Workshop

Comparative Mapping in Salicaceae

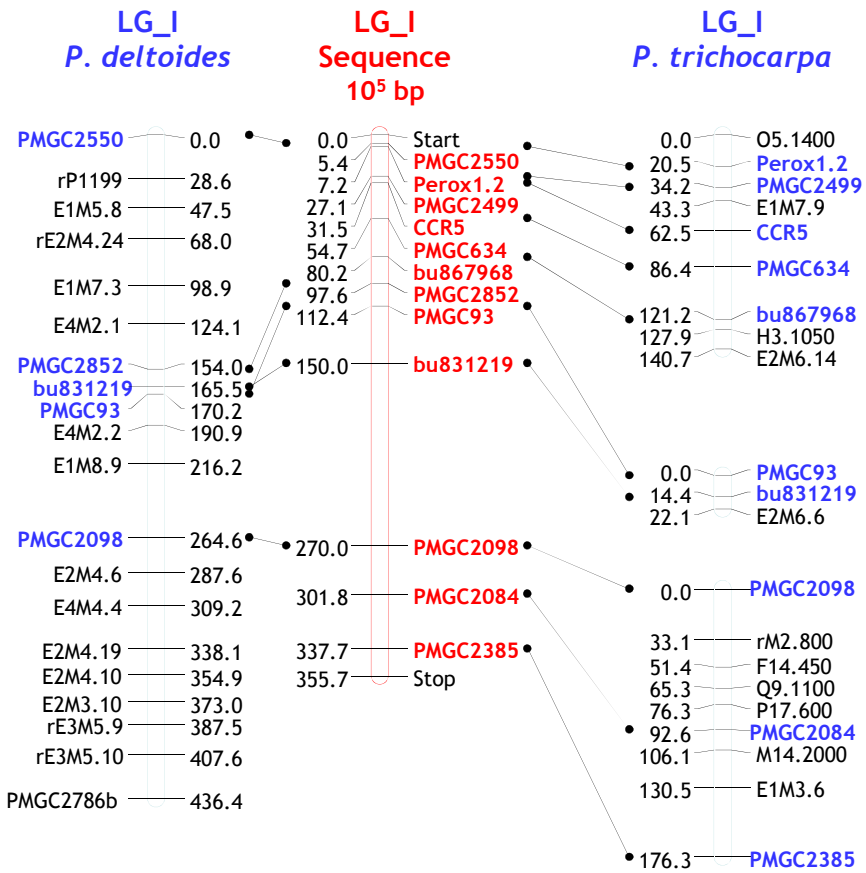
QTL projection on the genome and gene list analysis

Véronique Jorge
INRA Orléans, P1d

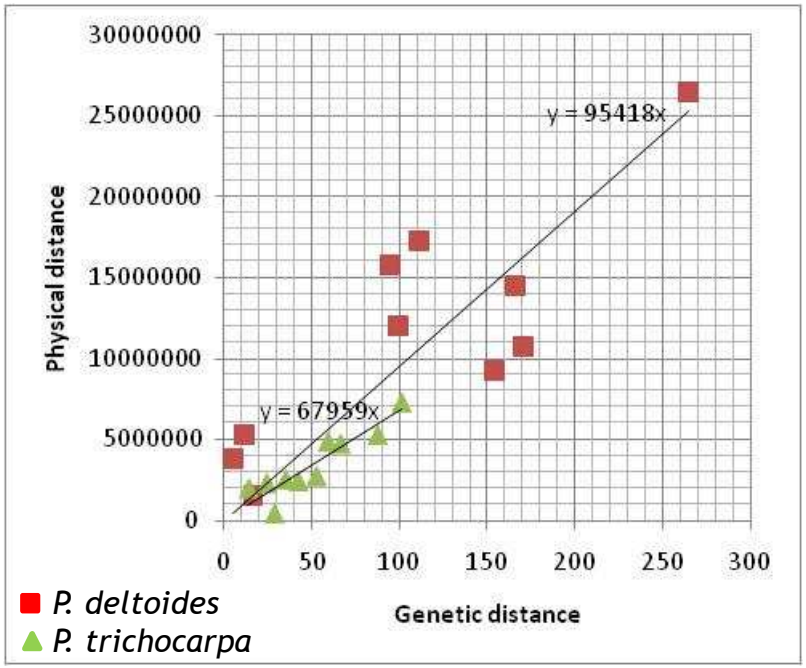
8-10 March 2010

Anchoring to the genome sequence and bp/cM ratio calculation

Example of POP2 map



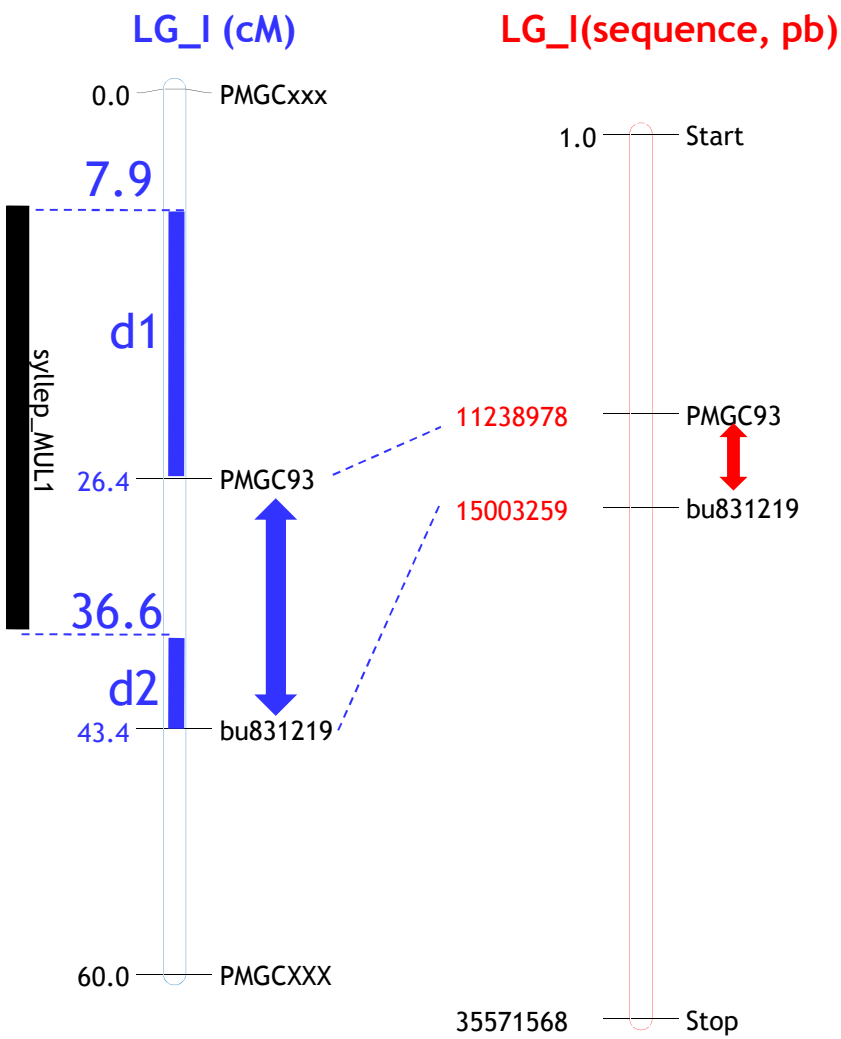
Relationship between pairwise genetic and physical distances between genome anchored markers



$$y = R\,x$$

Number of bp / cM

Calculation of start and stop of the QTL projected on the genome using the nearest markers and R (projection geometry)



$$\text{Start(pb)} = 11238978 - [|\overset{d1}{26.4 - 7.9}| \times R]$$
$$= 7\ 142\ 811 \text{ pb}$$

$$\text{Stop(pb)} = 15003259 - [|\overset{d2}{43.4 - 36.6}| \times R]$$
$$= 13\ 497\ 405 \text{ pb}$$

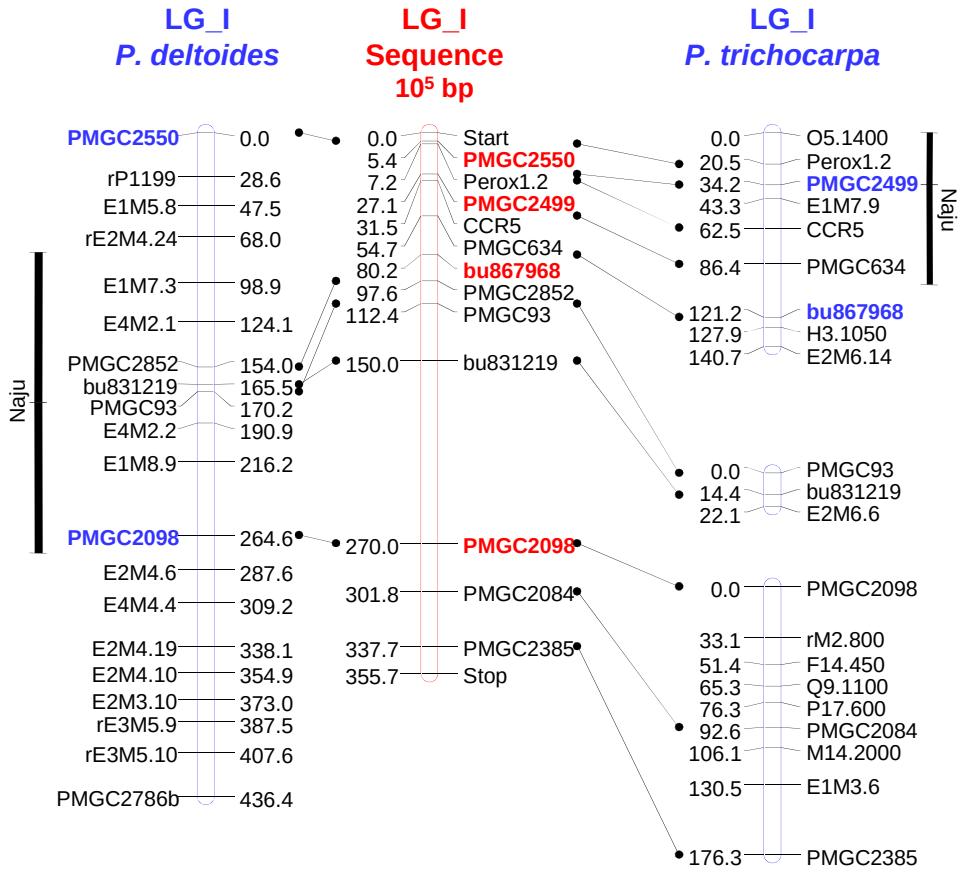
Example of physical QTL intervals calculation and corresponding number of genes

P. trichocarpa genome sequence v1.1, POP2 genetic maps

Nitrogen content:

8 QTL, 9.7% - 34.7% of the total PVE

197.8 cM
17 670 094 bp
1393 genes



34.19 cM
6 188 807 bp
724 genes

Analysis of gene lists

- Annotation use a classification based on Gene Ontology (GO)
  - → Molecular function (ex: oxidoreductase activity, …)
  - → Biological process (ex: photosynthesis light harvesting, …)
  - → Cellular component (ex: membrane, chloroplast …)

GO:0008150 Name: biological_process (root)

- GO:0007275 development pv: 1.00 (Query list: 11)
- GO:0007582 physiological process pv: 1.00 (Query list: 77)
- GO:0009987 cellular process pv: 1.00 (Query list: 72)
- GO:0050789 regulation of biological process pv: 1.00 (Query list: 14)
- GO:0050896 response to stimulus pv: 2.7e-03 (Query list: 31)
  - GO:0006950 response to stress pv: 1.2e-05 (Query list: 23)
  - GO:0009628 response to abiotic stimulus pv: 1.1e-03 (Query list: 25)
    - GO:0006970 response to osmotic stress pv: 1.4e-03 (Query list: 8)
    - GO:0009266 response to temperature stimulus pv: 1.1e-07 (Query list: 14)
    - GO:0009415 response to water pv: 3.7e-09 (Query list: 12)
    - GO:0042221 response to chemical stimulus pv: 1.00 (Query list: 11)
  - GO:0009719 response to endogenous stimulus pv: 0.40 (Query list: 11)

Analysis of gene lists

5 QTLs explaining 2.8 to 9% of the variation for nb of sylleptic branches

QTL set
2906 genes included
in the whole confidence intervals

Genome set
45 555 genes in
the genome sequence
(~60% have GO terms)

➢ Enrichment in GO terms is tested comparing the QTL set and genome set

ID	Name	p-Value	p-Value (Adj)	Study Count	Population Count
GO:0010422	regulation of brassinosteroid biosynthetic process	0,0000	0,0000	5	6
GO:0010023	proanthocyanidin biosynthetic process	0,0000	0,0000	6	10
GO:0009641	shade avoidance	0,0000	0,0000	6	11
GO:0016595	glutamate binding	0,0001	0,0001	4	5
GO:0009741	response to brassinosteroid stimulus	0,0003	0,0003	10	38
GO:0009788	negative regulation of abscisic acid mediated signaling pathway	0,0006	0,0006	5	11
GO:0006122	mitochondrial electron transport, ubiquinol to cytochrome c	0,0006	0,0006	5	11
GO:0004406	H3/H4 histone acetyltransferase activity	0,0008	0,0008	4	7
GO:0048830	adventitious root development	0,0008	0,0008	4	7

Analysis of gene lists

References

Bauer, S., S. Grossmann, M. Vingron, and P.N. Robinson. 2008. Ontologizer 2.0 - a multifunctional tool for GO term enrichment analysis and data exploration. Bioinformatics 24:1650-1651.

Huang, D.W., B.T. Sherman, and R.A. Lempicki. 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research 37:1-13.