



HAL
open science

Comparative mapping in Salicaceae: a tool for identifying important genes controlling adaptive traits

Véronique V. Jorge, Francesco Fabbrini, Patricia P. Faivre-Rampant, Matthias Fladung, Muriel Gaudet, Ulf Lagercrantz, Jean-Charles Leplé, Michele Morgante, Birte Pakull, Maurizio Sabatti, et al.

► To cite this version:

Véronique V. Jorge, Francesco Fabbrini, Patricia P. Faivre-Rampant, Matthias Fladung, Muriel Gaudet, et al.. Comparative mapping in Salicaceae: a tool for identifying important genes controlling adaptive traits. *Forest ecosystem genomics and adaptation*, Jun 2010, San Lorenzo de El Escorial, Spain. hal-02755815

HAL Id: hal-02755815

<https://hal.inrae.fr/hal-02755815>

Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Comparative mapping in *Salicaceae*: a tool for identifying important genes controlling adaptive traits

Véronique Jorge,

Francesco Fabbrini, Patricia Faivre-Rampant, Matthias Fladung, Muriel Gaudet, Ulf Lagercrantz, Jean-Charles Leplé, Michele Morgante, Birte Pakull, Maurizio Sabatti, Véronique Storme, Gail Taylor, Nicola Vitacolonna, Jennifer De Woody, Catherine Bastien



University of Udine (DISA)



Comparative QTL mapping

Definition of comparative mapping

Alignment of chromosomes of related species based on genetic mapping of common DNA markers

Goal

Identification of conserved/specific genomic regions.

Specific questions

How cumulated QTL data based on heterogeneous mapping experiments could help functional annotation of genomes and identification of candidate genes?

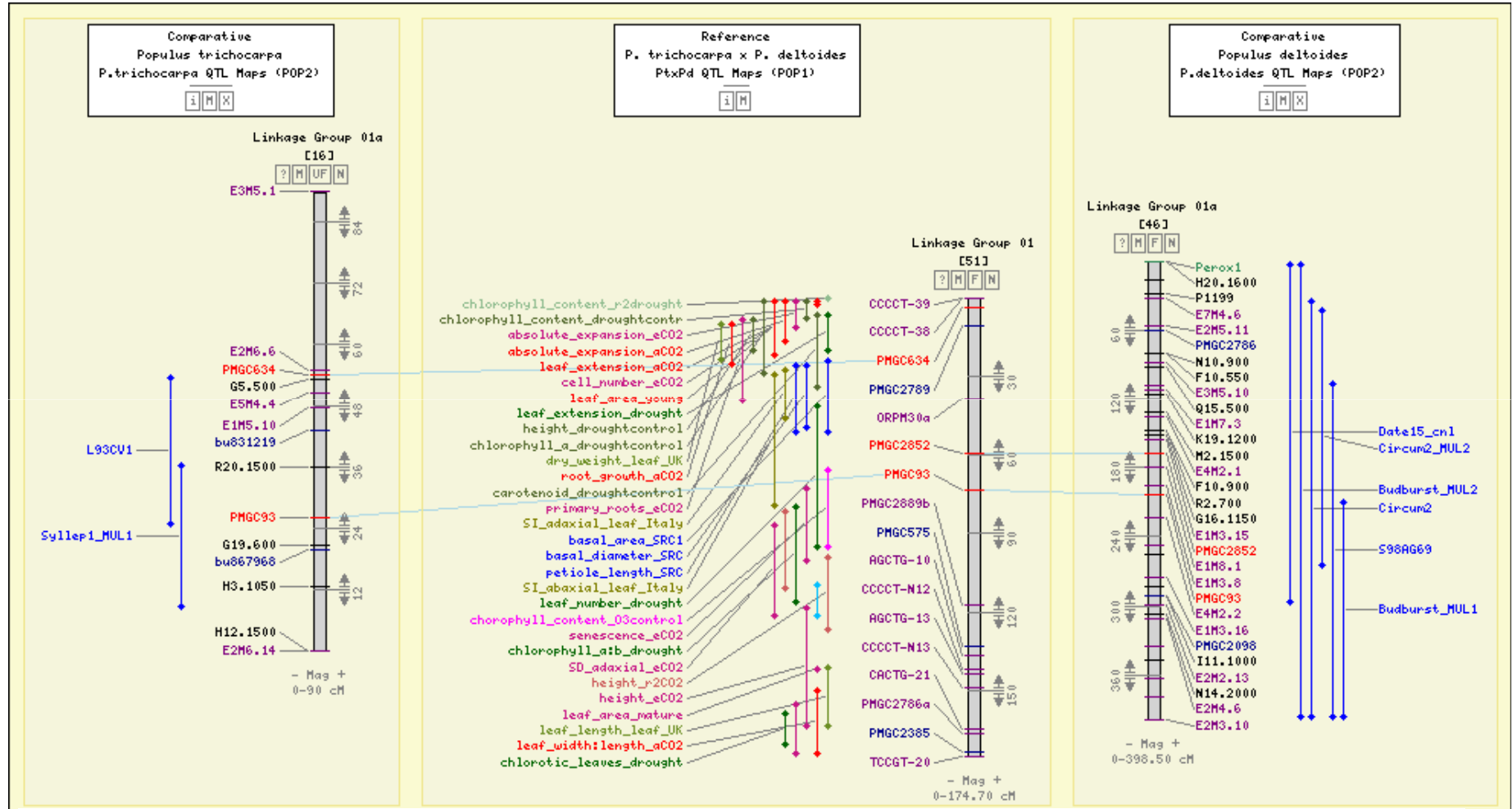
Strategies

- ✓ Simple graphical comparison
- ✓ Map construction and QTL analysis based on multiple mapping experiment (on raw data)
- ✓ Consensus and meta-analysis (on position data)



Methods for comparative mapping

Simple graphical comparison: matching marker names (Cmap)



➤ Rough identification of linkage groups/chromosomes of interest

Data from <http://www.evoltree.soton.ac.uk/cmap/>



Methods for comparative mapping

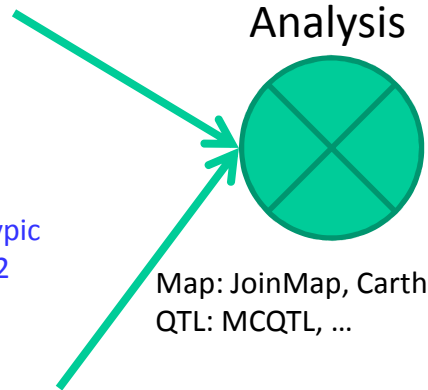
Map construction and QTL analysis based on multiple mapping experiment (on raw data)

Marker and phenotypic data on pedigree 1

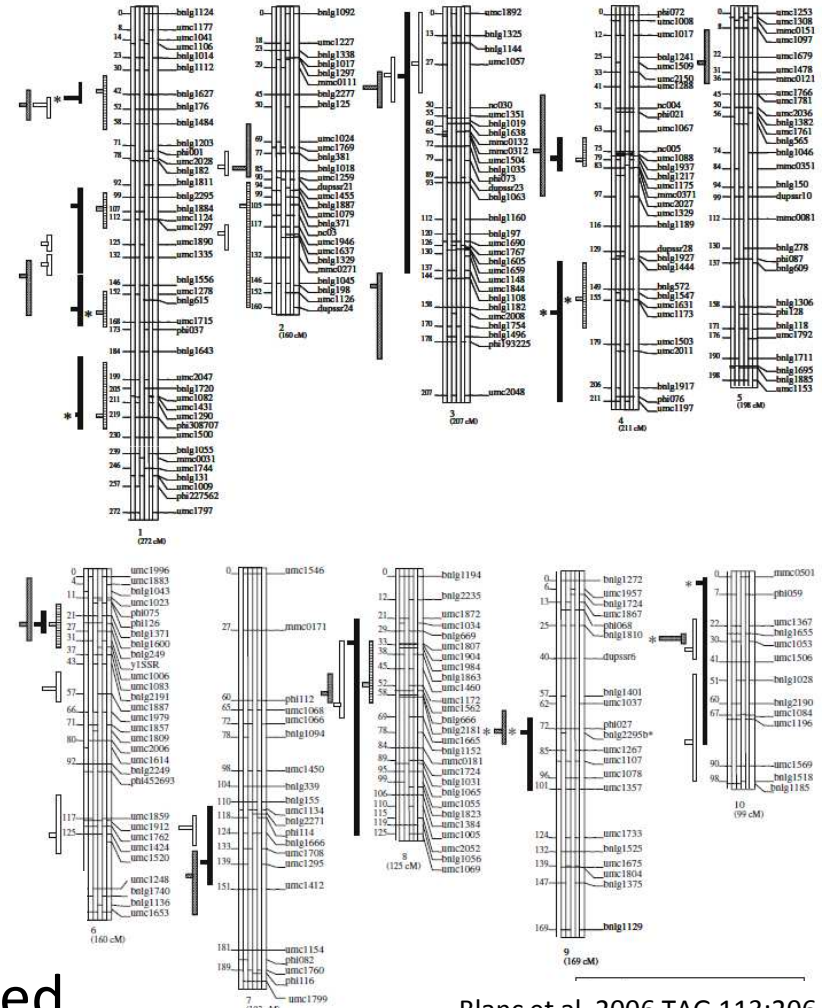
010001101110101
 110111101100110
 110110011010001
 110111101100110
 110111101100110
 25 35 21 18 20
 25 26 30 21 20

Marker and phenotypic data on pedigree 2

010001101110101
 110111101100110
 110110011010001
 110111101100110
 110111101100110
 25 35 21 18 20
 25 26 30 21 20



Map: JoinMap, Carthagene
 QTL: MCQTL, ...



- Raw data not easily available
- Connected pedigrees often needed

Blanc et al. 2006 TAG 113:206–224



Methods for comparative mapping

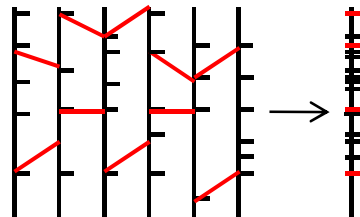
Consensus map and QTL meta-analysis based on position data

- From elaborated data (maps, position and QTL confidence intervals)
- Meta-analysis: combine data from different sources in a single study
- Widely used in medical and social science

Advantages: elaborated data retrieved from published data

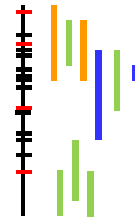
Steps for QTL meta-analysis

Construction of the consensus map

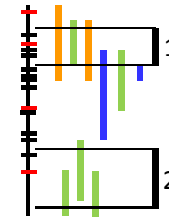


For each chromosome

Projection of QTLs on the consensus map

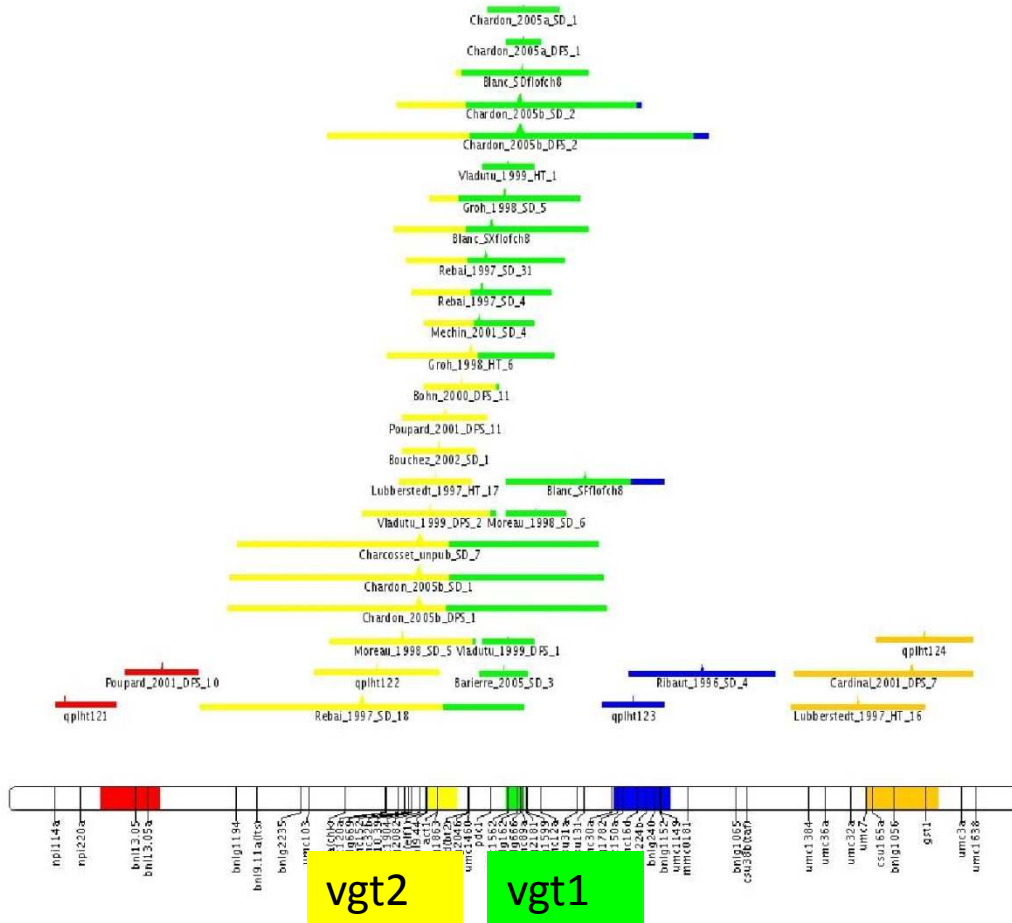


QTL meta-analysis



Example of a promising QTL meta-analysis

Simulations and analysis of real data sets showed that the **confidence intervals of the meta-QTLs are considerably reduced** compared to the observed QTL



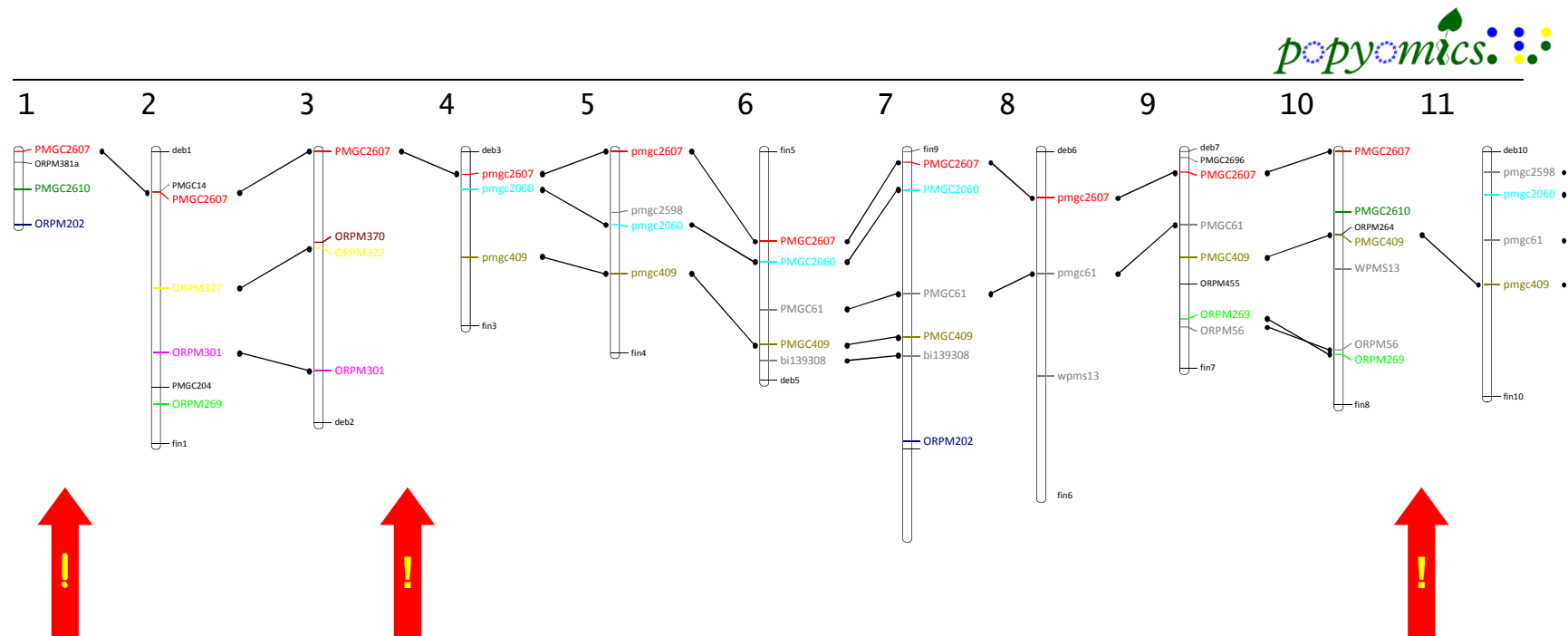
Maize example (Truntzler et al. submitted):

- 34 QTLs for early flowering components from 15 papers.
- The best model $k = 5$ meta-QTLs.
- Colors -> quantitative assignation of QTL -> 75% of observed QTLs correspond to vgt1 and/or vgt2.
- Confidence intervals reduced: vgt1 (~ 4cM) include a fine mapped QTL (Salvi et al., 2000).

Create a consensus map

In general, all genetic maps need to be **connected**.

Example in *Populus*: case of a LG well covered by SSR (LG_VIII)



Create a consensus map

Markers containing sequence and position information (genome):

1) SSR markers (http://www.ornl.gov/sci/ipgc/ssr_resources.htm):

4 166 primer pairs available.

2) SNP located in genes

- evenly distributed on the genome
- in specific region (QTL) involved in adaptive traits (Phenology, drought resistance ...)
- functional candidates (transcriptomics, ...).

Mapping data for *Salicaceae*

Genetic maps connected to the genome

8 different species involved in 13 intra and interspecific crosses:

(*Populus trichocarpa*, *P. deltoides*, *P. nigra*, *P. alba*, *P. angustifolia*, *P. fremontii*,
Salix viminalis, *S. schwerinii*) ...

Genus	Family name – Pedigree type	Progeny size	Reference(s)
Populus	Family 13 – BC1 TD x D	171	Yin et al 2004
	Family 331 - F2 TDxTD (POP1)	210	Tuskan et al.
	Family 545 – F1 TxD	101	Yin et al 2008
	Family XXX – D x DN	93	Yin et al 2008
	Family BCN - A x (FxA)	246	Yin et al 2008
	POP2 - DxT	330	Jorge et al. (2005, +in prep)*
	POP3a - DxN	180	Cervera et al.(2001)*
	POP3b - DxT	182	Cervera et al.(2001)*
	POP4 - AxA	141	Paolucci et al. (unpublished)*
	Family 52-124 - (TD)xD	396	Dorst et al 2009
Salix	S1 – [Vx(VS)] x V	463	Berlin et al 2010
	S3 - VxV	282	Berlin et al 2010
	K8 - [Vx(VS)] x [Vx(VS)]	471	Hanley et al 2006

Mapping data for *Salicaceae*

...but few framework genetic maps with genome anchoring markers
and QTL data (growth related traits)

Pedigree	Progeny size	Nb of anchoring markers	# of QTLs	Reference(s)
F2 TDxTD (POP1)	210	83	188, (37, 43)	Rae et al 2008, (2009)*
DxT (POP2)	330	74, 87	77	Jorge et al. (unpublished)
DxT, DxN (POP3)	180, 182	19, 14, 14, 16	21, 29, 48, 20	Dillen et al, 2009*
AxA (POP4)	141	55, 51	10, 11	Paolucci et al. (unpublished)*
(TxD)xD	396	178	63	Dorst et al 2009 Novaes et al. 2009

*Data from <http://www.evoltree.soton.ac.uk/cmap/>



Lack of transferability of markers between genetic maps

Populus spp.

	Novaes_2009	Rae_2009	Dillen_2009_D1	Dillen_2009_D2	Dillen_2009_N	Dillen_2009_T	Jorge_2010_D	Jorge_2010_T	Paolucci_2010_6K3	Paolucci_2010_14P11
Rae_2009	30									
Dillen_2009_D1	5	8								
Dillen_2009_D2	2	7	60							
Dillen_2009_N	4	5	4	5						
Dillen_2009_T	4	3	4	4	3					
Jorge_2010_D	19	28	10	7	8	7				
Jorge_2010_T	18	30	13	10	7	7	39			
Paolucci_2010_6K3	13	18	8	2	3	3	13	12		
Paolucci_2010_14P11	10	15	4	4	3	2	16	14	24	

Salix spp.

	Berlin_S1_2010	Berlin_S3_2010
Berlin_S3_2010	193	
Hanley_2006	48	37

Transferability
Pt genome -> maps

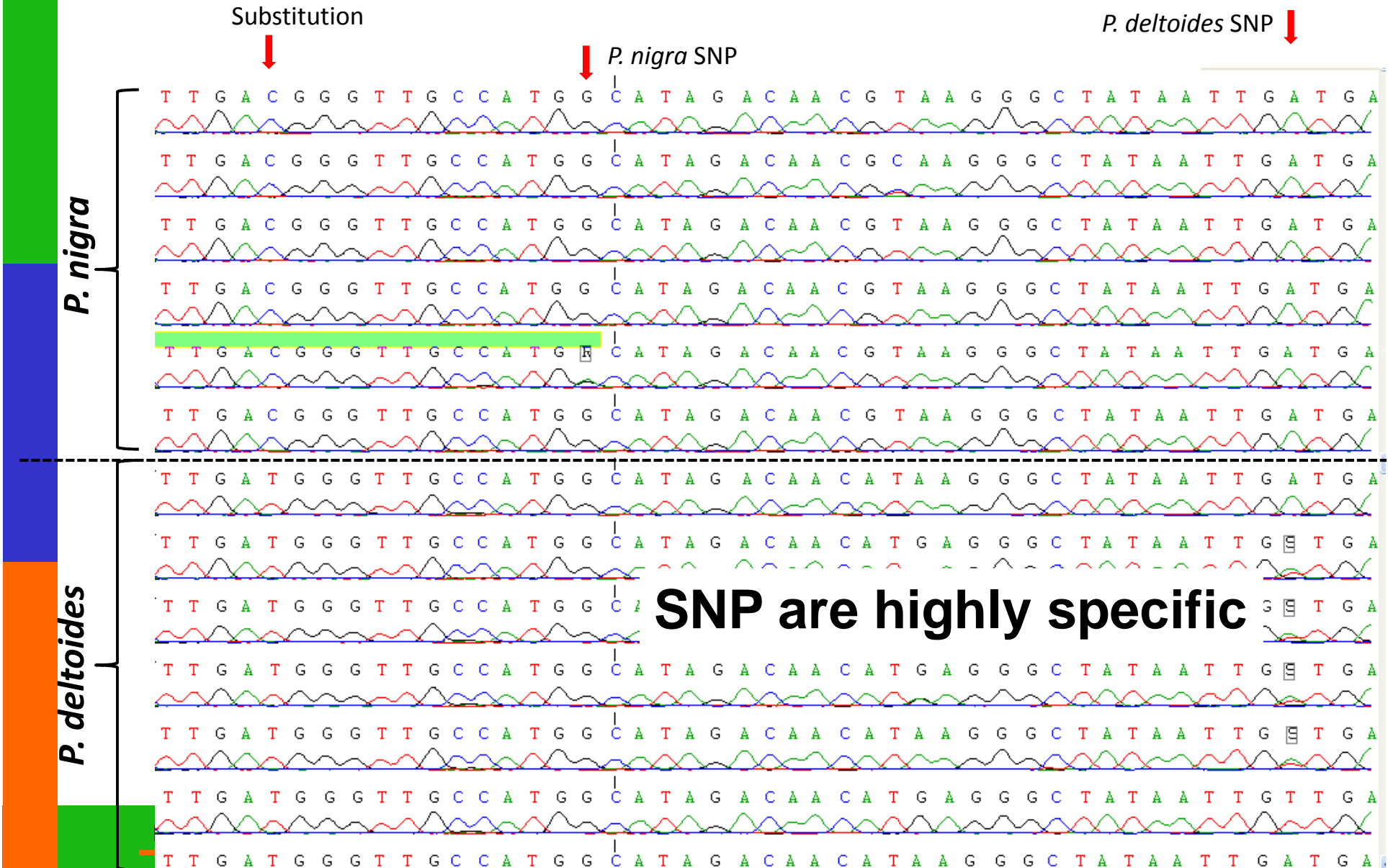
Populus spp. : on average 20% fewer for *P. alba* pedigree (distant taxus).

Salix spp.: 16%

Number of shared markers between genetic maps 10/05/2010

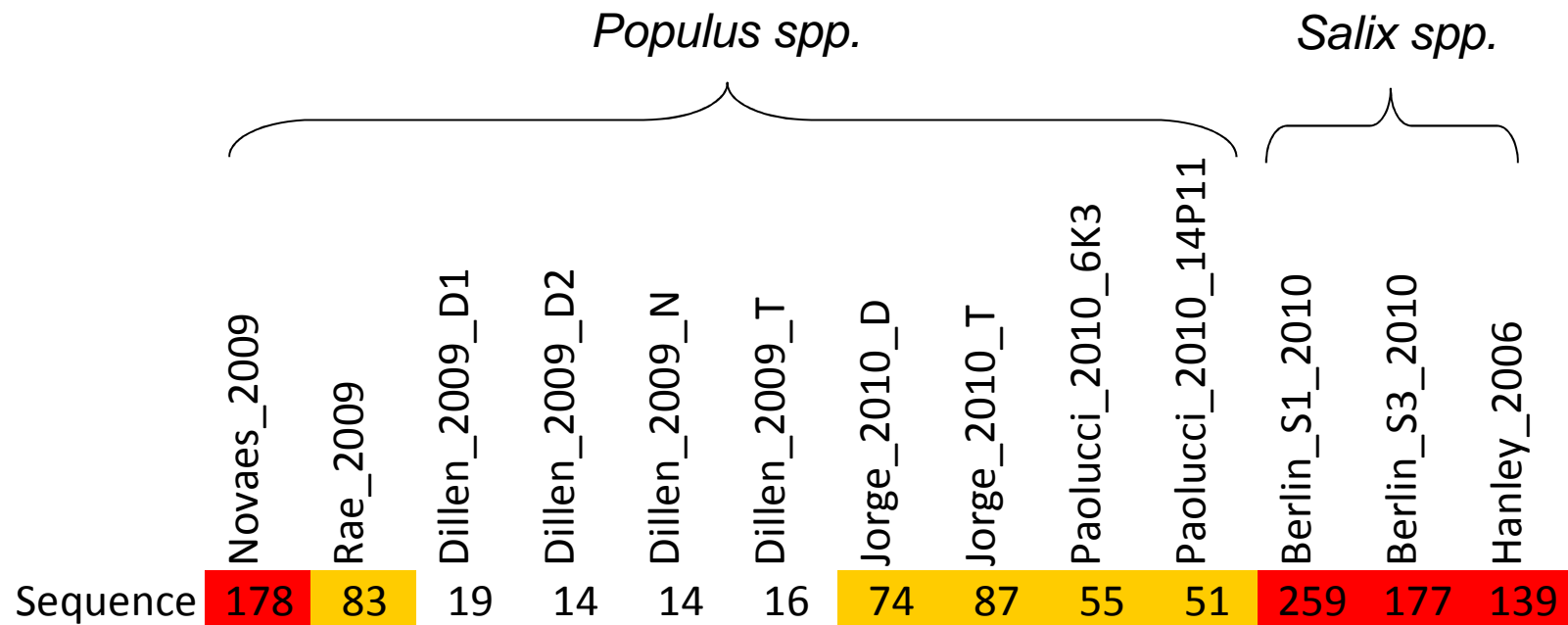


Lack of transferability of markers between genetic maps



Good anchoring on the genome sequence

Status of the alignment of Salicaceae genetic maps on the
Populus trichocarpa genome

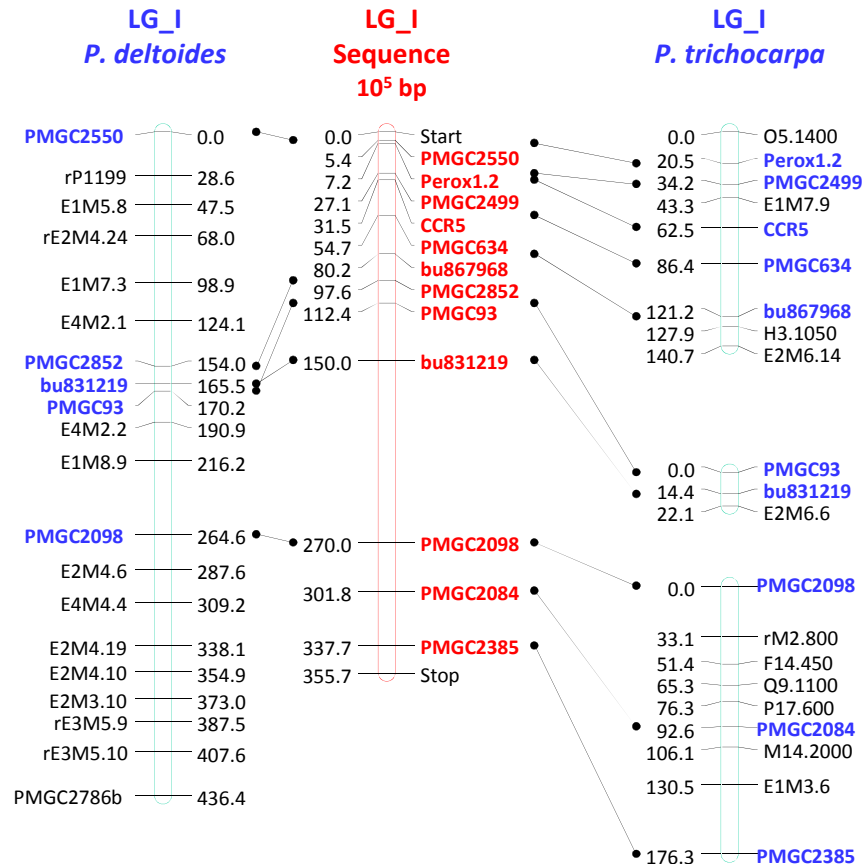


Number of anchoring markers between genetic maps
and the *P. trichocarpa* genome 10/05/2010

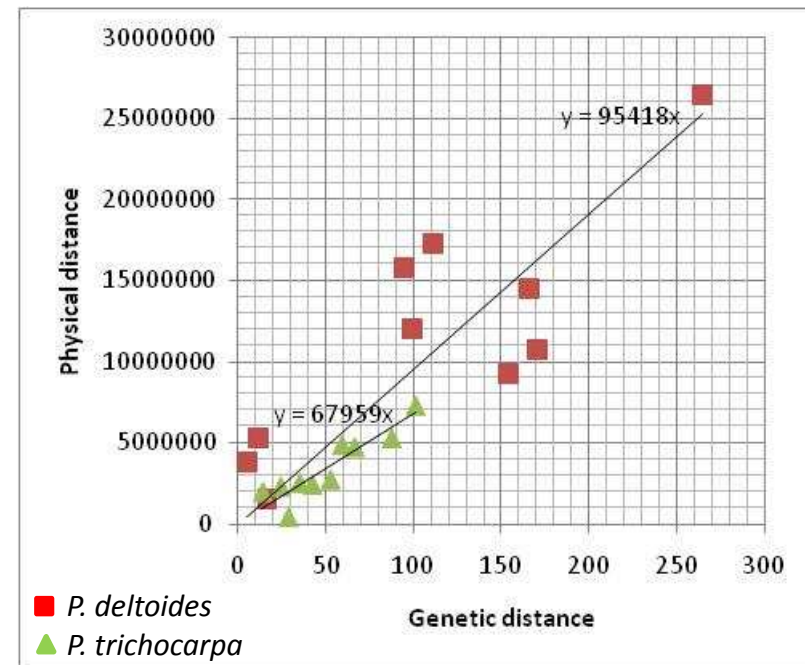


Strategy for comparative QTL mapping in *Salicaceae*

Anchoring to the genome sequence and bp/cM ratio calculation



Relationship between pair wise genetic and physical distances between genome anchored markers



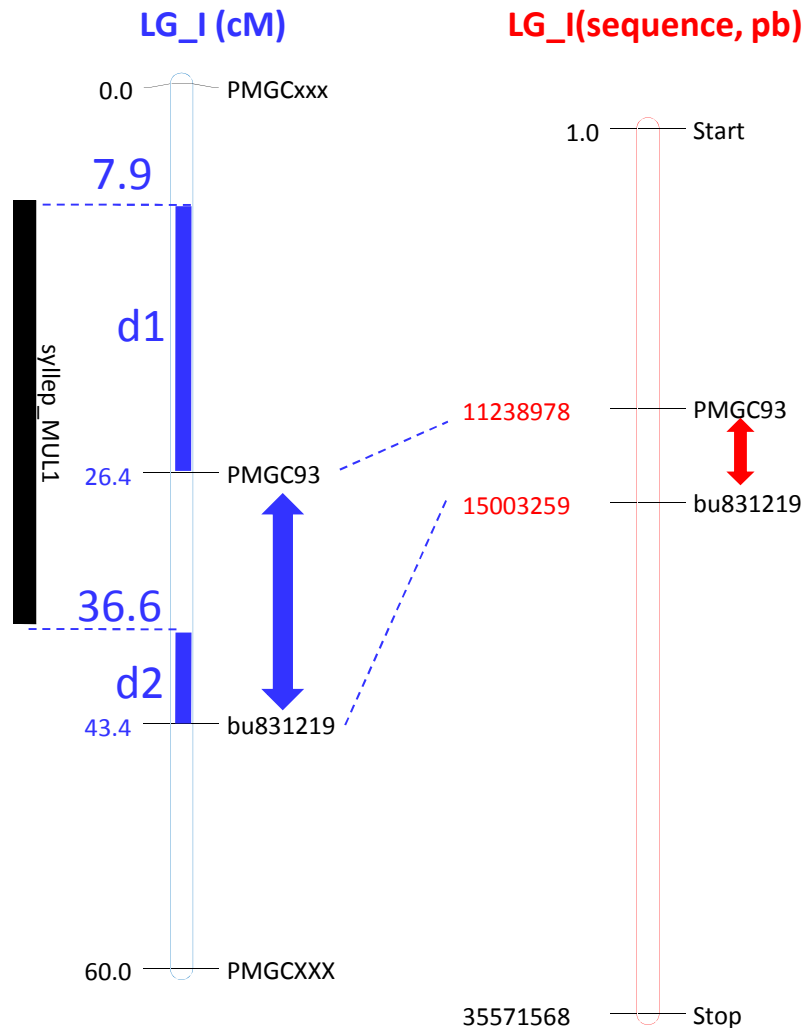
$$y = R x$$

Number of bp / cM



Strategy for comparative QTL mapping in *Salicaceae*

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



$$\text{Start(pb)} = 11238978 - [|26.4 - 7.9| \times R]$$

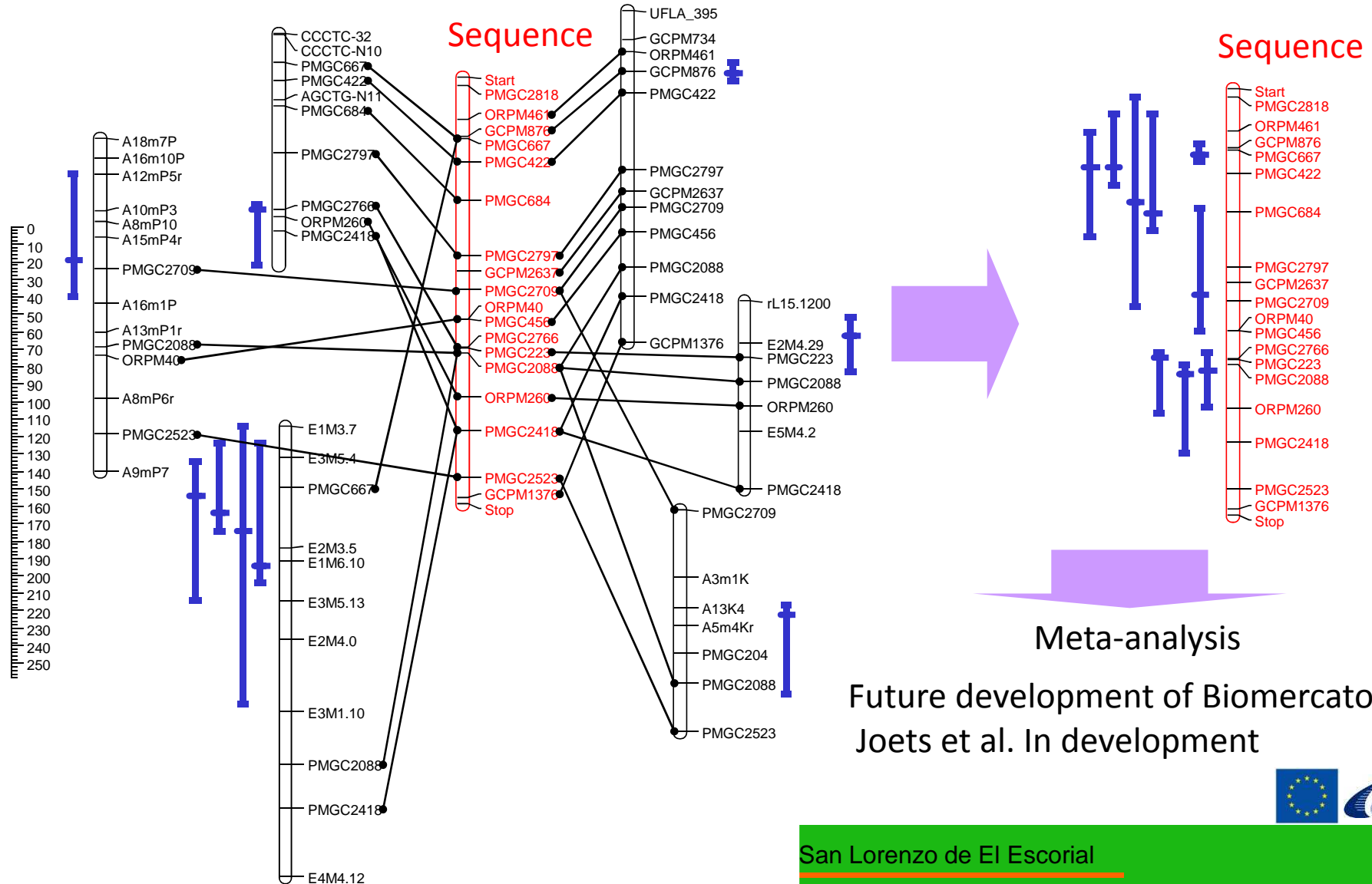
$$= 7\ 142\ 811\ \text{pb}$$

$$\text{Stop(pb)} = 15003259 - [|43.4 - 36.6| \times R]$$

$$= 13\ 497\ 405\ \text{pb}$$

Strategy for comparative QTL mapping in *Salicaceae*

QTL projection from each map to the genome sequence thanks to genome anchoring markers

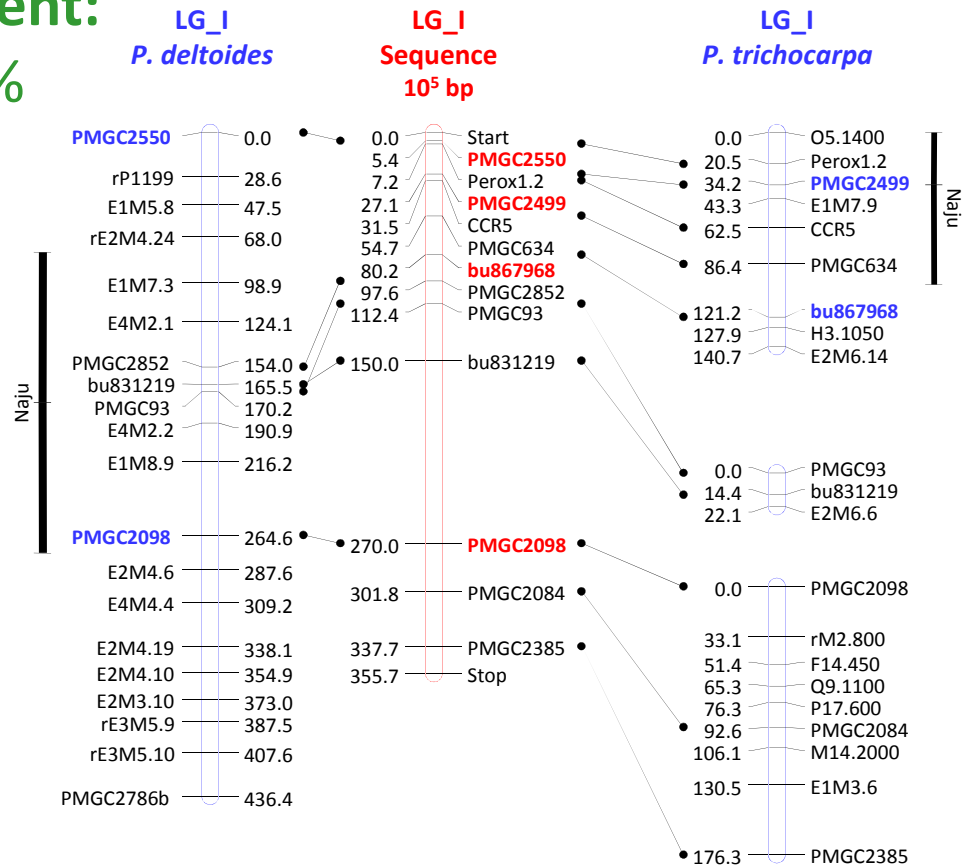


Example of physical QTL intervals calculation and corresponding number of genes

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

Leaf 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















Functional analysis of large gene lists

Functional analysis of large gene lists

Annotation use a classification based on Gene Ontology (GO)

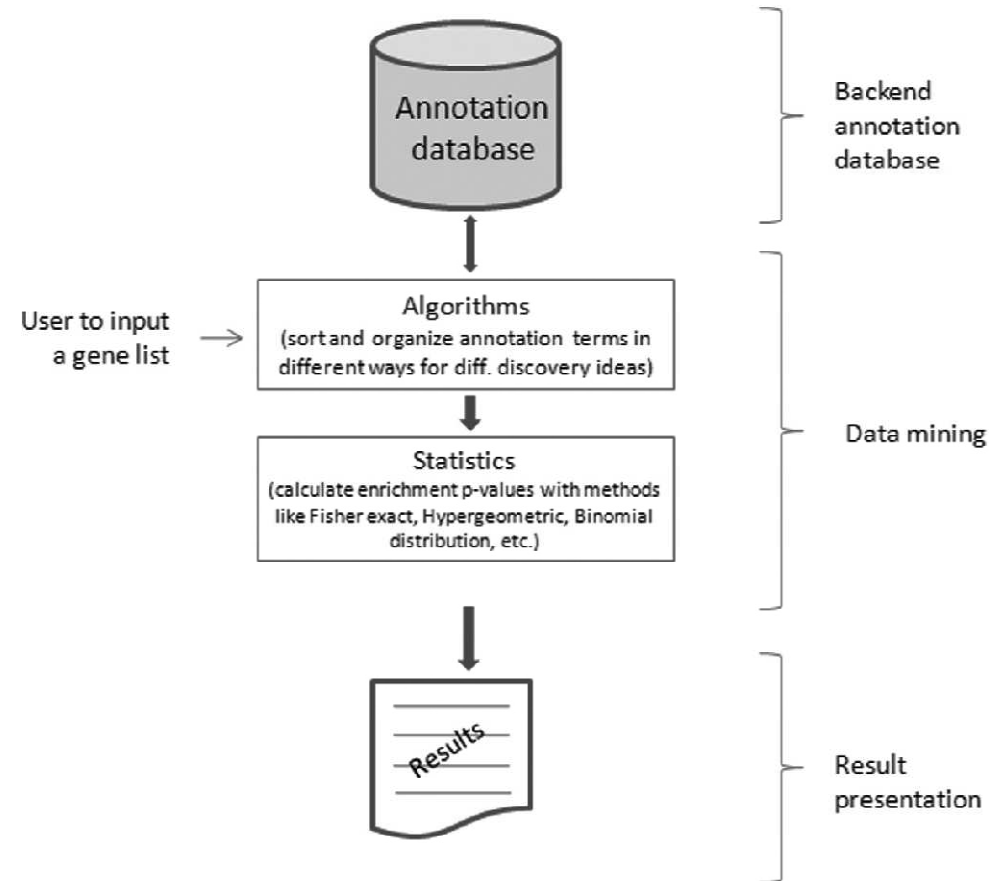
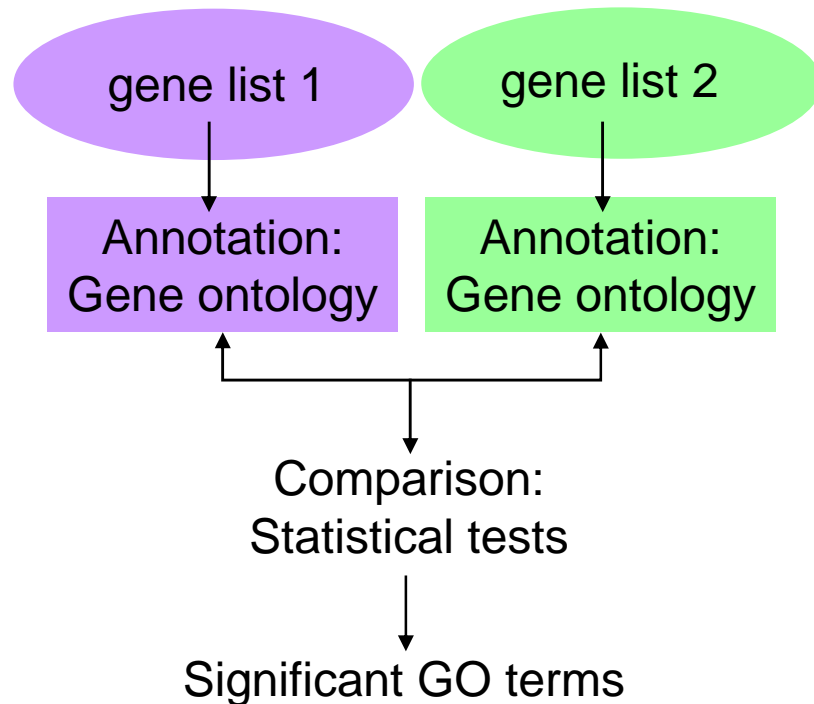
- Molecular function (ex: oxidoreductase activity, ...)
- Biological process (ex: photosynthesis light harvesting, ...)
- Cellular component (ex: membrane, chloroplast ...)

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)

Functional analysis of large gene lists

Methods: Bioinformatic Enrichment Tools



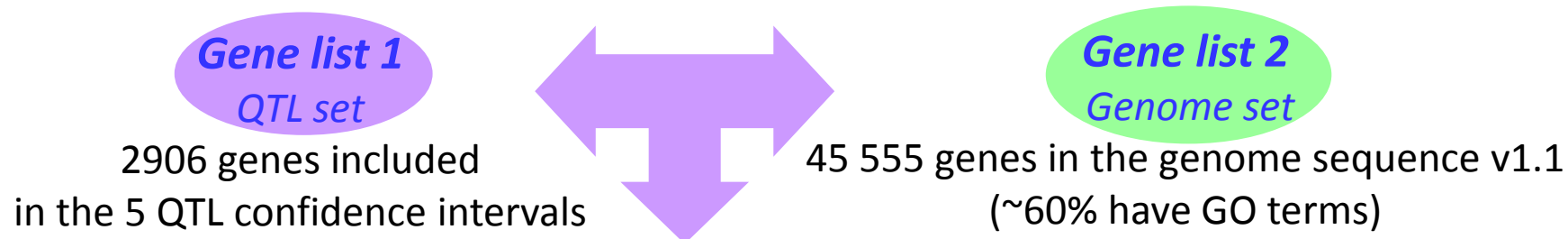
Strategy used in microarray experiment analysis

Bauer et al, 2008. Bioinformatics 24:1650-1651.

Huang et al, 2009. Nucleic Acids Research 37:1-13.

Functional analysis of large gene lists

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

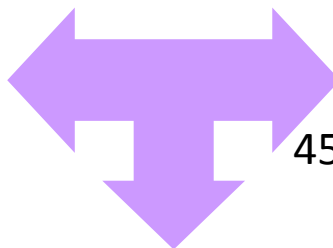


ID	Name	p-Value	p-Value (Adj)	Study Count	Population Count
GO:0010422	regulation of brassinosteroid biosynthetic process	0,0000	0,0000	5	6
GO:0010023	proanthocyanidin biosynthetic process	0,0000	0,0000	6	10
GO:0009641	shade avoidance	0,0000	0,0000	6	11
GO:0016595	glutamate binding	0,0001	0,0001	4	5
GO:0009741	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

Functional analysis of large gene lists

Gene list 1
QTL set

2337 genes included
in LG_X



Gene list 2
Genome set

45 555 genes in the genome sequence v1.1
(~60% have GO terms)

ID	Name	P value	P value (Adj.)	Study count	Pop. count
GO:0008825	cyclopropane-fatty-acyl-phospholipid synthase activity	5,88E-05	5,88E-05	5	9
GO:0005575	cellular_component	9,97E-05	9,97E-05	1258	21116
GO:0000287	magnesium ion binding	3,06E-04	3,06E-04	4	7
GO:0008150	biological_process	7,95E-04	7,95E-04	1433	24560
GO:0009536	plastid	0,001270407	0,001270407	232	3413
GO:0009507	chloroplast	0,001524634	0,001524634	226	3327
GO:0044444	cytoplasmic part	0,001912735	0,001912735	400	6246
GO:0044275	cellular carbohydrate catabolic process	0,002304709	0,002304709	19	164
GO:0005982	starch metabolic process	0,002975107	0,002975107	7	35
GO:0044265	cellular macromolecule catabolic process	0,003129754	0,003129754	38	422

Future developments

Combining meta QTL analysis and gene enrichment approaches

- Biomecator3: projection of QTLs on the genome and QTL meta-analysis
- Gene list analysis: tools crossing information from GO + bibliography + biochemical pathways + various omics ...



Thank you!