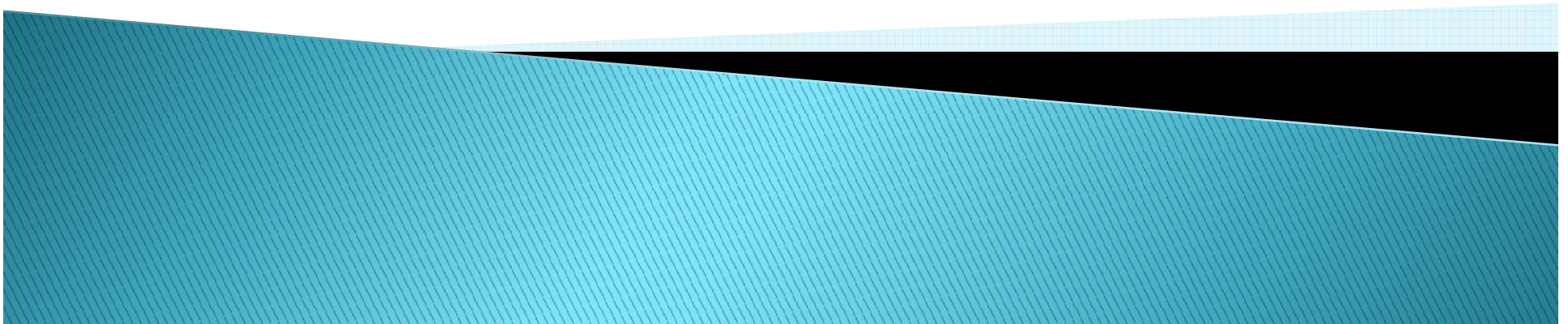


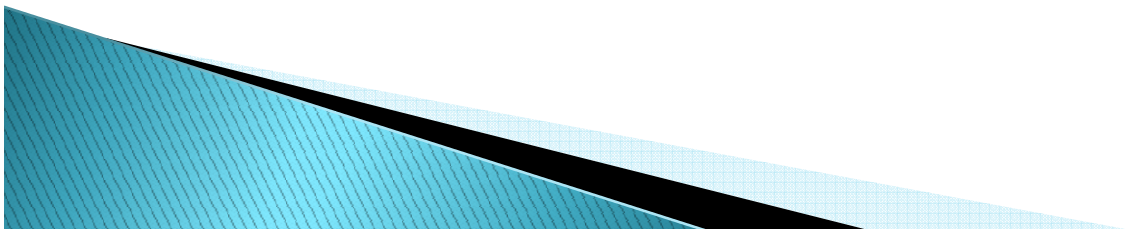
Graphical Genotypes (GGT) software for visualization and analysis of genetic data

OARDC
Nancy Huarachi Morejon
Email: huarachi.1@osu.edu



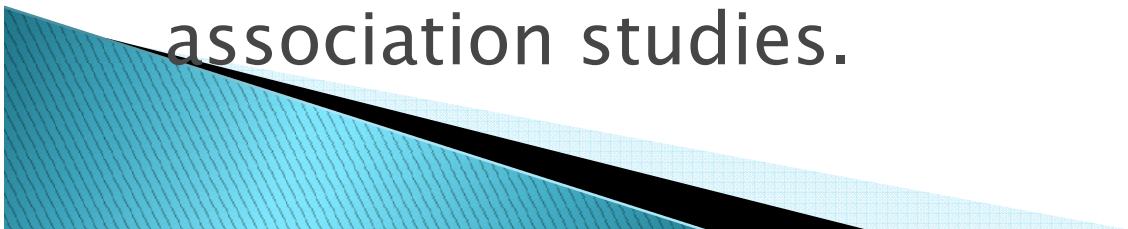
Purpose of this tutorial

- ▶ To demonstrate how to download and install “GGT” (an acronym for Graphical GenoTypes) software.
- ▶ Explain how to organize and input data.
- ▶ Explain the uses of GGT to visualize and analyze genetic data.



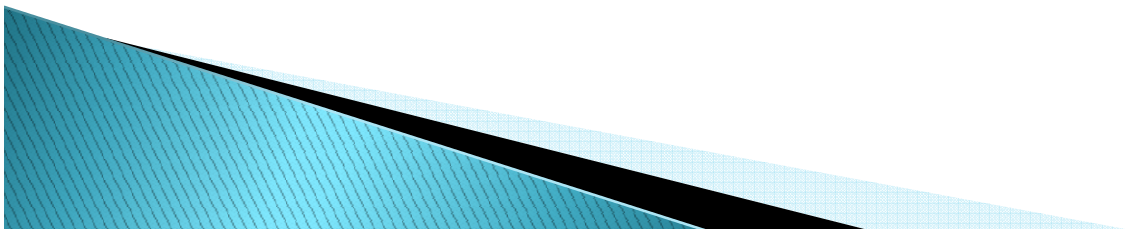
GGT 2.0 Graphical Genotypes

- ▶ Version 1 was released in 1999 as a free software package for visualization of molecular marker data (van Berloo, 1999).
- ▶ The concept of graphical genotypes was developed for plant breeding and plant genetics (Young, N. D. and S. D. Tanksley, 1989).
- ▶ The latest version (2010) has expanded options for genetic analysis of populations including diversity analyses and simple association studies.



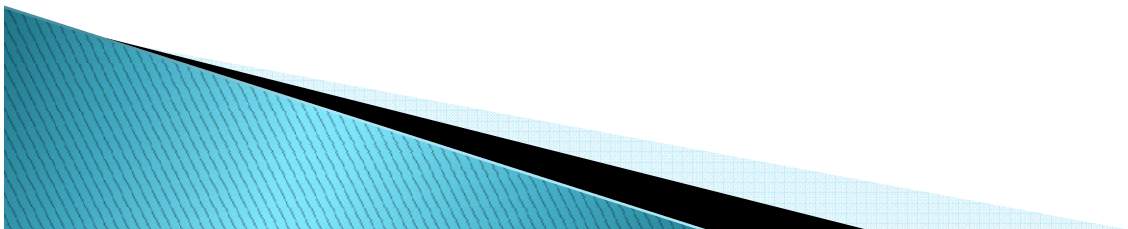
Why GGT ?

- ▶ A graphical representation of molecular marker data can be an important tool in the process of selection and evaluation of plant material.
- ▶ Enables the representation of molecular marker data by simple chromosome drawing in several ways.

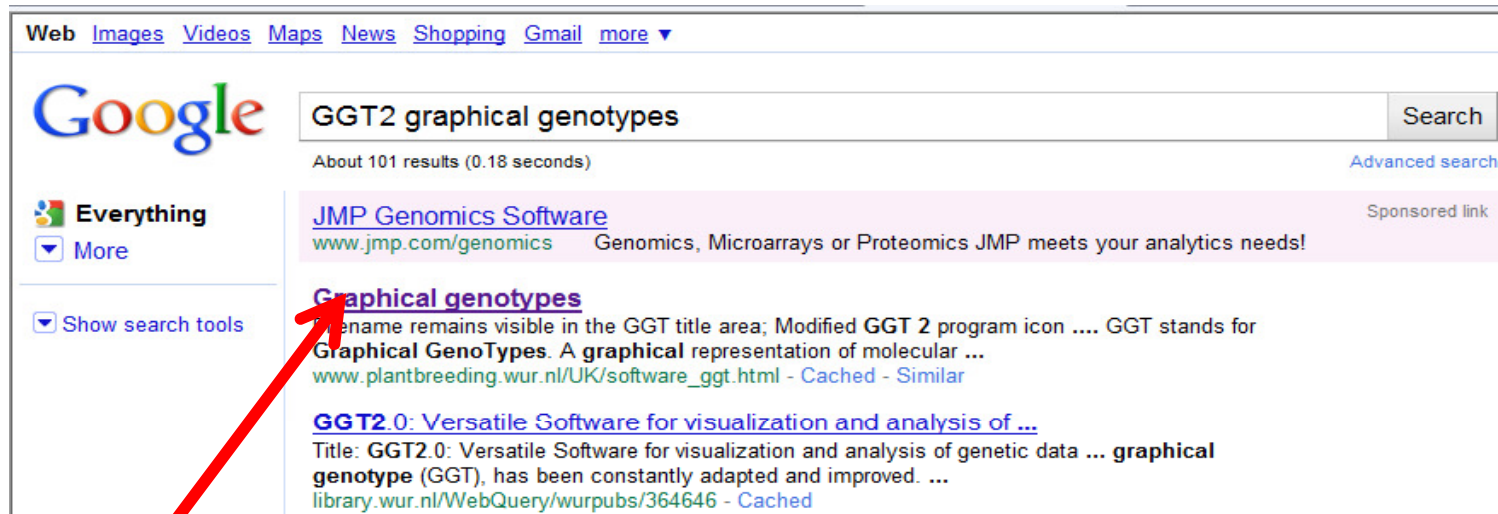


Requirements

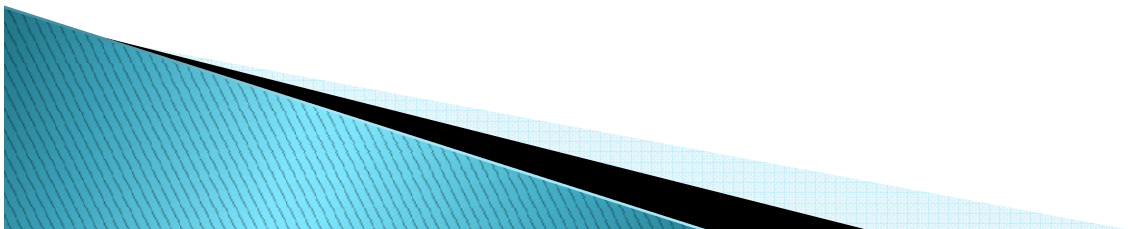
- ▶ Software is in English, developed by Laboratory of Plant Breeding, Wageningen University, The Netherlands (Van Berloo, 1999).
- ▶ System requirements
 - Computer running windows
 - Will take 5 – 10 MB of hard disk.



Where to find...



www.plantbreeding.wur.nl/UK/software_ggt.html



Plant Breeding

Education

Research

Publications

News & Calendar

About Plant Breeding


Work at

Phone book

Links

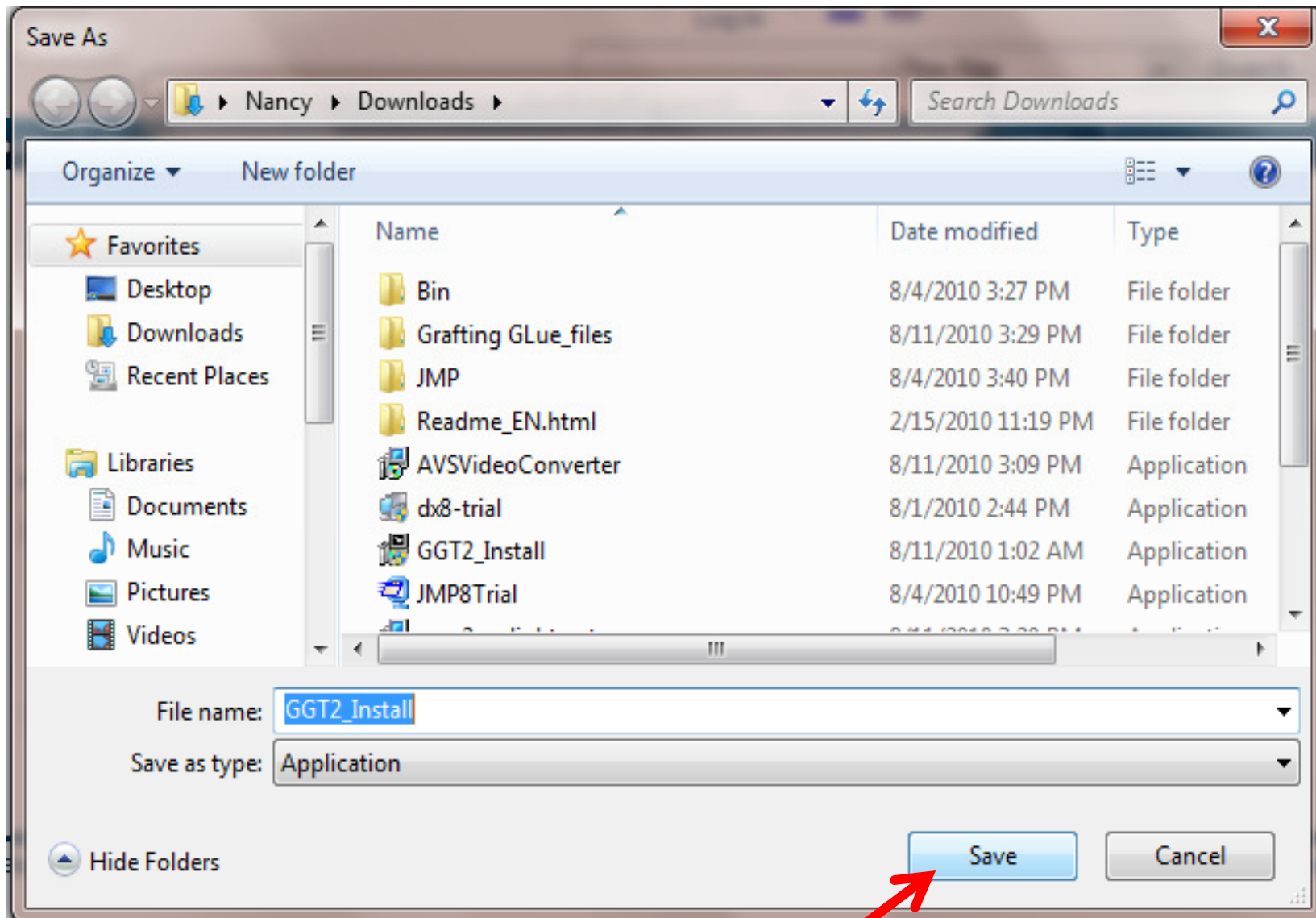
Content

GGT 2.0: graphical genotypes

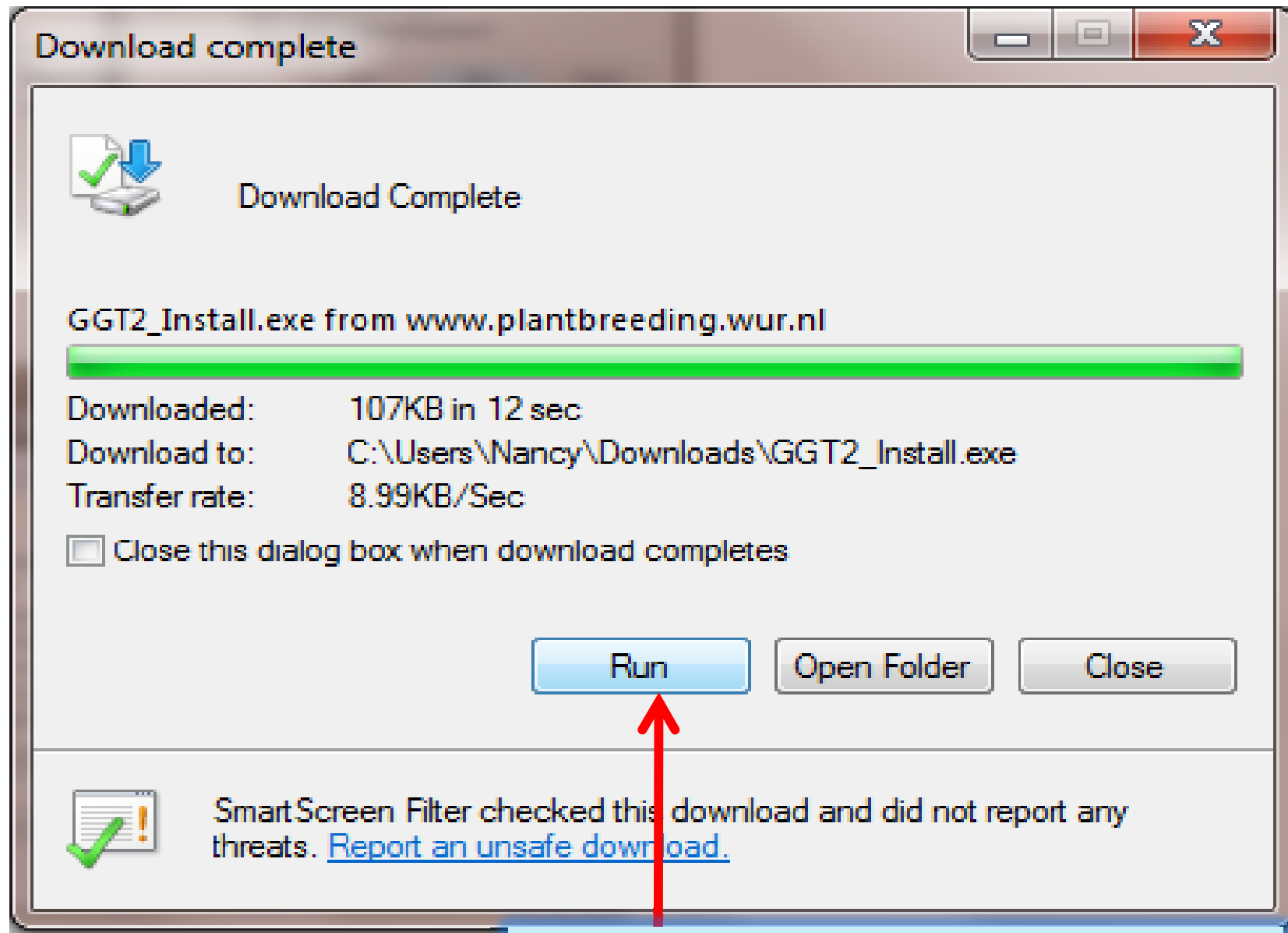
please click [this link](#)  to download the GGT 2.0 install package (2.5 MB)

[back to GGT homepage](#)

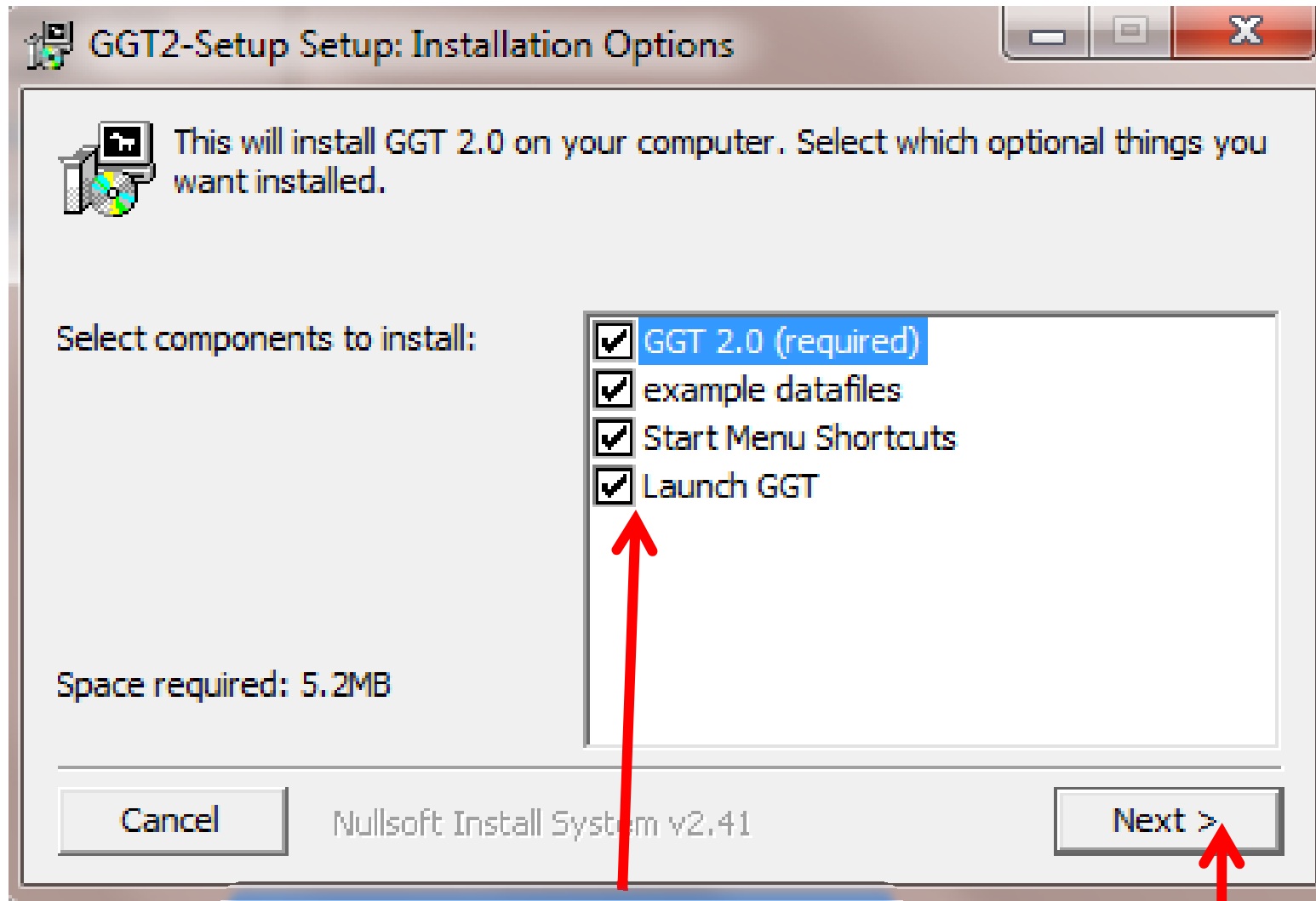
Click here to download the software



Save to file 'Downloads'

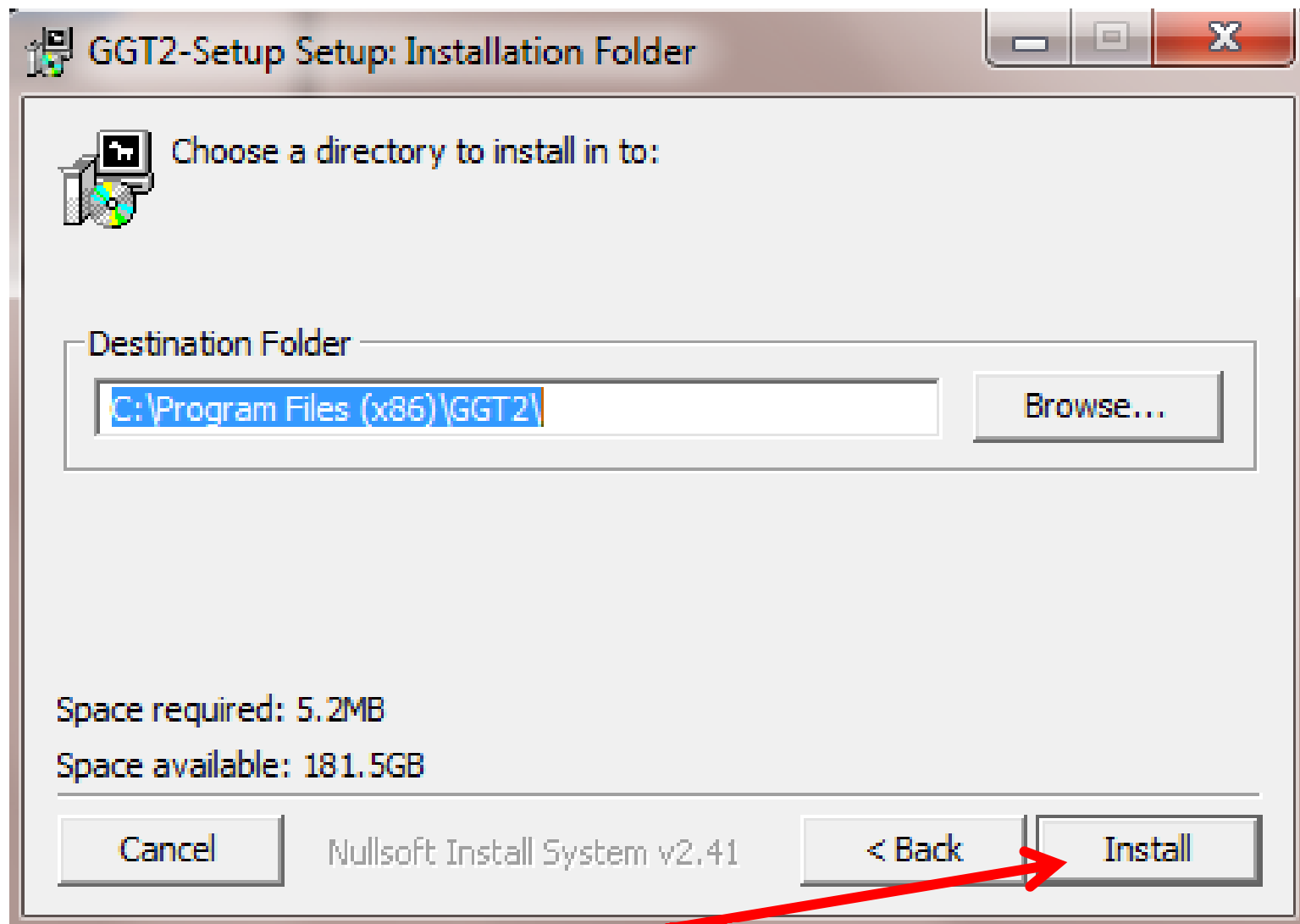


Click on run to start the installation!



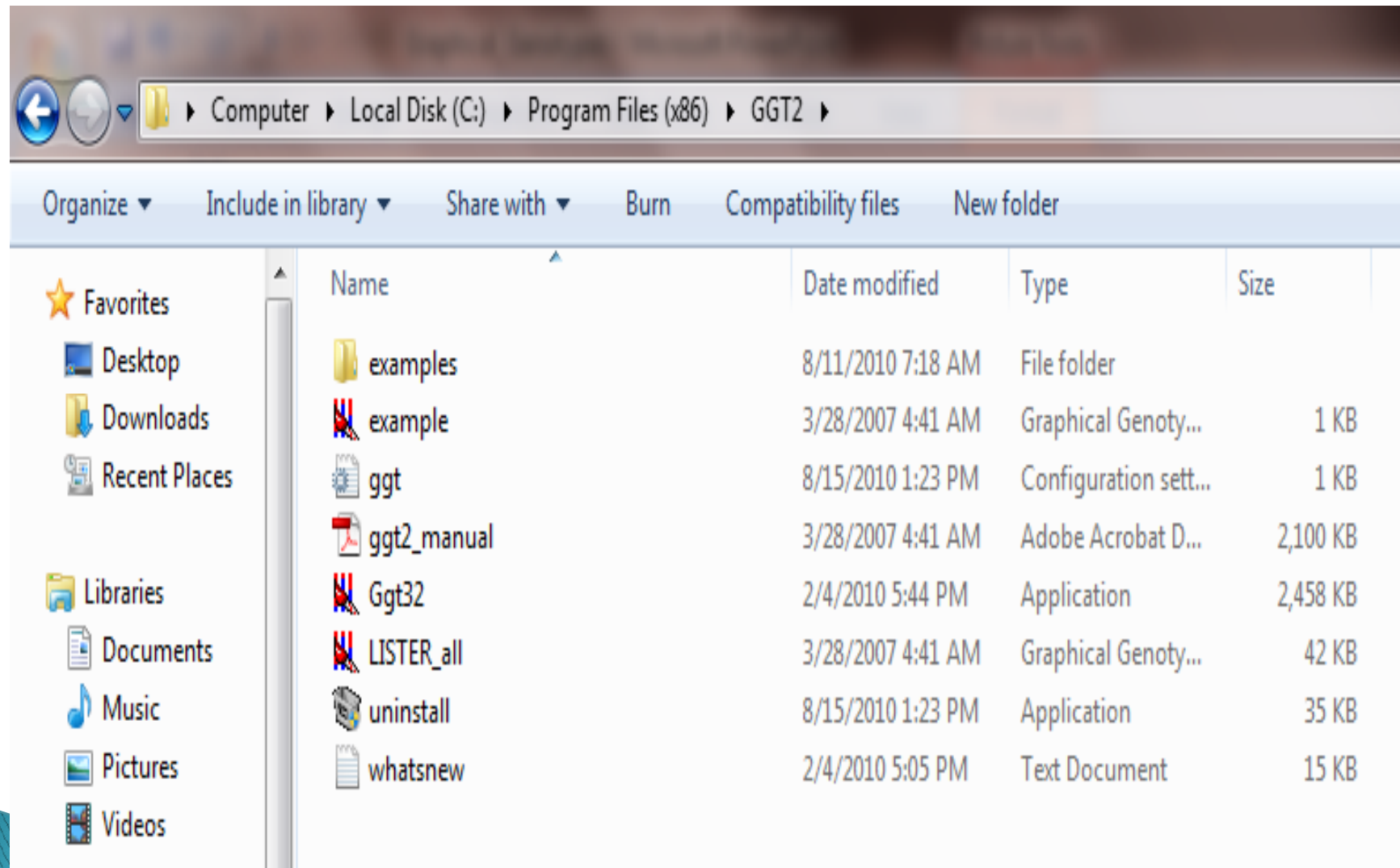
Check all boxes

Click here!

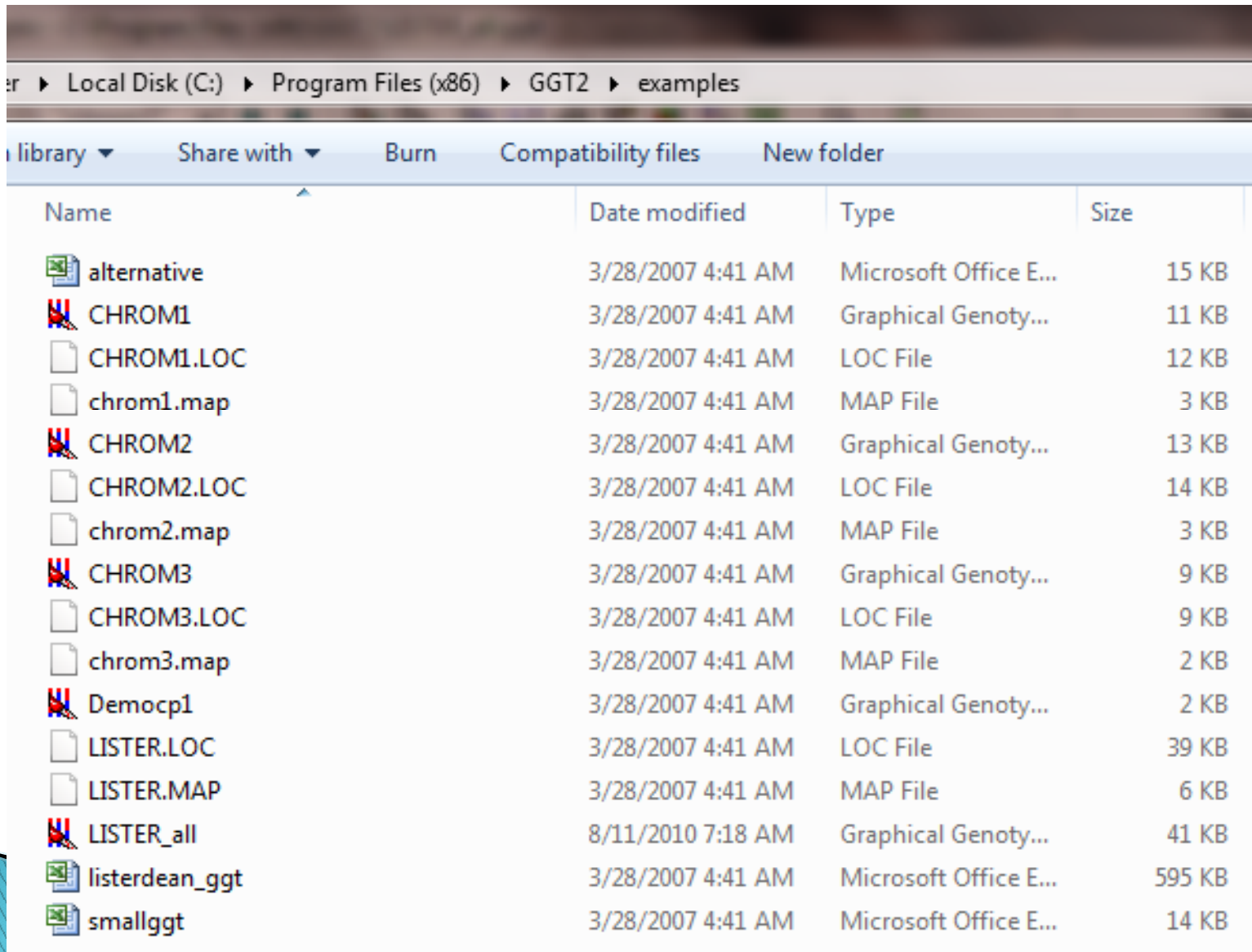


Last step click here!

















What files should you get ?



Example Data

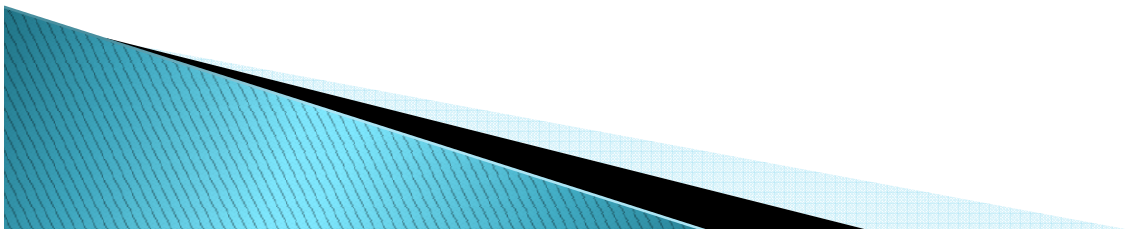


The screenshot shows a Windows Explorer window with the address bar displaying the path: Local Disk (C:) > Program Files (x86) > GGT2 > examples. The window has a menu bar with options: Library, Share with, Burn, Compatibility files, and New folder. Below the menu bar is a table listing files and folders in the directory.

Name	Date modified	Type	Size
 alternative	3/28/2007 4:41 AM	Microsoft Office E...	15 KB
 CHROM1	3/28/2007 4:41 AM	Graphical Genoty...	11 KB
 CHROM1.LOC	3/28/2007 4:41 AM	LOC File	12 KB
 chrom1.map	3/28/2007 4:41 AM	MAP File	3 KB
 CHROM2	3/28/2007 4:41 AM	Graphical Genoty...	13 KB
 CHROM2.LOC	3/28/2007 4:41 AM	LOC File	14 KB
 chrom2.map	3/28/2007 4:41 AM	MAP File	3 KB
 CHROM3	3/28/2007 4:41 AM	Graphical Genoty...	9 KB
 CHROM3.LOC	3/28/2007 4:41 AM	LOC File	9 KB
 chrom3.map	3/28/2007 4:41 AM	MAP File	2 KB
 Democp1	3/28/2007 4:41 AM	Graphical Genoty...	2 KB
 LISTER.LOC	3/28/2007 4:41 AM	LOC File	39 KB
 LISTER.MAP	3/28/2007 4:41 AM	MAP File	6 KB
 LISTER_all	8/11/2010 7:18 AM	Graphical Genoty...	41 KB
 listerdean_ggt	3/28/2007 4:41 AM	Microsoft Office E...	595 KB
 smallggt	3/28/2007 4:41 AM	Microsoft Office E...	14 KB

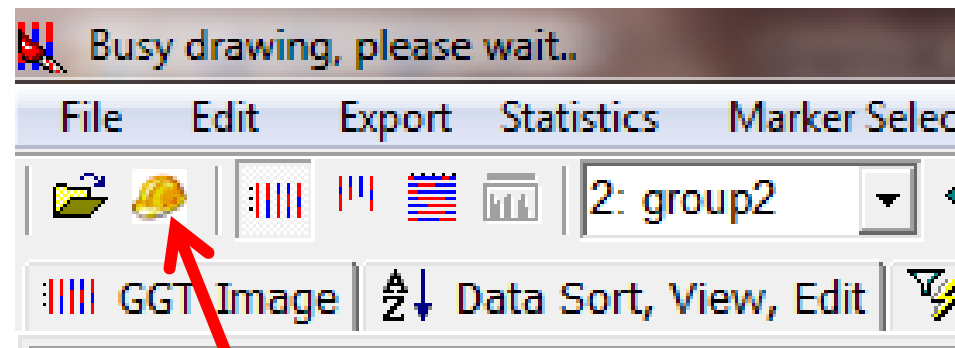
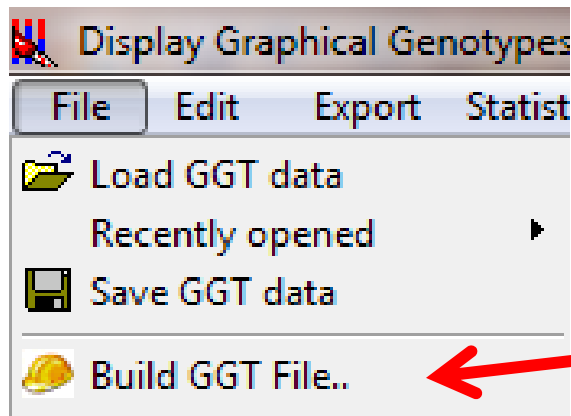
How to prepare data?

- ▶ There are many options to prepare data:
 - 1) Build GGT from two sources, locus and map data.
 - 2) Export using a EXCEL spreadsheet.
 - 3) Alternative layout.



Input Data Building GGT files

- ▶ From two sources:
 - Locus Data
 - Map Data



Build GGT file

Locus Data

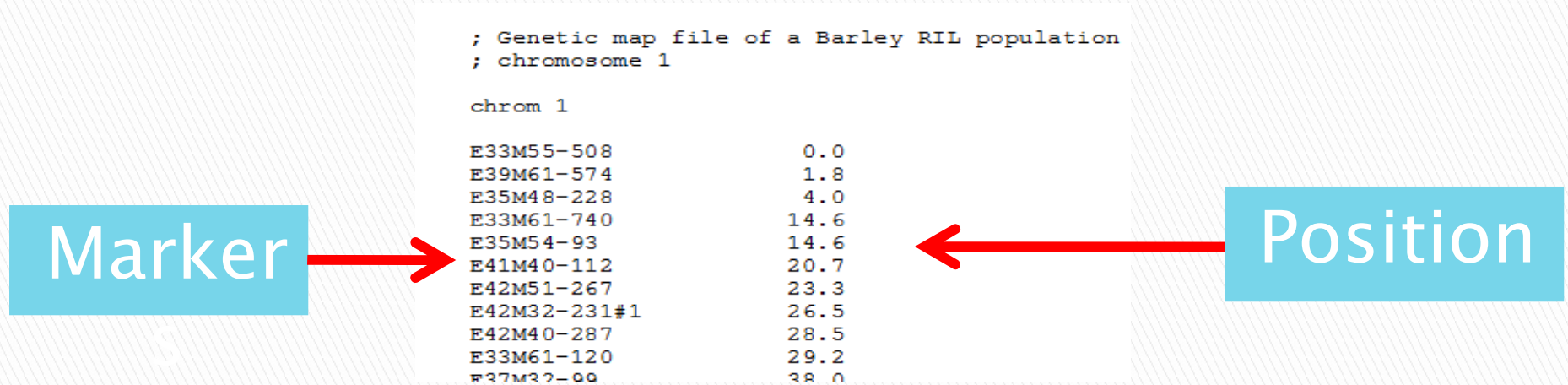
```
; This file was used as input for the JOINMAP mapping software
; use the BUIL GGT FILE option to merge '.loc' and '.map' files into a
'.ggt' file
;
; Fri, 10 Jan 1997, 11:54
; grouping file: mylvuniq.grp
; original file: mylvuniq.loc
; linkage group: 1

name = lvuniq-1|
popt = RI9
nloc = 72
nind = 103

E42M32-231#1
aaaaa -ca-a ccccc aacca cacac caaca aaaca cccca accaa accca
caaac aaaaa accaa acccc accaa cccac acaaa cacc caaaa ccaac
ccc
E33M61-740
aaaaa aaaca ccccc aacca aacac caaca caaca acaaa acacc accaa
cacaa acaaa accca ccccc accaa accac accaa caaca caaaa caaac
```

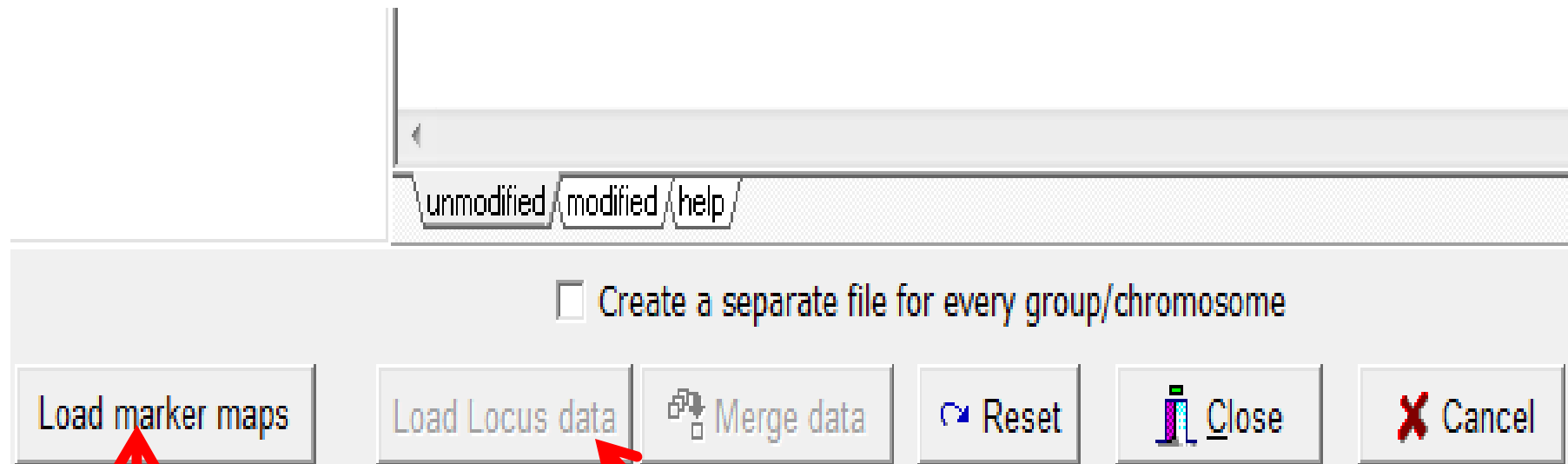
- ▶ *A locus file*, the locus file is a plain text file.
- ▶ *A = Parent 1; B = Parent 2; H = heterozygote; C = not A; D = not B (MAPMAKER FORMAT)*

Data Map



- ▶ The map file is a plain text file as well (MAPMAKER Output). It lists, in two columns, markers and their map positions. GGT needs information on both the allelic composition of a locus as well as the map position.

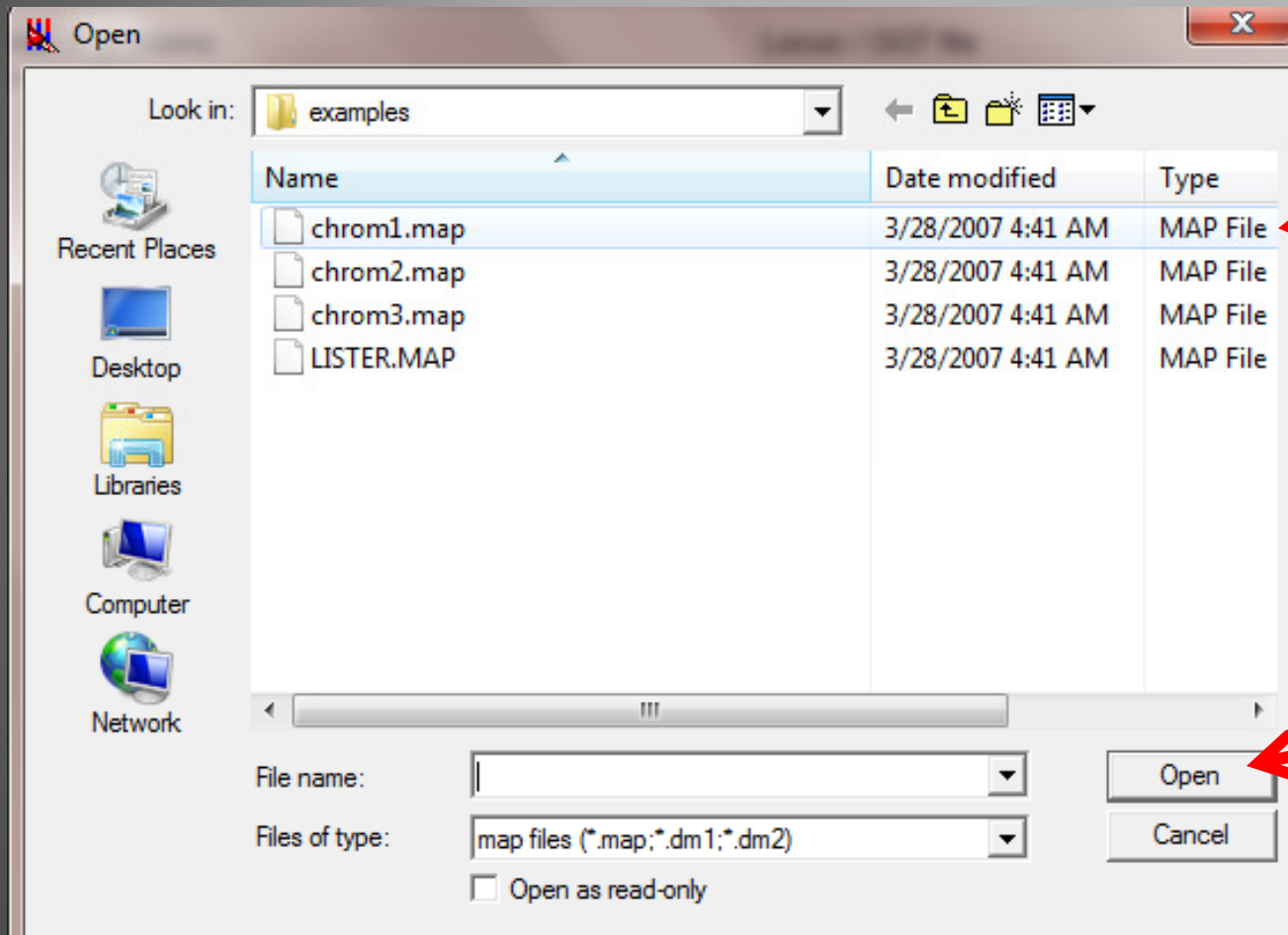
To merge both sources of data



Click here Map data

Click here for locus Data

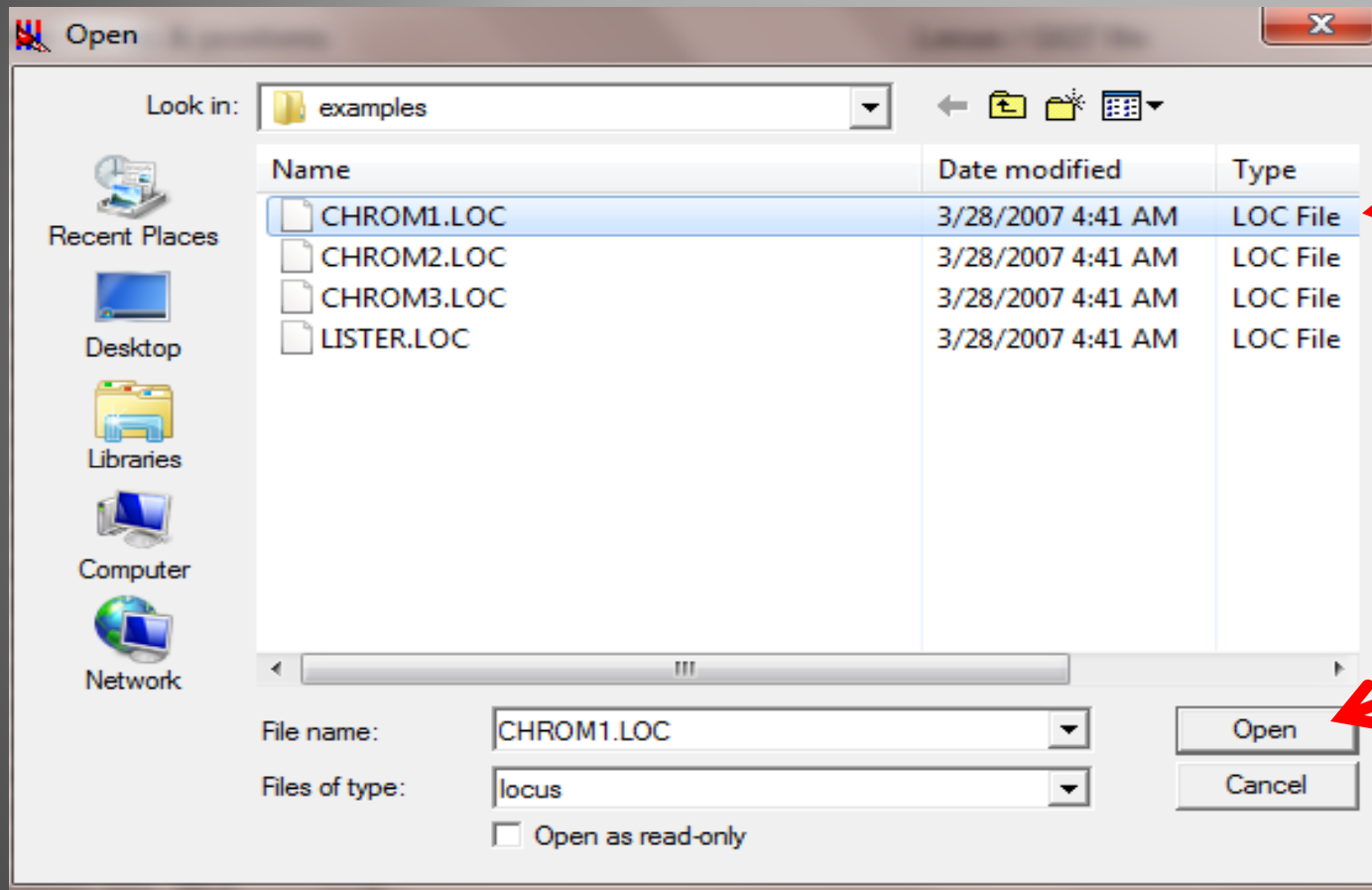
Loading marker maps



Select
one file

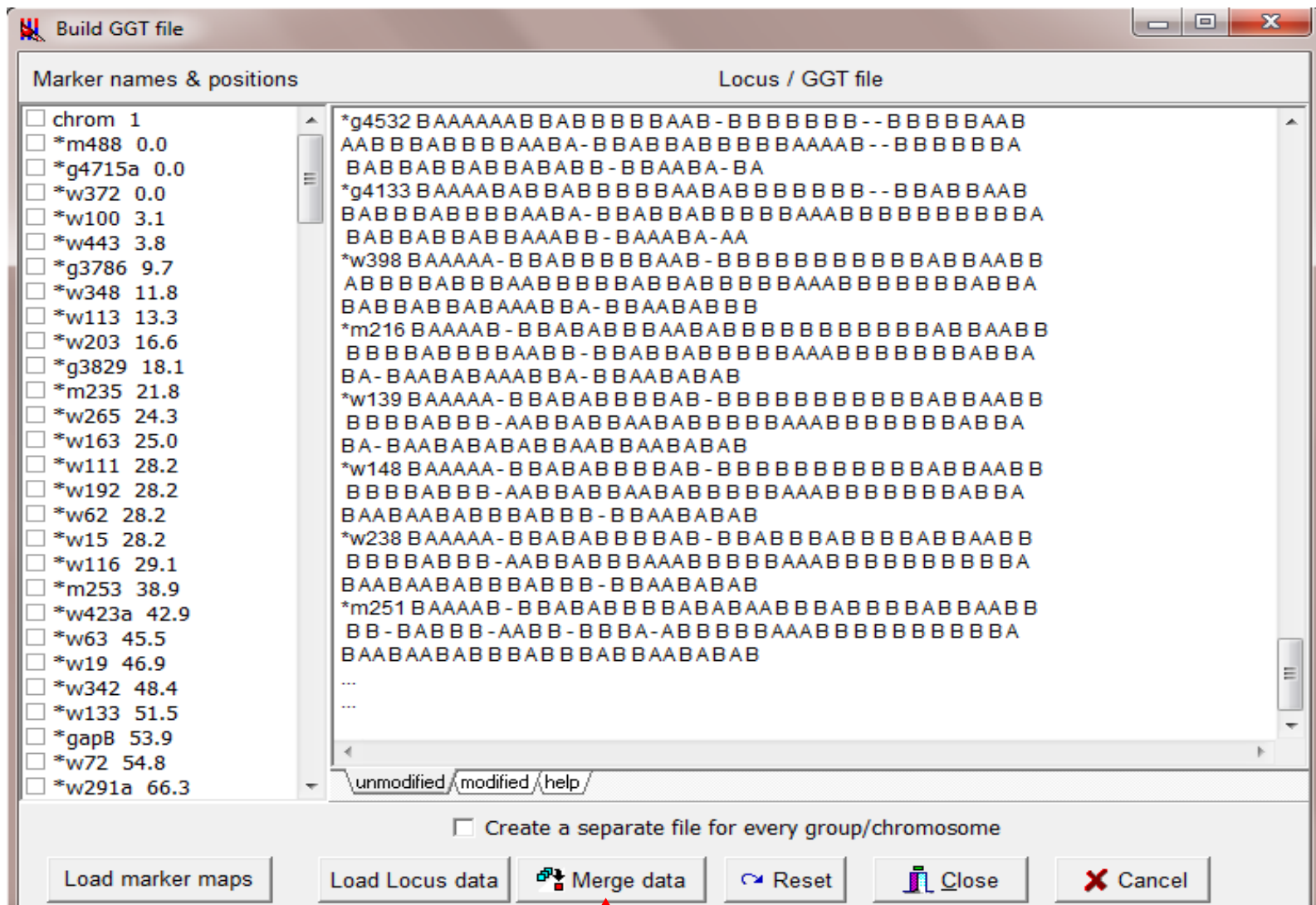
click on
open

Loading locus data

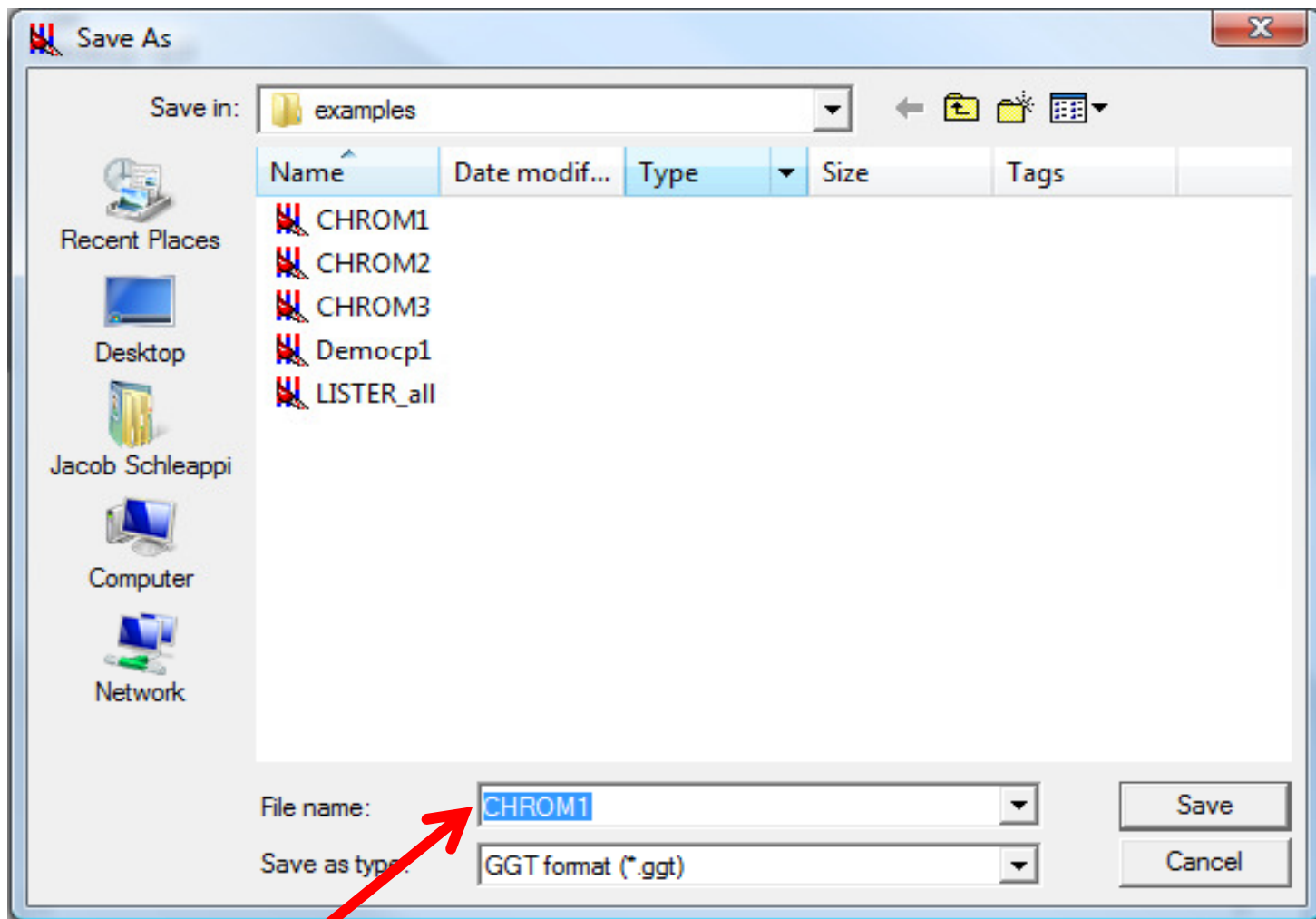


Select
one file

And
click on
open



Click on merge data



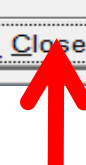
Save in the same folder as your original data files using *.ggt format

Build GGT file

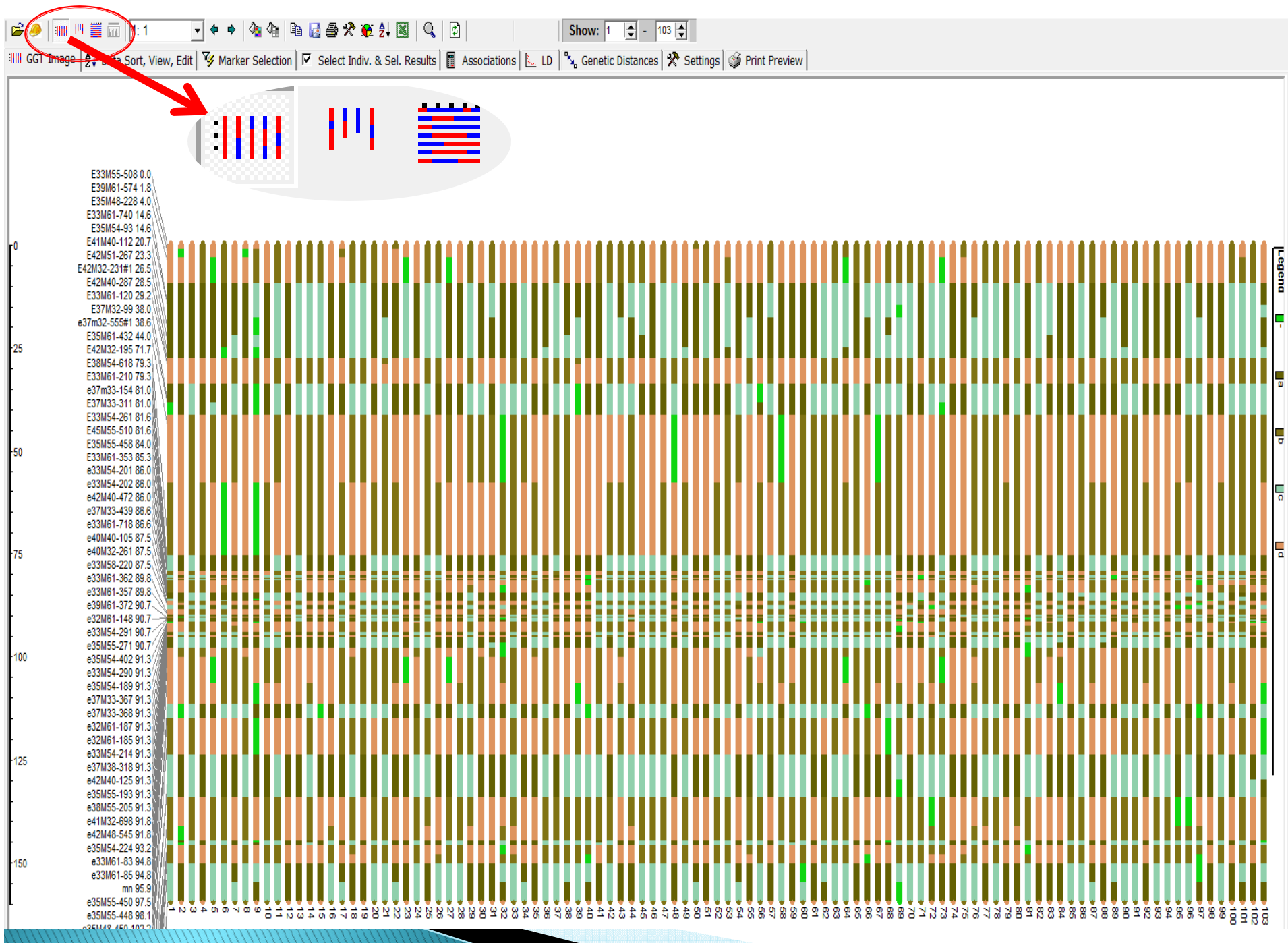
Marker names & positions	Locus / GGT file
<input checked="" type="checkbox"/> chrom 1 <input checked="" type="checkbox"/> E33M55-508 0.0 <input checked="" type="checkbox"/> E39M61-574 1.8 <input checked="" type="checkbox"/> E35M48-228 4.0 <input checked="" type="checkbox"/> E33M61-740 14.6 <input checked="" type="checkbox"/> E35M54-93 14.6 <input checked="" type="checkbox"/> E41M40-112 20.7 <input checked="" type="checkbox"/> E42M51-267 23.3 <input checked="" type="checkbox"/> E42M32-231#1 26.5 <input checked="" type="checkbox"/> E42M40-287 28.5 <input checked="" type="checkbox"/> E33M61-120 29.2 <input checked="" type="checkbox"/> E37M32-99 38.0 <input checked="" type="checkbox"/> e37m32-555#1 38.6 <input checked="" type="checkbox"/> E35M61-432 44.0 <input checked="" type="checkbox"/> E42M32-195 71.7 <input checked="" type="checkbox"/> E38M54-618 79.3 <input checked="" type="checkbox"/> E33M61-210 79.3 <input checked="" type="checkbox"/> e37m33-154 81.0 <input checked="" type="checkbox"/> E37M33-311 81.0 <input checked="" type="checkbox"/> E33M54-261 81.6 <input checked="" type="checkbox"/> E45M55-510 81.6 <input checked="" type="checkbox"/> E35M55-458 84.0 <input checked="" type="checkbox"/> E33M61-353 85.3 <input checked="" type="checkbox"/> e33M54-201 86.0 <input checked="" type="checkbox"/> e33M54-202 86.0 <input checked="" type="checkbox"/> e42M40-472 86.0 <input checked="" type="checkbox"/> e37M33-439 86.6 <input checked="" type="checkbox"/> e33M61-718 86.6	<pre> nchrom=1 nind=103 popt=ri9 name= 1 nloc=72 E33M55-508 0.0 dddbd bdddb bdbbb dbbbb dbddb bdddb bdddb bdbbb dbddb bdddb dbddd ddbbb bdbbb bdddb dbbdb bdddb bdbdb bdddb dbd E39M61-574 1.8 d-dbd bd-bd bdbbb dbbbb ddddb bdddb bdddb dbddd bdbbb dbddd bdddb dbddd ddbbb bdbbb bdddb dbbdb dbddd bdbdb bdbdb bdddb dbd E35M48-228 4.0 dddb- bdddb bdbbb ddbbb dd-db b-dbd bdddb dbddd bdbbb dbddd bdbdb dbddd ddb-b bdbbb bd-dd dbbdb dbddd bdbdb bdbdb bdddb bbd E33M61-740 14.6 aaaaa aaaca ccccc aacca aacac caaca caaca acaaa acacc accaa caaa acaaa accca ccccc accaa accac accaa caaca caaaa caaac cca </pre>

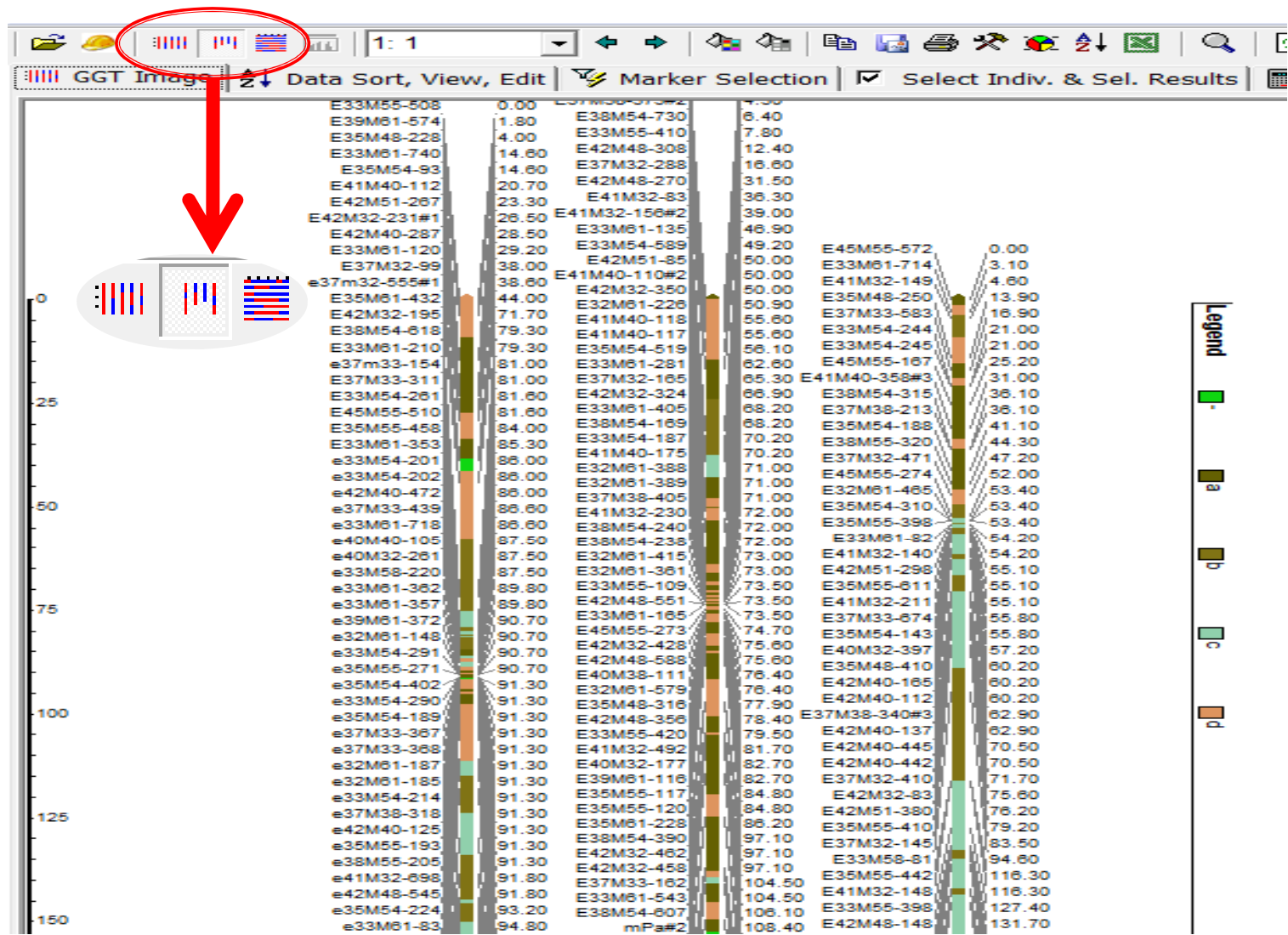
unmodified modified help

☐ Create a separate file for every group/chromosome



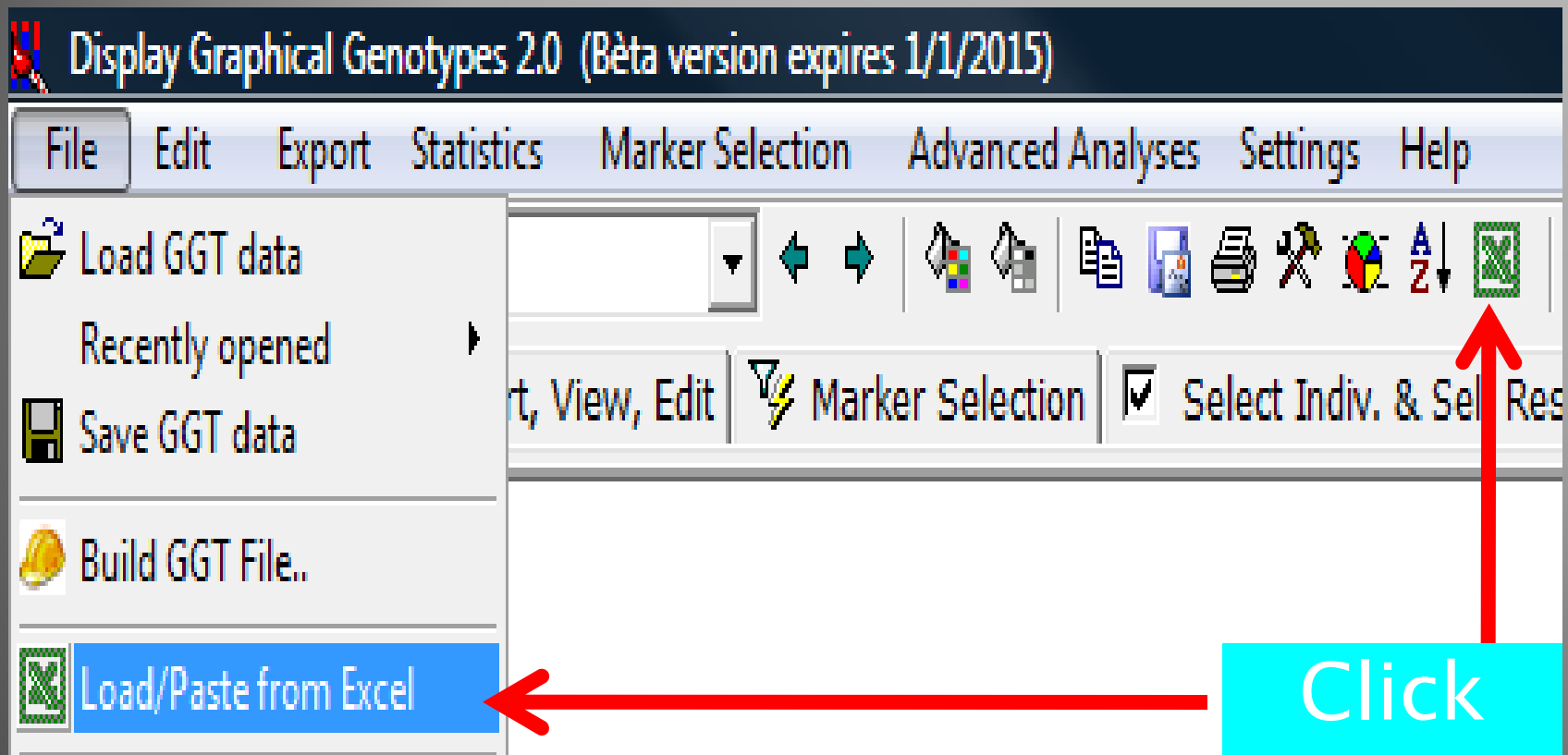
Click here!







Input Data using EXCEL



Organization of data in EXCEL spread sheet

	A	B	C	D	E	F	G	H	I	J	K
1	alias		1	2	3	4	5	6	7	8	9
2	popt	ri8									
3	nchrom	10									
4											
5											
6											
7											
8	name="chrom1"										
9	m488	0.00	B	A	A	B	B	A	U	B	A
10	g4715a	0.00	B	A	A	B	B	A	U	B	A
11	w372	0.00	B	A	A	B	B	A	U	B	A
12	w100	3.10	B	A	A	B	B	A	U	B	U
13	w443	3.80	B	A	A	B	B	A	U	B	A
55	w103	105.20	A	A	A	A	A	A	U	B	
56	w425	112.50	B	A	A	A	A	A	U	B	
57	m532	114.00	B	A	A	A	A	B	U	B	
58	g17311	120.70	B	A	A	A	A	B	U	B	
59	w157	121.50	B	A	A	A	A	B	U	B	
60											
61	name="chrom2"										
62	w51	0.00	B	B	B	A	A	B	U	A	
63	w204	3.30	A	B	B	A	A	B	U	A	
64	w122	4.00	A	B	B	A	A	B	U	A	
65	g3843	6.90	A	B	B	A	A	B	U	A	
66	w301	7.60	A	B	B	A	A	B	U	A	B

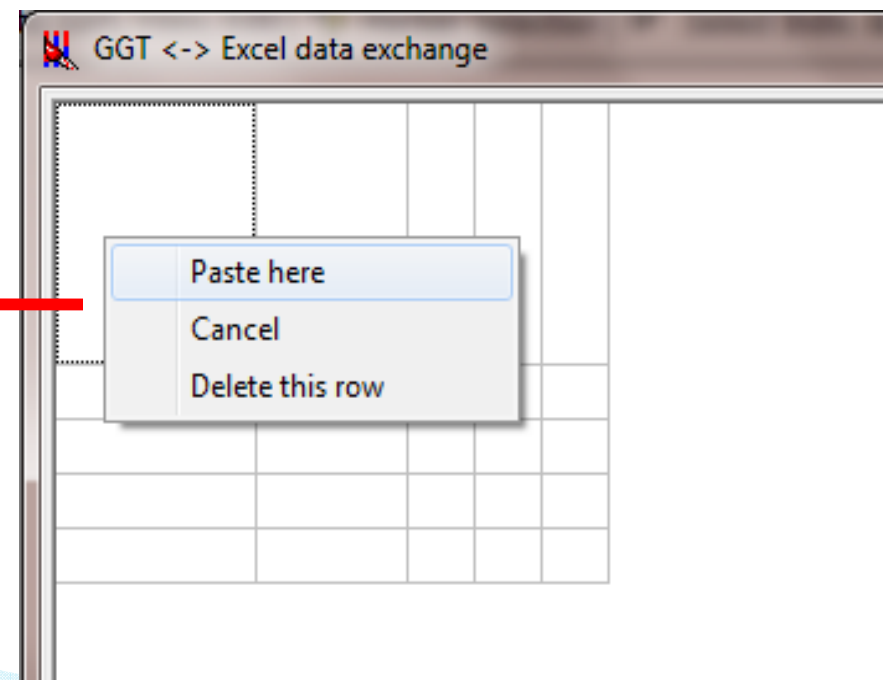
Space
between
each
chromosome



	A	B	C	D	E	F	G	H
1	alias		geno1	geno2	geno3	geno4	geno5	
2								
3	nchrom	2						
4								
5	popt	F2						
6	locus	pos						
7								
8	name=group1							
9	marker1	0	33	40	33	56	33	
10	marker2	25.8	A	?	A	H	B	
11	marker3	76.9	all1	all2	?	all2	all3	
12	marker4	101.3	A	B	H	A	A	
13								
14	name=group2							
15	marker21	15.5	H	B	A	H	A	
16	marker22	43.7	all1	all2	all3	all4	all2	
17	marker23	81	A	?	A	H	B	
18								
19								
20								

This is how to organize your data in an Excel spread Sheet.





Right Click



Input Data using EXCEL

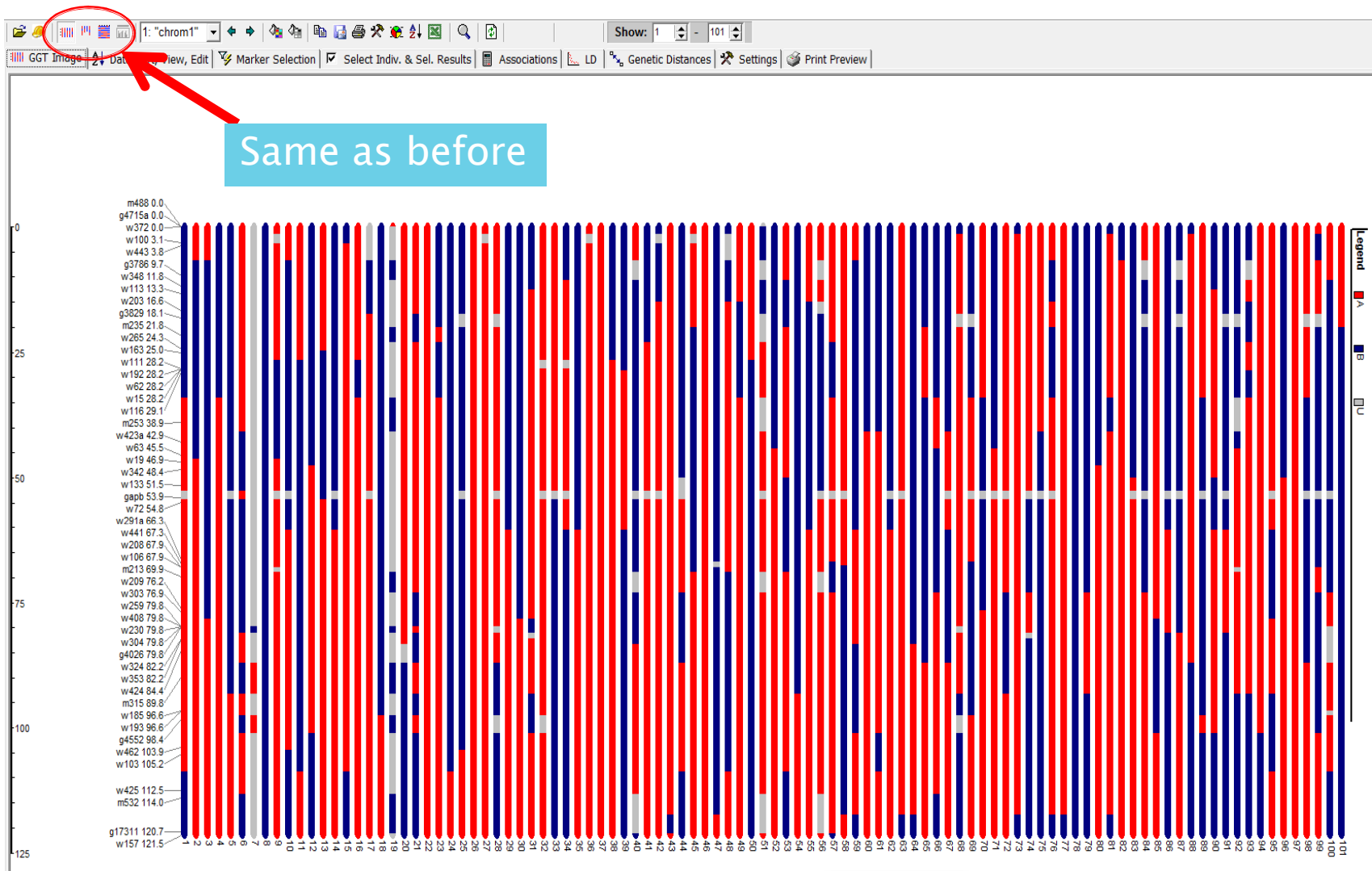
w265	11.00	A	A	A	A	A	A	A	A	A	A	A	A
w163	12.00	A	A	A	A	A	A	A	A	A	A	A	A

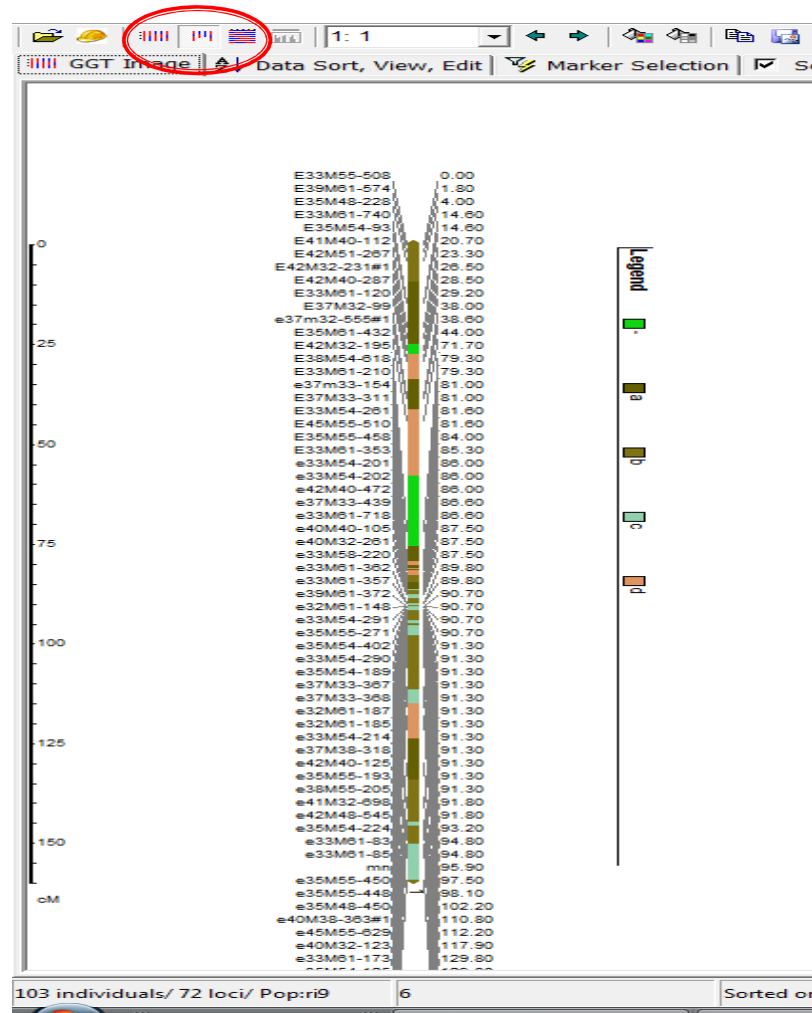
Show Headers: ☒ Yes ☐ No ☐ Impute map positions

 Open Excel File  Paste Excel data  Recode  Import into GGT

Open file or paste. Either way works

Import

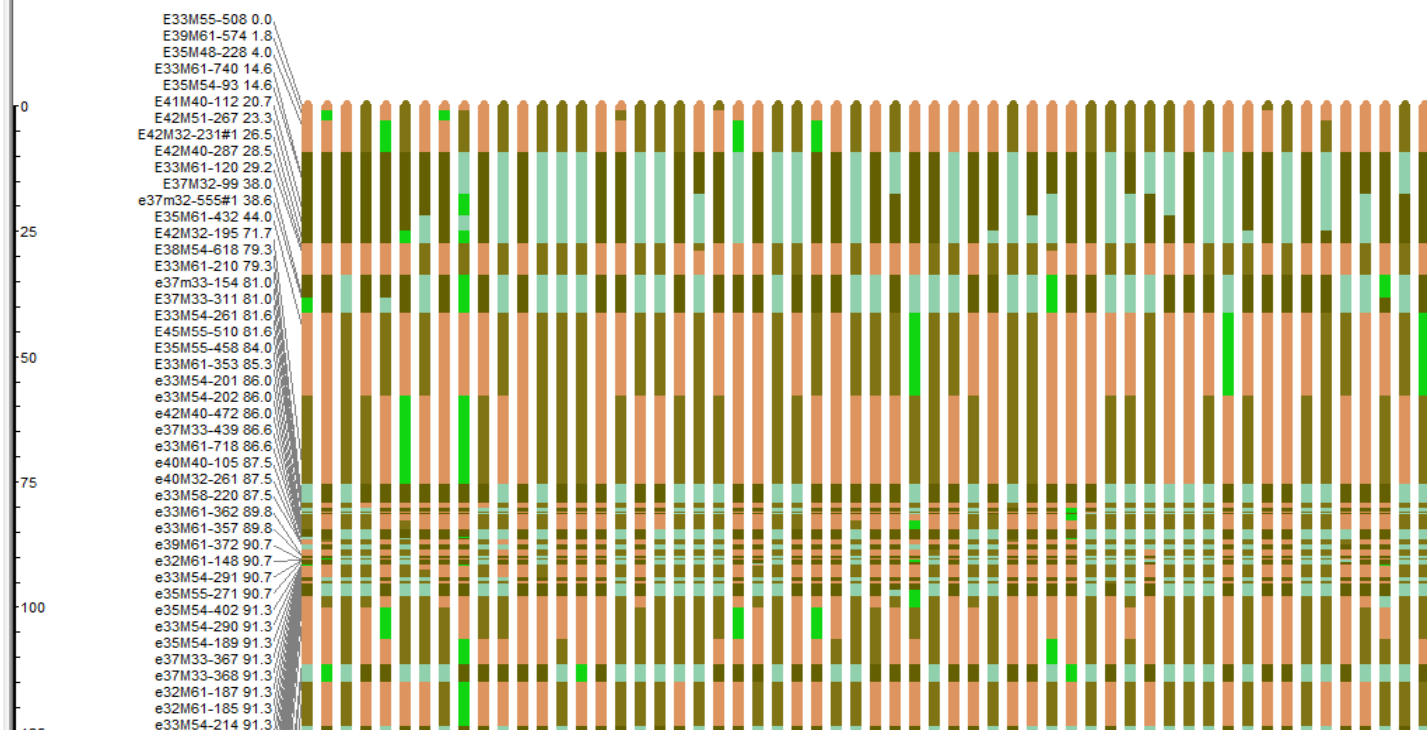




Data for one chromosome, one individual. Used to check a specific genotype for all markers.



GGT Image Data Sort, View, Edit Marker Selection ☒ Select Indiv. & Sel. Results Associations LD Genetic Distances Settings Print Pre



Manipulation of data

The screenshot shows the GGT Image software interface. The 'Data Sort, View, Edit' menu is circled in red. The 'Sort individuals' dialog box is open, showing options for sorting. The 'Selected group' option is selected. The 'overall allele content' dropdown menu is open, showing alleles: -, a, b, c, and d. The 'Sort' button is highlighted. The 'Refresh' button is also visible. The 'Update' button is at the bottom right.

File Edit Export Statistics Marker Selection Advanced Analyses Settings Help

GGT Image Data Sort, View, Edit Marker Selection Select Indiv. & Sel. Results Associations

2: 1 Ind: 11

Show & edit data in gridview

Sort individuals

- ☐ No sort /undo sort
- ☐ All markers
- ☐ Selected marker
- ☒ Selected group
- ☐ Alias
- ☐ overall allele content:

Note: Updates will be visible

Markername

E33M61-740

Position (cM)

14.6

QTL interval

- -

Update

Sort

Refresh

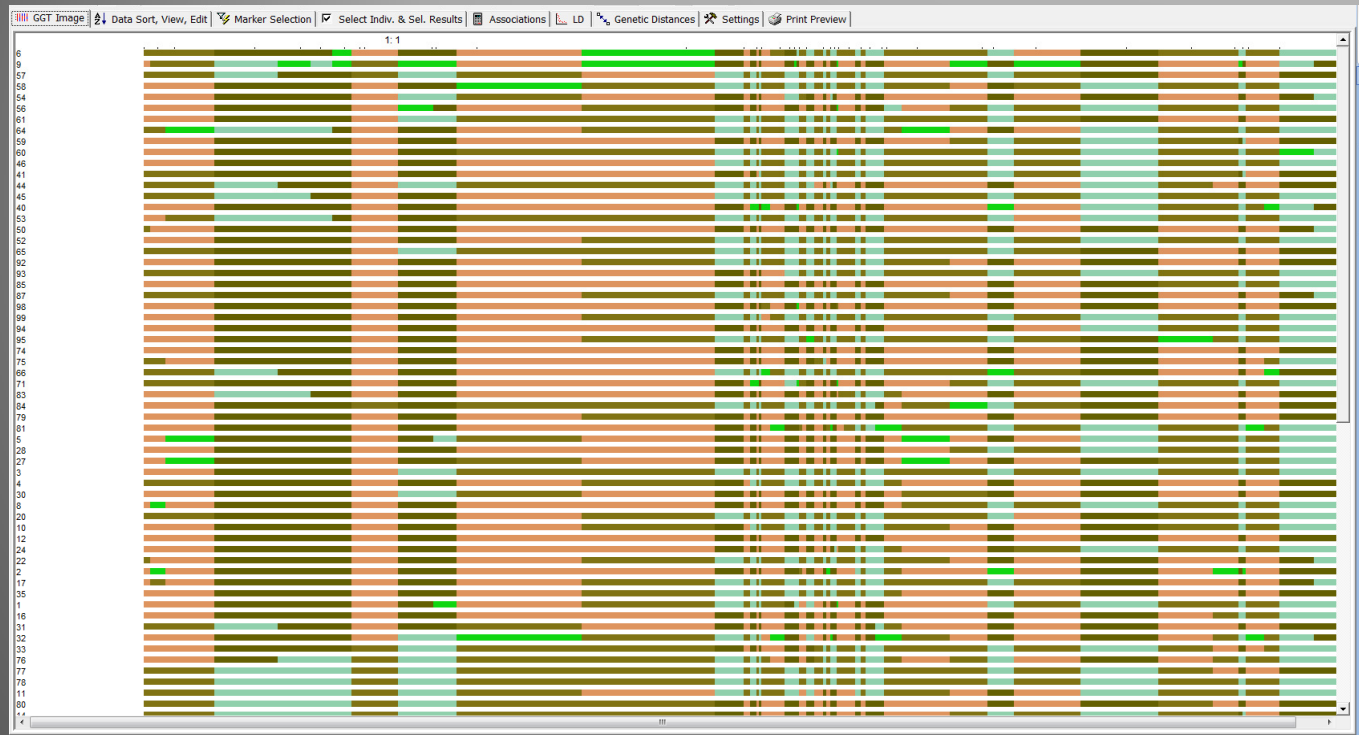
Options to view and edit data

First, decide how you want to sort data

Second, choose allele to sort by

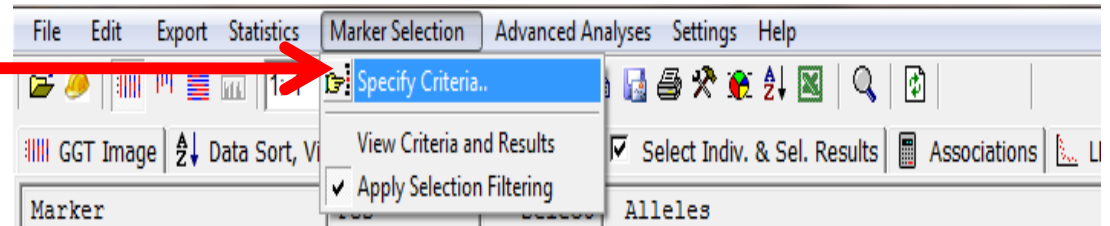
Three, Sort and Refresh

Results after data are sorted



Manipulation of data (continued)

Setting selection Criteria



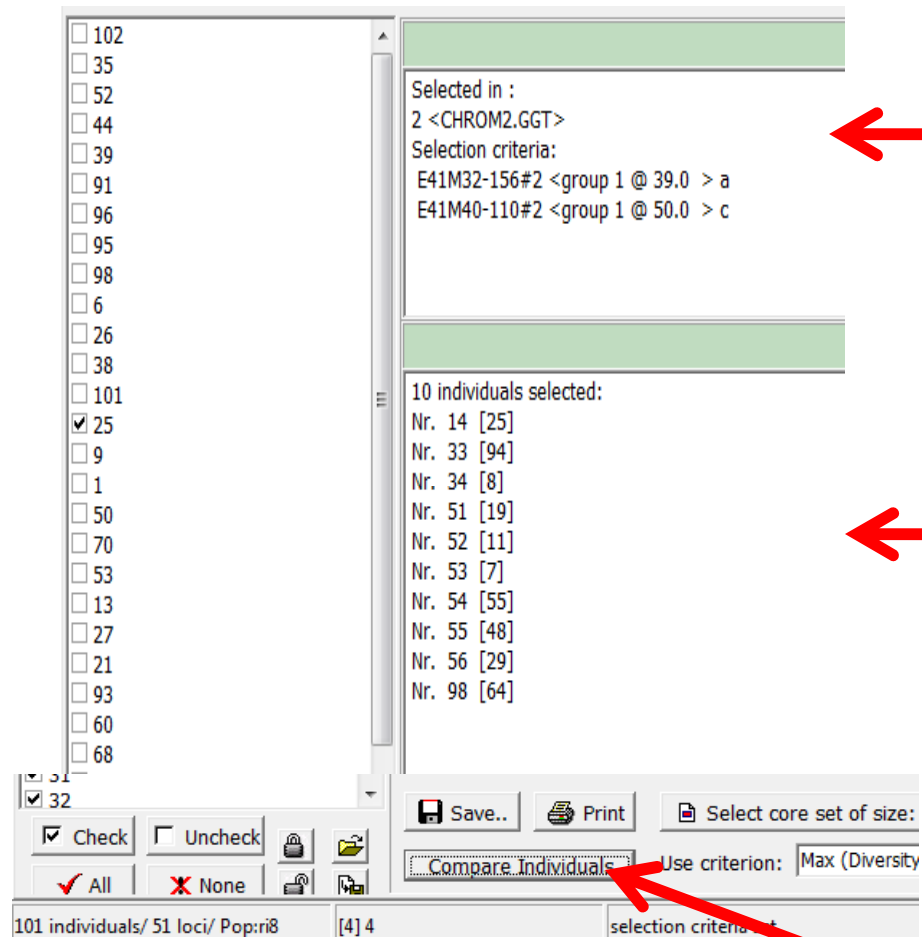
*w113	13.3	A11	BBBBBA-BABABABAABB-AAABBBAAABBAABBB
*w203	16.6	A11	BBBBBA-BABABABAABB-AAABBBAAABBAABBB
✓ *g3829	18.1	A	BBBBBA-BABABABAABB-AAABBBAAABBAABBB
*m235	21.8	A11	BBBBBA-BABABABAABB-AAABBBAAABBAABBB
*w265	24.3	A11	BBBBBA-BABABABAABB-AAABBBAAABBAABBB
*w163	25.0	A11	BBBBBA-BABABABAABB-AAABBBAAABBAABBB
*w111	28.2	A11	BBBBBA-BBBBBBBABAB-AAABBBAAABBAABBB
✗ *w192	28.2	!B	BBBBBA-BBBBBBBABAB-AAABBBAAABBAABBB
*w62	28.2	A11	BBBBBA-BBBBBBBABAB-AAABBBAAABBAABBB
*w15	28.2	A11	BBBBBA-BBBBBBBABAB-AAABBBAAABBAABBB
*w116	29.1	A11	BBBBBA-BBBBBBBABAB-AAABBBAAABBAABBB
*w192	28.2	A11	BBBBBA-BBBBBBBABAB-AAABBBAAABBAABBB

A red arrow points from the 'A' allele in the row for *g3829 to the 'Select individual with the A Allele for this marker' text box.

Select individual with the A Allele for this marker

Do not have B allele for this marker

Manipulation of data (continued)

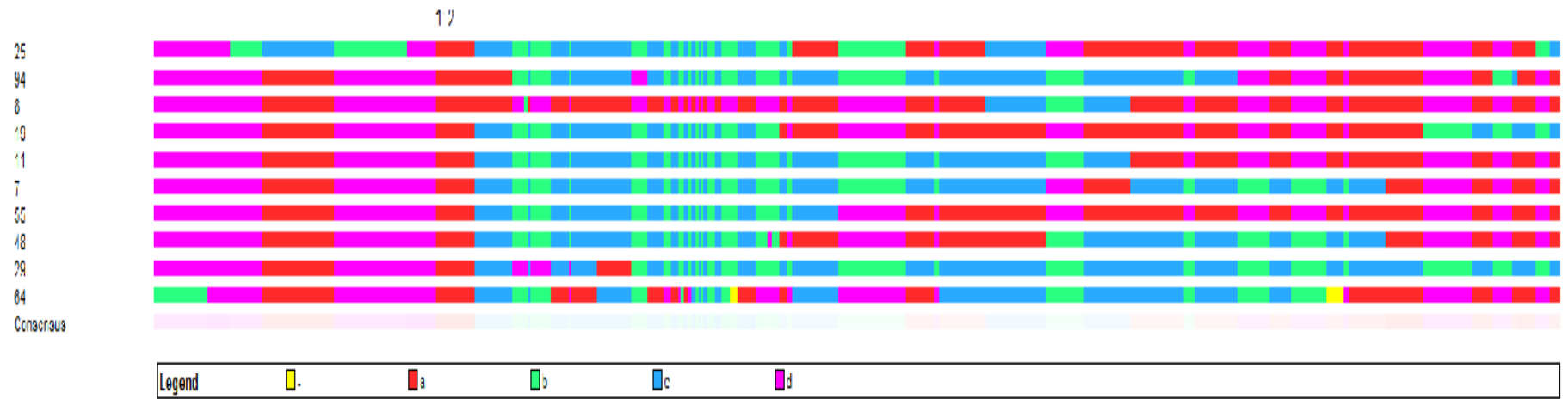


Criteria
selected

Individual
selected

And Click here!

Results (unsorted)



Sort Data

Show & edit data in gridview

Genome

group 1: "2"

Sort individuals

☐ No sort /undo sort

☐ All markers

☒ Selected marker

☐ Selected group

☐ Alias

☐ overall allele content: a (a)

Sort !

Note: Updates will be visible after redrawing

Markername

E41M32-156#2

Refresh tree

Position (cM)

39.0

QTL interval

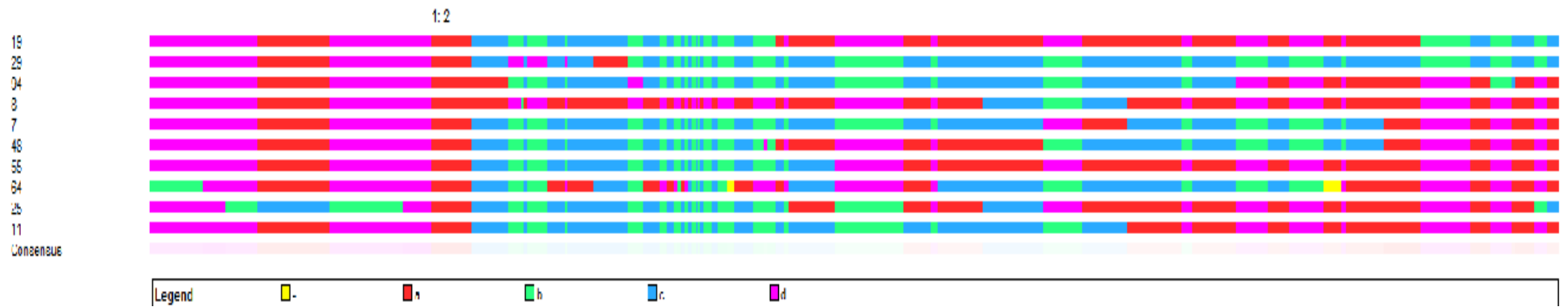
-

-

Update

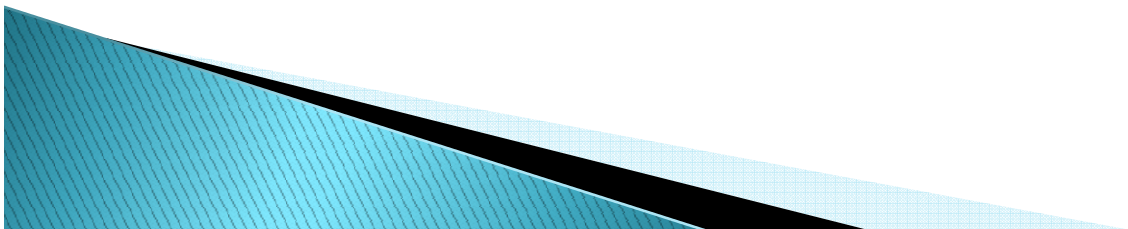
Save GGT data..

Results after data are sorted



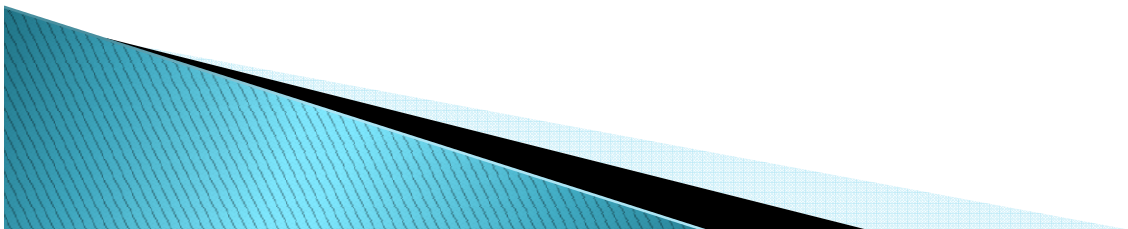
Conclusions

- ▶ In many cases the GGT software package can assist in the exploration, visualization and analysis of molecular marker data in applications in plant (and animal) science.



What else can GGT do

- ▶ Genetic distance
- ▶ Linkage disequilibrium (LD)
- ▶ Summary statistics by Individual or Marker (background genome selection)



References

- ▶ Young, N. D. and S. D. Tanksley. 1989. Restriction fragment length polymorphisms maps and the concept of graphical genotypes. Theoretical and Applied Genetics. 77(1): 95–101.
- ▶ Van Berloo, Ralph. 2008. GGT 2.0: Versatile software for visualization and analysis of genetic data. Journal of Heredity. 99(2):232–236.
- ▶ Van Berloo, Ralph. 1999. GGT: software for the display of graphical genotypes. J Hered 90: 328–329.



Thank you ...Questions?

Show & edit data in gridview

GGT <-> Excel data exchange

alias

rightclick to paste data in place

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
popt	ri9																									
nchrom	1																									
nind	103																									
autoformatted GGT																										
LOCUS	POS																									
name=1																										
E33M55-508	0.00	d	d	d	b	d	b	d	d	d	b	d	b	b	b	d	d	b	b	b	d	b	d	d	b	
E39M61-574	1.80	d	-	d	b	d	b	d	-	b	d	b	d	b	b	b	d	b	b	b	b	d	d	d	d	b
E35M48-228	4.00	d	d	d	b	-	b	d	d	b	d	b	d	b	b	b	d	d	b	b	b	d	d	-	d	b
E33M61-740	14.60	a	a	a	a	a	a	a	a	c	a	c	a	c	c	c	a	a	c	c	a	a	a	c	a	c
E35M54-93	14.60	a	a	a	a	a	a	a	a	c	a	c	a	c	c	c	a	a	c	c	a	a	a	c	a	c
E41M40-112	20.70	a	a	a	a	a	a	a	a	-	a	c	a	c	c	c	a	a	c	c	a	c	a	c	a	c
E42M51-267	23.30	a	a	a	a	a	a	c	a	c	a	c	a	c	c	c	a	a	c	c	a	c	a	c	a	c
E42M32-231#1	26.50	a	a	a	a	a	-	c	a	-	a	c	a	c	c	c	a	a	c	c	a	c	a	c	a	c
E42M40-287	28.50	d	d	d	d	d	b	d	b	d	b	d	b	b	b	d	d	b	b	d	b	d	d	d	d	b
E33M61-120	29.20	d	d	d	d	d	d	b	d	b	d	b	d	b	b	b	d	d	b	b	d	d	d	d	d	b
E37M32-99	38.00	a	a	c	a	a	a	c	a	-	a	c	a	c	c	c	a	a	c	c	a	c	a	a	a	c
e37m32-555#1	38.60	-	a	c	a	c	a	c	a	-	a	c	a	c	c	c	a	a	c	c	a	c	a	a	a	c
E35M61-432	44.00	d	d	b	d	b	d	b	d	d	b	d	b	b	b	d	d	b	b	d	b	d	d	d	d	b

Show Headers: ☐ Yes ☒ No

☐ Impute map positions

Open Excel File

Paste Excel data

Recode

Import into GGT

Save Excel File

Clear

Close