Application of Next Generation Sequencing to Study the Genetics of Idiopathic Neonatal Arterial Ischemic Stroke

Jonathan F. Olival T.

UNIVERSITAT DE BARCELONA

Doctoral Thesis

Application of Next Generation Sequencing to Study the Genetics of Idiopathic Neonatal Arterial Ischemic Stroke

Jonathan F. Olival T.

Barcelona, 2022

UNIVERSITAT DE BARCELONA

Doctoral Program in Genetics

Application of Next Generation Sequencing to Study the Genetics of Idiopathic Neonatal Arterial Ischemic Stroke

Thesis submitted by Jonathan F. Olival T. in fulfillment of the requirements to obtain a Doctor of Philosophy degree from the University of Barcelona

Author:	**Director**:	**Co-Director**:	**Tutor**:
Jonathan F. Olival T.	Dr. Francesc Palau Martínez	Dr. Alfredo Garcia-Alix Pérez	Dra. Raquel Rabionet Janssen

Thesis performed in the Neurogenetics and Molecular Medicine Laboratory - *Instituto Pediátrico de Enfermedades Raras* (IPER) of the Institut de Recerca Sant Joan de Déu

ACKNOWLEDGMENTS

En la vida, a lo largo de ella, se presentan y se presentarán muchas oportunidades y siempre las más importantes son las que suelen producir ese miedo, ese vértigo y esa incertidumbre de si uno será capaz de afrontar el reto. Para mí, el haber llegado hasta este punto tan importante en mi vida, tanto personal como profesional, representa una de esas oportunidades que nunca me hubiese imaginado que disfrutaría.

Quiero agradecer a Paco por haber confiado en mi desde el principio y por haberme apoyado cuando te comenté que quería meter el proyecto para el que trabajaba como tesis para poder acceder al doctorado. Nunca se me olvidará la vez que me entrevistaste en la planta cero del docente en 2018. Sueles ser una persona muy ocupada pero siempre has estado allí en los momentos más decisivos y aportando infinidad de conocimiento clave en mi desarrollo. ¡Muchas gracias por todo!

Janet, al igual que sucede con Paco, nunca olvidaré cuando subí con él al laboratorio para hablar directamente contigo. Me sorprendió muchísimo saber que venias del mismo sitio que yo y que podías ser un ejemplo de superación, sobre todo estando en el extranjero. Recuerdo claramente cuando me preguntaste que por qué tenía escrito en mi CV que me consideraba una persona resiliente. Creo que en este tiempo que hemos compartido lo hemos descubierto. Muchas gracias por estar siempre allí, por apoyarme muchísimo en mi doctorado y, a pesar de no estar directamente relacionada con la tesis, por darme esos constante empujones. Gracias por la confianza puesta en mí, por tu sinceridad y muchas veces escucharme. Me has enseñado muchas cosas y al mismo tiempo me has reforzado la disciplina y la rigurosidad que hay que tener a nivel profesional.

Alfredo, eres otra de esas personas a las que tengo mucho que reconocer en toda esta etapa. Gracias por confiar en mis capacidades para desempeñar esa labor de investigador el en proyecto de la Marató de la TV3. No tienes idea lo que significó haber entrado en ese proyecto y la oportunidad que hoy en día representa a nivel profesional. Muchas gracias por todos tus conocimientos compartidos, tus ideas, tus puntos de vista, por todo el apoyo constante y el darme el visto bueno para meter el proyecto como tesis. Siempre has estado positivo y con buenos deseos. Incluso, en tus peores momentos, has estado presente y es algo que valoro increíblemente y que siempre tendré en cuenta.

Gracias a Lara, Jordi, María, Alba, Johana, Elena, Gemma. Fueron los primeros que conocí y relacioné. Incontables las veces que les he preguntado y pedido consejo sobre infinidad de experimentos y técnicas. Lara, algún día espero hacer las presentaciones como las haces, tienen un nivel que se disfruta. Jordi, luego de 4 años nunca te he visto agobiado y ni yo ni nadie sabe cómo lo haces. ¡Deberías de compartir tu secreto! Gracias María y Alba (la compi, el equipo NGS) por todos esos momentos no sólo científicos, sino también de reflexión, y espiritualidad. Muchas veces desconectamos nuestro ser por estar tan enfocado en lo laboral y mantener un equilibrio es fundamental. Gracias también por todos esos momentos únicos, las risas, la música electrónica recomendada (Alba gracias por los buenos DJs), los puntos de vista tan diferentes que tenemos pero que entendemos y aprendemos a tolerar, todas las conversaciones de vida, mindfulness, yoga, inteligencia emocial y un sinfín de cosas que podría escribir aquí. En general, cada uno de ustedes me ha aportado mucho conocimiento en diferentes áreas, cosas positivas, momentos únicos, risas y buena compañía. Los considero a todos unos cracks en lo que hacen, de verdad que son ejemplo para seguir para muchas cosas y lo que he aprendido de ustedes siempre formará parte de mí.

A mis amigos de toda la vida que, a pesar de estar esparcidos por el mundo, la distancia nunca ha sido excusa para no apoyarnos. Siempre estamos al tanto de todos nosotros y creemos ciegamente en nosotros. A ustedes, Deli, Andre (que fortuna tener en Barcelona), Esthy, Greis, Faby y Gus, ¡Muchísimas gracias!

¡A mi familia! La cual siempre ha estado allí aun estando tan lejos y con tantos años sin vernos. Todos mis logros van dedicados a ustedes, a mi mamá, mi papá, mi abuela ya que, gran parte de lo que soy hoy en día como persona, ha sido por el apoyo, el amor y los ánimos que me han dado desde siempre. A mi hermanita, que por suerte tú si no estas tan lejos y puedo verte más seguido. Gracias por tu aliento y motivación durante todo mi proceso de emigración. Desde un principio has sido un pilar fundamental en mi vida y un claro ejemplo de superación. ¡Los amo a todos!

Gracias a ti Yesica. Porque todo esto creo que no lo hubiese logrado sin ti, sin tu apoyo, sin tu amor, sin tu cariño, sin tu presencia en todos los momentos buenos y no tan buenos, por creer en mí y por decirme constantemente lo orgullosa que estás de mí.

Nunca lo había escuchado tanto. Te amo muchísimo y me siento muy afortunado de tenerte a mi lado y de que hayas formado parte de todo este proceso. Fuiste la que inició, indirectamente, todo esto y lo sabes. Esto es algo que te lo agradeceré de por vida. No tienes idea de lo que significa para mi saber que cuento con una persona como tú a mi lado. Ahora es que nos quedan aventuras juntos. ¡Eres increíble!

Por último, gracias a mí mismo. Por haber llegado hasta donde he llegado y no tirar la toalla en ningún momento. Por ser resiliente y perseverante frente a las adversidades, mantener siempre mis metas claras y luchar por ellas. Por enfrentarme a nuevos retos y exponerme a mis miedos (sin exposición no hay crecimiento) y saber que en este tiempo he podido dejar un granito de arena en la sociedad. Todo esto me hace, cuando lo pienso, me hace sentir orgulloso y rotundamente feliz.

ABSTRACT

Neonatal arterial ischemic stroke (NAIS) is a cerebrovascular disease that shows a focal disruption of cerebral blood flow due to artery obstruction. Due to its early onset, it can cause long-lasting outcomes such as cerebral palsy, loss of neurologic function, physical language, and visual impairments, and behavioural problems. The genetics of many complex diseases such as stroke is far from being resolved. Additionally, several investigations have focused on children, young adults, and adults have attempted to identify the missing heritability without success. At time, the role of genetics as a determinant in NAIS has not been investigated.

To decipher the genetics of idiopathic NAIS new approaches and strategies are needed to identify new candidate genes. Especially those focused on rare *de novo* variants. In this study we hypothesised that sporadic idiopathic NAIS in term newborns could be associated with *de novo* dominant variants in genes involved in relevant biological pathways of the pathophysiology of NAIS, such as endothelial rheology, coagulation, and cell adhesion. This based on the current evidence of *de novo* dominant mutations associated with sporadic cases of epileptic and developmental encephalopathies in newborns and infants. Furthermore, in case of not being able to demonstrate this hypothesis, we considered the autosomal recessive inheritance as a second hypothesis. To accomplish this, we applied whole-exome sequencing (WES) in trio (patient and its parents). This approach could allow us to identify a suitable candidate gene .

Furthermore, because clinically homogeneous subgroups are assumed to be genetically homogeneous as well, we evaluated homogenous clinical subgroup of NAIS according to the arterial territory of the middle cerebral artery (MCA) in the stroke event. This to increase the power of identification of candidate genes and gene-disease association with the disease.

In our study, 23 consecutive infants suffering from acute symptomatic idiopathic NAIS, and their parents, were prospectively recruited for WES-trio. We conducted a custom workflow analysis comprising variant annotation, filtering, pathogenicity impact prediction, and a knowledge-driven analysis (KDA) pipeline for scoring, ranking, and prioritization of the variants. Our results identified 28 *de novo* variants in 28 unique genes, meaning that the probands did not share neither common *de novo* variants nor genes. Moreover, our KDA pipeline identified the c.1292A>*G* (p.Gln431Arg) *de novo*

dominant variant in *PIK3CD* as the strongest autosomal-dominant candidate gene for NAIS association. This variant was carried by a patient who suffered a perforant stroke. However, in terms of the autosomal recessive inheritance as a second hypothesis, no candidate genes were found.

In silico analyses and functional studies of PIK3CD^{Gln431Arg} showed that this variant could affect the protein function due to the damaging predictions of the pathogenicity predictor tools, protein destabilization, and aberrant aggregations at the subcellular level. However, when studying its role in the PI3K/Akt pathway, a relevant cell-signalling pathway involved in stroke in which PIK3CD interacts, we found no significant changes in the expressions of total AKT and phosphorylated-AKT (Ser473).

Although we have not found conclusive evidence that demonstrate the genetic cause of idiopathic NAIS, we cannot completely rule out that the genetic cause of the disease is not due to the presence of genetic variations. In conclusion, because of the pertinence of knowing the etiopathogenesis of this major neonatal disorder and molecular mechanisms involved, the development of precision medicine, and safer tailored therapies, more studies and approaches are needed. Specially to clarify the genetic cause of the disease in term newborns and whether *PIK3CD* could be associated with perforant stroke in NAIS.

CONTENTS

TABLE OF CONTENT

INDEX OF FIGURES

INDEX OF TABLES

INDEX OF SUPPLEMENTARY DATA

ABBREVIATIONS

ACA	Anterior Cerebral Artery
AARS1	Alanyl-TRNA Synthetase 1
AChA	Anterior Choroidal Artery
ACMG	American College of Medical Genetics
AD	Autosomal Dominant
aEEG	Amplitude Electroencephalography
AFF3	AF4/FMR2 Family Member 3
ALOX5AP	Arachidonate 5-Lipoxygenase Activating Protein
APC	Activated Protein C
APoC	Adenomatous Polyposis Coli
APOE	Apolipoprotein E
AR	Autosomal Recessive
ATD	Arterial Territory Definition
BAI1	Brain-Specific Angiogenesis Inhibitor 1
BAM	Binary Alignment Map
BMI	Body Mass Index
BSITD-III	Bayley Scales of Infant and Toddler Development, Third Edition
CACNA1A	Calcium Channel, Voltage-Dependent, P/Q Type, Alpha-1a Subunit
CADASIL	Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy
CADD	Combined Annotation Dependent Depletion
CHD4	Chromodomain Helicase DNA-Binding Protein 4
CHRNA4	Cholinergic Receptor Nicotinic Alpha 4 Subunit
CLDN14	Claudin 14
CNOT3	CCR4-Not Transcription Complex, Subunit 3
CNV	Copy Number Variation
COL4A1	Collagen, Type Iv, Alpha-1
COL6A3	Collagen Type VI Alpha 3 Chain
COX-2	Cyclooxygenases-2
CP	Cerebral Palsy
CRP	C-Reactive Protein
CSF	Cerebrospinal Fluid
CSVS	Collaborative Spanish Variability Server
CT	Computed Tomography
CTGF	Connective Tissue Growth Factor
DAPI	4',6-diamidino-2-phenylindole
DENND4A	DENN Domain Containing 4A
DNV	*De novo* Variants
DR	Digenic Recessive
DWI	Diffusion Weighted Imaging
DYM	Dymeclin
DYRK1A	Dual-Specificity Tyrosine Phosphorylation-Regulated Kinase 1a
eBLT	Enrichment Bead-Linked Transposomes

EC	Endothelial Cells
ECMO	Embolism During Extracorporeal Membrane Oxygenation
EEG	Electroencephalography
EGF	Epidermal Growth Factor
EGFP	Enhanced Green Fluorescence Protein
eNOS	Endothelial Nitric Oxide Synthase
EPS15L1	Epidermal Growth Factor Receptor Pathway Substrate 15 Like 1
Erk	Extracellular Regulated Kinase
ESRRG	Estrogen Related Receptor Gamma
F13A1	Factor XIII, A1 Subunit
F2P	Factor II Polymorphism
F5	Coagulation Factor V
FAM193A	Family with Sequence Similarity 193 Member A
FAM83A	Family with Sequence Similarity 83 Member A
FAR2	Fatty Acyl-CoA Reductase 2
FASTKD5	FAST Kinase Domains 5
FATHMM	Functional Analysis through Hidden Markov Models
FGF	Fibroblast Growth Factor
FKHR	Forkhead Factors
FVL	Factor V Leiden
GATK	Genome Analysis Tool Kit
G-CSF	Granulocyte-Colony Stimulating Factor
GERP	Genomic Evolutionary Rate Profiling
GLA	Galactosidase, Alpha
GMFCS	Gross Motor Function Classification System
GO	Gene Ontology
GRCh37	Genome Reference Consortium Human Build 37
GWAS	Genome-Wide Association
HC	Head Circumference
Hcy	Homocysteine
HECTD1	HECT Domain E3 Ubiquitin Protein Ligase 1
HGF	Hepatocyte Growth Factor
HGMD	Human Genome Mutation Database
hHcy	Hyperhomocysteinaemia
HIF	Hypoxia-Inducible Factor
HSF	Human Splice Finder
ICAM-1	Intercellular Adhesion Molecule-1
IFN-α	Interferon-Alpha
IFN-γ	IFN-Gamma
IGSF5	Immunoglobulin Superfamily Member 5
ITGB3	Integrin, Beta-3
JAM1	Junctional Adhesion Molecule 1
JAM2	Junctional Adhesion Molecule 2
JAM3	Junctional Adhesion Molecule 3
KCNC2	Potassium Voltage-Gated Channel Subfamily C Member 2

KCNQ2	Potassium Voltage-Gated Channel Subfamily Q Member 2
KDA	Knowledge-Driven Analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
KRT6A	Keratin 6A
LIF	Leukaemia Inhibitory Factor
MAF	Minor Allele Frequency
MAPK	Mitogen-Activated Protein Kinase
MCA	Middle Cerebral Artery
MEF2A	Myocyte Enhancer Factor 2A
MIM	Mendelian Inheritance in Man
MMP	Matrix Metalloproteinase
M-NAIS	Massive NAIS
MRI	Magnetic Resonance Imaging
mRNA	mature RNA
MTHFR	Methylenetetrahydrofolate Reductase
NAIS	Neonatal Arterial Ischemic Stroke
NGS	Next-Generation Sequencing
NHS	NHS Actin Remodelling Regulator
NINJ2	Nerve Injury-Induced Protein 2
NO	Nitric Oxide
NOS3	Nitric Oxide Synthase 3
NOTCH3	Notch Receptor 3
NPC1	NPC Intracellular Cholesterol Transporter 1
NR2E1	Nuclear Receptor Subfamily 2 Group E Member 1
NSE	Neuronal-Specific Enolase
OCLN	Occludin
OMIM	Online Mendelian Inheritance in Man
OPN	Osteopontin
PA	Plasminogen Activator
PAI	Plasminogen Activator Inhibitor
PCA	Posterior Cerebral Artery
PCBP1	Poly(RC) Binding Protein 1
PDGF	Platelet-Derived Growth Factor
PEDF	Pigment Epithelium-Derived Factor
PF4	Platelet Factor-4
PGE2	Prostaglandin E2
PGI2	Prostacyclin
PI3K	Phosphatidylinositol 3'-Kinase
PIGF	Placental Growth Factor
PIK3CD	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Delta
PIP2	Phosphatidylinositol-4,5-bisphosphate
PIP3	Phosphatidylinositol-3,4,5-trisphosphate
PLCγ1	Phospholipase C
PPI	Protein-Protein Interaction
PRKCH	Protein Kinase C Eta

PROVEAN	Protein Variation Effect Analyzer
RA	Retinoic Acid
RBM7	RNA Binding Motif Protein 7
RIV	Relative Infarct Volume
RNF213	Ring Finger Protein 213
SCI	Silent Cerebral Infarction
SD	Standard Deviation
SDF1-1	Stromal Cell-Derived Factor-1
SETD5	Set Domain-Containing Protein 5
SHANK3	SH3 And Multiple Ankyrin Repeat Domains 3
SIFT	Sorting Intolerant from Tolerant
SLC2A10	Solute Carrier Family 2 Member 10
SNPs	Single Nucleotide Polymorphisms
SNTB2	Syntrophin Beta 2
SPTBN5	Spectrin Beta, Non-Erythrocytic 5
SYNE1	Spectrin Repeat Containing Nuclear Envelope Protein 1
SYNE2	Spectrin Repeat Containing Nuclear Envelope Protein 2
T1W	T1-Weighted
T2W	T2-Weighted
TAF	Tumour Angiogenesis Factor
TF4	Thrombocyte Factor 4
TGF	Transforming Growth Factors
TGF-β	Transforming Growth Factor-Beta
TIA	Transient Ischemic Attack
TIMPS	Tissue Inhibitors of Metalloproteinase Inhibitors
TJP2	Tight Junction Protein 2
TNF	Tumour Necrosis Factor
TOP2A	DNA Topoisomerase II Alpha
TRABD2B	TraB Domain Containing 2B
TSP-1	Thrombospondin-1
TTN	Titin
TxA2	Thromboxane A2
UNC13D	Unc-13 Homolog D
VCAM-1	Vascular Cell Adhesion Molecule 1
VEGF	Vascular Endothelial Growth Factor
VEGFR-1	Vascular Endothelial Growth Factor Receptor 1
VEGFR-2	Vascular Endothelial Growth Factor Receptor 2
VSMCs	Vascular Smooth Muscle Cells
VTE	Venous Thromboembolism
VUS	Variant of Uncertain Significance
vWF	Von Willebrand Factor
WES	Whole-Exome Sequencing
WGS	Whole-Genome Sequencing
WT	*Wild type*
YAP	Yes-Associated Protein

INTRODUCTION

1. Stroke as a Disease

1.1 Overview

It has been more than 100 years since Sigmund Freud published his third monograph related to cerebral palsy, *Infantile cerebral paralysis* (Freud, 1968). In this book, Freud recognized that many cases of infantile cases of cerebral palsy had a vascular component. A piece of the monograph states: "*It has further been determined that a large number of cases of infantile cerebral palsy is caused by the same factors that bring about the majority of cases of cerebral paralysis of adults: tearing, embolism, and thrombosis of cerebral vessels*" (Rothman, 2002). Thenceforth, great efforts have been made to understand the pathophysiological mechanism involved in the disease, clarify the epidemiology, associated risk factors and outcomes, and improve the definition and classification of cerebral infarction by neuroimaging.

Stroke is a cerebrovascular disease that displays a negative impact on social, public health and resources (Alvarez-Sabín et al., 2017; Díaz-Guzmán et al., 2012). In 2019 alone, there were 11.401 new cases of stroke in Spain, costing public health around 315 million euros (Atlas del Ictus, 2019). Mayo Clinic defines strokes as a focal disruption or reduction in the brain's blood supply, preventing brain tissue from getting oxygen and nutrients (Mayo Clinic Staff, 2021). Surviving stroke patients experience chronic sequelae that include focal brain damage, neurological deficits lasting more than 24 hours, loss of neurological function, physical impairment that compromises mobility, problems with vision, speech, mood, and cognitive and behavioural ability (McKevitt et al., 2011).

As a disease, it has a high prevalence and incidence and worldwide, tops the list of acquired disabilities in adults (Atlas del Ictus, 2019). In Spain, at present, around 27 thousand people die each year due to a stroke (Peña-Legazkue et al., 2019) and in Catalonia, according to the IBERICTUS study, the incidence of stroke is around 187.4 cases per 100.000 inhabitants (Díaz-Guzmán et al., 2012). Quality of life is tremendously affected by stroke and a national health survey in Catalonia based on the self-perceived state of health in the population confirmed this. The results showed that 76.9% of people who suffered a stroke considered that their state of health was between regular and very poor, compared to 41.2% of other chronic patients and 6.5% of the healthy population (Figure 1).

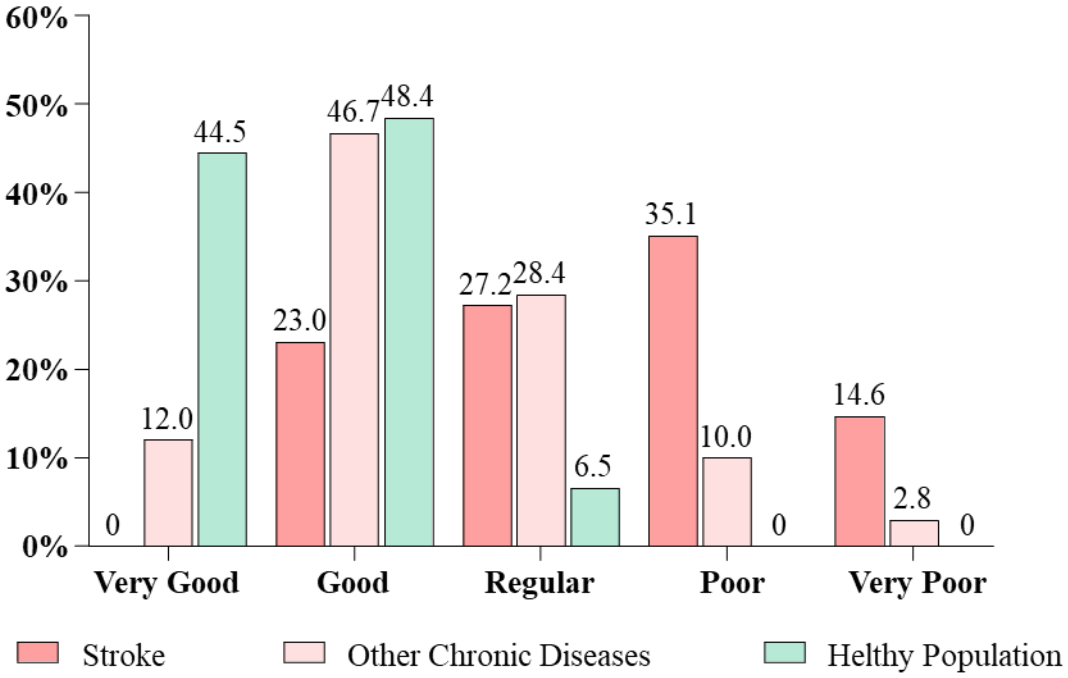


Figure 1. **Self-perceived health status data after stroke**. The graph shows the health status among healthy individuals, ones that had a stroke, and other chronic diseases. It ranges from very good to very poor. Adapted from Instituto Nacional de Estadística (2017).

The classification of stroke varies according to the clinical type and the onset of the infarction. According to the American Stroke Association, the clinical stroke subgroups are the following:

1. Ischemic, which includes cerebrovascular accidents where the vessels that supply blood to the brain are obstructed. This type of stroke occurs unnoticed and accounts for about 87% of all strokes (Silent cerebral infarction [SCI]).
2. Haemorrhagic stroke, which is caused by a weakened vessel that ruptures and bleeds into the brain, pooling and compressing the surrounding brain tissue. The haemorrhage can be either intracerebral (within the brain) or subarachnoid. It accounts for about 13% of all stroke cases.
3. Transient Ischemic Attack (TIA) defined as a temporary blockage of blood flow to the brain and does not cause permanent brain damage.
4. Cryptogenic strokes, which is a stroke with an unknown cause.
5. Brainstem stroke, which could produce a wide spectrum of symptoms, deficits, and recovery rates.

The classification of stroke onset time is based on the age of the affected individual: adult, child (1 to 18 years of age), and perinatal. The latter is known as "a heterogeneous disorder in which there is a focal disruption of cerebral blood flow secondary to arterial or venous thrombosis or embolization, between 20 weeks of foetal life through the 28th postnatal day and confirmed by neuroimaging or neuropathological studies" (Govaert et al., 2009b; Pulver et al., 2017; Raju et al., 2007).

Perinatal stroke encompasses three groups: (1) fetal stroke occurring after 20 weeks of gestation and 0 days of life (birth); (2) neonatal stroke ranging from day 0 to day 28 of life; (3) presumed perinatal stroke, which defines infants who, outside of the normal perinatal period, have a chronic focal stroke on neuroimaging that is believed to have occurred in the perinatal period (Mineyko et al., 2011). The goal of this categorization was to enable and help build new research protocols around the world (Chabrier et al., 2011).

1.2 Neonatal Arterial Ischemic Stroke

Experts in the field refer to Neonatal Arterial Ischemic Stroke (NAIS) as infants who present a focal interruption of cerebral blood flow due to obstruction that occurs between postnatal days 0 and 28 and with acute symptoms diagnosed by neuroimaging that confirms the arterial ischemic stroke (Núñez et al., 2021; Raju et al., 2007). In addition, symptomatic NAIS cases are the most diagnosed in the neonatal period because they present with clinical manifestations that lead to the suspicion of a focal brain lesion or to perform a neuroimaging study.

The exact time of NAIS onset was generally unclear, but advances in neuroimaging have provided insight into the origin of the stroke and the time of injury. By analysing the dynamic appearance of cytotoxic edema in brain images of newborns with stroke while they were developing symptoms using magnetic resonance imaging (MRI), researchers concluded that the approximate period of the lesion must have occurred around day 0 (Dudink et al., 2009; Govaert et al., 2009b; Pulver et al., 2017).

After the first month of life, ischemic stroke has an incidence of about 1/30.000 children per year, while symptomatic neonatal ischemic stroke is about 1/3,000 to 1/5,000 live births (Grunt et al., 2015; Nelson et al., 2004; Raju et al., 2007). Furthermore, studies on the predominance of sex in NAIS have shown a male majority in the disease (Golomb et

al., 2009; Kirton et al., 2008). NAIS is one of the main causes of severe long-term neurodisability with a wide range of possible outcomes, the most common being motor and language deficits, but also cognitive deficit, epilepsy, and chronic neuropsychological and behavioural disorders (Bosenbark et al., 2017; Chabrier et al., 2016; Wagenaar et al., 2018).

Depending on the affected blood vessel, different authors have defined methods that can help classify the subtypes of NAIS. For example, the classification based on the arteries of the cerebral blood bed (Figure 2) affected by the stroke event: (1) Middle Cerebral Artery (MCA) infarction, (2) Anterior Cerebral Artery (ACA) infarction, (3) Posterior Cerebral Artery (PCA) infarction, (4) Vertebrobasilar infarction, (5) Cerebellar infarction and (6) Lacunar infarction, which affects arteries located in the deep cerebral white matter, basal ganglia, or pons (Monroe, 2020).

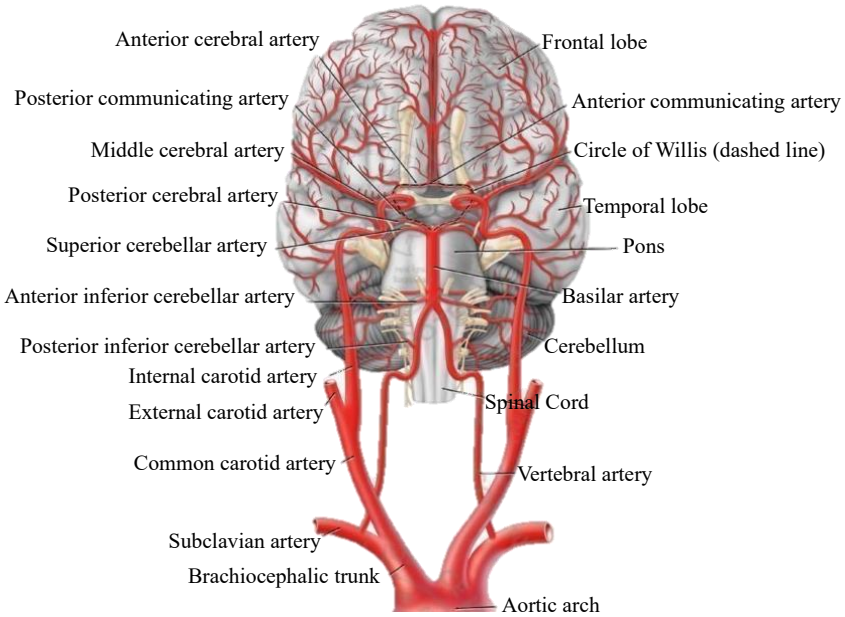


Figure 2. **Cerebral blood bed figure**. Illustration of the anteroinferior arteries of the brain and the major arterial structures that transport blood to the brain, including: the aortic arch, Circle of Willis, the common carotid arteries, vertebral arteries, internal carotid arteries, basilar artery, and multiple cerebellar and cerebral arteries. Taken and modified from (*Anteroinferior Arteries of the Brain*, 2021).

Currently, a more modern classification considers neuroimaging analysis of MRI to unify efforts and create an advanced definition of NAIS based on the time of both injury and presentation of the damage in relation to the age group (Figure 3). This consensus

allows clinicians to accurately diagnose and predict the neurodevelopmental outcome of the infant during the first two weeks after the ischemic event and makes it more suitable for the age group studied (Arca et al., 2020). This classification comprises two major groups: (1) focal or multifocal stroke, depending on whether there is a single or multiple affected territories, respectively, and (2) M1 to M4, which depends on the arterial territory affected *per se*.

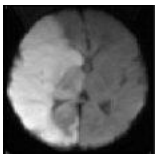
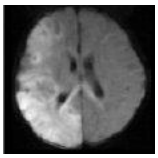
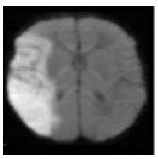
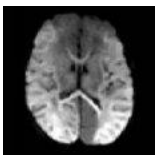
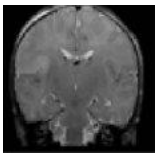
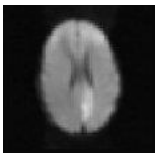
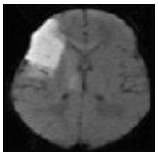
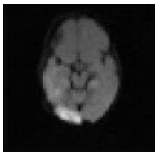
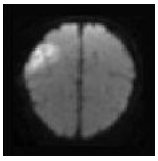
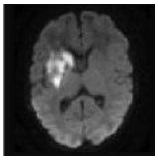
ATD	Scheme	MRI	ATD	Scheme	MRI
M1 Sphenoidal segment crossing over the limen insulae			**Multifocal**		
M1 Pre Involves almost the entire territory of the MCA including all or partially the basal ganglia			**Unilateral** Different arterial territory in one hemisphere		
M1 Post The basal ganglia were not found damaged			**Bilateral** Different arterial territory in both hemispheres		
M2 Insular segment			**ACA**		
M3 Opercular Segment			**PCA**		
M4 Cortical, pial segment			**Perforant**		

Figure 3. **Affected arterial territory definition classification**: classification of NAIS according to MRI neuroimaging analysis, which is divided into groups M1 to M4 and focal or multifocal stroke, depending on whether there is a single or multiple affected territories. This classification depends on the affected arterial territory. The scheme column represents a drawing of the corresponding MRI. ATD: Arterial Territory Definition, MRI: Magnetic Resonance Imaging. Information was taken and modified from Arca et al., 2020

1.3 Clinical Features of NAIS

The clinical hallmark of NAIS is seizures, especially focal clonic or hemicorporal seizures (Raju et al., 2007). In our experience, the second most frequent type of seizures frequently seen in NAIS are apnea or tonic seizures (Table 1). More subtle early findings associated with stroke on physical examination include poor feeding, hypotonia, lethargy, and asymmetrical limb tone and movements (Chalmers, 2005). Some infants may have asymmetrical tone and movement in the limb contralateral to the infarction site for the first few days to weeks. However, many infants may remain asymptomatic.

Seizures in NAIS characteristically begin within the first 48 hours, usually after 12 hours of life. In a large consecutive cohort of infants with NAIS (N = 48) a prospective study showed that the clinical onset of seizures was at a mean age of 30.5 hours of life (Table 1 – Unpublished data from Gemma et al). However, the clinical picture of newborns with symptomatic NAIS in terms of the time of clinical debut, clinical expressiveness, duration of seizures, and response to treatment is highly variable. This variability is generally attributed to the great variability in the lesion location, even within the same affected arterial territory, extension, and volume in patients with NAIS. Nevertheless, not all newborns with NAIS have seizures.

Table 1. Clinical data in a cohort of patients with consecutive NAIS

Clinical Feature	**All NAIS N = 48**
Clinical onset, hours (SD)	30.5 (24.3)
Seizures (yes/no)	47/1
Duration of seizures, days*	1.1 (0.5-2.5)
Type of clinic seizures:	
Clonic	35 (73%)
Tonic	4 (8.3%)
Apnea	6 (12.5%)
Apnea + tonic	2 (4.2%)
Only Electric	1 (2%)
Infants who needed more than one antiepileptic drug (yes/no)	20/28

SD: Standard deviation.
* Displayed as median (P25-P75)

2. Diagnosis of NAIS

The diagnosis of NAIS is based on radiological (or pathological) confirmation. Death is very rare in the term or near-term newborn with NAIS. Therefore, neuroimaging has become an essential technique for diagnosing neonatal stroke. Imaging criteria define a cerebrovascular accident as a focal brain lesion located in an arterial territory or cerebral images corresponding to an infarct in a recognized vascular territory.

2.1 Electroencephalography

Monitoring with complete electroencephalography (EEG) or an integrated amplitude EEG (aEEG) is mandatory in infants with symptomatic NAIS. The tracing of the seizures may suggest that the focal lesion, in one of the cerebral hemispheres, is due to the presence of palmar asymmetry in the tracing voltage background, that the seizures originate in only one hemisphere, or that the seizures predominate in one of the hemispheres (Figure 4). In a careful study, 17% of patients with NAIS showed persistent asymmetry for days in the background trace voltage (Mercuri et al., 1999). In a recent study, the greater the persistence of abnormal background activity on the affected side, the worse the cognitive development. This study suggests that aEEG or EEG may further help predict cognitive development (Wagenaar et al., 2019).

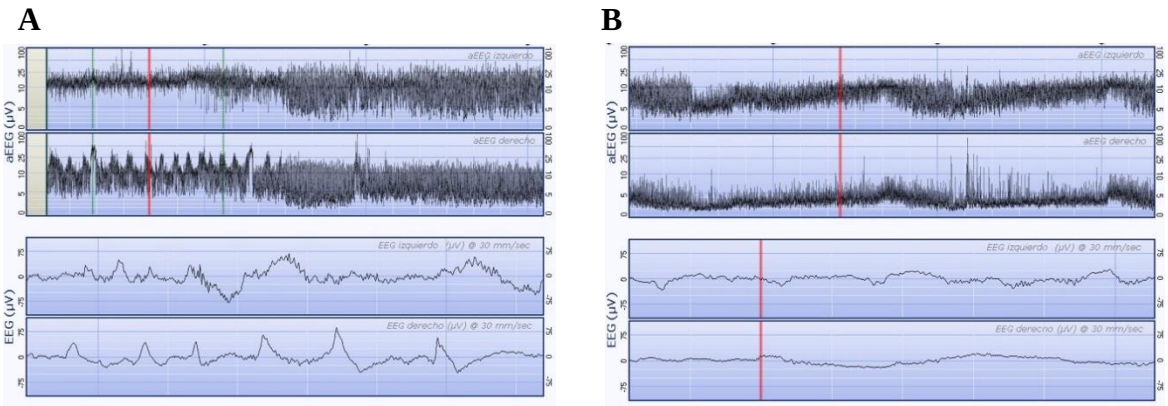


Figure 4. **Bi-channel aEEG images of NAIS**. The images show the first three hours after the admission of a newborn with cerebral arterial infarction in the M1 territory of the right Middle Cerebral Artery that debuted with hemicorporal clonic seizures in the first 24 hours. **A**) Note the epileptic status located exclusively on the right side (lower register). **B**) Tracing at four days of age shows a frank asymmetry in voltage. While the left side (top panel) shows a continuous normal voltage pattern, the right side shows a low voltage pattern.

2.2 Neuroimaging

Cranial ultrasonography, computed tomography (CT), and magnetic resonance imaging (MRI) are the three main imaging techniques available for imaging the neonatal brain. Although brain MRI is a gold standard tool for the diagnosis of stroke, brain ultrasonography and CT are two other effective diagnostic tools.

2.3 Ultrasonography

Ultrasonography allows rapid evaluation as it is readily available, inexpensive, portable, and can be performed at the bedside. It is usually the first-line brain-imaging technique performed in the neonatal unit and allows serial examinations to be performed without the need to transfer the sick newborn to the radiology service.

This technique shows focal hyperechogenicity, typically well-defined, in an arterial territory (Figure 5) and, occasionally, abnormalities of the arterial vascularization visualized by Doppler. It has been shown that the sensitivity of ultrasound increases during the first days of life (Cowan et al., 2005; Golomb et al., 2003), but these studies were performed more than 10 years ago. Recently, a clinical team evaluated the current sensitivity of cranial ultrasonography (CUS) to detect NAIS when performed by medical personnel with different training at a tertiary referral centre (Olivé et al., 2019). The results indicated that, in expert hands, CUS could detect neonatal MCA stroke in 87% of cases shortly after clinical onset. The sensitivity of this technique reported by Golomb et al. (2003) was 47.3% and slightly higher in the Cowan et al. (2005) study (68%). We can conclude that with modern CUS equipment and trained medical personnel, CUS is a very useful tool for NAIS detection until MRI can be performed.

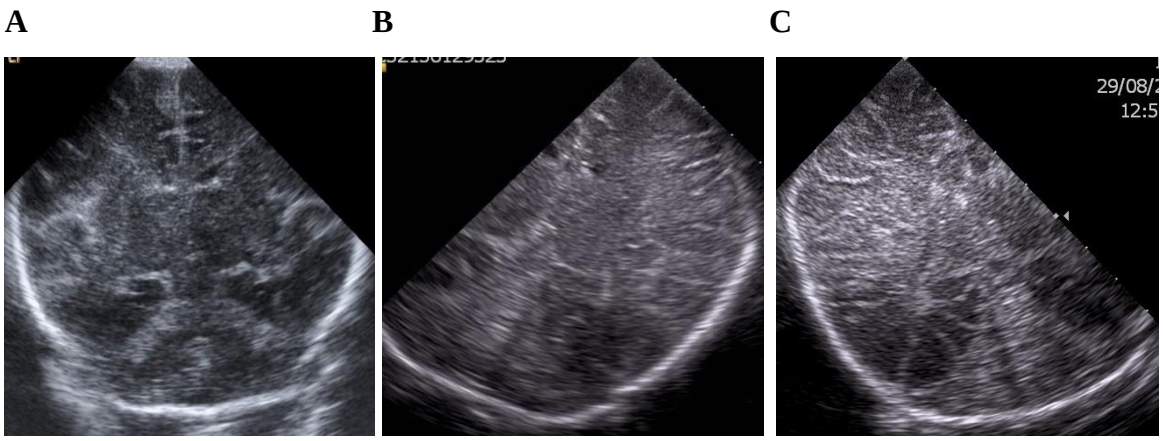


Figure 5. **Ultrasound scans images of NAIS**. **A**-**C**) Scans of three different patients in which the echo-dense cortico-subcortical lesion that corresponds to a NAIS is well appreciated.

2.4 Computed Tomography

Some centres, especially in developing countries, still use CT to diagnose NAIS, but it requires ionizing radiation and therefore has long-term risks in the newborn. Its diagnostic capacity has not been compared with the evaluation by ultrasonography or MRI of the brain. In our environment, this study has been abandoned and therefore we will not extend this information.

2.5 Magnetic Resonance Imaging

Magnetic resonance imaging is the gold standard diagnostic tool. This evaluation does not use ionizing radiation. The most frequently used sequences in MRI studies include T1-weighted (T1W), T2-weighted (T2W) and diffusion-weighted imaging (DWI) or ADC Map within the first week after clinical debut (Figure 6).

Not all MRI studies are equally informative about cerebral infarction at any given time within the first few days of life. The diagnostic efficiency for each sequence depends on the time elapsed since the infarction. Within the first week after clinical debut, while T1W barely shows alterations during the first 48 hours, T2W will show high signal intensity in the affected cortex and white matter from 24 to 48 hours onwards (Dudink et al., 2009). The reduced contrast between the cortex and white matter seen on both T1W and in T2W is called the ‘missing-cortex sign '. While T1W will show lower signal intensity in the cortex and white matter, T2W will display higher signal intensity. At the end of the first week, the pattern begins to change. T2W will show a lower signal

intensity in the cortex, while high signal intensity can be seen on T1W involving the cortex (cortical highlighting), but the white matter will show a lower signal intensity.

The magnetic resonance imaging-DWI has a high sensitivity to detect acute ischemic lesions and the changes precede those on conventional T1W, T2W and DWI. This diagnostic tool is usually the preferred study since it has a greater diagnostic capacity during the first 24 to 72 hours after clinical debut (Dudink et al., 2009).

The lesional changes of infarction show a hyperintense signal in the affected artery territory on DWI and low intensity on the ADC map. The lowest ADC can be observed around day three, after which ADC values will slowly increase (Dudink et al., 2009), and after, approximately 6-10 days pseudo-normalizing the ADC values.

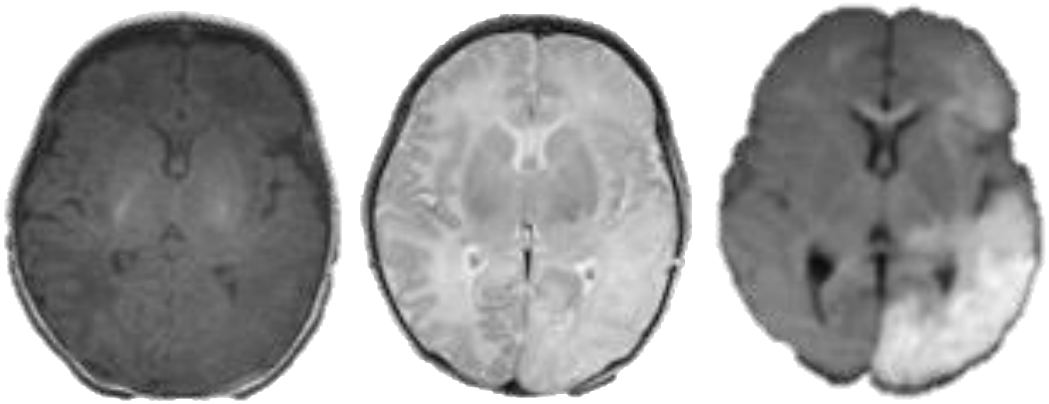


Figure 6. **Magnetic resonance imaging in an infant with NAIS**. T1W, T2W and DWI sequences at 60 hours of life and 48 hours after clinical debut in a patient with NAIS in the M3 territory of the left MCA.

Furthermore, nearly a third of patients have signs of space-occupying brain swelling, such as midline shift of the brain and/or extra-axial space and/or ventricular collapse. This infarction is called massive NAIS (M-NAIS). The relative volume of M-NAIS is greater than in the rest of the infarcts (29% vs. 4.9%). Newborns with M-NAIS always presented lesions that affected the M1 arterial territory of the MCA (Figure 7) and showed differences in the clinical presentation of seizures. Moderate to severe adverse neurodevelopmental outcomes are present in 79% of patients with M-NAIS vs. 6% in non-massive NAIS.

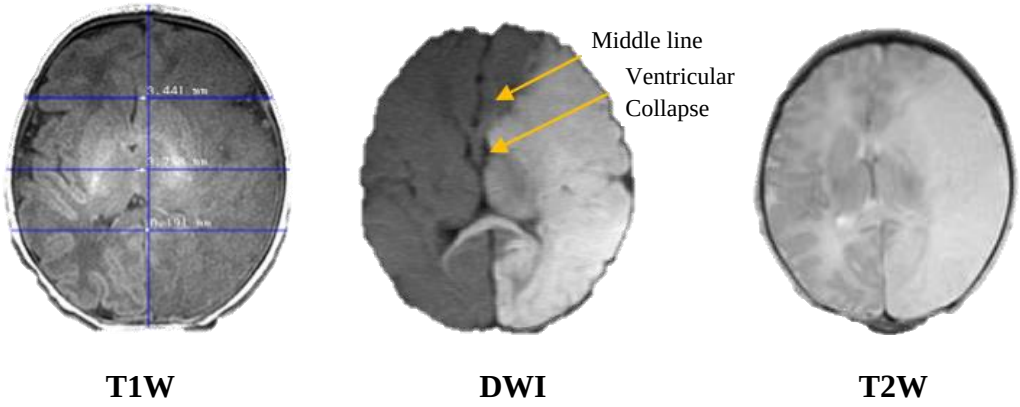


Figure 7. **Magnetic resonance imaging of M-NAIS**. T1W, DWI y T2W MRI images in a newborn with massive NAIS.

The location of neonatal arterial ischemic stroke has a considerable impact on long-term outcomes. While the distribution of stroke in adults has been extensively studied, until recently a map showing the spatial distribution of NAIS was lacking. Using a novel segmentation technique, a clinical group has been able to accurately localize brain infarcts in NAIS from early MRI data (Figure 8). The region posterior to the central sulcus is frequently the most affected, which suggests a greater vulnerability of this area (Stephan-Otto et al., 2017). Additionally, the left hemisphere is affected more frequently than the right, but both hemispheres show symmetrical patterns of involvement.

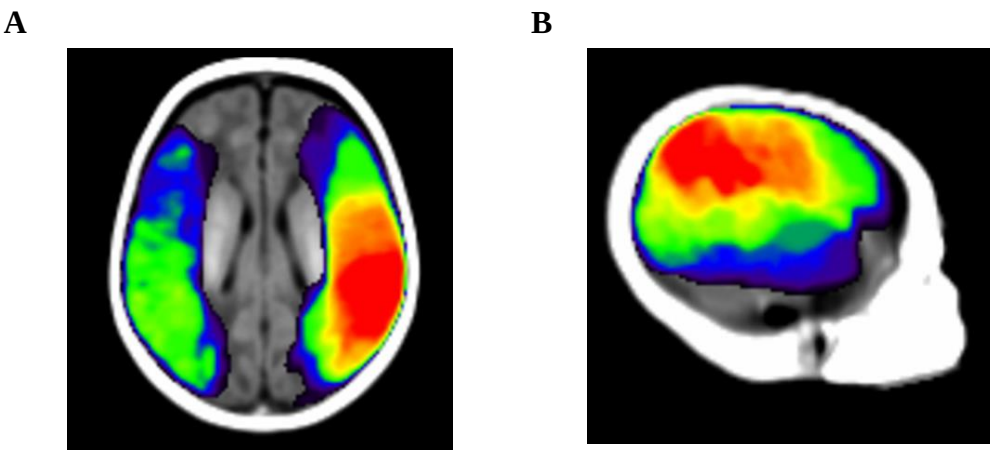
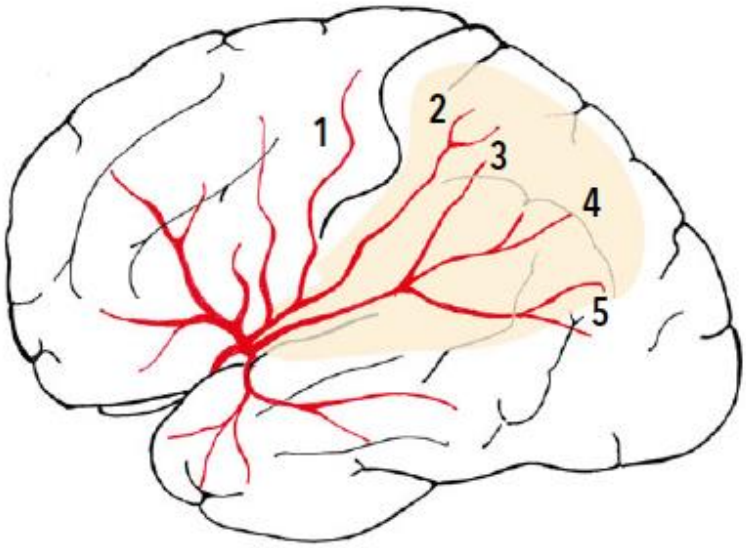


Figure 8. **Axial and sagittal MRI images**. **A**) Axial MRI cut. **B**) Sagittal MRI cut. The coloured regions are where at least 10% of the NAIS overlaps. Higher intensity colour means overlap in at least 40% of the subjects.

This preferential distribution of the NAIS in the tributary territory of the posterior branches (Figure 9, arteries 2-5) is directed backward in practically the same direction as the major axis of the Sylvian fissure. Because these vessels are straighter, they probably have a preferential and higher flow, favouring the destination of emboli in these territories.



1. Central Sulcus Artery
2. Anterior Parietal Artery
3. Posterior Parietal Artery
4. Angular Artery
5. Temporo-occipital Artery

Figure 9. **Territory distribution of the posterior artery branches**. Image taken from "*Neurología Neonatal de un Vistazo*" with copy rights (García-Alix A & Arnaez J, 2022).

2.6 Biochemical markers

Early diagnosis of NAIS is a major challenge for patient management, and the ability to accurately predict the outcome depends solely on neuroimaging biomarkers. However, some predictive motor outcome changes on MRI may take time to develop and may not yet be present 2 days after symptoms onset (Wagenaar et al., 2017). In addition, the ability to predict non-motor outcomes on MRI imaging is limited. The potential underestimation of the full extent of the injury, the difficulty in predicting intellectual disabilities, language impairment and epilepsy, as well as limitations to performing early MRIs in many hospitals and the unavailability of MRIs in some, lend themselves to supporting the search for other biomarkers of brain injury.

Neurobiochemical markers in adult stroke patients have attracted great attention due to their diagnostic and prognostic possibilities (Brouns et al., 2009; Foerch et al., 2009; Selakovic et al., 2005). To date, only one study has examined the role of the cerebrospinal fluid (CSF) neuronal-specific enolase (NSE) as a biomarker of NAIS (Arca et al., 2020). This study showed that CSF-NSE could be a surrogate indicator of cerebral infarction revealing a correlation between the volume of the lesion and the outcomes at 2 years. Having an accurate biomarker related to the volume of ischemic brain damage and prognosis within a few hours after establishing clinical suspicion is relevant to identify high-risk patients (García-Alix A & Arnaez J, 2022).

2.7 Neurodevelopmental outcomes after a NAIS

Both short-term and long-term outcomes mainly depend on the extent and location of the stroke. Mortality rates in infants suffering NAIS are very low. Recurrence of stroke in the child and siblings is rare. Adverse consequences of NAIS include cerebral palsy (CP), usually of the monoparetic or hemiparetic type, language delay or need for speech therapy, cognitive dysfunction, epilepsy, and visual deficits (Table 2). It should be noted that infants with adverse outcome develop sequelae in multiple domains.

Table 2. Outcomes prevalence of neurodevelopmental consequences in NAIS

Outcomes at seven years	Prevalence*
Neurodevelopmental disorder	≈ 40-50%
Motor disorder	< 35%
Cognitive dysfunction	< 10%
Low academic skills	< 30%
Language delay or problems	40-50%
Behavioural problems	30-50%
Epilepsy	<15%

*Based on data from (Bosenbark et al., 2017; Chabrier et al., 2016; Wagenaar et al., 2018)

In summary, NAIS has a complicated nature due to the uncertainty of when it occurs, the poorly established pathogenesis, the diversity of risk factors involved, and the heterogeneity of the lesion locations even within the same arterial territory. Besides, there is a wide variation in the size and volume of NAIS-derived lesions as well as in long-term neurological outcomes.

3. Physiopathology Pathophysiology of NAIS

3.1 Overview

The causes of ischemic stroke due to a clot in the cerebral circulation differ between age groups and vary considerably depending on factors such as geography, socioeconomics, and environment and, globally, it is the most common cause of morbidity and mortality (Arnaez et al., 2018a; Jackson, 2011). For instance, in the United States, tuberculous meningitis and neurocysticercosis are not common causes of stroke but are important in India and Central and South America (Caplan, 2009). Newborns, children, and young adults have different mechanisms involved due to the occurrence of stroke, with migraine, trauma, and cardiac diseases being important aetiologies in the latter two age groups, respectively (Caplan, 2009; Chabrier et al., 2011).

The components and functions of the arteries during infancy and childhood differ from those in older adults, because of the age group, the intracranial arteries are rarely affected by degenerative atherosclerotic changes (Caplan, 2009). Subsequently, the pathogenesis of NAIS is far from being established due to little knowledge about the mechanisms and pathways involved. Hence, some authors affirm the presence of a black box in NAIS due to the absence of critical information (Arca et al., 2020; Chabrier et al., 2011; Kirton et al., 2011; Mineyko et al., 2011) such as the internal mechanisms involved in the process.

In addition, in most cases, the interactions of the maternal–placental and foetal–neonatal related pathways are complex to study and the mechanism and causes of the arterial obstruction remain the subject of debate (Chabrier et al., 2011; Dashti et al., 2017; Mineyko et al., 2011). Furthermore, the explanation of the functional changes that are occurring within an infant having a stroke and the exact processes and pathways involved, are far from established (Govaert et al., 2009b; Lehman et al., 2014).

Although thromboembolism is detected in only a minority of all NAIS cases, it is by far the most common cause of stroke in infants due to the risk of emboli entering the cerebral circulation (Arnaez et al., 2017; Fluss et al., 2016; Govaert et al., 2009b). For example, emboli that enter the infant´s brain through placental mechanisms do so through the foetal circulation while the placenta separates at birth, leading to a direct embolization (Ichord, 2020). Additionally, placental pathology can cause an

inflammatory and prothrombotic state that promotes thrombus formation in the placenta and foetus (Elbers et al., 2011).

Other mechanisms described by different studies include the crossing of venous clots through the patent foramen ovale and redirection to the cerebral arterial vessels, a right-to-left shunt that produces emboli, and permanence of catheters in the umbilical vessels. Furthermore, the inflammatory response, coagulation, endothelium, and cell adhesion are also considered molecular mechanisms with a fundamental contribution in the generation of a stroke event (Arnaez et al., 2017; Elbers et al., 2011; Fluss et al., 2016; Govaert et al., 2009b; Ramaswamy et al., 2004; Ruff et al., 1979; Sreenan et al., 2000).

3.2 Molecular mechanisms involved

Recent advances in the basic biology of many signalling pathways, have provided powerful insights into the molecular mechanisms and connectomes involved in the disease, how they integrate into the pathophysiological condition, and the association with stroke events, especially with thrombus formation. As expected, the endothelium plays an important role in this due to its influence on the regulation of vascular tone, platelet activation and function, and endothelial cell (EC) proliferation.

3.2.1 Angiogenesis

Endothelium formation begins during embryogenesis through a physiological process called neovascularization, an important process for organ maturation and differentiation (Matsumoto et al., 2003). A *de novo* synthesis of EC (vasculogenesis) and the growth or derivation of vascular structures from other pre-existing blood vessels (angiogenesis) are the two main processes of this neovascularization (Chen et al., 2012; Risau et al., 1995). The development of this new vascular supply occurs in three different ways: (1) intussusceptive vascular growth (2) arteriogenesis and (3) sprouting angiogenesis. All three processes require molecular interactions among, cytokines, growth factors, proteases, extracellular matrix, EC, pericytes, and inflammatory cells (Carmeliet, 2000, 2003), which are defined as key elements of the pathophysiology in NAIS (Cosentino et al., 2005).

Angiogenesis (angios: vessels; genesis: formation), is a complex dynamic process and the most predominant mechanism of blood vessel formation in late embryonic development (Folkman et al., 1992). However, when this regulation is disrupted, it has a greater impact on our health and forms part of many pathophysiological processes and diseases (Carmeliet, 2003). The new vascular bed generated from the angiogenesis process relies on many tightly regulated molecular mechanisms to expand into a network of stable sprouting vessels (Chen et al., 2012) (Figure 10). For instance, some studies have determined that the loss of a single allele of the *VEGFA* gene, which code for vascular endothelial growth factor (VEGF) protein, causes embryonic vascular defects (Carmeliet et al., 1996; Ferrara et al., 1996) and that only a 25% reduction in VEGF levels results in motor neurons degeneration and amyotrophic lateral sclerosis (Oosthuyse et al., 2001). This establishes the vital role of molecular mechanisms and the strict dose of factors involved in angiogenesis.

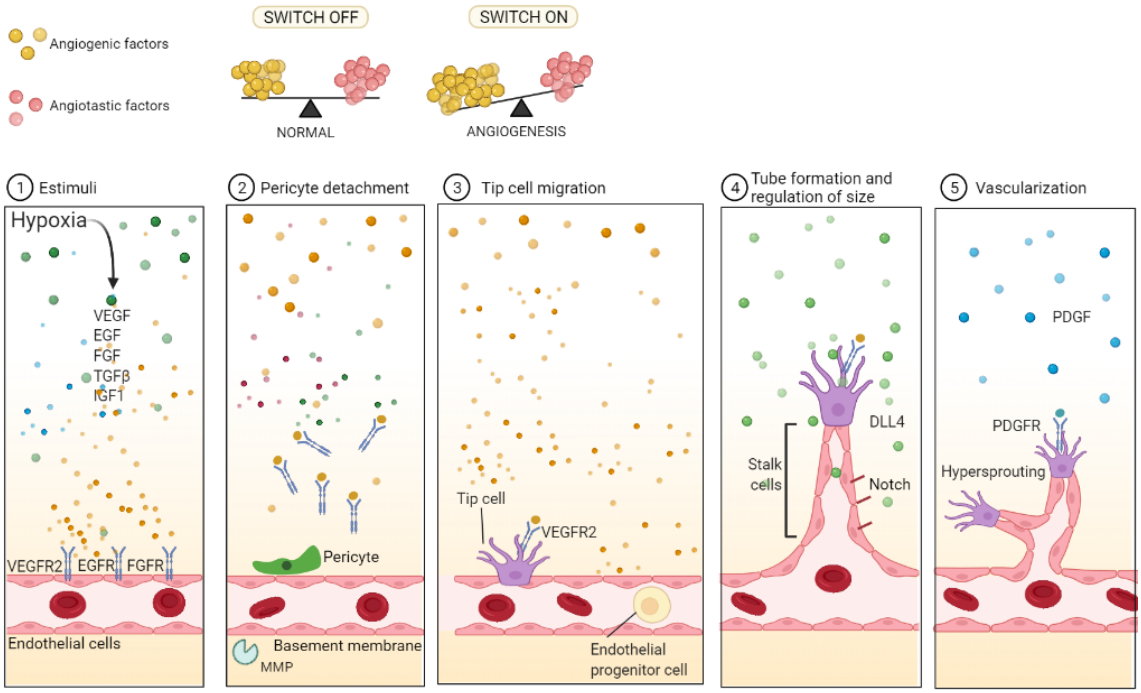


Figure 10. **Steps involved in angiogenesis during development**. **A**) Induction release of angiogenic factors, by hypoxia stimuli and upregulation of the MMP protease expression. **B**) This leads to basement membrane degradation and pericyte detachment. **C**) Tip cells (specialized EC) migrate along angiogenic factors and differentiate into proliferative stack cells, which will be the main body of the new vessel. **D**) The angiogenic factor VEFG activates DLL4 secretion, which binds to Notch-1 receptors, downregulating VEGFR and subsequently downregulating proliferation. **E**) Vascularisation occurs when PDGFbeta stimulates pericyte attachment and reduces VEGF sensitivity. The concept has been taken and modified from Cancer Research Product Guide Edition 3 (Bioscience, 2015).

The formation of blood vessels has been the subject of studies for many years. In 1948, Isaac Michaelson first described the presence of a diffusible factor associated with ischemic retina and neovascularization (Michaelson, 1948). Today we know that the activity of angiogenesis is controlled by angiogenic and angiostatic factors (Table 3) (Folkman et al., 1971; Hoeben et al., 2004). Although these factors have been identified as relevant players in many diseases, only a few of them have been associated with ischemic stroke (Chen et al., 2012). Therefore, the need for further investigation and identification of NAIS-specific roles of these endogenous factors in many pathways, such as Wnt signalling (Jun Olsen et al., 2017), VEGF signalling (Cross et al., 2003), and PI3k/Akt (P. R. Somanath et al., 2006).

Table 3. Angiogenic and angiostatic factors involved in angiogenesis

Angiogenic	Angiostatic
Tumour Angiogenesis Factor (TAF)	Angiostatin
Fibroblast Growth Factor (FGF)	Endostatin
Vascular Endothelial Growth Factor (VEGF)	Thrombospondin-1 (TSP-1)
Platelet-Derived Growth Factor (PDGF)	Angiopoietin 2
Transforming Growth Factors (TGF)	Pigment Epithelium-Derived Factor (PEDF)
Tumour Necrosis Factor (TNF)	Interferon
Placental Growth Factor (PIGF)	Kringle 5
Matrix Metalloproteinase (MMP)	Tissue Inhibitors of Metalloproteinase Inhibitors (TIMPS)
Plasminogen Activator (PA)	Plasminogen Activator Inhibitor (PAI)
Epidermal Growth Factor (EGF)	Platelet Factor-4 (PF4)
Hepatocyte Growth Factor (HGF)	Thrombocyte Factor 4 (TF4)
Granulocyte-Colony Stimulating Factor (G-CSF)	Transforming Growth Factor-Beta (TGF-β)
Connective Tissue Growth Factor (CTGF)	Retinoic Acid (RA)
Stromal Cell-Derived Factor-1 (SDF1-1)	Interferon-Alpha (IFN-α)
Hypoxia-Inducible Factor (HIF)	IFN-Gamma (IFN-γ)
Cyclooxygenases-2 (COX-2)	Il-4
Prostaglandin E_2 (PGE_2)	Il-12
Nitric Oxide (NO)	Il-13
Angiopoietins	Leukaemia Inhibitory Factor (LIF)
Angiotropin	Neuropilin 1
Integrins	Vascular Endothelial Growth Factor Receptor 1 (VEGFR-1)
Ephrins	Osteopontin (OPN)
Cadherins	Brain-Specific Angiogenesis Inhibitor 1 (BAI1)

3.2.2 Wnt Signalling Cascade

Wnt factors are involved in a wide variety of functions, and it is involved in angiogenesis (Miller, 2002). The name Wnt is a consensus hybrid created from the names Wingless and Int-1 or by the Wingless-related integration site 1. participates in at

least three cell signalling processes, the canonical Wnt/β-catenin, the two non-canonical Wnt/Ca^{2+}, and the Wnt/planar cell polarity pathway (Wnt/PCP) (Miller, 2002). These Wnt pathways are involved in the normal development and the pathogenesis of several diseases (Goodwin et al., 2002; Klein et al., 2009). Recent studies have shown that non-canonical pathways are present in vascular endothelial cells and related to angiogenesis (Figure 11A) in normal and pathological conditions through the regulators Fzd7 and R-spondin3 (Scholz et al., 2016; Tojais et al., 2014; Vanhollebeke et al., 2015). Additionally, the transcriptional regulation of the angiogenic factor VEFG by Wnt/β-catenin signalling has been demonstrated (Figure 11B) (Easwaran et al., 2003).

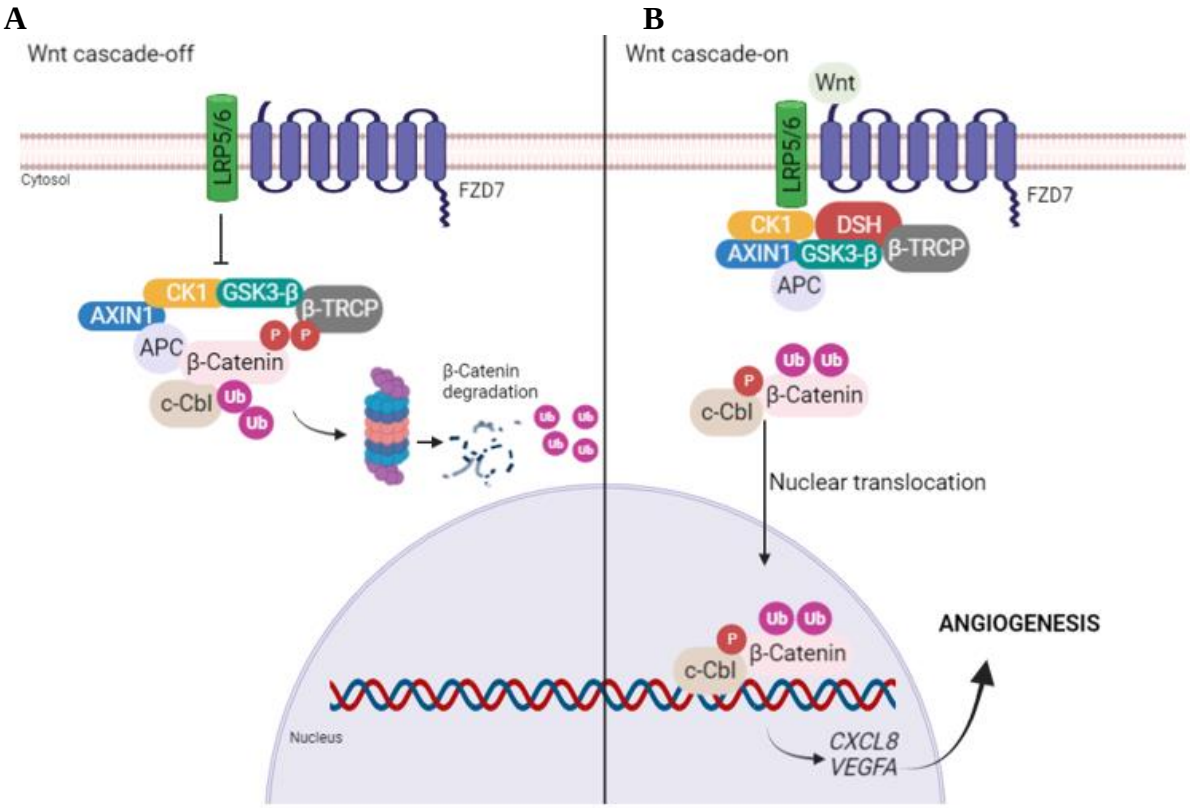


Figure 11. **Wnt signalling pathway**. Angiogenesis regulation through off and off Wnt signalling cascade. **A**) Wnt cascade-off showing c-Cbl regulating angiogenesis. Constitutively Wnt signalling is suppressed in endothelial cells. In the absence of a Wnt ligand, β-Catenin is downregulated by a cytosol complex consisting of Adenomatous Polyposis Coli (APoC). Phosphorylation of β-Catenin by GSK3β leads to its ubiquitination and proteasomal degradation by E3 ligases c-Cbl and β-TrCP. **B**) Wnt cascade-on showing the activation of the Frz receptors by Wnt ligand. This activation allows β-catenin to undergo nuclear translocation and activate target genes (*CXCL8* and *VEGFA*) to induce angiogenesis.

3.2.3 Vascular Endothelial Growth Factor Pathway

Vascular Endothelial Growth Factor (VEGF) comprehends seven members: VEGF-A, VEGF-B, VEGF-C, VEGF-D, VEGF-E, VEGF-F, and PLGF. These factors play a vital role in the development and functioning of the nervous and circulatory systems, hence their involvement in the pathophysiology of stroke (Greenberg et al., 2013). Moreover, they share a common VEGF homology domain and three receptors tyrosine kinase (VEGFR-1/2/3) (Hoeben et al., 2004).

Nowadays, there is much evidence to consider that VEGF and VEGFR-2 have a crucial role in the regulation of angiogenesis in both physiological and pathological settings. This notion is supported by the observation that VEGF is one of the most potent known agents regarding vascular permeability (Cross et al., 2003) and that VEGFR-2 is the first molecule expressed in mesodermal cells that gives rise to angioblast (Risau et al., 1995). Moreover, it is the main mediator in the response of endothelial cells (Waltenberger et al., 1994; Zachary, 1998).

Binding of VEGF to its receptor VEGFR-2 initiates a cascade of different interacting pathways (Figure 12) that up-regulates a set of genes related to vascular permeability, proliferation, survival, and migration of ECs. For example, the Ras pathway, which is activated after VEGFR-2 binds to phospholipase C (PLCγ1) (Takahashi et al., 2001), leads to activation of the mitogen-activated protein kinase (MAPK) pathway and the extracellular regulated kinase (Erk) pathway (Hoeben et al., 2004). Additionally, activation of the phosphatidylinositol 3'-kinase (PI3K)-Akt pathway occurs and leads to increased EC survival and cell migration signalling (Fulton et al., 1999). A deeper understanding of this cell signalling pathway is explained elsewhere (Hoeben et al., 2004).

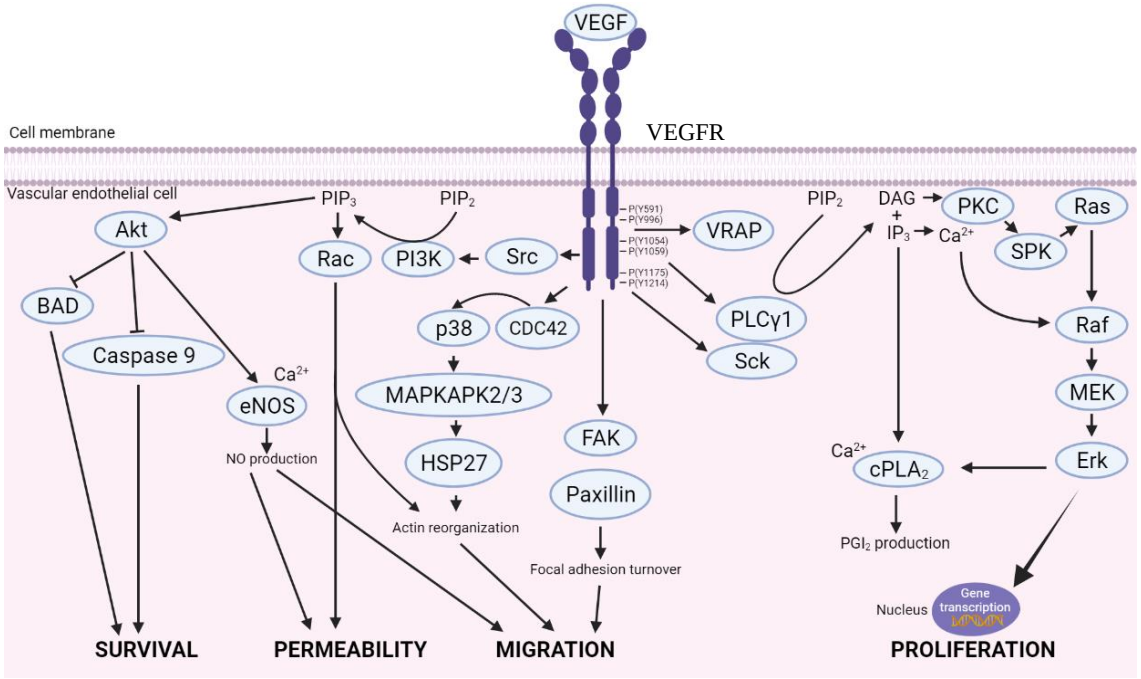


Figure 12. **VEGF-VEGFR signalling pathway**. The binding of VEGF to VEGFR-2 leads to a cascade of different signalling pathways that up-regulate genes involved in the mediation, proliferation, and migration of EC, promoting their vascular permeability and survival. VEFG/VEGFR2 binding allows dimerization of the receptor and further activation of the PLCγ-PKC-Raf kinase-MEK-Erk pathway resulting in EC proliferation due to initiation of DNA synthesis and cell growth. On the other hand, VEFG/VEGFR2 binding activates PI3K-Akt pathway, which leads to increased EC survival, permeability, and migration. Adapted and edited from Hoeben et al. (2004).

3.2.4 PI3K-Akt signalling

Reports of the PI3K-Akt signalling establish that it is a pathway with a versatile nature related to EC functions due to the numerous substrates that are phosphorylated within the pathway (Shiojima et al., 2002). The activation of protein kinase B (known as Akt) in this pathway is quick and regulates cellular processes related to EC survival, cell cycle, and proliferation (M. Lee et al., 2014). Akt is part of a family of serine-threonine protein kinases that share a highly similar N-terminal pleckstrin homology domain, a serine/threonine-specific kinase domain and a C-terminal regulatory domain. Stimulated cells express Akt protein in the cytoplasm and the two non-phosphorylated regulatory sites (threonine (Thr) at 308 and serine (Ser) at 473) (Shiojima et al., 2002) will subsequently be phosphorylated when lipid products of PI3K bind. This Thr308 and Ser473 phosphorylation occurs by PDK1 and PDK2, respectively (Hemmings, 1997)

and is a key component of many signalling pathways involved in normal and pathological processes (Downward, 1998).

In mammalian cells, there are three Akt isoforms referred as Akt1, 2, and 3 (also known as PKBα, β, and γ) and are encoded by different genes (Bellacosa et al., 2004). However, Akt1 is considered to have the greatest impact on EC due to its involvement in the VEGF signalling pathway (Zhuang et al., 2013), the impairment of endothelial nitric oxide synthase (eNOS), integrin activation, and cultured $Akt1^{-/-}$ EC proliferation (Ackah et al., 2005; Somanath et al., 2008). The regulation of eNOS activity by AKT happens after the phosphorylation of Ser1177 where the development of angiogenesis occurs in the endothelium and there is subsequent control of NO production and vasodilation (Lee et al., 2014; Somanath et al., 2006). Furthermore, Akt is vital for postnatal angiogenesis due to the many targets in its signalling cascade, including Bad (Rusiñol et al., 2004), Forkhead factors (FKHR) (Daly et al., 2004), caspase-9 (Alladina et al., 2005), IKKα (Bai et al., 2009), Mdm2 (Fang et al., 2005), Yes-associated protein (YAP) (Basu et al., 2003). All of them with important repercussions on pathophysiological processes.

3.2.5 Platelet regulation and function

Platelets are anucleate cells with an essential role in haemostasis and thrombi function formation in pathologic conditions. In recent years, strong evidence has emerged demonstrating the involvement of platelets, which have been critically involved in acute ischemic stroke. Recanalization of the occluded brain artery to reconstitute cerebral blood flow is the main goal in the treatment of adult stroke patients and has been implicated in the progression of infarct size (Cevik et al., 2016). Therefore, proteins related to the structure and function of platelets are considered future markers of ischemic stroke or therapeutic drug targets (Gawryś et al., 2022).

The molecular mechanisms that regulate platelet activation and function, and their role in arterial thrombosis, occur most frequently after exposure to potent thrombogenic elements in the bloodstream due to vascular endothelium rupture or erosion. The generated thrombi are mainly composed of aggregated platelets, which are adhered to the wall of the injured artery and other activated platelets (Jackson, 2011). During vascular injury, von Willebrand factor (vWF), fibrillar collagens, fibronectin, and

laminin become accessible to the blood. These proteins are part of the subendothelial matrix (Figure 13) and promote platelet adhesion through interactions with receptors; vWF with the GPIb-V-IX platelet complex, interacts with GPVI glycoprotein receptor and $\alpha_2\beta_1$ integrin, and fibronectin interacts with $\alpha_6\beta_1$ integrin (Moroi et al., 1996; Nieswandt et al., 2003).

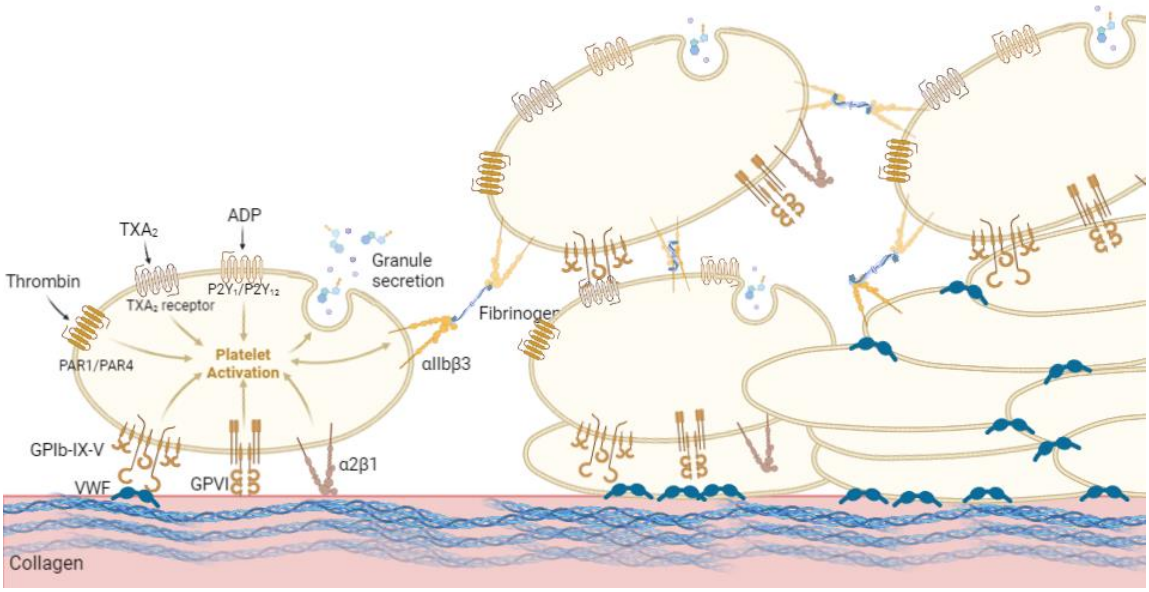


Figure 13. **Molecular mechanism in platelet adhesion and activation**. Injured vessel walls expose VWF in the sub-endothelium, and flowing blood platelets interacts via a rapid on-off specific interaction between the GPIb-V-IX platelet complex and the vWF collagen-bound. Platelet GPVI binding to fibrillar collagen, the interaction of α2β1 collagen, and the fibronectin engagement of α5β1 allow for stable platelet adhesion. This strong cell adhesion induces a series of biochemical changes that leads to the activation of αIIbβ3 and subsequently a high-affinity interaction with vWF, fibrinogen and fibronectin. Taken and modified from Jackson (2011).

A stable platelet-platelet adhesion formation is based on a second interactive step mediated by the binding of integrin $\alpha_{IIb}\beta_3$ to vWF, fibrinogen, fibrin and fibronectin (Bennett et al., 1979; Ruggeri et al., 1982). Once the platelet has adhered to the vessel wall or the surface of a thrombus, morphological and biochemical changes occur leading to platelet activation, where potent contents of platelet granules such as thromboxane A_2 (TxA_2) and dense-granule ADP, are released. Further molecular mechanisms involved in the interaction of platelet and thrombus are explained elsewhere (Jackson, 2011).

How endothelium regulates platelet function, as well as the proliferative state, contraction, and vasodilation of vascular smooth muscle cells (VSMCs), is through potent derived factors including nitric oxide (NO), prostacyclin (PGI_2), endothelin, TxA_2, prostaglandin H_2, and angiotensin II (Cosentino et al., 2005; Lüscher et al., 1993). Furthermore, the cyclooxygenase cascade is known to mediate endothelium-dependent contractions through the generation of superoxide anions ($O2^-$), which enhances NO degradation and stimulates VSMCs (Cosentino et al., 1994). Additionally, the endothelium can regulate the adhesiveness of platelets on the vessel wall through a strong negative charge and thrombin generation through the synthesis of NO, PGI_2 (Cheng et al., 2002), and CDD39 (Enjyoji et al., 1999).

3.2.6 Adhesion and Endothelium

Adhesion molecules, which are known to play a fundamental role in pathophysiological processes, including stroke (Papayianni et al., 2002), are considered candidates in vascular events. Intercellular adhesion molecule-1 (ICAM-1) and vascular cell adhesion molecule 1 (VCAM-1) are adhesion molecules normally expressed by ECs that play an important role in the pathophysiology of stroke (Simats et al., 2016). ICAM-1 promotes the apical endothelial cups assembly through activation of ARHGEF26/SGEF and RHOG, while VCAM-1 mediates cell-cell recognition and leukocyte adhesion through interaction with alpha-4/beta-1 integrin (ITGA4/ITGB1) on leukocytes mediating both adhesion and signal transduction. Studies have shown that during a stroke and other diseases such as cancer and cardiopathies, there is a high level of VCAM-1 in the plasma (Brondani et al., 2007; Papayianni et al., 2002). This corroborates the role of this molecule in directing the immune response to ischemia through the downstream signal transduction triggered after endothelium activation (Simats et al., 2016). Dysfunctionality and disruption of the molecular mechanisms involved in the endothelium impair the normal stimulus and responses of VSMCs, which alters the signalling pathways that regulate NO activity, cell adhesion, the inflammatory process and has an impact on the progression of ischemic strokes (Volpe et al., 1996).

3.2.7 Shear Stress and Blood Flow Dynamics

Whole blood is under constant and complex flow dynamics and the presence of bifurcations and junctions in the vascular network has a direct effect on the velocity profile of the blood (Zilberman-Rudenko et al., 2017). Considering that the low flow regions are prone to accumulation of blood cells and procoagulant proteases (Mackman, 2012), studies have shown that blood flow is a vital factor that intervenes in thrombus formation (Zilberman-Rudenko et al., 2017), especially at points of stagnation where acceleration and deceleration of fluids occur (Nesbitt et al., 2009; Westein et al., 2013).

The neonatal brain does not have the same self-regulatory capability as the adult brain, which means that it is more vulnerable to changes in blood pressure (Caplan, 2009). Additionally, the cerebral blood flow in preterm newborns is lower compared to term newborns (20 mL/100 g/minute vs. 50-60 mL/100 g/minute respectively) (Rorke et al., 1992). Once cerebral blood flow is low, platelet adhesion and thrombin generation have been demonstrated (Bark et al., 2013; Nesbitt et al., 2009). In addition, ECs modify their morphology and switch from an anticoagulant to a procoagulant phenotype (Zilberman-Rudenko et al., 2017). This evidence shows that the biorheology of blood flow is closely related to blood vessel maintenance (Tsai et al., 2012; Westein et al., 2013) and thus could have a potential role in the pathophysiology of NAIS.

3.2.8 Immune Response and Stroke

During a stroke, when blood flow is interrupted by thrombi, a lack of blood supply occurs with a subsequent severe reduction of glucose and oxygen in cerebral neurons, other supporting cells, and brain structure networks (El Amki et al., 2017). In the absence of these vital constituents, a series of metabolic and biochemical changes initiate at the cellular level, bringing massive cell death, which is responsible for the long-term impairments in motor skills, vision, speech, mood, cognitive, and behaviour in the infant.

The dying core, which is the area where this cell death occurs in the brain during ischemic stroke, is under constant apoptosis (Simats et al., 2016). This programmed cell death releases endogenous molecules that are signals recognized by microglia and inactivate the immune system (Liesz et al., 2015). Microglial cells are the local immune cells of the brain and exert diverse functions in physiological or pathological conditions;

various populations of microglia with different functional characteristics could be targeted for disease prevention or treatment after stroke. These immune cells are activated after alarming recognition and, in concordance with peripheral leukocytes, infiltrate the brain within a short time (Fumagalli et al., 2015).

The interactions between peripheral and local leukocytes produce an exacerbated release of pro-inflammatory cytokines that play a role in endothelial cells, upregulating the expression of leukocyte adhesion molecules, chemokines, and the complement (D'Ambrosio et al., 2001; Simats et al., 2016). Evidence shows that there is collateral damage due to immune cells infiltration and production of cytotoxic mediators, contributing to an increase in brain damage and other secondary complications that influence the outcome of stroke in the newborn (Lakhan et al., 2009). Furthermore, there has been a correlation between the extent of brain lesions and the strength of the neuroinflammatory response (Nishijima et al., 2016).

Aside from the adverse side effects of the pro-inflammatory reaction, the immune system has a second vital function in the suppression of inflammation (Simats et al., 2016). After triggering the pro-inflammatory network, this dual function (Figure 14) starts to execute an anti-inflammatory response by releasing the anti-inflammatory cytokines IL-4, IL-10, and TGF-β, which transform the M1 macrophages (pro-inflammatory macrophages) into M2 macrophages (anti-inflammatory macrophages) (Benakis et al., 2015). These M2 macrophages act as an antagonist of the inflammation process through the clearance of necrotic debris and both recovery and neuroregeneration of the ischemic brain through the production of neuroprotective molecules such as insulin-like growth factor and fibroblast growth factor (Shichita et al., 2014).

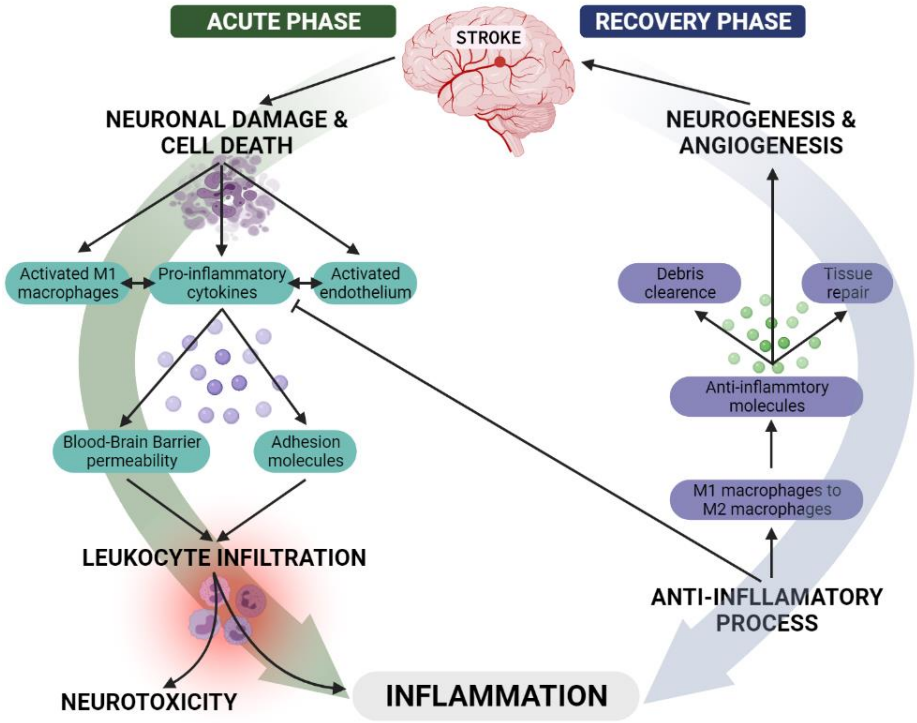


Figure 14. **Acute and recovery phases in inflammation response**. Natural role of the immune system response and its dual function during a stroke event. Taken and edited from Simats et al (2016).

It is important to understand that the consequence of a stroke is neurological, but the cause is vascular, and many reports show that any of the arteries that comprehend the cerebral vascular bed can be affected in newborns with ischemic strokes. Studies have shown that the left MCA is the most affected artery, accounting for more than 50% of unilateral cortical strokes in NAIS (Estan et al., 1997; Koelfen et al., 1995; Mercuri et al., 1999; Perlman et al., 1994). This may be caused by differences in the haemodynamic factors of the patent ductus arteriosus or the direct common carotid route (Sreenan et al., 2000). The rest of the arteries, including the internal carotid artery (ICA), ACA, PCA, and the anterior choroidal artery (AChA), comprise the remaining NAIS cases (Govaert et al., 2009a; Kirton et al., 2011; Ramaswamy et al., 2004).

Preterm infants are more likely to have multiple arteries involved during infarction, producing unilateral or bilateral multifocal stroke, whereas term infants tend to have main branch occlusion (Nelson, 2007). Multifocal stroke can be linked to several risk factors such as meningitis, embolism, thrombophilia and arteriopathy. Additionally, the thrombus generated during the event could break down into several emboli causing arterial spasms and ultimately leading to multiple separate strokes (Govaert et al.,

2009b). Neonatal air embolism is also likely to cause multiple infarcts because air enters the infant's vascular system, bypasses the carotid arteries, and embolizes multiple cerebral arteries (Sivan et al., 1990).

4. Risk Factors and Stroke

4.1 Neonatal Risk Factors

Several factors appear to be related to thrombosis or embolism of a major cerebral artery, indicating problems at different levels and times during late pregnancy, delivery, and the neonatal period (Harteman et al., 2012). Many efforts have been made to find a relationship between the pathophysiology of NAIS and several complex multifactorial risk factors. However, many clinicians suggest that the associations cannot be recognised and that the definitive cause of the stroke remains elusive (Arnaez et al., 2018a; Govaert et al., 2009b; Lehman et al., 2014; Mineyko et al., 2011; Raju et al., 2007).

A wide variety of risk factors for neonatal arterial ischemic stroke have been and continue to be studied. For example, family history, maternal factors, placental conditions, prothrombotic states, and the influence of genetics are important factors that allow researchers to comprehend and identify their mechanism of association with NAIS (Arnaez et al., 2018a). Additionally, sickle cell disease, cardioembolism, arterial dissection, Moya Moya syndrome, steno-occlusive cerebral arteriopathy, infection, Fabry disease, and mitochondrial disorders are also identified as potential risks factors for NAIS. However, these factors and the mechanisms involved are not mutually exclusive, and a combination of them may coexist and contribute to a stroke event (Arnaez et al., 2018a; Ichord, 2020).

Adequate descriptive, epidemiological, case-controlled, or prospective studies of risk factors will allow a definitive aetiology of the disease to be found, leading to a better understanding of the causal pathways and the possibility of calling for better recognition and treatment of risk/susceptibility of the stroke. The proposed risk factors in newborns (Table 4) have been selected from case series evidence and anecdotal case reports (Raju et al., 2007). On the other hand, the association of these factors with NAIS and the understanding of the mechanisms involved is not yet defined in many risk factors due to

the complexity of the phenotype and onset of the events (Caplan, 2009; Wraige E & Pohl KR, 2005).

Table 4. Neonatal risk factors that have been associated with the occurrence of NAIS

Group	Risk Factor
Heart and Vessels	Endothelial injury and/or inflammation from the meninges (Fitzgerald et al., 2007).
	Arteritis and⁄or phlebitis from within the meninges (Govaert, et al., 2009b).
	Left heart (exchange transfusion, Rashkind procedure, cardiac surgery, arrhythmia) (McQuillen et al., 2006).
	Torn basilar or vertebral artery; dissection of the intima of ICA or MCA by breech or instrumental delivery (Govaert et al., 1993; Lequin et al., 2004).
	Underneath fracture or dehiscence between skull bones (Govaert et al., 1993)
	Lobar, epidural or subdural haematoma due to trauma, with compression of one or both PCAs by the uncus (Remillard et al., 1974); possibly compression of basal vein
	Spasm and⁄or compression of arteries near a subdural or subarachnoid haematoma following difficult delivery (Govaert et al., 1993).
	In utero high spinal cord infarction with hyperextended breech presentation (Bhagwanani et al., 1973).
	Infantile arterial calcinosis (Van Der Sluis et al., 2006).
Thromboembolism	Embolism from or hypoperfusion of a canulated carotid artery (de Vries et al., 1997).
	Paradoxal embolism: Open foramen ovale (Scher et al., 1986), Umbilical vein (Ruff et al., 1979), portal vein (Parker et al., 2002), central venous line with thrombus (Pellicer et al., 1992), monochorionic twinning (Larroche et al., 1990), subclavian steal syndrome with ipsilateral brain infarction (Beattie et al., 2006), umbilical vein thrombosis after placenta or cord disorder (Redline et al., 2008).
	Common or internal carotid or vertebral artery thrombosis (Alfonso et al., 2001); anomaly of the aortic arch or common carotid artery
	Air embolism, documented in the brain (Sivan et al., 1990).
	Temporal artery catheter (Prian et al., 1978).
	Embolism during extracorporeal membrane oxygenation (ECMO) (Luisiri et al., 1988).
	PHACE síndrome (Prada et al., 2010).
	Carotid hypoplasia by drugs, toxins, or infectious agents (Afifi, 1987).
	Fibromuscular dysplasia, carotid elastin hyperplasia (Thompson et al., 1975).
	22q11 deletion (Curry et al., 2007).

Table 4. Continued		
Coagulation	Prothrombotic states	Alteration levels of protein C, factor V, protein S, homocysteine
		Dysfunctional haemostasis (Günther et al., 2000; Lynch et al., 2005).
	Genetics	*F5* Leiden. Heterozygous 5-10 times risk increased. Homozygous 50-100 times increased risk
		F2 c.*96C>T , *MTHFR* C677T, *VEGF* c.936C>T
		Anomaly of karyotype (Curry et al., 2007)
Infections		Rubella, CMV, toxoplasmosis (Govaert et al., 2009b), Parvovirus, Influenza A
		Meningitis, Encephalitis
		Brain abscess
		Septic shock, tamponade, acute massive blood loss associated with Disseminated Intravascular Coagulation (DIC) without meningitis (Smith et al., 1991).
Miscellaneous	Metabolic	Inborn error of metabolism (hyperhomocysteinaemia, primary hyperoxaluria, galactosialidosis, molybdenum cofactor deficiency, mitochondrial disease) (Govaert et al., 2009b).
		Incontinentia pigmenti (Maingay-de Groof et al., 2008).
		Polycythaemia (Amit et al., 1980).
		Nephrotic syndrome (Sran et al., 1988).
		Profound asphyxia with DIC and venous stasis in absence of other risk factors (Govaert et al., 2009b).
		Spasm induced by the presence of subarachnoid haemorrhage (Reynolds et al., 2007).
		Aicardi–Goutières syndrome (Govaert et al., 2009b).
	Drugs	Carotid hypoplasia (Afifi, 1987).
		Spasm (Reynolds et al., 2007).
	Hormones	High levels of endogenous testosterone (Golomb et al., 2009; Kirton et al., 2008).
Other		Apgar score of <7 (Harteman et al., 2012).
		Unknown or combined risk factors.

4.2 Family History and Maternal Factors

Family history of vascular disease is well studied in stroke in young and adults, but there is a gap regarding arterial stroke in infants. This is due to the difficulties in identifying the onset of the stroke (Arnaez et al., 2018a; Jerrard-Dunne et al., 2003; Schulz et al., 2004). Several potential maternal risk factors have been identified to be involved during a stroke of the newborn. If the number of risk factors of the pregnant mother increases, the risk of an infant suffering from NAIS also increases. (J. Lee et al., 2005). Numerous studies have associated several maternal risk factors with NAIS and the mechanisms involved in the pathophysiology of stroke (Table 5). Additionally, some of these factors have been independently associated with stroke in infants. For example, maternal smoking (Darmency-Stamboul et al., 2012), gestational hypertension and/or preeclampsia (Kirton et al., 2011), and maternal fever (Harteman et al., 2012). However, other researchers did not find any maternal risk factors associated with NAIS (Arnaez et al., 2018a).

Table 5. Maternal risk factors that have been associated with NAIS

Thrombophilia (James et al., 2006)
Autoimmune disorders
Drug abuse (Chasnoff et al., 1986)
Diabetes (James et al., 2006).
Infertility and its treatment
Prolonged rupture of membranes
Preeclampsia (Golomb et al., 2001)
Intrauterine growth retardation (Lehman et al., 2014)
Infections (Golomb et al., 2001)
Smoking (Darmency-Stamboul et al., 2012).
Intrapartum maternal fever (Harteman et al., 2012; Y. W. Wu et al., 2004)
Placental conditions (Lehman et al., 2014)

Abnormal placental conditions are of special interest as a risk factor implicated in NAIS. Studies evaluating placental pathology have determined that severe foetal chorioamnionitis, diffuse chorioamniotic hemosiderosis, and extensive avascular villi are independently associated with cerebral palsy and neurological impairment (Redline & O'Riordan, 2000). Placental infarcts and placental weight less than the 10^{th} percentile have also been identified as potential risk factors for NAIS (Laugesaar et al., 2007; Lehman et al., 2014). Additionally, there have been cases of infants with hemiplegic cerebral palsy presumably caused by placental thrombosis in mothers who were documented to have coagulation abnormalities (Kraus et al., 1999). Another study found

that placental abnormalities are a common risk factor in infants with NAIS compared to those with presumed perinatal NAIS (Laugesaar et al., 2007). However, in preterm newborns, one study concluded that there was no difference in placental pathology between cases and controls (Benders et al., 2007).

On the other hand, diabetes as another possible maternal risk factor associated with NAIS is due to the increased level of haemoglobin concentration in the blood and thus to increase in the size of the foetus, a condition that favours a difficult delivery (Darmency-Stamboul et al., 2012). However, there is not enough evidence to support the previous hypothesis. The association with NAIS of these complex risk factors, mechanisms and interactions, has insufficient evidence even in large multinational studies where true causality is lacking in most cases (Golomb et al., 2001; Kirton et al., 2011).

4.3 Infection

Newborns exposed to early-onset infections such as sepsis, meningitis, ventriculitis, brain abscess, maternal fever during labour, and chorioamnionitis can lead to ischemic stroke (Harteman et al., 2012). For example, chorioamnionitis, which is an infection of the placental membranes, may cause direct thromboembolism in the placenta that could release an embolus with access to the foetal brain circulation. It is important to remark that the clinical diagnosis of chorioamnionitis has high rates of false positives and false negatives (Redline et al., 2000). In addition, some cases of meningitis caused by group B Streptococcus have been reported as a cause of stroke in infants due to a focal inflammatory response and thrombosis in the central nervous system (Fitzgerald et al., 2007; Ment et al., 1986).

During the early stages of infection, thrombi formation occurs and there are two main mechanisms involved in the pathophysiology of arterial ischemic stroke. First, there is a shift in the coagulation system that leads to a prothrombotic state due to the release of many pro-inflammatory cytokines, complement components, decreased anticoagulant proteins, and platelet activation (Redline et al., 2008), and second, a direct injury to the endothelium that initiates endothelial activation, platelet adhesion and subsequently, the development of thrombi (Harteman et al., 2012). Another study supports this theory

where infections and inflammatory conditions are risk factors to consider in NAIS (James et al., 2006).

4.4 Foetal distress

The Apgar score, established by Dr. Virginia Apgar in 1952, is a system that works as a convenient method to assess the clinical status of the newborn immediately after birth and establishes whether there is a need for resuscitation (American Academy of Pediatrics Committee on Fetus and Newborn, 2015). The score comprises five essential components: colour, heart rate, reflexes, muscle tone, and breathing. Each of these criteria has a given score of zero, one or two, and quantifies clinical signs of neonatal depression such as apnoea, hypotonia, gasping for breath, and bradycardia, among others. Reports determine that a 5-minute Apgar score of 7-10 is classified as reassuring, 4-6 as moderate, and a score of 0-3 as low in the term and late-preterm infant (Neonatal encephalopathy and neurologic outcome, 2014).

Adverse events during childbirth have been studied due to the important effects they could have on the newborn. One study determined that a 5 min Apgar score of <7 probably shows that these adverse events are occurring and playing a fundamental role in foetal distress and the generation of subsequent hypoxia-ischemia (Harteman et al., 2012). Another study in mice found that hypoxic-ischemic was a strong and rapid stimulus for involuntary coagulation (Adhami et al., 2006). The consequence of having hypoxia-ischemia is a hypercoagulability state due to the decrease of physiological inhibitors of coagulation, including antithrombin III, protein C, protein S, and other mechanisms related to the pathophysiology of NAIS (El Beshlawy et al., 2004; Nowak-Göttl et al., 1997). However, the study could not answer whether the aforementioned Apgar score would be involved in the causal pathway.

4.5 Thrombophilia and thrombotic states

Thrombophilia is a state of hypercoagulability of blood constituents that can be caused by factors that influence thrombus formation such as injury to the vessel wall, stasis of blood flow (Nelson et al., 2004), and the triad of hypercoagulability. This triad described in 1856 by Virchow, emphasizes vascular stasis and vascular trauma as the harbingers of vascular thrombosis (Kumar et al., 2010). In NAIS, a hypercoagulable

state describes a role in which it confers a propensity to develop clot formation and subsequently arterial obstruction (Nakashima et al., 2014).

Thrombophilia has been considered for several decades as a risk factor for NAIS due to its mechanisms of affectation during pregnancy, the puerperium, and its contribution to complications of cerebral infarcts in infants (James et al., 2006; Mineyko et al., 2011; Nelson, 2007). For normal blood flow, a balance between pro-coagulant and anti-coagulant factors is essential, especially for the coagulation status of the infant (Simats et al., 2016). Activators and inhibitors, the functional properties of their products and their final regulation and degradation, are factors that influence coagulation homeostasis. These dynamic interactions are constantly changing within gestational and postnatal ages, and coagulation is triggered in both the mother and the foetus (Darmency-Stamboul et al., 2012; Suarez et al., 1988). Nevertheless, deregulation of these processes or a disrupted interaction can generate a pathological state in which thromboembolisms initiate (Byrnes et al., 2017; Esmon, 2009).

Although fifteen years ago it was proposed to evaluate states of hypercoagulability in plasma and DNA in an infant with stroke and its mother (Raju et al., 2007), today this proposal cannot continue. Scientific evidence has shown that some prothrombotic polymorphisms, such as the coagulation factor V G1691A, factor II G20210A and methylenetetrahydrofolate reductase C677T, do not play a major role in symptomatic NAIS (Arnaez et al., 2018b). However, these results are more controversial in neonatal cerebral sinovenous thrombosis (Garrido-Barbero et al., 2019). Many attribute the interpretation of these results to a consequence of the complexity of the phenotype (Hassan et al., 2000; Nowak-Göttl et al., 1999; Pfeiffer et al., 2019). For the identification and association of possible candidate genes with cerebrovascular diseases, genetic research is playing an important role by functioning as a group of hopeful approaches in individuals who have suffered a stroke. (Collins, 2011).

5. Genetics and Stroke

5.1 Genetic Influence

Genetic studies in recent decades have helped to reveal and identify loci that predispose to various human diseases and traits, such as loci on chromosome 5q12 loci and its association with stroke susceptibility in adults (Gretarsdottir et al., 2002). Additionally, this information allows comprehending the impact of genetic variants on the pathophysiology of the disease and the interaction in important biological pathways (Zuk et al., 2014). Thus, the potential use of common and rare variant studies for the identification of candidate genes and association with phenotypes.

High-throughput sequencing, also known as next-generation sequencing (NGS), describes technologies that sequence DNA and RNA rapidly and cost-effectively. These techniques allow large-scale collaborations that help the search for new genetic variants and subsequent validation and association with biological pathways of stroke (Falcone et al., 2014). This approach has worked with complex traits including lipid metabolism (Teslovich et al., 2010), diabetes (Qi et al., 2012), coronary artery disease (Prins et al., 2012), schizophrenia (Ripke et al., 2013), Parkinson's disease (J. F. Guo et al., 2018), and epilepsy (Perucca et al., 2020).

The use of NGS technology has increased over the last decade, specifically Whole-Exome Sequencing (WES) (Figure 15). WES is a technique widely used that involves the determination of the nucleotide sequence of the protein-coding regions of an individual's genome. These coding regions are named Exome, which only represent around 2% of the genome but contains approximately 85% of the variants identified and associated with diseases (van Dijk et al., 2014). Regarding stroke, it has provided benefits in the association with human monogenic disorders (Table 6). Thanks to many epidemiological and animal studies, supported by genome-wide associations (GWAS) and family analyses, there is strong evidence for genetic influence on the pathogenesis of multifactorial ischaemic stroke events in children and young adults (Falcone et al., 2014; Günther et al., 2000; Hamzi et al., 2011; Ikram et al., 2009; Junker et al., 1999). Additionally, individual investigators and larger organizations such as the International Stroke Genetics Consortium have helped increase the pace of research to accelerate the identification of stroke-related genes. For example, the METASTROKE study was conducted in 2012 (Traylor et al., 2012).

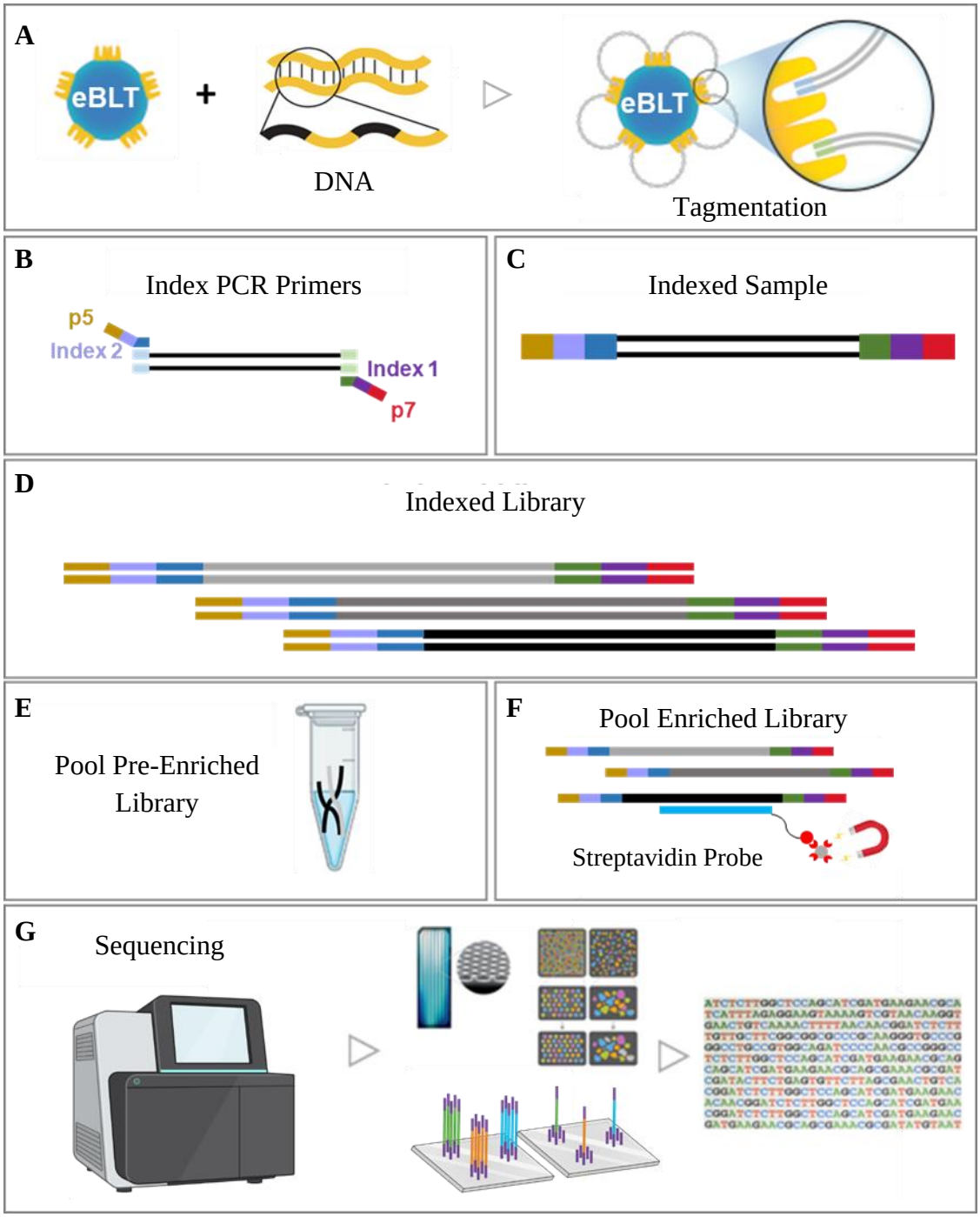


Figure 15. **Whole-Exome Sequencing workflow with DNA enrichment**. **A-F**) Pool library preparation. **A**) Enrichment Bead-Linked Transposomes (eBLT) used for tagmentation (tag and fragment) DNA with adapter sequences. **B**) Amplification step of the tagmented DNA adds prepared adapter indexes i72 and i5 for sequencing cluster generation. **C**) Indexed sample ready for pooling and enrichment. **D**) Example of indexed sample libraries. **E**) Pool pre-enriched library ready for sequencing. **F**) Streptavidin hybridizing probes for DNA capture and enrichment. **G**) NextSeq 500 sequencing platform and flow cell where sequencing occurs.

Table 6. Human single-gene disorders associated with stroke

Stroke mechanism	Disease/trait	
Cardioembolic	Cardiomyopathies	Primary or secondary
	Familial atrial myxoma	
	Familial dysrhythmias	
Thromboembolic	Metabolic	Homocysteinuria
		Dyslipidaemias
	Hemoglobinopathies	Sickle cell disease
	Prothrombotic states	*F5* Leiden
		F2 G20210A
		MTHFR C677T
		Antithrombin III deficiency
		Protein C deficiency
		Protein S deficiency
Small vessel disease	CADASIL	
	Fabry disease	
Mitochondrial disorder	MELAS	
Arterial dissection	Marfan syndrome	
	Ehlers–Danlos syndrome	
Channelopathies	Familial hemiplegic migraine	
Other	Familial amyloid angiopathy	

5.2 Identification of Candidate Genes

The genetic influence on the aetiology of ischemic stroke likely affects different signalling pathways involved in the pathophysiology of NAIS. A minority of stroke cases could be explained by rare monogenic Mendelian disorders, but most cases have a multifactorial aetiology and genetic influence remains unidentified (Sharma et al., 2013). However, even if a polygenic theory is considered, an isolated gene may still be relevant due to the capacity to interact with several factors involved in the stroke event. For instance, researchers identified cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL) as a cause of familial subcortical stroke due to mutations in the *NOTCH3* gene (Sharma et al., 2013). Other studies that applied the GWAS strategy concluded that *NINJ2*, *NOS3*, *COL4A1* and *DYRK1A* genes are associated with an increased risk of stroke (Ikram et al., 2009; Malik et al., 2018). Furthermore, in NAIS it is important to consider genes not included in the workflow analyses that might have the ability to modulate and influence stroke pathophysiology through synergistic interactions.

The approach of identifying a candidate genetic factor in NAIS, even with current sequencing technology, is a challenge because most cases are multifactorial and complex. In addition, sometimes classical patterns of inheritance cannot be demonstrated due to the nature of the genetic factors involved at the molecular level (Zuk et al., 2014). Therefore, assignment of probable causality to a new genetic variant identified by WES, Whole-Genome Sequencing (WGS) or other approaches, depends on many filtering aspects, such as inheritance filtering, which plays an important role for strong genetic support (MacArthur et al., 2014).

Furthermore, the application of an inheritance model in NAIS works as a useful way to reduce the numbers of variants to be analysed downstream. For example, in an autosomal dominant mode of inheritance, all affected individuals carry a candidate gene with a heterozygous variant, whilst in an autosomal recessive mode of inheritance, the pattern could be a biallelic compound heterozygous or homozygous variant. Additionally, a *de novo* inheritance could be applied when investigating a disorder that occurs due to a spontaneously arising variant, where the allele in the affected individual is not a private mutation.

In NAIS, regardless of the model applied for the identification and association of candidate gene, when there is no clear candidate gene, we must recognise some aspects that could be affecting the analysis such as (i) the early onset of the stroke event, which makes difficult the genetic comparison among patients; (ii) different stroke subtypes which could reflect a different genetic influence; (iii) genetic heterogeneity where variants in several genes could result in similar phenotypes; (iv) the presence of phenocopies; (v) variable penetrance caused by gene dose, gene-environment, and epistatic interaction; (vi) cofounders due to the presence of risk factors not previously assessed in the patient (Dashti et al., 2017).

Identifying the disease-causing mutation is not an easy task. The main reason is that a typical human individual, compared to the reference human genome, differs by around 4.1-5.0 million sites (Auton et al., 2015). To excel in this gene-disease association, experts must rely on and crosscheck several databases of variants and *in silico* predictors of pathogenicity. In addition to this, the use of bioinformatics algorithms is fundamental, which can process the vast information that raw data can generate and detect variants of great interest (Ehrhart et al., 2021).

5.3 Genetic Polymorphisms and Stroke

A genetic polymorphism could be defined as a difference in the DNA sequence between individuals, groups, or populations, where the minor allele has a frequency ≥1%. These genetic variations provide the key elements for evolutionary changes and processes including single nucleotide polymorphisms (SNPs), tandem repeats, recombination, copy number variations (CNV), deletions, and insertions (Traylor et al., 2015).

Congenital thrombophilic states caused by different polymorphisms have been associated with many diseases such as stroke. For example, the coagulation *F5* Leiden, *F2* c.*96C>T, the homozygous *MTHFR* polymorphism c.665C>T, and the antithrombin III, have been identified in young adults as a genetic risk factor for stroke, but its role in the NAIS pathophysiology remains unclear (Günther et al., 2000; Ioannidis et al., 2004). Another study that analysed the risk related to family history, maternal diseases, and common genetic variants of coagulation factors, determined that neither the genetic influence of these polymorphisms nor familial predisposition to thrombotic events were associated with NAIS (Arnaez et al., 2018a). However, they found a tendency for the coexistence of maternal relatives with a history of thromboembolic events in the NAIS group compared to controls. On the other hand, it has been proposed that the influence that genetics has on NAIS mechanisms and pathways probably differs between ischemic stroke subtypes (Schulz et al., 2004) and a study based on subsets of clinical criteria would be recommended to increase the power of the gene-disease association (Traylor et al., 2015).

5.3.1 Factor V Leiden

The factor V protein encoded by the *F5* gene, which is mapped to 1q24.2, is an essential cofactor of the blood coagulation cascade and is converted to its active form during coagulation by the protein thrombin. Factor V Leiden (FVL) is a polymorphism (rs6025) of human coagulation factor V, consisting of a p.Gln534Arg/c.1601G>A change and is the most common hypercoagulability genetic polymorphism inherited among Europeans (Gregg et al., 1997) accounting for 40-50% of hereditary thrombophilia (Kujovich, 2011). FVL is named after the Dutch city of Leiden, where it was first identified by Rogier Maria Bertina in 1994. Carrying this variant causes a defect in the downregulation of thrombin formation due to a poor anticoagulant

response to the activated protein C (APC), which is an anticoagulant protein in charge of cleaving and inactivating procoagulant factor Va and VIIIa (Dahlbäck, 2008). Despite that, some investigators speculate that the polymorphism confers some advantage on the heterozygous state, reducing mortality from bleeding associated with childbirth (Dahlbäck, 2008), trauma (Zivelin et al., 1997), cardiac surgery (Donahue et al., 2003), and haemophiliacs (D. H. Lee et al., 2000). Although, this speculation of evolutionary benefit has not been fully confirmed.

The role of this polymorphism in ischemic stroke varies. It may participate in the pathophysiology of stroke in certain subgroups of individuals and populations (Kujovich, 2011) or it may interact with other genes and risk factors associated with stroke. For example, many studies of unselected adult populations found no association between ischemic stroke and FVL (Cushman et al., 1998; Lalouschek et al., 2005; Marcucci et al., 2005) and no significant association was found in one study that focused on NAIS and its possible association with different genetic risk factors (Arnaez et al., 2018b). However, it has been demonstrated that this genetic variant in individuals younger than 50 years has a 3-fold increased risk factor for a stroke event (Margaglione et al., 1999), and the risk is higher among women of the same age (Aznar et al., 2004).

The clinical manifestation of FVL is variable. It is predisposed by coexisting genetic background, acquired thrombophilic disorders, and risk factors, including obesity (Pomp et al., 2007), oral contraceptives (Legnani et al., 2002), hormone replacement therapy (O. Wu et al., 2005), prothrombin 20210 G>A (Emmerich et al., 2001), and hyperhomocysteinemia (Ridker et al., 1997a). Its primary clinical manifestation is venous thromboembolism (VTE) and individuals carrying this polymorphism may never develop VTE or may experience their first thrombotic event in adulthood (Kujovich, 2011). It all depends on whether the individual is homozygous or heterozygous for FVL (Zoller et al., 1994).

The presence of the polymorphisms has been controversial when associated with other thrombotic events including arterial thrombosis and stroke. Information is debatable due to conflicting results and some studies finding no association between having the polymorphism and a higher risk of stroke (Cushman et al., 1998; Linnemann et al., 2008). However, a large meta-analysis showed that it confers a moderate risk of myocardial infarction and coronary diseases (Ye et al., 2006). Furthermore, additional

evidence shows that it is not a major risk factor for arterial thrombosis and that its contribution to complex diseases is small (Kujovich, 2011).

5.3.2 Coagulation Factor II

Concomitantly with FVL, the coagulation factor II polymorphism (F2P) commonly referred to as G20210A (rs1799963), is the second most frequent genetic risk factor studied in venous thromboembolism and prothrombotic states (Emmerich et al., 2001). This variant is in the *F2* gene, location 11p11.2, and the DNA change is at position 96 of the 3'-untranslated region (c.*96C>T). Prothrombin, which is the protein encoded by this gene, is cleaved proteolytically in several steps to form thrombin, the activated form. It plays an important role in thrombosis and haemostasis by converting fibrinogen to fibrin during blood clot formation, stimulating platelet aggregation, and activating factors V, VII, VIII, XIII, when complexed with thrombomodulin and protein C.

When considered a genetic risk factor, F2P is a common but mild risk factor for thrombosis (Ferraresi et al., 1997; Poort et al., 1996). However, a different effect of F2P has been found between men and women in which there is a greater risk in the group of men (Lalouschek et al., 2005). The mechanisms that might be involved in this different effect are unclear, but some say it may be related to protein and hormone levels as seen with FVL (Henkens et al., 1995; Ossei-Gerning et al., 1988). Another study focused on childhood thrombophilia and genetic risk factors showed that the heterozygous F2P was more prevalent in patients compared to the control group (4.2% *vs*. 1.1%) (Nowak-Göttl et al., 1999, Junker et al., 1998), showing the importance of this variant and the conferred risk of infantile thrombosis. Furthermore, data are suggesting that the combination of F2P and FVL increases the risk of thrombosis and that F2P alone tends to increase the risk of cryptogenic stroke in adults (Aznar et al., 2004; Junker et al., 1999). In contrast, the association with NAIS is not clear and genetic studies have concluded that there is no association between them (Arnaez et al, 2018b).

5.3.3 Methylenetetrahydrofolate Reductase

The predisposition to arteriothrombotic events, such as stroke, is a consequence of many factors, including acquired or inherited hypercoagulable states (Green, 2003; Ng et al., 2011). For example, hyperhomocysteinaemia (hHcy), which is defined as a medical condition characterized by homocysteine levels above 15 μmol/L in the blood (H. Guo et al., 2009), is one of those risk factors associated with cardiovascular diseases (Annus et al., 2020). A strong predisposition to hHcy is the presence of polymorphisms in the methylenetetrahydrofolate reductase (*MTHFR)* gene. *MTHFR* is mapped to 1p36.22 (B. O. Choi et al., 2003), and encodes the protein of the same name, which is a vital enzyme in folate metabolisms and a key determinant of the methylation cycle (Ganguly et al., 2015). Additionally, MTHFR catalyzes the conversion of N^5,N^{10}-methylenetetrahydrofolate to N^5-methyltetrahydrofolate, a cosubstrate required for homocysteine remethylation to methionine (Manolescu et al., 2010)(Manolescu et al., 2010)(Manolescu et al., 2010)(Manolescu et al., 2010)(Manolescu et al., 2010)(Manolescu et al., 2010)(Manolescu et al., 2010)(Manolescu et al., 2010).

The increase in plasma homocysteine levels in affected individuals, is caused by the genetic alteration of the *MTHFR* gene and generates a reduction in the conversion of the N^5,N^{10}-methylene tetrahydrofolate molecule to N^5-methyl tetrahydrofolate (Somarajan et al., 2011). This is the case, for example, of the two most studied polymorphisms associated with hHcy and arteriothrombotic events, the commonly known C677T (rs1801133) and the A1298C (rs1801131). The aftermath of having the *MTHFR* rs1701133 in heterozygosity (CT genotype) is a 30% reduction in enzyme activity, while homozygocity for rs1801133 and rs1801131 (TT and CC genotype respectively) accounts for a 60-70% reduction in enzyme activity (Hiraoka et al., 2004; Qin et al., 2020; van der Put et al., 1998).

The role of hHcy and stroke is unclear. Many authors doubt the relationship among hHcy, stroke, and other cardiovascular diseases (Faeh et al., 2006). However, the mechanisms involved in the formation of cardiovascular diseases are well explored. For example, hHcy increases vascular smooth muscle cell proliferation and collagen synthesis, further damaging the elastic material of the arterial wall, altering endothelial functions, and causing oxidative damage (Ganguly et al., 2015). Additionally, a study of the homocysteine effect on VSMCs showed that there is induction of mRNA and protein expression of c-reactive protein (CRP) both *in vitro* and *in vivo* (Pang et al.,

2014). This suggests the capacity to induce an inflammatory response in VSMCs through the production and mediation of CRP through the NMDAr-ROS-ERK1/2/p38-NF-κB signalling pathway.

The association between these two *MTHFR* polymorphisms and stroke remains debatable as different studies found association with ischemic stroke while others did not (Dong et al., 2021; Jiménez-González et al., 2021; Lalouschek et al., 1999; Pezzini et al., 2002; Soriente et al., 1998). Concomitantly, when studying the homozygous *MTHFR* rs180133 and NAIS, other studies have found no association (Arnaez et al, 2018a; Arnaez et al., 2018b).

5.3.4 Protein C and Protein S deficit

Protein C (PC), encoded by the *PROC* gene, is a 62-kD vitamin K-dependent plasma glycoprotein, which is a key factor in the anticoagulation system. Binding of the thrombin-thrombomodulin complex on endothelial cells cleaves and activates the protein C. Activated protein C (APC) functions as a serine protease that degrades the activated forms of factors V and VIII (Majid et al., 2020). For PC to convert to APC, the presence of nonenzymatic cofactor protein S (PS) is required, another vitamin K-dependent plasma protein that acts as a blood clotting inhibitor. PS is encoded by the *PROS1* gene and in plasma, about 40% is present as the active free form of the protein, whilst the remaining 60%, is present as the inactive form bound to C4BPA protein (Dahlbäck, 1991; Dahlbäck et al., 1981).

Both PC and PS act as natural anticoagulants that have been associated with autosomal dominant or recessive thrombophilia in presence of *PROC* and/or *PROS1* variants. Heterozygous PC deficiency is characterized by a reduced amount of protein (type I deficiency) or normal amounts with functionally defective protein (type II deficiency) (Bertina et al., 1984). On the other hand, the heterozygous PS is subclassified into type I (low levels of both free PS and total PS with decreased activity), type II (normal plasma PS levels with decreased activity), and type III (low levels of free PS and decreased activity with normal levels of total PS) (Bertina, 1990).

Autosomal recessive deficiency of both proteins may be the result of homozygous or compound heterozygous variants. Recessive PC deficiency can manifest as a severe neonatal disorder or a milder late-onset thrombophilia (Millar et al., 2000) whereas

recessive PS manifests as a very rare and severe hematologic disorder causing thrombosis and secondary haemorrhages (Fischer et al., 2010). Additionally, associations of stroke with deficiency of these two proteins have been investigated. For example, two studies found no significant association between isolated PC deficiency with stroke in adults (Hankey et al., 2001; Zakai et al., 2018) whereas other authors have speculated on this association (Camerlingo et al., 1991; Millar et al., 2000; Ueki et al., 2020). Concomitantly, low PS levels have been associated with recurrent ischemic stroke in a 16-year-old girl (Hooda et al., 2009) and with three individuals carrying a missense variant in the *PROS1* gene who suffered an ischemic stroke in their forties (Leung et al., 2010). The study of the mechanisms underlying the associations of PC and PS with stroke is vital to improve our knowledge of cerebral infarction.

5.3.5 APOE, F13A1 and ITGB3

Other polymorphisms have been highlighted for their possible role as risk or protective factors in ischemic stroke, including the *APOE* ε4 allele (rs429358), *APOE* ε2 allele (rs7412), (rs5985) and *ITGB3* (rs5918) (Konialis et al., 2016; Martijn Pruissen et al., 2008). However, there has been some confusion regarding the precise role of these polymorphisms in ischemic stroke in adults due to not enough statistically significant results. Additionally, there are not enough data where these polymorphisms have been studied in newborns with arterial ischemic stroke.

APOE gene, mapped to 19q13.32, exhibits three major alleles ε2, ε3, ε4. The ε3/3 genotype is represented in approximately 60% of the population, followed by the ε3/4 (22%), ε2/3 (12%), ε4/4 (3%), ε2/4 (2%), and ε2/2 (1%) (A. L. Gould et al., 1998). *APOE* encodes apolipoprotein E, which is the major chylomicron apoprotein and is essential for metabolism of triglycerides-rich lipoprotein constituents. Many researchers have studied the heterogeneity of *APOE* and stroke, but the results are debatable (Souza et al., 2003). For example, some studies demonstrated a significantly higher frequency of the *APOE* ε4 allele in patients with a family history of ischemic stroke (McCarron et al., 1999; Pedro-Botet et al., 1992; Peng et al., 1999), whilst other studies did not (Catto et al., 2000; Ferrucci et al., 1997; Souza et al., 2003). Moreover, these investigations have been conducted only on adults and not on other age groups.

Regarding *F13A1* gene, it has a cytogenetic location on 6p25.1 and encodes the A subunit of coagulation factor XIII. It is the last zymogen to be activated in the blood coagulation cascade for further fibrin clot stabilization and is involved in fibrin cross-linking by transglutaminase reactions (Martijn Pruissen et al., 2008). A study focused on the p.Val35Leu variant (rs5985) showed a higher rate of activation of factor XIII which alters the fibrin clot structure and results in an association with a lower risk of ischemic stroke (Anwar et al., 1999). On the other hand, the p.Tyr205Phe variant (rs3024477), has been associated with low levels of factor XIII activity (Anwar et al., 1999) and low levels in plasma (Gallivan et al., 1999). To some extent, there is an increased risk of ischemic stroke in young women (Martijn Pruissen et al., 2008; Reiner et al., 2002). As with the *APOE* gene and its alleles, there are no studies to date that reflect the *F13A1* variants and their role in neonatal ischemic stroke.

Another gene that could be studied in NAIS is *ITGB3*. It is located on chromosome 17q21.32 and encodes Platelet glycoprotein IIb/IIIa (GpIIb-IIIa) protein, a platelet membrane receptor, and a member of the integrin family. When GpIIb-IIIa is activated, it acts as a receptor for fibrinogen and von Willebrand factor, promoting platelet aggregation and thrombosis (Phillips et al., 1988), thus an alteration in the protein could affect the normal processes of clot formation. The p.Leu59Pro (rs5918) polymorphism in this gene has been identified as a risk factor associated with the pathogenesis of acute coronary syndromes (Weiss et al., 1996). However, its relationship with stroke has not been well established and some studies state that there is no association (Carllsson et al., 1997; Ridker et al., 1997b). In addition, another study in which this polymorphism was analysed in young women with cerebral infarction was not conclusive (Wagner et al., 1998).

6. Rare *De novo* Variants and Disease

In recent years, medical genetics has focused on studies of inherited variations, including GWAS for complex diseases and the family-based studies for Mendelian diseases. However, thanks to advances in technology and genetics, researchers can focus on *de novo* variants (DNV) to study mutational processes in many diseases (Veltman et al., 2012). DNV are variants that appear in an individual and are not present in their parents. They represent an extreme form of rare genetic variation and tend not to

be described in the population frequency databases or have an allele frequency less than <0.01 (Eyre-Walker et al., 2007). Therefore, DNVs are the main culprit in many diseases that occur irregularly, which makes it possible to explain their occurrence and the aetiology mechanisms involved in the disease.

Currently, thanks to genome-wide NGS studies, we know that the germline *de novo* SNV mutation rate in humans is 1.0 to 1.8×10^{-8} per nucleotide per generation and with significant variation between families (Rahbari et al., 2016; Roach et al., 2010). This rate can be translated to 44-82 *de novo* SNV in an individual's genome, with one or two variants affecting the coding sequence (Francioli et al., 2015; Gilissen et al., 2014; Kong et al., 2012). This overall rate of DNV might be higher in individuals with genetic diseases (Veltman et al., 2012). Furthermore, it has been proven that DNVs have a paternal relationship in which, the higher the paternal age at conception, the greater the possibility of *de novo* variants in the offspring (Devlin et al., 2012; Kong et al., 2012).

A clear example of the important role of NGS in the identification of pathogenic DNV is the successful application of parent-child WES-trio or multiple family members. As is the case with pathogenic DNV in cases of anencephaly (Wang et al., 2019) and many neurodevelopmental disorders such as intellectual disability, autism spectrum disorder, and schizophrenia (Devlin et al., 2012; Girard et al., 2011; Rauch et al., 2012; Yates et al., 2017). Regarding cerebrovascular diseases and NGS, there has been identification and association of DNV with diseases that lead to or predispose to stroke event. These DNV associations include *RNF213*, *CHD4*, *CNOT3*, and *SETD5* genes with Moya-Moya disease (Harel et al., 2015; Pinard et al., 2020, 2021), p.Leu1692Gln in *CACNA1A* with early stroke and intractable epilepsy in a 4-year-old girl (Gudenkauf et al., 2020), *GLA* and p.Phe229Leufs*11 associated with Fabry Disease presented with young stroke (Ling-Chih et al., 2019), and DNVs in the *NOTCH3* gene causing CADASIL (Stojanov et al., 2014).

In sequencing studies with access to a large case-control cohort, rare variants can be assessed according to different published gene-based rare variant methods (Table 7) allowing genetic hypotheses to be tested in complex diseases. However, these types of studies lead to substantial analytical challenges (Moutsianas et al., 2015). On the other hand, there is no current evidence in the literature for the use of NGS and WES-trio for the identification and association of candidate genes/variants with NAIS. Furthermore, when investigating the presence of a genetic aetiology in NAIS, the results have not

been successful. The reason may be due to the small number of samples analysed which is related to the inherent complexity of the disease, newborns not properly screened and diagnosed with this type of stroke, lack of research funding, or parents of the affected infant without a notion of how important a thorough investigation of the disease can be.

.**Table 7**. Gene-based association methods for large control-case cohorts

Method	Method Description
Mixed Effects Score Test (Sun et al., 2013)	Categorized regression model which uses two independent statistic tests to quantify heterogeneity and variant size effect
SKAT-O (S. Lee et al., 2012)	Unidirectional burden test and variance-component SKAT test in an adaptive linear combination
Sequence Kernel Association Test (M. C. Wu et al., 2011)	C-ALPHA version that considers allele frequency for weighting
C-ALPHA (Purcell et al., 2007)	Case-control variants analysed to detect deviation of observed from expected binomial distribution
UNIQ test (Purcell et al., 2007)	Permutations to assess significance and straightforward count of total case-unique rare alleles
BURDEN method (Purcell et al., 2007)	Case-control permutation-based test comparing raw allele counts
Kernel-Based Adaptive Cluster (B. Li et al., 2008)	Individuals are assessed using a weighted sum of allele counts. Variant weights are established adaptively and are based on observed effect sizes.
Weighted Sum Method (Madsen et al., 2009)	Inverse frequency-weighted rare variants and phenotypes scores are summed
Weighted Sum Statistic (Madsen et al., 2009)	Inverse-frequency-weighted rare variant counts per individual in cases and controls are compared using a permutation-based technique
Variable threshold (Price et al., 2010)	Case-control sum of rare allele count including an allele frequency threshold
Combined Multivariate and Collapsing (B. Li et al., 2008)	Individual dose for the collapsed 'variant' is regressed against phenotype; all rare variants are condensed into a single variant

6.1 COL4A1

To understand how *COL4A1* is involved in vascular defects, it is important to highlight the critical role of type IV collagen in the integrity of the vasculature. Collagen alpha-1 (IV) chain protein, which is encoded by the *COL4A1* gene, is expressed in all tissues, including the vessels (Urabe et al., 2002). This type IV collagen is the major structural component of the glomerular basement membranes. It forms a mesh-like structure and functions as part of a heterotrimer, which interacts with other components of the extracellular matrix, such as laminins, proteoglycans, and entactin/nidogen (Yurchenco et al., 2004).

Mutations in *COL4A1* have been shown to cause structural alterations in the basement membrane of endothelial cells, which can lead to vascular defects (van der Knaap et al., 2006). These vascular alterations have been associated with the following documented phenotypes: brain small vessel disease with or without ocular anomalies (MIM #175780), susceptibility to intracerebral haemorrhage (MIM #614519), tortuosity of retinal arteries (MIM #180000), hereditary angiopathy with nephropathy, aneurysms, and muscle cramps (MIM #611773), and autosomal dominant pontine microangiopathy and leukoencephalopathy (MIM #618564).

Brain small-vessel disease accounts for 20-30% of ischemic strokes and a higher proportion of intracerebral haemorrhage (D. B. Gould et al., 2006). Additionally, intracerebral haemorrhage underlies 10-15% of strokes with a high rate of death and long-term disability (Thom et al., 2006). However, it is common for both diseases to occur in the absence of risk factors, which leads to the intervention of other unidentified predisposing factors (D. B. Gould et al., 2006).

Nearly all reported mutations in *COL4A1* that have been associated with the disease are missense, but there have also been cases of frameshifts mutations. For example, a study from Belgium identified two *de novo* variants, a frameshift leading to a premature stop codon (p.Gly696fs) and a splice site mutation c.2194-1G>A, which most likely led to exon skipping (Lemmens et al., 2013). Clinical assessment and brain imaging of a 1-year-old female carrier of the p.Gly696fs mutation showed porencephaly in the right lateral ventricle potentially caused by in utero stroke, whereas the adult carrier of the splice site mutation had a right thalamic haemorrhage and white matter changes. Interestingly, other heterozygous missense variants (p.Gly1236Arg and p.Gly749Ser) have been identified with individuals presenting similar phenotypes (D. B. Gould et al., 2005).

Mutations in *COL4A1* have an autosomal dominant pattern of inheritance and are usually located in a highly conserved region of the triple helix domain of the gene (Lemmens et al., 2013). The proposed pathogenic mechanism that could explain the contribution of *COL4A1* mutations to haemorrhagic stroke is haploinsufficiency and the dominant-negative mechanism (D. B. Gould et al., 2006; Pöschl et al., 2004). These findings show the important implications of *COL4A1* sequencing analysis for the evaluation of vascular diseases and the adequate measures that could be taken in

individuals at risk, for example, newborns, where it could increase the likelihood of survival and decrease the severity of a possible cerebral haemorrhage.

6.2 JAM3

The regulation of the endothelial permeability is through a specialized cell-cell tight junction or zonula occludens, which is involved in the transduction of different signals. At the same time, it forms continuous seals around cells that function as a diffusion barrier that prevents solutes and water from passing through the paracellular space (Mochida et al., 2010). These cell-cell tight junctions are composed of transmembrane proteins such as claudins, occludins, tricellulins, JAMS (junctional adhesion molecules), scaffold adaptors, and regulatory proteins (Ebnet, 2008).

The Junctional Adhesion Molecule 3 (*JAM3*) gene is part of a superfamily immunoglobulin, which encodes the junctional adhesion-molecule protein C, a vital tight junction of the blood-brain barrier in the cerebrobascular endothelium. Mutations in this gene have been associated with the autosomal recessive haemorrhagic destruction of the brain (MIM #613730) (Akawi et al., 2013). Additionally, mutations in other genes encoding tight junction proteins have been reported to cause many genetic disorders (Burdon et al., 2003). For example, the familial hypercholanemia (MIM #607748) is caused by mutations in the *TJP2* gene (Carlton et al., 2003), autosomal-recessive deafness 29 (MIM #605608) by mutations in *CLDN14* (Wilcox et al., 2001), and X-Linked dominant syndrome Nance-Horan (MIM #302350), which is caused by mutations in the *NHS* gene (Burdon et al., 2003).

Although *JAM3* has not been associated with stroke per se, it is associated with an autosomal recessive cerebrovascular haemorrhagic disorder with a fatal outcome in childhood. For example, one study demonstrated how the homozygous p.Cys219Tyr and p.Glu116Lys variants in *JAM3* were the cause of serious autosomal recessive disorder in which individuals were born with congenital cataracts and developed severe progressive haemorrhagic destruction of cerebral white matter and the basal ganglia (Mochida et al., 2010). Considering that missense mutations in *JAM3* are likely to result in complete loss of protein function and cause severe syndromes with a high chance of mortality in infancy, we recommend including not only *JAM3* but also other tight-

junction genes (*JAM1*, *JAM2*, *IGSF5* (previously known as *JAM4*), *TJP2*, *CLDN14*, *OCLN* and *NHS)* to gene panels designed to assess cerebrovascular disorders.

Summarising, NAIS is a complex disease due to its variability in clinical presentation and aetiology, in which the genetic part is unclear. In this thesis, we aimed in a series of patients with a very well-defined idiopathic NAIS, to unveil whether there is any gene(s) involved that can partially or totally explain the pathophysiology and vulnerability of the disease.

HYPOTHESIS AND OBJECTIVES

Hypothesis

In the last years, it has been demonstrated that some early-onset disorders in the newborn or infant period of life are caused by *de novo* dominant mutations in several genes, such as developmental and epileptic encephalopathies. We hypothesised that sporadic idiopathic NAIS in term newborns is associated with rare *de novo* variants in genes involved in endothelial rheology, coagulation, and cell adhesion. The WES-trio genetic approach will allow us to identify these variants and understand the underlying pathways involved in the disease in homogeneous subgroups of patients.

Objectives

1. To identify rare *de novo* pathogenic variants that might lead to the cause of NAIS.

2. To evaluate *in silico* the putative pathogenic impact of the variants identified as possible candidates.

3. To discover a major common genetic determinant that could influence the development of the disease.

4. To discern if a specific subclinical homogeneous phenotype in NAIS could be conditioned by a certain genetic determinant.

5. To decipher the association between the selected candidate gene with NAIS by identifying the relevant biological mechanisms involved.

MATERIALS & METHODS

1. Clinical Methodology

1.1 Patient Selection

1.1.1 Ethical Considerations

The written informed consent was obtained from the parents of the patients and the procedures performed complied with the ethical guidelines by Sant Joan de Déu Children's Hospital (HSJD) Clinical Research Ethics Committee (references PIC-30-17 and PIC-223-19) and with the Helsinki Declaration of the World Medical Association.

1.1.2 Subjects Selection

This is a sub-study of a multicentre prospective observational study, which included 48 infants with symptomatic NAIS located in the MCA. The study was conducted in six paediatric university hospitals in Spain from January 2010 to December 2017. We decided to study a well-defined homogeneous sample of patients with NAIS meeting inclusion and exclusion criteria. The inclusion criteria comprised the following: a) patients must have at least 35 weeks' gestation, b) presenting with symptomatic NAIS within the first 28 days of life, and c) NAIS confirmed by magnetic resonance imaging (MRI). Symptomatic NAIS was defined as the finding of an acute ischemic lesion by MRI study in the territory of the main cerebral arteries (middle cerebral artery [MCA], anterior cerebral artery [ACA] and posterior cerebral artery [PCA]) in infants presenting with seizures, recurrent apnoea, or acute neurological deficit during the first days of life. The exclusion criteria comprised the following: a) MRI suggestive of old infarction occurring before birth, b) vascular malformation, c) moderate to severe hypoxic-ischemic encephalopathy (HIE), d) congenital or chromosomal anomalies, e) metabolic (including persistent hypoglycaemia) or infectious diseases, f) complex congenital heart anomalies, or g) had genetic thrombophilia: *F5* G1691A, *F2* G20210A, or homozygosity for *MTHFR* C677T genotypes, as well as those whose family refused to participate or continue in the study or when a false detected paternity/maternity.

Of the 48 infants with symptomatic NAIS, it was only possible to have blood samples from 26 trios. Three of these 26 were removed due to one false maternity and two infants with subjacent conditions. Thus, the final clinical cohort consisted of 23 trios.

1.1.3 Family History and Maternal Diseases

Family information is soughed through a specific interview of both parents of the infants and the clinical records of their mothers. The following maternal data were collected: age, weight, and height to calculate body mass index (BMI), hypertension, spontaneous abortions, antiphospholipid syndrome, connective tissue, or autoimmune disease such as lupus or dermatomyositis, thrombocytosis, idiopathic thrombocytopenia, polycystic ovary syndrome, migraine, epilepsy, and ingestion of oral contraceptives. A three-generation pedigree was obtained for each family regarding 1) myocardial infarction, 2) pulmonary embolism, 3) cerebrovascular event, and 4) deep vein thrombosis.

1.2 Clinical Phenotype

All infants recruited with NAIS had a careful clinical evaluation during the neonatal period and at discharge. In addition, continuous neurophysiological evaluation during the first days after clinical debut, determination of brain damage biomarkers in CSF (neuronal specific enolase) and multimodal neuroimaging studies with MRI (T1W, T2W, DWI, ADC, DTI, TOF) during the neonatal period were performed.

1.2.1 Cerebral Magnetic Resonance Imaging

All MRI studies were performed using a 1.5 T Signa Excite scanner (General Electric). MRI was performed at a median age of 3 days (P25, P75; 2.75). The acquisition protocol has been described elsewhere (Stephan-Otto et al., 2017). Diffusion-weighted imaging (DWI) data was available in all cases.

Infarcts were classified by 2 observers blind to the clinical data according to the segments of the middle cerebral artery and the branches involved. Differences in the anatomical location of NAIS were resolved by consensus among three observers, as previously described (Arca et al., 2020).

1.2.2 Lesion Volume

Two observers performed the volumetric segmentation of the lesions by MRI independently: both were unaware of the neuroimaging diagnosis and the clinical data. NAIS volume included only primary lesions and excluded distant changes by pre-Wallerian degeneration (Srivastava et al., 2019). The process was performed through the multimodal analysis of MR images with ITK-Snap software, version 3.0 (www.itk-snap.org) (Yushkevich et al., 2006). The whole procedure has been published elsewhere (Stephan-Otto et al., 2017). Once a satisfactory segmentation of the lesion was achieved in the evolution step, it was manually corrected by simultaneously inspecting its features in all the imaging modalities, as well as its 3D visualization. In this step, the consensus of 3 experts blinded to the radiological diagnosis was reached.

Total infarct volumes were expressed in cubic centimetres (cm3), and the relative infarct volume (RIV) in percentage. The latter was obtained as a ratio to the total brain volume excluding the cerebellum.

1.2.3 Outcomes

All patients were followed up. Systematic visits were performed at the ages of 6 months, 1 and 2 years in the different study centres participating. Serial head circumference (HC) measurements were made up to 24 months, and microcephaly was defined as an HC more than two standard deviations below the mean for sex and age (Van Den Broeck et al., 2009).

The Gross Motor Function Classification System (GMFCS) was used to graduate functional impairment at two years of age. We considered motor impairment when the infants had GMFCS scores of at least I. Further, the outcome at 24 months was assessed using the Bayley Scales of Infant and Toddler Development, Third Edition (BSITD-III). 19 Mild adverse outcomes were considered when there was a Bayley-III score between minus 2 SD and minus 1 SD (70 to 85) in any of the domains and/or any cerebral palsy (GMFCS I). Moderate-severe adverse outcome was defined as minus 2 SD (MDI or PDI <70) in any of the Bayley-III domains and/or cerebral palsy (GMFCS II-V).

2. Molecular and Genomic Methodology

2.1 DNA Extraction

DNA samples from the recruited individuals and their parents were extracted at the Sant Joan de Deu Hospital Biobank, Barcelona, Spain using the MasterPure™ DNA extraction kit (Ecogen SL, Spain).

2.2 Paternity Test

Identification of true paternity/maternity of the trios involved in this study was carried out using the PowerPlex® 16 HS System (Promega, WI, USA). This system co-amplifies the loci D18S51, D21S11, TH01, D3S1358, Penta E, FGA, TPOX, D8S1179, vWA, Amelogenin, CSF1PO, D16S539, D7S820, D13S317, D5S818 and Penta D. One trio from our cohort could not go further in the investigation due to false maternity (ovodonation not mentioned).

2.3 Exome Capture and Sequencing

Next generation sequencing technology for the whole-exome sequencing of 23 trios was used to sequence the protein-coding regions of the genome through exome enrichment and library preparation (DNA Prep with Enrichment, Illumina, San Diego, CA). Subsequent sequencing was performed by the Illumina NextSeq 500 System.

2.4 Variant Annotation

The bioinformatics team of the Hospital Sant Joan de Dèu performed the analysis using a custom pipeline. The quality of the reads was determined (FastQC version 0.11.5; Babraham Institute, Cambridge, UK), the low-quality adapters and reads were removed (Cutadapt version 1.13; National Bioinformatics Infrastructure Sweden, Uppsala, Sweden), and subsequently aligned to the human reference genome [GRCh37 version (BWA-MEM version 0.7.15; Wellcome Trust Sanger Institute, Cambridge, UK)]. Low mapping reads and duplicates [BEDtools version 2.26.0 (University of Utah, Salt Lake City, UT) and Picard version 2.9.0 (Broad Institute, Cambridge, MA)] were removed. Variants are detected with four different programs [SAMtools version 1.5 (Wellcome

Trust Sanger Institute, Cambridge, UK), GATK version 3.7 (Broad Institute, Cambridge, MA), FreeBayes version 1.1.0 (Boston College, Boston, MA), and VarScan version 2.4.0 (Washington University, St. Louis, MO)], and the annotation (SnpEff version 4.3k; Wayne State University, Detroit, MI) included the databases 1000 Genomes, dbSnp, ExAc, and ClinVar. Variants with frequency >1% and classified as likely benign or benign in Clinvar were removed.

2.5 Variant Filtering

Annotated variants followed a filtering workflow analysis (Supplementary Figure S1) with the following criteria: (a) total and European non-Finish population variant frequency <0,01 (<1%) using Genome Aggregation Database (gnomad.broadinstitute.org) and Centro de Investigación Biomédica en Red de Enfermedades Raras-Spanish Variant Server (csvs.babelomics.org), (b) *de novo* autosomal dominant, (c) variant located in coding exons, canonical splice site regions, and deep intronic ±70 nucleotides, (d) variant interpretation as Pathogenic, Likely Pathogenic and VUS according to the ACMG standards and rules using Varsome (varsome.com/), and (e) visualized as *de novo* using integrative genome viewer (IGV).

Based on the biological pathways that might be associated with NAIS, a list termed "NAIS gene list" was created with 1550 genes (Supplementary Table S3) using OMIM database (https://www.omim.org/), IOBIO (https://iobio.io/), and literature review using PubMed (https://pubmed.ncbi.nlm.nih.gov/). The list included genes related to cell adhesion (544), coagulation (268), platelet activation (63), inflammation (160), stroke (257), and endothelium (258). Additionally, recessive mode of inheritance (compound heterozygous and homozygous variants) and prothrombotic state polymorphisms approach were applied as a secondary approach.

2.6 Pathogenicity Prediction

The remaining post-filtering variants went through several *in silico* tools to assess and predict the level of pathogenicity/impact of the variants. These tools comprise the following: (a) sequence and structure-based analysis: FATHMM (fathmm.biocompute.org.uk/), MutationTaster (www.mutationtaster.org/), Polyphen-2 (genetics.bwh.harvard.edu/pph2/), (b) sequence homology-based tool for pathogenicity

prediction: SIFT (sift.bii.a-star.edu.sg/www/SIFT_seq_submit2.html) PROVEAN (provean.jcvi.org/seq_submit.php), CADD (cadd.gs.washington.edu/snv); (c) splice sites analysis: SpliceAI (https://spliceailookup.broadinstitute.org/); (d) constraint scores according to Genome Aggregation Database (https://gnomad.broadinstitute.org/); (e) variant localization at the protein and genome level: Mutation Mapper (www.cbioportal.org/mutation_mapper) and IGV, respectively.

2.7 Detection, validation, and primer design for *De novo* Variants

Variants detected as possible *de novo* were visualised using IGV. Possible *de novo* variants identified presented the alternative allele as heterozygous in the infant and the reference allele as homozygous in both parents. Validation of findings were performed using Sanger sequencing using a Bioanalyser (Applied Biosystems 3130XL Genetic Analyser), which is still considered the gold standard for validation of NGS genetic variants due to absence of current NGS guidelines to define specific guidance and quality parameters for confirmatory results.

Flanking primers (Supplementary Table S1) were designed for each DNA target using Genome browser (https://genome.ucsc.edu) for proper RefSeq DNA sequence selection, primer3 tool (https://primer3.ut.ee) for forward and reverse primer design, and the *in silico* PCR tool of Genome Browser (https://genome.ucsc.edu/cgi-bin/hgPcr) to verify that the chosen primers only amplified the target DNA. Endpoint PCR was used as the system to amplify the DNA sequences and PCR products were analysed manually by agarose gel electrophoresis. The amplicons matching the product base pairs followed a purification step with ExoSAP (Applied Biosystem, Foster City, CA, USA) and subsequently sequenced with the BigDye Terminator version 3.1 (Invitrogen Carlsbad, CA, USA) according to the manufacturer's instruction and analysed using a Genetic Analyser 3130XL (Applied Biosystems, Foster City, CA, USA). Chromatograms were visualised and analysed using SnapGene Viewer software (GSL Biotech LLC).

3. Identification of NAIS Candidate Genes

3.1 Knowledge-driven Analysis and Prioritization

The identification and association of candidate genes with NAIS were carried out performing a Knowledge-driven Analysis (KDA) (Dashti et al., 2017), which was adapted for the study. We designed a pipeline based on a set of qualitative and quantitative information related to each identified gene/variant with a customised score based on the results. The information and scores comprises the following: (b) Variant rarity using gnomAD and CSVS: variant not found (3 points), Frequency = 0 (2 points), Frequency <0.01 (1 point); (c) Variant/phenotype reported in databases using ClinVar, HGMD, and OMIM: yes (1 point), no (0 points); (d) Pathogenicity prediction using CADD scores ≥30 (2 points), 15-29.9 (1 point), <15 (0 points); (e) ACMG variant classification according to Varsome: Pathogenic (2 points), Likely Pathogenic (1 point), VUS (0 points); (f) Nucleotide conservation score using GERP score >2 (1 point), score <2 (0 points9; (g) Identified phenotypes with mode of inheritance information: autosomal dominant (1 point), autosomal recessive (0 points); (h) Gene constraint score using gnomAD: z-score >3.09 (1 point), z-score <3.09 (0 points), pLI score <0.5 (1 point), pLI score >0.5 (0 points); (i) Gene molecular/biological ontology related to NAIS pathophysiology using Gene Ontology, AmiGO, Enrichr: yes (1 point), no (0 points); (j) Gene expression in relevant tissue/organ using Human Protein Atlas and Expression Atlas: yes (1 point), no (0 points); (k) Encoded protein in a pathway associated with NAIS using KEGG, Reactome, PathwayCommons, GeneMania: yes (1 point), no (0 points); (l) Protein interactions with other proteins with possible relationship with NAIS using StringDB, Signor 2.0, IntAct, PathwayCommons, and GeneMania: yes (1 point), no (0 points). Additionally, for information purposes only, we analysed the mutational architecture of the variant and whether the genes identified between the probands have relevant similarities among them. The maximum score for this strategy is 21 therefore; the variant with the highest score is considered the most pathogenic variant. Furthermore, for a purely visual purpose, we added a gradient colour scale scheme in which the green colour and its gradient towards yellow indicates the section where the variants considered more benign are found, while the gradient between yellow and red indicates the presence of the most pathogenic variants.

This custom KDA pipeline design allowed us to further rank and prioritize our gene set to select the most suitable candidate gene for performing functional studies and gene-disease association.

3.2 Gene Ontology and Pathway Enrichment

Assessing the biological and molecular roles of the genes was a crucial step in identifying which gene might be directly or indirectly associated with the pathophysiology of NAIS. We looked for any term related to angiogenesis, endothelium, cell adhesion, inflammatory response, and coagulation processes. Equally important was the use of KEGG, Reactome, PathwayCommons, GeneMania, and Interactome3D for finding relevant signalling pathways.

4. Cloning

4.1 Plasmids generation

For functional studies, overexpression of PIK3CD *wild type* and mutant proteins were performed. The entire cDNA of *PIK3CD* was PCR-amplified from the pHAGE-PIK3CD vector (Addgene plasmid # 116773, Watertown, MA, USA) and cloned into the pEGFP-N1 vector using recombination cloning with the NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's recommendation (Supplementary Figure S2). The *PIK3CD* PCR amplification conditions were as follows: 98°C for 30 s, then 35 cycles of 98°C for 10 s, 60°C for 48 s, and 72°C for 30 s, and followed by one cycle of 72°C for 5 min. The missense variant of *PIK3CD* p.Gln431Arg was generated by site-directed mutagenesis using the Q5® Site-Directed Mutagenesis Kit (New England Biolabs, Ipswich, MA, USA) following the manufacturer's indications. The reaction conditions were as follows: 98°C for 30 s, then 25 cycles of 98°C for 10 s, 57°C for 30 s, and 72°C for 3.3 min, and followed by one cycle of 72°C for 2 min. Primers used (Supplementary Table S2) were recombinant to both donor and recipient expression plasmids.

4.2 Cell Transformation

Recombinant *wild type* and mutant plasmids were transformed using 50 µl of *Escherichia coli* DH5α cells followed by 30 min incubation on ice, 30 s at 42 °C and 5 min on ice. Reconstitution was carried out adding 950 µl of Luria-Broth (LB) (Sigma-Aldrich St. Louis, MO) medium without antibiotics (NaCl 5 g/L, Tryptone 10 g/L, Yeast Extract 5 g/L) and incubating for 1h at 37 °C in agitation at 300 r.p.m. The suspension was centrifuged 5 min at 5000g for supernatant removal. The remaining volume (≈100 µl) was spread on a LB plate with Kanamycin (50mg/mL) as antibiotic and grown overnight at 37 °C for colonies analysis.

4.3 Plasmid Preparation, Sequencing and Analysis

Plasmids were extracted from a 50 ml LB-Kanamycin (50mg/mL) overnight colony culture with the Midiprep kit (ZymoPURE^{TM} II Plasmid Midiprep Kit, Zymoresearch, Irvine, CA, US) following manufacturer's indications. Plasmid DNA of pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{Gln431Arg} were sequenced with BigDye Terminator version 3.1 (Invitrogen Carlsbad, CA, USA) according to manufacturer's instructions using a Bioanalyser (Applied Biosystems 3130XL Genetic Analyser). Chromatograms were analysed using SnapGene Viewer software (GSL Biotech LLC).

4.4 Cell Culture, Transfection, and stimulation.

Fibroblast-like cell lines derived from monkey kidney tissue (COS-7) were obtained from ATCC. Cells were cultured in Dulbecco's modified Eagle medium (Sigma-Aldrich, St. Louis, MO) supplemented with 10% v/v fetal bovine serum (FBS, Gibco, Thermo Fisher Scientific, Inc), 2 mmol/L L-glutamine (Sigma-Aldrich St. Louis, MO), and 100 mg/mL penicillin-streptomycin (Sigma-Aldrich St. Louis, MO). Cells cultures were maintained at 37°C in a 5% CO2 humidified atmosphere. Cells were transfected with FuGene HD (Promega, WI, USA) and Gibco™ Opti-MEM™ (Thermo Fisher Scientific, Inc) according to the manufacturer's instructions. COS-7 transfected were serum-starved for 24h and subsequently, the cells were stimulated with 100 nM insulin (Gibco, Thermo Fisher Scientific, Inc) for 1 h at 37ºC.

5. Western Blot Analysis

The soluble fraction of COS-7 cells was obtained with RIPA buffer, quantified by BCA Protein Assay Kit (Pierce; Thermo Fisher Scientific, Inc), separated using 10% SDS-PAGE gels, and transferred to a PVDF membrane (Millipore, Sigma-Aldrich St. Louis, MO). Membrane blocking was performed with 5% skim milk for 1 hour and subsequently incubated with the primary antibodies α-Phospho-Akt (Ser473) (1:1000, #4058, Cell Signalling, Danvers, MA, USA), and α-Akt1 Antibody (B-1) (1:500, sc-5298, Santa Cruz Biotechnology, Dallas, TX, USA) at 4℃ overnight and under agitation. The next day, after washing with TBS-T1× buffer, the membrane was incubated with α-mouse IgG-HRP secondary antibody (1:10.000, sc2537, Santa Cruz Biotechnology) for 2 h at room temperature. The ECL kit detection system (Cytiva Amersham Pl, UK) was used for semi-quantification. Western Blot images were analysed using ImageJ software version 1.46 (National Institutes of Health, Bethesda, MD, USA).

6. Confocal Microscopy Imaging

Confocal microscopy analysis was performed by Leica TCS SP8 equipped with a white light laser, HyVolution and Hybrid spectral detectors (Leica Microsystems GmbH, Mannheim, Germany). Confocal images of transfected COS-7 cells were acquired using an HC x PL APO 20x/0.75 dry and an HC x PL APO 63x/1.4 oil immersion objective. DAPI was excited with a blue diode laser (405 nm) and detected in the 425–480 nm. Green fluorescence protein (GFP) was excited with an argon laser (488 nm) and detected in the 505-550 nm. In COS-7 cells, Z stacks of 8-10 sections were acquired every 0.5-0.8 μm along with the sample thickness. Appropriate negative controls were used to adjust confocal settings to avoid non-specific fluorescence artefacts. Maximum intensity projections were performed using LAS X software (Leica Microsystems GmbH, Mannheim, Germany) and ImageJ software version 1.46.

7. Statistical Analysis

Data analysis is expressed as means ± SEM, column bars or scatter plots graphs, showing error bars and individual values. P values less than 0.05 were considered significant and are indicated by asterisks: $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$. P values above 0.05 were not considered statistically significant and are designated "ns." Two-tailed Student *t*-test (parametric test) was used in Western Blot for comparison of phospho-AKT (pAKT) and total-AKT (tAKT) protein expression in pEGFP-PIK3CD^{WT}, pEGFP-PIK3CD^{Gln431Arg} and Mock (control pEGFP-N1). Additionally, two-tailed Student t-test (parametric test) was used to compare the two-protein patterns of PIK3CD in COS-7 transfected cells. SAMtools version 1.5 (Wellcome Trust Sanger Institute, Cambridge, UK) was used for the comparison between the sequencing coverage of the exonic target regions of the WES-trio at a sequencing depth of 1X, 10X, 20X, and 30X. Statistical graphs and computation were performed using GraphPad Prism version 8.0.1 (GraphPad Software, Inc., La Jolla, CA) and RStudio software version 1.1.447 (RStudio Inc., Boston, MA).

RESULTS

1. WES-trio for *De Novo* Variant Identification

In our study 26 infants were recruited but 3 did not meet the inclusion criteria. One infant presented a false maternity and two were diagnosed with NAIS due to underlying risk factors (Figure 16A). The remaining 23 infants and their parents (Table 8) continued with the WES-trio. On average, 94.4% (±1.8) of the target regions, based on the Genome Reference Consortium Human Build 37 (GRCh37), were covered at least 30-fold (Supplementary Figure S3). This coverage at 30-fold allowed us to detect heterozygous DNVs without any bias due to the uniform sequencing coverage between parents and probands (p= 0.9811).

In total, 1.966 variants from the 23 WES-trio met the chosen *de novo* bioinformatic criteria for variant annotation (Figure 16B). Additionally, we did not find annotated variants in our NAIS gene list that met the bioinformatic filter criteria. However, 912 variants in other unlisted genes did (Figure 16C). Validation of variants visualized as *de novo* with the IGV software (Figure 17A) was carried out by Sanger sequencing (Figure 17B) to discard false positive variants. These false *de novo* variants can be incorrectly annotated due to being under-called in the parents or due to artefacts related to the sequence itself. Subsequent analysis of all 912 variants (Figure 16D), confirmed 28 unique DNV in 28 genes (Table 9). These means that the probands did not share any common *de novo* variant gene.

Despite having a mean of 85.5 ± 20.4 annotated variants per proband (Figure 18A), our results showed that, of all 28 unique DNV identified, 9 (39%) probands carried at least one DNV and 8 (35%) carried >1 DNV (Figure 18B). However, in six probands (26%) we could not continue the analysis workflow. Five (21.7%) had variants that met our filter criteria, but they were located in regulatory regions, deep intronic variants greater than ±70, or classified as synonymous variants and one proband (4.3%) had no DNV identified. These findings were confirmed with a less restrictive bioinformatic re-analysis (Figure 16B) to ensure that no variant was left behind by bioinformatic variables or manual errors during the filtering workflow (data no shown).

All 28 DNV were distributed across the majority of all the chromosomes (Figure 18C), except in six autosomal chromosomes (3,5,7,9,10,13) and the sexual chromosomes (X,Y). The difference between the number of annotated variants and the validated DNV is due to

the presence of false *de novo* variants because of the flexible bioinformatic analysis carried out.

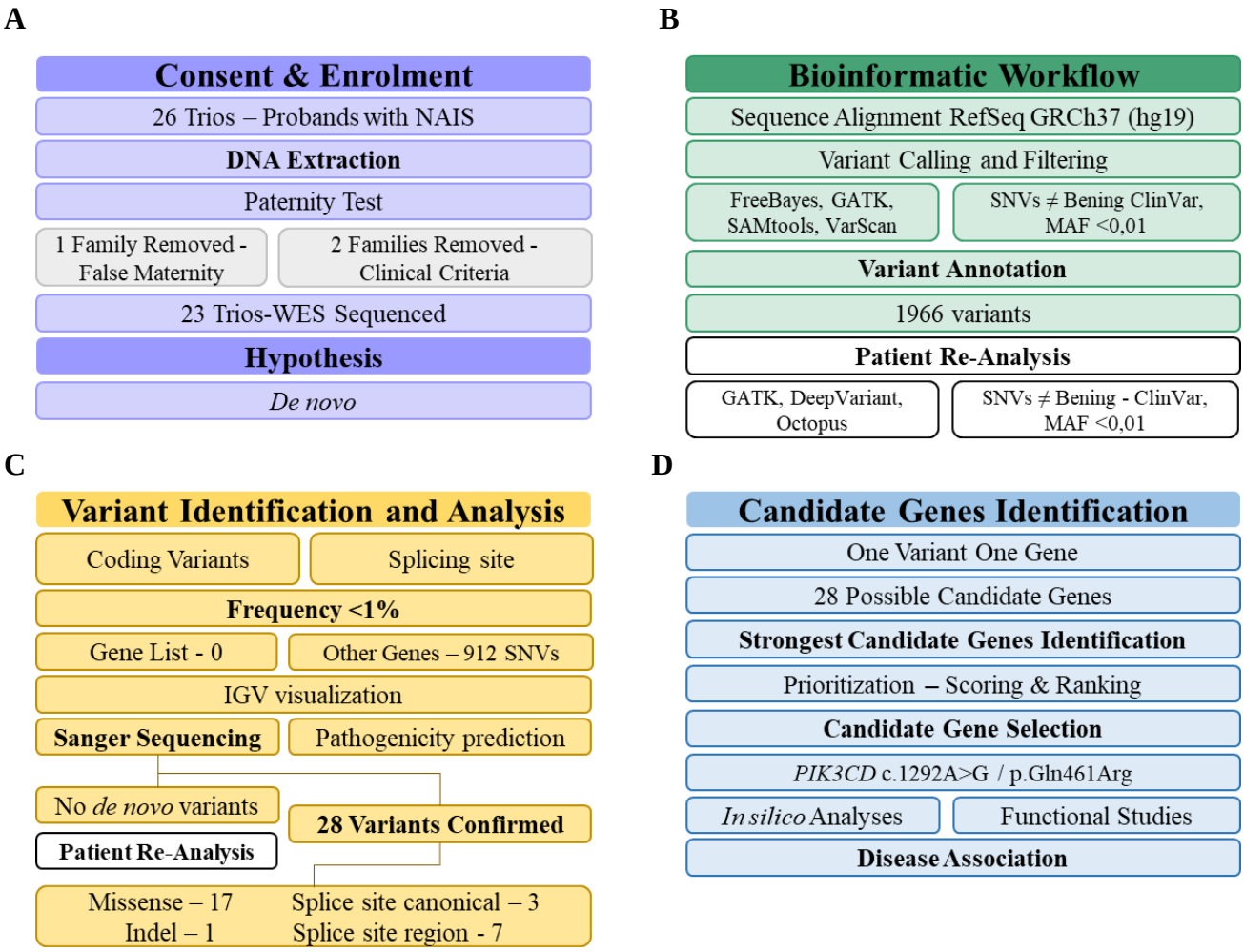


Figure 16. **Workflow sum up of rare *de novo* variant analysis**. **A**) Recruitment of subjects for WES-trio and hypothesis analysis. **B**) Bioinformatic pipeline for variant filtering and annotation before variant identification and analysis. **C**) Personalised variant filtering was performed for the analysis and identification of possible candidate genes. **D**) Workflow for the selection of the strongest candidate genes for disease association studies. ACMG: American College of Medical Genetics; P: Pathogenic, LP: Likely Pathogenic, VUS: Variant of Uncertain Significance, MAF: Minor Allele Frequency, SNV: Single Nucleotide Variant, IGV: Integrative Genome Viewer, RefSeq: Reference Sequence.

Table 8. Characteristics of the 23 Infants with Neonatal Arterial Ischemic Stroke

Characteristic	**Values**[a]
Male-to-female ratio	2.8:1
Age (years)	
Infant	5 ± 4
Mother	32 ± 4
Father	38 ± 7
Ethnicity	
Caucasian	23/23 (100)
NAIS Criteria	
Seizures	16/23 (70)
Multifocal Seizure	1/23 (4)
Neurological Clinic	3/23 (13)
Apnoea	3/23 (13)
Hospitalization (days)	11.8 ± 3.0
Perinatal Factors	
pH	7.19 ± 0.10
Fetal Presentation	
Cephalic	23/23 (100)
Delivery	
Eutocic	5/23 (21)
Instrumental	2/23 (9)
Caesarean	
Urgent	13/23 (57)
Scheduled	3/23 (13)
Apgar score ≤5 at 5 minutes	2/23 (9)
Arterial Territory Involvement of MCA[c]	
M1 pre	6/23 (27)
M1 post	4/23 (17)
M2	7/23 (30)
M3-M4	4/23 (17)
Perforant	2/23 (9)
Full Brain Infarct Volume[d]	
Absolute (cm^3)	91 - 261
Relative	2.3 - 66.2

[a] Categorical variables are expressed as n/N (%)

[b] Not Applicable

[c] Middle Cerebral Artery

[d] Expressed in value range

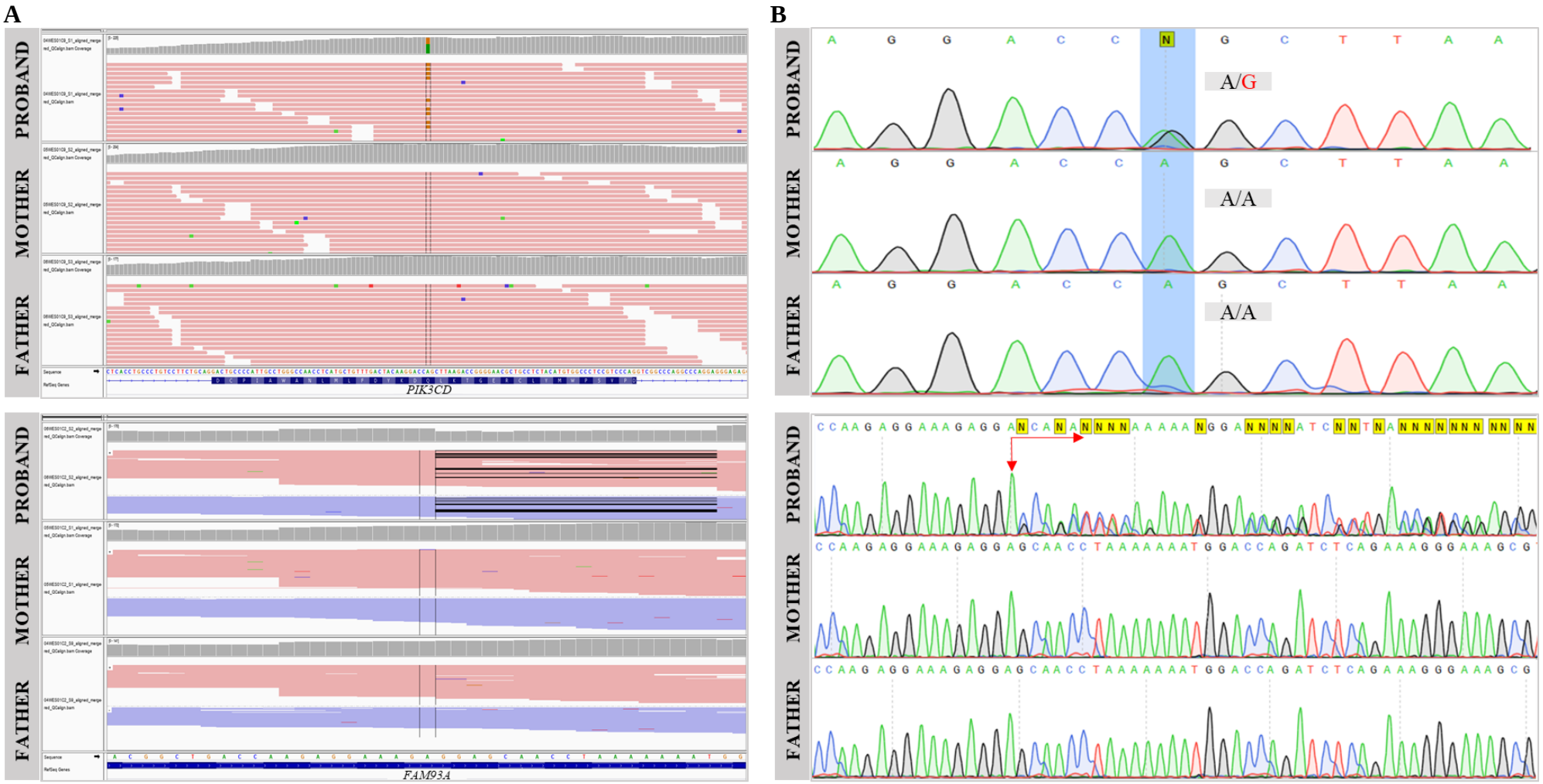


Figure 17. **Sanger sequencing for validation of candidate *de novo* variants. A**) Visualization of two candidate *de novo* variants in different families using the IGV browser. The top panel shows a *de novo* SNV localized in the *PIK3CD* gene with the coverage from the proband, mother, and father (grey)**,** the forward (red) and reverse (blue) reads, and the reference sequence GRCh37 (bottom thicker blue line) with its respective translated amino acids. Bottom panel shows a *de novo* deletion in *FAM193A* gene. **B**) Chromatograms of Sanger sequencing showing the validation of the two candidate *de novo* variants.

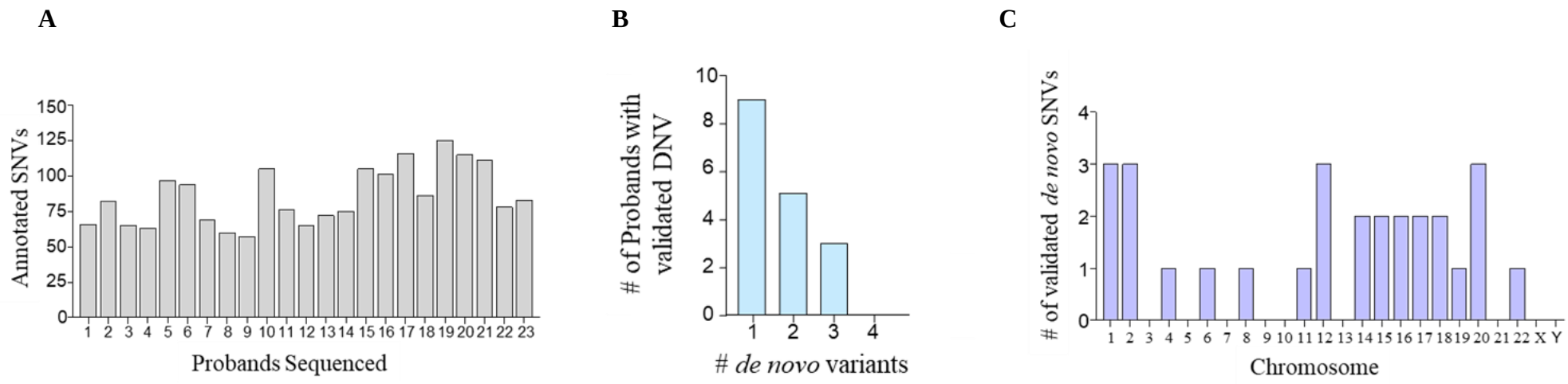


Figure 18. **Annotated, filtered, and validated *de novo* variants data**. **A**) Number of possible *de novo* annotated variants per proband meeting the bioinformatics workflow filter criteria. **B**) Number of probands with the number of validated *de novo* variants carried. **C**) Distribution of all 28 confirmed *de novo* variants among chromosomes according to the autosomal dominant analysis.

Table 9. Detailed information of all 28 confirmed *de novo* variants

Gene	RefSeq[1]	Variant Details			RefSNP[2]
		Genomic Change	Nucleotide Change	Amino Acid Change	
EPS15L1	NM_001258374.1	19:16503267:C:T	c.1967-16G>A	N/A	rs367891749
SYNE2	NM_182914.2	14:64449435:G:A	c.1924G>A	p.Val642Met	-
FAM193A	NM_001256666.1	4:2698166:AGGAGCAACCTAAAAAAAT:A	c.2484_2501del	p.Glu828_Met833del	-
HECTD1	NM_015382.3	14:31613253:G:A	c.2778+64C>T	N/A	rs1041312283
TOP2A	NM_001067.3	17:38573165:A:G	c.22-18T>C	N/A	-
RBM7	NM_001286045.1	11:114271419:A:G	c.26A>G	p.Asp9Gly	-
KRT6A	NM_005554.3	12:52883718:T:A	c.1203+9A>T	N/A	rs752406296
UNC13D	NM_199242.2	17:73836958:C:T	c.615-47G>A	N/A	-
FAR2	NM_001271783.1	12:29464062:G:T	c.870G>T	p.Trp290Cys	-
SHANK3	NM_001372044.1	22:51142694:C:T	c.1954+18C>T	N/A	-
PCBP1	NM_006196.3	2:70315206:T:C	c.331T>C	p.Ser111Pro	-
KCNC2	NM_139137.3	12:75444970:A:G	c.815T>C	p.Leu272Pro	rs973994418
CHRNA4	NM_000744.6	20:61981022:C:T	c.1741G>A	p.Glu581Lys	rs752289948
SPTBN5	NM_016642.3	15:42147738:T:C	c.9218+9A>G	N/A	-
TRABD2B	NM_001194986.1	1:48260308:G:A	c.938C>T	p.Ala313Val	rs781498332
FAM83A	NM_032899.6	8:124195260:G:A	c.164G>A	p.Ser55Asn	-
NR2E1	NM_001286102.1	6:108502797:G:A	c.1048G>A	p.Ala350Thr	rs200730208
FASTKD5	NM_021826.4	20:3127550:G:A	c.2167C>T	p.Arg723Trp	rs1347412310
PIK3CD	NM_005026.3	1:9780028:A:G	c.1292A>G	p.Gln431Arg	-
SNTB2	NM_006750.3	16:69304210:GA:G	c.1148+24delA	N/A	-
NPC1	NM_000271.4	18:21121278:G:A	c.2365C>T	p.Arg789Cys	rs1555633697
DENND4A	NM_001320835.1	15:65993539:T:G	c.2560-9A>C	N/A	-
DYM	NM_017653.3	18:46784867:T:C	c.1252-4A>G	N/A	-
SLC2A10	NM_017653.3	20:45354351:G:T	c.676G>T	p.Asp226Tyr	-
AARS1	NM_001605.2	16:70287258:G:T	c.2634C>A	p.Phe878Leu	-
ESRRG	NM_001243518.1	1:216850763:G:T	c.142C>A	p.Pro48Thr	rs773155899
COL6A3	NM_004369.3	2:238249568:A:C	c.7991T>G	p.Val2664Gly	-
AFF3	NM_001025108.1	2:100199274:C:T	c.2854G>A	p.Gly952Ser	rs199884902

[1] RefSeq: Reference Sequence
[2] RefSNP: polymorphisms reference number database

2. Prediction of the Pathogenicity Impact for the Assessment of Variants

To assess the pathogenicity of all 28 DNV identified, the variants were classified according to their location in the gene and pathogenicity. Our results show 18 exonic variants (17 SNV (60.7%) and 1 in frame deletion variant (3.6%)), 3 variants (10.7%) located at the canonical splice site regions, and 7 deep intronic variants (25%) located ± 70 of the canonical splice site regions. Regarding pathogenicity of the variants, we first used Combined Annotation Dependent Depletion (CADD) for scoring the deleteriousness of single nucleotide variants. For this purpose, we select a CADD threshold score of >15, which turns out to be the mean value of all possible canonical splice site changes and non-synonymous variants in the human genome. The results (Table 10) showed variants (64.3%) with a CADD score >15 (maximum score 33) and 10 variants (35.7%) with a score <15 (minimum score 0.061).

Additionally, according to Varsome which is based on the ACMG standards and guidelines for sequence variant interpretation (Kopanos et al., 2019; Richards et al., 2015), 1 variant (3.5%) had a pathogenic verdict, 22 variants (78.6%) were classified as likely pathogenic, and 5 variants (17.9%) as Variant of Uncertain Significance (VUS). Furthermore, to have an in-depth analysis of the probable variant pathogenicity impact, we used other *in silico* predictors, which are mentioned in the pathogenicity prediction section of Material and Methods. This allowed us to better evaluate the variants and apply further strategies to select the strongest candidate gene.

Table 10. Detailed pathogenicity prediction scores of confirmed *de novo* variants

Gene RefSeq	Variant	*In silico* pathogenicity predictor tool															
		CADD[1]	Varsome[2]	MT[3]		SIFT[4]		F-MKL[5]		PROV[6]		PP-2[7]		SpliceAI[8]		Constraint[9]	GERP[10]
		S[11]	P[12]	S	P	S	P	S	P	S	P	S	P	S	P	S	S
EPS15L1 NM_001258374.1	19:16503267:C:T	0.06	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	-2.77
SYNE2 NM_182914.2	14:64449435:G:A	21	VUS	0.99	POL	0.12	T	0.783	D	0.092	N	0.85	P	-	-	-1.45	5.06
FAM193A NM_001256666.1	4:2698166:AGGAGCA ACCTAAAAAAAT:A	22.2	P	0.99	DC	-	-	-	-	-20.86	D	-	-	-	-	-	2.67
HECTD1 NM_015382.3	14:31613253:G:A	1.7	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	0.85
TOP2A NM_001067.3	17:38573165:A:G	1.2	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	-6.02
RBM7 NM_001286045.1	11:114271419:A:G	25.4	VUS	1	DC	0.001	D	0.800	D	-4.29	D	0.99	P	-	-	0.83	4.67
KRT6A NM_005554.3	12:52883718:T:A	0.56	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	-8.56
UNC13D NM_199242.2	17:73836958:C:T	1.5	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	-3.3
FAR2 NM_001271783.1	12:29464062:G:T	33	P	1	DC	0	D	0.993	D	-10.8	D	1	P	-	-	1.7	4.48
SHANK3 NM_001372044.1	22:51142694:C:T	0.54	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	-4.29
PCBP1 NM_006196.3	2:70315206:T:C	23	VUS	1	DC	0.002	D	0.444	N	-4.45	D	0.104	B	-	-	3.2	4.03
KCNC2 NM_139137.3	12:75444970:A:G	17	VUS	0.99	POL	0.2	T	0.567	N	-1.26	N	0	B	-	-	2.98	3.17
CHRNA4 NM_000744.6	20:61981022:C:T	26.9	VUS	1	DC	0.26	T	0.971	D	-1.882	N	0.999	D	-	-	0.34	4.72
SPTBN5 NM_016642.3	15:42147738:T:C	5.4	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	2.38
TRABD2B NM_001194986.1	1:48260308:G:A	23	LP	0.76	DC	0.1	T	0.949	D	-2.05	N	0.692	P	-	-	2.99	4.6
FAM83A NM_032899.6	8:124195260:G:A	16.2	VUS	1	POL	0.119	T	0.246	N	-1.43	N	0.134	B	-	-	0.48	3.43
NR2E1 NM_001286102.1	6:108502797:G:A	23.2	VUS	1	DC	0.5	T	0.998	D	-0.783	N	0.803	P	-	-	2.6	5.97
FASTKD5 NM_021826.4	20:3127550:G:A	27.1	LP	0.99	DC	0	D	0.942	D	-6.64	D	1	P	-	-	0.56	3.7

Table 10. Continued

PIK3CD NM_005026.3	1:9780028:A:G	26.7	P	0.99	DC	0.009	D	0.981	D	-2.52	D	0.992	P	-	-	4.27	5.5
SNTB2 NM_006750.3	16:69304210:GA:G	1.7	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	-2.5
NPC1 NM_000271.4	18:21121278:G:A	28.6	P	1	DC	0	D	0.897	D	-7.78	D	1	P	-	-	1.09	5.0
DENND4A NM_001320835.1	15:65993539:T:G	19.7	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	3.3
DYM NM_017653.3	18:46784867:T:C	14	VUS	-	-	-	-	-	-	-	-	-	-	0	NSA	-	1.0
SLC2A10 NM_017653.3	20:45354351:G:T	32	P	0.99	DC	0.002	D	0.999	D	-4.89	D	1	P	-	-	-0.75	5.86
AARS1 NM_001605.2	16:70287258:G:T	16.6	VUS	0.99	DC	0.663	T	0.929	D	-0.404	N	0.001	B	-	-	0.05	5.4
ESRRG NM_001243518.1	1:216850763:G:T	26.6	P	1	DC	0.002	D	0.999	D	-4.63	D	0.999	P	-	-	2.56	6.1
COL6A3 NM_004369.3	2:238249568:A:C	26.2	P	1	DC	0	D	0.999	D	-5.49	D	1	P	-	-	-0.61	5.1
AFF3 NM_001025108.1	2:100199274:C:T	12.9	VUS	1	POL	0.72	T	0.041	N	1.236	N	0	B	-	-	1.71	5.9

Colour Scheme: deleterious pathogenic results (red), likely pathogenic (orange), uncertain significance (grey), benign pathogenic results (green).

[1] **CADD**. Combined Annotation Dependent Depletion. The variant is pathogenic above 15 (Kircher et al., 2014).

[2] **Varsome**. American College of Medical Genetics (ACMG) classification: Pathogenic (P), Likely Pathogenic (LP), VUS (Variant of Uncertain Significance), Likely Benign (LB), Benign (B) (Kopanos et al., 2019; Richards et al., 2015).

[3] **MutationTaster**. Scores close to 1 indicate a high 'security' of the damaging or polymorphism prediction. Polymorphisms (POL) and Disease Causing (DC) (Schwarz et al., 2014).

[4] **SIFT**. Sorting Intolerant from Tolerant. Scores range between 0 [Damaging (D)] and 1 [Tolerant (T)] (Sim et al., 2012).

[5] **FATHMM**. Functional Analysis through Hidden Markov Models. Scores range -16.13 (benign) and 10.64 (pathogenic) (Shihab et al., 2015).

[6] **PROVEAN**. Protein Variation Effect Analyzer. If the PROVEAN score is equal to or below a predefined threshold (-2.5), the protein variant is predicted to have a "deleterious" (D) effect. If the PROVEAN score is above the threshold, the variant is predicted to have a "neutral" (N) effect. (Y. Choi et al., 2015).

[7] **PolyPhen2**. Polymorphism Phenotyping v2. Scores range between 0 [benign (B)] and 1 [pathogenic (P)] (Adzhubei et al., 2010).

[8] **SpliceAI**: probability of the variant being splice-altering (SA) or not splice-altering (NSA). Cut-offs range from 0.2 (high recall), 0.5 (recommended), and 0.8 (high precision)(Gelfand et al., 2013).

[9] **Constraint**: Positive Z-scores indicate more constraint (fewer observed variants than expected), and negative scores indicate less constraint (more observed variants than expected). Threshold score >3.09 is equivalent to constrained transcript (Karczewski et al., 2020; Lek et al., 2016).

[10] **GERP**: Genomic Evolutionary Rate Profiling. Scores range between -12.3 and 6.17. Scores higher than 2 can be considered conserved (Cooper et al., 2005).

[11] Score

[12] Prediction

3. Analysis of Autosomal Recessive Inheritance and Polymorphisms

As we could not demonstrate the rare *de novo* hypothesis, we reformulated the hypothesis and performed a data analysis considering a recessive mode of inheritance. This strategy allowed us to explore homozygous and bi-allelic compound heterozygous variants as other possibilities of the genetic cause of NAIS. Due to the recessive approach, variant identification and filtering were performed following the same established workflow as in *de novo* autosomal dominant hypothesis but considering the nature of the homozygous and bi-allelic compound heterozygous variants (Figure 19). Our results showed 168 homozygous and 240 compound heterozygous variants annotated with a mean of 10.4 ± 4.8 and 7.3 ± 4.5 from homozygous and compound heterozygous variants, respectively (Figure 20). However, none of the variants met all the filter criteria in variant identification and analysis; thus, no variant continued the pipeline for candidate gene selection (data no shown).

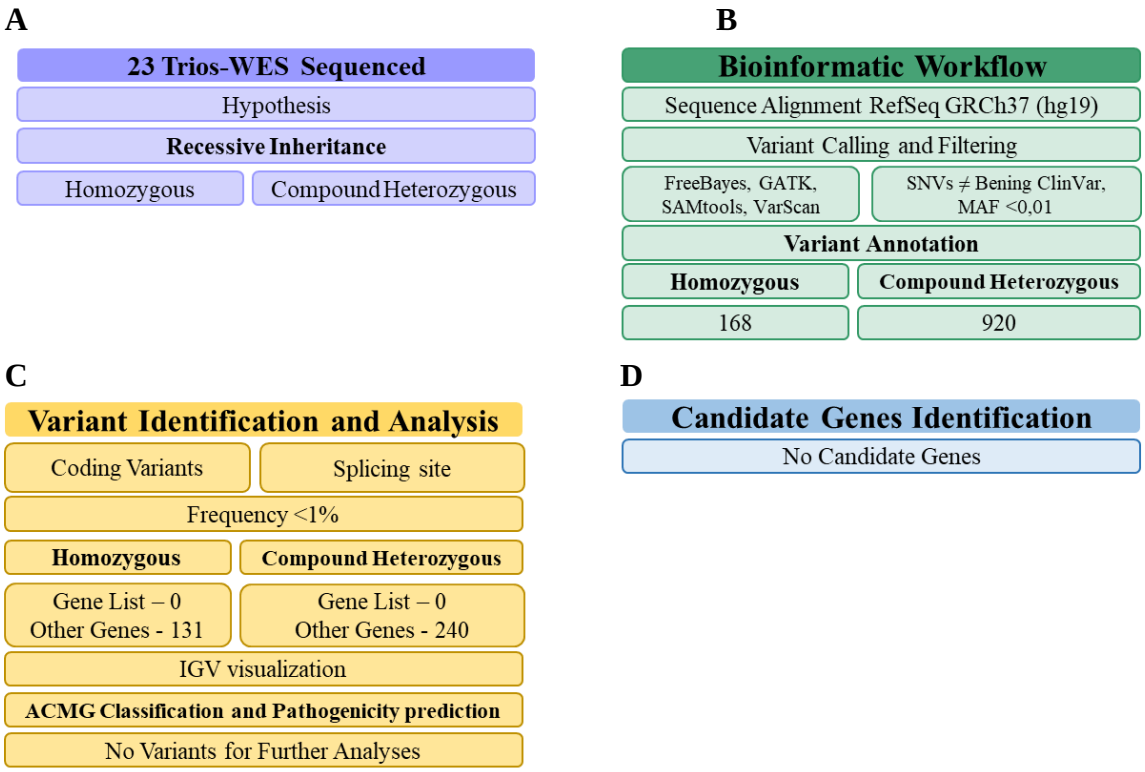


Figure 19. **Workflow results for the recessive inheritance approach**. **A**) Subjects sequenced for analysis of homozygous and compound heterozygous variants. **B**) Bioinformatic pipeline for variant filtering and annotation prior variant identification and analysis. **C**) Custom variant filtering performed for recessive inheritance analysis and identification of possible candidate genes. **D**) No candidate gene was selected for the candidate gene workflow application.

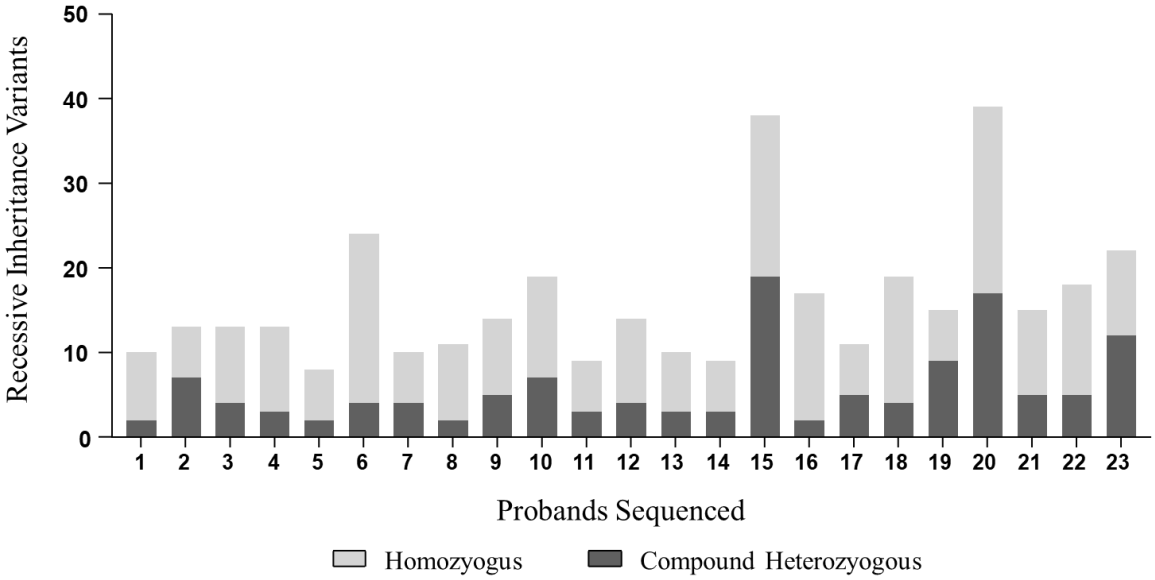


Figure 20. **Recessive mode of inheritance variants analysed by filter criteria**. Number of homozygous variants (black) and compound heterozygous variants (grey) present in the 23 probands.

Regarding the polymorphisms that have been studied in prothrombotic states and individuals with stroke (Table 11), the genotyping results showed that there were no cases of homozygous variants for the *F2*, *F5* Leiden, *APOE*, *F13A1*, *ITGB3,* and *MTHFR* (rs1801133) polymorphisms (Table 12). However, there was one homozygous case for the *MTHFR* rs1801131 polymorphism, which is considered abnormal. Furthermore, to avoid false positive results the genes in both autosomal dominant and autosomal recessive inheritances were compared with frequently mutated genes (FLAGS) in public exomes (Shyr et al., 2014). No relevant results were obtained that could be used to totally or partially explain the genetics of NAIS (data no shown).

Table 11. Detailed genetic information of prothrombotic state polymorphisms

Gene RefSeq	RefSNP	Variant (GRCh37)	Protein Change	DNA Change
F2 NM_000506.5	rs1799963	11:46761055:G:A	N/A[1]	c.*96C>T [2]
F5 NM_000130.5	rs6025	1:169519049:T:C	p.Gln534Arg	c.1601G>A[3]
MTHFR NM_005957.5	rs1801133	1:11856378:G:A	p.Ala222Val	c.665C>T[4]
MTHFR NM_005957.5	rs1801131	1:11854476:T:G	p.Glu429Ala	c.1286A>C
APOE NM_001302688.2	rs7412	19:45412079:C:T	p.Arg202Cys	c.604C>T
APOE NM_001302688.2	rs429358	19:45411941:T:C	p.Cys156Arg	c.466T>C
F13A1 NM_000129.4	rs5985	6:6318795:C:A	p.Val35Leu	c.103G>T
F13A1 NM_000129.4	rs3024477	6:6251120:T:A	p.Tyr205Phe	c.614A>T
ITGB3 NM_000212.3	rs5918	17:45360730:T:C	p.Leu59Pro	c.176T>C

[1] Not Applicable
[2] Commonly known as factor II G20210A
[3] Commonly known as factor V G1691A
[4] Commonly known as *MTHFR* C677T

Table 12. Genotyping results of prothrombotic state polymorphisms

Polymorphism Gene/RefSNP	Genotype NAIS Cases n (%)			MAF[1] %	
F2	**GG**	**GA**	**AA**	**G**	**A**
rs1799963	23 (100)	0 (0)	0 (0)	99.2	0.8
F5	**CC**	**CT**	**TT**	**C**	**T**
rs6025	23 (100)	0 (0)	0 (0)	98	2
MTHFR	**GG**	**GA**	**AA**	**G**	**A**
rs1801133	11 (47.8)	12 (52.2)	0 (0)	69.2	30.8
MTHFR	**TT**	**TG**	**GG**	**T**	**G**
rs1801131	15 (65.2)	7 (30.4)	1 (4.4)	71.4	44010
APOE	**CC**	**CT**	**TT**	**C**	**T**
rs7412	20 (86.9)	3 (13.1)	0 (0)	93.5	6.5
APOE	**TT**	**TC**	**CC**	**T**	**C**
rs429358	20 (86.9)	3 (13.1)	0 (0)	85.8	14.2
F13A1	**CC**	**CA**	**AA**	**C**	**A**
rs5985	13 (56.5)	10 (43.5)	0 (0)	79.4	20.6
F13A1	**TT**	**TA**	**AA**	**T**	**A**
rs302447	22 (95.6)	1 (4.4)	0 (0)	97.6	2.4
ITGB3	**TT**	**TC**	**CC**	**T**	**C**
rs5918	20 (86.9)	3 (13.1)	0 (0)	87.8	12.2

[1] Minor Allele Frequency

4. Identification of Common Features among Candidate Genes

Determining the strongest candidate gene possibly associated with NAIS requires complementary strategies to collect relevant evidence. We decided to analyse all 28 genes identified with confirmed *de novo* variants (*EPS15L1, SYNE2, FAM193A, HECTD1, TOP2A, RBM7, KRT6A, UNC13D, FAR2, SHANK3, PCBP1, KCNC2, CHRNA4, SPTBN5, TRABD2B, FAM83A, NR2E1, FASTKD5, PIK3CD, SNTB2, NPC1, DENND4A, DYM, SLC2A10, AARS1, ESRRG, COL6A3, AFF3*) to find any indication of similar processes involved, biological and molecular function reported in databases, protein-protein interactions (PPI), and cell signalling pathways.

Furthermore, to explore a connection between the cohorts' gene set and genes associated with hypercoagulability states, either directly or indirectly, the six most relevant genes related to hypercoagulability states (*F2, F5, MTHFR, APOE, F13A1*, and *ITGB3*) were added for PPI analysis. The PPI analysis results obtained by StringDB, based on experimental data, gene fusion, neighbourhood, co-expression, and co-occurrence, showed no reported relevant interaction between all 28 proteins apart from the COL6A3-ITGB3 interaction (Figure 21A). PathwayCommons, another database that collects and disseminates pathways and biological interactions (Cerami et al., 2011), revealed, from Molecular Signatures Database, expression data between PIK3CD and EPS15L1 (Figure 21B) and one interaction between AFF3 and PCBP1 according to the Reactome database. Moreover, further analysis performed with GeneMania (Warde-Farley et al., 2010), revealed co-expression results without any significant relevance (Figure 21C).

Likewise, signalling pathways and gene ontologies (GO) terms related to NAIS were analysed to find possible relationships between all 28 possible candidate genes (data no shown). Our results indicated that regulation of gene expression was the most common feature shared by *DENND4A, SLC2A10, PCBP1, PIK3CD, ESRRG, AFF3*, in addition to cell migration by *PIK3CD* and *SYNE2*. Both *PIK3CD* and *COL6A3* shared the PI3K-Akt pathway and, together with the *SPTBN5*, the Platelet-derived Growth Factor signalling pathway. *HECTD1* and *SLC2A10* showed artery development as the main molecular GO terminus. Due to the absence of *FAM193A* and *KRT6A* information from different gene ontology databases, this GO analysis was somewhat limited.

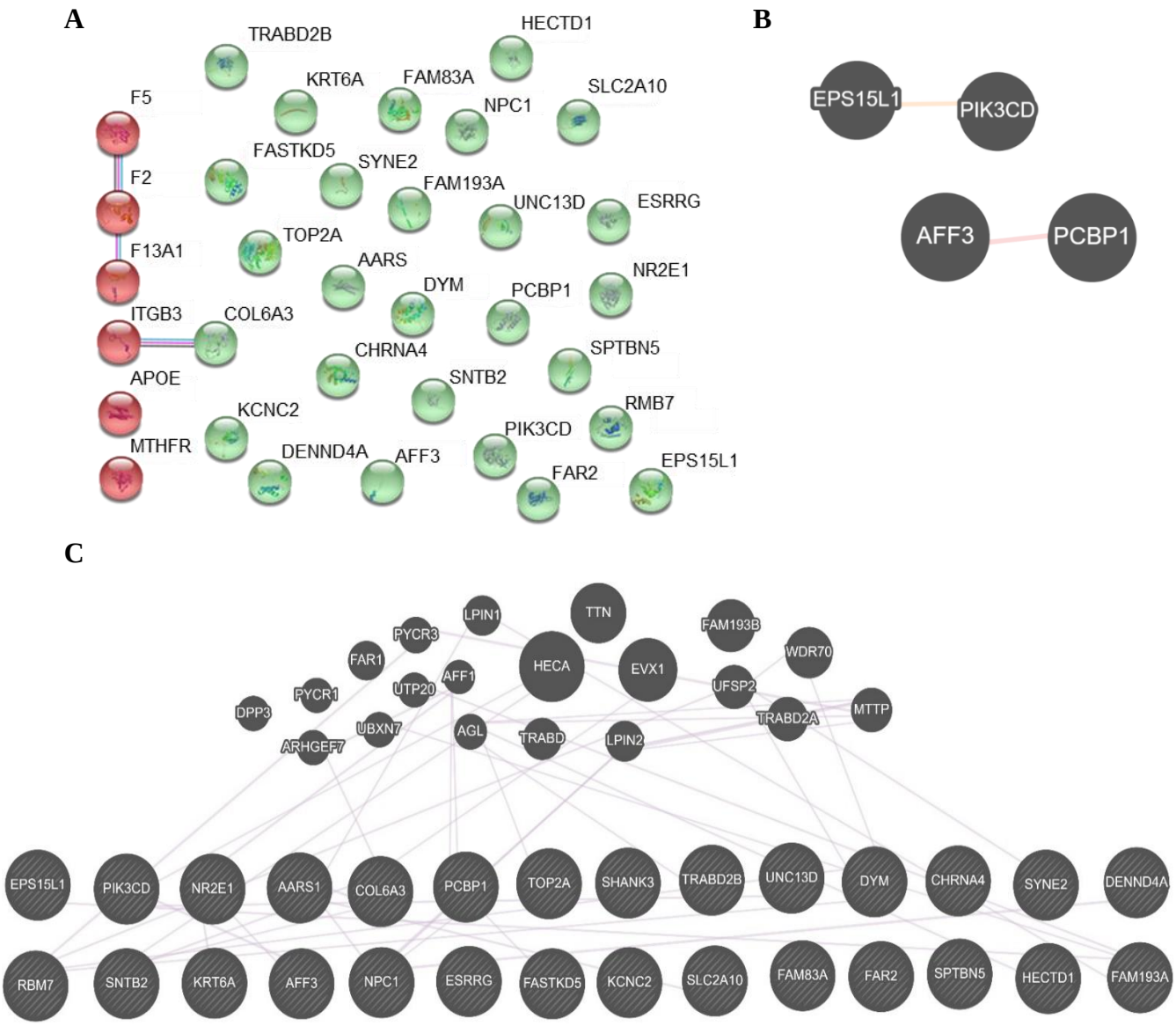


Figure 21. **Protein-protein interaction networks**. Interaction prediction between all 28 possible gene candidates identified and the genes associated with prothrombotic states. **A**) Representation of protein-protein interaction between the set of proteins analysed by StringDB. Green nodes represent the proteins encoded by all 28 genes identified in our study (SHANK3 is missing from the database). Red nodes represent the proteins encoded by the genes associated with prothrombotic states. The pink lines represent evidence of experimental interaction. Black lines represent co-expression interactions. Blue lines represent evidence from databases. **B**) Protein-protein interaction performed using Pathway Commons. The pink line represents expression data from the Molecular Signatures Database. The orange line represents the Reactome modification data **C**) GeneMania Interaction network showing co-expression results (violet lines) between all 28 genes identified (big circles below) and other genes that are not part of the study (small circles above).

5. Analysis of Clinically Homogeneous Subsets of NAIS

Given that the evaluation of different scenarios is useful for gene-disease associations we grouped all 23 NAIS patients in subgroups according to the definition of the affected arterial territory during the stroke event. This NAIS homogenous subgroup analysis was carried out to increase the power of the gene-disease association and identify a potential common genetic determinant. Additionally, clinically homogeneous subgroups are assumed to be genetically homogeneous as well. The results revealed that the most frequent subgroup of NAIS was M1 (44%), which involves the territory before and after the MCA bifurcation (27%-M1 pre + 17% M1 post) and followed by infarcts involving the territory M2 (30%). Therefore, 74% of the patients had infarcts of considerable size (Table 8).

Regarding the analysis of cell signalling pathways of each clinical subgroups (Table 13), we did not find any relevant relationship that could totally or partially explain the physiopathology behind each affected arterial territory. These results involve the genes/proteins identified in each subset, the volume of the lesion, and the outcome at two years of age. Although further analyses are required, this corroborates the exclusivity and heterogeneity of each accidental genetic finding in all 23 infants with NAIS. Moreover, the single infant with M3 infarct was one of the probands without DNV that met the established filter criteria.

Table 13. Identified genes, protein functions, and cell signalling pathways based on clinical subgroups of NAIS

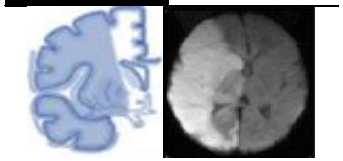
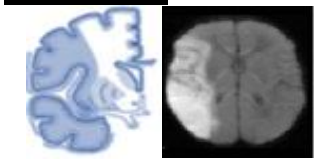
ATD[1] Scheme / MRI[2]	Gene Protein	Protein Function	Pathways
M1 Pre 	***FAM193A*** Protein FAM193A	No Data Available	No Data Available
	HECTD1 E3 ubiquitin-protein ligase HECTD1	Ligase which mediates 'Lys-63'-linked polyubiquitination of HSP90AA1	- Antigen processing
	TOP2A DNA topoisomerase 2-alpha	Decatenating enzyme that alters DNA topology by binding to two double-stranded DNA molecules	- SUMOylation - DNA replication - G0 and Early G1
	PCBP1 Poly(rC)-binding protein 1	Single-stranded nucleic acid binding protein that binds preferentially to oligo dC. In case of infection by poliovirus, plays a role in initiation of viral RNA replication in concert with the viral protein 3CD	- Processing of Capped Intron Containing - Pre-mRNA pre-mRNA splicing
	SNTB2 Beta-2-syntrophin	Adapter protein organizes the subcellular localization of a variety of membrane proteins. May link various receptors to the actin cytoskeleton and the dystrophin glycoprotein complex.	No Data Available
	ESRRG Estrogen Related receptor gamma	Orphan receptor that acts as transcription activator in the absence of bound ligand. Binds specifically to an estrogen response element and activates reporter genes controlled by estrogen response elements	- Nuclear Receptor transcription pathway
	AARS1 Fatty Acyl-CoA Reductase 2	Catalyzes the reduction of saturated but not unsaturated C16 or C18 fatty acyl-CoA to fatty alcohols.	- Metabolisms - Wax biosynthesis
M1 Post 	***FAR2*** Alanine-tRNA ligase, cytoplasmic	Catalyses the attachment of alanine to tRNA(Ala) in a two-step reaction	- Gene Expression -tRNA Aminoacylation
	SHANK3 SH3 and multiple ankyrin repeat domains protein 3	Major scaffold postsynaptic density protein which interacts with multiple proteins and complexes to orchestrate the dendritic spine and synapse formation, maturation, and maintenance	- Signalling events regulated by Ret tyrosine kinase - RET signalling - Neurexins and neuroligins Ionotropic glutamate receptor pathway
	CHRNA4 Neuronal acetylcholine receptor subunit α4	Nicotinic acetylcholine receptor, which belongs to a superfamily of ligand-gated ion channels that play a role in fast signal transmission at synapses	- CREB Pathway - Activation of Nicotinic Acetylcholine Receptors
	SPTBN5 Spectrin β chain, non-erythrocytic 5	Heterodimeric molecule that is the central component in a ubiquitous and complex system linking membrane proteins	- NCAM signalling for neurite out-growth - RAF/MAP kinase cascade - COPI-mediated anterograde transport

Table 13. Continued (1)

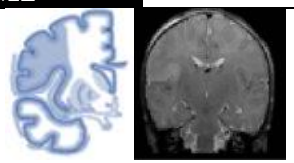
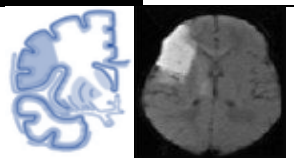
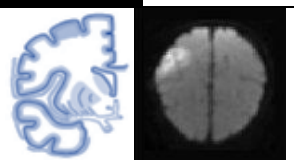
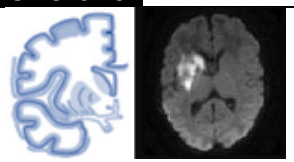
	COL6A3 Collagen alpha-3 (VI) chain	Cell-binding protein	- Signalling by PDGF - Collagen chain trimerization - Integrin signalling pathway
	AFF3 AF4/FMR2 family member 3	Putative transcription activator that may function in lymphoid development and oncogenesis	- Signalling by FGFR2
M2 	***EPS15L1*** Epidermal growth factor receptor substrate 15-like 1	Constitutive component of clathrin-coated pits that is required for receptor-mediated endocytosis	- Signalling by EGFR - Clathrin-mediated endocytosis
	SYNE2 Nesprin-2	Multi-isomeric modular protein, which is a component of the LINC. Forms a linking network between organelles and the actin cytoskeleton to maintain the subcellular spatial organization.	- Meiotic synapsis
	KCNC2 Potassium voltage-gated channel subfamily C member 2	Voltage-gated potassium channel that mediates transmembrane potassium transport in excitable membranes, primarily in the brain	- Voltage gated Potassium channels
	TRABD2B Metalloprotease TIKI2	Metalloprotease that acts as a negative regulator of the Wnt signalling pathway	No Data Available
	FAM83A Protein FAM83A	Probable proto-oncogene that functions in the EGFR signalling pathway.	No Data Available
	NR2E1 Nuclear receptor subfamily 2 group E member 1	Orphan receptor that binds DNA as a monomer to hormone response. May be required to pattern anterior brain differentiation	- Regulation of PTEN gene transcription - Nuclear Receptor transcription pathway - PI3K/AKT Signalling
	NPC1 NPC intracellular cholesterol transporter 1	Intracellular cholesterol transporter which may play a role in vesicular trafficking in glia	- LDL clearance
M3 	Without DNV meeting the filtering criteria		

Table 13. Continued (2)

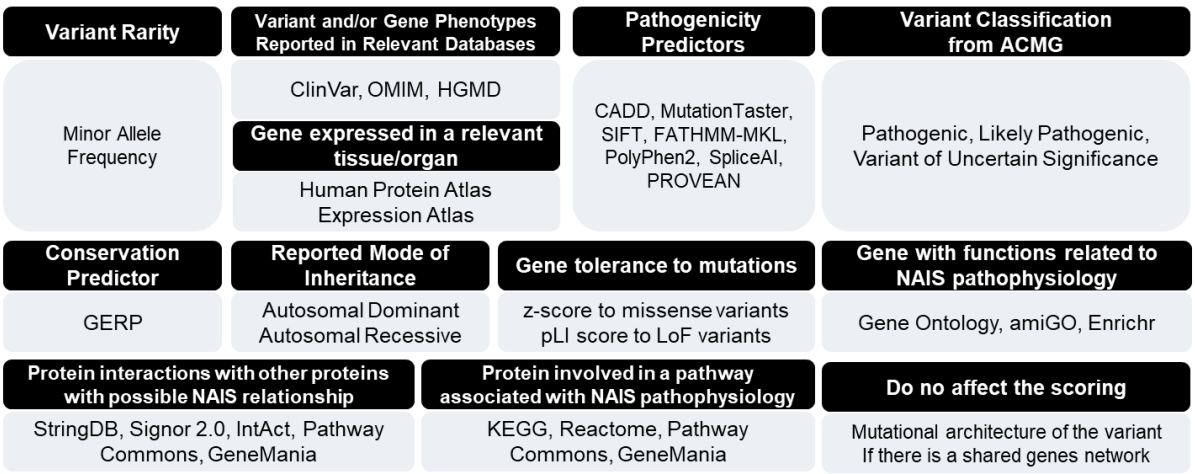
M4 	***RBM7*** RNA-binding protein 7	RNA-binding subunit of the trimeric nuclear exosome targeting (NEXT) complex, a complex that functions as an RNA exosome cofactor	No Data Available
	KRT6A Keratin, type II cytoskeletal 6A	Epidermis-specific type I keratin involved in wound healing	- EGFR1 - Keratinization
	UNC13D Protein unc-13 homolog D	Cytotoxic granule exocytosis in lymphocytes. Required for both granule maturation and granule docking and priming at the immunologic synapse	- Synaptic vesicle trafficking - Neutrophil degranulation
	DENND4A C-myc promoter-binding protein	Probable guanine nucleotide exchange factor (GEF) which may activate RAB10. Promotes the exchange of GDP to GTP	- RAB GEFs exchange GTP for GDP on RABs
	DYM Dymeclin	Necessary for correct organization of Golgi apparatus. Involved in bone development.	No Data Available
	SLC2A10 Solute carrier family 2, facilitated glucose transporter member 10	Facilitative glucose transporter required for the development of the cardiovascular system	- Cellular hexose transport - NRF2 pathway
Perforant 	***FASTKD5*** FAST kinase domain-containing protein 5, mitochondrial	Plays an important role in the processing of non-canonical mitochondrial mRNA precursors	No Data Available
	PIK3CD Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit delta isoform	Phosphorylates phosphatidylinositol (PI) and its phosphorylated derivatives at position 3 of the inositol ring to produce 3-phosphoinositides. Mediates immune responses	- PI3K/AKT Signalling - Inositol phosphate metabolism - Immune system - Angiogenesis

[1] Arterial Territory Definition
[2] Magnetic Resonance Imaging

6. Knowledge-Driven Analysis Pipeline

In many cases, even using a strong experimental designs and variant filtering workflows, it is not feasible to verify through functional studies all the possible candidate genes obtained in a study. As we cannot perform functional studies in all 28 DNV we applied the Knowledge-driven analysis (KDA), which considers individual and common quantitative and qualitative information about the genes studied. In addition, we established a customised score specific for our study that allowed us to prioritise further our gene set (Table 14). This score ranges from 1 to 21, therefore, the variant with the highest score is considered the most pathogenic variant and consequently the most suitable gene for further functional studies. Furthermore, for a purely visual purpose, we added a gradient colour scale scheme in which the green colour and its gradient towards yellow indicates the section where the variants considered more benign are found, while the gradient between yellow and red indicates the presence of the most pathogenic variants (Figure 22).

A



B

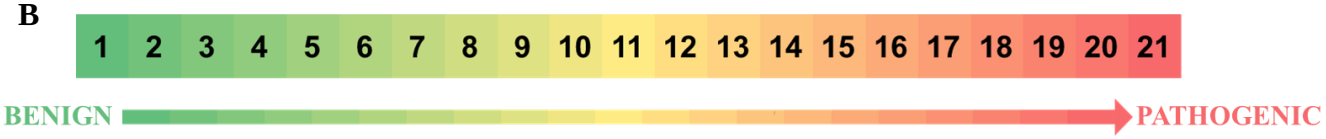


Figure 22. Established criteria and score for the Knowledge-Driven Analysis Pipeline. Customised pipeline for the application of KDA in candidate gene selection. **A**) Information used for the analysis, score settling, and selection of the most suitable candidate gene within all 28 DNV identified. **B**) Analysis score range and its respective colour gradient.

Table 14. Criteria and custom scoring used in KDA for ranking and prioritization of variants

Criteria	*In silico* tool	Result	Score
Variant rarity (MAF)	gnomAD, CSVS	Variant Not Found	3
		Frequency 0	2
		Frequency <0.01	1
Mutational architecture of the variant	Type of variant position within the gene	No scoring. For information only	-
Variant and/or phenotype reported in relevant databases	ClinVar, HGMD, OMIM	Yes	1
		No	0
Pathogenicity predictors	CADD	≥30	2
		15-29,9	1
		<15	0
	MutationTaster	Disease Causing	1
		Polymorphism	0
	SIFT	Intolerant	1
		Tolerant	0
	FATHMM-MKL	Damaging	1
		Neutral	0
	PROVEAN	Damaging	1
		Neutral	0
	PolyPhen-2	Possibly Damaging	1
		Benign	0
	Splice AI	Splicing alteration	1
		No alteration	0
Variant classification ACMG	Varsome	Pathogenic	2
		Likely Pathogenic	1
		VUS	0
Position conserved	GERP	>2	1
		<2	0
Mode of inheritance	Autosomal dominant or Autosomal recessive	AD	1
		AR	0
Gene tolerant/intolerant to mutations	Z-score	>3.09	1
	or	<3.09	0
	pLI score	<0.5	1
		>0.5	0
Gene with functions related to NAIS pathophysiology	Gene Ontology, AMIGO Enrichr	Yes	1
		No	0
Gene expressed in a relevant tissue/organ	Human Protein Atlas, Expression Atlas	Yes	1
		No	0
Encoded protein involved in a pathway associated with NAIS pathophysiology	KEGG, Reactome, PathwayCommons, GeneMania	Yes	1
		No	0
Protein interactions with other proteins with possible relationship with NAIS	StringDB, Signor 2.0, IntAct	Yes	1
		No	0
Genes identified among probands with relevant similarities among them	Enrichr	No scoring. For information only	-

7. Selection of Candidate Gene

The evaluation of individual genes is an important step to manage further validation and exploration of genetic variants. All 28 potential candidate genes were comprehensively screened using our established custom KDA pipeline. The evidence showed that the c.1292A>G variant in *PIK3CD*, carried by a patient who suffered a perforant stroke, obtained the highest score (18) among the entire set (Table 15) followed by the c.7991T>G variant in *COL6A3* (score of 16). Although *COL6A3* is only two points downstream of *PIK3CD* and involved in some cell signalling pathways possibly linked to NAIS, it was omitted as a candidate gene due to the already associated muscular phenotypes (MIM #158810, MIM #616411, and MIM #254090), the different inheritance patterns reported in databases, the gene ontology of its direct interactors, and the lack of evidence to support its possible association with NAIS (Table 16).

Regarding the direct interactors of PIK3CD, many of them have biological gene ontology terms associated with endothelium and coagulation (Table 17). In conclusion, the compendium of all the evidence on PIK3CD regarding protein functions and protein-protein interactions, potential cell signalling pathways associated with the pathophysiology of NAIS (Figure 23), and the scoring through the KDA pipeline, indicates that PIK3CD is the most suitable candidate gene for carrying out functional studies. This could allow us to discern whether the pathogenicity impact of the variant has a relevant disruptive effect on the protein function and whether there is a possible association with the pathophysiology or susceptibility of NAIS.

In addition, the KDA pipeline analysis indicated that the frequency of 19 (67.8%) DNV were not reported in the gnomAD database, 8 (28.6%) had a frequency in the European (non-Finnish) population of <0.0001, and 1 (3.6%) was reported in the database but only in the Ashkenazi Jewish population (data no shown).

Table 15. Candidate gene ranking score results of all 28 confirmed DNVs after KDA analysis

Patient[1]	Gene	F[2]	CADD	VAR	MT	SIFT	F-MKL	PVN	PP-2	SAI	DB[3]	GERP	AD/AR	Z/pLI	GO[4]	TE[5]	PW[6]	PPI[7]	Score
16	***PIK3CD***	**3**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**-**	**1**	**1**	**1**	**1**	**1**	**0**	**1**	**1**	**18**
22	*COL6A3*	3	1	2	1	1	1	1	1	-	1	1	1	0	0	0	1	1	**16**
19	*SLC2A10*	3	2	2	1	1	1	1	1	-	1	1	0	0	1	0	0	1	**16**
5	*FAR2*	3	2	2	1	1	1	1	1	-	0	1	-	0	0	1	0	0	**14**
18	*NPC1*	3	1	2	1	1	1	1	1	-	1	1	0	0	0	1	0	0	**14**
21	*ESRRG*	1	1	2	1	1	1	1	1	-	0	1	-	0	0	1	1	1	**13**
3	*RBM7*	3	1	0	1	1	1	1	1	-	0	1	-	0	0	1	-	0	**11**
1	*SYNE2*	3	1	0	1	0	1	0	1	-	1	1	1	0	0	1	0	0	**11**
20	*AARS1*	3	1	0	1	0	1	0	0	-	1	1	1	0	0	0	0	1	**10**
13	*TRABD2B*	2	1	1	1	0	1	0	1	-	0	1	-	0	1	1	-	0	**10**
16	*FASTKD5*	1	1	1	1	1	1	1	1	-	0	1	-	0	0	1	-	0	**10**
2	*FAM193A*	3	1	2	1	-	-	0	0	-	0	1	-	1	0	1	-	0	**10**
6	*PCBP1*	3	1	0	1	1	0	1	0	-	0	1	-	1	0	0	0	0	**9**
5	*SHANK3*	3	0	0	-	-	-	-	-	0	1	0	1	-	1	1	0	1	**8**
10	*CHRNA4*	1	1	0	1	0	1	0	1	-	0	1	-	0	0	1	0	0	**7**
15	*NR2E1*	1	1	0	1	0	1	0	1	-	0	1	-	0	0	0	0	1	**7**
9	*KCNC2*	3	1	0	0	1	0	0	0	-	0	1	-	0	0	0	0	0	**6**
19	*DENND4A*	3	1	0	-	-	-	-	-	0	0	1	-	-	0	0	1	0	**6**
1	*EPS15L1*	3	0	0	-	-	-	-	-	0	0	0	-	-	0	1	1	0	**5**
15	*FAM83A*	3	1	0	0	0	0	0	0	-	0	1	-	0	0	0	-	0	**5**
12	*SPTBN5*	3	0	0	-	-	-	-	-	0	0	1	-	-	1	0	0	0	**5**
2	*TOP2A*	3	0	0	-	-	-	-	-	0	1	0	-	-	0	0	-	1	**5**
19	*DYM*	3	0	0	-	-	-	-	-	0	1	0	0	0	0	0	-	0	**4**
3	*UNC13D*	3	0	0	-	-	-	-	-	0	1	0	0	-	0	0	-	0	**4**
2	*HECTD1*	1	0	0	-	-	-	-	-	0	0	0	-	-	1	1	0	1	**4**
17	*SNTB2*	3	0	0	-	-	-	-	-	0	0	0	-	-	0	0	-	0	**3**
3	*KRT6A*	1	0	0	-	-	-	-	-	0	1	0	1	-	0	0	-	0	**3**
22	*AFF3*	1	0	0	0	0	0	0	0	-	0	0	-	0	0	0	-	0	**1**

[1] Consecutive number given to the patient sequenced
[2] Frequency
[3] Databases (OMIM, HGMD, ClinVar)
[4] Gene Ontology
[5] Tissue Expression
[6] Pathway
[7] Protein-Protein Interactions

Table 16. Additional information on the top two ranked genes for candidate gene selection

Gene	Protein Encoded			Gene Ontology Functions		Pathways	Reported Phenotypes	
	Name	Subcellular Localization	Tissue Expression	Molecular	Biological		Phenotype MIM#[5]/Inheritance	HGMD[6]
PIK3CD	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic	Vesicles & Cytokinetic bridge	Testis Appendix Lymph node	Phosphatidylinositol phosphate kinase activity (GO[1]:0016307)	Phosphatidylinositol-3-phosphate biosynthetic process (GO:0036092)	VEGF signalling pathway Ras signalling pathway	?Roifman-Chitayat syndrome (613328/ DR[2])	Activated PI3Kδ syndrome
		Plasma membrane	Tonsil Nasopharynx Bronchus Gallbladder	Phosphatidylinositol kinase activity (GO:0052742)	Positive regulation of cell migration by vascular endothelial growth factor signalling pathway (GO:0038089)	ErbB signalling pathway PI3K-Akt signalling pathway	Immunodeficiency 14A (615513/AD[3])	Macrophage activation syndrome/haemophagocytic lymphohistiocytosis
			Epididymis Spleen Soft tissue	Phosphotransferase activity, alcohol group as acceptor (GO:0016773)	Axonogenesis (GO:0007409)	Phosphatidylinositol signalling system	Immunodeficiency 14B (619281/AR[4])	T cell senescence & immunodeficiency
					B cell activation (GO:0042113)	Longevity regulating pathway		Immunodeficiency, primary
					Inflammatory response (GO:0006954)	Toll-like receptor signalling pathway		Hyper-IgM syndrome
					Positive regulation of endothelial cell migration (GO:0010595)	HIF-1 signalling pathway		Roifman-Chitayat syndrome
					Positive regulation of endothelial cell proliferation (GO:0001938)	Cholinergic synapse		Impaired B cell maturation
					Positive regulation of vasculature development (GO:1904018)	Neurotrophin signalling pathway		T cell senescence & immunodeficiency
					Regulation of endothelial cell migration (GO:0010594)	AMPK signalling pathway		
					Vascular endothelial growth factor signalling pathway (GO:0038084)	Platelet activation		
						Fluid shear stress and atherosclerosis cAMP signalling pathway		

Table 16. Continued

Gene	Protein			Gene Ontology Functions		Pathways	Phenotypes	
	Encoded	**Subcellular Localization**	**Tissue Expression**	**Molecular**	**Biological**		**Phenotype MIM#/Inheritance**	**HGMD**
COL6A3	Collagen Type VI Alpha 3 Chain	Endoplasmic reticulum Vesicles	Ovary Bone marrow Caudate Stomach Cervix Smooth muscle Tonsil	Serine-type endopeptidase inhibitor activity Extracellular matrix structural constituent conferring tensile strength	Muscle organ development (GO:0007517) Collagen fibril organization (GO:0030199) Extracellular structure organization (GO:0043062) External encapsulating structure organization (GO:0045229) Extracellular matrix organization (GO:0030198)	ECM-receptor interaction Protein digestion and absorption Focal adhesion Human papillomavirus infection PI3K-Akt signalling pathway Assembly of collagen fibrils and other multimeric structures NCAM signalling	Bethlem myopathy (158810/AD, AR) Dystonia 27 (616411/AR) Ullrich congenital muscular dystrophy (254090/AD, AR)	Ullrich congenital muscular dystrophy Collagen VI myopathy Arthrogryposis Bethlem myopathy Atrioventricular septal defect, Down-syndrome-associated Muscular dystrophy, limb girdle 1A Intellectual disability/Myopathy/Muscular dystrophy Dystonia-27, isolated, early-onset Neuromuscular disorder

[1] Gene Ontology number
[2] Digenic Recessive
[3] Autosomal Dominant
[4] Autosomal Recessive
[5] Mendelian Inheritance in Man Number
[6] Human Mutation Database

Table 17. Gene ontology of PIK3CD interaction partners

Interacting Partner	Pathway and Ontology Annotation Evidence
HRAS	**VEGF signalling pathway**, MAPK signalling pathway, Pathways in cancer, positive regulation of actin cytoskeleton reorganization (GO[1]:2000251) positive regulation of JNK cascade (GO:0046330) regulation of MAP kinase activity (GO:0043405) positive regulation of ERK1 and ERK2 cascade (GO:0070374) positive regulation of MAPK cascade (GO:0043410)
NRAS	**VEGF signalling pathway**, MAPK signalling pathway, Pathways in cancer, **positive regulation of endothelial cell proliferation** (GO:0001938) **regulation of endothelial cell proliferation** (GO:0001936)
KRAS	**VEGF signalling pathway**, MAPK signalling pathway, Pathways in cancer, small GTPase mediated signal transduction (GO:0007264) stimulatory C-type lectin receptor signalling pathway (GO:0002223) cellular response to lectin (GO:1990858) innate immune response activating cell surface receptor signalling pathway (GO:0002220) Ras protein signal transduction (GO:0007265) positive regulation of protein modification process (GO:0031401) regulation of protein phosphorylation (GO:0001932) MAPK cascade (GO:0000165) positive regulation of protein phosphorylation (GO:0001934)
RRAS2	MAPK signalling pathway, Ras signalling pathway, Cellular senescence, Regulation of actin cytoskeleton, regulation of neuron death (GO:1901214) regulation of cell death (GO:0010941) small GTPase mediated signal transduction (GO:0007264) Ras protein signal transduction (GO:0007265)
RRAS	MAPK signalling pathway, **positive regulation of angiogenesis** (GO:0045766) **positive regulation of vasculature development** (GO:1904018) **regulation of angiogenesis** (GO:0045765) regulation of ERK1 and ERK2 cascade (GO:0070372) regulation of MAPK cascade (GO:0043408)
CDC42	**VEGF signalling pathway**, MAPK signalling pathway, Adherens junction, Pathways in cancer, **Wnt signalling pathway**, planar cell polarity pathway (GO:0060071) **non-canonical Wnt signalling pathway** (GO:0035567) establishment of Golgi localization (GO:0051683) viral genome replication (GO:0019079) dendrite morphogenesis (GO:0048813) regulation of plasma membrane bounded cell projection assembly (GO:0120032) interleukin-12-mediated signalling pathway (GO:0035722)
RHOG	Rho protein signal transduction (GO:0007266) vesicle-mediated transport (GO:0016192) regulation of small GTPase mediated signal transduction (GO:0051056) positive regulation of cellular process (GO:0048522) cell morphogenesis (GO:0000902) Rac protein signal transduction (GO:0016601)
PIK3R3 **PIK3R2** **PIK3R1**	**VEGF signalling pathway**, MAPK signalling pathway, **Platelet activation**, Pathways in cancer, Prolactin signalling pathway. **vascular endothelial growth factor receptor signalling pathway** (GO:0048010) positive regulation of intracellular protein transport (GO:0090316) MAPK cascade (GO:0000165) positive regulation of focal adhesion disassembly (GO:0120183) **cell migration involved in sprouting angiogenesis** (GO:0002042) **blood vessel endothelial cell migration** (GO:0043534) ERBB2 signalling pathway (GO:0038128) **sprouting angiogenesis** (GO:0002040) negative regulation of MAPK cascade (GO:0043409)
PIK3CA **PIK3CB** **PIK3CG**	**Platelet activation pathway**, **PI3K-Akt signalling pathway, VEGF signalling pathway**, Toll-like receptor signalling pathway, phosphatidylinositol-3-phosphate biosynthetic process (GO:0036092) **platelet aggregation** (GO:0070527) **vascular endothelial growth factor receptor signalling pathwa**y (GO:0048010) **positive regulation of endothelial cell migration** (GO:0010595) MAPK cascade (GO:0000165) **negative regulation of vascular endothelial growth factor signalling pathway** (GO:1900747) **negative regulation of sprouting angiogenesis** (GO:1903671) **vasculature development** (GO:0001944) ERBB signalling pathway (GO:0038127)

[1] Gene Ontology accession number

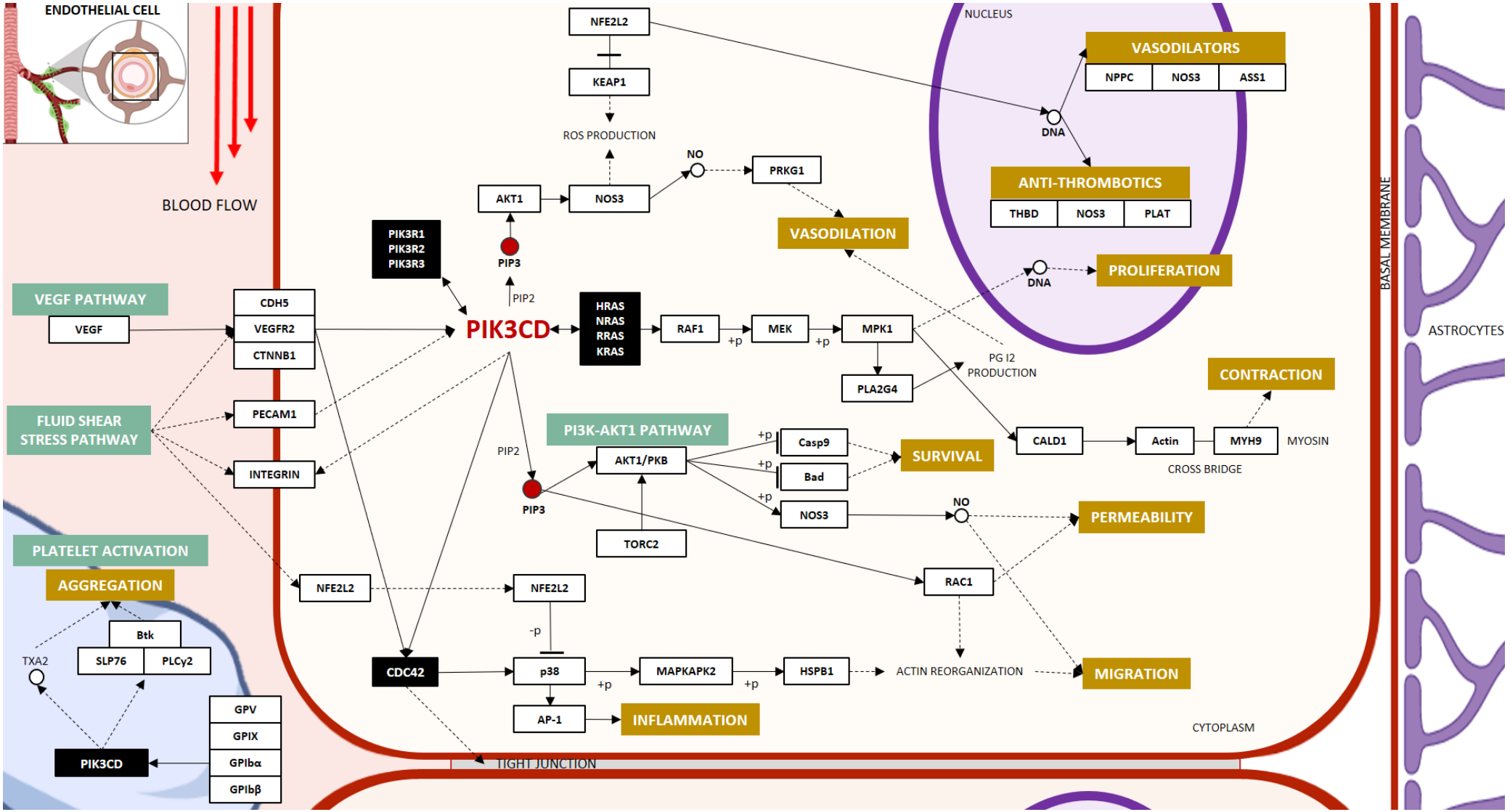


Figure 23. **Compendium of the described cell signalling pathways of PIK3CD**. Endothelial cell, blood flow, and a platelet with their respective PIK3CD-associated signalling pathways and the resulting biological functions (brown boxes) of the cascade processes. PIK3CD is shown in capital red letters, its main interactors in black boxes, its PIP3 product is showed in a red circle, and the rest of the proteins involved in the pathways are in white boxes.

8. The p.Gln431Arg variant *PIK3CD* and its Pathogenicity Impact

To decipher the pathogenicity impact of the p.Gln431Arg variant, we used several in silico tools that resulted in damaging predictions (Figure 24A). Varsome's verdict classifies the variant as Pathogenic due to the applied rules of PM2 (variant not found in frequency databases), PP3 (17 predictors with damaging results compared to 5 benign, data no shown) and PS2 (variant confirmed as *de novo*) (Figure 24B). Furthermore, the p.Gln431Arg variant frequency is missing in gnomAD, ClinVar, and in the CSVS databases (Figure 24C). In terms of amino acid conservation, glutamine (Gln) at position 431 is a well-conserved amino acid in many species (Figure 24D) and its substitution with arginine (Arg) changes the charges of the position (Gln is neutral and Arg is positive charged) (Figure 24E).

To delve into the pathogenicity prediction of the variant and assess whether it affects the PIK3CD protein, we conducted additional *in silico* analyses and functional studies. First, for the *in silico* analyses, we used mutation mapper (www.cbioportal.org/mutation_mapper) to analyse the position of the variant within the protein domains together with other reported variants in HGMD. The results showed that, of the five protein domains described in PIK3CD, the p.Gln431Arg variant is located as other variants reported within the phosphoinositide 3-kinase (PI3K_C2) domain region (Figure 25A). A protein domain which is involved in targeting proteins to cell membranes and corresponds to amino acid residues 338–466 of the protein.

Subsequently, we predicted the amino acid sequence disorder through PONDR VL-XT (Xue et al., 2010). The results of the prediction of the amino acid sequence disorder showed that the region near the amino acid residue 431 is ordered except for approximately 25 residues before this position (Figure 25B, upper panel). Additionally, the 287-312 amino acid region is predicted to be disordered. Something reported also in UniProt via the UniProt's Automatic Annotation pipeline. Additionally, because it has been described that proteins with intrinsically disordered regions may undergo liquid-liquid phase separation and that it is a sequence dependent dynamic (Tesei et al., 2021), we performed an aggregation propensity analysis using Aggrescan (Conchillo-Solé et al., 2007). The results indicated that the p.Gln431Arg variant does not increase the aggregation propensity of the protein (Figure 25B, middle panel). Finally, the ExPasy ProtScale platform analysis (Kyte & Doolittle, 1982) was used to evaluate hydrophobicity changes caused by Arg in the 431 position. The results showed no changes in the hydrophobicity of the protein (Figure 25B, lower panel).

Furthermore, secondary structure analysis of the protein is an essential step to predict additional alterations due to the presence of a new amino acid. For this purpose, we used the DynaMut2 platform (Rodrigues et al., 2021) for both the *wild type* and the p.Gln431Arg variant (Figure 26). The results showed that, based on Gibbs Free Energy (Kcal/mol), substitution of the glutamine residue for an arginine (Arg) at position 431, has a predicted stability change of -0.4 Kcal/mol (ΔΔGStability) indicating protein destabilization (Figure 26A). Hence, the p.Gln431Arg variant has a relevant pathogenic impact on the secondary structure of the PIK3CD protein. The mechanism by which this variant destabilises the protein is due to many changes in amino acid interactions. More precisely, the acquisition of an extra hydrophobic interaction between Arg431 and Tyr484, the substitution of a hydrogen bond with Val483 for two polar interactions, the substitution of a hydrogen bond for an ionic interaction with Asp427, and the elimination of two polar interactions (Figure 26B,C).

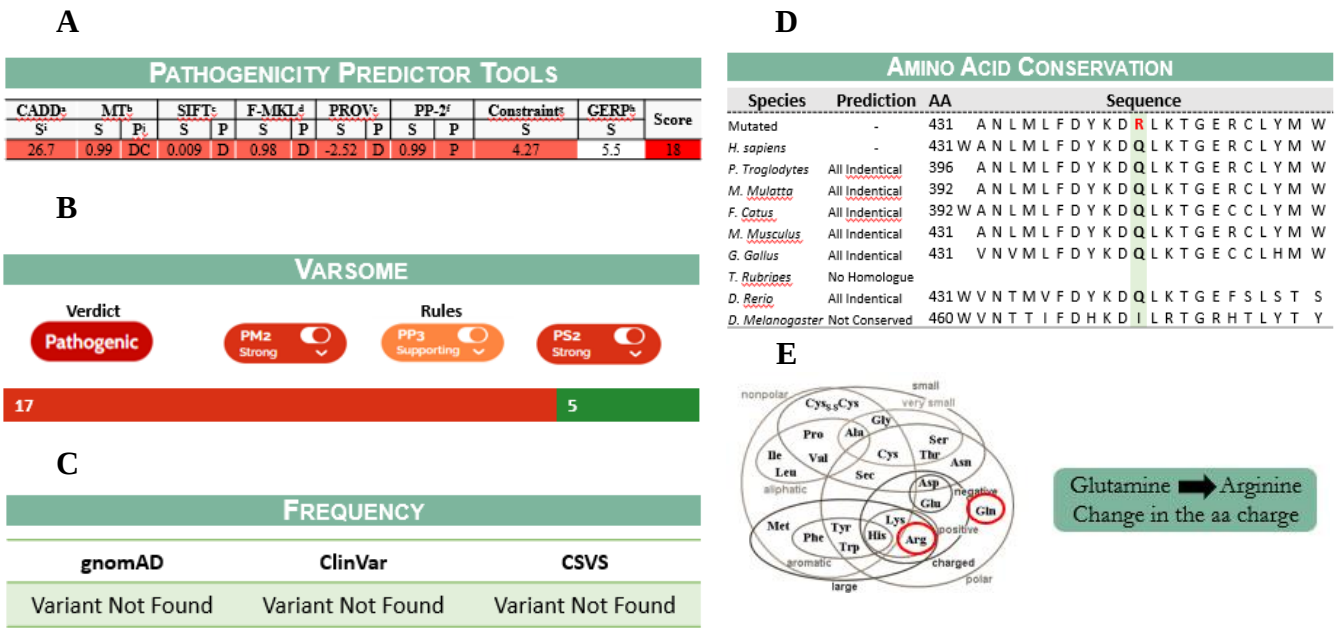
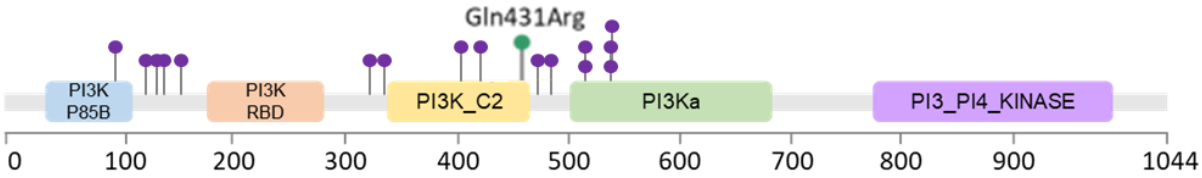


Figure 24: ***In silico*** **prediction analysis used for assessment of the pathogenicity impact of the candidate variant p.Gln431Arg on *PIK3CD*.** **A**) Pathogenic impact predictors used for the *in silico* evaluation of the identified candidate variant and total score after Knowledge-Driven A analysis. **B**) The verdict of Varsome according to the rules established by the ACMG. **C**) Missing frequency information of the variant in three different databases. **D**) Conservation of Amino acid according to MutationTaster results. **E**) Scheme showing the change of amino acid charge between Glutamine and Arginine (red circles). CSVS: Collaborative Spanish Variant Server, CADD: Combined Annotation Dependent Depletion MT: MutationTaster, F-MKL: FATHMM-MKL, PROV: PROVEAN, GERP: Genomic Evolutionary Rate Profiling, S: Score, P: Prediction, DC: Disease Causing, D, Damaging, P: Pathogenic, AA: amino acid. Arg: Arginine, Gln: Glutamine.

A



B

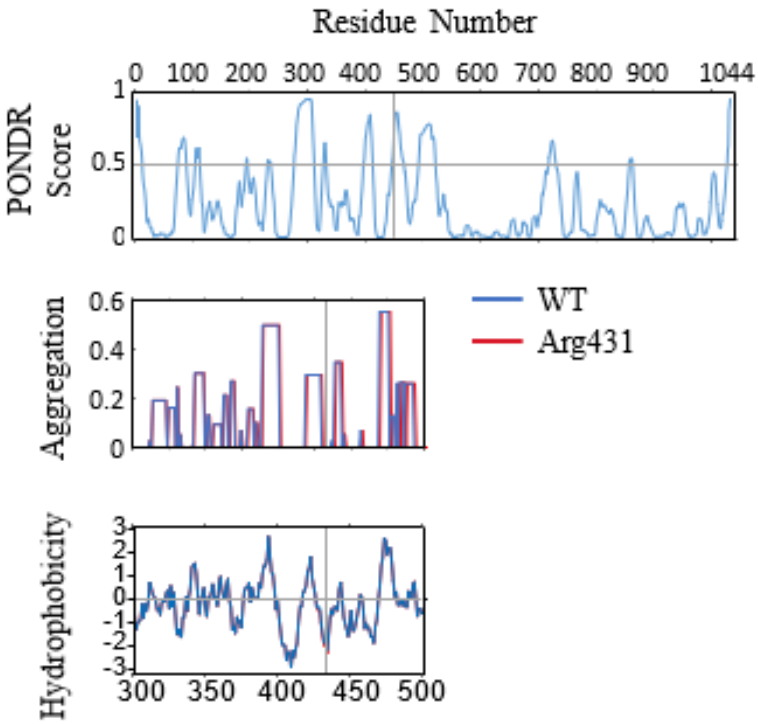


Figure 25. ***In silico*** **analysis of PIK3CD**. Position of the identified PIK3CD mutation and *in silico* analysis of the biological characteristics of two PIK3CD sequences (WT and p.Gln431Arg). **A**) Position of the p.Gln431Arg variant of *PIK3CD* (green lollipop) withing the protein domains together with other variants reported in the HGMD database (purple lollipops) using Mutation mapper. **B**) Protein disorder and hydrophobicity prediction. Upper panel shows the disorder prediction of the full-length amino acid sequence using the PONDR VL-XT algorithm. The middle panel shows the aggregation propensity of 200 amino acids containing the substituted residue calculated with Aggrescan. The bottom panel shows hydrophobicity prediction of 200 amino acids containing the substituted Arg431 residue using ExPasy ProtScale. A vertical grey line indicates the position of the mutated amino acid (Gln431 in WT).

A

Predicted Stability Change ($\Delta\Delta G^{Stability}$)

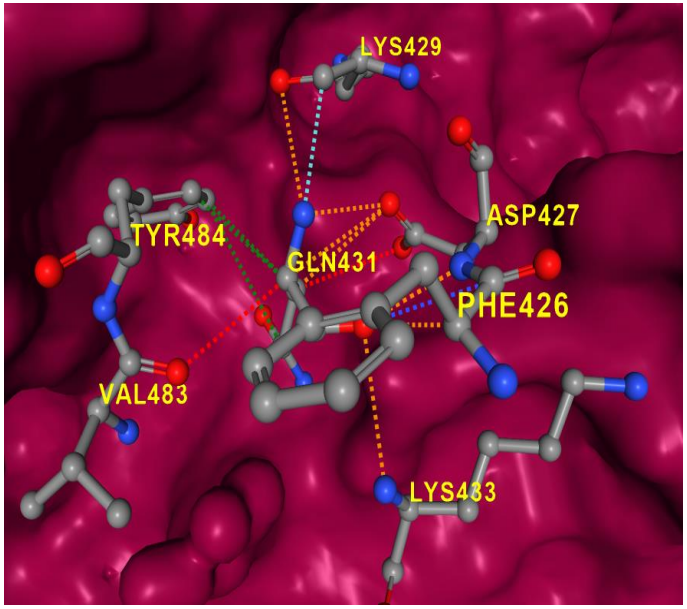
-0.4 kcal/mol

(Destabilising)

Mutation Details

Chain: **A**
Position: **431**
Wild-type: **Q**
Mutant: **R**

B



C

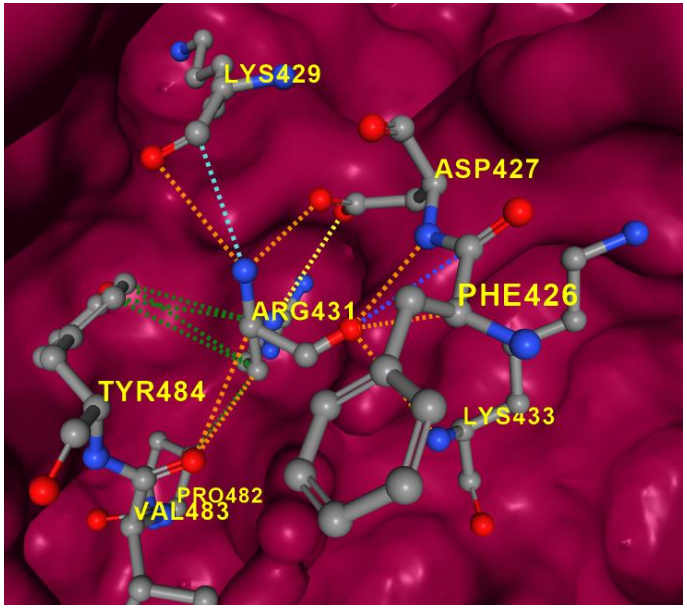


Figure 26. **Predicted stability change analysis of PIK3CD^{WT} and PIK3CD^{Gln431Arg}**. **A**) Result of the predicted stability Change (ΔΔGStability) of the Arginine residue at position 431 according to the variation in Gibbs free energy using DynaMut2. **B**,**C**) Interactions between neighbouring residues of PIK3CD^{WT} (Gln431) and PIK3CD^{Gln431Arg} predicted stability change of, respectively using DynaMut2. The coloured dashed lines indicate the following interactions: van der Waals (light blue), hydrophobic (green), hydrogen bond (red), carbonyl (blue, ionic (yellow), and polar interaction (orange). Q: Glutamine, R: Arginine.

Regarding functional studies, following *in silico* analyses, we investigated whether the p.Gln431Arg variant had a detrimental effect on the subcellular localization of the PIK3CD protein. For this, we transfected COS-7 cells with the pEGFP-PIK3CD^{WT} and the p.EGFP-PIK3CD^{Gln431Arg} plasmids. Confocal microscopy imaging results revealed that the overexpression of both PIK3CD *wild type* and mutant remains in the cytoplasm. Although, we observed the presence of punctiform and gross aggregation patterns in both *wild type* and mutant (Figure 27A). To delve into the observed patterns, quantification of the aggregates was carried out. The results showed that there is a significant difference when comparing the subcellular localization patterns of the *wild type* with the mutant (Figure 27B, left panel). The gross aggregate is the most predominant pattern in the p.Gln431Arg variant (64%, p= 0.006) unlike the *wild type*, in which the punctiform aggregate prevails (64%, p= 0.006). However, when comparing the size of these aggregates there is no significant difference between both constructs (Figure 27B, middle panel, p= 0.276). Additionally, the total aggregated area of both patterns ranged from less than 0.1 µm^2 to 1026 µm^2 (Figure 27B, right panel, p= 0.113 WT *vs*. Gln431Arg).

Currently, *PIK3CD* variants have been associated with autosomal dominant (Angulo et al., 2013) and recessive immunodeficiency (Sogkas et al., 2018), and Roifman-Chitayat syndrome (Roifman & Chitayat, 2009). However, the implication of the impact of *PIK3CD* variants within many pathways involved with NAIS has not been published. To study whether this aberrant pattern predominance of the mutated protein (Figure 28A) affects directly the PIK3-Akt pathway and indirectly the second messenger PIP3, we performed western blot analysis of the COS-7 cells transfected with both pEGFP-PIK3CD constructs (Figure 28B, upper panel). The detection was performed using specific antibodies against total-AKT1 (tAKT) and phosphorylated-AKT (pAKT) at Serine 437. However, under the tested conditions, the quantification results did not show significant difference in the pAKT/tAKT ratio (Figure 28B, bottom panel; p= 0.700 WT *vs*. Gln431Arg, p= 0.609 Mock *vs*. Gln431Arg, p= 0.992 Mock *vs*. WT).

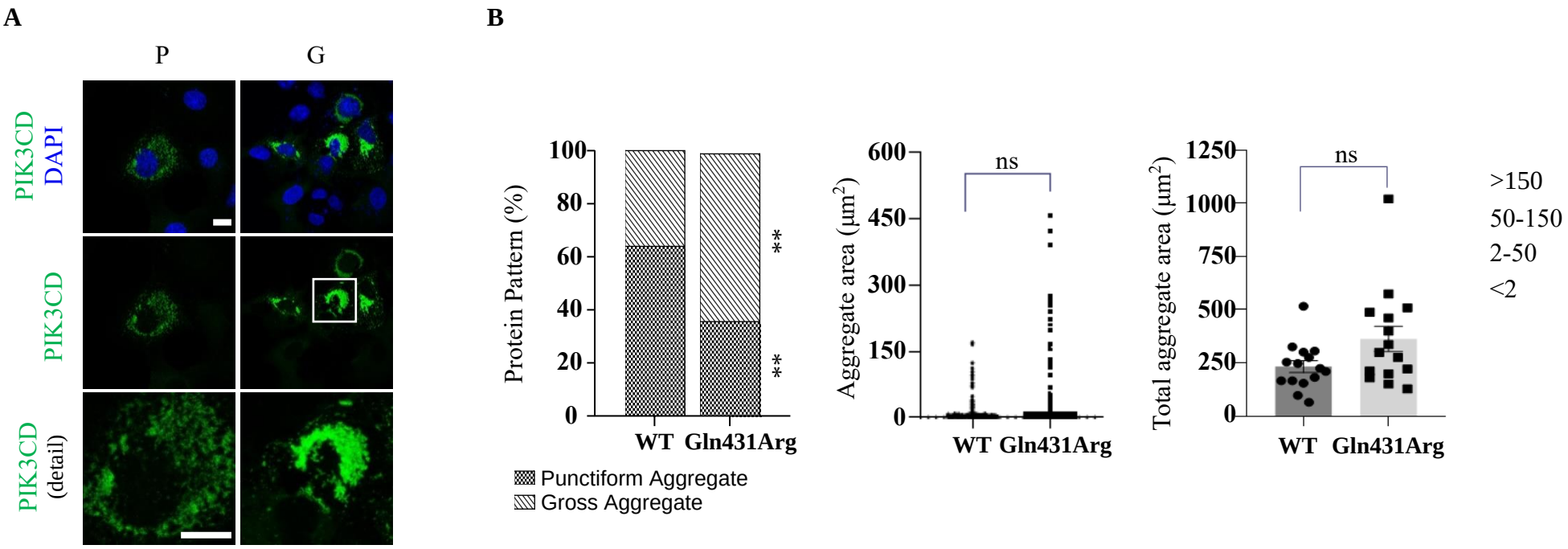


Figure 27. **Analysis of protein patterns in PIK3CD^{WT} and PIK3CD^{Gln431Arg}**. **A**) Representative fluorescent imaging of punctiform and gross protein aggregate patterns (green) observed in PIK3CD^{WT} and PIK3CD^{Gln431Arg} in COS-7 transfected cells. Nuclei were stained with DAPI (blue). Scale bar = 25 µm. **B**) Left panel shows the quantification of the two protein patterns observed in PIK3CD^{WT} and PIK3CD^{Gln431Arg}. The middle panel represents the quantified area of the two PIK3CD aggregates observed in the PIK3CD^{WT} and PIK3CD^{Gln431Arg}. Right panel shows the total aggregate area of the PIK3CD aggregates present in the PIK3CD^{WT} and PIK3CD^{Gln431Arg}. Values are means ± SEM. Two-tailed Student's t-test and Mann-Whitney test were used for comparisons between PIK3CD^{Gln431Arg} and PIK3CD^{WT} protein patterns. P: Punctiform aggregate. G: Gross aggregate.

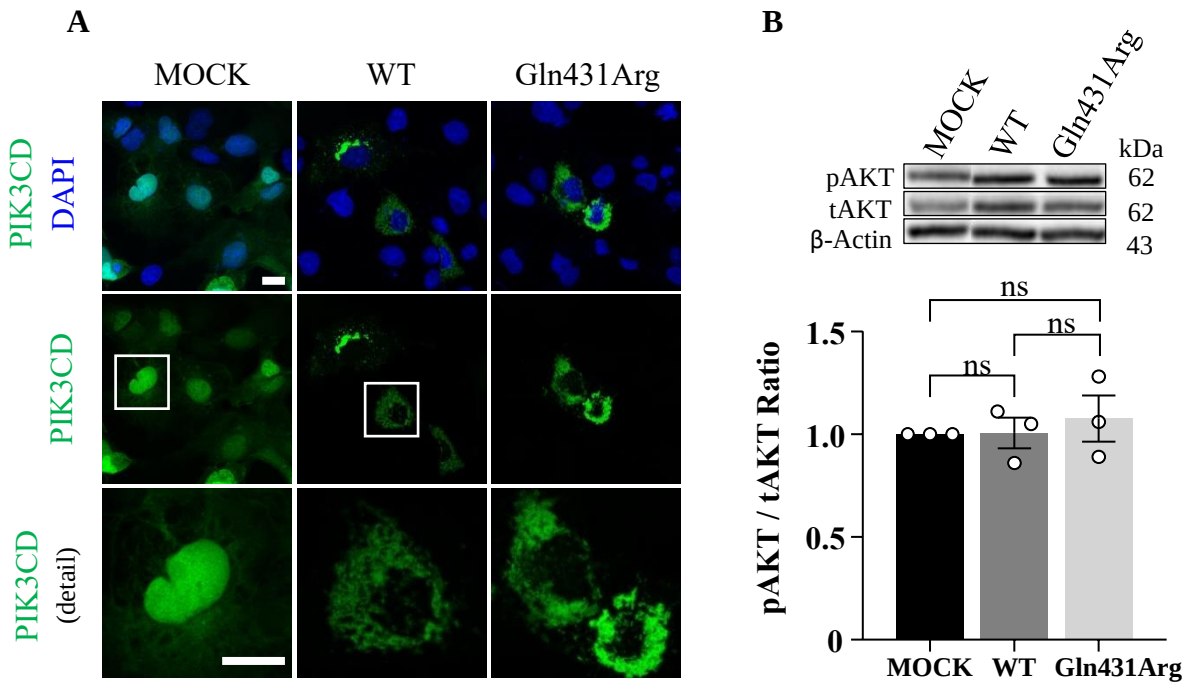


Figure 28. **Functional studies of PIK3CD^{WT} and PIK3CD^{Gln431Arg}**. **A**) Representative fluorescent images of GFP (green) and DAPI nuclear staining (blue) of COS-7 cells transfected with pEGFP-PIK3CD^{WT}, pEGFP-PIK3CD^{Gln431Arg} and control pEGFP-N1 (MOCK). Scale bar = 25 µm. **B**) Western blot of control (MOCK), pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{Gln431Arg} COS-7 transfected cells. Blots were probed for phosphor-AKT (pAKT), total-AKT (tAKT), and β-Actin. Quantification of protein expression in WT and Gln461Arg was normalized to the mock. Values are means ± SEM. Two-tailed Student's t-test was used for comparisons against mock and PIK3CD^{WT} (ns: not significant).

DISCUSSION

The insufficient evaluation of rare genetic variants might be the cause of the missing heritability of many diseases (Zuk et al., 2014). However, application of next generation sequencing (NGS) has allowed the identification of several genes with complex human diseases. This is the case of the studies based on the genetics of stroke which have been mainly focused on the adult and young adults leaving out the neonatal group.

Furthermore, the core genetic foundations in a growing number of major neonatal diseases that were initially considered non-genetic, have been established thanks to modern molecular biology methods (Liu et al., 2010; Suratanee et al., 2018). This is the case of newborns and infants with epileptic and developmental encephalopathies associated with *de novo* dominant mutations in sporadic cases (autosomal recessive and X-linked forms have also been described) (Happ et al., 2020). Based on this recent knowledge, we hypothesized that sporadic idiopathic NAIS in term newborns could be associated with *de novo* dominant variants in genes involved in biological pathways that could be related to the pathophysiology of NAIS, such as endothelial rheology, coagulation, and cell adhesion. In case of negative findings, we hypothesized autosomal recessive inheritance as a second option.

In our work, the use of NGS, more specifically the WES-trio in 23 infants with idiopathic NAIS, allowed us to identify 28 unique *de novo* variants in 28 genes. Subsequently, for variant prioritization, we used several pathogenicity predictors together with the establishment of a personalized knowledge-based analysis pipeline specific for this type of study. Our results identified, among these 28 variants, the p.Gln431Arg variant in the *PIK3CD* gene as the most suitable candidate for elucidating whether it is the genetic cause of idiopathic NAIS in an infant with perforating stroke. However, even though subsequent analyses of this variant have shown that the pathogenicity impact could affect protein function through secondary structure destabilization, further studies and approaches are needed to clarify the possibility of NAIS association.

The generation and offer of specific and potential therapies that can favour the development of precision medicine should be a realistic goal for infants with NAIS. This is where the relevance of our study lies since knowledge of the etiopathogenesis of this major neonatal disorder could allow us to achieve this goal. Moreover, deciphering the genetic gap in NAIS, together with the search for a better understanding of the molecular mechanisms involved in idiopathic NAIS, is something that responds to the

current yearning to identify individual genetic profiles, specific therapeutic targets, and safer tailored therapies.

Finally, focusing on studies based on the genetics of NAIS would make it possible to establish genetic screening, accurate prognosis, and adequate use of the information for the rehabilitation and prophylaxis of comorbidities. Crucial factors for affected families because it would help provide reliable prenatal and postnatal genetic counselling.

1. Rare *De novo* Variants and Its Association with NAIS

The study of WES of patients has grown exponentially, enabling rapid and cost-effective large-scale germline DNA sequencing to resolve the hidden heritability in many complex diseases. The main purpose is to generate a detailed set of genetic variants and separate those that truly cause diseases from the broader non-pathogenic ones present in the human genome (MacArthur et al., 2014). However, at present, there is a current absence of genetic studies that try to identify several loci of susceptibility to NAIS and that are based on well-defined homogeneous groups of subjects.

High-throughput sequencing approaches for gene discovery and disease association are challenging tasks in biomedicine and human genetics. These strategies involve comparative analysis of individual common variants or sets of rare variants between cases and controls, and the results are often unclear (Zuk et al., 2014). For instance, NGS-based studies have focused especially on those diseases with complex traits such as schizophrenia (Ripke et al., 2013), Parkinson's disease (J. F. Guo et al., 2018), epilepsy (Perucca et al., 2020), and stroke in adults (Falcone et al., 2014).

Many cardiovascular diseases and phenotypes have a broad multifactorial genetic background, such as susceptibility to ischemic stroke (MIM #601367) which has been associated with mutations in *F5*, *NOS3*, *F2*, *ALOX5AP*, and *PRKCH* and with the documented 5q12 locus (MIM %606799) (Gretarsdottir et al., 2002). Moreover, stroke can also be inherited in a classic Mendelian pattern, with low prevalence but high risk for mutation carriers. For example, *NOTCH3* which causes CADASIL (MIM #125310) (Stojanov et al., 2014), *COL4A1* associated with Small-Vessel disease (MIM #175780) and haemorrhagic stroke susceptibility (MIM #614519) (D. B. Gould et al., 2006), and *APP*, *BRI*, and *CST3* associated with autosomal dominant amyloid angiopathies (Tournier-Lasserve, 2002).

In this work, the application of WES-trio and a systematic workflow filtering criterion in 23 patients allowed us to identify 28 *de novo* rare private variants in 28 genes (*EPS15L1, SYNE2, FAM193A, HECTD1, TOP2A, RBM7, KRT6A, UNC13D, FAR2, SHANK3, PCBP1, KCNC2, CHRNA4, SPTBN5, TRABD2B, FAM83A, NR2E1, FASTKD5, PIK3CD, SNTB2, NPC1, DENND4A, DYM, SLC2A10, AARS1, ESRRG, COL6A3, AFF3*). In our cohort, five infants carried more than one *de novo* variant and one had no variants that met our filter criteria. These results are consistent with the calculated rate of *de novo* SNVs in an individual's genome, which could oscillate from 44-82 *de novo* SNVs in the entire genome with nearly one or two variants located in the protein-coding regions (Francioli et al., 2015; Gilissen et al., 2014). Furthermore, this rate of *de novo* variants could change because it also depends on the age of the father (Kong et al., 2012).

Therefore, the identification of a major candidate gene behind NAIS was far from being achieved. All 28 variants identified in 28 genes were isolated and exclusive to each individual sequenced in the study. It has been reported that isolated genes could have a relevant impact on the development of many particular diseases, especially in those in which the pathophysiology could result in cerebrovascular accidents (Sharma et al., 2013). For instance, *NOTCH3* with the subcortical stroke CADASIL (Sharma et al., 2013), *NINJ2*, *NOS3*, *COL4A1* and *DYRK1A* with an increased risk of stroke (Ikram et al., 2009; Malik et al., 2018), and *RNF213, CHD4, CNOT3, SETD5, CACNA1A,* and *GLA* with cerebrovascular diseases (Gudenkauf et al., 2020; Harel et al., 2015; Pinard et al., 2020, 2021).

When searching in databases for stroke association studies based on our set of identified genes, the only evidence found was on the *SPTBN5* gene. The study associated the polymorphism rs4923918 with ischemic stroke in Japanese adults (Yoshida et al., 2010). However, this polymorphism was not present in our dataset. In addition, comprehensive *in silico* analyses of the identified *SPTBN5* variant (c.938C>T /p.Ala313Val) in a patient with a M1 post stroke were not considered to have a potential pathogenicity impact (see results section 2: Prediction of the Pathogenicity Impact for the Assessment of Variants, section 6: Selection of Candidate Genes, Table 10, and Table 15).

2. Complementary Inheritance Approaches in NAIS

Next generation sequencing technology generates a vast amount of useful information, and to take advantage of it, a recessive inheritance model was applied. This made it possible to explore other possibilities that could help uncover the genetic influence of NAIS. When the parents are not consanguineous, the genetic pattern in diseases with autosomal recessive inheritance is usually heterozygous for two different pathogenic bi allelic variants in the affected individual, or, less frequently, homozygous (Dashti et al., 2017). This is something that has been documented, for example, with homozygous mutations in the tight-junction protein *JAM3* and its association with haemorrhagic destruction of the brain (Akawi et al., 2013; Mochida et al., 2010). On the contrary, our results showed that there was no relevant candidate gene with bi-allelic compound heterozygous nor homozygous variant. Additionally, as was the case with the *de novo* hypothesis, all the variants that were filtered step by step, were exclusive to each individual.

It is important to explore non-Mendelian inheritance risk factors associated with cerebrovascular diseases. For this we decided to analyse polymorphisms that have been associated with the development of prothrombotic states in individuals with stroke. We found no infants carrying the homozygous polymorphisms *F2* (rs1799963), *F5* Leiden (rs6025), *APOE* (rs7412 and rs429358), *F13A1* (rs5985 and rs302447), and the *ITGB3* (rs5918). These results are consistent with other studies conducted in newborns where genetic factors of thrombophilia could not be associated with NAIS (Arnaez et al., 2018a; Arnaez et al., 2018b).

3. Variant Analysis Predicts Putative Pathogenicity Impacts

The adequate assessment of genetic variants is relevant for the identification of the predisposing genetic risk factor within any disease. Regardless of the inheritance model applied in the study, understanding how variants could be affecting gene/protein functions is a vital aspect to consider for candidate gene selection. Especially when it has been shown that damaging *de novo* rare variants often contribute to phenotypic variations and susceptibility to many diseases, particularly rare diseases (Roca et al., 2018). Currently, computational algorithms are used to predict the potential functional impact and relevance of variants, but due to the surpassing volume of data, cautious

analysis should be considered when weighing the significance of these predictions (Dashti & Gamieldien, 2017).

NGS approaches are potent methods for elucidating new biological insights into human diseases, but it is a challenge to manage many datasets of NGS data produced and slows down the efforts of gene-disease associations (Mardis E. R., 2016). The tools used to establish the pathogenicity impact of the variants are designed on the conservation of the nucleotides and amino acids, and on the differences between several physicochemical features of the change implicated (Roca et al., 2018). However, there is still a current lack of information and computer predictors that allow the identification of the relevant disease-causing variant. Pathogenic variant prediction has no actual gold standard (M. H. Guo et al., 2018) and the tools needed for variant filtering should be discussed within the research team for proper evaluation. Something that has been recommended by other different studies (Carson et al., 2014; J. F. Guo et al., 2018; Pabinger et al., 2014; Tang & Thomas, 2016).

In our study, we used an extensive list of prediction tools to obtain a truthful result of the analysis of each of all 28 confirmed *de novo* variants. This prediction tools were based on protein structure, combined sequence and structural approaches, sequence conservation, and a combination of meta-predictors that integrate results of other platforms. This combination proved to be a useful way for *in-silico* pathogenicity prediction analysis of exonic variants, splicing site variants, and to separate the obvious damage/disease-causing variants from those neutral/benign or of uncertain significance. This approach has been done previously with data from other functional genomics studies and showed that nearly half of the *de novo* or rare missense mutations identified by genome sequencing were practically neutral mutations (Miosge et al., 2015).

4. Candidate Gene Selection Strategies in NAIS

After NGS analysis, the relationship between candidate genes and clinical phenotypes is not always obvious. Many analyses allow a mechanistic view of omics data to be obtained, facilitating the interpretation of the gene list. Furthermore, understanding these interactions might show common attributes of these genes, including whether they are enriched in known pathways, complexes and functions, possible causal mechanisms, predict new roles for genes, functional redundancy, relevant protein-protein interactions,

improve the statistical power, and ease candidate gene selection for disease-phenotype correlation.

A shared interaction network is a useful way to interpret a set of potential candidate genes for further evaluation. This interaction network can be defined as a set of noodles (genes/proteins) that are connected by edges (interactions) that represent demonstrated or inferred functional relationships among the noodles (Bebek, 2012). Virtually all these interactions discovered through experimental studies or identified through data mining are placed in public databases, allowing direct access to physical interaction, co-expression, co-occurrence, gene fusion, and neighbourhood data.

Many studies showed that the application of a shared interaction network has the power to associate genes with diseases (Navlakha & Kingsford, 2010). For instance, *MEF2A* with coronary artery disease (MIM #6083209 (Dashti et al., 2017), associations of genes and diseases with the immune response (Dong et al., 2015), and inferred associations of genes and diseases related to inborn errors, inflammation, lipid metabolism, cerebellar ataxia, cognitive deterioration, and malignant tumours (Suratanee & Plaimas, 2018). In our case, this approach did not have the power to identify an enriched candidate gene that had interactions that could fully or partially explain the underlying molecular mechanisms involved in the pathophysiology of the stroke or whether there was a relevant interaction with known stroke-related genes.

When necessary, other strategies can complement gene set analysis by increasing the power of association and identification. A phenotypically homogeneous subgroup is a resource that, if available, allows for a finer analysis of clinical data in genetic disease studies. Generally, gene-disease association investigations of clinically defined cases versus controls are primarily conducted using the GWAS approach. At present, these studies are improving in terms of re-analysing the already existing data but a reorganization of the cohort in homogenous subsets is essential (Traylor et al., 2015).

The clinical heterogeneity of NAIS is mostly attributed to the great variation in the stroke localization, extent, and volume, which does not automatically reflect the structures involved during the infarct event (Kirton et al., 2007). Therefore, to improve the genetic analysis, a subgroup analysis of homogeneous phenotypes is recommended. Our data was subclassified into five subsets regarding the definition of affected arterial territory (ATD) among infants. We use the classification proposed by Gemma Arca and

colleagues (Arca et al., 2020), in which the M1 pre, M1 post, M2, M3 and the perforant ATD of the MCA. Although there is a demonstrated increase in the power of this methodology for gene-disease associations (Traylor et al., 2015), our results did not reproduce the same conclusions after the analysis strategy. More specifically, no relationship could be found between all 28 genes and the ATDs regarding the signalling pathways involved or the function of the encoded proteins.

A thorough analysis of complex diseases is a recommendation due to the intricate pathophysiology involved. It evaluates the possible candidates individually or together to establish a gene prioritization strategy which is considered an essential step in gene-disease associations (Bamshad et al., 2011). Additionally, it allows intense scrutiny of variants at different biological levels, which is important in rare disorders due to distinct phenotypes caused by different types of mutations that affect different functional regions of the gene (Gussow et al., 2016). For example, whilst *TTN* and *SYNE1* are susceptible to nonsense variants and tolerate missense variants, *SETBP1* is sensitive to variant localization and results in a severe phenotype when the variant is within a specific position of the encoded protein.

More recently, Knowledge-Driven Analysis (KDA) is terminology designed for more comprehensive variant analysis (Dashti & Gamieldien, 2017) that complements the mutational architecture and tolerance analysis (Roca et al., 2018). KDA integrates variant rarity, mutational tolerance and architecture, variant frequency, amino acid conservation, mode of inheritance, gene function, tissue/organ expression, cell signalling pathways, protein-protein interactions, biological and molecular gene ontologies, human phenotype ontologies, and common features among candidate genes.

In our study, the application of a specific and personalized score according to the evidence of the results obtained by the KDA was a new strategy for the identification of the most suitable candidate gene. Generally, the approaches used for this purpose are based on the use of RNAseq as a technique to obtain a network analysis of differential expression of genes for subsequent prioritization (W. Li et al., 2021; Nitsch et al., 2010; Ransbotyn et al., 2015). The results of this in-house workflow analysis identified *PIK3CD* as the strongest candidate gene for further functional studies and disease association. This gene with its rare *de novo* variant c.1292A>G/p.Gln431Arg is present in an infant that suffered a perforant ischemic stroke. Studies using similar gene prioritization approximations have been able to identify associations between genes and

diseases, such as *MEF2A* with coronary artery disease (Dashti & Gamieldien, 2017) and novel mutations with neurodevelopmental disorders (Hoischen et al., 2014).

5. *PIK3CD* and Its Implication in NAIS as a Candidate Gene

Many pathways are involved in stroke pathophysiology, especially those that comprehend endothelial cell rheology, the coagulation cascade, cell adhesion, platelet aggregation, and the inflammatory response. To date, NAIS gene candidate association studies have typically involved the investigation of classic genetic risk factors, such as the *MTHFR* C667T and A1298C polymorphisms, FVL, *F2* G20210A, *TNF-α*-308, *IL-6*, and polymorphisms of the ε2 and ε4 alleles of apolipoprotein E (Gelfand et al., 2013). However, these studies usually analyse paediatric and perinatal cases with a positive family history of risk factors and related syndromic conditions (Grossi et al., 2020).

Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Delta (*PIK3CD*), is a gene that encodes a class I PI3K protein that binds p85 adapter and GTP-bound RAS proteins. A relevant feature of PIK3CD is that it phosphorylates itself and converts phosphatidylinositol-4,5-bisphosphate (PIP2) into phosphatidylinositol-3,4,5-trisphosphate (PIP3) that recruits and activates downstream proteins containing the pleckstrin homology domains (Balasuriya et al., 2018). Moreover, PIK3CD and its production of PIP3 acts as a second messenger that helps activate the alpha serine/threonine-protein kinase (AKT) (Zhuang et al., 2013). This activation occurs at the plasma membrane, where AKT is phosphorylated at serine^{473} (Ser473) and threonine^{308} (Thr308) residues of the protein by upstream kinases. Phosphorylation of both residues allows greater control of several cellular processes (Shiojima et al., 2002). Therefore, evaluating the PI3K-Akt pathway is a useful approach to directly study the pathway itself and indirectly the function of the protein through the second messenger PIP3,

According to many expression databases, the *PIK3CD* gene is ubiquitously expressed in major human tissues including immune, nervous, muscular, secretory, and reproductive tissues. PIK3CD is part of the phosphatidylinositol 3'-kinase (PI3K)-Akt cell-signalling pathway, which is activated by many types of cellular stimuli or toxic insults. Additionally, it is involved in other important pathways previously associated with the

pathophysiology of stroke, such as the vascular endothelial growth factor (VEGF), fluid shear stress, and platelet activation pathway (Liao et al., 2019).

Many studies emphasize the role of PI3K-Akt signalling in processes such as vascular homeostasis and angiogenesis (Ackah et al., 2005; M. Lee et al., 2014; Shiojima et al., 2002), endothelial nitric oxide synthase activity (Papapetropoulos et al., 1997), regulation of endothelial cell migration (Morales-Ruiz et al., 2000), statins (Kureishi et al., 2000), anti-thrombotic and vasodilators molecules (Luo et al., 2000), the integrated control of organ growth (Verdu et al., 1999), and cell permeability, survival and proliferation (Gerber et al., 1998). In addition, two studies highlight the relevant role of the *PIK3CA* gene genotype in maternal and fetal PI3K signalling and the regulating maternal-fetal resource allocation (Jansson, 2016; Sferruzzi-Perri et al., 2016). However, information on the involvement of PI3K-Akt signalling and the development of neonatal arterial ischemic stroke is lacking. The study of this cell-signalling pathway should be considered a relevant step for understanding the mechanisms involved in many diseases.

On the other hand, *PIK3CD* has been associated with the autosomal dominant immunodeficiency 14A (MIM #615513) previously known as activated PI3Kδ syndrome 1, autosomal recessive immunodeficiency 14B (MIM #619281), and to the digenic Roifman-Chitayat syndrome (MIM #613328), hence the challenge of associating *PIK3CD* with, specifically idiopathic neonatal arterial ischemic perforant stroke. To investigate this *PIK3CD* association, we decided to perform *in silico* and functional analysis to study the pathogenic impact that the p.Gln431Arg variant could have and improve our understanding of the underlying mechanisms in NAIS through the PI3K-Akt cell signalling pathway.

Our results, however, showed that although *PIK3CD* was flagged as a candidate gene, p.Gln431Arg does not have a relevant impact on protein structure, helicity, aggregation propensity, and expression levels of total AKT and phosphorylated-AKT (Ser473). Regarding the localization of the protein overexpressed, we observed differences in the subcellular pattern of $PIK3CD^{WT}$ and $PIK3CD^{p.Gln431Arg}$ models. It is important to consider that the enzymatic activity of kinases is often modulated by post-translational modification events that do not alter protein abundance or localization. Additionally, these aberrant patterns could be due to overexpression of the protein itself. Therefore, having the possibility of carrying out functional studies with cells from the patient under

study is something that would allow better investigation of the impact of the identified variant on the endogenous protein. However, this was not possible in our case.

In conclusion, our data suggest that NAIS might not be a single-gene disorder with autosomal dominant or recessive inheritance. Therefore, it seems that rare *de novo* variants are not related to the stroke event in NAIS, and despite extensive genetic analysis of the patients, the genetic cause of neonatal arterial ischemic stroke remains unclear. This observation is relevant because it redirects the vision towards the role of the interactions of the maternal-placental and fetal-neonatal pathways in the genesis of thrombi that manage to reach the cerebral circulation of the newborn.

In addition, our results do not indicate the need for genetic studies of autosomal *de novo* or recessive inheritance variants in patients with idiopathic NAIS. Henceforth, it is necessary to consider alternative hypothesis strategies for future studies that may involve an imbalance in susceptibility and protective genetic factors, oligogenic or polygenic inheritance and/or unknown external stimuli with epigenetic influence. Ultimately, more research options and additional analyses are warranted to achieve a better understanding of the genetic aetiology in NAIS and delve into other molecular aspects involved in the disease.

Although in our case our findings did not support our hypothesis, the results are important for others. Some limitations that could influence our results are:

- *Small sample size:* our exclusion and inclusion criteria, in addition to what idiopathic NAIS entails, left us with a small number of infants recruited. We considered that this limitation is relevant due to the importance to find a significant common gene that could be associated with NAIS or more specifically, with a stroke subtype according to the arterial territory definition and volume.
- *Variants not captured during sequencing:* whole-exome sequencing relies on thousands of probes and enriched capture methods that are designed based on the annotated reference sequence obtained from databases such as the consensus coding sequence (CCDS) and the RefSeq. As a result, unknown genes, or exons yet to be annotated cannot be captured and sequenced; therefore, the causative variant is missing from the target area and is not identified. Moreover, the application of WES is because exons are the most likely region of the genome to carry pathogenic

variants, which excludes the sequencing of the non-coding regions that have been shown in some cases to contain disease-causing variants.

- *Sequencing Coverage:* inadequate coverage of the region confers an added challenge to the study. Depth of coverage is required for adequate detection of those variations in the individual exome that differs from the reference sequence. In our case, we had excellent coverage depth with 10X-30X without showing biased results among the trios. Although, we must consider that GC-rich regions and regions with repetitive sequences are not well characterized due to the prevention of probe inclusion and reads not easily mapped.

  Additionally, if the disease requires a mosaicism analysis, a depth coverage of 200X is the minimum required to reveal a 3% mosaic mutation, with 500X being a more stable coverage for detection. However, in cases where mosaicism is reduced to 0.1%, even 10,000X depth coverage could not detect the mutation (Dai et al., 2020).

- *Lack of reliable evidence*: There are some cases in which a gene/protein is absent from the databases used, generating information gaps that hamper the collection of relevant features, leaving researchers to draw their conclusions about the candidate gene involvement in signalling pathways, ontologies, functions, and protein-protein interactions.

Despite these limitations, we believe that our study involves a series of parameters that make the results relevant, valid, with impact, and with real and pertinent consequences for the public health system and the scientific community. These considerations include:

- *Sampling*: getting the big picture is a useful approach when trying to explore the genetics of complex diseases. The infants recruited in our research reflect this by being an ecological sample. Thus, we believe that the 23 infants studied with neonatal arterial ischemic stroke of unknown cause portray what is appreciated in clinical practice regarding idiopathic stroke.

- *Newborn-focused*: essentially all genetic stroke studies focus on adult or young adults and those dedicated to the newborn are scarce. Furthermore, if we consider the well-defined sampling criteria, the methodology, the approaches

applied, and the homogenous subgroup of subjects, we can say that our specific type of study is currently absent form scientific databases.

- *Trio approach*: *de novo* variants are genetic changes that represent an extreme form of rare genetic variation and are often the main cause of many diseases. The application of trio whole-exome sequencing (index case and parents) is a highly effective strategy that makes variant screening and analysis more practical for the identification of these *de novo* events in the coding regions of the genome.
- *Completeness:* in this research, a great effort was made from sampling to analysis and genetic interpretation. Our intended goal was to cover as many aspects as possible that could shed light on the genetic cause of NAIS. From different models of inheritance to a detail-focused analysis of homogeneous subclinical groups. This is always based on previous relevant studies and deep clinical understanding of NAIS.

CONCLUSION

1. Comprehensive genetic analysis under *de novo* autosomal dominant and autosomal recessive inheritance of the coding regions of 23 parent-proband trios with idiopathic NAIS did not identify a major common gene or pathway that could explain, totally or partially, the genetic cause of the disease.

2. The study of classifications of homogeneous subclinical groups according to the affected arterial territory, volume, and evolution at two years of age did not reveal a common cell-signalling pathway that could be associated between them.

3. In a patient with perforating stroke, the functional study of the candidate gene *PIK3CD* and its variant c.1292A>G/p.Gln431Arg has shown that this variant could affect the protein function, yet further studies and approaches are needed to elucidate whether this gene could be associated with NAIS.

4. Neonatal arterial ischemic stroke remains a disease with no demonstrated etiology, either genetic or acquired. Other possible genetic causes to be considered include genetic variants that are not detected by exome sequencing i.e., small recurrent rearrangements such as gene inversion, oligogenic inheritance and/or epigenetic modifications.

5. We cannot exclude that the presence of genetic variations in the genome is the cause of idiopathic NAIS due to the possibility that these variants are located outside the coding regions, but to explore this, other strategies and technological approaches designed to target non-coding DNA regions are necessary.

References

Ackah, E., Yu, J., Zoellner, S., Iwakiri, Y., Skurk, C., Shibata, R., Ouchi, N., Easton, R. M., Galasso, G., Birnbaum, M. J., Walsh, K., & Sessa, W. C. (2005). Akt1/protein kinase Bα is critical for ischemic and VEGF-mediated angiogenesis. *Journal of Clinical Investigation*, *115*(8), 2119–2127. https://doi.org/10.1172/JCI24726

Adhami, F., Liao, G., Morozov, Y. M., Schloemer, A., Schmithorst, V. J., Lorenz, J. N., Dunn, R. S., Vorhees, C. V., Wills-Karp, M., Degen, J. L., Davis, R. J., Mizushima, N., Rakic, P., Dardzinski, B. J., Holland, S. K., Sharp, F. R., & Kuan, C. Y. (2006). Cerebral ischemia-hypoxia induces intravascular coagulation and autophagy. *American Journal of Pathology*, *169*(2), 566–583. https://doi.org/10.2353/ajpath.2006.051066

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Afifi, A. K. (1987). Cerebral Hemiatrophy, Hypoplasia of Internal Carotid Artery, and Intracranial Aneurysm. *Archives of Neurology*, *44*(2), 232. https://doi.org/10.1001/archneur.1987.00520140090024

Akawi, N. A., Canpolat, F. E., White, S. M., Quilis-Esquerra, J., Morales Sanchez, M., Gamundi, M. J., Mochida, G. H., Walsh, C. A., Ali, B. R., & Al-Gazali, L. (2013). Delineation of the Clinical, Molecular and Cellular Aspects of Novel JAM3 Mutations Underlying the Autosomal Recessive Hemorrhagic Destruction of the Brain, Subependymal Calcification, and Congenital Cataracts. *Human Mutation*, *34*(3), 498–505. https://doi.org/10.1002/humu.22263

Alfonso, I., Prieto, G., Vasconcellos, E., Aref, K., Pacheco, E., & Yelin, K. (2001). Internal carotid artery thrombus: an underdiagnosed source of brain emboli in neonates? *J Child Neurol*, *16*(6), 446–447. https://doi.org/10.1177/088307380101600612

Alladina, S. J., Song, J. H., Davidge, S. T., Hao, C., & Easton, A. S. (2005). TRAIL-induced apoptosis in human vascular endothelium is regulated by phosphatidylinositol 3-kinase/Akt through the short form of cellular FLIP and Bcl-2. *Journal of Vascular Research*, *42*(4), 337–347. https://doi.org/10.1159/000086599

Alvarez-Sabín, J., Quintana, M., Masjuan, J., Oliva-Moreno, J., Mar, J., Gonzalez-Rojas, N., Becerra, V., Torres, C., & Yebenes, M. (2017). Economic impact of patients admitted to stroke units in Spain. *European Journal of Health Economics*, *18*(4), 449–458. https://doi.org/10.1007/s10198-016-0799-9

American Academy of Pediatrics Committee on Fetus and Newborn, A. C. of O. and G. C. on O. P. (2015). The Apgar Score. *Pediatrics.*, *136*(4), 819–822. https://doi.org/10.1542/peds.2015-2651

Amit, M., & Camfield, P. R. (1980). Neonatal Polycythemia Causing Multiple Cerebral Infarcts. *Archives of Neurology*, *37*(2), 109–110. https://doi.org/10.1001/archneur.1980.00500510067013

Angulo, I., Vadas, O., Garçon, F., Banham-Hall, E., Plagnol, V., Leahy, T. R., Baxendale, H., Coulter, T., Curtis, J., Wu, C., Blake-Palmer, K., Perisic, O., Smyth, D., Maes, M., Fiddler, C., Juss, J., Cilliers, D., Markelj, G., Chandra, A., … Nejentsev, S. (2013). Phosphoinositide 3-kinase δ gene mutation predisposes to respiratory infection and airway damage. *Science (New York, N.Y.)*, *342*(6160), 866–871. https://doi.org/10.1126/science.1243292

Annus, Á., Juhász, L. Á., Szabó, E., Rárosi, F., Szpisjak, L., Vécsei, L., & Klivényi, P. (2020). Connection between small vessel disease related stroke and the MTHFR C677T polymorphism in a Hungarian population. *Heliyon*, *6*(11). https://doi.org/10.1016/j.heliyon.2020.e05305

Anteroinferior Arteries of the Brain. (2021). Anatomical Justice. https://anatomicaljustice.com/product/anteroinferior-arteries-of-the-brain-no-text/

Anwar, R., Gallivan, L., Edmonds, S. D., & Markham, A. F. (1999). Genotype/phenotype correlations for coagulation factor XIII: specific normal polymorphisms are associated with high or low factor XIII specific activity. *Blood*, *93*(3), 897–905.

Arca, G., Arnaez, J., Agut, T., Núñez, C., Stephan-Otto, C., Valls, A., & García-Alix, A. (2020). Neuron-specific enolase is correlated with lesion topology, relative infarct volume and outcome of symptomatic NAIS. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *105*(2), F132–F137. https://doi.org/10.1136/archdischild-2018-316680

Arnaez, J., Arca, G., Martín-Ancel, A., Agut, T., & Garcia-Alix, A. (2018a). Neonatal arterial ischemic stroke: Risk related to family history, maternal diseases, and Genetic Thrombophilia. *Clinical and Applied Thrombosis/Hemostasis*, *24*(1), 79–84. https://doi.org/10.1177/1076029617736383

Arnaez, J., Arca, G., Martín-Ancel, A., & Garcia-Alix, A. (2018b). Coagulation factor V G1691A, factor II G20210A and methylenetetrahydrofolate reductase C677T gene mutations do not play a major role in symptomatic neonatal arterial ischaemic stroke. In *British Journal of Haematology* (Vol. 180, Issue 2, pp. 290–292). Blackwell Publishing Ltd. https://doi.org/10.1111/bjh.14308

Arnaez, J., & Garcia-Alix, A. (2017). Extracerebral thrombosis in symptomatic neonatal arterial ischemic stroke. *European Journal of Paediatric Neurology*, *21*(4), 687–688. https://doi.org/10.1016/j.ejpn.2017.05.004

Atlas del Ictus, C. (2019). *EL ATLAS DEL ICTUS CATALUÑA*.

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., … Schloss, J. A. (2015). A global reference for human genetic variation. In *Nature* (Vol. 526, Issue 7571, pp. 68–74). Nature Publishing Group. https://doi.org/10.1038/nature15393

Aznar, J., Mira, Y., Vayá, A., Corella, D., Ferrando, F., Villa, P., & Estellés, A. (2004). Factor V Leiden and prothrombin G20210A mutations in young adults with cryptogenic ischemic stroke. *Thrombosis and Haemostasis*, *91*(5), 1031–1034. https://doi.org/10.1160/th03-11-0690

Bai, D., Ueno, L., & Vogt, P. (2009). Akt-mediated regulation of NFkB for the oncogenicity of PI3K and Akt. *International Journal of Cancer*, *125*(12), 2863–2870. https://doi.org/10.1002/ijc.24748.Akt-mediated

Balasuriya, N., McKenna, M., Liu, X., Li, S. S. C., & O'Donoghue, P. (2018). Phosphorylation-dependent inhibition of Akt1. *Genes*, *9*(9). https://doi.org/10.3390/genes9090450

Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A., & Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Reviews. Genetics*, *12*(11), 745–755. https://doi.org/10.1038/nrg3031

Bark, D. L., & Ku, D. N. (2013). Platelet transport rates and binding kinetics at high shear over a thrombus. *Biophysical Journal*, *105*(2), 502–511. https://doi.org/10.1016/j.bpj.2013.05.049

Basu, S., Totty, N. F., Irwin, M. S., Sudol, M., & Downward, J. (2003). Akt Phosphorylates the Yes-Associated Protein, YAP, to Induce Interaction with 14-3-3 and Attenuation of p73-Mediated Apoptosis some of the metabolic effects of Akt, whereas BAD can account for some of the survival-promoting effects. The most convincingly. *Molecular Cell*, *11*, 11–23. https://doi.org/10.1016/S1097-2765(02)00776-1

Beattie, L. M., Butler, S. J., & Goudie, D. E. (2006). Pathways of neonatal stroke and subclavian steal syndrome. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *91*(3), 204–207. https://doi.org/10.1136/adc.2005.079830

Bebek, G. (2012). Identifying gene interaction networks. *Methods in Molecular Biology (Clifton, N.J.)*, *850*, 483–494. https://doi.org/10.1007/978-1-61779-555-8_26

Bellacosa, A., Test, a J., Moore, R., & Larue, L. (2004). A portrait of AKT kinases: human cancer and animal models depict a family with strong individualities. *Cancer Biol Ther*, *3*(3), 268–275. https://doi.org/10.4161/cbt.3.3.703

Benakis, C., Garcia-Bonilla, L., Iadecola, C., & Anrather, J. (2015). The role of microglia and myeloid immune cells in acute cerebral ischemia. *Frontiers in Cellular Neuroscience*, *8*(JAN), 1–16. https://doi.org/10.3389/fncel.2014.00461

Benders, M. J. N. L., Groenendaal, F., Uiterwaal, C. S. P. M., Nikkels, P. G. J., Bruinse, H. W., Nievelstein, R. A. J., & De Vries, L. S. (2007). Maternal and infant characteristics associated with perinatal arterial stroke in the preterm infant. *Stroke*, *38*(6), 1759–1765. https://doi.org/10.1161/STROKEAHA.106.479311

Bennett, J. S., & Vilaire, G. (1979). Exposure of platelet fibrinogen receptors by ADP and epinephrine. *The Journal of Clinical Investigation*, *64*(5), 1393–1401. https://doi.org/10.1172/JCI109597

Bertina, R. M. (1990). Nomenclature proposal for protein S deficiency. *Annual Meeting of the Scientific and Standardization Committee of the ISTH, Barcelona (Spain)*.

Bertina, R. M., Broekmans, A. W., Krommenhoek-van Es, C., & van Wijngaarden, A. (1984). The use of a functional and immunologic assay for plasma protein C in the study of the heterogeneity of congenital protein C deficiency. *Thrombosis and Haemostasis*, *51*(1), 1–5. http://www.ncbi.nlm.nih.gov/pubmed/6547008

Bhagwanani, S. G., Price, H. V., Laurence, K. M., & Ginz, B. (1973). Risks and prevention of cervical cord injury in the management of breech presentation with hyperextension of the fetal head. *American Journal of Obstetrics and Gynecology*, *115*(8), 1159–1161. https://doi.org/10.1016/0002-9378(73)90573-5

Bioscience, T. (2015). *Cancer Research Product Guide* (3rd ed.).

Bosenbark, D. D., Krivitzky, L., Ichord, R., Vossough, A., Bhatia, A., Jastrzab, L. E., & Billinghurst, L. (2017). Clinical Predictors of Attention and Executive Functioning Outcomes in Children After Perinatal Arterial Ischemic Stroke. *Pediatric Neurology*, *69*, 79–86. https://doi.org/10.1016/j.pediatrneurol.2017.01.014

Brondani, R., Rieder, C. R. M., Valente, D., Araújo, L. F., & Clausell, N. (2007). Levels of vascular cell adhesion molecule-1 and endothelin-1 in ischemic stroke: A longitudinal prospective study. *Clinical Biochemistry*, *40*(3–4), 282–284. https://doi.org/10.1016/j.clinbiochem.2006.09.014

Brouns, R., & De Deyn, P. P. (2009). The complexity of neurobiological processes in acute ischemic stroke. In *Clinical Neurology and Neurosurgery* (Vol. 111, Issue 6, pp. 483–495). https://doi.org/10.1016/j.clineuro.2009.04.001

Burdon, K. P., McKay, J. D., Sale, M. M., Russell-Eggitt, I. M., Mackey, D. A., Wirth, M. G., Elder, J. E., Nicoll, A., Clarke, M. P., FitzGerald, L. M., Stankovich, J. M., Shaw, M. A., Sharma, S., Gajovic, S., Gruss, P., Ross, S., Thomas, P., Voss, A. K., Thomas, T., … Craig, J. E. (2003). Mutations in a novel gene, NHS, cause the pleiotropic effects of Nance-Horan syndrome, including severe congenital cataract, dental anomalies, and mental retardation. *American Journal of Human Genetics*, *73*(5), 1120–1130. https://doi.org/10.1086/379381

Byrnes, J. R., & Wolberg, A. S. (2017). New findings on venous thrombogenesis. *Hamostaseologie*, *37*(1), 25–35. https://doi.org/10.5482/HAMO-16-09-0034

Caplan, L. R. (2009). Stroke in Children and Young Adults. In *Caplan's Stroke* (pp. 523–542). Elsevier. https://doi.org/10.1016/b978-1-4160-4721-6.50016-x

Carllsson, L., Greinacher, A., Spitzer, C., Walther, R., & Kessler, C. (1997). Polymorphisms of the human platelet antigens HPA-1, HPA-2, HPA-3, and HPA-5 on the platelet receptors for fibrinogen (GPIIb/IIIa), von Willebrand factor (GPIb/IX), and collagen (GPIa/IIa) are not correlated with an increased risk of stroke. *Stroke*, *28*, 1392–1395. https://doi.org/10.1161/01.str.28.7.1392

Carlton, V. E. H., Harris, B. Z., Puffenberger, E. G., Batta, A. K., Knisely, A. S., Robinson, D. L., Strauss, K. A., Shneider, B. L., Lim, W. A., Salen, G., Morton, D. H., & Bull, L. N. (2003). Complex inheritance of familial hypercholanemia with associated mutations in TJP2 and BAAT. *Nature Genetics*, *34*(1), 91–96. https://doi.org/10.1038/ng1147

Carmeliet, P. (2000). Mechanisms of angiogenesis and arteriogenesis. *Nat Med*, *6*(4), 389–395. https://doi.org/10.1038/74651

Carmeliet, P. (2003). Angiogenesis in health and disease. *General Pharmacology: Vascular System*, *9*(6), 653–660. https://doi.org/10.1016/S0306-3623(01)00111-2

Carmeliet, P., Ferreira, V., Breier, G., Pollefeyt, S., Kieckens, L., Gertsenstein, M., Fahrig, M., Vandenhoeck, A., Harpal, K., Eberhardt, C., Declercq, C., Pawling, J., Moons, L., Collen, D., Risaut, W., & Nagy, A. (1996). Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. In *Nature* (Vol. 380, Issue 6573, pp. 435–439). https://doi.org/10.1038/380435a0

Carson, A. R., Smith, E. N., Matsui, H., Brækkan, S. K., Jepsen, K., Hansen, J. B., & Frazer, K. A. (2014). Effective filtering strategies to improve data quality from population-based whole exome sequencing studies. *BMC Nephrology*, *15*(1), 1–15. https://doi.org/10.1186/1471-2105-15-125

Catto, A., Mc Cormack, L., & Mansfield, MW, et al. (2000). Apolipoprotein E polymorphism in cerebrovascular disease. *Acta Neurol Scand*, *101*, 399–404. https://doi.org/10.1034/j.1600-0404.2000.90308a.x

Cerami, E. G., Gross, B. E., Demir, E., Rodchenkov, I., Babur, O., Anwar, N., Schultz, N., Bader, G. D., & Sander, C. (2011). Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Research*, *39*(Database issue), D685-90. https://doi.org/10.1093/nar/gkq1039

Cevik, O., Baykal, A. T., & Sener, A. (2016). Platelets Proteomic Profiles of Acute Ischemic Stroke Patients. *PLOS ONE*, *11*(6), e0158287. https://doi.org/10.1371/journal.pone.0158287

Chabrier, S., Husson, B., Dinomais, M., Landrieu, P., & Nguyen The Tich, S. (2011). New insights (and new interrogations) in perinatal arterial ischemic stroke. *Thrombosis Research*, *127*(1), 13–22. https://doi.org/10.1016/j.thromres.2010.10.003

Chabrier, S., Peyric, E., Drutel, L., Deron, J., Kossorotoff, M., Dinomais, M., Lazaro, L., Lefranc, J., Thébault, G., Dray, G., Fluss, J., Renaud, C., Nguyen The Tich, S., Darteyre, S., Dégano, C., Delion, M., Groeschel, S., Hertz-Pannier, L., Husson, B., … Vuillerot, C. (2016). Multimodal outcome at 7 years of age after neonatal arterial ischemic stroke. *Journal of Pediatrics*, *172*, 156-161.e3. https://doi.org/10.1016/j.jpeds.2016.01.069

Chalmers, E. A. (2005). Perinatal stroke - Risk factors and management. In *British Journal of Haematology* (Vol. 130, Issue 3, pp. 333–343). https://doi.org/10.1111/j.1365-2141.2005.05554.x

Chasnoff, I., Bussey, M., Savich, R., & Stack, C. (1986). Perinatal cerebral infarction and maternal cocaine use. *J Pediatr*, *86*(March), 80896–80904. https://doi.org/10.1016/s0022-3476(86)80896-4

Chen, Y., WU, J., Yang, S., Huang, C., Chang, C., Sun, G., & Lin, T. (2012). Stroke, angiogenesis and phytochemicals. *Front Biosci*, *1*(r), 599–610. https://doi.org/10.2741/s287

Cheng, Y., Austin, S. C., Rocca, B., Koller, B. H., Coffman, T. M., Grosser, T., Lawson, J. A., & FitzGerald, G. A. (2002). Role of prostacyclin in the cardiovascular response to thromboxane A2. *Science*, *296*(5567), 539–541. https://doi.org/10.1126/science.1068711

Choi, B. O., Kim, N. K., & Kim, S. H. (2003). Homozygous C677T mutation in the MTHFR gene as an independent risk factor for multiple small-artery occlusions. *Thromb. Res*, *111*, 39–44. https://doi.org/10.1016/j.thromres.2003.08.022

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*(16), 2745–2747. https://doi.org/10.1093/bioinformatics/btv195

Collins, F. S. (2011). Reengineering Translational Science: The Time Is Right. *Science Translational Medicine*, *3*(90), 1–6. https://doi.org/10.1126/scitranslmed.3002747

Conchillo-Solé, O., de Groot, N. S., Avilés, F. X., Vendrell, J., Daura, X., & Ventura, S. (2007). AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics*, *8*(1), 65. https://doi.org/10.1186/1471-2105-8-65

Cooper, G. M., Stone, E. A., Asimenos, G., Green, E. D., Batzoglou, S., & Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. *Genome Research*, *15*(7), 901–913. https://doi.org/10.1101/gr.3577405

Cosentino, F., Christopher Sill, J., & Katušić, Z. S. (1994). Role of superoxide anions in the mediation of endothelium-dependent contractions. *Hypertension*, *23*(2), 229–235. https://doi.org/10.1161/01.HYP.23.2.229

Cosentino, F., & Volpe, M. (2005). Hypertension, stroke, and endothelium. *Current Hypertension Reports*, *7*(1), 68–71. https://doi.org/10.1007/s11906-005-0057-5

Cowan, F., Mercuri, E., Groenendaal, F., Bassi, L., Ricci, D., Rutherford, M., & De Vries, L. (2005). Does cranial ultrasound imaging identify arterial cerebral infarction in term neonates? In *Archives of Disease in Childhood: Fetal and Neonatal Edition* (Vol. 90, Issue 3). https://doi.org/10.1136/adc.2004.055558

Cross, M. J., Dixelius, J., Matsumoto, T., & Claesson-Welsh, L. (2003). VEGF-receptor signal transduction. *Trends in Biochemical Sciences*, *28*(9), 488–494. https://doi.org/10.1016/S0968-0004(03)00193-2

Curry, C. J., Bhullar, S., Holmes, J., Delozier, C. D., Roeder, E. R., & Hutchison, H. T. (2007). Risk Factors for Perinatal Arterial Stroke: A Study of 60 Mother-Child Pairs. *Pediatric Neurology*, *37*(2), 99–107. https://doi.org/10.1016/j.pediatrneurol.2007.04.007

Cushman, M., Rosendaal, F. R., Psaty, B. M., Cook, E. F., Valliere, J., Kuller, L. H., & Tracy, R. P. (1998). Factor V Leiden is not a risk factor for arterial vascular disease in the elderly: results from the Cardiovascular Health Study. *Thrombosis and Haemostasis*, *79*(5), 912–915.

D'Ambrosio, A. L., Pinsky, D. J., & Connolly, E. S. (2001). The Role of the Complement Cascade in Ischemia/Reperfusion Injury: Implications for Neuroprotection. *Molecular Medicine*, *7*(6), 367–382. https://doi.org/10.1007/bf03402183

Dahlbäck, B. (1991). Protein S and C4b-binding protein: components involved in the regulation of the protein C anticoagulant system. *Thrombosis and Haemostasis*, *66*(1), 49–61. http://www.ncbi.nlm.nih.gov/pubmed/1833851

Dahlbäck, B. (2008). Advances in understanding pathogenic mechanisms of thrombophilic disorders. *Blood*, *112*(1), 19–27. https://doi.org/10.1182/blood-2008-01-077909

Dahlbäck, B., & Stenflo, J. (1981). High molecular weight complex in human plasma between vitamin K-dependent protein S and complement component C4b-binding protein. *Proceedings of the National Academy of Sciences*, *78*(4), 2512–2516. https://doi.org/10.1073/pnas.78.4.2512

Dai, C., Cheng, D., Li, W., Zeng, S., Lu, G., & Zhang, Q. (2020). Identification of paternal germline mosaicism by MicroSeq and targeted next-generation sequencing. *Molecular Genetics & Genomic Medicine*, *8*(9). https://doi.org/10.1002/mgg3.1394

Daly, C., Wong, V., Burova, E., Wei, Y., Zabski, S., Griffiths, J., Lai, K. M., Lin, H. C., Ioffe, E., Yancopoulos, G. D., & Rudge, J. S. (2004). Angiopoietin-1 modulates endothelial cell function and gene expression via the transcription factor FKHR (FOXO1). *Genes and Development*, *18*(9), 1060–1071. https://doi.org/10.1101/gad.1189704

Darmency-Stamboul, V., Chantegret, C., Ferdynus, C., Mejean, N., Durand, C., Sagot, P., Giroud, M., Bejot, Y., & Gouyon, J. B. (2012). Antenatal factors associated with perinatal arterial ischemic stroke. *Stroke*, *43*(9), 2307–2312. https://doi.org/10.1161/STROKEAHA.111.642181

Dashti, M. J. S., & Gamieldien, J. (2017). A practical guide to filtering and prioritizing genetic variants. *BioTechniques*, *62*(1), 18–30. https://doi.org/10.2144/000114492

de Vries, L., Groenendaal, F., Eken, P., van Haastert, I., Rademaker, K., & Meiners, L. (1997). Infarcts in the Vascular Distribution of the Middle Cerebral Artery in Preterm and Fullterm Infants. *Neuropediatrics*, *28*(02), 88–96. https://doi.org/10.1055/s-2007-973679

Devlin, B., Boone, B. E., Levy, S. E., Lihm, J., Buxbaum, J. D., Wu, Y., Lewis, L., Han, Y., Boerwinkle, E., Gibbs, R. A., Fromer, M., Shakir, K., Fennell, T., Garimella, K., Banks, E., Poplin, R., Gabriel, S., De Pristo, M., Sunyaev, S., & Daly, M. J. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. In *Nature* (Vol. 485, Issue 7397, pp. 242–246). https://doi.org/10.1038/nature11011

Díaz-Guzmán, J., Egido, J. A., Gabriel-Sánchez, R., Barberá-Comes, G., Fuentes-Gimeno, B., & Fernández-Pérez, C. (2012). Stroke and transient ischemic attack incidence rate in Spain: The IBERICTUS study. *Cerebrovascular Diseases*, *34*(4), 272–281. https://doi.org/10.1159/000342652

Donahue, B. S., Gailani, D., Higgins, M. S., Drinkwater, D. C., & George, A. L. (2003). Factor V Leiden protects against blood loss and transfusion after cardiac surgery. *Circulation*, *107*(7), 1003–1008. https://doi.org/10.1161/01.CIR.0000051864.28048.01

Dong, X., Jiang, Z., Peng, Y.-L., & Zhang, Z. (2015). Revealing Shared and Distinct Gene Network Organization in Arabidopsis Immune Responses by Integrative Analysis. *Plant Physiology*, *167*(3), 1186–1203. https://doi.org/10.1104/pp.114.254292

Dong, X., Wang, J., Wang, G., Wang, J., Wang, L., & Du, Y. (2021). MTHFR A1298C gene polymorphism on stroke risk: an updated meta-analysis. *Genes and Environment*, *43*(1). https://doi.org/10.1186/s41021-021-00208-z

Downward, J. (1998). Mechanisms and consequences of activation of protein kinase B/Akt. *Current Opinion in Cell Biology*, *10*(2), 262–267. https://doi.org/10.1016/S0955-0674(98)80149-X

Dudink, J., Mercuri, E., Al-Nakib, L., Govaert, P., Counsell, S. J., Rutherford, M. A., & Cowan, F. M. (2009). Evolution of unilateral perinatal arterial ischemic stroke on conventional and diffusion-weighted MR imaging. *American Journal of Neuroradiology*, *30*(5), 998–1004. https://doi.org/10.3174/ajnr.A1480

Easwaran, V., Lee, S. H., Inge, L., Guo, L., Goldbeck, C., Garrett, E., Wiesmann, M., Garcia, P. D., Fuller, J. H., Chan, V., Randazzo, F., Gundel, R., Warren, R. S., Escobedo, J., Aukerman, S. L., Taylor, R. N., & Fantl, W. J. (2003). B-Catenin Regulates Vascular Endothelial Growth Factor Expression in Colon Cancer. *Cancer Research*, *63*(12), 3145–3153.

Ebnet, K. (2008). Organization of multiprotein complexes at cell-cell junctions. *Histochemistry and Cell Biology*, *130*(1), 1–20. https://doi.org/10.1007/s00418-008-0418-7

Ehrhart, F., Willighagen, E. L., Kutmon, M., van Hoften, M., Curfs, L. M. G., & Evelo, C. T. (2021). A resource to explore the discovery of rare diseases and their causative genes. *Scientific Data*, *8*(1). https://doi.org/10.1038/s41597-021-00905-y

El Amki, M., & Wegener, S. (2017). Improving cerebral blood flow after arterial recanalization: A novel therapeutic strategy in stroke. *International Journal of Molecular Sciences*, *18*(12). https://doi.org/10.3390/ijms18122669

El Beshlawy, A., Hussein, H. A., Abou-Elew, H. H., & Kader, M. S. E. M. A. (2004). Study of protein C, protein S, and antithrombin III in hypoxic newborns. *Pediatric Critical Care Medicine*, *5*(2), 163–166. https://doi.org/10.1097/01.PCC.0000113261.13338.C3

Elbers, J., Viero, S., MacGregor, D., DeVeber, G., & Moore, A. M. (2011). Placental pathology in neonatal stroke. *Pediatrics*, *127*(3). https://doi.org/10.1542/peds.2010-1490

Emmerich, J., Rosendaal, F. R., Cattaneo, M., Margaglione, M., De Stefano, V., Cumming, T., Arruda, V., Hillarp, A., & Reny, J. L. (2001). Combined effect of factor V Leiden and prothrombin 20210A on the risk of venous thromboembolism--pooled analysis of 8 case-control studies including 2310 cases and 3204 controls. Study Group for Pooled-Analysis in Venous Thromboembolism. *Thrombosis and Haemostasis*, *86*(3), 809–816.

Enjyoji, K., Sévigny, J., Lin, Y., Frenette, P. S., Christie, P. D., Am Esch, J. S., Imai, M., Edelberg, J. M., Rayburn, H., Lech, M., Beeler, D. L., Csizmadia, E., Wagner, D. D., Robson, S. C., & Rosenberg, R. D. (1999). Targeted disruption of cd39/ATP diphosphohydrolase results in disordered hemostasis and thromboregulation. *Nature Medicine*, *5*(9), 1010–1017. https://doi.org/10.1038/12447

Esmon, C. T. (2009). Basic mechanisms and pathogenesis of venous thrombosis. *Blood Reviews*, *23*(5), 225–229. https://doi.org/10.1016/j.blre.2009.07.002

Estan, J., & Hope, P. (1997). Unilateral neonatal cerebral infarction in full term infants. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *76*(2). https://doi.org/10.1136/fn.76.2.F88

Eyre-Walker, A., & Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, *8*(8), 610–618. https://doi.org/10.1038/nrg2146

Faeh, D., Chiolero, A., & Paccaud, F. (2006). Homocysteine as a risk factor for cardiovascular disease: should we (still) worry about? *Swiss Med Wkly*, *136*(47–48), 745–756.

Falcone, G. J., Malik, R., Dichgans, M., & Rosand, J. (2014). Current concepts and clinical applications of stroke genetics. *The Lancet Neurology*, *13*(4), 405–418. https://doi.org/10.1016/S1474-4422(14)70029-8

Fang, J., Xia, C., Cao, Z., Zheng, J. Z., Reed, E., & Jiang, B.-H. (2005). Apigenin inhibits VEGF and HIF-1 expression via PI3K/AKT/p70S6K1 and HDM2/p53 pathways. *The FASEB Journal*, *19*(3), 342–353. https://doi.org/10.1096/fj.04-2175com

Ferrara, N., Carver-Moore, K., Chen, H., Dowd, M., Lu, L., O'Shea, K. S., Powell-Braxton, L., Hillan, K. J., & Moore, M. W. (1996). Heterozygous embryonic lethality induced by targeted inactivation of VEGF. *Nature*, *380*, 439–442. https://doi.org/10.1038/380439a0

Ferraresi, P., Marchetti, G., Legnani, C., Cavallari, E., Castoldi, E., Mascoli, F., Ardissino, D., Palareti, G., & Bernardi, F. (1997). The Heterozygous 20210 G/A Prothrombin Genotype Is Associated With Early Venous Thrombosis in Inherited Thrombophilias and Is Not Increased in Frequency in Artery Disease. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *17*(11), 2418–2422. https://doi.org/10.1161/01.atv.17.11.2418

Ferrucci, L., Guralnik, J. M., Pahor, M., Harris, T., Corti, M.-C., Hyman, B. T., Wallace, R. B., & Havlik, R. J. (1997). Apolipoprotein E ε2 Allele and Risk of Stroke in the Older Population. *Stroke*, *28*(12), 2410–2416. https://doi.org/10.1161/01.STR.28.12.2410

Fischer, D., Porto, L., Stoll, H., Geisen, C., & Schloesser, R. L. (2010). Intracerebral Mass Bleeding in a Term Neonate: Manifestation of Hereditary Protein S Deficiency with a New Mutation in the PROS1 Gene. *Neonatology*, *98*(4), 337–340. https://doi.org/10.1159/000298282

Fitzgerald, K. C., & Golomb, M. R. (2007). Neonatal arterial ischemisc stroke and sinovenous thrombosis associated with meningitis. *Journal of Child Neurology*, *22*(7), 818–822. https://doi.org/10.1177/0883073807304200

Fluss, J., Garcia-Tarodo, S., Granier, M., Villega, F., Ferey, S., Husson, B., Kossorotoff, M., Muehlethaler, V., Lebon, S., & Chabrier, S. (2016). Perinatal arterial ischemic stroke related to carotid artery occlusion. *European Journal of Paediatric Neurology*, *20*(4), 639–648. https://doi.org/10.1016/j.ejpn.2016.03.003

Foerch, C., Montaner, J., Furie, K. L., Ning, M. M., & Lo, E. H. (2009). Invited article: searching for oracles? Blood biomarkers in acute stroke. *Neurology*, *73*(5), 393–399. https://doi.org/10.1212/WNL.0b013e3181b05ef9

Folkman, J., Merler, E., Abernathy, C., & Williams, G. (1971). Isolation of a tumor factor responsible for angiogenesis. *Journal of Experimental Medicine*, *133*(2), 275–288. https://doi.org/10.1084/jem.133.2.275

Folkman, J., & Shing, Y. (1992). Angiogenesis. *J Biol Chem*, *267*(16), 10931–10934.

Francioli, L. C., Polak, P. P., Koren, A., Menelaou, A., Chun, S., Renkens, I., Van Duijn, C. M., Swertz, M., Wijmenga, C., Van Ommen, G., Slagboom, P. E., Boomsma, D. I., Ye, K., Guryev, V., Arndt, P. F., Kloosterman, W. P., De Bakker, P. I. W., & Sunyaev, S. R. (2015). Genome-wide patterns and properties of de novo mutations in humans. *Nature Genetics*, *47*(7), 822–826. https://doi.org/10.1038/ng.3292

Freud, S. (1968). Infantile cerebral paralysis: by Sigmund Freud. In *Translated by L. A. Russin, 376 pages, University of Miami Press, Coral Gables, Fla*.

Fulton, D., Gratton, J. P., McCabe, T. J., Fontana, J., Fujio, Y., Walsh, K., Franke, T. F., Papapetropoulos, A., & Sessa, W. C. (1999). Regulation of endothelium-derived nitric oxide production by the protein kinase Akt. *Nature*, *399*(6736), 597–601. https://doi.org/10.1038/21218.Regulation

Fumagalli, S., Perego, C., Pischiutta, F., Zanier, E. R., & De Simoni, M. G. (2015). The ischemic environment drive s microglia and macrophage function. *Frontiers in Neurology*, *6*(MAR), 1–19. https://doi.org/10.3389/fneur.2015.00081

Gallivan, L., Markham, A. F., & Anwar, R. (1999). The Leu564 factor XIIIA variant results in significantly lower plasma factor XIII levels than the Pro564 variant. *Thrombosis and Haemostasis*, *82*(4), 1368–1370.

Ganguly, P., & Alam, S. F. (2015). Role of homocysteine in the development of cardiovascular disease. In *Nutrition Journal* (Vol. 14, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/1475-2891-14-6

García-Alix A, Arnaez J, T. A. (2022). *Neurología Neonatal de un Vistazo* (1st ed.). Cabeza de Chorlito.

Garrido-Barbero, M., Arnaez, J., Loureiro, B., Arca, G., Agut, T., & Garcia-Alix, A. (2019). The Role of Factor V Leiden, Prothrombin G20210A, and MTHFR C677T Mutations in Neonatal Cerebral Sinovenous Thrombosis. *Clinical and Applied Thrombosis/Hemostasis*, *25*, 107602961983435. https://doi.org/10.1177/1076029619834352

Gawryś, K., Turek-Jakubowska, A., Gawryś, J., Jakubowski, M., Dębski, J., Szahidewicz-Krupska, E., Trocha, M., Derkacz, A., & Doroszko, A. (2022). Platelet-Derived Drug Targets and Biomarkers of Ischemic Stroke—The First Dynamic Human LC-MS Proteomic Study. *Journal of Clinical Medicine*, *11*(5), 1198. https://doi.org/10.3390/jcm11051198

Gelfand, A. A., Croen, L. A., Torres, A. R., & Wu, Y. W. (2013). Genetic Risk Factors for Perinatal Arterial Ischemic Stroke. *Pediatric Neurology*, *48*(1), 36–41. https://doi.org/10.1016/j.pediatrneurol.2012.09.016

Gerber, H.-P., McMurtrey, A., Kowalski, J., Yan, M., Keyt, B. A., Dixit, V., & Ferrara, N. (1998). Vascular Endothelial Growth Factor Regulates Endothelial Cell Survival through the Phosphatidylinositol 3′-Kinase/Akt Signal Transduction Pathway. *Journal of Biological Chemistry*, *273*(46), 30336–30343. https://doi.org/10.1074/jbc.273.46.30336

Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., Van De Vorst, M., Van Bon, B. W. M., Willemsen, M. H., Kwint, M., Janssen, I. M., Hoischen, A., Schenck, A., Leach, R., Klein, R., Tearle, R., Bo, T., Pfundt, R., Yntema, H. G., De Vries, B. B. A., Kleefstra, T., Brunner, H. G., … Veltman, J. A. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature*, *511*(7509), 344–347. https://doi.org/10.1038/nature13394

Girard, S. L., Gauthier, J., Noreau, A., Xiong, L., Zhou, S., Jouan, L., Dionne-Laporte, A., Spiegelman, D., Henrion, E., Diallo, O., Thibodeau, P., Bachand, I., Bao, J. Y. J., Tong, A. H. Y., Lin, C. H., Millet, B., Jaafari, N., Joober, R., Dion, P. A., … Rouleau, G. A. (2011). Increased exonic de novo mutation rate in individuals with schizophrenia. *Nature Genetics*, *43*(9), 860–863. https://doi.org/10.1038/ng.886

Golomb, M. R., Dick, P. T., MacGregor, D. L., Armstrong, D. C., & DeVeber, G. A. (2003). Cranial Ultrasonography Has a Low Sensitivity for Detecting Arterial Ischemic Stroke in Term Neonates. *Journal of Child Neurology*, *18*(2), 98–103. https://doi.org/10.1177/08830738030180021401

Golomb, M. R., Fullerton, H. J., Nowak-Gottl, U., & Deveber, G. (2009). Male predominance in childhood ischemic stroke: Findings from the international pediatric stroke study. *Stroke*, *40*(1), 52–57. https://doi.org/10.1161/STROKEAHA.108.521203

Golomb, M. R., MacGregor, D. L., Domi, T., Armstrong, D. C., McCrindle, B. W., Mayank, S., & Deveber, G. A. (2001). Presumed pre- or perinatal arterial ischemic stroke: Risk factors and outcomes. *Annals of Neurology*, *50*(2), 163–168. https://doi.org/10.1002/ana.1078

Goodwin, A., & D'Amore, P. (2002). Wnt signaling in the vasculature. *Experimental Cell Research*, *1*(9), 1–9. https://doi.org/10.1016/j.yexcr.2012.12.023

Gould, A. L., Rossouw, J. E., Santanello, N. C., Heyse, J. F., & Furberg, C. D. (1998). Cholesterol Reduction Yields Clinical Benefit. *Circulation*, *97*(10), 946–952. https://doi.org/10.1161/01.CIR.97.10.946

Gould, D. B., Phalan, F. C., Breedveld, G. J., Van Mil, S. E., Smith, R. S., Schimenti, J. C., Aguglia, U., Van Der Knaap, M. S., Heutink, P., & John, S. W. M. (2005). Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. *Science*, *308*(5725), 1167–1171. https://doi.org/10.1126/science.1109418

Gould, D. B., Phalan, F. C., van Mil, S. E., Sundberg, J. P., Vahedi, K., Massin, P., Bousser, M. G., Heutink, P., Miner, J. H., Tournier-Lasserve, E., & John, S. W. M. (2006). Role of COL4A1 in small-vessel disease and hemorrhagic stroke. *The New England Journal of Medicine*, *354*(14), 1489–1496. https://doi.org/10.1056/NEJMoa053727

Govaert, P., Ramenghi, L., Taal, R., de Vries, L., & DeVeber, G. (2009a). Diagnosis of perinatal stroke I: definitions, differential diagnosis and registration. *Acta Paediatrica*, *98*(10), 1556–1567. https://doi.org/10.1111/j.1651-2227.2009.01461.x

Govaert, P., Ramenghi, L., Taal, R., Dudink, J., & Lequin, M. (2009b). Diagnosis of perinatal stroke II: mechanisms and clinical phenotypes. *Acta Paediatrica*, *98*(11), 1720–1726. https://doi.org/10.1111/j.1651-2227.2009.01462.x

Govaert, P., Vanhaesebrouck, P., & Praeter, C. de. (1993). Traumatic neonatal intracranial bleeding and stroke. *Archives of Disease in Childhood*, *68*(3 SPEC NO), 339. https://doi.org/10.1136/adc.68.3_Spec_No.339-a

Green, D. (2003). Thrombophilia and Stroke. *Topics in Stroke Rehabilitation*, *10*(3), 21–33. https://doi.org/10.1310/L9KD-N5N8-69X0-08QK

Greenberg, D. A., & Jin, K. (2013). Vascular endothelial growth factors (VEGFs) and stroke. *Cellular and Molecular Life Sciences*, *70*(10), 1753–1761. https://doi.org/10.1007/s00018-013-1282-8

Gregg, J. P., Yamane, A. J., & Grody, W. W. (1997). Prevalence of the factor V-Leiden mutation in four distinct American ethnic populations. *American Journal of Medical Genetics*, *73*(3), 334–336. https://doi.org/10.1002/(SICI)1096-8628(19971219)73:3<334::AID-AJMG20>3.0.CO;2-J

Gretarsdottir, S., Sveinbjörnsdottir, S., Jonsson, H. H., Jakobsson, F., Einarsdottir, E., Agnarsson, U., Shkolny, D., Einarsson, G., Gudjonsdottir, H. M., Valdimarsson, E. M., Einarsson, O. B., Thorgeirsson, G., Hadzic, R., Jonsdottir, S., Reynisdottir, S. T., Bjarnadottir, S. M., Gudmundsdottir, T., Gudlaugsdottir, G. J., Gill, R., … Gulcher, J. R. (2002). Localization of a Susceptibility Gene for Common Forms of Stroke to 5q12. *The American Journal of Human Genetics*, *70*(3), 593–603. https://doi.org/10.1086/339252

Grunt, S., Mazenauer, L., Buerki, S. E., Boltshauser, E., Mori, A. C., Datta, A. N., Fluss, J., Mercati, D., Keller, E., Maier, O., Poloni, C., Ramelli, G. P., Schmitt-Mechelke, T., & Steinlin, M. (2015). Incidence and outcomes of symptomatic neonatal arterial ischemic stroke. *Pediatrics*, *135*(5), e1220–e1228. https://doi.org/10.1542/peds.2014-1520

Gudenkauf, F. J., Azamian, M. S., Hunter, J. V., Nayak, A., & Lalani, S. R. (2020). A novel CACNA1A variant in a child with early stroke and intractable epilepsy. *Molecular Genetics and Genomic Medicine*, *8*(10). https://doi.org/10.1002/mgg3.1383

Günther, G., Junker, R., Sträter, R., Schobess, R., Kurnik, K., Kosch, A., & Nowak-Göttl, U. (2000). Symptomatic ischemic stroke in full-term neonates: Role of acquired and genetic prothrombotic risk factors. *Stroke*, *31*(10), 2437–2441. https://doi.org/10.1161/01.STR.31.10.2437

Guo, H., Chi, J., Xing, Y., & Wang, P. (2009). Influence of folic acid on plasma homocysteine levels & arterial endothelial function in patients with unstable angina. *Indian J Med Res*, *129*(3), 279–284.

Guo, J. F., Zhang, L., Li, K., Mei, J. P., Xue, J., Chen, J., Tang, X., Shen, L., Jiang, H., Chen, C., Guo, H., Wu, X. L., Sun, S. L., Xu, Q., Sun, Q. Y., Chan, P., Shang, H. F., Wang, T., Zhao, G. H., … Tang, B. S. (2018). Coding mutations in NUS1 contribute to Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11567–11572. https://doi.org/10.1073/pnas.1809969115

Guo, M. H., Plummer, L., Chan, Y. M., Hirschhorn, J. N., & Lippincott, M. F. (2018). Burden Testing of Rare Variants Identified through Exome Sequencing via Publicly Available Control Data. *American Journal of Human Genetics*, *103*(4), 522–534. https://doi.org/10.1016/j.ajhg.2018.08.016

Gussow, A. B., Petrovski, S., Wang, Q., Allen, A. S., & Goldstein, D. B. (2016). The intolerance to functional genetic variation of protein domains predicts the localization of pathogenic mutations within genes. *Genome Biology*, *17*(1), 9. https://doi.org/10.1186/s13059-016-0869-4

Hamzi, K., Tazzite, A., & Nadifi, S. (2011). Large-scale meta-analysis of genetic studies in ischemic stroke: Five genes involving 152,797 individuals. *Indian Journal of Human Genetics*, *17*(3), 212–217. https://doi.org/10.4103/0971-6866.92105

Happ, H. C., & Carvill, G. L. (2020). A 2020 View on the Genetics of Developmental and Epileptic Encephalopathies. *Epilepsy Currents*, *20*(2), 90–96. https://doi.org/10.1177/1535759720906118

Harel, T., Posey, J. E., Graham, B. H., Walkiewicz, M., Yang, Y., Lalani, S. R., & Belmont, J. W. (2015). Atypical presentation of moyamoya disease in an infant with a de novo RNF213 variant. *American Journal of Medical Genetics, Part A*, *167*(11), 2742–2747. https://doi.org/10.1002/ajmg.a.37230

Harteman, J. C., Groenendaal, F., Kwee, A., Welsing, P. M. J., Benders, M. J. N. L., & De Vries, L. S. (2012). Risk factors for perinatal arterial ischaemic stroke in full-term infants: A case-control study. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *97*(6). https://doi.org/10.1136/archdischild-2011-300973

Hassan, A., & Markus, H. (2000). Genetics and ischaemic stroke. *Brain*, *123*(9), 1784–1812. https://doi.org/10.1136/jnnp-2012-304834

Hemmings, B. A. (1997). Akt signaling: Linking membrane events to life and death decisions. *Science*, *275*(5300), 628–630. https://doi.org/10.1126/science.275.5300.628

Henkens, C., Bom, V., Seinen, A., & van der Meer, J. (1995). Sensitivity to activated protein c: influence of oral contraceptives and sex. *Thromb Haemost*, *73*, 402–440.

Hiraoka, M., Kato, K., Saito, Y., Yasuda, K., & Kagawa, Y. (2004). Gene–nutrient and gene–gene interactions of controlled folate intake by Japanese women. *Biochemical and Biophysical Research Communications*, *316*(4), 1210–1216. https://doi.org/10.1016/j.bbrc.2004.02.174

Hoeben, A., Landuyt, B., Highley, M., Wildiers, H., Van Oosterom, A., & De Bruijn, E. (2004). Vascular Endothelial Growth Factor and Angiogenesis. *Pharmacol Rev*, *56*(4), 549–580. https://doi.org/10.1124/pr.56.4.3.549

Ichord, R. (2020). *Stroke in the newborn: Classification, manifestations, and diagnosis*. www.uptodate.com

Ikram, M. A., Seshadri, S., Bis, J. C., Fornage, M., DeStefano, A. L., Aulchenko, Y. S., Debette, S., Lumley, T., Folsom, A. R., van den Herik, E. G., Bos, M. J., Beiser, A., Cushman, M., Launer, L. J., Shahar, E., Struchalin, M., Du, Y., Glazer, N. L., Rosamond, W. D., … Wolf, P. A. (2009). Genomewide Association Studies of Stroke. *New England Journal of Medicine*, *360*(17), 1718–1728. https://doi.org/10.1056/nejmoa0900094

Instituto Nacional de Estadística. (2017). *Encuesta Nacional de Salud 2017. Cuestionario de Adultos.* https://www.sanidad.gob.es/estadEstudios/estadisticas/encuestaNacional/encuesta2017.htm

Ioannidis, J. P. A., Ntzani, E. E., & Trikalinos, T. A. (2004). "Racial" differences in genetic effects for complex diseases. *Nature Genetics*, *36*(12), 1312–1318. https://doi.org/10.1038/ng1474

Jackson, S. P. (2011). Arterial thrombosis-insidious, unpredictable and deadly. *Nature Medicine*, *17*(11), 1423–1436. https://doi.org/10.1038/nm.2515

James, A. H., Bushnell, C. D., Jamison, M. G., & Myers, E. R. (2006). Incidence and risk factors for stroke in pregnancy and the puerperium. *Obstetrical and Gynecological Survey*, *61*(1), 4–5. https://doi.org/10.1097/01.ogx.0000193837.87337.d7

Jansson, T. (2016). Placenta plays a critical role in maternal–fetal resource allocation. *Proceedings of the National Academy of Sciences*, *113*(40), 11066–11068. https://doi.org/10.1073/pnas.1613437113

Jerrard-Dunne, P., Cloud, G., Hassan, A., & Markus, H. S. (2003). Evaluating the genetic component of ischemic stroke subtypes: A family history study. *Stroke*, *34*(6), 1364–1369. https://doi.org/10.1161/01.STR.0000069723.17984.FD

Jiménez-González, M. C., Santiago-Germán, D., Castillo-Henkel, E. F., Alvarado-Moreno, J. A., Hernández-Juárez, J., Leaños-Miranda, A., Majluf-Cruz, A., & Isordia-Salas, I. (2021). Identification of genetic risk factors associated with ischaemic stroke in young Mexican patients. In *Neurología* (Vol. 36). https://doi.org/10.1016/j.nrleng.2018.01.011

Jun Olsen, J., Öther-Gee Pohl, S., Deshmukh, A., Visweswaran, M., Ward, N. C., Arfuso, F., Agostino, M., & Dharmarajan, A. (2017). The Role of Wnt Signalling in Angiogenesis. In *Clin Biochem Rev* (Vol. 131, Issue 3).

Junker, R., Koch, H. G., Auberger, K., Münchow, N., Ehrenforth, S., & Nowak-Göttl, U. (1999). Prothrombin G20210A gene mutation and further prothrombotic risk factors in childhood thrombophilia. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *19*(10), 2568–2572. https://doi.org/10.1161/01.ATV.19.10.2568

Junker, R., & Nowak-Göttl, U. (1998). The prothrombin G20210A mutation: a common cause of thrombophilia? *J Lab Med*, *22*, 472–482.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., … MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. https://doi.org/10.1038/s41586-020-2308-7

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*(3), 310–315. https://doi.org/10.1038/ng.2892

Kirton, A., Armstrong-Wells, J., Chang, T., DeVeber, G., Rivkin, M. J., Hernandez, M., Carpenter, J., Yager, J. Y., Lynch, J. K., & Ferriero, D. M. (2011). Symptomatic neonatal arterial ischemic stroke: The international pediatric stroke study. *Pediatrics*, *128*(6). https://doi.org/10.1542/peds.2011-1148

Kirton, A., DeVeber, G., Pontigon, A. M., Macgregor, D., & Shroff, M. (2008). Presumed perinatal ischemic stroke: Vascular classification predicts outcomes. *Annals of Neurology*, *63*(4), 436–443. https://doi.org/10.1002/ana.21334

Kirton, A., Shroff, M., Visvanathan, T., & deVeber, G. (2007). Quantified corticospinal tract diffusion restriction predicts neonatal stroke outcome. *Stroke*, *38*(3), 974–980. https://doi.org/10.1161/01.STR.0000258101.67119.72

Klein, D., Demory, A., Peyre, F., Kroll, J., Géraud, C., & Ohnesorge, N. (2009). Wnt2 acts as an angiogenic growth factor for non-sinusoidal endothelial cells and inhibits expression of stanniocalcin-1. *Angiogenesis*, *12*, 251–265. https://doi.org/10.1007/s10456-009-9145-5

Koelfen, W., & Freund, M. (1995). Neonatal Stroke Involving the Middle Cerebral Artery in Term Infants : Clinical Presentation, EEG and Imaging Studies , and Outcome. *Developmental Medicine and Child Neurology*, *37*, 204–212. https://doi.org/10.1111/j.1469-8749.1995.tb11993.x

Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S. A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., Wong, W. S. W., Sigurdsson, G., Walters, G. B., Steinberg, S., Helgason, H., Thorleifsson, G., Gudbjartsson, D. F., Helgason, A., Magnusson, O. T., … Stefansson, K. (2012). Rate of de novo mutations and the importance of fathers age to disease risk. *Nature*, *488*(7412), 471–475. https://doi.org/10.1038/nature11396

Konialis, C., Spengos, K., Iliopoulos, P., Karapanou, S., Gialafos, E., Hagnefelt, B., Vemmos, K., Zakopoulos, N., & Pangalos, C. (2016). The APOE E4 allele confers increased risk of Ischemic Stroke among Greek carriers. *Advances in Clinical and Experimental Medicine*, *25*(3), 471–478. https://doi.org/10.17219/acem/38841

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R., & Massouras, A. (2019). VarSome: the human genomic variant search engine. *Bioinformatics*, *35*(11), 1978–1980. https://doi.org/10.1093/bioinformatics/bty897

Kraus, F. T., & Acheen, V. I. (1999). Fetal thrombotic vasculopathy in the placenta: Cerebral thrombi and infarcts, coagulopathies, and cerebral palsy. *Human Pathology*, *30*(7), 759–769. https://doi.org/10.1016/S0046-8177(99)90136-3

Kujovich, J. L. (2011). Factor v Leiden thrombophilia. *Genetics in Medicine*, *13*(1), 1–16. https://doi.org/10.1097/GIM.0b013e3181faa0f2

Kumar, D. R., Hanlin, E. R., Glurich, I., Mazza, J. J., & Yale, S. H. (2010). Virchow's contribution to the understanding of thrombosis and cellular biology. *Clinical Medicine and Research*, *8*(3–4), 168–172. https://doi.org/10.3121/cmr.2009.866

Kureishi, Y., Luo, Z., Shiojima, I., Bialik, A., Fulton, D., Lefer, D. J., Sessa, W. C., & Walsh, K. (2000). The HMG-CoA reductase inhibitor simvastatin activates the protein kinase Akt and promotes angiogenesis in normocholesterolemic animals. *Nature Medicine*, *6*(9), 1004–1010. https://doi.org/10.1038/79510

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, *157*(1), 105–132. https://doi.org/10.1016/0022-2836(82)90515-0

Lakhan, S. E., Kirchgessner, A., & Hofer, M. (2009). Inflammatory mechanisms in ischemic stroke: Therapeutic approaches. *Journal of Translational Medicine*, *7*, 1–11. https://doi.org/10.1186/1479-5876-7-97

Lalouschek, W., Aull, S., Serles, W., Schnider, P., Mannhalter, C., & Lang, T. (1999). Genetic and nongenetic factors influencing plasma homocysteine levels in patients with ischemic cerebrovascular disease and in healthy control subjects. *J Lab Clin Med*, *133*, 575–582. https://doi.org/10.1016/s0022-2143(99)90187-7

Lalouschek, W., Schillinger, M., Hsieh, K., Endler, G., Tentschert, S., Lang, W., Cheng, S., & Mannhalter, C. (2005). Matched case-control study on factor V Leiden and the prothrombin G20210A mutation in patients with ischemic stroke/transient ischemic attack up to the age of 60 years. *Stroke*, *36*(7), 1405–1409. https://doi.org/10.1161/01.STR.0000170635.45745.b8

Larroche, J. C., Droullé, P., Delezoide, A. L., Narcy, F., & Nessmann, C. (1990). Brain damage in monozygous twins. *Neonatology*, *57*(5), 261–278. https://doi.org/10.1159/000243201

Laugesaar, R., Kolk, A., Tomberg, T., Metsvaht, T., Lintrop, M., Varendi, H., & Talvik, T. (2007). Acutely and retrospectively diagnosed perinatal stroke: A population-based study. *Stroke*, *38*(8), 2234–2240. https://doi.org/10.1161/STROKEAHA.107.483743

Lee, D. H., Walker, I. R., Teitel, J., Poon, M.-C., Ritchie, B., Akabutu, J., Sinclair, G. D., Pai, M., Wu, J. W. Y., Reddy, S., Carter, C., Growe, G., Lillicrap, D., Lam, M., & Blajchman, M. A. (2000). Effect of the Factor V Leiden Mutation on the Clinical Expression of Severe Hemophilia A. *Thrombosis and Haemostasis*, *83*(03), 387–391. https://doi.org/10.1055/s-0037-1613824

Lee, J., Croen, L. a, Backstrand, K. H., Yoshida, C. K., Henning, L. H., Lindan, C., Ferriero, D. M., Fullerton, H. J., & Wu, Y. W. (2005). *Maternal and infant characteristics associated with perinatal arterial stroke in the infant*. *293*(6), 723–729. https://doi.org/10.1001/jama.293.6.723

Lee, M., Luciano, A., Ackah, E., Rodriguez-Vitad, J., Bancroft, T. A., Eichmann, A., Simons, M., Kyriakides, T., Morales-Ruiz, M., & Sessa, W. C. (2014). Endothelial Akt1 mediates angiogenesis by phosphorylating multiple angiogenic substrates. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(35), 12865–12870. https://doi.org/10.1073/pnas.1408472111

Lee, S., Emond, M. J., Bamshad, M. J., Barnes, K. C., Rieder, M. J., Nickerson, D. A., NHLBI GO Exome Sequencing Project—ESP Lung Project Team, Christiani, D. C., Wurfel, M. M., & Lin, X. (2012). Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *American Journal of Human Genetics*, *91*(2), 224–237. https://doi.org/10.1016/j.ajhg.2012.06.007

Legnani, C., Palareti, G., Guazzaloca, G., Cosmi, B., Lunghi, B., Bernardi, F., & Coccheri, S. (2002). Venous thromboembolism in young women: Role of thrombophilic mutations and oral contraceptive use. *European Heart Journal*, *23*(12), 984–990. https://doi.org/10.1053/euhj.2001.3082

Lehman, L. L., & Rivkin, M. J. (2014). Perinatal arterial ischemic stroke: Presentation, risk factors, evaluation, and outcome. *Pediatric Neurology*, *51*(6), 760–768. https://doi.org/10.1016/j.pediatrneurol.2014.07.031

Lek, M., Karczewski, K. J., Minikel, E. V, Samocha, K. E., Banks, E., Fennell, T., O'Donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B., Tukiainen, T., Birnbaum, D. P., Kosmicki, J. A., Duncan, L. E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., … Exome Aggregation Consortium. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*(7616), 285–291. https://doi.org/10.1038/nature19057

Lemmens, R., Maugeri, A., Niessen, H. W. M., Goris, A., Tousseyn, T., Demaerel, P., Corveleyn, A., Robberecht, W., Knaap, V. der, Thijs, V. N., & Zwijnenburg, P. J. G. (2013). Novel COL4A1 mutations cause cerebral small vessel disease by haploinsufficiency. *Human Molecular Genetics*, *22*(2), 391–397. https://doi.org/10.1093/hmg/dds436

Lequin, M. H., Peeters, E. A. J., Holscher, H. C., de Krijger, R., & Govaert, P. (2004). Arterial infarction caused by carotid artery dissection in the neonate. *European Journal of Paediatric Neurology*, *8*(3), 155–160. https://doi.org/10.1016/j.ejpn.2004.02.001

Li, B., & Leal, S. M. (2008). Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *American Journal of Human Genetics*, *83*(3), 311–321. https://doi.org/10.1016/j.ajhg.2008.06.024

Li, W., Zhang, Y., Wang, Y., Rong, Z., Liu, C., Miao, H., Chen, H., He, Y., He, W., & Chen, L. (2021). Candidate gene prioritization for chronic obstructive pulmonary disease using expression information in protein–protein interaction networks. *BMC Pulmonary Medicine*, *21*(1), 280. https://doi.org/10.1186/s12890-021-01646-9

Liao, S. Y., Kuo, I. Y., Chen, Y. T., Liao, P. C., Liu, Y. F., Wu, H. Y., Lai, W. W., & Wang, Y. C. (2019). AKT-mediated phosphorylation enhances protein stability and transcription activity of ZNF322A to promote lung cancer progression. *Oncogene*, *38*(41), 6723–6736. https://doi.org/10.1038/s41388-019-0928-x

Liesz, A., Dalpke, A., Mracsko, E., Antoine, D. J., Roth, S., Zhou, W., Yang, H., Na, S. Y., Akhisaroglu, M., Fleming, T., Eigenbrod, T., Nawroth, P. P., Tracey, K. J., & Veltkamp, R. (2015). DAMP signaling is a key pathway inducing immune modulation after brain injury. *Journal of Neuroscience*, *35*(2), 583–598. https://doi.org/10.1523/JNEUROSCI.2439-14.2015

Ling-Chih, W., Chien-Ta, C., & Kun-Feng, L. (2019). A Case of Fabry Disease Presenting with Young Stroke and Fever. *Acta Neurol Taiwan*, *28*, 52–56.

Linnemann, B., Schindewolf, M., Zgouras, D., Erbe, M., Jarosch-Preusche, M., & Lindhoff-Last, E. (2008). Are patients with thrombophilia and previous venous thromboembolism at higher risk to arterial thrombosis? *Thrombosis Research*, *121*(6), 743–750. https://doi.org/10.1016/j.thromres.2007.07.014

Luisiri, A., Graviss, E. R., Weber, T., Silberstein, M. J., Tantana, S., Connors, R., & Brodeur, A. E. (1988). Neurosonographic changes in newborns treated with extracorporeal membrane oxygenation. *Journal of Ultrasound in Medicine*, *7*(8), 429–438. https://doi.org/10.7863/jum.1988.7.8.429

Luo, Z., Fujio, Y., Kureishi, Y., Rudic, R. D., Daumerie, G., Fulton, D., Sessa, W. C., & Walsh, K. (2000). Acute modulation of endothelial Akt/PKB activity alters nitric oxide–dependent vasomotor activity in vivo. *Journal of Clinical Investigation*, *106*(4), 493–499. https://doi.org/10.1172/JCI9419

Lüscher, T. F., & Tanner, F. C. (1993). Endothelial regulation of vascular tone and growth. *Am J Hypertens.*, *6*, 283S-293S. https://doi.org/10.1093/ajh/6.7.283s

Lynch, J. K., Han, C. J., Nee, L. E., & Nelson, K. B. (2005). Prothrombotic factors in children with stroke or porencephaly. *Pediatrics*, *116*(2), 447–453. https://doi.org/10.1542/peds.2004-1905

MacArthur, D. G., Manolio, T. A., Dimmock, D. P., Rehm, H. L., Shendure, J., Abecasis, G. R., Adams, D. R., Altman, R. B., Antonarakis, S. E., Ashley, E. A., Barrett, J. C., Biesecker, L. G., Conrad, D. F., Cooper, G. M., Cox, N. J., Daly, M. J., Gerstein, M. B., Goldstein, D. B., Hirschhorn, J. N., … Gunter, C. (2014). Guidelines for investigating causality of sequence variants in human disease. *Nature*, *508*(7497), 469–476. https://doi.org/10.1038/nature13127

Mackman, N. (2012). New insights into the mechanisms of venous thrombosis. *Journal of Clinical Investigation*, *122*(7), 2331–2336. https://doi.org/10.1172/JCI60229

Madsen, B. E., & Browning, S. R. (2009). A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genetics*, *5*(2), e1000384. https://doi.org/10.1371/journal.pgen.1000384

Maingay-de Groof, F., Lequin, M. H., Roofthooft, D. W., Oranje, A. P., de Coo, I. F., Bok, L. A., van der Spek, P. J., Mancini, G. M., & Govaert, P. P. (2008). Extensive cerebral infarction in the newborn due to incontinentia pigmenti. *European Journal of Paediatric Neurology*, *12*(4), 284–289. https://doi.org/10.1016/j.ejpn.2007.09.001

Majid, Z., Tahir, F., Ahmed, J., Bin Arif, T., & Haq, A. (2020). Protein C Deficiency as a Risk Factor for Stroke in Young Adults: A Review. *Cureus*. https://doi.org/10.7759/cureus.7472

Malik, R., Rannikmäe, K., Traylor, M., Georgakis, M. K., Sargurupremraj, M., Markus, H. S., Hopewell, J. C., Debette, S., Sudlow, C. L. M., & Dichgans, M. (2018). Genome-wide meta-analysis identifies 3 novel loci associated with stroke. *Annals of Neurology*, *84*(6), 934–939. https://doi.org/10.1002/ana.25369

Manolescu, B. N., Oprea, E., Farcasanu, I. C., Berteanu, M., & Cercasov, C. (2010). Homocysteine and vitamin therapy in stroke prevention and treatment: a review. *Acta Biochimica Polonica*, *57*(4). https://doi.org/10.18388/abp.2010_2432

Marcucci, R., Sofi, F., Fedi, S., Lari, B., Sestini, I., Cellai, A. P., Pulli, R., Pratesi, G., Pratesi, C., Gensini, G. F., & Abbate, R. (2005). Thrombophilic risk factors in patients with severe carotid atherosclerosis. *Journal of Thrombosis and Haemostasis*, *3*(3), 502–507. https://doi.org/10.1111/j.1538-7836.2005.01173.x

Mardis E. R. (2016). The challenges of big data. *Disease Models & Mechanisms*, *9*(5), 483–485. https://doi.org/10.1242/dmm.025585

Margaglione, M., D'Andrea, G., Giuliani, N., Brancaccio, V., De Lucia, D., Grandone, E., De Stefano, V., Tonali, P. A., & Di Minno, G. D. (1999). Inherited prothrombotic conditions and premature ischemic stroke: Sex difference in the association with factor V Leiden. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *19*(7), 1751–1756. https://doi.org/10.1161/01.ATV.19.7.1751

Martijn Pruissen, D. O., C Slooter, A. J., Rosendaal, F. R., van der Graaf, Y., & Algra, A. (2008). Coagulation factor XIII gene variation, oral contraceptives, and risk of ischemic stroke. *Blood*. https://doi.org/10.1182/blood-2007-08

Matsumoto, T., & Claesson-welsh, L. (2003). VEGF Receptor Signal Transduction. *Trends Biochem Sci*, *28*(9), 488–494. https://doi.org/10.1126/stke.2001.112.re21

Mayo Clinic Staff. (2021). *Stroke*. https://www.mayoclinic.org/diseases-conditions/stroke/symptoms-causes/syc-20350113#:~:text=Overview,brain damage and other complications.

McCarron, M. O., Delong, D., & Alberts, M. J. (1999). APOE genotype as a risk factor for ischemic cerebrovascular disease: A meta-analysis. *Neurology*, *53*(6), 1308–1308. https://doi.org/10.1212/WNL.53.6.1308

McKevitt, C., Fudge, N., Redfern, J., Sheldenkar, A., Crichton, S., Rudd, A. R., Forster, A., Young, J., Nazareth, I., Silver, L. E., Rothwell, P. M., & Wolfe, C. D. A. (2011). Self-reported long-term needs after stroke. *Stroke*, *42*(5), 1398–1403. https://doi.org/10.1161/STROKEAHA.110.598839

McQuillen, P. S., Hamrick, S. E. G., Perez, M. J., Barkovich, A. J., Glidden, D. V., Karl, T. R., Teitel, D., & Miller, S. P. (2006). Balloon atrial septostomy is associated with preoperative stroke in neonates with transposition of the great arteries. *Circulation*, *113*(2), 280–285. https://doi.org/10.1161/CIRCULATIONAHA.105.566752

Ment, L. R., Ehrenkranz, R. A., & Duncan, C. C. (1986). Bacterial meningitis as an etiology of perinatal cerebral infarction. *Pediatr Neurol*, *2*(5), 276–279. https://doi.org/10.1016/0887-8994(86)90019-6

Mercuri, E., Rutherford, M., Cowan, F., Pennock, J., Counsell, S., Papadimitriou, M., Azzopardi, D., Bydder, G., & Dubowitz, L. (1999). Early prognostic indicators of outcome in infants with neonatal cerebral infarction: A clinical, electroencephalogram, and magnetic resonance imaging study. *Pediatrics*, *103*(1), 39–46. https://doi.org/10.1542/peds.103.1.39

Michaelson, I. . (1948). The mode of development of the vascular system of the retina with some observations on its significance for certain retinal disorders. *Trans. Ophthalmol. Soc. U. K.*, *68*(137–180).

Millar, D. S., Johansen, B., Berntorp, E., Minford, A., Bolton-Maggs, P., Wensley, R., Kakkar, V., Schulman, S., Torres, A., Bosch, N., & Cooper, D. N. (2000). Molecular genetic analysis of severe protein C deficiency. *Human Genetics*, *106*(6), 646–653. https://doi.org/10.1007/s004390000315

Miller, J. R. (2002). The Wnts. *Genome Biology*, *3*(1), 1–15. https://doi.org/10.1186/gb-2001-3-1-reviews3001

Mineyko, A., & Kirton, A. (2011). The black box of perinatal ischemic stroke pathogenesis. *Journal of Child Neurology*, *26*(9), 1154–1162. https://doi.org/10.1177/0883073811408312

Miosge, L. A., Field, M. A., Sontani, Y., Cho, V., Johnson, S., Palkova, A., Balakishnan, B., Liang, R., Zhang, Y., Lyon, S., Beutler, B., Whittle, B., Bertram, E. M., Enders, A., Goodnow, C. C., & Andrewsa, T. D. (2015). Comparison of predicted and actual consequences of missense mutations. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(37), E5189–E5198. https://doi.org/10.1073/pnas.1511585112

Mochida, G. H., Ganesh, V. S., Felie, J. M., Gleason, D., Hill, R. S., Clapham, K. R., Rakiec, D., Tan, W. H., Akawi, N., Al-Saffar, M., Partlow, J. N., Tinschert, S., Barkovich, A. J., Ali, B., Al-Gazali, L., & Walsh, C. A. (2010). A homozygous mutation in the tight-junction protein JAM3 causes hemorrhagic destruction of the brain, subependymal calcification, and congenital cataracts. *American Journal of Human Genetics*, *87*(6), 882–889. https://doi.org/10.1016/j.ajhg.2010.10.026

Monroe, J. (2020). Ischemic stroke. *MEDSURG Nursing*, *29*(4), 278–279. https://doi.org/10.5005/jp/books/12952_6

Morales-Ruiz, M., Fulton, D., Sowa, G., Languino, L. R., Fujio, Y., Walsh, K., & Sessa, W. C. (2000). Vascular Endothelial Growth Factor–Stimulated Actin Reorganization and Migration of Endothelial Cells Is Regulated via the Serine/Threonine Kinase Akt. *Circulation Research*, *86*(8), 892–896. https://doi.org/10.1161/01.RES.86.8.892

Moroi, B. M., Jung, S. M., Shinmyozu, K., Tomiyama, Y., Ordinas, A., & Diaz-ricart, M. (1996). Analysis of Platelet Adhesion to a Collagen-Coated Surface Under Flow Conditions: The Involvement of Glycoprotein VI in the Platelet Adhesion. *Blood*, *88*(6), 2081–2092.

Moutsianas, L., Agarwala, V., Fuchsberger, C., Flannick, J., Rivas, M. A., Gaulton, K. J., Albers, P. K., McVean, G., Boehnke, M., Altshuler, D., & McCarthy, M. I. (2015). The Power of Gene-Based Rare Variant Methods to Detect Disease-Associated Variation and Test Hypotheses About Complex Disease. *PLoS Genetics*, *11*(4). https://doi.org/10.1371/journal.pgen.1005165

Nakashima, M. O., & Rogers, H. J. (2014). Hypercoagulable states: An algorithmic approach to laboratory testing and update on monitoring of direct oral anticoagulants. *Blood Research*, *49*(2), 85–94. https://doi.org/10.5045/br.2014.49.2.85

Navlakha, S., & Kingsford, C. (2010). The power of protein interaction networks for associating genes with diseases. *Bioinformatics (Oxford, England)*, *26*(8), 1057–1063. https://doi.org/10.1093/bioinformatics/btq076

Nelson, K. B. (2007). Perinatal ischemic stroke. *Stroke*, *38*(2 PART 2), 742–745. https://doi.org/10.1161/01.STR.0000247921.97794.5e

Nelson, K. B., Lynch, J. K., & The. (2004). Stroke in newborn infants. *Medicinski Arhiv*, *61*(1), 62–64. https://doi.org/10.1016/S1474-4422(04)00679-9

Neonatal encephalopathy and neurologic outcome, second edition. (2014). Neonatal Encephalopathy and Neurologic Outcome, Second Edition. *Pediatrics*, *133*(5), e1482–e1488. https://doi.org/10.1542/peds.2014-0724

Nesbitt, W. S., Westein, E., Tovar-Lopez, F. J., Tolouei, E., Mitchell, A., Fu, J., Carberry, J., Fouras, A., & Jackson, S. P. (2009). A shear gradient-dependent platelet aggregation mechanism drives thrombus formation. *Nature Medicine*, *15*(6), 665–673. https://doi.org/10.1038/nm.1955

Ng, K. W. P., Loh, P. K., & Sharma, V. K. (2011). Role of Investigating Thrombophilic Disorders in Young Stroke. *Stroke Research and Treatment*, *2011*, 1–9. https://doi.org/10.4061/2011/670138

Nieswandt, B., & Watson, S. P. (2003). Platelet-collagen interaction: Is GPVI the central receptor? *Blood*, *102*(2), 449–461. https://doi.org/10.1182/blood-2002-12-3882

Nishijima, D., K. Simel, D. L., Wisner, D. H., & Holmes, J. F. (2016). The Functional Role of Regulatory Lymphocytes in Stroke: Facts and Controversies. *Physiology & Behavior*, *176*(1), 139–148. https://doi.org/10.1161/STROKEAHA.114.008608.The

Nitsch, D., Gonçalves, J. P., Ojeda, F., de Moor, B., & Moreau, Y. (2010). Candidate gene prioritization by network analysis of differential expression using machine learning approaches. *BMC Bioinformatics*, *11*(1), 460. https://doi.org/10.1186/1471-2105-11-460

Nowak-Göttl, U., Sträter, R., Heinecke, A., Junker, R., Koch, H.-G., Schuierer, G., & von Eckardstein, A. (1999). Lipoprotein (a) and Genetic Polymorphisms of Clotting Factor V, Prothrombin, and Methylenetetrahydrofolate Reductase Are Risk Factors of Spontaneous Ischemic Stroke in Childhood. *Blood*, *94*(11), 3678–3682. https://doi.org/10.1182/blood.v94.11.3678

Nowak-Göttl, U., Von Kries, R., & Göbel, U. (1997). Neonatal symptomatic thromboembolism in Germany: Two year survey. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *76*(3), 163–167. https://doi.org/10.1136/fn.76.3.F163

Núñez, C., Stephan-Otto, C., Arca, G., Agut, T., Arnaez, J., Cordeiro, M., Benavente-Fernández, I., Boronat, N., Lubián-López, S. P., Valverde, E., Hortigüela, M., & García-Alix, A. (2021). Neonatal arterial stroke location is associated with outcome at 2 years: a voxel-based lesion-symptom mapping study. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, fetalneonatal-2020-320400. https://doi.org/10.1136/archdischild-2020-320400

Olivé, G., Agut, T., Echeverría-Palacio, C. M., Arca, G., & García-Alix, A. (2019). Usefulness of Cranial Ultrasound for Detecting Neonatal Middle Cerebral Artery Stroke. *Ultrasound in Medicine and Biology*, *45*(3), 885–890. https://doi.org/10.1016/j.ultrasmedbio.2018.11.004

Oosthuyse, B., Moons, L., Storkebaum, E., Beck, H., Nuyens, D., Brusselmans, K., Dorpe, J. Van, Hellings, P., Gorselink, M., Heymans, S., Theilmeier, G., Dewerchin, M., Laudenbach, V., Vermylen, P., Raat, H., Acker, T., Vleminckx, V., Bosch, L. Van Den, Cashman, N., … Carmeliet, P. (2001). Deletion of the hypoxia-response element in the vascular endothelial growth factor promoter causes motor neuron degeneration. *Nature Genetics*, *28*(2), 131–138. https://doi.org/10.1038/88842

Ossei-Gerning, N., Wilson, I., & Grant, P. (1988). Sex differences in coagulation and fibrinolysis in subjects with coronary artery disease. *Thromb Haemost*, *79*(4), 736–740.

Pabinger, S., Dander, A., Fischer, M., Snajder, R., Sperk, M., Efremova, M., Krabichler, B., Speicher, M. R., Zschocke, J., & Trajanoski, Z. (2014). A survey of tools for variant analysis of next-generation genome sequencing data. *Briefings in Bioinformatics*, *15*(2), 256–278. https://doi.org/10.1093/bib/bbs086

Pang, X., Liu, J., Zhao, J., Mao, J., Zhang, X., Feng, L., Han, C., Li, M., Wang, S., & Wu, D. (2014). Homocysteine induces the expression of C-reactive protein via NMDAr-ROS-MAPK-NF-κB signal pathway in rat vascular smooth muscle cells. *Atherosclerosis*, *236*(1), 73–81. https://doi.org/10.1016/j.atherosclerosis.2014.06.021

Papapetropoulos, A., García-Cardeña, G., Madri, J. A., & Sessa, W. C. (1997). Nitric oxide production contributes to the angiogenic properties of vascular endothelial growth factor in human endothelial cells. *Journal of Clinical Investigation*, *100*(12), 3131–3139. https://doi.org/10.1172/JCI119868

Papayianni, A., Alexopoulos, E., Giamalis, P., Gionanlis, L., Belechri, A. M., Koukoudis, P., & Memmos, D. (2002). Circulating levels of ICAM-1, VCAM-1, and MCP-1 are increased in haemodialysis patients: Association with inflammation, dyslipidaemia, and vascular events. *Nephrology Dialysis Transplantation*, *17*(3), 435–441. https://doi.org/10.1093/ndt/17.3.435

Parker, M. J., Joubert, G. I., & Levin, S. D. (2002). Portal vein thrombosis causing neonatal cerebral infarction. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *87*(2), 125–127. https://doi.org/10.1136/fn.87.2.f125

Pedro-Botet, J., Sentì, M., & Nogues, X. (1992). Lipoprotein and apolipoprotein profile in men with ischemic stroke: role of lipoprotein (a), triglyceride-rich lipoproteins and apoprotein E polymorphism. *Stroke*, *23*, 1556–1562. https://doi.org/10.1161/01.str.23.11.1556

Pellicer, A., Cabañas, F., García-Alix, A., Pérez-Higueras, A., & Quero, J. (1992). Stroke in neonates with cardiac right-to-left shunt. *Brain and Development*, *14*(6), 381–385. https://doi.org/10.1016/S0387-7604(12)80344-5

Peña-Legazkue, I., Guerrero, M., González-Pernía, J. L., & Montero, J. (2019). Informe GEM España 2019-2020. *Power*.

Peng, D., Zhao, S., & Wang, J. (1999). Lipoprotein (a) and apolipoprotein E epsilon 4 as independent risk factors for ischemic stroke. *J. Cardiovasc Risk*, *6*, 1–6. https://doi.org/10.1177/204748739900600101

Perlman, J. M., Rollins, N. K., & Evans, D. (1994). Neonatal stroke: Clinical characteristics and cerebral blood flow velocity measurements. *Pediatric Neurology*, *11*(4), 281–284. https://doi.org/10.1016/0887-8994(94)90002-7

Perucca, P., Bahlo, M., & Berkovic, S. F. (2020). The genetics of epilepsy. *Annual Review of Genomics and Human Genetics*, *21*, 205–230. https://doi.org/10.1176/ajp.103.4.457

Pezzini, A., Del Zotto, E., Archetti, S., Negrini, R., Bani, P., & Albertini, A. (2002). Plasma homocysteine concentration, C677T MTHFR genotype, and 844ins68bp CBS genotype in young adults with spontaneous cervical artery dissection and atherothrombotic stroke. *Stroke*, *33*, 664–669. https://doi.org/10.1161/hs0302.103625

Pfeiffer, D., Chen, B., Schlicht, K., Ginsbach, P., Abboud, S., Bersano, A., Bevan, S., Brandt, T., Caso, V., Debette, S., Erhart, P., Freitag-Wolf, S., Giacalone, G., Grau, A. J., Hayani, E., Jern, C., Jiménez-Conde, J., Kloss, M., Krawczak, M., … Lindgren, A. (2019). Genetic Imbalance Is Associated with Functional Outcome after Ischemic Stroke. *Stroke*, *50*(2), 298–304. https://doi.org/10.1161/STROKEAHA.118.021856

Phillips, D., Charo, I., Parise, L., & Fitzgerald, L. (1988). The platelet membrane glycoprotein IIb-IIIa complex. *Blood*, *71*, 831–843.

Pinard, A., Fiander, M. D. J., Cecchi, A. C., Rideout, A. L., Azouz, M., Fraser, S. M., McNeely, P. D., Walling, S., Novara, S. C., Hurst, A. C. E., Guo, D., Parkash, S., Bamshad, M. J., Nickerson, D. A., Vandersteen, A. M., & Milewicz, D. M. (2021). Association of De Novo RNF213 Variants With Childhood Onset Moyamoya Disease and Diffuse Occlusive Vasculopathy. *Neurology*, *96*(13), e1783–e1791. https://doi.org/10.1212/WNL.0000000000011653

Pinard, A., Guey, S., Guo, D., Cecchi, A. C., Kharas, N., Wallace, S., Regalado, E. S., Hostetler, E. M., Sharrief, A. Z., Bergametti, F., Kossorotoff, M., Hervé, D., Kraemer, M., Bamshad, M. J., Nickerson, D. A., Smith, E. R., Tournier-Lasserve, E., & Milewicz, D. M. (2020). The pleiotropy associated with de novo variants in CHD4, CNOT3, and SETD5 extends to moyamoya angiopathy. *Genetics in Medicine*, *22*(2), 427–431. https://doi.org/10.1038/s41436-019-0639-2

Pomp, E. R., Le Cessie, S., Rosendaal, F. R., & Doggen, C. J. M. (2007). Risk of venous thrombosis: Obesity and its joint effect with oral contraceptive use and prothrombotic mutations. *British Journal of Haematology*, *139*(2), 289–296. https://doi.org/10.1111/j.1365-2141.2007.06780.x

Poort, S. R., Rosendaal, F. R., Reitsma, P. H., & Bertina, R. M. (1996). A Common Genetic Variation in the 3'-Untranslated Region of the Prothrombin Gene Is Associated With Elevated Plasma Prothrombin Levels and an Increase in Venous Thrombosis. *Blood*, *88*, 3698–3703.

Pöschl, E., Schlötzer-Schrehardt, U., Brachvogel, B., Saito, K., Ninomiya, Y., & Mayer, U. (2004). Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development (Cambridge, England)*, *131*(7), 1619–1628. https://doi.org/10.1242/dev.01037

Prada, F., Mortera, C., Bartrons, J., Rissech, M., Jiménez, L., Carretero, J., Rovira, C., & Vicente, M. A. (2010). Coartación aórtica compleja y síndrome de PHACE. *Revista Espanola de Cardiologia*, *63*(11), 1367–1370. https://doi.org/10.1016/S0300-8932(10)70306-7

Prian, G. W., Wright, G. B., Rumack, C. M., & O'Meara, O. P. (1978). Apparent cerebral embolization after temporal artery catheterization. *The Journal of Pediatrics*, *93*(1), 115–118. https://doi.org/10.1016/S0022-3476(78)80619-2

Price, A. L., Kryukov, G. V, de Bakker, P. I. W., Purcell, S. M., Staples, J., Wei, L.-J., & Sunyaev, S. R. (2010). Pooled association tests for rare variants in exon-resequencing studies. *American Journal of Human Genetics*, *86*(6), 832–838. https://doi.org/10.1016/j.ajhg.2010.04.005

Prins, B. P., Lagou, V., Asselbergs, F. W., Snieder, H., & Fu, J. (2012). Genetics of coronary artery disease: genome-wide association studies and beyond. *Atherosclerosis*, *225*(1), 1–10. https://doi.org/10.1016/j.atherosclerosis.2012.05.015

Pulver, M., Juhkami, K., Loorits, D., Ilves, P., Kuld, J., Õiglane-Šlik, E., Metsvaht, T., & Laugesaar, R. (2017). Symptomatic Neonatal Arterial Ischemic Stroke With Prenatal and Postnatal Neuroimaging. *Child Neurology Open*, *4*, 2329048X1773046. https://doi.org/10.1177/2329048x17730460

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Qi, Q., & Hu, F. B. (2012). Genetics of type 2 diabetes in European populations. *Journal of Diabetes*, *4*(3), 203–212. https://doi.org/10.1111/j.1753-0407.2012.00224.x

Qin, X., Spence, J. D., Li, J., Zhang, Y., Li, Y., Sun, N., Liang, M., Song, Y., Zhang, Y., Wang, B., Cheng, X., Zhao, L., Wang, X., Xu, X., & Huo, Y. (2020). Interaction of serum Vitamin B12 and folate with MTHFR genotypes on risk of ischemic stroke. *Neurology*, *94*(11), e1126–e1136. https://doi.org/10.1212/WNL.0000000000008932

Rahbari, R., Wuster, A., Lindsay, S. J., Hardwick, R. J., Alexandrov, L. B., Al Turki, S., Dominiczak, A., Morris, A., Porteous, D., Smith, B., Stratton, M. R., & Hurles, M. E. (2016). Timing, rates and spectra of human germline mutation. *Nature Genetics*, *48*(2), 126–133. https://doi.org/10.1038/ng.3469

Raju, T. N. K., Nelson, K. B., Ferriero, D., & Lynch, J. K. (2007). Ischemic perinatal stroke: Summary of a workshop sponsored by the National Institute of Child Health and Human Development and the National Institute of Neurological Disorders and Stroke. *Pediatrics*, *120*(3), 609–616. https://doi.org/10.1542/peds.2007-0336

Ramaswamy, V., Miller, S. P., Barkovich, A. J., Partridge, J. C., & Ferriero, D. M. (2004). Perinatal stroke in term infants with neonatal encephalopathy. *Neurology*, *62*(11), 2088–2091. https://doi.org/10.1212/01.WNL.0000129909.77753.C4

Ransbotyn, V., Yeger-Lotem, E., Basha, O., Acuna, T., Verduyn, C., Gordon, M., Chalifa-Caspi, V., Hannah, M. A., & Barak, S. (2015). A combination of gene expression ranking and co-expression network analysis increases discovery rate in large-scale mutant screens for novel Arabidopsis thaliana abiotic stress genes. *Plant Biotechnology Journal*, *13*(4), 501–513. https://doi.org/10.1111/pbi.12274

Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N., Dufke, A., Cremer, K., Hempel, M., Horn, D., Hoyer, J., Joset, P., Röpke, A., Moog, U., Riess, A., … Strom, T. M. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: An exome sequencing study. *The Lancet*, *380*(9854), 1674–1682. https://doi.org/10.1016/S0140-6736(12)61480-9

Redline, R. W., Sagar, P., King, M. E., Krishnamoorthy, K. S., Grabowski, E. F., & Roberts, D. J. (2008). Case 12-2008: A newborn infant with intermittent apnea and seizures. *The New England Journal of Medicine*, *358*(16), 1713–1723. https://doi.org/10.1056/NEJMcpc0801164

Redline, R. W., Wilson-Costello, D., Borawski, E., Fanaroff, A. A., & Hack, M. (2000). The relationship between placental and other perinatal risk factors for neurologic impairment in very low birth weight children. *Pediatr Res*, *47*(6), 721–726. https://doi.org/10.1203/00006450-200006000-00007

Reiner, A. P., Frank, M. B., Schwartz, S. M., Linenberger, M. L., Longstreth, W. T., Teramura, G., Rosendaal, F. R., Psaty, B. M., & Siscovick, D. S. (2002). Coagulation factor XIII polymorphisms and the risk of myocardial infarction and ischaemic stroke in young women. *British Journal of Haematology*, *116*(2), 376–382. https://doi.org/10.1046/j.1365-2141.2002.03265.x

Remillard, G., Ethier, R., & Andermann, F. (1974). Temporal lobe epilepsy and perinatal occlusion of the posterior cerebral artery. A syn- drome analogous to infantile hemiplegia and a demonstrable etiology in some patients with temporal lobe epilepsy. *Neurology*, *24*, 1001–1009. https://doi.org/10.1212/wnl.24.11.1001

Reynolds, E. W., Riel-Romero, R. M. S., & Bada, H. S. (2007). Neonatal abstinence syndrome and cerebral infarction following maternal codeine use during pregnancy. *Clinical Pediatrics*, *46*(7), 639–645. https://doi.org/10.1177/0009922807300795

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (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. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Ridker, P. M., Hennekens, C. H., Lindpaintner, K., Stampfer, M. J., & Miletich, J. P. (1997a). Arterial and Venous Thrombosis Is Not Associated With the 4G/5G Polymorphism in the Promoter of the Plasminogen Activator Inhibitor Gene in a Large Cohort of US Men. *Circulation*, *95*(1), 59–62. https://doi.org/10.1161/01.CIR.95.1.59

Ridker, P. M., Hennekens, C. H., Schmitz, C., Stampfer, M. J., & Lindpaintner, K. (1997b). PIA1/A2 polymorphism of platelet glycoprotein IIIa and risks of myocardial infarction, stroke, and venous thrombosis. *The Lancet*, *349*(9049), 385–388. https://doi.org/10.1016/S0140-6736(97)80010-4

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kähler, A. K., Akterin, S., Bergen, S. E., Collins, A. L., Crowley, J. J., Fromer, M., Kim, Y., Lee, S. H., Magnusson, P. K., Sanchez, N., Stahl, E. A., Williams, S., Wray, N. R., Xia, K., & Bettella, P. F. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature Genetics*, *45*(10), 1150–1159. https://doi.org/10.1016/j.physbeh.2017.03.040

Risau, W., & Flamme, I. (1995). Vasculogenesis. *Annual Review of Cell and Developmental Biology*, *11*(1), 73–91. https://doi.org/10.1146/annurev.cb.11.110195.000445

Roach, J. C., Glusman, G., Smit, A. F. A., Huff, C. D., Hubley, R., Shannon, P. T., Rowen, L., Pant, K. P., Goodman, N., Bamshad, M., Shendure, J., Drmanac, R., Jorde, L. B., Hood, L., & Galas, D. J. (2010). Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing. *Science*, *328*(5978), 636–639. https://doi.org/10.1126/science.1186802

Roca, I., Fernández-Marmiesse, A., Gouveia, S., Segovia, M., & Couce, M. L. (2018). Prioritization of variants detected by next generation sequencing according to the mutation tolerance and mutational architecture of the corresponding genes. In *International Journal of Molecular Sciences* (Vol. 19, Issue 6). MDPI AG. https://doi.org/10.3390/ijms19061584

Roifman, C. M., & Chitayat, D. (2009). Combined immunodeficiency, facial dysmorphism, optic nerve atrophy, skeletal anomalies and developmental delay: a new syndrome. *Clinical Genetics*, *76*(5), 449–457. https://doi.org/10.1111/j.1399-0004.2009.01239.x

Rorke, L., & Zimmerman, R. (1992). Prematurity, postmaturity, and destructive lesions in utero. *AJNR Am J Neuroradiol*, *13*(2), 571–536.

Rothman, S. M. (2002). Stroke in children: Freud's first analysis See. *Lancet*, *360*(9345), 1526. https://doi.org/10.1016/S0140-6736(02)11575-3

Ruff, R. L., Shaw, C. -M, Beckwith, J. B., & Iozzo, R. V. (1979). Cerebral infarction complicating umbilical vein catheterization. *Annals of Neurology*, *6*(1), 85–85. https://doi.org/10.1002/ana.410060125

Ruggeri, Z. M., Bader, R., & De Marco, L. (1982). Glanzmann thrombasthenia: Deficient binding of von Willebrand factor to thrombin-stimulated platelets. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(19 I), 6038–6041. https://doi.org/10.1073/pnas.79.19.6038

Rusiñol, A. E., Thewke, D., Liu, J., Freeman, N., Panini, S. R., & Sinensky, M. S. (2004). AKT/Protein Kinase B Regulation of BCL Family Members during Oxysterol-induced Apoptosis. *Journal of Biological Chemistry*, *279*(2), 1392–1399. https://doi.org/10.1074/jbc.M308619200

Scher, M. S., Klesh, K. W., Murphy, T. F., & Guthrie, R. D. (1986). Seizures and infarction in neonates with persistent pulmonary hypertension. *Pediatric Neurology*, *2*(6), 332–339. https://doi.org/10.1016/0887-8994(86)90073-1

Scholz, B., Korn, C., Wojtarowicz, J., Mogler, C., Augustin, I., Boutros, M., Niehrs, C., & Augustin, H. G. (2016). Endothelial RSPO3 Controls Vascular Stability and Pruning through Non-canonical WNT/Ca2+/NFAT Signaling. *Developmental Cell*, *36*(1), 79–93. https://doi.org/10.1016/j.devcel.2015.12.015

Schulz, U. G. R., Flossmann, E., & Rothwell, P. M. (2004). Heritability of Ischemic Stroke in Relation to Age, Vascular Risk Factors, and Subtypes of Incident Stroke in Population-Based Studies. *Stroke*, *35*(4), 819–824. https://doi.org/10.1161/01.STR.0000121646.23955.0f

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. *Nature Methods*, *11*(4), 361–362. https://doi.org/10.1038/nmeth.2890

Selakovic, V., Raicevic, R., & Radenovic, L. (2005). The increase of neuron-specific enolase in cerebrospinal fluid and plasma as a marker of neuronal damage in patients with acute brain infarction. *Journal of Clinical Neuroscience*, *12*(5), 542–547. https://doi.org/10.1016/j.jocn.2004.07.019

Sferruzzi-Perri, A. N., López-Tello, J., Fowden, A. L., & Constancia, M. (2016). Maternal and fetal genomes interplay through phosphoinositol 3-kinase(PI3K)-p110α signaling to modify placental resource allocation. *Proceedings of the National Academy of Sciences*, *113*(40), 11255–11260. https://doi.org/10.1073/pnas.1602012113

Sharma, P., Yadav, S., & Meschia, J. F. (2013). Genetics of ischaemic stroke. *Journal of Neurology, Neurosurgery and Psychiatry*, *84*(12), 1302–1308. https://doi.org/10.1136/jnnp-2012-304834

Shichita, T., Ito, M., & Yoshimura, A. (2014). Post-ischemic inflammation regulates neural damage and protection. *Frontiers in Cellular Neuroscience*, *8*(OCT), 1–8. https://doi.org/10.3389/fncel.2014.00319

Shihab, H. A., Rogers, M. F., Gough, J., Mort, M., Cooper, D. N., Day, I. N. M., Gaunt, T. R., & Campbell, C. (2015). An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics*, *31*(10), 1536–1543. https://doi.org/10.1093/bioinformatics/btv009

Shiojima, I., & Walsh, K. (2002). Role of Akt signaling in vascular homeostasis and angiogenesis. *Circulation Research*, *90*(12), 1243–1250. https://doi.org/10.1161/01.RES.0000022200.71892.9F

Shyr, C., Tarailo-Graovac, M., Gottlieb, M., Lee, J. J., van Karnebeek, C., & Wasserman, W. W. (2014). FLAGS, frequently mutated genes in public exomes. *BMC Medical Genomics*, *7*(1), 64. https://doi.org/10.1186/s12920-014-0064-y

Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research*, *40*(W1), W452–W457. https://doi.org/10.1093/nar/gks539

Simats, A., García-Berrocoso, T., & Montaner, J. (2016). Neuroinflammatory biomarkers: From stroke diagnosis and prognosis to therapy. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, *1862*(3), 411–424. https://doi.org/10.1016/j.bbadis.2015.10.025

Sivan, Y., Nelson, M. D., Lee, S., & Wood, B. P. (1990). Radiological Cases of the Month. *Archives of Pediatrics & Adolescent Medicine*, *144*(12), 1351. https://doi.org/10.1001/archpedi.1990.02150360077025

Smith, C. D., & Baumann, R. J. (1991). Clinical Features and Magnetic Resonance Imaging in Congenital and Childhood Stroke. *Journal of Child Neurology*, *6*(3), 263–272. https://doi.org/10.1177/088307389100600311

Sogkas, G., Fedchenko, M., Dhingra, A., Jablonka, A., Schmidt, R. E., & Atschekzei, F. (2018). Primary immunodeficiency disorder caused by phosphoinositide 3-kinase δ deficiency. *The Journal of Allergy and Clinical Immunology*, *142*(5), 1650-1653.e2. https://doi.org/10.1016/j.jaci.2018.06.039

Somanath, P., Chen, J., & Byzova, T. (2008). Akt1 is necessary for the vascular maturation and angiogenesis during cutaneous wound healing. *Angiogenesis*, *11*(3), 277–288. https://doi.org/10.1007/s10456-008-9111-7.Akt1

Somanath, P. R., Razorenova, O. V, Chen, J., & Byzova, T. V. (2006). Akt1 in Endothelial Cell and Angiogenesis. *NIH Public Access*.

Somarajan, B. I., Kalita, J., Mittal, B., & Misra, U. K. (2011). Evaluation of MTHFR C677T polymorphism in ischemic and hemorrhagic stroke patients. A case-control study in a Northern Indian population. *J. Neurol. Sci*, *304*, 67–70. https://doi.org/10.1016/j.jns.2011.02.010

Soriente, L., Coppola, A., Madonna, P., Cerbone, A., Di Minno, G., & Orefice, G. (1998). Homozygous C677T mutation of the 5,10 methylenetetrahydrofolate reductase gene and hyperhomocysteinemia in Italian patients with a history of earlyonset ischemic stroke. *Stroke*, 869–871. https://doi.org/10.1161/01.str.29.4.869

Souza, D. R. S., Campos, B. F., Arruda, É. F. de, Yamamoto, L. J., Trindade, D. M., & Tognola, W. A. (2003). Influence of the polymorphism of apolipoprotein E in cerebral vascular disease. *Arquivos de Neuro-Psiquiatria*, *61*(1), 7–13. https://doi.org/10.1590/S0004-282X2003000100002

Sran, S. K., & Baumann, R. J. (1988). Outcome of Neonatal Strokes. *American Journal of Diseases of Children*, *142*(10), 1086–1088. https://doi.org/10.1001/archpedi.1988.02150100080031

Sreenan, C., Bhargava, R., & Robertson, C. M. T. (2000). Cerebral infarction in the term newborn: Clinical presentation and long-term outcome. *Journal of Pediatrics*, *137*(3), 351–355. https://doi.org/10.1067/mpd.2000.107845

Srivastava, R., Rajapakse, T., Carlson, H. L., Keess, J., Wei, X. C., & Kirton, A. (2019). Diffusion Imaging of Cerebral Diaschisis in Neonatal Arterial Ischemic Stroke. *Pediatric Neurology*, *100*, 49–54. https://doi.org/10.1016/j.pediatrneurol.2019.04.012

Stephan-Otto, C., Núñez, C., Arca, G., Agut, T., & García-Alix, A. (2017). Three-Dimensional Map of Neonatal Arterial Ischemic Stroke Distribution from Early Multimodal Brain Imaging. *Stroke*, *48*(2), 482–485. https://doi.org/10.1161/STROKEAHA.116.014186

Stojanov, D., Grozdanović, D., Petrović, S., Benedeto-Stojanov, D., Stefanović, I., Stojanović, N., & Ilić, D. N. (2014). De novo mutation in the NOTCH3 gene causing CADASIL. *Bosnian Journal of Basic Medical Sciences*, *14*(1), 48–50.

Suarez, C. R., Gonzalez, J., Menendez, C., Fareed, J., Fresco, R., & Walenga, J. (1988). Neonatal and maternal platelets: Activation at time of birth. *American Journal of Hematology*, *29*(1), 18–21. https://doi.org/10.1002/ajh.2830290105

Sun, J., Zheng, Y., & Hsu, L. (2013). A unified mixed-effects model for rare-variant association in sequencing studies. *Genetic Epidemiology*, *37*(4), 334–344. https://doi.org/10.1002/gepi.21717

Suratanee, A., & Plaimas, K. (2018). Network-based association analysis to infer new disease-gene relationships using large-scale protein interactions. *PLOS ONE*, *13*(6), e0199435. https://doi.org/10.1371/journal.pone.0199435

Takahashi, T., Yamaguchi, S., Chida, K., & Shibuya, M. (2001). A single autophosphorylation site on KDR/Flk-1 is essential for VEGF-A-dependent activation of PLC-γ and DNA synthesis in vascular endothelial cells. *EMBO Journal*, *20*(11), 2768–2778. https://doi.org/10.1093/emboj/20.11.2768

Tang, H., & Thomas, P. D. (2016). Tools for predicting the functional impact of nonsynonymous genetic variation. *Genetics*, *203*(2), 635–647. https://doi.org/10.1534/genetics.116.190033

Tesei, G., Schulze, T. K., Crehuet, R., & Lindorff-Larsen, K. (2021). Accurate model of liquid–liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. *Proceedings of the National Academy of Sciences*, *118*(44). https://doi.org/10.1073/pnas.2111696118

Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J. P., Ripatti, S., Chasman, D. I., Willer, C. J., Johansen, C. T., Fouchier, S. W., Isaacs, A., Peloso, G. M., Barbalic, M., Ricketts, S. L., Bis, J. C., Aulchenko, Y. S., Thorleifsson, G., … Kathiresan, S. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Naure*, *466*(7307), 707–713. https://doi.org/10.1038/nature09270.Biological

Thom, T., Haase, N., Rosamond, W., Howard, V. J., Rumsfeld, J., Manolio, T., Zheng, Z.-J., Flegal, K., O'Donnell, C., Kittner, S., Lloyd-Jones, D., Goff, D. C., Hong, Y., Adams, R., Friday, G., Furie, K., Gorelick, P., Kissela, B., Marler, J., … Wolf, P. (2006). Heart Disease and Stroke Statistics—2006 Update. *Circulation*, *113*(6). https://doi.org/10.1161/CIRCULATIONAHA.105.171600

Thompson, J. A., Grunnet, M. L., & Anderson, R. E. (1975). Carotid arterial elastic hyperplasia in a newborn. *Stroke*, *6*(4), 391–394. https://doi.org/10.1161/01.STR.6.4.391

Tojais, N. F., Peghaire, C., Franzl, N., Larrieu-Lahargue, F., Jaspard, B., Reynaud, A., Moreau, C., Couffinhal, T., Duplàa, C., & Dufourcq, P. (2014). Frizzled7 controls vascular permeability through theWnt-canonical pathway and cross-talk with endothelial cell junction complexes. *Cardiovascular Research*, *103*(2), 291–303. https://doi.org/10.1093/cvr/cvu133

Tournier-Lasserve, E. (2002). New players in the genetics of stroke. *The New England Journal of Medicine*, *347*(21), 1711–1712. https://doi.org/10.1056/NEJMcibr022035

Traylor, M., Farrall, M., Holliday, E. G., Sudlow, C., Hopewell, J. C., Cheng, Y.-C., Fornage, M., Ikram, M. A., Malik, R., Bevan, S., Thorsteinsdottir, U., Nalls, M. A., Longstreth, W., Wiggins, K. L., Yadav, S., Parati, E. A., Destefano, A. L., Worrall, B. B., Kittner, S. J., … Markus, H. S. (2012). Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *The Lancet. Neurology*, *11*(11), 951–962. https://doi.org/10.1016/S1474-4422(12)70234-X

Traylor, M., Markus, H., & Lewis, C. M. (2015). Homogeneous case subgroups increase power in genetic association studies. *European Journal of Human Genetics*, *23*(6), 863–869. https://doi.org/10.1038/ejhg.2014.194

Tsai, M., Kita, A., Leach, J., Rounsevell, R., Huang, J. N., Moake, J., Ware, R. E., Fletcher, D. A., & Lam, W. A. (2012). In vitro modeling of the microvascular occlusion and thrombosis that occur in hematologic diseases using microfluidic technology. *Journal of Clinical Investigation*, *122*(1), 408–418. https://doi.org/10.1172/JCI58753

Urabe, N., Naito, I., Saito, K., Yonezawa, T., Sado, Y., Yoshioka, H., Kusachi, S., Tsuji, T., Ohtsuka, A., Taguchi, T., Murakami, T., & Ninomiya, Y. (2002). Basement membrane type IV collagen molecules in the choroid plexus, pia mater and capillaries in the mouse brain. *Archives of Histology and Cytology*, *65*(2), 133–143. https://doi.org/10.1679/aohc.65.133

Van Den Broeck, J., Willie, D., & Younger, N. (2009). The World Health Organization child growth standards: Expected implications for clinical and epidemiological research. *European Journal of Pediatrics*, *168*(2), 247–251. https://doi.org/10.1007/s00431-008-0796-9

van der Knaap, M. S., Smit, L. M. E., Barkhof, F., Pijnenburg, Y. A. L., Zweegman, S., Niessen, H. W. M., Imhof, S., & Heutink, P. (2006). Neonatal porencephaly and adult stroke related to mutations in collagen IV A1. *Annals of Neurology*, *59*(3), 504–511. https://doi.org/10.1002/ana.20715

van der Put, N., Gabreëls, F., Stevens, E., Smeitink, J., Trijbels, F., Eskes, T., van den Heuvel, L., & Blom, H. (1998). A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects? *Am J Hum Genet*, *62*(5), 1044–1051. https://doi.org/10.1086/301825

Van Der Sluis, I. M., Boot, A. M., Vernooij, M., Meradji, M., & Kroon, A. A. (2006). Idiopathic infantile arterial calcification: Clinical presentation, therapy and long-term follow-up. *European Journal of Pediatrics*, *165*(9), 590–593. https://doi.org/10.1007/s00431-006-0146-8

van Dijk, E. L., Auger, H., Jaszczyszyn, Y., & Thermes, C. (2014). Ten years of next-generation sequencing technology. *Trends in Genetics*, *30*(9), 418–426. https://doi.org/10.1016/j.tig.2014.07.001

Vanhollebeke, B., Stone, O. A., Bostaille, N., Cho, C., Zhou, Y., Maquet, E., Gauquier, A., Cabochette, P., Fukuhara, S., Mochizuki, N., Nathans, J., & Stainier, D. Y. R. (2015). Tip cell-specific requirement for an atypical Gpr124- and Reck-dependent Wnt/β-catenin pathway during brain angiogenesis. *ELife*, *4*(JUNE), 1–25. https://doi.org/10.7554/eLife.06489

Veltman, J. A., & Brunner, H. G. (2012). De novo mutations in human genetic disease. In *Nature Reviews Genetics* (Vol. 13, Issue 8, pp. 565–575). https://doi.org/10.1038/nrg3241

Verdu, J., Buratovich, M. A., Wilder, E. L., & Birnbaum, M. J. (1999). Cell-autonomous regulation of cell and organ growth in Drosophila by Akt/PKB. *Nature Cell Biology*, *1*(8), 500–506. https://doi.org/10.1038/70293

Volpe, M., Iaccarino, G., Vecchione, C., Rizzoni, D., Russo, R., Rubattu, S., Condorelli, G., Ganten, U., Ganten, D., Trimarco, B., & Lindpaintner, K. (1996). Association and cosegregation of stroke with impaired endothelium- dependent vasorelaxation in stroke prone, spontaneously hypertensive rats. *Journal of Clinical Investigation*, *98*(2), 256–261. https://doi.org/10.1172/JCI118787

Wagenaar, N., Martinez-Biarge, M., Van Der Aa, N. E., Van Haastert, I. C., Groenendaal, F., Benders, M. J. N. L., Cowan, F. M., & De Vries, L. S. (2018). Neurodevelopment after perinatal arterial ischemic stroke. *Pediatrics*, *142*(3). https://doi.org/10.1542/peds.2017-4164

Wagenaar, N., van den Berk, D. J. M., Lemmers, P. M. A., van der Aa, N. E., Dudink, J., van Bel, F., Groenendaal, F., de Vries, L. S., Benders, M. J. N. L., & Alderliesten, T. (2019). Brain activity and cerebral oxygenation after perinatal arterial ischemic stroke are associated with neurodevelopment. *Stroke*, *50*(10), 2668–2676. https://doi.org/10.1161/STROKEAHA.119.025346

Wagenaar, N., van der Aa, N. E., Groenendaal, F., Verhage, C. H., Benders, M. J. N. L., & de Vries, L. S. (2017). MR imaging for accurate prediction of outcome after perinatal arterial ischemic stroke: Sooner not necessarily better. *European Journal of Paediatric Neurology*, *21*(4), 666–670. https://doi.org/10.1016/j.ejpn.2017.04.002

Wagner, K. R., Giles, W. H., Johnson, C. J., Ou, C.-Y., Bray, P. F., Goldschmidt-Clermont, P. J., Croft, J. B., Brown, V. K., Stern, B. J., Feeser, B. R., Buchholz, D. W., Earley, C. J., Macko, R. F., Mccarter, R. J., Sloan, M. A., Stolley, P. D., Wityk, R. J., Wozniak, M. A., Price, T. R., & Kittner, S. J. (1998). Platelet Glycoprotein Receptor IIIa Polymorphism P1A2 and Ischemic Stroke Risk The Stroke Prevention in Young Women Study. *Stroke*. http://stroke.ahajournals.org/

Waltenberger, J., Claesson-Welsh, L., Siegbahn, A., Shibuya, M., & Heldin, C. H. (1994). Different signal transduction properties of KDR and Flt1, two receptors for vascular endothelial growth factor. *Journal of Biological Chemistry*, *269*(43), 26988–26995. https://doi.org/10.1016/s0021-9258(18)47116-5

Wang, L., Ren, A., Tian, T., Li, N., Cao, X., Zhang, P., Jin, L., Li, Z., Shen, Y., Zhang, B., Finnell, R. H., & Lei, Y. (2019). Whole-Exome Sequencing Identifies Damaging de novo Variants in Anencephalic Cases. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.01285

Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C. T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G. D., & Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research*, *38*(suppl_2), W214–W220. https://doi.org/10.1093/nar/gkq537

Weiss, E., Bray, P., Tayback, M., Schulman, S., Kickler, T., Becker, L., Weiss, J., Gerstenblith, G., & Goldschmidt-Clermont, P. (1996). A polymorphism of a platelet glycoprotein receptor as an inherited risk factor for coronary thrombosis. *N Engl J Med*, *334*, 1090–1094. https://doi.org/10.1056/NEJM199604253341703

Westein, E., Van Der Meer, A. D., Kuijpers, M. J. E., Frimat, J. P., Van Den Berg, A., & Heemskerk, J. W. M. (2013). Atherosclerotic geometries exacerbate pathological thrombus formation poststenosis in a von Willebrand factor-dependent manner. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(4), 1357–1362. https://doi.org/10.1073/pnas.1209905110

Wilcox, E. R., Burton, Q. L., Naz, S., Riazuddin, S., Smith, T. N., Ploplis, B., Belyantseva, I., Ben-Yosef, T., Liburd, N. A., Morell, R. J., Kachar, B., Wu, D. K., Griffith, A. J., Riazuddin, S., & Friedman, T. B. (2001). Mutations in the gene encoding tight junction claudin-14 cause autosomal recessive deafness DFNB29. *Cell*, *104*(1), 165–172. https://doi.org/10.1016/s0092-8674(01)00200-8

Wraige E, Pohl KR, G. V. (2005). A proposed classification for subtypes of arterial ischaemic stroke in children. *Dev Med Child Neurol.*, *47*, 252–256. https://doi.org/10.1017/s0012162205000484

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., & Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *American Journal of Human Genetics*, *89*(1), 82–93. https://doi.org/10.1016/j.ajhg.2011.05.029

Wu, O., Robertson, L., Langhorne, P., Twaddle, S., Lowe, G. D. O., Clark, P., Greaves, M., Walker, I. D., Brenkel, I., Regan, L., & Greer, I. A. (2005). Oral contraceptives, hormone replacement therapy, thrombophilias and risk of venous thromboembolism: a systematic review. *Thrombosis and Haemostasis*, *94*(07), 17–25. https://doi.org/10.1160/th04-11-0759

Wu, Y. W., March, W. M., Croen, L. A., Grether, J. K., Escobar, G. J., & Newman, T. B. (2004). Perinatal stroke in children with motor impairment: A population-based study. *Pediatrics*, *114*(3), 612–619. https://doi.org/10.1542/peds.2004-0385

Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K., & Uversky, V. N. (2010). PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1804*(4), 996–1010. https://doi.org/10.1016/j.bbapap.2010.01.011

Yates, C. L., Monaghan, K. G., Copenheaver, D., Retterer, K., Scuffins, J., Kucera, C. R., Friedman, B., Richard, G., & Juusola, J. (2017). Whole-exome sequencing on deceased fetuses with ultrasound anomalies: Expanding our knowledge of genetic disease during fetal development. *Genetics in Medicine*, *19*(10), 1171–1178. https://doi.org/10.1038/gim.2017.31

Ye, Z., Liu, E. H. C., Higgins, J. P. T., Keavney, B. D., Lowe, G. D. O., Collins, R., & Danesh, J. (2006). Seven haemostatic gene polymorphisms in coronary disease: Meta-analysis of 66 155 cases and 91 307 controls. *Lancet*, *367*(9511), 651–658. https://doi.org/10.1016/S0140-6736(06)68263-9

Yoshida, T., Kato, K., Yokoi, K., Oguri, M., Watanabe, S., Metoki, N., Yoshida, H., Satoh, K., Aoyagi, Y., Nozawa, Y., & Yamada, Y. (2010). Association of genetic variants with ischemic stroke in Japanese individuals with or without metabolic syndrome. *International Journal of Molecular Medicine*, *25*(2), 281–286. https://doi.org/10.3892/ijmm_00000342

Yurchenco, P. D., Amenta, P. S., & Patton, B. L. (2004). Basement membrane assembly, stability and activities observed through a developmental lens. *Matrix Biology : Journal of the International Society for Matrix Biology*, *22*(7), 521–538. https://doi.org/10.1016/j.matbio.2003.10.006

Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, *31*(3), 1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015

Zachary, I. (1998). Vascular endothelial growth factor: How it transmits its signal. *Experimental Nephrology*, *6*(6), 480–487. https://doi.org/10.1159/000020560

Zhuang, G., Yu, K., Jiang, Z., Chung, A., Yao, J., Ha, C., Toy, K., Soriano, R., Haley, B., Blackwood, E., Sampath, D., Bais, C., Lill, J. R., & Ferrara, N. (2013). Phosphoproteomic analysis implicates the mTORC2-FoxO1 axis in VEGF signaling and feedback activation of receptor tyrosine kinases. *Science Signaling*, *6*(271), 1–12. https://doi.org/10.1126/scisignal.2003572]

Zilberman-Rudenko, J., Sylman, J. L., Lakshmanan, H. H. S., McCarty, O. J. T., & Maddala, J. (2017). Dynamics of Blood Flow and Thrombus Formation in a Multi-Bypass Microfluidic Ladder Network. *Cellular and Molecular Bioengineering*, *10*(1), 16–29. https://doi.org/10.1007/s12195-016-0470-7

Zivelin, A., Griffin, J. H., Xu, X., Pabinger, I., Samama, M., Conard, J., Brenner, B., Eldor, A., & Seligsohn, U. (1997). A single genetic origin for a common Caucasian risk factor for venous thrombosis. *Blood*, *89*(2), 397–402. https://doi.org/10.1182/blood.v89.2.397

Zoller, B., Svensson, P. J., He, X., & Dahlback, B. (1994). Identification of the same factor V gene mutation in 47 out of 50 thrombosis-prone families with inherited resistance to activated protein C. *Journal of Clinical Investigation*, *94*(6), 2521–2524. https://doi.org/10.1172/JCI117623

Zuk, O., Schaffner, S. F., Samocha, K., Do, R., Hechter, E., Kathiresan, S., Daly, M. J., Neale, B. M., Sunyaev, S. R., & Lander, E. S. (2014). Searching for missing heritability: Designing rare variant association studies. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(4). https://doi.org/10.1073/pnas.1322563111

SUPPLEMENTARY DATA

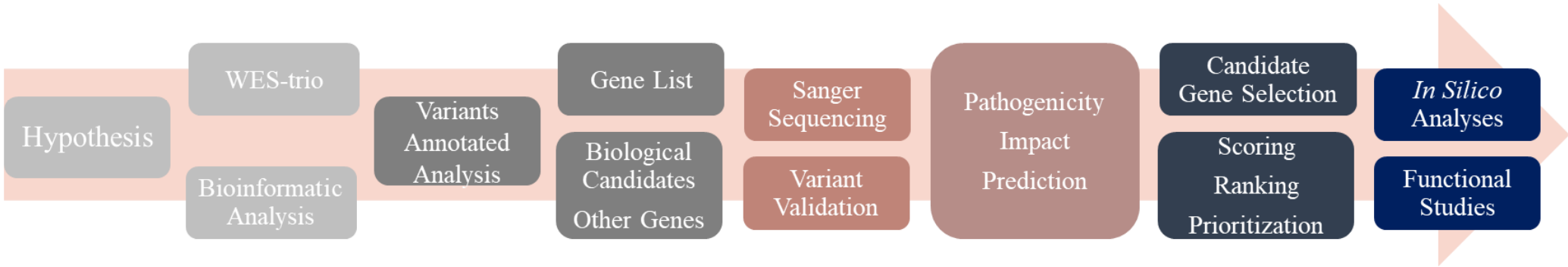


Figure S1. **Schematic showing the general filtering analysis approach applied in this study.** Custom post-WES-trio workflow for variant screening, analysis, validation, pathogenicity impact prediction, and selection of candidate gene(s) for gene-disease association according to the established hypothesis.

Table S1. Summary of primers used for validation of *de novo* variants by Sanger sequencing

Gene RefSeq	Forward Primer (5'- 3')	Reverse Primer (5'- 3')	Amplification gDNA	Product Size (bp)
AARS1 NM_001605.2	CTGCTCTGGGAAGGGGTAG	GCAGGCTGTACACACACAGG	chr16:70287095-70287396	302
AFF3 NM_001025108.*1*	TTCCTTTTTCAGTGTGGCAAT	TCTCAAGGGACCTCAAGCAT	chr2:100199151-100199473	323
CHRNA4 NM_000744.6	AAGGAGCCCTCTTCGGTGT	CATAGCAGGCTTGGGAAGAG	chr20:61980837-61981160	324
COL6A3 NM_004369.3	CCACCCTGTTCCAGTTCAAT	AAGACGTCGTTTGGCTCACT	chr2:238249178-238249670	493
DENND4A NM_001320835.1	CCACTCCTGGCCTACTTTGA	TGGTGGTTTTTAAATGGGAAG	chr15:65993235-65993634	400
DYM NM_017653.3	GCCTCTGTCTGTGCTGCTTA	CATCCTAAAAATTGGGATTGC	chr18:46784666-46784991	326
EPS15L1 NM_001258374.1	TGTTTTCCGTCACTAGCGTTC	CACCCTTACCACACAGAGCA	chr19:16503056-16503381	326
ESRRG NM_001243518.1	TTTCTGTCTCTCTTTGTATCTCCTCA	CAGTCTCTTGGGCATCGAGT	chr1:216850509-216850866	358
FAM193A NM_001256666.1	CCACTTTTGGGTGACTTGCT	GGCTGGCTCTCAGAACAAAG	chr4:2697908+2698470	563

Table S1. Continued (1)

FAM83A NM_032899.6	GGAAGCGTCTGGAAGATGTC	TTCAGGTAGGGCTTCTCAGC	chr8:124195128+124195485	358
FAR2 NM_001271783.1	GCTGGGAAGGTAGGATAGGC	ACGCACTTTCACACATTTGG	chr12:29463783+29464196	414
FASTKD5 NM_021826.4	GCCCATGGAGTTGGAGAATA	AATGCCCTTCAGAGAGCAGA	chr20:3127414-3127758	345
HECTD1 NM_015382.3	CTATGTTTATGTTTTTGATTGCCTTC	CATAAAACCTAAAGGCTGAAGAAA	chr14:31613162-31613511	350
KCNC2 NM_139137.3	TTTGCCTGGAAACACATGAA	GAGGAAGCCAAGCACATCTT	chr12:75444744-75445048	305
KRT6A NM_005554.3	CTCTCTTTCAGGTGCGGTTC	ACGGACATGGGTTGTTAGGA	chr12:52885200-52885531	332
NPC1 NM_000271.4	TCTCTCTCTTTGCGGGATTG	CTGGACGCTTGTTCCATCTT	chr18:21121113-21121364	252
NR2E1 NM_001286102.1	TCCCAGATGCAATTTCTATGC	TTAACTGGCTGGCAACAAGA	chr6:108502696+108502975	280
PCBP1 NM_006196.3	ATCGACAAGCTGGAGGAAGA	TGCTTGACACACTCGGTGAC	chr2:70315101+70315357	257
PIK3CD NM_005026.3	GGAGCCCTCAGAGGAAAGAG	TAGAGGGGTGACCAGCATCT	chr1:9779794+9780144	351

Table S1. Continued (2)

RBM7 NM_001286045.1	GGGCGTTTTCGTTTGTGA	CCTAGGCGAGACGGAAAAC	chr11:114271281+114271539	259
SHANK3 NM_001372044.1	GGCCAGGTATGAAGAGGTCA	AAGGCAGGGAGGGAGTGT	chr22:51142435+51142795	361
SLC2A10 NM_030777.3	AATCCCTCAGCCTCCTCTTC	GGTAGCTGCCACCTTCACTG	chr20:45354208+45354527	320
SNTB2 NM_006750.3	GGAAGACAGCAATGGAGACC	AAGGTACACACAAAAACGCTCA	chr16:69304061+69304410	350
SPTBN5 NM_016642.3	ACAGTGCTGAAGCCACACAG	CGTGAGGTTTGCTCAGGAGT	chr15:42147395-42147879	485
SYNE2 NM_182914.2	TGTGTTGCTTGTGTATCATGTAGG	GCTGCAACCAATCAAAAGAA	chr14:64449325+64449537	213
TOP2A NM_001067.3	CTGACACCTGGGGACCTG	AACTGCCTTTGATGAGCTTGA	chr17:38572965-38573337	373
TRABD2B NM_001194986.1	GATGGCTTGAGCCAGAGAGT	AGCACTGTGTGGCTTTTCCT	chr1:48260130-48260559	430
UNC13D NM_199242.2	TGGGCTGTGGTCACTTACTG	AGCATCCAGTGTGCATGTTG	chr17:73836766-73837160	395

Table S2. List of primers used for the generation of the *wild type* and mutant mammalian expression vectors

Primer Name	Primer Sequence (5'- 3')	Use
Fw_PIK3CD_HindIII	TAAGCAAAGCTTATGCCCCCTGGGGTGGACTG	Amplification of *PIK3CD* from pHAGE-PIK3CD and insertion of HindIII restriction enzyme target
Rv_PIK3CD_BamHI	TGCTTAGGATCCAACTGCCTGTTGTCTTTGGACACGT	Amplification of *PIK3CD* from pHAGE-PIK3CD and insertion of BamHI restriction enzyme target
Recomb_PIK3CD_Fw	GGACTCAGATCTCGAGCTCAAGCTTATGCCCCCTGGGGTGG	Generation of pEGFP-PIK3CD^{WT} with NEBuilder® HiFi DNA Assembly Master Mix
Recomb_PIK3CD_Rv	GGTGGCGACCGGTGGATCCGCCTGCCTGTTGTCTTTGGAC	Generation of pEGFP-PIK3CD^{WT} with NEBuilder® HiFi DNA Assembly Master Mix
Fw_PIK3CD_Mut_Q431R	GGCTTAAGACCGGGGAACGCTG	Nucleotide change c.1292A>G for generation of pEGFP-PIK3CD^{MUT} with Q5® Site-Directed Mutagenesis Kit
Rv_PIK3CD_Mut_Q431R	GGTCCTTGTAGTCAAACAGCAT	Amplification of pEGFP-PIK3CD^{WT} with Q5® Site-Directed Mutagenesis Kit
PIK3CD cDNA-1 Fw	ATGCCCCCTGGGGTGGAC	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-2 Rv	CACCAGCGGCTGCCGGAA	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-3 Fw	GCAGCCGGAAGACTACACG	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-4 Rv	GCTATCCGTGTTGGGGTTAC	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-5 Fw	GCCGCTGCCCTGCTCATC	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-6 Rv	GCCTCCTGCCGCATGCAC	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-7 Fw	CTACCTAGAGGCCCTCTC	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-8 Rv	CGGAACCGTTCAAATTTCTC	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-9 Fw	TATGTGCTGGGCATTGGCGA	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}
PIK3CD cDNA-10 Rv	CTGCCTGTTGTCTTTGGACA	Verification of *PIK3CD* cDNA sequence in pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{MUT}

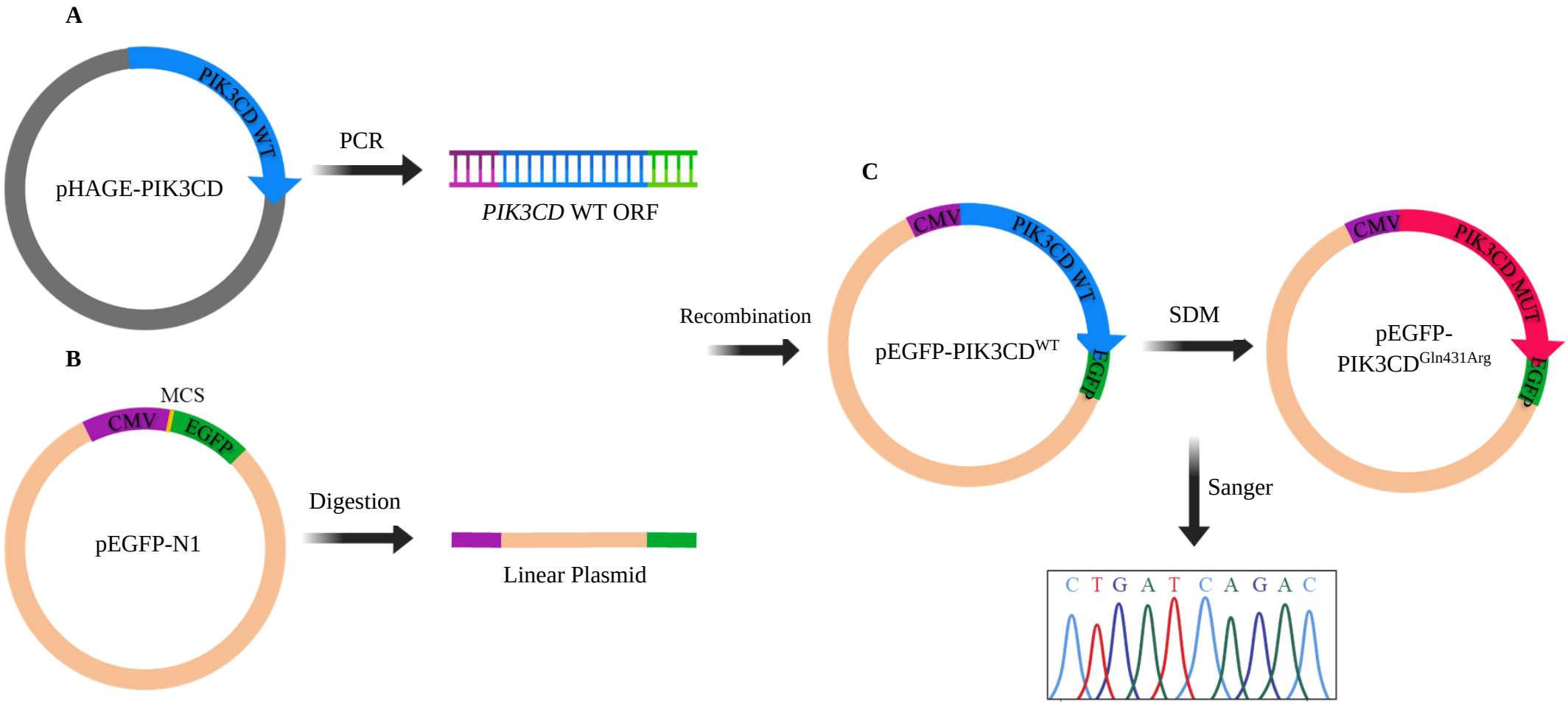


Figure S2. **Schematic representation of the PIK3CD cloning workflow**. **A**) Insert preparation using the plasmid pHAGE-PIK3CD, which contains the wild type open reading frame of PIK3CD. **B**) Preparation of the vector using the recipient plasmid pEGFP-N1 after digestion with restriction enzymes. **C**) Recombination technique for the generation of the recombinant plasmid pEGFP-PIK3CD^{WT} and subsequent Site-Directed Mutagenesis for the generation of the mutated recombinant plasmid pEGFP-PIK3CD^{Gln431Arg} carrying the Nucleotide change c.1292A>G. Both pEGFP-PIK3CD^{WT} and pEGFP-PIK3CD^{Gln431Arg} sequences were confirmed by Sanger sequencing. ORF: open reading frame. CMV: Cytomegalovirus promoter. EGFP: Enhanced Green Fluorescent Protein. MCS: Multiple Cloning Site. SDM: Site-Directed Mutagenesis. WT: Wild type.

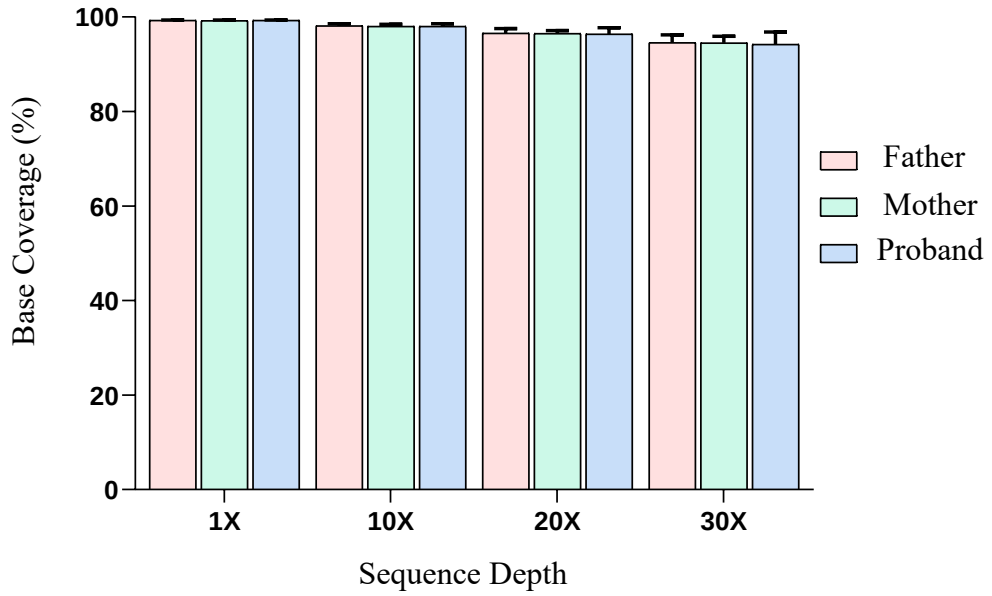


Figure S3. **Base coverage plot of 23 WES-trio**. Sequencing coverage of exonic target regions of all 23 Wes-trios (father, mother, and proband) sequenced according to the RefSeq GRCh37. Each trio shows the average base coverage totals (%) at 1X, 10X, 20X, and 30X sequencing depth.

Table S3. Biological genes included in the NAIS gene list for variant filtering

Biological process	Genes
Cell Adhesion	*AAMP, ABI3BP, ABL1, ADAM10, ADAM15, ADAM17, ADAM19, ADAM7, ADAM8, ADAMTS12, ADAMTSL4, ADGRE2, ADGRE5, ADGRL1, ADGRL3, ADRM1, AJAP1, ALCAM, AMIGO2, AMIGO3, ANOS1, ANP32A, ANTXR1, ANTXR2, AP1AR, APBA1, APBA2, APBB1IP, APC, APC2, ARHGAP12, ARHGAP26, ARHGAP9, ARMC3, ASTL, ATP2A2, AZGP1, B3GAT2, BARX2, BCAM, BOC, BVES, C1QTNF5, C2CD4A, C2CD4B, CADM1, CADM2, CADM3, CADM4, CALR, CAMSAP3, CAPNS1, CARD14, CASS4, CCDC80, CCM2L, CCN1, CCN2, CD151, CD164, CD177, CD209, CD22, CD226, CD302, CD33, CD36, CD44, CD47, CD6, CD72, CD84, CD9, CD93, CD99, CD99L2, CDH1, CDH10, CDH11, CDH12, CDH13, CDH15, CDH16, CDH18, CDH19, CDH2, CDH20, CDH22, CDH23, CDH24, CDH26, CDH3, CDH4, CDH5, CDH6, CDH7, CDH8, CDH9, CDHR1, CDHR2, CDHR3, CDHR4, CDON, CEACAM1, CEACAM16, CEACAM18, CEACAM19, CEACAM20, CEACAM21, CEACAM3, CEACAM4, CEACAM5, CEACAM6, CEACAM7, CEACAM8, CELSR1, CELSR2, CERCAM, CFDP1, CHL1, CIB1, CLDN1, CLDN14, CLDN16, CLDN19, CLDN3, CLDN6, CLEC10A, CLEC12A, CLEC14A, CLEC18A, CLEC1A, CLEC2B, CLEC4A, CLEC4C, CLEC4D, CLEC4E, CLEC4G, CLEC4M, CLEC5A, CLMP, CNN2, CNTN1, CNTN2, CNTN4, CNTN5, CNTN6, CNTNAP2, CNTNAP3B, CNTNAP4, CNTNAP5, CORO2B, CRELD1, CRISP2, CSF3R, CTNNA1, CTNNA3, CTNNB1, CTNND1, CTTN, CXADR, CXCL10, CYFIP2, DAAM1, DCC, DCHS1, DCHS2, DDR1, DDR2, DENND6A, DGCR2, DIAPH3, DKC1, DNAJB4, DOCK3, DSC1, DSC3, DSCAM, DSCAML1, DSG2, DSG3, DSG4, DUSP22, EFNB1, EGFLAM, EMCN, EMILIN2, EMP2, ENAH, EPB41L4B, EPCAM, EPDR1, EPHB6, ERG, ERMAP, ESAM, EVL, EZR, F11R, FAM107A, FAT1, FAT2, FBLIM1, FBLN5, FER, FERMT2, FERMT3, FGB, FGG, FGR, FLRT1, FLRT2, FLRT3, FMN2, FN1, FNDC3A, FREM3, FRMD5, FSCN1, FUCA2, FUT3, FXYD5, FYB2, GIT1, GIT2, GNE, GP1BB, GRB7, HAPLN1, HAPLN3, HAPLN4, HEPACAM, HMCN2, HOXD3, HS6ST2, HSPG2, ICAM1, ICAM2, ICAM5, IGDCC3, IGFBP7, IGSF11, IGSF5, IGSF9, IGSF9B, IMPG1, IQGAP1, IQGAP2, ISLR, ITGA1, ITGA10, ITGA2, ITGA2B, ITGA3, ITGA4, ITGA5, ITGA6, ITGA8, ITGA9, ITGAD, ITGAM, ITGAV, ITGB1, ITGB1BP1, ITGB2, ITGB3, ITGB4, ITGB5, ITGB6, ITGB7, IZUMO1R, JAM2, JAM3, JAML, KAZN, KIAA0319, KIRREL2, KIRREL3, KISS1, L1CAM, LAMA1, LAMA4, LAMA5, LAMB1, LAMB2, LAMC1, LAMC2, LAMC3, LAMP2, LAYN, LGALS3, LGALS8, LGTN, LIMD1, LIMS1, LIMS2, LIN7A, LMLN2, LOC102724770, LPP, LPXN, LRFN3, LRG1, LRRC66, LRRC8A, LRRN2, LRSAM1, LTBP2, LYPD3, MADCAM1, MCAM, MDGA1, MEGF10, MEGF11, MFAP4, MIA, MKX, MPZL2, MSLN, MTF2, MUC1, MYBPH, MYO7B, NAT8, NCAM1, NCAM2, NCAN, NDUFB7, NEBL, NECTIN1, NECTIN2, NECTIN3, NECTIN4, NEDD9, NEGR1, NEO1, NEXN, NFASC, NID2, NINJ1, NINJ2, NLGN4Y, PHP1, NPHS1, NPNT, NRCAM, NRXN2, NRXN3, NTM, NTN4, NUAK1, OLFM4, OLFML3, OMD, OPCML, OSBPL3, OSCAR, PALLD, PARVA, PARVB, PCDH1, PCDH12, PCDH15, PCDH19, PCDH7, PCDH8, PCDH9, PCDHA1, PCDHA10, PCDHA11, PCDHA12, PCDHA13, PCDHA2, PCDHA3, PCDHA4, PCDHA5, PCDHA6, PCDHA7, PCDHA8, PCDHA9, PCDHAC1, PCDHAC2, PCDHB1, PCDHB10, PCDHB11, PCDHB12, PCDHB13, PCDHB14, PCDHB15, PCDHB16, PCDHB2, PCDHB3, PCDHB4, PCDHB5, PCDHB6, PCDHB7, PCDHB8, PCDHB9, PCDHGA1, PCDHGA10, PCDHGA11, PCDHGA12, PCDHGA2, PCDHGA3, PCDHGA4, PCDHGA5, PCDHGA6, PCDHGA7, PCDHGA8, PCDHGA9, PCDHGB1, PCDHGB2, PCDHGB3, PCDHGB4, PCDHGB5, PCDHGB6, PCDHGB7, PCDHGC3, PCDHGC4, PCDHGC5, PDE3B, PDLIM2, PEAK1, PECAM1, PHLDB2, PKP3, PLEKHA2, PLEKHA7, PLET1, PLXNB1, PLXNC1, PNN, POSTN, PPFIA1, PPFIA3, PRKCA, PRKD1, PRPH2, PSG6, PSG9, PSTPIP1, PTK2, PTK2B, PTK7, PTPN12, PTPRA, PTPRB, PTPRF, PTPRM, PTPRT, PTPRU, PVR, PXN, RAB21, RADIL, RAP1A, RAP1B, RAP1GAP, RASSF5, RELL2, RHOJ, RHPN1, RNASE10, RNF5, RPSA, RRAS, S1PR1, SCN2B, SCYL3, SDCBP, SDK1, SDK2, SELE, SELL, SELP, SELPLG, EMA3C, SEMA4D, SERPINB8, SH3KBP1, SH3PXD2B, SHTN1, SIGLEC1, SIGLEC9, SIPA1L3, SKAP1, SMAGP, SNED1, SORBS3, SPA17, SPACA4, SPACA7, SPAM1, SPG7, SPON2, SRCIN1, ST3GAL5, ST6GALNAC2, ST8SIA1, ST8SIA2, ST8SIA4, STAB1, STAB2, STARD8, SVIL, SYNPO2, TAX1BP3, TBCEL-TECTA, TENM2, TES, TGFBI, THBS2, THY1, TIAM1, TIGIT, TM4SF19, TM4SF20, TM9SF4, TMEM102, TMEM204, TMEM8B, TNFSF4, TNN, TNR, TNS1, TNS2, TNS3, TPBG, TRIP6, TRO, TROAP, TRPM7, TSPAN16, TTN, UNC5D, VASP, VCAM1, VCAN, VIT, VTN,*

	VWC2, XIRP2, ZDHHC2, ZYX

Table S3. Continued (1)

Coagulation	*A2ML1, ABCB1, ABCC6, ABCC8, ABL1, ACD, ACE, ACTB, ACTG1, ACTN1, ACVRL1, ADAMTS13, ADAMTS2, ADH1B, ADRB1, ADRB2, ADRB3, ALDH2, ALK,, ANK1, ANKRD26, ANO6, AP3B1, APC, APOE, ARHGAP31, ASPA, ASXL1, ATM, ATP7B, BBS2, BCHE, BCKDHA, BCKDHB, BLM, BLOC1S3, BLOC1S6, BMS1, BPGM, BRAF, CALR, CBL, CBS, CCR5, CD36, CDC42, CEBPA, CFTR, CLRN1, COL1A1, COL1A2, COL4A3, COMT, CPT2, CSF3R, CYCS, CYP1A2, CYP1B1, CYP2A6, CYP2B6, CYP2C19, CYP2C9, CYP2D6, CYP3A4, CYP3A5, DHCR7, DHDDS, DICER1, DLD, DLL4, DNMT3A, DOCK6, DPYD, DTNBP1, EGLN1, ELP1, ENG, ENPP1, EOGT, EPAS1, EPB41, EPB42, EPCAM, EPOR, ETV6, EZH2, F10, F11, F12, F13A1, F13B, F2, F5, F7, F8, F9, FAH, FANCC, FERMT3, FGA, FGB, FGG, FKTNFLG, FLI1, FLNA, FLT3, FMR1, G6PC, G6PD, GALT, GATA1, GATA2, GBA, GDF2, GFI1B, GGCX, GP1BA, GP1BB, GP6, GP9, HABP2, HBA1, HBA2, HBB, HEXA, HIF1A, HOXA11, HPS1, HPS3, HPS4, HPS4, HPS5, HPS6, HRAS, HRG, HTR2A, IDH1, IDH2, ITGA2, ITGA2B, ITGB3, JAK2, KAT6B, KCNN4, KIT, KLKB1, KNG1, KRAS, LARS1, LMAN1, LYST, LZTR1, MAP2K1, MAP2K2 MASTL, MCFD2, MCOLN1, MEN1, MLH1, MPL, MRAS, MSH2, MSH6, MTHFR, MTR, MTRR, MTTP, MYH9, NAT2, NBEA, NBEAL2, NEB, NF1, NF2, NOTCH1, NPM1, NRAS, NPRM1, P2RX1, P2RY12, PARN, PCDH15, PEX2, PHF6, PHGDH, PHOX2B, PIEZO1, PIGM, PKHD1, PKLR, PLA2G4A, PLAT, PLAU, ,PLG, PMM2, PMS2, PON1, POT1, PPP1CB, PRKAR1A, PROC, PROCR, PROS1, PTCH1, PTEN, PTPN11, ,RAF1,RASA1, RASA2, RBM8A, RBPJ, RHAG, RIT1, RPSA, RRAS, RTEL1, RUNX1, SAR1B SERPINA10, SERPINC1, SERPIND1, SERPINE1, SERPINF2, SETBP1, SF3B1, SH2B3, SHOC2, SLC19A1, SLC35A3, SLC4A1, SLCO1B1, SLFN14, SMAD4, SMARCA4, SMARCB1, SMARCE1, SMN1, SMPD1, SOD2, SOS1, SOS2, SPRED1, SPTA1, SPTB, SRC, SRSF2, STXBP1, SUFU, SUMF1, TBXA2R, TBXAS1, TET2, TET3, THBD, THPO, TMEM216, TP53, TPMT, TSC1, TSC2, TUBB1, U2AF1, UGT1A1, VHL, VKORC1, VWF, WAS, WT1*
Platelet Activation	*ACTB, ADCY1, AKT3, APBB1IP, ARHGAP35, ARHGEF12, BTK, COL1A1, F2, F2R, FCER1G, FCGR2A FERMT3, FGA, GNA13, GNAI1, GNAQ, GNAS, GP1BA, GP1BB, GP5, GP6, GP9, GUCY1A2, ITGA2, ITGA2B, ITPR1, LCP2, LYN, MAPK1, MAPK14, MYL12B, MYLK4, NOS3, ORAI1, P2RX1, P2RY1, P2RY12, PIK3CA, PIK3R6, PLA2G4B, PLCB1, PLCG2, PPP1R12A, PRKACA, PRKCI, PRKG1, PTGIR, PTGS1, RAP1A, RASGRP1, RHOA, ROCK1, SNAP23, SRC, STIM1, SYK, SYK, TBXA2R, TBXAS1, TLN1, VAMP8, VASP, VWF*
Inflammation	*ABCF1, ABHD12, ACKR2, ADAM32, ADAMTS10, ADGRE5, AGER, AIMP1, ALOX15, ANGPT2, ANXA5, AXL, BDKRB1, BDKRB2, BGN, BRINP3, C15orf48, C1R, C3, C4A, C4B, C5, CABS1, CASP1, CCL13, CCL27, CCL7, CCL8, CCR1, CCR10, CCR7, CCR9, CCRL2, CD163, CD274, CDH26, CDX2, CERK, CHI3L1, CKLF, CLC, CLEC10A, CLEC12A, CLEC14A, CLEC1A, CLEC2B, CLEC4A, CLEC4C, CLEC4D, CLEC4E, CLEC5A, CSF2, CST9, CTSG, CXCL1, CXCL12, CXCL14, CXCL2, CXCL3, CXCL8, CXCR2, CYP4F2, CYP4F3, DCN, EFB4A, DEFB4B, ENPP7,EREG, ERG, F2RL3, FBXL19, FCSK, FETUB, FPR1, FUT3, GBP5, GDF15, GGCX, GPR32, GPR84, GSK3B, GZMB, HMGB1, HPGD, HRH4, HS6ST3, HSP90AB1, IKBKG, IL10, IL17RB, IL1RN, IL4, IL6, ILRUN, ITGA1, JAK1, JAK2, KLF2, KLKB1, KNG1, LTB4R, MEP1B, MGST1, MGST2, MGST3, MIF, MMP25, MROCKI, NFE2L2, NFKB2, NLRP10, NLRP12, NLRP3, NOD1, NR1H2, NR1H3, NSMAF, OPTN, ORM1, ORM2, OTOP1, OTULIN, OXGR1, P2RY2, PADI4, PAPPA, PLA2G2D, PPARA, PROKR1, PTGDR, PTGDR2, PTGS2, PTX3, REG3A, REL, RIPK1, SAA1, SAA2, SCGB1A1, SELE, SELENOS, SELL, SELPLG, SERPINB4, SERPIND1, SERPINI2, SIRT6, SLC39A8, TECPR2, TNFAIP6, TNFAIP8L2, TNFRSF1A, TNFRSF21, TREM2, TREML1, TRIM40, TSLP, TYROBP, UBD, VEGFA*
Stroke	*ABCA1, ABCC8, ACE, ACTA2, ACVR1, ADA2, ADAMTS13, ADIPOQ, ADORA2A, AGT, AGTR1, AKT1, ALOX5AP, ANGPT1, ANGPT2, ANKRD1, APOA1, APOB, APOE, APOLD1, APP, AQP1, AQP4, AR, ARAP3, ARHGEF15, ARL6IP6, ASB10, ASIC1, ASIC3, ATOH7, ATP2C1, ATP7B, BAK1, BCL2L11, BEST1, BMX, C3, C3AR1, CA4, CANT1, CBS, CCR1, CD109, CD4, CDH5, CDK5, CFAP20, CFH, CLDN5, CLU, COL18A1, COL3A1, COL4A1, COL4A2, COL8A1, COL8A2, CRP, CRPPA, CRYAA, CST3, CTNNA3, CTSL, CXCL12, CYP1B1, CYP2C19, CYP4F2, DAPK1, DDIT4, DJ1, DLG4, DLL4, EDN1, EGFL7, EGLN1, ELN, EMCN, ENPP1, EPHX2, EPO, ESAM, ESM1, ESR1, F10, F11, F12, F13A1, F13B, F2, F5, F7, F8, F9, FABP2, FGA, FGB, FGD5, FGF2, FGG, FOXC1, G6PD, GABRA5, GABRD, GAS6, GGCX, GH1,*

	GHR, GLDC, GLYCTK, GP1BA, GP6, GPR17, GPR182, GRIN1, GRIN2A, GRIN2B, GRIN3A, GUCY1A3, HABP2, HAVCR1, HBB, HIF1A, HMGCR,
Table S3. Continued (2)	
Stroke	*HRG, HSF1, HTRA1, HYOU1, IL18, IL1RL1, IL20, IL4, IL6, INSIG1, ITGA2, ITGB3, ITM2B, KCNE2, KCNK2, KCNMB1, KDR, KLKB1, KNG1, KRIT1, LDLR, LGMN, LIPG, LMAN1, LMNA, LMX1B, LOXL1, LPA, LPL, LTA, LTBP2, LYL1, MAPK10, MAPKAPK2, MBL2, MCFD2, MFRP, MIF, MMP2, MMP3, MMP9, MMRN1, MPL, MTHFR, MYBPC3, MYOC, NDNF, NFKB1, NGB, NOS1, NOS3, NOTCH1, NOTCH3, NPPB, NRROS, NTF4, OLFM2, OPA1, OPTC, OPTN, P2RX5, PARK2, PAX6, PCDH1, PCSK9, PDE4D, PDGFC, PDGFRA, PGF, PITX2, PITX3, PLA2G7, PLAT, PLG, PLVAP, POMT1, PPARGC1A, PRKCE, PRKCH, PROC, PROCR, PROS1, PROZ, PRSS56, PTGER2, PTGS1, PTPRB, RAPGEF3, RASIP1, RHOJ, RNF213, RPGRIP1, RPS19, RRM2B, RYR1, SBF2, SCG2, SELE, SELENOS, SELP, SERPINC1, SERPIND1, SERPINE1, SERPINF2, SH3PXD2B, SHMT2, SIRT2, SLC1A2, SLC4A4, SLC6A11, SOX17, SPP1, SREBF1, STC1, TEK, TET2, THBD, THBS1, TIE1, TIMD4, TLR4, TMEM204, TNF, TNFRSF1A, TNFRSF1B, TRIM2, TTR, UCP2, VCAM1, VEGFA, VKORC1, VSX2, VWF, WDR36*
Endothelium	*AAMP, ACKR2, ADGRA2, ADGRG6, AGGF1, AIMP1, AKAP12, AMOT, AMOTL1, AMOTL2, ANG, ANGPT1, ANGPT2, ANGPT4, ANGPTL1, ANGPTL2, ANGPTL4, ANKRD1, ANTXR1, ANXA13, AP000866.3, APOBR, APOL3, APOLD1, ARHGEF17, ATG9B, ATOH8, B9D1, BCAR1, BMPER, BMX, CADM4, CCM2, CCN1, CCN2, CCR10, CCR4, CD109, CD151, CD177, CD200, CD226, CD248, CD48, CDC42EP1, CDH13, CDH5, CERCAM, CHST4, CLDN1, CLDN10, CLDN11, CLDN12, CLDN14, CLDN15, CLDN16, CLDN17, CLDN18, CLDN20, CLDN22, CLDN23, CLDN24, CLDN25, CLDN3, CLDN5, CLDN6, CLDN7, CLDN8, CLDN9, CLEC4G, CLEC4M, CLIC5, CLMP, COL8A1, COL8A2, COLEC12, CRE3, CSPG4, CXCL11, CXCL8, CXCR2, DCBLD2, DNAJA4, DNAJB9, ECSCR, EDF1, EDIL3, EGFL7, EGFL8, EGLN3, EGR3, EMC10, EMP2, EOLA1, EPAS1, ESAM, ESM1, EXOC3L2, F11R, FBLN5, FGF1, FGF7, FLNA, FLT1, FLT4, FLVCR2, FOXJ2, GATA2, GATA3, GLG1, GP1BB, GPER1, GPIHBP1, GRHL3, HECW2, HERC5, HSP90B1, HSPG2, HYI, ICAM1, IGSF11, IL17D, IL17F, IL33, ITLN1, ITLN2, JAM2, JAM3, JCAD, KDR, KLF4, KNDC1, KRIT1, LCP1, LETR1, LIPG, LOC105375809, LOC109609705, LOC117600004, LOC121132685, LPAR1, LPAR2, LPAR3, LYVE1, MADCAM1, MALL, MAP3K6, MCAM, MED28, MMRN1, MT1G, MYADM, NECTIN1, NOS3, NOSIP, NOSTRIN, NRP1, NRP2, NTN4, NUCB2, OSM, OVOL2, P2RX6, PCDH12, PDGFA, PDGFB, PEAR1, PECAM1, PGF, PHACTR1, PIANP, PIR-FIGF, PLIN2, PLPP3, PLS3, PLVAP, PLXDC1, PLXDC2, PRKCB, PROCR, PROK1, PTGS1, PTPRB, PTX3, RAP1GAP2, RAPGEF3, RAPGEF6, RASD2, RCAN2, RCN1, REM1, RGS5, RHOJ, ROBO4, S1PR1, S1PR2, S1PR3, S1PR4, S1PR5, SASH1, SCARF1, SCARF2, SCO2, SCUBE1, SELE, SELL, SELP, SEMA3F, SEMA3G, SEMA4A, SENCR, SERPINE1, SETSIP, SHB, SHROOM2, SLC4A11, SLC7A14, SMOC2, SPARCL1, SPRY2, STAB1, STAB2, SVBP, SYNJ2BP, TCF4, TEK, THBD, THSD7A, TIE1, TIMP2, TJP1, TJP2, TNFAIP1, TNFAIP2, TNFSF12, TNFSF15, TNFSF18, TNFSF4, TNIP2, TNMD, TNS3, TRARG1, TRPC4, TXNDC5, TYMP, VASH1, VCAM1, VEGFA, VEGFB, VEGFC, VEGFD, VEZF1, VEZF1P1, ZEB1, ZNF124, ZNF444, ZNF580, ZNF71*