# MendelGeneticCounseling ASHG Workshop

### last update: October 12 2019, 8pm.

The purpose of this tutorial is to demonstrate how to calculate genetic risks for individuals using their family histories and covariate information.

## Installation instructions (not needed during workshop)

`MendelGeneticCounseling.jl#ASHG2019` currently supports Julia versions 1.0, 1.1 and 1.2, but it is currently an unregistered package. To install, press `]` to invoke the package manager mode and install these packages by typing:

```
add https://github.com/OpenMendel/SnpArrays.jl
add https://github.com/OpenMendel/MendelSearch.jl
add https://github.com/OpenMendel/MendelBase.jl
add https://github.com/OpenMendel/MendelGeneticCounseling.jl#ASHG2019
build SpecialFunctions
```

Be sure also to download the BRCA and Cholesterol data from https://github.com/OpenMendel/GeneticCounseling_ASHG2019 (https://github.com/OpenMendel/GeneticCounseling_ASHG2019).

### NOTE: When finished with this notebook, go to the File tab and first select save and checkpoint to save your results. Then under the file tab, select close and halt to prevent copies of Jupyter notebook from running indefinitely in the background.

## Navigating a Jupyter Notebook

The jupyter notebook has a menu with tabs. You can use them to insert or delete cells, change cells from coding to markdown, move up or down the notebook, and to save your notebook. For example, pressing the + button adds a new cell and the scissors deletes a cell.

## Check Julia version:

For reproducibility, check the machine information below. To execute a notebook command, hold down`Shift Enter` within the box. This tutorial and corresponding modules have been checked with Julia versions 1.0, 1.1 and 1.2. Please report any issues running the tutorial or the module to Janet Sinsheimer PhD. (jsinshei@g.ucla.edu).

In [1]:
```
versioninfo()
```

```
Julia Version 1.2.0
Commit c6da87ff4b (2019-08-20 00:03 UTC)
Platform Info:
  OS: macOS (x86_64-apple-darwin18.6.0)
  CPU: Intel(R) Core(TM) i7-6567U CPU @ 3.30GHz
  WORD_SIZE: 64
  LIBM: libopenlibm
  LLVM: libLLVM-6.0.1 (ORCJIT, skylake)
```

It's a good idea to check for updates although it can take time so don't do so during the workshop. In the future remove "#" to run.

In [ ]:
```
# ] update
```

## When to use MendelGeneticCounseling

`MendelGeneticCounseling.jl` is capable of calculating the risk of an underlying genotype (e.g. Homozygous_Normal, Heterozygous, Homozygous_Mutant) given a family history and individual risk factors including closely linked genetic markers. It can use parametric models or a penetrance file.

Example 1 uses a parametric model. The parametric models are currently restricted to the following generalized linear model distributions, binomial, exponential, gamma, inverse gaussian, logistic, lognormal, negative binomial, Poisson, and of course, the normal distributions but more distributions can be added to the `apply_dist.jl` function. The inverse link functions available are for the links: log, logit, cauchit, complementary log log, inverse (1/x), probit, and square root. A more experienced user can add more models to `apply_inverse_link_new.jl`. We will be adding the capability to run censored survival models soon.

`MendelGeneticCounseling.jl` is also capable of using a penetrance file that provides the probability that an individual is affected conditional their genotype and risk factors (Example 2).

This module is a prototype and features will be added. Currently we are working on models that can handle censoring, data in snp binary or vcf files, and mutation at the risk locus.

Check working directory. For convenience have the julia notebook in the same directory as your files.

In [2]:
```
pwd()
```

Out[2]:
```
"/Users/janets/Documents/lap_top_janet/janet/short_courses_2019/ASHGproposal/Gen
eticCounseling_ASHG2019-master10062019"
```

In [ ]:
```
#If you need to change directories you can use cd(). The exact syntax depends on w
hether you are using a Mac or a PC (windows).

#For the Mac an example is:  cd("/Users/janets/GeneticCounseling/Chol")

#For the PC, an example is: cd("C:\\Users\\Janet Sinsheimer\\Documents\\Julia_file
s")
```

### Analysis keywords needed for a glm penetrance function.

| Keyword | Default Value | Allowed value | Description |
|---|---|---|---|
| glm_mean | 0.0 | symbolic expression | provides the form of x^b |
| glm_response | "Normal" | GLM distribution | One of the following distribution choices: Binomial, exponential, gamma, inverse Gaussian, logistic, lognormal, negative binomial, Poisson, normal |
| glm_link | "IdentityLink" | Inverse Link Function for the appropriate GLM link | CauchitLink,CloglogLink, IdentityLink, InverseLink, LogitLink, LogLink, ProbitLink, Sqrt |
| glm_trait | Affected | String | Defines the column in the pedigree file that contains the trait phenotype when using a glm for the penetrance. Note: Individuals' values must be quantitative at this time |
| glm_scale | 1.0 | Positive real number | measure of the spread of the trait - in the case of the normal distribution, the standard deviation |
| glm_trials | 1 | Positive Integer | For logistic regression glm_trials = 1 |

### Analysis keywords if using a penetrance file

| Keyword | Default Value | Allowed value | Description |
|---|---|---|---|
| disease_status | "" | String | Defines the column in the pedigree file that contains the trait phenotype when including a penetrance file. Values affected = 1, unaffected = 0, unknown = -1 |
| penetrance_file | "" | String | The absence of a penetrance file automatically results in a parametric penetrance file. The presence of a penetrance file results in discrete, user defined values based on risk classes |

### Analysis keywords - input and output common to both options.

| Keyword | Default Value | Allowed value | Description |
|---|---|---|---|
| locus_file | "" | String | Provides the population specific locus names, allele names and allele frequencies |
| output_file | Mendel_Output.txt | String | OpenMendel generated output file with table of kinship coefficients |
| pedigree_file | "" | String | Numerator and Denominator pedigree |
| phenotype_file | "" | String | provides genetic model for the underlying genetic locus |

A list of OpenMendel keywords common to most analysis package can be found here (https://openmendel.github.io/MendelBase.jl/#keywords-table)

# Example 1: Probability that an individual is heterozygous given pedigree and covariate information.

## Using a gamma distribution to model the penetrance

### Data used in Example 1:

The input files for all examples in this tutorial can be obtained from https://github.com/OpenMendel/GeneticCounseling_ASHG2019 (https://github.com/OpenMendel/GeneticCounseling_ASHG2019)

The data are from an example pedigree used in the Mendel version 16.0 release. The pedigree structure and phenotypes are originally from Schrott et al. (1972) Ann Int Med 76:711–720. We have used the pedigree to provide a slightly contrived example in which Mother III13, who has cholesterol value 440 at age 21 is concerned that her young son might also be affected with extreme hypercholesterolemia. This first analysis calculates the probability that person IV11 has a heterozygous genotype given his covariates and their relatives' information.

### Step 1: Examine the pedigree file:

Recall what is needed in a valid pedigree structure (https://openmendel.github.io/MendelBase.jl/#pedigree-file).

'MendelGeneticCounseling.jl' calculates the conditional probability of an individual being affected given the pedigree information by calculating the joint probability of the family and the individual's affection status (numerator pedigree) divided by the probability of the family (denominator pedigree). Accordingly the user needs to provide the program with a pedigree file with a numerator pedigree that includes the genotype of the individual of interest and the denominator pedigree doesn't. In our example we have named the pedigrees Top and Bottom but the choice of two names is left to the user as long as they are distinct.

If you like using the terminnal on the Mac you can do so without leaving the notebook by typing ";". On a Mac you can view the pedigree file by typing "; cat PedChol.csv"

Here however we will use Julia commands that will work for both the Mac and the PC. The command readlines displays the contents of a file.

```
In [3]: readlines("Cholestrol/PedChol.csv")
```

```
Out[3]: 219-element Array{String,1}:
 "Pedigree,Person,Father,Mother,Sex,HC,Age ,lnChol,Chol,High_Chol"
 "  TOP ,  III11   ,  II2     ,  II1     ,2,,22,6.39,595,1"
 "  TOP ,  II16    ,  I2      ,  I1      ,1,,38,6.17,479,1"
 "  TOP ,  III56   ,  II19    ,  II18    ,2,,4,6.12,454,1"
 "  TOP ,  II18    ,  I2      ,  I1      ,2,,36,6.1,448,1"
 "  TOP ,  III13   ,  II2     ,  II1     ,2,,21,6.09,440,1"
 "  TOP ,  III50   ,  II16    ,  II17    ,2,,10,6.07,433,1"
 "  TOP ,  III7    ,  II2     ,  II1     ,1,,26,6.05,425,1"
 "  TOP ,  III1    ,  II2     ,  II1     ,1,,31,6.05,422,1"
 "  TOP ,  II29    ,  I2      ,  I1      ,1,,27,6.03,416,1"
 "  TOP ,  III58   ,  II20    ,  II21    ,1,,6,6.02,413,1"
 "  TOP ,  III55   ,  II19    ,  II18    ,1,,10,6.02,412,1"
 "  TOP ,  III15   ,  II2     ,  II1     ,1,,20,5.97,391,1"
 ⋮
 "BOTTOM,  III67   ,  II32    ,  II33    ,1,,2,5.08,160,0"
 "BOTTOM,  II8     ,,,2,,38,5.03,153,0"
 "BOTTOM,  II33    ,,,2,,19,4.97,144,0"
 "BOTTOM,  I1      ,,,2,,62,NA,NA,NA"
 "BOTTOM,  II1     ,  I2      ,  I1      ,2,,38,NA,NA,NA"
 "BOTTOM,  II10    ,,,1,,38,NA,NA,NA"
 "BOTTOM,  II26    ,,,1,,28,NA,NA,NA"
 "BOTTOM,  II5     ,  I2      ,  I1      ,1,,41,NA,NA,NA"
 "BOTTOM,  II7     ,  I2      ,  I1      ,1,,38,NA,NA,NA"
 "BOTTOM,  II9     ,  I2      ,  I1      ,2,,38,NA,NA,NA"
 "BOTTOM,  III14   ,,,1,,21,NA,NA,NA"
 "BOTTOM,  III30   ,,,1,,20,NA,NA,NA"
```

The top line provides the column names. The first $5$ columns indicate the necessary pedigree information including sex. The $6$th column label, HC, is the risk locus genotype, which is unknown for most individuals. The $7$th column label, Age, is age in years. The $8$th column label, lnChol, is log base e of the cholesterol value, the $9$th column label, Chol, is total cholesterol in mg/dl, and $10$th column label is a dichotomize trait where total cholesterol greater than or equal to 225 is denoted as 1 for affected and less than 225 is denoted as 0. Missing values are denoted as NA.

If you examine the file carefully you will see that all the information in the numerator pedigree is repeated in the denominator pedigree except that we specify individual IV11's genotype in the numerator pedigree but not the denominator pedigree. The likelihood of the numerator pedigree divided by the likelihood of denominator pedigree gives us $P(G_{IV11} = 1/2|\mathbf{G, Chol, Age, Sex})$, the "risk."

### Step 2: Examine the control file

A control file gives specific instructions to `MendelGeneticCounseling`. We specify the dependent variable with the keyword `glm_trait`=Chol. The trait name must correspond exactly to the column name in the pedigree file.

In this example we treat the cholesterol values, which are highly right tailed, as gamma distributed. This information is specified by the keyword `glm_response` = GammaDist and `glm_link`= LogLink.

The glm_mean is the linear equation, $x^t b$.

Specifically, $f(y) = (\frac{\sigma}{\mu})^{\sigma} \frac{y^{\sigma-1} e^{\frac{-\sigma y}{\mu}}}{\Gamma(\mu)}$ where the scale $\sigma = 44.68$ and the mean, $\mu = g(x^t p) = e^{(4.691+0.562(\max(allele1,allele2))+0.00194Age+0.036Sex)}$ and $g$ is the exponential function which is the inverse link function for $\ln$. Note that allele1/allele2 represents the current genotype presented to the penetrance function.

We specify the "normal allele" as $1$ and the "mutant allele" as $2$ according to their order of appearance in the locus file. By using the expression $\max(allele1, allele2)$, we are specifying that genotype $1/1$ adds $0.562$, and genotype $1/2$ or $2/2$ adds $1.124$ to $\ln \mu$. In other words, we are modeling the locus as dominant. Sex is coded here as $1$ for males and $2$ for females. Male adds $0.036$ to and female adds $0.072$ to $\ln \mu$. Age is measured in years; each year adds $0.00194$ to $\ln \mu$. Thus a 50 year old female with genotype $1/1$ has an expected cholesterol value of $e^{(4.691+0.562+0.00194(50)+0.036(2))} = 226.3$.

The keyword `glm_trials` is ignored in this analysis. It is used when running logistic regression mean models, where `glm_response` = BinomialDist and `glm_link` = LogitLink, binomial and negative binomial models.

```
In [4]: readlines("Cholestrol/ControlParametricPenetranceExample.txt")
```

```
Out[4]: 16-element Array{String,1}:
         "#"
         "# Input and Output files."
         "#"
         "locus_file = LocusChol.txt"
         "pedigree_file = PedChol.csv"
         "phenotype_file = PhenoChol.txt"
         "output_file = CholHeterozygousRisk.txt"
         "#"
         "#  Analysis parameters for Genetic Counseling option."
         "#"
         "glm_mean = 4.691+0.562(max(allele1,allele2))+0.00194Age+0.036Sex"
         "glm_response = GammaDist"
         "glm_link = LogLink"
         "glm_trait = Chol"
         "glm_scale = 44.68"
         "glm_trials = 1"
```

### Step 3: Examine the Locus and Phenotype Files.

```
In [5]: readlines("Cholestrol/LocusChol.txt")
```

```
Out[5]: 3-element Array{String,1}:
         "Locus,Allele,Chromosome,European"
         "HC,-,autosome,0.9600"
         "HC,+,autosome,0.0400"
```

```
In [6]: readlines("Cholestrol/PhenoChol.txt")
```

```
Out[6]: 4-element Array{String,1}:
          "Locus,Phenotype,Genotypes"
          "HC,Homozygous_Normal,\"-/-\""
          "HC,Homozygous_Mutant,\"+/+\""
          "HC,Heterozygous,\"+/-\""
```

The locus file provides the name of the putative disease locus. The name must match exactly the column name in the pedigree file. In this simple example, this locus is unobserved and so no one has a genotype in the pedigree file except person IV11. The locus HC has two alleles "+" and "-". The locus is on an autosomal chromosome. The "-" allele has frequency 0.96 and the "+" allele has frequency 0.04. The phenotype file provides the genetic model for this locus. The "-/-" genotype has phenotype "Homozygous_Normal", the "+/-" genotype has phenotype "Heterozygous", and the "+/+" genotype has phenotype "Homozygous_Mutant."

### Step 4: Run the analysis in the Julia REPL or directly in notebook

The first command `using MendelGeneticCounseling` loads the MendelGeneticCounseling module. This command needs to be issued just once during this tutorial. The next command `GeneticCounseling` reads the control file, reads in the data and runs the analysis.

In [7]:
```
# first call of the GeneticCounseling function takes longer because of JIT compili
ng
using MendelGeneticCounseling
GeneticCounseling("Cholestrol/ControlParametricPenetranceExample.txt")
```

```
┌ Info: Recompiling stale cache file /Users/janets/.julia/compiled/v1.2/MendelGe
neticCounseling/l0fXi.ji for MendelGeneticCounseling [5eee5fa4-c0d8-5591-aecf-5d
586585de4b]
└ @ Base loading.jl:1240


     Welcome to OpenMendel's
 Genetic Counseling Analysis Option


Reading the data.

The current working directory is "/Users/janets/Documents/lap_top_janet/janet/sh
ort_courses_2019/ASHGproposal/GeneticCounseling_ASHG2019-master10062019/Cholestr
ol".

Keywords modified by the user:

  control_file = Cholestrol/ControlParametricPenetranceExample.txt
  glm_link = LogLink
  glm_mean = 4.691+0.562(max(allele1,allele2))+0.00194Age+0.036Sex
  glm_response = GammaDist
  glm_scale = 44.68
  glm_trait = Chol
  glm_trials = 1
  locus_file = LocusChol.txt
  output_file = CholHeterozygousRisk.txt
  pedigree_file = PedChol.csv
  phenotype_file = PhenoChol.txt

 no penetrance file

Analyzing the data.

 The risk =  0.27557.


Mendel's analysis is finished.
```

### Step 5: Interpreting the result

`MendelGeneticCounseling` should have generated a file `CholHeterozygousRisk.txt` in your local directory. The value is the conditional probability that person IV11 is heterozygous given the information provided regarding the family and the individuals' age and genotype. The probability is $0.27557$.

### Step 6: Test yourself.

(1) Modify the pedigree file and the control file to calculate the probability that individual IV11 has the homozygous normal genotype. Then rerun MendelGeneticCounseling with these new data. You should find the probabiity is $0.71335$.

(2) Modify the pedigree file to determine the probability individual III13 has the heterozygous genotype. You should find the probability is $0.96698$.

**Help on modifying the jupyter notebook so you can rerun MendelGeneticCounseling without destroying your original results:**

First go to the file tab and press "Save and checkpoint" then go to the insert and click "insert cell below". Then insert command to run GeneticCounseling with your new control file.

# Example 2: Using a penetrance file.

In this example we illustrate how to use a penetrance file. We thank Brian Shirts for pointing us to the "Analyze My Variant" website (http://analyzemyvariant.com (http://analyzemyvariant.com)) for a realistic example. We use the penetrance classes provided for BRCA1 (http://analyzemyvariant.com/brca1-info (http://analyzemyvariant.com/brca1-info)).

BrianspedigreeMay162019.jpg

### Step 1: Examine the Penetrance file

We have set up the penetrance file to have 5 columns. The first column is for the risks for homozygous wild type genotype (1/1) for each sex and risk decade (penetrance class). The second column is for the risks for heterozyous genotype (1/2) by penetrance class. The third column is for the risks for the homozygous high risk genotype (2/2) by penetrance class. The next column corresponds to the sex of the individual. The final column corresponds to the risk decade of the individual (1: $0 \leq age < 20$, 2: $20 \leq age < 30$, 3: $30 \leq age < 40$, 4: $40 \leq age < 50$, 5: $50 \leq age < 60$, 6: $60 \leq age < 70$, and 7: $70 \leq age$).

To view the file type:

```
In [8]: readlines("BRCA/PenBRCAExample.csv")
```

```
Out[8]: 15-element Array{String,1}:
          "Homozygous_Normal,Heterozygous,Homozygous_Mutant,Sex,Risk_decade"
          "0.000000885,0.001025896,0.001025896,female,1"
          "0.000040997,0.047524,0.047524,female,2"
          "0.00189916,0.18042,0.18042,female,3"
          "0.00878848,0.3736,0.3736,female,4"
          "0.0275136,0.5752,0.5752,female,5"
          "0.05646,0.6889,0.6889,female,6"
          "0.0793,0.785,0.785,female,7"
          "7.58E-08,1.07E-05,1.07E-05,male,1"
          "0.0000012,0.00017,0.00017,male,2"
          "0.000019,0.0012,0.0012,male,3"
          "0.000085,0.003,0.003,male,4"
          "0.00027,0.0062,0.0062,male,5"
          "0.00067,0.012,0.012,male,6"
          "0.0012,0.018,0.018,male,7"
```

### Step 2: Examine the pedigree file

In this example we are interested in determining the probability that individual 18, a 38 year old female who is currently unaffected with breast cancer is heterozygous for a BRCA1 mutation. In this case, some members of her family have genotypes at the locus.

Like the previous example, the pedigree is present in two copies. The first copy (called Top) has the genotype of individual 18 as heterozygous. The second copy (called Bottom) has her genotype missing. The likelihood of the first pedigree divided by the likelihood of second pedigree gives us $P(G_{18} = 1/2|\mathbf{G, CancerStatus, Age, Sex})$, the "risk."

```
In [9]: readlines("BRCA/PedBRCAExample.csv")
```

```
Out[9]: 41-element Array{String,1}:
         "Pedigree,Person,Father,Mother,Sex,BRCA,Age,Risk_decade,Proband,Cancer"
         "TOP,1,0,0,male,,79,7,0,-1"
         "TOP,2,0,0,female,,78,7,0,-1"
         "TOP,3,1,2,female,Heterozygous,40,4,0,1"
         "TOP,4,1,2,female,,79,7,0,0"
         "TOP,5,1,2,female,,85,7,0,1"
         "TOP,6,1,2,female,,43,4,0,1"
         "TOP,7,0,0,male,,80,7,0,0"
         "TOP,8,7,3,male,,73,7,0,0"
         "TOP,9,7,3,male,,41,4,0,0"
         "TOP,10,0,0,female,,89,7,0,-1"
         "TOP,11,7,3,male,Heterozygous,30,3,0,0"
         "TOP,12,0,0,female,,80,7,0,0"
         ⋮
         "BOTTOM,9,7,3,male,,41,4,0,0"
         "BOTTOM,10,0,0,female,,89,7,0,-1"
         "BOTTOM,11,7,3,male,Heterozygous,30,3,0,0"
         "BOTTOM,12,0,0,female,,80,7,0,0"
         "BOTTOM,13,9,10,female,Heterozygous,41,4,0,1"
         "BOTTOM,14,9,10,male,,60,6,0,0"
         "BOTTOM,15,9,10,female,Heterozygous,50,5,0,1"
         "BOTTOM,16,9,10,female,Heterozygous,60,6,0,0"
         "BOTTOM,17,11,12,female,Heterozygous,49,4,1,1"
         "BOTTOM,18,11,12,female,,38,3,0,0"
         "BOTTOM,19,11,12,male,Heterozygous,36,3,0,0"
         "BOTTOM,20,11,12,female,Heterozygous,48,4,0,1"
```

### Step 3: Examine the Locus and Phenotype File

These files are very similar to the ones we used in the first example.

```
In [10]: readlines("BRCA/LocusBRCAExample.txt")
```

```
Out[10]: 3-element Array{String,1}:
          "Locus,Allele,Chromosome,European"
          "BRCA,\"1\",Autosome,0.998"
          "BRCA,\"2\",Autosome,0.002"
```

```
In [11]: readlines("BRCA/PhenoBRCAExample.txt")
```

```
Out[11]: 4-element Array{String,1}:
          "Locus,Phenotype,Genotypes"
          "BRCA,Homz_rare,\"2/2\""
          "BRCA,Heterozygous,\"1/2\""
          "BRCA,Homz_common,\"1/1\""
```

### Step 4: Examine the control file

This control file has some of the same features as parametric ones but it doesn't need to specify the GLM values because there is a penetrance file. Besides the input and output file names, the only needed information is the name of the Column in the pedigree that contains the information regarding breast cancer status.

In [12]:
```
readlines("BRCA/ControlBRCAExample.txt")
```

Out[12]:
```
10-element Array{String,1}:
 "#"
 "# Input and Output files."
 "#"
 "locus_file = LocusBRCAExample.txt  "
 "pedigree_file = PedBRCAExample.csv"
 "phenotype_file = PhenoBRCAExample.txt"
 "penetrance_file = PenBRCAExample.csv"
 "output_file = BRCAExampleOut.txt"
 "#"
 "disease_status = Cancer"
```

### Step 5: Running the analysis

In [13]:
```
using MendelGeneticCounseling
GeneticCounseling("BRCA/ControlBRCAExample.txt")
```

```
    Welcome to OpenMendel's
 Genetic Counseling Analysis Option


Reading the data.

The current working directory is "/Users/janets/Documents/lap_top_janet/janet/sh
ort_courses_2019/ASHGproposal/GeneticCounseling_ASHG2019-master10062019/BRCA".

Keywords modified by the user:

  control_file = BRCA/ControlBRCAExample.txt
  disease_status = Cancer
  locus_file = LocusBRCAExample.txt
  output_file = BRCAExampleOut.txt
  pedigree_file = PedBRCAExample.csv
  penetrance_file = PenBRCAExample.csv
  phenotype_file = PhenoBRCAExample.txt

this problem has 2 factors called Symbol[:Sex, :Risk_decade]

Analyzing the data.

 The risk =  0.45091.


Mendel's analysis is finished.
```

### Step 6: Interpreting the Result

Again you should get a file with the results. The probability of that individual 18 is heterozygous is 0.45091.

### Step 7: Test yourself.

(1) Modify the pedigree file and the control file to calculate the probability that individual 18 has the homozygous normal genotype. Then rerun MendelGeneticCounseling with these new files.

(2) How would the risk change if individual 18 were 20 years old instead 38 years old? What if she were 68 years old? (To determine, change her risk decade in the pedigrees and rerun the analysis).

# Final Comments

The Julia version of Mendel, provides an opportunity for the user to easily modify the code to suit their own needs. All the source code is provided and Julia is both accessible and very fast.

### Reference

For publication please cite:

OPENMENDEL: a cooperative programming project for statistical genetics. Zhou H, Sinsheimer JS, Bates DM, Chu BB, German CA, Ji SS, Keys KL, Kim J, Ko S, Mosher GD, Papp JC, Sobel EM, Zhai J, Zhou JJ, Lange K. Hum Genet. 2019 Mar 26. doi: 10.1007/s00439-019-02001-z

### NOTE: When Finished with this notebook. Go to the File tab and first select save and checkpoint to save your results. Then under the file tab, select close and halt to prevent copies of Jupyter notebook from running indefinitely in the background