SE 315 – SOFTWARE PROJECT MANAGEMENT
Spring 2019-2020
Project Proposal

Problem Definition

We are living in a globalized world, and as a result of this, there is an increase in the number of inter-race couples. Thus, the gene combinations that cause genetical diseases also increased. Since it is a long process to find treatment for genetic diseases, we are aiming to prevent them from begining. In order to do that we combined software technology and genetical knowledge together to create an algorithm to find the possibility of these genetichal diseases.

Background Information

After the discovery of independent assortments and segregations of genes by Gregor Mendel, modern genetic have gained momentum since the 19th century. Nowadays scientists developed this knowledge into advance. At present, while a couple planning to marry, they check their blood types to prevent later complications such as RH disease. What we are planning to do is to bring forward this process and estimate the couples possible offsprings by combining the features of software computing and genetical knowledge.

Objectives

- Generating synthetic genomes
- Taking genetic information from partners
- Checking the similarities between participants genome and ex clinical results
- Determining mutations in participants genome
- Simulating crossing over
- Creating possible offsprings of both participants
- Giving the possibility of participants' children having a genetical disease

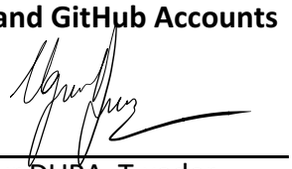
Scope

During our design, As a team, we decided to scale down the size of the human genome. The reason being is the human genome contains 4.5 billion base pairs and around 20,449 different genes(Ensemble genome browser release). It may create some implantation difficulties. Instead of checking each one of them, we just focused on the common genes which can generate major problems such as haemoglobin beta gene (HBB) or cellular tumour antigen p53. We decided to generate a synthetic genome which we defined the genes inside already. We choose GRCh38 reference genome as our template and we add inside of this template Cystic fibrosis transmembrane conductance regulator gene (CFTR-In the case of mutation of this gene causes Cystic Fibrosis), Hemoglobin beta gene (HBB-In the case of mutation of this gene causes Sickle cell anaemia), Oculocutaneous albinism II gene (OCA2-In the case of mutation of this gene causes Albinism), Phenylalanine hydroxylase gene (PAH-In the case of mutation of this gene causes Phenylketonuria), Huntington gene (HTT-In the case of mutation of this gene causes Huntington disease), Neurofibromin gene (NF1- In the case of mutation of this gene causes Neurofibromatosis type I) and Tumor protein p53(TP53- In most cancer cells, this gene found as damaged). We have ongoing discussions about adding one more chromosome which also includes some basic genes which define eye or hair colour or stature near this synthetic chromosome which contains potential genetical disease genes. While we were working on the synthetic genome, we added telomer sides to the both ending part, some noncoding parts and also crossing over sites into the genome to increase accuracy. Besides, instead of using basic algorithms to compare which checks

SE 315 – SOFTWARE PROJECT MANAGEMENT
Spring 2019-2020
Project Proposal

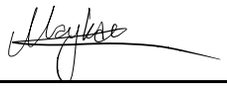
only the difference with the reference genome, we decided to combine clinical results to increase presision. The reason for this is examining ex clinical results with software computation is easier and advisable when we compare to understanding whole signalling pathways, genetic correlation and originate connection between them.
The program starts the process by taking genetic information from partners. Then, the program check against differences between reference genes. If there any mismatch, the program starts to check against ex clinical results to find a similarity between participants genome and ex clinical results. After the determination of the mutations of partners, programs start to simulate the crossing over to create possible offsprings of both participants. In the final stage, the program combines these offsprings and examines the new possible genomes with the reference genome to give the result of their child according to the possibility of getting a genetic disease.
This Project may lead the way for designer baby tools. In case of participants want to have a baby, they will have an opportunity to change damaged parts of their offspring to prevent later complications.

Approval Signatures and GitHub Accounts



Ugur DURA, Twodor

Zeynep Ovgu YAYCI , zovgu



Mustafa AYKAC, Maykac

Cem OZCAN, cem1101

Burak ERDOGAN, burkerdgn