Variant Calling for Primary Angle Closure Glaucoma

Introduction:

Glaucoma is a group of heterogeneous optic neuropathies that affects approximately 70 million individuals worldwide. It is the second leading cause of irreversible blindness worldwide. Types of Glaucoma are characterized by progressive and irreversible destruction of the optic nerve and degeneration of retinal ganglion cells (RGCs). In 2010, an estimation 60.5 million of people were diagnosed with glaucoma (primary open angle glaucoma (POAG) and primary angle closure glaucoma (PACG) combined) worldwide, which will increase significantly in the future.

Primary angle-closure glaucoma is a complex heterogeneous disease, with the genetic susceptibility under investigation. Due to the high prevalence among Inuit's and Asians compared to Caucasians, suggesting a genetic predisposition for the disorder.

Also, there is an unusually high incidence of PACG among siblings of affected patients, it was suggested that genetic factors were involved in its pathology and the action of a large number of grouped or independently inherited genes along with environmental factors result in anatomical abnormalities of PACG.

Aim:

A pilot study will be performed to identify the possible genetic variants that involved in Angle Closure Glaucoma (PACG) pathology.

Methods:

1. Three samples and 2 controls from "Exome-Seq of Homosapiens: PACG" will be selected as SRA data type from NCBI https://www.ncbi.nlm.nih.gov/sra?linkname=bioproject_sra_all&from_uid=394051.

2. Each member of our team will use a subset of 8 million paired end reads to achieve about 46x coverage of the whole exome. The exome length is about 1.1% of the total genome, or about 30 mega bases of DNA and our reads are 93*2, that is due to our limited computational resources as the coverage should be about 80x for WES.

3. The quality of the data will be tested by FastQC then the data will be trimmed if needed.

4. The alignment step will be performed by an appropriate aligner, probably hisat2 as a splicing aware aligner. One chromosome will be used as reference for the alignment. We will choose the chromosome according to literatures which defined the chromosomes that have gene mutations for PCAG, also it's the smallest in size, as our computational resources are limited.

5. After the alignment, we will use the GATK for the variant calling and joint variant calling using HaplotypeCaller. Needed resources, such as the known PACG variants for the chosen chromosome, will be used. The final expected results will be the PACG SNPs and INDELs.

Due to our limited resources, we expect that we may face some technical issues and the results may not be accurate enough.

All team members will have an equal contribution in this project.

Team Members:

Samah Magdy Ahmed Mahdy (teamleader)	ID:191066
Dalia Mahmoud Ezzat	ID:191067
Fayrouz Ali El-Baioumy Muhammad	ID:191065
Marwa Ahmed Abd Al-Azim	ID:191068
Marwa Mohamed Hassan Aswa	ID:191069