

## GWAS - Genome wide associated studies

### Step 1 : Clone Tassel from GitHub

Command :-

```
git clone https://bitbucket.org/tasseladmin/tassel-5-standalone.git  
cd tassel-5-standalone/  
git pull  
.start_tassel.pl -Xmx6g
```

```
~/Downloads/300subset/Demo/Demo_GWAS/Demofn  
> git clone https://bitbucket.org/tasseladmin/tassel-5-standalone.git  
Cloning into 'tassel-5-standalone'...  
remote: Enumerating objects: 739, done.  
remote: Counting objects: 100% (739/739), done.  
remote: Compressing objects: 100% (690/690), done.  
remote: Total 739 (delta 368), reused 101 (delta 46), pack-reused 0 (from 0)  
Receiving objects: 100% (739/739), 202.01 MiB | 3.18 MiB/s, done.  
Resolving deltas: 100% (368/368), done.
```

### Step 2 : Convert the ped file into a tped

```
plink --file 300_GWAS --transpose --allow-extra-chr --const-fid --cow --recode --out mytransposeddata
```

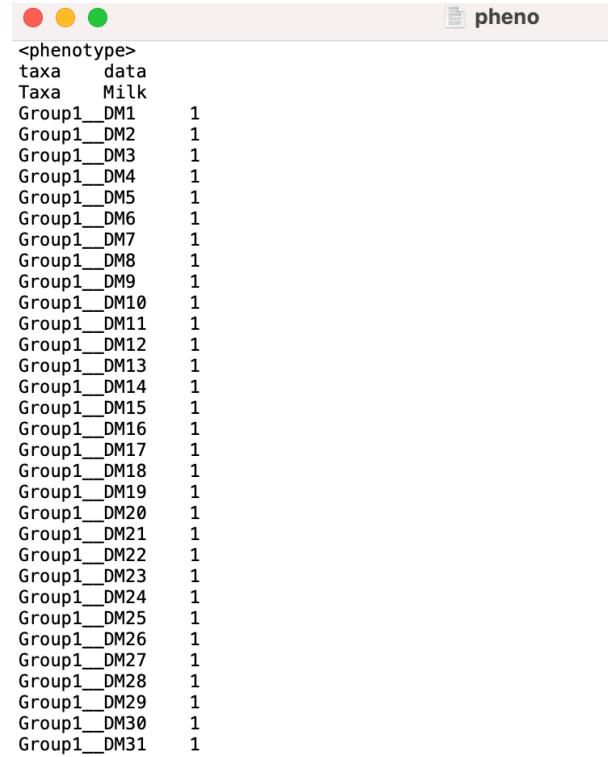
```
~/Downloads/300subset/Demo/Demo_GWAS/Demofn  
> plink --file 300_GWAS --transpose --allow-extra-chr --const-fid --cow --recode --out mytransposeddata  
PLINK v1.90b6.21 64-bit (19 Oct 2020) www.cog-genomics.org/plink/1.9/  
(C) 2005-2020 Shaun Purcell, Christopher Chang GNU General Public License v3  
Logging to mytransposeddata.log.  
Options in effect:  
--allow-extra-chr  
--const-fid  
--cow  
--file 300_GWAS  
--out mytransposeddata  
--recode  
--transpose  
  
Note: --transpose flag deprecated. Use "--recode transpose ...".  
32768 MB RAM detected; reserving 16384 MB for main workspace.  
.ped scan complete (for binary autoconversion).  
Performing single-pass .bed write (737689 variants, 300 cattle).  
--file: mytransposeddata-temporary.bed + mytransposeddata-temporary.bim +  
mytransposeddata-temporary.fam written.  
737689 variants loaded from .bim file.  
300 cattle (0 males, 0 females, 300 ambiguous) loaded from .fam.  
Ambiguous sex IDs written to mytransposeddata.nosex .  

```

### Step 3 : Convert the tped file into a hapmap format

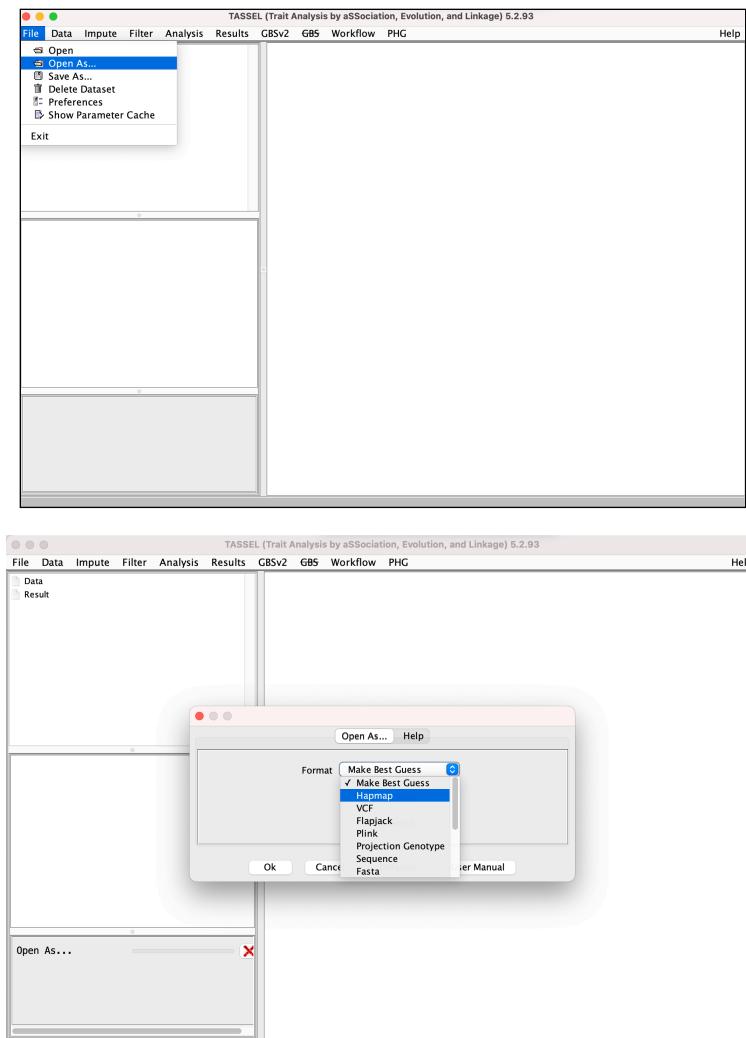
```
perl convert_tped_to_hapamp.pl --tped mytransposeddata.tped --tfam  
mytransposeddata.tfam -build=ncbi_36
```

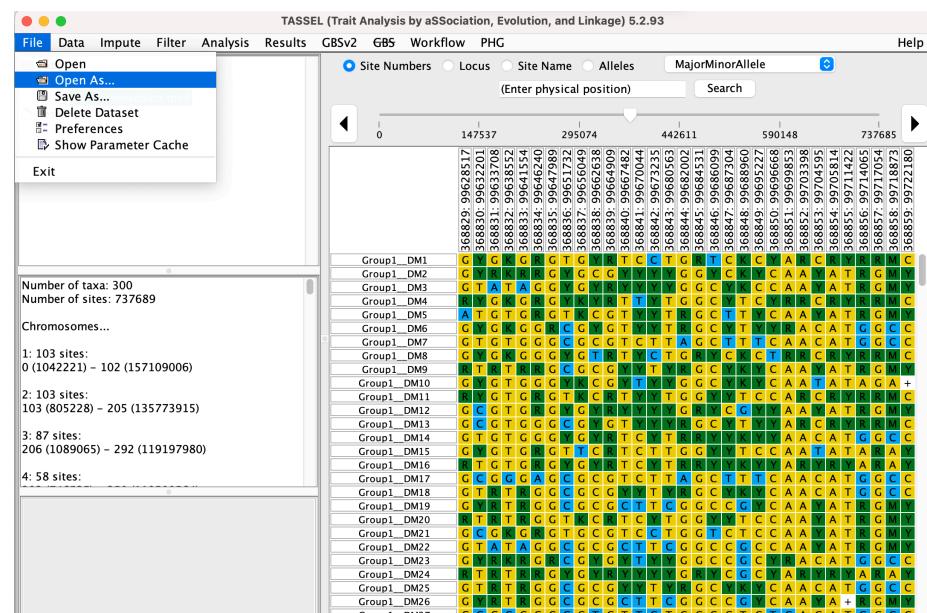
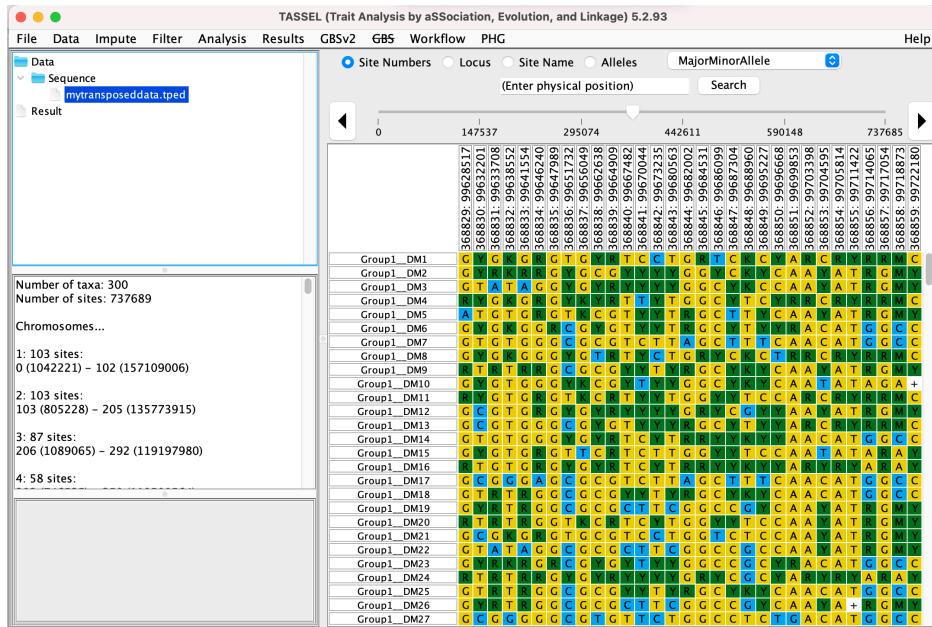
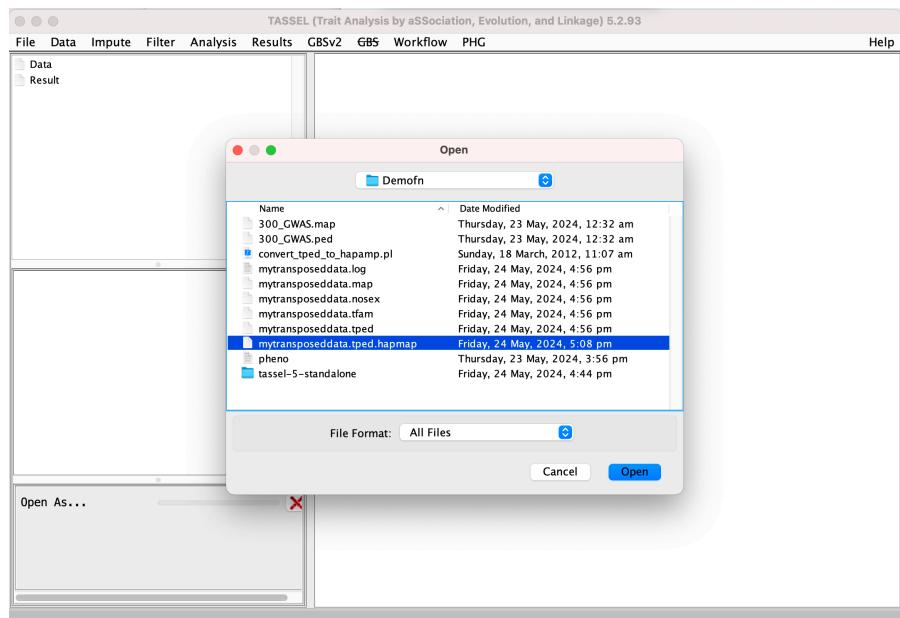
## Step 4 : Prepare the pheno file



```
<phenotype>
taxa    data
Taxa    Milk
Group1_DM1    1
Group1_DM2    1
Group1_DM3    1
Group1_DM4    1
Group1_DM5    1
Group1_DM6    1
Group1_DM7    1
Group1_DM8    1
Group1_DM9    1
Group1_DM10   1
Group1_DM11   1
Group1_DM12   1
Group1_DM13   1
Group1_DM14   1
Group1_DM15   1
Group1_DM16   1
Group1_DM17   1
Group1_DM18   1
Group1_DM19   1
Group1_DM20   1
Group1_DM21   1
Group1_DM22   1
Group1_DM23   1
Group1_DM24   1
Group1_DM25   1
Group1_DM26   1
Group1_DM27   1
Group1_DM28   1
Group1_DM29   1
Group1_DM30   1
Group1_DM31   1
```

## Step 5 : Call in the pheno file and the hap map file into Tassel





TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93

**Data**   **Sequence**   **Result**

Number of taxa: 300  
Number of sites: 737689

Chromosomes...

1: 103 sites:  
0 (104221) – 102 (157109006)

2: 103 sites:  
103 (805228) – 205 (135773915)

3: 87 sites:  
206 (1089065) – 292 (119197980)

4: 58 sites:

Open As... X

**File Format:** All Files Open

**Demofn**

Name	Date Modified
300_GWAS.map	Thursday, 23 May, 2024, 12:32 am
300_GWAS.ped	Thursday, 23 May, 2024, 12:32 am
convert_tped_to_hapmap.pl	Sunday, 18 March, 2024, 11:07 am
mytransposeddata.log	Friday, 24 May, 2024, 4:56 pm
mytransposeddata.map	Friday, 24 May, 2024, 4:56 pm
mytransposeddata.nosex	Friday, 24 May, 2024, 4:56 pm
mytransposeddata.fam	Friday, 24 May, 2024, 4:56 pm
mytransposeddata.tped	Friday, 24 May, 2024, 4:56 pm
mytransposeddata.hapmap	Friday, 24 May, 2024, 5:08 pm
<b>pheno</b>	<b>Thursday, 23 May, 2024, 3:56 pm</b>
tassel-5-standalone	Friday, 24 May, 2024, 4:44 pm

**File Format:** All Files Open

**TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93**

**Data**   **Numerical**   **Sequence**   **Result**

Table Title: pheno  
Number of columns: 2  
Number of rows: 300  
Matrix size (excludes row headers): 300

Taxa	Milk
Group1_DM1	1
Group1_DM2	1
Group1_DM3	1
Group1_DM4	1
Group1_DM5	1
Group1_DM6	1
Group1_DM7	1
Group1_DM8	1
Group1_DM9	1
Group1_DM10	1
Group1_DM11	1
Group1_DM12	1
Group1_DM13	1
Group1_DM14	1
Group1_DM15	1
Group1_DM16	1
Group1_DM17	1
Group1_DM18	1
Group1_DM19	1
Group1_DM20	1
Group1_DM21	1
Group1_DM22	1
Group1_DM23	1
Group1_DM24	1
Group1_DM25	1
Group1_DM26	1
Group1_DM27	1
Group1_DM28	1
Group1_DM29	1
Group1_DM30	1
Group1_DM31	1
Group1_DM32	1
Group1_DM33	1
Group1_DM34	1
Group1_DM35	1
Group1_DM36	1
Group1_DM37	1
Group1_DM38	1
Group1_DM39	1
Group1_DM40	1
Count	1

## Step 6 : Calculate the distance matrix from the hapmapfile and go for getting the PCs in MDS in Tassel

**Analysis**

**Distance Matrix**

**Relatedness**

**Diversity**

**GBSv2**   **GBS**   **Workflow**   **PHG**

**Data**   **Numerical**   **Sequence**   **Result**

Number of taxa: 300  
Number of sites: 737689

Chromosomes...

1: 103 sites:  
0 (104221) – 102 (157109006)

2: 103 sites:  
103 (805228) – 205 (135773915)

3: 87 sites:  
206 (1089065) – 292 (119197980)

4: 58 sites:

**Remove NaN From Distance Matrix**  
**Subtract Distance Matrix**  
**Add Distance Matrix**  
**Taxa Identity Recognition**

**TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93**

**Data**   **Numerical**   **Sequence**   **Result**

Table Title: pheno  
Number of columns: 2  
Number of rows: 300  
Matrix size (excludes row headers): 300

Taxa	Milk
Group1_DM7	1
Group1_DM8	1
Group1_DM9	1
Group1_DM10	1
Group1_DM11	1
Group1_DM12	1
Group1_DM13	1
Group1_DM14	1
Group1_DM15	1
Group1_DM16	1
Group1_DM17	1
Group1_DM18	1
Group1_DM19	1
Group1_DM20	1
Group1_DM21	1
Group1_DM22	1
Group1_DM23	1
Group1_DM24	1
Group1_DM25	1
Group1_DM26	1
Group1_DM27	1

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.9.3

File Data Impute Filter Analysis Results GBSv2 GBS Workflow PHC Help

**Data**

- Matrix
  - Matrix: mytransposedata.tped
- Numerical
  - pheno
- Sequence
  - mytransposedata.tped
- Result

**Table Title:** Distance Matrix  
**Number of columns:** 301  
**Number of rows:** 300  
**Matrix size (excludes row headers):** 90000

**Annotations...**  
 IBS\_Distance\_Matrix.AverageTotalSites: 737689.0  
 IBS\_Distance\_Matrix.NumAlleles: 3  
 IBS\_Distance\_Matrix.TruelIBS: false  
 Matrix\_Type: IBS\_Distance\_Matrix

**Distance Matrix**

Taxa	Group1..._0								
Group1_DM1	0	0.361	0.372	0.384	0.378	0.38	0.38	0	0
Group1_DM2	0.361	0	0.368	0.383	0.383	0.381	0.381	0	0
Group1_DM3	0.372	0.368	0	0.385	0.383	0.382	0.381	0	0
Group1_DM4	0.384	0.383	0.385	0	0.388	0.385	0.388	0	0
Group1_DM5	0.378	0.382	0.383	0.388	0	0.379	0.379	0	0
Group1_DM6	0.38	0.381	0.382	0.385	0.379	0	0.379	0	0
Group1_DM7	0.38	0.381	0.381	0.388	0.379	0.375	0	0	0
Group1_DM8	0.381	0.38	0.383	0.314	0.386	0.383	0.385	0	0
Group1_DM9	0.381	0.383	0.383	0.347	0.386	0.383	0.385	0	0
Group1_DM10	0.381	0.382	0.382	0.388	0.385	0.382	0.382	0	0
Group1_DM11	0.366	0.347	0.363	0.384	0.383	0.382	0.382	0	0
Group1_DM12	0.381	0.38	0.381	0.384	0.383	0.38	0.38	0	0
Group1_DM13	0.37	0.352	0.372	0.387	0.382	0.382	0.385	0	0
Group1_DM14	0.382	0.381	0.382	0.386	0.382	0.382	0.383	0	0
Group1_DM15	0.381	0.38	0.381	0.389	0.378	0.376	0.378	0	0
Group1_DM16	0.381	0.383	0.383	0.331	0.386	0.385	0.386	0	0
Group1_DM17	0.379	0.379	0.381	0.385	0.383	0.38	0.383	0	0
Group1_DM18	0.383	0.384	0.382	0.387	0.386	0.384	0.385	0	0
Group1_DM19	0.382	0.38	0.382	0.385	0.383	0.383	0.382	0	0
Group1_DM20	0.381	0.38	0.381	0.387	0.383	0.383	0.381	0	0
Group1_DM21	0.335	0.348	0	0.383	0.381	0.381	0.383	0	0
Group1_DM22	0.379	0.381	0.382	0.387	0.381	0.379	0.379	0	0
Group1_DM23	0.38	0.381	0.379	0.387	0.383	0.382	0.382	0	0
Group1_DM24	0.379	0.381	0.379	0.387	0.383	0.382	0.382	0	0
Group1_DM25	0.373	0.366	0.372	0.383	0.383	0.381	0.381	0	0
Group1_DM26	0.382	0.381	0.382	0.385	0.382	0.382	0.382	0	0
Group1_DM27	0.38	0.38	0.383	0.314	0.386	0.383	0.385	0	0
Group1_DM28	0.38	0.383	0.381	0.385	0.378	0.378	0.379	0	0
Group1_DM29	0.381	0.381	0.383	0.386	0.386	0.383	0.382	0	0
Group1_DM30	0.373	0.37	0.327	0.386	0.382	0.383	0.381	0	0
Group1_DM31	0.369	0.363	0.358	0.383	0.384	0.381	0.382	0	0
Group1_DM32	0.383	0.38	0.382	0.387	0.384	0.385	0.382	0	0
Group1_DM33	0.392	0.394	0.393	0.396	0.394	0.394	0.396	0	0
Group1_DM34	0.384	0.384	0.386	0.392	0.387	0.386	0.386	0	0
Group1_DM35	0.368	0.36	0.373	0.389	0.386	0.383	0.382	0	0
Group1_DM36	0.392	0.391	0.394	0.395	0.392	0.392	0.395	0	0
Group1_DM37	0.384	0.384	0.385	0.389	0.385	0.388	0.387	0	0
Group1_DM38	0.374	0.372	0.375	0.39	0.385	0.387	0.386	0	0
Group1_DM39	0.386	0.385	0.386	0.39	0.388	0.385	0.387	0	0
Group1_DM40	0.386	0.384	0.386	0.386	0.380	0.387	0.384	0.386	0

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.9.3

File Data Impute Filter Analysis Results GBSv2 GBS Workflow PHC Help

**Data**

- Matrix
  - Matrix: mytransposedata.tped
- Numerical
  - pheno
- Sequence
  - mytransposedata.tped
- Result

**Table Title:** Distance Matrix  
**Number of columns:** 301  
**Number of rows:** 300  
**Matrix size (excludes row headers):** 90000

**Annotations...**  
 IBS\_Distance\_Matrix.AverageTotalSites: 737689.0  
 IBS\_Distance\_Matrix.NumAlleles: 3  
 IBS\_Distance\_Matrix.TruelIBS: false  
 Matrix\_Type: IBS\_Distance\_Matrix

**Distance Matrix**

Taxa	Group1..._0								
Group1_DM1	0	0.361	0.372	0.384	0.378	0.38	0.38	0	0
Group1_DM2	0.361	0	0.368	0.383	0.382	0.381	0.381	0	0
Group1_DM3	0.372	0.368	0	0.385	0.383	0.382	0.381	0	0
Group1_DM4	0.384	0.383	0.385	0	0.388	0.385	0.388	0	0
Group1_DM5	0.378	0.382	0.383	0.388	0	0.379	0.379	0	0
Group1_DM6	0.38	0.381	0.382	0.385	0.379	0	0.375	0	0
Group1_DM7	0.38	0.381	0.381	0.388	0.379	0.375	0	0	0
Group1_DM8	0.381	0.38	0.383	0.314	0.386	0.383	0.385	0	0
Group1_DM9	0.381	0.383	0.383	0.347	0.386	0.383	0.385	0	0
Group1_DM10	0.381	0.382	0.382	0.388	0.385	0.382	0.382	0	0
Group1_DM11	0.366	0.347	0.363	0.384	0.384	0.383	0.38	0	0

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.9.3

File Data Impute Filter Analysis Results GBSv2 GBS Workflow PHC Help

**Data**

- Matrix
  - Matrix: mytransposedata.tped
- Numerical
  - pheno
- Sequence
  - mytransposedata.tped
- Result

**Table Title:** Distance Matrix  
**Number of columns:** 301  
**Number of rows:** 300  
**Matrix size (excludes row headers):** 90000

**Annotations...**  
 IBS\_Distance\_Matrix.AverageTotalSites: 737689.0  
 IBS\_Distance\_Matrix.NumAlleles: 3  
 IBS\_Distance\_Matrix.TruelIBS: false  
 Matrix\_Type: IBS\_Distance\_Matrix

**Distance Matrix**

Taxa	Group1..._0								
Group1_DM1	0	0.361	0.372	0.384	0.378	0.38	0.38	0	0
Group1_DM2	0.361	0	0.368	0.383	0.382	0.381	0.381	0	0
Group1_DM3	0.372	0.368	0	0.385	0.383	0.382	0.381	0	0
Group1_DM4	0.384	0.383	0.385	0	0.388	0.385	0.388	0	0
Group1_DM5	0.378	0.382	0.383	0.388	0	0.379	0.379	0	0
Group1_DM6	0.38	0.381	0.382	0.385	0.379	0	0.375	0	0
Group1_DM7	0.38	0.381	0.381	0.388	0.379	0.375	0	0	0
Group1_DM8	0.381	0.38	0.383	0.314	0.386	0.383	0.385	0	0
Group1_DM9	0.381	0.383	0.383	0.347	0.386	0.383	0.385	0	0
Group1_DM10	0.381	0.382	0.382	0.388	0.385	0.382	0.382	0	0
Group1_DM11	0.366	0.347	0.363	0.384	0.384	0.383	0.38	0	0

**MDS**

Number of Axes: 20

Remove NaNs:

OK Cancel Defaults User Manual

Taxa	Group1..._0								
Group1_DM7	0.36	0.369	0.373	0.374	0.378	0.378	0.378	0.379	0
Group1_DM8	0.38	0.383	0.381	0.385	0.378	0.378	0.378	0.379	0
Group1_DM9	0.381	0.381	0.383	0.386	0.386	0.383	0.382	0.381	0
Group1_DM10	0.373	0.37	0.327	0.386	0.382	0.382	0.383	0.381	0
Group1_DM11	0.369	0.363	0.358	0.383	0.384	0.384	0.382	0.381	0
Group1_DM12	0.383	0.38	0.382	0.387	0.389	0.385	0.386	0.387	0
Group1_DM13	0.392	0.394	0.393	0.396	0.394	0.394	0.396	0.396	0
Group1_DM14	0.384	0.384	0.386	0.386	0.392	0.387	0.386	0.386	0
Group1_DM15	0.368	0.36	0.373	0.389	0.386	0.387	0.387	0.387	0
Group1_DM16	0.384	0.384	0.385	0.389	0.387	0.388	0.385	0.387	0
Group1_DM17	0.379	0.379	0.381	0.385	0.383	0.384	0.385	0.387	0
Group1_DM18	0.383	0.384	0.382	0.387	0.386	0.384	0.385	0.387	0
Group1_DM19	0.382	0.38	0.383	0.388	0.387	0.386	0.385	0.387	0
Group1_DM20	0.381	0.38	0.381	0.387	0.383	0.383	0.381	0.381	0
Group1_DM21	0.335	0.348	0	0.383	0.381	0.381	0.383	0.383	0
Group1_DM22	0.379	0.381	0.382	0.387	0.381	0.382	0.379	0.379	0
Group1_DM23	0.38	0.381	0.379	0.384	0.384	0.381	0.382	0.382	0
Group1_DM24	0.379	0.381	0.382	0.383	0.382	0.382	0.382	0.382	0
Group1_DM25	0.373	0.366	0.372	0.383	0.383	0.381	0.381	0.381	0
Group1_DM26	0.382	0.381	0.382	0.385	0.382	0.382	0.382	0.382	0
Group1_DM27	0.38	0.38	0.383	0.385	0.386	0.378	0.378	0.379	0
Group1_DM28	0.38	0.383	0.381	0.385	0.382	0.382	0.382	0.382	0
Group1_DM29	0.381	0.381	0.383	0.386	0.386	0.383	0.383	0.382	0
Group1_DM30	0.373	0.37	0.327	0.386	0.382	0.382	0.383	0.381	0
Group1_DM31	0.369	0.363	0.358	0.383	0.384	0.384	0.382	0.382	0
Group1_DM32	0.383	0.38	0.382	0.387	0.384	0.385	0.382	0.382	0
Group1_DM33	0.392	0.394	0.393	0.396	0.394	0.394	0.396	0.396	0
Group1_DM34	0.384	0.384	0.386	0.386	0.392	0.387	0.386	0.386	0
Group1_DM35	0.368	0.36	0.373	0.389	0.386	0.387	0.387	0.387	0
Group1_DM36	0.392	0.391	0.394	0.395	0.395	0.392	0.392	0.395	0
Group1_DM37	0.384	0.384	0.385	0.389	0.385	0.385	0.388	0.387	0
Group1_DM38	0.374	0.372	0.375	0.39	0.385	0.387	0.387	0.386	0
Group1_DM39	0.386	0.385	0.386	0.39	0.388	0.385	0.385	0.387	0
Group1_DM40	0.386	0.384	0.386	0.386	0.380	0.387	0.384	0.386	0

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93										
File	Data	Impute	Filter	Analysis	Results	GBSv2	GBS	Workflow	PHG	Help
Matrix				Taxa	PC1	PC2	PC3	PC4	PC5	PC6
Matrix	Matrix:mytransposeddata.tped			Group_1_DM1	-0.013	-0.001	-0.05	-0.065	0.032	-0.036
Numerical				Group_1_DM2	-0.015	0	-0.052	-0.082	0.049	-0.06
pheno				Group_1_DM3	-0.012	0.004	-0.04	-0.039	0.012	-0.014
MDS_PCs_Matrix:mytransposeddata.tped				Group_1_DM4	0.113	-0.002	0.021	-0.02	0.04	0.029
MDS_Eigenvalues_Matrix:mytransposeddata.tpe				Group_1_DM5	-0.018	-0.005	-0.014	0.01	-0.003	-0.005
Sequence				Group_1_DM6	-0.015	-0.008	-0.015	0.019	0.011	0.002
mytransposeddata.tped				Group_1_DM7	-0.017	-0.004	-0.015	0.018	0.006	-0.001
Result				Group_1_DM8	0.156	-0.005	0.021	-0.021	0.037	0.038
				Group_1_DM9	0.107	-0.004	0.019	-0.013	0.038	0.025
				Group_1_DM10	-0.028	0.085	0.036	-0.001	0	0.005
				Group_1_DM11	-0.014	-0.001	-0.057	-0.083	0.046	-0.056
				Group_1_DM12	-0.022	0.061	0.023	0.001	0.002	0.001
				Group_1_DM13	-0.014	0.003	0.046	-0.006	0.026	-0.034
				Group_1_DM14	-0.019	0.056	0.019	-0.001	0.001	0.002
				Group_1_DM15	-0.023	-0.011	-0.023	0.039	0.032	0.008
				Group_1_DM16	0.123	-0.007	0.021	-0.018	0.035	0.03
				Group_1_DM17	-0.023	0.061	0.022	-0.003	-0.001	0.002
				Group_1_DM18	-0.014	0.074	0.028	0.001	0.001	-0.001
				Group_1_DM19	-0.019	0.061	0.023	0.002	0.002	0.006
				Group_1_DM20	0.028	0.094	0.041	-0.004	0.003	0.001
				Group_1_DM21	-0.014	-0.001	-0.058	0.084	0.047	-0.058
				Group_1_DM22	-0.035	-0.052	0.044	-0.005	-0.001	0.002
				Group_1_DM23	-0.022	0.062	0.021	-0.001	0	0.001
				Group_1_DM24	-0.022	0.069	0.024	-0.005	0.002	0.002
				Group_1_DM25	-0.017	0.023	-0.016	-0.032	0.019	-0.021
				Group_1_DM26	-0.025	0.078	0.033	-0.001	0.003	0.005
				Group_1_DM27	0.103	-0.004	0.017	-0.014	0.034	0.024
				Group_1_DM28	-0.015	-0.021	-0.002	0.011	-0.001	0.006
				Group_1_DM29	-0.027	0.087	0.039	-0.005	0.001	0.004
				Group_1_DM30	0.012	0.004	-0.041	-0.004	0.007	-0.011
				Group_1_DM31	-0.011	0.002	-0.036	0.039	0.011	-0.005
				Group_1_DM32	-0.031	0.104	0.05	-0.005	0.005	0.004
				Group_1_DM33	-0.007	-0.006	-0.032	0.006	-0.062	-0.028
				Group_1_DM34	-0.032	0.092	0.039	-0.004	0.001	0.002
				Group_1_DM35	-0.016	0.006	-0.047	-0.073	0.041	-0.048
				Group_1_DM36	-0.03	0.106	0.057	-0.003	0.006	0.005
				Group_1_DM37	-0.003	0.101	0.047	-0.003	0.004	0.003
				Group_1_DM38	-0.015	0.005	-0.043	-0.043	0.013	-0.017
				Group_1_DM39	-0.026	0.073	0.056	-0.002	0	0.002
				Group_1_DM40	-0.076	0.114X	0.123X	-0.061	0.002	-0.003

## **Step 7 : Calculate the kinship matrix using the hapmapfile in Tassel**

The screenshot shows the Tassel software interface. At the top, the menu bar includes File, Data, Impute, Filter, Analysis, Results, GB5v2, GB5, Workflow, PHG, and Help. The main window displays a file browser on the left with entries like 'Matrix:mytransposedata.tped' and 'Numerical:pheno'. On the right, there's a search interface for site numbers, locus, site name, alleles, and major/minor allele, with a search button and a text field '(Enter physical position)'. Below this is a kinship matrix visualization with chromosomes as columns and samples as rows. A context menu is open over the matrix with options: Kinship, Help, Centered\_IBS (selected), Max Alleles (6), Observed\_Allele\_Freq, Ok, Cancel, Defaults, and User Manual.

## Step 8 : Intersect the pheno, hapmap file and PCs

**TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93**

**Data**

- Matrix
  - Matrix:mytransposedata.tped
  - Centered\_IBS\_mytransposedata.tped
- Numerical
  - pheno
  - MDS\_PCs\_Matrix:mytransposedata.tped
  - MDS\_Eigenvalues\_Matrix:mytransposedata.tped
- Sequence
  - mytransposedata.tped

**Result**

Table Title: pheno  
Number of columns: 2  
Number of rows: 300  
Matrix size (excludes row headers): 300

Taxa	Milk
Group1_DM1	1
Group1_DM2	1
Group1_DM3	1
Group1_DM4	1
Group1_DM5	1
Group1_DM6	1
Group1_DM7	1
Group1_DM8	1
Group1_DM9	1
Group1_DM10	1
Group1_DM11	1
Group1_DM12	1
Group1_DM13	1
Group1_DM14	1
Group1_DM15	1
Group1_DM16	1
Group1_DM17	1
Group1_DM18	1
Group1_DM19	1
Group1_DM20	1
Group1_DM21	1
Group1_DM22	1
Group1_DM23	1
Group1_DM24	1
Group1_DM25	1
Group1_DM26	1
Group1_DM27	1
Group1_DM28	1
Group1_DM29	1
Group1_DM30	1
Group1_DM31	1
Group1_DM32	1
Group1_DM33	1
Group1_DM34	1
Group1_DM35	1
Group1_DM36	1
Group1_DM37	1
Group1_DM38	1
Group1_DM39	1
Group1_DM40	1
Group1_DM41	1

class net.maizegenetics.phenotype.CorePhenotype

**TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93**

**Data**

- Get Taxa List
- Get Position List
- Sort Taxa Alphabetically
- Sort Genotype File
- P-> Synonymizer
- Intersect Join
- Union Join
- Merge Genotype Tables
- Separate
- Homozygous Genotype
- SetLowDepthGenosToMissing
- Change Indels to Unknown
- Transform Phenotype
- 0..1 Numerical Genotype
- ABH Genotype
- Thin Sites by Position
- Cluster Genotypes
- Mask Genotype
- Find Inversions
- Create Hybrid Genotypes
- Geno Summary

**Result**

Table Title: pheno  
Number of columns: 23  
Number of rows: 300  
Matrix size (excludes row headers): 6600

Taxa	Milk
Group1_DM1	1
Group1_DM2	1
Group1_DM3	1
Group1_DM4	1
Group1_DM5	1
Group1_DM6	1
Group1_DM7	1
Group1_DM8	1
Group1_DM9	1
Group1_DM10	1
Group1_DM11	1
Group1_DM12	1
Group1_DM13	1
Group1_DM14	1
Group1_DM15	1
Group1_DM16	1
Group1_DM17	1
Group1_DM18	1
Group1_DM19	1
Group1_DM20	1
Group1_DM21	1
Group1_DM22	1
Group1_DM23	1
Group1_DM24	1
Group1_DM25	1
Group1_DM26	1
Group1_DM27	1
Group1_DM28	1
Group1_DM29	1
Group1_DM30	1
Group1_DM31	1
Group1_DM32	1
Group1_DM33	1
Group1_DM34	1
Group1_DM35	1
Group1_DM36	1
Group1_DM37	1
Group1_DM38	1
Group1_DM39	1
Group1_DM40	1
Group1_DM41	1

class net.maizegenetics.phenotype.CorePhenotype

**TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93**

**Data**

- Matrix
  - Matrix:mytransposedata.tped
  - Centered\_IBS\_mytransposedata.tped
- Numerical
  - pheno
  - MDS\_PCs\_Matrix:mytransposedata.tped
  - MDS\_Eigenvalues\_Matrix:mytransposedata.tped
  - pheno + MDS\_PCs\_Matrix:mytransposedata.tped
- Sequence
  - mytransposedata.tped

**Result**

Table Title: Phenotype\_with\_genotypes  
Number of columns: 23  
Number of rows: 300  
Matrix size (excludes row headers): 6600

Intersect Join

Taxa	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Group1_DM1	1	-0.013	-0.001	-0.05	-0.065	0.032	-0.036
Group1_DM2	1	-0.028	0.085	0.036	-0.001	0	0.005
Group1_DM10	1	-0.045	-0.077	0.067	-0.016	-0.006	0.004
Group1_DM11	1	-0.035	-0.052	0.038	-0.011	0.003	-0.001
Group1_DM12	1	-0.04	-0.068	0.057	-0.012	-0.002	0.002
Group1_DM13	1	-0.039	-0.062	0.057	-0.009	-0.002	0.002
Group1_DM14	1	-0.04	-0.062	0.051	-0.01	0.002	0.002
Group1_DM15	1	-0.035	-0.056	0.044	-0.009	-0.003	0.004
Group1_DM16	1	0.116	-0.006	0.013	-0.001	0.002	0.003
Group1_DM17	1	0.092	-0.003	0.006	0.01	-0.017	-0.012
Group1_DM18	1	0.001	0.007	-0.011	0.03	-0.038	-0.032
Group1_DM19	1	0.005	0.003	-0.014	0.029	-0.037	-0.033
Group1_DM20	1	-0.014	-0.001	-0.057	-0.083	0.046	-0.056
Group1_DM21	1	0.001	0.006	-0.011	0.027	-0.031	-0.029
Group1_DM22	1	0.001	-0	-0.016	0.035	-0.039	-0.037
Group1_DM23	1	0.003	-0.003	-0.016	0.038	-0.047	-0.048
Group1_DM24	1	0.118	-0.007	0.015	-0.003	0.004	0.006
Group1_DM25	1	0.132	-0.008	0.02	-0.001	0.007	0
Group1_DM26	1	0.113	-0.004	0.012	-0.002	0.007	0.01
Group1_DM27	1	0	0.007	-0.01	0.025	-0.027	-0.028
Group1_DM28	1	0.102	-0.003	0.007	0.001	-0.011	-0.007
Group1_DM29	1	0.004	0.004	-0.012	0.029	-0.034	-0.027
Group1_DM30	1	0.002	0.006	-0.01	0.031	-0.033	-0.034
Group1_DM31	1	-0.022	0.061	0.023	0.001	0.002	0.001
Group1_DM32	1	0.003	0.006	-0.009	0.03	-0.034	-0.03
Group1_DM33	1	0.002	0.002	-0.016	0.037	-0.049	-0.05
Group1_DM34	1	0.001	0.002	-0.01	0.026	-0.031	-0.029
Group1_DM35	1	0.088	-0.005	0.001	0.006	-0.031	-0.02
Group1_DM36	1	0.116	-0.006	0.012	-0.006	0.002	-0.004
Group1_DM37	1	0	0.001	-0.012	0.04	-0.049	-0.047
Group1_DM38	1	0.099	-0.005	0.011	0.005	-0.008	-0.005
Group1_DM39	1	0	0.002	-0.014	0.028	-0.033	-0.027
Group1_DM40	1	0.135	-0.005	0.024	-0.003	0.01	-
Group1_DM41	1	0.119	-0.007	0.018	-0.002	0.005	0.009
Group1_DM42	1	-0.014	0.003	-0.046	-0.006	0.026	-0.034
Group1_DM43	1	0.009	-0.006	0.007	0.008	-0.017	-0.013
Group1_DM44	1	0.014	-0.01	-0.039	0.025	-0.105	-0.055
Group1_DM45	1	0.002	0.001	-0.016	0.031	-0.039	-0.035
Group1_DM46	1	0.002	0.003	-0.013	0.03	-0.038	-0.037
Group1_DM47	1	0.002	0.006	-0.011	0.03	-0.033	-0.032

class net.maizegenetics.phenotype.GenotypePhenotype

## Step 9 : Do MLM with IBS correction by selecting the files as shown in the figure below

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93

**Analysis**

- Results GBsV2 GB5 Workflow PHC Help
- Data
- Matrix
- Centered\_IBS\_mytransposedata.tped
- Numerical
- pheno
- MDS\_PCs\_Matrix:mytransposedata.tped
- MDS\_Eigenvalues\_Matrix:mytransposedata.tped
- pheno + MDS\_PCs\_Matrix:mytransposedata.tped
- Sequence
- mytransposedata.tped
- Result

Table Title: Phenotype\_with\_genotypes  
Number of columns: 23  
Number of rows: 300  
Matrix size (excludes row headers): 6600  
Intersect Join

MLM

Taxa	Milk	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Group1_DM1	1	-0.013	-0.001	-0.05	-0.065	0.032	-0.036	0
Group1_DM10	1	-0.028	0.085	0.036	-0.001	0	0.005	0
Group1_DM100	1	-0.045	-0.077	0.067	-0.016	-0.006	0.004	0
Group1_DM101	1	-0.035	-0.052	0.038	-0.011	0.003	-0.001	0
Group1_DM102	1	-0.04	-0.068	0.051	-0.012	-0.002	0.002	0
Group1_DM103	1	-0.039	-0.062	0.053	-0.009	-0.002	0.002	0
Group1_DM104	1	-0.04	-0.062	0.051	-0.012	-0.002	0.002	0
Group1_DM105	1	-0.035	-0.056	0.044	-0.009	-0.002	0.004	-0
Group1_DM106	1	-0.116	-0.006	0.012	-0.006	0.002	-0.004	-0
Group1_DM107	1	0	0.007	-0.01	0.025	-0.027	-0.028	-0
Group1_DM108	1	0.002	0.003	-0.014	0.029	-0.037	-0.033	-0
Group1_DM109	1	-0.014	-0.001	-0.057	-0.083	0.046	-0.056	0
Group1_DM110	1	0.001	0.000	-0.011	0.027	-0.031	-0.029	-0
Group1_DM111	1	0.001	0	-0.016	0.035	-0.039	-0.037	-0
Group1_DM112	1	0.003	-0.003	-0.016	0.038	-0.047	-0.048	-0
Group1_DM113	1	0.118	0.075	0.015	0.075	-0.004	0.016	0
Group1_DM114	1	0.132	-0.008	0.021	0.001	0.007	0.007	0
Group1_DM115	1	0.113	-0.004	0.012	-0.002	0.001	0.001	0
Group1_DM116	1	0	0.007	-0.01	0.025	-0.027	-0.028	-0
Group1_DM117	1	0.102	-0.003	0.007	0.001	-0.011	-0.007	0
Group1_DM118	1	0.004	0.000	-0.012	0.029	-0.034	-0.027	-0
Group1_DM119	1	0.002	0.000	-0.01	0.031	-0.033	-0.034	-0
Group1_DM120	1	-0.022	0.061	0.023	0.001	0.002	0.001	-0
Group1_DM121	1	0.003	0.006	-0.009	0.03	-0.034	-0.03	-0
Group1_DM122	1	0.002	0.016	-0.016	0.037	-0.037	-0.035	-0
Group1_DM123	1	0.001	0.002	-0.01	0.026	-0.031	-0.039	-0
Group1_DM124	1	0.088	-0.005	0.001	0.006	0.031	-0.02	0
Group1_DM125	1	0.116	-0.006	0.012	-0.006	0.002	-0.004	-0
Group1_DM126	1	0	0.001	-0.012	0.04	-0.049	-0.047	-0
Group1_DM127	1	0.099	-0.005	0.011	0.005	-0.008	-0.005	0
Group1_DM128	1	0.135	-0.005	0.024	-0.003	0.019	0.011	-
Group1_DM129	1	0.119	-0.007	0.018	-0.002	0.009	0.009	-0
Group1_DM130	1	-0.044	0.036	-0.016	0.006	0.026	-0.04	-0
Group1_DM131	1	0.009	-0.006	0.007	0.008	0.047	-0.013	-0
Group1_DM132	1	0.014	0.01	-0.039	0.025	-0.105	-0.055	0
Group1_DM133	1	0.002	0.001	-0.016	0.031	-0.039	-0.035	-0
Group1_DM134	1	0.002	0.003	-0.013	0.03	-0.038	-0.037	-0
Group1_DM135	1	0.003	0.006	-0.011	0.03	-0.033	-0.032	-0

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93

**Analysis**

- Results GBsV2 GB5 Workflow PHC Help
- Data
- Matrix
- Centered\_IBS\_mytransposedata.tped
- Numerical
- pheno
- MDS\_PCs\_Matrix:mytransposedata.tped
- MDS\_Eigenvalues\_Matrix:mytransposedata.tped
- pheno + MDS\_PCs\_Matrix:mytransposedata.tped
- Sequence
- mytransposedata.tped
- Result

Table Title: Phenotype\_with\_genotypes  
Number of columns: 23  
Number of rows: 300  
Matrix size (excludes row headers): 6600  
Intersect Join

MLM Options

Compression Level

- Optimum Level
- Custom Level:
- No Compression

Variance Component Estimation

- P3D (estimate once)
- Re-estimate after each marker

Run Cancel Help Me Choose

Taxa	Milk	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Group1_DM1	1	-0.013	-0.001	-0.05	-0.065	0.032	-0.036	0
Group1_DM10	1	-0.028	0.085	0.036	-0.001	0	0.005	0
Group1_DM100	1	-0.045	-0.077	0.067	-0.016	-0.006	0.004	0
Group1_DM101	1	-0.035	-0.052	0.038	-0.011	0.003	-0.001	0
Group1_DM102	1	-0.04	-0.068	0.051	-0.012	-0.002	0.002	0
Group1_DM103	1	-0.039	-0.062	0.053	-0.009	-0.002	0.002	0
Group1_DM104	1	-0.04	-0.062	0.051	-0.012	-0.002	0.002	0
Group1_DM105	1	-0.035	-0.056	0.044	-0.009	-0.002	0.004	-0
Group1_DM106	1	-0.116	-0.006	0.012	-0.001	0.003	-0.003	-0
Group1_DM107	1	0.092	-0.003	0.006	0.01	-0.017	-0.012	0
Group1_DM123	1	0.088	-0.005	0.01	0.004	-0.026	-0.029	-0
Group1_DM124	1	0.116	-0.006	0.012	-0.006	0.002	-0.004	-0
Group1_DM125	1	0	0.001	-0.012	0.04	-0.049	-0.047	-0
Group1_DM126	1	0.099	-0.005	0.011	0.005	-0.008	-0.005	0
Group1_DM127	1	0	0.002	-0.012	0.028	-0.033	-0.027	-0
Group1_DM128	1	0.135	-0.005	0.024	-0.003	0.019	0.011	-
Group1_DM129	1	0.119	-0.007	0.018	-0.002	0.011	-0.007	-0
Group1_DM130	1	0.09	-0.006	0.007	0.008	-0.017	-0.013	-0
Group1_DM131	1	0.014	-0.001	-0.039	0.025	-0.105	-0.055	0
Group1_DM132	1	0.002	0.001	-0.016	0.031	-0.039	-0.035	-0
Group1_DM133	1	0.002	0.003	-0.013	0.03	-0.038	-0.037	-0
Group1_DM134	1	0.002	0.006	-0.011	0.03	-0.032	-0.032	-0

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.93

**Analysis**

- Results GBsV2 GB5 Workflow PHC Help
- Data
- Matrix
- Centered\_IBS\_mytransposedata.tped
- Numerical
- pheno
- MDS\_PCs\_Matrix:mytransposedata.tped
- MDS\_Eigenvalues\_Matrix:mytransposedata.tped
- pheno + MDS\_PCs\_Matrix:mytransposedata.tped
- Sequence
- mytransposedata.tped
- Result

Table Title: Phenotype\_with\_genotypes  
Number of columns: 23  
Number of rows: 300  
Matrix size (excludes row headers): 6600  
Intersect Join

Choose Output Format

By default all output will be stored in memory. If the output is large enough it could exceed the memory space available to TASSEL and cause TASSEL to crash. If you think this could be a problem, you can choose to write the output to a text file instead. Enter only the file base name. The dataset name and file type will be appended when the report is saved.

Write output to file

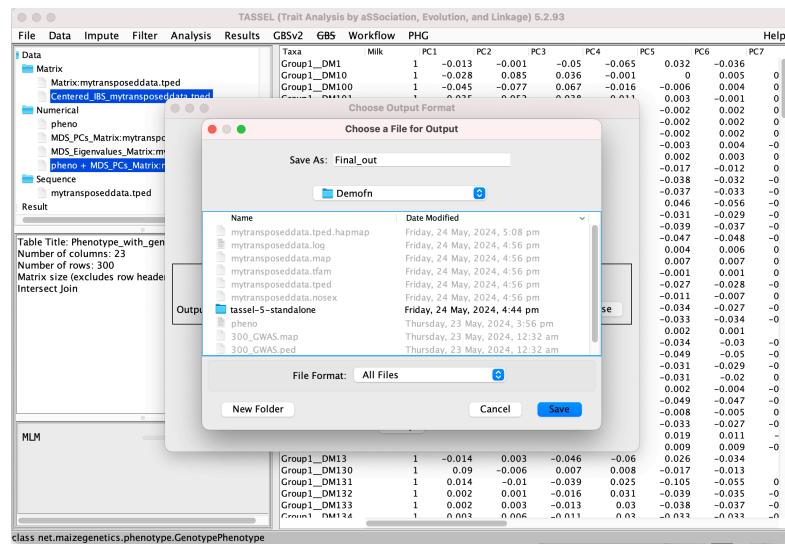
Output base file name:  Browse

Filter output on p-value

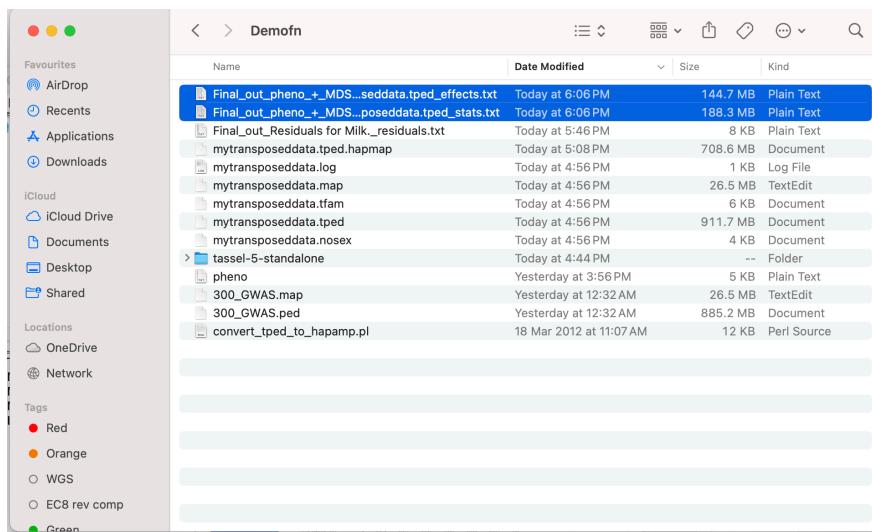
Do not keep p-values larger than  1e-3

Okay

Taxa	Milk	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Group1_DM1	1	-0.013	-0.001	-0.05	-0.065	0.032	-0.036	0
Group1_DM10	1	-0.028	0.085	0.036	-0.001	0	0.005	0
Group1_DM100	1	-0.045	-0.077	0.067	-0.016	-0.006	0.004	0
Group1_DM101	1	-0.035	-0.052	0.038	-0.011	0.003	-0.001	0
Group1_DM102	1	-0.04	-0.068	0.051	-0.012	-0.002	0.002	0
Group1_DM103	1	-0.039	-0.062	0.053	-0.009	-0.002	0.002	0
Group1_DM104	1	-0.04	-0.062	0.051	-0.012	-0.002	0.002	0
Group1_DM105	1	-0.035	-0.056	0.044	-0.009	-0.002	0.004	-0
Group1_DM106	1	-0.116	-0.006	0.012	-0.001	0.003	-0.003	-0
Group1_DM107	1	0.092	-0.003	0.006	0.01	-0.002	0.002	0
Group1_DM123	1	0.088	-0.005	0.01	0.004	-0.026	-0.029	-0
Group1_DM124	1	0.116	-0.006	0.012	-0.006	0.002	-0.004	-0
Group1_DM125	1	0	0.001	-0.012	0.04	-0.049	-0.047	-0
Group1_DM126	1	0.099	-0.005	0.011	0.005	-0.008	-0.005	0
Group1_DM127	1	0	0.002	-0.012	0.028	-0.033	-0.027	-0
Group1_DM128	1	0.135	-0.005	0.024	-0.003	0.019	0.011	-
Group1_DM129	1	0.119	-0.007	0.018	-0.002	0.009	0.009	-0
Group1_DM130	1	0.09	-0.006	0.007	0.008	-0.017	-0.013	-0
Group1_DM131	1	0.014	-0.001	-0.039	0.025	-0.105	-0.055	0
Group1_DM132	1	0.002	0.001	-0.016	0.031	-0.039	-0.035	-0
Group1_DM133	1	0.002	0.003	-0.013	0.03	-0.038	-0.037	-0
Group1_DM134	1	0.002	0.006	-0.011	0.03	-0.032	-0.032	-0



**Step 10 : Two Output files are generated, rename the stats file as Stats.txt, and pull out the required fields as in step 11**



**Step 11 : Awk the required columns**

```
awk 'BEGIN{FS="\t";}{print $2,$3,$4,$7;}' Stats.txt > pvalues.txt
```

```
~/Downloads/300subset/Demo/Demo_GWAS/Demofn
```

```
> awk 'BEGIN{FS="\t";}{print $2,$3,$4,$7;}' Stats.txt > pvalues.txt
```

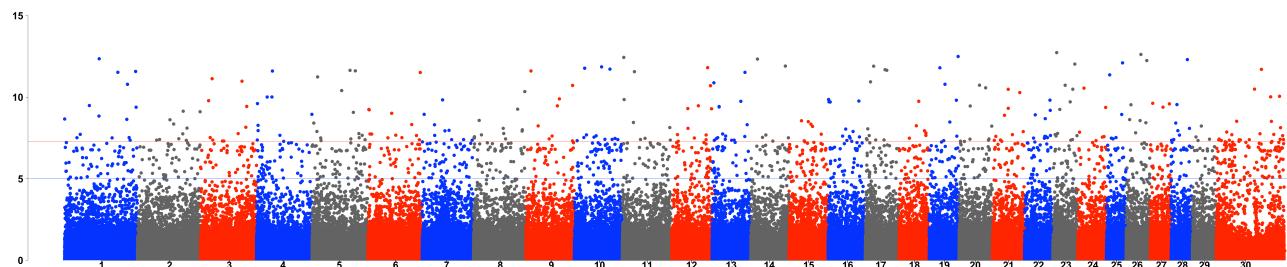
pvalues.txt				
Marker	Chr	Pos	P	
None			NaN	
AX-509427750	1	1042221	0.5172675067081663	
AX-171452203	1	1148684	0.828691180785311	
AX-169487117	1	1564289	0.5702650770136717	
AX-509806720	1	1623206	0.6821314085170298	
AX-117090525	1	2257731	0.046728492240382	
AX-169169554	1	3323416	0.7711770499150251	
AX-509657683	1	3575435	0.12313687426234	
AX-430054288	1	3700335	0.4811574897425122	
AX-312423467	1	3894961	0.4114124642673877	
AX-509640297	1	3977122	0.4214338733410613	
AX-320768980	1	4491418	0.02650551913414925	
AX-429583508	1	6045529	0.8888498647955103	
AX-429587127	1	7753872	0.06973446130107802	
AX-310540716	1	8649341	0.7195593498762846	
AX-429340279	1	10632901	0.384815165780763	
AX-310552422	1	13626703	0.2813299718491153	
AX-428847643	1	14226390	0.1439524996702932	
AX-169381377	1	15752303	0.24041107000000002	
AX-370677578	1	18822104	0.937702514453092	
AX-430080474	1	19942562	0.7808212702811722	
AX-169381377	1	20555044	0.05110076825477238	
AX-169514563	1	22639749	0.22639749	
AX-171467326	1	24647973	0.6770527853634355	
AX-428870833	1	25351726	0.7315146643592952	
AX-509311215	1	27915261	0.26348840731852397	
AX-509515935	1	29293699	0.2079917583645916	
AX-169402280	1	30369063	0.8223755451711551	
AX-519783484	1	30614172	0.13436012582240617	

**Step 12 : Sort the pvalues, remove the NaNs, decode the chromosomes and use the values for Manhattan and QQ plot and lambda estimation**

```
sort -k4n pvalues.txt > pvalue_sorted.txt
```

**Step 12 : Manhattan**

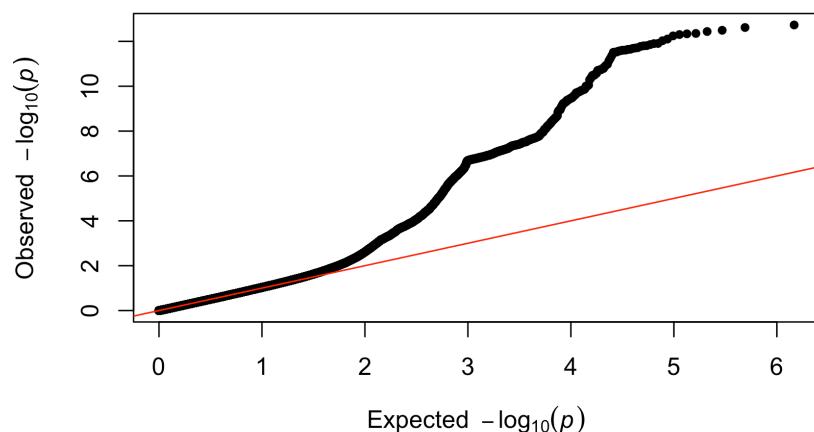
```
library(qqman)
library(QCEWAS)
tiff(res = 300, "Example.tiff", units = "cm", width = 145, height = 35)
man <- read.table("pvalue_decoded.txt", header=T)
head(man)
P_lambda (man$p)
as.data.frame(table(man$CHR))
d <- manhattan(man, chr =
'CHR', bp='Pos', p='p', snp='Marker', col=c("blue","#666666","red"), lo
gp =
TRUE, ylab='', xlab='', cex=2.5, cex.lab=1.25, cex.axis=2.6, ylim=c(0.0,
17.0), font.axis=2, lwd=2)
dev.off()
```



Lambda was 1.001997

**Step 13 : QQ-plot**

```
library(qqman)
tiff(res = 300, "QQ.tiff", units = "cm", width = 15, height = 10 )
man <- read.table("pvalue_decoded.txt", header=T)
qq(man$p)
dev.off()
```



Note : the data taken for the analysis is a hypothetical and may not reflect the Manhattan and qq-plot as per the standards