



Project started in december 2009

OptiMAS: a decision support tool to conduct Marker Assisted Selection (MAS) programs



F. Valente, F. Gauthier, J. Joets

A. Charcosset & L. Moreau



PAG Meeting

San Diego, CA, USA

January 14-18, 2012





Maize



Rice



Wheat



Sorghum

Initial user cases: 8 crops across 32 countries

Cereals

Maize Africa: Angola, Ethiopia, Kenya, Malawi, Mozambique, Tanzania, Uganda, Zambia, Zimbabwe

Asia: China, India, Indonesia, Thailand, The Philippines, Vietnam

Rice Africa: Benin, Burkina Faso, Ethiopia, Gambia, Ghana, Guinea, Liberia, Madagascar, Mali, Mozambique, Nigeria, Rwanda, Senegal, Tanzania, Uganda

Asia: Bangladesh, Cambodia, China, India, Indonesia, Laos, Nepal, Pakistan, Sri Lanka, Vietnam

Sorghum Mali

Wheat Africa: Ethiopia, Kenya

Asia: China, India

Legumes

Beans Ethiopia, Kenya, Tanzania, Malawi

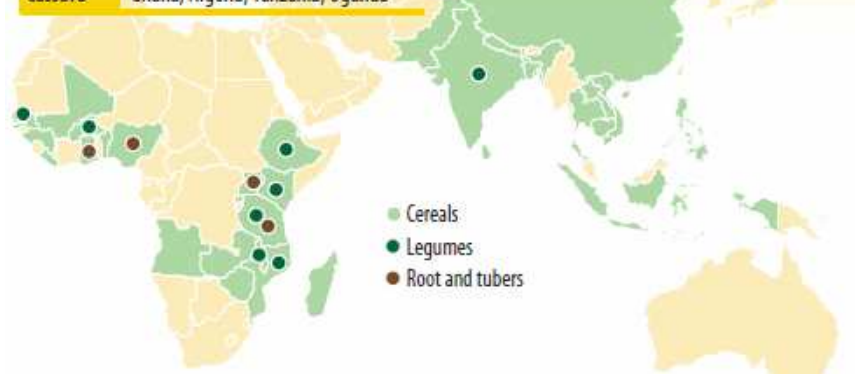
Chickpeas Africa: Ethiopia, Kenya

Asia: India

Cowpeas Burkina Faso, Mozambique, Senegal

Roots and tubers

Cassava Ghana, Nigeria, Tanzania, Uganda



Chickpeas



Beans



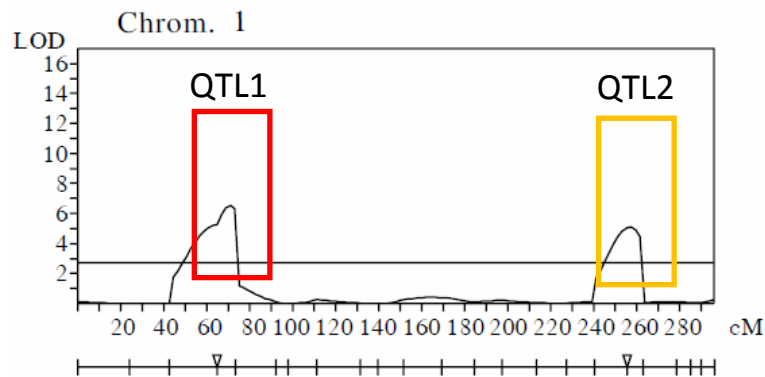
Cassava



Cowpeas

Breeding decisions adapted to different crops / projects

P1 (allele « a ») × *P2 (allele « b »)*



- QTL will be detected for different traits of interest and favorable alleles will be found

- Aim: to create a target genotype with all the favorable alleles at the QTL positions

Target genotype
(ideotype)

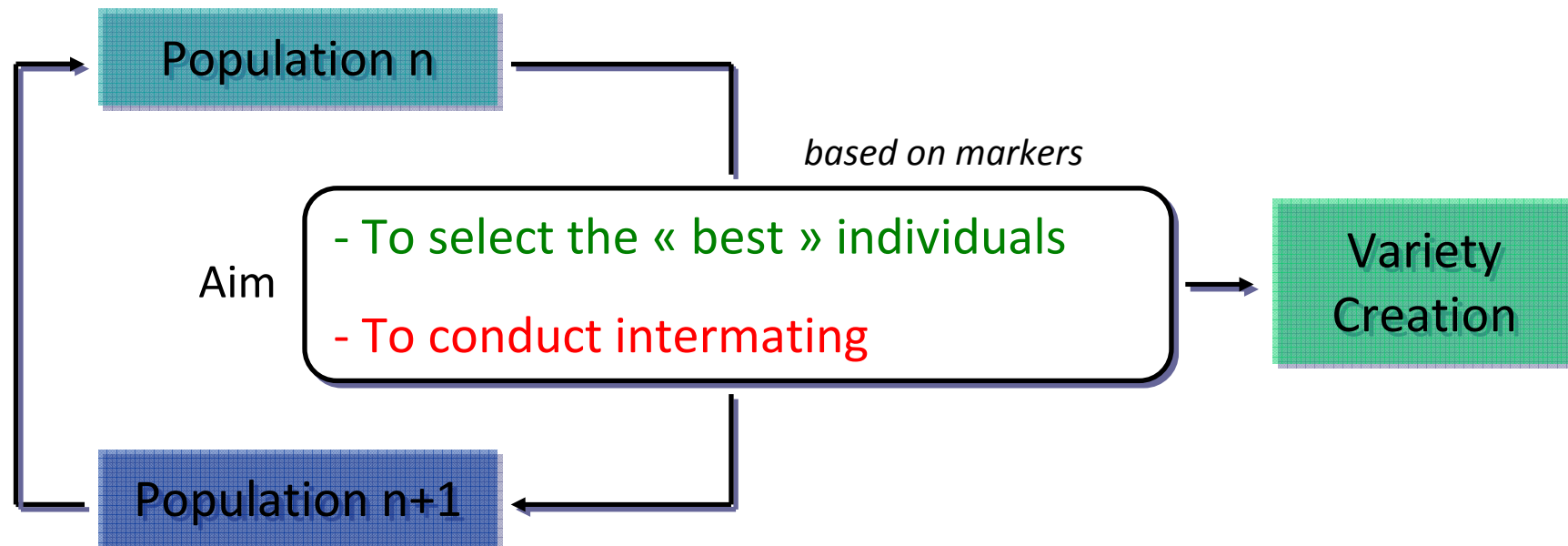


Mostly biparental population(s), but several cases where populations are connected through common parent(s)

Common objective: to create new genetic materials assembling favorable QTL alleles from 2 parents or several parents (get challenging as the No. of QTLs increases!)

→ Need user oriented tool to assemble favorable alleles through Marker Assisted Recurrent Selection (MARS) strategies

Implementation of Marker Assisted Recurrent Selection in OptiMAS

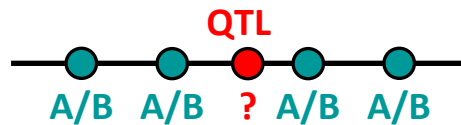


- ✓ **Objective 1:** to develop algorithms to compute probabilities of allele transmission through generations
- ✓ **Objectif 2:** to identify the best individuals
- ✓ **Objective 3:** to develop methodology to identify the best intermating scheme to accumulate favourable alleles and to extract varieties

Background elements on Objective 1: probabilities of allele transmission at QTLs

Aim: to evaluate individuals, we need to know **which parental alleles are transmitted at the QTL position**, based on the information from neighboring markers

QTL position rarely located at a marker → QTL alleles are unknown and must be **inferred from flanking markers**

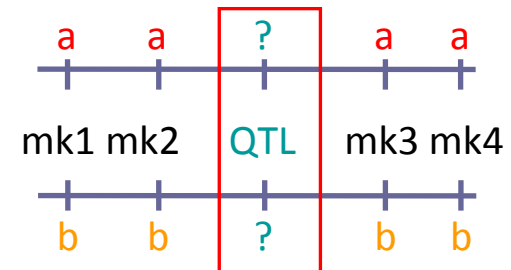


e.g 4 markers / QTL

(bi/multi allelic context)

Selection of a subset of **informative markers** to follow the favorable parental alleles (**based on haplotypes**)

Identification of all the possible **phased genotypes** for each QTL and computation of their **probabilities**



probabilities of allele transmission

- Information available (input files: pedigree/genotype & genetic map)
 - ✓ Pedigree
 - ✓ Molecular markers (observed genotypes)

<i>Id</i>	<i>P1</i>	<i>P2</i>	<i>Step</i>	<i>Cycle</i>	<i>Group</i>	<i>mrk1</i>	<i>mrk2</i>	<i>mrk3</i>
Fr2	a	a	IL	IL		A	A	A
Fr252	b	b	IL	IL		B	B	B
f1	Fr2	Fr252	CR	F1		A/B	A/B	A/B
ind1	f1	f1	S2	F3		A	A	A
ind2	f1	f1	S2	F3		A	A	A
ind3	f1	f1	S2	F3		A/B	-	A
ind4	f1	f1	S2	F3		B	B	A/B
ind5	f1	f1	S2	F3		B	B	B
ind6	f1	f1	S2	F3				
ind7	f1	f1	S2	F3		A	A	A

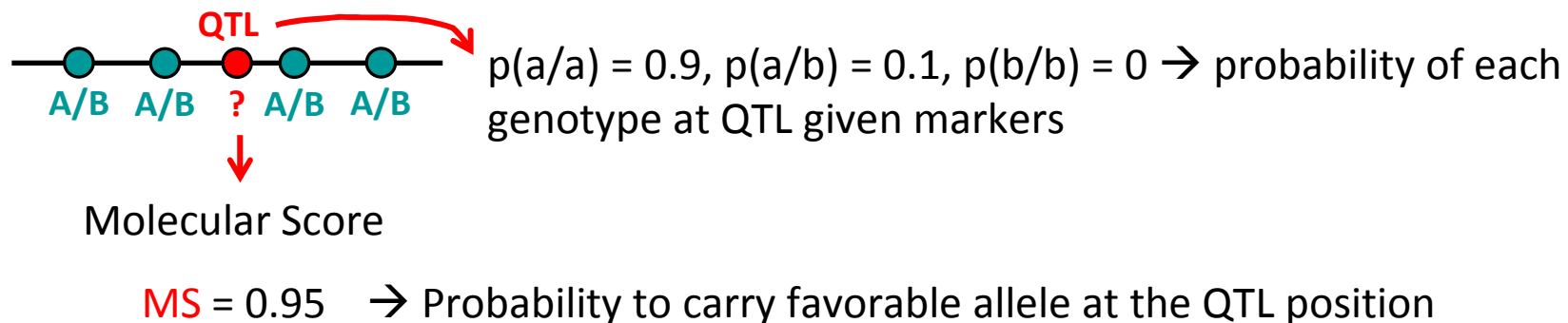
probabilities of allele transmission

- **Information available** (input files: pedigree/genotype & genetic map)
 - ✓ Pedigree
 - ✓ Molecular markers (observed genotypes)
 - ✓ **QTL position**
 - ✓ **Distance between loci** (Haldane → **recombination rate**)

Locus	Chr	QTL	Pos	All+
marker1	1	1	42.2	a
marker2	1	1	64.0	
qtl1	1	1	70.0	
marker3	1	1	72.5	
marker4	1	1	90.8	b
marker5	1	2	237.1	
marker6	1	2	252.2	
qtl2	1	2	254.0	
marker7	1	2	259.5	
marker8	1	2	274.8	

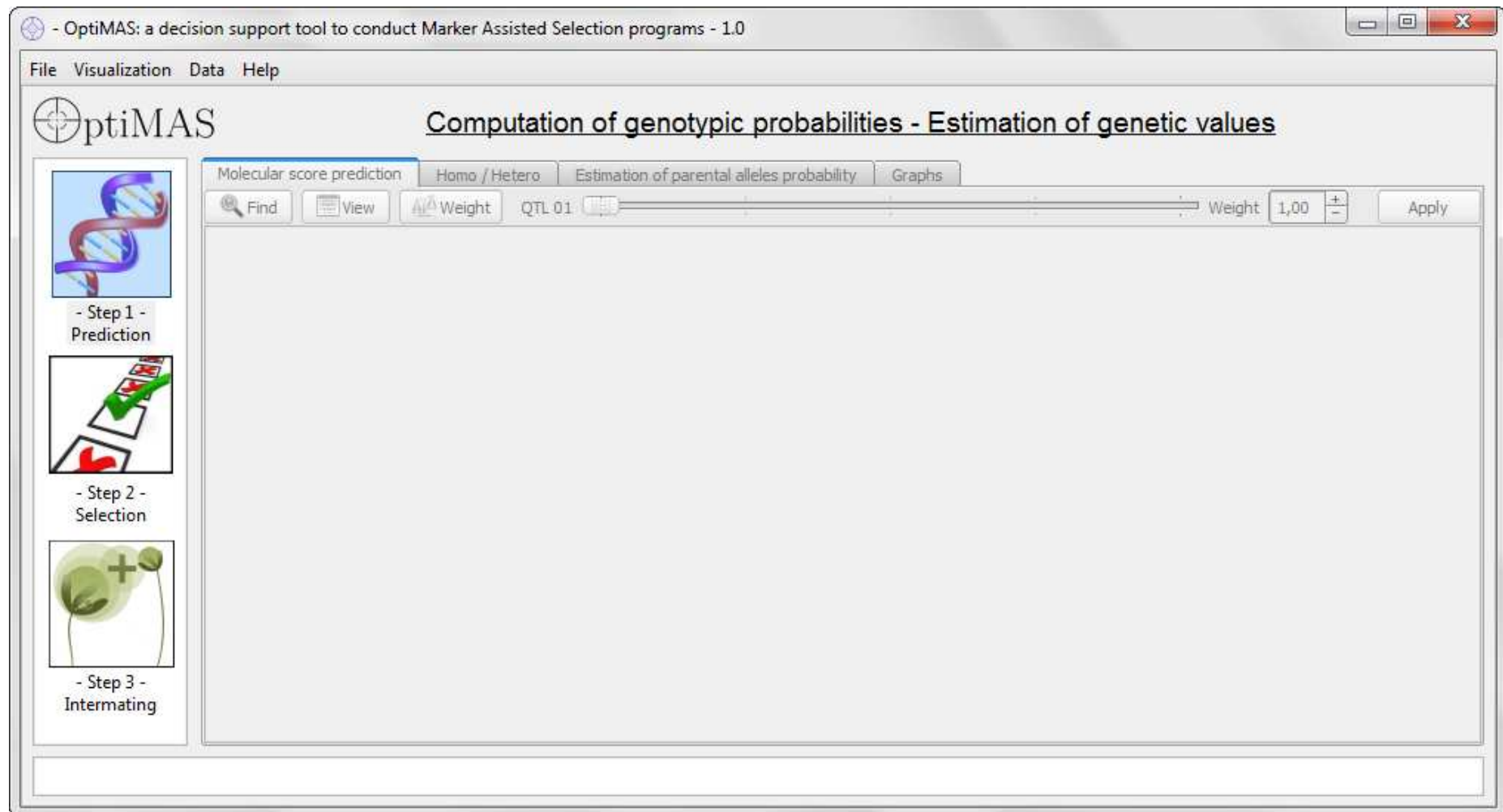
Objective 1: probabilities of allele transmission

- **Information available** (input files: pedigree/genotype & genetic map)
 - ✓ Pedigree
 - ✓ Molecular markers (observed genotypes)
 - ✓ QTL position
 - ✓ Distance between loci (Haldane \rightarrow recombination rate)
- **Algorithm: compute the probabilities of allele transmission** in different MAS schemes and mating designs (intercrossing, selfing, backcrossing, DH, RIL)
 - ✓ 2 versions: command line \rightarrow integrated in a new GUI









Presentation of mean features of OptiMAS

□ New Graphical User Interface (GUI) organized in 3 modules corresponding to the \neq steps of the selection program



Step 1: Computation of genotypic probabilities – Estimation of genetic values

Molecular score prediction																		
Homo / Hetero					Estimation of parental alleles probability							Graphs						
 Find	 View	 Weight	QTL 01 										Weight 1,00  		Apply			
Id	P1	P2	Cycle	Group	MS ^	Weight	UC	No.(+/+)	No.(+/-)	No.(+/-)	No.(?)	QTL1	QTL2	QTL3	QTL4	QTL5	QTL6	QTL7
B8	A1005	A1005	C2	-	0.8366	0.8366	10	8	1	0	2	0.0000	0.8598	0.9761	0.8752	0.8959	0.9745	0.9960
B158	A251	A1005	C2	G1	0.8024	0.8024	10	7	1	1	2	0.0000	0.8792	0.8925	0.9405	0.9731	0.9693	0.9960
B28	A1006	A251	C2	G1	0.7740	0.7740	9,7071	6	1	1	3	0.0000	0.9716	0.4938	0.9492	0.9137	0.9199	0.9797
B13	A1006	A1005	C2	-	0.7609	0.7609	9,7071	6	1	2	2	0.0000	0.8844	0.9559	0.9179	0.9731	0.4963	0.9895
B38	A1040	A1005	C2	-	0.7494	0.7494	9,366	6	1	1	3	0.0000	0.9349	0.4907	0.9429	0.5822	0.9260	0.9860
B37	A1040	A1005	C2	-	0.7433	0.7433	9	7	1	1	2	0.0000	0.9528	0.2334	0.9172	0.8735	0.9718	0.9860
B40	A1040	A1040	C2	-	0.7404	0.7404	9,366	7	1	2	1	0.0000	0.9775	0.4376	0.9592	0.4974	0.9231	0.9759
B242	A37	A1005	C2	-	0.7305	0.7305	9,366	6	1	1	3	0.0000	0.7259	0.8959	0.8879	0.2896	0.9722	0.9860
B124	A23	A167	C2	-	0.7223	0.7223	9,366	6	1	1	3	0.4812	0.6698	0.9559	0.9429	0.2787	0.9614	0.9924
B125	A23	A167	C2	-	0.7032	0.7032	8,7071	6	2	1	2	0.4812	0.6536	0.9559	0.9429	0.0058	0.9616	0.9924
A1040_ sf683	df37	C1	-	-	0.6615	0.6615	8,618	5	1	5	0	0.0000	0.9771	0.4842	0.9592	0.4974	0.9231	0.9759
A1040	A1040_	A1040_	-	-	0.6615	0.6615	8,618	5	1	0	5	0.0000	0.9771	0.4842	0.9592	0.4974	0.9231	0.9759
A1005_ df108	sf650	C1	-	-	0.6418	0.6418	8,618	5	1	4	1	0.0000	0.8928	0.4896	0.8751	0.4878	0.9747	0.9960
A1005	A1005_	A1005_	-	-	0.6417	0.6417	8,618	5	1	0	5	0.0000	0.8927	0.4895	0.8751	0.4878	0.9747	0.9960
A1003	A1003_	A1003_	-	-	0.6413	0.6413	8	6	3	0	2	0.0000	0.9757	0.9738	0.9135	0.0000	0.4986	0.9759
A1003_ df108	df37	C1	-	-	0.6413	0.6413	8	6	3	1	1	0.0000	0.9757	0.9738	0.9135	0.0000	0.4987	0.9759
A251_ sf683	df108	C1	-	-	0.6364	0.6364	8,618	3	1	5	2	0.0000	0.5014	0.4896	0.9544	0.4974	0.9090	0.9960

□ Individuals in lines, MS value [0-1] → 1 = ideotype

Step 1: Computation of genotypic probabilities – Estimation of genetic values

Molecular score prediction																		
Homo / Hetero																		
Estimation of parental alleles probability																		
Graphs																		
Find View Weight QTL 01 Weight 1,00 Apply																		
Id	P1	P2	Cycle	Group	MS ^	Weight	UC	No.(+/+)	No.(-/-)	No.(+/-)	No.(?)	QTL1	QTL2	QTL3	QTL4	QTL5	QTL6	QTL7
B8	A1005	A1005	C2	-	0.8366	0.8366	10	8	1	0	2	0.0000	0.8598	0.9761	0.8752	0.8959	0.9745	0.9960
B158	A251	A1005	C2	G1	0.8024	0.8024	10	7	1	1	2	0.0000	0.8792	0.8925	0.9405	0.9731	0.9693	0.9960
B28	A1006	A251	C2	G1	0.7740	0.7740	9,7071	6	1	1	3	0.0000	0.9716	0.4938	0.9492	0.9137	0.9199	0.9797
B13	A1006	A1005	C2	-	0.7609	0.7609	9,7071	6	1	2	2	0.0000	0.8					
B38	A1040	A1005	C2	-	0.7494	0.7494	9,366	6	1	1	3	0.0000	0.9					
B37	A1040	A1005	C2	-	0.7433	0.7433	9	7	1	1	2	0.0000	0.9					
B40	A1040	A1040	C2	-	0.7404	0.7404	9,366	7	1	2	1	0.0000	0.9					
B242	A37	A1005	C2	-	0.7305	0.7305	9,366	6	1	1	3	0.0000	0.7					
B124	A23	A167	C2	-	0.7223	0.7223	9,366	6	1	1	3	0.4812	0.6					
B125	A23	A167	C2	-	0.7032	0.7032	8,7071	6	2	1	2	0.4812	0.6					
A1040_ sf683	df37	C1	-	-	0.6615	0.6615	8,618	5	1	5	0	0.0000	0.9					
A1040_ A1040_	A1040_	-	-	-	0.6615	0.6615	8,618	5	1	0	5	0.0000	0.9					
A1005_ df108	sf650	C1	-	-	0.6418	0.6418	8,618	5	1	4	1	0.0000	0.8					
A1005_ A1005_	A1005_	-	-	-	0.6417	0.6417	8,618	5	1	0	5	0.0000	0.8					
A1003_ A1003_	A1003_	-	-	-	0.6413	0.6413	8	6	3	0	2	0.0000	0.9					
A1003_ df108	df37	C1	-	-	0.6413	0.6413	8	6	3	1	1	0.0000	0.9					
A251_ sf683	df108	C1	-	-	0.6364	0.6364	8,618	3	1	5	2	0.0000	0.5014	0.4896	0.9544	0.4974	0.9090	0.9960

Id: B28

QTL: 2

All+: f

Genotype

Homo(+/-)=0.943913

f:f=0.943913

Hetero(+/-)=0.055302

s:f=0.029221 d:f=0.026081

Homo(-/-)=0.000782

d:s=0.000418 s:s=0.000186 d:d=0.000177

Founders

d=0.013427

f=0.971564

s=0.015006

x=0.000000

- Prob(++/--/+-) at the QTL position + detailed genotype in terms of parental alleles
- Estimation (probability) of parental allele transmission

Step 1: Computation of genotypic probabilities – Estimation of genetic values

Visualization of genotypes

The probabilities to be homozygous / heterozygous, at the QTL positions, have been computed according to favourable / unfavourable grouping of founder alleles.

Set a threshold and select a color to display a new view of the molecular score table based on genotypes.

Customize cut-off/colours:

Prob(+/-) \geq 0,75

Prob(+/-) \geq 0,75

Prob(-/-) \geq 0,75

The rest: uncertain genotypes (?)

Reset Cancel Apply

Id	P1	P2	Cycle	Group	MS	Weight	UC	No.(+/+)	No.(-/-)	No.(+/-)	No.(?)	QTL1	QTL2	QTL3	QTL4	QTL5	QTL6	QTL7
B8								8	1	0	2	0.0000	0.8598	0.9761	0.8752	0.8959	0.9745	0.9960
B158								7	1	1	2	0.0000	0.8792	0.8925	0.9405	0.9731	0.9693	0.9960
B28								6	1	1	3	0.0000	0.9716	0.4938	0.9492	0.9137	0.9199	0.9797
B13								6	1	2	2	0.0000	0.8844	0.9559	0.9179	0.9731	0.4963	0.9895
B38								6	1	1	3	0.0000	0.9349	0.4907	0.9429	0.5822	0.9260	0.9860
B37								7	1	1	2	0.0000	0.9528	0.2334	0.9172	0.8735	0.9718	0.9860
B40								7	1	2	1	0.0000	0.9775	0.4376	0.9592	0.4974	0.9231	0.9759
B242								6	1	1	3	0.0000	0.7259	0.8959	0.8879	0.2896	0.9722	0.9860
B124								6	1	1	3	0.4812	0.6698	0.9559	0.9429	0.2787	0.9614	0.9924
B125								6	2	1	2	0.4812	0.6536	0.9559	0.9429	0.0058	0.9616	0.9924
A1040_								5	1	5	0	0.0000	0.9771	0.4842	0.9592	0.4974	0.9231	0.9759
A1040								5	1	0	5	0.0000	0.9771	0.4842	0.9592	0.4974	0.9231	0.9759
A1005_								5	1	4	1	0.0000	0.8928	0.4896	0.8751	0.4878	0.9747	0.9960
A1005								5	1	0	5	0.0000	0.8927	0.4895	0.8751	0.4878	0.9747	0.9960
A1003								6	3	0	2	0.0000	0.9757	0.9738	0.9135	0.0000	0.4986	0.9759
A1003_								6	3	1	1	0.0000	0.9757	0.9738	0.9135	0.0000	0.4987	0.9759
A251_								3	1	5	2	0.0000	0.5014	0.4896	0.9544	0.4974	0.9090	0.9960

- Visualization of genotypes via colors for each QTL (how many QTL are fixed ?)

Step2: Selection of individuals

Manual selection

Molecular score prediction Homo / Hetero Estimation of parental alleles probability Graphs

Find View Weight QTL 01 Weight 1,00 Apply

Id	P1	P2	Cycle	Group	MS	Weight	UC	No.(+/+)	No.(+/-)	No.(+/-)	No.(?)	QTL1	QTL2	QTL3	QTL4	QTL5	QTL6	QTL7
B8	A1005	A1005	C2	-	0.8366	0.8366	10	8	1	0	2	0.0000	0.8598	0.9761	0.8752	0.8959	0.9745	0.9960
B158	A251	A1005	C2	G1	0.8024	0.8024	10	7	1	1	2	0.0000	0.8792	0.8925	0.9405	0.9731	0.9693	0.9960
B28	A1006	A251	C2	G1	0.7740	0.7740	9,7071	6	1	1	3	0.0000	0.9716	0.4938	0.9492	0.9137	0.9199	0.9797
B13	A1006	A1005	C2	-	0.7609	0.7609	9,7071	6	1	2	2	0.0000	0.8844	0.9559	0.9179	0.9731	0.4963	0.9895
B38	A1040	A1005	C2	-	0.7494	0.7494	9,366	6	1	1	3	0.0000	0.9349	0.4907	0.9429	0.5822	0.9260	0.9860
B37	A1040	A1005	C2	-	0.7433	0.7433	9,366	7	1	1	2	0.0000	0.9528	0.2334	0.9172	0.8735	0.9718	0.9860
B40	A1040	A1040	C2	-	0.7404	0.7404	9,366	7	1	2	1	0.0000	0.9775	0.4376	0.9592	0.4974	0.9231	0.9759

Add to list...

Add to list

List Selection 1
List Selection 2
List1_Manual_Selection

OK Cancel

Selection of individuals Graphs Pedigree

List of selected individuals List1_Manual_Selection

	Id	P1	P2	Cycle	Group	MS
1	B13	A1006	A1005	C2	-	0.7609
2	B40	A1040	A1040	C2	-	0.7404
3	B37	A1040	A1005	C2	-	0.7433
4	B158	A251	A1005	C2	G1	0.8024
5	B8	A1005	A1005	C2	-	0.8366

Step2: Selection of individuals

- ❑ Manual selection
- ❑ Truncation selection based on:
 - ✓ Molecular score (MS)

Truncation selection (MTS)

N_{sel} Criterion List Option... Run

Selection of individuals

List of selected individuals

List1_Manual_Selection
List2_Truncation_MS_Selection
List3_Complementation_Selectio

Save
Add
Remove
Reset

View

List2_Truncation_MS_Selection

	Id	P1	P2	Cycle	Group	MS ^	Weight	UC	No.(+/+)	No.(-/-)	No.(+/-)	No.(?)	QTL1	QTL2	QTL3
1	B8	A1005	A1005	C2	-	0.8366	0.8366	10	8	1	0	2	0.0000	0.8598	0.9761
2	B158	A251	A1005	C2	G1	0.8024	0.8024	10	7	1	1	2	0.0000	0.8792	0.8925
3	B28	A1006	A251	C2	G1	0.7740	0.7740	9,7071	6	1	1	3	0.0000	0.9716	0.4938
4	B13	A1006	A1005	C2	-	0.7609	0.7609	9,7071	6	1	2	2	0.0000	0.8844	0.9559
5	B38	A1040	A1005	C2	-	0.7494	0.7494	9,366	6	1	1	3	0.0000	0.9349	0.4907
6	B37	A1040	A1005	C2	-	0.7433	0.7433	9	7	1	1	2	0.0000	0.9528	0.2334
7	B40	A1040	A1040	C2	-	0.7404	0.7404	9,366	7	1	2	1	0.0000	0.9775	0.4376
8	B242	A37	A1005	C2	-	0.7305	0.7305	9,366	6	1	1	3	0.0000	0.7259	0.8959
9	B246	A37	A1040	C2	-	0.7303	0.7303	9,366	6	1	3	1	0.0000	0.6036	0.8755
10	B293	A9	A1040	C2	-	0.7268	0.7268	9,366	6	1	3	1	0.0000	0.9549	0.9741

No. ind = 10
Mean = 0.75947

No. group = 1
Var = 0.00115008

Step2: Selection of individuals

- ❑ Manual selection
- ❑ Truncation selection based on:
 - ✓ **Molecular score (MS)**
 - ✓ **Weighted MS**: give more or less importance to the different QTL

QTLs Weights

Different weights for each QTL can be assigned:

1. Select the QTLs that have to be weighted
2. Set a weight
3. Click update
4. Click Apply

All None Weight 1,00 Update

	QTL	Weight
1	QTL1	3,00
2	QTL2	1,00
3	QTL3	1,00
4	QTL4	1,00
5	QTL5	1,00
6	QTL6	1,00
7	QTL7	1,00

Reset Cancel Apply

List3_Truncation_Weight_Selection

	Id	P1	P2	Cycle	Group	MS ^	Weight	UC	No.(+/+)	No.(+/-)	No.(+/-)	No.(?)	QTL1	QTL2	QTL3
1	B8	A1005	A1005	C2	-	0.8366	0.7079	10	8	1	0	2	0.0000	0.8598	0.9761
2	B158	A251	A1005	C2	G1	0.8024	0.6790	10	7	1	1	2	0.0000	0.8792	0.8925
3	B28	A1006	A251	C2	G1	0.7740	0.6550	9,7071	6	1	1	3	0.0000	0.9716	0.4938
4	B13	A1006	A1005	C2	-	0.7609	0.6438	9,7071	6	1	2	2	0.0000	0.8844	0.9559
5	B124	A23	A167	C2	-	0.7223	0.6852	9,366	6	1	1	3	0.4812	0.6698	0.9559
6	B125	A23	A167	C2	-	0.7032	0.6691	8,7071	6	2	1	2	0.4812	0.6536	0.9559
7	B110	A212	A1005	C2	-	0.6863	0.6481	9,2247	4	0	4	3	0.4376	0.7319	0.4876
8	B57	A167	A1040	C2	-	0.6860	0.6545	8,8229	4	0	6	1	0.4812	0.6066	0.4230
9	B123	A212	A91	C2	-	0.6717	0.6390	8,618	4	1	3	3	0.4590	0.6119	0.4879
10	B47	A166	A167	C2	-	0.6309	0.6570	8,366	4	2	2	3	0.8010	0.4966	0.8733

No. ind = 10
Mean = 0.727447

No. group = 1
Var = 0.00372879

Step2: Selection of individuals

- ❑ Manual selection
- ❑ Truncation selection based on:
 - ✓ **Molecular score (MS)**
 - ✓ **Weighted MS:** give more or less importance to the different QTL
 - ✓ **Utility criterion:** select candidates based on the possibility of obtaining superior genotype in their progeny (favor heterozygous)
- ❑ QTL complementation selection

(Hospital et al., 2000)

Step2: Selection of individuals – QTL complementation selection (QCS) *(Hospital et al., 2000)*

List4_Complementation_Selection																			
	Id	P1	P2	Cycle	Group	MS ^	Weight	UC	No.(+/+)	No.(-/-)	No.(+/-)	No.(?)	QTL1	QTL2	QTL3	QTL4	QTL5	QTL6	QTL7
1	B8	A1005	A1005	C2	-	0.8366	0.7079	10	8	1	0	2	0.0000	0.8598	0.9761	0.8752	0.8959	0.9745	0.9960
2	B158	A251	A1005	C2	G1	0.8024	0.6790	10	7	1	1	2	0.0000	0.8792	0.8925	0.9405	0.9731	0.9693	0.9960
3	B28	A1006	A251	C2	G1	0.7740	0.6550	9,7071	6	1	1	3	0.0000	0.9716	0.4938	0.9492	0.9137	0.9199	0.9797
4	B13	A1006	A1005	C2	-	0.7609	0.6438	9,7071	6	1	2	2	0.0000	0.8844	0.9559	0.9179	0.9731	0.4963	0.9895
5	B38	A1040	A1005	C2	-	0.7494	0.6341	9,366	6	1	1	3	0.0000	0.9349	0.4907	0.9429	0.5822	0.9260	0.9860
6	B37	A1040	A1005	C2	-	0.7433	0.6290	9	7	1	1	2	0.0000	0.9528	0.2334	0.9172	0.8735	0.9718	0.9860
7	B40	A1040	A1040	C2	-	0.7404	0.6265	9,366	7	1	2	1	0.0000	0.9775	0.4376	0.9592	0.4974	0.9231	0.9759
8	B242	A37	A1005	C2	-	0.7305	0.6181	9,366	6	1	1	3	0.0000	0.7259	0.8959	0.8879	0.2896	0.9722	0.9860
+																			
9	B124	A23	A167	C2	-	0.7223	0.6852	9,366	6	1	1	3	0.4812	0.6698	0.9559	0.9429	0.2787	0.9614	0.9924
10	B125	A23	A167	C2	-	0.7032	0.6691	8,7071	6	2	1	2	0.4812	0.6536	0.9559	0.9429	0.0058	0.9616	0.9924

Take into account complementarities between candidates regarding the favorable alleles they carry

Prevent the loss of rare favourable alleles and fixation of unfavourable alleles at QTL with small effects.

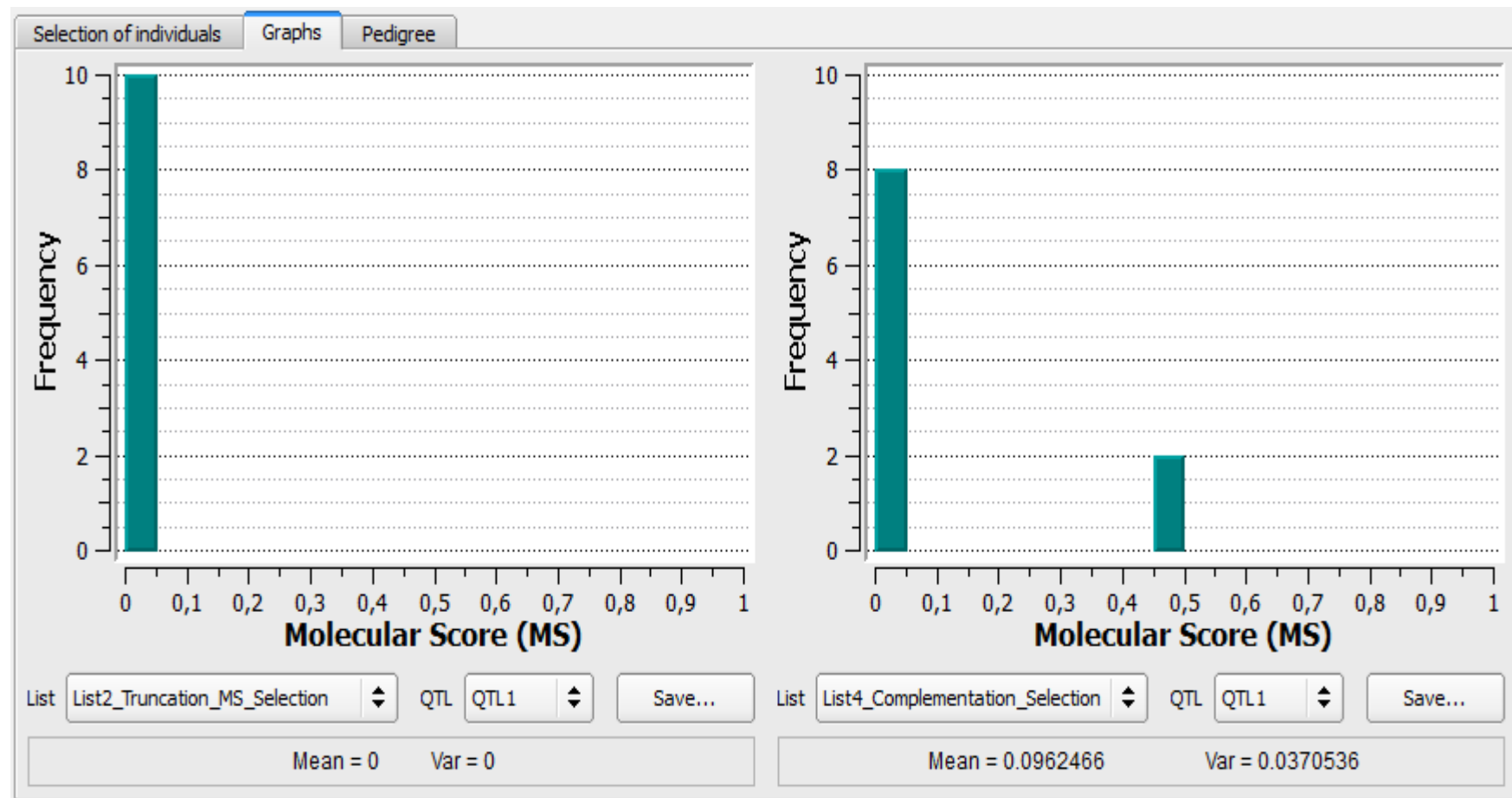
A QCS strategy is described by four parameters: (Hospital et al., 2000)			
θ_{MS}	\geq	<input type="text" value="6,47"/>	The threshold for Molecular Score (MS), above which a favourable QTL allele is declared 'present'.
n_T	$=$	<input type="text" value="2"/>	Each QTL is requested to be 'present' in at least n_T selected individuals.
MS_{min}	\geq	<input type="text" value="0,70"/>	The minimum threshold value (Molecular Score) for the addition of an individual.
N_{max}	$=$	<input type="text" value="10"/>	Maximum number of individuals selected at the end of the complementation process.

Step2: Selection of individuals – Comparison between lists of selected individuals



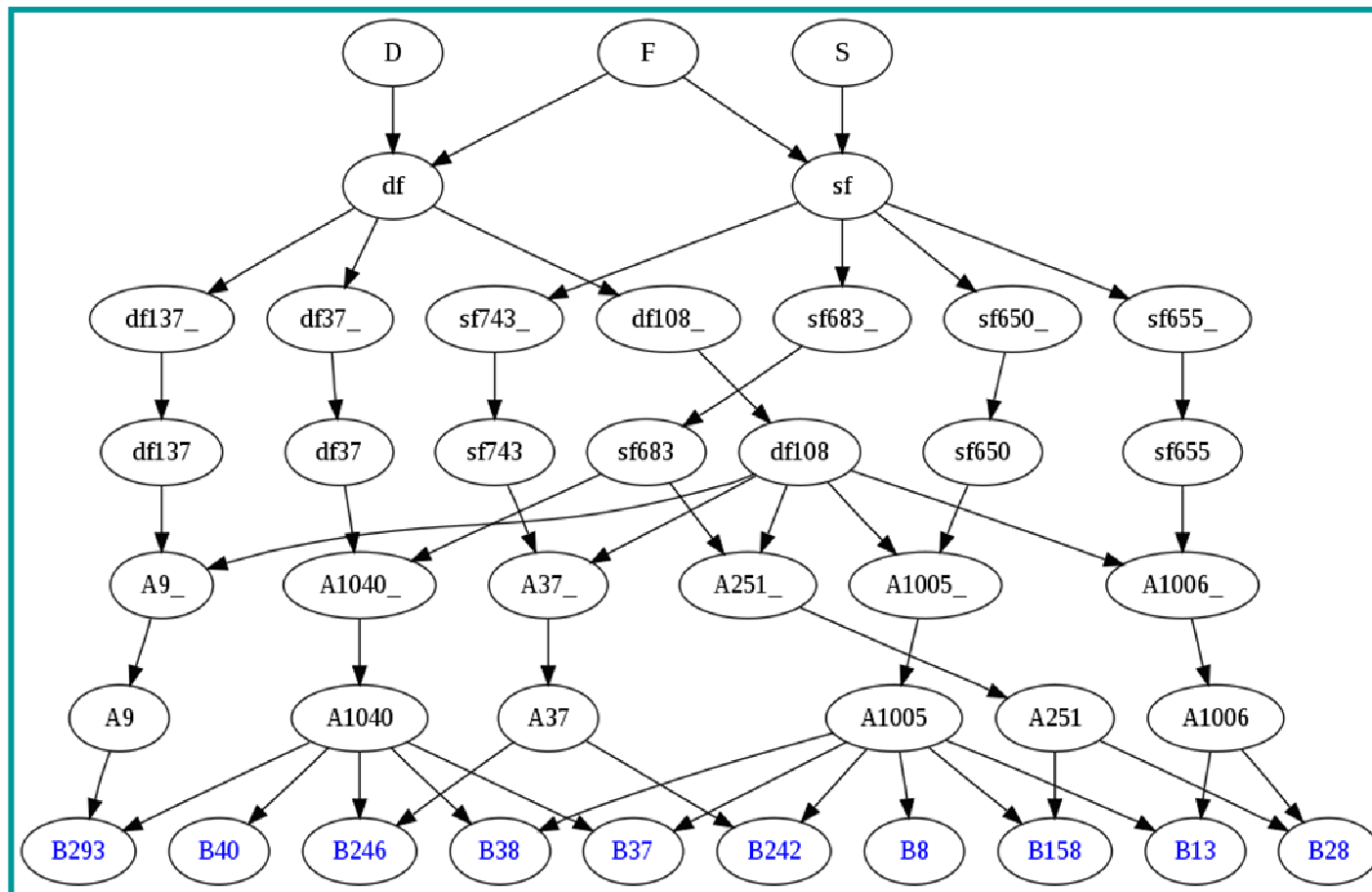
List2_Truncation_MS_Selection									List4_Complementation_Selection								
	Id	P1	P2	Cycle	Group	MS ^	Weight	UC		Id	P1	P2	Cycle	Group	MS ^	Weight	UC
1	B8	A1005	A1005	C2	-	0.8366	0.7079	10	1	B8	A1005	A1005	C2	-	0.8366	0.7079	10
2	B158	A251	A1005	C2	G1	0.8024	0.6790	10	2	B158	A251	A1005	C2	G1	0.8024	0.6790	10
3	B28	A1006	A251	C2	G1	0.7740	0.6550	9,7071	3	B28	A1006	A251	C2	G1	0.7740	0.6550	9,7071
4	B13	A1006	A1005	C2	-	0.7609	0.6438	9,7071	4	B13	A1006	A1005	C2	-	0.7609	0.6438	9,7071
5	B38	A1040	A1005	C2	-	0.7494	0.6341	9,366	5	B38	A1040	A1005	C2	-	0.7494	0.6341	9,366
6	B37	A1040	A1005	C2	-	0.7433	0.6290	9	6	B37	A1040	A1005	C2	-	0.7433	0.6290	9
7	B40	A1040	A1040	C2	-	0.7404	0.6265	9,366	7	B40	A1040	A1040	C2	-	0.7404	0.6265	9,366
8	B242	A37	A1005	C2	-	0.7305	0.6181	9,366	8	B242	A37	A1005	C2	-	0.7305	0.6181	9,366
9	B246	A37	A1040	C2	-	0.7303	0.6180	9,366	9	B124	A23	A167	C2	-	0.7223	0.6852	9,366
10	B293	A9	A1040	C2	-	0.7268	0.6150	9,366	10	B125	A23	A167	C2	-	0.7032	0.6691	8,7071
<div> <div></div> <div> </div> <div></div> </div>									<div> <div></div> <div> </div> <div></div> </div>								
No. ind = 10					No. group = 1				No. ind = 10					No. group = 1			
Mean = 0.75947					Var = 0.00115008				Mean = 0.756312					Var = 0.00140282			

Step2: Selection of individuals – Comparison between lists of selected individuals



- ❑ Distribution of individuals for each list regarding the molecular score (MS)
- ❑ Favorable allele (QTL1) lost without the QTL Complementation Selection (QCS)

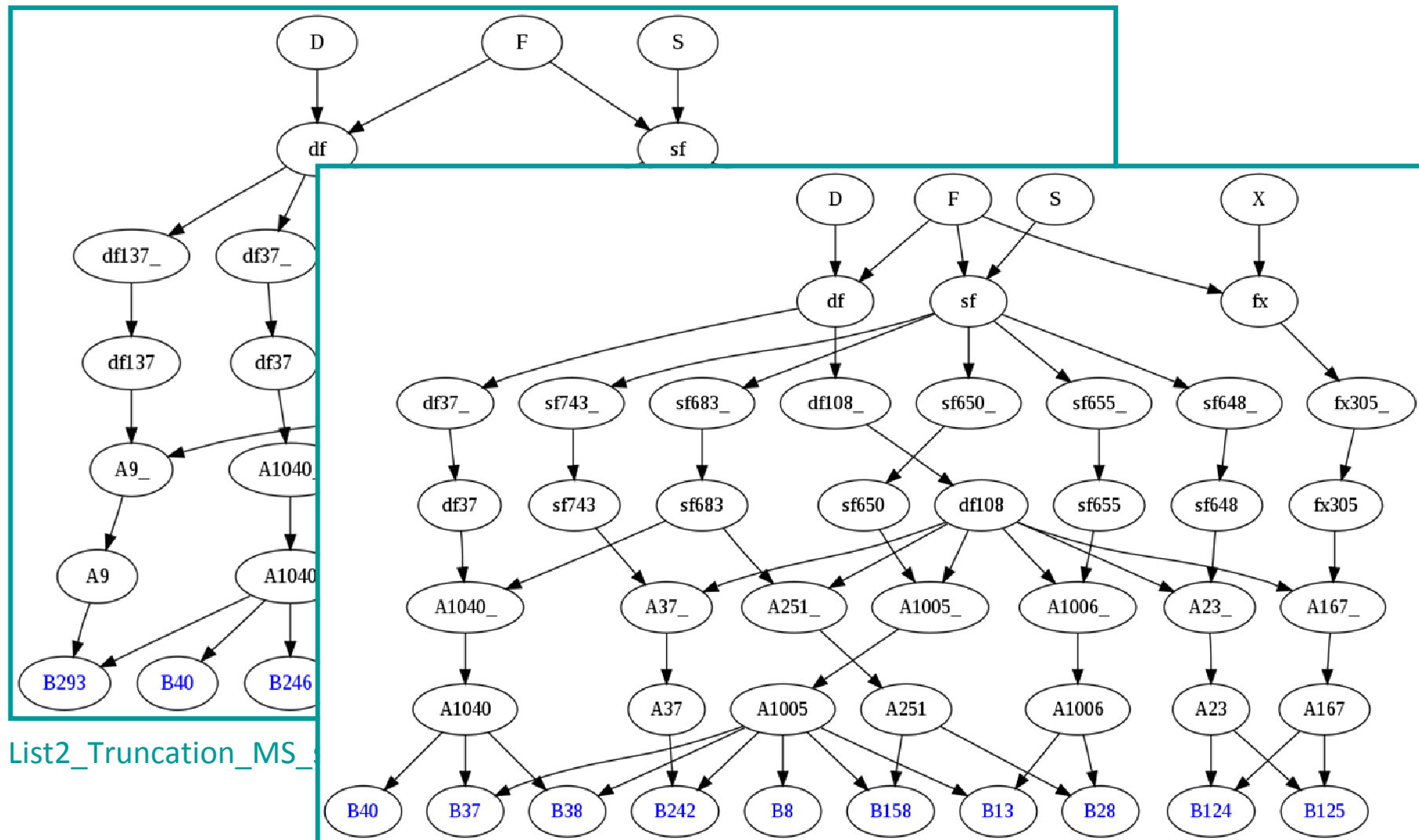
Step2: Selection of individuals – Comparison between lists of selected individuals (pedigree)



List2_Truncation_MS_selection

- To follow the contribution of selected individuals over generations and prevent possible bottlenecks

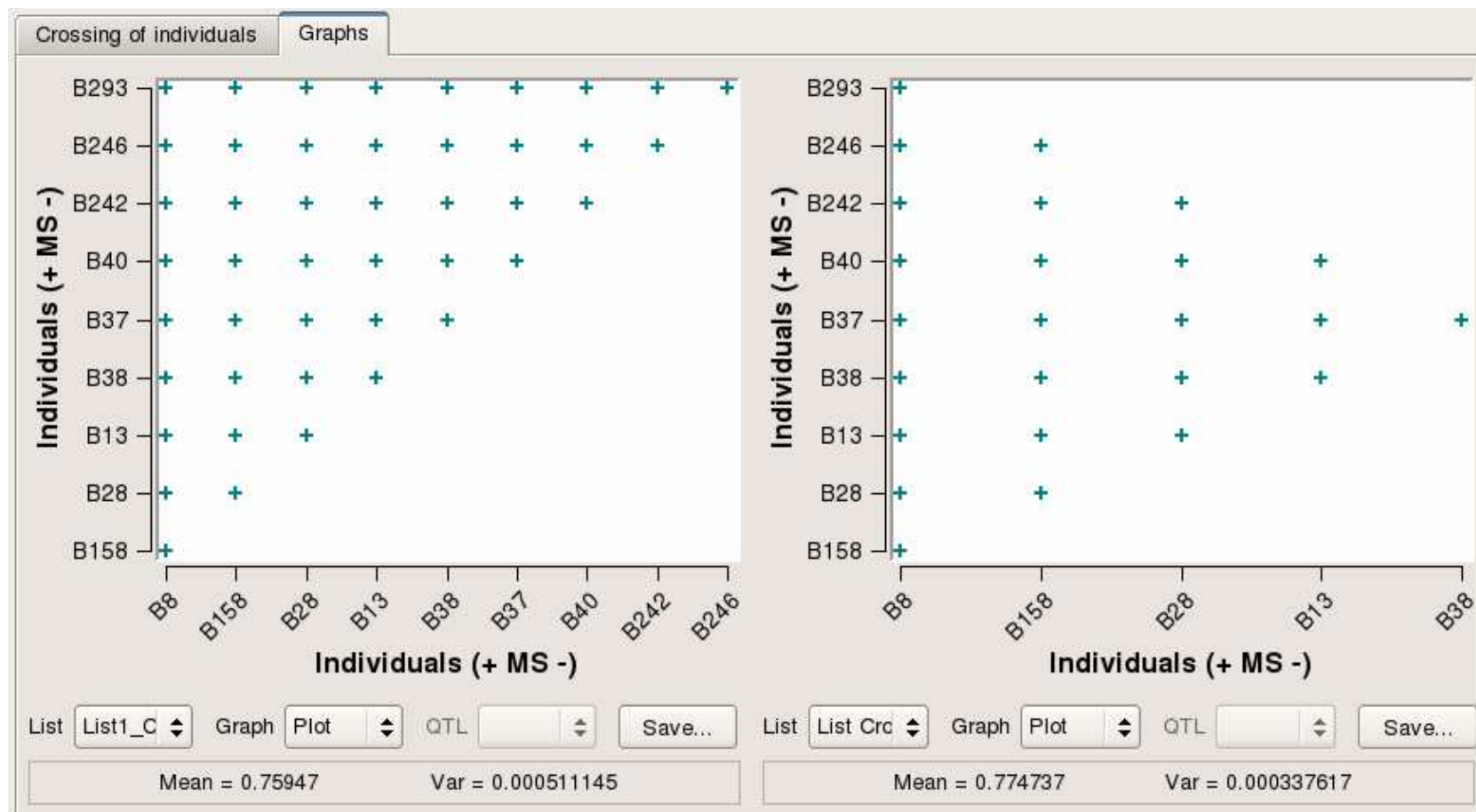
Step2: Selection of individuals – Comparison between lists of selected individuals (pedigree)



List4_Complementation_selection: 4 parental alleles present in the next generation

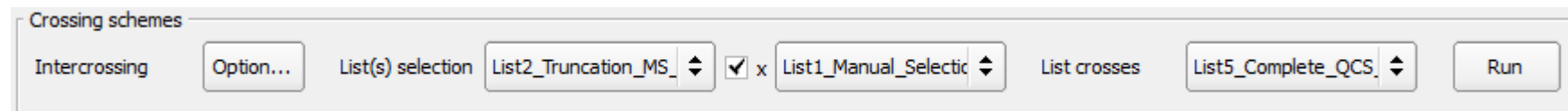
Step3: Identification of crosses to be made among selected individuals

- Half-diallel between selected candidates
- « Better-half » strategy which consists of avoiding crosses between the «worst» selected individuals (*Bernardo et al., 2006*) → optimization of selection intensity



Step3: Identification of crosses to be made among selected individuals

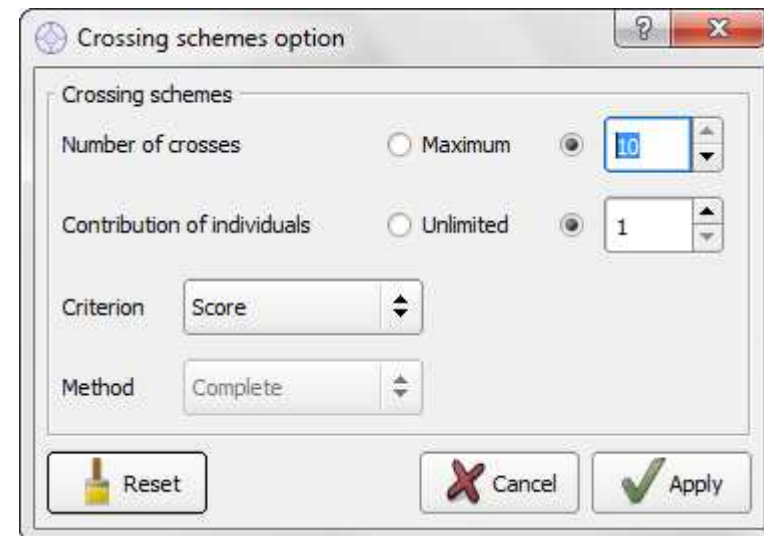
- Half-diallel between selected candidates
- « Better-half » strategy which consists of avoiding crosses between the worst selected individuals (*Bernardo et al., 2006*) → optimization of selection intensity
- 2 lists: factorial design



Crossing schemes

Intercrossing Option... List(s) selection List2_Truncation_MS_ ☒ x List1_Manual_Selectic List crosses List5_Complete_QCS_ Run

- Constraints:
 - ✓ Maximum number of crosses
 - ✓ Contribution of individuals



Crossing schemes option

Crossing schemes

Number of crosses ☐ Maximum ☒ 10

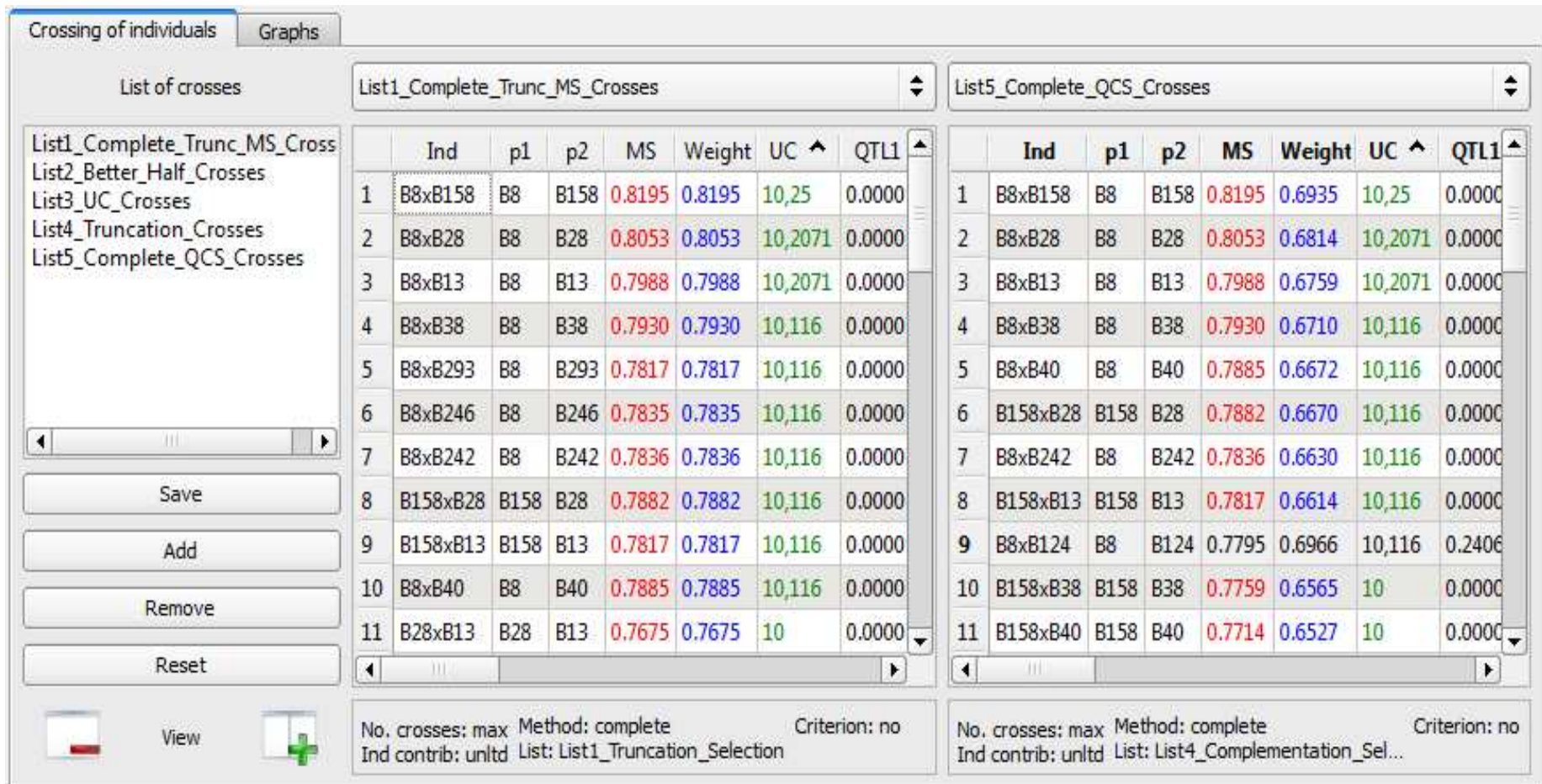
Contribution of individuals ☐ Unlimited ☒ 1

Criterion Score

Method Complete

Reset Cancel Apply

Step3: Identification of crosses to be made among selected individuals



The screenshot shows a software interface for managing crosses. It has two main panels: 'List1_Complete_Trunc_MS_Crosses' and 'List5_Complete_QCS_Crosses'. Both panels display a table of crosses with columns: Ind, p1, p2, MS, Weight, UC, and QTL1. The MS column is highlighted in red, indicating the expected MS of the progeny.

List1_Complete_Trunc_MS_Crosses

	Ind	p1	p2	MS	Weight	UC	QTL1
1	B8xB158	B8	B158	0.8195	0.8195	10,25	0.0000
2	B8xB28	B8	B28	0.8053	0.8053	10,2071	0.0000
3	B8xB13	B8	B13	0.7988	0.7988	10,2071	0.0000
4	B8xB38	B8	B38	0.7930	0.7930	10,116	0.0000
5	B8xB293	B8	B293	0.7817	0.7817	10,116	0.0000
6	B8xB246	B8	B246	0.7835	0.7835	10,116	0.0000
7	B8xB242	B8	B242	0.7836	0.7836	10,116	0.0000
8	B158xB28	B158	B28	0.7882	0.7882	10,116	0.0000
9	B158xB13	B158	B13	0.7817	0.7817	10,116	0.0000
10	B8xB40	B8	B40	0.7885	0.7885	10,116	0.0000
11	B28xB13	B28	B13	0.7675	0.7675	10	0.0000

List5_Complete_QCS_Crosses

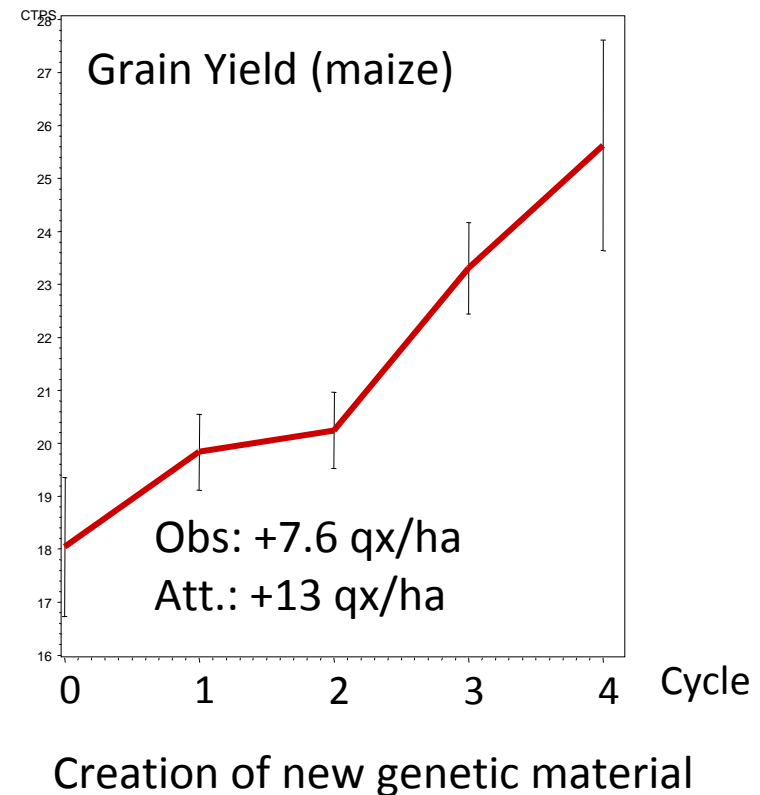
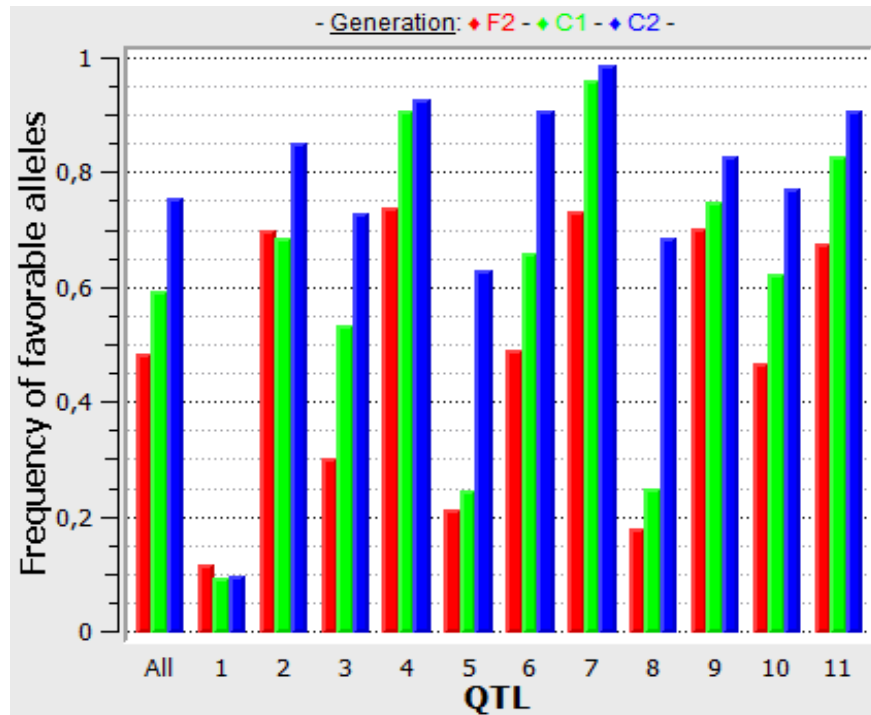
	Ind	p1	p2	MS	Weight	UC	QTL1
1	B8xB158	B8	B158	0.8195	0.6935	10,25	0.0000
2	B8xB28	B8	B28	0.8053	0.6814	10,2071	0.0000
3	B8xB13	B8	B13	0.7988	0.6759	10,2071	0.0000
4	B8xB38	B8	B38	0.7930	0.6710	10,116	0.0000
5	B8xB40	B8	B40	0.7885	0.6672	10,116	0.0000
6	B158xB28	B158	B28	0.7882	0.6670	10,116	0.0000
7	B8xB242	B8	B242	0.7836	0.6630	10,116	0.0000
8	B158xB13	B158	B13	0.7817	0.6614	10,116	0.0000
9	B8xB124	B8	B124	0.7795	0.6966	10,116	0.2406
10	B158xB38	B158	B38	0.7759	0.6565	10	0.0000
11	B158xB40	B158	B40	0.7714	0.6527	10	0.0000

At the bottom of each panel, there is a summary: 'No. crosses: max Method: complete Criterion: no Ind contrib: unltl List: List1_Truncation_Selection' and 'No. crosses: max Method: complete Criterion: no Ind contrib: unltl List: List4_Complementation_Sel...'.

- In each case: computation of the expected MS of the progeny

Conclusion

- OptiMAS tested on a multiparental connected design (*Blanc et al., 2008*)



- Development of 2 version of the tool:
 - ✓ Command line: C-ANSI language
 - ✓ Graphical Interface (GUI): C++ (Qt & Qwt libraries)
- OptiMAS release: [available online soon](#) for Windows & Linux then Mac OSX

Integrated Breeding Platform Website

The screenshot displays the Integrated Breeding Platform (IBP) website in a Firefox browser window. The address bar shows the URL <https://www.integratedbreeding.net/ib-tools/breeding-decision/optimas>. The website header features the IBP logo, navigation links (Home, Crop information, Breeding activities, Capacity building, My Community, News & Events, Help), and a search bar. The main content area is titled "OptiMAS" and includes a sidebar with navigation options: Project planning and queries, Germplasm management, Field trial management, Genotypic data management, Data analysis, and Breeding decision (highlighted). The main content area provides details about the OptiMAS tool, including its category (Marker assisted recurrent selection tool), platform (Windows), version (Alpha), website (OptiMAS information), developers (Alain Charcosset, Laurence Moreau, Fabio Valente (INRA, France)), and technical support (Delphine Fleury). A table of data is visible in the background. The page concludes with a "beta-version to be released in 2012" note and a "This tool is in development" banner.

Firefox | OptiMAS | Integrated Breeding Platform | +

integratedbreeding.net | <https://www.integratedbreeding.net/ib-tools/breeding-decision/optimas> | Google

Integrated Breeding Platform
Accessible cutting-edge crop breeding tools and services

Register for an account | Log in | Search IBP

Home | Crop information | Breeding activities | Capacity building | My Community | News & Events | Help

IB tools | Breeding services & schemes | Traits & protocols | Supplementary toolbox

OptiMAS

Project planning and queries >
Germplasm management >
Field trial management >
Genotypic data management >
Data analysis >
Breeding decision >

Category Marker assisted recurrent selection tool

Platform

Version Alpha

Website [OptiMAS information](#)

Developers Alain Charcosset
Laurence Moreau
Fabio Valente (INRA, France)

Technical Support Delphine Fleury

OptiMAS is a software for marker-assisted recurrent selection (MARS) and gene pyramiding. It aids decision making in a MARS selection program in order to assemble favorable QTL alleles.

Features:

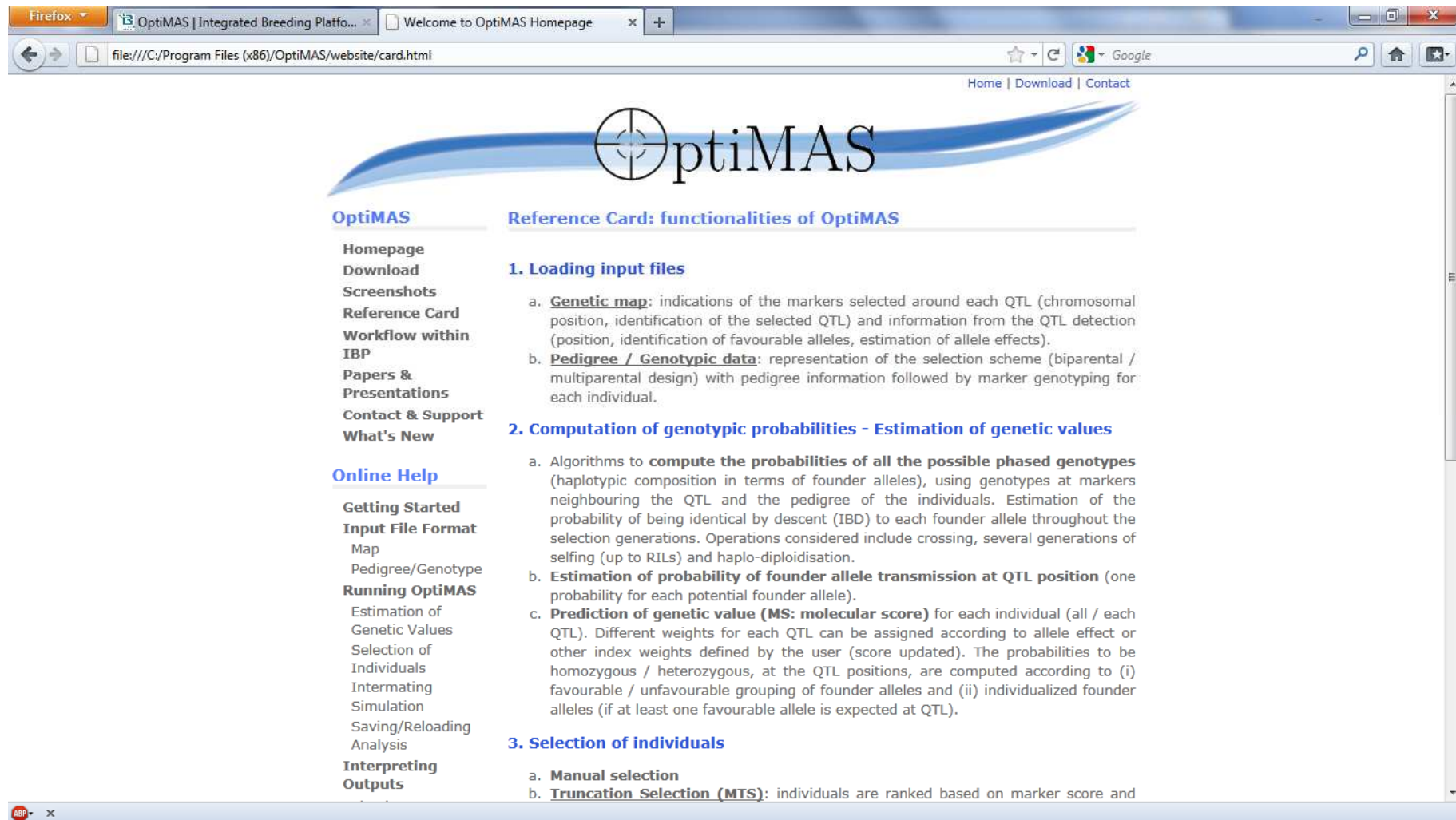
- Bi-parental or multi-parental selection schemes
- Trace parental QTL throughout selection generations by estimating their genetic values
- Select the best individuals
- Identify crosses to be made
- Draw pedigree of selected plants

beta-version to be released in 2012

[Login](#) or [register](#) to post comments | [Breeding decision](#) |

This tool is in development

Tutorial/Documentation



Firefox | OptiMAS | Integrated Breeding Platform | Welcome to OptiMAS Homepage

file:///C:/Program Files (x86)/OptiMAS/website/card.html

Home | Download | Contact

OptiMAS

- Homepage
- Download
- Screenshots
- Reference Card
- Workflow within IBP
- Papers & Presentations
- Contact & Support
- What's New

Online Help

- Getting Started
- Input File Format
 - Map
 - Pedigree/Genotype
- Running OptiMAS
 - Estimation of Genetic Values
 - Selection of Individuals
 - Intermating
 - Simulation
 - Saving/Reloading
 - Analysis
- Interpreting Outputs

Reference Card: functionalities of OptiMAS

1. Loading input files

- Genetic map**: indications of the markers selected around each QTL (chromosomal position, identification of the selected QTL) and information from the QTL detection (position, identification of favourable alleles, estimation of allele effects).
- Pedigree / Genotypic data**: representation of the selection scheme (biparental / multiparental design) with pedigree information followed by marker genotyping for each individual.

2. Computation of genotypic probabilities - Estimation of genetic values

- Algorithms to **compute the probabilities of all the possible phased genotypes** (haplotypic composition in terms of founder alleles), using genotypes at markers neighbouring the QTL and the pedigree of the individuals. Estimation of the probability of being identical by descent (IBD) to each founder allele throughout the selection generations. Operations considered include crossing, several generations of selfing (up to RILs) and haplo-diploidisation.
- Estimation of probability of founder allele transmission at QTL position** (one probability for each potential founder allele).
- Prediction of genetic value (MS: molecular score)** for each individual (all / each QTL). Different weights for each QTL can be assigned according to allele effect or other index weights defined by the user (score updated). The probabilities to be homozygous / heterozygous, at the QTL positions, are computed according to (i) favourable / unfavourable grouping of founder alleles and (ii) individualized founder alleles (if at least one favourable allele is expected at QTL).

3. Selection of individuals

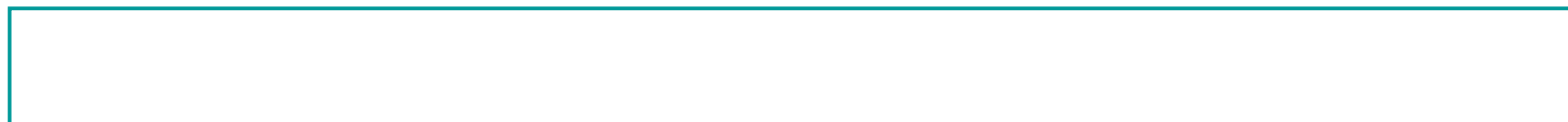
- Manual selection**
- Truncation Selection (MTS)**: individuals are ranked based on marker score and

Acknowledgments

- ❑ Bill & Melinda Gates Foundation
- ❑ Guylaine Blanc & J.B Veyrieras for developing the first prototype of the program
- ❑ Nicolas Bardol & Delphine Fleury (GCP) for beta testing

Thank you
Questions welcome





Identification of crosses to be made among selected individuals

- ❑ **Complete** : all possible crosses between the selected individuals.
- ❑ **Better Half** : avoid crosses between the last selected genotypes.
 - Optimization of selection intensity
- ❑ **Predefined number of crosses** to be made based on :
 - ✓ **Molecular score**.
 - ✓ **Utility criterion** [*in progress*] : expected mean and variance at the next generation (favor heterozygous...).

- Favor couples that will produce the best individuals as possible in the next generation.

$$U_{\text{couple}} = \text{Mean of the parental values of the couple} + i \sigma_{\text{couple}}$$

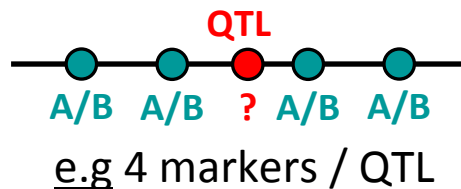
Variance = only heterozygotes genotypes at QTL contribute to variance in the next generation.

Selection of individuals

- **Aim:** to obtain an ideotype combining all the favorable alleles at QTL positions.
- **Prediction of genetic value** based on the information from neighboring markers.

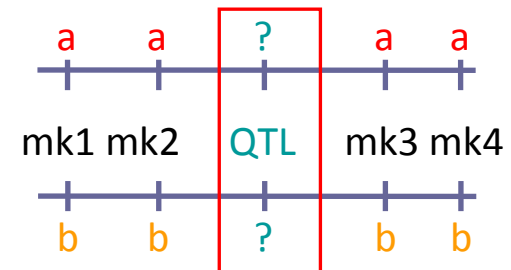
⇒ Which parental alleles are transmitted at the QTL position ?

- QTL position rarely located at a marker → QTL alleles are unknown and must be **inferred from flanking markers**.



Selection of a subset of **informative markers** to follow the favourable parental alleles (based on haplotypes).

Identification of all the possible **phased genotypes** for each QTL and computation of their **probabilities**



Following « phased genotypes »

□ Information available

- ✓ Pedigree
- ✓ Distance between loci (Haldane → recombination rate)
- ✓ Molecular markers (observed genotypes)

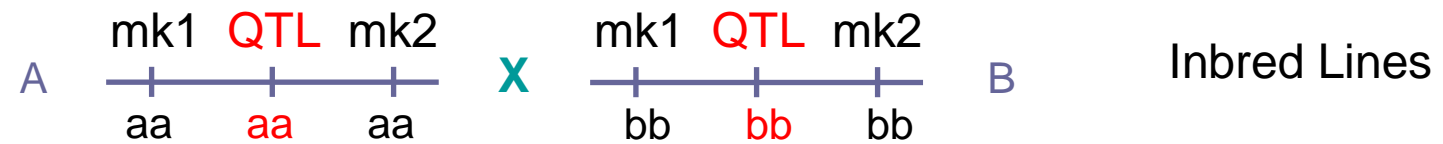
⇒ Identification of all the possible phased genotypes for each QTL and computation of their probabilities.



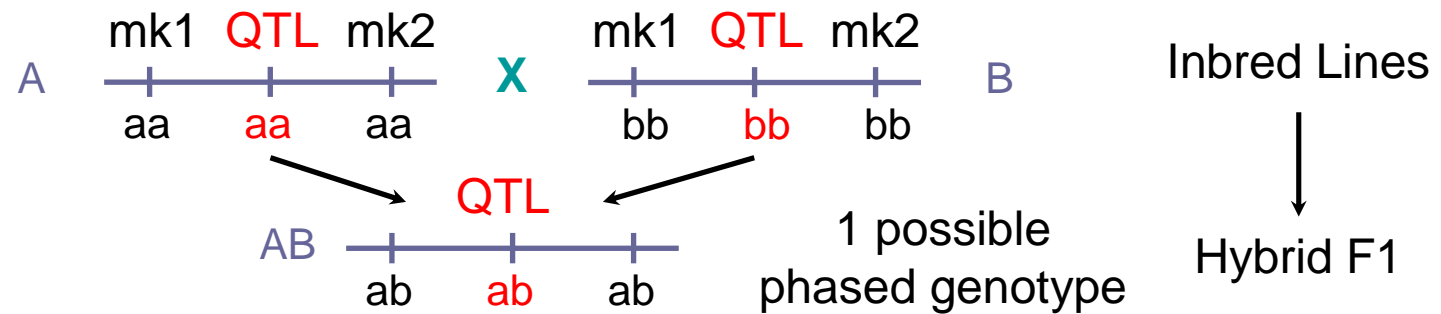
- ✓ Phase (haplotypes) unknown → all possibilities considered

Phased genotype: contains no ambiguity on alleles transmitted from parents and the phase.

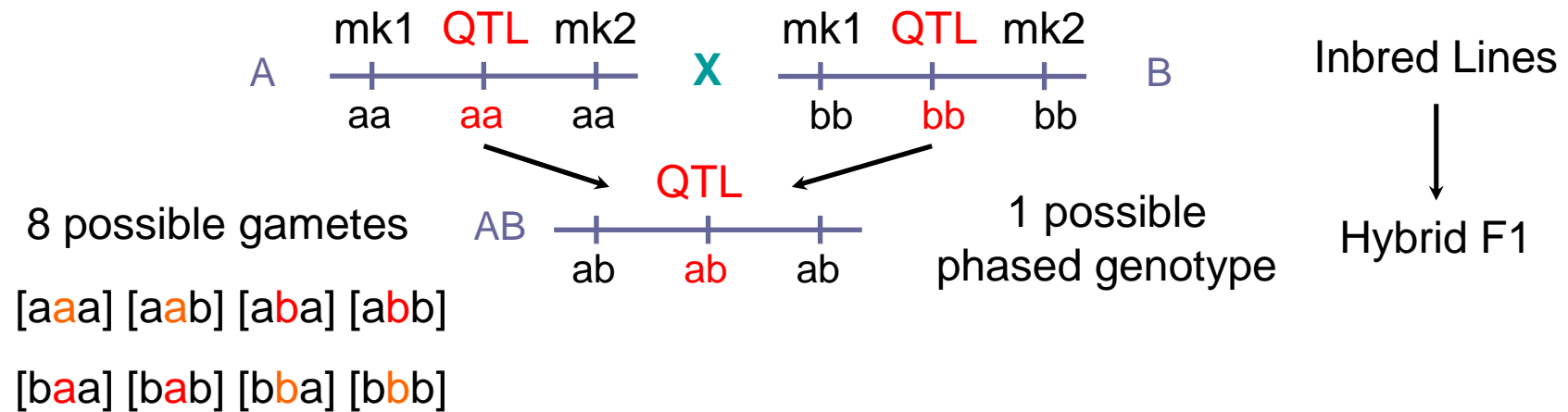
New algorithms: adaptability to different marker selection schemes and mating systems



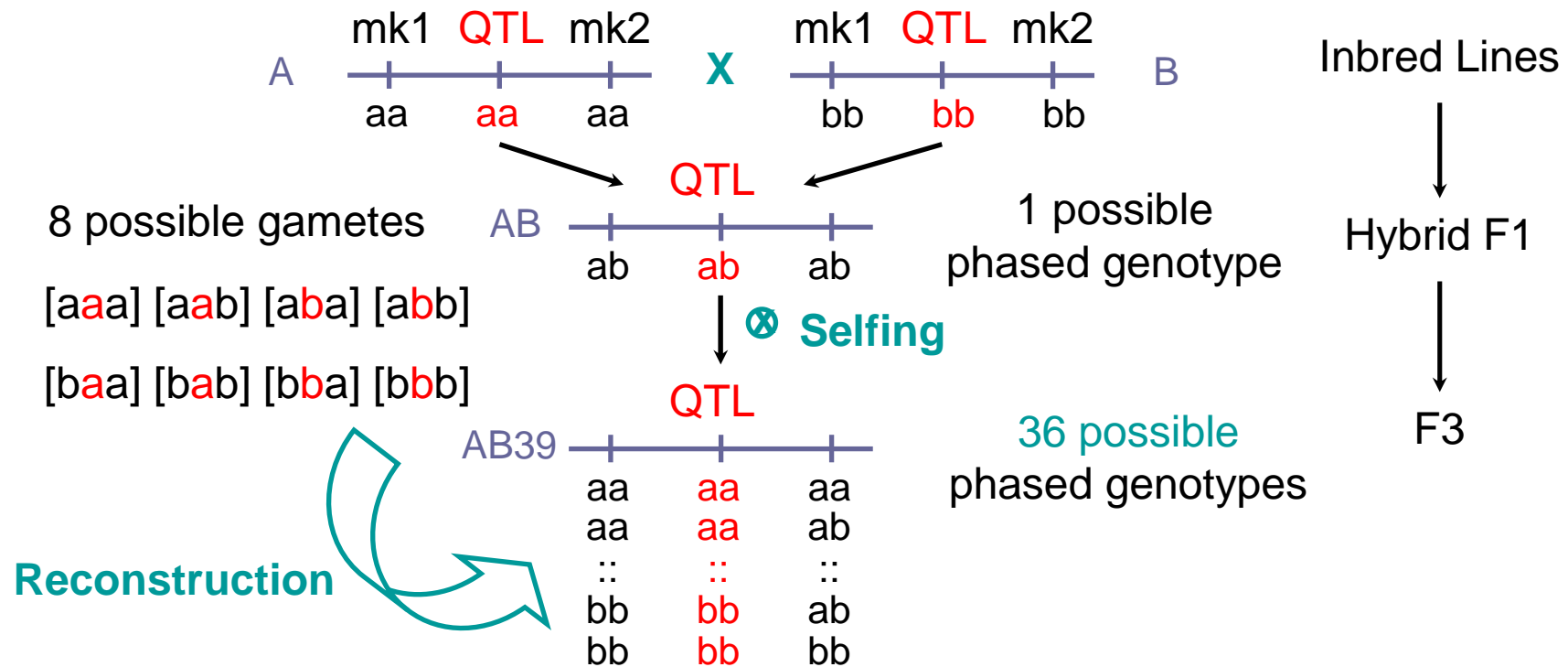
New algorithms: adaptability to different marker selection schemes and mating systems



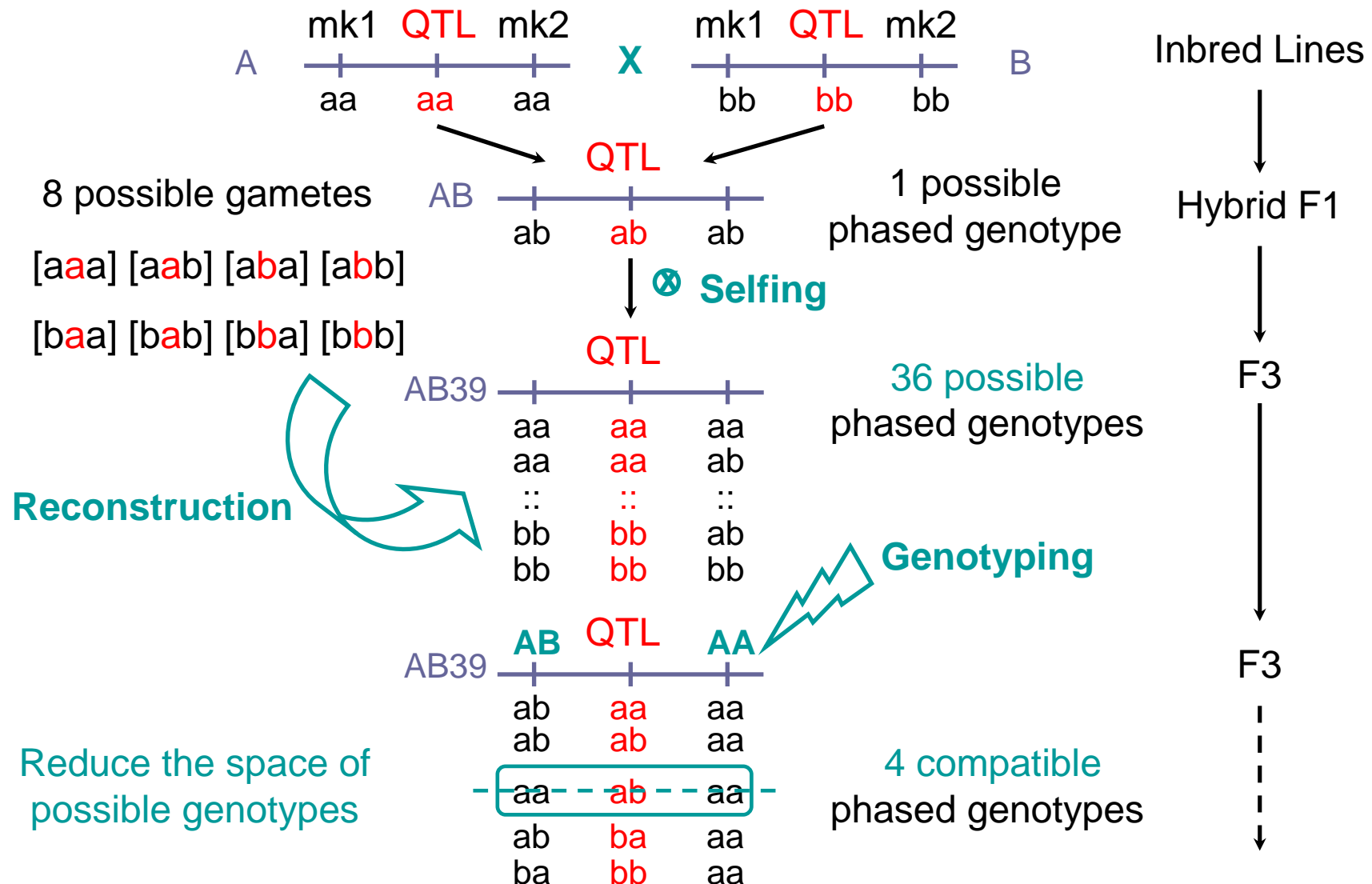
New algorithms: adaptability to different marker selection schemes and mating systems



New algorithms: adaptability to different marker selection schemes and mating systems



New algorithms: adaptability to different marker selection schemes and mating systems



Estimation of genetic values

□ `_scores`: genetic value (molecular score) for each individual

✓ Each QTL: expected dose of favorable allele (0 or 1) for all possible phased genotypes, weighted by the probability of each phased genotype.

$$M_{each} = \frac{\sum \theta_q \cdot p_G}{2}$$

$$M_{all} = \frac{\sum M_{each}}{nb_qtl}$$

2 classes:

- Favorable

- Not favorable

θ_q : genotype (homo ^{-/+}, hetero) at QTL position (0, 1, 2)

p_G : probability of the current genotype

→ Option: effect associated with parental allele(s) at QTL

→ QTL are considered as being independent (unlinked)

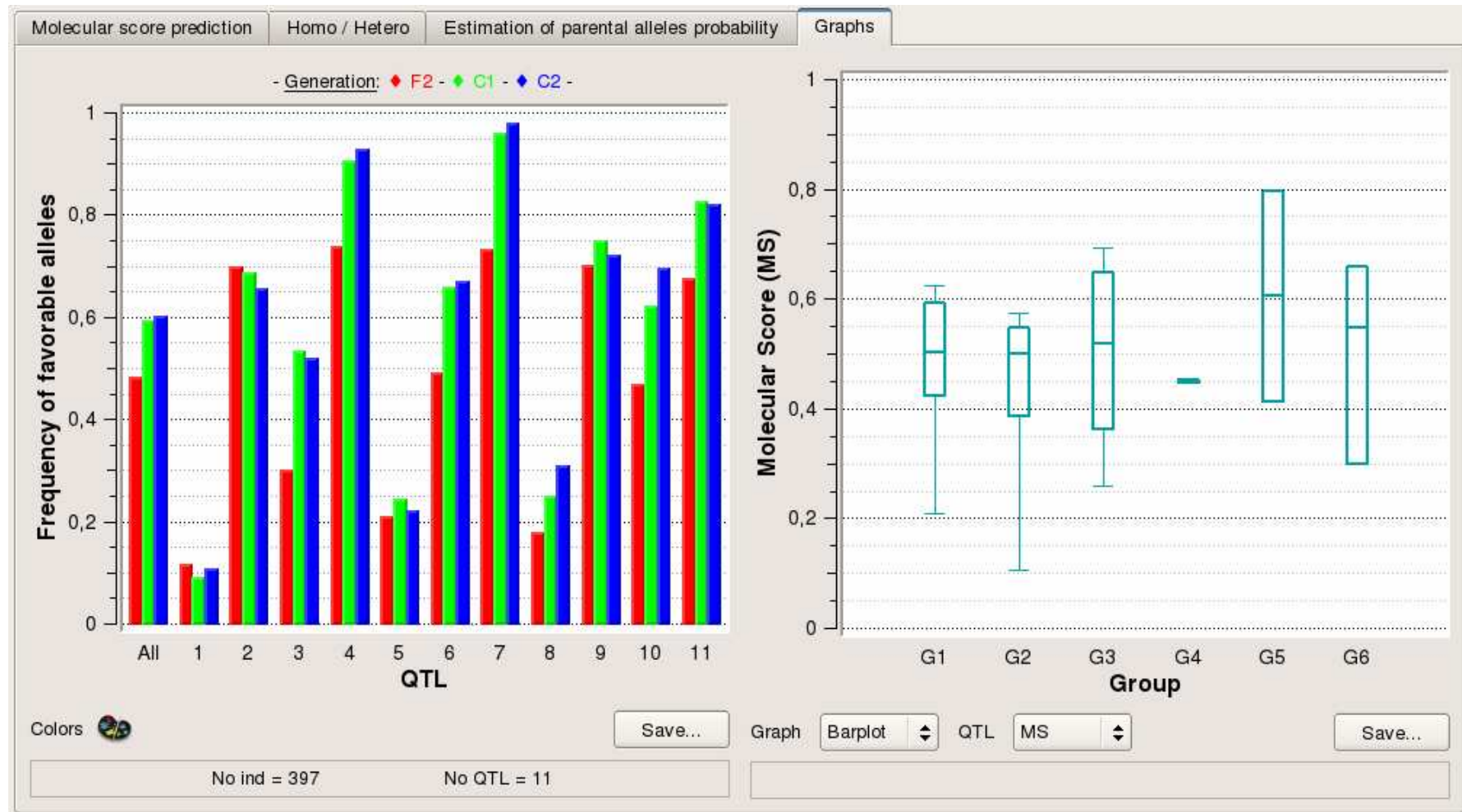
Geno	θ
------	----------

Homo ⁻	0
-------------------	---

Hetero	1
--------	---

Homo ⁺	2
-------------------	---

Step 1: Computation of genotypic probabilities – Estimation of genetic values (graphs)



- Mean frequency of favorable alleles at all/each QTL through generation of selection
- Intra-group variation (ears)