

TASSEL 3.0 / 4.0

www.maizegenetics.net/tassel

Terry Casstevens¹, Peter Bradbury^{2,3}, Zhiwu Zhang¹,
Yang Zhang¹, Edward Buckler^{1,2,4}

¹ *Institute for Genomic Diversity, Cornell University, Ithaca, NY*

² *USDA-ARS*

³ *Cornell Theory Center, Cornell University*

⁴ *Dept. of Plant Breeding and Genetics, Cornell University*

TASSEL

- **Tools for phenotype to genotype analysis**
- **Specialty is association mapping of structured populations**

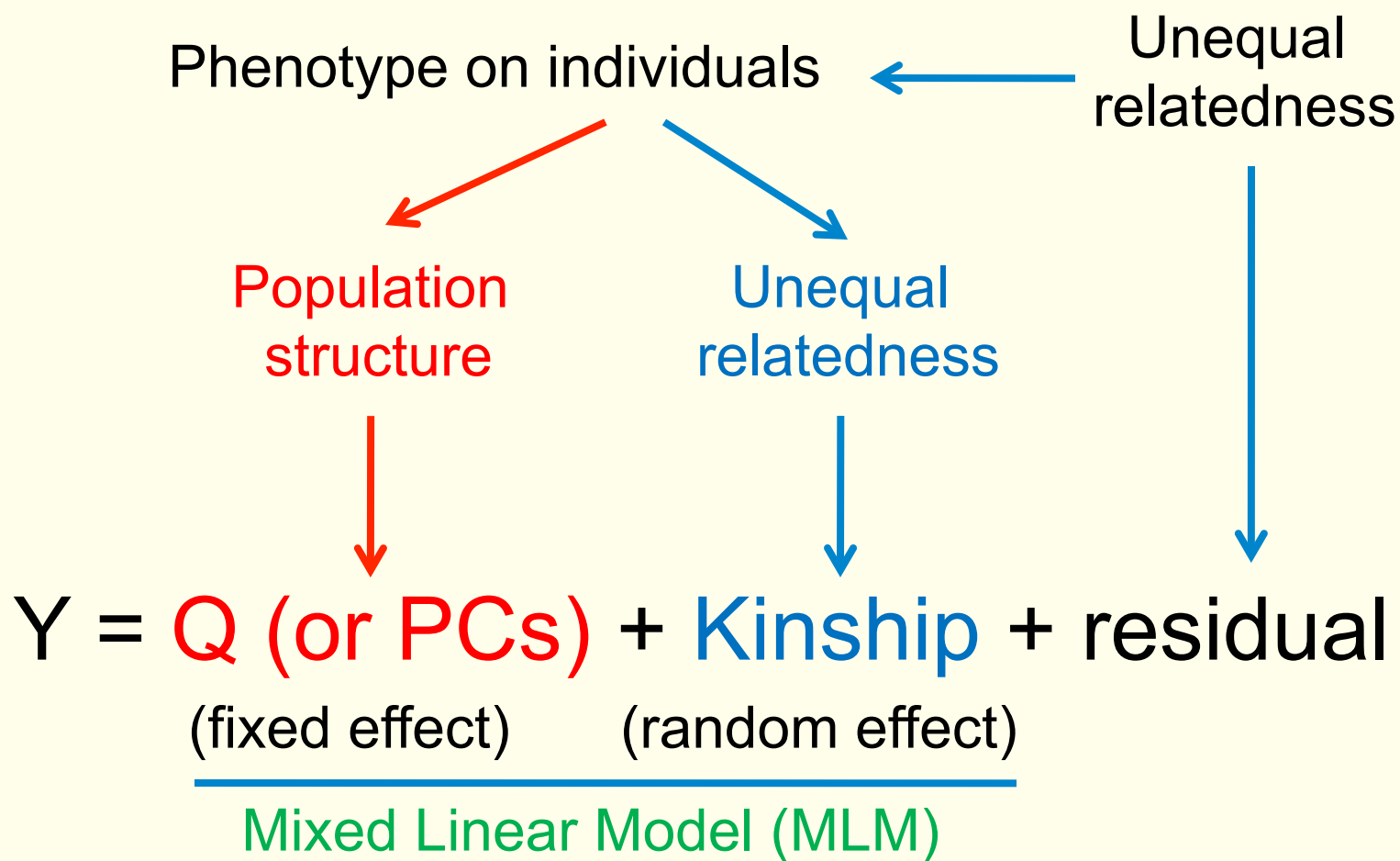
TASSEL : Tools for Genetic Research

- **Association Analysis (General Linear Model; GLM and Mixed Linear Model; MLM)**
- **Genomic Selection (Ridge Regression)**
- **Linkage Disequilibrium Analysis**
- **Missing Data Imputation (genotype and phenotype)**
- **SNP Extraction, filtering, numericalization, formatting (Hapmap, Plink, Flapjack, and Phylip)**
- **Diversity Analysis**
- **Kinship and Principal Component Analysis**

What's New in TASSEL?

	2.1	3.0
Marker Number	1000s	1,000,000s
PCA	✓	✓
TASSEL can now handle millions of SNPs and it 1000X faster for key association analyses.		
MLM	✓	Faster & More Data
EMMA	✓	✓
3PD & Compression		✓
Pipeline	✓	Many Improvements

GLM/MLM for GWAS



New algorithms for MLM

- **EMMA**: Convert optimization on two dimensions (genetic and residual variance components) to one dimension (their ratio), faster. By Kang et al (2007, Genetics)
- **Compression**: To group individuals into group to reduce size of MLM equations. Better speed and better power. By Zhang et al (2010, Nature Genetics)
- **P3D/EMMAx**: Population parameters (such as variance components) optimized only once and fixed in screening SNPs, Faster. By Zhang et al (2010, Nature Genetics, named as P3D) and Kang et al (2010, Nature Genetics, named EMMAx).

Demonstration

- **How to start?**
- **TASSEL Graphic User Interface (GUI)**
- **Data formats**
- **GLM as example**

www.maizegenetics.net/tassel

Buckler Lab for Maize Genetics and Diversity
A USDA-ARS Lab with Cornell Institute for Genomic Diversity

Main Menu
[Home](#)
[People](#)
[Research](#)
[Bioinformatics](#)
[Publications](#)
[Germplasm](#)
[Jobs](#)
[About Us](#)
[Links](#)
[Contact Us](#)
[Site Map](#)
[Repository](#)

Search Site

TASSEL

Tassel Version 3.0 (Build: January 12, 2012 Requires: Java 1.6)
[Launch TASSEL 3.0](#)
[Launch TASSEL 3.0 \(950Mb Heap Size\)](#)
[Launch TASSEL 3.0 \(512Mb Heap Size\)](#)

Tassel Version 3.0 Standalone (Build: January 12, 2012 Requires: Java 1.6)
[TASSEL 3.0 Standalone](#)

Tassel Version 4.0 (Build: January 12, 2012 Requires: Java 1.6)
[Launch TASSEL 4.0](#)
[TASSEL 4.0 Standalone](#)

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.59

File Tools Help GUPC

Analysis Results Delete Wizard Show Memory

GUPC Load Export Sites Taxa Traits Impute SNPs ?+5 Transform p+q Synonymizer Join Join

Data Result

Phenotype

<Trait>	EarHT	dpoll	EarDia
811	59.5	-999	-999
33-16	64.75	64.5	-999
38-11	92.25	68.5	37.897
4226	65.5	59.5	32.71933
4722	81.1	70.5	35.221
A188	27.5	62	31.419
A214N	65	69	32.006
A239	47.88	61	36.064
A272	35.63	70	-999
A441-5	53.5	67.5	35.008
A554	38.5	66	33.41775
A556	28	65	31.929
A6	109.5	70.5	31.5175
A619	36	61	40.63
A632	60	61	35.953
A634	54	59	35.601
A635	37	64	35.2005

Population structure

<Covariate>	Q1	Q2	Q3
33-16	0.014	0.972	0.014
38-11	0.003	0.993	0.004
4226	0.007	0.11	0.012
4722	0.035	0.854	0.111
A188	0.013	0.982	0.005
A214N	0.072	0.017	0.221
A239	0.035	0.963	0.002
A272	0.000	0.000	0.000
A441-5	0.000	0.000	0.000
A554	0.000	0.000	0.000
A556	0.000	0.000	0.000
A6	0.000	0.000	0.000
A619	0.000	0.000	0.000
A632	0.000	0.000	0.000
A634	0.000	0.000	0.000
A635	0.000	0.000	0.000

Kinship

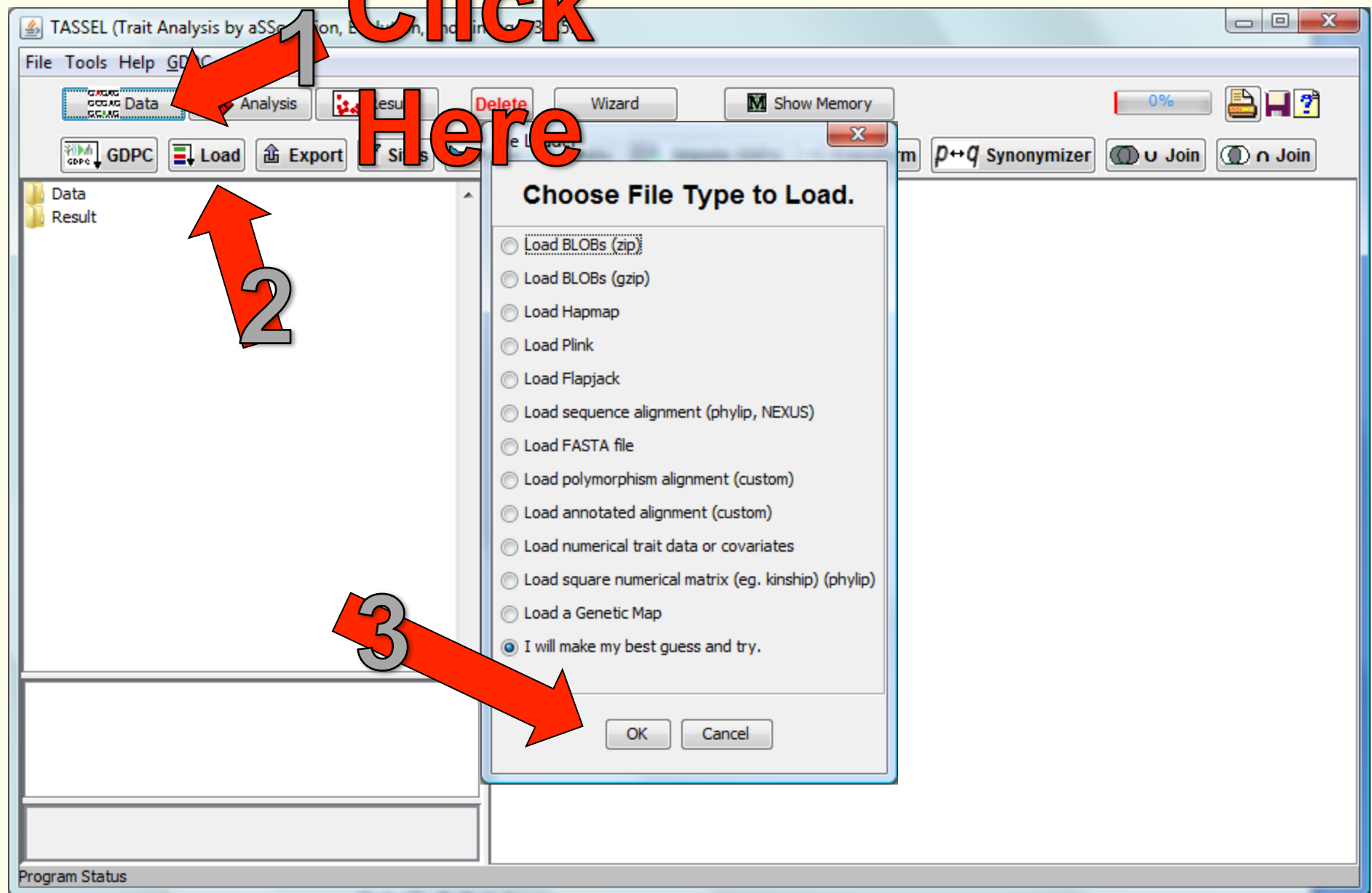
	33-16	38-11	4226	4722	A188
33-16	2.00	0.18	0.02	0.00	0.15
38-11	0.18	2.00	0.00	0.11	0.00
4226	0.02	0.00	2.00	0.00	0.00
4722	0.00	0.11	0.00	2.00	0.00
A188	0.15	0.00	0.00	0.00	2.00
A214N	0.00	0.00	0.00	0.00	0.00
A239	0.02	0.00	0.00	0.00	0.00
A272	0.00	0.00	0.00	0.00	0.00
A441-5	0.18	0.00	0.00	0.00	0.00
A554	0.17	0.03	0.06	0.05	0.00
A556	0.02	0.13	0.10	0.00	0.16
A6	0.10	0.00	0.00	0.04	0.00
A619	0.00	0.05	0.02	0.10	0.00
A632	0.00	0.08	0.00	0.00	0.07
A634	0.00	0.00	0.00	0.00	0.78
A635	0.00	0.00	0.00	0.00	0.00

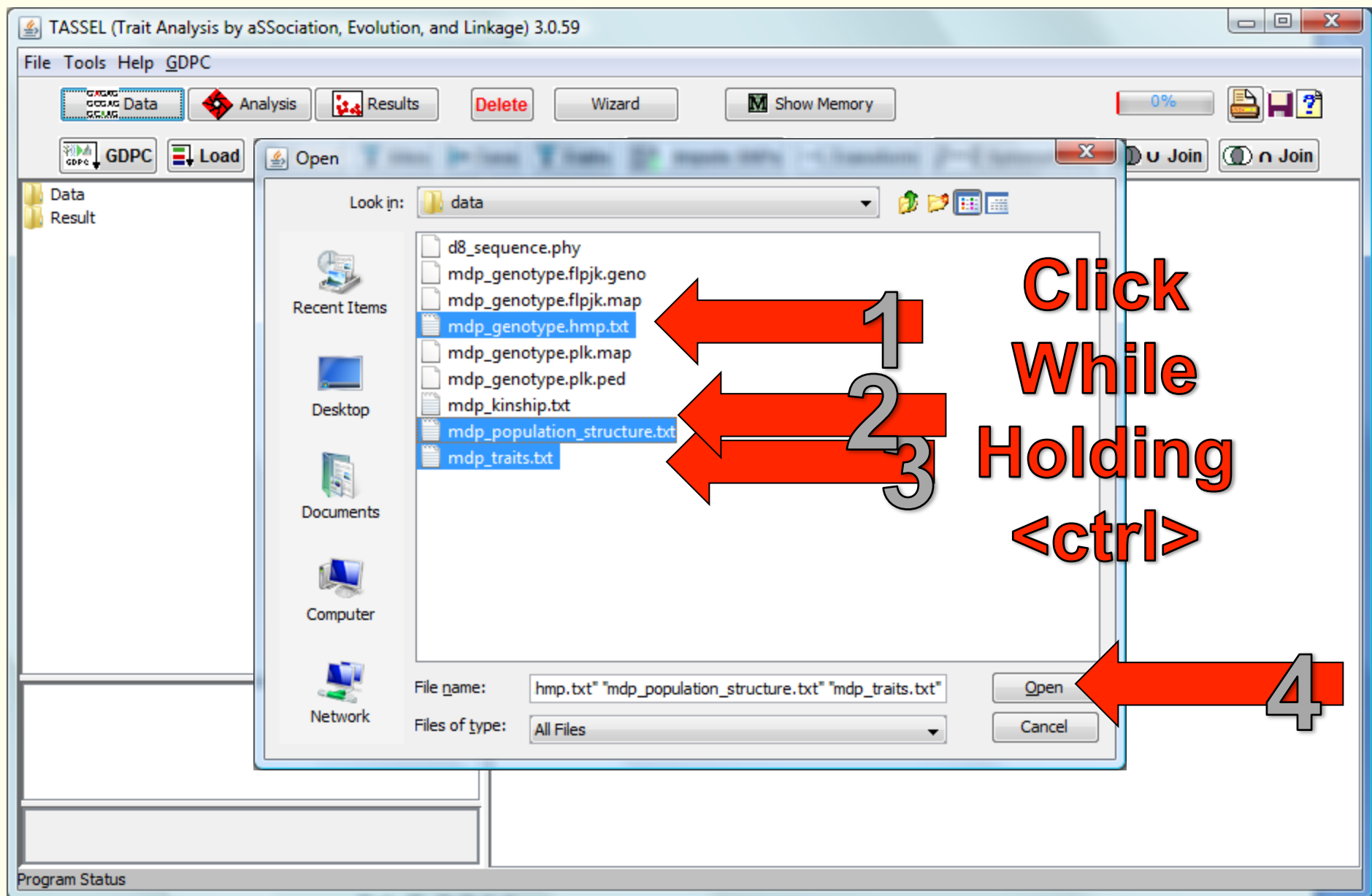
Association?

rs#	alleles	chrom	pos	strand	assembly#	center	protLSID	assayLSID	panel	QCcode	33-16	38-11	4226	4722	A188
PZB00859.1	A/C	1	157104	+	AGPv1	Panzea	NA	NA	maize282	NA	CC	CC	CC	CC	AA
PZA01271.1	C/G	1	194108	+	AGPv1	Panzea	NA	NA	maize282	NA	CC	GG	CC	GG	CC
PZA03613.2	G/T	1	291466	+	AGPv1	Panzea	NA	NA	maize282	NA	GG	GG	GG	GG	GG
PZA03614.1	A/T	1	2914171	+	AGPv1	Panzea	NA	NA	maize282	NA	TT	TT	TT	TT	TT
PZA03614.2	A/G	1	2915078	+	AGPv1	Panzea	NA	NA	maize282	NA	GG	GG	GG	GG	GG
PZA03614.1	A/T	1	2915242	+	AGPv1	Panzea	NA	NA	maize282	NA	TT	TT	TT	TT	TT
PZA00258.3	C/G	1	2973508	+	AGPv1	Panzea	NA	NA	maize282	NA	GG	CC	CC	CG	CC
PZA02962.13	A/T	1	3205252	+	AGPv1	Panzea	NA	NA	maize282	NA	TT	TT	TT	TT	TT
PZA02962.14	C/G	1	3205262	+	AGPv1	Panzea	NA	NA	maize282	NA	CC	CC	CC	CC	CC
PZA00599.25	C/T	1	3206090	+	AGPv1	Panzea	NA	NA	maize282	NA	CC	TT	CC	TT	TT

Genotype

Program Status





TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.59

File Tools Help GDPC

GDPC Analysis Results Delete Wizard Show Memory 0%

GDPC Load Export Sites Taxa Traits Impute SNPs ?+5 Transform p+q Synonymizer Join n Join

Data

- Sequence
 - mdp_genotype
- Polymorphisms
- Numerical
 - mdp_population_structure
 - mdp_traits
- Matrix
- Tree
- Fusions
- Synonymizer
- Result

Number of sequences: 201
Number of sites: 3093
Data type: IUPACNucleotide
Loci: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Site Numbers Locus Alleles (Enter physical position)

0 773 1546 2319 3092

	3072: 143676157	3073: 143676207	3074: 144025574	3075: 144172190	3076: 144172316	3077: 144239494	3078: 144239564	3079: 144548839	3080: 144548910	3081: 144549995	3082: 145850233	3083: 145850362	3084: 145850422	3085: 146999711	3086: 147762925	3087: 147762964	3088: 147819958	3089: 148332885	3090: 148488692	3091: 148816538	3092: 148907116
33-16	C	T	G	T	C	A	G	C	T	A	A	G	C	C	A	R	A	C	G	A	G
38-11	C	C	G	T	A	A	C	C	C	G	A	G	C	C	A	A	A	C	G	A	G
4226	C	C	G	T	A	A	C	C	C	G	A	T	G	C	G	G	A	G	G	A	G
4722	C	Y	N	T	N	G	N	C	C	A	A	T	G	C	G	G	A	C	G	A	G
A188	C	C	G	C	C	G	G	C	T	A	A	T	G	C	G	G	A	C	G	A	G
A214N	C	C	G	T	A	G	G	C	C	A	A	T	G	C	G	G	A	G	G	A	G
A239	C	T	G	T	C	A	G	C	C	A	A	G	G	C	A	N	G	C	G	A	G
A272	C	T	C	C	C	R	C	C	C	N	N	N	G	C	A	R	A	N	N	A	G
A441-5	C	C	C	T	C	G	G	C	C	A	A	T	G	C	A	A	G	C	G	A	G
A554	C	T	G	T	C	A	G	C	T	A	A	G	G	C	G	G	A	C	G	A	G
A556	C	T	G	T	C	A	G	C	C	A	N	T	G	C	G	G	A	C	C	A	G
A6	C	T	G	T	C	A	C	C	T	A	A	G	C	C	A	A	G	G	G	A	G
A619	C	T	G	T	C	A	C	C	C	A	A	T	G	C	G	G	A	C	C	A	G
A632	C	T	G	T	A	G	G	C	C	A	A	T	G	C	G	G	A	G	G	A	G
A634	C	T	G	T	A	G	G	C	C	A	A	T	G	C	G	G	A	G	G	A	G
A635	C	T	G	T	N	G	G	C	C	A	A	T	G	C	G	G	A	G	G	A	G

Program Status

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.50

File Tools Help GDCP

Data Analysis Results Wizard Show Memory 0%

GDPC Load Export Sites Taxa Traits Impute SNPs Transform Synonymizer Join

Data

- Sequence
 - mdp_genotype
- Polymorphisms
- Numerical
 - mdp_population
 - mdp_traits
- Matrix
- Tree
- Fusions
- Synonymizer
- Result

Filter Alignment

Filter Alignment

Minimum Count: 210 out of 281 sequences

Minimum Frequency: 0.05

Position Type: Position index

Start Position: 0

End Position: 3092 of 3092 sites

☐ Extract Indels

☐ Remove minor SNP states

☐ Generate haplotypes via sliding window

Haplotype Length

Step Length

Filter Select Chromosomes... Cancel

Program Status

Number of sites: 3092

Data type: IUPAC

Loci: 1, 2, 3, 4

Position	Sequence
3087:147762964	A C G A G
3088:147819958	A C G A G
3089:148332885	A G G A G
3090:148488692	A C G A G
3091:148816538	A G G A G
3092:148907116	A G G A G

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.59

File Tools Help GDPC

Data Analysis Results Delete Wizard Show Memory 0%

GDPC Load Export Sites Taxa Traits Impute SNPs ?+5 Transform p+q Synonymizer Join n Join

Data

- Sequence
 - mdp_genotype
 - mdp_genotype_chr1_2_3_4_5_6_7_8_157104- 
- Polymorphisms
- Numerical
 - mdp_population_structure
 - mdp_traits
- Matrix
- Tree
- Fusions
- Synonymizer
- Result

Site Numbers ☒ Locus ☐ Alleles (Enter physical position)

640 1280 1920 2560

2542: 143269455	2543: 143269563	2544: 143676007	2545: 143676207	2546: 144025574	2547: 144172190	2548: 144172316	2549: 144239494	2550: 144239564	2551: 144548839	2552: 144548910	2553: 144549995	2554: 145850362	2555: 145850422	2556: 146999711	2557: 147762925	2558: 147762964	2559: 147819958	2560: 148332885	2561: 148488692	2562: 148907116	
33-16	T	C	G	T	G	T	C	A	G	C	T	A	G	C	C	A	R	A	C	G	G
38-11	C	C	C	C	G	T	A	A	C	C	C	G	G	C	C	A	A	A	C	G	G
4226	C	C	C	C	G	T	A	A	C	C	C	G	T	G	C	G	G	A	G	G	G
4722	N	A	G	Y	N	T	N	G	N	C	C	A	T	G	C	G	G	A	C	G	G
A188	N	C	C	C	G	C	C	G	G	C	T	A	T	G	C	G	G	A	C	G	G
A214N	T	A	G	C	G	T	A	G	G	C	C	A	T	G	C	G	G	A	G	G	G
A239	T	A	G	T	G	T	C	A	G	C	C	A	G	G	C	A	N	G	C	G	G
A272	N	C	G	T	C	C	C	R	C	C	C	N	N	G	C	A	R	A	N	N	G
A441-5	C	C	C	C	C	T	C	G	G	C	C	A	T	G	C	A	A	G	C	G	G
A554	C	C	G	T	G	T	C	A	G	C	T	A	G	G	C	G	G	A	C	G	G
A556	N	A	G	T	G	T	C	A	G	C	C	A	T	G	C	G	G	A	C	C	G
A6	C	C	G	T	G	T	C	A	C	C	T	A	G	C	C	A	A	G	G	G	G
A619	T	C	G	T	G	T	C	A	C	C	C	A	T	G	C	G	G	A	C	C	G
A632	T	C	G	T	G	T	A	G	G	C	C	A	T	G	C	G	G	A	G	G	G
A634	T	C	G	T	G	T	A	G	G	C	C	A	T	G	C	G	G	A	G	G	G
A635	T	C	G	T	G	T	N	G	G	C	C	A	T	G	C	G	G	A	G	G	G

Site Filter
Base Type: CombineAlignment
Point Poly.

Program Status

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.59

File Tools Help GDPC

Data Analysis Results Delete Wizard Show Memory 0%

GDPC Load Export Sites Taxa Traits Impute SNPs ?+5 Transform p+q Synonymizer Join Join

Data

- Sequence
 - mdp_genotype
 - mdp_genotype_chr1_2_3_4_5_6_7_8_157104-1
- Polymorphisms
- Numerical
 - mdp_population_structure**
 - mdp_traits
- Matrix
- Tree
- Fusions
- Synonymizer
- Result

Taxa	Q1	Q2	Q3
33-16	1.4E-2	0.972	1.4E-2
38-11	3.0E-3	0.993	4.0E-3
4226	7.1E-2	0.917	1.2E-2
4722	3.5E-2	0.854	0.111
	1.3E-2	0.982	5.0E-3
	0.762	1.7E-2	0.221
A239	3.5E-2	0.963	2.0E-3
A272	1.9E-2	0.122	0.859
A441-5	5.0E-3	0.531	0.464
A554	1.9E-2	0.979	2.0E-3
A556	4.0E-3	0.994	2.0E-3
A6	3.0E-3	3.0E-2	0.967
A619	9.0E-3	0.99	1.0E-3
A632	0.993	4.0E-3	3.0E-3
A634	0.897	0.1	3.0E-3
A635	0.825	0.171	4.0E-3
A641	0.517	0.481	2.0E-3
A654	8.3E-2	0.915	2.0E-3
A659	6.0E-3	0.991	3.0E-3
A661	0.111	0.852	3.7E-2
A679	0.862	0.127	1.1E-2
A680	0.993	4.0E-3	3.0E-3
A682	2.0E-3	0.997	1.0E-3
AB28A	2.0E-3	0.776	0.222
B10	0.429	0.57	1.0E-3
B103	0.163	0.829	8.0E-3
B104	0.694	0.305	1.0E-3
B105	0.397	0.566	3.7E-2

Number of columns: 4
Number of rows: 286
Number of elements: 1144

Program Status



TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.59

File Tools Help GDCP

Data
 Analysis
 Results
 Delete
 Wizard
 Show Memory

GDCP
 Load
 Export
 Sites
 Taxa
 Traits
 Impute SNPs
 ?+5 Transform
 p+q Synonymizer
 u Join
 n Join

Data

- Sequence
 - mdp_genotype
 - mdp_genotype_chr1_2_3_4_5_6_7_8_157104-1
- Polymorphisms
- Numerical
 - mdp_population_structure
 - mdp_traits
 - Filtered_mdp_population_structure**
- Matrix
- Tree
- Fusions
- Synonymizer
- Result

Taxa	Q1	Q2
33-16	1.4E-2	0.972
38-11	3.0E-3	0.993
4226	7.1E-2	0.917
4722	3.5E-2	0.854
A188	1.3E-2	0.982
A214N	0.762	1.7E-2
	-2	0.963
	-2	0.122
A441-5	5.0E-3	0.531
A554	1.9E-2	0.979
A556	4.0E-3	0.994
A6	3.0E-3	3.0E-2
A619	9.0E-3	0.99
A632	0.993	4.0E-3
A634	0.897	0.1
A635	0.825	0.171
A641	0.517	0.481
A654	8.3E-2	0.915
A659	6.0E-3	0.991
A661	0.111	0.852
A679	0.862	0.127
A680	0.993	4.0E-3
A682	2.0E-3	0.997
AB28A	2.0E-3	0.776
B10	0.429	0.57
B103	0.163	0.829
B104	0.694	0.305
B105	0.397	0.566

Number of columns: 3
 Number of rows: 286
 Number of elements: 858

Program Status

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.59

File Tools Help GDPC

GDPC Data Analysis Results Delete Wizard Show Memory 0%

GDPC Load Export Sites Taxa Traits Impute SNPs ?+5 Transform p+q Synonymizer Join Join

Data

- Sequence
 - mdp_genotype
 - mdp_genotype_chr1_2_3_4_5_6_7_8_157104
- Polymorphisms
- Numerical
 - mdp_population_structure
 - mdp_traits
 - Filtered_mdp_population_structure
- Matrix
- Tree
- Fusions
- Synonymizer
- Result

Site Numbers Locus Alleles (Enter physical position)

Click While Holding <ctrl>

1

2

3

4

	640	1280	1920	2560
33-16	T	A	G	C
38-11	C	C	C	C
4226	C	C	C	C
4722	N	A	G	Y
A188	N	C	C	C
A214N	T	A	G	C
A239	T	A	G	T
A272	N	C	G	T
A441-5	C	C	C	C
A554	C	C	G	T
A556	N	A	G	T
A6	C	C	G	T
A619	T	C	G	T
A632	T	C	G	T
A634	T	C	G	T
A635	T	C	G	T

Site Filter
Base Type: CombineAlignment
Point Poly.

Program Status

TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.59

File Tools Help GDPC

GDPC Analysis Results Delete Wizard Show Memory 0%

GDPC Load Export Sites Taxa Traits Impute SNPs ?+5 Transform p+q Synonymizer Join n Join

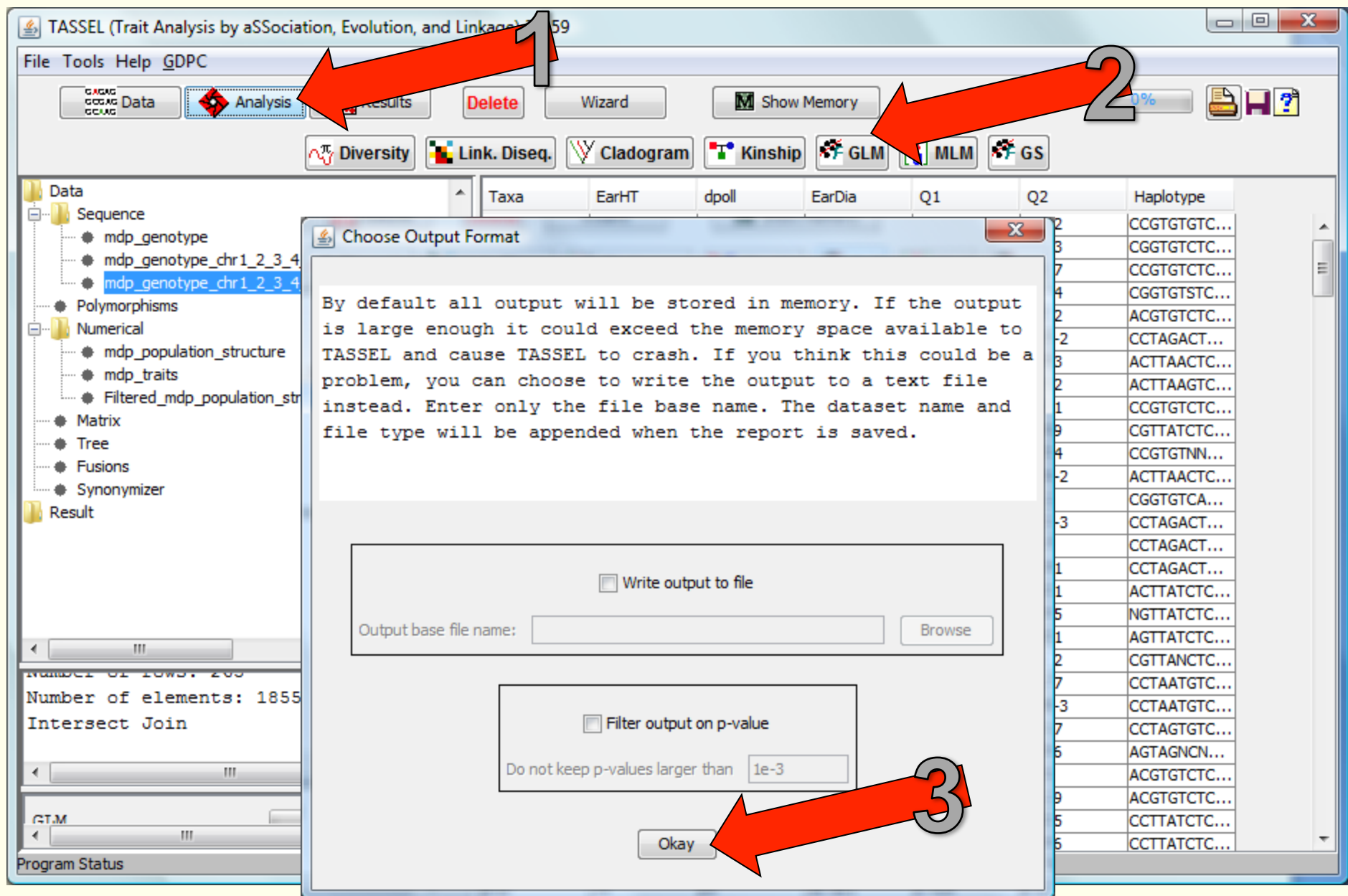
Data

- Sequence
 - mdp_genotype
 - mdp_genotype_chr1_2_3_4_5_6_7_8_15710
 - mdp_genotype_chr1_2_3_4_5_6_7_8_157
- Polymorphisms
- Numerical
 - mdp_population_structure
 - mdp_traits
 - Filtered_mdp_population_structure
- Matrix
- Tree
- Fusions
- Synonymizer
- Result

Taxa	EarHT	dpoll	EarDia	Q1	Q2	Haplotype
33-16	64.75	64.5	NaN	1.4E-2	0.972	CCGTGTGTC...
38-11	92.25	68.5	37.897	3.0E-3	0.993	CGGTGTCTC...
		59.5	32.21933	7.1E-2	0.917	CCGTGTCTC...
		71.5	32.421	3.5E-2	0.854	CGGTGTSTC...
A188	27.5	62	31.419	1.3E-2	0.982	ACGTGTCTC...
A214N	65	69	32.006	0.762	1.7E-2	CCTAGACT...
A239	47.88	61	36.064	3.5E-2	0.963	ACTTAAGTC...
A272	35.63	70	NaN	1.9E-2	0.122	ACTTAAGTC...
A441-5	53.5	67.5	35.008	5.0E-3	0.531	CCGTGTCTC...
A554	38.5	66	33.41775	1.9E-2	0.979	CGTTATCTC...
A556	28	65	31.929	4.0E-3	0.994	CCGTGTNN...
A6	109.5	80.5	31.5175	3.0E-3	3.0E-2	ACTTAAGTC...
A619	36	61	40.63	9.0E-3	0.99	CGGTGTCA...
A632	60	61	35.953	0.993	4.0E-3	CCTAGACT...
A634	54	59	35.601	0.897	0.1	CCTAGACT...
A635	37	64	35.3005	0.825	0.171	CCTAGACT...
A641	54.5	66	33.727	0.517	0.481	ACTTATCTC...
A654	39	64	NaN	8.3E-2	0.915	NGTTATCTC...
A659	46.5	58.5	38.84617	6.0E-3	0.991	AGTTATCTC...
A661	51.5	59	39.323	0.111	0.852	CGTTANCTC...
A679	65	66	42.471	0.862	0.127	CCTAATGTC...
A680	68	65.5	41.152	0.993	4.0E-3	CCTAATGTC...
A682	47	57.5	35.928	2.0E-3	0.997	CCTAGTCTC...
AB28A	73.5	78	32.504	2.0E-3	0.776	AGTAGNCN...
B10	74	69	36.561	0.429	0.57	ACGTGTCTC...
B103	37	57.5	NaN	0.163	0.829	ACGTGTCTC...
B104	56.25	64.5	44.773	0.694	0.305	CCTTATCTC...
B105	67.5	68	39.857	0.397	0.566	CCTTATCTC...

Number of rows: 200
Number of elements: 1855
Intersect Join

Program Status



TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 3.0.59

File Tools Help GDPC

Data Analysis Results Delete Wizard Show Memory 0%

Diversity Link. Diseq. Cladogram Kinship GLM MLM GS

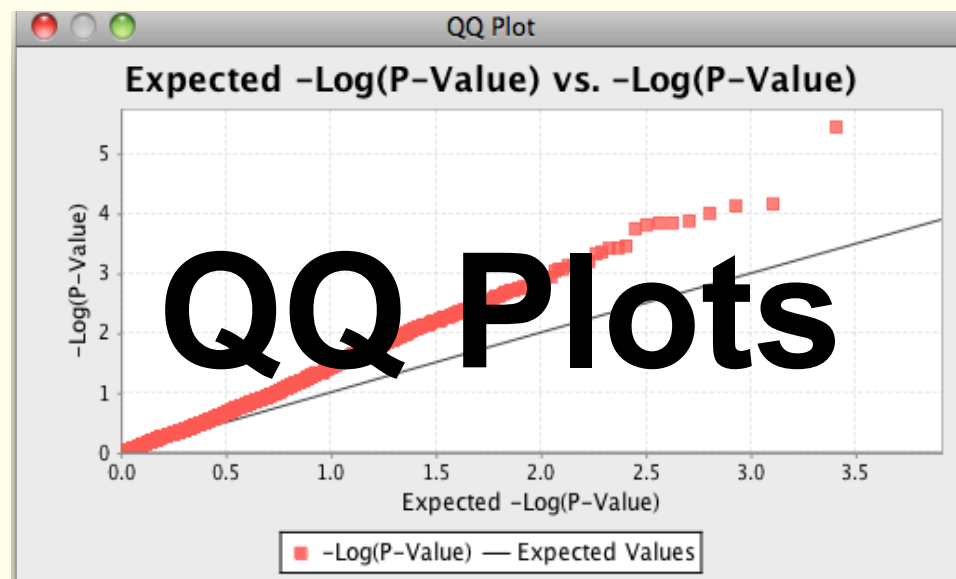
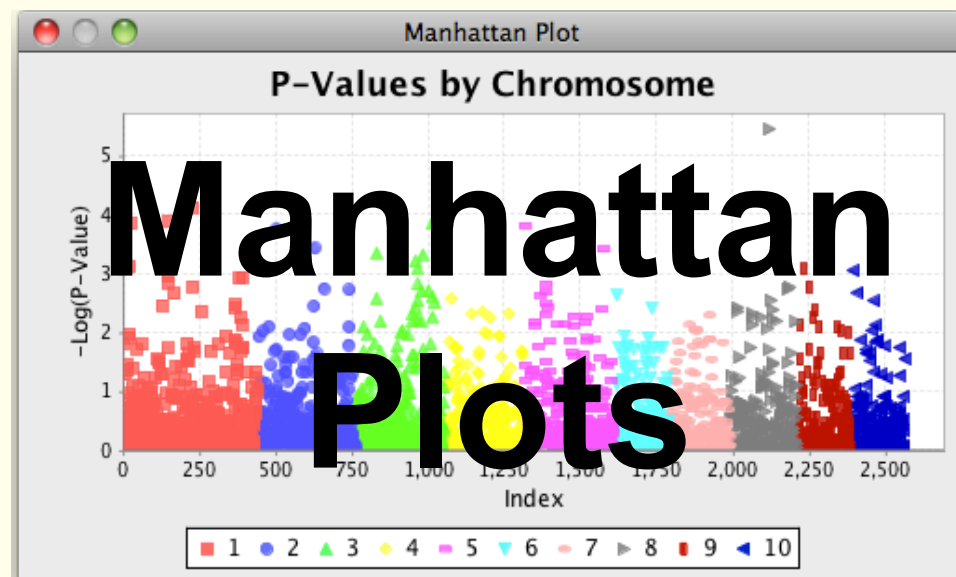
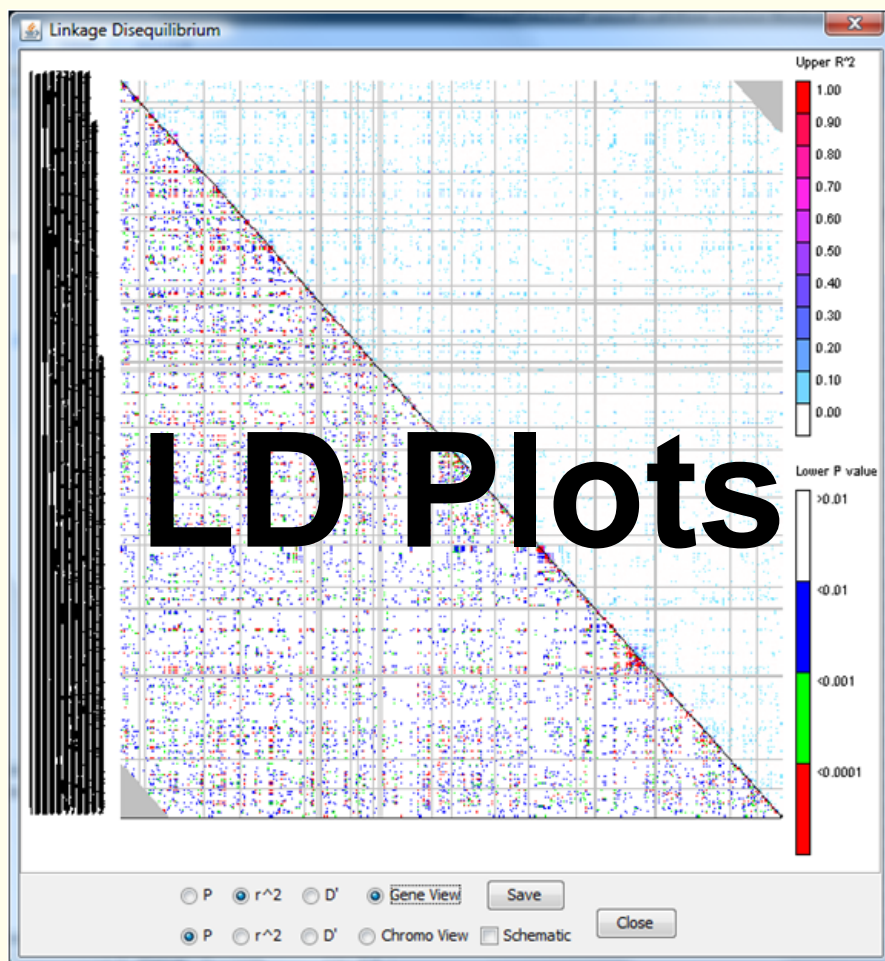
mdp_genotype_chr1_2_3_4_5_6_7_8_157104-14
 mdp_genotype_chr1_2_3_4_5_6_7_8_157104-14
 Polymorphisms
 Numerical
 mdp_population_structure
 mdp_traits
 Filtered_mdp_population_structure
 Matrix
 Tree
 Fusions
 Synonymizer
 Result
 Diversity
 SNP Assays
 LD
 Association
 GLM_marker_test_mdp_genotype_chr1_2_3_4_5_6_7_8_157104-14
 GLM_allele_estimates_for_mdp_genotype_chr1_2_3_4_5_6_7_8_157104-14
 Variances
 Stepwise

Trait	Marker	Locus	Locus_pos	marker_F	marker_p	markerR2	markerDF
dpoll	PZA03591.1	8	134813437	22.43615	3.5963E-6	5.3662E-2	1
EarDia	PZA00189.21	8	170014413	19.8951	1.2872E-5	7.8775E-2	1
EarHT	PZA02945.10	10	33619582	19.26773	1.6724E-5	5.9009E-2	1
EarHT	PZA03214.3	1	245136244	18.74896	2.1302E-5	5.6985E-2	1
EarHT	PZA00332.5	4	194749287	10.89055	2.8892E-5	6.5957E-2	2
EarHT	PZA00332.3	4	194749230	10.33673	4.8693E-5	6.3655E-2	2
EarHT	PZA03591.1	8	134813437	16.92435	5.2242E-5	5.1685E-2	1
dpoll	PZB00149.1	1	188047564	9.91778	7.0941E-5	4.7491E-2	2
dpoll	PZA03194.1	1	212244425	16.16438	7.6371E-5	3.9306E-2	1
EarHT	PZA00274.7	1	41505016	9.85381	7.6441E-5	6.2368E-2	2
EarHT	PZA00116.2	9	125052222	9.80008	7.9559E-5	6.0864E-2	2
EarHT	PZA03305.2	1	286642765	15.81122	9.1392E-5	5.0287E-2	1
EarHT	PZA00460.8	8	163565334	15.73719	9.6026E-5	5.2769E-2	1
dpoll	PZB0077.1	6	80877074	15.68435	9.7444E-5	3.8868E-2	1
		1	174183331	15.33716	1.1529E-4	4.8073E-2	1
		3	55946639	15.26376	1.2444E-4	6.3117E-2	1
dpoll	PZA01267.3	1	76054042	15.09123	1.3122E-4	3.7826E-2	1
EarHT	PZA00447.8	1	9024005	15.0015	1.3946E-4	5.0486E-2	1
dpoll	PZA03146.1	3	214272814	14.93327	1.4122E-4	3.669E-2	1
dpoll	PZA00447.8	1	9024005	14.97689	1.4164E-4	4.0607E-2	1
EarDia	PZA00793.2	8	64421988	14.95819	1.4478E-4	6.2288E-2	1
dpoll	PZB01352.6	5	3142800	14.83396	1.531E-4	4.0698E-2	1
dpoll	PZA02814.2	2	17268906	8.93843	1.7914E-4	4.5363E-2	2
EarDia	PZB02033.1	5	184588958	14.41939	1.8882E-4	5.9483E-2	1
EarHT	PZD00032.1	10	30830484	14.31457	1.9689E-4	4.9912E-2	1
EarHT	PZA03188.4	1	280719882	8.81266	1.9981E-4	5.5159E-2	2
EarDia	PZA00453.7	4	166281276	14.00703	2.3057E-4	5.6541E-2	1

ests of Marker-Phenotype AssociationG
 ata set: mdp_genotype_chr1_2_3_4_5_6_7_8_157104-14
 odel: trait = mean + Q1 + Q2 + marker

Program Status

Visualization Tools

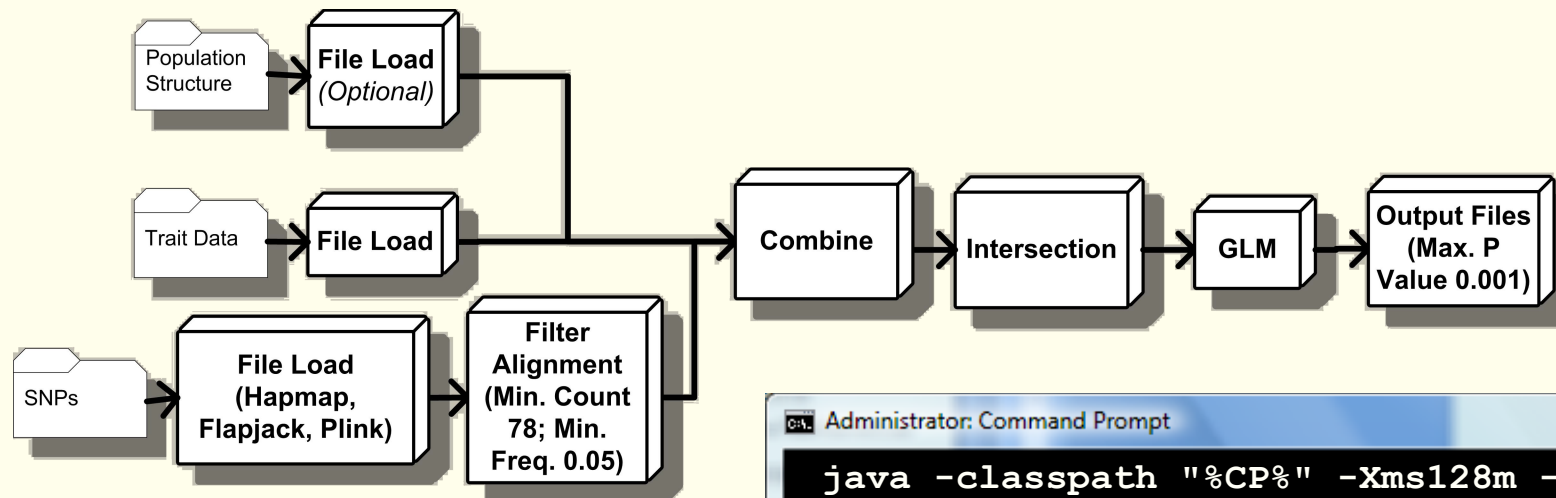


Tassel Pipeline

www.maizegenetics.net/tassel/docs/TasselPipelineCLI.pdf

- **Automates Complex Analyses.**
- **Don't need to write Java Code.**
- **Threaded (Pipeline segments run simultaneously).**
- **Works from web site Tassel launch.**
- **Works from Command Line Interface.**
- **Can produce same graphs as GUI.**
- **XML Configuration Files**

Example Pipeline: GLM Analysis



Evaluated
2.4 Billion
GLM Analyses in
14 CPU Hours!

```
Administrator: Command Prompt

java -classpath "%CP%" -Xms128m -Xmx1024m
net.maizegenetics.pipeline.TasselPipeline
-fork1 -h "mdp_genotype.hmp.txt"
-filterAlign -filterAlignMinCount 78
-filterAlignMinFreq 0.05 -fork2
-r "mdp_traits.txt" -fork3
-q "mdp_population_structure.txt"
-combine4 -input1 -input2 -input3
-intersect -glm -glmOutputFile glm_output
-glmMaxP 0.001 -runfork1 -runfork2
-runfork3
```


www.gramene.org/diversity/tassel_launch.html

TASSEL Web Launch

LAUNCH TASSEL with SNP data sets

Clicking on a link below will launch TASSEL 3.0-beta with one or all chromosomes. Your computer must have Java 6 installed.

-  **Maize**
 - **Maize Diversity** Hapmap1, AGPv2 coordinates, 1.6 million SNP x 27 NAM lines.[chr: [unmapped 1 2 3 4 5 6 7 8 9 10 11 12](#)]
-  **Arabidopsis**
 - NSF 2010 Project, Magnus Nordborg, <http://walnut.usc.edu> NSF Arabidopsis 2010 Project 250K chip, v3.06; 1307 ecotypes x 214,052 SNPs; TAIR10.[chr: [1 2 3 4 5](#)]
 - Richard Mott; WTCHG(Wellcome Trust Centre for Human Genetics) 1001 Genomes WTCHG/Mott set from dbSNP; 17 ecotypes x 2698797 SNPs; TAIR10.[chr: [1 2 3 4 5](#)]
 - Atwell S, et. al. Atwell et. al, Nature 2010, GWAS of 107 phenotypes in Arabidopsis thaliana inbred lines; TAIR10.[chr: [1 2 3 4 5 ALL](#)]
 - D. Weigel NSF Arabidopsis 2010 Project SNP Discovery; 20 ecotypes, 637522 SNPs; Clark et. al. 2007; TAIR10.[chr: [1 2 3 4 5](#)]
 - M. Nordborg NSF Arabidopsis 2010 Project 250K chip, v3.04; 1179 ecotypes x 214,545 SNPs; TAIR10.[chr: [1 2 3 4 5](#)]
-  **Rice**
 - Yulin Jia, USDA-ARS DBNRR B-Gross/J-Yulin SNP Pi-ta study in AA Oryza; 143 accessions x 632 SNP.[chr: [12](#)]
 - Rice HaplotypeMap Project, National Center for Gene Research, CAS; www.ncgr.ac.cn/RiceHapMap Huang X, et al, Nat Genetics 2010 Nov. 842,474 SNPs* x 374 landraces, 14 phenotypes. *IRGSP4 coordinates..[chr: [1 2 3 4 5 6 7 8 9 10 11 12](#)]
 - Zhao K et. al.; ricediversity.org '1536 chip' - 1311 SNPs x 395 rice accessions (Zhao et. al, PLoS, 2010; ricediversity.org).[chr: [1 2 3 4 5 6 7 8 9 10 11 12](#)]
 - McNally et. al.; oryzaSNP.org ~160K SNPs x 20 accessions (McNally 2009 PNAS/Oryzasnp Project).[chr: [1 2 3 4 5 6 7 8 9 10 11 12](#)]
 - Susan McCouch RiceDiversity 44K SNP assay, 34 trait GWAS study, Nat Comm Sept 2011 .[chr: [1 2 3 4 5 6 7 8 9 10 11 12](#)]

TASSEL 4.0 – It's All About the Bits!

TASSEL 4.0 introduces the next major improvements since TASSEL 3.0! Our new design puts emphasis on Efficiency, Performance, and Consistency. This is accomplished largely due to our bit-wise data storage and analysis algorithms.

- 1.6 Million LD Calculations Executes 30 times faster in TASSEL 4.0!
- Distance Matrix for 1320 Taxa Executes 25 Times faster in TASSEL 4.0!

Join the TASSEL Community

- ~3000 Users
- TASSEL Documentation, Tutorial Data Sets
<http://www.maizegenetics.net/tassel>
- Discussion Group:
<http://groups.google.com/group/tassel>
- Source Code:
<https://sourceforge.net/projects/tassel>
- Visit Poster 993 (TASSEL 4.0: Faster, Larger Diversity Analysis: It's All About the Bits)
- Email developers listed on the poster