axons provide somatosensory signals to the central nervous system (CNS). The lowering of amplitude of somatosensory signals via DRG during pain is targeted in pharmacological therapy to reduce pain. DRG neurons are classified according to their sensory modes, size, and the expression of particular receptors and ion channels, including nicotinic acetylcholine receptors. DRG neurons express many nAChR subtypes, according to functional investigations, however the makeup of these receptors is only partially recognized. The existence of receptors with α7-, α3β4*-, and α4β2* subtypes is supported by pharmacological data (Genzen *et al*, 2001). Because of their expression in DRG neurons, nAChRs are being studied as therapeutic targets for NP. Most of the early research work was centred on creating a potent agonist for α4β2 subtype but most of these agonists have not been accepted as they cause several side-effects (Jain, 2004). As a result, other subtypes, including α6β4*, α7-, α9/10-consisting subtypes, have been recognised as possible novel therapeutic targets.

The α6β4* subtypes are expressed in DRG neurons, specifically in small and medium fibres. Also, they co-expressed with other nAChR subtypes such as the one containing the α7 and β2 subunits (Hone *et al.*, 2012). Wieskopf *et al.* demonstrated that the loss-of-function variants in *CHRNA6* increase allodynia, whereas gain-of-function mutations decrease pain. They also reported that the expression of the α6β4 subtype desensitises the P2X2/3 receptors, which inhibits them and diminishes pain.

Another nAChR is α7-subtype, observed in supraspinal and spinal pain-transmission pathways; α7 nAChR ligands are effective against visceral pain. This indicates that modulation of these receptors might ameliorate pain and provide more effective treatment options (Bagdas *et al.*, 2018).

The *CHRNA10* gene is a part of this pathway and was shared among familial and sporadic cohorts as mutated genes. The α10 nAChRs are also expressed in DRG neurons, skin keratinocytes, leukocytes, and the pituitary. A study reported that α9/10 nAChR antagonists showed instant pain relief and that daily administration of analgesia had a sustained effect without any tolerance (Vincler & McIntosh, 2007). Provided this information, studying the mutations in the *CHRNA10* gene found in our cohort could elucidate the mechanistic role of this gene in NP.

On the other hand, the nAChRs assist in nociceptive neuron excitability. Some of these nAChRs are expressed in unmyelinated C-fibers, and these receptors contribute to C-fiber neuronal excitability (Lang *et al.*, 2003). A recent study analysed the mechanistic role of nAChRs in nociceptive neurons.

They found that α3β4, α3β2, α6, and possibly α7 nAChRs participated in the modulation of nociceptive neurons (Shreckengost *et al*, 2021). This study strengthens the role of these receptors in nociceptive sensitisation and pain transmission.

The nAChR genes observed in our cohort are: *CHRNA2*, *CHRNA3*, *CHRNA7*, *CHRNA10,* and *CHRNE*. The *CHRNA7* and *CHRNA10* genes, which encode for α7-and α10- consisting subtypes, have been explored as therapeutic targets for pain, but other genes encoding for α2-, α3- and ε-containing subtypes might potentially be considered as therapeutic targets as well. Moreover, the activity of these receptors can be studied with the incorporation of the studied variants through a functional study. This functional study could help understand the effect of these variants on pain transmission. Consequently, the notion that these receptors play a role in pain-related pathways may be supported.

For more than 30 years, researchers have known that nAChR ligands would be a better alternative to opioid drugs. But, to date, they have not been able to develop an approved nAChR-related analgesic (Hone & McIntosh, 2018). We expect that the mutations discovered in nAChRs in this study may provide insight on and contribute to the development of targeted therapies for alleviating pain.

Structural variants

In parallel to detecting mutations in the coding region of the genome, the WES approach also offers the possibility to investigate variants that could have a structural impact at the genomic level. Due to this fact, we reviewed different SV calling tools to incorporate them into our workflow.

There are several SV tools available to help in the identification of SVs from WES analysis results. We explored a variety of tools, either alone or in combination. We discovered that each SV calling tool used a unique set of SV identifying methods. We observed that each SV calling tool was only capable of detecting one type of SV. The conifer tool, for example, could only detect long duplications, but the manta tool detected deletions and break-ends better than any other type. Because a study of the literature on SV analysis confirmed this concept, our findings show that the output of multiple SV calling techniques should be merged rather than evaluated separately.

According to the Pfundt *et al*, providing healthy controls to these SV callers might aid in the identification of SVs in cases . In further detail, the SV callers use the reference dataset to train their algorithms for exact SV recognition, owing to the differences found in the healthy control group against the case group. Furthermore, it is recognized that the identification of SVs is more robust when more samples are supplied in the control and case groups. In our study, 80 cases were compared to 40 unrelated healthy controls. Unfortunately, our dataset was not well-balanced enough to identify likely SV associated with the phenotype. In the future, a larger cohort of participants will be required to deepen the field of SV in PN. Another option for investigating the SV in PN is to use accessible WES samples from the extended database, such as GnomAD (Karczewski *et al.*, 2020). This GnomAD database contains a large number of reference datasets that may be compared to sequenced patients. Furthermore, because whole genome sequencing (WGS) offers a richer view of SVs at the genomic level, the more informative samples could be re-sequenced with the WGS approach, in order to compare these data with the whole genome samples already available from the GnomAD database as a reference cohort.

Due to the lack of studies and databases needed for SV prioritisation, we could not confidently stratify a SV region. Also, the SV regions found in chromosome X were observed in male samples and not in females, probably since reference samples supplied to these callers consisted mostly of females (>50%) than males. Due to this gender bias, these callers detected SV regions as potential SVs, but further steps need to be performed to confirm these results. Therefore, a high number of reference samples for precise variant calling is necessary and should be balanced according to gender type. Alternatively, some studies also suggest

processing samples with different gender separately, to reduce gender bias in a cohort (Collins *et al.*, 2019). Furthermore, technical concerns such as insufficient variation coverage depth for a confident SV region selection (the variant region's sequencing coverage was less than 25%), suboptimal SV databases to enable prioritisation of SVs, and low power of control set were ascribed to non-detection of SVs in our cohort.

Conclusions and future prospects

This study attempts to decode the genetic heterogeneity and complexity of painful PN by using the WES approach. WES experiments appear to be effective in detecting mutations and deciphering novel genotype-phenotype relationships. The data generated from the WES were comparable to other exome sequencing studies in terms of sequencing coverage and variant quality. Thanks to the approach used in this study, we prioritised mutations in 113 genes in the analysed cohort, and we propose that this list of genes could be used to prioritize and interpret data provided by NGS technologies such as WES or targeted-sequencing in future research investigations where family history or phenotypic information is indicative of painful PN.

Additionally, our data suggests genetic heterogeneity of the painful PN in which the VGSCs genes and other genes are involved, mediating the phenotype and altering pain signalling that could contribute to disease onset and pain modulation. For example, we found variants in TRP genes, specifically in *TRPM2,* functionally linked with NP. The WES results also intimate the involvement of abnormal copper metabolism that could contribute to the symptoms. In particular, if the *ATP7B* gene mutations will be further validated and replicated in other cohorts, it could be interesting to assess the levels of copper in the plasma.

The pathway analysis revealed the nicotinic acetylcholine receptor signalling pathway, which was found to be enriched in individuals from familial and sporadic cohorts, pointing to a critical mechanism of synaptic transmission shared by the two studied cohorts. The present project suggests that most of the genes mutated in patients participate in neurotransmission. Future studies need to be performed to further replicate the result in another cohort. Proper functional studies need to be considered to elucidate the role of nicotinic acetylcholine receptor signalling in the context of the studied disease and pain transmission.

All the qualifying variants detected in this study and classified as 'VUS', 'Likely pathogenic' and 'Pathogenic' will be further validated by Sanger sequencing. Mutations found in

CACNA1E, Potassium Channels, VGSCs and *TRP* genes will be further evaluated for the electrophysiological study.

The results from the SV variant analysis were not up to the mark, and we need further evaluations. This could be achieved by increasing the power of reference samples provided to the tools (Pfundt *et al.*, 2016) or performing this analysis in a larger cohort of at least 500 samples that could be recruited thanks to the effort performed by different national and international members involved in a similar project.

A strong heterogeneity across families and also within sporadic samples has been observed in the present project. In order to try to address this critical point, we plan to perform in the next future a gene-based burden testing of rare variants, to assess the global and weighted contribution of the studied genes. Specifically, the burden test will be carried out in 592 candidate genes against the control dataset (n=123,136) publically available from the GnomAD database. This test might resolve the genetic heterogeneity observed in this study and validate some candidate genes.

With the advent of technology, the clinical applicability of WES has become more apparent over the years. The databases and the tools applied in WES data analysis will undoubtedly evolve to provide more information and validate the new genetic variants and genes found in this study. One of the advantages of the WES application is that the samples can be re-evaluated with the evolution of new databases or tools. Also, the candidate genes can be expanded or reduced with the development of new biological insights.

In the near future, diagnosis through genetic screening will become a standard and provide new genetic markers that could translate into therapeutic targets for PN.

The results achieved from this study could be a precursor to other future genetic screening projects. As the quantity of data dictates quality, if the same study is replicated in a larger cohort, it could improve the clinical diagnostic process and precise stratification of high-risk individuals, which are for example the ones belonging to families with PN. In clinical practice, for example, a gene panel of 113 genes may be assessed for genetic screening of painful PN. The pathogenicity of novel mutations in these genes may be assessed in specific contexts based on the clinical phenotypic description provided and the relevant biological inference. This approach could help in the identification of subjects at higher risk, allowing a proper longitudinal assessment of these subjects in specialized centre. Moreover, the discoveries from the present project could help in the identification of variants that could be considered also as potential novel targets, for the development of new drugs that could be beneficial for a large population of patients.

Materials and Methods

Recruitment of subjects

In this genetic study, Italian samples were recruited in the Neuroalgology Unit at IRCCS Foundation 'Carlo Besta' Neurological Institute (CBNI). San Raffaele Hospital (OSR), and in particular the Laboratory of Human Genetics of Neurological Disorders, was involved in performing WES experiments of sporadic and familial subjects affected by painful PN and healthy subjects belonging to families with affected relatives, with the final aim of exploring the genetic component underlying the disease.

Whole blood or saliva samples were collected from these subjects. The whole blood was the preferred biological material used for DNA extraction. At the same time, saliva was collected only from subjects, mainly belonging to the familial cohort, who could not go to the hospital for geographical distance or health problems. The Ethics Committee approved the project of CBNI and OSR, and each individual signed the informed consent.

All subjects were recruited according to the same inclusion/exclusion criteria, and WES experimental procedure was carried out for all the samples. The total number of people who were recruited was 96, and they included both familial and sporadic patients, as summarised in Figure 40.
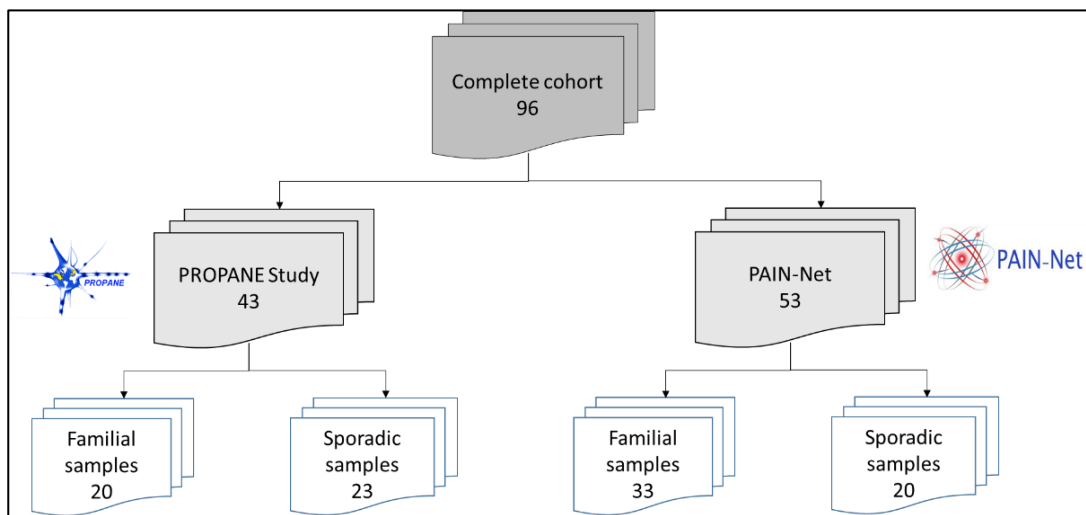

Figure 40: A schematic representation of the selected cohort of subjects for WES experiments from two different projects (PROPANE Study and PAIN-Net).

Familial Cohort

In this study, 12 families were included: 6 from the PROPANE project and 6 from the PAIN-Net project, 53 samples. Among them, 37 were the affected and 16 were the unaffected members. The inclusion criteria for familial samples are as follows:

- presence of at least three members belonging to the family with at least one (the proband) or more subjects affected with PN and NP
- the pain intensity was measured by Numerical Rating Score (NRS) scale, NRS ≥ 4 for at least one subject in the family;
- a positive skin biopsy indicative of sensory nerve involvement for the proband;
- evidence of positive neurological history for the proband.

The 16 unaffected members had normal results of the neurological tests and had a negative neurological medical history.

Sporadic Cohort

43 samples were recruited for the sporadic cohort; 23 belong to the PROPANE project, whereas 20 were enrolled within the PAIN-Net project. These are the inclusion criteria:

- the pain intensity score: NRS ≥ 4;
- positive skin biopsy indicative of sensory nerve involvement;
- evidence of positive neurological history;
- age at onset of symptoms ≤ 45 years.

Whole exome sequencing

DNA extraction and DNA quality control

Whole blood collected by CBNI was stored in anticoagulant tubes, and a proper identification number was allotted to each tube. According to the manufacturer's instruction manual, genomic DNA was extracted from whole blood at CBNI using QIAamp DNA Blood Maxi Kit by QIAGEN. The manufacturer's instruction manual was referred to collect the saliva samples in the Oragene DNA OG-500 kit by DNA Genoteck, and genomic DNA was extracted using the Oragene DNA OG-500 kit. Genomic DNA from both sources was additionally purified using Agencourt AMPure XP magnetic beads by Beckman Coulter to eliminate every kind of contaminants in the extraction buffer and increase the DNA quality for further steps. The extracted DNA was eluted in deionised water, and a quality and quantity control check was performed. An agarose gel at 1% checked the DNA quality was run to evaluate the presence of degraded DNA.

DNA quantity check was done by checking DNA concentration by two instruments: NanoDrop 8000 UV-Vis Spectrophotometer by Thermo Scientific and Qubit® 2.0 Fluorometer by Thermo Scientific, using Qubit dsDNA and HS assay kit by Life Technologies, ensuing manufacturer's instruction manual. All the isolated DNA were then stored at -20°C.

WES experiments

The DNA extracted from the blood of subjects from the PAIN-Net cohort, after the quality control step, was shipped to the Centre for Advanced Studies, Research and Development in Sardinia (CRS4). CRS4 processed the samples for DNA fragmentation, adapter and index tagging and hybridisation using the Sure Select QXT Human All Exon V5 kit (Agilent genomics). Sure Select QXT Human All Exon V5 kit has RNA probes designed to enrich the exome component of the genome. Because the two families were recruited later and the probes of the Sure-Select QXT Human All Exon V5 kit were discontinued by the manufacturer, they were hybridised using the Sure Select Human All Exon V7 target enrichment kit (Agilent Technologies). After the hybridisation, captured exomes were amplified to obtain the libraries to proceed with the sequencing. Sample libraries were sequenced on the HiSeq 3000 platform by Illumina running a paired-end (2x150 bp) in rapid sequencing modality. The steps mentioned above were only carried out for samples recruited under the PAIN-Net project. In contrast, the samples recruited under the PROPANE project were hybridised using the same Sure-Select QXT Human All Exon V5 kit but sequenced on the HiSeq 2500 platform by Illumina with paired-end sequencing (2x101 bp) modality (Illumina Technologies).

The library preparation and sample sequencing for the PROPANE project were carried out in part by researchers of the Laboratory of Human Genetics of Neurological Disorders and in part from the genomic facility both at OSR.

WES data analysis pipeline

Raw reads of the samples from the PROPANE project were already available in the Laboratory of Human Genetics of Neurological Disorders (OSR). At the same time, CRS4 provided the raw reads of samples enrolled in the PAIN-Net project. The raw reads from both studies were then analysed, applying the same pipeline for consistency.

The raw reads, also known as fastq files, were used as input in the WES analysis workflow. In this workflow, the initial step was to check the quality of raw reads.

Specifically, the raw reads were checked by the FastQC (version 0.11.8) tool, which evaluated different aspects of sequences of each sample as the per-base sequence quality, sequence duplication levels, over-represented sequences, per sequence GC content and per sequence quality scores. (Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data).

135

The index, a nucleotide sequence that serves to uniquely designate a sample in a complex of samples sequenced at the same time, had already been extracted by the Illumina sequencing machine software, and the fastq data file was matched to the subject.

The next important step was to remove the adapters from each read. The adapter removal step was done with two different tools: Trimmomatic (version: 0.39), to remove the low-quality bases and adapters, and Cutadapt (version: 2.7), to remove the partially remaining adapter while maintaining the read quality (Bolger *et al.*, 2014; Martin, 2011). Adapter-trimmed fastq files were aligned to the human reference genome (version GRch37/hg19) using Burrows-wheeler aligner-MEM (version 0.7.17) tool (Li, 2013).

These aligned reads were converted to Binary Alignment Map (BAM) files using Samtools (version 1.9). These BAM files were also processed with Samtools and Picard (version 2.22.5) to obtain alignment and coverage reports of the samples. The BAM files were then processed via Genome Analysis Toolkit (GATK) workflow (version 4.1.7) for the identification of variants (Li *et al.*, 2009; Poplin *et al.*, 2018). The identified variants were then annotated using the SnpEff suite (version 4.3t) into an annotated Variant Call Format (VCF). The following step was the "variant interpretation", as reported below.

Regarding the structural variants (SV), a specific pipeline was adopted. The SV calling tools are sensitive towards three features of a cohort: i) the target-enrichment kit used for the WES experiments; ii) the average coverage reached per sample; and iii) the number of reference samples provided for the analysis. To minimise these biases in our cohort, samples produced only by the Sure Select V5 kit were utilised, and samples with a coverage depth greater than 50X were selected. BAM files from the GATK tool were processed with four different tools for performing structural variant analysis. The tools used for structural variant analysis are as follows: GATK-Germline CNV caller (version: 4.1.7), Cnvkit (version: 0.9.8), conifer (version: 0.2.1), and manta (version: 1.6) (Poplin *et al.*, 2018; Talevich *et al.*, 2016; Krumm *et al.*, 2012; Chen *et al.*, 2016).

The overlapping segments of the four SV calling tools were calculated with a minimum cut-off of 70% overlap and merged before variant identification. Unique regions from each caller were also retained in the final file. These files were then annotated using 'AnnotSV' (version: 3.1) (Geoffroy et al., 2021). The genotype information is not provided by the selected SV calling tools. As a result, no genetic segregation analysis was done on the familial cohort. Due to the availability of in-house 40 unrelated healthy controls processed with the Sure Select V5 kit, a case-control strategy was used for SV detection in both cohorts. These samples were used as reference samples for the SV calling tools. The control samples were used in two ways: i) as a training set for SV callers to learn how to make

precise SV detection specific for a particular target enrichment kit, and ii) as a control set for variant filtering.

The SV regions observed in the affected members per family for the familial cohort were compared against unaffected members and the control samples. For the sporadic cohort, SVs regions in the sporadic cases were compared with the 40 control samples. The SV regions unique to the individual sporadic case was then processed for downstream filtering. The variants were then filtered for quality, the allele frequency of the SV region and known benign variants seen in the global population. The following step was the "variant interpretation", as reported below.

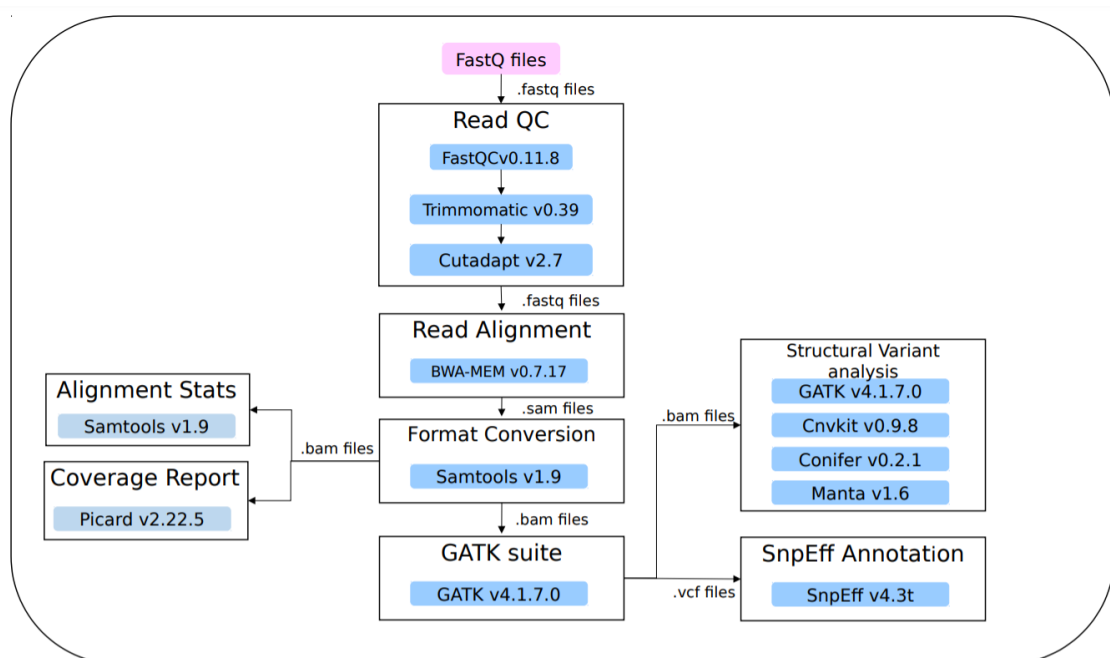The steps performed in the WES data workflow have been summarised in Figure 41.

Figure 41: A schema for WES data workflow from raw reads to annotated file

Variant interpretation

Variant filtering

For ensuring proper selection of SNVs and short INDELs, the following steps were applied on each variant and performed for familial and sporadic cohorts. The Variant Call Format file was converted to a tabulated text file to facilitate the filtering for potentially putative variants causing a deleterious effect on the resultant protein.

The filtering criteria was based on:

- high or moderate impact on protein translation and at RNA splicing level;
- sequencing depth quality ≥ 10 and alternate allele depth percentage ≥ 20 %;
- the quality score given by GATK tool ≥ 100;
- alternate allele frequency of ≤ 5 % or absent in GnomAD population database for non-Finnish European population;
- alternate allele frequency of ≤ 2 % or absent in the in-house internal Italian population;
- Phylopway gives evolutionary conservation score of 100 vertebrates ≥ 1 or undefined score;
- variants segregating with the disease under a transmission model according to the pedigree in familial cases or less than 25% frequent among sporadic patients.

These filtering criteria could not be applied for SVs as the data obtained from the SV callers or annotation tool does not have the same information as observed for SNVs and short INDELs. The SV regions were filtered as follows:

- the quality score given by GATK-Germline CNV caller ≥100 or absent;
- allele frequency ≤ 5 % or absent in GnomAD SV database;
- SV region not marked as known benign variant by the annotation tool.

These qualifying SV regions were then selected if the region consisted of one or more genes that were also found in the candidate genes.

Segregation model in families

After variant filtering, a genetic variant segregation model was applied in familial samples, according to the features of each family and following the 'Law of segregation' by Mendel (Rosenberg & Rosenberg, 2012). In addition, a causative or protective segregation model was then applied to each family. A causative model suggests that the ensuing qualifying variant is likely to have a pathogenic effect on the phenotype; on the other hand, a protective model says that the consequent qualifying variant has a resilient effect on the phenotype. For example, in a causative transmission model, if a variant is present in the affected member but not in the unaffected member, this suggests a causative influence. In the protective transmission model, on the contrary, if the variant is lacking in the affected member but present in the unaffected member, it suggests that a protective effect on the phenotype.

Sharing model in sporadic patients

After the variant filtering step of each sporadic sample, a summary table of qualifying variants in the heterozygous or homozygous state was tabulated.

At that level, the sharing model for variants was applied to identify variants shared among more than one sporadic subject but in less than 25% of the sporadic cohort (so present in less than 11 sporadic subjects) to avoid the selection of common variants in the studied population. The sharing model was also applied at the gene level to identify genes commonly mutated in the sporadic group if variants were not shared among individuals.

Comparison among families and sporadic

The list of qualifying variants and genes was tabulated separately for the two cohorts. The list of qualifying genes accumulated from the two cohorts was clustered according to 'BRITE functional hierarchies' through the KEGG database (Kanehisa *et al.*, 2012). The tool classified these genes into four categories: 'Ion channel', 'Neurotransmission', 'Metabolism' and 'Immune response'. The qualifying genes were categorized in order to summarize them according to their functional biological mechanism.

The qualifying genes were then gathered and compared to see if any gene or variant was shared across families or within families, as well as among at least one or more cases from the sporadic cohort. The schema of variant interpretation workflow is described in Figure 42.
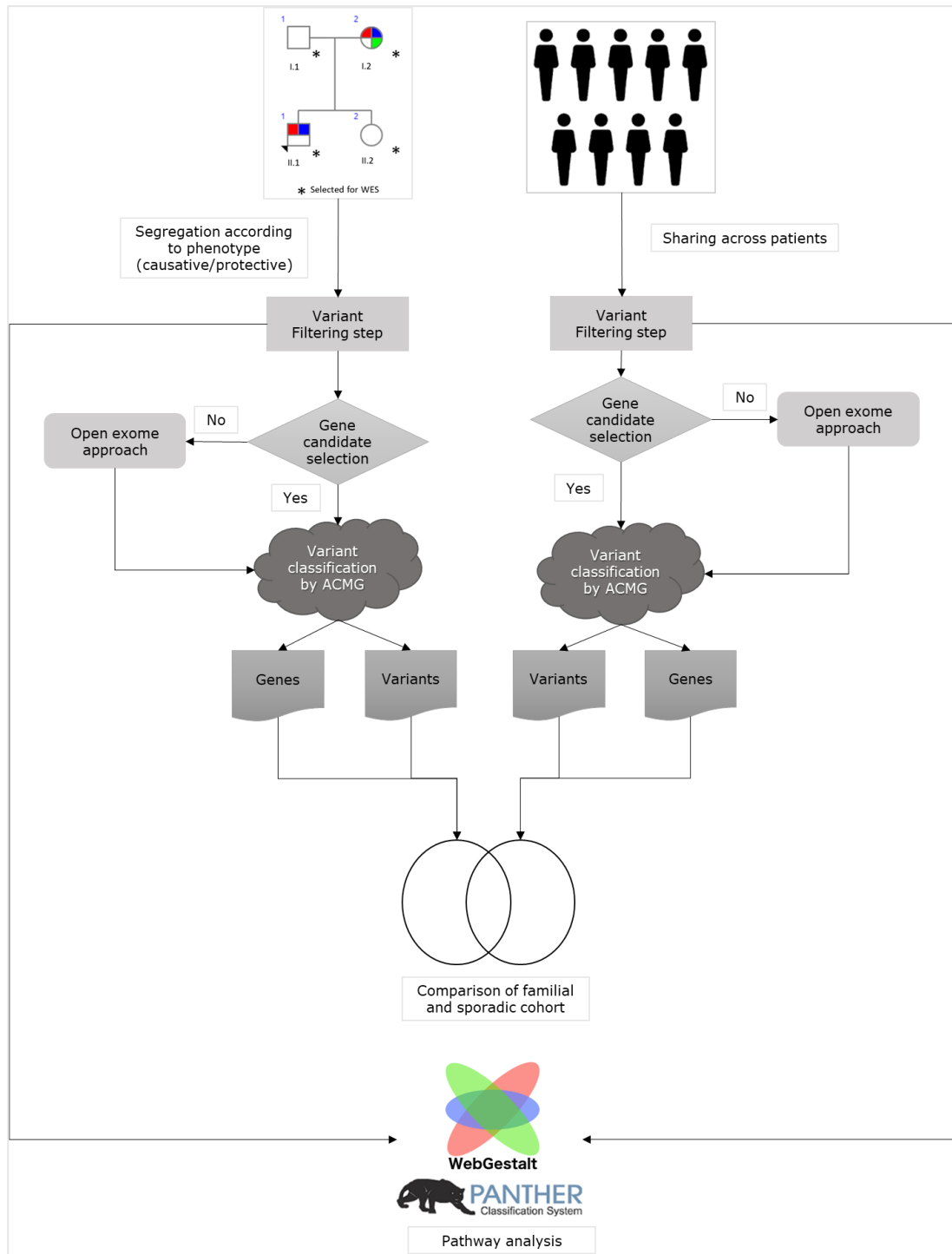
Figure 42: A schematic representation of the variant interpretation workflow applied to the familial and sporadic cohorts

Variant prioritisation

All filtered variants were then prioritised in two different steps: i) application of 'candidate gene' filter, and ii) classifying each variant according to the classification criteria of the American College of Medical Genetics (ACMG) guidelines 2015 (Richards et al., 2015).

For prioritising the qualifying variants located in 'candidate genes', which are the genes related to the phenotype. A phenotype-driven virtual gene panel was created.

This gene panel was curated using multiple databases as follows:

- International Association for Study of Pain (Pain Gene Resource | Pain Research Forum);
- Human Pain Genetics database where the phenotype description was related to 'Pain' and not to 'Migraine', 'Cancer pain' and 'Temporomandibular disorders' (Meloto et al., 2018);
- Genes associated with keywords as mentioned in Table 30 were extracted from Human Phenotype Ontology (HPO) database and Online Mendelian Inheritance in Man (OMIM). (Köhler et al., 2021; Amberger & Hamosh, 2017).

Phenotypes and keywords used for the candidate gene selection	
Pain (Burning, shooting, prickling or itching)	Hot flushes
Paraesthesia	Orthostatic dizziness
Allodynia	Cardiac palpitations
Thermal sensory loss	Pruritus, Itching
Pinprick loss	Neuropathic itch
Restless leg syndrome	Pain perception
Sicca syndrome	Pain hypersensitivity
Hyperhidrosis or hyperhidrosis	Neuropathic pain
Impotence	Hyperalgesia
Bowel disturbances	Peripheral neuropathy

Table 30: Phenotypes and keywords selected as input to create the virtual gene panel

If no variants remained after the applied filters, the last prioritisation according to the location of the variant in the candidate genes panel was skipped, and the analysis was performed on the entire WES data (applying the so-called 'open exome' strategy).

The remaining variant were then processed through various variant descriptors by following the ACMG guidelines, as mentioned in Table 2 of the Introduction and Appendix A.2. Only the variants classified as 'VUS', 'Likely Pathogenic' and 'Pathogenic' were retained.

In the case of structural variants, the annotation tool applied to the SV regions provided ACMG ranking for each SV region based on the type of SV is observed. Similar to SNVs and short INDELs, we categorised the SV regions into the classifications mentioned in Table 2 of the Introduction.

Pathway analysis on resultant genes

The pathway analysis was carried out to better understand the functional characteristics of qualifying genes and the underlying biological pathways. The analysis was performed using the WebGestalt tool (Liao *et al.*, 2019), in which an over-representation method was applied (Khatri *et al.*, 2012). Using the Benjamin-Hochberg method, this method uses a hypergeometric test to calculate p-values and false discovery rates (FDR). The input data were the filtered genes separated for the familial and sporadic cohorts. The Protein ANalysis THrough Evolutionary Relationships (PANTHER) database was used to assign to the submitted list of genes their enrichment in a specific pathway (Kanehisa *et al.*, 2012; Mi *et al.*, 2021; Jassal *et al.*, 2020). This database was selected as the data in their database is manually curated from published experimental studies and other associated literature resources (Chowdhury & Sarkar, 2015).

The reference genome selected was 'genome protein-coding' and 'human' as the organism of interest. The other parameters were used as default conditions as given by the tool.

In the tool's section called 'Gene List', input genes filtered according to the criteria above described were provided as gene symbols.

After the 'Variant filtering' stage, one gene list was acquired and updated by certain criteria for each cohort before being used as an input for pathway analysis.

The following steps were performed to attain a list of genes from the two cohorts:

- For the familial cohort, genes presented with qualifying variants from causative segregation models were accrued into one list. Duplicated genes were removed.
- For the sporadic cohort, genes presented with qualifying variants from the sharing model approach were accrued into one list. Duplicated genes were removed.
- To exclude genes anticipated to tolerate a functional alteration, the LoFtool tool score filter was applied to each gene list. This tool ranks each gene based on its tolerance to loss of function and tissue expression specificity (Fadista et al., 2017). To replace the 'Variant prioritisation' phase, the LoFtool score filter was included.

The gene list obtained from the two cohorts was then supplied to the WebGestalt tool's 'Gene list' section. This WebGestalt tool has a functionally annotated gene list for the PANTHER database. The tool initially maps the input dataset unambiguously to protein-coding genes found in the human genome. The input genes that survive the initial mapping step intersect with each database's annotated gene list. The output gene list is retained for further association to biological pathways.

One pathway analysis per cohort was performed, and the related output was examined. The obtained output report is a table with different pathways matching the set of input genes along with the following parameters:

- pathway name;
- gene set: number of reference genes in the pathway;
- mapped input: number of genes in the gene set and also in the pathway;
- the ratio of the enrichment;
- p-value: raw p-value by the hypergeometric test;
- FDR: p-value adjusted by the multiple test adjustment.

A pathway was selected using the following criteria:

- p-value < 0.05
- false discovery rate (FDR) < 0.1
- number of genes belonging to 'Gene List' overlapping with the pathway ≥ 2

Appendix A.1

Gene name	Type of IPN	Gene name	Type of IPN	Gene name	Type of IPN	Gene name	Type of IPN
PMP22	HMSN	*SPTLC1*	HSN/HSAN	*HSPB8*	HMN	*GJB1*	HMSN/HMN
MPZ	HMSN	*SPTLC2*	HSN/HSAN	*HSPB1*	HMN	*PRPS1*	HMSN/HMN
LITAF	HMSN	*ATL1*	HSN/HSAN	*HSPB3*	HMN	*PDK3*	HMSN/HMN
EGR2	HMSN	*DNMT1*	HSN/HSAN	*FBXO38*	HMN	*ATP7A*	HMSN/HMN
NEFL	HMSN	*ATL3*	HSN/HSAN	*HARS*	HMN	*DRP2*	HMSN/HMN
PMP2	HMSN	*WNK1*	HSN/HSAN	*AARS*	HMN		
ARHGEF10	HMSN	*FAM134B*	HSN/HSAN	*FBLN5*	HMN		
GDAP1	HMSN	*KIF1A*	HSN/HSAN	*SLC12A6*	HMN		
MTMR2	HMSN	*SCN9A*	HSN/HSAN	*GARS*	HMN		
SBF1	HMSN	*IKBKAP*	HSN/HSAN	*BSCL2*	HMN		
SBF2	HMSN	*NTRK1*	HSN/HSAN	*REEP1*	HMN		
SH3TC2	HMSN	*NGF*	HSN/HSAN	*SETX*	HMN		
NDRG1	HMSN	*DST*	HSN/HSAN	*IGHMBP2*	HMN		
PRX	HMSN	*SCN11A*	HSN/HSAN	*SIGMAR1*	HMN		
HK1	HMSN	*PRDM12*	HSN/HSAN	*PLEKHG5*	HMN		
FGD4	HMSN	*ZFHX2*	HSN/HSAN	*DNAJB2*	HMN		
FIG4	HMSN	*TRPA1*	HSN/HSAN	*HSPB1*	HMN		
SURF1	HMSN	*SCN10A*	HSN/HSAN	*SLC5A7*	HMN		
DNM2	HMSN	*COL6A5*	HSN/HSAN	*DCTN1*	HMN		
YARS	HMSN	*PIEZO2*	HSN/HSAN	*WARS*	HMN		
INF2	HMSN	*CRLF1*	HSN/HSAN	*TRPV4*	HMN		
GNB4	HMSN	*CLCF1*	HSN/HSAN	*DYNC1H1*	HMN		
MFN2	HMSN			*BICD2*	HMN		
RAB7A	HMSN						
TRPV4	HMSN						
GARS	HMSN						
HSPB1	HMSN						
HSPB8	HMSN						
AARS	HMSN						
DYNC1H1	HMSN						
LRSAM1	HMSN						
DHTKD1	HMSN						
MME	HMSN						
MARS	HMSN						
NAGLU	HMSN						
HARS	HMSN						
VCP	HMSN						
MORC2	HMSN						
NEFH	HMSN						
ATP1A1	HMSN						
TFG	HMSN						
DGAT2	HMSN						

Gene name	Type of IPN
DCAF8	HMSN
HSPB3	HMSN
SEPTIN9	HMSN
CHCHD10	HMSN
FBLN5	HMSN
LMNA	HMSN
PNKP	HMSN
TRIM2	HMSN
IGHMBP2	HMSN
SPG11	HMSN
HINT1	HMSN
DNAJB2	HMSN
REEP1	HMSN
MCM3AP	HMSN
SGPL1	HMSN
KARS	HMSN
PLEKHG5	HMSN
COX6A1	HMSN

Table A.1: List of genes that have been reported to be involved in a type of IPN. This table has been adapted from Scherer et al, 2020
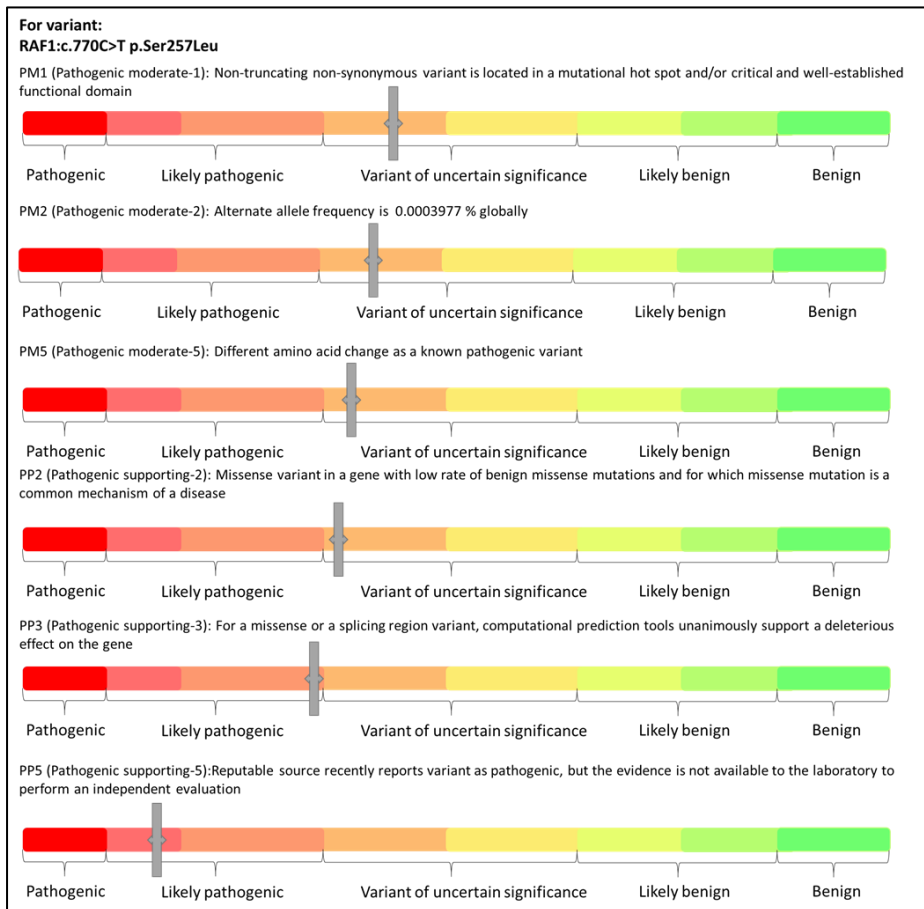
Appendix A.2

Figure 1a: A schematic representation of the attributes given to a particular variant and its ranking along with its classification result

Appendix A.3

Gene symbol	Entrez gene ID	Gene name
ABCB1	5243	ATP-binding cassette, subfamily B, member 1 (P-glycoprotein-1/multiple drug resistance-1)
ABCB11	8647	ATP-binding cassette, subfamily B, member 11 (bile salt export pump)
ABCB4	5244	ATP-binding cassette, subfamily B, member 4 (P-glycoprotein-3/multiple drug resistance-3)
ABCC2	1244	ATP-binding cassette, subfamily C, member 2 (canalicular multispecific organic anion transporter)
ABCC4	10257	ATP-binding cassette, subfamily C, member 4
ABHD12	26090	Abhydrolase domain-containing protein 12, lysophospholipase
ACAN	176	Aggrecan (chondroitin sulfate proteoglycan-1, large aggregating proteoglycan, antigen identifies by monoclonal antibody A0122)
ACAP2	23527	Centaurin, beta-2
ACCN1	40	Amiloride-sensitive cation channel 1, neuronal (degenerin)
ACCN2	41	Cation channel, amiloride-sensitive, neuronal, 2
ACE	1636	Angiotensin I converting enzyme (dipeptidyl carboxypeptidase-1)
ACE2	59272	Angiotensin I converting enzyme 2
ACO1	48	Aconitase, soluble (iron-responsive element-binding protein-1)
ACTA1	58	Actin, alpha-1, skeletal muscle
ACTG1	71	Actin, gamma-1
ACTG2	72	Actin, gamma-2, smooth muscle, enteric
ACTN4	81	Actinin, alpha-4
ADAM28	10863	ADAM metallopeptidase domain 28
ADAMTS4	9507	ADAM metallopeptidase with thrombospondin type 1 motif 4 (aggrecanase 1)
ADAMTS5	11096	ADAM metallopeptidase domain with thrombospondin type 1 motif, 5 (aggrecanase 2)
ADAMTSL1	92949	ADAMTS-like protein 1
ADAMTSL4	54507	ADAMTS-like 4
ADARB2	105	Adenosine deaminase, RNA-specific, B2 (homolog of rat BLUE)
ADCYAP1R1	117	ADCYAP receptor, type I
ADGRB3	577	Adhesion G protein-coupled receptor B3
ADH1B	125	Alcohol dehydrogenase IB (class I), beta polypeptide
ADORA2A	135	Adenosine A2a receptor
ADRA1A	148	Adrenergic, alpha-1C-, receptor
ADRA1D	146	Adrenergic, alpha-1D-, receptor
ADRA2A	150	Adrenergic, alpha-2A-, receptor
ADRA2C	152	Adrenergic, alpha-2C-, receptor

Gene symbol	Entrez gene ID	Gene name
ADRB1	153	Adrenergic, beta-1-, receptor
ADRB2	154	Adrenergic, beta-2-, receptor, surface
AGA	175	Aspartylglucosaminidase
AGBL1	123624	ATP/GTP-binding protein-like 1
AGER	177	Advanced glycosylation end product-specific receptor
AIP	9049	Aryl hydrocarbon receptor-interacting protein
AIRE	326	Autoimmune regulator
AJAP1	55966	Adherens junction-associated protein 1
AKAP6	9472	A-kinase anchor protein 6
ALMS1	7840	ALMS1 centrosome and basal body associated protein
ALOXE3	59344	Arachidonate lipoxygenase 3
AMBP	259	Alpha-1-microglobulin/bikunin precursor (inter-alpha-trypsin inhibitor, light chain; protein HC)
ANG	283	Angiogenin
ANK3	288	Ankyrin-3, node of Ranvier
ANKK1	255239	Ankyrin repeat and kinase domain containing 1
ANO1	55107	Anoctamin 1, calcium activated chloride channel
ANO3	63982	Anoctamin 3
ANXA2	302	Annexin A2 (lipocortin I)
AOC1	26	Amiloride-binding protein-1
APC	324	APC regulator of WNT signaling pathway
APOE	348	Apolipoprotein E
APP	351	Amyloid beta (A4) precursor protein
AQP5	362	Aquaporin-5
AR	367	Androgen receptor (dihydrotestosterone receptor)
ARMS2	387715	LOC387715 gene
ARRB2	409	Arrestin, beta 2
ARSA	410	Arylsulfatase A
ARTN	9048	Artemin
ARVCF	421	ARVCF delta catenin family member
ASIC3	9311	Cation channel, amiloride-sensitive, neuronal 3
ASTN1	460	Astrotactin 1
ASTN2	23245	Astrotactin 2
ATL1	51062	Atlastin GTPase 1
ATL3	25923	Atlastin GTPase 3
ATOH1	474	Atonal, Drosophila, homolog of, 1
ATP1A1	476	ATPase, Na+K+ transporting, alpha-1 polypeptide
ATP1A2	477	ATPase, Na+K+ transporting, alpha-2 polypeptide
ATP1A3	478	ATPase, Na+K+ transporting, alpha-3 polypeptide
ATP1B1	481	ATPase, Na+K+ transporting, beta-1 polypeptide
ATP2A2	488	ATPase, Ca++ transporting, slow-twitch, cardiac muscle-2
ATP2C2	9914	ATPase, Ca(2+)-transporting, type 2C, member 2
ATR	545	ATR serine/threonine kinase

Gene symbol	Entrez gene ID	Gene name
ATXN1	6310	Ataxin-1
ATXN3	4287	Ataxin-3 (josephin)
ATXN8OS	6315	Ataxin 8 opposite strand
AVIL	10677	Advillin
AVPR1A	552	Arginine vasopressin receptor-1A
B2M	567	Beta-2-microglobulin
BAHCC1	57597	BAH domain- and coiled-coil domain-containing protein 1
BAP1	8314	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)
BCKDHA	593	Branched chain keto acid dehydrogenase E1, alpha polypeptide
BCORL1	63035	BCL6 corepressor-like 1
BDKRB1	623	Bradykinin receptor B1
BDKRB2	624	Bradykinin receptor B2
BDNF	627	Brain-derived neurotrophic factor
BHLHE22	27319	Basic helix-loop-helix family, member E22
BMP2	650	Bone morphogenetic protein-2
BMP6	654	Bone morphogenetic protein-6
BORCS5	118426	BLOC1-related complex, subunit 5
BRCA1	672	Breast cancer-1 gene
BTBD9	114781	BTB/POZ domain-containing protein 9
CACNA1A	773	Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit
CACNA1B	774	Calcium channel, voltage-dependent, L type, alpha 1B subunit
CACNA1E	777	Calcium channel, voltage-dependent, alpha 1E subunit
CACNA1G	8913	Calcium channel, voltage-dependent, T type, alpha-1G subunit
CACNA1H	8912	Calcium channel, voltage-dependent, T type, alpha-1H subunit
CACNA1S	779	Calcium channel, voltage-dependent, L type, alpha 1S subunit
CACNA2D1	781	Calcium channel, voltage-dependent, L type, alpha 2/delta subunit
CACNA2D2	9254	Calcium channel, voltage-dependent, alpha-2/delta subunit 2
CACNA2D3	55799	Calcium channel, voltage-dependent, alpha-2/delta subunit 3
CACNA2D4	93589	Calcium channel, voltage-dependent, alpha-2/delta subunit 4
CACNB2	783	Calcium channel, voltage-dependent, beta 2 subunit
CACNB3	784	Calcium channel, voltage-dependent, beta 3 subunit
CACNG2	10369	Calcium channel, voltage-dependent, gamma-2 subunit (stargazin)
CALCA	796	Calcitonin/calcitonin-related polypeptide, alpha

Gene symbol	Entrez gene ID	Gene name
CALCRL	10203	Calcitonin receptor-like
CALM1	801	Calmodulin-1 (phosphorylase kinase, delta)
CAMK4	814	Ca(2+)-calmodulin-dependent protein kinase type IV of brain
CARF	79800	Amyotrophic lateral sclerosis 2 chromosome region, candidate 8 (calcium response factor)
CASP9	842	Caspase 9, apoptosis-related cysteine protease
CAT	847	Catalase
CBS	875	Cystathionine beta-synthase
CCL2	6347	Small inducible cytokine A2 (monocyte chemotactic protein, homologous to mouse Sig-je)
CCR2	729230	Chemokine (C-C) receptor 2
CCT5	22948	chaperonin containing TCP1, subunit 5 (epsilon)
CD247	919	CD247 antigen
CDH1	999	Cadherin-1 (E-cadherin; uvomorulin)
CDH12	1010	Cadherin-12 (N-cadherin 2)
CDH18	1016	Cadherin 18
CDH23	64072	Cadherin-23 (otocadherin)
CDK4	1019	Cyclin-dependent kinase 4
CDK5	1020	Cyclin-dependent kinase 5
CDKN2A	1029	Cyclin-dependent kinase inhibitor 2A (p16, inhibits CDK4)
CDSN	1041	Corneodesmosin
CFLAR	8837	CASP8- and FADD-like apoptosis regulator
CHCHD2	51142	Coiled-coil-helix-coiled-coil-helix domain-containing protein 2
CHD7	55636	Chromodomain helicase DNA binding protein 7
CHRM2	1129	Cholinergic receptor, muscarinic, 2
CHRM3	1131	Cholinergic receptor, muscarinic, 3
CHRNA1	1134	Cholinergic receptor, nicotinic, alpha polypeptide-1, muscle
CHRNA10	57053	Cholinergic receptor, neuronal nicotinic, alpha polypeptide 10
CHRNA9	55584	Cholinergic receptor, neuronal nicotinic, alpha polypeptide 9
CLDN10	9071	Claudin 10
CLIC5	53405	Chloride intracellular channel 5
CLOCK	9575	Circadian locomotor output cycles kaput
CNR1	1268	Cannabinoid receptor-1, brain
CNR2	1269	Cannabinoid receptor 2
CNTN1	1272	Contactin 1
CNTN2	6900	Contactin 2 (transiently-expressed axonal glycoprotein)
CNTNAP2	26047	Contactin-associated protein-like 2
COCH	1690	Cochlin
COG6	57511	Component of oligomeric golgi complex 6

Gene symbol	Entrez gene ID	Gene name
COL11A2	1302	Collagen XI, alpha-2 polypeptide
COL17A1	1308	Collagen XVII, alpha-1 polypeptide
COL1A1	1277	Collagen I, alpha-1 polypeptide
COL1A2	1278	Collagen I, alpha-2 polypeptide
COL3A1	1281	Collagen III, alpha-1 polypeptide
COL5A1	1289	Collagen V, alpha-1 polypeptide
COL6A5	256076	Collagen, type VI, alpha-5
COL7A1	1294	Collagen VII, alpha-1 polypeptide
COMT	1312	Catechol-O-methyltransferase
COX6A1	1337	Cytochrome c oxidase, subunit VIa, polypeptide-1
CP	1356	Ceruloplasmin
CPT2	1376	Carnitine palmitoyltransferase II
CREB1	1385	cAMP-response element-binding protein-1
CRHBP	1393	Corticotropin releasing hormone-binding protein
CRHR2	1395	Corticotropin releasing hormone receptor-2
CRLF1	9244	Cytokine-like factor 1
CRLF2	64109	Cytokine receptor-like factor 2
CRYAB	1410	Crystallin, alpha B
CSF3R	1441	Colony-stimulating factor-3 receptor (granulocyte)
CSNK1D	1453	Casein kinase-1, delta
CTSC	1075	Cathepsin C
CTSG	1511	Cathepsin G
CX3CR1	1524	Chemokine (C-X3-C) receptor 1 (G protein-coupled receptor-13)
CXCL8	3576	Chemokine, CXC motif, ligand 8
CYB561	1534	Cytochrome b-561
CYP19A1	1588	Cytochrome P450, family 19, subfamily A, polypeptide 1 (aromatization of androgens)
CYP1A2	1544	Cytochrome P450, subfamily I, aromatic compound-inducible, polypeptide 2
CYP2D6	1565	Cytochrome P450, subfamily IID, polypeptide 6
CYP3A4	1576	Cytochrome P450, subfamily IIIA (nifedipine oxidase) polypeptide 4
DBH	1621	Dopamine-beta-hydroxylase
DCC	1630	Deleted in colorectal carcinoma
DES	1674	Desmin
DICER1	23405	dicer 1, ribonuclease III
DMD	1756	Dystrophin
DMPK	1760	Dystrophia myotonica-protein kinase
DNMT1	1786	DNA methyltransferase 1
DOCK4	9732	Dedicator of cytokinesis 4
DPYSL2	1808	Dihydropyrimidinase-like 2
DRD2	1813	Dopamine receptor D2

Gene symbol	Entrez gene ID	Gene name
DRD3	1814	Dopamine receptor D3
DRP2	1821	Dystrophin-related protein 2
DST	667	Dystonin (bullous pemphigoid antigen 1)
DURS1	10674	Duane retraction syndrome 1
DUSP6	1848	Dual-specificity phosphatase-6
DYNC1H1	1778	Dynein, cytoplasmic-1, heavy chain-1
ECM1	1893	Extracellular matrix protein-1
EDN1	1906	Endothelin-1
EDNRA	1909	Endothelin receptor type A
EDNRB	1910	Endothelin receptor type B
EEC1	1913	Ectrodactyly, ectodermal dysplasia, cleft lip/palate, 1
ELAC2	60528	elaC ribonuclease Z 2
ELAVL3	1995	ELAV-like RNA binding protein 3
ELAVL4	1996	ELAV-like RNA binding protein 4 (Hu antigen D)
EMD	2010	Emerin
ENPP1	5167	Ectonucleotide pyrophosphatase/phosphodiesterase 1 (Ly-41 antigen, mouse, homolog of)
EPOR	2057	Erythropoietin receptor
EREG	2069	Epiregulin
ERGIC1	57222	Endoplasmic reticulum-golgi intermediate compartment protein 1
ESR1	2099	Estrogen receptor 1
ESR2	2100	Estrogen receptor-2 (ER beta)
ESRRB	2103	Estrogen-related receptor beta
F2	2147	Coagulation factor II (thrombin)
F5	2153	Coagulation factor V (proaccelerin, labile factor)
FAAH	2166	Fatty acid amide hydrolase
FAH	2184	Fumarylacetoacetase
FAM134B	54463	Family with sequence similarity 134, member B
FAM173B	134145	ATP synthase C subunit lysine N-methyltransferase
FANCM	57697	FANCM gene
FGF10	2255	Fibroblast growth factor-10
FGF12	2257	Fibroblast growth factor-12
FGF13	2258	Fibroblast growth factor-13
FGF14	2259	Fibroblast growth factor-14
FGF6	2251	Fibroblast growth factor-6
FGFR3	2261	Fibroblast growth factor receptor-3
FHL5	9457	Four-and-a-half LIM domains 5
FIGLA	344018	Folliculogenesis specific bHLH transcription factor
FKBP5	2289	FK506-binding protein 5
FLNA	2316	Filamin A, alpha (actin-binding protein-280)
FMR1	2332	FMRP translational regulator 1
FOXL1	2300	Forkhead box L1
FSHR	2492	Follicle stimulating hormone receptor

Gene symbol	Entrez gene ID	Gene name
FTL	2512	Ferritin, light chain
FTO	79068	FTO alpha-ketoglutarate dependent dioxygenase
FXN	2395	Frataxin
G6PD	2539	Glucose-6-phosphate dehydrogenase
GABBR2	9568	Gamma-aminobutyric acid B receptor 2
GABR	2550	Gamma-aminobutyric acid B receptor 1
GABRA2	2555	Gamma-aminobutyric acid (GABA) A receptor, alpha-2
GABRA3	2556	Gamma-aminobutyric acid (GABA) A receptor, alpha-3
GABRB1	2560	Gamma-aminobutyric acid (GABA) A receptor, beta-1
GABRB3	2562	Gamma-aminobutyric acid (GABA) A receptor, beta-3
GABRG2	2566	Gamma-aminobutyric acid (GABA) A receptor, gamma-2
GABRQ	55879	Gamma-aminobutyric acid receptor, theta
GABRR1	2569	Gamma-aminobutyric acid (GABA) A receptor, rho-1
GAL	51083	Galanin
GBA	2629	Glucosidase, acid beta
GBE1	2632	Glycogen branching enzyme
GBP1	2633	Guanylate binding protein 1, interferon-inducible, 67kD
GCH1	2643	GTP cyclohydrolase 1
GDAP1	54332	Ganglioside-induced differentiation-associated protein 1
GDF5	8200	Growth/differentiation factor-5 (cartilage-derived morphogenetic protein-1)
GDNF	2668	Glial cell line derived neurotrophic factor
GFAP	2670	Glial fibrillary acidic protein
GFRA2	2675	GDNF family receptor alpha-2
GJA1	2697	Gap junction protein, alpha-1, 43kD (connexin 43)
GJB2	2706	Gap junction protein, beta-2, 26kD (connexin 26)
GLA	2717	Galactosidase, alpha
GLO1	2739	Glyoxalase I
GNAO1	2775	Guanine nucleotide-binding protein (G protein), alpha-activating activity
GNAS	2778	GNAS complex locus (guanine nucleotide binding protein (G protein), alpha stimulating activity polypeptide 1)
GPD2	2820	Glycerol-3-phosphate dehydrogenase 2 (mitochondrial)
GRIA1	2890	Glutamate receptor, ionotropic, AMPA 1
GRIA2	2891	Glutamate receptor, ionotropic, AMPA 2
GRIA3	2892	Glutamate receptor, ionotropic, AMPA 3
GRIA4	2893	Glutamate receptor, ionotropic, AMPA 4
GRIC1	2897	Glutamate receptor, ionotropic, kainate 1
GRIC2	2898	Glutamate receptor, ionotropic, kainate 2
GRIC3	2899	Glutamate receptor, ionotropic, kainate 3
GRIC4	2900	Glutamate receptor, ionotropic, kainate 4
GRIC5	2901	Glutamate receptor, ionotropic, kainate 5
GRIN2A	2903	Glutamate receptor, ionotropic, N-methyl D-aspartate 2A
GRK5	2869	G protein-coupled receptor kinase 5

Gene symbol	Entrez gene ID	Gene name
GRM7	2917	Glutamate receptor, metabotropic, 7
GRP	2922	Gastrin-releasing peptide
GRPR	2925	Gastrin-releasing peptide receptor
GUCA1A	2978	Guanylate cyclase activator 1A, retina
H19	283120	Imprinted maternally expressed non-coding transcript
HBB	3043	Hemoglobin beta
HCK	3055	Hemopoietic cell kinase
HCN1	348980	Hyperpolarization-activated cyclic nucleotide-gated potassium channel 1
HCN2	610	Hyperpolarization-activated cyclic nucleotide-gated potassium channel 2
HCRTR1	3061	Hypocretin receptor 1
HCRTR2	3062	Hypocretin receptor 2
HEXB	3074	Hexosaminidase B, beta polypeptide
HEY2	23493	HES-related family bHLH transcription factor with YRPW motif 2
HFE	3077	Homeostatic iron regulator
HFM1	164045	Helicase for meiosis 1
HJURP	55355	Holliday junction recognition protein
HJV	148738	Hemojuvelin
HK1	3098	Hexokinase-1
HLA-B	3106	Major histocompatibility complex, class I, B
HLA-DQ	3119	Major histocompatibility complex, class II, DQ beta-1
HLA-DQA1	3117	Major histocompatibility complex, class II, DQ alpha-1
HLA-DRA	3122	Major histocompatibility complex, class II, DR alpha
HOMER1	9456	Homer scaffold protein 1
HOXB8	3218	Homeo box-B8
HPGD	3248	Hydroxyprostaglandin dehydrogenase 15-(NAD)
HPRT1	3251	Hypoxanthine phosphoribosyltransferase 1
HPSE2	60495	Heparanase 2
HSPB1	3315	Heat-shock 27kD protein-1
HTR1A	3350	5-hydroxytryptamine (serotonin) receptor-1A
HTR1B	3351	5-hydroxytryptamine (serotonin) receptor-1B
HTR2A	3356	5-hydroxytryptamine (serotonin) receptor-2A
HTR2C	3358	5-hydroxytryptamine (serotonin) receptor-2C
HTR3A	3359	5-hydroxytryptamine (serotonin) receptor-3
HTR3B	9177	5-hydroxytryptamine receptor 3B
HTR7	3363	5-hydroxytryptamine (serotonin) receptor-7, adenylate cyclase-coupled
HTRA1	5654	HTRA serine peptidase 1
IARS2	55699	Isoleucyl-tRNA synthetase 2
ICAM1	3383	Intercellular adhesion molecule-1
IFNG	3458	Interferon, gamma

Gene symbol	Entrez gene ID	Gene name
IFRD1	3475	Interferon-related developmental regulator 1
IGF1	3479	Insulin-like growth factor-1, or somatomedin C
IGF2	3481	Insulin-like growth factor-2, or somatomedin A
IGLL1	3543	Immunoglobulin lambda-like polypeptide 1 (immunoglobulin omega peptide)
IGSF9B	22997	Immunoglobulin superfamily, member 9B
IKBKAP	8518	Elongator complex protein 1
IKBKG	8517	Inhibitor of nuclear factor kappa B kinase, regulatory subunit gamma
IL10	3586	Interleukin-10
IL10RB	3588	Cytokine receptor, family II, member 4
IL12	3593	Interleukin-12B (natural killer cell stimulatory factor-2, cytotoxic lymphocyte maturation factor-2, p40)
IL13	3596	Interleukin-13
IL16	3603	Interleukin 16
IL18	3606	Interleukin-18
IL18R1	8809	Interleukin 18 receptor 1
IL1A	3552	Interleukin-1, alpha
IL1B	3553	Interleukin-1, beta
IL1R1	3554	Interleukin-1 receptor, type I
IL1R2	7850	Interleukin-1 receptor, type II
IL1RN	3557	Interleukin-1 receptor antagonist
IL31	386653	Interleukin 31
IL31RA	133396	Interleukin 31 receptor A
IL33	90865	Interleukin 33
IL37	27178	Interleukin 37
IL4	3565	Interleukin-4
IL6	3569	Interleukin-6 (interferon, beta-2)
IL7R	3575	Interleukin-7 receptor
IL9	3578	Interleukin-9
IMMP2L	83943	Inner mitochondrial membrane peptidase, subunit 2, S. cerevisiae, homolog of
INADL	10207	PALS1-associated tight junction protein
INSR	3643	Insulin receptor
ISCU	23479	Iron-sulfur cluster scaffold, E. coli, homolog of
ITCH	83737	Itchy, mouse, homolog of
ITPK1	3705	Inositol 1,3,4-trisphosphate 5/6-kinase
JAG1	182	Jagged 1
JPH2	57158	Junctophilin 2
JUN	3725	Jun proto-oncogene, AP-1 transcription factor subunit
JUNB	3726	Jun B proto-oncogene, AP-1 transcription factor subunit
KARS1	3735	Lysyl-tRNA synthetase 1
KCNA1	3736	Potassium voltage-gated channel, shaker-related subfamily, member 1

Gene symbol	Entrez gene ID	Gene name
KCNA2	3737	Potassium channel, voltage-gated, Shaker-related subfamily, member 2
KCNA4	3739	Potassium voltage-gated channel, shaker-related subfamily, member 4
KCNA5	3741	Potassium voltage-gated channel, shaker-related subfamily, member 5
KCNB2	9312	Potassium channel, voltage-gated, shab-related subfamily, member 2
KCND2	3751	Potassium voltage-gated channel, Shal-related subfamily, member 2
KCNE2	9992	Potassium voltage-gated channel, Isk-related family, member 2
KCNIP3	30818	Potassium channel-interacting protein 3
KCNJ10	3766	Potassium inwardly-rectifying channel, subfamily J, member 10
KCNJ3	3760	Potassium inwardly-rectifying channel, subfamily J, member 3
KCNJ6	3763	Potassium inwardly-rectifying channel, subfamily J, member 6
KCNK1	3775	Potassium channel, subfamily K, member 1
KCNK12	56660	Potassium channel, subfamily K, member 12
KCNK18	338567	Potassium channel, subfamily K, member 18
KCNK2	3776	Potassium channel, subfamily K, member 2
KCNK5	8645	Potassium channel, subfamily K, member 5
KCNK9	51305	Potassium channel, subfamily K, member 9
KCNN3	3782	Potassium channel, calcium-activated, intermediate/small conductance, subfamily N, member 3
KCNQ1	3784	Potassium voltage-gated channel, KQT-like subfamily, member 1
KCNQ1OT1	10984	KCNQ1-overlapping transcript 1
KCNQ2	3785	Potassium voltage-gated channel, KQT-like subfamily, member 2
KCNQ3	3786	Potassium voltage-gated channel, KQT-like subfamily, member 3
KCNQ4	9132	Potassium voltage-gated channel, KQT-like subfamily, member 4
KCNQ5	56479	Potassium channel, voltage-gated, KQT-like subfamily, member 5
KCNS1	3787	Potassium channel, voltage-gated, delayed-rectifier, subfamily S, member 1
KCNS3	3790	Potassium voltate-gated channel, delayed-rectifier, subfamily S, member 3
KCNT2	343450	Potassium channel, subfamily T, member 2
KDM4C	23081	Lysine demethylase 4C
KIAA1456	57604	tRNA Methyl Tranferase 9-like
KIF1A	547	Kinesin family member 1A
KIF20B	9585	M-phase phosphoprotein 1
KIF5A	3798	Kinesin family member 5A
KLHL24	54800	Kelch-like 24
KLK5	25818	Kallikrein-related peptidase 5

Gene symbol	Entrez gene ID	Gene name
KRIT1	889	KREV interaction trapped 1
KRT16	3868	Keratin 16
KRT6A	3853	Keratin 6A
LAMB3	3914	Laminin, beta-3 (nicein, 125kD; kalinin, 140kD; BM600, 125kD)
LDLRAD3	143458	Low density lipoprotein receptor class A domain-containing protein 3
LEP	3952	Leptin (murine obesity homolog)
LIF	3976	Leukemia inhibitory factor (cholinergic differentiation factor)
LIPG	9388	Lipase G, endothelial
LMNA	4000	Lamin A/C
LRATD1	151354	Family with sequence similarity 84, member A
LRP1	4035	Low density lipoprotein-related protein-1 (alpha-2-macroglobulin receptor)
LTA	4049	Lymphotoxin alpha (formerly tumor necrosis factor beta)
LYZ	4069	Lysozyme
LZTR1	8216	Leucine-zipper-like regulator-1
MACC1	346389	MET transcriptional regulator MACC1
MAOA	4128	Monoamine oxidase A
MAOB	4129	Monoamine oxidase B
MAP1B	4131	Microtubule-associated protein-1B
MAP2K1	5604	Mitogen-activated protein kinase kinase 1
MAPK1	5594	Mitogen-activated protein kinase 1
MAPK14	1432	Mitogen-activated protein kinase 14 (cytokine suppressive anti-inflammatory drug binding protein 1)
MAPK3	5595	Mitogen-activated protein kinase 3
MAPK9	5601	Mitogen-activated protein kinase 9
MAPT	4137	Microtubule-associated protein tau
MARCH4	57574	Membrane-associated RING-CH finger protein 4
MARS1	4141	Methioninyl-tRNA synthetase 1
MATR3	9782	Matrin 3
MC1R	4157	Melanocortin-1 receptor (alpha melanocyte-stimulating hormone receptor)
MECP2	4204	Methyl-CpG-binding protein-2
MEF2D	4209	MADS box transcription enhancer factor 2, polypeptide D (myocyte enhancer factor 2D)
MEFV	4210	MEFV innate immunity regulator, pyrin
MEIS1	4211	Meis homeobox 1
MEN1	4221	Menin
METTL21A	151194	Methyltransferase-like 21A
MFN2	9927	Mitofusin 2
MIR183	406959	Micro RNA 183
MITF	4286	Microphthalmia-associated transcription factor
MLH1	4292	DNA mismatch repair protein MLH1

Gene symbol	Entrez gene ID	Gene name
MMP16	4325	Matrix metalloproteinase 16 (membrane-inserted)
MMP17	4326	Matrix metalloproteinase 17
MMP2	4313	Matrix metalloproteinase 2 (gelatinase A, 72kD type IV collagenase)
MMP9	4318	Matrix metalloproteinase 9 (gelatinase B, 92kD type IV collagenase)
MPDZ	8777	Multiple PDZ domain protein
MPPED2	744	Metallophosphoesterase domain-containing protein 2
MPZ	4359	Myelin protein zero
MRGPRX1	259249	Mas-related G protein-coupled receptor family, member X1
MRVI1	10335	Murine retrovirus integration site 1, homolog of
MSN	4478	Moesin
MT-CYB	4519	Cytochrome b of Complex III
MTDH	92140	Metadherin
MTHFD1	4522	Methylenetetrahydrofolate dehydrogenase 1
MTHFR	4524	Methylenetetrahydrofolate reductase
MTOR	2475	Mechanistic target of rapamycin
MTR	4548	5-methyltetrahydrofolate-homocysteine methyltransferase 1
MT-RNR1	4549	Mitochondrial 12S ribosomal RNA
MTRR	4552	Methionine synthase reductase
MT-TL1	4567	Mitochondrial tRNA for leucine
MT-TS1	4574	Mitochondrial tRNA for serine
MYCL	4610	MYCL protooncogene, bHLH transcription factor
MYD88	4615	MYD88 innate immune signal transduction adaptor
MYH14	79784	Myosin, heavy chain 14, nonmuscle
MYH9	4627	Myosin, heavy polypeptide-9, nonmuscle
MYO7A	4647	Myosin VIIA
MYORG	57462	Myogenesis-regulating glycosidase
MYT1L	23040	Myelin transcription factor 1-like
NAGLU	4669	N-acetylglucosaminidase, alpha-
NALCN	259232	Sodium leak channel, nonselective
NCOA7	135112	Nuclear receptor coactivator 7
NECAB2	54550	N-terminal EF-hand calcium-binding protein 2
NECTIN4	81607	Nectin 4
NEDD4	4734	Neural precursor cell expressed, developmentally downregulated-4
NEDD4L	23327	Ubiquitin protein ligase NEDD4-like
NEFM	4741	Neurofilament protein, medium polypeptide
NF1	4763	Neurofibromin (neurofibromatosis, type I)
NF2	4771	Merlin
NFASC	23114	Neurofascin
NFKB1A	4790	Nuclear factor kappa-B, subunit 1 (p105, p50)
NFKBIA	4792	Nuclear factor kappa-B inhibitor, alpha
NFKBIL1	4795	Nuclear factor-kappa-B inhibitor-like protein 1

Gene symbol	Entrez gene ID	Gene name
NGF	4803	Nerve growth factor, beta
NGFR	4804	Nerve growth factor receptor
NLRP1	22861	NLR family, pyrin domain containing 1
NNMT	4837	Nicotinamide N-methyltransferase
NOS1	4842	Nitric oxide synthase 1, neuronal
NOS2	4843	Nitric oxide synthase 2A
NOS3	4846	Nitric oxide synthase 3, endothelial cell
NOTCH1	4851	Notch receptor 1
NOTCH3	4854	Notch, Drosophila, homolog of, 3
NPPB	4879	Natriuretic peptide precursor B
NPSR1	387129	Neuropeptide S receptor 1
NPSR1-AS1	404744	NPSR1 antisenase RNA 1
NPY1R	4886	Neuropeptide Y receptor
NR0B1	190	Nuclear receptor subfamily 0, group B, member 1
NR1	2902	Glutamate receptor, ionotropic, N-methyl D-aspartate 1
NR2	2904	Glutamate receptor, ionotropic, N-methyl D-aspartate 2B
NR3C1	2908	Nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)
NRCAM	4897	Neuronal cell adhesion molecule
NRG1	3084	Neuregulin 1 (heregulin, alpha, 45kD; ERBB2 p185-activator)
NRIP1	8204	Nuclear receptor interacting protein 1 (receptor interacting protein 140)
NRXN3	9369	Neurexin 3
NSRP1	84081	Nuclear speckle splicing regulatory protein 1
NTRK1	4914	Neurotrophic tyrosine kinase, receptor, type 1
NYS4	317685	Nystagmus 4, congenital, autosomal dominant
OPA1	4976	OPA1 mitochondrial dynamin-like GTPase
OPRD1	4985	Opioid receptor, delta-1
OPRK1	4986	Opiate receptor, kappa-1
OPRL1	4987	Opioid receptor-like 1
OPRM1	4988	Opioid receptor, mu-1
OTOF	9381	Otoferlin
OTOG	340990	Otogelin
P2RX2	22953	Purinergic receptor P2X, ligand-gated ion channel, 2
P2RX3	5024	Purinergic receptor P2X, ligand-gated ion channel, 3
P2RX4	5025	Purinergic receptor P2X, ligand-gated ion channel, 4
P2RX6	9127	Purinergic receptor P2X-like 1
P2RX7	5027	Purinergic receptor P2X, ligand-gated ion channel, 7
P2RY2	5029	Purinergic receptor P2Y, G-protein coupled, 2
P2RY4	5030	Pyrimidinergic receptor P2Y, G-protein coupled, 4
PACERR	103752588	PTGS2 antisense NFKB1 complex-mediated expression regulator RNA, noncoding

Gene symbol	Entrez gene ID	Gene name
PAFAH1B1	5048	Platelet-activating factor acetylhydrolase, isoform 1B, alpha subunit
PAH	5053	Phenylalanine hydroxylase
PARK2	5071	Parkin
PATL2	197135	PAT1 homolog 2
PAX3	5077	Paired box homeotic gene-3
PBX1	5087	Pre-B cell leukemia transcription factor-1
PCDHA3	56145	Protocadherin-alpha 3
PCSK6	5046	Proprotein convertase subtilisin/kexin type 6
PDE5A	8654	Phosphodiesterase 5A
PDGFC	56034	Platelet-derived growth factor C
PDYN	5173	Prodynorphin
PDZD2	23037	PDZ domain containing 2
PENK	5179	Proenkephalin
PGR	5241	Progesterone receptor
PHACTR1	221692	Phosphatase and actin regulator 1
PIEZO2	63895	PIEZO-type mechanosensitive ion channel component 2
PIK3C2G	5288	Phosphatidylinositol 3-kinase, class 2, gamma
PINK1	65018	PTEN-induced putative kinase 1
PIRT	644139	Phosphoinositide-interacting regulator of transient receptor potential channels
PKD2	5311	Polycystin-2
PLA2	8398	Phospholipase A2, group VI
PLAUR	5329	Plasminogen activator, urokinase, receptor
PLCE1	51196	Phospholipase C, epsilon-1
PLCG1	5335	Phospholipase C, gamma 1 (formerly subtype 148)
PLEKHA4	57664	Pleckstrin homology domain-containing protein, family A, member 4
PMP2	5375	Peripheral myelin protein-2
PMP22	5376	Peripheral myelin protein-22
PNOC	5368	Prepronociceptin
PNPLA6	10908	Patatin-like phospholipase domain-containing protein 6
POLE	5426	Polymerase, DNA, epsilon
POLG	5428	Polymerase (DNA directed), gamma
POMC	5443	Proopiomelanocortin (adrenocorticotropin/beta-lipotropin)
PON1	5444	Paraoxonase-1
POT1	25913	Protection of telomeres 1
PPARG	5468	Peroxisome proliferator activated receptor, gamma
PPOX	5498	Protoporphyrinogen oxidase
PPP1R1C	151242	Protein phosphatase 1, regulatory subunit 1C
PRDM12	59335	PR domain-containing protein 12
PRDM16	63976	PR domain-containing protein 16
PRKACA	5566	Protein kinase, cAMP-dependent, catalytic, alpha
PRKCA	5578	Protein kinase C, alpha polypeptide

Gene symbol	Entrez gene ID	Gene name
PRKCB	5579	Protein kinase C, beta 1 polypeptide
PRKCE	5581	Protein kinase C, epsilon
PRKCZ	5590	Protein kinase C, zeta form
PRLHR	2834	G protein-coupled receptor-10
PRND	23627	Prion-like protein doppel
PRNP	5621	Prion protein
PROK2	60675	Prokineticin 2
PRPS1	5631	Phosphoribosyl pyrophosphate synthetase-1
PRRT2	112476	Proline-rich transmembrane protein 2
PRX	57716	Periaxin
PSEN1	5663	Presenilin 1
PTEN	5728	Phosphatase and tensin homolog (mutated in multiple advanced cancers 1)
PTGER1	5731	Prostaglandin E receptor 1, EP1 subtype, 42kD
PTGER2	5732	Prostaglandin E receptor 2, EP2 subtype, 53kD
PTGIR	5739	Prostaglandin I2 (prostacyclin) receptor (IP)
PTGS1	5742	Prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)
PTGS2	5743	Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)
PTPRD	5789	Protein tyrosine phosphatase, receptor type, delta polypeptide
PTPRZ1	5803	Protein-tyrosine phosphatase, receptor-type, zeta-1, polypeptide
PTX3	5806	Pentraxin-3
RAB7A	7879	Ras-associated protein RAB7
RAMP1	10267	Receptor activity-modifying protein 1
RB1	5925	Retinoblastoma-1
RBFOX1	54715	RNA-binding protein FOX1, C. elegans, homolog of, 1
REST	5978	RE1-silencing transcription factor
RGS12	6002	Regulator of G protein signaling 12
RHBDF2	79651	Rhomboid 5, Drosophila, homolog of, 2
RHO	6010	Rhodopsin (opsin 2)
RIC3	79608	RIC3 acetylcholine receptor chaperone
RNF11	26994	Ring finger protein 11
RNF170	81790	RING finger protein 170
RNF213	57674	Ring finger protein 213
RPE65	6121	RPE65 retinoid isomerohydrolase
RUNX2	860	Runt-related transcription factor 2
RXFP2	122042	Leucine-rich repeat-containing G protein-coupled receptor 8
SCLT1	132320	Sodium channel and clathrin linker 1
SCN10A	6336	Sodium voltage-gated channel, alpha subunit 10
SCN11A	11280	Sodium voltage-gated channel, alpha subunit 11
SCN1A	6323	Sodium voltage-gated channel, alpha subunit 1
SCN1B	6324	Sodium voltage-gated channel, beta subunit 1

Gene symbol	Entrez gene ID	Gene name
SCN2A	6326	Sodium voltage-gated channel, alpha subunit 2
SCN2B	6327	Sodium voltage-gated channel, beta subunit 1
SCN3A	6328	Sodium voltage-gated channel, alpha subunit 3
SCN3B	55800	Sodium voltage-gated channel, beta subunit 3
SCN4A	6329	Sodium voltage-gated channel, alpha subunit 4
SCN4B	6330	Sodium voltage-gated channel, beta subunit 4
SCN5A	6331	Sodium channel, voltage-gated, type V, alpha polypeptide
SCN7A	6332	Sodium voltage-gated channel, alpha subnit 7
SCN8A	6334	Sodium voltage-gated channel, alpha subunit 8
SCN9A	6335	Sodium voltage-gated channel, alpha subunit 9
SCNN1A	6337	Sodium channel, epithelial 1, subunit alpha
SDHD	6392	Succinate dehydrogenase complex, subunit D, integral membrane protein
SELENON	57190	Selenoprotein N
SEMA3A	10371	Semaphorin 3A
SEPT7	989	Septin 7
SEPT9	10801	Septin 9
SERPINA6	866	Corticosteroid-binding globulin
SERPINB6	5269	Protease inhibitor 6 (placental thrombin inhibitor)
SERPINB7	8710	Serpin peptidase inhibitor, clade B (ovalbumin), member 7
SGCA	6442	Sarcoglycan, alpha (50kD dystrophin-associated glycoprotein; adhalin)
SHH	6469	Sonic hedgehog signaling molecule
SHMT1	6470	Serine hydroxymethyltransferase (soluble)
SIN3B	23309	Sin3, yeast, homolog of, B
SLC11A2	4891	Solute carrier family 11 (proton-coupled divalent metal ion transporter) member 2
SLC17A6	57084	Solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 6
SLC17A8	246213	Solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 8 (vesicular glutamate transporter 3)
SLC22A5	6584	Solute carrier, family 22 (organic cation transporter), member 5
SLC24A3	57419	Solute carrier family 24 (sodium/potassium/calcium exchanger), member 3
SLC25A46	91137	Solute carrier family 25, member 46
SLC26A4	5172	Solute carrier family 26 (sulfate transporter), member 4
SLC2A1	6513	Solute carrier family 2 (facilitated glucose transporter), member 1
SLC6A2	6530	Solute carrier family 6 (neurotransmitter transporter, noradrenalin), member 2, cocaine- and antidepressant-sensitive
SLC6A3	6531	Solute carrier family 6 (neurotransmitter transporter, dopamine), member 3
SLC6A4	6532	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4
SLC8A1	6546	Solute carrier family 8, member 1 (sodium-calcium exchanger-1)

Gene symbol	Entrez gene ID	Gene name
SLC8A2	6543	Solute carrier family 8, member 2 (sodium-calcium exchanger 2)
SLC9A9	285195	Solute carrier family 9 (sodium/hydrogen exchanger), member A9
SMAD3	4088	SMAD family member 3
SMAD4	4089	SMAD family member 4
SMARCA4	6597	SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily A, member 4
SMARCAD1	56916	SWI/SNF-related, matrix-associated actin-dependent regulator of chromatin, subfamily A, DEAD/H box-containing, 1
SMARCB1	6598	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1
SMN1	6606	Survival of motor neuron 1, telomeric
SMN2	6607	Survival of motor neuron 2, centromeric
SNAP25	6616	Synaptosomal-associated protein, 25kD
SNCA	6622	Synuclein, alpha (non A4 component of amyloid precursor)
SOD1	6647	Superoxide dismutase-1, soluble
SOD2	6648	Superoxide dismutase-2, mitochondrial
SP4	6671	Sp4 transcription factor
SPG16	57760	Spastic paraplegia 16, X-linked, complicated
SPINK2	6691	Serine protease inhibitor, Kazal-type, 2
SPON1	10418	F-spondin, Rat, homolog of
SPTBN4	57731	Spectrin, beta, nonerythrocytic, 4 (quivering, mouse, homolog of)
SPTLC1	10558	Serine palmitoyltransferase, long-chain base subunit 1
SPTLC2	9517	Serine palmitoyltransferase, long-chain base subunit 2
STAT6	6778	Signal transducer and activator of transcription-6, interleukin-4 induced
STK11	6794	Serine/threonine protein kinase-11
STOML3	161003	Stomatin-like protein 3
STUB1	10273	STIP1 homologous and U box-containing protein 1
STX1A	6804	Syntaxin 1A, brain
SUGCT	79783	Chromosome 7 open reading frame 10
SV2C	22987	Synaptic vesicle glycoprotein 2C
SVEP1	79987	Sushi, von willebrand factor type A, EGF, and pentraxin domain-containing 1
SYN2	6854	Synapsin II
SYN3	8224	Synapsin III
SYNE1	23345	Spectrin repeat-containing nuclear envelope protein 1 (nesprin 1)
SYT16	83851	Synaptotagmin 16
TAAR1	134864	Trace amine-associated receptor 1
TAAR2	9287	Trace amine-associated receptor 2
TAC1	6863	Tachykinin 1 (substance K; neurokinin A; neurokinin 2; neuromedin L; neuropeptide gamma; tachykinin 2)

Gene symbol	Entrez gene ID	Gene name
TACR1	6869	Tachykinin receptor 1 (substance P receptor; neurokinin-1 receptor)
TAOK3	51347	TAO kinase 3
TCL1A	8115	T-cell lymphoma/leukemia 1A
TCN1	6947	Transcobalamin I
TECTA	7007	Tectorin, alpha
TERT	7015	Telomerase reverse transcriptase
TF	7018	Transferrin
TFRC	7037	Transferrin receptor
TG	7038	Thyroglobulin
TGFB1	7040	Transforming growth factor, beta-1
TGFBR2	7048	Transforming growth factor, beta receptor II, 70-80kD
TGM1	7051	Transglutaminase-1 (K polypeptide epidermal type I, protein-glutamine gamma-glutamyltransferase)
TH	7054	Tyrosine hydroxylase
THBS2	7058	Thrombospondin 2
TIPIN	54962	Timeless-interacting protein
TLR2	7097	Toll-like receptor-2
TLR4	7099	Toll-like receptor-4
TMOD2	29767	Tropomodulin 2, neuronal
TNC	3371	Tenascin C (hexabrachion)
TNF	7124	Tumor necrosis factor (cachectin)
TNFRSF11A	8792	Tumor necrosis factor receptor superfamily, member 11A
TNFRSF11B	4982	Tumor necrosis factor receptor superfamily, member 11B (osteoprotegerin)
TNFRSF1B	7133	Tumor necrosis factor receptor superfamily, member 1B
TNR	7143	Tenascin R (restrictin, janusin)
TP53	7157	Tumor protein p53
TPH1	7166	Tryptophan hydroxylase 1 (tryptophan-5-monooxygenase)
TPH2	121278	Tryptophan hydroxylase 2
TPK1	27010	Thiamine pyrophosphokinase
TPM3	7170	Tropomyosin 3
TPP1	1200	Tripeptidyl peptidase 1
TRIM21	6737	Tripartite motif-containing protein 21
TRPA1	8989	Transient receptor potential cation channel, subfamily A, member 1
TRPC1	7220	Transient receptor potential channel 1
TRPC3	7222	Transient receptor potential cation channel, subfamily C, member 3
TRPC4	7223	Transient receptor potential cation channel, subfamily C, member 4
TRPC6	7225	Transient receptor potential channel-6
TRPM1	4308	Transient receptor potential cation channel, subfamily M, member 1 (melastatin)

Gene symbol	Entrez gene ID	Gene name
TRPM2	7226	Transient receptor potential cation channel, subfamily M, member 2
TRPM3	80036	Transient receptor potential cation channel, subfamily M, member 3
TRPM4	54795	Transient receptor potential cation channel, subfamily M, member 4
TRPM5	29850	Transient receptor potential cation channel, subfamily M, member 5 (MLSN1- and TRP-related gene 1)
TRPM6	140803	Transient receptor potential cation channel, subfamily M, member 6
TRPM7	54822	Transient receptor potential cation channel, subfamily M, member 7
TRPM8	79054	Transient receptor potential cation channel, subfamily M, member 8
TRPP	5310	Polycystin-1
TRPV1	7442	Transient receptor potential cation channel, subfamily V, member 1 (vanilloid receptor 1; capsaicin receptor)
TRPV2	51393	Transient receptor potential cation channel, subfamily V, member 2
TRPV3	162514	Transient receptor potential cation channel, subfamily V, member 3
TRPV4	59341	Transient receptor potential cation channel, subfamily V, member 4 (vanilloid receptor-related osmotically activated channel)
TRPV5	56302	Transient receptor potential cation channel, subfamily V, member 5
TRPV6	55503	Transient receptor potential cation channel, subfamily V, member 6
TSC1	7248	Hamartin (tuberous sclerosis 1 gene)
TSC2	7249	Tuberin (tuberous sclerosis 2 gene)
TSLP	85480	Thymic stromal lymphopoietin
TSPAN2	10100	Tetraspanin 2
TSPEAR	54084	Thrombospondin-type laminin G domain and EAR repeats
TSPO	706	Translocator protein, 18kD (benzodiazepine receptor, peripheral type)
TTR	7276	Transthyretin (prealbumin)
TUSC5	286753	Tumor suppressor candidate 5
TXNRD2	10587	Thioredoxin reductase 2
UGT2B7	7364	UDP-glucuronyltransferase, family 2, beta-7
USP12	219333	Ubiquitin-specific protease 12
UTS2	10911	Urotensin II
VDR	7421	Vitamin D (1,25-dihydroxyvitamin D3) receptor
VEGFA	7422	Vascular endothelial growth factor
VHL	7428	von Hippel-Lindau tumor suppressor
VWA5A	4013	von Willebrand factor A domain-containing protein 5A
WFS1	7466	Wolframin
WNK1	65125	WNK lysine deficient protein kinase 1

Gene symbol	Entrez gene ID	Gene name
WNK2	65268	WNK lysine deficient protein kinase 2
WNT10A	80326	Wingless-type MMTV integration site family, member 10A
XRCC3	7517	X-ray repair corss complementing 3
YAP1	10413	YES1-associated transcriptional regulator
ZNF235	9310	Zinc finger protein 93, mouse, homolog of
ZSCAN20	7579	Zinc finger- and SCAN domain-containing protein 20

Table A.3: List of candidate genes used in this project as 'candidate gene' filter that was curated by using multiple databases in a multi-step procedure.

Appendix A.4

Gene symbol	Gene name
ABCB4	ATP-binding cassette, subfamily B, member 4
AGL	Amylo-1,6-glucosidase, 4-alpha-glucanotransferase (glycogen debranching enzyme)
ANO1	Anoctamin 1, calcium activated chloride channel
AOC1	Amiloride-binding protein-1
ATAD3A	ATPase family, AAA domain-containing, member 3A
ATL1	Atlastin GTPase 1
ATP2A2	ATPase, Ca++ transporting, slow-twitch, cardiac muscle-2
ATP7B	ATPase, Cu++ transporting, beta polypeptide
BBS1	Bardet-Biedl syndrome 1 gene
BDKRB1	Bradykinin receptor B1
CACNA1A	Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit
CACNA1E	Calcium channel, voltage-dependent, alpha 1E subunit
CACNA1G	Calcium channel, voltage-dependent, T type, alpha-1G subunit
CACNA1S	Calcium channel, voltage-dependent, L type, alpha 1S subunit
CACNB1	Calcium channel, voltage-dependent, beta 1 subunit
CACNB2	Calcium channel, voltage-dependent, beta 2 subunit
CARF	Calcium response factor
CAST	Calpastatin
CHRNA10	Cholinergic receptor, neuronal nicotinic, alpha polypeptide 10
CHRNA2	Cholinergic receptor, nicotinic, alpha polypeptide-2
CHRNA6	Cholinergic receptor, neuronal nicotinic, alpha polypeptide 6
CHRNA9	Cholinergic receptor, neuronal nicotinic, alpha polypeptide 9
CLCN1	Chloride channel-1, skeletal muscle
CLTCL1	Clathrin, heavy polypeptide-like 1
COL1A2	Collagen I, alpha-2 polypeptide
COL3A1	Collagen III, alpha-1 polypeptide
COL5A1	Collagen V, alpha-1 polypeptide
COL6A2	Collagen VI, alpha-2 polypeptide
COL6A5	Collagen, type VI, alpha-5
COL7A1	Collagen VII, alpha-1 polypeptide
CRHR2	Corticotropin releasing hormone receptor-2
DNMT1	DNA methyltransferase 1
ESR2	Estrogen receptor-2
FSHR	Follicle stimulating hormone receptor
GABRD	Gamma-aminobutyric acid (GABA) A receptor, delta
GBA	Glucosidase, acid beta
GLA	Galactosidase, alpha
GRIA2	Glutamate receptor, ionotropic, AMPA 2
HCRTR2	Hypocretin receptor 2
KCNA2	Potassium channel, voltage-gated, Shaker-related subfamily, member 2
KCNA5	Potassium voltage-gated channel, shaker-related subfamily, member 5
KCND3	Potassium voltage-gated channel, Shal-related subfamily, member 3

Gene symbol	Gene name
KCNE2	Potassium voltage-gated channel, Isk-related family, member 2
KCNH1	Potassium voltage-gated channel, subfamily H, member 1
KCNH8	Potassium voltage-gated channel, subfamily H, member 8
KCNJ12	Potassium inwardly-rectifying channel, subfamily J, member 12
KCNK18	Potassium channel, subfamily K, member 18
KCNK5	Potassium channel, subfamily K, member 5
KCNQ2	Potassium voltage-gated channel, KQT-like subfamily, member 2
KCNQ3	Potassium voltage-gated channel, KQT-like subfamily, member 3
KCNQ4	Potassium voltage-gated channel, KQT-like subfamily, member 4
KIF1A	Kinesin family member 1A
KIF21A	Kinesin family member 21A
KRT16	Keratin 16
LAMB3	Laminin, beta-3 (nicein, 125kD; kalinin, 140kD; BM600, 125kD)
LMNA	Lamin A/C
LZTR1	Leucine-zipper-like regulator-1
MARS	Methionyl-tRNA Synthetase 1
MEFV	MEFV innate immunity regulator, pyrin
MEN1	Menin 1
MITF	Microphthalmia-associated transcription factor
MMP9	Matrix metalloproteinase 9
MYLK2	Myosin light chain kinase 2
MYT1L	Myelin transcription factor 1-like
NAGLU	N-acetylglucosaminidase, alpha-
NECAB2	N-terminal EF-hand calcium-binding protein 2
NEFH	Neurofilament, heavy polypeptide
NF1	Neurofibromin 1
NF2	Neurofibromin 2
NOTCH3	Notch, Drosophila, homolog of, 3
NRXN3	Neurexin 3
NTRK1	Neurotrophic tyrosine kinase, receptor, type 1
NTRK3	Neurotrophic tyrosine kinase, receptor, type 3
OPRL1	Opioid receptor-like 1
OPRM1	Opioid receptor, mu-1
P2RX3	Purinergic receptor P2X, ligand-gated ion channel, 3
PIEZO2	PIEZO-type mechanosensitive ion channel component 2
PMP22	Peripheral myelin protein-22
PNOC	Prepronociceptin
PRDM16	PR domain-containing protein 16
RPE65	Retinoid Isomerohydrolase RPE65
RYR1	Ryanodine receptor-1, skeletal
SCN10A	Sodium voltage-gated channel, alpha subunit 10
SCN11A	Sodium voltage-gated channel, alpha subunit 11
SCN3A	Sodium voltage-gated channel, alpha subunit 3
SCN3B	Sodium voltage-gated channel, beta subunit 3
SCN5A	Sodium channel, voltage-gated, type V, alpha polypeptide

Gene symbol	Gene name
SCN8A	Sodium voltage-gated channel, alpha subunit 8
SCN9A	Sodium voltage-gated channel, alpha subunit 9
SLC45A1	Solute carrier family 45, member 1
SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1
SPTBN2	Spectrin, beta, nonerythrocytic, 2
STOML3	Stomatin-like protein 3
STUB1	STIP1 homologous and U box-containing protein 1
SV2C	Synaptic vesicle glycoprotein 2C
TERT	Telomerase reverse transcriptase
TLR4	Toll-like receptor-4
TNFRSF1A	Tumor necrosis factor receptor superfamily, member 1A
TRPA1	Transient receptor potential cation channel, subfamily A, member 1
TRPC6	Transient receptor potential channel-6
TRPM1	Transient receptor potential cation channel, subfamily M, member 1 (melastatin)
TRPM2	Transient receptor potential cation channel, subfamily M, member 2
TRPM3	Transient receptor potential cation channel, subfamily M, member 3
TRPM4	Transient receptor potential cation channel, subfamily M, member 4
TRPM5	Transient receptor potential cation channel, subfamily M, member 5
TRPM7	Transient receptor potential cation channel, subfamily M, member 7
TRPM8	Transient receptor potential cation channel, subfamily M, member 8
TRPV1	Transient receptor potential cation channel, subfamily V, member 1 (vanilloid receptor 1; capsaicin receptor)
TRPV2	Transient receptor potential cation channel, subfamily V, member 2
WARS	Tryptophanyl-tRNA synthetase
WFS1	Wolframin ER Transmembrane Glycoprotein
WNK1	WNK lysine deficient protein kinase 1
ZSCAN20	Zinc finger- and SCAN domain-containing protein 20

Table A.4: List of 113 prioritised genes observed in the entire cohort

Appendix A.5

Permissions

Figure 1: License for Figure 6 in the Introduction section.

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Feb 14, 2022

This Agreement between Ms. Kaalindi Misra ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5247560336426
License date	Feb 14, 2022
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Cellular and Molecular Neurobiology
Licensed Content Title	Opioid Receptor Regulation of Neuronal Voltage-Gated Calcium Channels
Licensed Content Author	Norbert Weiss et al
Licensed Content Date	Jun 8, 2020
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Student
Institution name	Ospedale San Raffaele
Expected presentation date	Mar 2022
Portions	Figure 1 c
Requestor Location	Ms. Kaalindi Misra Via Olgettina 60 Milano, 20132 Italy Attn: Ospedale San Raffaele
Total	**0.00 USD**
Terms and Conditions	

Springer Nature Customer Service Centre GmbH

Figure 2: License for Figure 33b in the Discussion section

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Feb 14, 2022

This Agreement between Ms. Kaalindi Misra ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5247560594582
License date	Feb 14, 2022
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	TRPs: Truly Remarkable Proteins
Licensed Content Author	Veit Flockerzi, Bernd Nilius
Licensed Content Date	Jan 1, 2014
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Student
Institution name	Ospedale San Raffaele
Expected presentation date	Mar 2022
Portions	Figure 1 pg. 406
Requestor Location	Ms. Kaalindi Misra
	Via Olgettina 60
	Milano, 20132
	Italy
	Attn: Ospedale San Raffaele
Total	**0.00 USD**
Terms and Conditions	

Figure 3: License for Figure 37 in the Discussion section

THE AMERICAN PHYSIOLOGICAL SOCIETY ORDER DETAILS

Feb 14, 2022

Order Number	501710161
Order date	Feb 14, 2022
Licensed Content Publisher	The American Physiological Society
Licensed Content Publication	Physiological Reviews
Licensed Content Title	Neuropathic Pain: From Mechanisms to Treatment
Licensed Content Author	Nanna Brix Finnerup, Rohini Kuner, Troels Staehelin Jensen
Licensed Content Date	Jan 1, 2021
Licensed Content Volume	101
Licensed Content Issue	1
Type of use	Thesis/Dissertation
Requestor type	non-profit academic/educational
Readers being charged a fee for this work	No
Format	print and electronic
Portion	figures/tables/images
Number of figures/tables/images	1
Will you be translating?	no
World Rights	no
Title	Student
Institution name	Ospedale San Raffaele
Expected presentation date	Mar 2022
Portions	Figure 5
Requestor Location	Ms. Kaalindi Misra Via Olgettina 60 Milano, 20132 Italy Attn: Ospedale San Raffaele
Total	**0.00 USD**

Figure 4: License for Figure 4 in the Introduction section

References

Abrams B (2006) Electromyography and Nerve Conduction Velocity. *Pain Manag* 1: 179–191

Agilent Technologies I Discover More with Greater Performance and Speed.

Al-Shekhlee A, Chelimsky TC & Preston DC (2002) Review: small-fiber neuropathy. *Neurologist* 8: 237–253

Alsaloum M, Labau JIR, Sosniak D, Zhao P, Almomani R, Gerrits M, Hoeijmakers JGJ, Lauria G, Faber CG, Waxman SG, *et al* (2021) A novel gain-of-function sodium channel b2 subunit mutation in idiopathic small fiber neuropathy. *J Neurophysiol* 126: 827–839

Amano K, Nishizawa D, Mieda T, Tsujita M, Kitamura A, Hasegawa J, Inada E, Hayashida M & Ikeda K (2016) Opposite Associations Between the rs3845446 Single-Nucleotide Polymorphism of the CACNA1E Gene and Postoperative Pain-Related Phenotypes in Gastrointestinal Surgery Versus Previously Reported Orthognathic Surgery. *J pain* 17: 1126–1134

Amarilyo G, Dori A, Fleitman IS, Nevo Y, Harel L, Agmon-Levin Md N, Kachko L, Gal |, Hoffer Z, Dabby R, *et al* (2020) Small fiber neuropathy associated with autoinflammatory syndromes in children Small-fiber neuropathy associated with autoinflammatory syndromes in children and adolescents.

Amberger JS & Hamosh A (2017) Searching Online Mendelian Inheritance in Man (OMIM): A Knowledgebase of Human Genes and Genetic Phenotypes. *Curr Protoc Bioinforma* 58: 1.2.1-1.2.12

Andersson DA, Gentry C, Moss S & Bevan S (2008) Transient receptor potential A1 is a sensory receptor for multiple products of oxidative stress. *J Neurosci*

Attal N, Fermanian C, Fermanian J, Lanteri-Minet M, Alchaar H & Bouhassira D (2008) Neuropathic pain: Are there distinct subtypes depending on the aetiology or anatomical lesion? *Pain*

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, *et al* (2013) From FastQ data to high confidence variant calls: the GenomeAnalysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 11: 11.10.1

Axelrod FB, Gold-von Simson G & Oddoux C (1993) NTRK1 Congenital Insensitivity to Pain with Anhidrosis.

Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data

Bagdas D, Gurun MS, Flood P, Papke RL & Damaj MI (2018) New Insights on Neuronal Nicotinic Acetylcholine Receptors as Targets for Pain and Inflammation: A Focus on α7 nAChRs. *Curr Neuropharmacol* 16: 415

Baker MD & Nassar MA (2020) Painful and painless mutations of SCN9A and SCN11A voltage-gated sodium channels. *Pflügers Arch - Eur J Physiol 2020 4727* 472: 865–880

Bakkers M, Faber CG, Hoeijmakers JGJ, Lauria G & Merkies ISJ (2014) Small fibers, large impact: quality of life in small-fiber neuropathy. *Muscle Nerve* 49: 329–336

Bandell M, Story GM, Hwang SW, Viswanath V, Eid SR, Petrus MJ, Earley TJ & Patapoutian A (2004) Noxious cold ion channel TRPA1 is activated by pungent compounds and bradykinin. *Neuron*

Bandmann O, Weiss KH & Kaler SG (2015) Wilson's disease and other neurological copper disorders. *Lancet Neurol* 14: 103

Basso L & Altier C (2017) Transient Receptor Potential Channels in neuropathic pain. *Curr Opin Pharmacol* 32: 9–15

Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Bursteinas B, *et al* (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 49: D480–D489

Bienias K (2020) Postural control in patients with hereditary motor and sensory neuropathy. Literature review. *Adv Rehabil* 34: 42–47

Binder A, May D, Baron R, Maier C, Tölle TR, Treede RD, Berthele A, Faltraco F, Flor H, Gierthmühlen J, *et al* (2011a) Transient receptor potential channel polymorphisms are associated with the somatosensory function in neuropathic pain patients. *PLoS One*

Binder A, May D, Baron R, Maier C, Tölle TR, Treede RD, Berthele A, Faltraco F, Flor H, Gierthmühlen J, *et al* (2011b) Transient receptor potential channel polymorphisms are associated with the somatosensory function in neuropathic pain patients. *PLoS One* 6

Birnbaum J, Lalji A, Saed A & Baer AN (2019) Biopsy-Proven Small-Fiber Neuropathy in Primary Sjögren's Syndrome: Neuropathic Pain Characteristics, Autoantibody Findings, and Histopathologic Features. *Arthritis Care Res* 71: 936–948

Bitzi LM, Lehnick D & Wilder-Smith EP (2021) Small fiber neuropathy: Swiss cohort characterization. *Muscle Nerve* 64: 293

Bolger AM, Lohse M & Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120

Brandt MR, Beyer CE & Stahl SM (2012) TRPV1 antagonists and chronic pain: Beyond thermal perception. *Pharmaceuticals* 5: 114–132 doi:10.3390/ph5020114 [PREPRINT]

C. H, J.G.J. H, S. L, M.M. G, R.H.M. TM, G. L, S.D. D-H, J.P.H. D, C.G. F, I.S.J. M, *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*

Calculated consequences

Calvo M, Davies AJ, Hébert HL, Weir GA, Chesler EJ, Finnerup NB, Levitt RC, Smith BH, Neely GG, Costigan M, *et al* (2019) The Genetics of Neuropathic Pain from Model Organisms to Clinical Application. *Neuron* doi:10.1016/j.neuron.2019.09.018 [PREPRINT]

Caraceni A & Shkodra M (2019) Cancer Pain Assessment and Classification. *Cancers 2019, Vol 11, Page 510* 11: 510

De Caro C, Russo R, Avagliano C, Cristiano C, Calignano A, Aramini A, Bianchini G, Allegretti M & Brandolini L (2018) Antinociceptive effect of two novel transient receptor potential melastatin 8 antagonists in acute and chronic pain models in rat. *Br J Pharmacol* 175: 1691–1706

Carson AR, Smith EN, Matsui H, Brækkan SK, Jepsen K, Hansen JB & Frazer KA (2014) Effective filtering strategies to improve data quality from population-based whole exome sequencing studies. *BMC Bioinformatics* 15: 125

Castelli G, Desai KM & Cantone RE (2020) Peripheral Neuropathy: Evaluation and Differential Diagnosis. *Am Fam Physician* 102: 732–739

Cazzato D & Lauria G (2017) Small fibre neuropathy. *Curr Opin Neurol* 30: 490–499

Chandra SR, Karru VR, Mudabbir MAM, Ramakrishnan S & Mahadevan A (2018) Immune-mediated Neuropathies Our Experience over 3 Years. *J Neurosci Rural Pract* 9: 30

Chapman B (2014) Validated whole genome structural variation detection using multiple callers | Blue Collar Bioinformatics.

Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, Cox AJ, Kruglyak S & Saunders CT (2016) Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* 32: 1220–1222

Chowdhury S & Sarkar RR (2015) Comparison of human cell signaling pathway databases—evolution, drawbacks and challenges. *Database* 2015: 126

Chung MK, Asgar J, Lee J, Shim MS, Dumler C & Ro JY (2015) THE ROLE OF TRPM2 IN HYDROGEN PEROXIDE-INDUCED EXPRESSION OF INFLAMMATORY CYTOKINE AND CHEMOKINE IN RAT TRIGEMINAL GANGLIA. *Neuroscience* 297: 160

Chung T, Prasad K & Lloyd TE (2014) Peripheral neuropathy. Clinical and electrophysiological considerations. *Neuroimaging Clin N Am* 24: 49–65

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X & Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* 6: 80

Cintra VP, Dohrn MF, Tomaselli PJ, Figueiredo FB, Marques SE, Camargos ST, Barbosa LSM, P. Rebelo A, Abreu L, Danzi M, *et al* (2021) Rare mutations in ATL3, SPTLC2 and SCN9A explaining hereditary sensory neuropathy and congenital insensitivity to pain in a Brazilian cohort. *J Neurol Sci* 427: 117498

Clarkson YL, Perkins EM, Cairncross CJ, Lyndon AR, Skehel PA & Jackson M (2014) β-III spectrin underpins ankyrin R function in Purkinje cell dendritic trees: protein complex critical for sodium channel activity is impaired by SCA5-associated mutations. *Hum Mol Genet* 23: 3875–3882

Cohen MJ, Jangro WC & Neff D (2018) Pathophysiology of Pain. *Challenging Neuropathic Pain Syndr Eval Evidence-Based Treat*: 1–5

Collins RL, Brand H, Karczewski KJ, Zhao X, Alföldi J, Khera A V., Francioli LC, Gauthier LD, Wang H, Watts NA, *et al* (2019) An open resource of structural variation for medical and population genetics. *bioRxiv*

Cox JJ, Sheynin J, Shorer Z, Reimann F, Nicholas AK, Zubovic L, Baralle M, Wraige E, Manor E, Levy J, *et al* (2010) Congenital Insensitivity to Pain: Novel SCN9A Missense and In-Frame Deletion Mutations. *Hum Mutat* 31: 1670

Cregg R, Cox JJ, Bennett DLH, Wood JN, Werdehausen R & Cregg R (2014) Mexiletine as a treatment for primary erythromelalgia: normalization of biophysical properties of mutant L858F NaV1.7 sodium channels. *Wiley Online Libr* 171: 4455

Cregg R, Laguda B, Werdehausen R, Cox JJ, Linley JE, Ramirez JD, Bodi I, Markiewicz M, Howell KJ, Chen YC, *et al* (2013) Novel mutations mapping to the fourth sodium channel domain of nav1.7 result in variable clinical manifestations of primary erythromelalgia. *NeuroMolecular Med* 15: 265–278

Cruccu G & Truini A (2017) Neuropathic Pain: The Scope of the Problem. *Pain Ther*

Cummins TR, Dib-Hajj SD & Waxman SG (2004) Electrophysiological Properties of Mutant Nav1.7 Sodium Channels in a Painful Inherited Neuropathy. *J Neurosci* 24: 8232–8236

Daniels RL, Takashima Y & McKemy DD (2009) Activity of the neuronal cold sensor TRPM8 is regulated by phospholipase C via the phospholipid phosphoinositol 4,5-bisphosphate. *J Biol Chem* 284: 1570–1582

Daousi C, MacFarlane IA, Woodward A, Nurmikko TJ, Bundred PE & Benbow SJ (2004) Chronic painful peripheral neuropathy in an urban community: a controlled comparison of people with and without diabetes. *Diabet Med* 21: 976–982

Davis JS, Hassanzadeh S, Winitsky S, Lin H, Satorius C, Vemuri R, Aletras AH, Wen H & Epstein ND (2001) The overall pattern of cardiac contraction depends on a spatial gradient of myosin regulatory light chain phosphorylation. *Cell* 107: 631–641

Davis RE, Swiderski RE, Rahmouni K, Nishimura DY, Mullins RF, Agassandian K, Philp AR, Searby CC, Andrews MP, Thompson S, *et al* (2007) A knockin mouse model of the Bardet–Biedl syndrome 1 M390R mutation has cilia defects, ventriculomegaly, retinopathy, and obesity. *Proc Natl Acad Sci U S A* 104: 19422

Detta N, Frisso G & Salvatore F (2015) The multi-faceted aspects of the complex cardiac Nav1.5 protein in membrane function and pathophysiology. *Biochim Biophys Acta - Proteins Proteomics* 1854: 1502–1509

Devigili G, Eleopra R, Pierro T, Lombardi R, Rinaldo S, Lettieri C, Faber CG, Merkies ISJ, Waxman SG & Lauria G (2014a) Paroxysmal itch caused by gain-of-function Nav1.7 mutation. *Pain*

Devigili G, Eleopra R, Pierro T, Lombardi R, Rinaldo S, Lettieri C, Faber CG, Merkies ISJ, Waxman SG & Lauria G (2014b) Paroxysmal itch caused by gain-of-function Nav1.7 mutation. *Pain* 155: 1702–1707

Devigili G, De Filippo M, Ciana G, Dardis A, Lettieri C, Rinaldo S, Macor D, Moro A, Eleopra R & Bembi B (2017) Chronic pain in Gaucher disease: skeletal or neuropathic origin? *Orphanet J Rare Dis* 12

Devigili G, Tugnoli V, Penza P, Camozzi F, Lombardi R, Melli G, Broglio L, Granieri E & Lauria G (2008) The diagnostic criteria for small fibre neuropathy: From symptoms to neuropathology. *Brain*

Dyck PJ (2010) The Causes, Classification, and Treatment of Peripheral Neuropathy. *http://dx.doi.org/101056/NEJM198207293070504* 307: 283–286

Eijkelkamp N, Linley JE, Baker MD, Minett MS, Cregg R, Werdehausen R, Rugiero F & Wood JN (2012) Neurological perspectives on voltage-gated sodium channels. *Brain* 135: 2585–2612

Eijkenboom I, Sopacua M, Hoeijmakers JGJ, De Greef BTA, Lindsey P, Almomani R, Marchi M, Vanoevelen J, Smeets HJM, Waxman SG, *et al* (2019) Yield of peripheral sodium channels gene screening in pure small fibre neuropathy. *J Neurol Neurosurg Psychiatry* 90: 342–352

Eilbeck K, Lewis SE, Mungall CJ, Yandell M, Stein L, Durbin R & Ashburner M (2005) The Sequence Ontology: a tool for the unification of genome annotations. *Genome Biol* 6: 1–12

El-Abassi RN, Soliman M, Levy MH & England JD (2022) Treatment and Management of Autoimmune Neuropathies. *Neuromuscul Disord*: 312–344

England JD, Gronseth GS, Franklin G, Carter GT, Kinsella LJ, Cohen JA, Asbury AK, Szigeti K, Lupski JR, Latov N, *et al* (2009) Practice Parameter: The Evaluation of Distal Symmetric Polyneuropathy: The Role of Autonomic Testing, Nerve Biopsy, and Skin Biopsy (An Evidence-Based Review). *PM&R* 1: 14–22

Estacion M, Vohra BPS, Liu S, Hoeijmakers J, Faber CG, Merkies ISJ, Lauria G, Black JA & Waxman SG (2015) Ca2+ toxicity due to reverse Na+/Ca2+ exchange contributes to degeneration of neurites of DRG neurons induced by a neuropathy-associated Nav1.7 mutation. *J Neurophysiol* 114: 1554–1564

Ewing B & Green P (1998) Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Res* 8: 186–194

Ewing B, Hillier LD, Wendl MC & Green P (1998) Base-Calling of Automated Sequencer Traces UsingPhred.   I. Accuracy Assessment. *Genome Res* 8: 175–185

Faber CG, Hoeijmakers JGJ, Ahn HS, Cheng X, Han C, Choi JS, Estacion M, Lauria G, Vanhoutte EK, Gerrits MM, *et al* (2012a) Gain of function Na V1.7 mutations in idiopathic small fiber neuropathy. *Ann Neurol*

Faber CG, Lauria G, Merkies ISJ, Cheng X, Han C, Ahn H-S, Persson A-K, Hoeijmakers JGJ, Gerrits MM, Pierro T, *et al* (2012b) Gain-of-function Nav1.8 mutations in painful neuropathy. *Proc Natl Acad Sci*

Faber CG, Lauria G, Merkies ISJ, Cheng X, Han C, Ahn HS, Persson AK, Hoeijmakers JGJ, Gerrits MM, Pierro T, *et al* (2012c) Gain-of-function Nav1.8 mutations in painful neuropathy. *Proc Natl Acad Sci U S A* 109: 19444–19449

Fadista J, Oskolkov N, Hansson O & Groop L (2017) LoFtool: a gene intolerance score based on loss-of-function variants in 60 706 individuals. *Bioinformatics* 33: 471–474

Farhat NM & Yezback KL (2016) Treatment of Diabetic Peripheral Neuropathy. *J Nurse Pract* 12: 660–666

Farooqi AA, Javeed MK, Javed Z, Riaz AM, Mukhtar S, Minhaj S, Abbas S & Bhatti S (2011) TRPM channels: same ballpark, different players, and different rules in immunogenetics. *Immunogenetics* 63: 773–787

Farschtschi SC, Mainka T, Glatzel M, Hannekum AL, Hauck M, Gelderblom M, Hagel C, Friedrich RE, Schuhmann MU, Schulz A, *et al* (2020) C-Fiber Loss as a Possible Cause of Neuropathic Pain in Schwannomatosis. *Int J Mol Sci* 21: 3569

Fertleman CR, Baker MD, Parker KA, Moffatt S, Elmslie F V., Abrahamsen B, Ostman J, Klugbauer N, Wood JN, Gardiner RM, *et al* (2006) SCN9A mutations in paroxysmal extreme pain disorder: allelic variants underlie distinct channel defects and phenotypes. *Neuron* 52: 767–774

Finnerup NB, Kuner R & Jensen TS (2021) Neuropathic pain: From mechanisms to treatment. *Physiol Rev* 101: 259–301

Foo JN, Liu JJ & Tan EK (2012) Whole-genome and whole-exome sequencing in neurological diseases. *Nat Rev Neurol 2012 89* 8: 508–517

Frésard L & Montgomery SB (2018) Diagnosing rare diseases after the exome. *Cold Spring Harb Mol Case Stud* doi:10.1101/mcs.a003392 [PREPRINT]

Gees M, Colsoul B & Nilius B (2010) The role of transient receptor potential cation channels in Ca2+ signaling. *Cold Spring Harb Perspect Biol*

Genzen JR, Van Cleve W & Mcgehee DS (2001) Dorsal root ganglion neurons express multiple nicotinic acetylcholine receptor subtypes. *J Neurophysiol* 86: 1773–1782

Glatte P, Buchmann SJ, Hijazi MM, Illigens BMW & Siepmann T (2019) Architecture of the Cutaneous Autonomic Nervous System. *Front Neurol* 10: 970

Gold MS & Gebhart GF (2010) Nociceptor sensitization in pain pathogenesis. *Nat Med 2010 1611* 16: 1248–1257

González A, Ugarte G, Restrepo C, Herrera G, Piña R, Gómez-Sánchez JA, Pertusa M, Orio P & Madrid R (2017) Role of the Excitability Brake Potassium Current IKD in Cold Allodynia Induced by Chronic Peripheral Nerve Injury. *J Neurosci* 37: 3109

de Greef BTA, Hoeijmakers JGJ, Gorissen-Brouwers CML, Geerts M, Faber CG & Merkies ISJ

(2018) Associated conditions in small fiber neuropathy – a large cohort study and review of the literature. *Eur J Neurol* 25: 348

Gross F & Üçeyler N (2020) Mechanisms of small nerve fiber pathology. *Neurosci Lett* 737: 135316

Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U & Shaw JE (2014) Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract* 103: 137–149

Gwathmey KG, Burns TM, Collins MP & Dyck PJB (2014) Vasculitic neuropathies. *Lancet Neurol* 13: 67–82

Han C, Hoeijmakers JGJ, Ahn HS, Zhao P, Shah P, Lauria G, Gerrits MM, Te Morsche RHM, Dib-Hajj SD, Drenth JPH, *et al* (2012a) Nav1.7-related small fiber neuropathy: Impaired slow-inactivation and DRG neuron hyperexcitability. *Neurology* 78: 1635–1643

Han C, Hoeijmakers JGJ, Liu S, Gerrits MM, Te Morsche RHM, Lauria G, Dib-Hajj SD, Drenth JPH, Faber CG, Merkies ISJ, *et al* (2012b) Functional profiles of SCN9A variants in dorsal root ganglion neurons and superior cervical ganglion neurons correlate with autonomic symptoms in small fibre neuropathy. *Brain* 135: 2613–2628

Haraguchi K, Kawamoto A, Isami K, Maeda S, Kusano A, Asakura K, Shirakawa H, Mori Y, Nakagawa T & Kaneko S (2012) TRPM2 Contributes to Inflammatory and Neuropathic Pain through the Aggravation of Pronociceptive Inflammatory Responses in Mice. *J Neurosci* 32: 3931

Harel T, Yoon WH, Garone C, Gu S, Coban-Akdemir Z, Eldomery MK, Posey JE, Jhangiani SN, Rosenfeld JA, Cho MT, *et al* (2016) Recurrent De Novo and Biallelic Variation of ATAD3A, Encoding a Mitochondrial Membrane Protein, Results in Distinct Neurological Syndromes. *Am J Hum Genet* 99: 831–845

Hasan NM, Gupta A, Polishchuks E, Yu CH, Polishchuks R, Dmitriev OY & Lutsenko S (2012) Molecular Events Initiating Exit of a Copper-transporting ATPase ATP7B from the Trans-Golgi Network. *J Biol Chem* 287: 36041

Van Hecke O, Austin SK, Khan RA, Smith BH & Torrance N (2014) Neuropathic pain in the general population: A systematic review of epidemiological studies. *Pain* doi:10.1016/j.pain.2013.11.013 [PREPRINT]

Hedayat KM & Lapraz J-C (2019) The autonomic nervous system. *The Theory of Endobiogeny*: 31–43

Heida A, Van Der Does LJME, Ragab AAY & De Groot NMS (2019) A Rare Case of the Digenic Inheritance of Long QT Syndrome Type 2 and Type 6. *Case Rep Med* 2019

Helbig KL, Lauerer RJ, Bahr JC, Souza IA, Myers CT, Uysal B, Schwarz N, Gandini MA, Huang S, Keren B, *et al* (2018) De Novo Pathogenic Variants in CACNA1E Cause Developmental and Epileptic Encephalopathy with Contractures, Macrocephaly, and Dyskinesias. *Am J Hum Genet* 103: 666

Hereditary Motor Sensory Neuropathies: Charcot-Marie-Tooth

Hoitsma E, Marziniak M, Faber CG, Reulen JPH, Sommer C, De Baets M & Drent M (2002) Small fibre neuropathy in sarcoidosis. *Lancet* 359: 2085–2086

Hone AJ & McIntosh JM (2018) Nicotinic Acetylcholine Receptors in Neuropathic and Inflammatory Pain. *FEBS Lett* 592: 1045

Hone AJ, Meyer EL, McIntyre M & McIntosh JM (2012) Nicotinic acetylcholine receptors in dorsal root ganglion neurons include the α6β4* subtype. *FASEB J* 26: 917–926

Hovaguimian A & Gibbons CH (2011) Diagnosis and treatment of pain in small-fiber neuropathy. *Curr Pain Headache Rep* doi:10.1007/s11916-011-0181-7 [PREPRINT]

Huang J, Han C, Estacion M, Vasylyev D, Hoeijmakers JGJ, Gerrits MM, Tyrrell L, Lauria G, Faber CG, Dib-Hajj SD, *et al* (2014) Gain-of-function mutations in sodium channel NaV1.9 in painful neuropathy. *Brain*

Huang J, Vanoye CG, Cutts A, Goldberg YP, Dib-Hajj SD, Cohen CJ, Waxman SG & George AL (2017) Sodium channel NaV1.9 mutations associated with insensitivity to pain dampen neuronal excitability. *J Clin Invest* 127: 2805

Hughes RAC (2002) Peripheral neuropathy. *BMJ* 324: 466–469

Huster D, Khne A, Bhattacharjee A, Raines L, Jantsch V, Noe J, Schirrmeister W, Sommerer I, Sabri O, Berr F, *et al* (2012) Diverse Functional Properties of Wilson Disease ATP7B Variants. *Gastroenterology* 142: 947

Illumina Technologies I Applications and Methods for HiSeq 3000/HiSeq 4000 Systems.

Incecik F, Bisgin A & Yılmaz M (2018) MEDNIK syndrome with a frame shift causing mutation in AP1S1 gene and literature review of the clinical features. *Metab Brain Dis* 33: 2065–2068

Jain KK (2004) Modulators of nicotinic acetylcholine receptors as analgesics. *Curr Opin Investig Drugs* 5: 76–81

Jain S, Farías GG & Bonifacino JS (2015a) Polarized sorting of the copper transporter ATP7B

in neurons mediated by recognition of a dileucine signal by AP-1. *Mol Biol Cell* 26: 218–228

Jain S, Farías GG & Bonifacino JS (2015b) Polarized sorting of the copper transporter ATP7B in neurons mediated by recognition of a dileucine signal by AP-1. *Mol Biol Cell* 26: 218–228

Jaquemar D, Schenker T & Trueb B (1999) An ankyrin-like protein with transmembrane domains is specifically lost after oncogenic transformation of human fibroblasts. *J Biol Chem*

Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, *et al* (2020) The reactome pathway knowledgebase. *Nucleic Acids Res* 48: D498–D503

Jiang YQ, Sun Q, Tu HY & Wan Y (2008) Characteristics of HCN channels and their participation in neuropathic pain. *Neurochem Res* 33: 1979–1989 doi:10.1007/s11064-008-9717-6 [PREPRINT]

Jobsis GJ, Keizers H, Vreijling JP, De Visser M, Speer MC, Wolterman RA, Baas F & Bolhuis PA (1996) Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. *Nat Genet* 14: 113–115

Johnson SA, Shouman K, Shelly S, Sandroni P, Berini SE, Dyck PJB, Hoffman EM, Mandrekar J, Niu Z, Lamb CJ, *et al* (2021) Small Fiber Neuropathy Incidence, Prevalence, Longitudinal Impairments, and Disability. *Neurology*: 10.1212/WNL.0000000000012894

Julius D (2013) TRP channels and pain. *Annu Rev Cell Dev Biol* 29: 355–384 doi:10.1146/annurev-cellbio-101011-155833 [PREPRINT]

Jung KH, Ahn TB & Jeon BS (2005) Wilson Disease With an Initial Manifestation of Polyneuropathy. *Arch Neurol* 62: 1628–1631

Kanehisa M, Goto S, Sato Y, Furumichi M & Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40: D109

Kaneko Y & Szallasi A (2014) Transient receptor potential (TRP) channels: A clinical perspective. *Br J Pharmacol* 171: 2474–2507 doi:10.1111/bph.12414 [PREPRINT]

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, *et al* (2019) Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, *et al* (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nat 2020 5817809* 581: 434–443

Kardon AP, Polgár E, Hachisuka J, Snyder LM, Cameron D, Savage S, Cai X, Karnup S, Fan CR, Hemenway GM, *et al* (2014) Dynorphin Acts as a Neuromodulator to Inhibit Itch in the Dorsal Horn of the Spinal Cord. *Neuron* 82: 573–586

Kenney SM & Cox DW (2007) Sequence variation database for the Wilson disease copper transporter, ATP7B. *Hum Mutat* 28: 1171–1177

Kharkovets T, Hardelin JP, Safieddine S, Schweizer M, El-Amraoui A, Petit C & Jentsch TJ (2000) From the Cover: KCNQ4, a K+ channel mutated in a form of dominant deafness, is expressed in the inner ear and the central auditory pathway. *Proc Natl Acad Sci U S A* 97: 4333

Khatri P, Sirota M & Butte AJ (2012) Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol* 8

Köhler S, Gargano M, Matentzoglu N, Carmody LC, Lewis-Smith D, Vasilevsky NA, Danis D, Balagura G, Baynam G, Brower AM, *et al* (2021) The Human Phenotype Ontology in 2021. *Nucleic Acids Res* 49: D1207–D1217

Kono T, Kaneko A, Omiya Y, Suzuki T, Hira Y, Chisato N, Watanabe T & Kohgo Y (2010) M1852 Activation of Transient Receptor Potential A1 Expressed in Intestinal Epithelial Cell Increases Intestinal Blood Flow via Release of Adrenomedullin. *Gastroenterology* 138

Kremeyer B, Lopera F, Cox JJ, Momin A, Rugiero F, Marsh S, Woods CG, Jones NG, Paterson KJ, Fricker FR, *et al* (2010) A Gain-of-Function Mutation in TRPA1 Causes Familial Episodic Pain Syndrome. *Neuron*

Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP, Quinlan AR, Nickerson DA & Eichler EE (2012) Copy number variation detection and genotyping from exome sequence data. *Genome Res* 22: 1525–1532

Kruse M, Schulze-Bahr E, Corfield V, Beckmann A, Stallmeyer B, Kurtbay G, Ohmert I, Schulze-Bahr E, Brink P & Pongs O (2009) Impaired endocytosis of the ion channel TRPM4 is associated with human progressive familial heart block type I. *J Clin Invest* 119: 2737–2744

Kumar R (2020) Endogenous and Exogenous Vanilloids Evoke Disparate TRPV1 Activation to Produce Distinct Neuronal Responses. *Front Pharmacol* 11: 903

Kumari S, Bonnet MC, Ulvmar MH, Wolk K, Karagianni N, Witte E, Uthoff-Hachenberg C, Renauld JC, Kollias G, Toftgard R, *et al* (2013) Tumor necrosis factor receptor signaling in keratinocytes triggers interleukin-24-dependent psoriasis-like skin inflammation in mice. *Immunity* 39: 899–911

Laedermann CJ, Cachemaille M, Kirschmann G, Pertin M, Gosselin RD, Chang I, Albesa M, Towne C, Schneider BL, Kellenberger S, *et al* (2013) Dysregulation of voltage-gated sodium channels by ubiquitin ligase NEDD4-2 in neuropathic pain. *J Clin Invest* 123: 3002–3013

Lafrenière RG, Cader MZ, Poulin JF, Andres-Enguix I, Simoneau M, Gupta N, Boisvert K, Lafrenière F, McLaughlan S, Dubé MP, *et al* (2010) A dominant-negative mutation in the TRESK potassium channel is linked to familial migraine with aura. *Nat Med* 16: 1157–1160

Lampert A, Eberhardt M & Waxman SG (2014) Altered Sodium Channel Gating as Molecular Basis for Pain: Contribution of Activation, Inactivation, and Resurgent Currents. *Handb Exp Pharmacol* 221: 91–110

Lang PM, Burgstahler R, Sippel W, Irnich D, Schlotter-Weigel B & Grafe P (2003) Characterization of neuronal nicotinic acetylcholine receptors in the membrane of unmyelinated human C-fiber axons by in vitro studies. *J Neurophysiol* 90: 3295–3303

Lapointe TK & Altier C (2011) The role of TRPA1 in visceral inflammation and pain. *Channels* 5: 525

Latorre R, Brauchi S, Madrid R & Orio P (2011) A cool channel in cold transduction. *Physiology* 26: 273–285

Lauria G, Cornblath DR, Johansson O, McArthur JC, Mellgren SI, Nolano M, Rosenberg N & Sommer C (2005) EFNS guidelines on the use of skin biopsy in the diagnosis of peripheral neuropathy. *Eur J Neurol* 12: 747–758

Lauria G & Devigili G (2007) Skin biopsy as a diagnostic tool in peripheral neuropathy. *Nat Clin Pract Neurol* doi:10.1038/ncpneuro0630 [PREPRINT]

Lehmann HC, Wunderlich G, Fink GR & Sommer C (2020) Diagnosis of peripheral neuropathy. *Neurol Res Pract* 2: 1–7

Leipold E, Liebmann L, Korenke GC, Heinrich T, Gießelmann S, Baets J, Ebbinghaus M, Goral RO, Stödberg T, Hennings JC, *et al* (2013) A de novo gain-of-function mutation in SCN11A causes loss of pain perception. *Nat Genet* 45: 1399–1407

Lelieveld SH, Spielmann M, Mundlos S, Veltman JA & Gilissen C (2015) Comparison of

Exome and Genome Sequencing Technologies for the Complete Capture of Protein-Coding Regions. *Hum Mutat* 36: 815–822

Lenders M, Duning T, Schelleckes M, Schmitz B, Stander S, Rolfs A, Brand SM & Brand E (2013) Multifocal white matter lesions associated with the D313Y mutation of the α-galactosidase A gene. *PLoS One* 8

De Lera Ruiz M & Kraus RL (2015) Voltage-Gated Sodium Channels: Structure, Function, Pharmacology, and Clinical Indications. *J Med Chem* 58: 7093–7118

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G & Durbin R (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079

Li QS, Cheng P, Favis R, Wickenden A, Romano G & Wang H (2015) SCN9A Variants May be Implicated in Neuropathic Pain Associated With Diabetic Peripheral Neuropathy and Pain Severity. *Clin J Pain* 31: 976

Li Y, Tatsui CE, Rhines LD, North RY, Harrison DS, Cassidy RM, Johansson CA, Kosturakis AK, Edwards DD, Zhang H, *et al* (2017) Dorsal root ganglion neurons become hyperexcitable and increase expression of voltage-gated T-type calcium channels (Cav3.2) in paclitaxel-induced peripheral neuropathy. *Pain* 158: 417

Liao Y, Wang J, Jaehnig EJ, Shi Z & Zhang B (2019) WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res* 47: W199–W205

Liu W, Deng J, Wang G, Zhang C, Luo X, Yan D, Su Q & Liu J (2014) KCNE2 modulates cardiac L-type Ca(2+) channel. *J Mol Cell Cardiol* 72: 208–218

Löscher W, Gillard M, Sands ZA, Kaminski RM & Klitgaard H (2016) Synaptic Vesicle Glycoprotein 2A Ligands in the Treatment of Epilepsy and Beyond. *CNS Drugs* 30: 1055

Machado AAC, Deguti MM, Genschel J, Cançado ELR, Bochow B, Schmidt H & Barbosa ER (2008) Neurological manifestations and ATP7B mutations in Wilson's disease. *Parkinsonism Relat Disord* 14: 246–249

Marchi M, Provitera V, Nolano M, Romano M, Maccora S, D'Amato I, Salvi E, Gerrits M, Santoro L & Lauria G (2018) A novel SCN9A splicing mutation in a compound heterozygous girl with congenital insensitivity to pain, hyposmia and hypogeusia. *J Peripher Nerv Syst* 23: 202

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing

reads. *EMBnet.journal* 17: 10–12

Martinelli-Boneschi F, Colombi M, Castori M, Devigili G, Eleopra R, Malik RA, Ritelli M, Zoppi N, Dordoni C, Sorosina M, *et al* (2017) COL6A5 variants in familial neuropathic chronic itch. *Brain*: aww343

Matsumoto K, Takagi K, Kato A, Ishibashi T, Mori Y, Tashima K, Mitsumoto A, Kato S & Horie S (2016) Role of transient receptor potential melastatin 2 (TRPM2) channels in visceral nociception and hypersensitivity. *Exp Neurol* 285: 41–50

McDermott LA, Weir GA, Themistocleous AC, Segerdahl AR, Blesneac I, Baskozos G, Clark AJ, Millar V, Peck LJ, Ebner D, *et al* (2019) Defining the Functional Role of NaV1.7 in Human Nociception. *Neuron* 101: 905

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P & Cunningham F (2016) The Ensembl Variant Effect Predictor. *Genome Biol* 17

Meents JE, Ciotu CI & Fischer MJM (2019) Trpa1: A molecular view. *J Neurophysiol* doi:10.1152/jn.00524.2018 [PREPRINT]

Meents JE, Fischer MJM & McNaughton PA (2017) Sensitization of TRPA1 by Protein Kinase A. *PLoS One* 12

Meloto CB, Benavides R, Lichtenwalter RN, Wen X, Tugarinov N, Zorina-Lichtenwalter K, Chabot-Doré AJ, Piltonen MH, Cattaneo S, Verma V, *et al* (2018) Human pain genetics database: A resource dedicated to human pain genetics research. *Pain* 159: 749–763

Messeguer A, Planells-Cases R & Ferrer-Montiel A (2005) Physiology and Pharmacology of the Vanilloid Receptor. *Curr Neuropharmacol* 4: 1–15

Mi H, Ebert D, Muruganujan A, Mills C, Albou LP, Mushayamaha T & Thomas PD (2021) PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res* 49: D394

Minde DP, Anvarian Z, Rüdiger SGD & Maurice MM (2011) Messing up disorder: how do missense mutations in the tumor suppressor protein APC lead to cancer? *Mol Cancer* 10: 101

Montpetit A, Côté S, Brustein E, Drouin CA, Lapointe L, Boudreau M, Meloche C, Drouin R, Hudson TJ, Drapeau P, *et al* (2008) Disruption of AP1S1, causing a novel neurocutaneous syndrome, perturbs development of the skin and spinal cord. *PLoS Genet* 4

Moore C, Gupta R, Jordt SE, Chen Y & Liedtke WB (2018) Regulation of Pain and Itch by TRP Channels. *Neurosci Bull* 34: 120

De Moraes Vieira ÉB, Garcia JBS, Da Silva AAM, Mualem Araújo RLT & Jansen RCS (2012) Prevalence, Characteristics, and Factors Associated With Chronic Pain With and Without Neuropathic Characteristics in São Luís, Brazil. *J Pain Symptom Manage* 44: 239–251

Morgan K, Sadofsky LR & Morice AH (2015) Genetic variants affecting human TRPA1 or TRPM8 structure can be classified in vitro as 'well expressed', 'poorly expressed' or 'salvageable'. *Biosci Rep* 35: 255

Motley W, Chaudry V & Lloyd TE (2022) Treatment and Management of Hereditary Neuropathies. *Neuromuscul Disord*: 278–311

Mykytyn K, Nishimura DY, Searby CC, Shastri M, Yen H jan, Beck JS, Braun T, Streb LM, Cornier AS, Cox GF, *et al* (2002) Identification of the gene (BBS1) most commonly involved in Bardet-Biedl syndrome, a complex human obesity syndrome. *Nat Genet* 31: 435–438

Nahorski MS, Al-Gazali L, Hertecant J, Owen DJ, Borner GHH, Chen YC, Benn CL, Carvalho OP, Shaikh SS, Phelan A, *et al* (2015) A novel disorder reveals clathrin heavy chain-22 is essential for human pain and touch development. *Brain* 138: 2147–2160

Nakajima R, Takao K, Huang SM, Takano J, Iwata N, Miyakawa T & Saido TC (2008) Comprehensive behavioral phenotyping of calpastatin-knockout mice. *Mol Brain* 1: 7

NGS Workflow Steps | Illumina sequencing workflow

Nielsen C, Knudsen G & Steingrímsdóttir Ó (2012) Twin studies of pain. *Clin Genet* 82: 331–340

Nielsen JB, Bentzen BH, Olesen MS, David JP, Olesen SP, Haunsø S, Svendsen JH & Schmitt N (2014) Gain-of-function mutations in potassium channel subunit KCNE2 associated with early-onset lone atrial fibrillation. *Biomark Med* 8: 557–570

Nilius B & Flockerzi V (2014a) Handbook of Experimental Pharmacology 222 Nilius B & Flockerzi V (eds) Berlin, Heidelberg: Springer

Nilius B & Flockerzi V (2014b) What do we really know and what do we need to know: Some controversies, perspectives, and surprises. *Handb Exp Pharmacol*

Noda Y, Mamiya T, Nabeshima T, Nishi M, Higashioka M & Takeshima H (1998) Loss of antinociception induced by naloxone benzoylhydrazone in nociceptin receptor-knockout mice. *J Biol Chem* 273: 18047–18051

Nolano M, Tozza S, Caporaso G & Provitera V (2020) Contribution of Skin Biopsy in Peripheral Neuropathies. *Brain Sci 2020, Vol 10, Page 989* 10: 989

Oaklander AL & Nolano M (2019) Scientific Advances in and Clinical Approaches to Small-Fiber Polyneuropathy: A Review. *JAMA Neurol* 76: 1240–1251

Okuda-Ashitaka E, Minami T, Tachibana S, Yoshihara Y, Nishluchi Y, Kimura T & Ito S (1998) Nocistatin, a peptide that blocks nociceptin action in pain transmission. *Nature* 392: 286–289

Pabinger S, Dander A, Fischer M, Snajder R, Sperk M, Efremova M, Krabichler B, Speicher MR, Zschocke J & Trajanoski Z (2014) A survey of tools for variant analysis of next-generation genome sequencing data. *Brief Bioinform* 15: 256–278

Pain Gene Resource | Pain Research Forum

Pandey P & Balekar N (2018) Target-specific delivery: An insight. *Drug Target Stimuli Sensitive Drug Deliv Syst*: 117–154

Paulsen CE, Armache JP, Gao Y, Cheng Y & Julius D (2015) Structure of the TRPA1 ion channel suggests regulatory mechanisms. *Nature*

Pavinato L, Nematian-Ardestani E, Zonta A, De Rubeis S, Buxbaum J, Mancini C, Bruselles A, Tartaglia M, Pessia M, Tucker SJ, *et al* (2021) Kcnk18 biallelic variants associated with intellectual disability and neurodevelopmental disorders alter tresk channel activity. *Int J Mol Sci* 22: 6064

Perraud AL, Fleig A, Dunn CA, Bagley LA, Launay P, Schmitz C, Stokes AJ, Zhu Q, Bessman MJ, Penner R, *et al* (2001) ADP-ribose gating of the calcium-permeable LTRPC2 channel revealed by Nudix motif homology. *Nat 2001 4116837* 411: 595–599

Peters MJH, Bakkers M, Merkies ISJ, Hoeijmakers JGJ, Van Raak EPM & Faber CG (2013) Incidence and prevalence of small-fiber neuropathy: A survey in the Netherlands. *Neurology*

Petrovski S, Todd JL, Durheim MT, Wang Q, Chien JW, Kelly FL, Frankel C, Mebane CM, Ren Z, Bridgers J, *et al* (2017) An exome sequencing study to assess the role of rare genetic variation in pulmonary fibrosis. *Am J Respir Crit Care Med* 196: 82–93

Pfundt R, Del Rosario M, Vissers LELM, Kwint MP, Janssen IM, De Leeuw N, Yntema HG, Nelen MR, Lugtenberg D, Kamsteeg EJ, *et al* (2016) Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. *Genet Med 2017 196* 19: 667–675

Picciotto MR, Higley MJ & Mineur YS (2012) Acetylcholine as a Neuromodulator: Cholinergic Signaling Shapes Nervous System Function and Behavior. *Neuron* 76: 116–129

Pingle SC, Matta JA & Ahern GP (2007) Capsaicin receptor: TRPV1 a promiscuous TRP

channel. *Handb Exp Pharmacol* 179: 155–171 doi:10.1007/978-3-540-34891-7_9 [PREPRINT]

Politei JM, Durand C & Schenone AB (2019) Small Fiber Neuropathy in Fabry Disease: a Review of Pathophysiology and Treatment. *J Inborn Errors Metab Screen* 4

Polli A, Godderis L, Ghosh M, Ickmans K & Nijs J (2020) Epigenetic and miRNA Expression Changes in People with Pain: A Systematic Review. *J Pain* 21: 763–780

Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, *et al* (2018) Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*: 201178

Probst V, Hoorntje TM, Hulsbeek M, Wilde AAM, Escande D, Mannens MMAM & Le Marec H (1999) Cardiac conduction defects associate with mutations in SCN5A. *Nat Genet* 23: 20–21

Quasthoff S & Hartung HP (2002) Chemotherapy-induced peripheral neuropathy. *J Neurol 2002 2491* 249: 9–17

Ramirez JD, Habib AM, Cox JJ, Themistocleous AC, McMahon SB, Wood JN & Bennett DLH (2014) Null mutation in SCN9A in which noxious stimuli can be detected in the absence of pain. *Neurology* 83: 1577

Raouf R, Quick K & Wood JN (2010) Pain as a channelopathy Find the latest version: Pain as a channelopathy. *J Clin Invest* 120: 3745

Rebelo AP, Abrams AJ, Cottenie E, Horga A, Gonzalez M, Bis DM, Sanchez-Mejias A, Pinto M, Buglo E, Markel K, *et al* (2016) Cryptic Amyloidogenic Elements in the 3' UTRs of Neurofilament Genes Trigger Axonal Neuropathy. *Am J Hum Genet* 98: 597–614

Review: How Long-Read Sequencing Is Revealing Unseen Genomic Variation - PacBio

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, *et al* (2015) Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*

Rosenberg LE & Rosenberg DD (2012) Transmission of Genes. *Hum Genes Genomes*: 51–73

Rotthier (A, Baets J, Timmerman V, Janssens K, Rotthier A, Baets J, Timmerman V & Janssens K (2012) Mechanisms of disease in hereditary sensory and autonomic neuropathies. *Nat Rev Neurol 2012 82* 8: 73–85

Saegusa H, Matsuda Y & Tanabe T (2002) Effects of ablation of N- and R-type Ca2+

channels on pain transmission. *Neurosci Res* 43: 1–7

Safaei R, Adams PL, Maktabi MH, Mathews RA & Howell SB (2012) The CXXC motifs in the metal binding domains are required for ATP7B to mediate resistance to cisplatin. *J Inorg Biochem* 110: 8–17

Santello M & Nevian T (2015) Dysfunction of cortical dendritic integration in neuropathic pain reversed by serotoninergic neuromodulation. *Neuron* 86: 233–246

Scherer SS, Kleopa KA & Benson MD (2020) Peripheral neuropathies. *Rosenberg's Mol Genet Basis Neurol Psychiatr Dis*: 345–375

Schwartzlow C & Kazamel M (2019a) Hereditary Sensory and Autonomic Neuropathies: Adding More to the Classification. *Curr Neurol Neurosci Rep* 19: 1–11

Schwartzlow C & Kazamel M (2019b) Hereditary Sensory and Autonomic Neuropathies: Adding More to the Classification. *Curr Neurol Neurosci Rep* 19: 1–11

Seaby EG, Pengelly RJ & Ennis S (2016) Exome sequencing explained: a practical guide to its clinical application. *Brief Funct Genomics* 15: 374–384

Seleman M, Hoyos-Bachiloglu R, Geha RS & Chou J (2017) Uses of next-generation sequencing technologies for the diagnosis of primary immunodeficiencies. *Front Immunol* 8: 847

Sène D (2018) Small fiber neuropathy: Diagnosis, causes, and treatment. *Jt Bone Spine* 85: 553–559

Shen X, Liu Y, Xu S, Zhao Q, Wu H, Guo X, Shen R & Wang F (2014) Menin regulates spinal glutamate-GABA balance through GAD65 contributing to neuropathic pain. *Pharmacol Rep* 66: 49–55

Shendure J & Ji H (2008) Next-generation DNA sequencing. *Nat Biotechnol 2008 2610* 26: 1135–1145

Shreckengost J, Halder M, Mena-Avila E, Garcia-Ramirez DL, Quevedo J & Hochman S (2021) Sensory Processing: Nicotinic receptor modulation of primary afferent excitability with selective regulation of Aδ-mediated spinal actions. *J Neurophysiol* 125: 568

Shtilbans A, Choi SG, Fowkes ME, Khitrov G, Shahbazi M, Ting J, Zhang W, Sun Y, Sealfon SC & Lange DJ (2011) Differential gene expression in patients with amyotrophic lateral sclerosis. *http://dx.doi.org/103109/174829682011560946* 12: 250–256

Soeda M, Ohka S, Nishizawa D, Hasegawa J, Nakayama K, Ebata Y, Ichinohe T, Fukuda KI &

Ikeda K (2021) Cold pain sensitivity is associated with single-nucleotide polymorphisms of PAR2/ F2RL1 and TRPM8. *Mol Pain* 17

Srour M, Shimokawa N, Hamdan FF, Nassif C, Poulin C, Al Gazali L, Rosenfeld JA, Koibuchi N, Rouleau GA, Al Shamsi A, *et al* (2017) Dysfunction of the Cerebral Glucose Transporter SLC45A1 in Individuals with Intellectual Disability and Epilepsy. *Am J Hum Genet* 100: 824–830

Stokes C, Treinin M & Papke RL (2015) Looking below the surface of nicotinic acetylcholine receptors. *Trends Pharmacol Sci* 36: 514–523

Strutz-Seebohm N, Seebohm G, Fedorenko O, Baltaev R, Engel J, Knirsch M & Lang F (2006) Functional coassembly of KCNQ4 with KCNE-beta- subunits in Xenopus oocytes. *Cell Physiol Biochem* 18: 57–66

Sturniolo GC, Lazzarini D, Bartolo O, Berton M, Leonardi A, Fregona IA, Parrozzani R & Midena E (2015) Small Fiber Peripheral Neuropathy in Wilson Disease: An In Vivo Documentation by Corneal Confocal Microscopy. *Invest Ophthalmol Vis Sci* 56: 1390–1395

Sun Y, Ruivenkamp CAL, Hoffer MJV, Vrijenhoek T, Kriek M, van Asperen CJ, den Dunnen JT & Santen GWE (2015) Next-Generation Diagnostics: Gene Panel, Exome, or Whole Genome? *Hum Mutat* 36: 648–655

Swanwick RS, Pristerá A & Okuse K (2010) The trafficking of NaV1.8. *Neurosci Lett* 486: 78–83

Szolcsányi J & Sándor Z (2012) Multisteric TRPV1 nocisensor: A target for analgesics. *Trends Pharmacol Sci* 33: 646–655 doi:10.1016/j.tips.2012.09.002 [PREPRINT]

T S, K A, M O, T Y, S F & Y N (1987) Ehlers-Danlos syndrome. A variant characterized by the deficiency of pro alpha 2 chain of type I procollagen. *Arch Dermatol* 123: 76–79

Tahara T, Shibata T, Nakamura M, Yamashita H, Yoshioka D, Hirata I & Arisawa T (2010) Homozygous TRPV1 315C influences the susceptibility to functional dyspepsia. *J Clin Gastroenterol* 44: e1–e7

Takahashi N, Mizuno Y, Kozai D, Yamamoto S, Kiyonaka S, Shibata T, Uchida K & Mori Y (2008) Molecular characterization of TRPA1 channel activation by cysteine-reactive inflammatory mediators. *Channels*

Talevich E, Shain AH, Botton T & Bastian BC (2016) CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol* 12

Tan PL, Barr T, Inglis PN, Mitsuma N, Huang SM, Garcia-Gonzalez MA, Bradley BA, Coforio

S, Albrecht PJ, Watnick T, *et al* (2007) From the Cover: Loss of Bardet–Biedl syndrome proteins causes defects in peripheral sensory innervation and function. *Proc Natl Acad Sci U S A* 104: 17524

Terkelsen AJ, Karlsson P, Lauria G, Freeman R, Finnerup NB & Jensen TS (2017) The diagnostic challenge of small fibre neuropathy: clinical presentations, evaluations, and causes. *Lancet Neurol* 16: 934–944

Tesfaye S & Selvarajah D (2012) Advances in the epidemiology, pathogenesis and management of diabetic peripheral neuropathy. *Diabetes Metab Res Rev* 28: 8–14

Tétreault M, Gonzalez M, Dicaire MJ, Allard P, Gehring K, Leblanc D, Leclerc N, Schondorf R, Mathieu J, Zuchner S, *et al* (2015) Adult-onset painful axonal polyneuropathy caused by a dominant NAGLU mutation. *Brain* 138: 1477

Thaisetthawatkul P, Fernandes Filho JAM & Herrmann DN (2013) Contribution of QSART to the diagnosis of small fiber neuropathy. *Muscle Nerve* 48: 883–888

Theadom A, Roxburgh R, Macaulay E, O'grady G, Burns J, Parmar P, Jones K & Rodrigues M (2019) Prevalence of Charcot-Marie-Tooth disease across the lifespan: a population-based epidemiological study. *BMJ Open* 9: 29240

Themistocleous AC, Ramirez JD, Serra J & Bennett DLH (2014) The clinical approach to small fiber neuropathy and painful channelopathy. *Pract Neurol* 14: 368–379

Tong Q, Zhang W, Conrad K, Mostoller K, Cheung JY, Peterson BZ & Miller BA (2006) Regulation of the transient receptor potential channel TRPM2 by the Ca2+ sensor calmodulin. *J Biol Chem* 281: 9076–9085

Torpy JM, Kincaid JL & Glass RM (2010) Peripheral Neuropathy. *JAMA* 303: 1556–1556

Trivedi JR, Silvestri NJ & Wolfe GI (2013) Treatment of Painful Peripheral Neuropathy. *Neurol Clin* 31: 377–403

Tsai PC, Soong BW, Mademan I, Huang YH, Liu CR, Hsiao CT, Wu HT, Liu TT, Liu YT, Tseng YT, *et al* (2017) A recurrent WARS mutation is a novel cause of autosomal dominant distal hereditary motor neuropathy. *Brain* 140: 1252–1266

Vandewauw I, De Clercq K, Mulier M, Held K, Pinto S, Van Ranst N, Segal A, Voet T, Vennekens R, Zimmermann K, *et al* (2018) A TRP channel trio mediates acute noxious heat sensing. *Nature* 555: 662–666

Vautrin J & Barker JL (2003) Presynaptic quantal plasticity: Katz's original hypothesis revisited. *Synapse* 47: 184–199

Veluchamy A, Hébert HL, Meng W, Palmer CNA & Smith BH (2018) Systematic review and meta-analysis of genetic risk factors for neuropathic pain. *Pain* 159: 825–848

Vernon H (2020) OMIM Entry - # 609313 - MENTAL RETARDATION, ENTEROPATHY, DEAFNESS, PERIPHERAL NEUROPATHY, ICHTHYOSIS, AND KERATODERMA; MEDNIK.

Vincler M & McIntosh JM (2007) Targeting the α9α10 nicotinic acetylcholine receptor to treat severe pain. *http://dx.doi.org/101517/14728222117891* 11: 891–897

Wang B, Li X, Huang S, Zhao H, Liu J, Hu Z, Lin Z, Liu L, Xie Y, Jin Q, *et al* (2019) A novel WARS mutation (p.Asp314Gly) identified in a Chinese distal hereditary motor neuropathy family. *Clin Genet* 96: 176–182

Wang Q, Shashikant CS, Jensen M, Altman NS & Girirajan S (2017) Novel metrics to measure coverage in whole exome sequencing datasets reveal local and global non-uniformity. *Sci Reports 2017 71* 7: 1–11

Watson JC & Dyck PJB (2015a) Peripheral Neuropathy: A Practical Approach to Diagnosis and Symptom Management. *Mayo Clin Proc* 90: 940–951

Watson JC & Dyck PJB (2015b) Peripheral Neuropathy: A Practical Approach to Diagnosis and Symptom Management. *Mayo Clin Proc* 90: 940–951

Weiss N & Zamponi GW (2021) Opioid Receptor Regulation of Neuronal Voltage-Gated Calcium Channels. *Cell Mol Neurobiol* 41: 839–847

Wieskopf JS, Mathur J, Limapichat W, Post MR, Al-Qazzaz M, Sorge RE, Martin LJ, Zaykin D V., Smith SB, Freitas K, *et al* (2015) The nicotinic α6 subunit gene determines variability in chronic pain sensitivity via cross-inhibition of P2X2/3 receptors. *Sci Transl Med* 7

Wu S, Yang S, Bloe CB, Zhuang R, Huang J & Zhang W (2021) Identification of Key Genes and Pathways in Mouse Spinal Cord Involved in ddC-Induced Neuropathic Pain by Transcriptome Sequencing. *J Mol Neurosci* 71: 651–661

Xu L, Ding X, Wang T, Mou S, Sun H & Hou T (2019) Voltage-gated sodium channels: structures, functions, and molecular modeling. *Drug Discov Today* 24: 1389–1397

Yeasmin S, Begum N, Begum S & Rahman SMH (2007) Sensory Neuropathy in Hypothyroidism: Electrophysiological and Clinical Findings. *J Bangladesh Soc Physiol* 2: 1–6

Yuan J-H, Estacion M, Mis MA, Tanaka BS, Schulman BR, Chen L, Liu S, Dib-Hajj FB, Dib-Hajj SD & Waxman SG (2021) KCNQ variants and pain modulation: a missense variant in Kv7.3 contributes to pain resilience. *Brain Commun* 3

Zafeiriou DI, Lehmann-Horn F, Vargiami E, Teflioudi E, Ververi A & Jurkat-Rott K (2009) Episodic ataxia type 2 showing ictal hyperhidrosis with hypothermia and interictal chronic diarrhea due to a novel CACNA1A mutation. *Eur J Paediatr Neurol* 13: 191–193

Zeberg H, Dannemann M, Sahlholm K, Tsuo K, Maricic T, Wiebe V, Hevers W, Robinson HPC, Kelso J & Pääbo S (2020) A Neanderthal Sodium Channel Increases Pain Sensitivity in Present-Day Humans. *Curr Biol* 30: 3465-3469.e4

Zhou L (2021) Small Fiber Neuropathy in the Elderly. *Clin Geriatr Med* 37: 279–288

Zhu M, Ni W, Dong Y & Wu ZY (2013) EGFP Tags Affect Cellular Localization of ATP7B Mutants. *CNS Neurosci Ther* 19: 346