Selective Mapping

Jose Luis Zambrano Mendoza

Selective Mapping

Content

- What is selective mapping, and why we would want to use it.
- Explanation of how to do it using MAPPOP
- Provide a case study based on maize virus resistance.

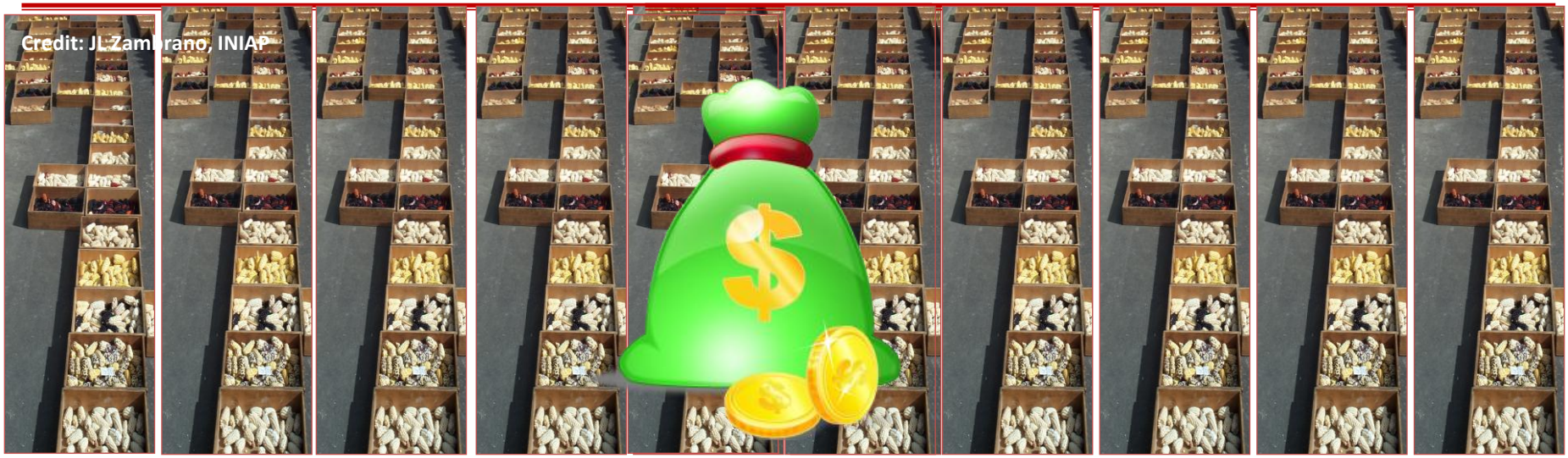
Selective Mapping

What is selective mapping ?

- ✓An approach that has been used mainly for Association and QTL mapping to reduce the number of individuals to be genotyped and phenotyped from a large population.
- ✓The lines in a mapping population can be defined as random samples from an infinitely large population.

Selective Mapping

Why, When ?



Can we save resources without losing experimental or selection power by concentrating on the most informative individuals in a population?

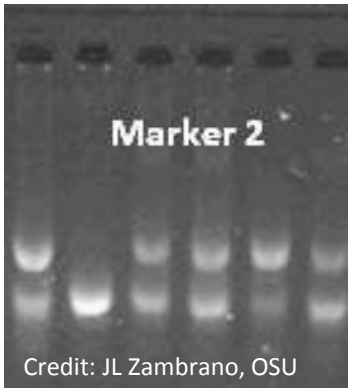
Selective Mapping

Choose individuals based on phenotype

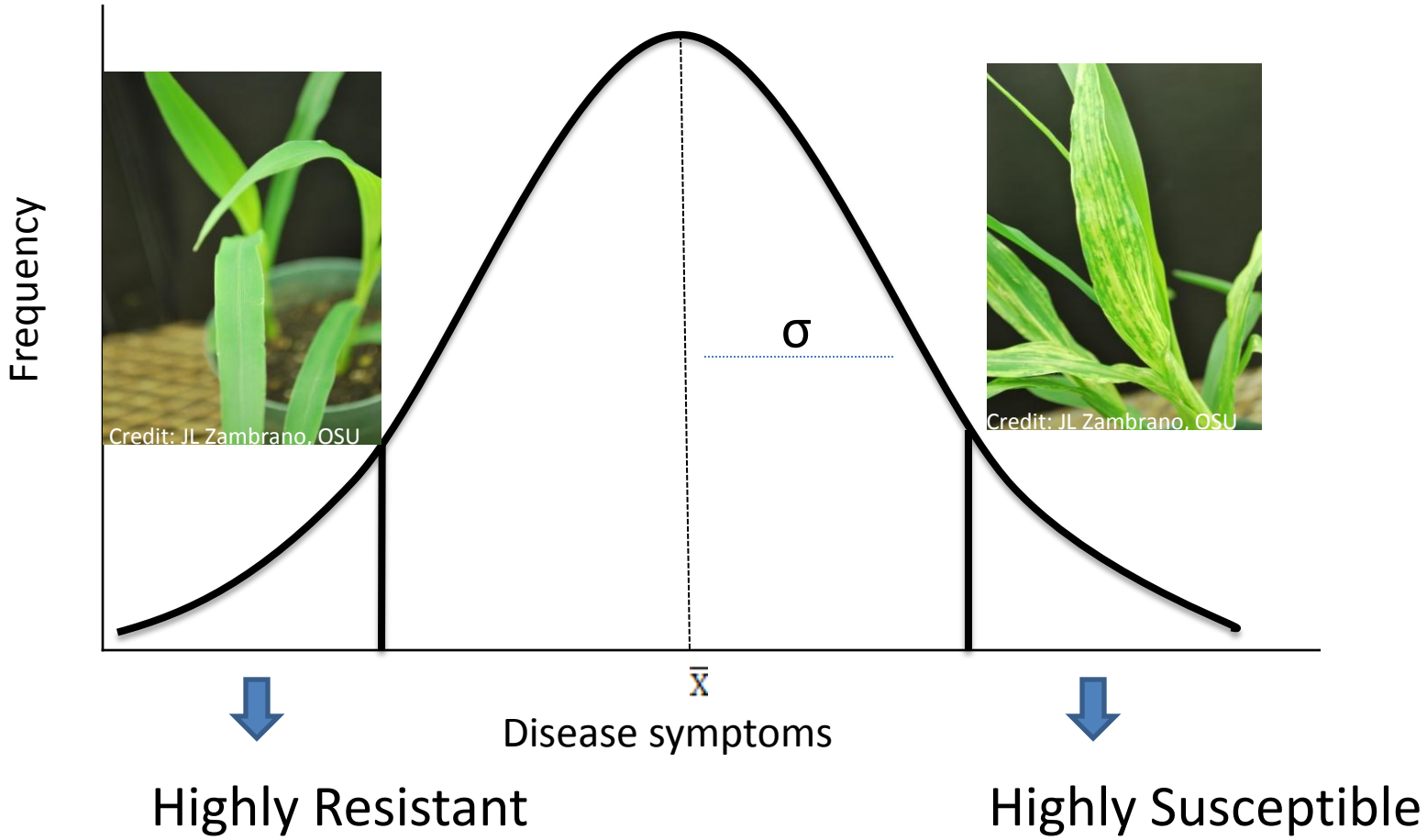
- We need Quantitative data (trait of interest)
- Excel

Choose individuals based on genotype

- We need Qualitative data (genetic markers)
- MapPop v. 1.0

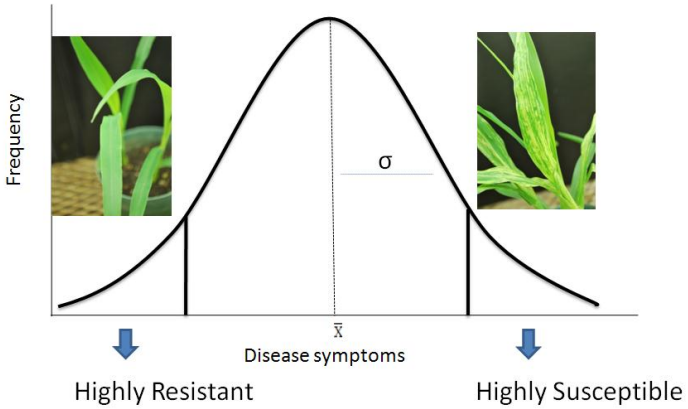


Selection Based on Phenotype



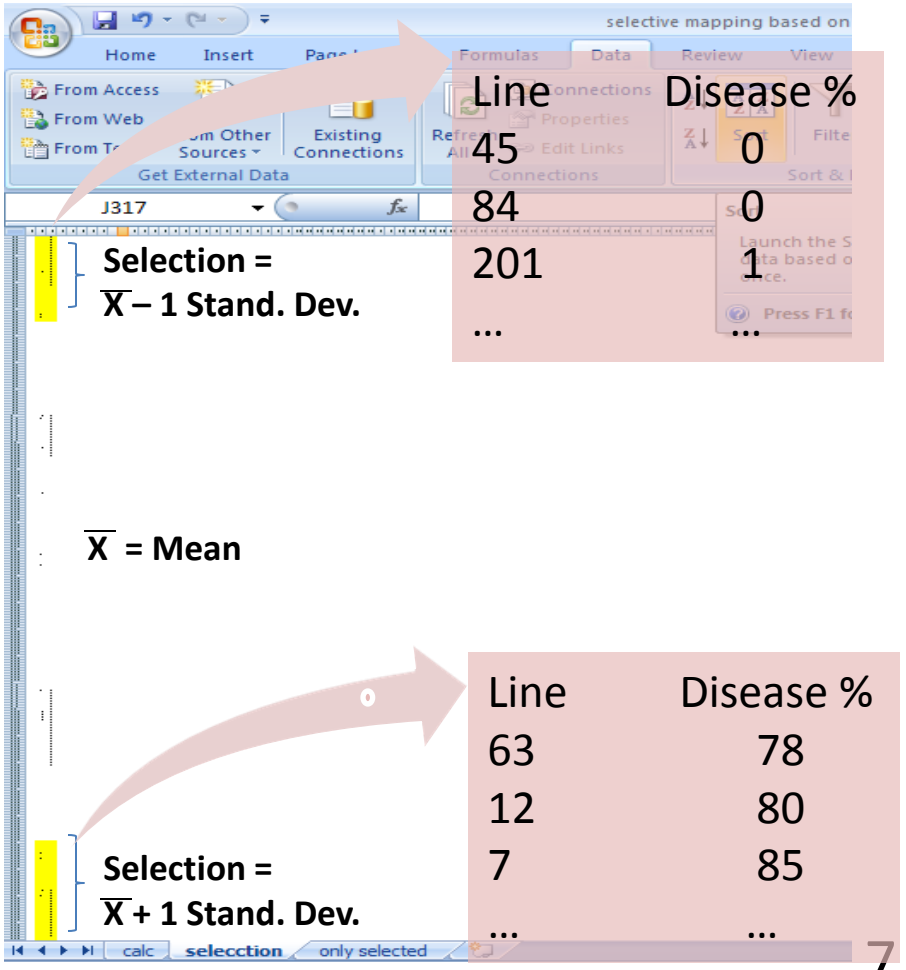
Selection Based on Phenotype

How to make selections?



- Be sure that your data is normally distributed
- Apply a measure of strength of selection, for instance: 1 Standard Deviation

Excel, Sort your data



Line	Disease %
45	0
84	0
201	1
…	…

Line	Disease %
63	78
12	80
7	85
…	…

Selection Based on Phenotype

- Individuals in extreme high and low range of distribution tend to have accumulated more positive and negative alleles

AABBCCDD… **aabbccdd…**

- More information: Lander and Bostein, 1989; Gallais et al., 2007; Sun et al., 2010)

Selection Based on Genotype (Selective Mapping)

- This method concentrates on:

    1. Individuals with maximum information (recombinant individuals)
    2. Different combinations of individuals with more desirable combinations of breakpoints

- MapPop provides a tool to select the most informative individuals

- This program quickly and consistently finds much-reduced samples with map resolution approaching that of the larger populations from which they are derived (Vision et al., 2000)

Selection Based on Genotype (Selective Mapping)

Before we use MapPop, we need some genotypic information for the entire population to feed into the program.

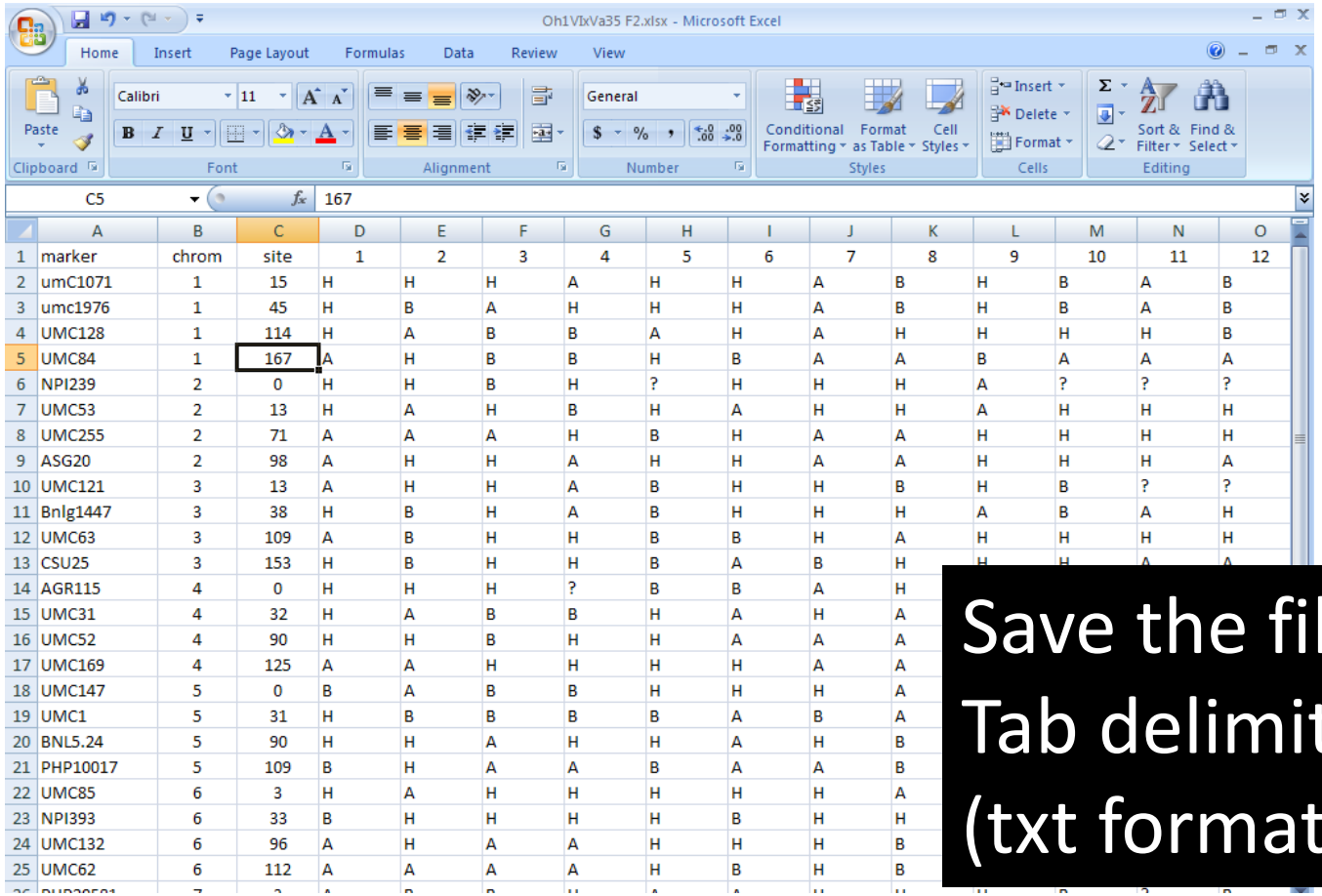
- Usually genetic information from a couple of markers from each chromosome arm is enough to build framework data.

- Markers and their genetic positions can be found in databases or consensus genetic maps. (http://www.maizegdb.org/map.php#)

Selection Based on Genotype Steps (Selective Mapping)

1. Build the file of the marker framework data as shown

First, be sure that marker segregation fits your expectation according to population type



Oh1VIxVa35 F2.xlsx - Microsoft Excel

C5: 167

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	marker	chrom	site	1	2	3	4	5	6	7	8	9	10	11	12
2	umC1071	1	15	H	H	H	A	H	H	A	B	H	B	A	B
3	umc1976	1	45	H	B	A	H	H	H	A	B	H	B	A	B
4	UMC128	1	114	H	A	B	B	A	H	A	H	H	H	H	B
5	UMC84	1	167	A	H	B	B	H	B	A	A	B	A	A	A
6	NPI239	2	0	H	H	B	H	?	H	H	H	A	?	?	?
7	UMC53	2	13	H	A	H	B	H	A	H	H	A	H	H	H
8	UMC255	2	71	A	A	A	H	B	H	A	A	H	H	H	H
9	ASG20	2	98	A	H	H	A	H	H	A	A	H	H	H	A
10	UMC121	3	13	A	H	H	A	B	H	H	B	H	B	?	?
11	Bnlg1447	3	38	H	B	H	A	B	H	H	H	A	B	A	H
12	UMC63	3	109	A	B	H	H	B	B	H	A	H	H	H	H
13	CSU25	3	153	H	B	H	H	B	A	B	H	H	H	A	A
14	AGR115	4	0	H	H	H	?	B	B	A	H				
15	UMC31	4	32	H	A	B	B	H	A	H	A				
16	UMC52	4	90	H	H	B	H	H	A	A	A				
17	UMC169	4	125	A	A	H	H	H	H	A	A				
18	UMC147	5	0	B	A	B	B	H	H	H	A				
19	UMC1	5	31	H	B	B	B	B	A	B	A				
20	BNL5.24	5	90	H	H	A	H	H	A	H	B				
21	PHP10017	5	109	B	H	A	A	B	A	A	B				
22	UMC85	6	3	H	A	H	H	H	H	H	A				
23	NPI393	6	33	B	H	H	H	H	B	H	H				
24	UMC132	6	96	A	H	A	A	H	H	H	B				
25	UMC62	6	112	A	A	A	A	H	B	H	B				

Columns = Line number

Rows = Markers

Parent 1 = A

Parent 2 = B

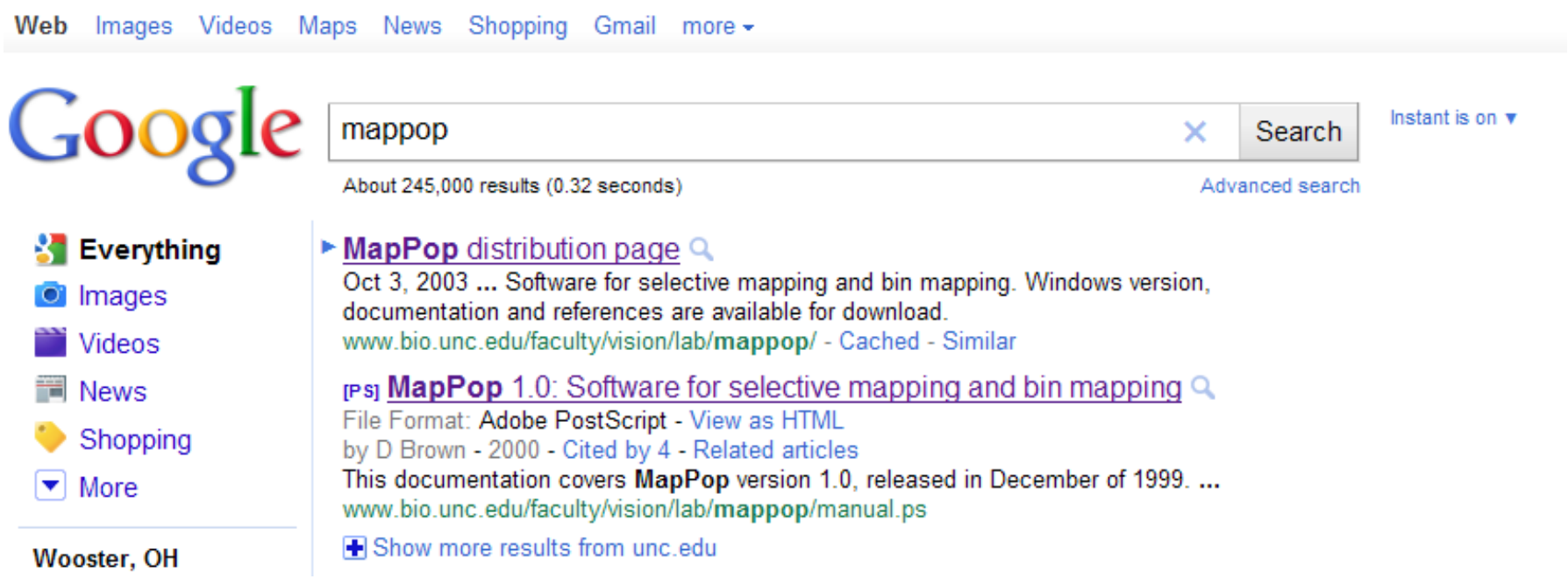
Hetero = H

Missing = ?

Save the file as Tab delimited (txt format)

Based on Genotype Steps (Selective Mapping)

2. Where do we find and how do we install MapPop ?



Based on Genotype Steps (Selective Mapping)

2. Where to find and how to install MapPop

http://www.bio.unc.edu/faculty/vision/lab/mappop/

MapPop

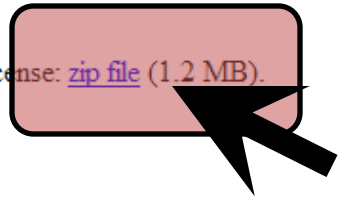
Software for selective mapping and bin mapping

This is the main source of MapPop, software for choosing good samples from mapping populations and for locating new markers on pre-existing maps.

If you download MapPop, please inform us by email so that we can alert users as updates occur.

Available on this page are:

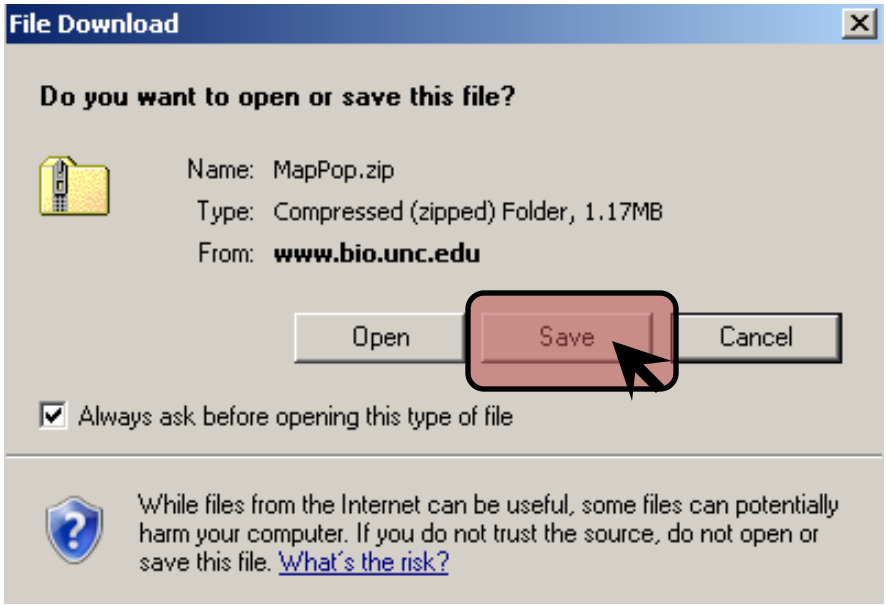
- MapPop executables (for Windows) and its documentation, along with sample files and license: zip file (1.2 MB).
- MapPop documentation a la carte: .ps version (235 k), .pdf version (110 k).
- Sample data files in text format: framework data (3 k) and new marker data (5 k)
- The MapPop license: text file (2 k).
- Papers which describe the underlying ideas in MapPop:
  - Vision TJ, Brown DG, Shmoys DB, Durrett RT, Tanksley SD (2000) nbsp; *Selective Mapping: A strategy for optimizing the construction of high-density linkage maps*. Genetics 155, 407-420.
  - Brown DG, Vision TJ, Tanksley SD (2000) *Selective mapping: A discrete optimization approach to selecting a population subset for use in a high-density genetic mapping project*. Proceedings of the 11th Annual {ACM-SIAM} Symposium on Discrete Algorithms: 419-428.
  - Xu Z, Zou F, Vision TJ (2005) Improving QTL mapping resolution in experimental crosses by the use of genotypically selected samples. Genetics 170, 401-408.



MapPop is a Free Program!

Based on Genotype Steps (Selective Mapping)

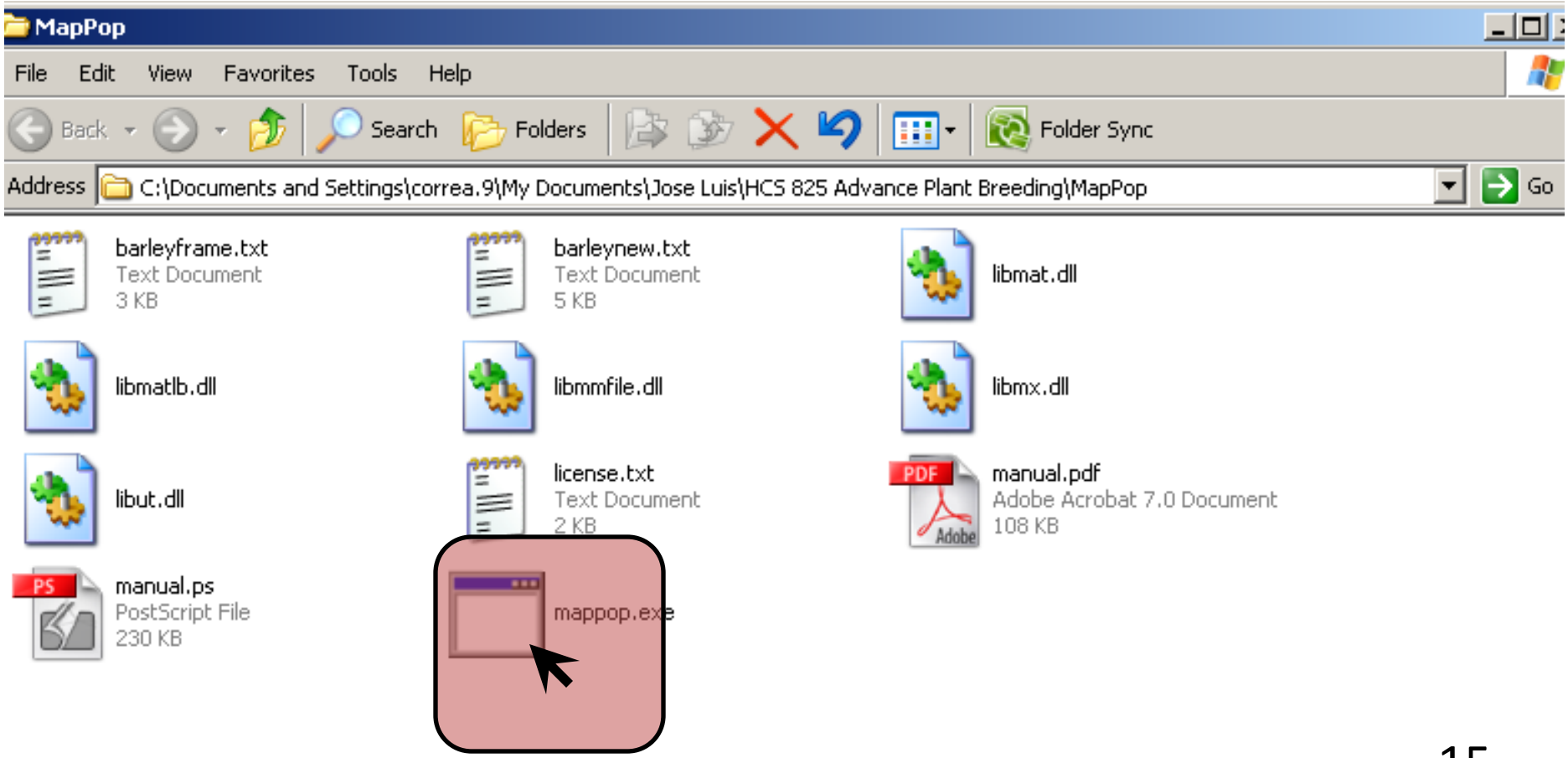
2. Where to find and how to install MapPop



Then, extract all the files from the zip folder.

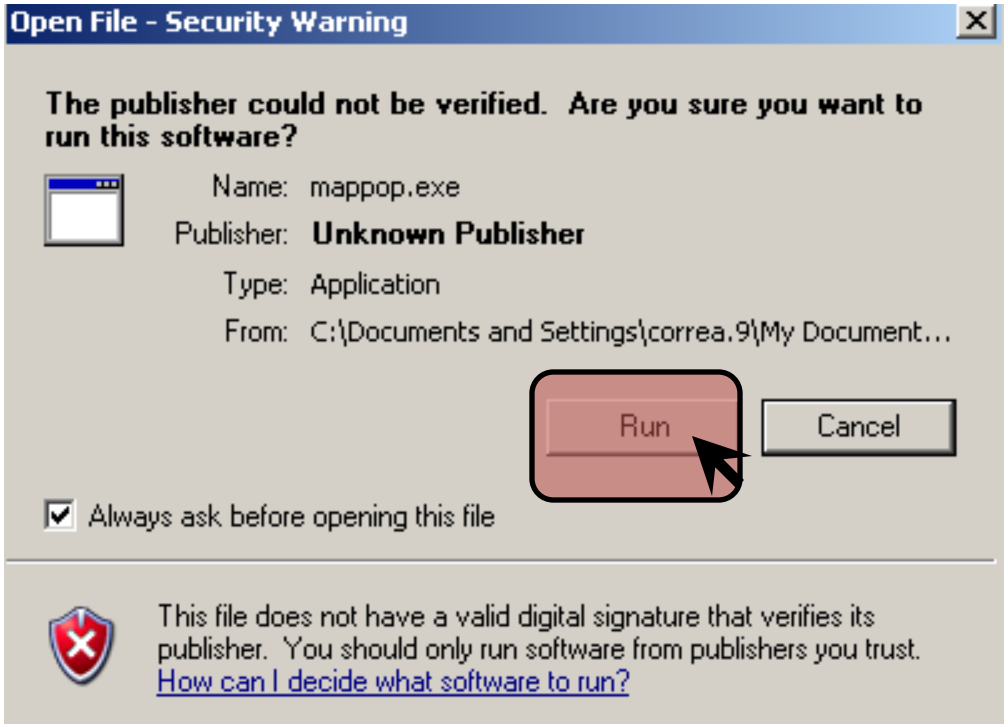
Based on Genotype Steps (Selective Mapping)

2. Where to find and how to install MapPop



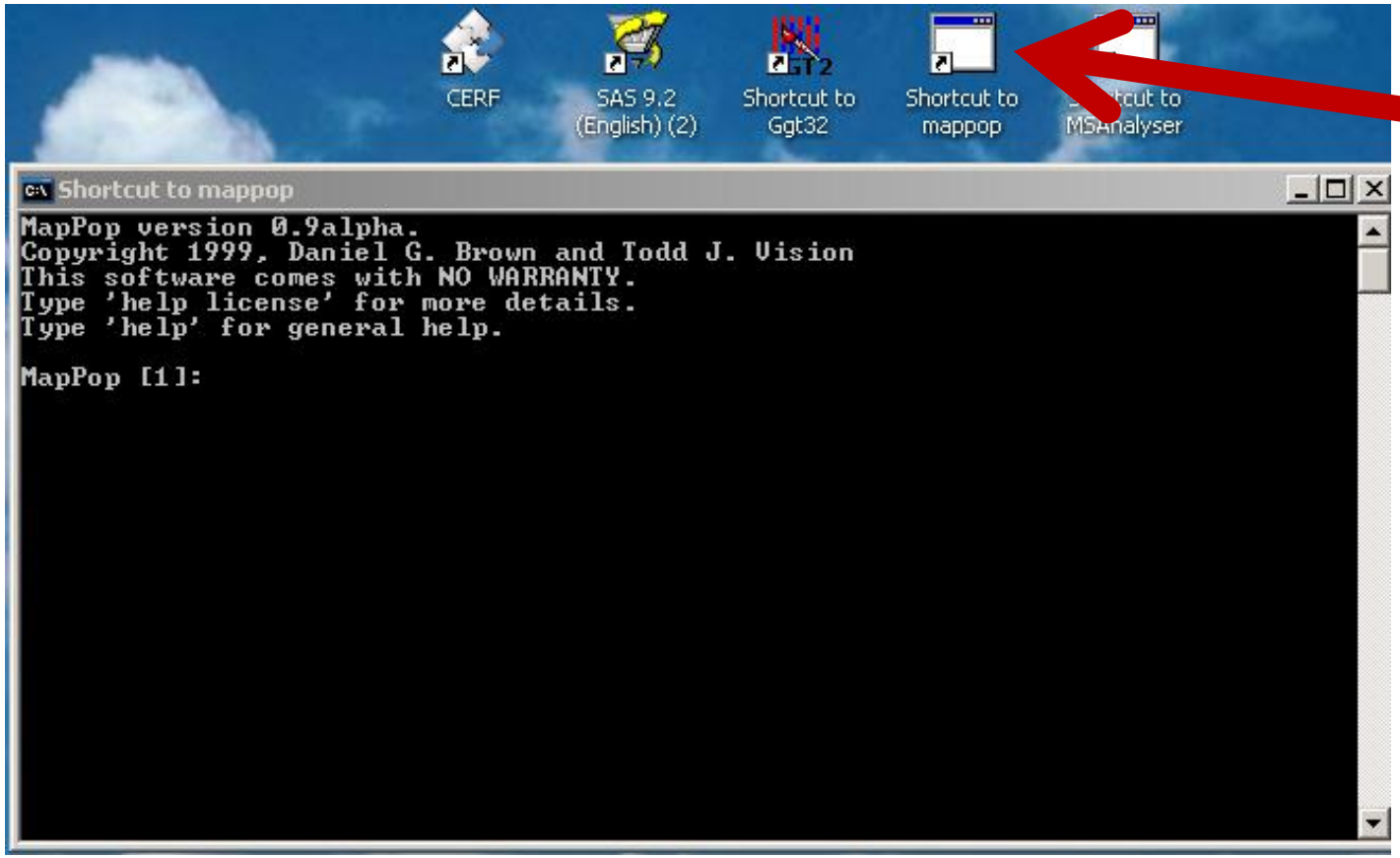
Based on Genotype Steps (Selective Mapping)

2. Where to find and how to install MapPop



Based on Genotype Steps (Selective mapping)

3-1. MapPop “Welcome” windows

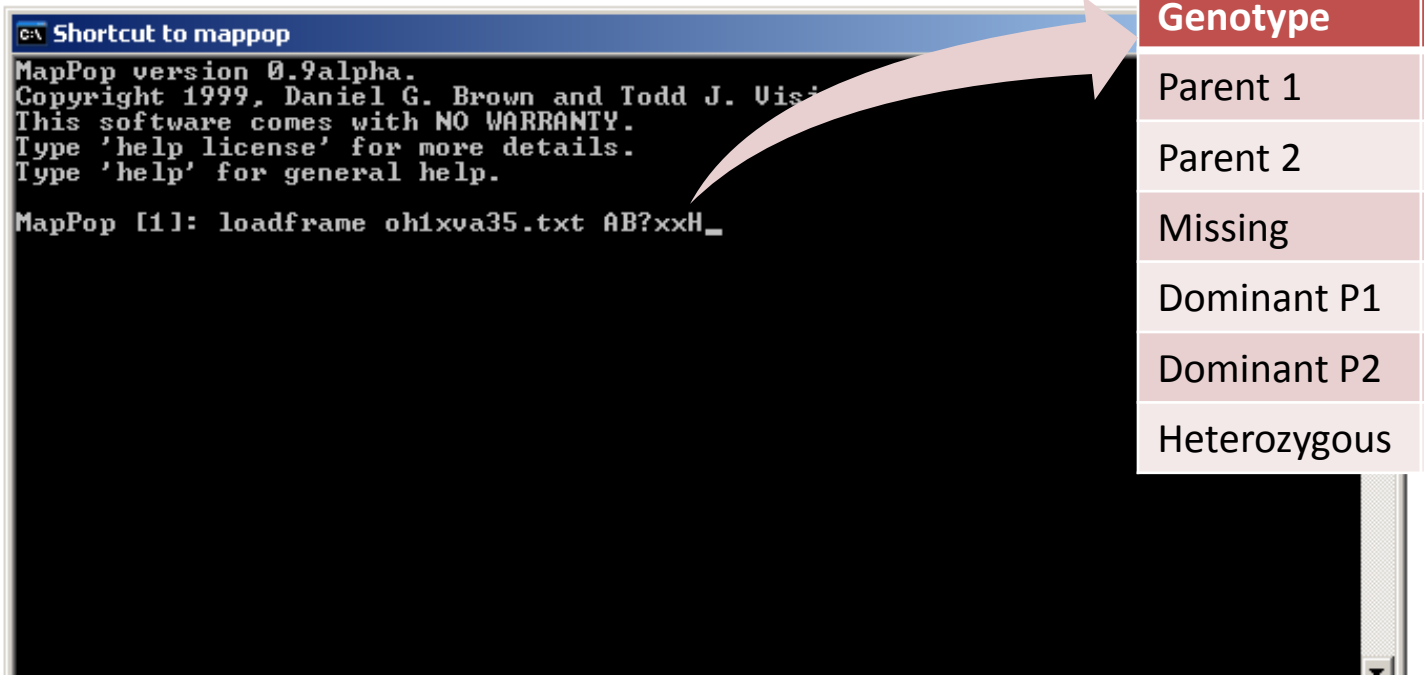


You can create a shortcut to MapPop on your desktop

Based on Genotype Steps (Selective Mapping)

3-2. To run MapPop: insert code and then Enter:

loadframe “name of the file.txt” “genotyping code”



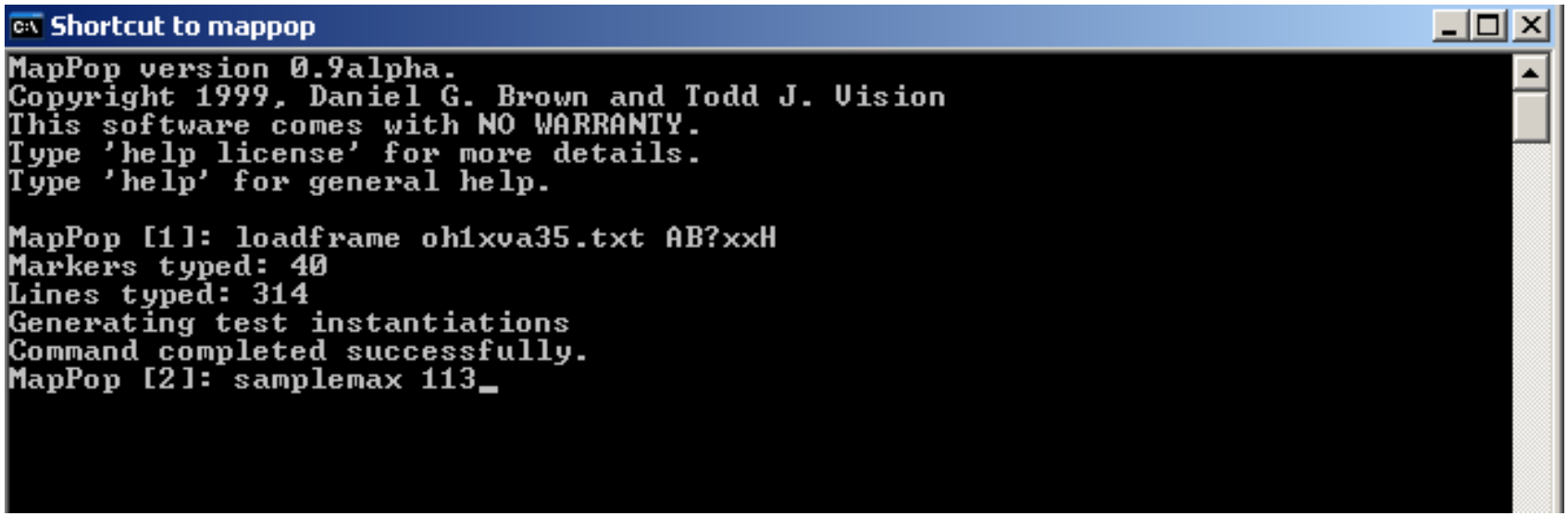
Genotype	Code
Parent 1	A
Parent 2	B
Missing	?
Dominant P1	X
Dominant P2	X
Heterozygous	H

Make sure that the txt file is in the same directory as MapPop (C:/MapPop/oh1xva35.txt)

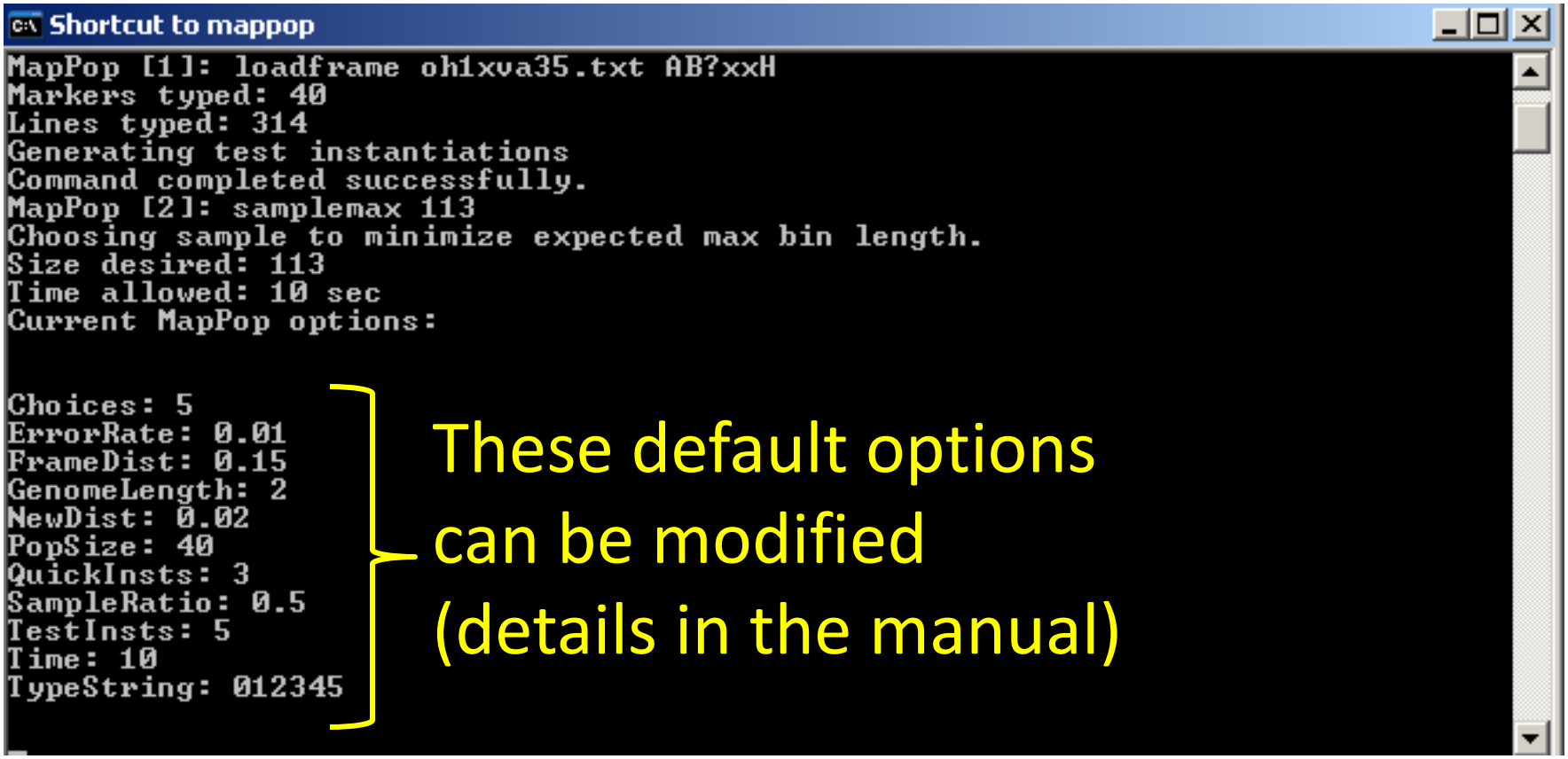
Based on Genotype Steps (Selective Mapping)

3-3. To run MapPop using the default options, insert either code: samplemax or sampleexp, then type the desired population size as shown (More details in the manual)

samplemax 113 → 113 ≈ 1/3 of my population size



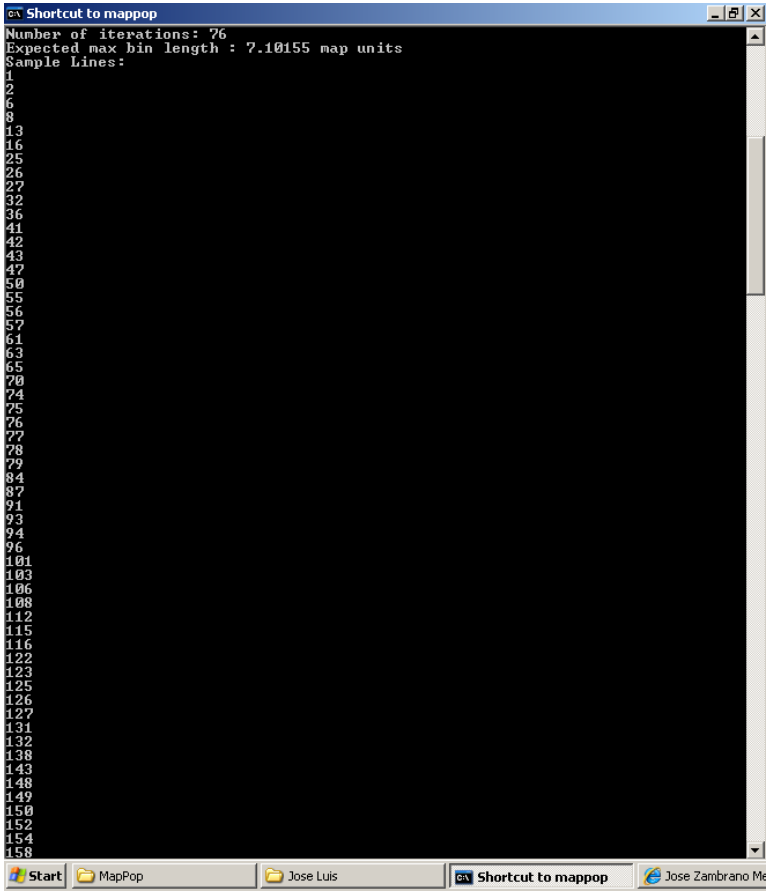
Based on Genotype Steps (Selective Mapping)

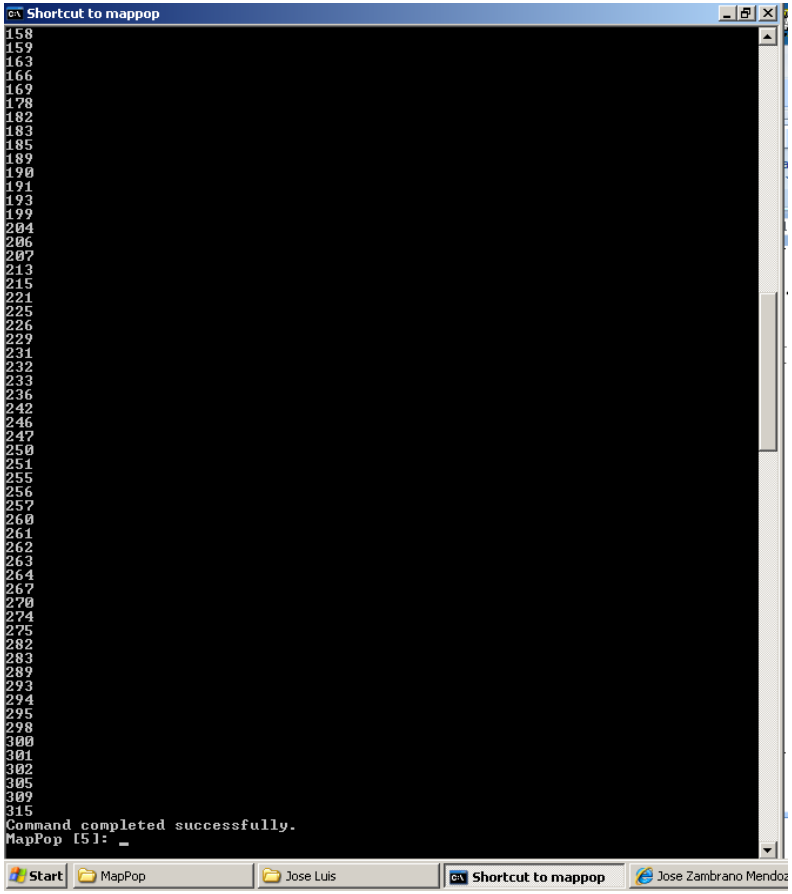


This screen appears while the computer is performing the analysis

Based on Genotype Steps (Selective Mapping)

MapPop Results: (may take a couple of minutes to display results)





Selective Mapping

Case Study

A Post-Hoc analysis was conducted using the same data that was analyzed and published in Theoretical and Applied Genetics (2004) 110: 48-57

Identification of quantitative trait loci controlling resistance to maize chlorotic dwarf virus

Mark W. Jones, Margaret G. Redinbaugh, Robert J. Anderson, R. Louie

Selective Mapping

Published Results

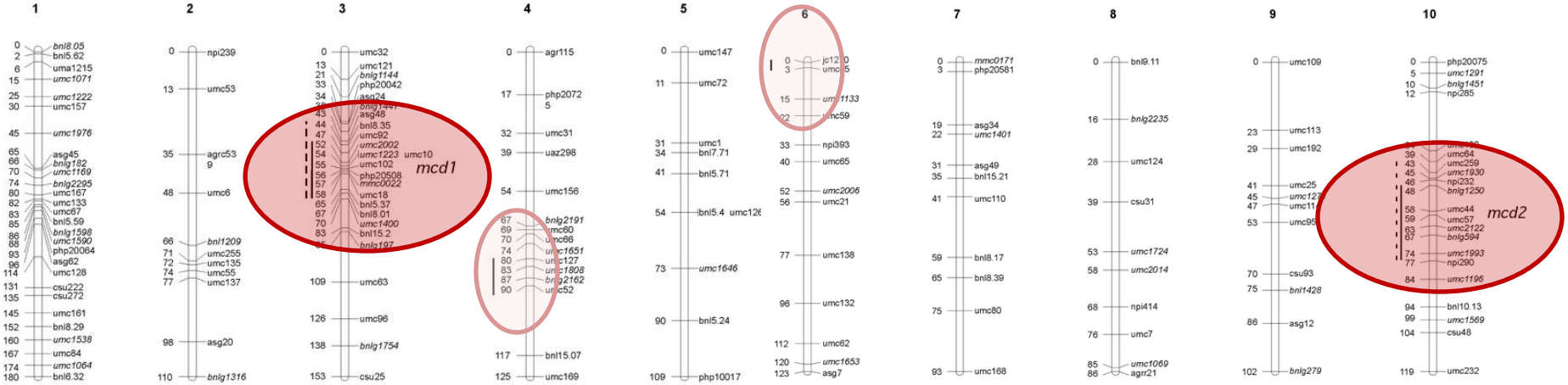


Figure modified from Jones et al., 2004

Linkage map for 314 Va35 × Oh1VI F_2 progeny. The positions for the two major QTL associated with MCDV resistance, *mcd1* and *mcd2*, are noted to the *right* of chromosomes 3 and 10, respectively. Two minor QTL on chromosomes 4 and 6 are also indicated.

Selective Mapping

Case Study - Objective

To compare the results using the whole population (314 lines) versus selective mapping based on phenotype and genotype

Questions

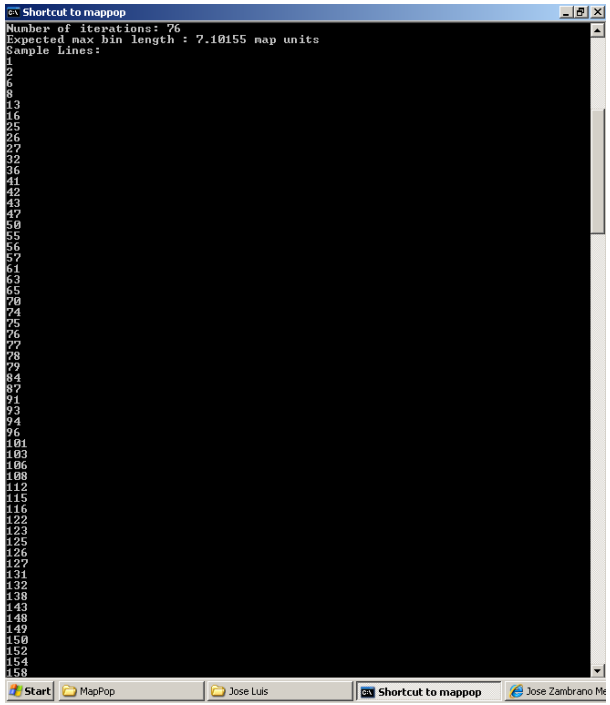
- How does selective mapping affect the genetic map?

- How does selective mapping affect QTL detection?

Selective Mapping

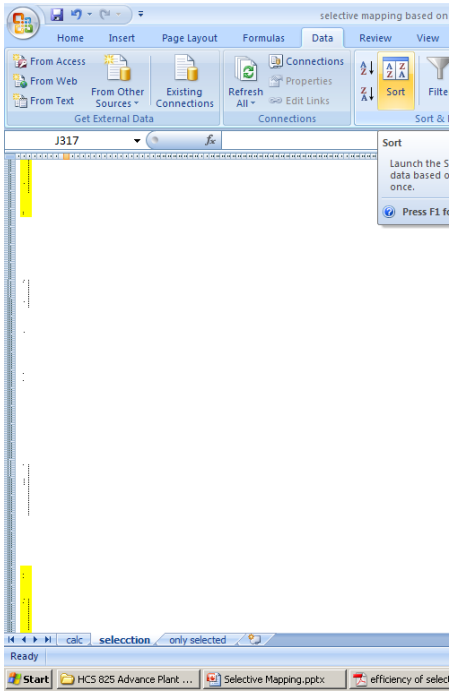
Questions ? – Case Study

Based on Genotype



Vs.

Based on Phenotype



Compared with the published study

Selective Mapping

Results :

Based on Genotype	Vs.	Based on Phenotype
113 Lines		114 Lines

The two sets of selected lines share 40 lines (35% of the selected population)

Notes:

✓Selection represents $\bar{X}$ ± 1σ

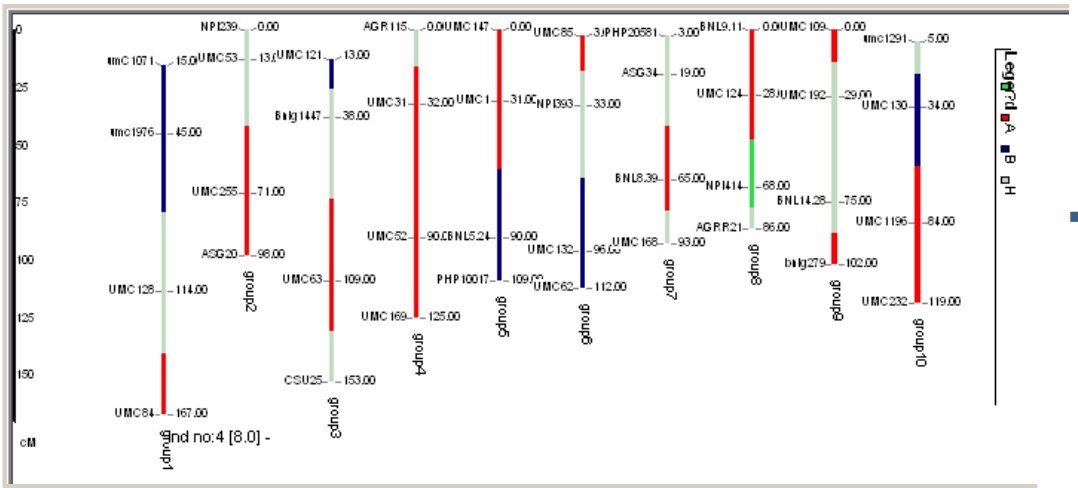
✓Quantitative data has normal distribution

✓Markers segregate as expected for an F_2 Population (1:2:1)

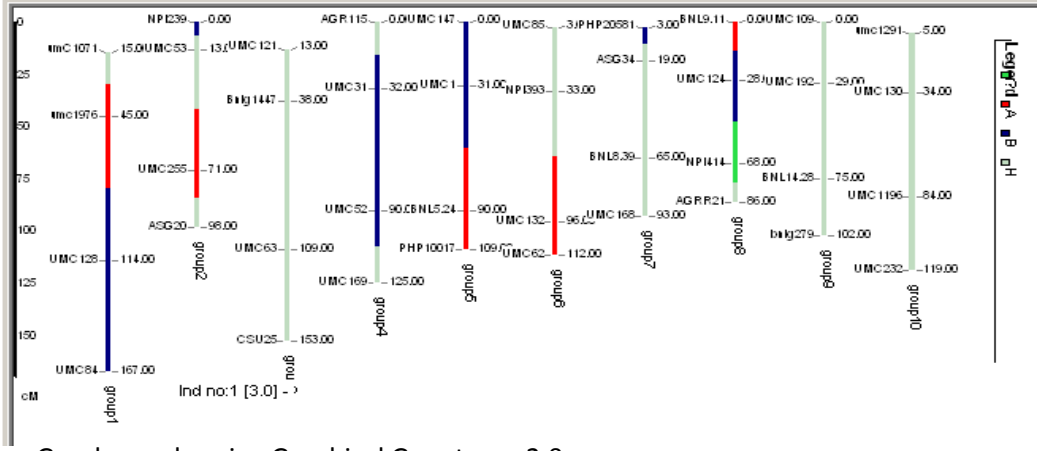
✓Linkage mapping and QTL analysis were conducted using the same procedures for the 2 sets of lines.

Selective Mapping

Results based on genotypic selection (Genetic map of individual lines)



→ **Selected line**



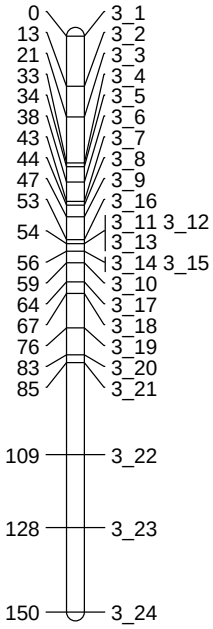
→ **Unselected line**

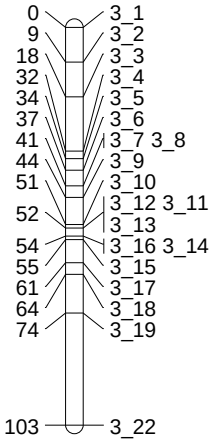
More recombination or breakpoints are observed in a selected line compared with an unselected line

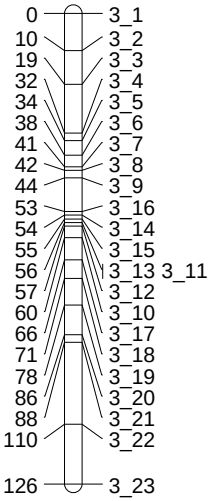
Graphs made using Graphical Genotypes 2.0

Results in Linkage Map, Group 3

	Whole Population 314 F_2	Selection Phenotype 114 F_2	Selection Genotype 113F_2
Length cM	150.1	102.6	126.4
Similar Order %		54	79
Av Density	6.5	5.4	5.7



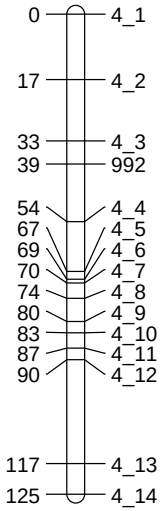


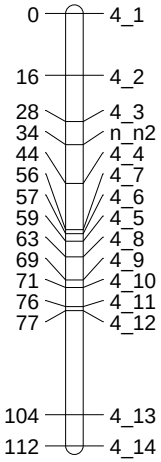


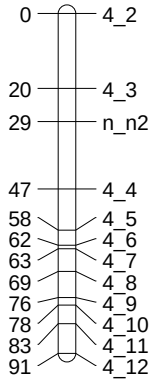
There were no major differences for groups 1, 2, 5, 7, and 9

Results in Linkage Map, Group 4

	Whole Population 314 F_2	Selection Phenotype 114 F_2	Selection Genotype 113F_2
Length cM	125.3	111.8	90.9
Similar Order %		87	100
Av Density	8.9	8.0	8.3



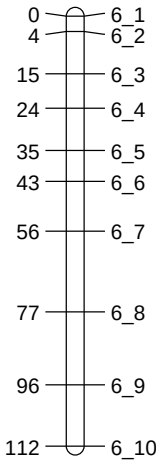


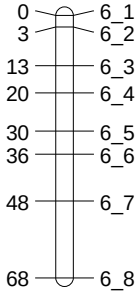


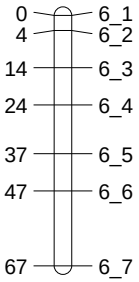
There were no major differences for groups 1, 2, 5, 7, and 9

Results in Linkage Map, Group 6

	Whole Population 314 F_2	Selection Phenotype 114 F_2	Selection Genotype 113F_2
Length cM	112.2	68.0	67.4
Similar Order %		100	100
Av Density	12.5	9.7	11.2



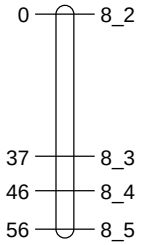


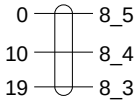


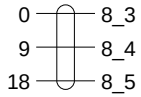
There were no major differences for groups 1, 2, 5, 7, and 9

Results in Linkage Map, Group 8

	Whole Population 314 F_2	Selection Phenotype 114 F_2	Selection Genotype 113F_2
Length cM	55.9	19.4	18.1
Similar Order %		100	100
Av Density	18.6	9.7	9.0



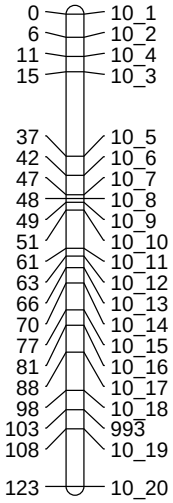


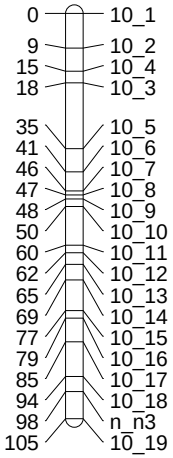


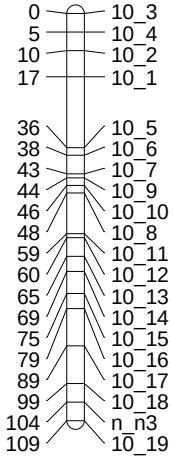
There were no major differences for groups 1, 2, 5, 7, and 9

Results in Linkage Map, Group 10

	Whole Population 314 F_2	Selection Phenotype 114 F_2	Selection Genotype 113F_2
Length cM	122.9	104.7	109.1
Similar Order %		100	65
Av Density	6.1	5.5	5.7







There were no major differences for groups 1, 2, 5, 7, and 9

Selective Mapping

Results – QTLs Detected:

Published study (Jones et al., 2004) — **Post-Hoc analysis**

314 F2			
map cM	Lod	%expl. Var.	locus
Linkage group 3			
93.7	20.1	21.1	3_15 !
Linkage group 4			
107.8	3.34	3.2	4_11
Linkage group 6			
108.6	3.44	3	6_2
Linkage group 10			
66.6	17.96	20.7	10_11 !

114 Selective Phenot.			
map cM	Lod	%expl. Var.	locus
Linkage group 3			
59.7	11.13	26.4	3_15 !
Linkage group 4			
16.4	3.22	5.2	4_2
Linkage group 5			
14.8	3.64	8.4	5_2
Linkage group 10			
54.6	10.04	23.5	10_11 !

113 Selective Genot.			
map cM	Lod	%expl. Var.	locus
Linkage group 3			
54.8	11.34	31.7	3_15 !
Linkage group 10			
58.3	5.31	13.1	10_11 !

Selective Mapping

Results – QTLs:

SUMMARY

	Whole Population 314 F_2	Selective Phenotyping 114 F_2	Selective Genotyping 113F_2
Number of QTLs detected	4	4	2
Number of Major QTLs detected	2	2	2
Variance explained by QTLs	48 %	64 %	45 %
Variance explained only by major QTLs	42 %	50 %	45 %
Variance explained by the largest QTL	21 %	26 %	32 %

Conclusions

- Selective mapping may affect the length of the linkage map
- Power of QTL detection based on selective mapping was not affected for major QTL. Power to detect minor QTL (<10%) was possibly reduced
- Mapping based on phenotypic selection explained a higher proportion of variance; mapping based on selective genotyping explained the same variance relative to the whole population

Conclusions

- Selective mapping based on phenotype is marginally more powerful for detecting QTL than selective mapping based on genotype
- Selective mapping based on genotype can have similar QTL detection power relative to mapping using the whole population
- Selective mapping is a valid strategy when resources are limited

References Cited and External Links

References Cited

- Gallais, A., L. Moreau, and L. Charcosset. 2007. Detection of marker-QTL association by studying changing in marker frequencies with selection. Theoretical and Applied Genetics 114: 669-681. (Available online at: http://dx.doi.org/10.1007/s00122-006-0467-z) (verified 29 Nov 2011).
- Jones, M. W., M. G. Redinbaugh, R. J. Anderson, and R. Louie. 2004. Identification of quantitative trait loci controlling resistance to maize chlorotic dwarf virus. Theoretical and Applied Genetics 110: 48-57. (Available online at: http://dx.doi.org/10.1007/s00122-004-1757-y) (verified 29 Nov 2011).
- Lander, E. S., and D. Bostein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185-199.
- Sun, Y. P., J. K. Wang, J. H. Crouch, and Y. B. Xu. 2010. Efficiency of selective genotyping for genetic analysis of complex traits and potential applications in crop improvements. Molecular Breeding 26: 493-511. (Available online at: http://dx.doi.org/10.1007/s11032-010-9390-8) (verified 29 Nov 2011).
- Vision, T. J., D. G. Brown, D. B. Shmoys, R. T. Durrett, and S. D. Tanksley. 2000. Selective mapping: A strategy for optimizing the construction of high-density linkage maps. Genetics 155: 407-420.

External Links

- Brown, D., and T. Vision. MapPop 1.0: Software for selective mapping and bin mapping [Online]. Department of Biology, The University of North Carolina at Chapel Hill. Available at: http://www.bio.unc.edu/faculty/vision/lab/mappop/manual.pdf (verified 29 Nov 2011).
- MaizeGDB: Maize genetics and genomics database [Online]. Available at: www.maizegdb.org (verified 29 Nov 2011).

Acknowledgments

- Mark Jones
- David Francis
- Margaret Redinbaugh
- Sung-Woo Lee
- Heather Merk, Carlos Hernandez, and Zhifen Zhang (tutorial reviewers)

Photo: J.L. Zambrano. San Jose de Minas, Ecuador