Research Proposal

Tentative Proposed Start Date:
June 15th
Where will I be conducting research?
I will be in Toronto, Canada for the duration of my research.
Introduction

Cerebral Palsy (CP) is a group of developmental conditions characterized by injury to the developing brain causing motor, speech, vision, or hearing impairments, and intellectual disabilities[1]. In most industrialized countries prevalence is approximately 1 in 500 live births[2]. CP is thought to be caused by both environmental and genetic factors creating a brain injury before or during birth, or within the first few years of life[3]. Environmental risk factors include prematurity, birth asphyxia, and exposure to infection. Population studies have also shown evidence for a genetic contribution to risk[4]; however, these contributions are poorly understood.

The Centre for Applied Genomics [5](TCAG) at The Hospital for Sick Children (SickKids) is a world-leading genetic research centre. With the Ontario Brain Institute funded “CP-NET” program[6], the group has collected DNA from 216 patients with CP and their parents, and completed whole genome sequencing (WGS) on these families. These genomes are being analyzed to detect clinically relevant changes: copy number variations (CNVs), single nucleotide variants (SNVs), and more complex changes, all of which may be relevant to CP. Other studies have identified variants through microarray[7] or exome analysis[8,9], but no studies have used WGSin a large cohort of CP families.

The aim of my research is to identify the occurrence and frequency of CNVs and SNVs in selected genes associated with CP and brain function. I would compare these variants to existing databases to identify genes recurrently changed in CP patients, and to identify associations with relevant clinical risk factors. This research would contribute to the broader research question: what are the genetic underpinnings of Cerebral Palsy?

This research combines my interest in molecular genetics with my personal connection to SickKids, where as a toddler I underwent surgical repair of a cleft palate, graduating from the

---

[1] CP-Net. (2021). About Cerebral Palsy. *CanChild*. Retrieved from http://cpnet.canchild.ca/en/about-cp/about-cerebral-palsy
[2] Colledge N. (2001). A Guide To Cerebral Palsy. *The Ontario Federation for Cerebral Palsy*. Retrieved from https://www.ofcp.ca/pdf/Web-Guide-To-CP.pdf
[3] Nelson KB. Causative factors in cerebral palsy. (2008). *Clin Obstet Gynecol*, *51*(4):749-62. doi: 10.1097/GRF.0b013e318187087c.
[4] Tollånes Mette C, Wilcox Allen J, Lie Rolv T, Moster Dag. (2014). Familial risk of cerebral palsy: population based cohort study. *BMJ, 349*
[5] TCAG. (2021). The Centre For Applied Genomics. *SickKids*. Retrieved from http://www.tcag.ca/
[6] CP-Net. (2021). CP-Net. *CanChild*. Retrieved from http://cpnet.canchild.ca/en/
[7] Zarrei M, Fehlings DL, Mawjee K,...Scherer SW. (2018). De novo and rare inherited copy-number variations in the hemiplegic form of cerebral palsy. *Genet Med, 20*(2):172-180. doi: 10.1038/gim.2017.83.
[8] Oskoui M, Gazzellone MJ, Thiruvahindrapuram B,...Scherer SW. (2015). Clinically relevant copy number variations detected in cerebral palsy. *Nat Commun, 6*:7949. doi: 10.1038/ncomms8949.
[9] Corbett MA, van Eyk CL, Webber DL,...Gecz J. (2018). Pathogenic copy number variants that affect gene expression contribute to genomic burden in cerebral palsy. *Genom. Med, 3*(33). https://doi.org/10.1038/s41525-018-0073-4

Cleft Palate program. Through this research, I hope to learn more about the biological and clinical application of molecular genetics, and contribute to the scientific understanding of the mechanisms of a common condition. I am eager to work in a well-renowned group, and learn from researchers from many scientific backgrounds.

Methods

Overview

There are two possible scenarios for this project. Current COVID-19 guidelines at the SickKids Research Institute would allow me to work in the lab. If this remains the case, I will biologically confirm findings of SNVs identified bioinformatically from WGS data. To investigate SNVs, I will design primers (short stretches of synthetic DNA, acting as start- and end-points) and amplify through polymerase chain reaction (PCR) short regions of DNA containing candidate variants. I will then purify the PCR products and submit them to the TCAG Sanger sequencing facility. I will follow up with a literature investigation to predict possible functional impact(s) of each variant by searching up the gene and specific mutation in Online Mendelian Inheritance in Man[10], NCBI Gene[11], and PubMed databases. This will include exploring the functions of my chosen genes. While the exact genes that I will be exploring are not finalized, I will likely be looking at about 12 genes that have previously been linked to CP.

Should I need to complete this project remotely, I will still design and order primers, but the PCR and qPCR processes described above will be completed by others in the lab. I will then focus more on the latter part of the process and investigate functional impacts of variants. With the additional time, I will further explore these variants, and look for associations with specific CP clinical subtypes, using information from the CP-NET Clinical Database. If time allows, I will also look for associations between the relevant SNV findings and known CP clinical risk factors including advanced maternal age, fetal exposure to toxins, brain bleeds, or prematurity[12].

Training and Certification

Since the lab in which I will be working is at SickKids hospital, I am required to undergo SickKids summer student training. This involves online modules on Occupational Health and Safety, Workplace Violence, Privacy, Accessibility, Fire Safety, Error Prevention, and Lab Safety. I have started these modules, and they total to about 3 days worth of training.

Involved Organizations and Supervisors

The organizations involved in my work are:

- SickKids: The hospital where my lab is based and where my training comes from.
- TCAG: My lab (The Centre for Applied Genomics).
- Dr. Stephen Scherer: The director of TCAG.
- Dr. Richard Wintle: The assistant director of TCAG, my more direct contact as a supervisor.
- Dr. Mehdi Zarrei: A research associate at TCAG who works on Cerebral Palsy and will be guiding me through the process of literature searches.

Timeline

---

[10] Online Mendelian Inheritance in Man. (2021). OMIM. *Johns Hopkins University*. Retrieved from https://www.omim.org/

[11] National Center for Biotechnology Information. (2021). Gene. *U.S. National Library of Medicine*. Retrieved from https://www.ncbi.nlm.nih.gov/gene/

[12] Nelson KB. Causative factors in cerebral palsy. (2008). *Clin Obstet Gynecol*, *51*(4):749-62. doi: 10.1097/GRF.0b013e318187087c.

- Week 1: Primer Design, Training (Practice literature search)
    - For my chosen genes, I will design primers using software such as UCSC Genome Browser, UCSC In-Silico PCR, and UCSC Blast. I will then order these primers, which usually take 2-4 days to arrive. In my first week I will also be meeting frequently with my supervisors for training and guidance. This will likely involve me completing a practice literature search of 1 gene that will be checked by my supervisors to ensure that I am doing the search correctly.
- Week 2: Genes 1-3 Literature Search, Primer Testing
    - In the second week, I will begin my literature search for the first few genes, looking thoroughly through databases such as OMIM, ClinVar, and NCBI gene and taking notes of any information I find on my mutation of interest or related mutations. I will also be testing the primers that I designed through PCR to ensure that they work. If I am working remotely, the primer testing will be replaced by a deeper literature search on gene function and connections to specific types of CP.
- Week 3: Genes 4-6 Literature Search, PCR and sequencing of genes 1-4
    - In the third week, I will continue my literature search on the next set of genes. I will also perform PCR on the first few genes and prepare them for sequencing. If I am working remotely, once again I will do a more thorough literature search. However, I will still be doing the prep work for the PCR (calculating volumes etc.) and analyzing the sequencing results, so this will still take up a significant amount of my time.
- Week 4: Genes 7-9 Literature Search, PCR and sequencing of genes 5-8
    - In the fourth week, I aim to continue the work from the third week but expand to cover more genes.
- Week 5: Genes 11 and 12 PCR and sequence genes 8-12
- Week 6: Sequencing result analysis and connection to literature information
    - In the sixth week, I will finish up any PCR or sequencing and then focus on analysis of the sequencing results. Analysis of sequencing results involves identifying the exact location of any mutations as well as whether the mutation is homozygous or heterozygous.
- Week 7: Finish any left work
    - Although I hope to stick to this schedule, I anticipate that I will inevitably have some leftover work at the end of 6 weeks. In the seventh week I will finish any loose ends and check over my work with my supervisors.

Research advisors

My primary research advisor is Dr. Stephen Scherer, the Scientific Director of TCAG and UofT's McLaughlin Centre, and a Senior Scientist at SickKids. Leading the CP project is Dr. Richard Wintle, TCAG's Assistant Director and Associate Scientific Director of CP-NET. TCAG has a world-leading, multi-disciplinary team with whom I would work, to determine which genes to prioritize, and to review previous findings. Specifically, Dr. Mehdi Zarrei has been leading the analysis of genome sequencing data in various neurodevelopmental conditions, and will supervise me[13]. Relevant research ethics approvals are in place and participants have consented for analysis of their genomes. If I complete the project remotely, TCAG will also provide data on variants and relevant clinical and demographic information of the participants. I will attend bi-weekly meetings with Drs. Scherer, Wintle, and Zarrei to monitor progress, and agree on benchmarks.

International Focus

This research is globally relevant because CP is the most common childhood physical disability in the world, with estimated incidence ranging from 1-4 per 1000 live births worldwide[14]. Motor impairment and other comorbidities such as seizures, intellectual disabilities, pain, sleep disorders, and problems with vision and hearing make CP extremely debilitating[15]. CP is also extremely costly to both families and the healthcare system: in 2005, average medical costs of children in the United States with CP were 10 times greater than average costs for unaffected children[16]. In many countries, the high cost of CP means that mortality is drastically increased for children with CP, as families cannot afford the necessary medical care[17].

Outcomes

The main outcome of my project will be the identification and exploration of clinically relevant mutations that may be connected to CP in selected genes. This includes identifying the function of the genes, the possible functional impacts of the mutations, the zygosity of the mutations (whether they come from both parents or just one), and whether these variants have been identified by others. If time allows, additional results may include associations between variants and specific types of CP specific clinical risk factors. On a larger scale, I hope that this experience will help me better understand lab-based research methods as well as how to effectively search literature.

Impact

An improved understanding of the genetic causes of CP will greatly benefit CP patients, families, and healthcare systems. These benefits include identification of at-risk mothers, earlier and more accurate diagnoses, and streamlining of patients to more specific therapeutic regimens. Additionally, understanding the genetic components of CP can inform future investigations of novel therapies based on each individual's genetic makeup (personalized, or precision, medicine). Overall, this research will increase understanding of CP and help children with CP receive rapid diagnoses and potentially novel treatments.

My research project is a small step in understanding the genetic underpinnings of CP. This begins with identifying relevant CNVs and SNVs from the genomes of 216 families,

---

[13] Zarrei M, Burton CL, Engchuan W,...Scherer SW. (2019). A large data resource of genomic copy number variation across neurodevelopmental disorders. NPJ *Genom Med*, *4*(26). doi: 10.1038/s41525-019-0098-3.

[14] Centers for Disease Control and Prevention. (2020). Cerebral Palsy Data and Statistics. *U.S. Department of Health & Human Services*. Retrieved from https://www.cdc.gov/ncbddd/cp/data.html

[15] World CP Day. (2017). What is Cerebral Palsy? *The Allergan Foundation*. Retrieved from https://ofcp.ca/wp-content/uploads/2017/09/WorldCPDay_WhatisCP_poster_World_2017.pdf

[16] Centers for Disease Control and Prevention. (2020). Cerebral Palsy Data and Statistics. *U.S. Department of Health & Human Services*. Retrieved from https://www.cdc.gov/ncbddd/cp/data.html

[17] Khandaker G, Jahan I, Das M,...Muhit M. (2018). Mortality among children with cerebral palsy (CP) in rural Bangladesh: results from the Bangladesh CP Register. *Dev Med Child Neurol, 60*(3) 13-13. https://doi.org/10.1111/dmcn.17_14017/

comparing variants to previous findings, and associating them with each participant’s clinical risk factors.