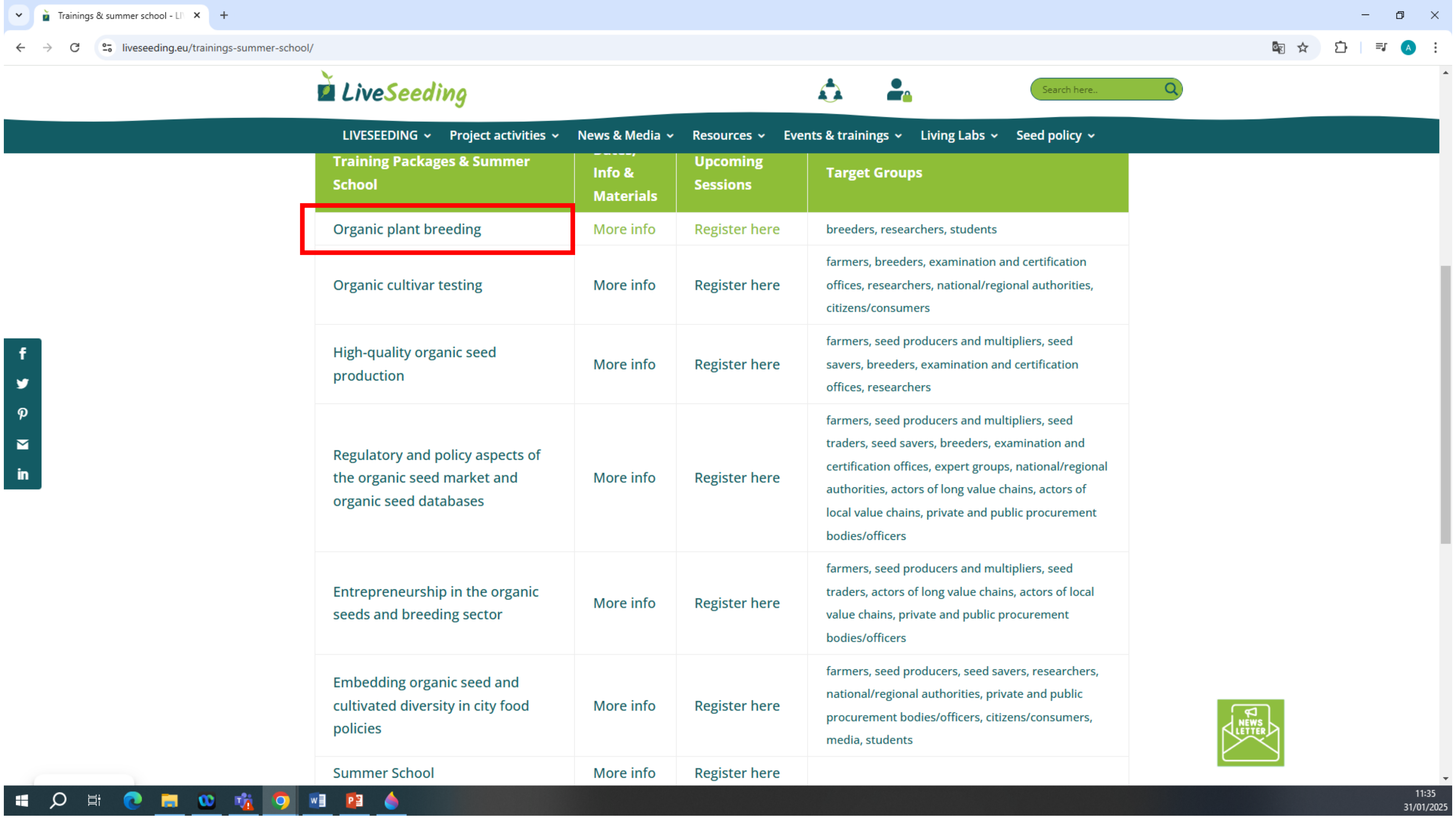
TRAINING IN ORGANIC BREEDING!

LiveSeeding

CONTEXT: Training in LIVESEEDING project

https://liveseeding.eu/trainings-summer-school/



Training Packages & Summer School	Info & Materials	Upcoming Sessions	Target Groups
Organic plant breeding	More info	Register here	breeders, researchers, students
Organic cultivar testing	More info	Register here	farmers, breeders, examination and certification offices, researchers, national/regional authorities, citizens/consumers
High-quality organic seed production	More info	Register here	farmers, seed producers and multipliers, seed savers, breeders, examination and certification offices, researchers
Regulatory and policy aspects of the organic seed market and organic seed databases	More info	Register here	farmers, seed producers and multipliers, seed traders, seed savers, breeders, examination and certification offices, expert groups, national/regional authorities, actors of long value chains, actors of local value chains, private and public procurement bodies/officers
Entrepreneurship in the organic seeds and breeding sector	More info	Register here	farmers, seed producers and multipliers, seed traders, actors of long value chains, actors of local value chains, private and public procurement bodies/officers
Embedding organic seed and cultivated diversity in city food policies	More info	Register here	farmers, seed producers, seed savers, researchers, national/regional authorities, private and public procurement bodies/officers, citizens/consumers, media, students
Summer School	More info	Register here	

Training in organic breeding organized in 5 Modules

1. **Module 1** - Plant Genetic Resources (PGRs): collection, conservation and exchange to support the increase of agrobiodiversity in farming systems
2. **Module 2** - Phenomics: approaches and tools for genetic resources and breeding material characterization - FEBRUARY 3rd 2025, 9:00 to 17:30 CET
3. **Module 3** - Breeding methods fundamentals - FEBRUARY 13th 2025, 9:00 to 18:00 CET
4. **Module 4** - Development and application of molecular methods in organic breeding - MARCH 4th 2025, 9:00 to 18:00 CET
5. **Module 5** - Organic heterogeneous material (OHM) design and development - MARCH 7th 2025, 9:00 to 18:00 CET

LiveSeeding

Training in organic breeding organized in 5 Modules

1. **Module 1** - Plant Genetic Resources (PGRs): collection, conservation and exchange to support the increase of agrobiodiversity in farming systems
2. **Module 2** - Phenomics: approaches and tools for genetic resources and breeding material characterization - FEBRUARY 3rd 2025, 9:00 to 17:30 CET
3. **Module 3** - Breeding methods fundamentals - FEBRUARY 13th 2025, 9:00 to 18:00 CET
4. **Module 4** - Development and application of molecular methods in organic breeding - MARCH 4th 2025, 9:00 to 18:00 CET
5. **Module 5** - Organic heterogeneous material (OHM) design and development - MARCH 7th 2025, 9:00 to 18:00 CET

LiveSeeding

Module 4 - Development and application of molecular methods in organic breeding

MARCH 4th 2025 9:00 to 18:00 CET

Unit 4.1: Genome Wide Association and its potential for understanding the genetic basis of complex traits

9:00-11:00 Partner in charge: UPV (Neus Ortega Albero)

11:00-11:30 Break

Unit 4.2: Genomic Selection and its potential in organic breeding

11:30-13:00 Partner in charge: FiBL (Michael Schneider)

13:00-14:30 Lunch Break

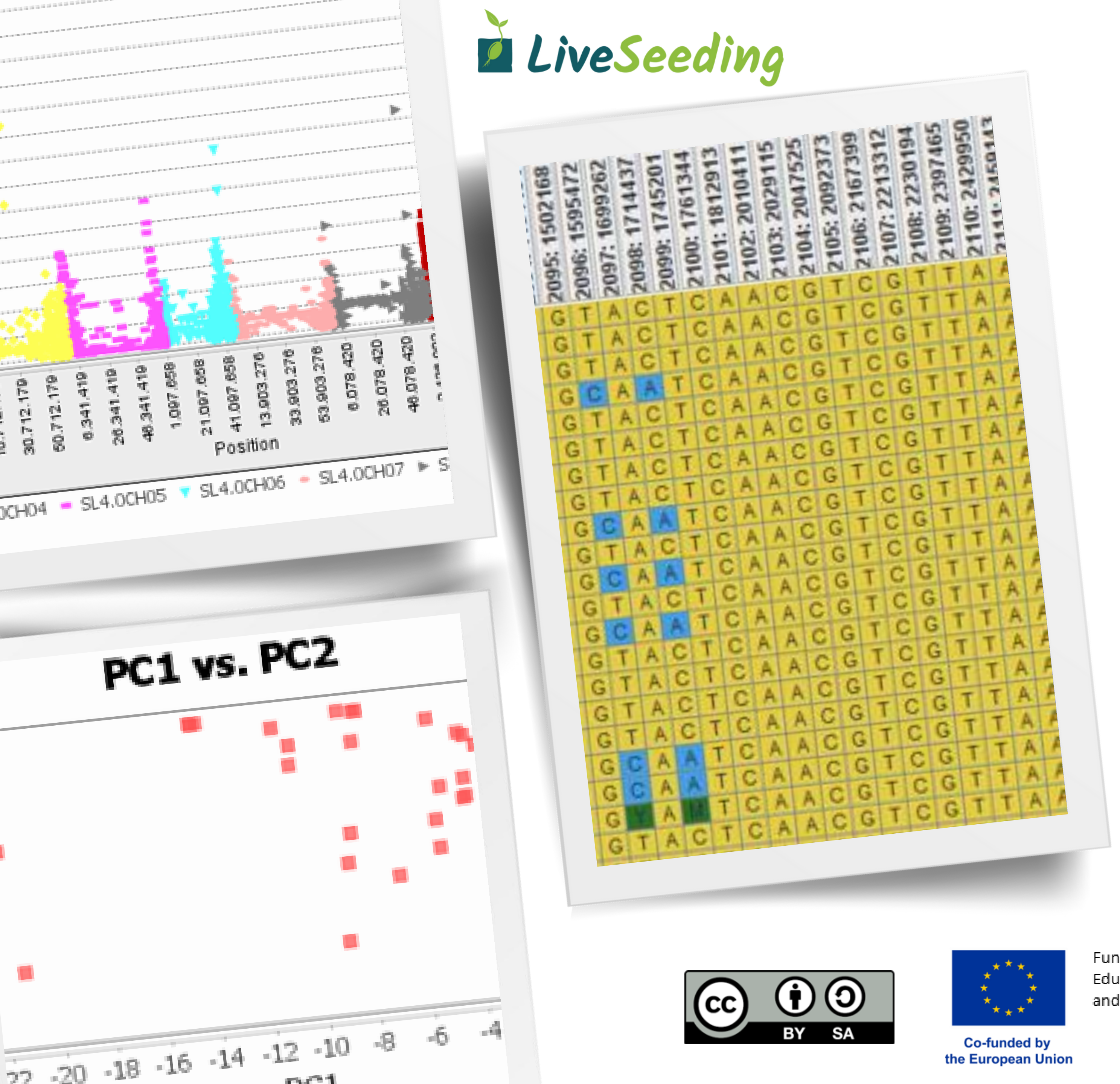
Unit 4.3: Application of molecular marker-based selection in practical breeding

14:30-16:00 Partners in charge: KIS (Barbara Pipan) + FiBL (Lupins; Teresa Lazzaro)

16:00-16:30 Break

Unit 4.4: Overview on Experimental populations (e.g. NILs, RILs, introgression families, MAGIC) and their role in molecular breeding approaches

16:30-18:00 Partner in charge: UPV (Neus Ortega-Albero, Adrian Rodríguez-Burruezo) + KIS (Vladimir Meglic)



LiveSeeding

Training in organic breeding

Module 4: **Development and application of molecular methods in organic breeding**

Unit 4.1: Genome Wide Association and its potential for understanding the genetic basis of complex traits

Author: **Neus Ortega Albero (UPV)**

Co-funded by the European Union

Funded by the European Union, the Swiss State Secretariat for Education, Research and Innovation (SERI) and UK Research and Innovation (UKRI).

UK Research and Innovation

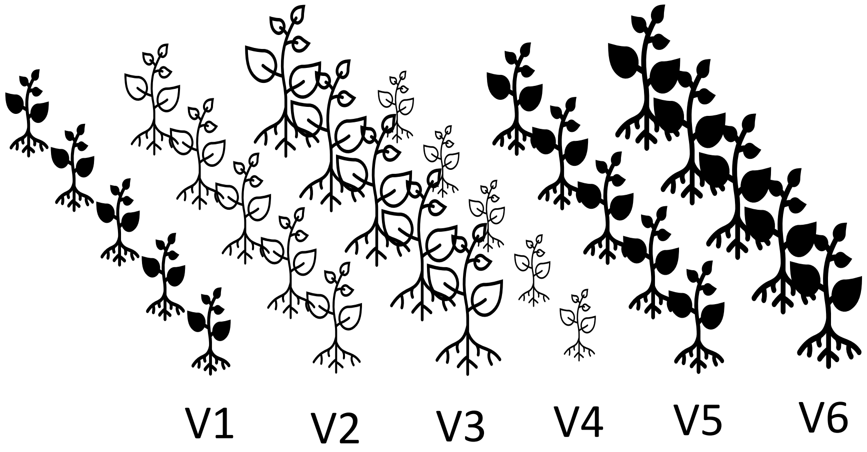
Relationship Gene - Trait

Genotype → Phenotype

***Meeting the causing genes*.**

a) Develop markers to select interesting phenotypes easily (Unit 3)

a) Establishing correlations between traits and finding if the genetic base is the same to select one complex trait through another easy-to-phenotype trait

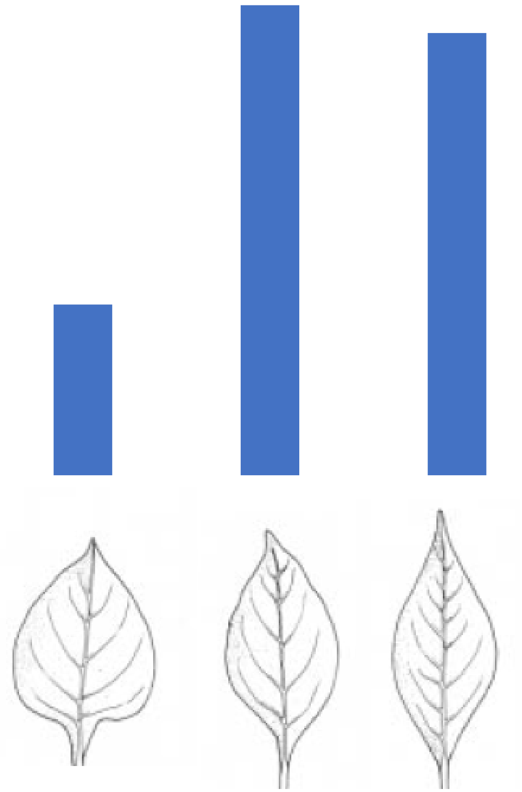


LiveSeeding

Qualitative vs. Quantitative traits

Qualitative traits → categories

Controlled by 1 or few genes



Quantitative traits → continuous expression

Controlled by many genes along the genome

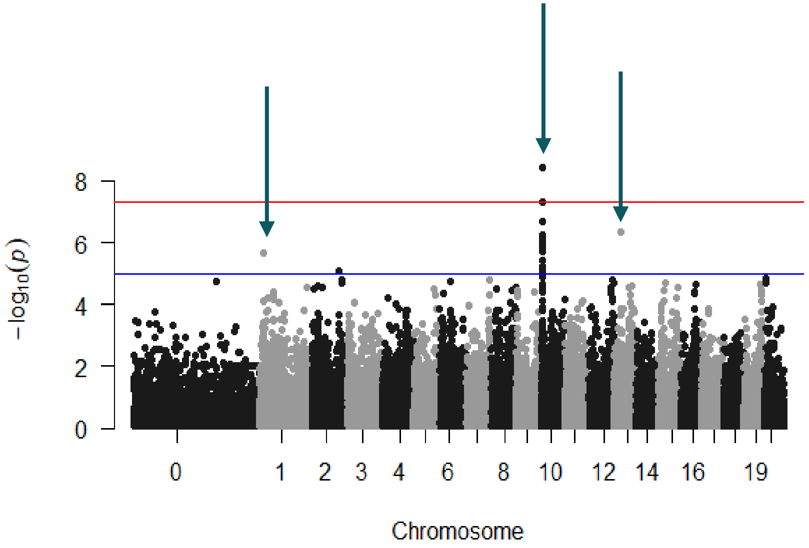
GWAS

Genome-Wide Association Analysis (GWAS)

Quantitative Trait Loci (QTL) → all the points (markers) in the genome associated statistically with the trait. The accumulative effect of these QTLs is the phenotype.

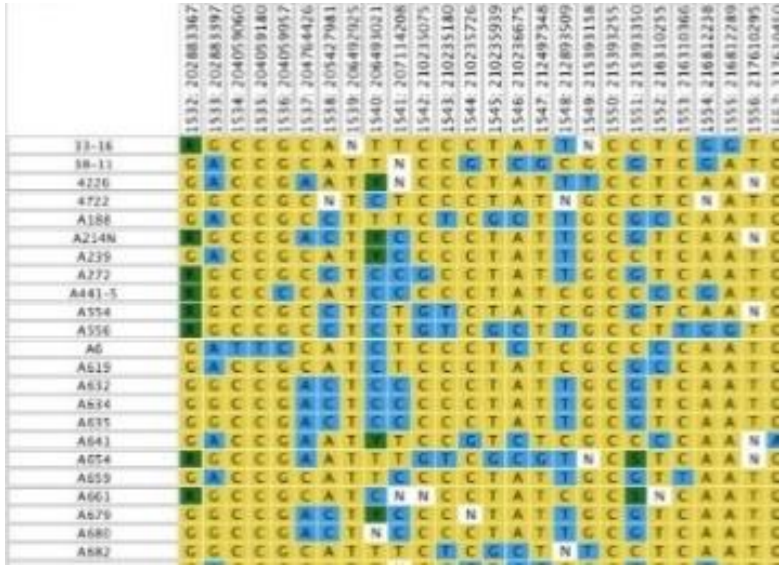
Exploit existing natural variability in a population accumulating recombination during decades.

Used in a **collection of varieties** or a **breeding population** embedding variation for the studied trait.

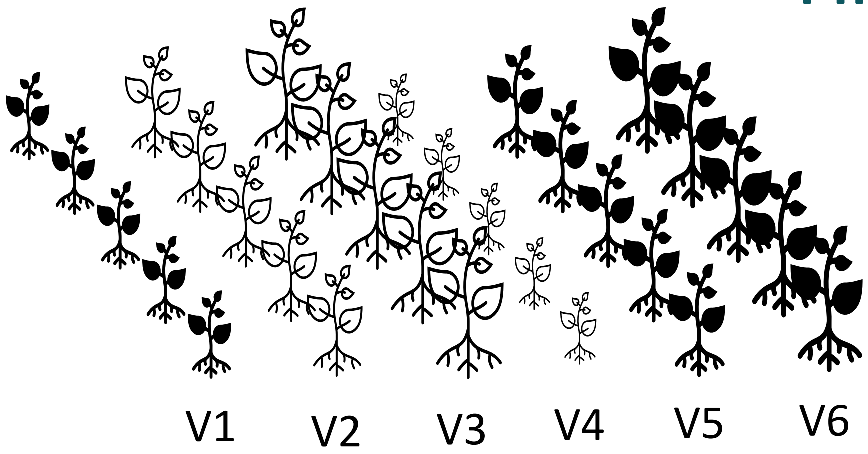


GWAS

Genotype



Phenotype



GWAS

GWAS

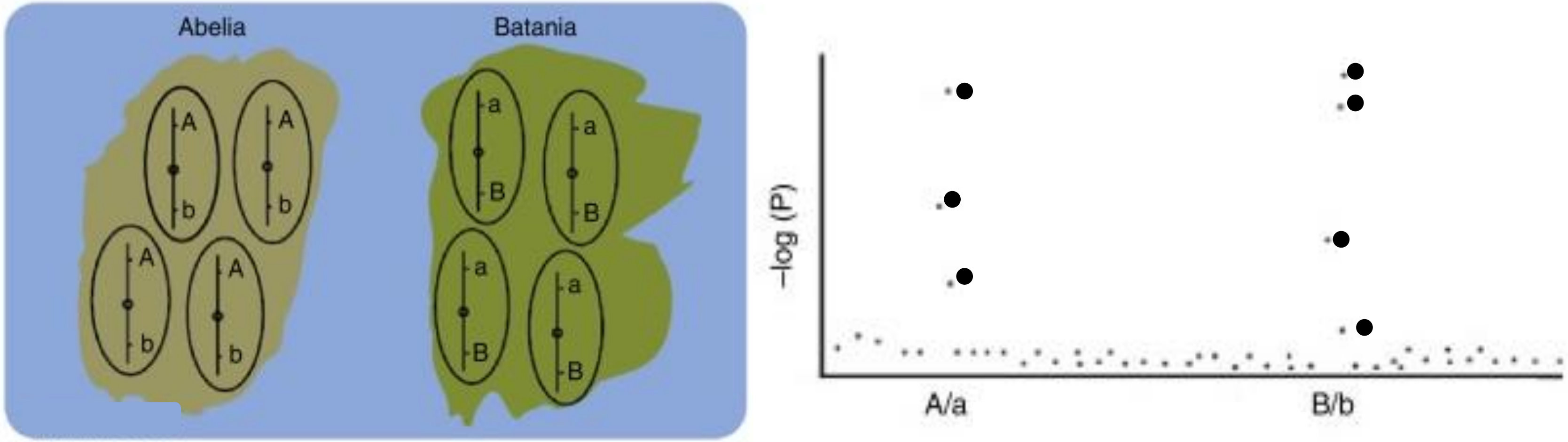
ADVANTAGES

- No need to create a breeding population
- Analyze a wider range of diversity
- Useful for any type of trait
- Work on several traits in very variable populations
- High precisión in gene mapping
- Association for complex traits like adaptation to low input conditions

DISADVANTAGES

- Need variant availability
- The population of study must show variability for the trait
- Avoid population stratification

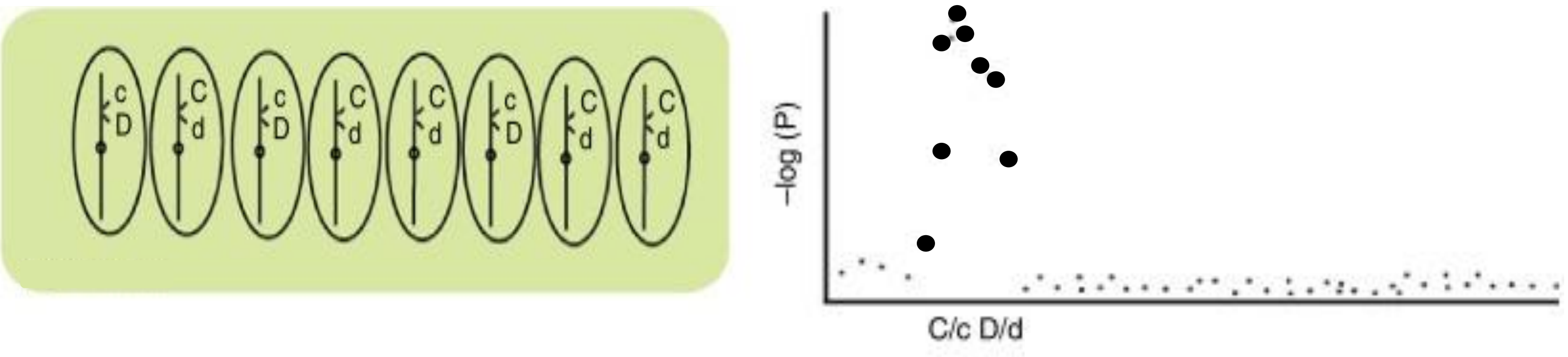
Fake Associations



Fixed alleles in the population

When only two combination of haplotypes are fixed in the population, both genes will appear as associated with the trait, even only one is really associated.

Fake Associations

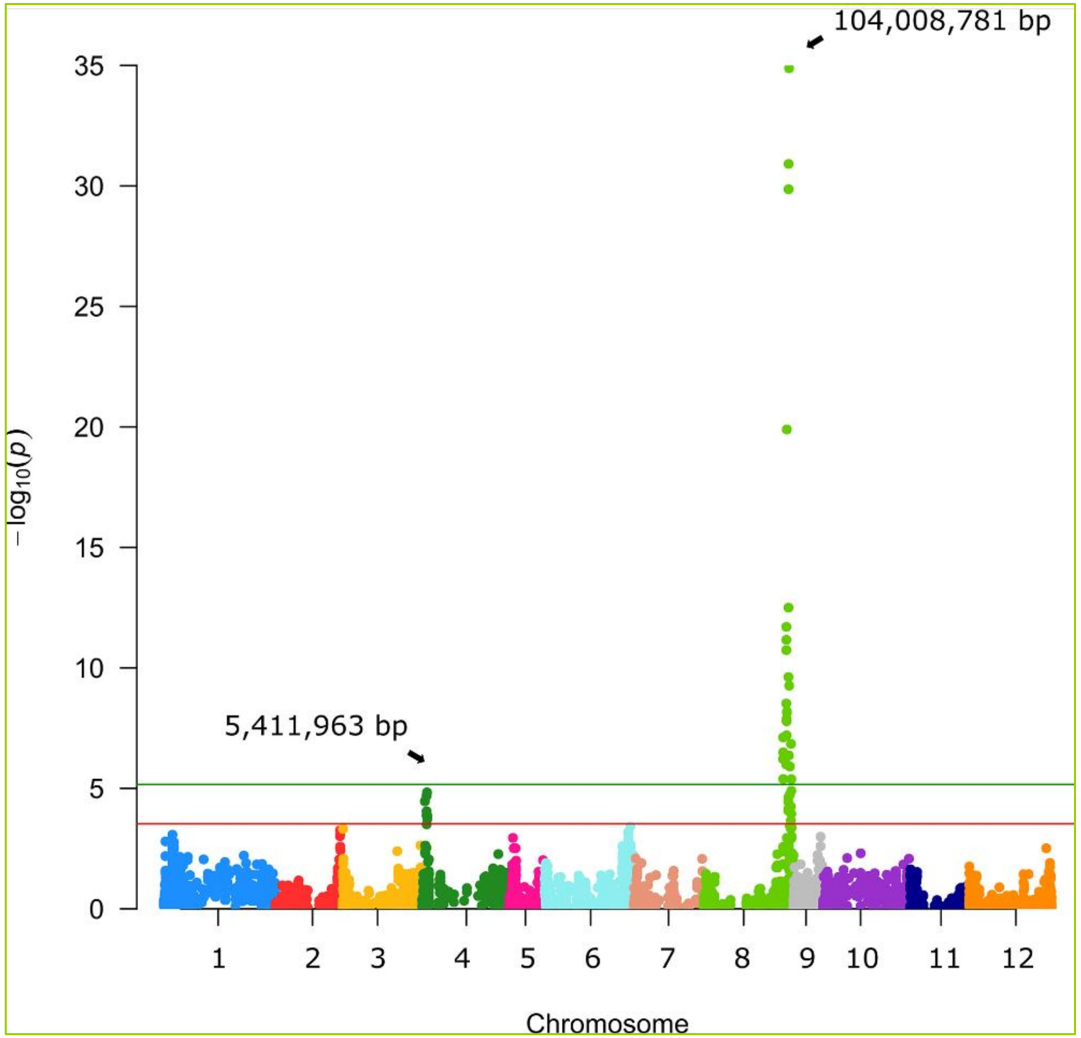


Ligated Genes or Linkage Disequilibrium

Two proximal genes always inherited together will be both associated to the trait, even only one of them is the cause of the phenotype.

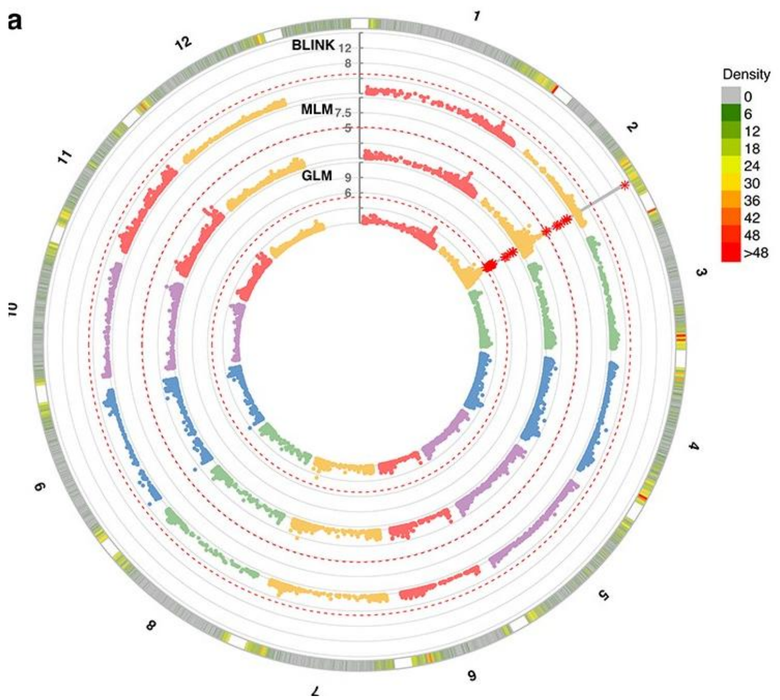
GWAS examples

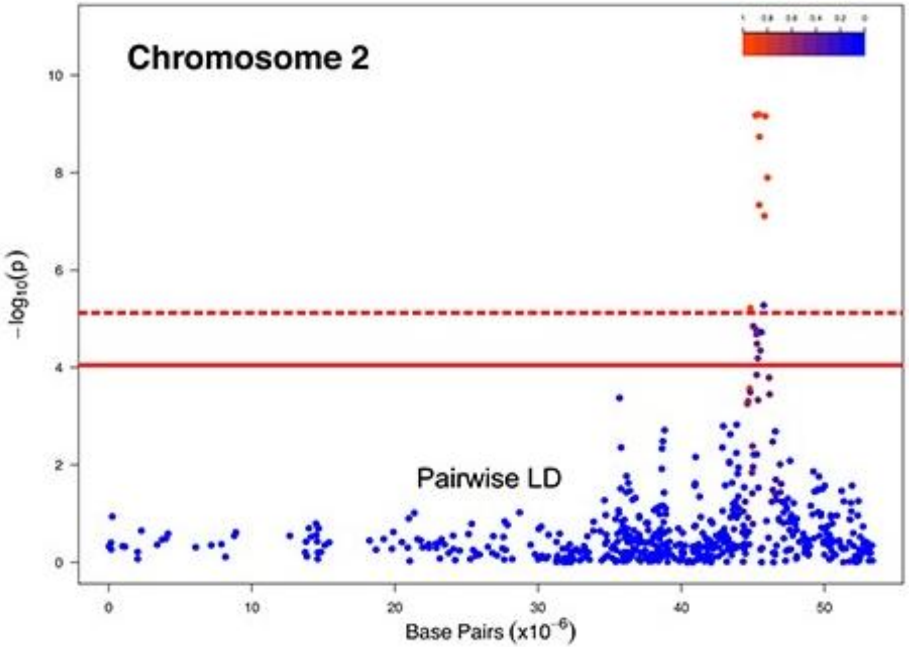
Eggplant → pigmentation of the peel



GWAS examples

Tomato → Locule number

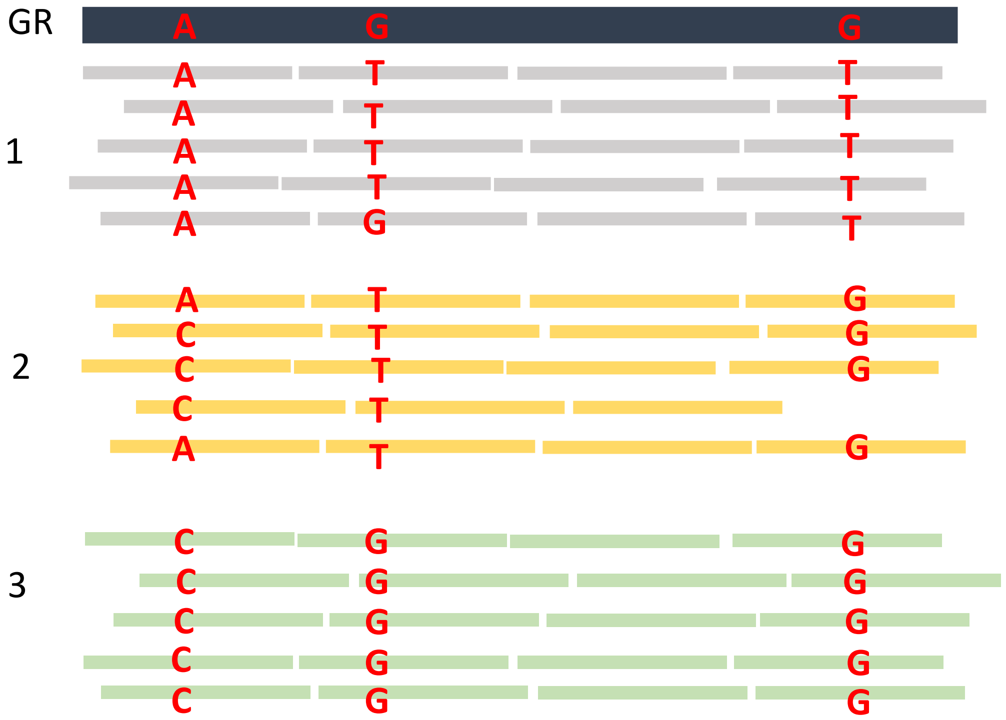




PRACTICAL EXAMPLE

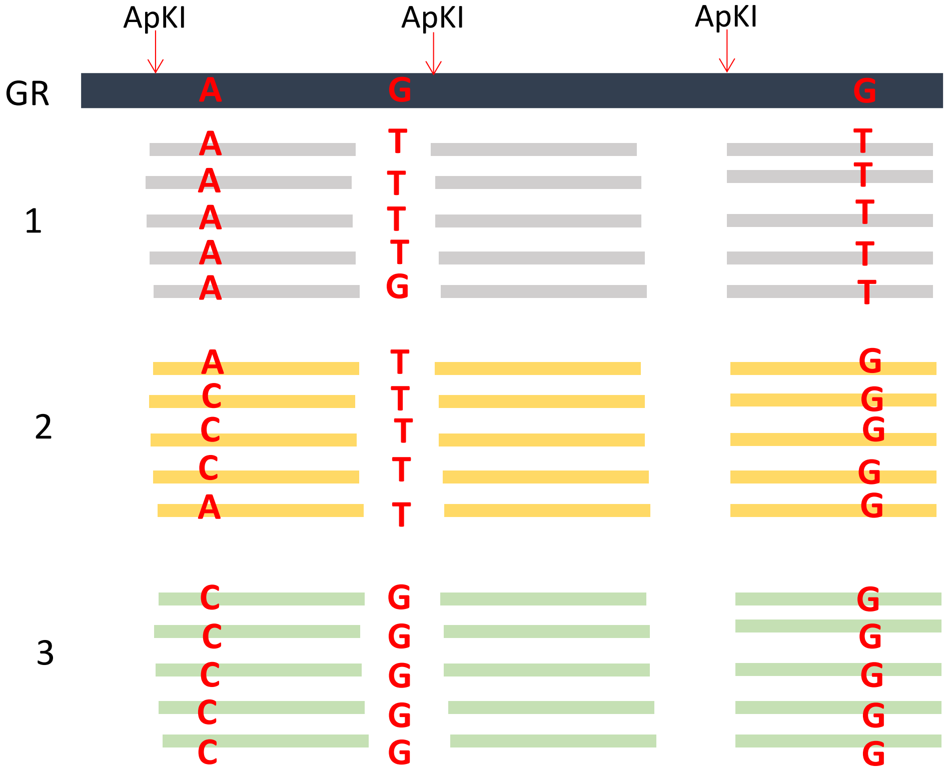
WGS vs. GBS

Whole Genome Sequencing (WGS)



1	0/0	0/1	1/1
2	0/1	1/1	0/0
3	1/1	0/0	0/0

Genotyping By Sequencing (GBS)



1	0/0	1/1
2	0/1	0/0
3	1/1	0/0

LiveSeeding

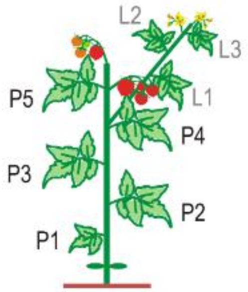
Practical Example

GWAS analysis in a collection of Solanum lycopersicum

Growth habit in tomato

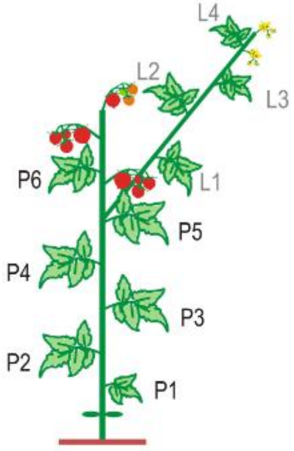
Determinate

1



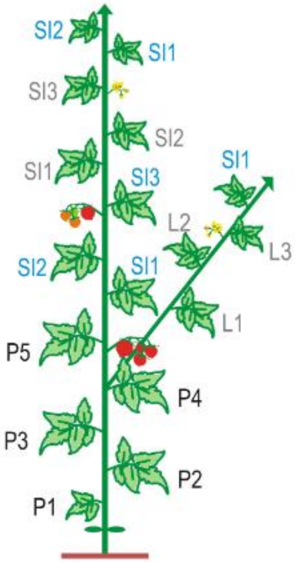
Semi - Determinate

2



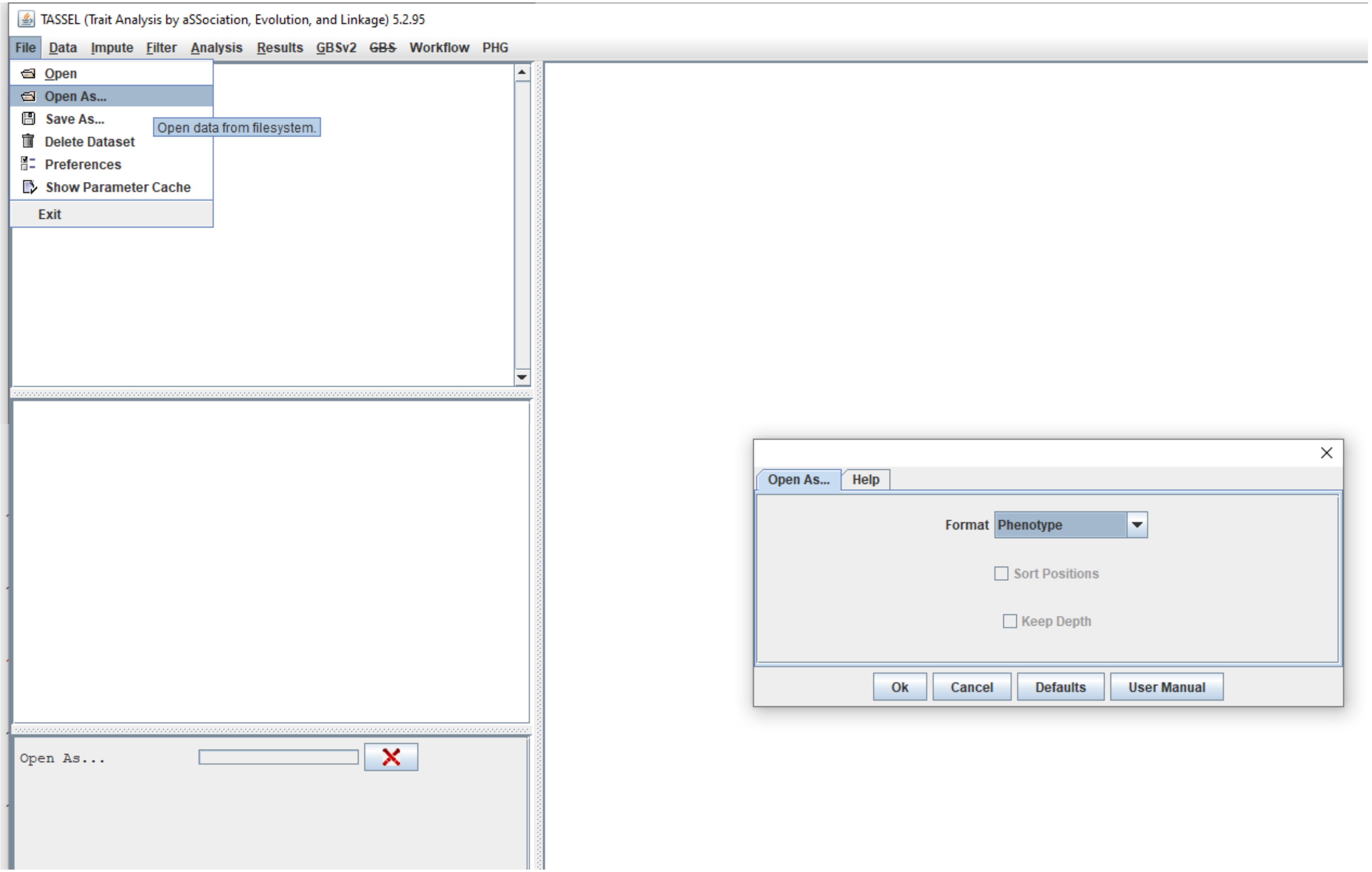
Indeterminate

3



Practical Example

1. Open phenotypic data file



Open data from filesystem.

Format Phenotype

☐ Sort Positions

☐ Keep Depth

```
Open As...
```

Practical Example

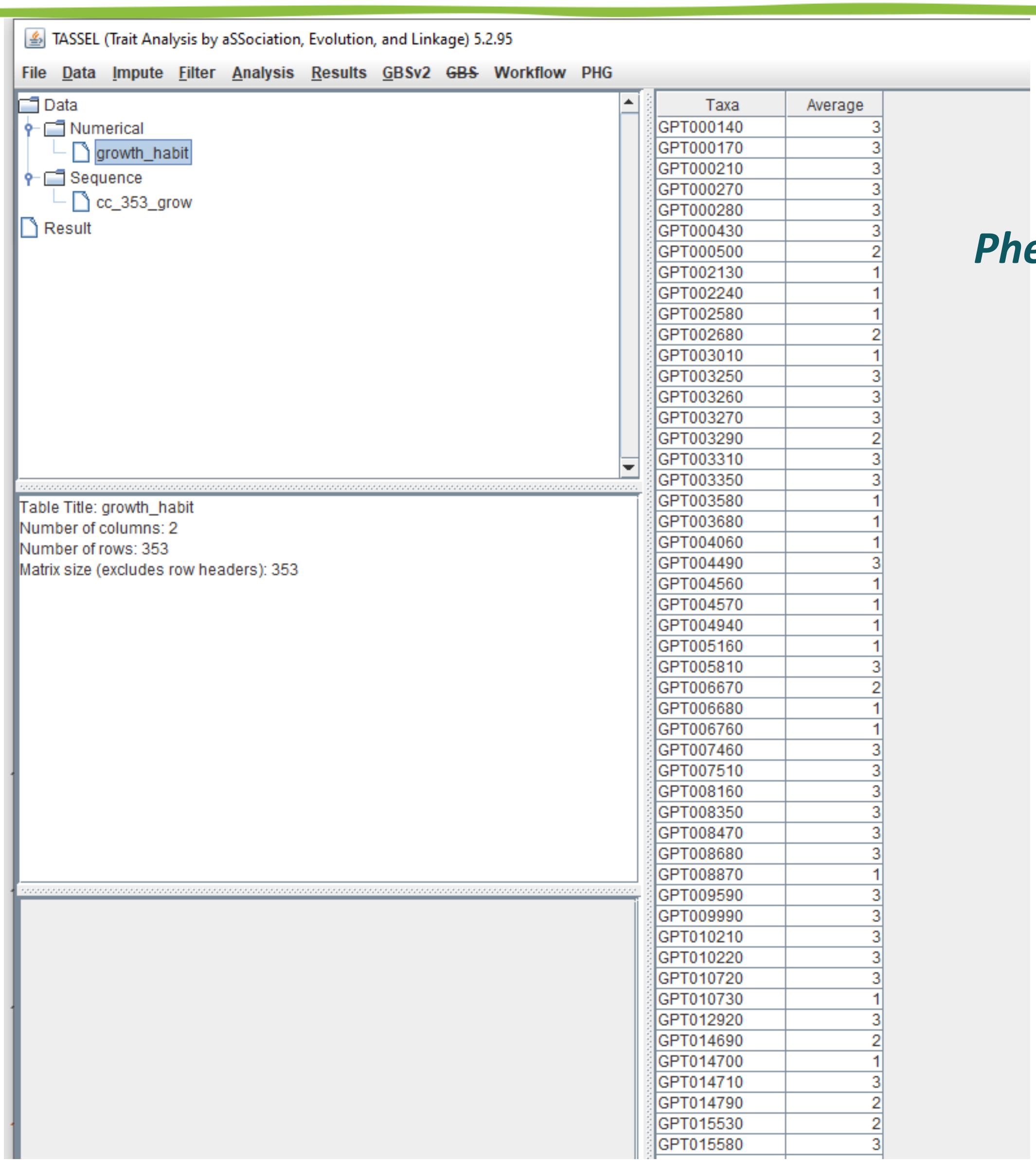


Table Title: growth_habit
Number of columns: 2
Number of rows: 353
Matrix size (excludes row headers): 353

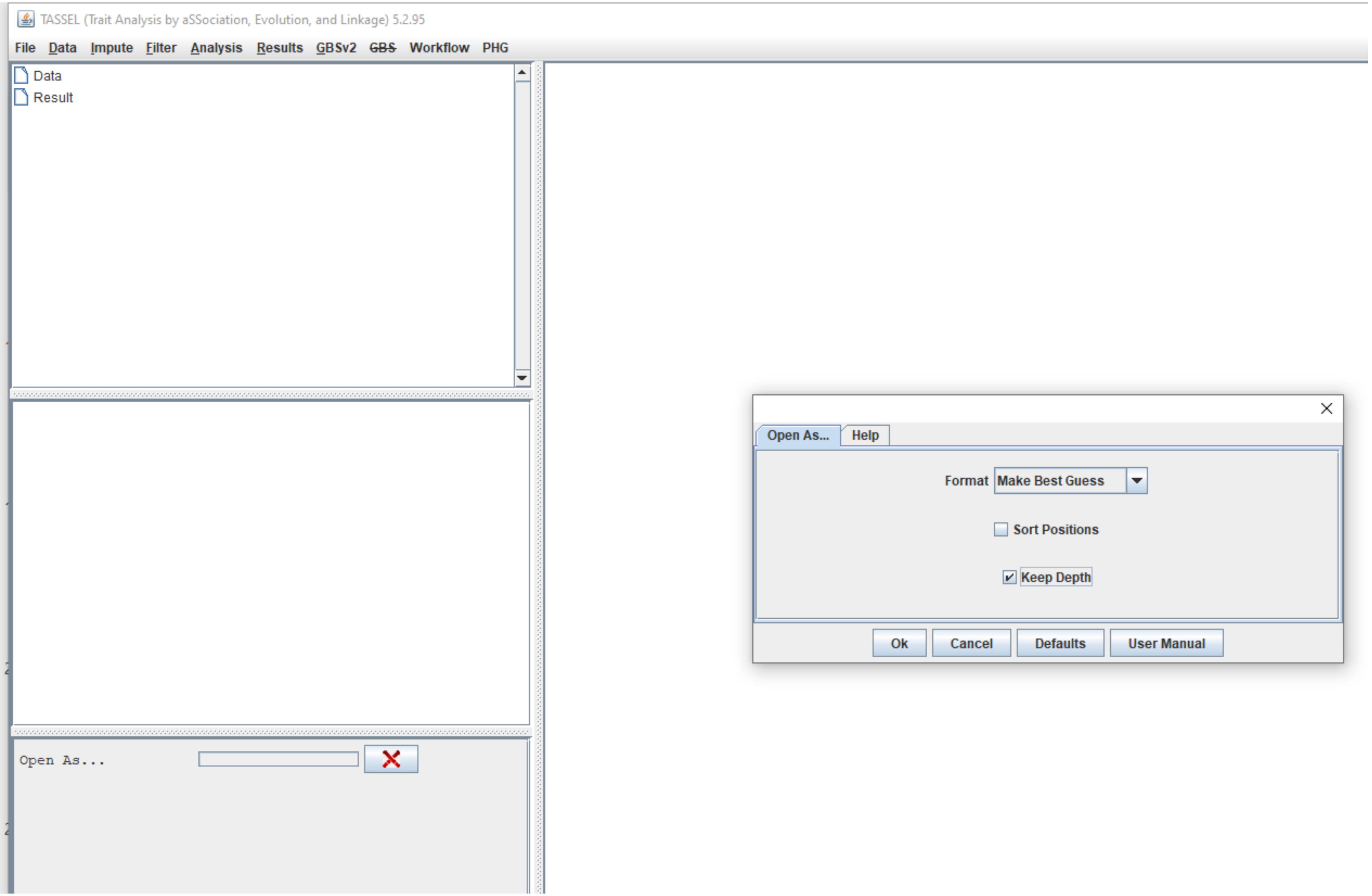
Taxa	Average
GPT000140	3
GPT000170	3
GPT000210	3
GPT000270	3
GPT000280	3
GPT000430	3
GPT000500	2
GPT002130	1
GPT002240	1
GPT002580	1
GPT002680	2
GPT003010	1
GPT003250	3
GPT003260	3
GPT003270	3
GPT003290	2
GPT003310	3
GPT003350	3
GPT003580	1
GPT003680	1
GPT004060	1
GPT004490	3
GPT004560	1
GPT004570	1
GPT004940	1
GPT005160	1
GPT005810	3
GPT006670	2
GPT006680	1
GPT006760	1
GPT007460	3
GPT007510	3
GPT008160	3
GPT008350	3
GPT008470	3
GPT008680	3
GPT008870	1
GPT009590	3
GPT009990	3
GPT010210	3
GPT010220	3
GPT010720	3
GPT010730	1
GPT012920	3
GPT014690	2
GPT014700	1
GPT014710	3
GPT014790	2
GPT015530	2
GPT015580	3

Phenotypic data displaying

LiveSeeding

Practical Example

2. Open GBS data file



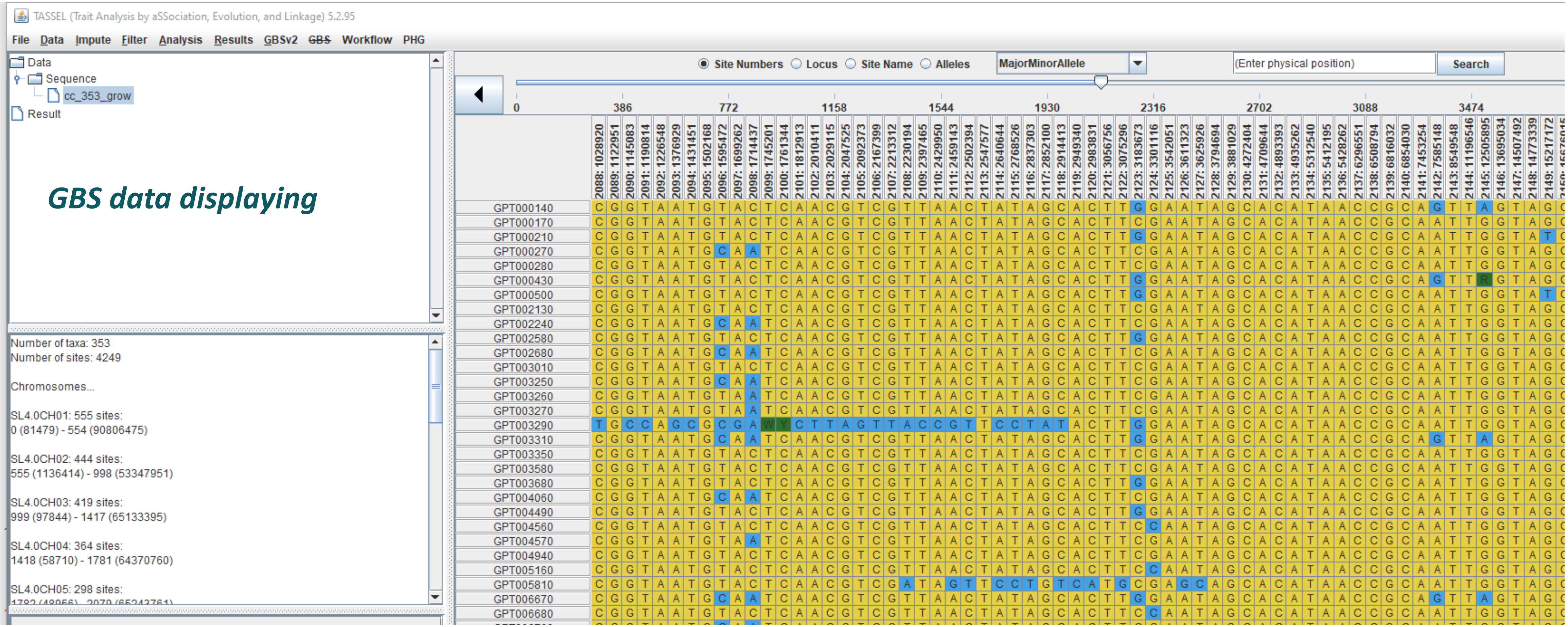
TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.95

Format Make Best Guess

☐ Sort Positions

☑ Keep Depth

Practical Example



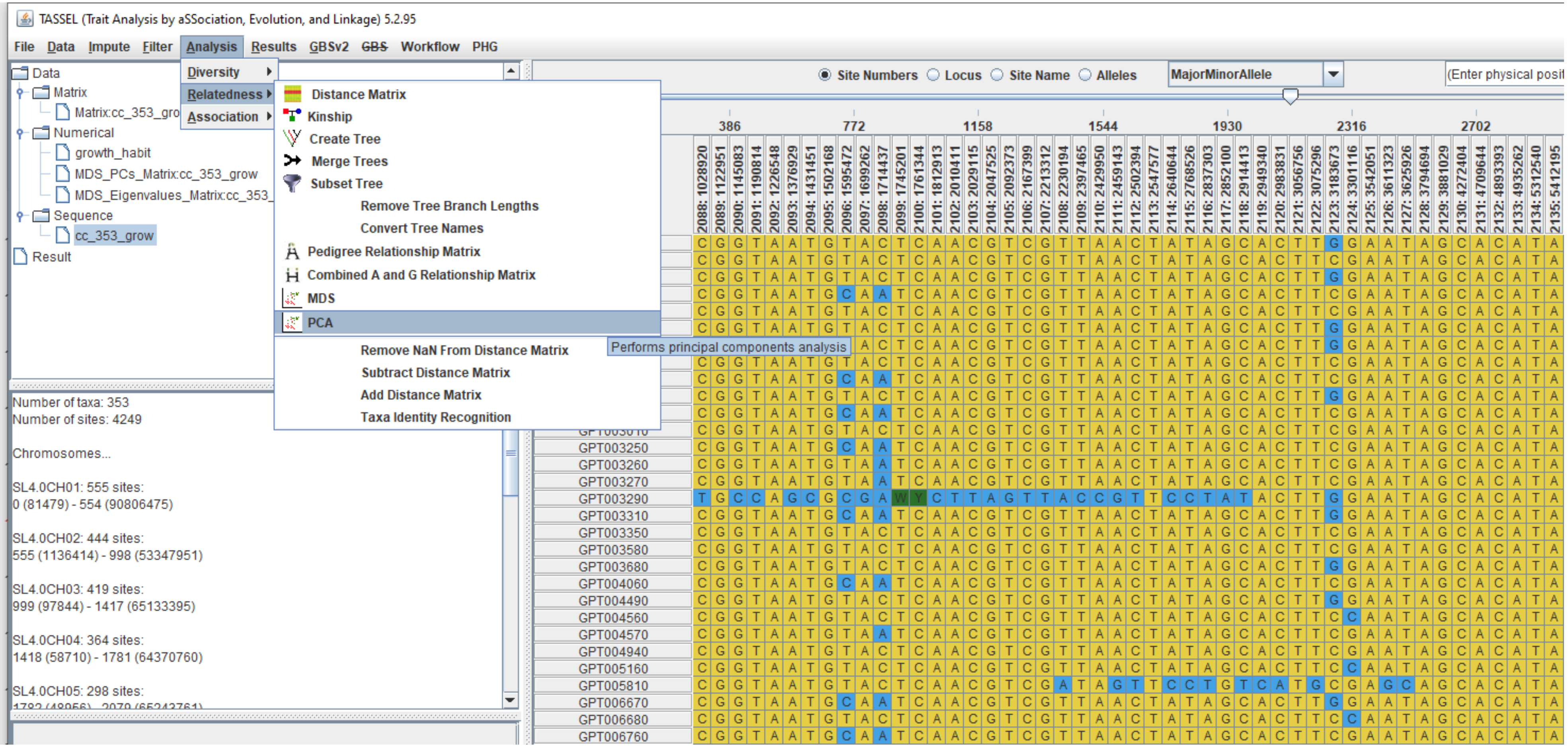
GBS data displaying

```
Number of taxa: 353
Number of sites: 4249

Chromosomes...

SL4.0CH01: 555 sites:
0 (81479) - 554 (90806475)

SL4.0CH02: 444 sites:
555 (1136414) - 998 (53347951)

SL4.0CH03: 419 sites:
999 (97844) - 1417 (65133395)

SL4.0CH04: 364 sites:
1418 (58710) - 1781 (64370760)

SL4.0CH05: 298 sites:
```

	2088: 1028920	2089: 1122951	2090: 1145083	2091: 1190814	2092: 1226548	2093: 1376929	2094: 1431451	2095: 1502168	2096: 1595472	2097: 1699262	2098: 1714437	2099: 1745201	2100: 1761344	2101: 1812913	2102: 2010411	2103: 2029115	2104: 2047525	2105: 2092373	2106: 2167399	2107: 2213312	2108: 2230194	2109: 2397465	2110: 2429950	2111: 2459143	2112: 2502394	2113: 2547577	2114: 2640644	2115: 2768526	2116: 2837303	2117: 2852100	2118: 2914413	2119: 2949340	2120: 2983831	2121: 3056756	2122: 3075296	2123: 3183673	2124: 3301116	2125: 3542051	2126: 3611323	2127: 3625926	2128: 3794694	2129: 3881029	2130: 4272404	2131: 4709644	2132: 4893393	2133: 4935262	2134: 5312540	2135: 5412195	2136: 5428262	2137: 6296551	2138: 6508794	2139: 6816032	2140: 6854030	2141: 7453254	2142: 7585148	2143: 8549548	2144: 11196546	2145: 12505895	2146: 13695034	2147: 14507492	2148: 14773339	2149: 15217172
GPT000140	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	G	T	T	A	G	T	A	G
GPT000170	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT000210	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	T
GPT000270	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT000280	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT000430	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	G	T	T	R	G	T	A	G
GPT000500	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	T
GPT002130	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT002240	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT002580	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT002680	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT003010	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT003250	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT003260	C	G	G	T	A	A	T	G	T	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT003270	C	G	G	T	A	A	T	G	T	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT003290	T	G	C	C	A	G	C	G	C	G	A	W	Y	C	T	T	A	G	T	T	A	C	C	G	T	T	C	C	T	A	T	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT003310	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	G	T	T	A	G	T	A	G
GPT003350	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT003580	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT003680	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT004060	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT004490	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT004560	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	C	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT004570	C	G	G	T	A	A	T	G	T	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT004940	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT005160	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	C	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT005810	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	A	T	A	G	T	T	C	C	T	G	T	C	A	T	G	C	G	A	G	C	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G
GPT006670	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	G	T	T	A	G	T	A	G
GPT006680	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	C	A	A	T	A	G	C	A	C	A	T	A	A	C	C	G	C	A	A	T	T	G	G	T	A	G

Practical Example

3. Checking population structure with PCA



Number of taxa: 353
Number of sites: 4249

Chromosomes...

SL4.0CH01: 555 sites:
0 (81479) - 554 (90806475)

SL4.0CH02: 444 sites:
555 (1136414) - 998 (53347951)

SL4.0CH03: 419 sites:
999 (97844) - 1417 (65133395)

SL4.0CH04: 364 sites:
1418 (58710) - 1781 (64370760)

SL4.0CH05: 298 sites:

Practical Example

3. Checking population structure with PCA

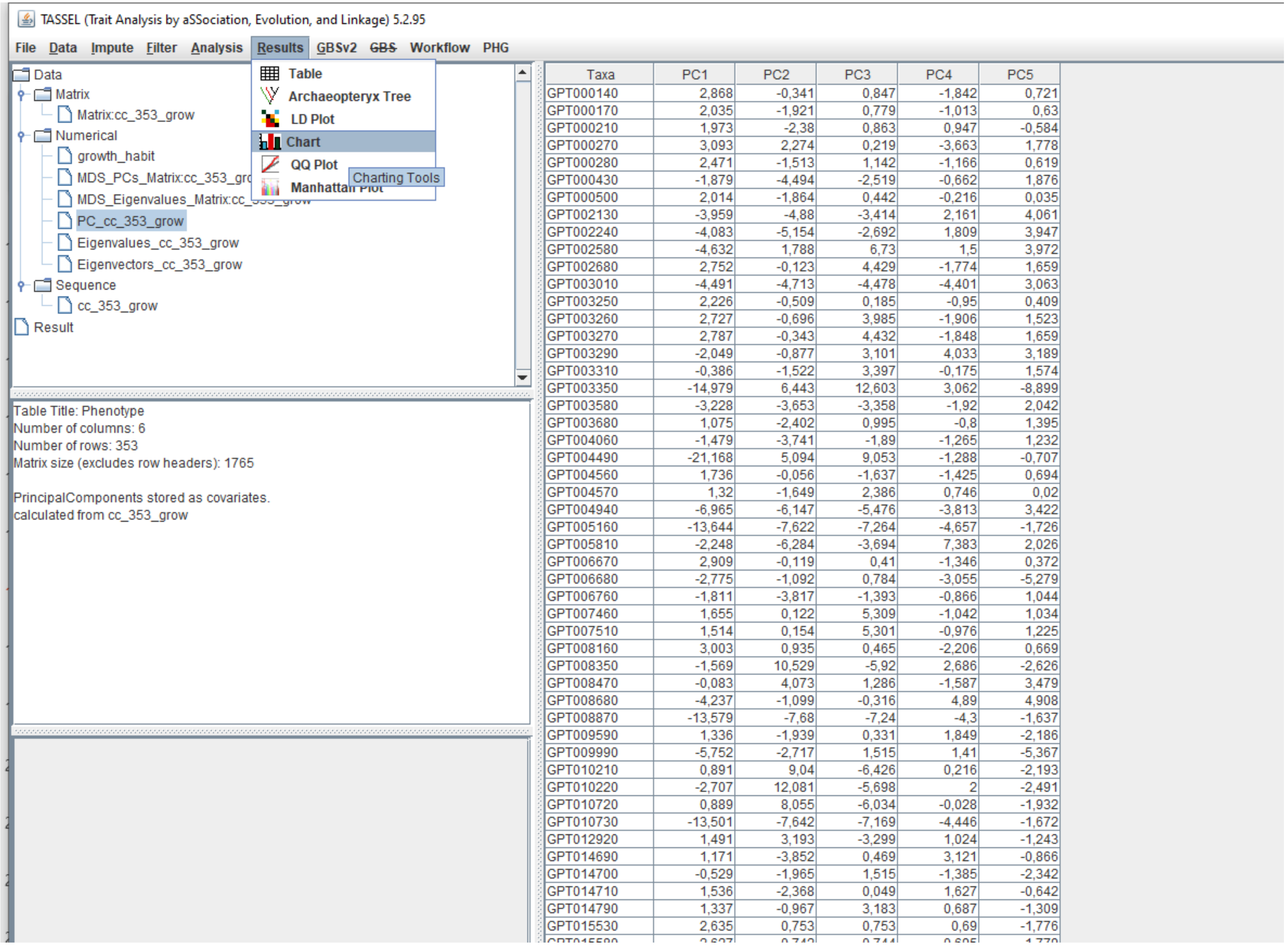


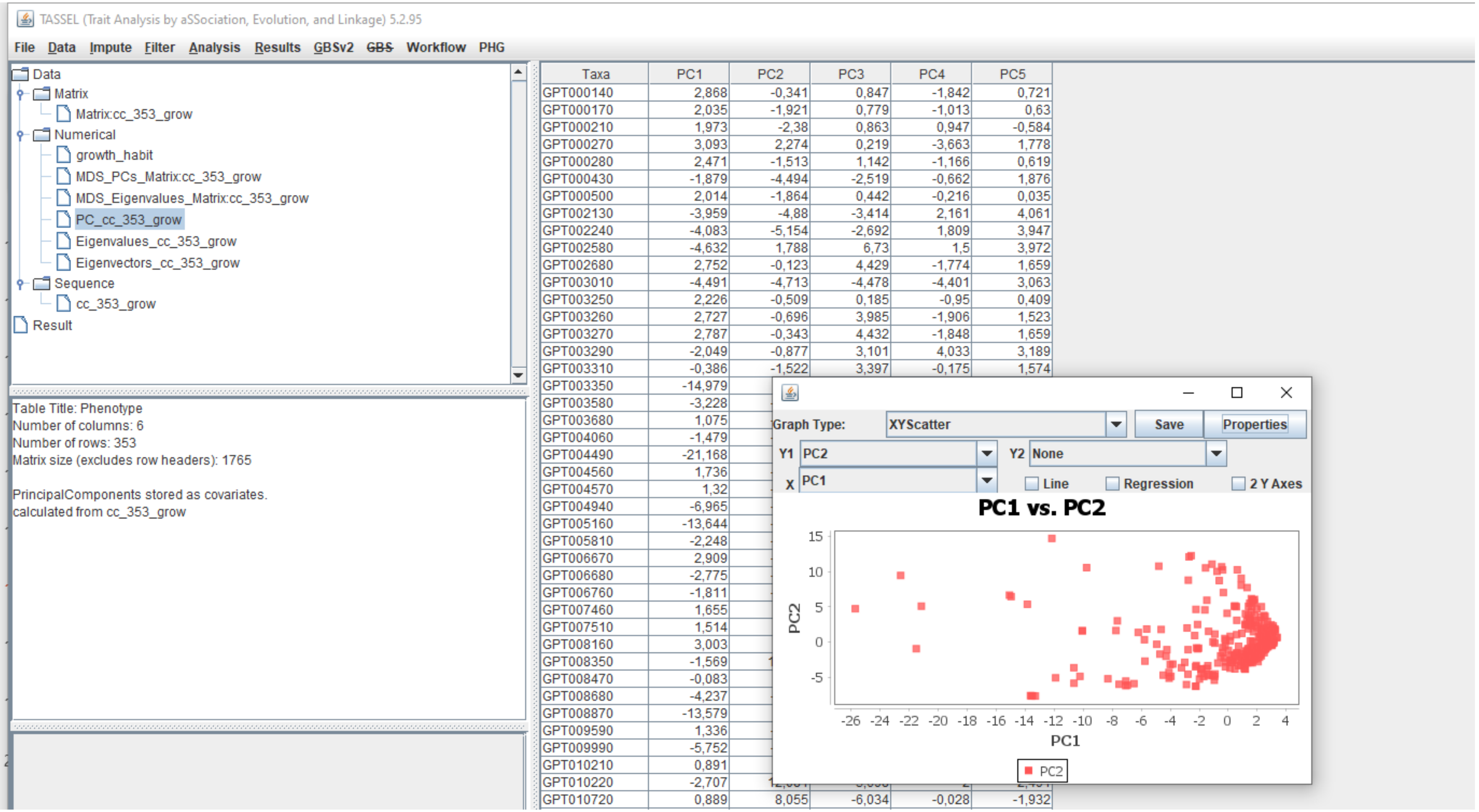
Table Title: Phenotype
Number of columns: 6
Number of rows: 353
Matrix size (excludes row headers): 1765

PrincipalComponents stored as covariates.
calculated from cc_353_grow

Taxa	PC1	PC2	PC3	PC4	PC5
GPT000140	2,868	-0,341	0,847	-1,842	0,721
GPT000170	2,035	-1,921	0,779	-1,013	0,63
GPT000210	1,973	-2,38	0,863	0,947	-0,584
GPT000270	3,093	2,274	0,219	-3,663	1,778
GPT000280	2,471	-1,513	1,142	-1,166	0,619
GPT000430	-1,879	-4,494	-2,519	-0,662	1,876
GPT000500	2,014	-1,864	0,442	-0,216	0,035
GPT002130	-3,959	-4,88	-3,414	2,161	4,061
GPT002240	-4,083	-5,154	-2,692	1,809	3,947
GPT002580	-4,632	1,788	6,73	1,5	3,972
GPT002680	2,752	-0,123	4,429	-1,774	1,659
GPT003010	-4,491	-4,713	-4,478	-4,401	3,063
GPT003250	2,226	-0,509	0,185	-0,95	0,409
GPT003260	2,727	-0,696	3,985	-1,906	1,523
GPT003270	2,787	-0,343	4,432	-1,848	1,659
GPT003290	-2,049	-0,877	3,101	4,033	3,189
GPT003310	-0,386	-1,522	3,397	-0,175	1,574
GPT003350	-14,979	6,443	12,603	3,062	-8,899
GPT003580	-3,228	-3,653	-3,358	-1,92	2,042
GPT003680	1,075	-2,402	0,995	-0,8	1,395
GPT004060	-1,479	-3,741	-1,89	-1,265	1,232
GPT004490	-21,168	5,094	9,053	-1,288	-0,707
GPT004560	1,736	-0,056	-1,637	-1,425	0,694
GPT004570	1,32	-1,649	2,386	0,746	0,02
GPT004940	-6,965	-6,147	-5,476	-3,813	3,422
GPT005160	-13,644	-7,622	-7,264	-4,657	-1,726
GPT005810	-2,248	-6,284	-3,694	7,383	2,026
GPT006670	2,909	-0,119	0,41	-1,346	0,372
GPT006680	-2,775	-1,092	0,784	-3,055	-5,279
GPT006760	-1,811	-3,817	-1,393	-0,866	1,044
GPT007460	1,655	0,122	5,309	-1,042	1,034
GPT007510	1,514	0,154	5,301	-0,976	1,225
GPT008160	3,003	0,935	0,465	-2,206	0,669
GPT008350	-1,569	10,529	-5,92	2,686	-2,626
GPT008470	-0,083	4,073	1,286	-1,587	3,479
GPT008680	-4,237	-1,099	-0,316	4,89	4,908
GPT008870	-13,579	-7,68	-7,24	-4,3	-1,637
GPT009590	1,336	-1,939	0,331	1,849	-2,186
GPT009990	-5,752	-2,717	1,515	1,41	-5,367
GPT010210	0,891	9,04	-6,426	0,216	-2,193
GPT010220	-2,707	12,081	-5,698	2	-2,491
GPT010720	0,889	8,055	-6,034	-0,028	-1,932
GPT010730	-13,501	-7,642	-7,169	-4,446	-1,672
GPT012920	1,491	3,193	-3,299	1,024	-1,243
GPT014690	1,171	-3,852	0,469	3,121	-0,866
GPT014700	-0,529	-1,965	1,515	-1,385	-2,342
GPT014710	1,536	-2,368	0,049	1,627	-0,642
GPT014790	1,337	-0,967	3,183	0,687	-1,309
GPT015530	2,635	0,753	0,753	0,69	-1,776

Practical Example

Displaying PCA



- Data
  - Matrix
    - Matrix:cc_353_grow
  - Numerical
    - growth_habit
    - MDS_PCs_Matrix:cc_353_grow
    - MDS_Eigenvalues_Matrix:cc_353_grow
    - PC_cc_353_grow
    - Eigenvalues_cc_353_grow
    - Eigenvectors_cc_353_grow
  - Sequence
    - cc_353_grow
- Result

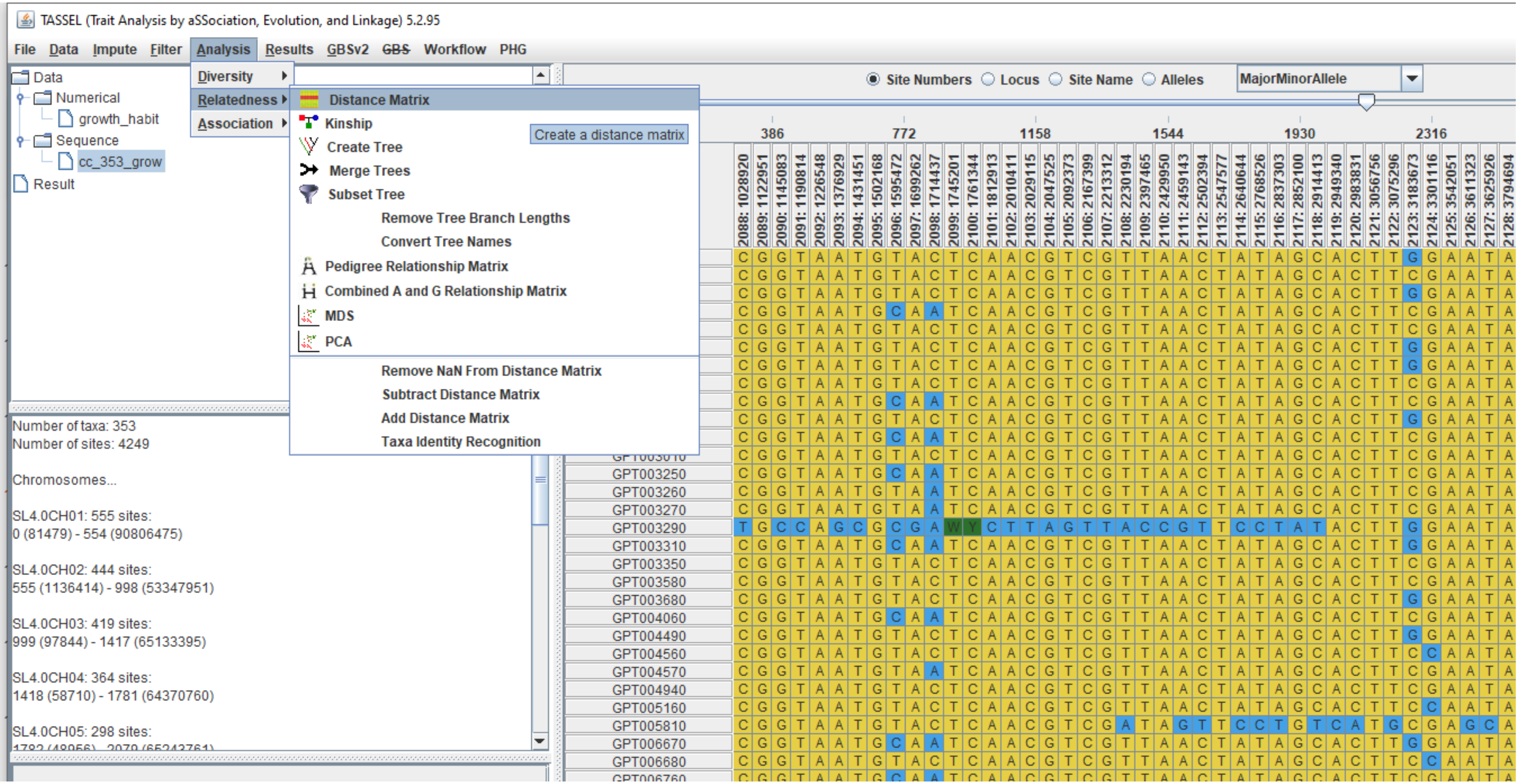
Table Title: Phenotype
Number of columns: 6
Number of rows: 353
Matrix size (excludes row headers): 1765

PrincipalComponents stored as covariates.
calculated from cc_353_grow

Taxa	PC1	PC2	PC3	PC4	PC5
GPT000140	2,868	-0,341	0,847	-1,842	0,721
GPT000170	2,035	-1,921	0,779	-1,013	0,63
GPT000210	1,973	-2,38	0,863	0,947	-0,584
GPT000270	3,093	2,274	0,219	-3,663	1,778
GPT000280	2,471	-1,513	1,142	-1,166	0,619
GPT000430	-1,879	-4,494	-2,519	-0,662	1,876
GPT000500	2,014	-1,864	0,442	-0,216	0,035
GPT002130	-3,959	-4,88	-3,414	2,161	4,061
GPT002240	-4,083	-5,154	-2,692	1,809	3,947
GPT002580	-4,632	1,788	6,73	1,5	3,972
GPT002680	2,752	-0,123	4,429	-1,774	1,659
GPT003010	-4,491	-4,713	-4,478	-4,401	3,063
GPT003250	2,226	-0,509	0,185	-0,95	0,409
GPT003260	2,727	-0,696	3,985	-1,906	1,523
GPT003270	2,787	-0,343	4,432	-1,848	1,659
GPT003290	-2,049	-0,877	3,101	4,033	3,189
GPT003310	-0,386	-1,522	3,397	-0,175	1,574
GPT003350	-14,979				
GPT003580	-3,228				
GPT003680	1,075				
GPT004060	-1,479				
GPT004490	-21,168				
GPT004560	1,736				
GPT004570	1,32				
GPT004940	-6,965				
GPT005160	-13,644				
GPT005810	-2,248				
GPT006670	2,909				
GPT006680	-2,775				
GPT006760	-1,811				
GPT007460	1,655				
GPT007510	1,514				
GPT008160	3,003				
GPT008350	-1,569				
GPT008470	-0,083				
GPT008680	-4,237				
GPT008870	-13,579				
GPT009590	1,336				
GPT009990	-5,752				
GPT010210	0,891				
GPT010220	-2,707				
GPT010720	0,889	8,055	-6,034	-0,028	-1,932

Practical Example

4. Calculating relationships within the population with the Distance Matrix



```
Number of taxa: 353
Number of sites: 4249

Chromosomes...

SL4.0CH01: 555 sites:
0 (81479) - 554 (90806475)

SL4.0CH02: 444 sites:
555 (1136414) - 998 (53347951)

SL4.0CH03: 419 sites:
999 (97844) - 1417 (65133395)

SL4.0CH04: 364 sites:
1418 (58710) - 1781 (64370760)

SL4.0CH05: 298 sites:
```

LiveSeeding

Practical Example

5. Incorporating the relationships to the analysis with the MDS

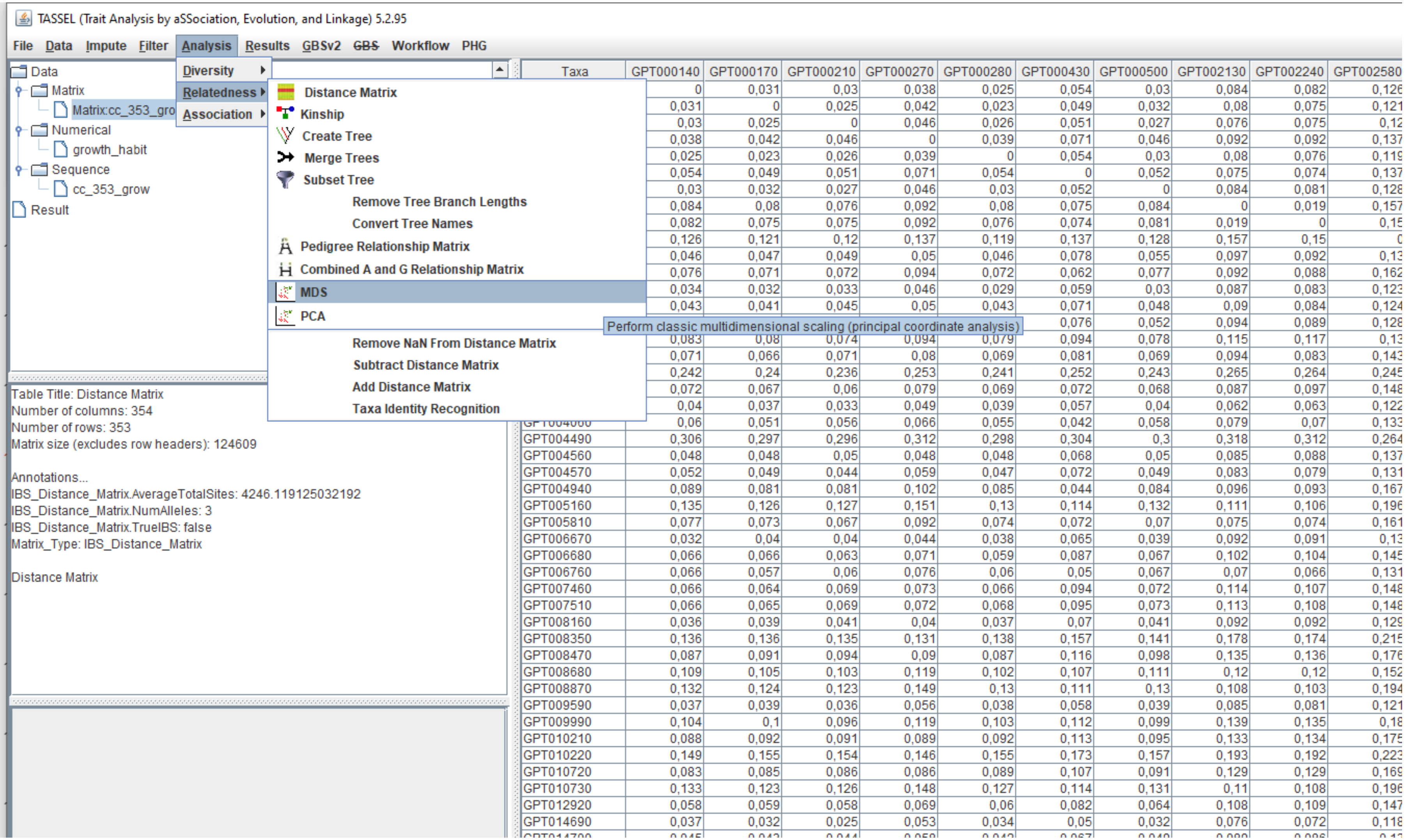


Table Title: Distance Matrix
Number of columns: 354
Number of rows: 353
Matrix size (excludes row headers): 124609

Annotations...
IBS_Distance_Matrix.AverageTotalSites: 4246.119125032192
IBS_Distance_Matrix.NumAlleles: 3
IBS_Distance_Matrix.TrueIBS: false
Matrix_Type: IBS_Distance_Matrix

Distance Matrix

Taxa	GPT000140	GPT000170	GPT000210	GPT000270	GPT000280	GPT000430	GPT000500	GPT002130	GPT002240	GPT002580
	0	0,031	0,03	0,038	0,025	0,054	0,03	0,084	0,082	0,126
	0,031	0	0,025	0,042	0,023	0,049	0,032	0,08	0,075	0,121
	0,03	0,025	0	0,046	0,026	0,051	0,027	0,076	0,075	0,12
	0,038	0,042	0,046	0	0,039	0,071	0,046	0,092	0,092	0,137
	0,025	0,023	0,026	0,039	0	0,054	0,03	0,08	0,076	0,119
	0,054	0,049	0,051	0,071	0,054	0	0,052	0,075	0,074	0,137
	0,03	0,032	0,027	0,046	0,03	0,052	0	0,084	0,081	0,128
	0,084	0,08	0,076	0,092	0,08	0,075	0,084	0	0,019	0,157
	0,082	0,075	0,075	0,092	0,076	0,074	0,081	0,019	0	0,15
	0,126	0,121	0,12	0,137	0,119	0,137	0,128	0,157	0,15	0
	0,046	0,047	0,049	0,05	0,046	0,078	0,055	0,097	0,092	0,13
	0,076	0,071	0,072	0,094	0,072	0,062	0,077	0,092	0,088	0,162
	0,034	0,032	0,033	0,046	0,029	0,059	0,03	0,087	0,083	0,123
	0,043	0,041	0,045	0,05	0,043	0,071	0,048	0,09	0,084	0,124
						0,076	0,052	0,094	0,089	0,128
	0,083	0,08	0,074	0,094	0,079	0,094	0,078	0,115	0,117	0,13
	0,071	0,066	0,071	0,08	0,069	0,081	0,069	0,094	0,083	0,143
	0,242	0,24	0,236	0,253	0,241	0,252	0,243	0,265	0,264	0,245
	0,072	0,067	0,06	0,079	0,069	0,072	0,068	0,087	0,097	0,148
	0,04	0,037	0,033	0,049	0,039	0,057	0,04	0,062	0,063	0,122
GPT004060	0,06	0,051	0,056	0,066	0,055	0,042	0,058	0,079	0,07	0,133
GPT004490	0,306	0,297	0,296	0,312	0,298	0,304	0,3	0,318	0,312	0,264
GPT004560	0,048	0,048	0,05	0,048	0,048	0,068	0,05	0,085	0,088	0,137
GPT004570	0,052	0,049	0,044	0,059	0,047	0,072	0,049	0,083	0,079	0,131
GPT004940	0,089	0,081	0,081	0,102	0,085	0,044	0,084	0,096	0,093	0,167
GPT005160	0,135	0,126	0,127	0,151	0,13	0,114	0,132	0,111	0,106	0,196
GPT005810	0,077	0,073	0,067	0,092	0,074	0,072	0,07	0,075	0,074	0,161
GPT006670	0,032	0,04	0,04	0,044	0,038	0,065	0,039	0,092	0,091	0,13
GPT006680	0,066	0,066	0,063	0,071	0,059	0,087	0,067	0,102	0,104	0,145
GPT006760	0,066	0,057	0,06	0,076	0,06	0,05	0,067	0,07	0,066	0,131
GPT007460	0,066	0,064	0,069	0,073	0,066	0,094	0,072	0,114	0,107	0,148
GPT007510	0,066	0,065	0,069	0,072	0,068	0,095	0,073	0,113	0,108	0,148
GPT008160	0,036	0,039	0,041	0,04	0,037	0,07	0,041	0,092	0,092	0,129
GPT008350	0,136	0,136	0,135	0,131	0,138	0,157	0,141	0,178	0,174	0,215
GPT008470	0,087	0,091	0,094	0,09	0,087	0,116	0,098	0,135	0,136	0,176
GPT008680	0,109	0,105	0,103	0,119	0,102	0,107	0,111	0,12	0,12	0,152
GPT008870	0,132	0,124	0,123	0,149	0,13	0,111	0,13	0,108	0,103	0,194
GPT009590	0,037	0,039	0,036	0,056	0,038	0,058	0,039	0,085	0,081	0,121
GPT009990	0,104	0,1	0,096	0,119	0,103	0,112	0,099	0,139	0,135	0,18
GPT010210	0,088	0,092	0,091	0,089	0,092	0,113	0,095	0,133	0,134	0,175
GPT010220	0,149	0,155	0,154	0,146	0,155	0,173	0,157	0,193	0,192	0,223
GPT010720	0,083	0,085	0,086	0,086	0,089	0,107	0,091	0,129	0,129	0,169
GPT010730	0,133	0,123	0,126	0,148	0,127	0,114	0,131	0,11	0,108	0,196
GPT012920	0,058	0,059	0,058	0,069	0,06	0,082	0,064	0,108	0,109	0,147
GPT014690	0,037	0,032	0,025	0,053	0,034	0,05	0,032	0,076	0,072	0,118

LiveSeeding

Practical Example

6. Select available data to the analysis (phenotype, genotype and MDS) and Intersect them

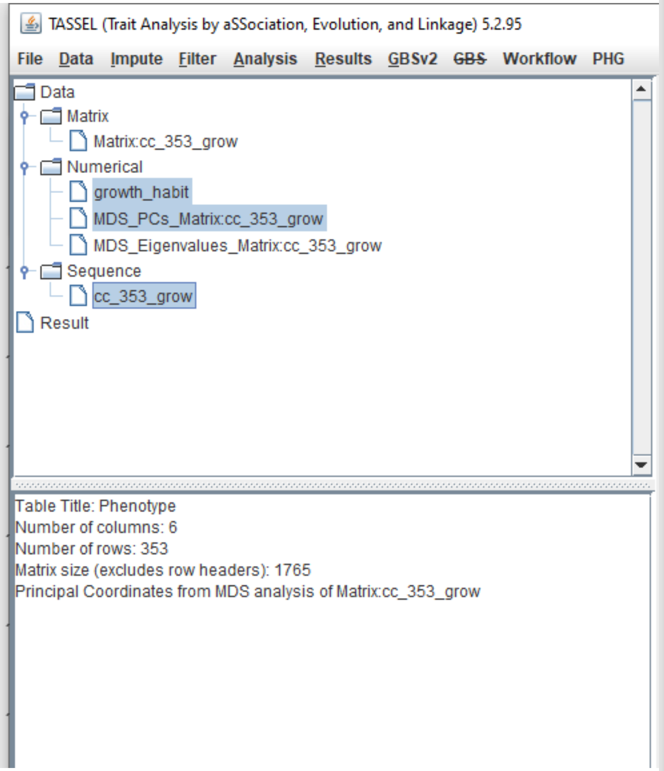
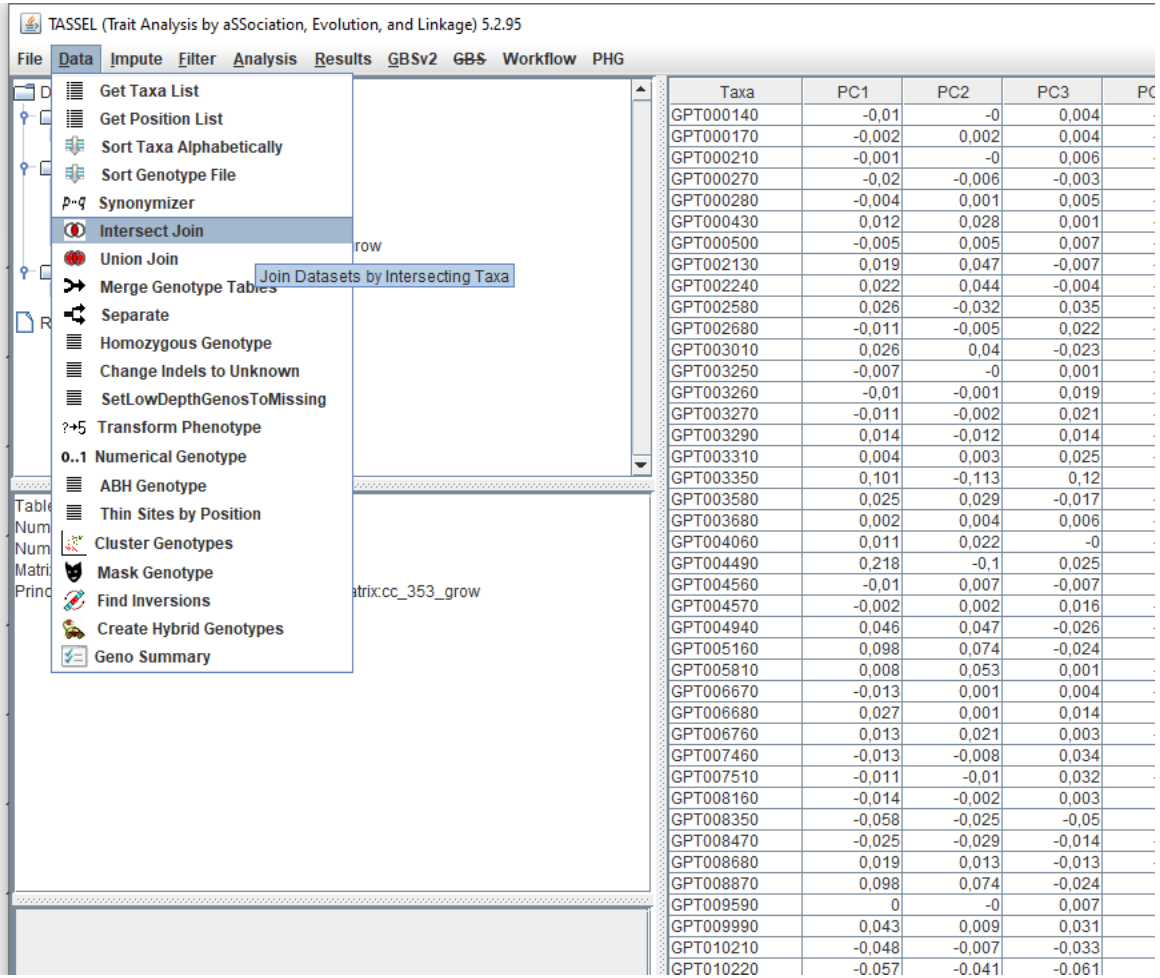


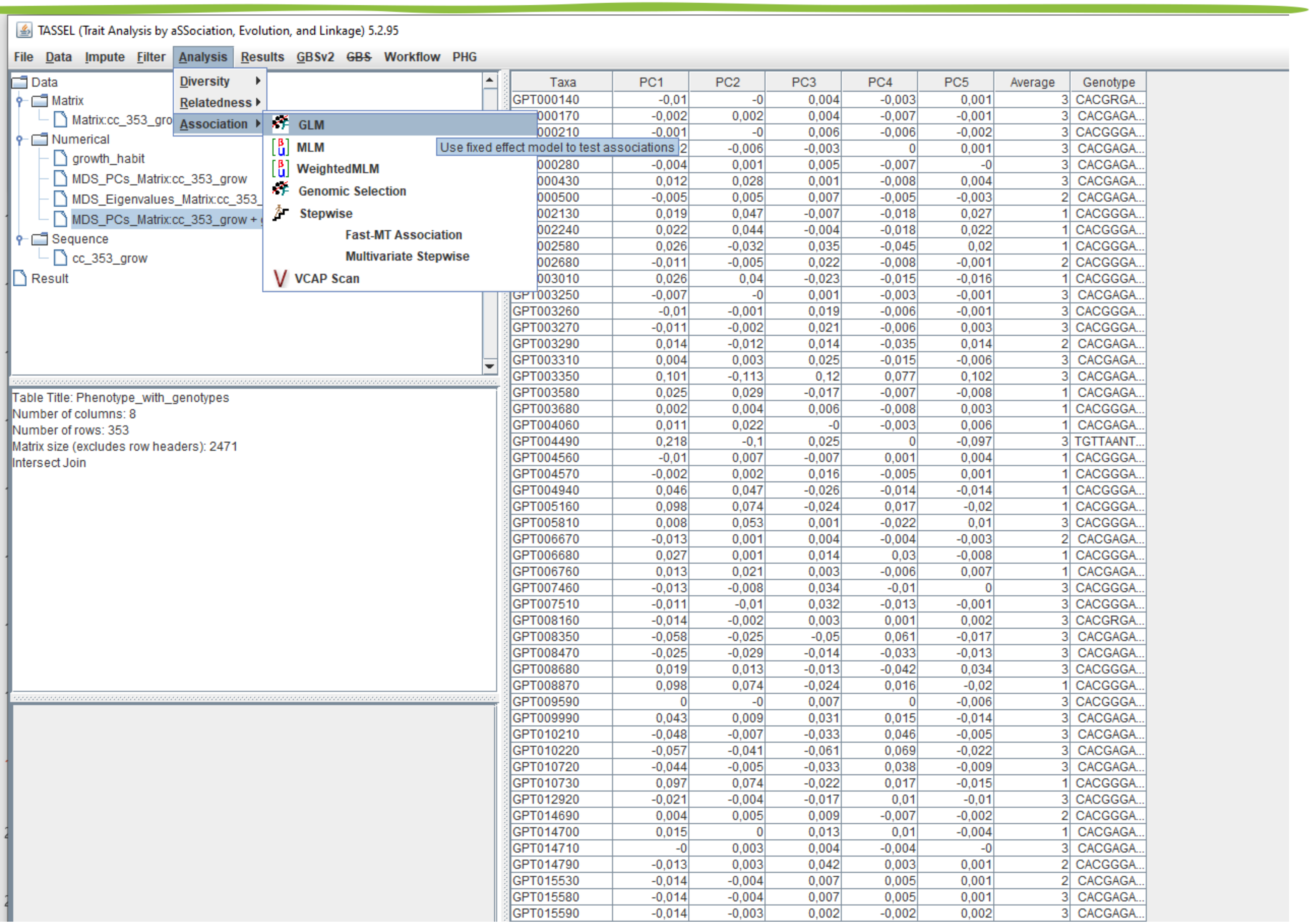
Table Title: Phenotype
Number of columns: 6
Number of rows: 353
Matrix size (excludes row headers): 1765
Principal Coordinates from MDS analysis of Matrix:cc_353_grow



Taxa	PC1	PC2	PC3
GPT000140	-0,01	-0	0,004
GPT000170	-0,002	0,002	0,004
GPT000210	-0,001	-0	0,006
GPT000270	-0,02	-0,006	-0,003
GPT000280	-0,004	0,001	0,005
GPT000430	0,012	0,028	0,001
GPT000500	-0,005	0,005	0,007
GPT002130	0,019	0,047	-0,007
GPT002240	0,022	0,044	-0,004
GPT002580	0,026	-0,032	0,035
GPT002680	-0,011	-0,005	0,022
GPT003010	0,026	0,04	-0,023
GPT003250	-0,007	-0	0,001
GPT003260	-0,01	-0,001	0,019
GPT003270	-0,011	-0,002	0,021
GPT003290	0,014	-0,012	0,014
GPT003310	0,004	0,003	0,025
GPT003350	0,101	-0,113	0,12
GPT003580	0,025	0,029	-0,017
GPT003680	0,002	0,004	0,006
GPT004060	0,011	0,022	-0
GPT004490	0,218	-0,1	0,025
GPT004560	-0,01	0,007	-0,007
GPT004570	-0,002	0,002	0,016
GPT004940	0,046	0,047	-0,026
GPT005160	0,098	0,074	-0,024
GPT005810	0,008	0,053	0,001
GPT006670	-0,013	0,001	0,004
GPT006680	0,027	0,001	0,014
GPT006760	0,013	0,021	0,003
GPT007460	-0,013	-0,008	0,034
GPT007510	-0,011	-0,01	0,032
GPT008160	-0,014	-0,002	0,003
GPT008350	-0,058	-0,025	-0,05
GPT008470	-0,025	-0,029	-0,014
GPT008680	0,019	0,013	-0,013
GPT008870	0,098	0,074	-0,024
GPT009590	0	-0	0,007
GPT009990	0,043	0,009	0,031
GPT010210	-0,048	-0,007	-0,033
GPT010220	-0,057	-0,041	-0,061

Practical Example

7. Association analysis with GLM



TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.95

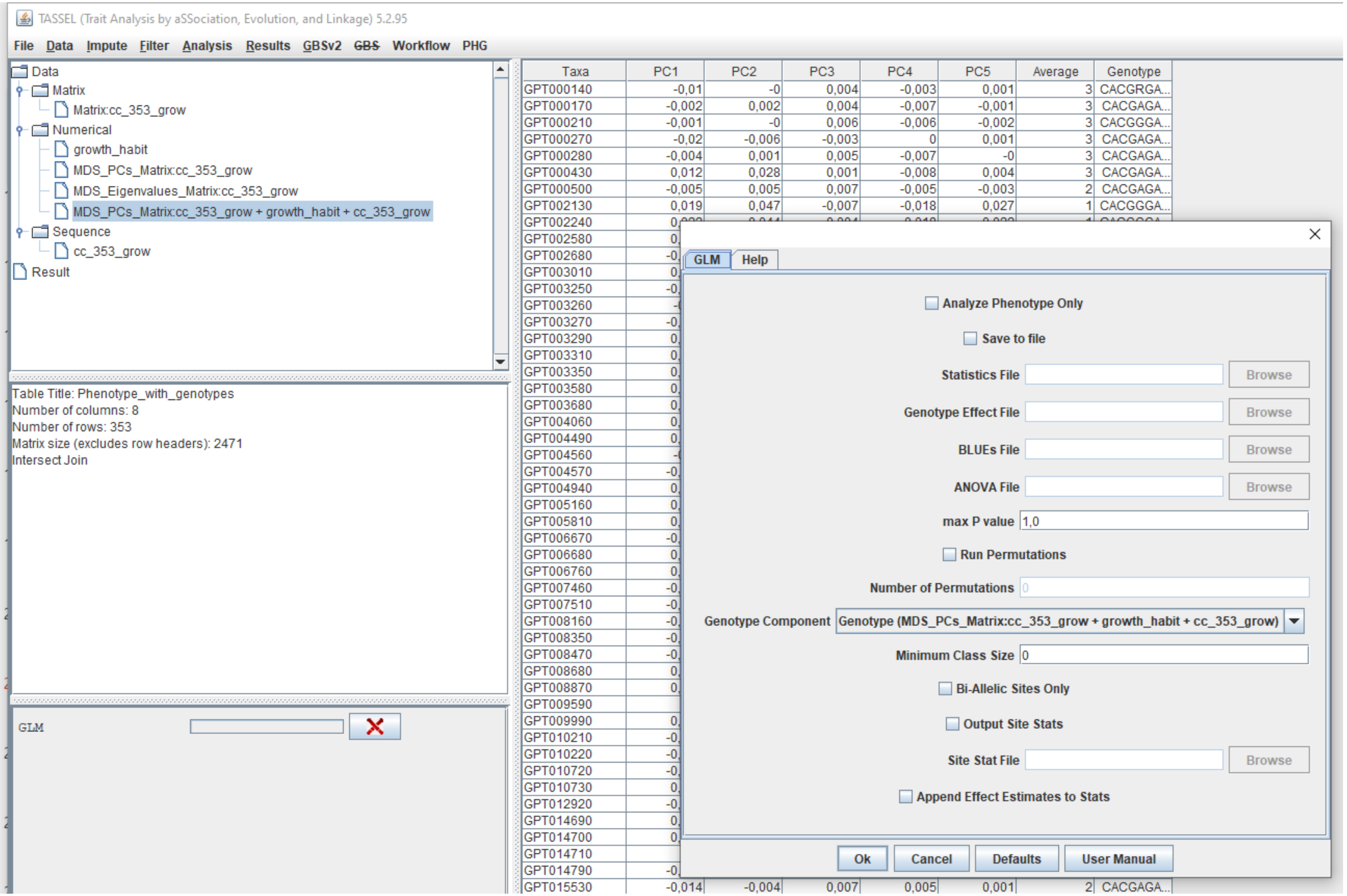
Use fixed effect model to test associations

Table Title: Phenotype_with_genotypes
Number of columns: 8
Number of rows: 353
Matrix size (excludes row headers): 2471
Intersect Join

Taxa	PC1	PC2	PC3	PC4	PC5	Average	Genotype
GPT000140	-0,01	-0	0,004	-0,003	0,001	3	CACGRGA...
000170	-0,002	0,002	0,004	-0,007	-0,001	3	CACGAGA...
000210	-0,001	-0	0,006	-0,006	-0,002	3	CACGGGA...
	2	-0,006	-0,003	0	0,001	3	CACGAGA...
000280	-0,004	0,001	0,005	-0,007	-0	3	CACGAGA...
000430	0,012	0,028	0,001	-0,008	0,004	3	CACGAGA...
000500	-0,005	0,005	0,007	-0,005	-0,003	2	CACGAGA...
002130	0,019	0,047	-0,007	-0,018	0,027	1	CACGGGA...
002240	0,022	0,044	-0,004	-0,018	0,022	1	CACGGGA...
002580	0,026	-0,032	0,035	-0,045	0,02	1	CACGGGA...
002680	-0,011	-0,005	0,022	-0,008	-0,001	2	CACGGGA...
003010	0,026	0,04	-0,023	-0,015	-0,016	1	CACGGGA...
GPT003250	-0,007	-0	0,001	-0,003	-0,001	3	CACGAGA...
GPT003260	-0,01	-0,001	0,019	-0,006	-0,001	3	CACGGGA...
GPT003270	-0,011	-0,002	0,021	-0,006	0,003	3	CACGGGA...
GPT003290	0,014	-0,012	0,014	-0,035	0,014	2	CACGAGA...
GPT003310	0,004	0,003	0,025	-0,015	-0,006	3	CACGAGA...
GPT003350	0,101	-0,113	0,12	0,077	0,102	3	CACGAGA...
GPT003580	0,025	0,029	-0,017	-0,007	-0,008	1	CACGAGA...
GPT003680	0,002	0,004	0,006	-0,008	0,003	1	CACGGGA...
GPT004060	0,011	0,022	-0	-0,003	0,006	1	CACGAGA...
GPT004490	0,218	-0,1	0,025	0	-0,097	3	TGTTAANT...
GPT004560	-0,01	0,007	-0,007	0,001	0,004	1	CACGGGA...
GPT004570	-0,002	0,002	0,016	-0,005	0,001	1	CACGGGA...
GPT004940	0,046	0,047	-0,026	-0,014	-0,014	1	CACGGGA...
GPT005160	0,098	0,074	-0,024	0,017	-0,02	1	CACGGGA...
GPT005810	0,008	0,053	0,001	-0,022	0,01	3	CACGGGA...
GPT006670	-0,013	0,001	0,004	-0,004	-0,003	2	CACGAGA...
GPT006680	0,027	0,001	0,014	0,03	-0,008	1	CACGGGA...
GPT006760	0,013	0,021	0,003	-0,006	0,007	1	CACGAGA...
GPT007460	-0,013	-0,008	0,034	-0,01	0	3	CACGGGA...
GPT007510	-0,011	-0,01	0,032	-0,013	-0,001	3	CACGGGA...
GPT008160	-0,014	-0,002	0,003	0,001	0,002	3	CACGRGA...
GPT008350	-0,058	-0,025	-0,05	0,061	-0,017	3	CACGAGA...
GPT008470	-0,025	-0,029	-0,014	-0,033	-0,013	3	CACGAGA...
GPT008680	0,019	0,013	-0,013	-0,042	0,034	3	CACGGGA...
GPT008870	0,098	0,074	-0,024	0,016	-0,02	1	CACGGGA...
GPT009590	0	-0	0,007	0	-0,006	3	CACGGGA...
GPT009990	0,043	0,009	0,031	0,015	-0,014	3	CACGAGA...
GPT010210	-0,048	-0,007	-0,033	0,046	-0,005	3	CACGAGA...
GPT010220	-0,057	-0,041	-0,061	0,069	-0,022	3	CACGAGA...
GPT010720	-0,044	-0,005	-0,033	0,038	-0,009	3	CACGAGA...
GPT010730	0,097	0,074	-0,022	0,017	-0,015	1	CACGGGA...
GPT012920	-0,021	-0,004	-0,017	0,01	-0,01	3	CACGGGA...
GPT014690	0,004	0,005	0,009	-0,007	-0,002	2	CACGGGA...
GPT014700	0,015	0	0,013	0,01	-0,004	1	CACGAGA...
GPT014710	-0	0,003	0,004	-0,004	-0	3	CACGAGA...
GPT014790	-0,013	0,003	0,042	0,003	0,001	2	CACGGGA...
GPT015530	-0,014	-0,004	0,007	0,005	0,001	2	CACGAGA...
GPT015580	-0,014	-0,004	0,007	0,005	0,001	3	CACGAGA...
GPT015590	-0,014	-0,003	0,002	-0,002	0,002	3	CACGAGA...

Practical Example

7. Association analysis with GLM



TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) 5.2.95

- Data
  - Matrix
    - Matrix:cc_353_grow
  - Numerical
    - growth_habit
    - MDS_PCs_Matrix:cc_353_grow
    - MDS_Eigenvalues_Matrix:cc_353_grow
    - MDS_PCs_Matrix:cc_353_grow + growth_habit + cc_353_grow
  - Sequence
    - cc_353_grow
- Result

Table Title: Phenotype_with_genotypes
Number of columns: 8
Number of rows: 353
Matrix size (excludes row headers): 2471
Intersect Join

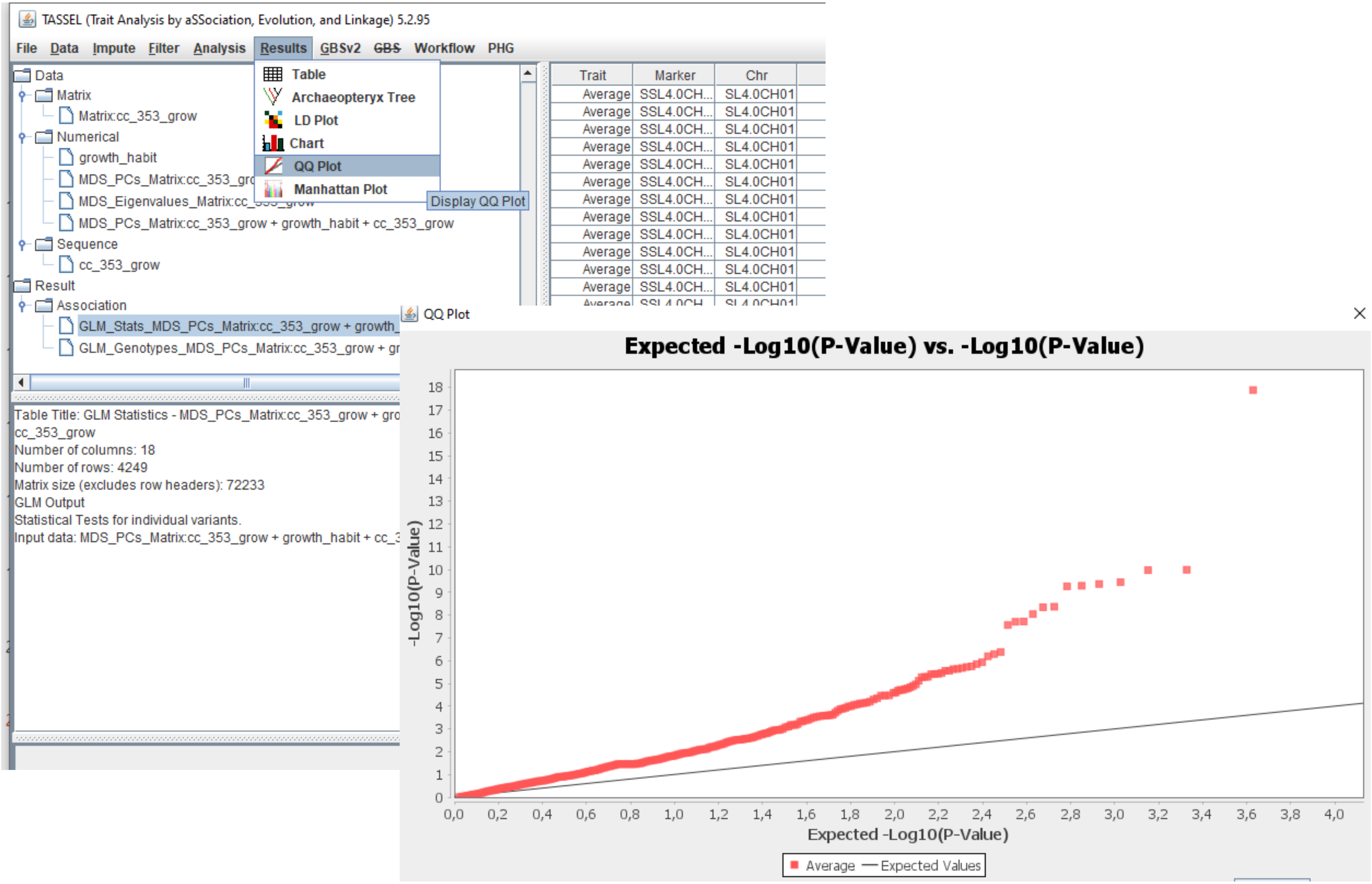
Taxa	PC1	PC2	PC3	PC4	PC5	Average	Genotype
GPT000140	-0,01	-0	0,004	-0,003	0,001	3	CACGRGA...
GPT000170	-0,002	0,002	0,004	-0,007	-0,001	3	CACGAGA...
GPT000210	-0,001	-0	0,006	-0,006	-0,002	3	CACGGGA...
GPT000270	-0,02	-0,006	-0,003	0	0,001	3	CACGAGA...
GPT000280	-0,004	0,001	0,005	-0,007	-0	3	CACGAGA...
GPT000430	0,012	0,028	0,001	-0,008	0,004	3	CACGAGA...
GPT000500	-0,005	0,005	0,007	-0,005	-0,003	2	CACGAGA...
GPT002130	0,019	0,047	-0,007	-0,018	0,027	1	CACGGGA...
GPT002240							
GPT002580							
GPT002680							
GPT003010							
GPT003250							
GPT003260							
GPT003270							
GPT003290							
GPT003310							
GPT003350							
GPT003580							
GPT003680							
GPT004060							
GPT004490							
GPT004560							
GPT004570							
GPT004940							
GPT005160							
GPT005810							
GPT006670							
GPT006680							
GPT006760							
GPT007460							
GPT007510							
GPT008160							
GPT008350							
GPT008470							
GPT008680							
GPT008870							
GPT009590							
GPT009990							
GPT010210							
GPT010220							
GPT010720							
GPT010730							
GPT012920							
GPT014690							
GPT014700							
GPT014710							
GPT014790							
GPT015530	-0,014	-0,004	0,007	0,005	0,001	2	CACGAGA...

☐ Analyze Phenotype Only
☐ Save to file
Statistics File: [Browse]
Genotype Effect File: [Browse]
BLUEs File: [Browse]
ANOVA File: [Browse]
max P value: 1,0
☐ Run Permutations
Number of Permutations: 0
Genotype Component: Genotype (MDS_PCs_Matrix:cc_353_grow + growth_habit + cc_353_grow)
Minimum Class Size: 0
☐ Bi-Allelic Sites Only
☐ Output Site Stats
Site Stat File: [Browse]
☐ Append Effect Estimates to Stats

Practical Example

8. Represent QQPlot

The distribution of the markers is according to the expected if any QTL is associated with the trait



Trait	Marker	Chr
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01
Average	SSL4.0CH...	SL4.0CH01

```
Table Title: GLM Statistics - MDS_PCs_Matrix:cc_353_grow + gro
cc_353_grow
Number of columns: 18
Number of rows: 4249
Matrix size (excludes row headers): 72233
GLM Output
Statistical Tests for individual variants.
Input data: MDS_PCs_Matrix:cc_353_grow + growth_habit + cc_3
```

Practical Example

9. Represent Manhattan plot

The position of the markers with their association to the trait

Trait	Marker
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...
Average	SSL4.0CH...

Table Title: GLM Statistics - MDS_PCs_Matrix:cc_353_grow + growth_habit + cc_353_grow
Number of columns: 18
Number of rows: 4249
Matrix size (excludes row headers): 72233
GLM Output
Statistical Tests for individual variants.
Input data: MDS_PCs_Matrix:cc_353_grow + growth_habit + cc_353_grow

LiveSeeding

Next steps

a) Select the most significant QTLs and establish a region upstream and downstream

b) Use available genomic information of the species to check annotated genes in this region

c) Analyze the function of these genes to know if they could be involved in the expression of the trait

d) Search information in other species and in other databases

Practical Example

GWAS analysis in a collection of Solanum lycopersicum

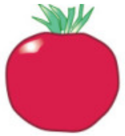
Fruit shape in tomato

Flattened

1

Slightly flattened

2

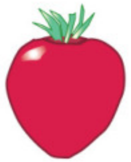


Rounded

3

High-rounded

4



Heart-shaped

5

Cylindrical

6

Pyriform

7

Ellipsoid

8

Practical Example

1. Open phenotypic data file

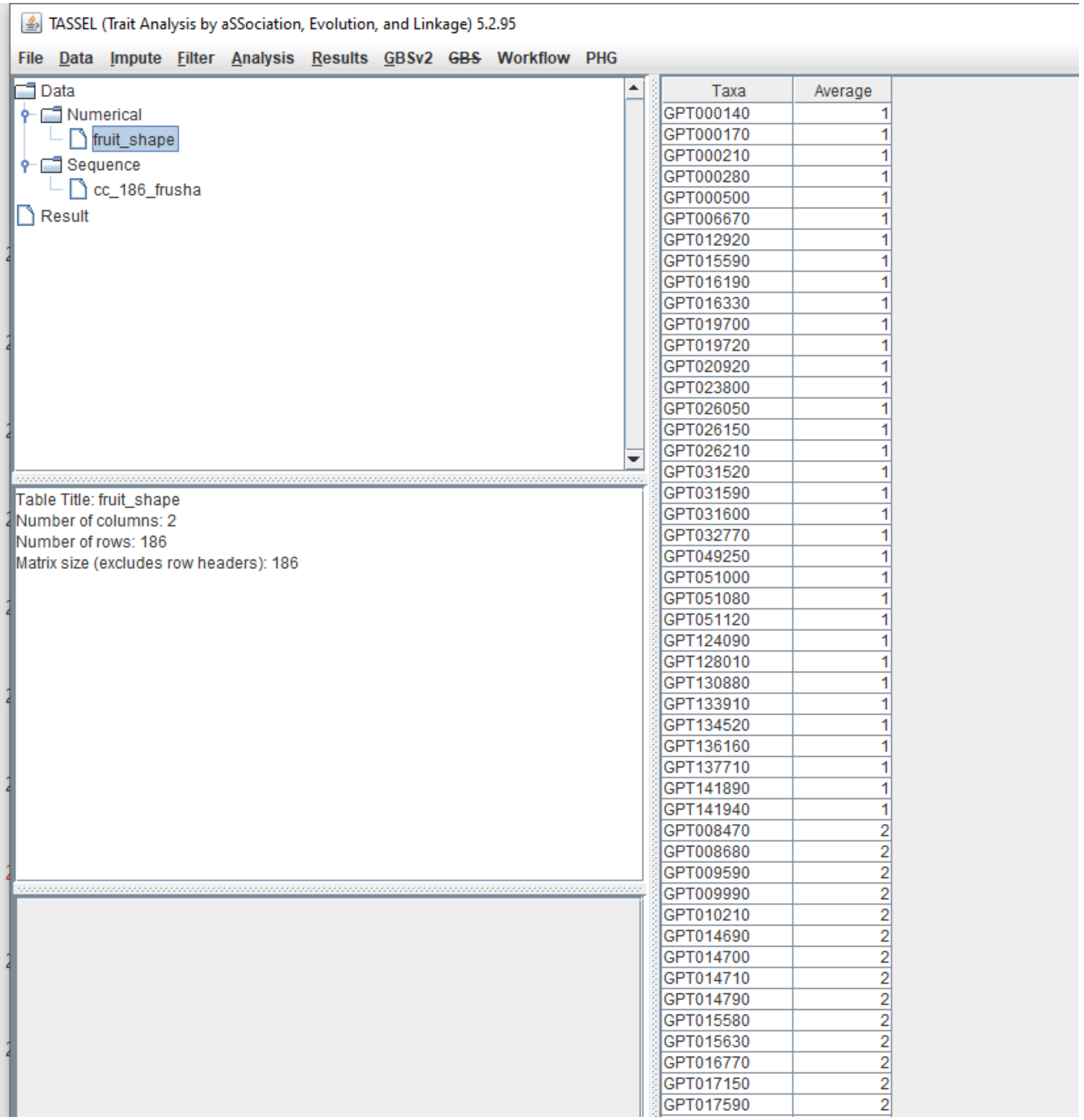
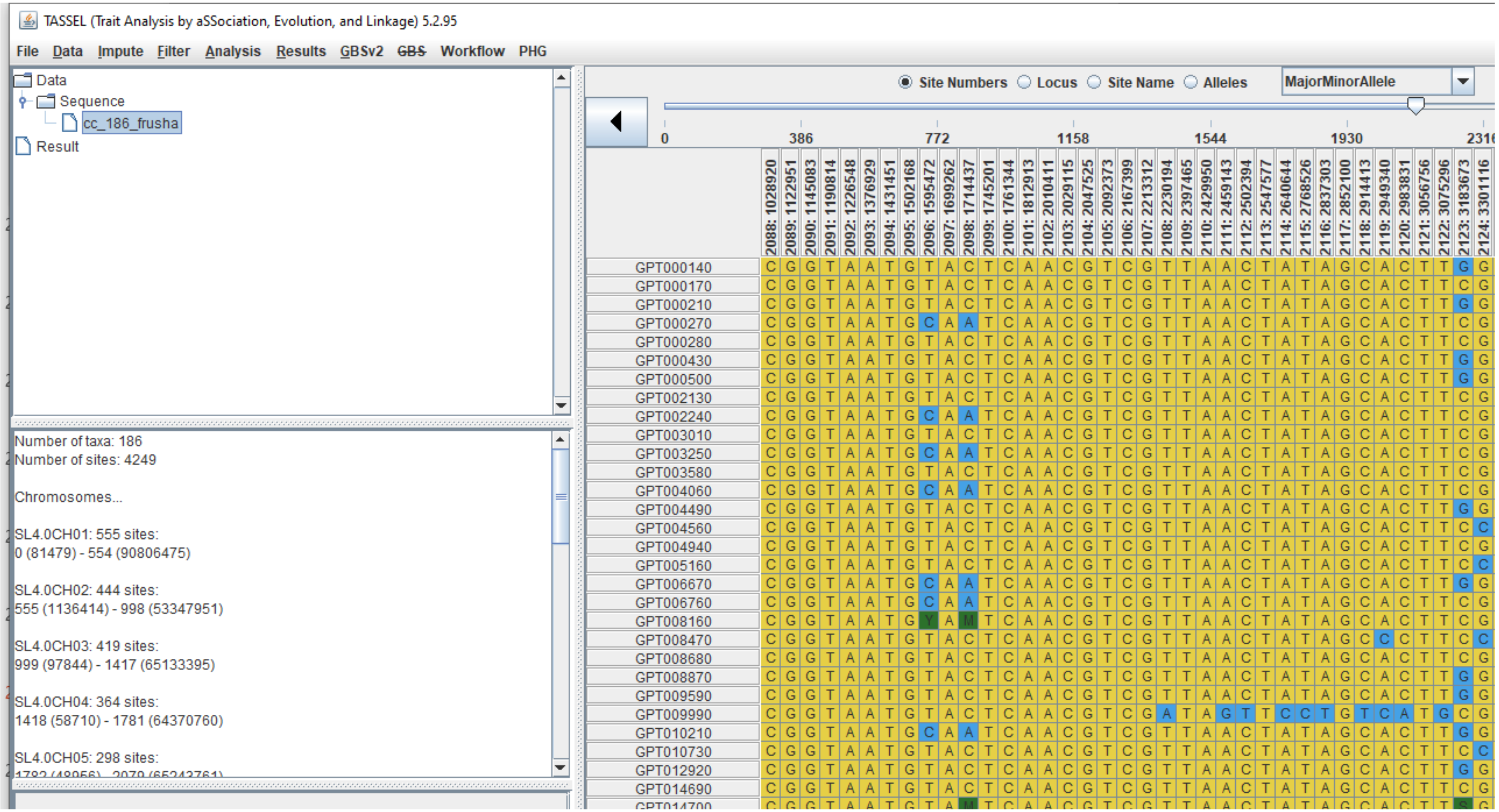


Table Title: fruit_shape
Number of columns: 2
Number of rows: 186
Matrix size (excludes row headers): 186

Taxa	Average
GPT000140	1
GPT000170	1
GPT000210	1
GPT000280	1
GPT000500	1
GPT006670	1
GPT012920	1
GPT015590	1
GPT016190	1
GPT016330	1
GPT019700	1
GPT019720	1
GPT020920	1
GPT023800	1
GPT026050	1
GPT026150	1
GPT026210	1
GPT031520	1
GPT031590	1
GPT031600	1
GPT032770	1
GPT049250	1
GPT051000	1
GPT051080	1
GPT051120	1
GPT124090	1
GPT128010	1
GPT130880	1
GPT133910	1
GPT134520	1
GPT136160	1
GPT137710	1
GPT141890	1
GPT141940	1
GPT008470	2
GPT008680	2
GPT009590	2
GPT009990	2
GPT010210	2
GPT014690	2
GPT014700	2
GPT014710	2
GPT014790	2
GPT015580	2
GPT015630	2
GPT016770	2
GPT017150	2
GPT017590	2

Practical Example

2. Open GBS data file

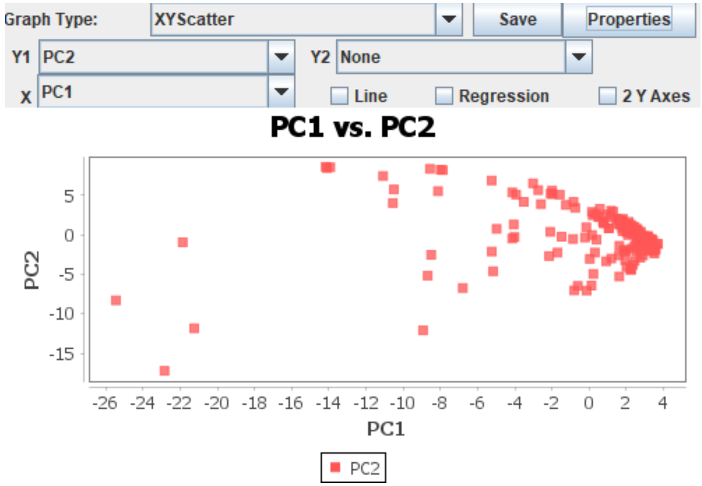


```
Number of taxa: 186
Number of sites: 4249

Chromosomes...

SL4.0CH01: 555 sites:
0 (81479) - 554 (90806475)

SL4.0CH02: 444 sites:
555 (1136414) - 998 (53347951)

SL4.0CH03: 419 sites:
999 (97844) - 1417 (65133395)

SL4.0CH04: 364 sites:
1418 (58710) - 1781 (64370760)

SL4.0CH05: 298 sites:
```

	2088: 1029920	2089: 1122951	2090: 1145083	2091: 1190814	2092: 1226548	2093: 1376929	2094: 1431451	2095: 1502168	2096: 1595472	2097: 1699262	2098: 1714437	2099: 1745201	2100: 1761344	2101: 1812913	2102: 2010411	2103: 2029115	2104: 2047525	2105: 2092373	2106: 2167399	2107: 2213312	2108: 2230194	2109: 2397465	2110: 2429950	2111: 2459143	2112: 2502394	2113: 2547577	2114: 2640644	2115: 2768526	2116: 2837303	2117: 2852100	2118: 2914413	2119: 2949340	2120: 2983831	2121: 3056756	2122: 3075296	2123: 3183673	2124: 3301116
GPT000140	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT000170	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT000210	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT000270	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT000280	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT000430	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT000500	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT002130	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT002240	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT003010	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT003250	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT003580	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT004060	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT004490	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT004560	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	C
GPT004940	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT005160	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	C
GPT006670	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT006760	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT008160	C	G	G	T	A	A	T	G	Y	A	M	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT008470	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	C	C	T	T	C	C
GPT008680	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT008870	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT009590	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT009990	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	A	T	A	G	T	T	C	C	T	G	T	C	A	T	G	C	G
GPT010210	C	G	G	T	A	A	T	G	C	A	A	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT010730	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	C
GPT012920	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	G	G
GPT014690	C	G	G	T	A	A	T	G	T	A	C	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	C	G
GPT014700	C	G	G	T	A	A	T	G	T	A	M	T	C	A	A	C	G	T	C	G	T	T	A	A	C	T	A	T	A	G	C	A	C	T	T	S	G

Practical Example

3. PCA



Graph Type: XYScatter

Y1 PC2

Y2 None

X PC1

☐ Line ☐ Regression ☐ 2 Y Axes

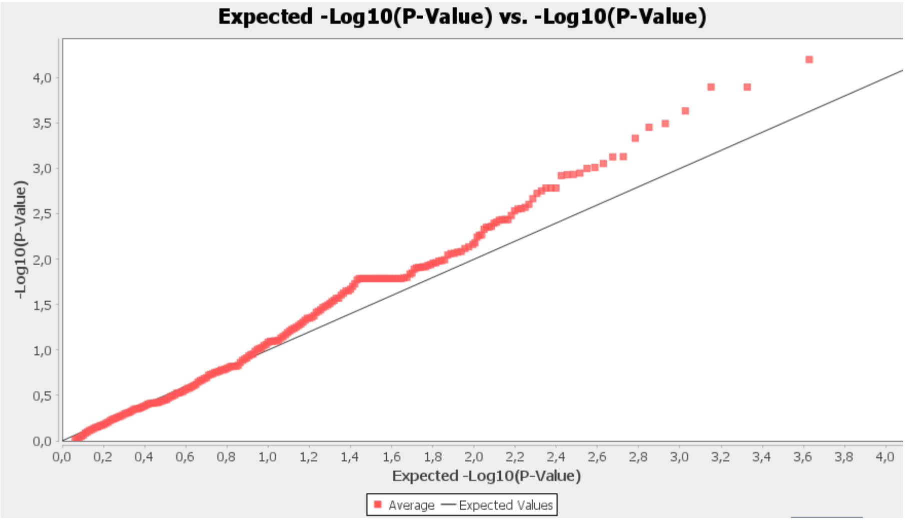
4. Distance Matrix

5. MDS

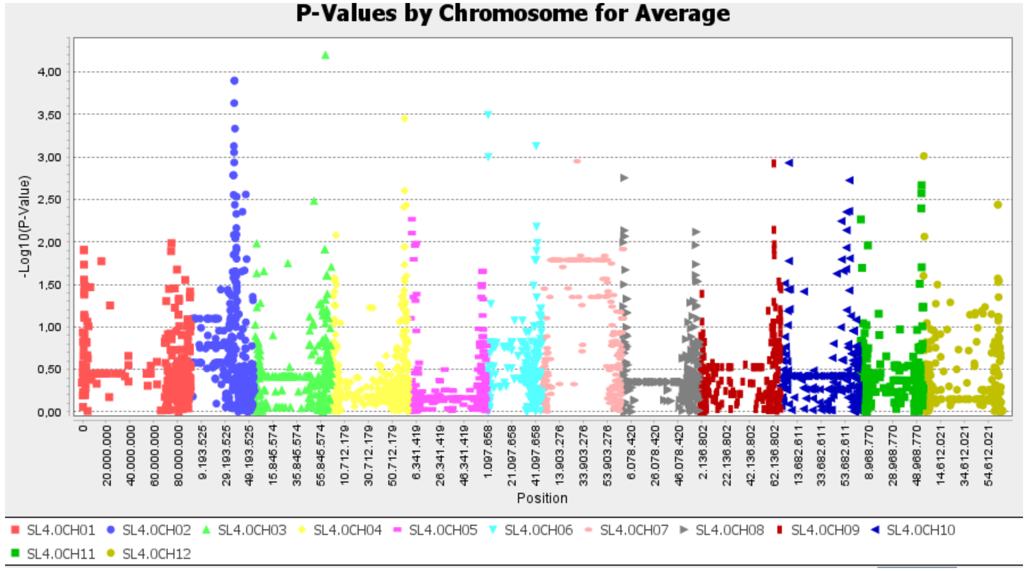
6. Intersect analysis information

7. Association analysis

8. QQPlot



9. Manhattan plot



LiveSeeding

Fast Quiz

- ☐ How can the results be affected by the existence of two subpopulations?
- ☐ Which could be the reason why in this case we do not obtain any associated QTL in the second case? Be aware of the data.

Send your answers to the chat

LiveSeeding

Take Home Messages

- ✓ GWAS is used to associate traits with genomic regions
- ✓ We need a non-structured and variable study population
- ✓ Specialized programs are required to intersect phenotypic and genotypic data
- ✓ Significant markers need to be analyzed to find the candidate genes responsible for the trait

Proposed Homework

Try to perform your own analysis with Tassel5 and the data in additional resources (genotyping_data.vcf, phenotyping_data.txt).

Send your results with screenshots of the QQplots and Manhattan plots and explain the obtained results.

Send your results to neuoral@doctor.upv.es, adrodbur@upvnet.upv.es, petra.jelincic@ips-konzalting.hr

Additional Resources

Tassel5

- https://www.maizegenetics.net/tassel
- https://www.youtube.com/watch?v=73hD5n-OOJc&list=PLdf-U83sN48OPSQ-DqL3AgqKOvu8nUoMT

Novel High throughput phenotyping platforms:

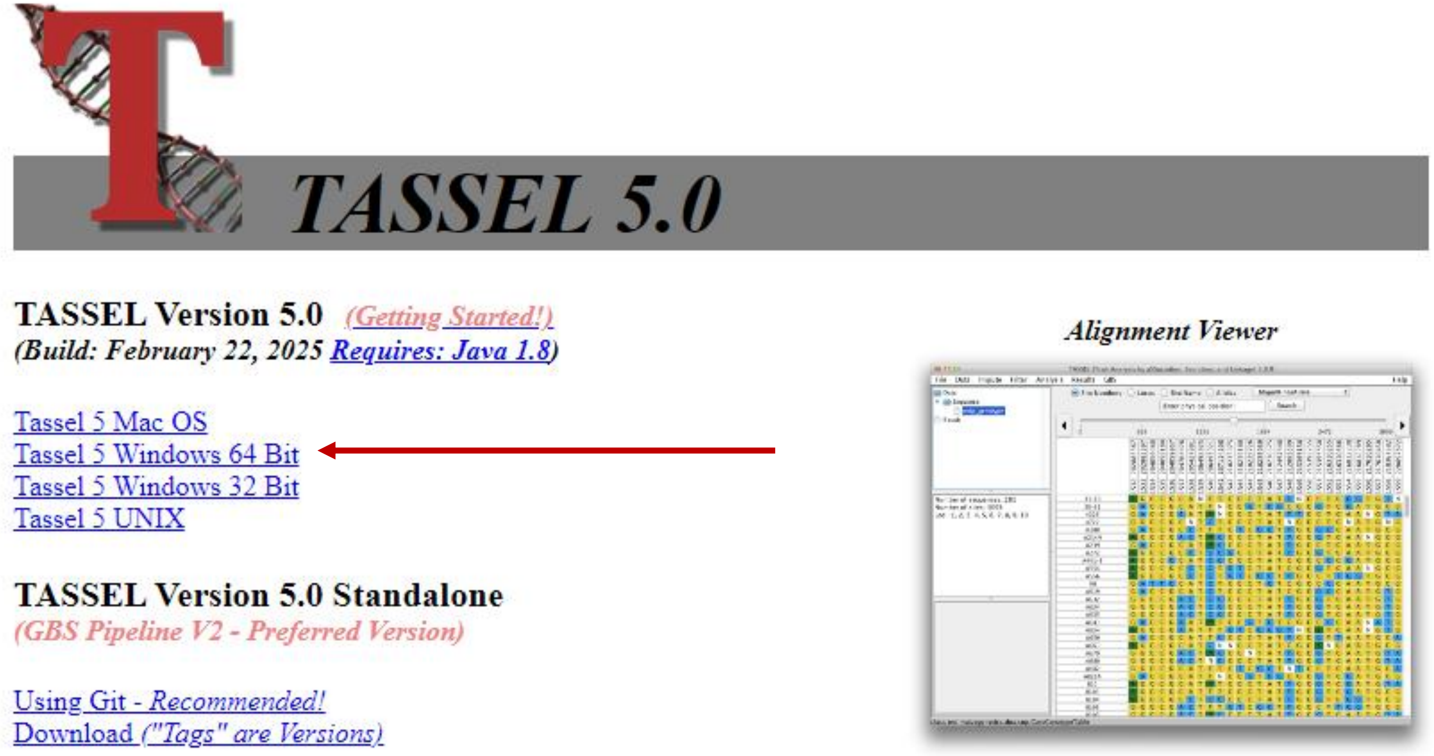
- https://www.youtube.com/watch?v=zw02B_oWL9k

Data for homework and presentation:

- https://github.com/jmartinezpoq/LiveSeeding_Module4/tree/master/Unit1_PracticalExercices

Additional Resources

How to Install Tassel5



TASSEL 5.0

TASSEL Version 5.0 *(Getting Started!)*
(Build: February 22, 2025 Requires: Java 1.8)

- Tassel 5 Mac OS
- Tassel 5 Windows 64 Bit
- Tassel 5 Windows 32 Bit
- Tassel 5 UNIX

TASSEL Version 5.0 Standalone
(GBS Pipeline V2 - Preferred Version)

- Using Git - *Recommended!*
- Download *("Tags" are Versions)*

LiveSeeding

SEMENTES VIVAS

Rete Semi Rurali

ARCOIRIS

bingenheimer saatgut

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

AEDIT SRL

PIN POLSKA IZBA NASIENNA

FiBL Europe

Kmetijski inštitut Slovenije

ÖMKi

INRAE la science pour la vie, l'humain, la terre

ΑΙΓΙΛΟΠΑΣ AEGILOPS

GEVES Expertise & Performance

phr POZNAŃSKA HODOWLA ROŚLIN

FEDERBIO

FiBL Switzerland

itab l'Institut de l'agriculture et de l'alimentation biologiques

IFOAM

KWS

Dottenfelderhof

FUNDULEA

LA UNIÓ DE LLAURADORS I RAMADERS

RED DE MUNICIPIOS POR LA AGROECOLOGÍA

IPS KONZALTING

ORGANIC RESEARCH CENTRE

crea Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

LET'S LIBERATE DIVERSITY!

Ramiro Arnedo semillas

organic arable

FiBL Deutschland

WAGENINGEN UNIVERSITY & RESEARCH

Louis Bolk Instituut

RGA®

Politécnico de Coimbra

pro specie rara

LiveSeeding

Co-funded by the European Union

Funded by the European Union, the Swiss State Secretariat for Education, Research and Innovation (SERI) and UK Research and Innovation (UKRI). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA, nor SERI or UKRI.

UK Research and Innovation

LiveSeeding