Appendix B - Using WinQTLCart resources for Simulation

WinQTLCart is a package widely used for adjustment of the models in genetic mapping as Simple Marker Mapping, Composite Interval Mapping, Multiple Interval Mapping for controlled crossing as backcross and F2 Inbred designs.

In this work, WinQTLCart was used to generate genetic mapping data (lines and markers), for simulation and also for analysis of real data (for instance, the INCOR project). We adopted mapping function of Haldane, a single quantitative trait, average trace equal to 130, design F2, number of QTL equal two additive effects and epistasis. Six sets of data were used in 8 simulated scenarios considering the following parameters: distribution of trace $N(\mu, \sigma^2)$, whose mean and variance of the trait for the simulated data sets are presented in Table B.1, heritability equal to 0.8, Vi / Va equal to 0.2, different sample sizes, numbers of chromosomes, markers per chromosome and genome size, distance between markers equal to 4cm as shown in Figures B.1, B.2 and Table B.2.

Tabela B.1. Sample size (n), mean and variance of the trace of the simulated data sets

Scenarios	n	mean	variance
1	50	130,2178	1,9102
2	50	130,2178	1,9102
3	50	130,2178	1,9102
4	200	130,0344	2,3747
5	200	130,0679	2,3611
6	200	130,0426	2,0412
7	200	130,0971	2,3584
8	200	129,8134	2,2654

Note in Table B.1 that Scenarios 1, 2 and 3 were obtained from a same data set simulated by using WinQTLCart but changing the following genetic algorithm parameters: the probability of mutation, 1to 0.4 (Scenarios 1 and 2) and number of generations per solution from 100 to 1000 (Scenarios 2 and 3), as shown in Table B.2.

Table B.2. Data considered in simulation studies.

SITUATIONS	MAPA	Genoma (cM)	n	pm	ns	ng
1	**2Chr 4M/Chr**	24	50	0,1	100	100
2	**2Chr 4M/Chr**	24	50	0,4	100	100
3	**2Chr 4M/Chr**	24	50	0,4	100	1.000
4	**2Chr 4M/Chr**	24	200	0,4	100	1.000
5	**2Chr 4M/Chr**	24	200	0,4	100	1.000
6	**2Chr 4M/Chr**	24	200	0,4	100	1.000
7	**10Chr 4M/Chr**	120	200	0,4	100	1.000
8	**10Chr 10M/Chr**	720	200	0,4	100	1.000

For Situations 4, 5 and 6 different replicates were generated from the same data set without changing the model parameters but using only different "seeds" and, as expected, it is observed in Table B.1 only slight changes to the mean and the variance of the trait.

In Situations 7 and 8 larger genomes were used only changing the number of chromosomes from 2 to 10 (Situation 6 to 7) and the number of markers per chromosome from 4 to 10 (Situation 7 to 8).

Figure B.1 illustrates the first step in setting WinQTLCart so one can generate a dataset to be used in simulations.

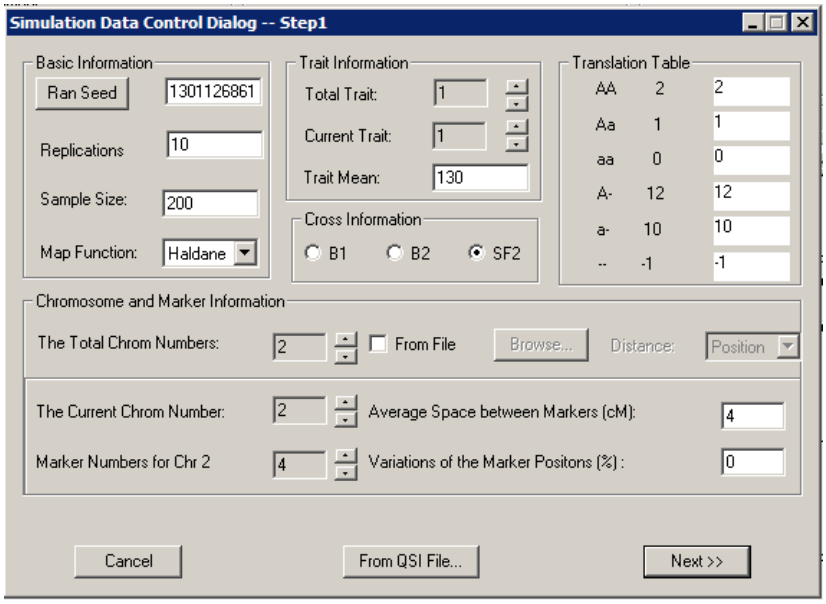

Figure B.1. Parameters in the WinQTLCart – part 1

Note in Figure B.1 that this example is being made to sample size 200, two chromosomes with four markers per chromosome, with distance between markers of 4cM and the other parameters previously mentioned. In Figure B.2 application is presented in another graphical

application window showing the values for some parameters adopted in the simulations. The parameters considered were two QTL additive main effects and their interaction, signal same direction, normal distribution, heritability 0.8, Vi / Va = 0.2.

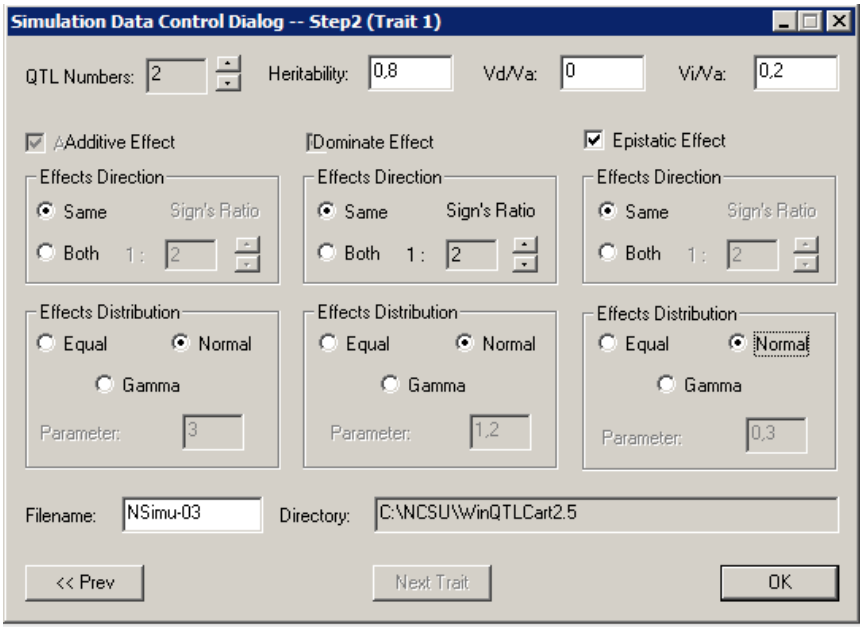

Figure B.2. Parameters do WinQTLCart – part 2.

Figure B.3 shows other parameters adopted in the simulations. It can be seen in Figure B.3 that the results obtained for the additive main effects were 0.8374 and -0.7950 at positions 2.6 cM on chromosome 1 (QTL 1) and 8.4 cM on chromosome 2 (QTL 2) respectively and the effect of epistasis was 0.6527.

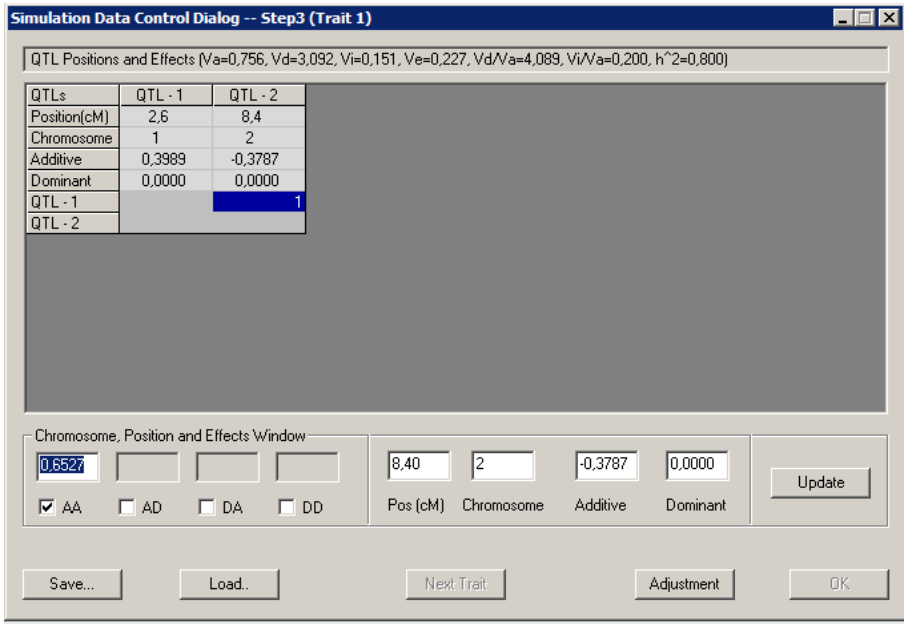

Figure B.3. Results WinQTLCart obtained by application to the additive effects and interaction.

These commands generate the data file shown below that contains data markers, quantitative trait and positions according to the parameters established in the simulation, creating a program file that contains the data mapping, markers, distance, sample size, considered design, which in this case was F2, and the data set used as an example. Table B.2 shows the output of the program considering the data simulated in Scenario 3

Tabela B.2. Programe Exemple

```
#FileID 1847488217
#bychromosome
-type position
-function 1
-Units cM
-chromosomes 2
-maximum 4
-named yes
-start
-Chromosome C1
MK-1-1        0,0000
MK-1-2        4,0000
MK-1-3        8,0000
MK-1-4        12,0000
-Chromosome C2
MK-2-1        0,0000
MK-2-2        4,0000
MK-2-3        8,0000
MK-2-4        12,0000
-stop
----------------------------------------------------
#bycross
-SampleSize 200
```

-Cross SF2
-traits 1
-missingtrait .
-case yes
-TranslationTable
AA 2 2
Aa 1 1
aa 0 0
A- 12 12
a- 10 10
-- -1 -1
-start markers
MK-1-1 1 2 0 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 2 2 1 1 0 0 1 1 1 0 2 1 0 0 1 2 1 1 1 1 2 1 1 2 0 0 2 2 2 1 1 1 1 1 1 1 1 0 0 1 0 1 0 1 2 1 1
1 1 2 12 1 1 2 2 2 2 0 0 1 1 1 1 1 1 1 2 2 1 0 1 1 2 1 0 1 2 0 0 1 1 0 0 1 11 0 0 1 1 1 2 2 1 0 1 1 1 1 2 1 1 2 2 1 0 1 0 0 1 2 2 1 1 2 1 1
2 2 0 1 0 0 0 2 1 1 2 0 1 1 2 1 1 1 1 2 1 1 1 0 2 1 1 1 1 1 1 1 1 1 1 0 2 2 1 2 1 2 1 0 1 1 1 2 1 2 2 1 0 0 0 0 1 0 0 1 0 1 1 2
MK-1-2 1 2 1 1 1 1 1 2 1 1 1 1 1 1 1 1 0 1 1 2 2 1 1 0 1 1 1 1 0 2 1 0 0 2 2 2 1 1 1 2 1 1 2 0 0 2 2 2 1 1 1 1 1 1 1 1 1 0 1 1 0 1 0 1 2 1 1
1 1 2 12 0 1 2 2 1 2 0 0 1 1 1 1 1 1 1 2 2 1 0 1 1 2 1 1 1 2 0 0 1 1 0 0 1 11 0 0 1 1 1 1 2 1 0 1 0 1 1 1 1 0 2 2 1 0 1 0 0 1 1 2 1 1 2 1 1
1 2 0 1 0 0 0 2 1 1 2 0 1 1 2 1 1 1 1 2 1 1 1 0 2 1 1 1 1 1 1 1 1 1 1 0 2 2 12 1 2 1 0 1 0 1 2 1 2 2 1 0 0 0 1 1 0 0 1 0 1 2 2
MK-1-3 1 2 1 1 1 1 1 2 1 1 1 1 1 1 1 1 0 1 2 1 2 1 0 0 1 2 1 1 0 2 1 0 1 2 2 2 1 1 1 2 1 1 2 0 0 2 2 2 1 1 1 1 1 1 1 1 1 0 0 2 0 1 0 1 2 1 1
2 1 2 12 0 1 2 2 1 2 0 0 1 0 1 1 1 1 1 2 2 1 0 1 1 2 1 1 0 2 1 1 1 1 0 0 1 2 1 0 0 1 1 1 1 2 1 0 1 0 1 1 1 1 0 2 2 1 0 1 0 0 2 1 2 1 1 2 1 0
1 2 0 2 0 0 0 2 11 2 0 1 1 2 1 1 1 1 2 1 1 1 0 2 1 1 1 1 1 1 1 1 1 0 2 2 12 1 2 1 0 1 0 1 2 2 1 1 1 0 0 0 1 1 0 0 1 1 1 2 2
MK-1-4 1 1 1 1 1 1 1 2 1 1 1 1 1 0 1 1 0 1 2 1 2 1 0 0 1 2 1 2 0 2 0 0 1 2 2 2 1 1 1 2 1 1 2 0 0 2 2 2 1 1 1 1 1 1 1 1 1 0 0 2 0 1 0 1 2 1 1
2 1 2 12 0 1 2 2 1 2 0 0 1 0 1 1 1 1 1 2 2 1 0 1 1 2 1 1 0 2 1 1 1 1 0 0 2 2 1 0 0 1 1 1 1 2 1 0 1 0 1 1 1 1 0 2 2 1 0 1 0 0 2 1 2 1 1 2 1 0
1 2 0 2 0 0 0 2 11 2 0 1 1 2 1 1 1 1 2 1 1 1 0 2 1 1 1 1 1 1 1 1 1 0 2 2 1 2 1 2 1 0 1 0 1 2 2 1 1 1 0 0 0 1 1 1 0 1 1 1 2 2
MK-2-1 0 1 1 1 1 1 0 0 1 1 2 0 1 0 1 1 1 2 2 0 2 1 2 1 0 0 2 1 1 0 1 0 0 1 1 0 1 1 0 0 2 1 1 0 2 1 0 2 0 1 1 2 0 2 0 1 1 1 1 2 1 1 0 1 0 2
1 1 2 11 1 1 2 2 1 1 1 2 1 0 2 2 1 0 1 1 1 2 0 1 1 1 2 2 1 1 0 2 1 0 2 1 2 1 2 1 0 1 2 0 0 1 0 1 1 1 1 1 1 0 0 1 1 2 0 0 1 0 1 0 1 1 1 2 1 1
0 1 0 1 0 2 1 0 02 1 0 0 1 0 2 1 0 1 2 1 0 2 0 0 1 1 2 2 1 2 2 1 1 1 2 1 11 2 2 1 1 2 1 0 1 1 1 2 1 1 2 2 1 0 1 1 1 2 1 1 1
MK-2-2 0 1 1 1 1 1 0 0 1 1 2 0 1 0 1 1 0 2 2 0 2 1 2 1 0 0 2 1 1 0 1 0 0 1 10 1 1 0 0 1 2 1 0 2 1 0 2 0 1 1 2 0 2 0 1 1 1 1 2 1 1 0 1 0 2
1 1 1 11 1 1 2 2 1 1 1 2 1 0 2 2 1 0 1 1 1 2 0 1 1 1 2 2 1 1 0 2 1 0 2 1 2 12 1 0 1 2 0 0 1 0 1 1 1 1 1 1 0 0 1 1 2 0 0 1 0 0 0 1 1 1 2 1 1
0 1 0 1 0 2 0 0 0 2 1 0 1 2 0 2 1 0 1 2 1 0 2 0 0 1 1 2 2 1 2 2 1 1 1 1 1 1 1 2 2 1 1 2 1 1 1 1 1 2 2 2 2 2 1 0 1 1 1 2 1 1 1
MK-2-3 0 0 1 1 1 1 0 0 1 1 2 0 0 0 1 1 0 2 2 0 2 1 2 1 0 0 2 1 1 1 1 0 0 1 10 1 2 0 0 1 2 1 0 2 1 0 2 0 1 1 2 0 2 0 1 1 0 1 2 1 1 0 1 0 2
1 1 1 11 1 1 2 2 1 0 1 2 1 0 2 2 1 0 2 1 1 2 0 1 1 2 2 1 1 0 2 1 1 2 1 2 1 1 1 0 1 2 0 0 1 1 1 1 1 1 1 1 0 0 0 1 2 0 0 1 0 0 0 1 1 1 2 1 1
0 1 0 1 0 2 0 0 0 2 1 0 1 2 0 2 1 0 1 1 1 0 1 1 0 1 1 2 2 1 2 2 1 1 1 1 1 10 2 2 1 1 2 1 1 1 1 1 2 1 2 2 2 1 0 1 0 1 2 0 1 1
MK-2-4 0 0 1 1 1 0 0 0 1 1 2 1 0 0 1 1 0 2 2 0 1 0 2 1 0 0 2 0 1 1 1 0 0 1 1 0 1 2 0 0 1 2 1 0 2 1 1 1 0 1 1 2 0 2 0 1 1 0 1 2 1 1 0 1 0 2
1 1 1 1
1 2 1 2 2 1 0 1 1 1 0 2 2 1 0 2 1 1 2 0 1 1 0 2 2 1 1 0 2 1 1 2 1 2 11 1 0 1 2 0 0 1 1 1 1 1 1 1 1 0 0 0 1 2 0 0 1 0 0 0 1 1 1 2 1 1 0 1 0 1
0 2 0 0 0 2 1 0 1 2 0 2 0 0 1 2 1 0 1 1 0 1 1 2 2 1 2 2 1 1 1 1 1 10 2 2 1 1 2 1 1 1 1 1 2 1 2 2 2 1 0 1 0 1 2 0 1 1
-stop markers
-start traits
Trait_1 131,9940 130,7484 131,3521 128,5802 128,7472 130,3351 129,8283 132,0139 129,5869 130,9772 129,1793 128,5120
131,0448 131,2661 128,8886 130,6391 133,1426 129,0078 129,5822 131,6199 129,7404 132,0338 130,0165 128,3421 131,5324
129,2091 127,6299 129,3889 128,9833 130,0391 130,5834 131,4492 131,5198 130,3906 129,8420 132,2192 129,1380 129,8979
130,1959 129,7802 131,9444 129,6076 131,0014 131,2172 125,9070 130,7371 129,5586 130,8852 131,0147 131,0435 127,5215
128,7773 131,6037 128,1901 130,1424 128,9657 130,1825 130,4901 128,3934 127,7508 130,3243 130,0199 132,4366 130,9627
130,2404 130,8178 130,8646 131,2904 129,7937 130,7150 129,8875 129,6072 129,2912 131,6177 130,4768 131,3124 130,6107
129,1639 126,4407 130,0504 129,4699 129,1771 126,3168 129,1988 130,3805 128,5273 129,0172 131,0937 127,3754 131,3418
131,8140 129,2276 129,3152 129,0514 128,8757 130,8212 131,5413 130,6854 128,2522 132,5050 128,7144 128,4959 126,3029
129,1718 129,6085 128,9654 128,7651 129,3657 128,9095 129,8326 130,9783 133,1472 129,8994 130,4459 127,6629 128,3831
128,4493 130,3464 130,0399 130,5056 130,9018 130,6653 129,4594 131,6964 129,3750 133,0418 129,8345 127,9654 131,9436
130,2304 130,8582 132,1581 131,2746 131,1475 130,3875 129,3133 131,4878 130,7652 129,1631 132,0846 130,6342 131,9626
126,7441 132,9076 129,6490 130,2902 129,1181 133,0289 133,6041 129,7108 128,2045 130,0034 130,1895 129,5293 128,7901
129,8254 131,7827 131,4898 130,0283 130,4310 129,9360 130,9486 129,7418 129,3757 127,4970 128,2692 130,0324 130,3821
126,5187 130,2948 131,4497 130,4720 130,8047 131,6038 130,8219 129,9014 128,5378 131,7425 129,9561 130,3038 129,8178
129,3620 129,8667 129,1907 129,8779 130,4823 132,5341 130,1876 128,1727 126,1477 127,0707 128,7635 131,7183 128,9860
132,1795
131,0016 126,7163 130,4565 132,1242 131,1883
-stop traits
-quit
-end
#EReplicate 1

Data mapping and markers are saved in files of type text to be read and interpreted by the search for AG program coded in R as can be seen in Appendix C - Source Code Program Implemented in R.

The example program is generated in accordance with the map of markers shown in Figure B.4 below.

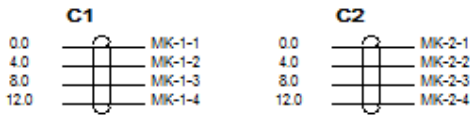

Figure B.4. Genetic Map of program example.

Note in Figure B.4 in this case was presented an example containing two chromosomes with four markers per chromosome and distance between markers of 4cm.