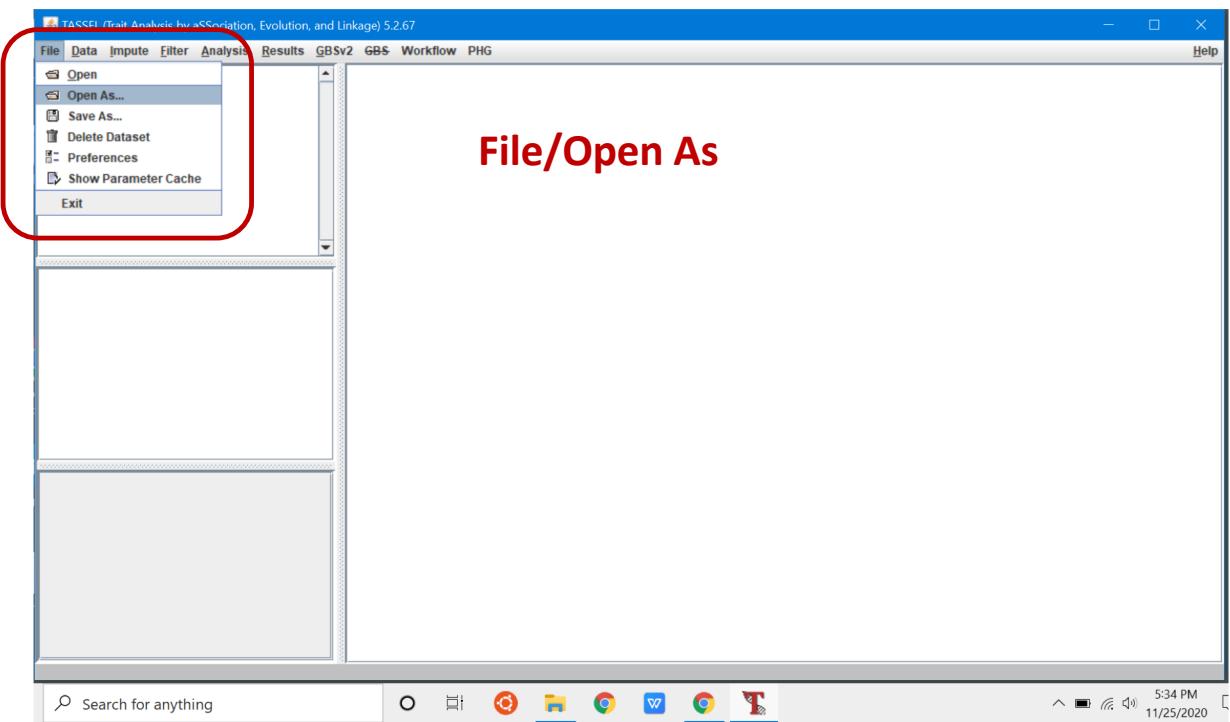
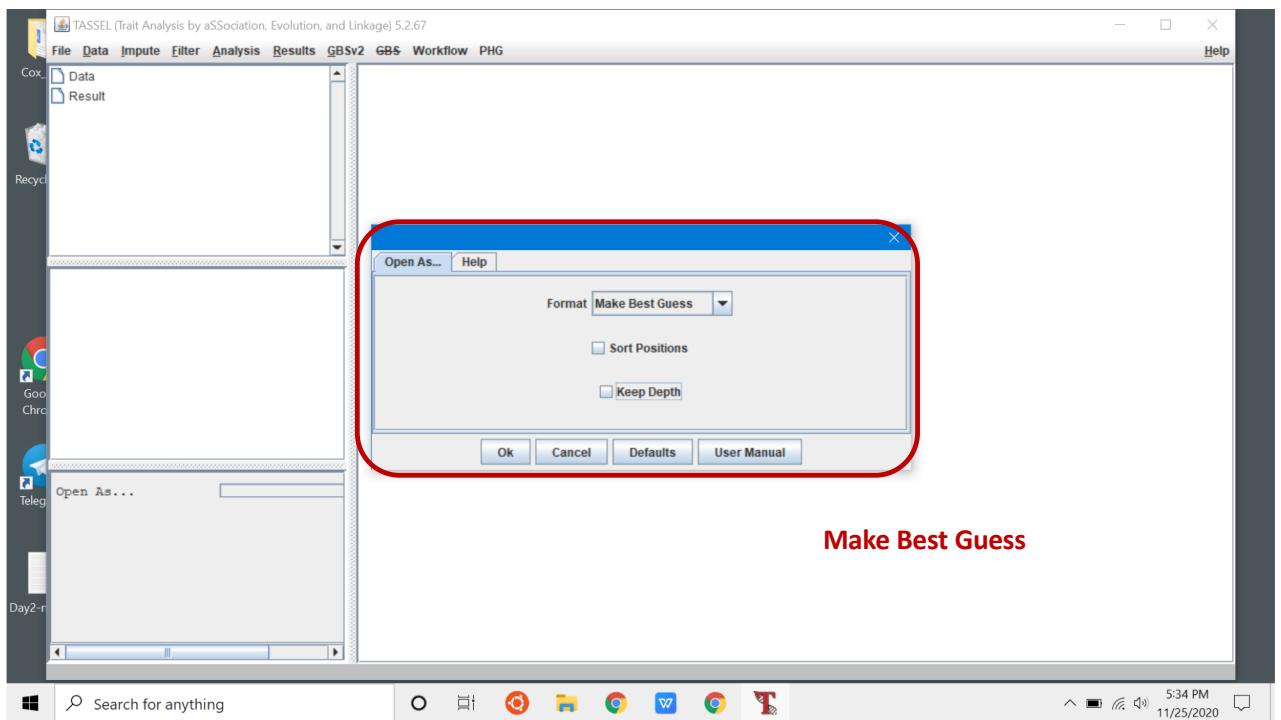
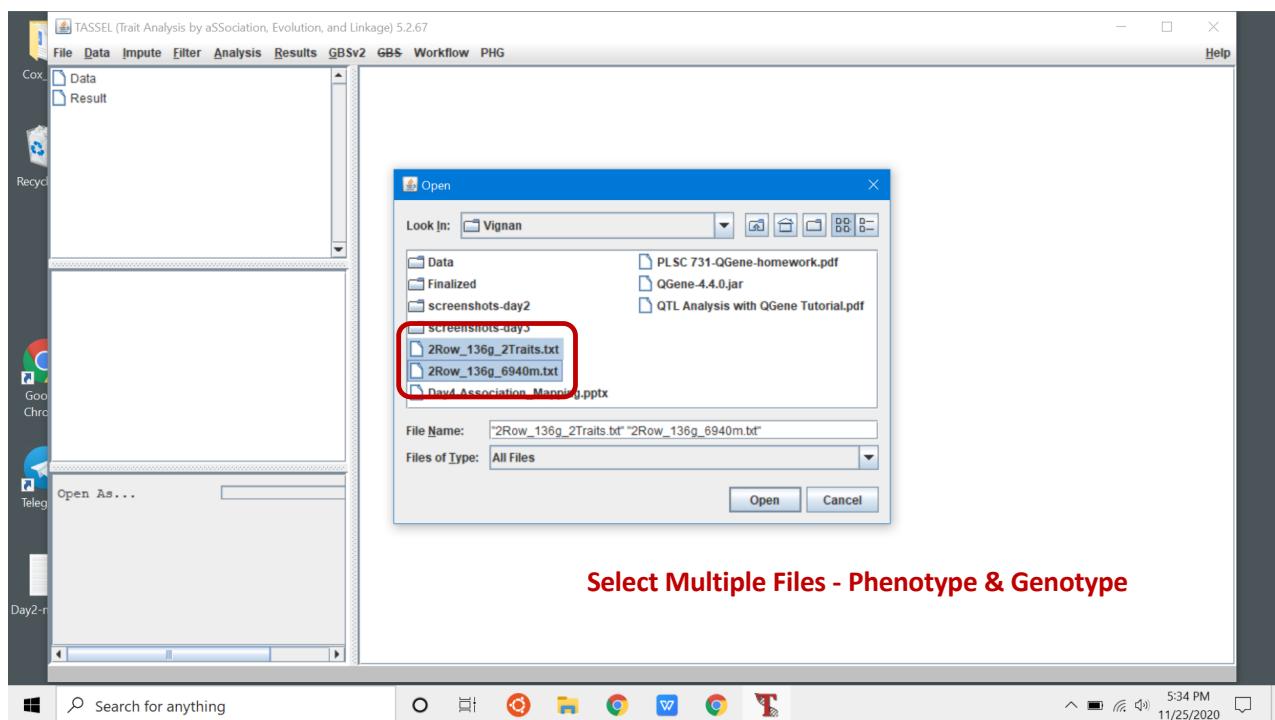


AM IN TASSEL

Sujan Mamidi







Phenotype File - Information
Impute Missing phenotype
Impute Missing Genotype

The screenshot shows the TASSEL (Trait Analysis by Association, Evolution, and Linkage) 5.2.67 software interface. The window title is "TASSEL (Trait Analysis by Association, Evolution, and Linkage) 5.2.67". The menu bar includes File, Data, Impute, Filter, Analysis, Results, GB, GBS, Workflow, DHC, and Help. The "Impute" menu is highlighted and has a red box around it. Sub-options under "Impute" include Extract Inbred Haplotypes by FILLIN, Impute By FILLIN, Impute By FSTHap, Numerical Impute, Remove indels for input to Beagle v.4, LD KNNI Imputation, and Evaluate Imputation Accuracy. A second red box highlights the "Extract Inbred Haplotypes by FILLIN" option. On the left, there's a file tree with "Data", "Results", and "Day2-r". The main workspace displays a table titled "Table Title: 2Row_136g_2Traits" with the following data:

Taxa	disease1	AUDPC
AM7	4	73
AM11	9	608
AM12	1	810
AM14	3	0
AM17	5	2,430
AM27	1	1,840
AM30	NaN	3,300
AM32	9	3,000
AM40	9	82.8
AM43	9	75
AM44	0	145
AM46	3	150
AM47	4	576
AM48	9	810
AM50	9	810
AM51	9	1,990
AM52	4	0
AM54	8	24.4
AM55	8	2,800
AM56	0	2,530
AM57	9	1,520
AM58	9	1,970
AM59	9	405
AM61	7	1,040
AM63	8	150
AM64	8	3,300
AM70	9	82.8
AM71	8	648
AM72	9	NaN
AM74	9	146
AM76	9	20.7
AM77	7	405
AM79	4	608
AM80	9	1,008
AM81	8	187.5
AM82	8	288

At the bottom, the status bar shows "Search for anything", the system tray icons, and the time "5:35 PM" and date "11/25/2020".

TASSEL (Trait Analysis by aSSocation, Evolution, and Linkage) 5.2.67

File Data Impute Filter Analysis Results GBSv2 GBS Workflow PHG Help

Cox Cox Recy Recy

Data Numerical Sequence Result

Number of taxa: 136 Number of sites: 6940 Chromosomes... UNKNOWN: 6940 sites: 0 (0) - 6939 (6939)

Diversity

Relatedness

Association

Physical Position

Distance Matrix Kinship Create Tree Merge Trees Subset Tree Remove Tree Branch Lengths Convert Tree Names Pedigree Relationship Matrix Combined A and G Relationship Matrix MDS PCA

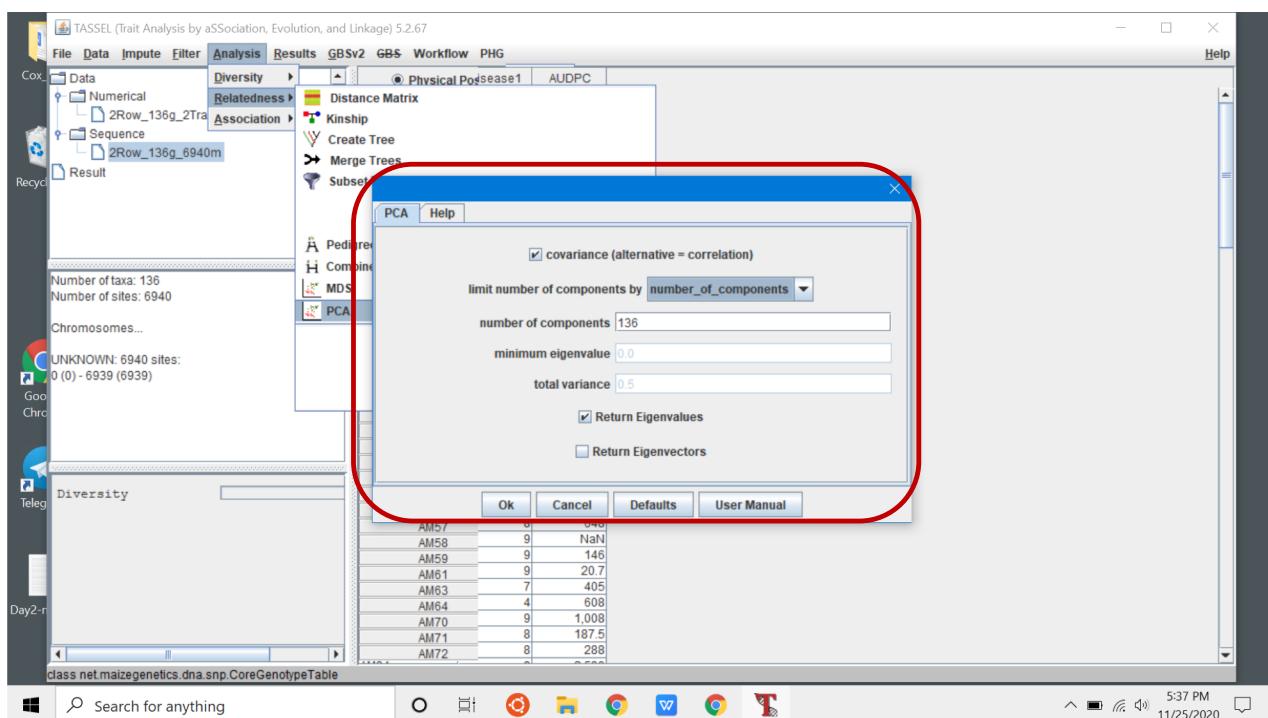
Remove NaN From Distance Matrix SubTRACT Distance Matrix Add Distance Matrix Taxa Identity Recognition

Perform principal components analysis

	AM48	AM50	AM51	AM52	AM54	AM55	AM56	AM57	AM58	AM59	AM61	AM63	AM64	AM66	AM70	AM71	AM72
	9	9	9	7	6	9	9	8	9	9	9	7	4	7	9	8	8
	1,320	1,970	405	1,040	150	3,300	82.8	648	NaN	146	20.7	405	608	1,008	187.5	288	288

Search for anything 5:37 PM 11/25/2020

- **Analysis/Relatedness**
- **S**
- **PCA**



**Set “number of components” to 10
Return Eigen values**

Results/Chart

Screenshot of TASSEL software interface showing results and charts.

The left sidebar shows project structure:

- Data
 - Numerical
 - 2Row_136g_2Traits
 - PC_2Row_136g_6940m
 - Eigenvalues_2Row_136g_6940m
 - Sequence
 - 2Row_136g_6940m
 - Result
- Cox...
- Recyc...
- Goo...
- Chro...
- Teleg...
- Day2-r

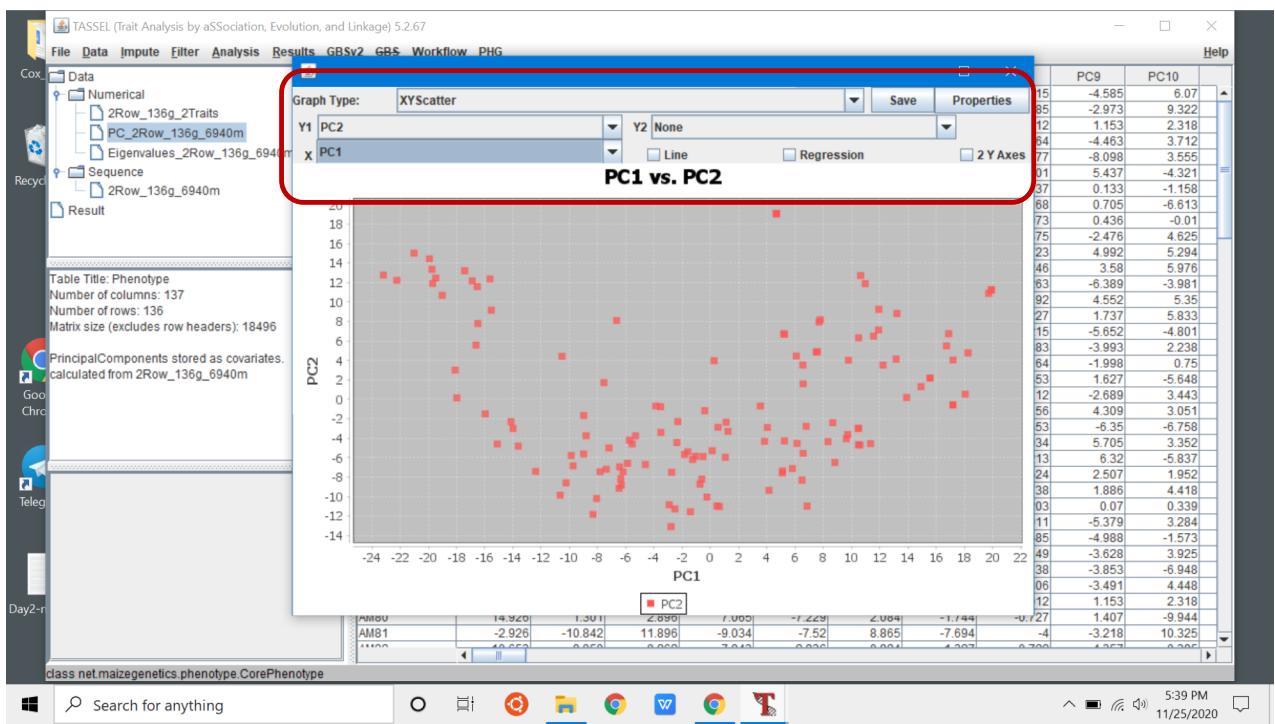
The main window displays a table of phenotype data with 136 rows and 137 columns. A red box highlights the "Charting Tools" section in the top menu bar, which includes options like Table, Archaeopteryx Tree, LD Plot, Chart, QQ Plot, and Manhattan Plot.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
6.086	4.463	-4.71	4.296	-1.745	0.764	-4.299	-5.715	-4.585	6.07	▲
-2.524	-11.26	13.234	-8.517	-7.669	7.334	-7.265	-2.585	-2.973	9.322	
10.49	-2.989	-12.14	8.621	-10.283	10.103	5.684	-1.012	1.153	2.318	
-1.259	-6.188	14.954	-10.203	-14.396	10.146	-7.589	-1.664	-4.463	3.712	
-4.613	-6.71	10.418	-11.44	-4.355	10.332	-3.089	-1.377	-8.098	3.555	
-1.05	-5.848	11.631	6.612	-9.602	-4.912	-2.757	6.101	5.437	-4.321	
-14.143	-2.296	-0.151	-7.718	-0.991	1.366	6.765	8.37	0.133	-1.158	
AM32	-15.971	-1.505	-6.38	-6.937	1.539	-7.673	2.26	0.168	0.705	-6.613
AM33	-12.394	-7.417	4.93	-4.955	4.653	4.297	3.456	6.073	0.436	-0.01
AM43	-16.534	11.594	-3.65	-7.088	-3.398	-2.508	-1.982	-1.475	-2.476	4.625
AM44	-16.888	12.17	1.109	1.466	-1.45	0.716	-6.362	-2.723	4.992	5.294
AM46	-6.657	8.105	-1.099	-8.368	-4.904	-7.226	2.451	-3.246	3.58	5.976
AM47	-19.739	13.361	-8.222	1.478	-5.526	0.34	-3.58	1.263	-6.389	-3.981
AM48	-17.423	13.225	1.305	0.898	-2.856	0.484	-7.381	-2.92	4.552	5.35
AM50	7.583	4.885	2.503	-13.001	-7.359	-12.606	8.369	-15.227	1.737	5.833
AM51	-21.012	15.035	-9.011	2.129	-4.353	0.461	-2.342	2.215	-5.652	-4.801
AM52	-15.535	9.166	-7.512	-7.261	2.28	0.549	-6.762	3.883	-3.993	2.238
AM54	-6.445	-6.937	4.11	-1.763	0.722	10.431	8.006	2.164	-1.998	0.75
AM55	19.913	11.243	10.388	-0.512	0.709	-7.872	-1.12	1.453	1.627	-5.648
AM56	-15.638	12.381	-4.79	-8.883	-2.703	-2.99	-2.172	-1.12	-2.689	3.443
AM57	0.25	3.981	-5.181	-12.09	1.137	-7.914	4.771	-1.56	4.309	3.051
AM58	-19.924	14.452	-8.948	1.378	-4.304	1.624	-1.312	1.653	-6.35	-6.758
AM59	-7.557	1.719	-2.444	-11.706	1.404	-1.669	-1.994	-1.034	5.705	3.352
AM61	-5.313	-3.765	1.245	-3.369	-3.162	-2.596	3.767	3.213	6.32	-5.837
AM63	-0.528	-5.893	5.393	2.704	-10.222	0.357	4.849	1.524	2.507	1.952
AM64	7.781	8.189	-0.863	-11.985	-6.395	-11.21	6.685	-13.738	1.886	4.418
AM70	3.536	-0.69	12.704	-0.726	-3.737	2.407	-4.454	8.203	0.07	0.339
AM71	0.125	-5.292	14.24	-11.251	-15.204	10.593	-7.703	-2.911	-5.379	3.284
AM72	6.139	-4.526	-3.887	6.813	0.313	-8.6	-0.78	-0.585	-4.988	-1.573
AM74	-1.829	-5.672	-2.388	1.825	0.726	2.215	14.115	-3.49	-3.628	3.925
AM76	11.921	7.124	8.271	-4.827	-2.492	-6.021	0.094	-3.138	-3.853	-6.948
AM77	8.827	-6.484	2.036	-1.911	1.873	-4.932	-2.789	12.806	-3.491	4.448
AM79	10.49	-2.989	-12.14	8.621	-10.283	10.103	5.684	-1.012	1.153	2.318
AM80	14.926	1.301	2.896	7.065	-7.229	2.084	-1.744	-0.727	1.407	-9.944
AM81	-2.926	-10.842	11.896	-9.034	-7.52	8.865	-7.694	-4	-3.218	10.325
AM82	10.662	-0.650	0.660	7.642	0.660	0.660	4.207	0.700	4.257	0.660

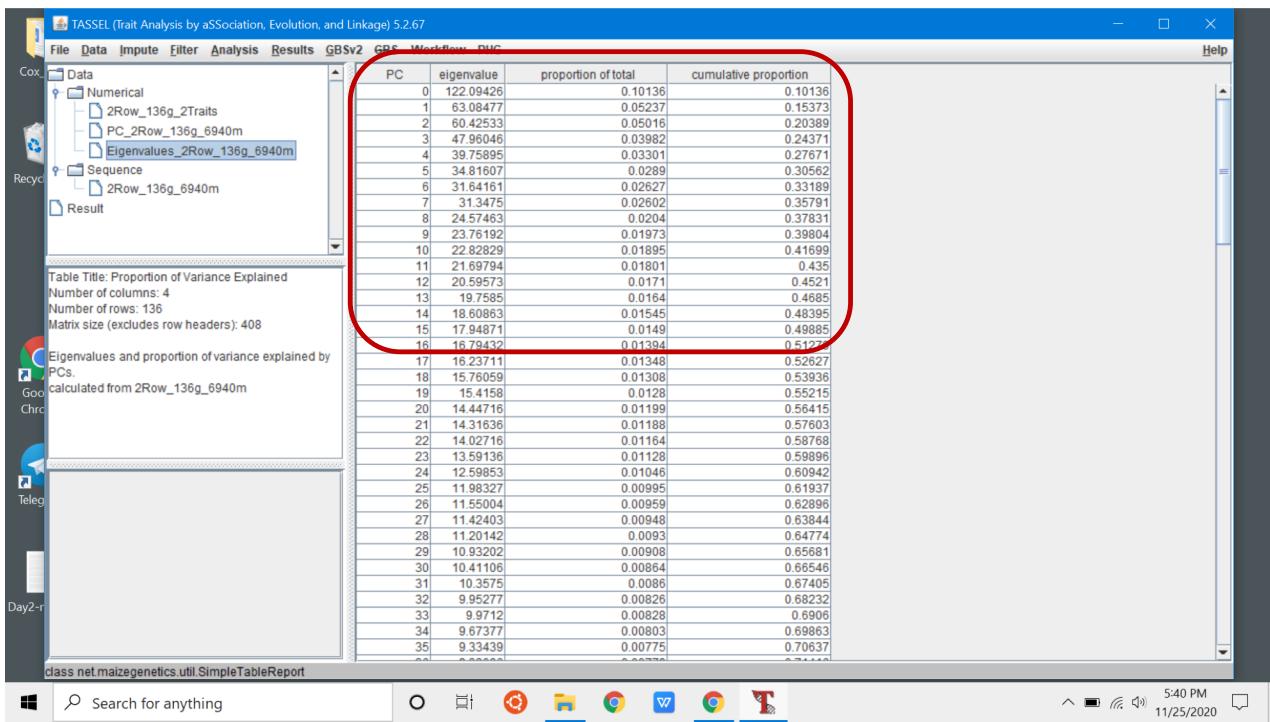
Bottom status bar: class net.maizegenetics.phenotype.CorePhenotype

Windows taskbar: Search for anything, File Explorer, Task View, Edge, Google Chrome, Microsoft Word, Microsoft Excel, Microsoft Powerpoint, Microsoft Word, 5:38 PM, 11/25/2020

PC Plot to understand population structure



PCA Eigen values - look for cumulative proportion



The screenshot shows the Tassel software interface. On the left, there's a file tree with 'Data' (Numerical and Sequence), 'Result', and other project files. In the center, a table titled 'Proportion of Variance Explained' is displayed. The table has four columns: PC, eigenvalue, proportion of total, and cumulative proportion. The first 16 rows are highlighted with a red box. The cumulative proportion for the 16th PC is 0.5127.

PC	eigenvalue	proportion of total	cumulative proportion
0	122.09426	0.10136	0.10136
1	63.08477	0.05237	0.15373
2	60.42533	0.05016	0.20389
3	47.96046	0.03982	0.24371
4	39.75895	0.03301	0.27671
5	34.81607	0.0289	0.30562
6	31.64161	0.02627	0.33189
7	31.3475	0.02602	0.35791
8	24.57463	0.0204	0.37831
9	23.76192	0.01973	0.39804
10	22.82829	0.01895	0.41699
11	21.69794	0.01801	0.435
12	20.59573	0.0171	0.4521
13	19.7585	0.0164	0.4685
14	18.60863	0.01545	0.48395
15	17.94871	0.0149	0.49885
16	16.79432	0.01394	0.5127
17	16.23711	0.01348	0.52627
18	15.76059	0.01308	0.53936
19	15.4158	0.0128	0.55215
20	14.44716	0.01199	0.56415
21	14.31636	0.01188	0.57603
22	14.02716	0.01164	0.58768
23	13.59136	0.01128	0.59896
24	12.59853	0.01046	0.60942
25	11.98327	0.00995	0.61937
26	11.55004	0.00959	0.62896
27	11.42403	0.00948	0.63844
28	11.20142	0.0093	0.64774
29	10.93202	0.00908	0.65681
30	10.41106	0.00864	0.66546
31	10.3575	0.0086	0.67405
32	9.95277	0.00826	0.68232
33	9.9712	0.00828	0.6906
34	9.67377	0.00803	0.69863
35	9.33439	0.00775	0.70637

Analysis/Kinship - Leave defaults

The screenshot shows the TASSEL software interface with the 'Analysis' tab selected. In the center, a 'Kinship' dialog box is open, highlighted by a red rectangle. The dialog box contains settings for 'Kinship method' (set to 'Centered_IBS'), 'Max Alleles' (set to 6), and 'Algorithm Variation' (set to 'Observed_Allele_Freq'). At the bottom of the dialog are 'Ok', 'Cancel', 'Defaults', and 'User Manual' buttons. Below the dialog, a table of genotypic data is visible, showing columns for Taxa, PC1, PC2, PC3, and PC4. The left sidebar displays project files and statistics: 'Number of taxa: 136', 'Number of sites: 6940', and 'UNKNOWN: 6940 sites: 0 (0) - 6939 (6939)'.

Taxa	PC1	PC2	PC3	PC4
AM6	-0.528	-5.893	5.393	2.04
AM64	7.781	8.189	-0.863	-11.985
AM70	3.536	-0.69	12.704	-0.726
AM71	0.125	-5.292	14.24	-11.251
AM72	6.139	-4.526	-3.887	6.813
AM74	-1.829	-5.672	-2.388	1.825
AM76	11.921	7.124	8.271	-4.827
AM77	8.827	-6.484	2.036	-1.911
AM79	10.49	-2.989	-12.14	8.621
AM80	14.926	1.301	2.896	7.065
AM81	-2.926	-10.842	11.896	-9.034
AM82	-10.653	-9.858	0.262	-7.843

The screenshot shows the TASSEL software interface. The main window displays a table with columns 'Taxa', 'disease1', and 'AUDPC'. The 'Analysis' menu is open, showing various options like 'Get Taxa List', 'Sort Genotype File', and 'Synonymizer'. The 'Intersect Join' option is highlighted with a red box. The bottom status bar shows the date and time: 6:00 PM 11/25/2020.

Taxa	disease1	AUDPC
AM7	4	73
AM11	9	608
AM12	1	810
AM14	3	0
AM17	5	2,430
AM27	1	1,840
AM30	NaN	3,300
AM32	9	3,000
AM40	9	82.8
AM43	9	75
AM44	0	145
AM46	3	150
AM47	4	576
AM48	9	810
AM50	9	810
AM51	9	1,990
AM52	4	0
AM54	8	24.4
AM55	8	2,800
AM56	0	2,530
AM57	9	1,520
AM58	9	1,970
AM59	9	405
AM61	7	1,040
AM63	8	150
AM64	8	3,300
AM70	9	82.8
AM71	8	648
AM72	9	NaN
AM74	9	146
AM76	9	20.7
AM77	7	405
AM79	4	608
AM80	9	1,008
AM81	8	187.5
AM82	8	288
AM84	9	2,530

- select Phenotype and genotype files
- data/intersect join - rename as "Naive"

TASSEL (trait Analysis by aSSociation, Evolution, and Linkage) 5.2.67

File Data Impute Filter Analysis Results GBSv2 GBS Workflow PHG Help

data

- Matrix
 - 2Row_136g_6940m_kinship
- Numerical
 - 2Row_136g_2Traits
 - 2Row_136g_2Traits_4PC
 - 2Row_136g_2Traits + 2Row_136g_6940m
- Sequence
 - 2Row_136g_6940m

Result

Table Title: 2Row_136g_2Traits
Number of columns: 3
Number of rows: 136
Matrix size (excludes row headers): 272

Taxa	disease1	AUDPC
AM7	4	73
AM11	9	608
AM12	1	810
AM14	3	0
AM17	5	2,430
AM27	1	1,840
AM30	NaN	3,300
AM32	9	3,000
AM40	9	82.8
AM43	9	75
AM44	0	145
AM46	3	150
AM47	4	576
AM48	9	810
AM50	9	810
AM51	9	1,990
AM52	4	0
AM54	8	24.4
AM55	8	2,800
AM56	0	2,530
AM57	9	1,520
AM58	9	1,970
AM59	9	405
AM61	7	1,040
AM63	8	150
AM64	8	3,300
AM70	9	82.8
AM71	8	648
AM72	9	NaN
AM74	9	146
AM76	9	20.7
AM77	7	405
AM79	4	608
AM80	9	1,008
AM81	8	187.5
AM82	8	288
AM84	9	2,530

Search for anything 6:00 PM 11/25/2020

- select Phenotype, genotype and PCA file
- data/intersect join - rename as "PC4"

TASSEL (Trait Analysis by association Evolution and Linkage) 5.2.67

Analysis menu is open, showing the following options:

- Diversity
- Relatedness
- Association
- GLM
- MLM
- Weighted MLM
- Genomic Selection
- Stepwise
- Fast-MT Association
- VCAP Scan

A red box highlights the "Association" option under the "Analysis" menu.

Table Title: 2Row_136g_2 Traits with genotypes

Number of columns: 4

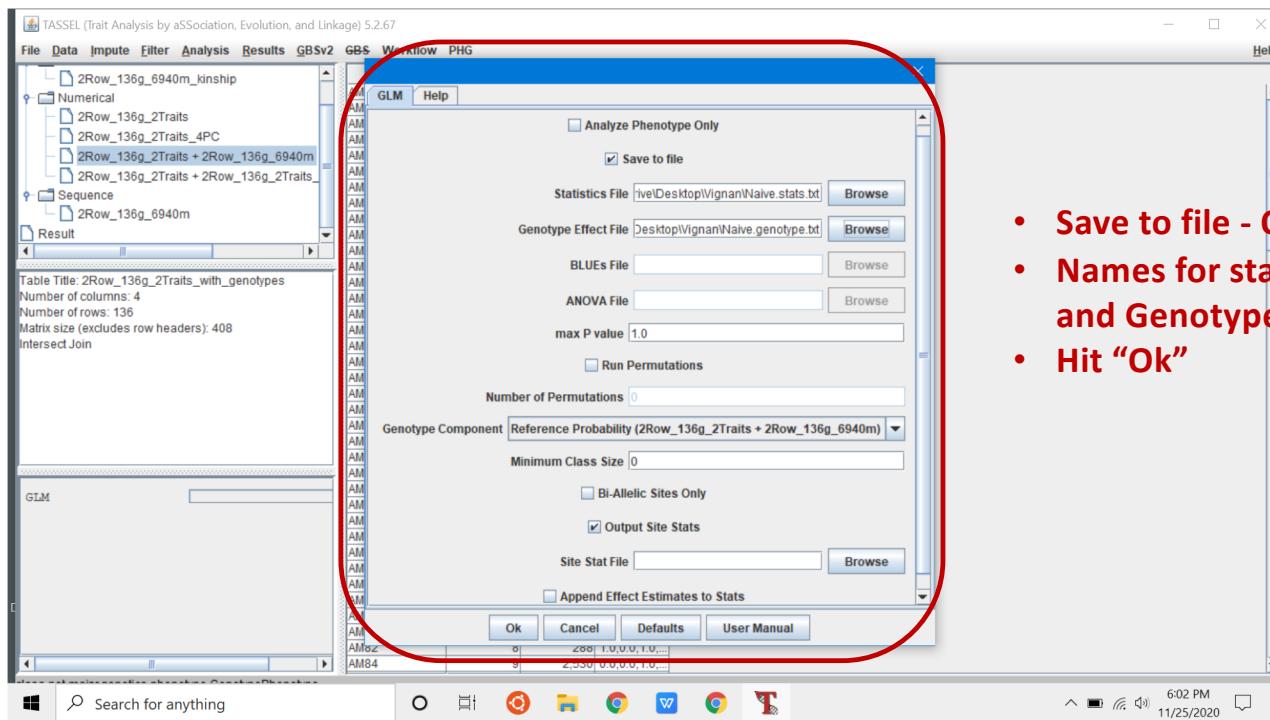
Number of rows: 136

Matrix size (excludes row headers): 408

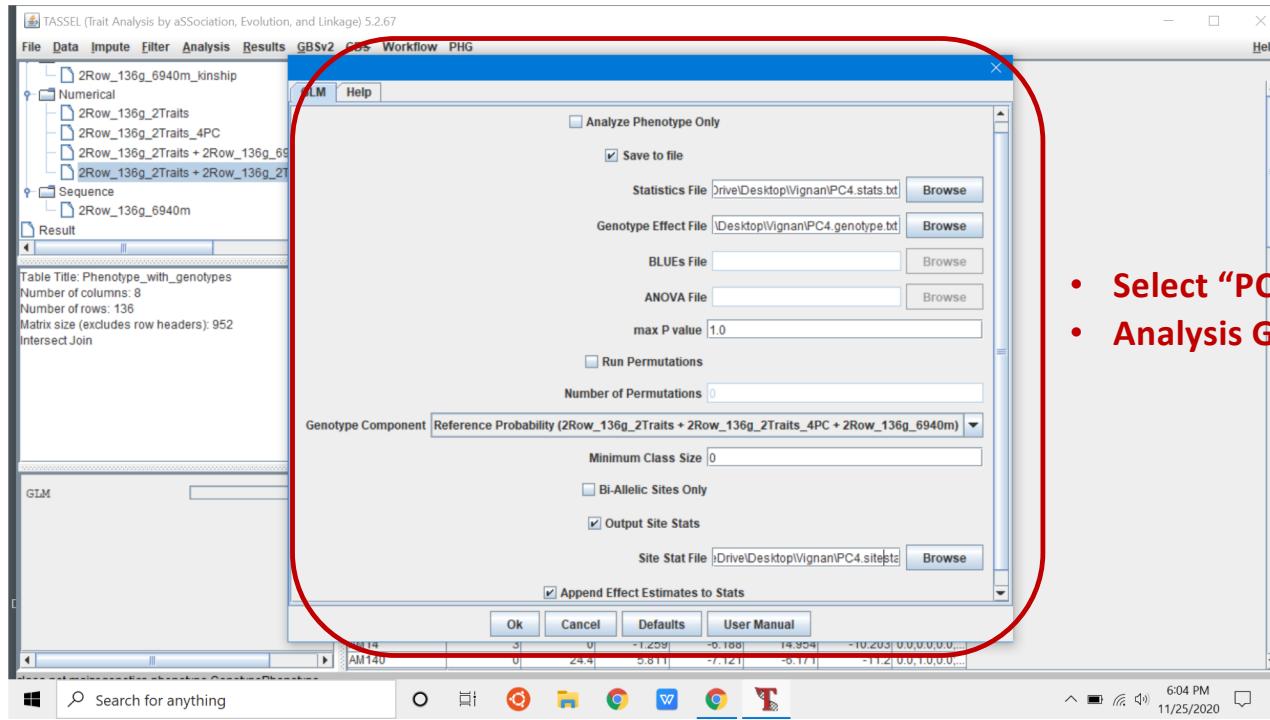
Intersect Join

	AM46	AM47	AM48	AM49	AM50	AM51	AM52	AM53	AM54	AM55	AM56	AM57	AM58	AM59	AM60	AM61	AM62	AM63	AM64	AM65	AM66	AM67	AM68	AM69	AM70	AM71	AM72	AM73	AM74	AM75	AM76	AM77	AM78	AM79	AM80	AM81	AM82	AM83	AM84	
Taxa	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
disease1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
ADPC	73	608	810	0	1.840	3,300	82.8	75	150	576	810	810	1,990	0	3,000	24.4	2,800	0	2,530	1,520	1,970	405	1,040	150	3,300	82.8	648	NaN	146	20.7	405	608	1,008	187.5	288	2,530				
Genotype	0.0.0.0.0....	0.0.0.0.0....	0.0.1.0.1....	0.0.0.0.0....	0.0.0.0.0....	0.0.1.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.1.0.0....	0.0.1.0.0....	0.0.1.0.0....	0.0.1.0.0....	0.0.1.0.0....	0.0.1.0.0....	0.0.1.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.1.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....	0.0.0.0.0....				

- Select “Naive” data
- Analysis GLM



- **Save to file - On**
- **Names for statistics file and Genotype effect file**
- **Hit "Ok"**



- Select “PC” data
- Analysis GLM

The screenshot shows the TASSEL software interface. On the left, there is a file tree with several entries under '2Row_136g_6940m_kinship'. A red box highlights this tree. In the center, there is a table viewer showing a dataset with columns 'Taxa', 'disease1', 'AUDPC', and 'Genotype'. A red box highlights this table. On the right, a 'MLM Options' dialog box is open, also with a red box around it. The dialog box contains sections for 'Compression Level' (with 'No Compression' selected), 'Variance Component Estimation' (with 'Re-estimate after each marker' selected), and buttons for 'Run', 'Cancel', and 'Help Me Choose'.

- Select “Naive” + kinship
- Analysis MLM
- “No Compression”
- “Re-estimate after each marker”
- “Run”
- Repeat with “PC4” + kinship

WPS Office Day4-AM_TasselSlides.pptx Book1

Home Insert Page Layout Formulas Data Review View Tools Q Cli... sujan ... Go Premium

Paste Cut Copy Format Painter Calibri 11 A A Merge and Center Wrap Text Conditional Formatting AutoSum AutoFilter Sort Format Fill Rows and Columns

N6

1	Trait	Marker	Chr	Pos	marker_F_p	marker_Rsq	marker_df	marker_MS	error_df	error_MS	model_df	model_MS
2	disease1	m1	UNKNOWI	0	0.004897	0.944324	3.89E-05	1	0.047508013	126	9.702438	1 0.047508
3	disease1	m2	UNKNOWI	1	2.203122	0.140229	0.017184623	1	21.00914094	126	9.536076	1 21.00914
4	disease1	m3	UNKNOWI	2	1.9235	0.16792	0.015036333	1	18.38273958	126	9.55692	1 18.38274
5	disease1	m4	UNKNOWI	3	0.026112	0.871887	2.07E-04	1	0.253303852	126	9.700805	1 0.253304
6	disease1	m5	UNKNOWI	4	2.580372	0.1107	0.020068162	1	24.53442558	126	9.508097	1 24.53443
7	disease1	m6	UNKNOWI	5	1.788193	0.183556	0.013993411	1	17.10770977	126	9.56704	1 17.10771
8	disease1	m7	UNKNOWI	6	0.569363	0.45192	0.004498424	1	5.49956939	126	9.659168	1 5.499569
9	disease1	m8	UNKNOWI	7	0.013142	0.908913	1.04E-04	1	0.127503034	126	9.701803	1 0.127503
10	disease1	m9	UNKNOWI	8	0.180746	0.671458	0.001432434	1	1.751228903	126	9.688916	1 1.751229
11	disease1	m10	UNKNOWI	9	0.013142	0.908913	1.04E-04	1	0.127503034	126	9.701803	1 0.127503
12	disease1	m11	UNKNOWI	10	0.013142	0.908913	1.04E-04	1	0.127503034	126	9.701803	1 0.127503
13	disease1	m12	UNKNOWI	11	0.013142	0.908913	1.04E-04	1	0.127503034	126	9.701803	1 0.127503
14	disease1	m13	UNKNOWI	12	0.138094	0.710808	0.001094782	1	1.338431342	126	9.692193	1 1.338431
15	disease1	m14	UNKNOWI	13	0.083581	0.772977	6.63E-04	1	0.810434626	126	9.696383	1 0.810435
16	disease1	m15	UNKNOWI	14	1.420603	0.235544	0.011148931	1	13.6301777	126	9.594639	1 13.63018
17	disease1	m16	UNKNOWI	15	0.247648	0.619604	0.001961602	1	2.398165761	126	9.683782	1 2.398166
18	disease1	m17	UNKNOWI	16	0.016253	0.898759	1.29E-04	1	0.157677533	126	9.701564	1 0.157678
19	disease1	m18	UNKNOWI	17	0.89331	0.346391	0.007039849	1	8.606600068	126	9.634509	1 8.6066
20	disease1	m19	UNKNOWI	18	2.288584	0.132833	0.017839357	1	21.80958946	126	9.529723	1 21.80959
21	disease1	m20	UNKNOWI	19	3.658944	0.058038	0.02821976	1	34.50020032	126	9.429004	1 34.5002
22	disease1	m21	UNKNOWI	20	0.455404	0.501016	0.003601302	1	4.402788766	126	9.667872	1 4.402789

Naive PC4 Kinship PC4+kin +

Search for anything 457 PM 11/26/2020

- **Copy results to Excel - All 4 Models**
- **Highlighted columns - remove**

- Kinship model
- Remove highlighted columns

I6

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Trait	Marker	Chr	Pos	df	F	P	add_effec	add_F	add_p	dom_effec	dom_F	dom_p	errordf	MarkerR2	Genetic V	Residual V	-2LnLikelihood
2	disease1	None			0	0.401651	0.527387	NaN	NaN	NaN	NaN	NaN	NaN	128	NaN	2.302313	8.766619	647.4115
3	disease1	m1	UNKNOWNI	0	1	2.346873	0.128042	NaN	NaN	NaN	NaN	NaN	NaN	128	0.003162	2.302313	8.766619	647.4115
4	disease1	m2	UNKNOWNI	1	1	1.893015	0.171303	NaN	NaN	NaN	NaN	NaN	NaN	128	0.014905	2.302313	8.766619	647.4115
5	disease1	m3	UNKNOWNI	2	1	2.105314	0.149273	NaN	NaN	NaN	NaN	NaN	NaN	128	3.87E-04	2.302313	8.766619	647.4115
6	disease1	m4	UNKNOWNI	3	1	0.00161	0.968056	NaN	NaN	NaN	NaN	NaN	NaN	128	0.016576	2.302313	8.766619	647.4115
7	disease1	m5	UNKNOWNI	4	1	0.927875	0.337261	NaN	NaN	NaN	NaN	NaN	NaN	128	0.007306	2.302313	8.766619	647.4115
8	disease1	m6	UNKNOWNI	5	1	0.730716	0.394274	NaN	NaN	NaN	NaN	NaN	NaN	128	0.005753	2.302313	8.766619	647.4115
9	disease1	m7	UNKNOWNI	6	1	0.00161	0.968056	NaN	NaN	NaN	NaN	NaN	NaN	128	1.27E-05	2.302313	8.766619	647.4115
10	disease1	m8	UNKNOWNI	7	1	1.177557	0.279926	NaN	NaN	NaN	NaN	NaN	NaN	128	0.009272	2.302313	8.766619	647.4115
11	disease1	m9	UNKNOWNI	8	1	1.192198	0.661844	NaN	NaN	NaN	NaN	NaN	NaN	128	0.001513	2.302313	8.766619	647.4115
12	disease1	m10	UNKNOWNI	9	1	1.192198	0.661844	NaN	NaN	NaN	NaN	NaN	NaN	128	1.27E-05	2.302313	8.766619	647.4115
13	disease1	m11	UNKNOWNI	10	1	0.00161	0.968056	NaN	NaN	NaN	NaN	NaN	NaN	128	0.00161	2.302313	8.766619	647.4115
14	disease1	m12	UNKNOWNI	11	1	1.177557	0.279926	NaN	NaN	NaN	NaN	NaN	NaN	128	1.27E-05	2.302313	8.766619	647.4115
15	disease1	m13	UNKNOWNI	12	1	0.020529	0.886299	NaN	NaN	NaN	NaN	NaN	NaN	128	1.62E-04	2.302313	8.766619	647.4115
16	disease1	m14	UNKNOWNI	13	1	1.03E-05	0.997441	NaN	NaN	NaN	NaN	NaN	NaN	128	8.13E-08	2.302313	8.766619	647.4115
17	disease1	m15	UNKNOWNI	14	1	1.120269	0.729321	NaN	NaN	NaN	NaN	NaN	NaN	128	0.00115	2.302313	8.766619	647.4115
18	disease1	m16	UNKNOWNI	15	1	1.120269	0.729321	NaN	NaN	NaN	NaN	NaN	NaN	128	8.20E-04	2.302313	8.766619	647.4115
19	disease1	m17	UNKNOWNI	16	1	1.120269	0.729321	NaN	NaN	NaN	NaN	NaN	NaN	128	9.47E-04	2.302313	8.766619	647.4115
20	disease1	m18	UNKNOWNI	17	1	1.177557	0.279926	NaN	NaN	NaN	NaN	NaN	NaN	128	0.009272	2.302313	8.766619	647.4115
21	disease1	m19	UNKNOWNI	18	1	1.192198	0.661844	NaN	NaN	NaN	NaN	NaN	NaN	128	0.015018	2.302313	8.766619	647.4115
22	disease1	m20	UNKNOWNI	19	1	2.863639	0.1267	NaN	NaN	NaN	NaN	NaN	NaN	128	0.01861	2.302313	8.766619	647.4115

WPS Office Results_edit_MSD.xlsx +

Home Insert Page Layout Formulas Data Review View Tools Q Cli... A1 1 sujan ... Go Premium

PivotTable AutoFilter Advanced Sort Show All Highlight Duplicates Manage Duplicates Reject Duplicate Entries Text to Columns Fill Validation Insert Drop-Down List Consolidate Form Group Ungroup S

A2 fx disease1

	A	B	C	D
1	Trait	Marker	p	marker_Rsq
2	disease1	m1	0.944324483	3.89E-05
3	disease1	m2	0.140228878	0.017184623
4	disease1	m3	0.167919961	0.015036333
5	disease1	m4	0.871886815	2.07E-04
6	disease1	m5	0.110699527	0.020068162
7	disease1	m6	0.183555607	0.013993411
8	disease1	m7	0.451920269	0.004498424
9	disease1	m8	0.908913404	1.04E-04
10	disease1	m9	0.671458455	0.001432434
11	disease1	m10	0.908913404	1.04E-04
12	disease1	m11	0.908913404	1.04E-04
13	disease1	m12	0.908913404	1.04E-04
14	disease1	m13	0.710807641	0.001094782
15	disease1	m14	0.772976915	6.63E-04
16	disease1	m15	0.235544116	0.011148931
17	disease1	m16	0.619603712	0.001961602
18	disease1	m17	0.898758632	1.29E-04
19	disease1	m18	0.346390915	0.007039849
20	disease1	m19	0.132832853	0.017839357
21	disease1	m20	0.058038271	0.02821976
22	disease1	m21	0.501016147	0.003601302

Naive PC4 Kinship PC4+kin

Sum=6388.934162 Average=0.23318129 Count=55520

Search for anything 5:01 PM 11/26/2020

The screenshot shows a WPS Office spreadsheet titled "Results_edit_MSD.xlsx". The active sheet is labeled "disease1". The data consists of four columns: Trait, Marker, p, and marker_Rsq. A red box highlights the "Sort" dialog box, which is open over the data range A1:D22. The dialog box shows two sort levels: first by "Trait" (values) in "A to Z" order, and then by "p" (values) in "A to Z" order. The status bar at the bottom indicates a sum of 6388.934162, an average of 0.23318129, and a count of 55520.

- Sort by Trait and then “p”

WPS Office Results_edit_MSD.xlsx

Home Insert Page Layout Formulas Data Review View Tools

Cut Copy Format Painter Calibri 11 Merge and Center Wrap Text General Conditional Formatting AutoSum AutoFilter Sort Format Fill Rows and Columns

A15 fx

Trait	Marker	p	marker_Rsq	Rank	expected p-value	Difference	Diff_Sq
disease1	m1847	4.50E-05	0.124218912	1	0.000144092	-9.91E-05	9.8E-09
disease1	m3467	2.12E-04	0.103535133	2	0.000288184	-7.57E-05	5.7E-09
disease1	m3659	2.12E-04	0.103535133	3	0.000432277	-0.00022	4.8E-08
disease1	m6253	3.54E-04	0.096678917	4	0.000576369	-0.000222	4.9E-08
disease1	m2333	7.17E-04	0.087149712	5	0.000720461	-3.19E-06	1E-11
disease1	m2798	7.17E-04	0.087149712	6	0.000864553	-0.000147	2.2E-08
disease1	m5150	7.17E-04	0.087149712	7	0.001008646	-0.000291	8.5E-08
disease1	m6501	7.17E-04	0.087149712	8	0.001152738	-0.000435	1.9E-07
disease1	m6832	7.68E-04	0.086219246	9	0.00129683	-0.000528	2.8E-07
disease1	m540	8.02E-04	0.085647118	10	0.001440922	-0.000639	4.1E-07
disease1	m4528	8.77E-04	0.084435219	11	0.001585014	-0.000708	5E-07
disease1	m539	9.96E-04	0.082709716	12	0.001729107	-0.000733	5.4E-07
disease1	m4083	0.001005107	0.082586619	13	0.001873199	-0.000868	7.5E-07
disease1	m4166	0.001159695	0.080650019	14	0.002017291	-0.000858	7.4E-07
disease1	m2981	0.001193477	0.080261216	15	0.002161383	-0.000968	9.4E-07
disease1	m1932	0.001321034	0.07888614	16	0.002305476	-0.000984	9.7E-07
disease1	m1871	0.001415219	0.077953218	17	0.002449568	-0.001034	1.1E-06
disease1	m3164	0.001462956	0.07750316	18	0.00259366	-0.001131	1.3E-06
disease1	m6604	0.001462956	0.07750316	19	0.002737752	-0.001275	1.6E-06
disease1	m5956	0.00158614	0.076408535	20	0.002881844	-0.001296	1.7E-06
disease1	m4024	0.001634159	0.076004414	21	0.003025937	-0.001392	1.9E-06

Naive PC4 Kinship PC4+kin

Search for anything 5:03 PM 11/26/2020

- Make 4 columns
- Rank - 1.2.3...n
- expected p-value = Rank/# of markers
- Difference = p- expected pvalue
- Diff_sq = Difference²

WPS Office | Results_edit_MSD.xlsx | +

Home Insert Page Layout Formulas Data Review View Tools Q Cli... sujan ... Go Premium

PivotTable AutoFilter Advanced Show All Sort Highlight Duplicates Manage Duplicates Reject Duplicate Entries Text to Columns Fill Validation Insert Drop-Down List Consolidate Form Group Ungroup S...

J1

Trait	Marker	p	marker_Rsq	Rank	expected p-value	Difference	Diff_Sq	MSD
disease1	m1847	4.50E-05	0.124218942	1	0.000144092	-9.91E-05	9.8E-09	
disease1	m3467	2.12E-04	0.103535183	2	0.000288184	-7.57E-05	5.7E-09	
disease1	m3659	2.12E-04	0.103535183	3	0.000432277	-0.00022	4.8E-08	
disease1	m6253	3.54E-04	0.096678957	4	0.000576369	-0.00022	4.9E-08	
disease1	m2333	7.17E-04	0.087149722	5	0.000720461	-3.19E-06	1E-11	
disease1	m2798	7.17E-04	0.087149722	6	0.000864553	-0.000147	2.2E-08	
disease1	m5150	7.17E-04	0.087149722	7	0.001008646	-0.000291	8.5E-08	
disease1	m6501	7.17E-04	0.087149722	8	0.001152738	-0.000435	1.9E-07	
disease1	m6832	7.68E-04	0.086219256	9	0.00129683	-0.000528	2.8E-07	
disease1	m540	8.02E-04	0.085647188	10	0.001440922	-0.000639	4.1E-07	
disease1	m4528	8.77E-04	0.084435249	11	0.001585014	-0.000708	5E-07	
disease1	m539	9.96E-04	0.082709746	12	0.001729107	-0.000733	5.4E-07	
disease1	m4083	0.001005107	0.082586629	13	0.001873199	-0.000868	7.5E-07	
disease1	m4166	0.001159695	0.080650079	14	0.002017291	-0.000858	7.4E-07	
disease1	m2981	0.001193477	0.080261296	15	0.002161383	-0.000968	9.4E-07	
disease1	m1932	0.001321034	0.078886144	16	0.002305476	-0.000984	9.7E-07	
disease1	m1871	0.001415219	0.077953278	17	0.002449568	-0.001034	1.1E-06	
disease1	m3164	0.001462956	0.07750386	18	0.00259366	-0.001131	1.3E-06	
disease1	m6604	0.001462956	0.07750386	19	0.002737752	-0.001275	1.6E-06	
disease1	m5956	0.00158614	0.076408535	20	0.002881844	-0.001296	1.7E-06	
disease1	m4024	0.001634159	0.076004414	21	0.003025937	-0.001392	1.9E-06	

Naive PC4 Kinship PC4+kin +

Search for anything

5:11 PM 11/26/2020

• **MSD - Calculate mean of Diff_sq - for each trait**

WPS Office | Results_edit_MSD.xlsx | sujan ... | Go Premium | X

Home Insert Page Layout Formulas Data Review View Tools Cli... +

Paste Cut Copy Format Painter Calibri 11 A A Merge and Center Wrap Text General Conditional Formatting AutoSum AutoFilter Sort Format Fill Rows and Columns

E16 fx

A B C D E F G H I J K L M N O P Q

2 trait	model	MSD
3 Disease1	kinship	0.000201566
4 Disease1	PC4+kin	0.000294704
5 Disease1	PC4	0.000312902
6 Disease1	Naive	0.004254349
7 Disease2	Naive	0.000101377
8 Disease2	kinship	0.000102634
9 Disease2	PC4+kin	0.000112044
10 Disease2	PC4	0.001129749
11		
12 p-value cutoffs	Bonferroni	7.39973E-06 Highly conservative 1.00E-03 Hypothetical cutoff
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		

Naive PC4 Kinship PC4+kin MSD +

Search for anything 5:16 PM 11/26/2020

- Sort by Trait and MSD
- decide on a p-value
 - Bonferroni
 - FDR/pFDR
 - cutoff value
 - Bootstraps

- Open Effect file to get Marker Effects

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.67

File Data Input Filter Analysis Results GBSv2 GBS Workflow PHG Help

Data

- Matrix
 - Day4-AM_kinship
 - Day4-AM_4PC
 - Day4-AM_phenotype_2Traits
- Numerical
- Sequence
- Day4-AM_genotype

Result

Number of taxa: 136
Number of sites: 6940

Chromosomes...

UNKNOWN: 6940 sites:
0 (0) - 6939 (6939)

Linkage Disequilibrium

Select LD type: Sliding Window

LD Window Size: 50
Sliding Window LD with 345725 comparisons.

How to treat heterozygous calls: Set to missing

Accumulate R2 Results

Number Of Intervals: 100

Run Close

AM7 1.029 -0.074 -0.390 1.029
AM11 0.921 7.124 8.271 -4.827
AM12 -8.827 -0.484 2.030 -1.911
AM14 -10.49 -2.989 -12.14 8.621
AM17 14.926 1.301 2.896 7.065
AM27 -2.926 -10.842 11.896 -9.034
AM30 10.653 -9.858 0.262 -7.843
AM32 4.154 -9.361 0.363 -10.089

• Calculate LD

• Phytozome - Jbrowse
for species

Phytozome v12.1: Info

JGI Phytozome 12 THE PLANT GENOMICS RESOURCE

Species Tools Info Download Help Cart Subscribe

Previous view Help with this page

Actions

- Revise query
- Launch Jalview
- Find related ...
- Add to cart
- Composite family

My Data (0)

- View cart
- Add to cart
- Upload user data
- Send to BioMart
- Send to PhytoMine

Genome Info Page at Phytozome v13
Hordeum vulgare at Phytozome v13

Overview

Cultivated barley, *Hordeum vulgare* ssp. *vulgare* L., is globally among the four most important cereal crops. It has a 5 Gbp genome and a haploid chromosome number of 7. Barley belongs to the tribe Triticeae, member of the family Poaceae and is closely related to other cereal crop species bread wheat, durum wheat, emmer wheat, einkorn wheat, spelt wheat, Triticale and rye. Most of barley is produced as animal feed (seeds or as forage crop) and a minor part is used for food including malt for beer brewing. Barley is globally grown and well adapted to a wide variety of climatic conditions and geographical latitudes and altitudes.

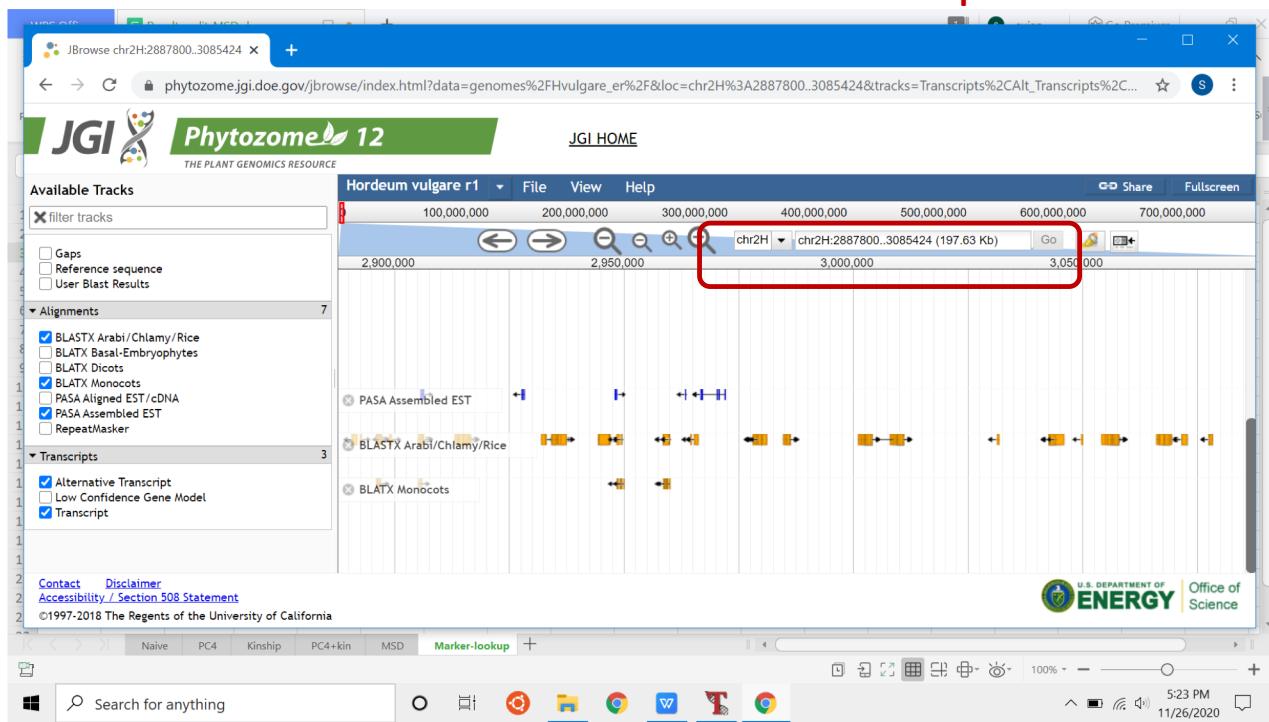
Marker-lookup

Search for anything

Naive PC4 Kinship PC4+kin MSD Marker-lookup +

5:20 PM 11/26/2020

- 100kb on either side of marker position



Gene description

The screenshot shows the Phytozome 12 genome browser interface for the *Hordeum vulgare* r1 genome. The main panel displays a genomic track for chromosome 2H, with a scale from 100,000,000 to 2,950,000. Several tracks are visible, including BLASTX results for Arabi/Chlamy/Rice, PASA Assembled EST, and BLATX Monocots. A detailed gene description is shown in a modal window for the gene **Spipo0G0020800.1**. The gene has a **Score** of 20.49 and is described as a **B3 domain-containing protein**. It is located at position **chr2H:2949436..2949944 (- strand)** and is 509 bp long. The gene is annotated with various attributes such as Alias (Spipo0G0020800), Description (B3 domain-containing protein Spipo0G0020800 Spipo0G0020800), Id (1017187.121177), Percentcover (43.78), Percentid (60.49), Score2 (26.49), Seq_id (chr2H), Source (Spolyrhiza_PAC2_0_290_peptide), Speciesab (Spipo), and Region sequence (available in a scrollable text area). The entire "Description" field is highlighted with a red box.

match Spipo0G0020800.1

Primary Data

Name	Spipo0G0020800.1
Type	match
Score	20.49
Description	B3 domain-containing protein Spipo0G0020800 Spipo0G0020800
Position	chr2H:2949436..2949944 (- strand)
Length	509 bp

Attributes

Alias	Spipo0G0020800
Description	B3 domain-containing protein Spipo0G0020800 Spipo0G0020800
Id	1017187.121177
Percentcover	43.78
Percentid	60.49
Score2	26.49
Seq_id	chr2H
Source	Spolyrhiza_PAC2_0_290_peptide
Speciesab	Spipo
Region sequence	(scrollable text area)