

## Table of Contents

Table of Contents	1
Introduction	2
Installation	2
Starting	3
Loading Images	4
Visual Phasing	5
Start of Visual Phasing	5
Renaming	6
Cousins	6
Finish of Visual Phasing	6
Recombination Points	6
FIR	6
Segment Lists	8
Appendix	8
User Callable Macros	9

# Introduction

This LibreOffice spreadsheet allows the user to perform Visual Phasing of the graphs from GEDmatch or any other source of segment graphs. This assumes the user is familiar with the Visual Phasing process (not described here). LibreOffice Calc is available on a variety of Operating Systems, including Windows, Linux, and MacOS. GEDmatch is the most common source for segment images, but any other source will also work. Scripts are provided to assemble and name files from GEDmatch.

## Installation

The graphs need to be organized in a specific way. The folder GEDmatchImages contains all of the graphs. The graphs should be named kit1\_kit2\_Chromosome\_Number.gif . For example, for the 1st image of Chromosome 13 for kit comparison of X0001 and C0001, the file is named X0001\_C0001\_13\_1.gif . There are 2 images for each case, the 1st is the Full/Half/NoMatch image of the SNP match and the second is the validity of segment indicator for the SNP. The graph can be linear in SNP or linear in Mbp. If they are linear in Mbp, there is a macro to determine MBP based on the width of the spreadsheet cell. This will not work for GEDmatch images since those images are not linear in Mbp.

The first step to use this spreadsheet is to generate all of the GEDmatch graphs. The user runs the One-to-One Autosomal and the One-to-One X tools in GEDmatch for each pair of kits (both Sibling and Cousin) and saves the Web Page from the browser into the working directory. This working directory also contains the spreadsheet. For 4 siblings, this is 6 cases. For the SampleData case, the folder looks like:

For the SampleData case, the folder looks like:

The contents of each of the folders is the graphs for that match case. Note that the naming is not the same as what this tool needs. An example is:

X0001\_X0004\_file:  
X0001\_X0004\_1.FBFCE\_1.gif X0001\_X0004\_15.FBFCE\_1.gif X0001\_X0004\_20.FBFCE\_1.gif X0001\_X0004\_29.FBFCE\_1.gif  
X0001\_X0004\_16.FBFCE\_2.gif X0001\_X0004\_16.FBFCE\_1.gif X0001\_X0004\_20.FBFCE\_2.gif X0001\_X0004\_5.FBFCE\_1.gif  
X0001\_X0004\_11.FBFCE\_1.gif X0001\_X0004\_18.FBFCE\_2.gif X0001\_X0004\_21.FBFCE\_1.gif X0001\_X0004\_5.FBFCE\_2.gif  
X0001\_X0004\_12.FBFCE\_1.gif X0001\_X0004\_18.FBFCE\_1.gif X0001\_X0004\_22.FBFCE\_1.gif X0001\_X0004\_6.FBFCE\_1.gif  
X0001\_X0004\_12.FBFCE\_2.gif X0001\_X0004\_18.FBFCE\_1.gif X0001\_X0004\_22.FBFCE\_2.gif X0001\_X0004\_7.FBFCE\_1.gif  
X0001\_X0004\_13.FBFCE\_1.gif X0001\_X0004\_18.FBFCE\_2.gif X0001\_X0004\_22.FBFCE\_1.gif X0001\_X0004\_7.FBFCE\_2.gif  
X0001\_X0004\_13.FBFCE\_2.gif X0001\_X0004\_19.FBFCE\_1.gif X0001\_X0004\_23.FBFCE\_1.gif X0001\_X0004\_8.FBFCE\_1.gif  
X0001\_X0004\_14.FBFCE\_1.gif X0001\_X0004\_19.FBFCE\_2.gif X0001\_X0004\_23.FBFCE\_2.gif X0001\_X0004\_8.FBFCE\_2.gif  
X0001\_X0004\_15.FBFCE\_1.gif X0001\_X0004\_1\_FBFCE\_1.gif X0001\_X0004\_3.FBFCE\_2.gif X0001\_X0004\_9.FBFCE\_1.gif  
X0001\_X0004\_15.FBFCE\_2.gif X0001\_X0004\_1\_FBFCE\_2.gif X0001\_X0004\_3.FBFCE\_2.gif X0001\_X0004\_9.FBFCE\_2.gif

Once the GEDmatch data is ready, the user then runs the command line command

*make GEDmatchImages*

to create the new folder and copy all of the graphs to this new folder. Then the user runs the command line command

*make rename*

to rename the graphs. This also makes similar files with the kit1 and kit2 order reversed. The consolidated GEDmatchImages folder which contains the graphs that are used by the spreadsheet after the graphs have been renamed is:

```
gedmatchImages:
C0001_X0001_10_1.gif C0010_X0002_22_2.gif C0010_X0004_19_1.gif X0002_C0001_10_2.gif X0003_C0013_22_1.gif
C0001_X0001_10_2.gif C0010_X0002_2_1.gif C0010_X0004_19_2.gif X0002_C0001_11_1.gif X0003_C0013_22_2.gif
C0001_X0001_11_1.gif C0010_X0002_2_2.gif C0010_X0004_1_1.gif X0002_C0001_11_2.gif X0003_C0013_2_1_0if
C0001_X0001_11_2.gif C0010_X0002_2_3.gif C0010_X0004_1_2.gif X0002_C0001_11_3.gif X0003_C0013_2_2_0if
C0001_X0001_12_1.gif C0010_X0002_3_1.gif C0010_X0004_20_1.gif X0002_C0001_12_2.gif X0003_C0013_3_1_0if
C0001_X0001_12_2.gif C0010_X0002_3_2.gif C0010_X0004_20_2.gif X0002_C0001_13_1.gif X0003_C0013_3_2_0if
C0001_X0001_13_1.gif C0010_X0002_3_3.gif C0010_X0004_20_3.gif X0002_C0001_13_2.gif X0003_C0013_4_1_0if
C0001_X0001_13_2.gif C0010_X0002_4_1.gif C0010_X0004_21_1.gif X0002_C0001_14_1.gif X0003_C0013_4_2_0if
C0001_X0001_14_1.gif C0010_X0002_4_2.gif C0010_X0004_21_2.gif X0002_C0001_14_2.gif X0003_C0013_5_1_0if
C0001_X0001_14_2.gif C0010_X0002_5_1.gif C0010_X0004_22_1.gif X0002_C0001_15_1.gif X0003_C0013_5_2_0if
C0001_X0001_15_1.gif C0010_X0002_5_2.gif C0010_X0004_23_1.gif X0002_C0001_15_2.gif X0003_C0013_6_1_0if
C0001_X0001_16_1.gif C0010_X0002_6_1.gif C0010_X0004_3_1.gif X0002_C0001_16_2.gif X0003_C0013_7_1_0if
C0001_X0001_16_2.gif C0010_X0002_6_2.gif C0010_X0004_3_2.gif X0002_C0001_17_1.gif X0003_C0013_7_2_0if
C0001_X0001_17_1.gif C0010_X0002_7_1.gif C0010_X0004_3_3.gif X0002_C0001_17_2.gif X0003_C0013_8_1_0if
C0001_X0001_17_2.gif C0010_X0002_7_2.gif C0010_X0004_3_4.gif X0002_C0001_18_1.gif X0003_C0013_8_2_0if
C0001_X0001_18_1.gif C0010_X0002_8_1.gif C0010_X0004_4_1.gif X0002_C0001_18_2.gif X0003_C0013_9_1_0if
C0001_X0001_18_2.gif C0010_X0002_8_2.gif C0010_X0004_4_2.gif X0002_C0001_19_1.gif X0003_C0013_9_2_0if
C0001_X0001_19_1.gif C0010_X0003_1_1.gif C0010_X0004_4_3.gif X0002_C0001_19_2.gif X0003_C0014_10_1_0if
C0001_X0001_19_2.gif C0010_X0003_1_2.gif C0010_X0004_4_4.gif X0002_C0001_1_1_0if X0003_C0014_10_2_0if
C0001_X0001_1_1_0if C0010_X0003_1_1_0if C0010_X0004_7_1.gif X0002_C0001_1_2_0if X0003_C0014_11_1_0if
```

If a new kit is added, the user enters the command

*make clobber*

(to remove the existing GEDmatchImages folder) followed by repeating the two above commands. This will recreate the GEDmatchImages folder with the new kit(s).

# Starting

The first step is to fill the Sibling and Cousin Tables with the kit information.

Cousin Table					
Code	Relationship	Short Name	Full Name	Kit Number	Notes
G1	Top line Grandparent 1	G1	Cousin G	C0001	Maternal line
G2	Top line Grandparent 2	G2	Cousin H	C0002	Thru Paternal GM
G3	Bottom line Grandparent 1	G3	Cousin I	C0003	Thru Paternal GM
G4	Bottom line Grandparent 2	G4	Cousin J	C0004	Thru Paternal GM
P1	Paternal Grandparent	P1	Cousin K	C0005	Thru Paternal GM
P2	Paternal Grandparent 2	P2	Cousin L	C0006	Thru Paternal GM
M1	Maternal Grandparent	M1	Cousin M	C0007	Thru Maternal GM
M2	Maternal Grandparent 2	M2	Cousin N	C0008	Thru Paternal GM
IP1	Paternal Grandfather	Paternal GF	Cousin O	C0009	Thru Paternal GM
IP2	Paternal Grandmother	Paternal GM	Cousin P	C0010	Thru Paternal GM
MP1	Maternal Grandfather	Maternal GF	Cousin Q	C0011	Thru Paternal GM
MP2	Maternal Grandmother	Maternal GM	Cousin R	C0012	Thru Paternal GM
S1	Sibling 1	S1	Cousin S	C0013	Thru Paternal GF
S2R	Sibling 2 Right	S2R	Cousin T	C0014	Thru Paternal GF
S2L	Sibling 2 Left	S2L	Cousin U	C0015	Thru Paternal GF
S3R	Sibling 3 Right	S3R	Cousin V	C0016	Thru Maternal GM
S3L	Sibling 3 Left	S3L	Cousin W	C0017	Thru Maternal GM
S4R	Sibling 4 Right	S4R	Cousin X	C0018	Thru Maternal GM
S4L	Sibling 4 Left	S4L	Cousin Y	C0019	Thru Maternal GM
S5R	Sibling 5 Right	S5R	Cousin Z	C0020	Thru Maternal GM
S5L	Sibling 5 Left	S5L	Cousin E	C0021	Thru Maternal GM
S6R	Sibling 6 Right	S6R	Cousin F	C0022	Thru Maternal GM

Once that is done, the user clicks the **Reset Sheets** button to remove any existing Chromosome sheets and to recreate based on the data in the Sibling Table. The results are:

LibreOffice File Edit View Insert Format Styles Sheet Data Tools Window Help VisualPhasingTemplate.ods

Calibri 11 pt B I U - A  $\Sigma$  fx

Q20

**Grandparent Table**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
8	G1	Top line Grandparent 1		Format												
9	G2	Top line Grandparent 2			G2											
10	G3	Bottom line Grandparent 1				G3										
11	G4	Bottom line Grandparent 2					G4									
12	P1	Paternal Grandparent 1						P1								
13	P2	Paternal Grandparent 2							P2							
14	M1	Maternal Grandparent 1								M1						
15	M2	Maternal Grandparent 2								M2						
16	Paternal GF	Paternal Grandfather									Paternal GF					
17	Maternal GF	Maternal Grandfather									Maternal GF					
18	Maternal GM	Maternal Grandmother									Maternal GM					
19	Maternal GM	Maternal Grandmother									Maternal GM					
20																
21	<b>Sibling Table</b>															
22	Total	Short Name	Full name	Kit Number	M/F											
23	A															
24	A	Sibling Kit A	Sibling Kit A	X0001	M											
25	B															
26	B	Sibling Kit B	Sibling Kit B	X0002	M											
27	C															
28	C	Sibling Kit C	Sibling Kit C	X0003	M											
29	D															
30	D	Sibling Kit D	Sibling Kit D	X0004	M											
31																
32																
33																
34																
35																
36																
37																
38																
39																
40																
41																
42																
43																
44																
45																

Reset Sheets Load Images Show Cousins Legend Format Update Load FTR data

Chromosome 13 looks like:

LibreOffice File Edit View Insert Format Styles Sheet Data Tools Window Help VisualPhasingTemplate.ods

Calibri 11 pt B I U - A  $\Sigma$  fx

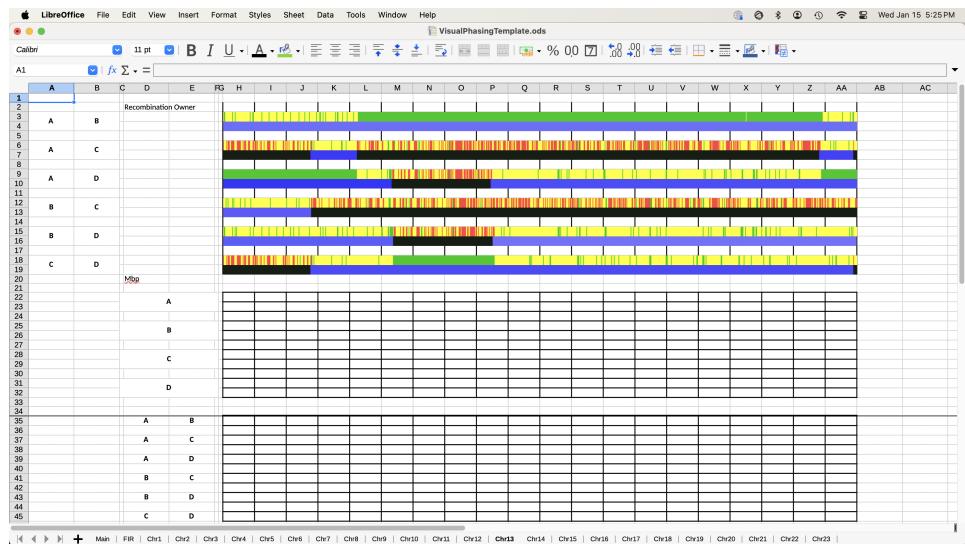
Main FIR Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16 Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 Chr23

**Recombination Owner**

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	
1																													
2																													
3	A	B																											
4																													
5	A	C																											
6																													
7	A	D																											
8																													
9	A	D																											
10																													
11																													
12	B	C																											
13																													
14	B	D																											
15																													
16	C	D																											
17																													
18	A	B																											
19																													
20	A	C																											
21																													
22																													
23																													
24																													
25																													
26																													
27																													
28																													
29																													
30																													
31																													
32																													
33																													
34																													
35																													
36																													
37																													
38																													
39																													
40																													
41																													
42																													
43																													
44																													
45																													

## Loading Images

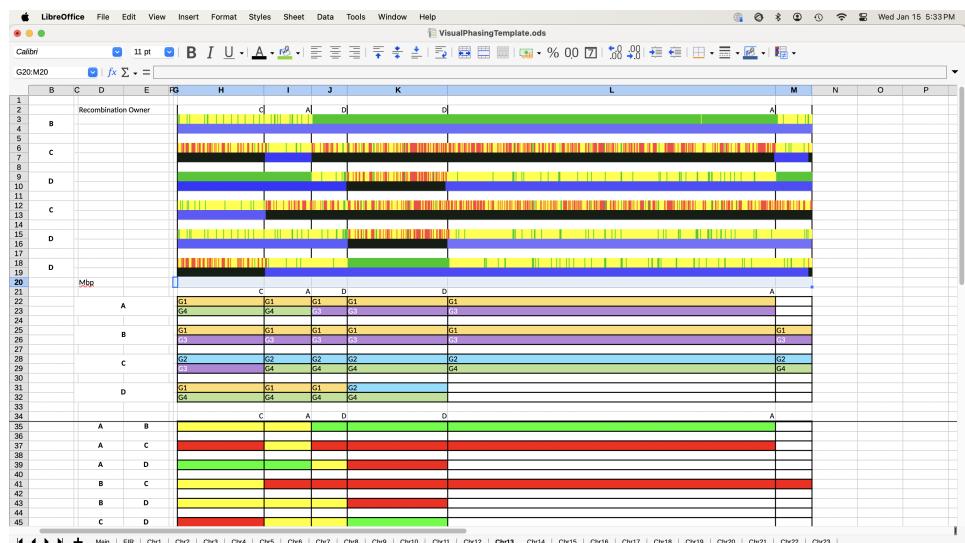
The next step is to load the images to each of the Chromosome sheets. Click the **Load Images** button. The result for Chromosome 13 is:



# Visual Phasing

## Start of Visual Phasing

Then proceed with moving the cell borders to identify recombination points and populate the owner of the recombination points. And start identifying the grandparent for each of the siblings segments.



# Renaming

To rename the grandparent in the painted area, use the **replacePaintValue** macro. Enter the existing text string in cell A1 of the sheet. Enter the new string value in cell A2. Then run the macro. This will replace all occurrences of the string in the painted area.

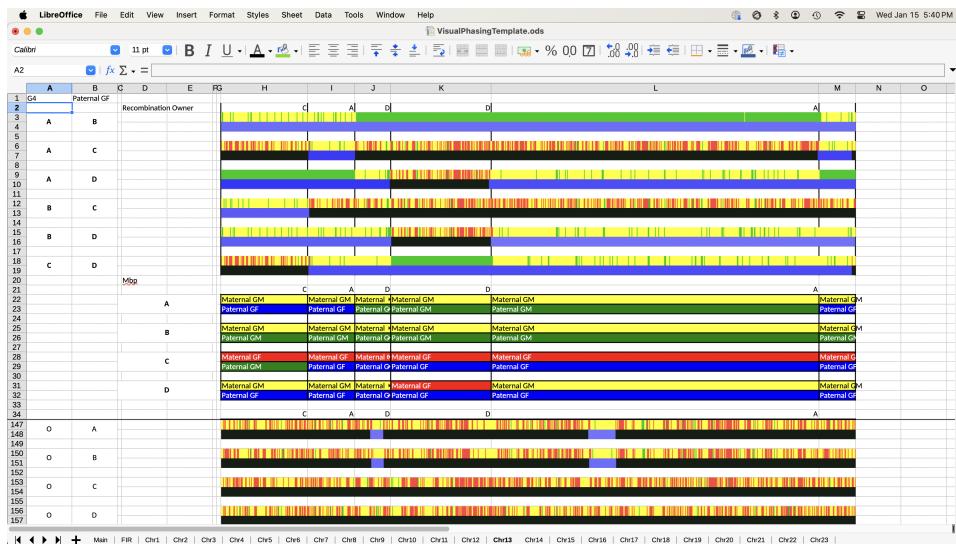
To start the painting process over, run the macro **resetPaintValue**. This will remove all of the strings in the painted area.

# Cousins

To add the Cousin graphs, go to the Main sheet and press the **Show Cousins** button. This adds all of the Cousin graphs to each of the Chromosome sheets. To omit a Cousin graph, put any character in the skip (1st) cell in the Cousins table.

# Finish of Visual Phasing

When visual phasing is complete, the image should be something like the below.



# Recombination Points

## FIR

FIR data can be generated in GEDmatch for each of the Sibling pairs. This is OPTIONAL. This data can be copied into the FIR sheet by clicking the **Load FIR Data** and then used as a reference when identifying the Mbp for a recombination point.

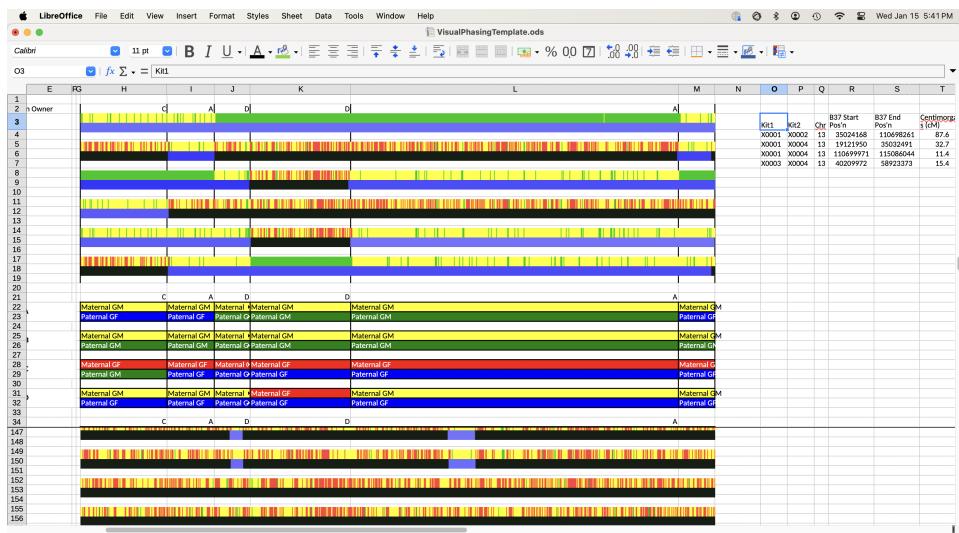
LibreOffice File Edit View Insert Format Styles Sheet Data Tools Window Help

VisualPhasingTemplate.ods

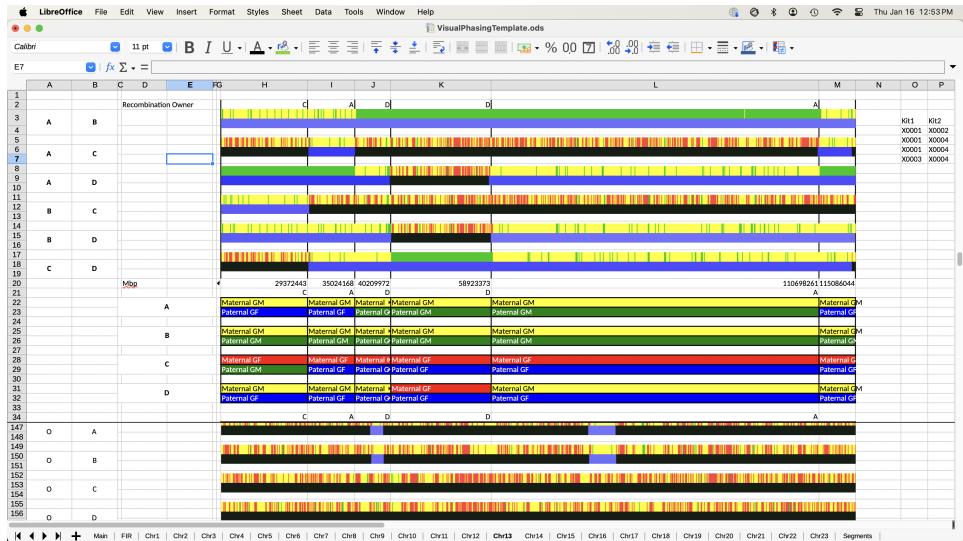
Calibri 11 pt B I U A fx Σ = Kit1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Kit1	Kit2	Chr	B37 Start	B37 End Pos'n	Chromosome	SNPs	Segment threshold	Stretch limit	SNP density											
2	X0002	X0004	1	4636767	94178912	111.3	13258	185	111	0.26											
3	X0002	X0004	1	158902352	20659510	53.5	6630	196	117	0.24											
4	X0002	X0004	2	12957400	13057400	62	2370	170	104	0.24											
5	X0002	X0004	2	228458917	236438781	21.4	2154	207	124	0.26											
6	X0002	X0004	3	4602847	64457067	73	9539	203	124	0.25											
7	X0002	X0004	3	4957500	5000000	10	1141	221	154	0.24											
8	X0002	X0004	4	71564	4276475	5.4	479	192	115	0.21											
9	X0002	X0004	4	5717807	11428521	14.6	1284	197	118	0.27											
10	X0002	X0004	4	2140000	2140000	43	4520	203	122	0.23											
11	X0002	X0004	5	38319	2172086	6.1	411	196	118	0.25											
12	X0002	X0004	5	3247420	3635334	5.3	626	190	116	0.25											
13	X0002	X0004	5	3283789	11673749	50	6819	209	125	0.24											
14	X0002	X0004	5	1494479	1500000	5.4	5300	210	127	0.27											
15	X0002	X0004	6	15561574	170821800	26.6	3127	201	121	0.28											
16	X0002	X0004	7	3246019	37466970	6	807	213	126	0.26											
17	X0002	X0004	7	1545460	1545460	5.3	270	210	125	0.29											
18	X0002	X0004	8	9550259	101862078	7.1	904	205	123	0.29											
19	X0002	X0004	8	12486264	139260468	25.6	2889	216	129	0.33											
20	X0002	X0004	9	4230000	4230000	13	150	190	117	0.27											
21	X0002	X0004	9	46587	10369718	25.6	2973	201	124	0.33											
22	X0002	X0004	10	138564	3077299	6.7	693	205	122	0.31											
23	X0002	X0004	10	1330564	1330564	4.9	427	227	134	0.3											
24	X0002	X0004	10	88832502	114518004	23.2	3450	218	131	0.27											
25	X0002	X0004	10	12913913	135368163	16.8	1610	225	134	0.31											
26	X0002	X0004	11	1254200	1254200	33	3627	217	130	0.29											
27	X0002	X0004	11	78873621	8348648	5	778	218	130	0.28											
28	X0002	X0004	11	12500573	13008638	9.3	1046	216	129	0.32											
29	X0002	X0004	12	19527826	20692607	1.3	1	220	130	0.28											
30	X0002	X0004	12	19527826	3835205	39.3	3451	217	130	0.31											
31	X0002	X0004	13	81192479	8861592	6.2	985	211	133	0.23											
32	X0002	X0004	14	10182846	10743294	7.8	461	204	122	0.22											
33	X0002	X0004	14	20182846	20182846	7.4	597	227	133	0.21											
34	X0002	X0004	15	3879088	58779441	18.5	2730	214	128	0.27											
35	X0002	X0004	16	7011318	74098542	73.6	8044	241	145	0.29											
36	X0002	X0004	16	7793453	8103823	3.4	347	227	133	0.29											
37	X0002	X0004	17	15346	1441448	5.9	323	216	129	0.31											
38	X0002	X0004	17	4938201	61614642	3.2	284	224	134	0.33											
39	X0002	X0004	17	2938201	3041200	4.4	4603	215	128	0.28											
40	X0002	X0004	17	7793453	8103823	6.9	263	205	123	0.14											
41	X0002	X0004	18	10096503	47153899	37.8	4821	231	138	0.29											
42	X0002	X0004	18	7387202	7418202	18.7	1863	225	130	0.26											
43	X0002	X0004	19	260912	6774089	1.3	1311	213	127	0.1											

Chromosome 13 looks like:

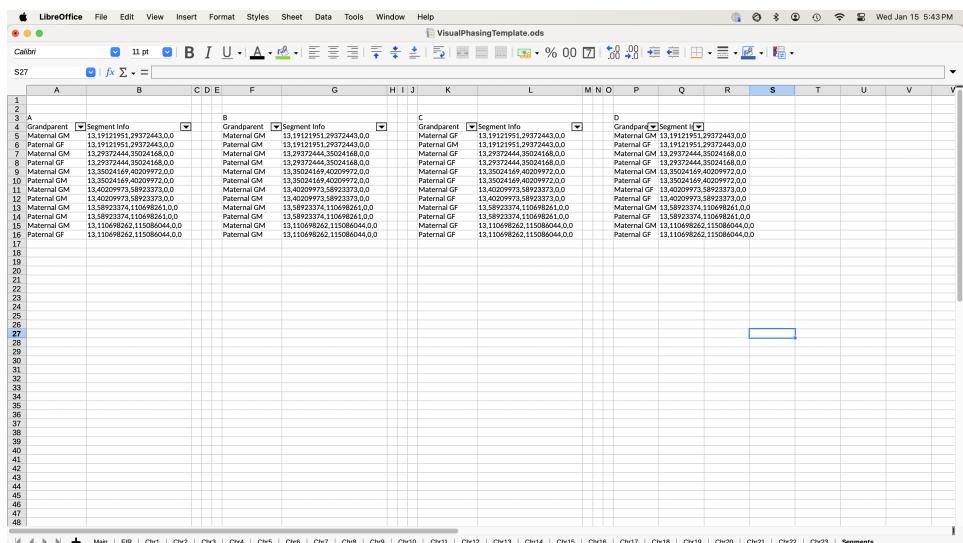


After the recombination points have been identified and the Mbp for each has been entered, the final sheet should be:



## Segment Lists

Segment lists can be generated by clicking the **Generate Segments** button on the Main sheet. For each Chromosome sheet, all of the segments that have an Mbp value and a segment grandparent labeled, an entry will be added to the Segment sheet. This is an OPTIONAL process.



## Appendix

# User Callable Macros

## Macros tied to buttons on the Main sheet

### **RebuildChromosomeSheets**

Remove all Chromosome sheets and regenerate Chromosome sheets. This should be done whenever there are changes to the Sibling Table.

### **LoadAllImages**

Loads all Chromosome sheets with the Sibling graphs. This needs to be done after resetting the Chromosome sheets.

### **LoadCousin**

Load the Cousin data for each Chromosome sheet.

### **UnloadCousin**

Remove the Cousin data for each Chromosome sheet.

### **GenerateSegmentList**

Create a new sheet labeled Segment and fill it with any segment information that has been defined. This data can be used in DNApainter, for instance.

### **Legend Map Changed**

To regenerate all of the painted area colors if the format in the Main sheet Grandparent Table has changed. This is tied to a button in the Main sheet

### **LoadFIR**

Load the FIR data for each Chromosome sheet.

## Macros not tied to buttons on the Main sheet

### **resetPaintValue**

Remove all of the text in the painted area. This allows the user to start the visual paint process again

### **replacePaintValue**

Replace an existing text string (like G1) in the painted area with a new text string (like M1)

### **RemoveChromosomeSheets**

Delete all Chromosome sheets

### **ShowRecombinationPoints**

Generates the recombination Mbp value assuming the graphs are linear in Mbp. The graphs generated by GEDmatch are not, so this is not useful for these. There are other applications that may generate graphs which are linear and can be used. In this case, the max Mbp need to be entered in Row 1, second Column after the painted area. If this is blank, the cell is painted red to indicate the error and the Recombination values are generated as a percentage.