Designing a program to detect mutations in the human gene using python, application software engineering principles to obtain a real life application while maintaining simplicity of design

Preface

We will implement an application that detects Mutation in Human DNA Codons, two types are to be investigated here: Point and Insertion mutations.

Following the system design, we use python as the implementation language, where dictionaries are the main structure that will meet our program needs.

The menu driven graphical user interface used in the application is the tkinter which comes with the python by default and doesn't need extra libraries or installations to use, the python version used is 3.5.1

Flow Charts

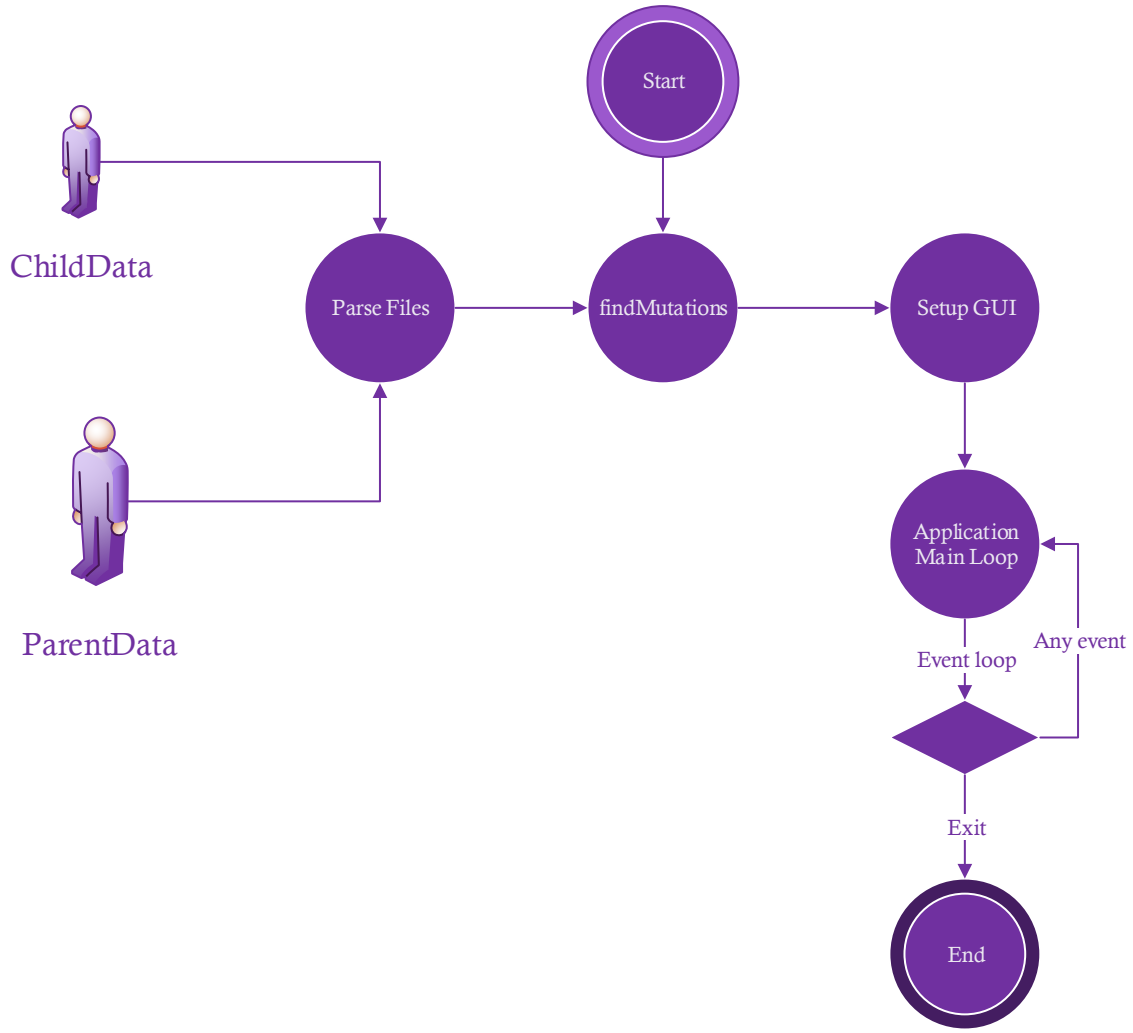

Figure 1: Overall Flow chart of the program.

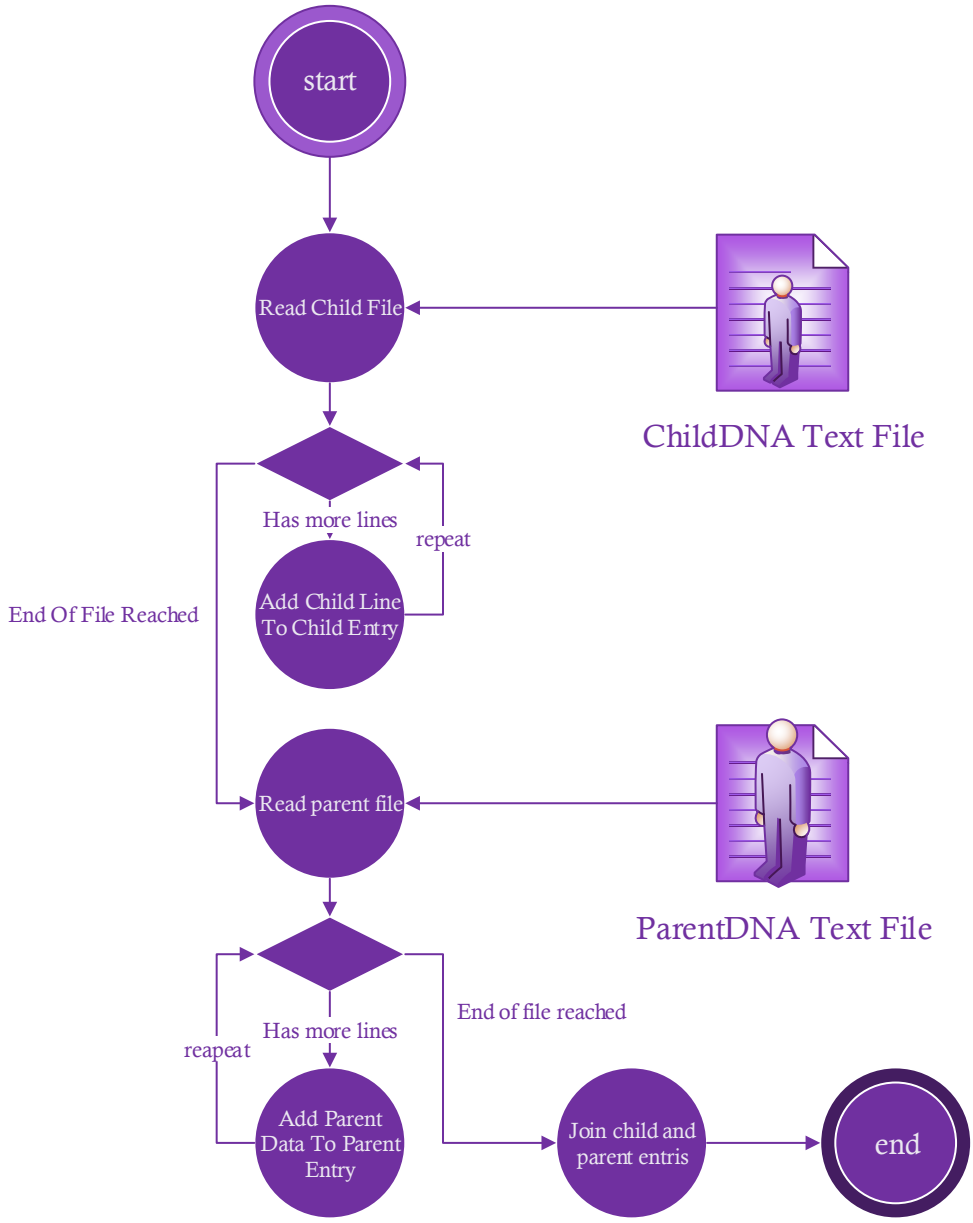

Figure 2: Parse Data Files

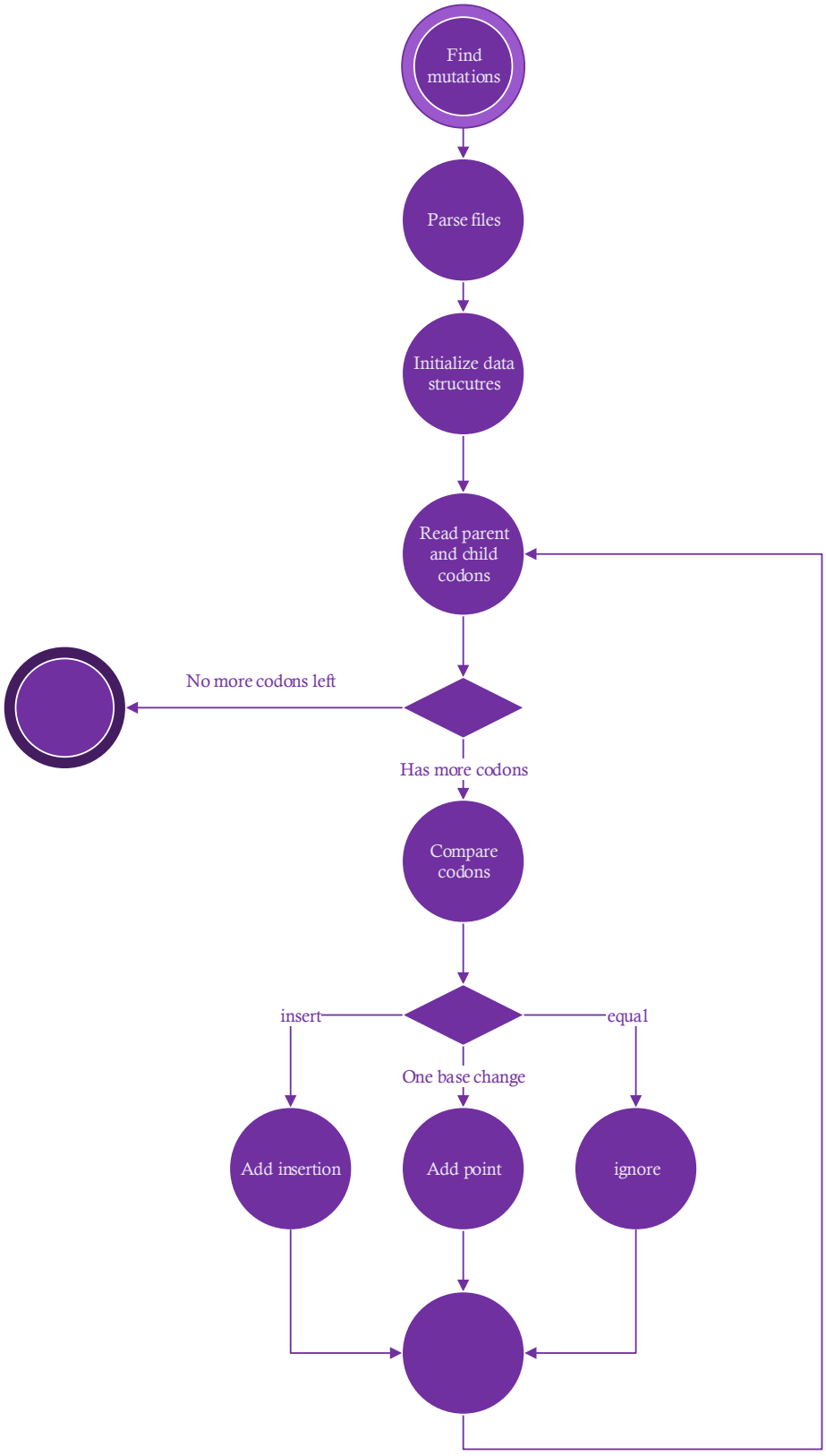

Figure 3: Find Mutations

Psudo code:

The main functions we have can be summarized as the following pseudo code

def parseFile ():

```
sampleData={} #contains two arrays, child and parent
open parent and child files
cData=[] #child array
pData=[] #parent array of dictionary values
for each line get child sample data
for each line get parent data
add data subsets to the data dictionary and return it
Close files
```

Def findMutations():

```
data=parseFiles()
cData=data["Child"]
pData=data["Parent"]
list=[] # dict of{name,sex,mutation index,desc} key is the name
#for each line in the data
  #read data
  # for each three chars in the dna sequence
      #if parent codon!= child codon then it is a mutation
          #if it is one base change then add a point mutation
          # Else add insertion mutation
# add remaining codons as insertion mutation
```

def setupGUI():

```
app = AppWindow(None)
app.title('DNA Mutations')
initialize application and setup controls and their events
app.mainloop()
```

Data Structures: all are dictionary bases data structures.

Tests and Results

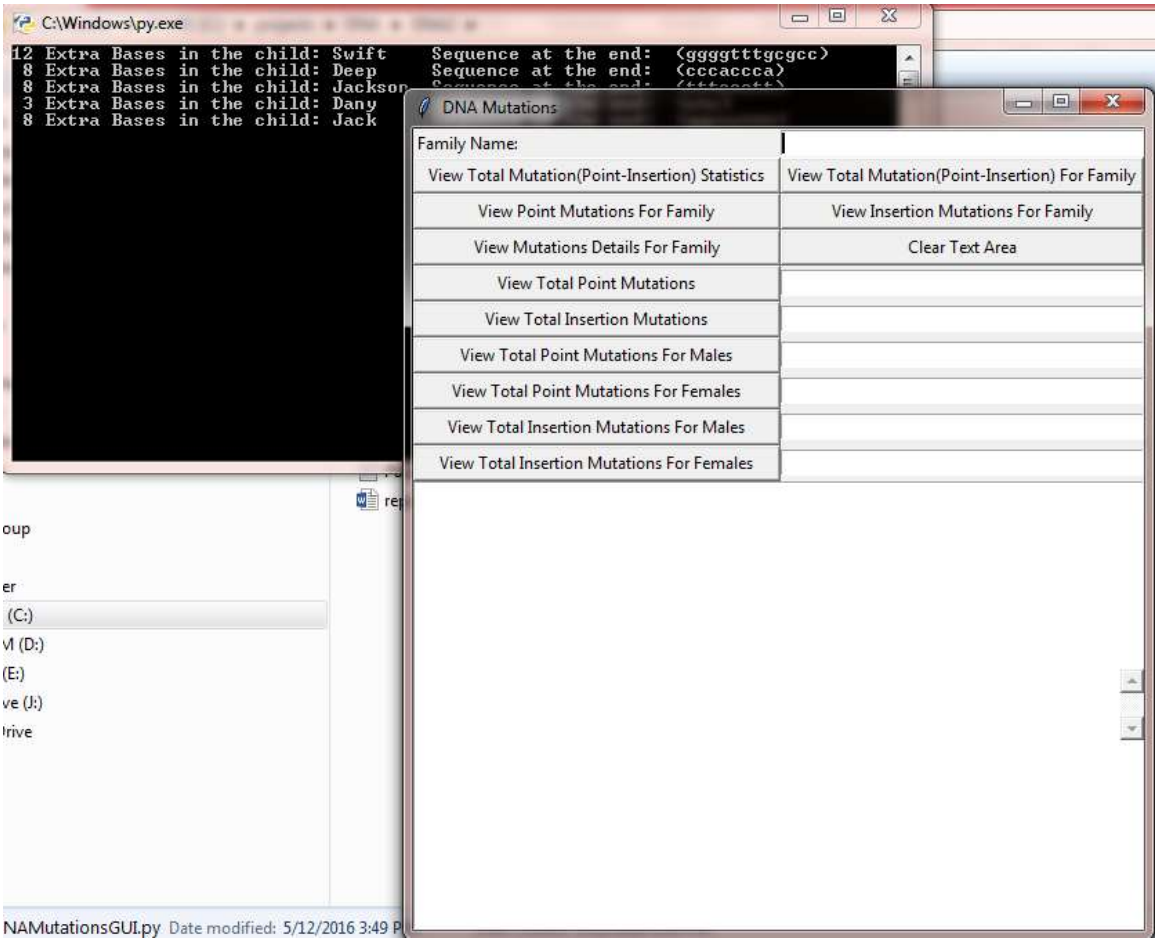

Figure 4: First Running the application

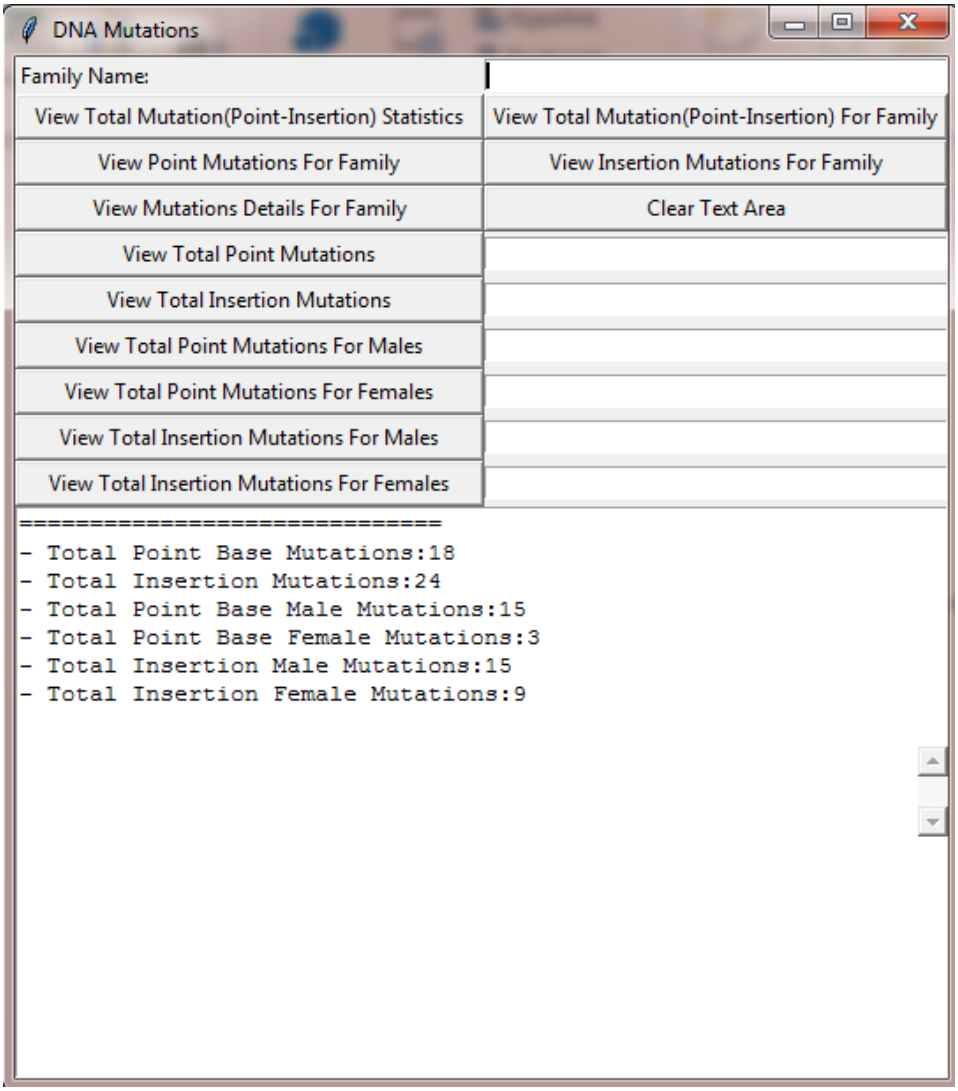

Figure 5: Viewing General Statistics about the sample files

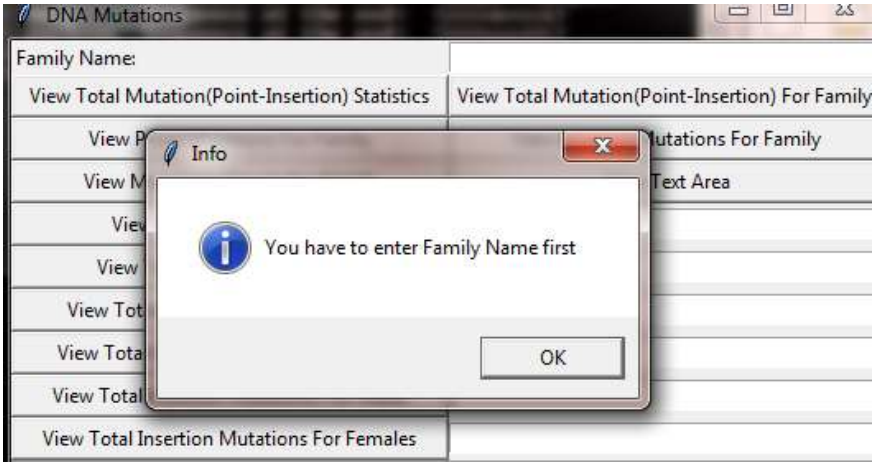

Figure 6: Error Messages, clicking a family related button without entering a family name

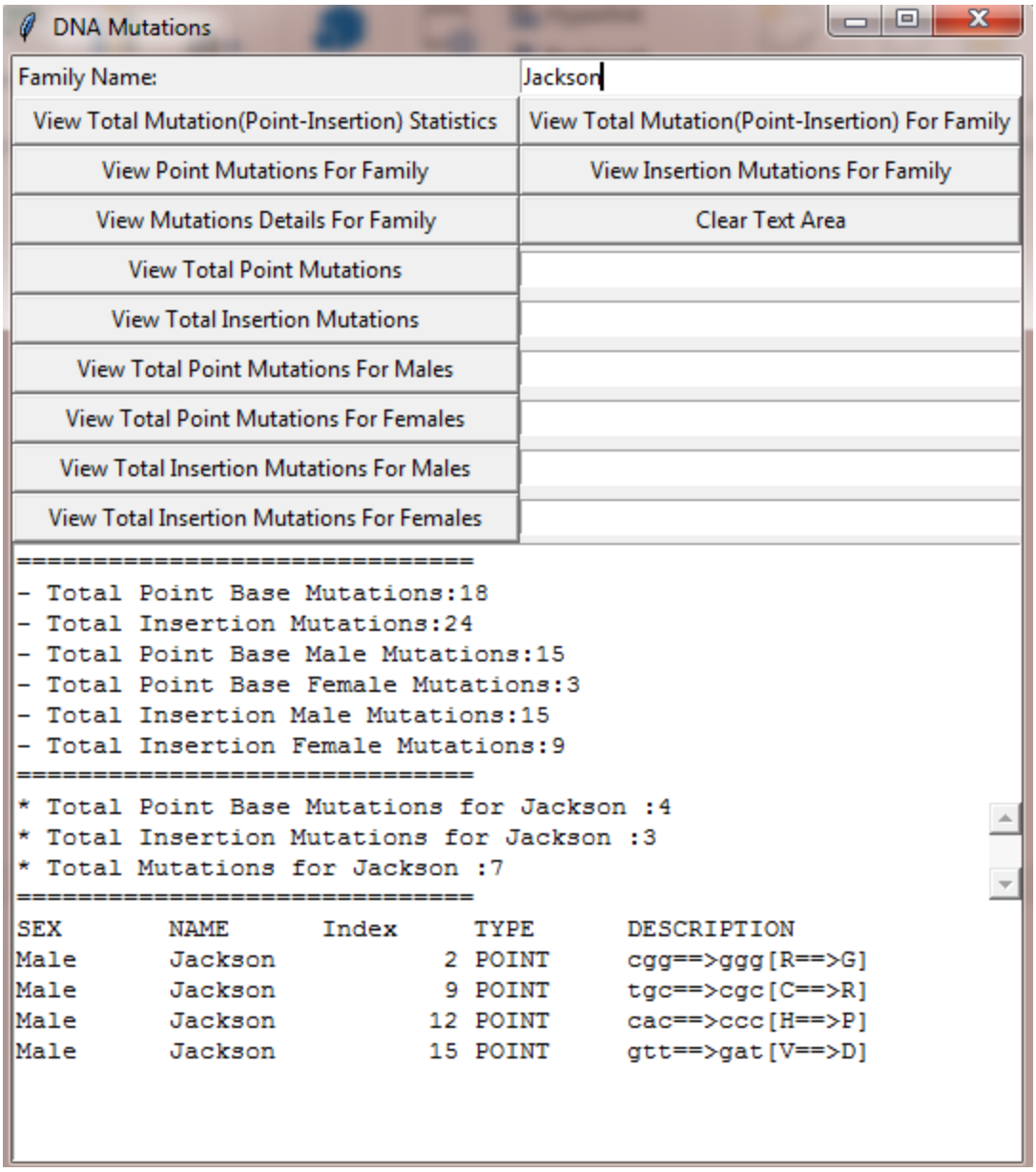

Figure 7: Point Mutations for the Jackson Family

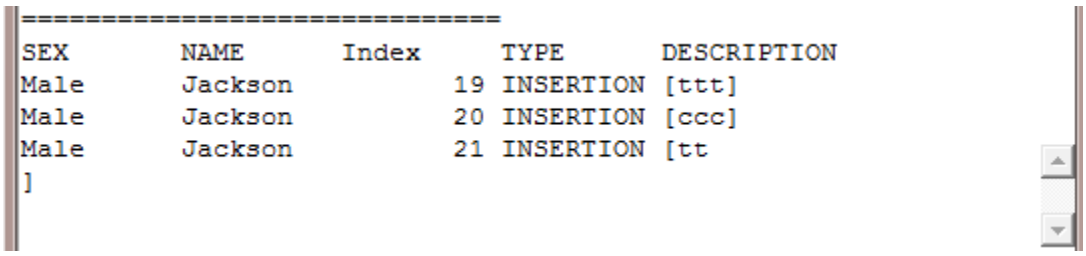

Figure 8: Insertion Mutations for the Jackson Family

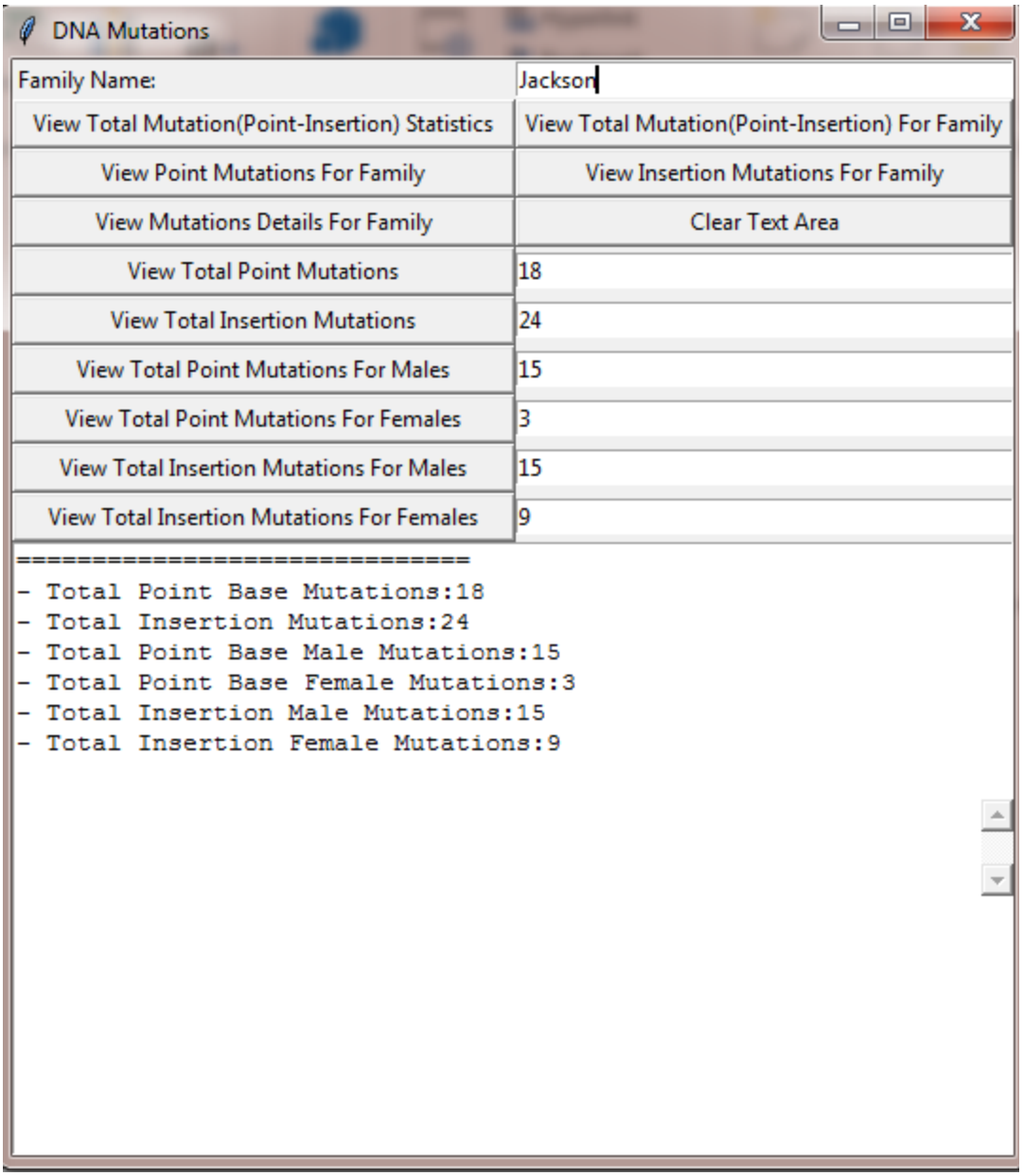

Figure 9: viewing individual totals for the files compared.