Conifer Translational Genomics Network
Coordinated Agricultural Project

Genomics in Tree Breeding and Forest Ecosystem Management

Module 9 – Mapping Quantitative Trait Loci (QTL)

Nicholas Wheeler & David Harry – Oregon State University
Requirements for identifying QTL

- An appropriate population
- Informative markers (genotypes)
- Framework map with complete genome coverage
- Good phenotypes
- Analytical tools
- Verification
QTL mapping: Conceptual steps

- Create mapping population
- Evaluate phenotypes
- Determine marker genotypes
- Examine phenotypic means among genotypic groups

Figure Credit: White, T. L., W. T. Adams, and D. B. Neale. 2007. Forest genetics. CAB International, Wallingford, United Kingdom. Used with permission.
Genotypes and phenotypes in QTL mapping

Figure Credit: Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics, Flint and Mott, 2001.
QTL mapping overview

- Basic idea is to associate quantitative trait phenotypes with the presence of genetic markers

- Use a population (full-sib family) segregating for markers and a quantitative trait
  - Marker genotypes are determined for progeny, as for linkage mapping
  - Same individuals must also be phenotyped
  - Offspring are grouped according to marker genotype
  - Phenotypes are averaged among offspring in different marker groups
  - If group means vary, then a QTL resides in the vicinity of the marker
QTL mapping overview

- Location and magnitude of QTL depend on
  - Marker density
  - Size of the mapping population
  - Likelihood of the offspring QTL genotype, given the marker genotype
  - Whether the trait is also affected by other QTL

- Analytical approaches
  - Single marker
  - Interval mapping
  - Composite interval mapping
Single marker approach

- Main advantage is simplicity
  - No map is needed (but is often used anyway)
  - Analyses can be done using standard statistical packages (e.g. t-test, ANOVA, regression)
  - Simplicity creates intuitive appeal

- Disadvantages
  - Map position of QTL lacks precision
  - Overall phenotypic effect is confounded with unknown recombination between marker and QTL
  - Cannot exclude the influence of other genomic regions on phenotype, so individual markers may overestimate QTL effects
  - No direct mechanism for including additional genetic information from adjacent markers
  - Comparisonwise-error rate
QTL profile: Single marker mapping significance testing

- Black diamonds depict the likelihood of a statistical difference between phenotypic means among marker classes.

- Tests are repeated for each marker along the chromosome.

- A linkage map is not needed, but they provide additional support, particularly when tests of adjacent markers approach significance.

- The significance threshold shown here (red line) depicts a genome-wide threshold given the number of markers assayed.

Figure Credit: Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics, Doerge, 2002.
Interval mapping

- Uses the mapped locations of other genetic markers to more accurately predict the location(s) of unseen QTL
  - Flanking markers predict probable haplotypes for nearby regions
  - Predictions are done in smaller intervals, essentially creating a sliding window analysis moving along a chromosome from one end to the other

- For each interval, individuals are grouped by their predicted genotypes and phenotypic variation is assessed

- Interval mapping groups individuals differently from single-marker analyses – phenotypic analyses are similar once individuals are grouped
Principles of interval mapping

Figure Credit: Reprinted, with permission, from the *Annual Review of Genomics and Human Genetics*, Volume 8, © 2007 by Annual Reviews, www.annualreviews.org
Interval vs. composite interval mapping (CIM)

- CIM is more robust to multiple QTL, particularly if they occur on the same chromosome.

- IM and CIM differ in how the phenotypic data are evaluated:
  - For interval mapping, phenotypic data are evaluated directly, without adjusting for possible genetic influences from outside the interval.
  - For CIM, intervals (genes) outside the interval are considered as well.
  - Other intervals are used as covariates, not unlike in multiple regression.
Comparing QTL mapping approaches

- Compare results from three mapping approaches: Single marker (black diamonds), interval mapping (blue curve), and CIM (green curve).

- CIM adjusts phenotypes for genetic influences outside the interval – QTL effects are less likely to be over-estimated.

- CIM identifies two QTL near the ends of the chromosome.

- The middle “ghost peak” QTL identified with interval mapping probably results from the combined influence of the true QTL on either side.

Figure Credit: Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics, Doerge, 2002.
Software for mapping QTL

- QTL Cartographer
  - http://statgen.ncsu.edu/qtlcart/manual/

- Mapmaker/QTL
  - www.broad.mit.edu/ftp/distribution/software/mapmaker3/

- QTL Express (GridQTL)
  - http://www.gridqtl.org.uk/
QTL mapping summary

- Basic strategy is to look for phenotypic differences among groups of individuals classified by genotype

- Prerequisites for QTL mapping include
  - A mapping population
  - Individuals that have been genotyped and phenotyped
  - Statistical analyses of genotypes and phenotypes

- Single-marker strategy
  - Simple to analyze using standard statistical tools
  - Analyses do not consider marker locations, but interpretation may be aided by knowing the map locations of the markers
QTL mapping summary

- **Interval mapping**
  - *Includes map information to infer marker genotypes within intervals flanked by markers*
  - *Intervals are processed as a “sliding-window” along the chromosome*
  - *Analysis requires specialized software*

- Both methods provide estimates for
  - *Additive (a) and dominance (d) quantitative genetic effects*
  - *Phenotypic effects (e.g. % phenotypic variation explained, PVE)*
QTL biology and utility

- QTL can be mapped – what else do we need to know?
  - How accurately have QTLs been mapped?
  - Do QTL exist in trees?
  - Can we find all QTL and estimate their effects?
  - Do large effect QTL exist?

- Are QTL interactions with environment, year, or genetic background important?

- Are QTL useful for breeding or other applications?

Figure Credit: Modified from Grattapaglia, 2007
Accuracy of QTL mapping

Figure Credit: Reprinted from Trends in Plant Science, Vol 11, Price, Believe it or not, QTLs are accurate! 213-216, Copyright 2006, with permission from Elsevier
# Accuracy of QTL mapping

Table 1. The distance between original QTL peak position and subsequently tagged or cloned genes in plant species

<table>
<thead>
<tr>
<th>Species</th>
<th>Trait</th>
<th>Gene or tagged locus</th>
<th>Mapping population</th>
<th>Distance to original LOD peak (cM)</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Major QTLs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tomato</td>
<td>Fruit size</td>
<td>fw2.2</td>
<td>264 BC₁ₛ</td>
<td>0.6</td>
<td>[9,10]</td>
</tr>
<tr>
<td>Tomato</td>
<td>Fruit shape</td>
<td>Ovate</td>
<td>82 F₂ₛ</td>
<td>0.0</td>
<td>[11,12]</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>Flowering time</td>
<td>FLW₁</td>
<td>98 RILₛ</td>
<td>0.0</td>
<td>[13]</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>Flowering time</td>
<td>CRY2</td>
<td>162 RILₛ</td>
<td>0.1ᵇ</td>
<td>[14,15]</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>Transpiration</td>
<td>ERECTA</td>
<td>100 RILₛ</td>
<td>&lt;1.0</td>
<td>[16]</td>
</tr>
<tr>
<td>Wheat</td>
<td>Frost tolerance</td>
<td>Cbf3</td>
<td>74 RILₛ</td>
<td>0.1ᵇ</td>
<td>[17]</td>
</tr>
<tr>
<td>Wheat</td>
<td>Grain protein</td>
<td>GPC</td>
<td>85 RICLₛ</td>
<td>0.2</td>
<td>[18,19]</td>
</tr>
<tr>
<td>Barley</td>
<td>Photoperiod response</td>
<td>Ppd-H1</td>
<td>94 DH</td>
<td>1.9</td>
<td>[20–22]</td>
</tr>
<tr>
<td>Soybean</td>
<td>Flowering time</td>
<td>FT₁</td>
<td>156 RILₛ</td>
<td>0.4</td>
<td>[23,24]</td>
</tr>
<tr>
<td>Brassica</td>
<td>Flowering time</td>
<td>COL₁</td>
<td>88 BC₁ₛ</td>
<td>0.4</td>
<td>[25,26]</td>
</tr>
<tr>
<td>Brassica</td>
<td>Eusic acid content</td>
<td>E₁</td>
<td>184 F₂ₛ</td>
<td>1.0</td>
<td>[27]</td>
</tr>
<tr>
<td><strong>Small QTLs</strong></td>
<td>Shoot morphology</td>
<td>tb₁</td>
<td>290 F₂ₛ</td>
<td>0.6ᵇ</td>
<td>[28–30]</td>
</tr>
<tr>
<td>Rice</td>
<td>Heading date</td>
<td>H₁</td>
<td>186 F₂ₛ</td>
<td>0.5</td>
<td>[31,33]</td>
</tr>
<tr>
<td>Rice</td>
<td>Heading date</td>
<td>H₂</td>
<td>186 F₂ₛ</td>
<td>0.3</td>
<td>[31,33]</td>
</tr>
<tr>
<td>Rice</td>
<td>Heading date</td>
<td>H₃</td>
<td>186 F₂ₛ</td>
<td>0.0</td>
<td>[31,33]</td>
</tr>
<tr>
<td>Rice</td>
<td>Heading date</td>
<td>H₄</td>
<td>186 F₂ₛ</td>
<td>0.2</td>
<td>[30,32]</td>
</tr>
<tr>
<td>Rice</td>
<td>Heading date</td>
<td>H₅</td>
<td>186 F₂ₛ</td>
<td>1.2</td>
<td>[30,32]</td>
</tr>
<tr>
<td>Rice</td>
<td>P uptake</td>
<td>Pup₁</td>
<td>98 BILₛ</td>
<td>1.0</td>
<td>[34,35]</td>
</tr>
<tr>
<td>Rice</td>
<td>Grain weight</td>
<td>gw₃.₁</td>
<td>258 BC₂F₂ₛ</td>
<td>&lt;1.6</td>
<td>[36,37]</td>
</tr>
<tr>
<td>Potato</td>
<td>Sugar content</td>
<td>inv/GE</td>
<td>146 F₁ₛ</td>
<td>&lt;3.0</td>
<td>[38,39]</td>
</tr>
</tbody>
</table>

*Abbreviations: BC₁, backcross 1; BC₂F₂, selfed backcross 2; BIL, backcross inbred lines; DH, double haploids; RICL, recombinant inbred chromosome lines; RIL, recombinant inbred lines. Note, because potato is inbreeding, an F1 is a segregating population.

ᵇPosition based on mean position of multiple traits or trait screens.

Figure Credit: Reprinted from Trends in Plant Science, Vol 11, Price, Believe it or not, QTLs are accurate! 213-216, Copyright 2006, with permission from Elsevier
Distribution of QTL effects

Figure Credit: Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics, Flint et al., 2005.
Verification of QTL stability and interaction effects

- To be useful, QTL must be stable across environments, years, and genetic backgrounds.

- Definition of verification: The repeated detection, at a similar position on the genetic map, of a QTL controlling a trait under more than one set of experimental conditions (Brown et al., 2003).

Clonally replicated verification trial of Douglas-fir, age 4 years from cutting. Image Credit: Nicholas Wheeler, Oregon State University.
3-generation pedigree and mapping populations

Maternal Grandmother (late flushing)  
Maternal Grandfather (early flushing)  
F₁ Parent  
(1991)

Paternal Grandmother (late flushing)  
Paternal Grandfather (early flushing)  
F₁ Parent  
(1994)

clonally replicated progeny linkage map (Jermstad et al. 1998)

Twin Harbors, WA test site (n=224) (Jermstad et al. 2001a, 2001b)

Turner, OR test site (n=78) (Jermstad et al. 2001a)

Bud flush experiment (n=429)

Winter chill (WC) hours
750 1500

Flushing temperature (FT) °C
10 15 20

Growth cessation experiment (357 < n < 407)

Daylength (DL)
NDL  EDL

Moisture stress (MS)
MS  NMS

Detection Population

Verification Population

Field Experiment

Longview, WA test site (n=408)
Springfield, OR test site (n=408)

Figure Credit: Modified from Jermstad et al., 2003
The other QTL mapping requirements

- Markers and genome coverage
  - 74 evenly spaced, highly informative RFLP markers
  - Map length of ~900 cM, marker density ~ every 12 cM

- Phenotypes
  - Bud flush etc (annually 1996-2001)

- Analytical Tools
  - Haley-Knott multiple marker interval mapping approach; scanned LG at 5 cM intervals, 1 and 2 QTL models; single marker approach
Bud flush QTLs in Douglas-fir

Figure Credit: Nicholas Wheeler, Oregon State University
Figures used with permission of the Genetics Society of America from "Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. III. QTL by environment interactions and verification" Jermstad et al. Genetics 165: 1489-1506. 2003; permission conveyed through Copyright Clearance Center, Inc
QTL maps and positional candidate genes

LG1

-10.0
3.7
4.2
9.1
29.4
45.1
60.0
62.0
66.5
70.0
77.0
103.0
106.0
106.1
110.0
117.0
118.1
121.5
131.0
138.0
146.3

Alpha tubulin

Cohort 1
Cohort 2

Pm1011_a
Pm1147_a
Pm1011_b
Pt2356_d

Pm1052_j

Pm1486_a
Pm1383_a
CABBP_1

Pm1174_a

40S_RPS2
DER1-like
Pm1592_a
CABBP_2

UGT
Pt2006_b
Pm1496_a

Pm1301_a
ACRE146

TBE
Pt2291_g

EF-1 (translation elongation factor -1)
CABBPI (chlorophyll a/b-binding protein type 1)
DER1-like (degradation of misfolded proteins)
CABBP2 (chlorophyll a/b-binding protein type 2)
F3H (flavanone-3-hydroxylase)
LEA-II (late embryogenesis abundant type II) dehydrin-like protein
MT-like (metallothionein-like protein)
SAHH (S-adenosyl-L-homocysteinase hydrolase)

LG2

0.0
16.6
26.7
33.3
66.4
76.0
78.6
80.0
94.1
147.2

Pt2957_a

Pm1504_b

Pm1052_c
Pm1611_b
Pm1301_a

Pm0343_a

F3H

Pm0123_a

MAD

Pm1611_b

Pm0343_a

Pm1301_a

Pm1174_a

Pm1052_c

Pm1052_j

40S_RPS2
DER1-like
Pm1592_a
CABBP_2

UGT
Pt2006_b
Pm1496_a

Pm1301_a
ACRE146

TBE
Pt2291_g

EF-1 (translation elongation factor -1)
CABBPI (chlorophyll a/b-binding protein type 1)
DER1-like (degradation of misfolded proteins)
CABBP2 (chlorophyll a/b-binding protein type 2)
F3H (flavanone-3-hydroxylase)
LEA-II (late embryogenesis abundant type II) dehydrin-like protein
MT-like (metallothionein-like protein)
SAHH (S-adenosyl-L-homocysteinase hydrolase)

LG4

7.0
20.0
38.0
43.2
47.0
48.0
70.0
75.0
79.6
88.0

Pt2957_a

Pm1486_e

Pm1480_a_MMIP

Pt2553_a

PRS

MT-like

LEA-II

ANT
SAHH

Formin-like

Bud flush

Fall cold hardiness (buds)
Fall cold hardiness (stem)
Fall cold hardiness (needles)
Spring cold hardiness (buds)
Spring cold hardiness (stem)
Spring cold hardiness (needles)

cold-induced
downregulated under the water deficit
cold-induced
downregulated under the water deficit
cold-induced
downregulated under the water deficit
cold-induced
stress-induced; downregulated under the water deficit

cold-induced
downregulated under the water deficit
cold-induced
stress-induced; downregulated under the water deficit

cold-induced
downregulated under the water deficit
cold-induced
stress-induced; downregulated under the water deficit

Figure Credit: Modified from Wheeler et al., 2005
Figure Credit: Modified from Brown et al., 2003
QTL studies are informative and useful

- Complex trait dissection and genetic architecture
  - Number of QTL influencing a trait
  - Size of the QTL effects (PVE)
  - Location of the QTL
  - Parental contribution of allelic effects
  - QTL by environment interactions

- Provide a foundation for MAS

- Identify positional candidate genes
QTL characteristics

- QTL in trees exist and they are common

- Population size affects both the number of QTL detected and the size of QTL effects (Beavis, 1995)
  - More individuals (~500) are better than fewer
  - Clonal studies increase reliability of phenotypic assessments, and increase detection sensitivity for QTL of small effect

- In trees, most QTL account for less than 5% of a trait’s phenotypic variation, although collectively, multiple QTL may account for a substantial amount of the total genetic variation for that trait
QTL characteristics

- QTL stability or expression is highly variable
- Some QTL are expressed repeatedly across years, environments, and even crosses (< 20% in our studies), but most are detected more sporadically
- This is a function of factors such as
  - Low heritability of a QTL
  - Poor experimental design
  - Not all QTL will segregate in every cross
Linkage disequilibrium

- Most outcrossing forest trees have genomes that are largely in linkage equilibrium
- QTL discovered in one cross may not exist in another cross, or if they do, marker phase may be different
References cited


References cited


References cited

External Links


Thank You.
Conifer Translational Genomics Network
Coordinated Agricultural Project

CTGN  CAP  UCDAVIS  USDA

United States Department of Agriculture
National Institute of Food and Agriculture