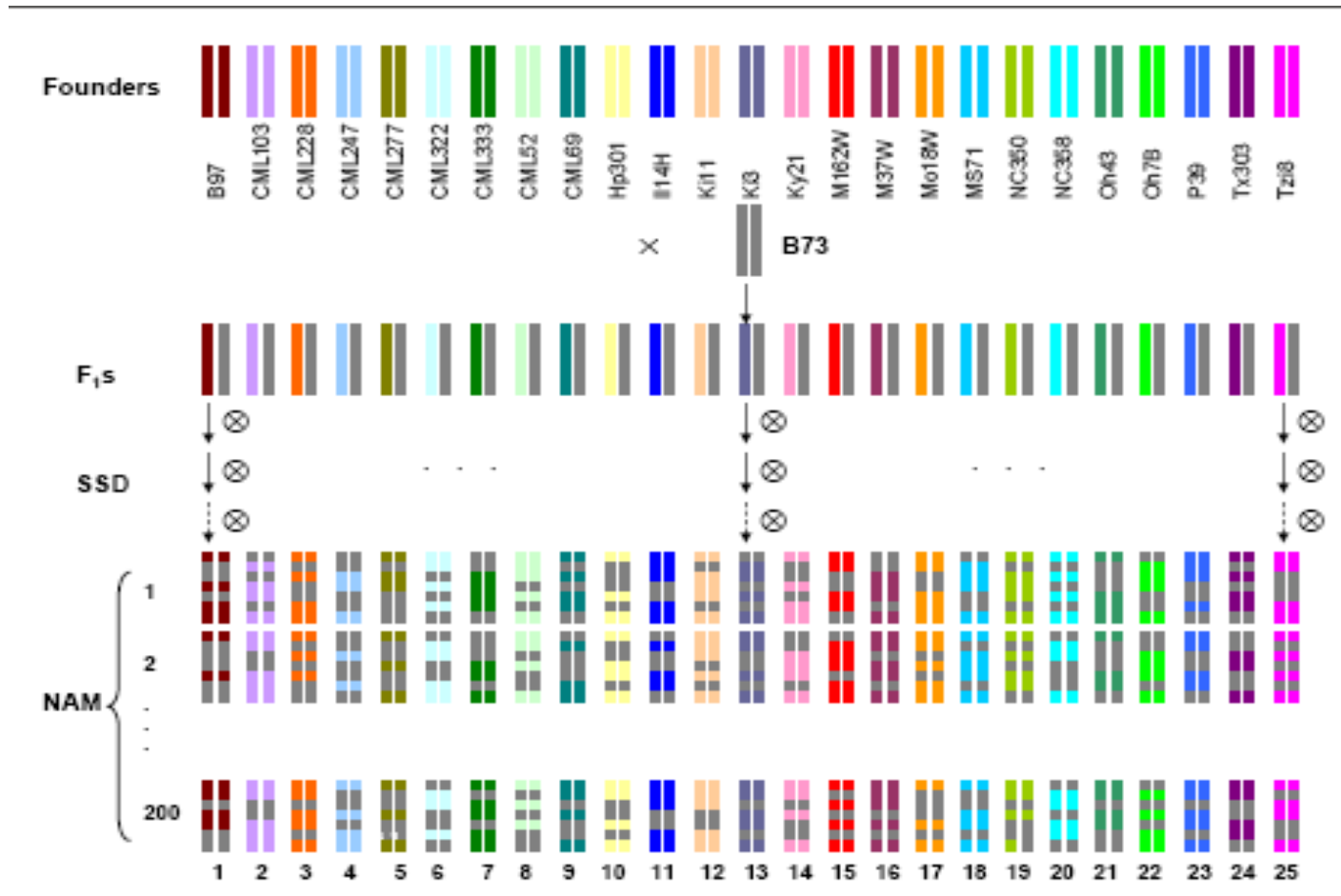


The 9th Workshop on QTL Mapping and Breeding Simulation
The University of Sydney, Cobbitty NSW, 7-9 March 2012

Joint Inclusive Composite Interval Mapping (JICIM) for the Nested Association (NAM) Mapping design

One NAM population in maize



Science (2009) 325: 714-718; Science (2009) 325: 737-740

The general linear model in JICIM

- General linear model (GLM)

$$\mathbf{Y} = b_0 + \boldsymbol{\alpha}\mathbf{u} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where \mathbf{Y} is the vector of phenotypic values; b_0 is the intercept; $\mathbf{u}=(u_1, u_2, \dots, u_F)$ is the effect vector indicating the cross effect of each founder with the common parent; $\boldsymbol{\alpha}$ is the $N \times F$ incidence matrix relating each u_f ($f=1, 2, \dots, F$) to \mathbf{Y} ; $\boldsymbol{\beta}$ is the $[(F+1) \cdot m] \times 1$ effect vector of the $N \times [(F+1) \cdot m]$ incidence matrix \mathbf{X} ; m is the number of markers; and $\boldsymbol{\varepsilon}$ is the vector of residual variance.

- Phenotype was adjusted by:
$$\Delta y_i = y_i - \alpha_{if} u_f - \sum_{j \neq k, k+1} \hat{b}_j x_{ij}$$

The mapping algorithm of JICIM

➤ Hypotheses

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_F = \mu_0$$

H_A : at least one of $\mu_1, \mu_2, \dots, \text{ and } \mu_F$ not equal to μ_0 .

➤ Parameter estimation

$$L_A = \sum_{f=1}^F \sum_{j=1}^4 \sum_{\substack{i=1 \\ i \in S_j}}^{n_f} \log \left[p_{ff} \Phi(\Delta y_i; \mu_f, \sigma_f^2) + (1 - p_{ff}) \Phi(\Delta y_i; \mu_0, \sigma_f^2) \right]$$

where S_j denotes the j^{th} marker type group ($j=1, 2, 3, 4$); p_{ff} is the proportion of QTL genotypes $Q_f Q_f$ in the j^{th} marker type group of the f^{th} family; μ_f and μ_0 are the means for QTL genotypes $Q_f Q_f$ and $Q_0 Q_0$, respectively, and $\Phi(\bullet; \mu, \sigma^2)$ represents the probability density of the normal distribution.

Genetic variance of each QTL in the NAM design

➤ Assuming there are F families and there is no segregation distortion, the genetic variance of the $F+1$ QTL genotypes in the NAM population can be calculated as,

$$V_{QTL} = 2 \sum_{f=1}^F \frac{n_f}{N} a_f^2 - \left(\sum_{f=1}^F \frac{n_f}{N} a_f \right)^2$$

where n_f is the population size in the f th family, a_f is the additive effect for QTL in the f th family, and N is the total population size.

Phenotypic variance explained (PVE) by Each QTL in the NAM design

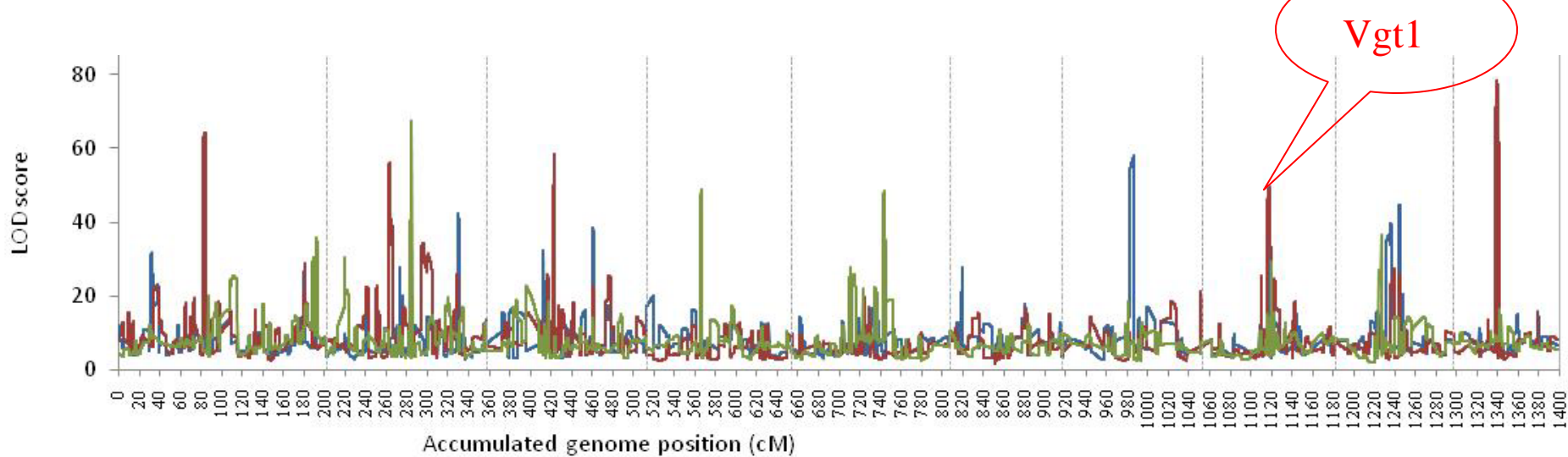
➤ PVE is

$$PVE = \frac{2 \sum_{f=1}^F \frac{n_f}{N} a_f^2 - \left(\sum_{f=1}^F \frac{n_f}{N} a_f \right)^2}{V_P}$$

where

$$V_P = \sum_{f=1}^F \frac{n_f}{N} V_{Pf}$$

LOD profiles of JICIM for three traits in the maize NAM population



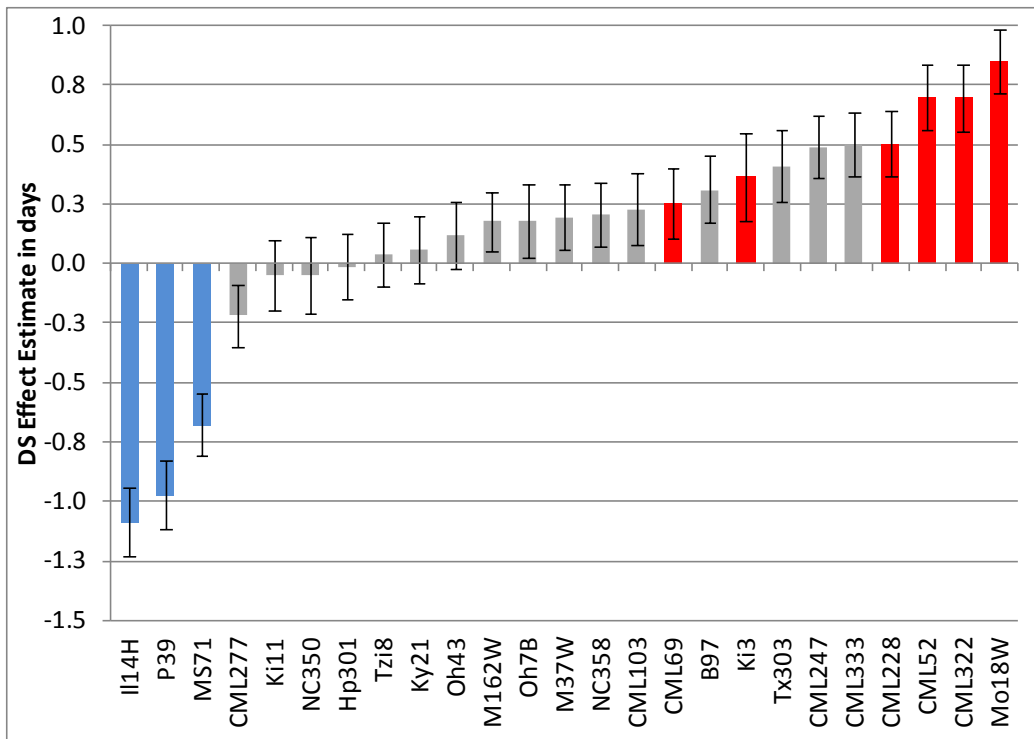
Blue: days to silk;

Red: days to anthesis;

Green: anthesis silking interval (ASI).

The vertical lines indicate the breaks between chromosomes.

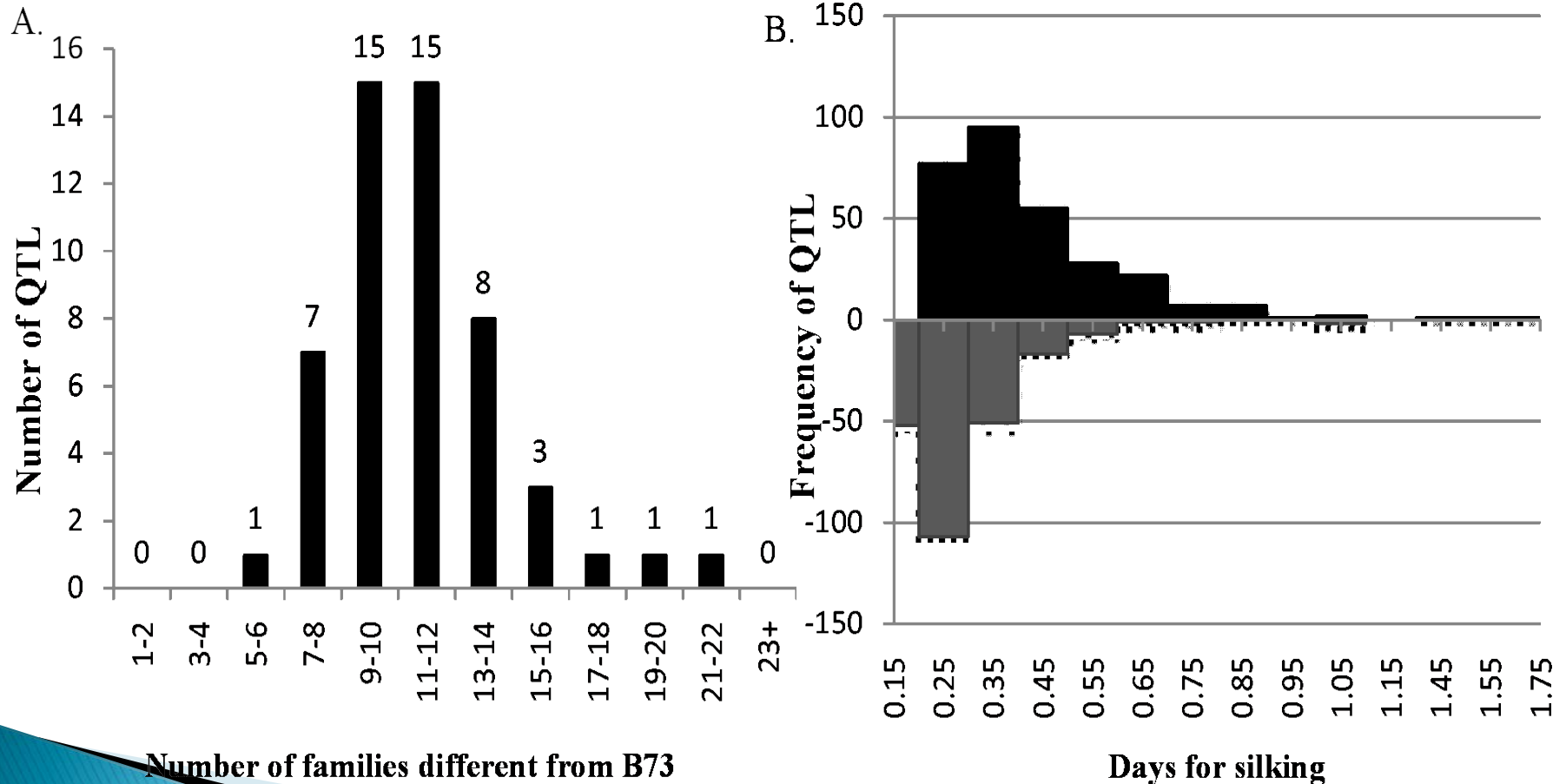
Association mapping at the candidate region of *Vgt1*



Estimated DS effects and standard errors for the *Vgt1* region of chromosome 8. Estimates are relative to B73 allele flowering. The blue alleles have the MITE at *Vgt1* (Mo17 also scored at the same time carries the polymorphism and equivalent effect). A simple t-test of founder effect estimate for MITE versus non-MITE was highly significant ($P=2 \times 10^{-8}$). The red alleles carry polymorphisms at *Vgt1* target gene *Rap2.7*, which is also significantly different ($P=7 \times 10^{-4}$).

Mapping results of JICIM

- JICIM detected 52 QTL controlling maize day to silk;
- 52 QTL explained 79% of the phenotypic variance in total.



Power analysis on identified QTL

(SI: support interval)

QTL	Chr.	Pos. (cM)	QTL overlapped with marker				QTL located in the middle of marker interval			
			SI=1 cM	SI=2 cM	SI=5 cM	SI=10 cM	SI=1 cM	SI=2 cM	SI=5 cM	SI=10 cM
<i>qZ1</i>	9	57.3	93.6	97.3	97.9	98.7	74.4	90.0	97.4	98.6
<i>qZ2</i>	10	41.9	96.0	97.7	98.4	99.1	76.7	89.6	98.0	98.9
<i>qZ3</i>	1	84.6	96.1	97.4	97.9	98.6	76.3	90.9	97.6	98.8
<i>qZ4</i>	8	69.7	95.8	97.4	98.3	98.8	78.8	91.3	97.8	98.7
<i>qZ5</i>	2	125.9	94.2	97.2	98.1	98.9	75.8	90.0	98.5	98.8
<i>qZ6</i>	3	105.0	93.6	97.0	98.1	98.9	74.3	89.0	97.4	98.8
<i>qZ7</i>	3	56.0	94.8	97.1	98.0	98.6	77.5	90.5	97.6	98.8
<i>qZ8</i>	2	76.5	93.9	96.7	98.1	98.8	74.2	88.3	97.6	98.8
<i>qZ9</i>	4	77.2	89.1	96.4	98.5	99.0	71.7	88.6	97.5	98.8
<i>qZ10</i>	7	76.5	89.8	96.1	98.1	98.7	72.0	89.4	96.7	98.7
<i>qZ11</i>	1	181.9	90.2	96.4	98.2	98.9	71.6	89.2	97.3	98.9
<i>qZ12</i>	5	78.4	88.8	96.2	98.6	99.2	70.3	87.8	97.8	98.8
<i>qZ13</i>	3	128.4	92.3	96.6	98.0	98.5	70.6	87.4	97.1	98.7
<i>qZ14</i>	1	60.8	88.0	95.1	98.9	99.0	70.4	86.4	97.0	98.6
<i>qZ15</i>	6	86.2	90.4	96.2	98.1	98.8	71.1	87.3	97.3	98.6
<i>qZ16</i>	9	62.9	95.2	97.5	98.0	98.8	76.9	91.4	98.1	98.7
<i>qZ17</i>	2	38.6	91.9	97.4	98.6	99.1	73.4	88.7	97.2	98.7
<i>qZ18</i>	3	70.6	93.8	97.9	98.6	99.1	75.7	90.1	97.6	98.6
<i>qZ19</i>	7	49.2	91.1	96.7	98.3	99.0	69.8	87.3	97.0	99.1
<i>qZ20</i>	1	137.6	91.0	96.7	98.2	98.7	73.7	89.2	97.4	98.9
<i>qZ21</i>	5	101.9	88.1	95.2	98.4	98.6	71.0	87.3	97.2	98.7
<i>qZ22</i>	1	31.7	89.1	95.8	98.5	99.4	71.2	86.6	97.0	98.7
<i>qZ23</i>	10	82.2	89.1	96.1	97.9	98.6	69.0	87.4	97.1	98.9
<i>qZ24</i>	4	111.5	88.0	95.4	98.6	98.9	70.4	86.4	97.0	98.6
<i>qZ25</i>	6	27.8	87.7	95.7	98.4	99.0	68.6	85.6	97.3	98.9
<i>qZ26</i>	2	107.8	89.4	95.6	98.2	99.2	71.2	87.7	97.2	98.9
<i>qZ27</i>	8	119.0	86.5	94.2	97.9	98.8	69.4	87.3	95.7	98.7
<i>qZ28</i>	5	0.0	88.0	95.0	98.0	98.6	68.3	86.7	96.8	98.6
<i>qZ29</i>	4	7.7	88.4	96.1	98.5	98.9	70.1	87.3	97.7	98.7
<i>qZ30</i>	8	18.3	87.7	95.8	98.5	99.3	69.4	86.4	96.3	98.4

One *Arabidopsis* NAM population

• *Ler* (*Landsberg erecta*)

× Antwerp (An-1)

64 markers 120 RILs

• *Ler* (*Landsberg erecta*)

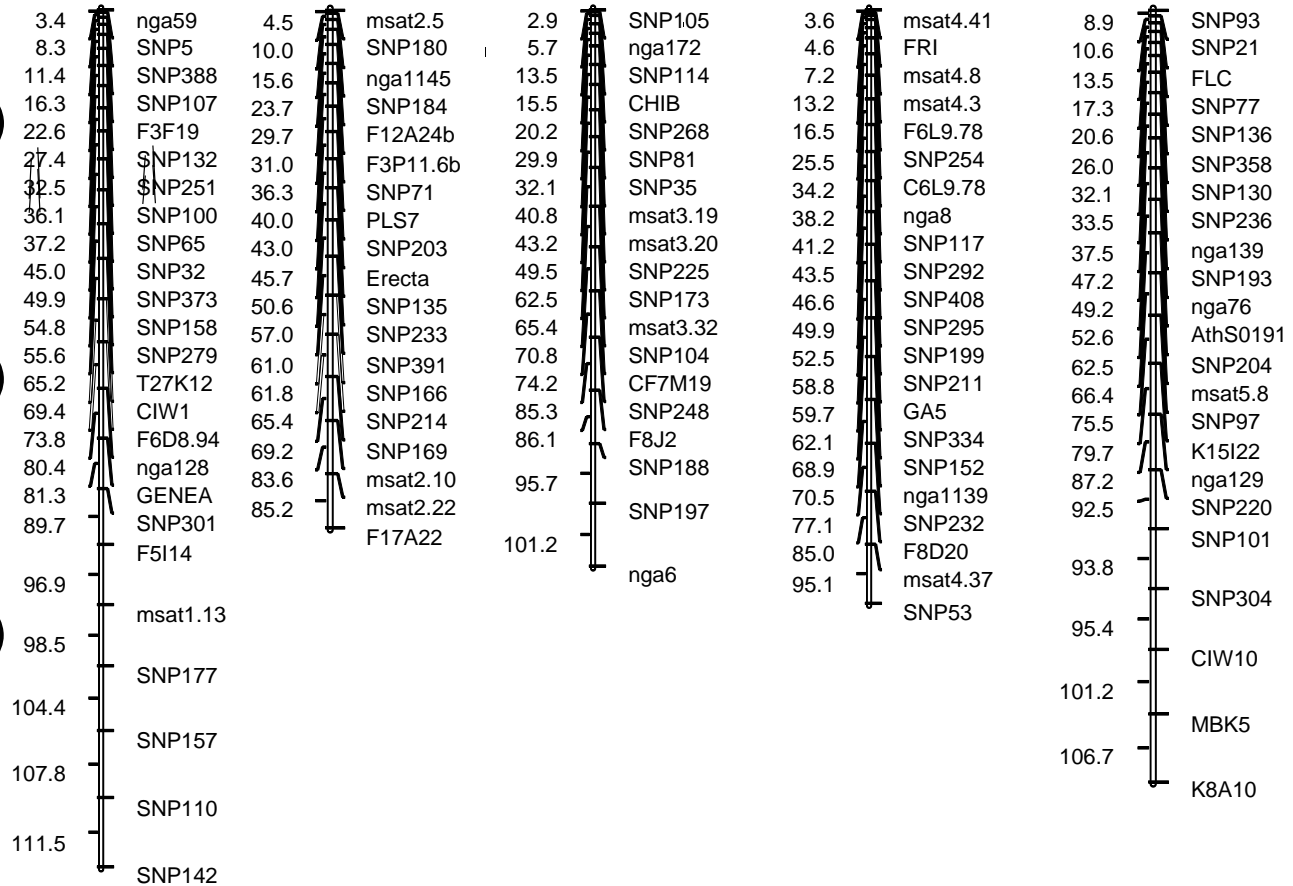
× Kashmir (Kas-2)

77 markers 164 RILs

• *Ler* (*Landsberg erecta*)

× Kondara (Kond)

75 markers 121 RILs



Chromosome 1

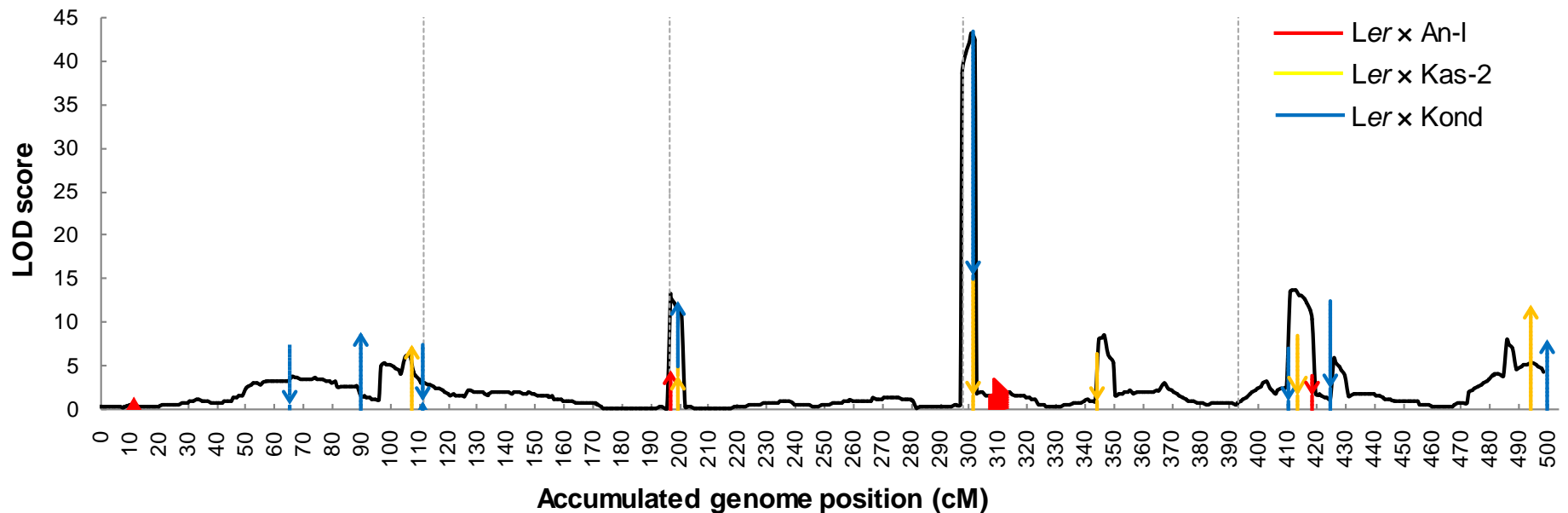
Chromosome 2

Chromosome 3

Chromosome 4

Chromosome 5

LOD profile of JICIM in the *Arabidopsis* NAM population



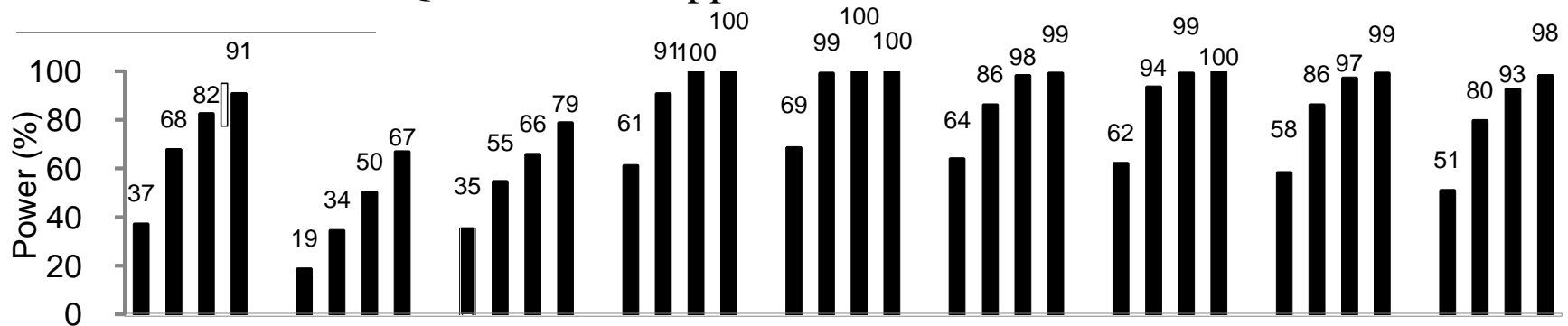
Mapping results of JICIM for the *Arabidopsis* NAM population

QTL	Chr.	Pos. (cM)	Left marker (cM)	Right marker (cM)	LOD score	Additive genetic effect (day) in each family ^b		
						Ler×An-I	Ler×Kas-2	Ler×Kond
<i>qA1</i>	1	66.0	CIW1* (65.2)	F6D8.94 (69.4)	3.75	-0.55	1.19	-2.64 (-3.00)
<i>qA2</i>	1	107.0	SNP157 (104.4)	SNP110* (107.8)	4.69	-0.12	1.32 (2.51)	1.00
<i>qA3</i>	2	0.0	msat2.5* (0.0)	msat2.5* (0.0)	3.05	-0.01	-0.38	-2.87 (-3.00)
<i>qA4</i>	3	0.0	SNP105* (0.0)	nga172* (2.9)	13.39	2.32 (1.80)	1.05 (2.00)	2.74 (3.00)
<i>qA5</i>	4	3.0	msat4.41 (0.0)	FRI* (3.6)	43.23	0.82	-5.60 (-6.00)	-11.24 (-11.40)
<i>qA6</i>	4	49.0	SNP295* (46.6)	SNP199 (49.9)	8.51	-0.06	-3.90 (-2.60)	-2.02
<i>qA7</i>	5	19.0	SNP136* (17.3)	SNP358* (20.6)	13.67	-1.50 (-1.60)	-2.78 (-3.40)	-3.48 (-2.80)
<i>qA8</i>	5	33.0	SNP236* (32.1)	nga139 (33.5)	5.87	0.05	-0.99	-4.18 (-5.00)
<i>qA9</i>	5	93.0	SNP101 (92.5)	SNP304 (93.8)	8.00	-1.12	2.76	-1.01

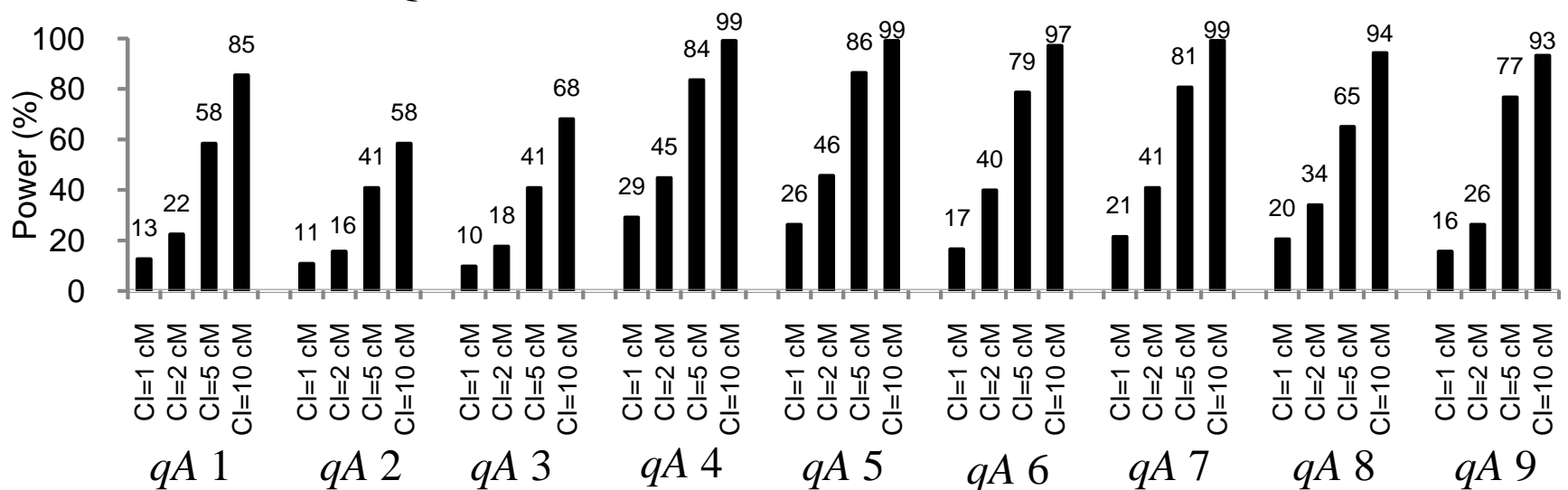
Power analysis on identified QTL

(SI: support interval)

A. Power when QTL was overlapped with marker

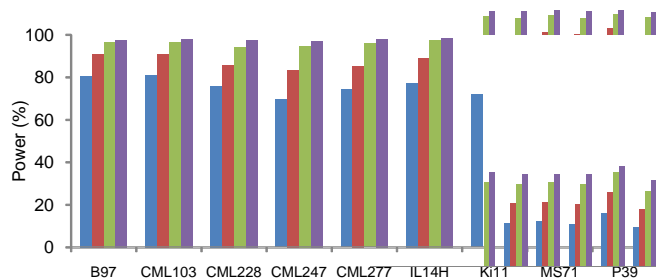


B. Power when QTL was located in the middle of marker interval

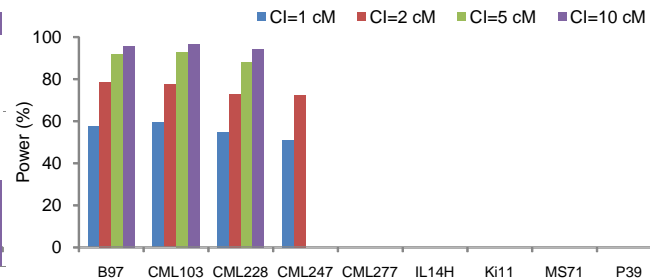


Power analysis of rare QTL

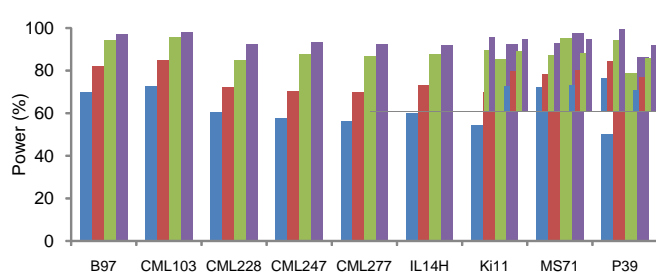
A. Additive effect of 1.5 days when QTL was overlapped with marker



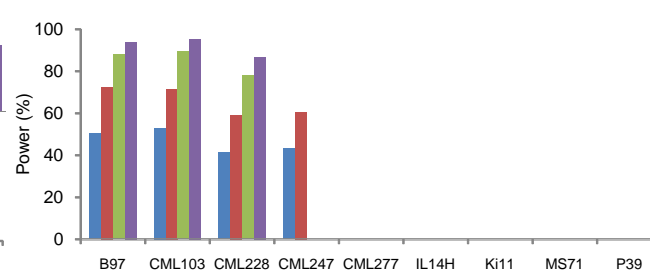
E. Additive effect of 1.5 days when QTL was located in the middle of marker interval



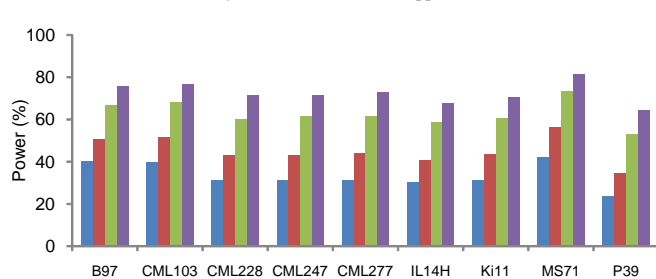
B. Additive effect of 1.0 day when QTL was overlapped with marker



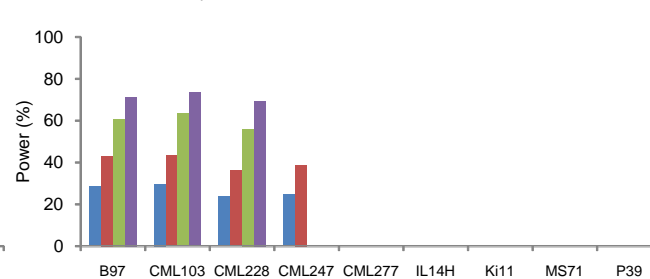
F. Additive effect of 1.0 day when QTL was located in the middle of marker interval



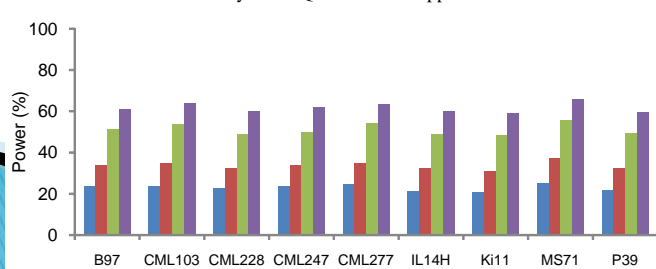
C. Additive effect of 0.5 day when QTL was overlapped with marker



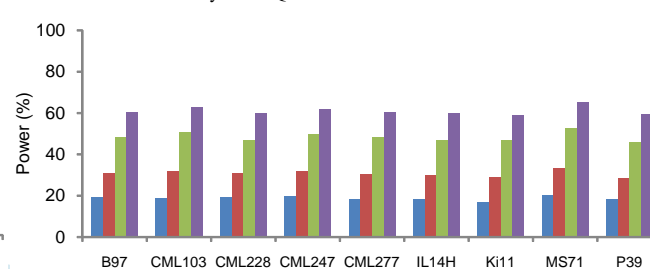
G. Additive effect of 0.5 day when QTL was located in the middle of marker interval



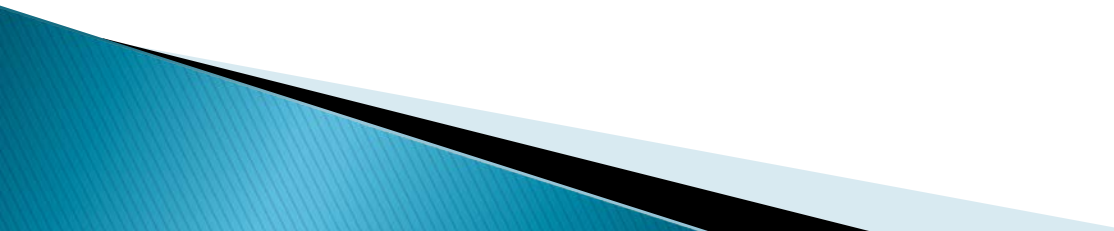
D. Additive effect of 0.25 day when QTL was overlapped with marker



H. Additive effect of 0.25 day when QTL was located in the middle of marker interval



**Whether it is more efficient to
use individual family mapping
for rare QTL?**

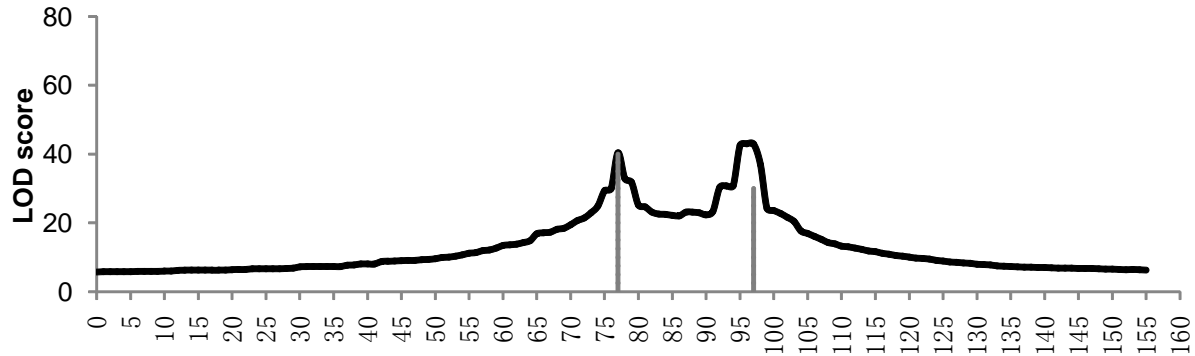


Power analysis for Rare QTL

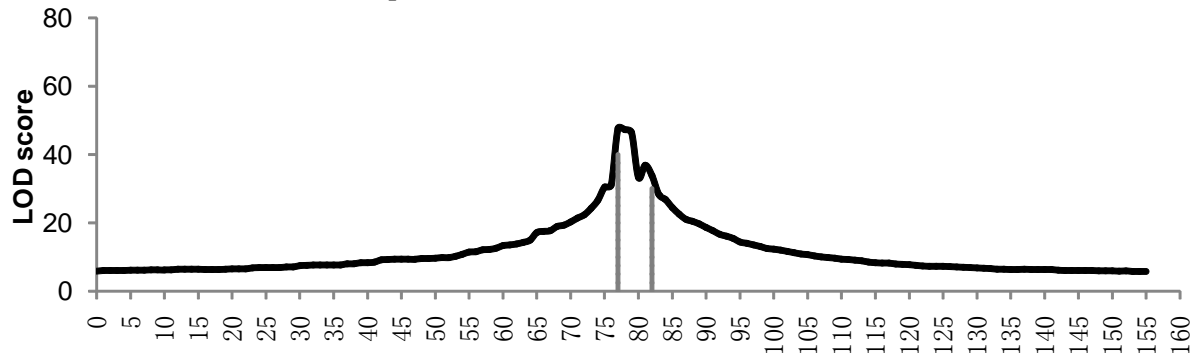
Method	Additive effect (day)	QTL overlapped with marker				QTL located in the middle of marker interval			
		1 cM	2 cM	5 cM	10 cM	1 cM	2 cM	5 cM	10 cM
JICIM in the maize population	1.50	77.4	89.1	97.5	98.6	55.2	75.0	91.0	96.4
	1.00	59.9	73.2	87.6	92.0	42.6	60.5	79.7	87.1
	0.50	30.4	40.8	58.8	67.7	22.6	35.1	53.3	64.9
	0.25	21.2	32.5	48.6	59.8	18.1	29.7	46.9	59.9
ICIM in family B73×II14H	1.50	84.7	92.2	94.8	94.8	51.6	71.2	87.1	93.5
	1.00	74.5	84.7	93.2	94.4	44.4	61.7	80.6	87.1
	0.50	25.3	42.4	53.2	61.0	21.9	33.3	49.5	57.3
	0.25	4.9	6.8	8.8	11.9	3.0	4.0	7.1	10.5

Dissection of linked QTL: LOD

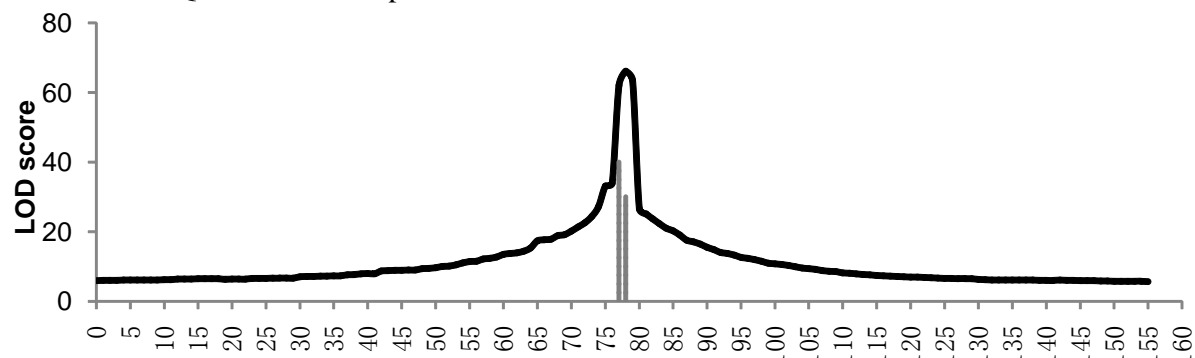
A. Two linked QTL were 20 cM apart



B. Two linked QTL were 5 cM apart



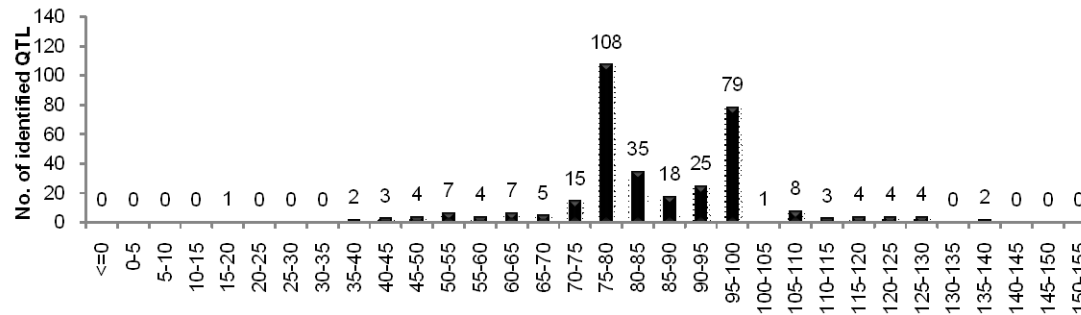
C. Two linked QTL were 1 cM apart



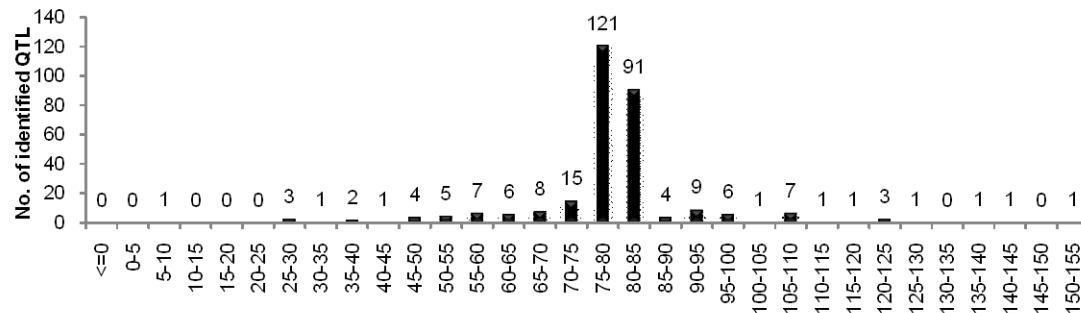
Position (cM) on chromosome 2

Dissection of linked QTL: power

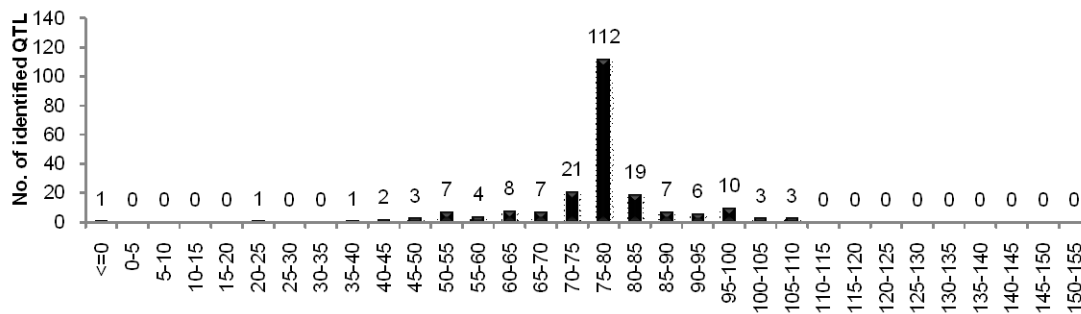
A. Two linked QTL were 20 cM apart



B. Two linked QTL were 5 cM apart



C. Two linked QTL were 1 cM apart



5 cM intervals on chromosome 2

Conclusions

- Mapping power of JICIM is high;
- JICIM can test the multiple alleles at the order of tens simultaneously;
- JICIM better avoids the confounding effect of population structure, and identified most genetic variation of flowering time;
- Simulation results showed that to map rare QTL with more than a 1-day effect, joint linkage mapping was not of merit.
- For small-effect QTL, the power of individual family mapping was slightly lower than the power of joint linkage mapping.

The NAM functionality in QTL IciMapping

One method available

- ▶ Joint inclusive composite interval mapping (JICIM) is available for NAM populations (Li et al., 2011. PLoS ONE 6(3): e17573)

Interface of the NAM functionality

The screenshot displays the 'Arabidopsis.nam' software interface. The top window shows a text editor with the following content:

```
*****Note: lines starting with "!" are remarks and will be ignored in the program*****
***** General Information *****
2      !Mapping Function (1 for Kosambi; 2 for Haldane; 3 for Morgan)
2      !Marker Space Type (1 for intervals; 2 for positions)
1      !Marker Space Unit(1 for centiMorgan; 2 for Morgan)
5      !Number of Chromosomes (or Linkage Group)
1      !Number of traits
3      !Number of families

***** Family information *****
!FamilyName  FamilySize  FamilyType (1 for BC1; 2 for BC2; 3 for DH; 4 for RIL)
An  120 4
Kas  164 4
Kond 121 4

***** Information for Chromosomes and Markers *****
!Chromosome;  NumMarkers on each chromosome
-Ch1  25
-Ch2  19
-Ch3  19
-Ch4  22
-Ch5  23

!MarkerName; Chromosome; position or interval as indicated by General Information
nga59  1    0
SNP5   1    3 4
```

The bottom panel, titled 'Parameters', contains the following settings:

- Missing Phenotype:** Deletion, Mean replacement
- Mapping Method:** ICIM-ADD(JICIM)
- Mapping Parameters:** Step (cM): 1.0000, Probability in stepwise regression: 0.0010
- LOD Threshold:** By manual input: 2.5000
- Selected Methods:** ICIM-ADD(JICIM)

LOD profile of ICIM additive mapping (ICIM-ADD in NAM, JICIM)

