Dissecting Quantitative Traits

Allen Van Deynze
Hamid Ashrafi

UC Davis
Tomato Breeders Round Table
June 28, 2009
What is being covered in this workshop?

What are Quantitative Trait Loci?

How can we identify and track these traits using molecular markers?

Different methods for QTL mapping (advantages/disadvantages)

Demo of WinQTL cartographer and R/QTL in tomato QTL mapping for different traits
Qualitative/Mendelian/Simple all refer to single gene traits

A trait that shows a discrete phenotype for each genotype

Typically controlled by a single gene

Little influence by the environment

Examples:
- Shriveled peas
- Flower color in roses
- Race specific disease resistance
A trait is **quantitative/polygenic** if its inheritance is controlled by two or more genes and their interaction with environment

1- A trait having a continuous distribution of phenotypes

2- Controlled by the interaction of many genes

Cont…
A trait is **quantitative/polygenic** if its inheritance is controlled by two or more genes and their interaction with the environment.

3- Influenced by the environment

Examples
- Yield
- Lycopene content in tomato
- Early Blight (EB) resistance in tomato
What are QTLs?

Genes are regions of DNA **controlling** a trait or phenotype.

Quantitative Trait Loci (QTLs) are regions of DNA **associated** with quantitative traits:
- They can be defined using molecular markers that highlight specific DNA segments or genes.

This can change how we manipulate these traits in breeding.
There are different QTL mapping methods

1- Maker Based Analysis (MBA)

2- Trait Based Analysis (TBA)
Detecting QTLs using MBA requires genotyping of individuals

Steps toward QTL analysis using MBA:

1- Segregating populations (F₂, BC, DH, RIL)

2- Measure phenotypic variation
   The more accurate measurement the more precise QTL identification.

3- Measure genotypic variation
   Genotype individuals
   Genetic map construction

4- Statistical analysis to find QTLs and QTL validation
Most popular QTL analysis methods for MBA are SMA, IM and CIM

- Single-Marker Analysis (SMA)
- Interval Mapping (IM)
  - Simple Interval Mapping (SIM)
    - MapMaker/QTL
    - QGENE
    - R/QTL
    - MapQTL5
  - Composite Interval Mapping (CIM)
    - QTL Cartographer
    - R/QTL
    - MapQTL5
    - QTL Network 2.0
- Multiple Interval Mapping (MIM)
  - WinQTL cartographer
Trait distribution in the progeny can indicate the presence of QTL

If we could observe directly the QTL we could see the 3 underlying trait distributions
SMA is done by analyzing one marker at a time

\[ y_j = \mu + f(A) + \varepsilon_i \]

A simple t-test
An analysis of variance
A linear regression
A likelihood ratio test
SMA is being used less nowadays due to its disadvantages and availability of more advanced statistical methods.

**Advantage:**
- Simplest approach for detecting QTLs
- Quick and less memory intensive

**Disadvantages:**
- May not detect loosely linked QTLs
- QTL effects are often underestimated
- Does not determine the distance between marker and QTL
- Does not determine the direction of QTL from marker
SIM is considering two markers at a time.

All methods are based on the expected co-segregation models among the two flanking markers:

1- Likelihood approach [Lander & Botstein (1986 and 1989)]

2- Regression approach
   - Nonlinear regression (Knapp et al. 1990)
   - Linear regression (Knapp et al. 1992)

3- A combination of 1 and 2 (Zeng 1994)

To read more about the models: Statistical Genomics, Linkage mapping and QTL analysis (B. H. Liu)
SIM was one the most commonly used methodology in QTL mapping

**Advantage:**
The likelihood map represents the position of the QTLs at various points of the genome

The probable position of the QTL is given by support intervals

Requires less progenies than the SMA

**Disadvantages:**
The number of QTLs cannot be resolved

The locations of the QTLs are sometimes not well resolved, the exact positions of the QTLs cannot be determined

The statistical power is still relatively low
CIM is a combination of SIM and multiple linear regression

The model for backcross population on a segment between marker $i$ and $i+1$

$$y_j = b_0 + b_i X_{ij} + \sum_{k \neq i, i+1} b_k X_{kj} + \varepsilon_j$$

The model for F$_2$ population on a segment between marker $i$ and $i+1$

$$y_j = \sum_{l=1}^{9} \theta_l X_{ilj} + \sum_{k \neq i, i+1} b_{ak} X_{akj} + \sum_{k \neq i, i+1} b_{dk} X_{dkj} + \varepsilon_j$$
CIM has at least four advantages over SMA and SIM

**Advantages:**

- By focusing on one genome region, the multidimensional search is reduced to one-dimensional search
- The resolution of QTL locations obtained is much higher than SMA and IM
- There are more variables in the model, making the model more efficient
- Markers can be used as boundary conditions to narrow down the most likely QTL position
Comparison of SMA, SIM and CIM for EB resistance in tomato
Comparison of SIM and CIM for EB resistance in tomato: scanning all chromosomes
Solanum lycopersicum $\times$ S. habrochaites
(EB Susceptible) $\downarrow$

F$_1$ $\times$ S. lycopersicum

BC$_1$

$\sim$850 BC$_1$ Plants

EB Resistant

EB Susceptible

Zhang et al 2003 Mol. Breed. 12: 3-19 Early blight resistance

Foolad et al 1997 Mol. Breed. 3: 269-277 Salt tolerance
Trait-based analysis (TBA) (Selective Genotyping)

Zhang et al. 2003 Mol. Breed. 12: 3-19 Early blight resistance
QTL mapping step by step: A practical example
A genetic map has to be first constructed
**QTL mapping step by step: A practical example**

A raw file is usually being used for the genetic map construction

<p>| | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data type ri self</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>172 294 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>#</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>*LC34L10a</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>-</td>
<td>B</td>
<td>H</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>5</td>
<td>*TG301</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>-</td>
<td>B</td>
<td>H</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>*TOS16116</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>-</td>
<td>B</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>7</td>
<td>*LC34L10b</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>H</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>8</td>
<td>*LEC71F17</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>H</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>9</td>
<td>*LEC72D1</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>10</td>
<td>*TG125</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>11</td>
<td>*CLE16L6</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>12</td>
<td>*TOF3A14</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>13</td>
<td>*TG70</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>14</td>
<td>*LC66G13b</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>15</td>
<td>*TOF20P4</td>
<td>B</td>
<td>-</td>
<td>B</td>
<td>B</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>16</td>
<td>*LEG6N2</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>17</td>
<td>*LEC6O2</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>18</td>
<td>*CT132PSU</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>19</td>
<td>*TG273</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>20</td>
<td>*LES9N20</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>H</td>
<td>H</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>21</td>
<td>*TG59</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>H</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>22</td>
<td>*CT191</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>H</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>23</td>
<td>*TOE6F10</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>H</td>
<td>B</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>24</td>
<td>*L66E21a</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>H</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>25</td>
<td>*LEG37F4</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>H</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>26</td>
<td>*TG559a</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>H</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>
QTL mapping step by step: A practical example
Phenotypic data should be collected and incorporated into the raw file
QTL mapping step by step: A practical example
A map file is also required for QTL analysis

*Chromosomes: 12
*chrom1 31 -1 0 0 -505.285
0 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 26 27 28 29 30
0.0004 0.1367 0.0614 0.0580 0.0103 0.0004 0.0004 0.0424 0.0818 0.0184
0.1190 0.0141 0.0069 0.0070 0.0544 0.0553 0.0146 0.0885 0.0639 0.0032
0.0004 0.0004 0.0818 0.0421 0.0439 0.1213 0.0244 0.0270 0.0555 0.0110
0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
*chrom2 32 -1 0 0 -394.876
31 32 33 34 35 36 37 38 39 40
41 42 43 44 45 46 47 48 49 50
51 52 53 54 55 56 57 58 59 60
61 62
0.0004 0.0004 0.0004 0.0004 0.0031 0.0004 0.0004 0.0004 0.0004 0.0032
0.0064 0.0732 0.0623 0.0033 0.1090 0.0674 0.0004 0.0727 0.0310 0.0788
0.0469 0.0004 0.0004 0.0385 0.0128 0.0274 0.0556 0.0278 0.1470 0.0438
0.0004
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
*chrom3 22 -1 0 0 -418.717
63 64 65 66 67 68 69 70 71 72
73 74 75 76 77 78 79 80 81
82 881
WinQTL cartographer and R/QTL Demo

A practical example of tomato QTL mapping using a RIL population

WinQTL cartographer
http://statgen.ncsu.edu/qtlcart/WQTLCart.htm

R/QTL
http://www.rqtl.org/