

Comparative genetics of quantitative seed quality traits in *Brassica napus*



Rod Snowdon

Department of Plant Breeding
Justus Liebig University
Giessen, Germany



Comparative genomics

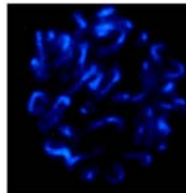
Use of comparative data from a “model” genome to identify genes responsible for important traits in another species



Arabidopsis
Genome size 120 Mbp
2n=10

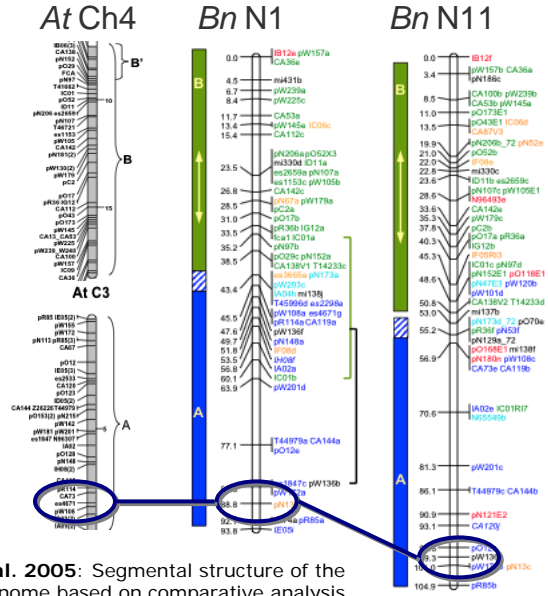


Brassica napus
Genome size
1200 Mbp
2n=38



Genome alignment At – Bn

Information on chromosome synteny between *Arabidopsis* and *Brassica* can potentially help us to find key genes underlying quantitative traits in *B. napus*



Parkin et al. 2005: Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171: 765-781

Integrated genome browsers
e.g. The Arabidopsis Integrated Database (www.atidb.org)

Arabidopsis chromosome region

Arabidopsis genes, transcripts and genomic clone sequences

Brassica ESTs & shotgun-sequences

Homologous Brassica BAC clones

The screenshot shows the ATIDB interface for a specific Arabidopsis chromosome region. It displays a genomic map with various annotations: Arabidopsis genes, transcripts, and genomic clone sequences (top); Brassica ESTs and shotgun-sequences (middle); and Homologous Brassica BAC clones (bottom). A search bar and navigation controls are visible at the top.

Exploiting synteny: Examples

1. Potentially gene-linked SSR markers for allele-trait association studies:

Identification of markers and key genes associated with seed glucosinolate content



2. QTL dissection and *Arabidopsis* resources for identification of candidate genes for a “phenotypically difficult” trait:

Seed colour and fibre content



Seed glucosinolate content

Around 30 genes are known to be involved in glucosinolate biosynthesis in *Arabidopsis* leaves and/or seeds. This corresponds to potentially hundreds of gene copies in the polyploid oilseed rape genome

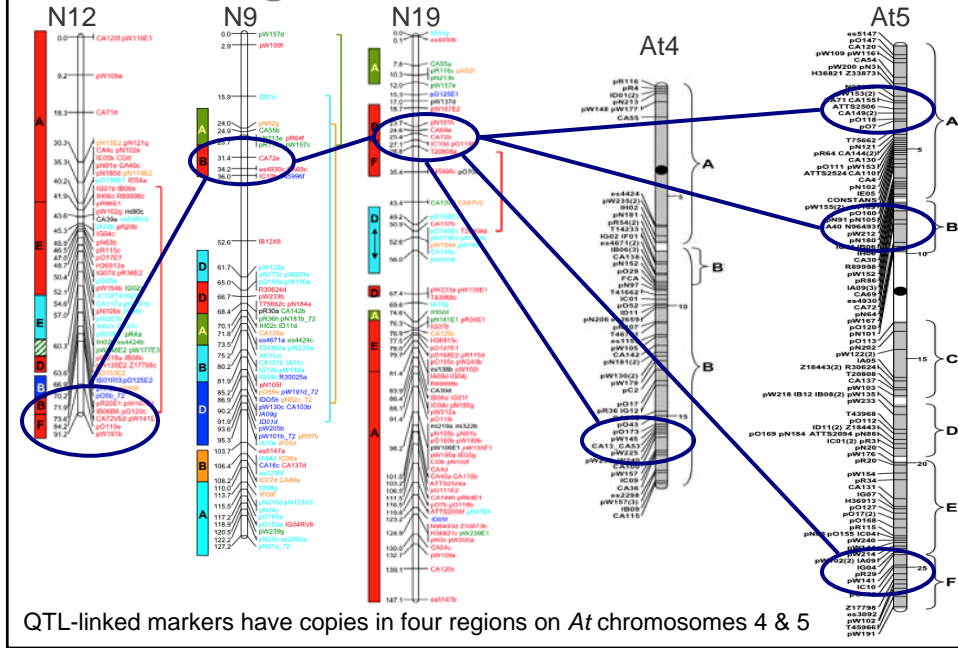
Which of these genes control seed GSL biosynthesis in *B. napus*?

What are the key genes that regulate the biosynthesis of aliphatic and indole GSLs?

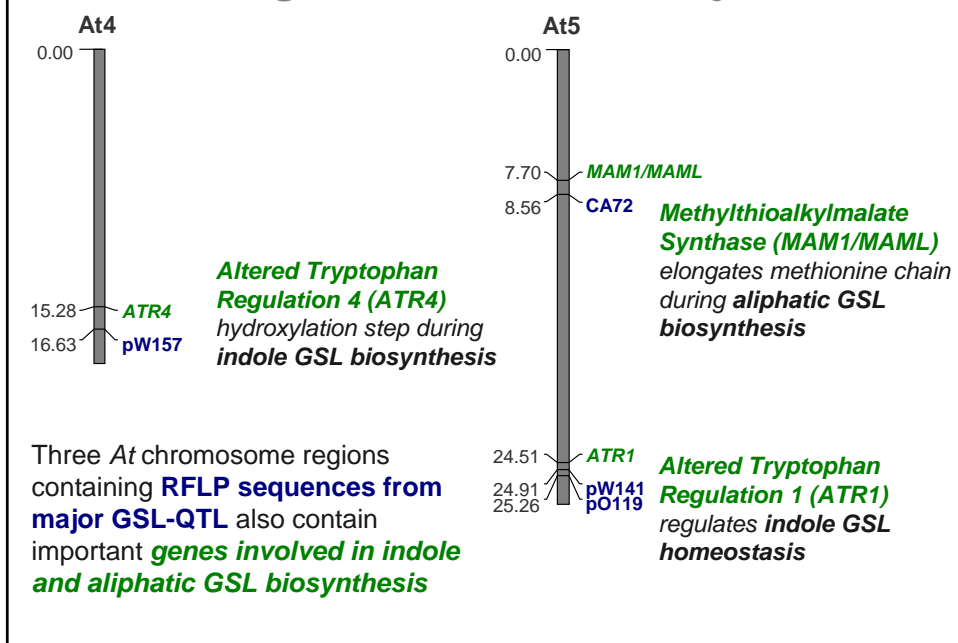
Can we find new allelic variation for different copies of these genes?



Homoeologous BnQTL for GSL content



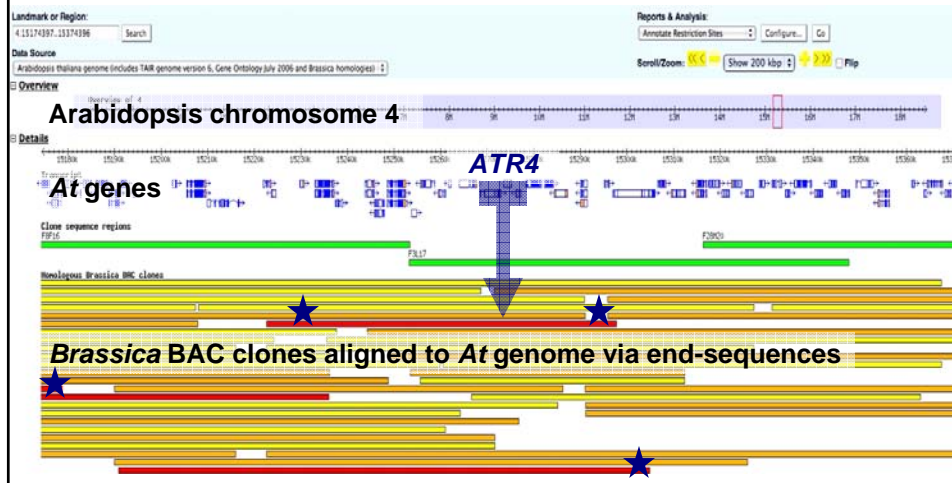
Candidate genes for GSL biosynthesis?



“Potentially gene-linked SSR markers”

Arabidopsis-Brassica integrated database

Search for SSR motifs (★) in *Brassica* sequences with high homology to regions containing *At* candidate genes



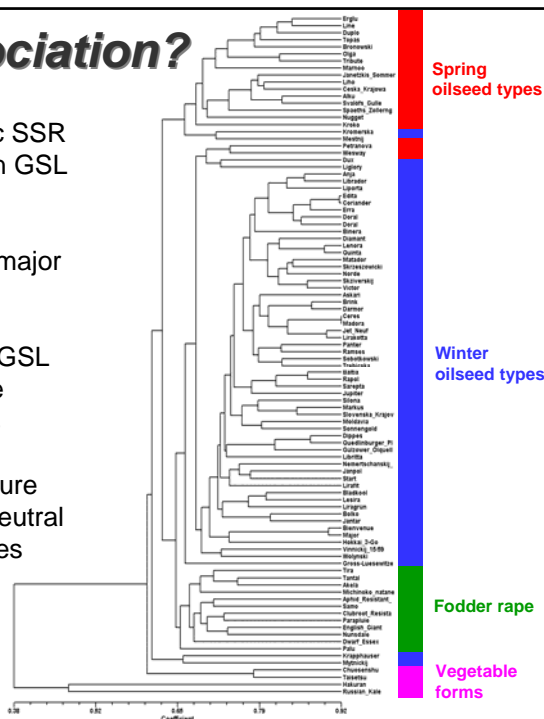
Marker-trait association?

12 well-functioning, polymorphic SSR primer combinations identified in GSL candidate gene regions

4 additional SSR markers from major QTL for seed GSL content

Tests for association with seed GSL content in 96 genetically diverse *B. napus* gene bank accessions

Adjustment for population structure using genome-wide data from neutral SSRs to control for false positives

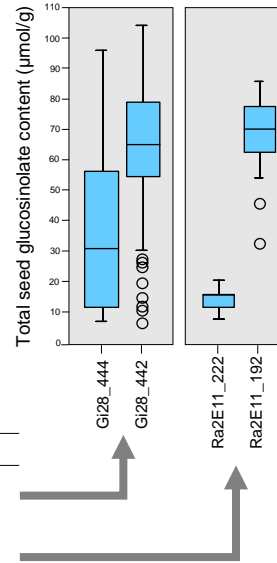


Trait association of gene-linked SSRs

For all of the candidate genes investigated we found flanking SSR sequences with significant associations to seed total GSL content

All of the QTL-linked SSR markers also showed significant trait associations in the gene bank accessions

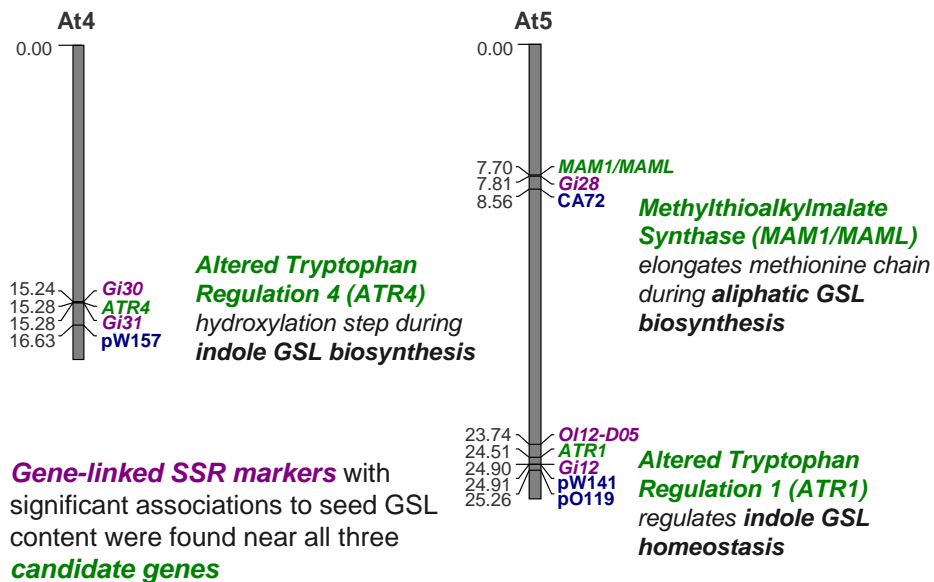
In most cases different alleles showed significant associations with increased and decreased GSL content, respectively



Examples

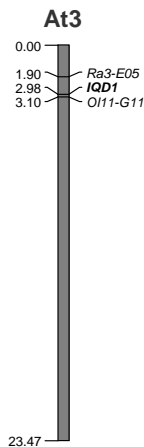
Marker	Frequency	Mean GSL (µmol/g)	P-value
Gi28_442	0.80	70.4	0.000
Gi28_444	0.18	37.7	0.000
Ra2E11_192	0.33	81.2	0.000
Ra2E11_222	0.53	17.8	0.006

Trait association of gene-linked SSRs



Identification of a new candidate gene

Two genome-wide SSR markers that were also significantly associated to seed GSL content in the gene bank accessions were found to be closely linked in Arabidopsis to another interesting candidate gene:



IQ-DOMAIN 1 (IQD1)
Transcription factor that
**modulates numerous
GSL pathway genes**

**Loss-of-function *IQD1* mutants in Arabidopsis
show a reduction in total seed GSL content**

New allelic diversity for GSL reduction?

All SSRs with significant associations to GSL content in the gene bank accessions were retested in a genetically diverse set of 46 old and new European winter rapeseed varieties

- Most GSL-associated markers were present at only very low frequency (<5%)
- Only three of the markers were also associated with GSL content in modern varieties
- Some of the markers are localised in *B. napus* genome regions that do not contain known QTL for GSL content

Conclusion: High-GSL *B. napus* gene bank accessions contain considerable untapped allelic variation for low seed GSL content

- Gene-linked markers may help to develop new low-GSL gene pools
- Further reduction of GSL in current 00 material may be possible by marker-assisted replacement of high-GSL associated alleles at individual loci

Seed colour and fibre content

The thin seed coat of yellow-seeded *B. napus* is associated with lower levels of antinutritive fibre and tannins in the seed meal

The yellow-seed trait can be extremely sensitive to the environment, so is difficult to use as a selection marker for fibre reduction



Development of gene-linked markers is an important aim for breeding of low-fibre, low-tannin varieties

QuickTime™ and a
H.264 codec
are required to view this picture.

Factors contributing to seed colour



One major QTL explaining >50% of the variation in seed colour and lignin content

Further QTL with additional epistatic effects on seed pigmentation

Plus possible epigenetic regulation?

Major factor: Seed coat thickness (thickness of palisade layer cell walls), which also has a strong influence on the antinutritive **lignin** content

Contributing factors: Dark-coloured **phenolic compounds** in the seed coat, particularly oxidised condensed tannins (proanthocyanidins)

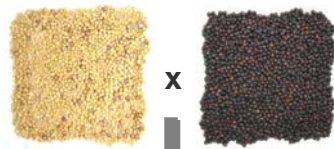
Transparent testa candidate genes



The thin seed coat of yellow-seeded *B. napus* is similar to the phenotype of many **transparent testa** mutants in *Arabidopsis*



Comparative mapping of major QTL

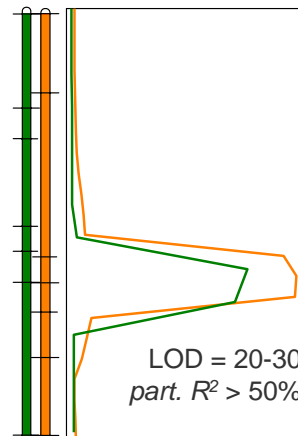


Large DH populations

Phenotype dissection

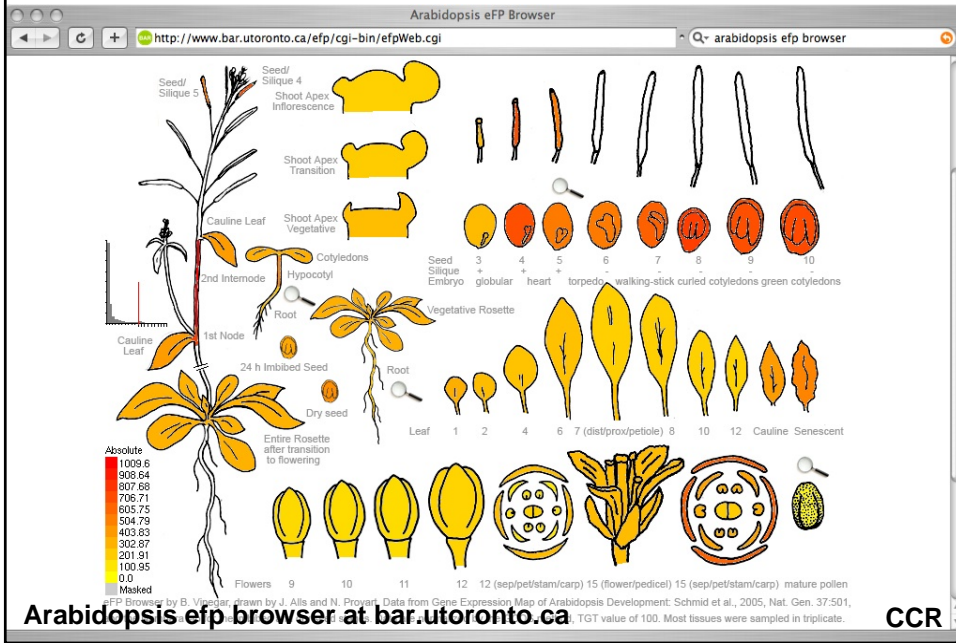
- Seed colour
- Fibre components
- Phenolic acids
- Flavonoids
- Germination
- Seedling development
- Hormone profiles

A major QTL for seed colour was co-localised in genetic maps from two different YS sources

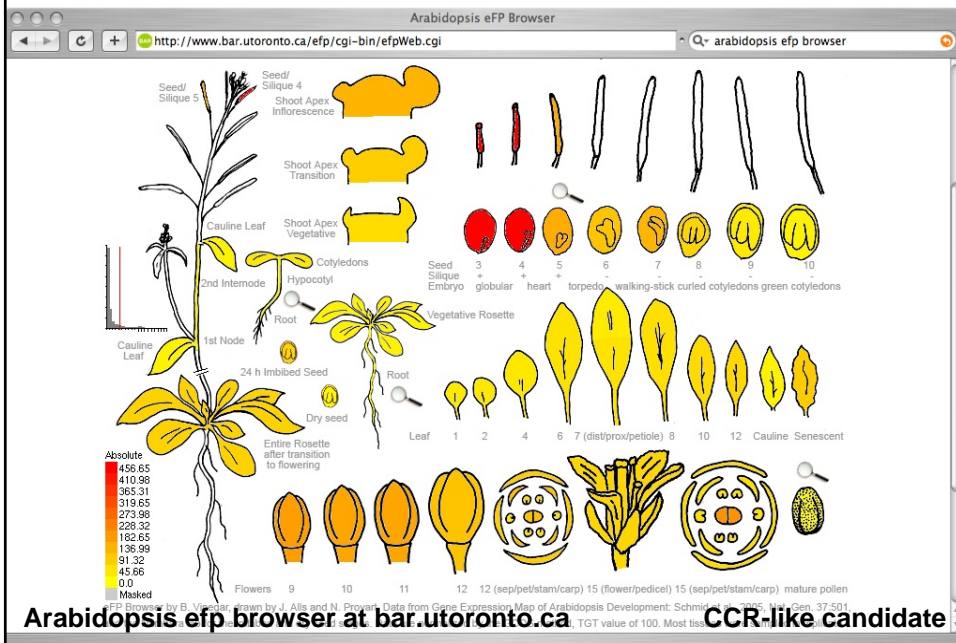


QTL-flanking markers sequenced and aligned to Arabidopsis

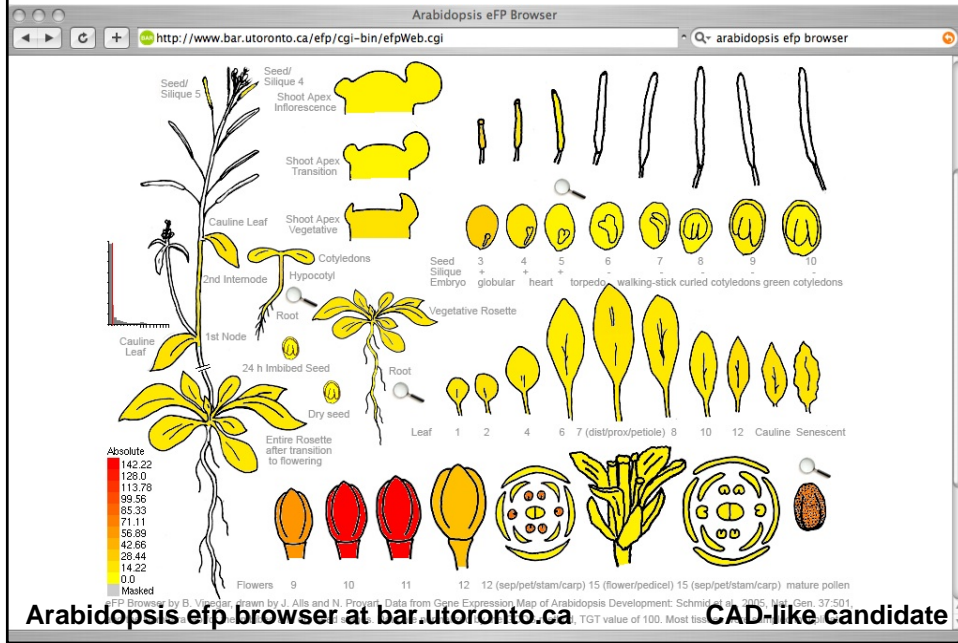
Tissue-specific expression?



Tissue-specific expression?

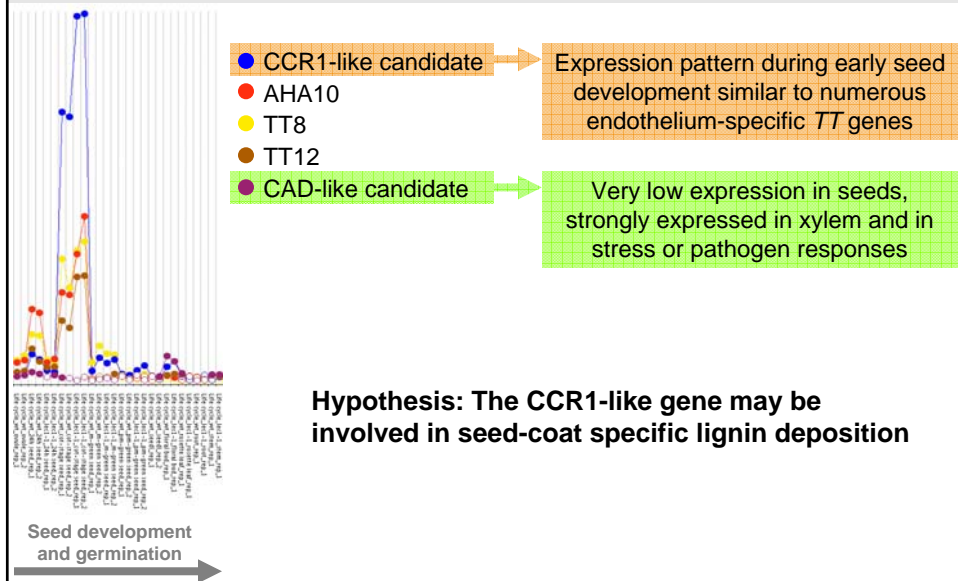


Tissue-specific expression?



Seed lignin candidates?

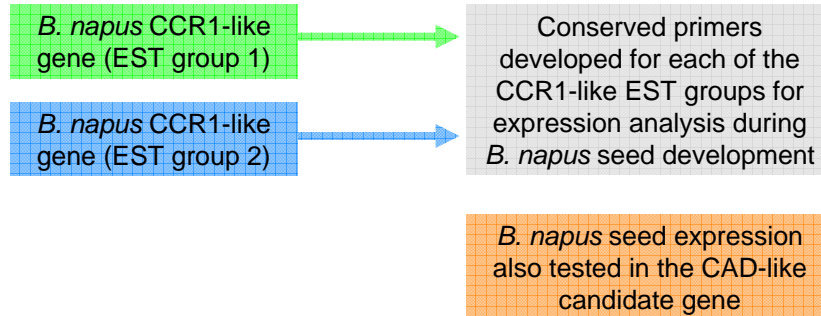
Expression patterns in Arabidopsis seeds (data from Genvestigator)



Expression in *B. napus*

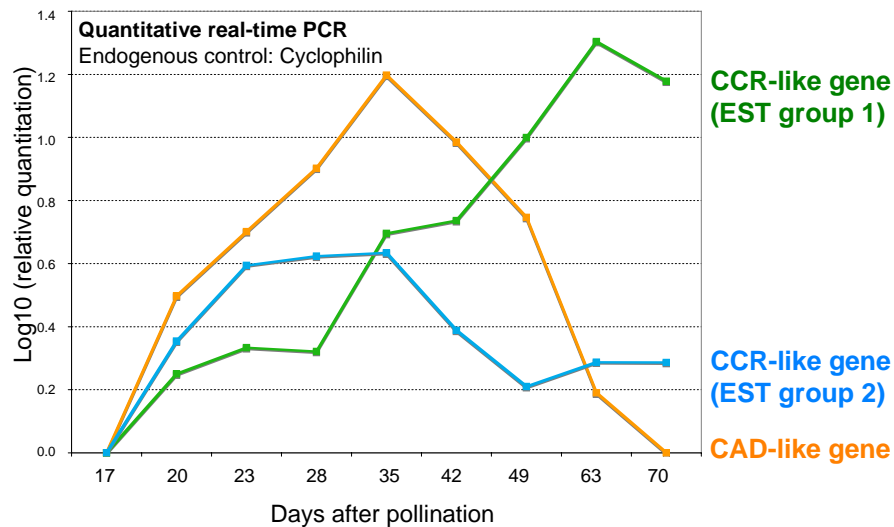
Sequence analysis of *B. napus* ESTs (particularly seed ESTs)

- The CCR1-like candidate is quite divergent from the real CCR1, but is still annotated to lignin biosynthesis
- *B. napus* ESTs with high homology to the CCR1-like candidate gene cluster into two distinct groups: Two similar homologs? Or two loci of same gene?



Expression in *Brassica napus*

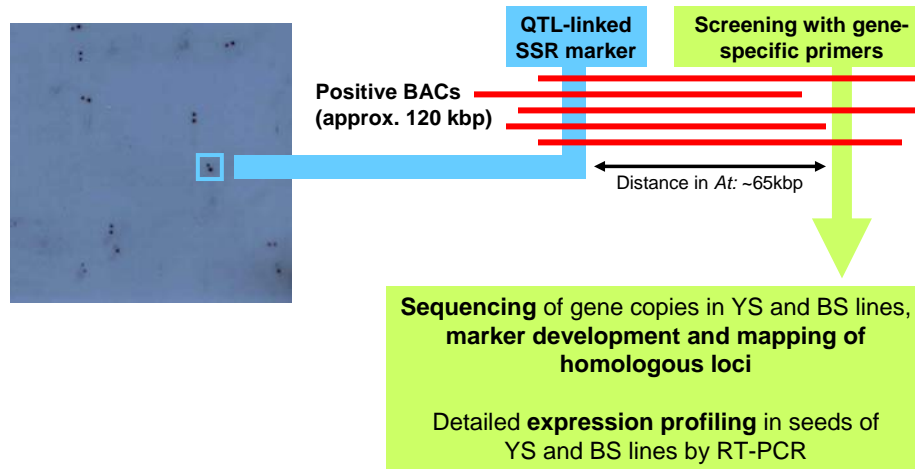
Gene expression in developing *B. napus* seeds (cv. Express)



Candidate validation

Isolation and characterisation of *B. napus* candidate gene copies

- Isolation of BACs containing the QTL-linked SSR marker together with copies of the *CCR1*-like candidate gene:



Summary and Outlook

- *Arabidopsis-Brassica* synteny tools enable development of useful gene-linked SSR markers for *B. napus* QTL, and identification of genes underlying quantitative traits
- As new-generation sequencing technologies become established, huge advances can be expected in the availability of *Brassica* sequence data and SNPs
- Microsyntenic alignment of high-density SNP maps and physical maps to the *Arabidopsis* genome will vastly increase the power of comparative analyses in *B. napus*
- Phenotype dissection into constituent traits can simplify the search for relevant candidate genes
- Epigenetic modifications and regulation of quantitative traits are still poorly understood and represent one of the major challenges with regard to understanding and using quantitative variation



