INVESTIGATING THE GENOMIC CONTRIBUTIONS TO FAMILIAL INTRACRANIAL ANEURYSMS IN A FIRST NATION FROM NORTHERN BRITISH COLUMBIA

by

Avery Newman-Simmons

B.Sc., Laurentian University, 2019

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis entitled:

Investigating the Genomic Contributions to Familial Intracranial Aneurysms in a First Nation from Northern British Columbia

submitted by Avery Newman-Simmons in partial fulfillment of the requirements for

the degree of Master of Science

in Medical Genetics

Examining Committee:

Dr. Laura Arbour, Professor, Department of Medical Genetics, UBC
Co-supervisor

Dr. Anna Lehman, Associate Professor, Department of Medical Genetics, UBC
Co-supervisor

Dr. William Gibson, Professor, Department of Medical Genetic, UBC
Supervisory Committee Member

Dr. Stefan Taubert, Associate Professor, Department of Medical Genetics, UBC
Additional Examiner

Dr. Alison Elliott, Associate Professor, Department of Medical Genetic, UBC
Additional Examiner

Additional Supervisory Committee Members:

Dr. Nadine Caron, Professor, Department of Surgery, UBC
Supervisory Committee Member

Abstract

INTRODUCTION: Familial intracranial aneurysms (FIA) may present with Mendelian inheritance and an increased risk for first-degree relatives to develop intracranial aneurysms, at a younger age, and with a higher risk of them rupturing in comparison to their sporadic counterparts. However, since genetic research on this non-syndromic condition began in the early 2000's only a few disease-gene associations have been discovered and much of the aetiology is still missing. Most of these genetic studies have focused primarily on European and East Asian populations although IA has been well documented in Canadian and Greenlandic Inuit. To our knowledge, we are presenting the first extensive whole genome sequencing study on FIA in several First Nation families from a community in Northwestern British Columbia, Canada.

METHODS: Whole genome sequencing was completed for 6 affected individuals, selected for having as distant a relationship as possible. To identify single nucleotide variations, small indels, mitochondrial variations, and structural variations that could cause FIA, various filtration strategies which included read depth and genotype quality controls, gnomAD minor allele frequencies, sequence ontology, and allele sharing between five or more participants were used. Filtered variants were subsequently prioritized based on extensive annotation information and biological significance thorough a literature search.

RESULTS: We found 25 single nucleotide variations that passed the initial filtration strategy. After prioritization, seven top variant candidates in the *HEMK1*, *CPT1A*, *LOC105371356*, *TUSC3*, *PLCB3*, *DKK3*, and *KRT8* genes remained. Extensive annotation and literature search results demonstrated that the *PLCB3* p.R874Q missense variant was the strongest among the seven top variant candidates. However, without a population specific variant

database to permit identification and removal of common population-specific variants no further prioritization could be completed.

CONCLUSION: From available data, rare variants were found in this family, but without adequate Indigenous representation in variant databases, interpretation accuracy is limited. Future studies of this family should utilize linkage to localize a single region if possible and also long-read sequencing potentially in order to ensure all variants in a linked region have been identified. The present study has largely ruled out currently known Mendelian disorders associated with intracranial aneurysms.

Lay Summary

Intracranial aneurysms are a common complex condition that occur when the cerebral arteries weaken and expand, which may lead to devastating results if they rupture. Familial intracranial aneurysms describes the occurrence of intracranial aneurysms is multiple family members, suggesting a genetic origin. However, after many years of research our genetic understanding of familial intracranial aneurysms is limited. In this study, we enrolled 6 First Nation individuals with familial intracranial aneurysms from the same remote community and completed whole genome sequencing to investigate the genetic contributions. To our knowledge, this is the first study to extensively describe and investigate the genetic contributions in a First Nation and adds to the current literature. We identified several candidate variants that were shared between most, if not all, of our participants. However, we found that without proper representation of Indigenous people in variant databases a confident and accurate interpretation of genetic variants is limited.

Preface

Through the UBC Community Genetics Research Program, familial intracranial aneurysms in the Gitxsan community was identified as a priority for research by the community in multiple engagement events. During a community engagement event in the spring of 2019, it was agreed upon to use pre-existing partnerships (Gitxsan Health Society Board and the Gitxsan Research Advisory Committee) to research the genetics of familial intracranial aneurysms through the newly established Silent Genomes project.

This study was conducted through Activity 2: *Precision Diagnosis for Indigenous Families with Genetic Conditions* of the Silent Genomes project which obtained harmonized ethics approval from the University of British Columbia Children's and Women's Research Ethics Board and Vancouver Island Health (REB Number: H18-00726). All participant data provided in this thesis is de-identified

The majority of the participants of this study were previously enrolled in the on-going Long QT Syndrome Study, except for 1 participant who came forward at a community information event and enrolled only this study. I selected potential participants from the Long QT Syndrome Study, with a reported intracranial aneurysm and who were consented for re-contact. Sarah McIntosh, a study genetic counselor with a long-standing relationship with the community assisted in re-contact and consented participants to this study. All participants provided informed consent through in-person or over-the-phone meetings.

Two samples were collected by transferring existing samples in storage from the Long QT Syndrome Study. The remaining samples were collected by having participants provide saliva or buccal samples or have blood samples collected through LifeLabs. Whole genome sequencing with completed by the Michael Smith Genome Sciences Centre.

Variant calling was completed by a genomic analyst, Jill Mwenifumbo or Arezoo Mohajeri, team members of the Silent Genome project. Variant annotations not found in VarSeq were curated by Jill Mwenifumbo, except for the gene candidate list which I curated. I, with the help of Jill Mwenifumbo, created the filters for single nucleotide, small indel, and mitochondrial variants. I completed the filtration of single nucleotide, small indels, mitochondrial, and structural variants. I completed the allele sharing analysis for all listed variants and subsequent candidate variant prioritization. Analysis of the remaining top candidate variants was completed by me.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAA – Abdominal aortic aneurysms

ACA – Anterior cerebral artery

ACoA – Anterior communicating artery

ACMG – American College of Medical Genetics

ACTA2 – Actin alpha 2

ADAMTS15 – ADAM metallopeptidase with thrombospondin type 1 motif 15

ADPKD – Autosomal dominant polycystic kidney disease

ANK – Ankyrin

ANGPTL6 – Angiopoietin like 6

ARHGEF – Rho Guanine Nucleotide Exchange Factor

C4orf6 – Chromosome 4 open read frame 6

CADD – Combined Annotation Dependent Depletion

CBPR – Community based participatory research

CCM2 – Cavernous malformation

CNV – Copy number variant

COL3A1 – Collagen type III alpha 1 chain

COVID19 – Coronavirus disease 2019

CPT – carnitine palmitoyltransferase

CT – Computerized tomography

CTA – Computerized tomography angiography

DECIPHER – DatabasE of genomiC varIation and Phenotype in Humans using Ensemble Resources

DKK3 – Dickkopf WNT signaling pathway inhibitor 3

DNA – Deoxyribonucleic acid

DNAH – Dynein axonemal heavy chain

DSA – Digital subtraction angiography

FATHMM – Functional Analysis Through Hidden Markov Models

FBN1 – Fibrillin 1

FDR – First-degree relatives

FHIT – Fragile histidine triad diadenosine triphosphatase

FIA – Familial intracranial aneurysm

FTAAD – Familial thoracic aortic aneurysm and dissection

GATK – Genome Analysis Toolkit

GERP++ - Genomic Evolutionary Rate Profiling++

gnomAD – Genome Aggregation Database

GTEx – Genotype-Tissue Expression

GWA – Genome-wide association

FIA – Familial intracranial aneurysm

HEMK1 – Hemk methyltransferase family member 1

HCM – Hypertrophic cardiomyopathy

HPO – Human Phenotype Ontology

HUVEC – Human umbilical vascular endothelial cell

IA – Intracranial aneurysm

ICA – Internal carotid artery

Indel – Insertion/deletion

KCNQ1 – Potassium voltage-gated channel subfamily Q member 1

KL - Klotho

KRT8 – Keratin 8

LDL-C – Low-density lipoprotein cholesterol

LOC – Long non-coding RNA

LOXL2 – Lysyl Oxidase Like 2

LQTS – Long QT syndrome

MAF – Minor allele frequency

MCA – Middle cerebral artery

MO – Morpholino oligomer

MOPD2 – Microcephalic osteodysplastic primordial dwarfism type 2

MRA – Magnetic resonance angiogram

MRI – Magnetic resonance imaging

mRNA – Messenger ribonucleic acid

MTRF1 – Mitochondrial translational release factor 1

NCBI – National Center for Biotechnology Information

NF1 – Neurofibromin 1

NFX1 – Nuclear transcription factor, x-box binding 1

NGS – Next-generation sequencing

OMIM – Online Mendelian Inheritance in Man

PCoA – Posterior communicating artery

PCNT - Pericentrin

PCR – Polymerase chain reaction

PKD – Polycystin

PLCB3 – Phospholipase C beta 3

RNF213 – Ring finger protein 213

ROS – Reactive oxygen species

SAH – Subarachnoid hemorrhage

SIFT – Sorting intolerant from tolerant

siRNA – Small interfering ribonucleic acid

SMAD3 – SMAD family member 3

SNP – Single nucleotide polymorphism

SNV – Single nucleotide variant

SPDYE1 – Speedy/RINGO cell cycle regulator family member 1

STARD13 – Star related lipid transfer domain containing 13

TAA – Thoracic aortic aneurysm

TGFB – Transforming growth factor beta

TGFBR – Transforming growth factor beta receptor

THSD1 – Thrombospondin type 1 domain containing 1

TMEM132B – Transmembrane protein 132B

TUSC3 – Tumor suppressor candidate 3

UBC – University of British Columbia

UTR – Untranslated region

VSMC – Vascular smooth muscle cells

WES – Whole exome sequencing

WGS – Whole genome sequencing

Acknowledgements

I would like to start by giving my greatest thanks to my co-supervisors Dr. Laura Arbour and Dr. Anna Lehman for all of guidance they provided throughout my graduate studies. Without their continuous encouragement, pushing me forward with my head held high, I would have not been able to finish my work. I would also like to thank my committee members Dr. Nadine Caron and Dr. William Gibson for their support when I was quickly finishing this thesis.

I would also like to extend my gratitude to the Sarah McIntosh, Jill Mwenifumbo, Arezoo Mohajeri, Laurie Montour, Dr. Leslie Alfaro and all the Silent Genomes project and Community Genetics Research program team members for of their help and work they put towards this project. These people provided a strong support and information network for me to rely on.

I would like to acknowledge funding provided by The University of British Columbia and Dofasco Inc. through the UBC Faculty of Graduate and Postdoctoral Studies Aboriginal Graduate Fellowship. Their generosity made living in British Columbia easier, and I am truly grateful for their help.

I would also like to thank and show appreciation for the Gitxsan community, the families, and the participants for bring this topic to our attention and participating in this study and others before. Their knowledge of their community and their support made this research project possible.

Finally, I would like to thank my entire family for all their love and support while living so far away from home. Their frequent phone calls and messages kept me grounded and pushed me forward.

Chapter 1: Introduction

1.1 Indigenous People

1.1.1 Indigenous Peoples and Health in Canada

Indigenous peoples in Canada are divided into 3 distinctive groups: First Nations, Métis, and Inuit. From the 2016 Statistics Canada Census,[1] they account for 4.9% of the total population. There are over 1.6 million Indigenous peoples across Canada,[1] but they experience health disparities when compared with non-Indigenous peoples including shorter life expectancies,[2] higher rates of chronic conditions,[3] and infant mortality.[4,5] These disparities are a linked to socioeconomic inequalities caused by colonization, systemic racism, and cultural oppression.[6] Efforts have been successful in improving some measurements of Indigenous health, but they are improving at a slower rate than what is seen in non-Indigenous people causing the health disparity gap to widen further.[7]

1.1.2 Medical Genetics Research and Indigenous Peoples

Medical genetics research has the potential to greatly benefit the health of populations. Currently, databases such as the Genome Aggregation Database (gnomAD) are commonly used by clinicians for genome reference information important in genomic analyses for patients with possible genetic conditions or diseases.[8] However, Indigenous populations are starkly underrepresented in such databases possibly leading to misinterpretations of genomic information.[8,9] The Silent Genomes Project[10] is currently underway with 4 distinct activities and many goals set in order to decrease the significant genomic information gap that exists. Activity 2 aims to address the lack of access to genetic services for families with genetic conditions. The information collected and engagement through the project and its activities will increase access

to genomic health care for Indigenous peoples in Canada, challenges with distance, increase variant interpretation accuracy, and ensure that participating Indigenous communities are equipped to participate in future genetic research fully and ethically.

Even though medical genetics research has the potential to do great good for Indigenous people, it can also be harmful. An often discussed negative experience with medical genetic research occurred when a geneticist from the University of British Columbia (UBC) collected over 800 blood samples from the Nuu-chah-nulth First Nation to investigate rheumatoid arthritis.[11] However, this researcher allowed additional studies to be conducted on migration[12] and retroviruses[13] without the knowledge of the Nuu-chah-nulth people.[14] This example of sample misuse supported the necessity and creation of the "DNA-on-Loan" concept, which states that any biological samples obtained, and data derived, are 'on-loan' to the researcher and are considered to be the property of the participant or community from which it was taken.[15] This methodology increases the autonomy of the patient and the communities, ensures trust and respect, and begins to equalize the power differential that exists between communities and the researcher.

1.1.3 The Gitxsan Community

Living in their territory since time immemorial, in what is now known as Northwestern British Columia,[16] most of the Gitxsan people currently reside in two provincial municipalities, Hazelton and New Hazelton and on 5 reserves: Gitanmaax, Gitsegukla, Glen Vowell, Kispoiox, and Kitwanga (Figure 1.1).[17,18]

Figure 1.1 - Gitxsan Territory

This figure depicts Gitxsan territory, highlighted in yellow, from the map of Canada and British Columbia.[17,18]

The Gitxsan community has participated in long QT syndrome research since 2003 with the UBC Community Genetics Research Program.[19] It was through on-going community engagement that the priority of assessing a possible genetic cause for intracranial aneurysms arose. Familial intracranial aneurysms were recognized broadly, and at least two multigenerational kindreds with what appeared to be an autosomal dominant pattern with reduced penetrance became apparent.

Affected persons and family members came forward expressing their concern, and suggested research be carried out. Community health leaders relayed that IA were a priority for exploration, and at several community meetings the topic came up. More than 20 persons enrolled in the long QT study reported they were affected with IA. This was especially notable in

the kindreds found to have the *ANK2* p.S646F variant contributing to AnkB-syndrome.[20] In spring 2019, at a community gathering, researchers from the Community Genetics Research Program presented the possibility that a research project could be carried out as part of the Silent Genomes Project. This research was welcomed, and a plan was set. The community has been updated as to the progress in their annual Community Genetics Research Program update newsletter, however a 2020 face to face meeting was postponed because of COVID-19. To maintain engagement and relationships with the Gitxsan community, in 2021 a virtual symposium was held to update the community of our progress in the Silent Genomes Project and the UBC Community Genetics Research Program, which included a presentation on this thesis (See Appendix C for details).

1.1.4 Congenital Long QT Syndrome

Long QT syndrome is characterised by a prolonged QT interval and arrhythmias causing syncope and in some cases sudden death.[21] Showing incomplete penetrance and variable expressivity, long QT syndrome is relatively rare occurring in approximately 1:2,000;[22] however in the Gitxsan community, the prevalence is much higher (~1:125).[23] Through the UBC Community Genetics Research Program, variants *KCNQ1* p.V205M[19] and *ANK2* p.S646F[20] have been shown to cause long QT syndrome in some community members. Although familial intracranial aneurysms had been long noted in this community, Swayne *et al*. (2017)[20] reported 2 unrelated Gitxsan families sharing the *ANK2* p.S646F variant also having intracranial aneurysms, but stated that it was too early to postulate a connection between the p.S646F variant and intracranial aneurysms.

1.2 Intracranial Aneurysm Background Information

Intracranial aneurysms (IA) occur due to weakening of the arterial walls in the cerebrovasculature resulting in bulging or a protrusion of the region.[24] In the general population, IA are predicted to occur in in 2-5%[24–26] of the population with an increasing prevalence after the age-of-30.[26] Due to their prevalence within the general populations, IA are considered a common complex condition that has both environmental and genetic contributors and are classified by their saccular or cylindrical shapes (Figure 1.2).[27,28] Saccular aneurysms account for most of all IA found, and are berry-like protrusions from cerebral arteries.[24,29–31] Fusiform aneurysms, known for their cylindrical shapes, are dilations along a small section of the vessel and account for 3-13% of IA.[27]

Figure 1.2 - Intracranial Aneurysm Shapes

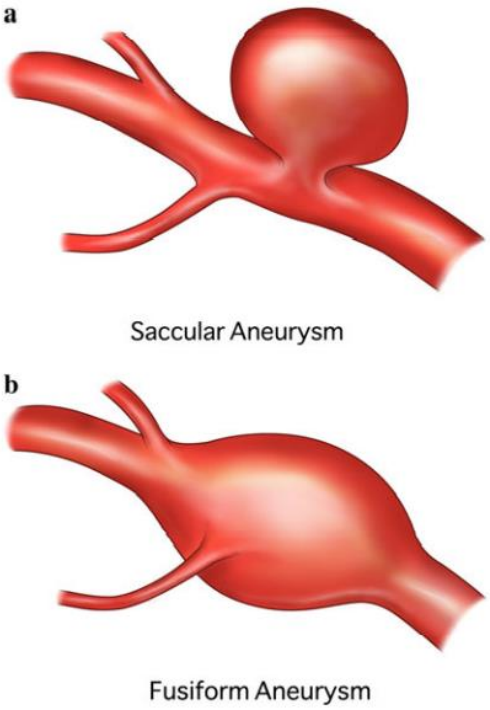


Figure of types of aneurysm shapes. (A) Intracranial saccular aneurysm or also commonly known as intracranial berry aneurysm and (B) fusiform aneurysm. This figure is from Withers *et al.* (2013) "Pipeline™ Embolization Device for the Treatment of Complex Intracranial Aneurysms" licensed for reuse under a Creative Commons Attribution-NonCommerical 4.0 license.[28]

IA are predominantly found within the anterior section of the circle of Willis[24,32,33] and commonly develop at major bifurcations in the cerebral vasculature.[34,35] In the case of intracranial berry aneurysms, the protrusion commonly ballooning in the direction of the blood flow.[36] They are common but the annual rupture rate is rare, occurring in less than 1% of intracranial aneurysms.[34,37,38] However, 80-85% of subarachnoid hemorrhages (SAH) are due to IA.[39] Patients with aneurysmal SAH have only a 35-50% survival rate.[40,41] Of those survivors, between 35-55% have permanent neurological deficits.[40,41] Detection of unruptured IA is difficult as most cases are asymptomatic and over 90% are discovered incidentally.[34,38] To aid in

detection, understanding the characteristics has shown that an individual may have an increased risk if they have hypertension,[38] are female,[26] have an advanced age,[26] or smoke,[38,42,43] and if the individual has a syndrome that predisposes them to IA[44,45] (Section 1.4.1.5); but the most significant risk factor is a family history.[24,25,33,46,47] However, the occurrence of multigenerational IA kindreds with multiple affected family members within each generation is rare.[37,38,48]

1.3 Intracranial Artery

The intracranial artery consists of 3 major layers: the intima, media, and adventitia.[36,49] The intima comes in direct contact with the blood and is composed of endothelium and subendothelial connective tissue.[36,49] The media surrounds the intima and is composed of vascular smooth muscle cells which is surrounded by elastin and collagen fibers.[36,49] Finally, the media is surrounded by a thin layer of connective tissue known as the adventitia.[36,49]

For neural signaling to occur, the central nervous system requires a very controlled environment.[50–52] Pericytes, astrocytic end-feet, and endothelial cells are the larger components of the blood-brain barrier that assist in controlling the substances that pass to and from the cerebrovascular system and the central nervous system's extracellular space.[50–52] Between the endothelial cells are the tight junctions which bind the endothelial cells together in order to prevent molecules from passing between them. The combination of tight junctions, gap junctions, and adherens junctions form the junctional complexes between the endothelial cells that prevents the free movement of molecules between the spaces and maintains the brain's delicate microenvironment.[53]

1.4 Familial Intracranial Aneurysms

Familial intracranial aneurysms (FIA) are hereditary, unlike sporadic IA, and occur when a family has two or more first- to third-degree family members with intracranial aneurysms.[54–57] When two or more first-degree family members have been diagnosed with IA, the rate of incidence increases significantly but the overall prevalence varies between studies ranging from ~9-20%.[39,41,58–60] Serial screening of first degree relatives can be done through computed tomography, magnetic resonance imaging and digital subtraction angiography.[55,56,61]

Patients with a positive IA and/or SAH family history tend to develop more serious phenotypes including a higher risk of developing multiple IA[62] and having a higher risk of rupturing.[58,63] Additionally, IA rupture at a younger age and at smaller sizes in those with a family history in comparison to those without.[25,58,64] In a long-term prospective study, from 1993-2013, Bor *et al.* (2014) invited patients ≥16-18 years old with a positive family history for serial screenings every 5 years.[65] Corroborating an overall IA prevalence of ~9% in first-degree relatives (FDR), they found IA in 11% of participants at the first screening, 8% at the second, 5% at the third, and 5% at the fourth.[65] This longitudinal study demonstrated that FDR may develop IA even after multiple negative screenings; showing why it is important to complete serial screenings throughout their life until the risks of treatment outweigh the risk of rupture.[65]

It has not been recommended,[61] nor is it economically feasible, to screen the general population. Recent evidence shows that the genetic architecture of unruptured and ruptured IA are similar[66] and the known heritability of SAH is 41%[67] which suggests that non-genetic factors have a significant role in the formation of IA. In the current literature, ~9% FDR ≥30 years old are found to have IA.[41,54,65] However, 19.1-20.6% of FDR ≥30 years old who have a history of smoking, hypertension, or both are found with IA.[58,68] These findings show that smoking and/or

hypertension appear to be important compounding modifiable factors in FIA. Moreover, the genetic contribution to FIA is prominent when compared with the general population with a 2- to 3-fold increase of risk in FDR. In addition to a positive family history, modifiable and non-modifiable risk factors should be taken into consideration in the screening and management of asymptomatic FIA first-degree family members.[61]

1.4.1 Genetics of Intracranial Aneurysms

Having a family history increases the risk of developing one or more IA. This suggests that there are one or more heritable factors. Many studies have implemented next-generation sequencing, genome-wide associations studies, and linkage analysis in families and population-wide projects to map the IA-gene regions. However, there is currently little known about the disease-gene associations.

1.4.1.1 Genetic Architecture of Common Complex Diseases and Conditions

Intracranial aneurysms are considered a common complex condition, meaning genetic and environmental factors can contribute. Mendelian diseases are suspected to be caused by rare variations with a high penetrance,[69] such as seen in a French-Canadian family with FIA (described in Section 1.4.2.3).[46] However, a common-complex condition that presents with apparent Mendelian inheritance could be a result of multiple genetic susceptibility variants – that together – confer the condition. In any case, a major gene may be responsible for FIA in some families.[46]

1.4.1.2 Genome-Wide Association Studies

Between individuals most of the genome is the same, but approximately every 1000 nucleotides there is genetic variation, these genetic variations are known as single nucleotide polymorphisms (SNPs).[70] Genome-Wide Association (GWA) studies use SNPs to identify regions of the genome with statistical significance that are associated with the condition at the population level. Since 2008,[71] GWA studies have aided in the identification of multiple loci along the genome that appear to be associated with the development of IA in predominantly Japanese, Dutch, Finnish, or other European descent populations (Table 1.1).[66,71–77] Chromosomal regions 9p21.3[66,71,73,77] and 8q11.23-12.1[66,71,77] have reached genome-wide significance in multiple studies which were replicated in more than one population. In the 9p21.3 locus, *CDKN2B-AS1*, an antisense non-coding RNA, is near the cyclin dependent kinase inhibitors *CDKN2A* and *CDKN2B*. The 9p21.3 locus has been associated with a few other vascular issues including: abdominal aortic aneurysms,[78] coronary heart disease,[79,80] and myocardial infarction.[81,82] SRY-box transcription factor 17, *SOX17*, is found at the 8q11.23 locus. This small gene, with only 2 exons, possesses a pathogenic heterozygous variation can cause congenital anomalies of the kidney and the urinary tract.[83,84] This locus has been associated with pulmonary arterial hypertension.[85–87] Recently, Bakker *et al*. (2020)[66] produced the largest GWA study to date with over 10,000 IA cases and over 306,000 controls from European and East Asian populations. Not only were they able to corroborate 6 previous associated loci, including 2 mentioned above, but they were also able to identify 11 new loci.

Table 1.1 – Loci Associated with Intracranial Aneurysms from Genome-Wide Association Studies

Chr.	Loci	RefSeq	Sequence Ontology[‡]	Populations Studied	Cohort Sizes (Case/Control)	OR Value	P-Value	Reference
2	q33.1	rs1429412	Regulatory Region	Dutch, Finnish, Japanese	2,196/8,085	1.22	5.8×10^{-7}	Bilguvar *et al.* (2008)[71]
		rs700651	Intronic			1.24	4.4×10^{-8}	Bilguvar *et al.* (2008)[71]
	q31.1	rs4667622	Regulatory Region	Portuguese	200/499	1.75	4.0×10^{-5}	Abrantes *et al.* (2015)[72]
4	q31.23	rs6842241	Intergenic	Japanese	2,431/12,696	1.25	9.6×10^{-9}	Low *et al.* (2012)[73]
		rs6841581	Intergenic	Dutch, Finnish and Japanese	5,891/14,181	1.22	2.2×10^{-8}	Yasuno *et al.* (2011)[74]
				European and East Asian	10,754/306,882	-	3.2×10^{-26}	Bakker *et al.* (2020)[66]
5	q31.1	rs4705938	Intronic	European and East Asian	10,754/306,882	-	2.5×10^{-10}	Bakker *et al.* (2020)[66]
6	q16.1	rs11153071	Intronic	European and East Asian	10,754/306,882	-	1.3×10^{-9}	Bakker *et al.* (2020)[66]
7	p21.1	rs10230207	Intronic	European and European Ancestry	4,133/7,869	1.21	9.91×10^{-10}	Foroud *et al.* (2014)[75]
8	q11.23	rs10958409	Intergenic	Finnish, Dutch, and Japanese.	2,196/8,085	1.36	1.4×10^{-10}	Bilguvar *et al.* (2008)[71]
		rs1072737	Intergenic	North American	1,483/1,683	1.25	<0.001	Foroud *et al.* (2012)[77]
		rs62516550	Exonic	European and East Asian	10,754/306,882	-	3.4×10^{-14}	Bakker *et al.* (2020)[66]
	q12	rs9298506	Intronic	Finnish, Dutch, and Japanese	2,196/8,085	1.29	1.8×10^{-9}	Bilguvar *et al.* (2008)[71]
9	p21	rs1333040	Intronic	Finnish, Dutch, and Japanese	2,196/8,085	1.29	1.4×10^{-10}	Bilguvar *et al.* (2008)[71]
		rs6475606	Intronic	North American	1,483/1,683	1.36	<0.001	Foroud *et al.* (2012)[77]
		rs10757272	Intronic	Japanese	2,431/12,696	1.21	1.6×10^{-6}	Low *et al.* (2012)[73]
	p21.3	rs1537373	Intronic	European and East Asian	10,754/306,882	-	2.9×10^{-29}	Bakker *et al.* (2020)[66]

Table 1.1 - Loci Associated with Intracranial Aneurysms from Genome-Wide Association Studies (Continued).

	q23.33	rs11187838	Exonic	European and East Asian	10,754/306,882	-	$1.6x10^{-8}$	Bakker *et al.* (2020)[66]
10	q24.3	rs12413409	Intronic	Dutch, Finnish, and Japanese	5,891/14,181	1.29	$1.2x10^{-9}$	Yasuno *et al.* (2010)[76]
		rs79780963	Intronic	European and East Asian	10,754/306,882	-	$2.3x10^{-14}$	Bakker *et al.* (2020)[66]
11	p15.5	rs2280543	3′ UTR	European and East Asian	10,754/306,882	-	$1.2x10^{-14}$	Bakker *et al.* (2020)[66]
12	p12.2	rs11044991	Intronic	European and East Asian	10,754/306,882	-	$1.7x10^{-8}$	Bakker *et al.* (2020)[66]
	q21.33	rs2681472	Intronic	European and East Asian	10,754/306,882	-	$6.7x10^{-9}$	Bakker *et al.* (2020)[66]
	q22	rs6538595	Intronic	Dutch, Finnish and Japanese	5,891/14,181	1.16	$1.1x10^{-7}$	Yasuno *et al.* (2011)[74]
		rs7137731	Intronic	European and East Asian	10,754/306,882	-	$4.9x10^{-14}$	Bakker *et al.* (2020)[66]
13	q13.1	rs9315204	Intronic	Dutch, Finnish, and Japanese	5,891/14,181	1.20	$2.5x10^{-9}$	Yasuno *et al.* (2010)[76]
		rs3742321	Exonic	European and East Asian	10,754/306,882	-	$5.5x10^{-15}$	Bakker *et al.* (2020)[66]
15	q25.1	rs8034191	Intronic	European and East Asian	10,754/306,882	-	$2.8x10^{-8}$	Bakker *et al.* (2020)[66]
16	q23.1	rs7184525	Intronic	European and East Asian	10,754/306,882	-	$5.6x10^{-15}$	Bakker *et al.* (2020)[66]
18	q11.2	rs11661542	Intergenic	Dutch, Finnish, and Japanese	5,891/14,181	1.22	$1.1x10^{-12}$	Yasuno *et al.* (2010)[76]
				European and East Asian	10,754/306,882	-	$3.2x10^{-17}$	Bakker *et al.* (2020)[66]
20	p11.23	rs4814863	Intronic	European and East Asian	10,754/306,882		$3.2x10^{-9}$	Bakker *et al.* (2020)[66]
	p12.1	rs1132274	Exonic	Dutch, Finnish and Japanese	5,891/14,181	1.20	$6.9x10^{-7}$	Yasuno *et al.* (2011)[74]
22	q12.1	rs39713	Intronic	European and East Asian	10,754/306,882	-	$4.1x10^{-8}$	Bakker *et al.* (2020)[66]

Reviewed in Kataoka 2015,[88] Hitchcock and Gibson 2017,[25] Zhou et al. 2018,[89] and Xu et al. 2019.[24] ǂSequence ontology collected using Ensembl Online Database (https://uswest.ensembl.org/index.html). *European ancestry includes: North America, Australia, and New Zealand. No provided odds ratio by corresponding study indicated by hyphen (-). OR: Odds ratio per allele; 3' UTR: 3 prime untranslated region.

1.4.1.3 Familial Gene Mapping Studies

Linkage is a term to describe how two genes are close enough in chromosomal position that they are more likely to be inherited together during recombination events. Linkage analysis in medical genetics uses our understanding of linkage to find genetic regions passed down after multiple recombination events that are associated or 'linked' with a phenotype.[90,91] Using our understanding of disease genetic architecture, when pedigrees present with Mendelian inheritance of intracranial aneurysms it is expected that the variant causing the condition will have a high penetrance, found co-segregated in the family, and rare in the global and subpopulations.

Linkage disequilibrium, a term coined by Lewontin and Kojima in 1960,[92] is the non-random association of alleles at two or more loci.[93,94] Originally microsatellite markers, which are multiallelic short tandem repeats, were used to identify 2 or more loci that are linked.[90] More recently, SNPs have been utilized decreasing the distance between markers that are statistically linked.[93,95] With advancing technologies, SNPs extracted from next-generation sequences can been used to study linkage throughout the entire genome.[96] The utilization of linkage analysis in studies of common-complex conditions, such as IA, have led to many loci being identified (Table 2.0).

Prior to discovering the connection between the *THSD1* gene and FIA, Santiago-Sim *et al.* (2009)[97] conducted linkage analysis on the French-Canadian family that harboured the deleterious *THSD1* variant.[97] Using a 250,000 SNP array, they were able to identify a 7-cM region with significant linkage in 13q14.12-21.1 with a maximum LOD (Logarithm of Odds) score of 4.56. *THSD1*, which in a subsequent study contained the pathogenic variant.[46] Additionally, they identified suggestive loci at 5q22.2-33.3, 9p23, 12p12, and 18p11 with LOD

scores of 2.91, 2.93, 3.10, and 3.15, respectively. Like GWA studies, linkage analysis only identifies a region and does not determine the exact variants causing the condition.

Table 1.2 – Loci Associated with Familial Intracranial Aneurysms from Linkage Studies

Chr.	Loci	Population Studied	Cohorts (Affected/Unaffected) ‡	LOD	Reference
1	p34.3-36.13	North American	12/8 from 1 family	4.20	Nahed *et al*. (2005)[98]
		Dutch	8/10 from 1 family	3.18	Ruigrok *et al.* (2008)[99]
4	q32.2	FIA Study	192 families†	2.50	Foroud *et al*. (2008)[100]
	q32.3	FIA Study	333 families†	2.60	Foroud *et al.* (2009)[101]
5	p15.2-14.3	French Canadian	9/3 from 1 family	3.57	Verlaan *et al.* (2006)[102]
	q22-33	Japanese	104 ASP from 85 families	2.24	Onda *et al.* (2001)[103]
7	q11	Japanese	104 ASP from 85 families	3.22	Onda *et al.* (2001)[103]
		North American	39/0 from 13 families	3.22	Farnham *et al.* (2004)[104]
8	p22.2	South Korean	9/22 from 5 families	3.61	Kim *et al.* (2011)[105]
11	q24-25	North American	2 families†	4.30	Ozturk *et al.* (2006)[106]
12	p13.3	FIA Study	333 families†	3.10	Foroud *et al.* (2009)[101]
	q21.33	FIA Study	192 families†	1.60	Foroud *et al*. (2008)[100]
13	q14.12-21.1	French-Canadian	10/25 from 1 family	4.56	Santiago-Sim *et al.* (2009)[97]
14	q22	Japanese	104 ASP from 85 families	2.31	Onda *et al.* (2001)[103]
	q23-31	North American	2 families†	3.00	Ozturk *et al.* (2006)[106]
17	cen	Japanese	93/27 from 29 families	3.00	Yamada *et al.* (2004)[107]
19	q13	Japanese	93/27 from 29 families	2.15	Yamada *et al.* (2004)[107]
		Finnish	48 ASP from 85 families	2.6	Olson *et al.* (2002)[108]
X	p22	Dutch	8/10 from 1 family	4.54	Ruigrok *et al.* (2008)[99]
		Japanese	93/27 from 29 families	2.08	Yamada *et al.* (2004)[107]
		Finnish	48 ASP from 85 families	2.16	Olson *et al.* (2002)[109]

Reviewed in Tromp et al. 2014[31], Hitchcock and Gibson 2017[25], Zhou et al. 2018[89], and Xu et al. 2019.[24] ‡Field values represent the number of affected/unaffected that participated in the linkage analysis, not the total number of affected/unaffected in the family. †The number of unaffected and affected participants used for the linkage analysis is unavailable. FIA Study: Familial Intracranial Aneurysm study with North American, Australian, and New Zealander populations. ASP: Affected sibling pair; LOD: Logarithm of Odds score.

1.4.1.4 Genomic Sequencing

The human haploid genome is consists of approximately 3 billion nucleotides.[110] As technologies advanced, we are now able to sequence whole exonic regions of the genome and the entire genome which is also known as next-generation sequencing (NGS). NGS provides critical information when diagnosing individuals and for disease-gene diagnosis.[111,112] Even though exonic regions compose approximately 2% of the genome, variants affecting protein coding regions account for approximately 85% of monogenic conditions.[113] However, a majority of the genome has biological importance,[114,115] making WGS studies important to consider when developing our understanding of conditions. Additionally, WGS is superior to WES due to its ability to interrogate intergenic, intronic, and exon regions, other non-coding biological elements, mitochondrial DNA, copy-number variations (CNV), and other structural variations. Genomic sequencing studies with variant filtration and analysis, found variations in *THSD1*,[46,116] *ARHGEF17*,[117] *ANGPTL6*,[116,118,119] *RNF213*,[116,120] *ADAMTS15*,[116,121] *TMEM132B*,[116,122] and *PCNT*[116,123] to be associated with FIA. Other FIA genomic sequencing studies have identified *EDIL3*,[116] LOXL2,[124] *ANK3*,[125] *NFX1*,[126] *C4orf6*,[127] and *SPDYE4*[127] as candidate genes but require further validation (Table 1.3).

Table 1.3 - Intracranial Aneurysm Genomic Sequencing Studies

Reference	Discovery Cohort	Sequenced (affected/ unaffected)*	Population	Type of Genomic Sequencing	Top Candidate Gene(s)†	Replication	Animal Models Study
Farlow *et al* (2015)[122]	7 families	(32/9)	N. American, New Zealander, Australian	WES	*TMEM132B*	No	No
Yan *et al.* (2015)[121]	12 families	(42/0)	Japanese	WES	*ADAMTS15*	Yes	No
Santiago-Sim *et al.* (2016)[46]	1 family	(9/13)	French-Canadian	WES	*THSD1*	Yes	Yes
Zhou *et al.* (2016)[120]	6 families	(26/0)	French-Canadian	WES	*RNF213*	Yes	No
Yang *et al.* (2018)[117]	20 familial and sporadic cases	(20/0)	Chinese	WES and WGS	*ARHGEF17*	Yes	Yes
Bourcier *et al.* (2018)[118]	5 families	(14/20)	French	WES	*ANGPTL6*	Yes	No
Lorenzo-Betancor *et al.* (2018)[123]	3 families	(5/8)	United States	WES	*PCNT*	Yes	No
Wu *et al.* (2018)[124]	1 family	(3/1)	Chinese	WES	*LOXL2*	Yes	No

Table 1.3 Intracranial Aneurysm Genomic Sequencing Studies (Continued)

Powell *et al.* (2019)[127]	2 families	(17/0)	Newfoundland Canadian	WES	*C4orf6*, *SPDYE4*	No	No
Ding *et al.* (2020)[126]	1 family	(7/8)	Chinese	WES	*NFX1*	No	No
Sauvigny *et al.* (2020)[116]	35 unrelated and 3 aff family members from 1 family	(38/0)‡	Northwestern German	WES	*PCNT*, *RNF213*, *THSD1*, *ANGPTL6*, *ADAMTS15*, *TMEM132B*, *EDIL3*	No	No
Hostettler *et al.* (2021)[119]	275 familial IA cases	(275/0)	United Kingdom	Targeted sequencing	*ANGPTL6*	No	No
Lui *et al.* (2021)[125]	3 families	(9/0)	Chinese	WES	*ANK3*	Yes	No

†Top candidate gene(s) identified by the study. *Includes individuals used for targeted sequencing validation. ‡Study participants had subarachnoid hemorrhages; included in this table due to our understanding that 80-85% of SAH are due to ruptured IA.[39] Aff: Affected; Unaff: Unaffected; WES: Whole exome sequencing; WGS: Whole genome sequencing; N. American: North American.

The first FIA disease-gene association made was by Santiago-Sim *et al.* in 2016.[46] In this study, WES was completed with two affected cousins, a part of a multigenerational family, identified a truncating variation (p.R450X) in the Thrombospondin Type 1 Domain Containing 1 (*THSD1)* gene on chromosome 13q14.3. This variant segregated in all 9 affected members in the family while not being present in 11 unaffected family members. To validate their findings, they performed targeted *THSD1* exon and exon-intron boundary sequencing in a cohort of 507 unrelated IA probands, testing against the ExAC database as well as 305 unaffected controls collected separately. They found in 8 of the 507 probands rare variants in the *THSD1* gene. In zebrafish and mouse animal models, they used morpholino oligomers (MO) to knockdown the *thsd1* zebrafish ortholog and a knock-in fluorescence reporter of mouse ortholog *thsd1*. This resulted in an in an increase of intracranial hemorrhage in both the zebrafish and mice. Santiago-Sim *et al.* (2016) *in vitro* studies used human umbilical vein endothelial cells (HUVEC) and completed a controlled knockdown of *THSD1* using small interfering RNA (siRNA) which showed a significant decrease in relative cell adhesion to collagen I. Using siRNA-resistant *THSD1* wildtype mRNA co-transcribed with the *THSD1* siRNA, HUVECs were rescued to the relative cell adhesion seen in the HUVECs transfected with a control siRNA. Patient-derived siRNA-resistant mRNA co-transcribed with the *THSD1* siRNA were not able to rescue the relative cell adhesion seen in the control HUVECs.

In a study done by Yang *et al*. in 2018, another gene -- Rho Guanine Nucleotide Exchange Factor (*ARHGEF17*) -- was found to have significant association with the development of intracranial hemorrhage.[117] A discovery cohort with 20 familial and sporadic IA patients of Chinese ancestry were studied using both WGS or WES prioritizing for candidate variants with a filtration strategy. Yang *et al.* (2018) prioritized variants that: had an allele

frequency less that 5% in the East Asian population of 1000 Genome Project and Exome Aggregation Consortium, were predicted to be damaging by some algorithms (GERP++, CADD, SIFT, Polyphen2, LRT and MutationTaster) and that the candidate variants were shared by 2 or more other families. This filtration identified 30 candidate variants from the discovery cohort. The variants were further investigated in 3 separated replication cohorts of Japanese, European-American, and French-Canadian ancestry. *ARHGEF17* was identified to have heavy deleterious variant burden in highly conserved regions. To further examine this in the zebrafish animal model, splice-blocking morpholinos (MO) were used to knockout the zebrafish *arhgef17* ortholog, which is 67% identical to the human *ARHGEF17*. This resulted in *arhgef17* deficiency and a significant increase in zebrafish intracranial hemorrhaging. Subsequently, mRNA rescue studies were conducted on zebrafish which were co-injected with the *arhgef17* MO and the human *ARHGEF17* mRNA wildtype or with a human-variant (rs2298808) *ARHGEF17* mRNA. In zebrafish co-injected with the MO and human *ARHGEF17* wildtype mRNA, there was no statistical difference from the control group; while zebrafish co-injected with the MO and the human variant *ARHGEF17* mRNA had a statistically significant increase in intracranial hemorrhage from the control group.

1.4.1.5 Syndrome-IA Associations

There are a several syndromes with an elevated risk of developing intracranial aneurysms, similar to that of a family history of IA.[25,89] Mendelian syndromes with association to intracranial aneurysms include: Autosomal Dominant Polycystic Kidney Disease (ADPKD), Vascular Ehlers-Danlos Syndrome Type IV, Loeys-Dietz Syndrome, Marfan Syndrome, Neurofibromatosis Type I and Microcephalic Osteodyplastic Primordial Dwarfism Type II

(Table 3.0). Other genetic conditions, such as pseudoxanthoma elasticum, hereditary hemorrhagic telangiectasia, and multiple endocrine neoplasia type I have been reported in case studies but have not been adequately studied to provide their respective IA prevalence.[25] These syndromes show that some genes can predispose an individual to IA. Understanding the roles of these genes in IA formation can provide insight into the genetic mechanism behind them.

Table 1.4 - Incidence of Intracranial Aneurysm in Genetic Syndromes

Syndrome	Mendelian Inheritance	Associated Genes	IA Prevalence
Autosomal Dominant Polycystic Kidney Disease (ADPKD)	AD	*PKD1* *PKD2*	8-12%[128–130]
Vascular Ehlers-Danlos Syndrome	AD	*COL3A1*	11%[45]
Loeys-Dietz Syndrome	AD	*TGFBR1*, *TGFBR2*, *TGFB2*, *SMAD3*	10-28%[45,131,132]
Marfan Syndrome	AD	*FBN1*	11-14%[45,133]
Neurofibromatosis Type 1	AD	*NF1*	9-11%[45,134]
Microcephalic Osteodysplastic Primordial Dwarfism, Type II (MOPD2)[a]	AR	*PCNT*	Up to ~50%[135,136]

[a]MOPD2 is a rare disease with an unknown prevalence (ORPHA:2637); IA, moyamoya disease, and other cerebrovascular condition studies are limited. AD: Autosomal dominant; AR: Autosomal recessive

1.4.1.6 Northern Indigenous Peoples and Intracranial Aneurysms

Kristensen (1983)[137] first described a higher prevalence of SAH in Greenlandic Inuit populations when compared to the Danish population in a retrospective study of SAH cases that were treated in Copenhagen, Denmark. In this time span, 436 Danish patients had been treated

while only 27 Greenlandic Inuit were transported for treatment. Additionally, she described SAHs happened at a younger age in the Greenlandic Inuit population.[137] However, she acknowledged that there is a significant selection bias due to stability of patients and the transport required.[137]

In an another retrospective study, Lindgaard *et al.* (2003)[138] confirmed that in comparison with the Danish population the Greenlandic Inuit had higher rate of SAH as well as newly identified higher prevalence of IA. Overlapping with the previously mentioned study[137] between 1978-1998, 120 Greenlandic Inuit were treated in Copenhagen, Denmark for SAH or IA, while 1,037 Danish patients were treated in the same hospital between 1978-83. They identified that the Indigenous population had a five-fold higher rate of having IA with a family history.[138]

1.4.1.7 Northern Indigenous Intracranial Aneurysm Genetics

Zhou *et al.* (2018) used GWA and WES methods to examine genetic variations in a French-Canadian FIA discovery cohort.[139] Using 621,983 SNPs with 172 French-Canadian IA cases and 1,772 French-Canadian controls, they completed a GWA study to identify loci of significance. After imputation, one locus reached genome-wide significance rs1554600 on 3p14.2 in gene Fragile Histidine Triad Diadenosine Triphosphatase (*FHIT)* with a p-value of 4.66×10^{-9} but a low odds ratio of 0.26. A FIA cohort of 34 Canadian Nunavik Inuit families were used to replicate their findings. From the WES data, the most significant of the GWA SNPs was rs780365 (intron variant) in the *FHIT* gene with a p-value of 2.839×10^{-3}, but lost significance after corrections for multiple testing. However, due to its expression in the brain, the significance in the French-Canadian population and suggestive significance in the replication Nunavik Inuit

cohort, this gene locus (3p14.2) is important to consider in our understanding of IA pathogenesis in an Indigenous population.

In a study of the genetic architecture in the Canadian Nunavik Inuit, they found that the geographically remote populations have distinct genetic architectures from any known present-day population.[140] Zhou *et al.* (2019)[140] conducted an association test on 8,291 SNPs within the coding regions of the genome found to have selection footprints using 62 Nunavik IA cases and 42 Nunavik controls. After the Bonferroni correction the genome-wide significance was set to p-value < 6.03×10^{-6}, they identified one SNP with genome-wide significance with p-value of 4.60×10^{-8} on the Olfactory Receptor Family 4 subfamily C member 3 (*OR4C3*) gene at rs7747087. The second most significant, without genome-wide significance, with a p-value of 6.71×10^{-5} was SH3 and Multiple Ankyrin Repeat Domain 3 (*SHANK3*) gene at rs116959666.

1.5 Objectives

The objectives of this study are to investigate the genetic variants that influence the development of familial intracranial aneurysms in a First Nation from Northwestern British Columbia and to aid in further closing the genomic information gap that exists between Indigenous and Non-Indigenous populations.

1.6 Study Questions

1. Is the condition within these families caused by a single rare variant (monogenic origin)?
2. Is there evidence of a multigenic origin to explain the condition within these families?

Chapter 2: Methods

2.1 Ethics Approval

This study on familial intracranial aneurysms in the Gitxsan Community was a sub-study in Precision Diagnosis for Indigenous Families with Genetic Conditions (Activity 2 of the Silent Genomes Project). Harmonized ethics approval for Activity 2 was obtained from the University of British Columbia Children's and Women's Research Ethics Board and Vancouver Island Health (REB Number: H18-00726).

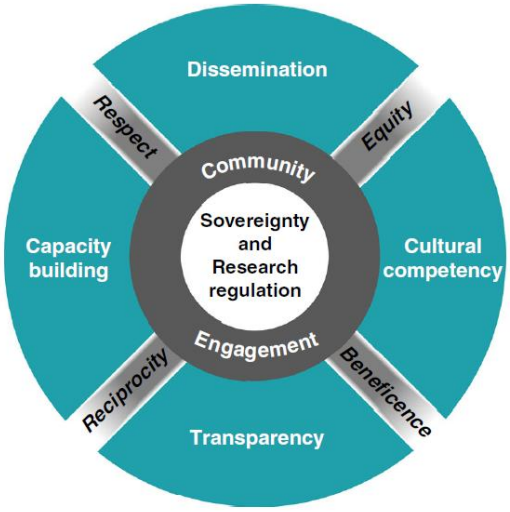
2.2 Community Engagement

Dr. Laura Arbour has been working with the Gitxsan since 2003 to address Long QT syndrome. On numerous occasions, the concern about IA arose, and they were asked to consider research that could address the problem. These conversations included the Gitxsan Health Executive Health Directors and the Gitxsan Health Board. It became apparent that the Silent Genomes Project, precision diagnosis project could be an opportunity. One of the three criteria for inclusion is: 'disorder is limited to a single system, is most likely due to a single genomic cause in the patient, but an appropriate targeted test is not available, or available tests were non-diagnostic' (https://www.bcchr.ca/silent-genomes-project/precision-diagnosis-study/health-professionals). At a community gathering in the spring of 2019, with the attendance of the Gitxsan Board Director and many board members, the Silent Genomes Project was presented as a way to address this problem. It had been announced that the genetics of IA would be on the agenda, therefore there were several affected individuals in the audience. It was agreed to work with the Long QT Syndrome (LQTS) study to identify cases (permission to re-contact for future research) and the Silent Genomes Project to enrol and carry out the genetic studies. Updates have

been provided through annual newsletters (2020 Annual Newsletter - Appendix A, 2021 Annual Newsletter Appendix B) and through a virtual research updated (spring 2021) where I presented the project to date (Appendix C).

Community-based participatory research (CBPR) is a research methodology that is recommended when working with minority groups, including Indigenous communities.[141–143] Informed by CBPR methodologies, Claw *et al.* (2018) proposed an ethical framework for engaging Indigenous communities for genome research (Figure 2.1).[143]

Figure 2.1 - Depiction of Indigenous Genetic Ethical Framework



A visual depiction of ethical framework when working with Indigenous people. This figure is from Claw *et al.* (2018) "A framework for enhancing ethical genomic research with Indigenous communities", licensed for reuse under a Creative Commons Attribution 4.0 license.[143]

Claw *et al.* (2018) placed the recognition of community sovereignty, self-determination, and community research regulations at the core of the framework.[143] Continuous community engagement in research protocol development until the dissemination of findings encompasses the community sovereignty and community research regulations. They state that researchers

involved must be culturally competent, be transparent with their goals and research progress, build capacity of knowledge and abilities of the community to better their understanding and involvement with the research; and disseminate the findings to the community to allow them to choose how to use and release this information that will be best for the community. The ethical values of reciprocity, beneficence, equity, and respect are placed in the framework as well because they ask that they be present throughout the research process. By updating the community about my progress, building their understanding of the condition, listening to their thoughts and understandings of the condition and by being transparent was how I followed this ethical framework and walked in a good way.

2.3 Patient Enrollment and Collection Clinical Characteristic Information

As stated in Section 1.1.3, these families were enrolled in this study because FIA was prioritized as a concern in the Gitxsan community and affected family members came forward. Building this research project on pre-existing relationships, affected community members were enrolled from previous community-research partnerships on long QT syndrome.[19,20] In this FIA study, participants were selected from 22 IA-affected participants of the long QT syndrome research with the exception of one participant who joined this study after attending a community gathering discussing the project. The LQTS project included those with the *KCNQ1* p.V205M and *ANK2* p.S646F contributing to LQTS, and some with IA were within these families.

A study Genetic Counsellor (Sarah McIntosh) contacted the participants who had previously provided re-contact consent. Over-the-phone or in-person interviews were completed to take their family history, gather verbal clinical information, consent to participate in the study (Appendix D: Silent Genomes Precision Diagnosis consent), and consent to an Authorization to

Release Healthcare Information. One individual, who was not a participant in the LQTS study, was enrolled only through the Silent Genomes Project consented by the same Genetic Counselor. Medical documentation retrieved from healthcare providers was used to confirm the patient's IA diagnosis, to determine if there were other underlying medical conditions that could be potentially contributory, their family history, and to better understand the patients' health in order to better to help interpret sequencing data.

Inclusion criteria for individuals of affected status:

1. Have a positive IA diagnosis by CT, MRI, or DSA;
2. Have one or more first-degree relatives (sibling, parent, or child) with a reported or documented IA.

Exclusion criteria for individual participation:

1. Have a syndrome with a predisposition to IA (ADPKD, Ehlers-Danlos Syndrome, Loeys-Dietz Syndrome, Marfan Syndrome, etc.);
2. Do not provide consent to the study.

A total of six affected individuals were selected within the families for sequencing. The selection was based on being as genetically distant as possible to limit the number of variants shared between the affected individuals. Only confirmed affected individuals were sequenced because the ostensible inheritance pattern appeared to be autosomal dominant with reduced penetrance. Unaffected individuals may carry candidate variant(s) but not present with the

phenotype. However, proceeding with the assumption that this condition in these families is caused by a single and large effect variant like the one presented by Santiago-Sim *et al.* (2016),[46] most, if not all, affected individuals would have the variant of interest.

2.4 Sample Collection and Storage

Blood, saliva, or buccal samples were collected for WGS. Buccal samples were collected using kits provided by LifeLabs. Saliva sample collection kits, DNA Genotek Oragene·DNA (OG-500), were sent to the participants homes and the samples were returned to the Michael Smith Genome Science Centre in Vancouver, British Columbia; except for two samples which were transferred from the Long QT Syndrome study to the Silent Genomes FIA study. Phlebotomy was performed by LifeLabs during an in-person visit to one of their clinics around British Columbia. Blood samples were preferred as the quality of variant calling, particularly for structural variants, continues to be higher from sterile tissue sources such as blood than from multi-organism samples such as obtained from the oral cavity.[144,145] All samples collected were governed by the concept of "DNA on Loan"[15], as described in section 1.1.2, and stored following the protocols agreed to in partnership with the community and the researchers.

2.5 Whole Genome Sequencing

DNA extraction, quality control testing, whole genome sequencing, and alignment were completed by the Michael Smith Genome Sciences Centre. Genomic DNA was extracted and then sheared through sonication (Covaris LE200), end-repaired targeting 300-400bp fractions with paramagnetic PCRClean DX beads (C-1003-450, Aline Biosciences), and after 3'A-tailing, full length TruSeq adapters were ligated. The PCR-free genome library concentrations were

tested using qPCR Library Quantification kit (KAPA, KK4824). Whole genome sequencing was completed using Illumina HiSeq X sequencers, generating 150 paired-end tag (PE150) reads at ~30-40X mean coverage. Read alignment was completed using the Burrows-Wheeler Aligner[146] (BWA-MEM) v0.7.6a to Genome Reference Consortium Human Builder 37 (hg19.GRCh37).

2.6 Variant Calling

Variant calling was completed by a Genomic Analyst (Drs. Jill Mwenifumbo and Arezoo Mohajeri). Single nucleotide variations and small indels were called using GATK Haplotype Caller[147] v3.8. Mitochondrial variants were called and annotated using MToolBox[148] v1.0. Structural variants were called using LUMPY Express[149] v0.2.13, Manta[150] v1.1.1, CNVnator[151] v.0.3.3, and ERDS[152] v1.1. Structural variants were then merged, annotated, and filtered with SURVIVOR[153] v1.0.7, ANNOVAR[154] v2018-04-16, and custom scripts respectively.

2.7 Variant List Annotation

I used VarSeq v2.2.3 by Golden Helix (https://www.goldenhelix.com/products/VarSeq/) to annotate the variants. VarSeq annotated the variants with information from the Online Mendelian Inheritance in Man (OMIM; https://omim.org/) database, ClinVar, gnomAD allele frequency databases, in silico predictor tools (CADD, Polyphen2, SIFT, etc.) and more.

Additionally, I manually prepared a gene candidate list using databases: Human Phenotype Ontology database (HPO; https://hpo.jax.org/app/), National Centre for Biotechnology Information Gene database (NCBI; https://www.ncbi.nlm.nih.gov/gene), and the Genotype-Tissue Expression (GTEx) database (https://www.gtexportal.org/home/) by searching key terms (Table 5.0). Genes were considered stronger candidates based on their genomic,

transcriptomic, and proteomic characteristics,[155] and current understanding of function collected from the NCBI Gene database and PubMed (https://pubmed.ncbi.nlm.nih.gov/advanced/) in association with the literature search terms (Table 2.1).

Table 2.1 - Literature Search Terms for Gene Candidate List

Intracranial Aneurysm Terms	Related Terms, Conditions, and Processes
• Intracranial Aneurysm‡ • Saccular Aneurysm • Fusiform Aneurysm • Subarachnoid hemorrhage • Hemorrhagic Stroke • Stroke*	• Thoracic Aortic Aneurysm (TAA)‡ • Abdominal Aortic Aneurysm (AAA)‡ • Aneurysm* • Cerebral • Vasculogenesis • Angiogenesis • Inflammation* • Inflammatory Response

MeSH terms denoted by asterisks (*). ǂ Genetic studies have identified overlapping loci and genes between various types of aneurysms[156,157] and studies have shown patients developing multiple types of aneurysms at the same time (IA, AAA, and TAA).[158,159]

2.8 Overall Analysis Strategy

2.8.1 Phenotyping Within Affected Families

SAH twin studies have identified a 41% heritable factor;[67] however, there is also evidence of concordance in aneurysm location.[160,161] Using available collected information of the location, size, number, age of diagnosis, symptoms, environmental factors (smoking, hypertension, etc.), and rupture incidence was compared across participants to determine if there was a consistency of phenotypes within these families.

2.8.2 Overall Filtration Strategy

When a disease affects individuals each generation of a family, with some cases of male-to-male transmission, Mendelian autosomal dominant inheritance is the likely explanation. However, a family from a relatively remote community may be experiencing pseudo-dominant inheritance due to bi-allelic variants in a recessive gene. To determine if the pathology has a monogenic origin, both autosomal models were considered. Filtering strategies were implemented to narrow down candidate SNV, small indels, mitochondrial variations, CNV, and structural variations.

2.8.2.1 Single Nucleotide and Small Indel Filtration

The filtering strategy for SNV and small indels was as follows:

1. **Read Depth** greater than or equal to 10 and **Genotype Quality score** greater than or equal to 20
2. **AND Minor Allele Frequency (MAF)** less than or equal to 1%[1] or unreported in the Broad Institute processed global genome and exome allele frequencies from Genome Aggregation Database (gnomAD)[162] v2.1.1 and less than or equal to 1% in gnomAD exome subpopulations v2.1.1 (European, African, Ashkenazi Jewish, East Asian, South Asian, Latino, & Other)
3. **AND Sequence Ontology and Splice-Site In-Silico Predictors** that indicate the variation is more likely to affect coding sequence[2]

[1] A less strict genome and exome gnomAD MAF used due to our understanding of complex disease architecture
[2] Filtration for coding sequencing effects due to our understanding that exonic variations causes ~85% of monogenic conditions,[113] therefore changes to the resulting protein are more likely to cause this Mendelian inheritance

a. RefSeq Genes[163] (v105.20201002) and Ensembl Genes[164] (v87): 5 prime UTR premature start codon gain variant, disruptive inframe indels, exon loss variant, frameshift variant, inframe indels, initiator codon variant, missense variant, splice donor/acceptor variant, splice region variant, stop gained/lost variant

b. **OR** Splice Altering Predictions[165] Ada and RF Score (v1.1) greater than or equal to 0.6

4. **AND** in-house database[166] **allele count** less than or equal to 10 (greater than or equal to 5%[3] allele frequency)[4]
5. **AND Allele Sharing** between greater than or equal to five affected individuals[5]

After filtration, the remaining variants were individually examined, noting the type of variation (Start/Stop, Splice site, etc.), the predicted impact from variation prediction algorithms, and the biological role of the impacted gene to prioritize top candidate variants. Zygosity was used to prioritize variants but homozygous and heterozygous variations were both considered due to the community's remote location. Additionally, the region or gene that the variants are harboured in were compared to the annotated gene candidate list as a means of prioritization. Prioritized variants were viewed in the Integrative Genomics Viewer[167] to identify if the variant was more likely to be an artifact.

[3] A loose allele count was used due to the small size of the database (~102). Richards *et al.* (2015)[195] published the ACMG guidelines for sequence variant interpretation which stated that a high MAF was 'stand-alone' evidence of a benign variant in Mendelian variant interpretation.
[4] An in-house database from an earlier study in the British Columbia Children's Hospital Research Institute, which was used to aid in removal of systemic artefacts
[5] IA are common (~3% of the general population), incomplete allele-sharing was allowed due to the possibility of sporadic IA phenocopies

2.8.2.2 Copy Number Variations and Structural Variation Filtration

The initial calling and annotations of copy number and structural variations was performed by the same Genomic Analysist. The filtration strategy utilized was as follows:

1. **AND** called by at least two algorithms (LUMPY Express[149] v0.2.13, Manta[150] v1.1.1, CNVnator[151] v.0.3.3, and ERDS[152] v1.1)
2. **AND** shared between both affected participants that provided blood samples

The remaining variants were then evaluated and prioritized in the same manner as SNV and small indels (Section 2.8.2.1). Only CNV and structural variations identified from blood samples were assessed for allele sharing due to the higher true positive and lower false positive rates that blood samples provide.[145]

2.8.2.3 Mitochondrial DNA Variation Filtration

Mitochondrial DNA is closed and circular with only ~16,500 nucleotides.[168,169] After variant calling a relatively limited number of variants remained. The filtration strategy was as follows:

1. **Read Depth** greater than or equal to 10 and **Genotype Quality** greater than or equal to 20;
2. **MAF** less than or equal to 1% from MitoMap[170] (curated by Dr. Jill Mwenifumbo on 06/NOV/2019);
3. **AND** in-house database **Allele Count** less than or equal to 10 (less than or equal to 5%[6] allele frequency);

[6] A loose allele count was used due to the small size of whole genomes in the database (~102). Richards *et al.* (2015)[195] published the ACMG guidelines for sequence variant interpretation which stated that a high MAF was 'stand-alone' evidence of a benign variant.

4. **AND Allele Sharing** between greater than or equal to five of the affected individuals

2.9 Modifiable Risk Factors

As the genetics of IA is being unwoven, it has been shown that this common complex condition has a multigenic architecture.[66] Additionally, a GWA study identified an increased risk of intracranial aneurysms associated with an additive effect of identified risk alleles.[71] Understanding this, risk alleles identified in previous GWA studies from other populations (Table 1.1) or from other Indigenous populations (Section 1.3.9.1) were investigated in these families to identify if they possess any risk alleles. Risk alleles found were filtered by allele sharing in greater than or equal to five affected participants.

2.10 Dissemination

Information necessary for patients will be relayed to them through healthcare professionals. As a partial requirement for a Master's degree from the University of British Columbia, collected data and results will be prepared for publication in a scientific journal in the form of a thesis. Additionally, information collected and discovered about the disease will be returned to the community through gatherings and update letters.

Chapter 3: Results

3.1 Overall Family Structure and Clinical Characteristics

This large multigenerational family has multiple separate families connected throughout the generations. The history of IA and SAH spans at least four generations with at least 40 individuals affected (Figure 3.1). Of the 40 reported individuals, 22 participated in the previously mentioned long QT syndrome (LQTS) study and one individual only participated in this study. Using only the aneurysm reports from LQTS study participants, the prevalence of IA in approximately 800 Gitxsan participants is close to the overall prevalence in the general population (2.75% vs. 3.2%[26]). All participants had at least one first-degree relative with 1 or more IA. FIA01, FIA02, and FIA03's exact biological relation with FIA04, FIA05, and FIA06 are not defined; however, it has been reported that FIA02's paternal grandmother and FIA04, FIA05, and FIA06's maternal grandfather were related (Table 3.1 and Figure 3.1).

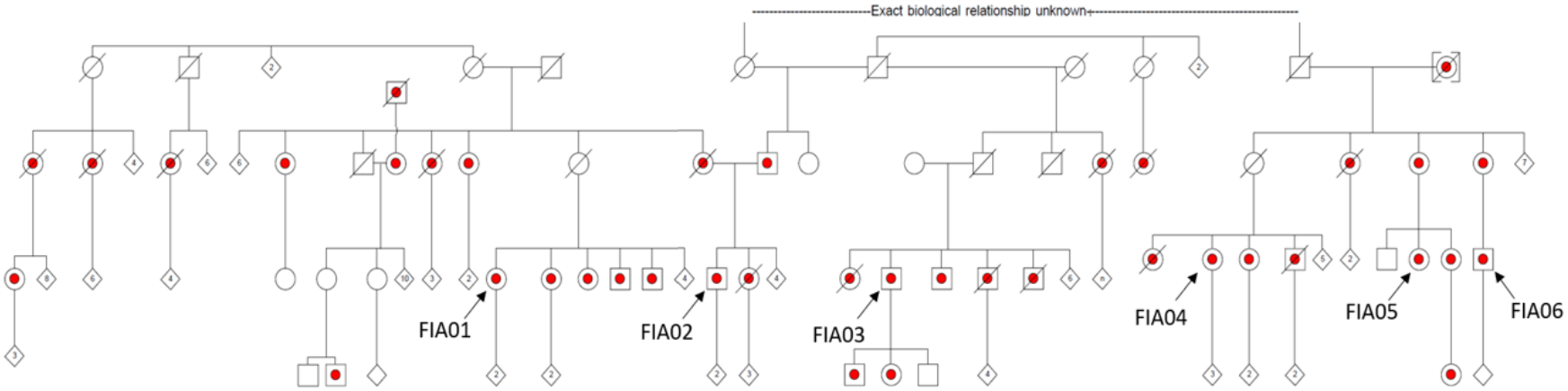
The LQTS study (n=17) had 13 females and four males with IA, while this FIA study had 3 affected females and 3 affected males for a total of 16 females and 7 males. The total age range was large ranging between 27 and 73 years old with the average age in the FIA study three years younger than the total average (45 and 48, respectively). The average number of IA in the FIA study was almost twice as high as the remaining LQTS study participants (2.5 and 1.40, respectively). Hypertension was reported in 13 out of 23 of the total participants. Variants found commonly in these families and throughout this community were reported with five individuals positive for *ANK2* p.S646F and four for the *KCNQ1* p.V205M. While all were positive for the *CPT1A* p.P479L variant, 10 participants were heterozygous and 12 homozygous for the P479L variant, while no homozygous wildtypes were found.

Table 3.1 - Overall Family Characteristics

		LQTS Study (n = 17)		**FIA Study** (n=6)		**Overall** (n=23)	
Proportion (F:M)		13:4		3:3		16:7	
Average age at diagnosis (range)		50 (27-73)		45 (34-60)		48 (27-73)	
Average number of IA		1.4		2.50		1.7	
Hypertension		10/17		4/6		14/23	
Common Variants[δ]	***ANK2* p.S646F**[†]	Positive 3	Negative 13	Positive 2	Negative 4	Positive 5	Negative 17
	***KCNQ1* p.V205M**	Positive 3	Negative 14	Positive 1	Negative 5	Positive 4	Negative 19
	***CPT1A* p.P479L**[†‡]	Het (PL) 9	Hom (LL) 7	Het (PL) 1	Hom (LL) 5	Het (PL) 10	Hom (LL) 12

[†]Testing for ANK2 p.S646F and *CPT1A* p.P479L was not completed for one participant due to insufficient sample. [‡]No participants with wildtype (PP) genotype. [δ]Variants found commonly in the Gitxsan community and these families. No smoking information collected from the LQTS Study; data for smoking in FIA Study found on Table 3.2. Het: Heterozygous; Hom: Homozygous; LQTS: Long QT syndrome; FIA: Familial intracranial aneurysm

Figure 3.1 – Large Multigenerational FIA Pedigree



This pedigree was generated using family histories collected from the Long QT Syndrome (LQTS) Study and through this study. The information presented may not perfectly reflect the up-to-date family structures and affected/unaffected status. There are reports of familial relations between the paternal grandmother of FIA02 and the maternal grandfather of FIA04/05/06; however, the exact biological relationship remained unclear: denoted by dashed-line (----). Circles: female; Squares: male; Diamond: sibship with no known IA/SAH diagnosis; Red dot: Affected; Slash: deceased.

3.2 FIA Study Participants' Characteristics

3.2.1 Non-Genetic Risk Factors

Cigarette smoking and hypertension are well-established modifiable risk factors for intracranial aneurysms.[38,42,43] With the exception of FIA01, all participants smoked tobacco at one point in time or are currently (Table 3.2). No history of hypertension was reported in FIA03 and FIA05 while the remaining participants had a positive history. FIA01, FIA04, and FIA06 reported having their hypertension treated and FIA01 and FIA04 had it under control while FIA06 has it under control with some fluctuations.

3.2.2 Intracranial Aneurysm History

Among the 6 participants, the average age at diagnosis was 45. This average may not be representative of when the aneurysm formed for several possible reasons, including the fact that screening of first-degree relatives is recommended after two or more familial cases are diagnosed. There was a total of 15 documented aneurysms across the six participants: four arose from the internal carotid artery, five from the middle cerebral artery, two were in unknown locations, and one aneurysm each found from the anterior communicating artery, posterior communicating artery, anterior cerebral artery, and the basilar artery (Table 3.2).

FIA01 was diagnosed at 60 years old with a possible 2.2 mm aneurysm on the left internal carotid artery (ICA) using a CT scan (Table 3.2). She has not received surgical or endovascular coiling treatment for the aneurysm and is undergoing serial screening to monitor the aneurysm.

With first diagnosis at the age of 38, FIA02 has developed four intracranial aneurysms to date. He presented with a rupture of an intracranial aneurysm (~5 mm) of the left middle cerebral

artery (MCA) at which point another aneurysm of the right MCA was also identified. Years later, a new aneurysm of the left posterior communicating artery (PCoA) ruptured and was treated with endovascular coiling while the previous aneurysm on the MCA increased in size and another was found arising from the anterior communicating artery (ACoA). The following year, FIA02 underwent elective clipping of the aneurysms on the right MCA and the ACoA.

FIA03 was first diagnosed at 52 with a 3mm aneurysm from the left MCA which was clipped shortly after its discovery. Later, he was diagnosed with two more aneurysms along the right MCA, and these were also surgically clipped. Recently, a CT angiogram revealed a small 1.5mm aneurysm of his right ICA.

At age 44 years, FIA04 required emergent aneurysm clipping. A year later, CT angiogram revealed three aneurysms of the basilar artery, left ICA, and left anterior cerebral artery (ACA). The aneurysms from the left ICA and the left ACA were clipped shortly after their discovery while the basilar artery aneurysm was treated with endovascular coiling a few months later.

From serial family screening, FIA05 was diagnosed via MRA with a 4.5 mm aneurysm of the right ICA at 39. This aneurysm is being monitored every 1-2 years.

FIA06 presented with severe headaches, photophobia, stiff neck, and episodes of nausea and vomiting. After an MRI, he was diagnosed with an aneurysm that was immediately treated surgically. Impairment of the short-term memory was reported after the surgery.

3.2.3 Common Variants Specific to this Population

The *ANK2* p.S646F variant was present in two family members (FIA04 and FIA 05) and neither individual was diagnosed with long QT syndrome type 4. In FIA01 carried the *KCNQ1*

p.V205M variant and had a diagnosis of long QT syndrome type 1. With five homozygous and one heterozygous, all participants carried the *CPT1A* p.P479L variant.

Table 3.2 – Characteristics of Affected Participants

Participant		FIA01	FIA02	FIA03	FIA04	FIA05	FIA06
Sex (M/F)		F	M	M	F	F	M
Sample Type		Buccal	Buccal	Blood	Saliva	Blood	Saliva
Age of First Diagnosis:		60	38	52	45	39	34
Method of Imaging		CT/MRI/MRA	CT/MRI	CTA/MRI	CTA	MRA/CTA	MRI
Number of IA:		1	4	4*	4*	1	1*
Location and Size:		L-ICA (2.2 mm)	L-MCA (5 mm) R-MCA (5 mm) ACoA (2 mm) L-PCoA (6 mm)	L-MCA (3 mm) R-MCA R-MCA R-ICA (1.5mm)	Basilar Artery L-ACA L-ICA	R-ICA (4.5mm)	-
Rupture		No	Yes	No	No	No	-
Surgical Treatment:		No	Yes	Yes	Yes	No	Yes
Hypertension (Y/N):		Yes	Yes	No	Yes	No	Yes
Under control (Y/N)?		Yes	-	N/A	Yes	N/A	Yes, with some fluctuations
Current or Prior Smoking		No	Yes	Yes	Yes	Yes	Yes
Common Variants$^{\delta}$	***ANK2* p.S646F**	Negative	Negative	Negative	Heterozygous	Heterozygous	Negative
	***KCNQ1* p.V205M**	Heterozygous	Negative	Negative	Negative	Negative	Negative
	***CPT1A* p.P479L**	Heterozygous	Homozygous	Homozygous	Homozygous	Homozygous	Homozygous

Table 3.1 Characteristics of Affected Participants (Continued).

Other Genetic Testing	Negative clinical HCM panel	Negative research WES and clinical TGFBR1/2 panel	Negative research WES	No	No	No

All participants had an affect status, confirmed through acquired medical records. Information that was not collected, unavailable, or unclear for respective fields denoted by a dash (-). * Number of aneurysms reported from the patients' medical documents that were reviewed; however, the locations of all aneurysms and/or their respective sizes were not clear. $^{\delta}$Variants found commonly in the Gitxsan community and these families. Het: Heterozygous; IA: Intracranial Aneurysm; WES: Whole exome sequencing; CT: Computerized tomography; CTA: Computerized tomography angiogram; MRI: Magnetic resonance imaging; MRA: Magnetic resonance angiogram; L-ICA: Left internal carotid artery; R-ICA: Right internal carotid artery; L-MCA: Left middle cerebral artery; R-MCA: Right middle cerebral artery; L-ACA: Left anterior cerebral artery; ACoA: Anterior communicating artery; L-PCoA: Left posterior communicating artery; HCM: Hypertrophic Cardiomyopathy; TGFBR: Transforming Growth Factor Beta Receptor.

3.3 Genomic Variants from WGS

From 6 individuals from the same community, 25 SNV variants passed the filters described from Chapter 2. After implementing mitochondrial and structural variant filtration strategies, no mitochondrial variants were identified, and one nuclear structural variant was identified. The structural variant, with an unclear size, was an interspersed duplication inversion, from an intronic region of the Rho Guanine Nucleotide Exchange Factor 4 (*ARHGEF4)* gene, in a non-coding region with no genes nearby.

3.3.1 Single Nucleotide Variations and Small Indels

In a previous master's thesis study, WES analysis was completed on FIA02.[171] A cerebral cavernous malformations 2 (*CCM2*) heterozygous missense variant (p.Q11R) was found to be the top candidate in this individual from the master's thesis, but it was not shared by any other participant in this study. In addition to investigating the genomic contributions in the FIA02 participant, Hitchcock compared heterozygous rare variants between five affected individuals from five families. Found in four of the five families, 38 variants were found affecting four genes. Each variant was examined in these families for allele sharing. Only one variant was shared between 3 or more of our participants. The dynein axonemal heavy chain 1 (*DNAH1*) p.S1728C variant was heterozygous in five individuals and homozygous in FIA04. Hitchcock identified this gene as a relatively poor candidate functionally, citing information on its role in sperm flagellum function. This variant did not pass the filtration steps in this study due to the gnomAD Exome MAF in the Latino subpopulation (MAF = 0.021632).

3.3.1.1 Allele Sharing Between All Participants

Whole genome sequencing, filtered following the methods in Chapter 2, produced 11 variants with a MAF less than or equal to 1% in gnomAD global and subpopulations and shared between all six participants. Each variant was evaluated using comprehensive annotations and a literature which resulted in a prioritized list of three variants in the *HEMK1*, *CPT1A*, and *LOC105371356* genes (Table 3.2A).

HEMK1 (HemK Methyltransferase Family Member 1) is expressed throughout the body and is responsible for the methylation of the conserved GGQ motif of the mitochondrial translation release factor (*MTRF1L*; OMIM:613542). OMIM does not identify *HEMK1* as disease causing but a reduced *HEMK1* gene expression has been associated with coronary artery disease.[172] Except for SIFT, *in silico* prediction algorithms scored the p.V68M variant on exon 2 with a more likely damaging effect: CADD: 25.0, SIFT: Tolerated, Polyphen2 (HVar): Possibly damaging, MutationTaster: Damaging, FATHMM: Damaging, GERP++: 4.84. The presentation of FIA in these families suggests autosomal dominant or pseudo-dominant inheritance. The zygosity (heterozygous) fits the expected autosomal dominant inheritance, and shows complete allele sharing.

The carnitine palmitoyltransferase 1A (*CPT1A*) encodes for an enzyme found on the outer membrane of the mitochondria is involved in long chain fatty acid oxidation in multiple tissues.[173] Bi-allelic *CPT1A* pathogenic variants cause CPTIA deficiency (OMIM: 255120) which can manifest as metabolic decompensation with hypoketotic hypoglycemia episodes. *CPT1A* p.P479L homozygosity has been associated with infant sudden death in Indigenous communities.[174–176] ClinVar[177] has eight submissions for this variant; six providers interpreted this variant pathogenic, one as having uncertain significance interpretation, and the last classified

it as benign. This variant changes a residue in the cytoplasmic region of the protein (R123-K773). Most *in silico* prediction algorithms predicted this variant as deleterious: CADD: 24.6, SIFT: Probably damaging, Polyphen2 (HVar): Damaging, MutationTaster: Damaging, FATHMM: missing, GERP++ 5.46. GnomAD v2.1.1 lists this variant as having ~$3.18x10^{-05}$ aggregate frequency; it was found in low frequency in the Latino and 'Other' populations with a frequency of $5.781x10^{-05}$ and $9.77x10^{-04}$, respectively. Opposite of the *HEMK1* heterozygous variant, this variant was found with almost complete homozygosity in these participants, except for FIA01 who was heterozygous. This variant fits with the possibility of a pseudo-dominant inheritance but requires significant scrutiny because the community is not well represented in gnomAD, and therefore our population frequency filter erroneously included it.

The *LOC105371356* splice site acceptor variant occurs in a long non-coding RNA found on 16q23.2 with 6 exons. This long non-coding RNA is not well characterized, and little information is known about its function. However, aspects of this variant including the MAF, sequence ontology, and an allele sharing and zygosity pattern that make this an interesting candidate for further research. All participants are heterozygous, with the exception of FIA04 who is homozygous. *In silico* prediction algorithms did not provide much information, CADD was the only algorithm that predicted the variant to be less likely to be present in the human genome (14.0) while all others including the splice site prediction algorithms did not provide predictions. This variant is a candidate for further research but without any functional information about this gene it is difficult to predict its biological significance in association with aneurysms.

Table 3.3 - SNV and Small Indel Top Candidates

Gene Name	Chr: Pos	dbSNP ID	Sequence Ontology‡	Ref/Alt Allele	Zygosity (Hetero/ Homo)	Protein Change	gnomAD Genome global MAF	OMIM	CADD
A. Shared Between 6/6									
HEMK1	3:50608737	rs762549255	Missense	G/A	6/0	p.V68M	$2.102x10^{-5}$	618609	25.0
CPT1A	11:68548130	rs80356779	Missense	G/A	1/5	p.P479L	$3.181x10^{-5}$	600528	24.6
LOC105371356	16.79748711	rs140521230	Splice Site Acceptor	G/C	5/1	-	0.002464	-	14.0
B. Shared Between 5/6									
TUSC3	8:15397767	-	5′ UTR Premature Start Codon Gain	C/A	5/0	p.Met1ext-57	-	601385	16.1
PLCB3	11:64031553	rs767552332	Missense	G/A	4/1	p.R874Q	-	600230	15.5
DKK3	11:12030459	rs140546479	5′ UTR Premature Start Codon Gain	C/A	4/1	p.Met1ext-23	0.007375	605416	8.03
KRT8	12:53300570	rs56925648	Intron	G/A	2/3	-	0.0003828	148060	6.98

‡ Sequence Ontology information collected from VarSeq annotations NCBI RefSeq Genes 105.20201022 v2 and https://www.ncbi.nlm.nih.gov/snp/. Hetero: heterozygous count; Homo: homozygous count; Chr: Pos: Chromosome and position; OMIM: Online Mendelian Inheritance in Man (https://www.omim.org/); CADD: Combined Annotation Dependent Depletion (https://cadd.gs.washington.edu/snv); HEMK: HemK methyltransferase; CPT: Carnitine palmitoyltransferase; LOC: Long non-coding RNA; PCSK: Proprotein convertase subtilisin/kexin; TUSC: Tumor suppressor candidate; PLC: Phospholipase C; KRT: Keratin; GPRIN: G-protein regulated inducer of neurite outgrowth.

3.3.1.2 Allele Sharing Between Five of Six Participants

Whole genome sequencing, filtered following the methods in Chapter 2, produced 14 variants shared between fix of six participants. Following the methods in Chapter 2, each variant was evaluated using comprehensive annotations and a literature search of each variant and the genes they were near or harboured on which resulted in a prioritized list of four variants in the *TUSC3*, *PLCB3*, *DKK3*, and *KRT8* (Table 3.2B).

TUSC3 (tumor suppressor candidate 3) localizes to the endoplasmic reticulum and has been associated with several processes including magnesium uptake, protein glycosylation and embryonic development. This gene is associated with an autosomal recessive form of non-syndromic mental retardation[178–181] (OMIM: 611093) but its phenotype does not have overlap with familial intracranial aneurysms. It does not appear to have any direct association with aneurysm formation, but the variant creates a new translation initiation site 57 nucleotides upstream, has a CADD score of 16.14, and the zygosity of the allele makes it a candidate.

PLCB3 (Phospholipase C Beta-3) spans approximately 17kb tightly packed with 31 exons and encodes for a phosphoinositide phospholipase C beta enzyme family member. This gene is associated with autosomal recessive spondylometaphyseal dysplasia with corneal dystrophy[182] (OMIM: 618961) but has no phenotypic overlap with intracranial aneurysms. With a CADD score above 10, this variant is in the top 10% of possible variants less likely to be found in the human genome. However, other prediction algorithms not only identified this variant as not damaging but SIFT, Polyphen2, MutationTaster and FATHMM identified this variant as tolerated, neutral, or benign.

DKK3 (Dickkopf WNT Signaling Pathway Inhibitor 3), with no reported disease associations and has roles in embryonic development via the WNT signaling pathway. Like the

TUSC3 variant, this variant is another 5′ UTR premature start codon gain where translation begins 23 nucleotides upstream. The variant is not reported in gnomAD exome global or any subpopulation but was found in gnomAD genome global with a relatively high MAF of 0.00737. Aside from the 8.027 CADD score provided, no other *in silico* prediction algorithms provided scores.

Keratin 8 (*KRT8*) is a member of the type II keratin family which forms intermediate sized filaments in epithelial cells. Pathogenic variants in *KRT8* cause cryptogenic cirrhosis and susceptibility to noncryptogenic cirrhosis (OMIM: 215600).[183,184] In this region of the genome (12q13.3), there is a cluster of other keratin genes that are classified in the type II keratin family.[185] *In silico* prediction algorithms provided conflicting reports with a CADD score of 6.98, SIFT as damaging, Polyphen2 (HVar) missing, MutationTaster with tolerated, FATHMM as damaging, and GERP++ with 2.49. This intronic variation was found in both the gnomAD genome and exome global MAF databases, with the highest subpopulation MAF of 0.00431280 in East Asians.

3.4 A Search for Potential Modifiers and GWAS Loci

Risk alleles identified in GWA studies with the Nunavik Inuit[139,140] were not shared in these families. Only two previously identified risks alleles from other populations were shared between five of six participants, while no risk alleles were shared between all participants (Table 3.3). Neither of the two reported risk alleles were found in FIA04 who had four intracranial aneurysms. FIA03, having four intracranial aneurysms, was homozygous for the T risk allele (rs9315204) and heterozygous for the A risk allele (rs11661542) possessing a total of three risk alleles. FIA01, having one intracranial aneurysm, was heterozygous for the rs9315204 risk allele

and homozygous for the rs11661542 risk allele, possessing a total of three risk alleles. The remaining participants FIA02, FIA05, FIA06 were heterozygous for both risk alleles, possessing a total of two risk alleles.

Table 3.4 - Modifier Risk Allele Sharing

Chr and Loci	Position	dbSNP ID	Suggested Genes[δ]	Ref/Alt Allele	RA	gnomAD Global/ E. Asian MAF	Context	OR	P-Value	RA Zygosity (Het/Hom)	Reference
13q13.1	33693837	rs9315204	KL, STARD13	C/T	T	0.2072/ 0.2349	Intronic	1.2	$2.5x10^{-9}$	4/1[ǂ†]	Yasuno *et al.* 2010[76]
18q11.2	20223695	rs11661542	-	C/A	A	0.5991/ 0.5289	Intergenic	1.22	$3.17x10^{-17}$	4/1[ǂ‡]	Bakker *et al.* 2020

Information not provided by corresponding study indicated by hyphen (-). δ Genes suggested by corresponding study. ǂ Risk alleles were missing in FIA04. † Heterozygous in all, except FIA03 who was homozygous. ‡ Heterozygous in all, except FIA01 who was homozygous. Yasuno *et al.* 2010 used European and Japanese populations; Bakker *et al.* 2020 used European, Chinese, and Japanese populations in their study. RA: Risk Allele; Het: heterozygous; Hom: homozygous; KL: Klotho; STARD13: StAR related lipid transfer domain containing 13.

Chapter 4: Discussion

Previous studies have documented high rates of familial IA and/or SAH in Northern Indigenous populations,[137,138] and our study adds to this literature. Over 800 participants enrolled in the LQTS study, and of those 22 were documented to have IA and/or SAH with another 18 family members, were reported to have one or more IA. Although, the IA prevalence in the LQTS cohort was ~2.75% which is close to the general IA prevalence (3.2%)[26] most of the 22 affected individuals were related through the three families. If including the 18 non-enrolled family members, the prevalence is ~5% which remains in the upper end of the 95% confidence interval of the general IA prevalence (3.2%, 95% CI = 1.9-5.9).[26] Sporadic cases, without a first or 2nd degree relative have not been reported in this study.

This study corroborated the findings from the previous retrospective studies[137,138] that certain Northern Indigenous communities have higher rates of familial IA/SAH in comparison to other non-Indigenous populations. Also, Lingdaard *et al.* (2003)[138] presented 12 family pedigrees going back three generations at most. Understanding that some Indigenous communities are relatively more homogeneous than many non-Indigenous populations, it is important to collect a detailed family history going back several generations to identify if these families are connected, like the families we have presented in this study.

4.1 Characteristics of Affected Individuals

The female to male affected ratio from the 23 affected participants enrolled in the LQTS and the FIA study is 2.28, similar to the female to male ratio found in the broader intracranial aneurysm population (2.2, 95% CI=1.3-3.6).[26] The average age of the participants diagnosed

with 1 or more intracranial aneurysms was three years younger than the general population average of approximately 50 years old.[26]

When screening first-degree relatives with a positive family history of intracranial aneurysms who smoked, had hypertension, or both, Broderick *et al.* (2009)[58] found intracranial aneurysms in 20.6% of the participants which is higher than the overall prevalence of intracranial aneurysms (~9%).[58,65] Smoking history was only collected from the six FIA study participants, five of whom were current or past smokers. From unpublished data, 74.6% out of 224 Gitxsan community members who participated in the Alliance study[186] were current or past smokers. Hypertension data was collected from both the FIA and LQT study, with over 50% of LQTS and FIA study participants with intracranial aneurysms having hypertension. Although these risk factors may have affected the number intracranial aneurysm cases in these families, the number of aggregate intracranial aneurysm cases found through the LQTS and this FIA study goes beyond what we would expect from either of these modifiable risk factors alone or together.

4.2 Genomic Variations

4.2.1 Structural Variation

The number of structural variations per genome has been reported as low as 4,442 SVs per genome[187], accounting for approximately 0.1% of variants, and as high as 27,622.[188] The interspersed duplication inversion with non-coding breakpoints was a duplication from an intronic region of the *ARHGEF4* gene which was then inverted and inserted into an intergenic region with no genes near it. Additionally, the size of the structural variation was unclear. Due to the structural variant's location and unclear size, this variant cannot be analyzed further. In future

research on FIA in these families, this structural variation should be investigated further to determine its size and variant sharing.

4.2.2 Candidate Single Nucleotide Variants

4.2.2.1 HemK Methyltransferase Family Member 1

Working with HeLa cell, Ishizawa *et al.* (2008)[189] completed a knockdown of *HEMK1* which resulted in loss of methylation of the glutamine residue on the GGQ motif of *MTRF1L* and a decreased mitochondrial translational activity. Knockdown of *MTRF1L* in HeLa cells has shown an increased production of reactive oxygen species (ROS) from the mitochondria.[190] Reactive oxygen species are mediators in inflammation which is already a documented process in intracranial aneurysms.[24,191] However, these observations do not provide strong evidence to substantially support the significance of this variant in the formation of aneurysms.

Information collected from UniProt[192] identifies this variant in an alpha-helix section of the protein that is not part of a functional domain. Missense constraint information, collected from DECIPHER,[193] identifies the region surrounding the variant as being more likely to tolerate missense variants. CADD (Combined Annotation Dependant Depletion) score studies have shown that most variants identified in the NCBI ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/) are in the top 1% of variant less likely to be found in the human genome, meaning they had a score greater than or equal to 20. With a CADD score of 25.0, being heterozygous in all affected individuals, and shared between all affected individuals, this variant seems like a strong candidate from this study. A key next step to evaluate this variant is to access an accurate minor allele frequency in this population in case the variant frequency is far too high to reasonably be a highly penetrant risk factor. Hopefully the Silent Genomes Project

(https://www.bcchr.ca/silent-genomes-project) and other variome projects will succeed in advancing inclusivity and equity in genomics with Indigenous engagement and leadership. Although the *HEMK1* p.V68M variant may be considered one of the strongest candidates per annotated information, the biological relevance to aneurysms is unclear and therefore cannot be considered a strong candidate to cause of this condition in these families.

4.2.2.2 Carnitine Palmitoyltransferase 1A

Other isoforms of the CPT1 protein are the heart and muscle *CPT1B* isoform and *CPT1C* brain isoform which is more closely related to the *CPT1A* isoform than the *CPT1B* isoform is.[194] Although the *CPT1A* p.P479L variant has fairly consistent damaging predictions by *in silico* algorithms and has a rare allele frequency documented in gnomAD making it a relatively stronger candidate, subpopulation minor allele frequencies not well-represented in gnomAD, make this candidate less likely to cause FIA with Mendelian pseudo-dominant inheritance.

The ACMG guidelines (2015)[195] for variant interpretation state that a population MAF is greater than 5% is considered a stand-alone indication of a benign variant. GnomAD identifies this variation as rare (MAF = 3.181×10^{-5}), but it is known to be common in multiple Northern Indigenous populations including Alaskan Natives, First Nations, Inuit, and Greenlandic Inuit.[175,196–198] From unpublished genotyping of 719 Gitxsan community members, 26.4% were homozygous and 49.4% were heterozygous for the p.P479L variant (leucine (L) variant frequency = 0.5141; proline (P) variant frequency = 0.4859). The *CPT1A* p.P479L variant appears to have Hardy-Weinberg equilibrium in this population cohort, which is evidence that the study population for the LQTS study did not have distinctly stratified subpopulations. Considering the ACMG guidelines (2015), due to its high prevalence in Northern Indigenous

communities, so high it has been dubbed the 'Arctic Variant', this variant is less likely to cause the Mendelian form of FIA in these families. This variant is more likely a filtration artifact due to underrepresentation of Indigenous peoples in variant databases. In two retrospective studies,[137,138] a higher prevalence of FIA in the Greenlandic Inuit populations was noted possibly suggesting a variant in common may be increasing genetic susceptibility to intracranial aneurysms. The *CPT1A* p.P479L variant is common in all three populations (West Coast First Nations,[175] Greenlandic Inuit,[197] and Nunavik Inuit[198]) and they share a common ancestry; it is possible other shared genetic risk factors for IA have not yet been discovered.

4.2.2.3 Phospholipase C Beta-3

Vascular smooth muscle cell (VSMC) induced vascular permeability is a documented biomarker of aneurysm formation and predictor of rupture.[199,200] *PLCB3* was observed to be a negative mediator of heat-induced VSMC-mediated vascular permeability in transgenic zebrafish, increasing vascular permeability more than heat-induced VSMC-mediated did alone.[201] However, when zebrafish were injected with the translation blocking morpholino without heating treatments, the change to vascular permeability was negligible when compared with the vascular permeability of controls (no *PLCB3* translation blocking morpholino or heat treatments). This work provided evidence that *PLCB3* is involved in processes related to aneurysm formation.

Vascular endothelial growth factor (VEGF) is known to have significant influence on angiogenesis, which is the process for developing new blood vessels, and vasculogenesis, which is the process for developing embryonic blood vessels.[202] Bhattacharya *et al.* (2009)[203] found when HUVECs with functional *PLCB3* were exposed to VEGF there was a significant migration

response (smaller tortuosity, fewer directional changes, higher velocity). However, in *PLCB3*-knockdown cells, there were no significant changes in migration response when exposed to VEGF, suggesting that *PLCB3* has an important role in the migration of endothelial cells. Interestingly, when assessing the impact of *PLCB3* on endothelial cell proliferation, they found that only *PLCB3*-knockdown in HUVEC cells without VEGF treatment resulted in a significant increase ($p=0.0001$) in proliferation when compared with the control, suggesting a negative regulator of endothelial cell proliferation. Affecting the efficacy of the subsequent enzymes produced from *PLCB3* may have an impact on vascular permeability, endothelial cell migration, and cell proliferation.

The amino acid change from arginine (R) with an electrically charged side chain, to a glutamine (Q) with a polar neutral side chain may influence the subsequent folding of the protein or the stability of the natural folding state. Like the *HEMK1* missense variant, this *PLCB3* variant is found in an alpha-helix region with no known functional domain significance and high tolerance to missense variants from UniProt and DECIPHER, respectively. Understanding that this variant exists in region of the gene with no known functional significance in the subsequent protein weakens the likelihood that variant is significant enough to cause a Mendelian condition such as FIA. Additionally, while this variant has a CADD score of 15.5 and no available gnomAD genome global allele frequency, other *in silico* algorithms do not predict this variant to be damaging. Having a CADD score above 10 means that the variant is in the top 10% of variants less likely to exist in the human genome. It was found in these participants with variable zygosity, missing in FIA06, homozygous in FIA05, and all others heterozygous. Though this variant an interesting candidate, it is most likely not the cause of FIA in these families.

4.2.2.4 Dickkopf WNT Signaling Pathway Inhibitor 3

DKK3 has been shown to be involved in cardiac remodeling,[204] and angiogenesis.[205] Using immunocytochemical and western blot analysis, Busceti *et al.* (2017)[205] began by demonstrating that when HUVECs were exposed to human recombinant Dkk3 there was a significant increase in mRNA VEGF levels. Once they showed an increase in mRNA VEGF, they investigated the possible effect on angiogenesis by examining the changes in endothelial cell proliferation in tubes. When the HUVECs were exposed to 10ng/mL of human recombinant Dkk3 for 6 and 18h there was a significant increase in cell proliferation in comparison to the controls. Like the *TUSC3* variant, the *DKK3* variant was a 5′ UTR premature start codon gain variant. The gnomAD global frequency (0.00737) was approaching the MAF filter of less than or equal to 1% and the CADD score was less than 10, both factors making this variant less likely to explain the inheritance pattern seen in these families when working under the hypothesis of a significant single genetic effect. However, the functional information demonstrates that affecting this gene influences angiogenesis and therefore is of interest for further studies.

4.2.2.5 Keratin 8

The *KRT8* intron variant had a weaker CADD score of 6.98, did not show complete allele sharing and it has inconsistent zygosity across the participants which does not follow the excepted zygosity of heterozygous for an autosomal dominant or homozygous for a pseudo-dominant inheritance pattern. Splice regions are located in the intronic regions, but this variant did not have any splice scores from either dbscSNV Splice Altering Predictions (ADA or RF Scores) or SpliceAI.[206] Additionally, there are no reported associations between *KRT8* and

aneurysm-related processes. Relative to the other variants, this intron variant identified in the *KRT8* gene is a poor candidate.

4.3 Potential Modifier Alleles

In a GWA study of IA in Finnish, Dutch, and Japanese populations, Bilguvar *et al.* (2008)[71] identified two new loci on 2q and 8q and corroborated previously identified 9p loci with significant association. They analyzed the cumulative effects of the risk alleles identified in their study on intracranial aneurysms using a simple logistic model. They found an additive effect of multiple risk alleles in each population, with the OR score increasing within each population as the number of risk alleles increased. To investigate the possible additive effect of multiple risk alleles identified in previous GWA studies, risk alleles shared between 5 or more participants were reported.

These two risk alleles are in two different loci. The risk allele at rs9315204 is in the intronic region of *STARD13* (StAR-related lipid transfer domain containing 13; also known as DLC2) and nearby *KL* (klotho). In DLC2-knockout mice, it has been observed that when exposed to Matrigel, an extraction from the Engelbreth-Holm-Swarm mouse sarcoma, that there was an increase in angiogenic responses from endothelial cells *in vivo* and *ex vivo*.[207] As a result, they suggested that DLC2 may regulate angiogenic responses. The other gene close to rs20223695 risk allele was *KL*. Highly expressed in the brain, *KL* has also been associated with angiogenesis.[208] The risk allele rs20223695 is in an intergenic region with no genes nearby; therefore, it is difficult to comment on its effect on aneurysm processes further.

Though these risk alleles were found in most of the participants, the number of risk alleles does not obviously appear to correlate with the phenotypes recorded. FIA02, FIA03, and

FIA04 had 12 intracranial aneurysms altogether, each with four aneurysms, but had different numbers of risk alleles, two, three and zero, respectively. The remaining participants, FIA01, FIA05, and FIA06, who each had one intracranial aneurysm had a total of three, two, and two risk alleles, respectively. Although there are five participants from these families that have at least two risk alleles, it was observed here was varying risk allele counts with FIA04 who had four aneurysms with zero risk alleles and FIA01 with one aneurysm who had three risk alleles. Therefore, in these participants risk allele counts do not obviously represent the likelihood of developing 1 or more intracranial aneurysms.

4.4 Summary of Findings

The findings presented in Chapter 3 are the preliminary findings of on-going research with these families and the Gitxsan community. After filtration and analysis, no variant found could be clearly associated with strong enough evidence to explain the aggregation of intracranial aneurysms in these families. The strongest variant including both annotations and biological information was the *PLCB3* p.R874Q missense variant, which was shared between most affected participants and had a CADD score of 15.5. However, this variant or the other variants examined were not compelling enough to explain the inheritance of FIA in these families.

These findings are not surprising as many families have participated in next-generation sequencing studies and little progress has been made in our understanding of the genetics behind FIA (Section 1.4.1.4). Additionally, another type of familial aneurysm, familial thoracic aortic aneurysms and dissections (FTAAD), has found some cases may be caused by the disruptive variants in the *ACTA2* gene which accounts for 12-21% of FTAAD cases.[209] To date, the

strongest FIA-gene association made was the discovery of the *THSD1* gene which from 507 IA probands, only 1.6% had heterozygous missense variants. Additionally, the genetic mechanism of aneurysm-related processes like angiogenesis and vasculogenesis are extremely complex with hundreds of genes associated with those processes. Finding connections between some of the identified genes in these families and angiogenesis was expected by chance alone.

4.5 Strengths

We used whole genome sequencing to assess variants found in these affected individuals. Whole genome sequencing has advantages over other next-generation sequencing technologies as it allows us to interrogate intergenic, intronic, and exon regions, other non-coding biological elements, mitochondrial DNA, copy-number variations (CNV), and structural variations. Even though we only interrogated genetic variants more likely to have an effect on protein coding, using WGS provided more information than the more often used WES.

Whole genome sequencing provides us with a significant amount of data to process, which is where WGS analysis becomes difficult. To aid in removing less likely variants causing this condition, we aimed to include the most distantly related individuals possible, with first cousins being the closest relation in the study, to examine the genome. First cousins should only share ~12.5% of the genome with one another. If this condition in these families was caused by one significant variant, having multiple more distantly related individuals would have helped remove rare variants that were unrelated to this condition in the filtration steps. Unfortunately, this allele-sharing approach also means that variants common to the broader community not indexed in a reference database will also appear.

The sample size of the cohort was relatively smaller when compared with other types of IA genetic studies, with only 6 affected individuals enrolled. Having a larger sample size would provide evidence of much stronger candidates if allele-sharing was strong as well; however, for a WGS study having 6 affected distantly related relatives is arguably acceptable, as filtering produced only a few variants that required consideration. Distantly related relatives should share only a small fraction of their genomes with one another. When investigating for a monogenic origin to a seemingly Mendelian condition, distantly related relatives would assist in removing insignificant sections of the genome and many variants, condensing massive amounts of genetic data to a few variants of significance. Having unaffected participants enrolled for FIA genetic studies can both be an advantage and a disadvantage. As previously stated, FIA is a multifactorial condition with both genetic and environmental factors that contribute to its development and progression. Intracranial aneurysms also form later in life which affects the concept of 'unaffected' in FIA studies. For example, if a 45-year-old female undergoes an MRA and no IA is found, she may be unaffected at that moment, but it does not preclude her from developing one in her life. An advantage of having unaffected participants is being able to assess co-segregation, and potentially remove numerous shared regions that arise from the potential distant relatedness between families but choosing unaffected individuals must be done with strictest of protocols to ensure that they are less likely to become affected due to significant genetic factors in the future. However, when investigating a Mendelian condition with reduced penetrance, some unaffected individuals may carry the variant without having the phenotype.

4.6 Limitations

In March 2020, the COVID-19 pandemic hit Canada,[210] delaying aspects of research or halting them entirely. When the first wave of COVID-19 occurred, our enrollment process and sample collection was delayed until protocols for safe enrollment and sample collection during a global epidemic were in place. This was a limitation to this study because it greatly slowed down research progress and community engagement was delayed.

A limitation to this study was how we enrolled the FIA participants from the pre-existing Gitxsan community LQTS study. In the LQTS, the data originally collected from the participants was limited in the context of intracranial aneurysms because the study focused on characteristics relevant to long QT syndrome. However, this limitation was mitigated by enrolling the FIA participants through the Silent Genomes Activity 2 (precision diagnosis) project and collecting data (including medical records and family history) specific to cerebral aneurysms.

One limitation in the sample size occurred when investigating structural variants. We only used the two available blood samples from our study to investigate these types of variants. Buccal and saliva samples were excluded from this part of the analysis due to a higher sensitivity of CNV detection in blood samples and higher false-positive rates in oral samples.[145] Though this exclusion was intentional it still reduced the number of samples available for analysis. However, even with only two samples only one structural variant with little biological significance was identified.

There are multiple ways to analyse genetic data and in WGS studies there are vast amounts of data that need to be processed and analysed. Using distantly related participants affected with intracranial aneurysms was a strength for filtering out variants more likely to be unrelated to this Mendelian condition. However, even with distantly related relatives, strict filters

were required to yield the most likely genetic causes of this Mendelian condition and reduce the number of variants to several top candidates. Also, a conscious choice was made to only analyze variants more likely to affect protein coding. These stricter filters were chosen for the purpose of exploring the possibility of a rare monogenic origin, like the *THSD1* p.R450X variant.[46] None the less, this filtration approach innately eliminates the analysis of SNV and small indel variants that are found more commonly (greater than 1%) and the analysis of SNV and small indels less likely to have protein coding effects.

In addition to missing SNV and small indel variants with non-protein coding effect, repeat expansions, mid-range CNVs, and "darker" parts of the genome were not assessed. Huntington's disease is a well-known example of a disease causing triplet repeat expansion but there are many other diseases that are associated with various types of repeat expansions.[211] The size of CNVs can range from less than 50 bp to several megabases and CNVs have been associated with multiple neurodevelopmental and psychiatric disorders.[212–214] Parts of the genome that cannot be properly assembled or aligned using short-read sequencing, like the one we used in this study, are known as "darker" parts of the genome.[215] Missing these type variants or variants from "darker" regions of the genome is another analytical limitation to the study as more aspects of the genome are missed.

In our opinion, the largest limitation is the underrepresentation of Indigenous peoples in Canada from genomic databases such as gnomAD. This limitation in such databases becomes a limitation of any next-generation sequencing study on Indigenous populations because the researchers can confirm if a variant is common in the global or subpopulations but not the population that they are studying. This limitation became very evident when examining the *CPT1A* p.P479L variant; rare in the global and subpopulation from gnomAD and common in the

Gitxsan community and many other Arctic Indigenous communities. This observation brings to question the strength of the remaining variants that were reported as unreported, rare, or uncommon by gnomAD. In the search for genetic causes of a Mendelian condition, the American College of Medical Genetics (ACMG) guidelines state that a high population frequency is stand-alone evidence of a benign variant. Without proper representation of Indigenous peoples in such databases a confident and accurate interpretation of genetic variants is limited.

4.7 Future Directions

4.7.1 Linkage Analysis

Since the variants identified in this study were not compelling enough to move immediately towards animal model studies, more information in necessary. The filtration strategies we used in this study were stricter, as we only examined variants more likely to affect protein-coding regions with a gnomAD MAF of less than or equal to 1%. Collecting saliva samples from a large portion of these families, completing a SNP array, and completing linkage analysis and linkage disequilibrium statistical analysis would provide tremendous assistance in narrowing down regions that require comprehensive investigation. Prior to discovery of the *THSD1*'s role in intracranial aneurysms, linkage analysis on the family narrowed down the regions that required interrogation (13q13.3 showed genome-wide significance; 5q22.2-33.3, 9p23, 12p12, and 18p11 showed suggestive significance).[97]

In our study, the enrollment criteria were looser having no requirements for age of diagnosis, number of IA, and/or number of subarachnoid hemorrhages. Future work with these families should have enrollment criteria that maximizes the likelihood of affected participants having a larger genetic effect and minimizes the likelihood of unaffected participants sharing

possible variants that contribute to this condition. To complete this, affected individuals would be family members who, from the literature, have the least likely chance of developing an aneurysm; a younger (less than 40) individual who does not smoke and/or have hypertension and has a positive diagnosis of an intracranial aneurysm, multiple intracranial aneurysms, and/or has had one or more subarachnoid hemorrhages. While minimizing the likelihood of a participant having the variants of significance future work would sequence unaffected individuals which would be family members of advanced age (greater than 60) who smoke and/or have hypertension and has had negative screen for intracranial aneurysms.

4.7.2 Animal Models

Though animal model studies should not be completed immediately after this study, this type of research would be a long-term goal. Animal models such as zebrafish have been used in the study of intracranial aneurysms to identify genes associated with its development.[46,117]

Zebrafish are useful in the study of the genetics of intracranial aneurysms because they are nearly transparent in the embryo, allowing researchers to see intracerebral hemorrhages that occur,[216] and the genetic homology between zebrafish and humans is remarkably similar.[217,218] Additionally, in the case of FIA, zebrafish are a proven example of associating genes with the intracerebral hemorrhage phenotype.[46,117] Using morpholino oligomers (MO), control gene-knockouts to test the association of a non-functional gene to intracranial aneurysms or subarachnoid hemorrhages can be completed.[46,117] Subsequently, mRNA rescue studies can be used to further assess not only the importance of the gene but also the effects of the human variant in question. By co-injecting the MO with the wildtype human transcript (mRNA) the phenotype should return to the normal (control) phenotype.[117] Additionally, by co-injecting the

MO with the human mutated transcript, the researcher would be able to assess the effects of the variant in the zebrafish animal model.[117]

4.7.3 Functional Studies

When animal models show an association between a particular gene and IA, it is important to assess the functional significance of the gene in IA formation. Santiago-Sim *et al.* (2016)[46] performed functional studies on human umbilical vein endothelial cells (HUVECs). To identify protein-protein interactions, they performed a co-immunoprecipitation which identified the talin protein interacting with the wildtype *THSD1*. They also performed a knockdown of *THSD1* using siRNA (small interfering RNA) which caused a significant loss of relative cell adhesion. Subsequently, rescue studies using the wildtype human mRNA restored the relative cell adhesion while the mutated human-derived vector did not restore the relative cell adhesion. This method can assess functional importance should be used when testing other genes for association with IA as this method has already been used to show shown an IA-gene association.

4.8 Significance of Research

This study may have not found one or more variants to explain the genetic cause of this condition, but to our knowledge, this study is the first to thoroughly describe and investigate the genomic contributions to FIA in First Nation families. We may be the first because of the lack of representation of Indigenous peoples in genomic databases or because of the lack of access of leading-edge genomic technologies for Indigenous peoples, but this study is the start to a long and continuous relationship and research partnership with the Gitxsan community and these families. Additionally, the topic of this thesis was identified and prioritized by the Gitxsan

community. By starting this research, we are facilitating what the community wants and enabling them to exercise their right to self-determination and sovereignty over research. This is one of many important components when researching in-partnership with Indigenous communities.[141–143]

Understanding the genetic contributions, the penetrance, and the variable expressivity is critical in understanding a piece to a likely larger molecular mechanism. As the genetic studies reviewed in Section 1.4.1 have shown, the genetic architecture of intracranial aneurysms and familial intracranial aneurysms are complex, showing associations with multiple genomic regions and with multiple genes. However, deciphering the complex genetic aetiology of familial intracranial aneurysms will provide invaluable information for possible therapeutic medicine targets and genetic screening of FIA susceptible family members.

4.8.1 Familial Serial Screening

The most optimal outcome of this study would have been finding a significant variant to explain the cause of FIA in these families. Understanding the genetic mechanism of familial intracranial aneurysms would help with screening of first-degree family members in these families and possibly unrelated families. At the moment, recommendation for familial serial screening by MRA or CTA of asymptomatic relatives only comes after two or more family members have been diagnosed with an intracranial aneurysm or a subarachnoid hemorrhage.[61] In the 2021 community update letter (Appendix B), we have recommendations for serial screening that take into consideration the familial clustering seen in these Gitxsan families. We recommend that serial screening, by CTA or MRA, occurs every 5 years, starting at age 18 years for asymptomatic Gitxsan community members with one first-degree relative with a confirmed

intracranial aneurysm. Comparing our recommendations for serial screening to the recommendations from the American Heart/Stroke Association (AHA, 2015)[61], we provided a definitive age to begin serial screening for unruptured intracranial aneurysms and a definitive period of time between screenings, while the AHA recommendations did not. Additionally, we recommended serial screening for community members with at least one first-degree relative lowering the number of affected family members required as for serial screening to begin.

There is clearly a substantial estimated inheritance of subarachnoid hemorrhage;[67] therefore understanding the genetic mechanism of familial intracranial aneurysms may help reduce the number of first-degree relatives that undergo serial screening. Additionally, this understanding would also help some children in these families as their risk of developing an intracranial aneurysm could be assessed well before the risk of developing an intracranial aneurysm increases after the age of 30. Identifying a genetic risk would more strongly justify beginning screening of those individuals early-on; therefore, aid in IA diagnosis and treatment to prevent rupture.

4.8.2 Treatments

Another long-term goal of studies such as this one, would be aid identifying possible pharmaceutical targets. In case of cardiovascular disease, our understanding of the genetic architecture has provided therapeutic targets like *PCSK9* inhibition by Evolobumab which, when given with statins, lowers LDL-C and decreases the risk of cardiovascular events for patients.[219,220] There are a few candidate drug therapies for unruptured intracranial aneurysms including aspirin, statins, calcium channel blockers, and angiotensin II receptor blockers.[221] Understanding the genetic mechanism of familial intracranial aneurysms could identify

molecular targets for pharmaceutical therapies which would help families and the general population affected with intracranial aneurysms.

References

1. Statistics Canada. Aboriginal peoples in Canada : Key results from the 2016 Census. https://www150.statcan.gc.ca/n1/daily-quotidien/171025/dq171025a-eng.htm?indid=14430-1&indgeo=0 (2017). Accessed 01 July 2021.
2. Tjepkema, M., Bushnik, T. & Bougie, E. Life expectancy of First Nations, Métis and Inuit household populations in Canada. *Heal. reports* **30**, 3–10 (2019) doi:10.25318/82-003-x201901200001-eng.
3. Statistics Canada. Health at a Glance: Select health indicators of First Nations people. https://www150.statcan.gc.ca/n1/pub/82-624-x/2013001/article/11763-eng.htm (2013). Accessed: 11 July 2021.
4. Luo, Z.-C. *et al.* Birth outcomes in the Inuit-inhabited areas of Canada. *Can. Med. Assoc. J.* **182**, 235–242 (2010) doi:10.1503/cmaj.082042.
5. Smylie, J., Fell, D. & Ohlsson, A. A Review of Aboriginal Infant Mortality Rates in Canada: Striking and Persistent Aboriginal/Non-Aboriginal Inequities. *Can. J. Public Heal.* **101**, 143–148 (2010) doi:10.1007/BF03404361.
6. Reading, C. & Wien, F. Health Inequalities and Social Determinants of Aboriginal Peoples' Health. *Natl. Collab. Cent. Aborig. Heal.* (2009).
7. First Nations Health Authority. Indigenous Health Improves But Health Status Gap with Other British Columbians Widens. https://www.fnha.ca/about/news-and-events/news/indigenous-health-improves-but-health-status-gap-with-other-british-columbians-widens (2018). Accessed: 11 July 2021.
8. Caron, N. R. *et al.* Indigenous Genomic Databases: Pragmatic Considerations and Cultural Contexts. *Front. Public Heal.* **8**, 1–7 (2020) doi:10.3389/fpubh.2020.00111.

9. Robertson, S. P. *et al.* Genomic medicine must reduce, not compound, health inequities: The case for hauora-enhancing genomic resources for New Zealand. *N. Z. Med. J.* **131**, 81–89 (2018).

10. GenomeCanada. Silent Genomes: Reducing health-care disparities and improving diagnostic success for Indigenous children with genetic disease. https://www.genomecanada.ca/en/silent-genomes-reducing-health-care-disparities-and-improving-diagnostic-success-indigenous-children (2017).

11. Atkins, C. *et al.* Rheumatic disease in the Nuu-Chah-Nulth native Indians of the Pacific Northwest. *J. Rheumatol.* **15**, 684—690 (1988). Accessed: 15 July 2021.

12. Ward, R. H., Frazier, B. L., Dew-Jager, K. & Pääbo, S. Extensive mitochondrial diversity within a single Amerindian tribe. *Proc. Natl. Acad. Sci.* **88**, 8720 LP – 8724 (1991) doi:10.1073/pnas.88.19.8720.

13. Peters, A. A. *et al.* HTLV type I/II in British Columbia Amerindians: A seroprevalence study and sequence characterization of an HTLV type IIa isolate. *AIDS Res. Hum. Retroviruses* **16**, 883–892 (2000) doi:10.1089/08892220050042828.

14. Dalton, R. Tribe blasts 'exploitation' of blood samples. *Nature* **420**, 111 (2002) doi:10.1038/420111a.

15. Arbour, L. & Cook, D. DNA on loan: Issues to consider when carrying out genetic research with aboriginal families and communities. *Community Genet.* **9**, 153–160 (2006) doi:10.1159/000092651.

16. Smith, M. J. *Placing Gitxsan Stories in Text: Returning the Feathers. Guuxs Mak'am Mik'aax*. [Doctoral Thesis] (University of British Columbia, 2004).

17. Huisman, L.-A. *Exploring the Impact of Long Qt Syndrome: Perspectives From a British*

Columbia First Nations Community. [Master's Thesis] (University of British Columbia, 2010).

18. Bene Watts, S. *The Impace of Congenital Long QT Syndrome on First Nations Children and Youth in Northern British Columbia*. [Master's Thesis] (University of Victoria, 2020).
19. Arbour, L. *et al.* A KCNQ1 V205M missense mutation causes a high rate of long QT syndrome in a First Nations community of northern British Columbia: A community-based approach to understanding the impact. *Genet. Med.* **10**, 545–550 (2008) doi:10.1097/GIM.0b013e31817c6b19.
20. Swayne, L. A. *et al.* Novel Variant in the ANK2 Membrane-Binding Domain Is Associated With Ankyrin-B Syndrome and Structural Heart Disease in a First Nations Population With a High Rate of Long QT Syndrome. *Circ. Cardiovasc. Genet.* **10**, 1–11 (2017) doi:10.1161/CIRCGENETICS.116.001537.
21. Wilde, A. A. M., Amin, A. S. & Postema, P. G. Diagnosis, management and therapeutic strategies for congenital long QT syndrome. *Heart* heartjnl-2020-318259 (2021) doi:10.1136/heartjnl-2020-318259 doi:10.1136/heartjnl-2020-318259.
22. Schwartz, P. J. *et al.* Prevalence of the congenital long-qt syndrome. *Circulation* **120**, 1761–1767 (2009) doi:10.1161/CIRCULATIONAHA.109.863209.
23. Jackson, H., Huisman, L.-A., Sanatani, S. & Arbour, L. T. Long QT syndrome. *Can. Med. Assoc. J.* **183**, 1272–1275 (2011) doi:10.1503/cmaj.100138.
24. Xu, Z., Rui, Y. N., Hagan, J. P. & Kim, D. H. Intracranial Aneurysms: Pathology, Genetics, and Molecular Mechanisms. *NeuroMolecular Med.* **21**, 325–343 (2019) doi:10.1007/s12017-019-08537-7.
25. Hitchcock, E. & Gibson, W. T. A Review of the Genetics of Intracranial Berry Aneurysms

and Implications for Genetic Counseling. *J. Genet. Couns.* **26**, 21–31 (2017) doi:10.1007/s10897-016-0029-8.

26. Vlak, M. H. M., Algra, A., Brandenburg, R. & Rinkel, G. J. E. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: A systematic review and meta-analysis. *Lancet Neurol.* **10**, 626–636 (2011) doi:10.1016/S1474-4422(11)70109-0.

27. Park, S. H., Yim, M. Bin, Lee, C. Y., Kim, E. & Son, E. I. Intracranial fusiform aneurysms: It's pathogenesis, clinical characteristics and managements. *J. Korean Neurosurg. Soc.* **44**, 116–123 (2008) doi:10.3340/jkns.2008.44.3.116.

28. Withers, K., Carolan-Rees, G. & Dale, M. Pipeline^{TM} embolization device for the treatment of complex intracranial aneurysms: A NICE medical technology guidance. *Appl. Health Econ. Health Policy* **11**, 5–13 (2013) doi:10.1007/s40258-012-0005-x.

29. Keedy, A. An overview of intracranial aneurysms. *McGill J. Med.* **9**, 141–146 (2006).

30. Toth, G. & Cerejo, R. Intracranial aneurysms: Review of current science and management. *Vasc. Med.* **23**, 276–288 (2018) doi:10.1177/1358863X18754693.

31. Tromp, G., Weinsheimer, S., Ronkainen, A. & Kuivaniemi, H. Molecular basis and genetic predisposition to intracranial aneurysm. *Ann. Med.* **46**, 597–606 (2014) doi:10.3109/07853890.2014.949299.

32. Ohashi, Y., Horikoshi, T., Sugita, M., Yagishita, T. & Nukui, H. Size of cerebral aneurysms and related factors in patients with subarachnoid hemorrhage. *Surg. Neurol.* **61**, 239–245 (2004) doi:10.1016/S0090-3019(03)00427-0.

33. Li, M. *et al.* Genetic polymorphisms and transcription profiles associated with intracranial aneurysm: A key role for NOTCH3. *Aging (Albany. NY).* **11**, 5173–5191 (2019)

doi:10.18632/aging.102111.

34. Hussain, S. *et al.* Search for Biomarkers of Intracranial Aneurysms: A Systematic Review. *World Neurosurg.* **84**, 1473–1483 (2015) doi:10.1016/j.wneu.2015.06.034.

35. Frösen, J. *et al.* Saccular intracranial aneurysm: Pathology and mechanisms. *Acta Neuropathol.* **123**, 773–786 (2012) doi:10.1007/s00401-011-0939-3.

36. Krings, T. *et al.* Intracranial aneurysms: From vessel wall pathology to therapeutic approach. *Nat. Rev. Neurol.* **7**, 547–559 (2011) doi:10.1038/nrneurol.2011.136.

37. International Study of Unruptured Intracranial Aneurysms Investigators. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. *Lancet* **362**, 103–110 (2003) doi:10.1056/NEJM199812103392401.

38. The UCAS Japan Investigators *et al.* The natural course of unruptured cerebral aneurysms in a Japanese cohort. *N Engl J Med* **366**, 2472–82 (2012) doi:10.1056/NEJMoa1113260.

39. Brown, R. D. & Broderick, J. P. Unruptured intracranial aneurysms: Epidemiology, natural history, management options, and familial screening. *Lancet Neurol.* **13**, 393–404 (2014) doi:10.1016/S1474-4422(14)70015-8.

40. Rinkel, G. J. E. & Algra, A. Long-term outcomes of patients with aneurysmal subarachnoid haemorrhage. *Lancet Neurol.* **10**, 349–356 (2011) doi:10.1016/S1474-4422(11)70017-5.

41. Ronkainen, A. *et al.* Risk of harboring an unruptured intracranial aneurysm. *Stroke* **29**, 359–362 (1998) doi:10.1161/01.STR.29.2.359.

42. Juvela, S. Risk factors for multiple intracranial aneurysms. *Stroke* **31**, 392–397 (2000) doi:10.1161/01.STR.31.2.392.

43. Juvela, S., Poussa, K. & Porras, M. Factors Affecting Formation and Growth of Intracranial Aneurysms. *Stroke* **32**, 485–491 (2001) doi:10.1161/01.STR.32.2.485.

44. Xu, H. W., Yu, S. Q., Mei, C. L. & Li, M. H. Screening for intracranial aneurysm in 355 patients with autosomal-dominant polycystic kidney disease. *Stroke* **42**, 204–206 (2011) doi:10.1161/STROKEAHA.110.578740.

45. Kim, S. T., Brinjikji, W. & Kallmes, D. F. Prevalence of intracranial aneurysms in patients with connective tissue diseases: A retrospective study. *Am. J. Neuroradiol.* **37**, 1422–1426 (2016) doi:10.3174/ajnr.A4718.

46. Santiago-Sim, T. *et al.* THSD1 (Thrombospondin Type 1 Domain Containing Protein 1) Mutation in the Pathogenesis of Intracranial Aneurysm and Subarachnoid Hemorrhage. *Stroke* **47**, 3005–3013 (2016) doi:10.1161/STROKEAHA.116.014161.

47. Li, B. *et al.* Associations among genetic variants and intracranial aneurysm in a Chinese population. *Yonsei Med. J.* **60**, 651–658 (2019) doi:10.3349/ymj.2019.60.7.651.

48. Ruigrok, Y. M., Rinkel, G. J. E., Wijmenga, C. & Van Gijn, J. Anticipation and phenotype in familial intracranial aneurysms. *J. Neurol. Neurosurg. Psychiatry* **75**, 1436–1442 (2004) doi:10.1136/jnnp.2003.025098.

49. Walmsley, J. G., Campling, M. R. & Chertkow, H. M. Interrelationships among wall structure, smooth muscle orientation, and contraction in human major cerebral arteries. *Stroke* **14**, 781–790 (1983) doi:10.1161/01.STR.14.5.781.

50. Daneman, R. & Prat, A. The Blood-Brain Barrier. *Cold Spring Harb Perspect Biol.* **7**, (2015) doi:10.1111/j.1469-8749.1961.tb10410.x.

51. Ballabh, P., Braun, A. & Nedergaard, M. The blood-brain barrier: An overview: Structure, regulation, and clinical implications. *Neurobiol. Dis.* **16**, 1–13 (2004)

doi:10.1016/j.nbd.2003.12.016.

52. Abbott, N. J., Patabendige, A. A. K., Dolman, D. E. M., Yusof, S. R. & Begley, D. J. Structure and function of the blood-brain barrier. *Neurobiol. Dis.* **37**, 13–25 (2010) doi:10.1016/j.nbd.2009.07.030.

53. Stamatovic, S. M., Johnson, A. M., Keep, R. F. & Andjelkovic, A. V. Junctional proteins of the blood-brain barrier: New insights into function and dysfunction. *Tissue Barriers* **4**, 1–12 (2016) doi:10.1080/21688370.2016.1154641.

54. Ronkainen, A. *et al.* Familial intracranial aneurysms. *Lancet* **349**, 380–384 (1997) doi:10.1016/S0140-6736(97)80009-8.

55. Bederson, J. B. *et al.* Recommendations for the management of patients with unruptured intracranial aneurysms: A statement for healthcare professionals from the Stroke Council of the American Heart Association. *Circulation* **102**, 2300–2308 (2000) doi:10.1161/01.CIR.102.18.2300.

56. Wang, M. C., Rubinstein, D., Kindt, G. W. & Breeze, R. E. Prevalence of intracranial aneurysms in first-degree relatives of patients with aneurysms. *Neurosurg. Focus* **13**, 13–16 (2002) doi:10.3171/foc.2002.13.3.3.

57. Wills, S. *et al.* Familial intracranial aneurysms: An analysis of 346 multiplex Finnish families. *Stroke* **34**, 1370–1374 (2003) doi:10.1161/01.STR.0000072822.35605.8B.

58. Broderick, J. P. *et al.* Greater rupture risk for familial as compared to sporadic unruptured intracranial aneurysms. *Stroke* **40**, 1952–1957 (2009) doi:10.1161/STROKEAHA.108.542571.

59. Zuurbier, C., Greving, J., Rinkel, G. & Ruigrok, Y. Higher risk of intracranial aneurysms and subarachnoid haemorrhage in siblings of families with intracranial aneurysms. *Eur.*

Stroke J. **5**, 73–77 (2020) doi:10.1177/2396987319868048.

60. Van Hoe, W. *et al.* Screening for Intracranial Aneurysms in Individuals with a Positive First-Degree Family History: A Systematic Review. *World Neurosurg.* **151**, 235-248.e5 (2021) doi:10.1016/j.wneu.2021.02.112.

61. Thompson, B. G. *et al.* Guidelines for the Management of Patients With Unruptured Intracranial Aneurysms. *Stroke* **46**, 2368–2400 (2015) doi:10.1161/STR.0000000000000070.

62. Mackey, J. *et al.* Unruptured intracranial aneurysms in the Familial Intracranial Aneurysm and International Study of Unruptured Intracranial Aneurysms cohorts: Differences in multiplicity and location - Clinical article. *J. Neurosurg.* **117**, 60–64 (2012) doi:10.3171/2012.4.JNS111822.

63. Mensing, L. A., Rinkel, G. J. E., Vlak, M. H. M., Van Der Schaaf, I. C. & Ruigrok, Y. M. Difference in aneurysm characteristics between patients with familial and sporadic aneurysmal subarachnoid haemorrhage. *PLoS One* **11**, 1–7 (2016) doi:10.1371/journal.pone.0154281.

64. Bacigaluppi, S. *et al.* Factors affecting formation and rupture of intracranial saccular aneurysms. *Neurosurg. Rev.* **37**, 1–14 (2014) doi:10.1007/s10143-013-0501-y.

65. Bor, A. S. E., Rinkel, G. J. E., van Norden, J. & Wermer, M. J. H. Long-term, serial screening for intracranial aneurysms in individuals with a family history of aneurysmal subarachnoid haemorrhage: A cohort study. *Lancet Neurol.* **13**, 385–392 (2014) doi:10.1016/S1474-4422(14)70021-3.

66. Bakker, M. K. *et al.* Genome-wide association study of intracranial aneurysms identifies 17 risk loci and genetic overlap with clinical risk factors. *Nat. Genet.* **52**, 1303–1313

(2020) doi:10.1038/s41588-020-00725-7.

67. Korja, M. *et al.* Genetic epidemiology of spontaneous subarachnoid hemorrhage: Nordic twin study. *Stroke* **41**, 2458–2462 (2010) doi:10.1161/STROKEAHA.110.586420.

68. Brown, R. D. *et al.* Screening for brain aneurysm in the Familial Intracranial Aneurysm study: Frequency and predictors of lesion detection. *J. Neurosurg.* **108**, 1132–1138 (2008) doi:10.3171/JNS/2008/108/6/1132.

69. McCarthy, M. I. *et al.* Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nat. Rev. Genet.* **9**, 356–369 (2008) doi:10.1038/nrg2344.

70. Syvänen, A. C. Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* **2**, 930–942 (2001) doi:10.1038/35103535.

71. Bilguvar, K. *et al.* Susceptibility loci for intracranial aneurysm in European and Japanese populations. *Nat. Genet.* **40**, 1472–1477 (2008) doi:10.1038/ng.240.

72. Abrantes, P. *et al.* Genetic variants underlying risk of intracranial aneurysms: Insights from a GWAS in Portugal. *PLoS One* **10**, 1–12 (2015) doi:10.1371/journal.pone.0133422.

73. Low, S. K. *et al.* Genome-wide association study for intracranial aneurysm in the Japanese population identifies three candidate susceptible loci and a functional genetic variant at EDNRA. *Hum. Mol. Genet.* **21**, 2102–2110 (2012) doi:10.1093/hmg/dds020.

74. Yasuno, K. *et al.* Common variant near the endothelin receptor type A (EDNRA) gene is associated with intracranial aneurysm risk. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19707–19712 (2011) doi:10.1073/pnas.1117137108.

75. Foroud, T. *et al.* Genome-wide association study of intracranial aneurysm identifies a new association on chromosome 7. *Stroke* **45**, 3194–3199 (2014) doi:10.1161/STROKEAHA.114.006096.

76. Yasuno, K. *et al.* Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat. Genet.* **42**, 420–425 (2010) doi:10.1038/ng.563.

77. Foroud, T. *et al.* Genome-wide association study of intracranial aneurysms confirms role of anril and SOX17 in disease risk. *Stroke* **43**, 2846–2852 (2012) doi:10.1161/STROKEAHA.112.656397.

78. Helgadottir, A. *et al.* The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat. Genet.* **40**, 217–224 (2008) doi:10.1038/ng.72.

79. McPherson, R. *et al.* A Common Allele on Chromosome 9 Associated with Coronary Heart Disease. *Science (80-.).* **316**, 1488–1491 (2007) doi:10.1126/science.1142447.

80. Hinohara, K. *et al.* Replication of the association between a chromosome 9p21 polymorphism and coronary artery disease in Japanese and Korean populations. *J. Hum. Genet.* **53**, 357–359 (2008) doi:10.1007/s10038-008-0248-4.

81. Helgadottir, A. *et al.* A Common Variant on Chromosome 9p21 Affects the Risk of Myocardial Infarction. *Science (80-.).* **316**, 1491–1493 (2007) doi:10.1126/science.1142842.

82. Helgadottir, A. *et al.* The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat. Genet.* **40**, 217–224 (2008) doi:10.1038/ng.72.

83. Gimelli, S. *et al.* Mutations in SOX17 are associated with congenital anomalies of the kidney and the urinary tract. *Hum. Mutat.* **31**, 1352–1359 (2010) doi:10.1002/humu.21378.

84. Heidet, L. *et al.* Targeted Exome Sequencing Identifies PBX1 as Involved in Monogenic

Congenital Anomalies of the Kidney and Urinary Tract. *J. Am. Soc. Nephrol.* **28**, 2901–2914 (2017) doi:10.1681/ASN.2017010043.

85. Rhodes, C. J. *et al.* Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *Lancet Respir. Med.* **7**, 227–238 (2019) doi:10.1016/S2213-2600(18)30409-0.

86. Wang, T.-M. *et al.* SOX17 Loss-of-Function Mutation Underlying Familial Pulmonary Arterial Hypertension. *Int. Heart J.* **62**, 566–574 (2021) doi:10.1536/ihj.20-711.

87. Zhu, N. *et al.* Rare variants in SOX17 are associated with pulmonary arterial hypertension with congenital heart disease. *Genome Med.* **10**, 56 (2018) doi:10.1186/s13073-018-0566-x.

88. Kataoka, H. Molecular mechanisms of the formation and progression of intracranial aneurysms. *Neurol. Med. Chir. (Tokyo).* **55**, 214–229 (2015) doi:10.2176/nmc.ra.2014-0337.

89. Zhou, S., Dion, P. A. & Rouleau, G. A. Genetics of intracranial aneurysms. *Stroke* **49**, 780–787 (2018) doi:10.1161/STROKEAHA.117.018152.

90. Pulst, S. M. Genetic linkage analysis. *Arch. Neurol.* **56**, 667–672 (1999) doi:10.1001/archneur.56.6.667.

91. Weiss, K. M. & Clark, A. G. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* **18**, 19–24 (2002) doi:10.1016/S0168-9525(01)02550-1.

92. Lewontin, R. C. & Dunn, L. C. The Evolutionary Dynamics of a Polymorphism in the House Mouse. *Genetics* **45**, 705–22 (1960).

93. Slatkin, M. Linkage disequilibrium - Understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* **9**, 477–485 (2008) doi:10.1038/nrg2361.

94. Ott, J., Wang, J. & Leal, S. M. Genetic linkage analysis in the age of whole-genome sequencing. *Nat. Rev. Genet.* **16**, 275–284 (2015) doi:10.1038/nrg3908.

95. Asuri, S. *et al.* Primary Biliary Cholangitis in British Columbia First Nations: Clinical features and discovery of novel genetic susceptibility loci. *Liver Int.* **38**, 940–948 (2018) doi:10.1111/liv.13686.

96. Smith, K. R. *et al.* Reducing the exome search space for Mendelian diseases using genetic linkage analysis of exome genotypes. *Genome Biol.* **12**, R85 (2011) doi:10.1186/gb-2011-12-9-r85.

97. Santiago-Sim, T. *et al.* Genomewide linkage in a large caucasian family maps a new locus for intracranial aneurysms to chromosome 13q. *Stroke* **40**, 57–60 (2009) doi:10.1161/STROKEAHA.108.534396.

98. Nahed, B. V. *et al.* Mapping a Mendelian form of intracranial aneurysm to 1p34.3-p36.13. *Am. J. Hum. Genet.* **76**, 172–179 (2005) doi:10.1086/426953.

99. Ruigrok, Y. M. *et al.* Genomewide linkage in a large Dutch family with intracranial aneurysms: Replication of 2 loci for intracranial aneurysms to chromosome 1p36.11-p36.13 and Xp22.2-p22.32. *Stroke* **39**, 1096–1102 (2008) doi:10.1161/STROKEAHA.107.495168.

100. Foroud, T. *et al.* Genome screen to detect linkage to intracranial aneurysm susceptibility genes: The familial intracranial aneurysm (FIA) study. *Stroke* **39**, 1434–1440 (2008) doi:10.1161/STROKEAHA.107.502930.

101. Foroud, T. *et al.* Genome screen in familial intracranial aneurysm. *BMC Med. Genet.* **10**, 1–10 (2009) doi:10.1186/1471-2350-10-3.

102. Verlaan, D. J. *et al.* A new locus for autosomal dominant intracranial aneurysm, ANIB4,

maps to chromosome 5p15.2-14.3. *J. Med. Genet.* **43**, 1–5 (2006) doi:10.1136/jmg.2005.033209.

103. Onda, H. *et al.* Genomewide-linkage and haplotype-association studies map intracranial aneurysm to chromosome 7q11. *Am. J. Hum. Genet.* **69**, 804–819 (2001) doi:10.1086/323614.

104. Farnham, J. M. *et al.* Confirmation of chromosome 7q11 locus for predisposition to intracranial aneurysm. *Hum. Genet.* **114**, 250–255 (2004) doi:10.1007/s00439-003-1044-z.

105. Kim, C. J. *et al.* Identification of an autosomal dominant locus for intracranial aneurysm through a model-based family collection in a geographically limited area. *J. Hum. Genet.* **56**, 464–466 (2011) doi:10.1038/jhg.2011.27.

106. Ozturk, A. K. *et al.* Molecular genetic analysis of two large kindreds with intracranial aneurysms demonstrates linkage to 11q24-25 and 14q23-31. *Stroke* **37**, 1021–1027 (2006) doi:10.1161/01.STR.0000206153.92675.b9.

107. Yamada, S. *et al.* Genome-wide scan for Japanese familial intracranial aneurysms: Linkage to several chromosomal regions. *Circulation* **110**, 3727–3733 (2004) doi:10.1161/01.CIR.0000143077.23367.18.

108. Olson, J. M. *et al.* Search for intracranial aneurysm susceptibility gene(s) using Finnish families. *BMC Med. Genet.* **3**, 7 (2002) doi:10.1186/1471-2350-3-7.

109. Olson, J. M. *et al.* Search for intracranial aneurysm susceptibility gene(s) using Finnish families. *BMC Med. Genet.* **3**, 7 (2002) doi:10.1186/1471-2350-3-7.

110. Craig Venter, J. *et al.* The sequence of the human genome. *Science (80-.).* **291**, 1304–1351 (2001) doi:10.1126/science.1058040.

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

112. Eilbeck, K., Quinlan, A. & Yandell, M. Settling the score: Variant prioritization and Mendelian disease. *Nat. Rev. Genet.* **18**, 599–612 (2017) doi:10.1038/nrg.2017.52.

113. Gilissen, C., Hoischen, A., Brunner, H. G. & Veltman, J. A. Disease gene identification strategies for exome sequencing. *Eur. J. Hum. Genet.* **20**, 490–497 (2012) doi:10.1038/ejhg.2011.258.

114. Gloss, B. S. & Dinger, M. E. Realizing the significance of noncoding functionality in clinical genomics. *Exp. Mol. Med.* **50**, (2018) doi:10.1038/s12276-018-0087-0.

115. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012) doi:10.1038/nature11247.

116. Sauvigny, T. *et al.* Exome sequencing in 38 patients with intracranial aneurysms and subarachnoid hemorrhage. *J. Neurol.* **267**, 2533–2545 (2020) doi:10.1007/s00415-020-09865-6.

117. Yang, X. *et al.* Rho Guanine Nucleotide Exchange Factor ARHGEF17 Is a Risk Gene for Intracranial Aneurysms. *Circ. Genomic Precis. Med.* **11**, e002099 (2018) doi:10.1161/CIRCGEN.117.002099.

118. Bourcier, R. *et al.* Rare Coding Variants in ANGPTL6 Are Associated with Familial Forms of Intracranial Aneurysm. *Am. J. Hum. Genet.* **102**, 133–141 (2018) doi:10.1016/j.ajhg.2017.12.006.

119. Hostettler, I. C. *et al.* ANGPTL6 genetic variants are an underlying cause of familial intracranial aneurysms. *Neurology* **96**, e947–e955 (2021) doi:10.1212/WNL.0000000000011125.

120. Zhou, S. *et al.* RNF213 Is Associated with Intracranial Aneurysms in the French-Canadian Population. *Am. J. Hum. Genet.* **99**, 1072–1085 (2016) doi:10.1016/j.ajhg.2016.09.001.

121. Yan, J. *et al.* Genetic Study of Intracranial Aneurysms. *Stroke* **46**, 620–626 (2015) doi:10.1161/STROKEAHA.114.007286.

122. Farlow, J. L. *et al.* Lessons Learned from Whole Exome Sequencing in Multiplex Families Affected by a Complex Genetic Disorder, Intracranial Aneurysm. *PLoS One* **10**, e0121104 (2015) doi:10.1371/journal.pone.0121104.

123. Lorenzo-Betancor, O. *et al.* PCNT point mutations and familial intracranial aneurysms. *Neurology* **91**, e2170–e2181 (2018) doi:10.1212/WNL.0000000000006614.

124. Wu, Y. *et al.* Exome Sequencing Identifies LOXL2 Mutation as a Cause of Familial Intracranial Aneurysm. *World Neurosurg.* **109**, e812–e818 (2018) doi:10.1016/j.wneu.2017.10.094.

125. Liu, J. *et al.* A Rare Variant of ANK3 Is Associated With Intracranial Aneurysm. *Front. Neurol.* **12**, (2021) doi:10.3389/fneur.2021.672570.

126. Ding, X. *et al.* Exome sequencing reveals a novel variant in NFX1 causing intracranial aneurysm in a Chinese family. *J. Neurointerv. Surg.* **12**, 221–226 (2020) doi:10.1136/neurintsurg-2019-014900.

127. Powell, A. E., Fernandez, B. A., Maroun, F., Noble, B. & Woods, M. O. Familial Intracranial Aneurysm in Newfoundland: Clinical and Genetic Analysis. *Can. J. Neurol. Sci. / J. Can. des Sci. Neurol.* **46**, 518–526 (2019) doi:10.1017/cjn.2019.230.

128. Irazabal, M. V. *et al.* Extended follow-up of unruptured intracranialaneurysms detected by presymptomatic screening inpatients with autosomal dominant polycystickidney disease. *Clin. J. Am. Soc. Nephrol.* **6**, 1274–1285 (2011) doi:10.2215/CJN.09731110.

129. Sanchis, I. M. *et al.* Presymptomatic screening for intracranial aneurysms in patients with autosomal dominant polycystic kidney disease. *Clin. J. Am. Soc. Nephrol.* **14**, 1151–1160 (2019) doi:10.2215/CJN.14691218.

130. Zhou, Z. *et al.* Is Regular Screening for Intracranial Aneurysm Necessary in Patients with Autosomal Dominant Polycystic Kidney Disease? A Systematic Review and Meta-analysis. *Cerebrovasc. Dis.* **44**, 75–82 (2017) doi:10.1159/000476073.

131. Rodrigues, V. J., Elsayed, S., Loeys, B. L., Dietz, H. C. & Yousem, D. M. Neuroradiologic Manifestations of Loeys-Dietz Syndrome Type 1. *AJNR Am J Neuroradiol* **30**, 1614–19 (2009) doi:10.3174/ajnr.A1651.

132. Loeys, B. L. *et al.* Aneurysm Syndromes Caused by Mutations in the TGF-β Receptor. *N. Engl. J. Med.* **355**, 788–798 (2006) doi:10.1056/NEJMoa055695.

133. Kim, J. H. *et al.* Intracranial Aneurysms Are Associated With Marfan Syndrome. *Stroke* **52**, 331–334 (2021) doi:10.1161/STROKEAHA.120.032107.

134. Schievink, W. I., Riedinger, M. & Maya, M. M. Frequency of incidental intracranial aneurysms in neurofibromatosis type 1. *Am. J. Med. Genet.* **134 A**, 45–48 (2005) doi:10.1002/ajmg.a.30475.

135. Duker, A. L. *et al.* Microcephalic osteodysplastic primordial dwarfism type II is associated with global vascular disease. *Orphanet J. Rare Dis.* **16**, 1–15 (2021) doi:10.1186/s13023-021-01852-y.

136. Waldron, J. *et al.* Multiple intracranial aneurysms and moyamoya disease associated with microcephalic osteodysplastic primordial dwarfism type II: surgical considerations. *Pediatr. Neurosurg* **4**, 439–444 (2009) doi:10.3171/2009.6.PEDS08137.

137. Ostergaard Kristensen, M. Increased incidence of bleeding intracranial aneurysms in

Greenlandic Eskimos. *Acta Neurochir. (Wien).* **67**, 37–43 (1983) doi:10.1007/BF01401666.

138. Lindgaard, L. *et al.* Familial aggregation of intracranial aneurysms in an Inuit patient population in Kalaallit Nunaat (Greenland). *Neurosurgery* **52**, 357–363 (2003) doi:10.1227/01.NEU.0000043695.77193.62.

139. Zhou, S. *et al.* Genome-wide association analysis identifies new candidate risk loci for familial intracranial aneurysm in the French-Canadian population. *Sci. Rep.* **8**, 1–8 (2018) doi:10.1038/s41598-018-21603-7.

140. Zhou, S. *et al.* Genetic architecture and adaptations of Nunavik Inuit. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 16012–16017 (2019) doi:10.1073/pnas.1810388116.

141. Laveaux, D. & Christopher, S. Contextualizing CBPR: Key Principles of CBPR meet the Indigenous research context. *Pimatisiwin* **7**, 1 (2009).

142. Cochran, P. A. L. *et al.* Indigenous ways of knowing: Implications for participatory research and community. *Am. J. Public Health* **98**, 22–27 (2008) doi:10.2105/AJPH.2006.093641.

143. Claw, K. G. *et al.* A framework for enhancing ethical genomic research with Indigenous communities. *Nat. Commun.* **9**, (2018) doi:10.1038/s41467-018-05188-3.

144. Yao, R. A., Akinrinade, O., Chaix, M. & Mital, S. Quality of whole genome sequencing from blood versus saliva derived DNA in cardiac patients. *BMC Med. Genomics* **13**, 1–10 (2020) doi:10.1186/s12920-020-0664-7.

145. Trost, B. *et al.* Impact of DNA source on genetic variant detection from human whole-genome sequencing data. *J. Med. Genet.* **56**, 809–817 (2019) doi:10.1136/jmedgenet-2019-106281.

146. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009) doi:10.1093/bioinformatics/btp324.

147. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010) doi:10.1101/gr.107524.110.

148. Calabrese, C. *et al.* MToolBox: a highly automated pipeline for heteroplasmy annotation and prioritization analysis of human mitochondrial variants in high-throughput sequencing. *Bioinformatics* **30**, 3115–3117 (2014) doi:10.1093/bioinformatics/btu483.

149. Layer, R. M., Chiang, C., Quinlan, A. R. & Hall, I. M. LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol.* **15**, R84 (2014) doi:10.1186/gb-2014-15-6-r84.

150. Chen, X. *et al.* Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220–1222 (2016) doi:10.1093/bioinformatics/btv710.

151. Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974–984 (2011) doi:10.1101/gr.114876.110.

152. Zhu, M. *et al.* Using ERDS to Infer Copy-Number Variants in High-Coverage Genomes. *Am. J. Hum. Genet.* **91**, 408–421 (2012) doi:10.1016/j.ajhg.2012.07.004.

153. Jeffares, D. C. *et al.* Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat. Commun.* **8**, 14061 (2017) doi:10.1038/ncomms14061.

154. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164–e164 (2010) doi:10.1093/nar/gkq603.

155. Manzoni, C. *et al.* Genome, transcriptome and proteome: The rise of omics data and their integration in biomedical sciences. *Brief. Bioinform.* **19**, 286–302 (2018) doi:10.1093/BIB/BBW114.

156. van't Hof, F. N. G. *et al.* Shared genetic risk factors of intracranial, abdominal, and thoracic aneurysms. *J. Am. Heart Assoc.* **5**, 1–19 (2016) doi:10.1161/JAHA.115.002603.

157. Regalado, E. S. *et al.* Exome Sequencing Identifies SMAD3 Mutations as a Cause of Familial Thoracic Aortic Aneurysm and Dissection With Intracranial and Other Arterial Aneurysms. *Circ. Res.* **109**, 680–686 (2011) doi:10.1161/CIRCRESAHA.111.248161.

158. Kurtelius, A. *et al.* Association of Intracranial Aneurysms With Aortic Aneurysms in 125 Patients With Fusiform and 4253 Patients With Saccular Intracranial Aneurysms and Their Family Members and Population Controls. *J. Am. Heart Assoc.* **8**, (2019) doi:10.1161/JAHA.119.013277.

159. Regalado, E. *et al.* Autosomal dominant inheritance of a predisposition to thoracic aortic aneurysms and dissections and intracranial saccular aneurysms. *Am. J. Med. Genet. Part A* **155**, 2125–2130 (2011) doi:10.1002/ajmg.a.34050.

160. Mackey, J. *et al.* Affected Twins in the Familial Intracranial Aneurysm Study. *Cerebrovasc. Dis.* **39**, 82–86 (2015) doi:10.1159/000369961.

161. ter Laan, M., Kerstjens-Frederikse, W. S., Metzemaekers, J. D. M., van Dijk, J. M. C. & Groen, R. J. M. Concordant Symptomatic Intracranial Aneurysm in a Monozygotic Twin: A Case Report and Review of the Literature. *Twin Res. Hum. Genet.* **12**, 295–300 (2009)

doi:10.1375/twin.12.3.295.

162. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020) doi:10.1038/s41586-020-2308-7.

163. Pruitt, K. D. *et al.* RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res.* **42**, D756–D763 (2014) doi:10.1093/nar/gkt1114.

164. Aken, B. L. *et al.* The Ensembl gene annotation system. *Database* **2016**, baw093 (2016) doi:10.1093/database/baw093.

165. Jian, X., Boerwinkle, E. & Liu, X. In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res.* **42**, 13534–13544 (2014) doi:10.1093/nar/gku1206.

166. Langlois, S. *et al.* De novo dominant variants affecting the motor domain of KIF1A are a cause of PEHO syndrome. *Eur. J. Hum. Genet.* **24**, 949–953 (2016) doi:10.1038/ejhg.2015.217.

167. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011) doi:10.1038/nbt.1754.

168. Anderson, S. *et al.* Sequence and organizing of the human mitochondrial genome. *Nature* **290**, 457–465 (1981) doi:https://doi.org/10.1038/290457a0.

169. Brandon, M. C. *et al.* MITOMAP : a human mitochondrial genome database — 2004 update. *Nucleic Acids Res.* **33**, 2004–2006 (2005) doi:10.1093/nar/gki079.

170. Lott, M. T. *et al.* mtDNA Variation and Analysis Using Mitomap and Mitomaster. *Curr. Protoc. Bioinforma.* **44**, (2013) doi:10.1002/0471250953.bi0123s44.

171. Hitchcock, E. C. *Next Generation Sequencing To Determine a Genetic Cause of Familial Intracranial Aneurysms*. [Master's Thesis] (The University of British Columia, 2017).

172. Balashanmugam, M. V., Shivanandappa, T. B., Nagarethinam, S., Vastrad, B. & Vastrad, C. Analysis of Differentially Expressed Genes in Coronary Artery Disease by Integrated Microarray Analysis. *Biomolecules* **10**, 35 (2019) doi:10.3390/biom10010035.

173. Collins, S. A. *et al.* Association of the CPT1A p . P479L Metabolic Gene Variant With Childhood Respiratory and Other Infectious Illness in Nunavut. *Front. Pediatr.* **9**, (2021) doi:10.3389/fped.2021.678553.

174. Collins, S. A. *et al.* Causes and risk factors for infant mortality in Nunavut, Canada 1999-2011. *BMC Pediatr.* **12**, (2012) doi:10.1186/1471-2431-12-190.

175. Sinclair, G. B. *et al.* Carnitine palmitoyltransferase I and sudden unexpected infant death in British Columbia First Nations. *Pediatrics* **130**, (2012) doi:10.1542/peds.2011-2924.

176. Gessner, B. D., Wood, T., Johnson, M. A., Richards, C. S. & Koeller, D. M. Evidence for an association between infant mortality and homozygosity for the arctic variant of carnitine palmitoyltransferase 1A. *Genet. Med.* **18**, 933–939 (2016) doi:10.1038/gim.2015.197.

177. National Center for Biotechnology Information. NM_001876.4(CPT1A:c.1436C>T (p.Pro479Leu). *ClinVar; [VCV000065644.10]* https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000065644.10.

178. Garshasbi, M. *et al.* A novel nonsense mutation in TUSC3 is responsible for non-syndromic autosomal recessive mental retardation in a consanguineous Iranian family. *Am. J. Med. Genet. Part A* **155**, 1976–1980 (2011) doi:10.1002/ajmg.a.34077.

179. Garshasbi, M. *et al.* A Defect in the TUSC3 Gene Is Associated with Autosomal Recessive Mental Retardation. *Am. J. Hum. Genet.* **82**, 1158–1164 (2008) doi:10.1016/j.ajhg.2008.03.018.

180. Al-Amri, A. *et al.* Homozygous single base deletion in TUSC3 causes intellectual disability with developmental delay in an Omani family. *Am. J. Med. Genet. Part A* **170**, 1826–1831 (2016) doi:10.1002/ajmg.a.37690.

181. Molinari, F. *et al.* Oligosaccharyltransferase-Subunit Mutations in Nonsyndromic Mental Retardation. *Am. J. Hum. Genet.* **82**, 1150–1157 (2008) doi:10.1016/j.ajhg.2008.03.021.

182. Ben-Salem, S. *et al.* Defect in phosphoinositide signalling through a homozygous variant in PLCB3 causes a new form of spondylometaphyseal dysplasia with corneal dystrophy. *J. Med. Genet.* **55**, 122–130 (2018) doi:10.1136/jmedgenet-2017-104827.

183. Ku, N.-O., Gish, R., Wright, T. L. & Omary, M. B. Keratin 8 Mutations in Patients with Cryptogenic Liver Disease. *N. Engl. J. Med.* **344**, 1580–1587 (2001) doi:10.1056/NEJM200105243442103.

184. Ku, N. O. *et al.* Keratin 8 and 18 mutations are risk factors for developing liver disease of multiple etiologies. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6063–6068 (2003) doi:10.1073/pnas.0936165100.

185. Jacob, J. T., Coulombe, P. A., Kwan, R. & Omary, M. B. Types I and II Keratin Intermediate Filaments. *Cold Spring Harb. Perspect. Biol.* **10**, a018275 (2018) doi:10.1101/cshperspect.a018275.

186. Anand, S. S. *et al.* Canadian Alliance for Healthy Hearts and Minds: First Nations Cohort Study Rationale and Design. *Prog. Community Heal. Partnerships Res. Educ. Action* **12**, 55–64 (2018) doi:10.1353/cpr.2018.0006.

187. Abel, H. J. *et al.* Mapping and characterization of structural variation in 17,795 human genomes. *Nature* **583**, 83–89 (2020) doi:10.1038/s41586-020-2371-0.

188. Chaisson, M. J. P. *et al.* Multi-platform discovery of haplotype-resolved structural

variation in human genomes. *Nat. Commun.* **10**, 1–16 (2019) doi:10.1038/s41467-018-08148-z.

189. Ishizawa, T., Nozaki, Y., Ueda, T. & Takeuchi, N. The human mitochondrial translation release factor HMRF1L is methylated in the GGQ motif by the methyltransferase HMPrmC. *Biochem. Biophys. Res. Commun.* **373**, 99–103 (2008) doi:10.1016/j.bbrc.2008.05.176.

190. Soleimanpour-Lichaei, H. R. *et al.* mtRF1a Is a Human Mitochondrial Translation Release Factor Decoding the Major Termination Codons UAA and UAG. *Mol. Cell* **27**, 745–757 (2007) doi:10.1016/j.molcel.2007.06.031.

191. Chalouhi, N. *et al.* Biology of intracranial aneurysms: Role of inflammation. *J. Cereb. Blood Flow Metab.* **32**, 1659–1676 (2012) doi:10.1038/jcbfm.2012.84.

192. Bateman, A. *et al.* UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021) doi:10.1093/nar/gkaa1100.

193. Firth, H. V. *et al.* DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.* **84**, 524–533 (2009) doi:10.1016/j.ajhg.2009.03.010.

194. Price, N. T. *et al.* A novel brain-expressed protein related to carnitine palmitoyltransferase I. *Genomics* **80**, 433–442 (2002) doi:10.1006/geno.2002.6845.

195. Richards, S., Aziz, N., Bale, S., Bick, D. & Das, S. ACMG Standards and Guidelines Standards and guidelines for the interpretation of sequence variants : a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015) doi:10.1038/gim.2015.30.

196. Collins, S. A. *et al.* Carnitine palmitoyltransferase 1A (CPT1A) P479L prevalence in live newborns in Yukon, Northwest Territories, and Nunavut. *Mol. Genet. Metab.* **101**, 200–204 (2010) doi:10.1016/j.ymgme.2010.07.013.

197. Rajakumar, C. *et al.* Carnitine palmitoyltransferase IA polymorphism P479L is common in Greenland Inuit and is associated with elevated plasma apolipoprotein A-I. *J. Lipid Res.* **50**, 1223–1228 (2009) doi:10.1194/jlr.P900001-JLR200.

198. Zhou, S. *et al.* Increased missense mutation burden of fatty acid metabolism related genes in Nunavik Inuit population. *PLoS One* **10**, 1–15 (2015) doi:10.1371/journal.pone.0128255.

199. Adams, L. C. *et al.* Noninvasive imaging of vascular permeability to predict the risk of rupture in abdominal aortic aneurysms using an albumin-binding probe. *Sci. Rep.* **10**, 3231 (2020) doi:10.1038/s41598-020-59842-2.

200. Amenta, P., Valle, E., Dumont, A. & Medel, R. Inflammation and intracranial aneurysms: mechanisms of initiation, growth, and rupture. *Neuroimmunol Neuroinflammation* **2**, 68–76 (2015) doi:10.4103/2347-8659.153975.

201. Hoeppner, L. H. *et al.* Revealing the role of phospholipase Cβ3 in the regulation of VEGF-induced vascular permeability. *Blood* **120**, 2167–2173 (2012) doi:10.1182/blood-2012-03-417824.

202. Liang, D. *et al.* The role of vascular endothelial growth factor (VEGF) in vasculogenesis, angiogenesis, and hematopoiesis in zebrafish development. *Mech. Dev.* **108**, 29–43 (2001) doi:10.1016/S0925-4773(01)00468-3.

203. Bhattacharya, R. *et al.* Distinct role of PLCβ3 in VEGF-mediated directional migration and vascular sprouting. *J. Cell Sci.* **122**, 1025–1034 (2009) doi:10.1242/jcs.041913.

204. Zhang, Y. *et al.* Dickkopf-3 attenuates pressure overload-induced cardiac remodelling. *Cardiovasc. Res.* **102**, 35–45 (2014) doi:10.1093/cvr/cvu004.

205. Busceti, C. L. *et al.* Dickkopf-3 Upregulates VEGF in Cultured Human Endothelial Cells by Activating Activin Receptor-Like Kinase 1 (ALK1) Pathway. *Front. Pharmacol.* **8**, (2017) doi:10.3389/fphar.2017.00111.

206. Jaganathan, K. *et al.* Predicting Splicing from Primary Sequence with Deep Learning. *Cell* **176**, 535-548.e24 (2019) doi:10.1016/j.cell.2018.12.015.

207. Lin, Y. *et al.* DLC2 modulates angiogenic responses in vascular endothelial cells by regulating cell attachment and migration. *Oncogene* **29**, 3010–3016 (2010) doi:10.1038/onc.2010.54.

208. Fukino, K. *et al.* Regulation of angiogenesis by the aging suppressor gene klotho. *Biochem. Biophys. Res. Commun.* **293**, 332–337 (2002) doi:10.1016/S0006-291X(02)00216-4.

209. Ostberg, N., Zafar, M., Ziganshin, B. & Elefteriades, J. The Genetics of Thoracic Aortic Aneurysms and Dissection: A Clinical Perspective. *Biomolecules* **10**, 182 (2020) doi:10.3390/biom10020182.

210. Vogel, L. & Eggertson, L. COVID-19: A timeline of Canada's first-wave response. *CMAJ* https://cmajnews.com/2020/06/12/coronavirus-1095847/ (2020).

211. Ellerby, L. M. Repeat Expansion Disorders: Mechanisms and Therapeutics. *Neurotherapeutics* **16**, 924–927 (2019) doi:10.1007/s13311-019-00823-3.

212. Alkan, C., Coe, B. P. & Eichler, E. E. Genome structural variation discovery and genotyping. *Nat. Rev. Genet.* **12**, 363–376 (2011) doi:10.1038/nrg2958.

213. Shaikh, T. H. Copy Number Variation Disorders. *Curr. Genet. Med. Rep.* **5**, 183–190

(2017) doi:10.1007/s40142-017-0129-2.

214. Aouiche, C., Shang, X. & Chen, B. Copy number variation related disease genes. *Quant. Biol.* **6**, 99–112 (2018) doi:10.1007/s40484-018-0137-6.

215. Ebbert, M. T. W. *et al.* Systematic analysis of dark and camouflaged genes reveals disease-relevant genes hiding in plain sight. *Genome Biol.* **20**, 97 (2019) doi:10.1186/s13059-019-1707-2.

216. Walcott, B. P. & Peterson, R. T. Zebrafish Models of Cerebrovascular Disease. *J. Cereb. Blood Flow Metab.* **34**, 571–577 (2014) doi:10.1038/jcbfm.2014.27.

217. Howe, K. *et al.* The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498–503 (2013) doi:10.1038/nature12111.

218. Bradford, Y. M. *et al.* Zebrafish Models of Human Disease: Gaining Insight into Human Disease at ZFIN. *ILAR J.* **58**, 4–16 (2017) doi:10.1093/ilar/ilw040.

219. Zhang, P.-Y. PCSK9 as a therapeutic target for cardiovascular disease. *Exp. Ther. Med.* **13**, 810–814 (2017) doi:10.3892/etm.2017.4055.

220. Sabatine, M. S. *et al.* Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *N. Engl. J. Med.* **376**, 1713–1722 (2017) doi:10.1056/NEJMoa1615664.

221. Shimizu, K. *et al.* Candidate drugs for preventive treatment of unruptured intracranial aneurysms: A cross-sectional study. *PLoS One* **16**, e0246865 (2021) doi:10.1371/journal.pone.0246865.

Appendices

Appendix A

MAY 2020

A great deal of work has been carried out developing educational tools for patients and community members, so we encourage you to check out the Silent Genomes web page:
https://www.bcchr.ca/silent-genomes-project

The Gitxsan are participating in the Silent Genomes project within Activities 2 and 3. As part of **Activity 2**, we are excited to introduce you to our new master's student, Avery Newman-Simmons (below right).

Avery's family is from the
Territory and he is currently enrolled in the Medical Genetics Program at the University of British Columbia. Avery is exploring whether there is an underlying genetic cause that can explain why there are so many Gitxsan people affected with **brain (cerebral) aneurysms**. (It has been brought to our attention at many gatherings that this is a priority to explore). To do this, he will use a genetic test called "whole genome sequencing" to see if any specific variants in the DNA of those affected, can be identified as variants predisposing to aneurysms. Identifying these DNA changes may help determine who is at higher risk of brain aneurysms and therefore could benefit from brain imaging (MRI or CT) for early diagnosis and prevention, **before** the aneurysm presents as a problem. This is complicated work, but with the help of many Gitxsan people who are affected, we are hoping to find some answers.

Activity 3 involves creation of an **IBVL** (see information box on the IBVL) and would help provide better care for Indigenous people who have a genetic disease. In order to create the IBVL, the Silent Genomes team is engaging with each community involved in the Alliance study (see page 7-9), to determine whether they will also be interested in participating in Silent Genomes' IBVL. Genetic samples stored as part of the Alliance study (with consents that indicate 'Biobanking for future research') could be used for the development of the IBVL, with return of the genomic data generated to the Alliance study. As the Gitxsan community was previously involved in the Alliance study,we presented this idea at the community gathering last year and there was expressed interest from the community in contributing their Alliance samples to the IBVL. With Dr. Caron's lead, we have now developed the Terms of Reference for an Indigenous Steering Committee.

What is an Indigenous Background Variant Library (IBVL)?

(adapted from https://www.bcchr.ca/silent-genomes-project/ibvl)

A Background Variant Library (BVL) is a list or collection of all the DNA variants in a group of people without severe genetic conditions, which helps to determine variants that are common or rare in the general population. If a variant is common in a group of 'healthy' people it is less likely to be responsible for a serious disease.

Genetic specialists and doctors use BVLs as references, to learn which DNA variants are common in people from various populations around the world.

When a patient with a suspected rare genetic disease has genetic testing, to try to find the DNA variant causing their condition, sorting through all the DNA variants found on their genetic test to find the 'suspect' disease variant that is causing their condition can be complicated.

Knowing which variants are present in Indigenous people can simplify the search for disease causing variants.

Appendix B

3. The *Silent Genomes* Project (‘*Silent Genomes*’)

The *Silent Genomes* project was first introduced at the community research gathering in 2019. *Silent Genomes* is a study led by Dr. Arbour and Dr. Nadine Caron of Prince George that aims to reduce barriers for diagnosis for Indigenous Children with Genetic Disease. The project includes 4 key activities:

- Activity 1: Integrating Indigenous-led governance, community engagement, community education, and student capacity building across all activities
- Activity 2: Precision genomic diagnosis for Indigenous patients and families with genetic disease
- Activity 3: Development of an **Indigenous background variant library (IBVL)**
- Activity 4: Economics of precision diagnosis for Indigenous children

A great deal of work has been carried out developing educational tools for patients and community members, so we encourage you to check out the Silent Genomes web page: https://www.bcchr.ca/silent-genomes-project

The Gitxsan are participating in the *Silent Genomes* project within Activities 2 and 3. As part of **Activity 2**, our team was joined by master’s student Avery Newman-Simmons, whose family is from the Territory.

Avery’s project has explored whether there is an underlying genetic cause that can explain why there are so many Gitxsan people affected with **brain (cerebral) aneurysms**. (It has been brought to our attention at many gatherings that this is a priority to explore). To do this, he used a genetic test called “whole genome sequencing” to see if any specific variants in the DNA of those affected, can be identified as variants predisposing to aneurysms. Based on his results, there does not appear to be a single genetic variant alone that explains why cerebral aneurysms are forming in many Gitxsan community members. This finding suggests that cerebral aneurysms in the community are likely multifactorial, forming due to a combination of both genetic and environmental factors. Screening is available for community members who have a first-degree relative (sibling, parent, child) with a cerebral aneurysm (see box below).

While Avery will be completing his master’s at the end of the summer, research in this area will continue. We are grateful for Avery’s work and wish him well as he starts medical school at the Northern Ontario School of Medicine this fall!

The following screening is recommended for Northern BC First Nations with a first-degree relative (sibling, parent, child) with a confirmed cerebral aneurysm: Brain imaging with enhancement (CTA or MRA) should be carried out and repeated every 5 years starting at age 18 years.

- It is important to ***repeat this test every 5 years even if the initial results are normal***
- Baseline screening at age 18 and every 5 years to follow
- This test can be arranged by your primary care provider, such as your family doctor

CTA – computed tomography angiography; MRA – magnetic resonance angiography

Appendix C


The Community Genetics Research Program presents
a zoom trainees research symposium:
Long QT syndrome and beyond
April 9th 7:00pm – 9:00pm
Zoom Registration Link:
Welcome and introduction:
Julie Morrison, Gitxsan Community Member
Laura Arbour MD: Overview- Long QT syndrome, Silent Genomes, Alliance - where are we now?
Simona Bene-Watts MSc: Long QT syndrome and CPT1A in children - a reassuring story
Nicole York BSc/MSc Candidate:
ANK2 and heart development
Juan Sanchez-Arias MD PhD: Understanding How ANK2 Affects Heart Rhythm
Avery Newman-Simmons BSc/MSc Candidate: Identifying a Genetic Cause for Intracranial Aneurysms in the Gitxsan
Guest: Dr. Jeff Brook, University of Toronto: Staying healthy in a time of environmental challenges
UBC
University of Victoria
University of Toronto

Appendix D

THE UNIVERSITY OF BRITISH COLUMBIA

Department of Medical Genetics

island health

PARTICIPANT INFORMATION AND CONSENT FORM

Title of Study: **Silent Genomes: Precision Diagnosis for Indigenous Families with Genetic Conditions**
(Activity 2 of *Silent Genomes: Reducing health care disparities and improving diagnostic success for children with genetic diseases from Indigenous populations*)

Principal Investigator and Victoria site lead:
Dr. Laura Arbour, MD, MSc, FRCPC FCCMG
Professor, Department of Medical Genetics
University of British Columbia / Children's & Women's Health Centre of BC

Co-Principal Investigator and Vancouver site lead:
Dr. Anna Lehman, MD, FRCPC FCCMG
Associate Professor, Department of Medical Genetics
University of British Columbia / Children's & Women's Health Centre of BC

Co-Investigators:
Dr. Maja Tarailo-Graovac- Alberta Children's Hospital Research Institute, University of Calgary
Dr. Francois Bernier- Dept Medical Genetics, University of Calgary
Dr. Nadine Caron- Dept Surgery, University of British Columbia Northern Medical Program (University Hospital of Northern BC)
Dr. Wyeth Wasserman- Dept Medical Genetics, University of British Columbia; BC Children's Hospital Research Institute
Dr. Marco Marra-Dept Medical Genetics, University of British Columbia

Sponsors: **Genome Canada; Canadian Institutes of Health Research; Provincial Health Services Authority; UBC Faculty of Medicine, BC Children's Hospital Foundation; BC Children's Hospital Research Institute, Michael Smith Foundation for Health Research; Illumina Inc.**

If you are a <u>parent or legal guardian of a child</u> who may take part in this study, this consent form pertains to your <u>child's</u> participation and, if acceptable, is to be signed by you. The assent (agreement) of your child may also be required.

If you are a <u>substitute decision-maker for an adult</u> who may take part in this study, this consent form is also to be signed by you, and the assent (agreement) of the adult for whom you are a substitute decision-maker may also be required.

If you are a parent joining this study along with your child, or other adult joining this study, this consent form is also to be signed by you as an adult participant.

When we say "you" or "your" in this consent form, we mean the research participant you are providing consent for. This may be you, a child, or an adult for whom you are the substitute decision-maker.

"We" means the study doctors and other research staff.

1. INVITATION

You are invited to take part in this research study because:

- You have a medical condition which is not yet understood, and is likely due to a genetic (single gene) cause
  OR
- You are a biological (blood-related) relative of a person who is affected with such a medical condition.

As part of this study you are invited to undergo a type of genetic testing called **'Whole Genome Sequencing' (WGS)**, a new technology which can find changes in DNA, sometimes referred to as "genetic changes" or "**variants**." The purpose of this study is to try to find the DNA variants causing the condition in your family and how best to manage your healthcare.

In most cases, the person affected with the condition as well as two additional blood-related relatives will be invited to take part, although sometimes more family members will be invited (depending on the condition). The reason for including relatives without the condition is to compare DNA variants that may be identified in the affected family member. When both parents of an affected child participate, we can determine whether the child's variants were passed down through the family, or if they are new genetic changes that happened for the first time in the child.

2. YOUR PARTICIPATION IS VOLUNTARY

Your participation is voluntary. You have the right to refuse to participate in this study. If you decide to participate, you may still choose to withdraw from the study at any time without any negative consequences to your medical care, education, or other services to which you are entitled or are presently receiving.

In order to decide whether or not you wish to be a part of this research study, you should understand what is involved and the potential risks and benefits. This form gives detailed information about the research study, which will be discussed with you. Once you understand the study, you will be asked to sign this form if you wish to participate.

Please take time to read the following information carefully and to discuss it with your family, friends, and doctor before you decide.

3. WHO IS CONDUCTING THE STUDY?

This study is being led by the Principal Investigators (study doctors), Drs. Laura Arbour and Anna Lehman, and the other investigators listed on the first page. This study is funded by Genome Canada, Canadian Institutes of Health Research, **Provincial Health Services Authority, UBC Faculty of Medicine, BC Children's Hospital Foundation, BC Children's Hospital Research Institute, and the Michael Smith Foundation for Health Research. Illumina Inc** is providing in-kind donation only.

4. BACKGROUND

Genetic conditions happen in **all populations around the world**. Individually, genetic conditions are rare, but if we group all genetic conditions together, the number is large. 'Rare' genetic conditions affect at least 1 in 50 people, and the number is even higher if we include more common conditions.

What are genes and DNA and variants?

We each have ~25,000 **genes** in every cell of our body, which we inherited from our parents. Genes are the 'instructions' for our bodies, which tell our cells, tissues, and organs how to work. Genes are made of **DNA** (deoxyribonucleic acid), a long string of 'genetic letters' that the cells can read. Every person has a slightly different combination of letters, due to changes (**'variants'**) in their DNA. You can think of variants as being similar to 'spelling mistakes' in the DNA. Most of these DNA variants are harmless (**non-disease-causing variants**), some have an unknown or uncertain effect (**variants of uncertain significance**), while others are known to play a role in disease (**disease-causing variants**).

What is Whole Genome Sequencing (WGS)?

The entire set of DNA in a person is called the human **genome.** The genome contains about 3 billion DNA 'letters' in total. **Whole Genome Sequencing (WGS)** is a new genetic technology that allows a person's entire genome to be read and studied at once. Since WGS has the ability to study all of your DNA letters at one time, it has the potential to replace many of the current genetic tests that are available which can only look at small sections of DNA (e.g. a single gene or a small number of genes) at one time.

Thousands of variants are found in each person undergoing WGS for a genetic condition. Although most of these will be harmless variants or variants of uncertain significance, there is about a **30% chance of finding a disease-causing variant** which leads to a new genetic diagnosis and explains the cause of the medical condition your family was being tested for.

Since WGS looks at your entire genome, it may unexpectedly reveal harmful variants in *other* genes which are not related to the reason you were tested. These types of variants are called **'incidental findings'**, because they were unexpectedly found while looking for something else. Sometimes incidental findings are **'medically actionable'**, meaning that there is an action that can be taken (for example, treatment or prevention plan) to reduce the chance the variant will affect health. For example, a disease-causing variant in a gene related to cancer risk could be unexpectedly found, and that information could lead to a recommendation for special cancer screening. How we will provide incidental finding results back to you is discussed in detail in section 8 (pages 7-8).

5. WHAT IS THE PURPOSE OF THIS STUDY?

The main purpose of this study is to **find the DNA variant(s) causing the suspected genetic**

condition in you or your family member. We will do this through WGS testing of family members with and without the condition.

This particular study focuses on Indigenous families in an effort to **reduce the barriers** many Indigenous communities face in accessing genetic testing and diagnosis.

6. WHO CAN PARTICIPATE IN THIS STUDY?

Your family may be able to take part in this study if you self-identify as Indigenous (First Nations, Inuit, or Métis) and/or have Indigenous ancestry and meet the following criteria:

- The affected person in the family has a suspected genetic (single-gene) condition which could present in childhood (childhood-onset)
- The condition has not been diagnosed through other available medical tests and investigations
- The condition is serious enough to affect the person's health or quality of life
- Learning the diagnosis is predicted to have a positive impact on the family's health and well-being (for example, the diagnosis may help improve health care delivery, help with treatment decisions, reduce uncertainty for family, etc.).

Eligibility to join the study will be decided on a case-by-case basis, after the study team has permission from the family to review the affected family member's medical records. The **Principal Investigator (study doctor)** at each site will make the final decision about eligibility, based on which families have the best chance of finding a disease-causing variant through WGS.

7. WHO SHOULD NOT PARTICIPATE IN THIS STUDY?

You will **not** be able to take part in this study if:

- The affected person in your family has already received a diagnosis through other medical tests or evaluations
- The affected person's condition is thought to be caused by something non-genetic, such as an infection, injury, or toxic exposure (either during pregnancy or after birth)
- The affected person is suspected to have a genetic condition for which there is a simpler and more cost-effective test available to the family for diagnosis
- The family is not able to provide informed consent or complete the steps required for this study

8. WHAT DOES THE STUDY INVOLVE?

If you agree to join this study, you will be asked to do the following steps. Each condition and each family is unique, so not every participant will do all the steps listed below.

Meeting with genetic counsellor (approximately 1 hour)

Once the study team has confirmed your family's eligibility to join the study, the genetic counsellor at your enrollment site will meet with you in person at the medical genetics clinic, by videoconference, or by telephone (depending on what your family prefers).

- The genetic counsellor will describe the study to you in detail, provide information about the different types of genetic results your family could receive from WGS testing, and discuss the possible pros and cons of joining this study.
- The genetic counsellor will ask you questions about the condition in your family and about your **family history**. She/he will draw your family tree to understand the biological

relationships in your family and who is affected by the condition.
- You will have the opportunity to have your questions answered.
- If you agree to join the study, you will be asked to sign this **consent form**

Looking at your medical records
We will ask you to sign an '*Authorization to Release Healthcare Information*' consent form to **review your medical records** so we can better understand the condition in your family, including how it might be affecting you. Looking closely at your medical records will also help us to interpret your WGS results – i.e. whether any DNA variants we may find through WGS match the features of the condition in your family. Your records will be requested from the healthcare provider who referred you to the study and/or from other medical facilities you have visited in the past. The types of records reviewed may include written reports from doctors, results of previous bloodwork (including any genetic testing), and/or imaging results (such as CT-scan or MRI results/images).

If **other family members' medical records** are important in understanding the condition in your family, we will ask you whether it would be OK for you to approach your relative(s) to provide an information letter about the research study and a form to sign if they give us permission to review their records too. This part is optional. We understand some family members may not agree to us reviewing their medical records, but you will still be eligible for the study.

Photography
We may ask to take photographs as a way of documenting the features of the condition in your family. Photos may include your face. Photos can help the study team remember and discuss the condition, especially if the photos show something that cannot be described as well in words. However, photos are optional. You can say 'no' to having photos taken and still take part in this study. **At the end of this consent form, you will be given the choice of whether or not you consent to photos.**

Any photos taken will be part of your research data and will be kept under the same secure conditions as the rest of your data, but separate from your other research data. They will only be available to the study team members. If it is important to include your photos in a future scientific presentation or publication to help other scientists and doctors understand this condition better, we will re-contact you to ask for **separate consent**, and you will have the right to say 'yes' or 'no'.

Digital photos will be kept as required for 5 years after the study is completed, and then they will be destroyed.

DNA sample collection
WGS testing will be done on a DNA sample collected from you. DNA can be collected in different ways, including through a blood, saliva, or cheek-swab sample. A blood sample is the preferred way to collect DNA for WGS, as DNA taken from blood is more likely to give the best results.

We will provide you with the paperwork to have a blood sample drawn at **LifeLabs** (a clinical laboratory with many different locations and partner labs), BC Children's and Women's Hospital, or Victoria General Hospital, depending on which location is most convenient for you.

The paperwork (requisition) provided to the lab and the original blood sample collected from you will have your name and other information that identifies you written on them, so we can be sure there is a correct match to you. Although the lab collecting your sample will know your identity, they will not have access to any of your other research information. A lab staff member will draw your blood sample in a location that is most convenient to you. The blood sample will be taken from a vein, most often the arm, in the usual way. Depending on your age and size, somewhere between 4mL – 10 mL (~1-2 teaspoons) of blood will be drawn. For children, no more than 1 mL (1/5 teaspoon) of blood per 1 pound of body weight will be taken.

In cases where blood collection is not possible, LifeLabs may be able to arrange saliva or cheek-swab collection instead.

DNA handling and storage

After your blood (or saliva or cheek-swab) sample has been collected, it will be sent to **LifeLabs in Toronto, Ontario** for DNA extraction. Once your DNA is extracted, LifeLabs will remove all the information that identifies you (i.e. your name, date of birth, personal health number, address) from your sample, and instead label your sample with a **unique study code and a lab code** which will not allow direct identification of you. Your **coded (de-identified) DNA sample** will then be sent to our research lab (the **Genome Sciences Centre, GSC**) in Vancouver, BC, where WGS testing will be done and your sample will be stored.

Only the study doctor, genetic counsellor and other designated study team members at your local enrollment site will have access to the master 'key' linking your name to your unique study code and research information. Your local enrollment site needs to be able to link back to your name since we will need to provide WGS results back to you and do further confirmation tests if possible disease-causing variants are found.

If a possible disease-causing variant is found through WGS, the GSC will send a **coded** sample of your DNA back to LifeLabs, who will do a **test to confirm the DNA variant** and issue a clinical report in your name.

You are the owner of your DNA, and it is considered to be "on loan" to us for the purpose of this research. Your DNA will only be used for the purposes of this particular study. We will *not* sell your DNA, we will *not* use it to make money, and we will *not* share it with others without your knowledge and permission. If we wish to use your DNA for any other research beyond the purpose of this study, we will come back to you and ask permission. If you want your DNA destroyed or sent back to you at any time, we will do that.

Once this study is complete and the results of the study have been published, any remaining DNA at the GSC will be destroyed unless you specifically request that your sample be returned to you or moved to a different study or biobank that you have provided separate consent for.

Whole Genome Sequencing (WGS) genetic testing

WGS will be done at our research lab (the GSC in Vancouver, BC), using your **coded (de-identified) DNA sample**. All electronic (computer) data resulting from WGS will be stored on a highly secure server at the GSC, and sent to the central study site at **BC Children's Hospital Research Institute (BCCHRI)** through a secure File Transfer Protocol (FTP). All data from your WGS testing will be securely backed up on the BCCHRI servers.

❖ *Types of WGS results*

Our research team will study your data from WGS and create a list of possible disease-causing DNA variants. Decisions about which variants might be disease-causing will be made by our research team, made up of doctors, genetic counsellors, data specialists, and lab staff.
There are four different types of possible results from WGS testing:

1) **A disease-causing variant is found that explains the condition that your family was originally being tested for ('positive' result):** This would provide an explanation for the condition in your family.
2) **No variant is found to explain the condition in your family ('negative' result)**: This does not mean that there is no genetic cause – it may be that we just can't find it with our current technology.
3) **Variant(s) of Uncertain Significance (VUS) are found**: Everyone has many variants in their DNA. Some variants cause disease, others do not. Sometimes there is not enough information available to decide if a variant might cause a health condition. When the meaning of a variant is unclear, we refer to it as a VUS.
4) **Incidental Findings (IFs) are found**: Sometimes, by chance, WGS finds a variant that may cause a completely *different* health condition –something we were not looking for. Please read the next section to learn more about IFs and the choices you will need to make about them.

It is important to understand the limitations of the WGS performed for this study. This study focuses on finding disease-causing variants *related to the specific condition* in the affected person in your family, and not on finding variants that may cause other *unrelated* diseases. Therefore, **the WGS performed for this study is not a full analysis of all your genes, and a 'negative' result does not mean that you do not have any genetic risk factors.**

❖ *Incidental (unexpected) findings*

DNA variants that cause health conditions *unrelated to the original reason for testing* are called **Incidental findings (IFs)**. Although we will not purposely look for variants related to *different* health conditions and will try to lower the chance of finding them, they may be found unexpectedly. IFs can be found in anyone who has WGS testing- even in healthy people.

Some IFs have a relatively high risk of causing health conditions where screening, prevention, or treatments are available to help improve health. These are called '**medically actionable' IFs**. It is not always clear which IFs fit this category, so our team as a whole will review each possible one and come to an agreement about it.

In this study, there are two types of medically actionable IFs that participants might have to face:

1) **Childhood-onset IFs**: These IFs predispose to health conditions where disease features can occur in childhood. Examples of conditions in this category are neurofibromatosis (a genetic condition of the skin and nervous system) and Long QT syndrome (a genetic condition affecting heart rhythm).
2) **Adult-onset IFs**: These IFs predispose to health conditions where features do not usually occur until adulthood. Examples include DNA variants that increase the risk for certain types of cancer or heart disease.

❖ *How will WGS results be given back to you?*

We will provide your WGS results to the healthcare provider who referred you to the study, and it will be his/her responsibility to share the results with you. Your healthcare provider will have the option of including the study genetic counsellor and/or study doctor in the results appointment, and follow-up genetic counselling will be available to you as part of this study.

We will inform your healthcare provider of your results from this study, even if nothing of significance is found. The types of results you may get back include:

- **Negative result,** which means that no DNA variants believed to be related to the condition in the family were found. This result does not guarantee there is no genetic cause- it may be that we just cannot find it with our current technology.

- **Disease-causing variant(s)** that explain the *original condition being tested for* or that explain *some of the features/symptoms* you have.

- **Certain Variants of Uncertain Significance (VUS) that are decided by the study team to be possibly disease-causing** for the *original condition being tested for* or for *some of the features/symptoms* you have.

- **Medically actionable IF** results will be handled in different ways, depending on whether the participant is a child, dependent adult (who has a legal substitute decision-maker), or competent adult who is capable of making his/her own decisions. Our approach, explained below, is based on current Canadian guidelines for reporting IFs:

  - **Child participant**: In the interest of the child, any **childhood-onset** medically actionable IFs found during this study will be given back to the child's referring healthcare provider. We will not usually report back any IFs causing adult-onset conditions; however, there may be some rare cases in which knowing about an adult-onset condition in childhood may be in the family's best interest.

  - **Dependent adult participant (requiring legal substitute decision-maker):** Given that a dependent adult is not able to make his/her own healthcare decisions, we **will** give back results for any medically actionable IFs that are found. This is in the best interest of the participant, to be sure that he/she has the opportunity to receive appropriate healthcare and take preventative actions to reduce harms related to the IF.

  - **Adult participant (competent)**: Adults who are able to make their own decisions have the option of whether or not they wish to be informed of any medically actionable IFs discovered in them. **At the end of this consent form, you will be given the choice of whether or not you wish to be told.** We will discuss the pros and cons of receiving IF results as part of this consent process, and will give you as much time as you need to come to a decision. Whatever your choice is now, you may change your mind in the future by contacting the study team.

    The decision of whether to learn about IFs is a personal one – what is right for one person may not be right for another. Some people like to know all they can about their health risks because they believe it will help them to plan and take preventative

actions. Others would rather not know what conditions they are at-risk for in case, for example, it causes them stress, strains family relationships, or puts them at risk for discrimination. This is further discussed in Section 10 ('*What are the possible harms and discomforts*?'), on pages 9-10 of this consent form.

Please note that if any information related to paternity (who a child's father is) or other unexpected family relationship is discovered during this research study, it will not be disclosed to participants.

❖ *Confirming disease-causing variants and medically actionable IFs*

If a **disease-causing variant**, a **VUS suspected to be disease-causing**, or a **medically actionable IF*** is found through WGS, your result will be confirmed by a repeat genetic test at LifeLabs before being given back to you and your referring healthcare provider. Since these types of results may affect your healthcare, it is important to make sure the result is correct by double-checking it in a **clinical** (non-research) lab before letting you know about it. Once confirmed, LifeLabs will give these results back to your healthcare provider as a **clinical report**, which will include your name and other identifying information.

If a **clinical report** of your genetic test result is issued, *it will no longer be only research data, but will become part of your medical record*, like all your other health records. These reports are also stored long-term in the LifeLabs clinical genetics database. Unlike your research data, clinical reports may be read by any healthcare providers involved in your care, either now or in the future.

**Please note: IFs will not be confirmed or reported back for those adult participants who have opted out of receiving IF results.*

❖ *Re-testing and follow-up*

Throughout the length of this study (approximately 4 years), we will likely re-test your DNA sample and/or re-analyze your data as our knowledge and WGS capabilities grow. Therefore, it is possible that you may receive new variant results at a later date, or that the interpretation of your results may change over the course of the study. New or changed results will be given back as explained in the *'How will WGS results be given back to you?'* section above.

Possible additional tests - optional

Throughout this study, we may suggest extra procedures to help understand the full nature of the condition in your family. This may involve the study of additional tissues such as another blood sample, hair sample, urine sample, skin biopsy, or tissue biopsy from a previous surgery. If this is the case, a **separate consent form** will be presented to you and you can decide whether you want to have the additional studies done.

If a variant is found that we suspect may be disease-causing but additional studies are required to be sure, we may wish to send your sample to a researcher who specializes in that particular gene for further analysis. In such cases, **a separate consent form** will be presented to you. If you consent to sending your sample for further analysis, only your de-identified (coded) sample will be shared with the other researchers and your identity will be protected.

Re-contact for future research studies - optional

At the end of this consent form, you will be given the choice of whether you would like to be re-

contacted about other research opportunities in the future.

9. WHAT ARE MY RESPONSIBILITIES?
Please keep in contact with us to let us know if you change your address or contact details. If there are significant health changes in the study participant related to the suspected genetic condition being tested for, it would be helpful for you to let us know.

10. WHAT ARE THE POSSIBLE HARMS AND DISCOMFORTS?
Discomfort of blood draw
The risks of drawing blood may include some minor local discomfort, light-headedness (dizziness or fainting), and/or minor infection or bruising around the area where the needle was inserted. If you are planning to have blood drawn as part of your routine clinical care, we will make every effort to coordinate your study blood draw so it can be done at the same time.

Risks related to WGS and Incidental Findings (IFs)
❖ *Potential psychological impact*:
When you donate your blood or tissue for genetic testing or research, you are sharing genetic information, not only about yourself, but also about biological (blood) relatives who share your DNA. If you learn that you have a genetic condition, there may be other family members, including those alive now or those who may be born in the future, who may also be at risk of having the same condition. This knowledge may provide you or your family with important information that could be used either to prevent the disease (if possible) or to inform other health care decisions. However, there is also a risk that simply having this knowledge may cause worry or stress. For some people, genetic results may influence their decision to have a child or affect other lifestyle decisions. This is one of the reasons that families have access to a genetic counsellor as part of the study. Genetic counselling helps each family to assess the pros and cons of the testing and understand what your results might mean for your wider family.

❖ *Potential risk of disclosure of your genetic information:*
There is a small risk of loss of privacy for you and your family. Despite our best efforts to protect your privacy, we cannot guarantee that your data will remain completely anonymous in all situations. If we decide to publish results from this study in a scientific medical journal, we will not include any information about your identity (such as your name, community, birth date, etc.). However, it is possible that somebody somewhere - perhaps, for example, someone who looked after you in hospital - may realize the unnamed person in the study could be you. Studies published in medical journals are distributed worldwide and are mainly read by doctors and scientists, but could be read by non-doctors too.

There is always a risk of computer systems or research offices being broken into and data stolen, although the chance is estimated to be very small. Every effort will be made to protect your privacy and the confidentiality of your genetic results.

If you decide to reveal your genetic test results to others, there is a chance it could lead to discrimination towards you and/or your blood relatives. A federal government law enacted in May 2017, called the *Genetic Non-Discrimination Act* bans discrimination based on genetic characteristics and makes it illegal for insurance companies (such as life insurance companies) and employers to require people to reveal their genetic test results. Insurance companies and employers are not allowed to use your *genetic test results* against you (for example, they cannot use this information to decide whether to offer you life insurance or

whether to hire you for a job). See *GNA* fact sheet available at: https://www.cagc-accg.ca/doc/S201%20fact%20sheet%20-%20final%20copy%20-%20May%2017%202017.pdf Although this law is helpful, it does not prevent insurance companies from using information about your current symptoms or family history of a health condition, or the results of *other types of medical tests* to decide if they will insure you (even if that health condition has a genetic basis). Also, laws sometimes change over time, so no one can guarantee that the *GNA* will always be in place to protect you.

As discussed in the '*Confirming disease-causing gene variants and medically actionable IFs*' section on page 9, if you receive a **clinical report** of your genetic test results from LifeLabs, it will become part of your medical record and may be read by any healthcare providers involved in your care, either now or in the future. These reports are also stored long-term in the LifeLabs clinical genetics database. Having genetic test results enter your medical record would remove your choice to keep your results private from your healthcare providers.

11. WHAT ARE THE POSSIBLE BENEFITS OF PARTICIPATING?

No one knows whether or not you will directly benefit from taking part in this study. It is possible that you will receive a genetic explanation for the condition in your family, which in some cases is helpful in guiding the medical care and support of affected family members. Some people find it helpful to learn the reason for the medical condition in their family. We hope that the information learned from this study can be used in the future to benefit other people with suspected genetic conditions.

12. WHAT ARE THE ALTERNATIVES?

It is important for you to know that you can choose not to take part in the study. Choosing not to participate will in no way affect your health care. The availability of genomic testing varies across Canada – If you choose not to join this study, WGS testing may or may not be available through the medical system in your region.

13. WHAT IF NEW INFORMATION BECOMES AVAILABLE THAT MAY AFFECT MY DECISION TO PARTICIPATE?

You will be told of any new information that becomes available that may affect your willingness to remain in the study. You may be asked to provide renewed consent if new information is learned that may affect your decision to take part in the study.

14. WHAT HAPPENS IF I DECIDE TO WITHDRAW MY CONSENT TO PARTICIPATE?

You may withdraw from this study at any time without giving reasons. If you choose to enter the study and then decide to withdraw at a later time, you have the right to request the withdrawal of your information (and/or samples) collected during the study. This request will be respected to the extent possible. Please note however that there may be exceptions where the data (and/or samples) will not be able to be withdrawn, for example where the data (and/or sample) is no longer identifiable (meaning it cannot be linked in any way back to your identity) or where your data have been combined with other data. If you would like to request the withdrawal of your data (and/or samples), please tell the study doctor at your site. If your participation in this study includes any optional studies or long-term follow-up, you will be asked whether you wish to withdraw from these as well.

15. CAN I BE ASKED TO LEAVE THE STUDY?

If you are not able to complete the steps of the study, we may ask you to withdraw from the

study. If you are asked to leave the study, the reasons for this will be explained to you and you will have the chance to ask questions about this decision. You could also be removed from the study if we are no longer able to contact you (e.g. moved, no forwarding address).

16. HOW WILL MY TAKING PART IN THIS STUDY BE KEPT CONFIDENTIAL?

Your confidentiality will be respected. However, research records and health or other source records identifying you may be inspected in the presence of the Investigator or his/her designate by representatives of the study sponsors, or the Research Ethics Board at the University of British Columbia or Island Health for the purpose of monitoring the research. No information or records that disclose your identity will be published without your consent, nor will any information or records that disclose your identity be removed or released without your consent unless required by law.

Your rights to privacy are legally protected by federal and provincial laws that require safeguards to ensure that your privacy is respected. You also have the legal right of access to the information about you that has been provided to the sponsor and, if need be, an opportunity to correct any errors in this information. Further details about these laws are available on request.

You will be assigned a unique study code as a participant in this study. This code will not include personal information that could identify you (e.g. it will not include your name, health card number, SIN, date of birth, address etc.). This code will be used on your data and samples collected during the course of this study, so that your identity will be kept confidential. Only the study doctors and other designated study team members at your local enrollment site will have access to the master 'key' linking your name to your unique study code and research-related information. Although LifeLabs and the lab involved in your blood collection will know your identity and unique study code, they will not have access to any of your research data. The key matching your name to your study code will not be released to anyone else without your consent, unless required by law.

Most of your study data (such as your WGS data, medical details about your condition, photographs, etc) will be labelled only with your unique study code and will be stored in password-protected files on computer servers protected by strict security measures. Any photographs taken of you will be stored in a separate file from your other study data. The 'key' matching your name to unique study code will be stored on a secure server at your local study doctor's institution, in a separate password-protected file, and will not be sent outside of your local enrollment site

Hard-copies of paperwork which could contain your personal information (i.e. medical records received for this study, signed consent form, family tree drawn by the genetic counsellor during the research visit, and records of the study team's communications with you) will be kept in a file in a lockable cabinet behind a locked door at your study doctor's institution and will only be accessed by the study doctor and genetic counsellor at your local site.

If we present or publish any study results for other scientists or health care providers, we will not include any personal information about you. If we wish to include your photographs in a future scientific presentation or publication, we will re-contact you to ask for separate consent, and you will have the right to say 'yes' or 'no'. Any scientific papers we write from this study will be reviewed by our Indigenous oversight committee before being published.

As discussed in previous sections, if you receive a **clinical report** of your genetic test results from LifeLabs, it will contain your name and other identifying information on it and will become part of your medical record. It will also be stored long-term in the LifeLabs clinical genetics database. As part of your medical record, it may be read by any healthcare providers involved in your care, either now or in the future. Safeguards are in place, as the privacy of medical records is protected by law.

17. WHAT HAPPENS IF SOMETHING GOES WRONG?
By signing this form, you do not give up any of your legal rights and you do not release the study doctor, participating institutions, or anyone else from their legal and professional duties. If you become ill or physically injured as a result of participation in this study, medical treatment will be provided at no additional cost to you. The costs of your medical treatment will be paid by your provincial medical plan and/or by the study sponsors.

18. WHAT WILL THE STUDY COST ME?
All research-related medical care and any tests that you have during your participation in this study will be provided at no cost to you.

19. WHO DO I CONTACT IF I HAVE ANY QUESTIONS ABOUT THE STUDY DURING MY PARTICIPATION?
If you have any questions or would like further information about the study at any time, or if you feel you have suffered any adverse effects from the study, **please contact:**

Victoria site – Sarah McIntosh (genetic counsellor) and Dr. Laura Arbour (study doctor) at:

Vancouver site – Karen Jacob (study coordinator/genetic counsellor) at:

20. WHO DO I CONTACT IF I HAVE ANY QUESTIONS OR CONCERNS ABOUT MY RIGHTS AS A PARTICPANT?
If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, contact the Research Participant Complaint Line in the University of British Columbia Office of Research Ethics by e-mail at . Please reference the study number (H18-00726) when contacting the Complaint Line so the staff can better assist you.
You may also contact the Vancouver Island Health Authority (VIHA) Research Ethics Board at

21. AFTER THE STUDY IS FINISHED:
This study is expected to last at least 4 years, although there is a possibility the study will be renewed for a longer period. As required, the data from this study will be stored for at least 5 years after the study is finished. Once all the testing and data analysis are complete and the findings are published, any remaining DNA samples will be destroyed, unless you specifically request that they be returned to you or transferred to another study or biobank that you have provided separate consent for. If you prefer that we return any remaining sample to you instead of destroying it at the end of the study, you may request this by contacting the study team (see section 19, above).

PRECISION DIAGNOSIS FOR INDIGENOUS FAMILIES WITH GENETIC CONDITIONS
PARTICIPANT CONSENT

My signature on this consent form means:

- I have read and understood the information on this consent form.
- I have had enough time to think about the information provided.
- I have been able to ask questions and have had satisfactory responses to my questions
- I understand that my participation in this study is voluntary.
- I understand that I am completely free at any time to refuse to participate or to withdraw from this study at any time, and that this will not change the quality of care that I receive.
- I authorize access to my health records as described in this consent form.
- I understand that I will be informed of any incidental findings (IFs) identified in a child participant that could alter his/her medical management during childhood.
- I understand that adult participants have the option of choosing whether or not they wish to be informed of any IFs identified in them that may alter their medical management.
- I understand that if a disease-causing or likely disease-causing variant is found in me, then my genetic test result will be confirmed in a clinical lab and will become part of my medical record, which all of my healthcare providers can look at.
- I understand that there is no guarantee that this study will provide any benefits to me.
- I understand that I am not waiving any of my legal rights as a result of signing this consent form.

I will receive a signed copy of this consent form for my own records. I consent to take part in this study.

ADULT PARTICIPANT consent:

_____	_____	_____
Printed name (adult participant)	Signature (adult participant)	Date
_____	_____	_____
Printed name & role of person obtaining consent	Signature	Date

OR

PARENTAL/GUARDIAN consent for CHILD participant (<18 years old) or LEGAL SUBSTITUTE DECISION-MAKER consent for dependent adult participant who is incapable of consent:

The parent/guardian or substitute decision-maker (legally authorized representative) and the investigator are satisfied that the information contained in this consent form was explained to the child/participant to the extent that he/she is able to understand it, that all questions have been answered, and that the child/participant assents to participating in the research.

Printed name of child: _____ OR adult participant (ward): _____

_____	_____	_____
Printed name of parent/guardian OR substitute decision-maker	Signature	Date
_____	_____	_____
Printed name & role of person obtaining consent	Signature	Date

OPTIONAL CONSENT CHECKBOXES – COMPETENT ADULT participant (Form 1)

Printed name of adult participant: _____

OPTION 1 –Photographs: Photographs of me, including photos of my face, may help the study doctors better understand the condition in my family and help with the interpretation of my Whole Genome Sequencing results. I understand that I have the option of deciding whether or not to have photos taken. Any photos taken will only be viewed by the study team members. They will have the same protections as the rest of my research data but will be stored separately from my other research data. If the study doctors feel it is important to include my photos in a future scientific presentation or publication, I will be re-contacted and asked for **separate consent**, and I will have the right to say 'yes' or 'no' to this.

Please check one box below.

☐ **YES, I agree** to have photographs taken of me for the purposes of this research study.

☐ **NO, I do NOT agree** to have photographs taken.

OPTION 2 - Incidental (Unexpected) Findings (IFs): I understand that I can choose whether or not I wish to be informed of any incidental findings that could alter my health management. Whatever I choose now, I can change my decision at any time by re-contacting the study team.

Please check one box below.

▫ **YES, I DO** wish to be informed of any 'medically actionable' incidental findings found in me.

▫ **NO, I do NOT** wish to be informed of any incidental findings found in me.

OPTION 3 - Re-Contact for Future Research Studies:

I understand that new research studies may be of interest to my family. I can choose whether I wish to be re-contacted about future research opportunities. If I agree to be re-contacted, it does not mean I have to participate in any future research projects. I will have the right to say 'yes' or 'no' to participating in any future studies that are presented to me.

Please check one box below

☐ **YES, I agree** to be contacted in the future to learn about a new research study.

☐ **NO, I do NOT agree** to be contacted in the future to learn about a new research study.

PRECISION DIAGNOSIS FOR INDIGENOUS FAMILIES WITH GENETIC CONDITIONS

OPTIONAL CONSENT CHECKBOXES–CHILD /DEPENDENT ADULT participant (Form 2)

Printed name of child participant: _____

OR

Printed name of dependent adult (ward): _____

OPTION 1 – Photographs: Photographs of my child/ward, including photos of his/her face, may help the study doctors better understand the condition in the family and help with the interpretation of Whole Genome Sequencing results. I understand that I have the option of deciding whether or not to have photos of my child/ward taken. Any photos taken will only be viewed by the study team members, and will have the same protections as the rest of my child's/ward's research data but will be stored separately from the other research data. If the study doctors feel it is important to include my child's/ward's photos in a future scientific presentation or publication, I will be re-contacted and asked for **separate consent**, and I will have the right to say 'yes' or 'no' to this.

Please check one box below.

☐ **YES, I agree** to have photographs of my child/ward taken for the purposes of this research study.

☐ **NO, I do NOT agree** to have photographs of my child/ward taken.

OPTION 2 – Re-Contact for Future Research Studies

I understand that new research studies may be of interest to my family. I can choose whether I wish to be re-contacted about future research opportunities for my child/ward. If I agree to be re-contacted, it does not mean my child/ward has to participate in any future research projects. I will have the right to say 'yes' or 'no' to participating in any future studies that are presented to me.

Please check one box below.

☐ **YES, I agree** to be contacted in the future to learn about a new research study.

☐ **NO, I do NOT agree** to be contacted in the future to learn about a new research study.