

## RESEARCH ARTICLE

## Open Access

# Organ specificity and transcriptional control of metabolic routes revealed by expression QTL profiling of source–sink tissues in a segregating potato population

Bjorn Kloosterman<sup>1,4</sup>, AM Anithakumari<sup>1,2</sup>, Pierre-Yves Chibon<sup>1,2</sup>, Marian Oortwijn<sup>1</sup>, Gerard C van der Linden<sup>1</sup>, Richard GF Visser<sup>1,3</sup> and Christian WB Bachem<sup>1\*</sup>

## Abstract

**Background:** With the completion of genome sequences belonging to some of the major crop plants, new challenges arise to utilize this data for crop improvement and increased food security. The field of genetical genomics has the potential to identify genes displaying heritable differential expression associated to important phenotypic traits. Here we describe the identification of expression QTLs (eQTLs) in two different potato tissues of a segregating potato population and query the potato genome sequence to differentiate between cis- and trans-acting eQTLs in relation to gene subfunctionalization.

**Results:** Leaf and tuber samples were analysed and screened for the presence of conserved and tissue dependent eQTLs. Expression QTLs present in both tissues are predominantly cis-acting whilst for tissue specific QTLs, the percentage of trans-acting QTLs increases. Tissue dependent eQTLs were assigned to functional classes and visualized in metabolic pathways. We identified a potential regulatory network on chromosome 10 involving genes crucial for maintaining circadian rhythms and controlling clock output genes. In addition, we show that the type of genetic material screened and sampling strategy applied, can have a high impact on the output of genetical genomics studies.

**Conclusions:** Identification of tissue dependent regulatory networks based on mapped differential expression not only gives us insight in tissue dependent gene subfunctionalization but brings new insights into key biological processes and delivers targets for future haplotyping and genetic marker development.

## Background

The field of associative genomics or genetical genomics attempts to combine the heritability of generated ~ omics data with phenotypic variation through genetic marker associations [1]. The aimed outcome of a genetical genomics study is to define genomic regions that control the expression of single or multiple genes, metabolites and/or proteins (eQTLs, mQTLs and pQTLs, respectively). The amount of population wide ~ omics data for non-model species has increased dramatically over the last few years as costs for data generation have

decreased and computational bottlenecks have been largely overcome. Successful application of genetical genomics has been demonstrated for a number of plant species including *Arabidopsis*, barley, wheat, eucalyptus and poplar [2-8]. In particular, studies with the model plant *Arabidopsis* have been crucial for the development of the concept of genetical genomics and have pushed software development forward to cope with the increasing amount of data generated by the different profiling platforms available [1,9]. Associations between ~ omics data, markers and phenotypes have highlighted the overall complexity of plant trait diversity and have led to the reconstruction of a number of regulatory networks including flowering [4], glucosinolates [10], flavonoids [11] and carotenoids [12]. The identification of eQTL

\* Correspondence: [christian.bachem@wur.nl](mailto:christian.bachem@wur.nl)

<sup>1</sup>Wageningen UR Plant Breeding, Wageningen University and Research Center, PO Box 386, 6700 AJ Wageningen, the Netherlands  
Full list of author information is available at the end of the article

hotspots may indicate the presence of major regulatory switches controlling the expression of many genes directly or indirectly [13]. The existence of eQTL hotspots was confirmed in a number of studies where an enrichment of gene functional categories was found for several of the identified regions [14,15]. However, network modelling still heavily relies on a-prior knowledge of the pathways targeted [16].

Despite the high potential of genetical genomics studies in plants, the association of differential gene expression or metabolite concentration with the phenotypic variation does not always disclose the responsible genetic factors, due to genetic linkage of additional genes with the genomic region. Identification of causative polymorphisms is important for the development of applicable genetic markers. For genetical genomics experiments, factors like population size, marker density, sampling strategy, environmental factors and screening platform used, greatly influence the outcome and the ability to detect small additive biological effects [reviewed in [17,18]]. The high costs for population-wide screening often limit genetical genomics experiments to a single tissue or developmental stage. Several studies indicate that the genetic architecture of gene expression is highly divergent between organs and/or developmental stages which will be reflected in eQTL tissue specificity [5,19-21]. In poplar it was found that less than one-third of genes with eQTLs have co-localizing eQTLs when comparing two different organs [5]. Gene duplication events and diversification through mutations, referred to as sub- and neofunctionalization, underlie gene expression differences and are hypothesized to be the driving forces for obtaining novel gene functions and consequently evolution of plant development [22,23].

Potato cultivars are highly heterozygous tetraploid outcrossing plants, which complicates genetic analysis aimed at understanding the molecular mechanisms underlying trait variation. Genetic studies therefore are often performed in diploid genotypes and populations. Precision breeding relies on the identification of causative polymorphisms in genes responsible for trait variation. Once a gene has been assigned a regulatory function for a targeted trait, the allelic diversity of the gene in tetraploid cultivars and wild type accessions can be determined and potentially exploited in marker assisted breeding [24].

The genome sequence of a dihaploid potato clone (DM) has recently been completed [25] and this will facilitate the rapid identification and cloning of genes associated to trait variation and thus haplotyping efforts. Genetical genomics experiments in potato opens up a new dimension with the availability of the genome sequence. With the physical as well as the genetic map

location of many genes known, QTL intervals can be more easily screened for candidate genes. Here we present the first genetical genomics study in potato with eQTL analysis in both leaves and tuber. A comparison between leaf and tuber tissues reveals conservation as well as tissue dependent variation of gene expression. Gene interactions and potential regulatory networks are identified and discussed.

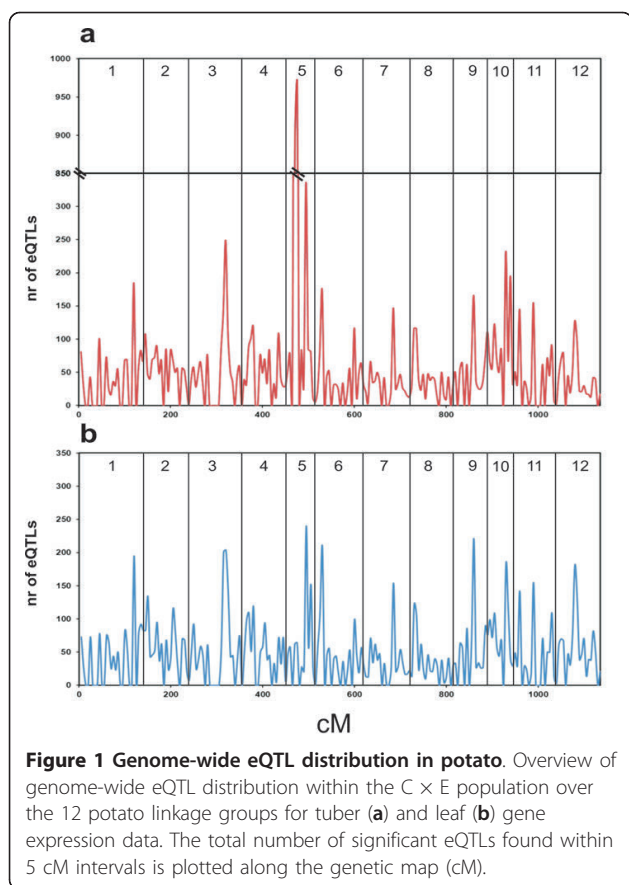
## Results and discussion

### QTL analysis and distribution

Expression profiles of leaf and tuber RNA extracts were obtained from two independent experiments. For tuber profiling, each sample was assayed twice in a two-colour dye experiment, allowing cross-validation of the observed variation and subsequent filtering based on correlation scores (Methods). The total number of array features exhibiting significant expression that could be used for QTL mapping was higher for leaf (22,193) than tuber samples (19,956). A wide overlap of expressed genes (19,590 features) was found between both organs, which was somewhat surprising considering their highly contrasting functions in terms of architecture and morphology, environmental exposure and developmental stage. However, the organs harvested consist of a multitude of different cell types thereby reducing the overall transcriptomic complexity and no distinction between basal or high gene expression was made. Similarly, the potato microarray is based on available EST sequences where temporal and spatially restricted transcripts have a lower chance of being captured and thus not profiled when using the microarray.

QTL analysis was performed based on the R/qtl package that allows high-throughput QTL mapping in an outbreeding species using a single integrated genetic map (Methods). A total of 17,764 QTLs were identified as significant for leaf and tuber data combined, corrected for gene redundancy present on the array where possible (i.e. similar blast scores for independent array features against the potato genome gene prediction sequences v3.2).

The distribution of QTLs over the different linkage groups is presented as a sliding window graph using 5 cM stepping size for tuber and leaf (Figures 1a and 1b, respectively). The number of significant QTLs for each linkage group is provided in Additional file 1. A clear over-representation of the number of QTLs for tuber expression was detected on linkage group 5 with over 1800 QTLs mapped in just two neighbouring bins (20-30 cM). For leaf expression data, the eQTLs appear to be more evenly distributed over the genome, although several genetic intervals can be identified in which the number of QTLs is high (> 200 on chromosomes 3, 5, 6, and 9). When comparing leaf and tuber eQTL



distribution overlapping regions with high eQTL numbers can be identified. These regions may indicate regions of high gene density or cover large physical sequence distances as a result of low recombination frequencies. Not all markers used in this study can be linked to the genomic scaffolds and not all scaffolds have been anchored onto the physical map, which needs to be resolved before answering these questions.

For the majority of features on the array (12,331; 70.1%), only a single QTL was detected while for 2,503 features (29.9%), two or more QTLs were found for both tissues combined (Table 1a). For two features on the array, a maximum of 5 QTLs were identified. One shows strong homology to a *tropinone reductase II* (TRII; EST *cSTB29M16TH*) [26] involved in the biosynthesis of tropane alkaloids which is known to be expressed in both below- and above-ground potato tissues. The second feature (BF\_LBCHXXX\_0032H04\_T3M.SCF) with 5 different QTLs has no significant homology to any protein present in the databases nor was its gene structure predicted in the current release of the potato genome annotation [25]. For both genes, four QTLs were detected in leaves and a fifth non-overlapping QTL in tubers. In a study from West et al. [2], the number of regions reported to control the expression of a gene in

*Arabidopsis* ranges from 0 to 11 with an average of 2.34 transcript per gene. In a barley study [3], over a third of regulated genes had only one identified eQTL while the rest of the genes showed 2-6 eQTLs. In our dataset we found relative few genes that were controlled by more than 2 genomic regions (2503 features) with a calculated average of 1.19 QTL per gene. Estimations of the average number of genomic regions (QTLs) controlling the expression of a single gene varies depending on the algorithm used, the population type and size [3,27]. The number of eQTLs per gene identified in this study is likely to be an underestimation, as genes with small additive effects may remain undetected due to the relative small population size ( $n = 94$ ) and marker density resulting in a high LOD threshold (4.35). Moreover, here we have targeted mature tubers and open leaves, to allow analysis of source-sink relationships, while more temporal and spatially restricted gene expression can be found during organ initiation and in young developing tissues.

#### Cis vs. Trans-acting eQTLs

One of the most exciting outputs from large-scale eQTL studies in species with a full genome sequence available is the ability to distinguish cis- and trans-acting regulation of gene expression, which holds the potential to reveal regulatory networks. With the availability of the potato draft genome sequence we were able to distinguish cis- and trans-acting QTLs for most genes. EST based unigenes representing the features on the array were blasted against the genome scaffolds as described in Methods. Putative chromosome locations could be assigned to the vast majority of genes (92%). Genes that could not be assigned a chromosome location were either similar to multiple regions on the genome, had similarity scores below the threshold, or represented significant allelic diversity.

Identified QTLs on the same linkage group as their physical map position are identified as cis-acting while QTLs on different linkage groups are defined as trans-acting. Our results show that almost twice as many cis-eQTLs were identified in comparison to trans-acting eQTLs for leaf and tuber data (Table 1b; > 10% explained variance). This may indicate that in potato there is a strong preference for cis- over trans-acting transcriptional control, which is in contrast with observations made in other studies where trans-acting QTLs are generally overrepresented [2,4,5,28]. Although the low mapping resolution and accompanying wide eQTL confidence intervals does not allow distinction between trans- or cis-eQTL for genes mapped on the same linkage group, this in itself cannot explain the large differences observed.

Using the threshold for significance, we have classified genes in groups based on the amount of explained

**Table 1 Overview of eQTL analysis results based on obtained population-wide gene expression profiles for potato tuber and leaf tissue**

A						
Nr of QTLs per Feature*	1	2	3	4	5	Total nr of QTLs
Nr of Features	12331	2269	213	19	2	<b>17594</b>
B						
QTL Expl. Var. %	> 10%	> 20%	> 30%	> 50%	> 70%	> 90%
Nr of QTLs Tuber	12192	10974	7306	3930	1856	293
Cis (%)	7550 (61.9)	7243(66.0)	5699 (78.0)	3279 (83.4)	1515 (81.6)	248 (84.6)
Trans (%)	3629 (29.8)	2792(25.4)	935 (12.8)	269 (6.8)	127 (6.8)	10 (3.4)
Unknown (%)	1013 (8.3)	939 (8.6)	672 (9.2)	382 (9.7)	214(11.5)	35 (11.9)
Nr of QTLs Leaf	10745	9629	6644	3884	1897	388
Cis (%)	7574 (70.5)	7211 (74.9)	5517 (83.0)	3239 (83.4)	1552 (81.8)	297 (76.5)
Trans (%)	2180 (20.3)	1536 (16.0)	497 (7.5)	255 (6.6)	118 (6.2)	27 (7.0)
Unknown (%)	991 (9.2)	882 (9.2)	630 (9.5)	390 (10.0)	227 (12.0)	64 (16.5)
C						
Tissue Comparison	Nr of eQTLs	Cis (%)	Trans (%)	Unknown (%)		
Leaf + Tuber eQTL	5818	4929 (84.7)	401 (6.9)	488 (8.4)		
Leaf Unique eQTL	5126	2839 (55.4)	1784 (34.8)	503 (9.8)		
Tuber unique eQTL	6650	2882 (43.3)	3243 (48.8)	525 (7.9)		
Tissue dependent**	1713	530 (30.9)	977 (57.0)	285 (12.0)		

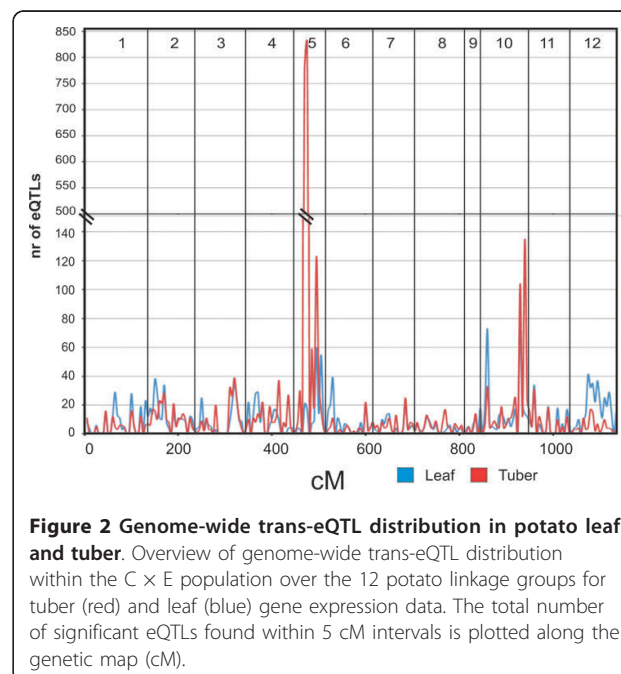
\* Leaf + Tuber data

\*\* eQTL both in leaf and tuber but non-overlapping

variance. Interesting to note is the reduction in the ratio between the numbers of cis- and trans-acting QTLs with increased explained variance (Table 1b). Several studies have reported a similar increase in local (cis) over trans-acting regulation, indicating that on average, cis-regulation results in stronger differential expression and thus genetic variation in comparison to trans-regulation [3,4,29-31].

One explanation for the relative high percentage of cis-eQTLs in potato could be its highly heterozygous nature and therefore possibly a tolerance to higher levels of sequence polymorphisms (SNPs and indels). Local (cis) control of transcriptional regulation may consequently be the preferred mechanism for transcriptional control in heterozygous plant species such as potato. Alternatively, the ability to detect trans-eQTLs with minor phenotypic effects is largely depending on the statistical power of the QTL study. Applying conservative filtering may result in many false negatives biasing detection towards cis-eQTLs as suggested by Petretto et al. [31]. One striking observation is the massive peak of trans-acting QTLs found on linkage group 5 (Figure 2). The question arises whether the high number of QTLs found in this small genomic region on linkage group 5 can be attributed to a single 'master switch' controlling the expression of many genes. The variation in timing of senescence and tuber formation has been phenotyped in the C × E population and has previously been

mapped as the earliness locus (*El*) in the same genomic region as the eQTL hotspot found for the tuber data on chromosome 5 [32]. The C × E population shows a strong variation in the onset of tuber formation (early-late) and to allow a comparison of gene expression in



**Figure 2 Genome-wide trans-eQTL distribution in potato leaf and tuber.** Overview of genome-wide trans-eQTL distribution within the C × E population over the 12 potato linkage groups for tuber (red) and leaf (blue) gene expression data. The total number of significant eQTLs found within 5 cM intervals is plotted along the genetic map (cM).

tubers of equivalent physiological state, a choice was made to extract RNA of 'mature' tubers harvested several weeks after haulm killing. Two scenarios may be envisaged to explain the large number of genes of which expression is regulated by the region that harbours the earliness locus. First, the initial assumption of a similar physiological state of the sampled tubers may be false; tubers formed late in the season may still be metabolically more active, even after haulm killing, than tubers formed earlier in the season. The difference in enzymatic activities in early and late tubers could be a consequence of the observed differential gene expression levels. In this scenario, the detected eQTLs are not under the direct control of a master switch regulator but merely reflect a different metabolic status of the tubers as a result of an earlier event (timing of tuberization). An alternative explanation could be that the locus controlling plant maturity and timing of tuberization is indeed directly regulating several metabolic pathways throughout the plant life cycle. Interestingly, within the leaf expression data this hotspot was not observed. The variation in gene expression explained by the *El* locus on chromosome 5, shows that regulation of gene expression in the tuber persists even after tuber harvest and throughout storage. These observations emphasize the impact of sampling strategies on the outcome of genetical genomics experiments.

In addition to the major QTL peak on chromosome 5, further comparison of trans-acting QTL distribution between leaf and tuber data reveals few other overlapping regions of high QTL frequency in comparison to the distribution plots containing all QTLs. Overlapping trans-eQTL hotspots for both tissues can be found on LG 5 and 10 and to a lesser extent on LG 3, 9 and 11.

A total of 5,818 overlapping QTLs (spanning the same genomic region) were found in leaf and tuber, with only 401 features (6.9%) exhibiting trans-acting transcriptional control (Table 1c). It has previously been suggested that trans-acting QTLs are more often tissue specific [5,28]. In this study, we find a similar result when considering only those QTLs that were uniquely found in one of the two tissues (leaf or tuber) for which the percentages of trans-acting QTLs are elevated to 34.8% for leaf and 48.8% for tuber, respectively. Trans-acting QTLs with different eQTLs identified for the same gene in the two tissues, may be crucial in establishing organ dependent transcriptional networks. The different mRNA extraction methods used, or differences in mRNA extractability for both organs, is not likely to impact on the identification of genetic regulation of gene expression.

#### Tissue dependency

Along with speciation, organ-specific gene regulation is likely to have evolved for the specific requirements of

new organs or organs that have acquired new functions. The activity of different metabolic routes is organ dependent and subject to change throughout the growing season under the influence of environmental factors and developmental stage. It is therefore not surprising that tissue or developmental stage-specific eQTLs can be identified in genetical genomics studies [5,19,20]. A common approach to analyse representation of genes active in particular pathways in biological data sets is to use Gene Ontology (GO) classification (Methods). In Table 2, we compared functional classification of features with matching eQTLs in both tuber and leaf tissue against tuber and leaf specific eQTLs. In the set of matching eQTLs (present in both tuber and leaf) only few trans-acting eQTLs were found, which is consistent with the idea that trans-regulation is an important driver of differentiation (Table 1c and [5]). This is exemplified by the lack of any trans-acting eQTLs in the development- and photosystem-associated GO classes. Although most functional classes are well represented in both tuber and leaf specific eQTLs, it is interesting to

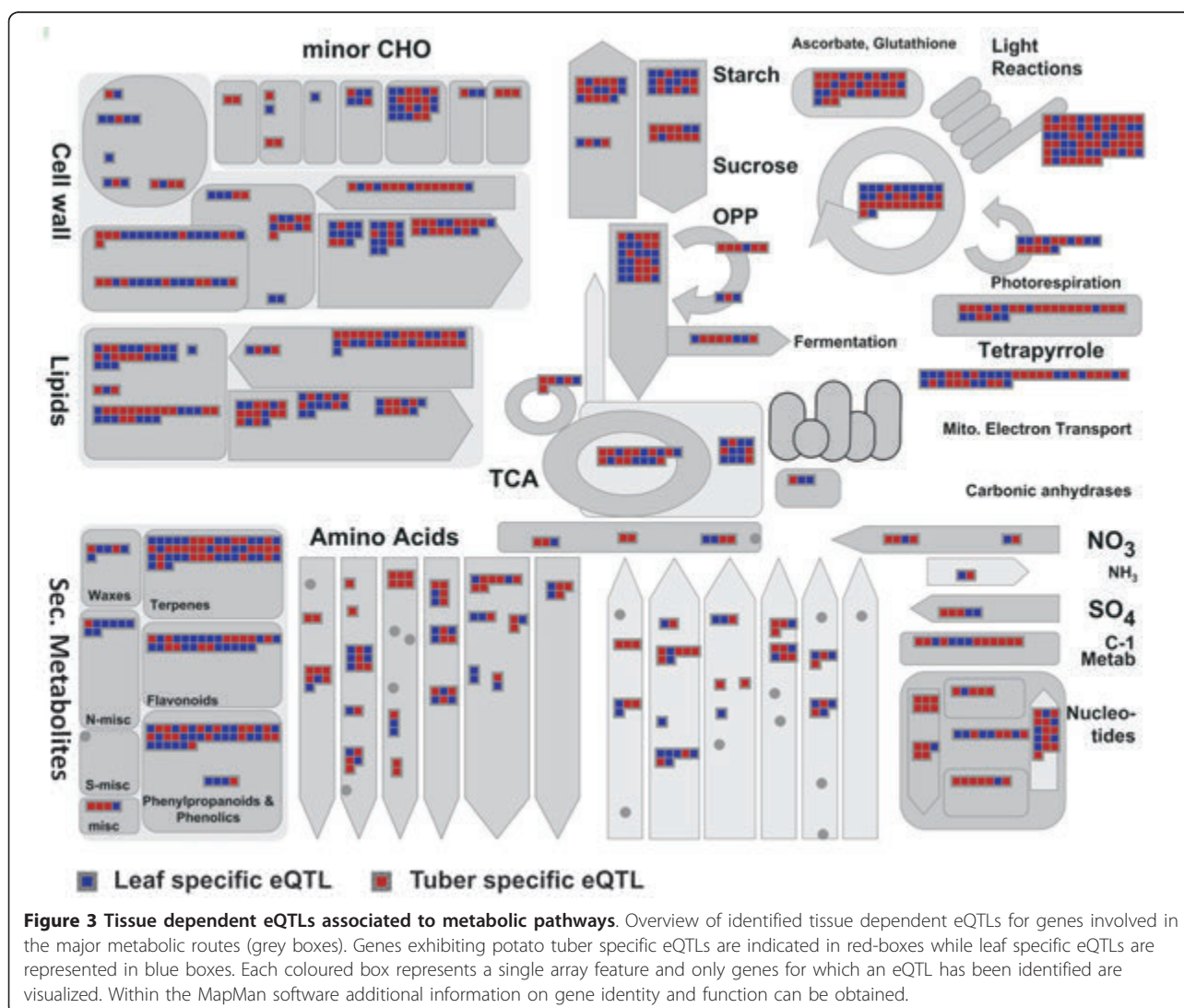
**Table 2 Functional classification of cis- and trans-acting eQTLs in both leaf and tuber (overlap) or uniquely present in leaf or tuber**

GO_Class	Overlap		Leaf Unique		Tuber Unique	
	Cis	Trans	Cis	Trans	Cis	Trans
<b>Protein</b>	639	30	297	170	332	411
<b>RNA</b>	361	19	174	125	260	279
<b>Signalling</b>	172	9	105	48	91	91
<b>Stress</b>	133	6	78	73	71	71
<b>Transport</b>	114	4	73	49	65	70
<b>Cell cycle and organisation</b>	101	10	45	25	58	68
<b>Lipid metabolism</b>	94	3	49	26	29	61
<b>Amino acid metabolism</b>	79	4	40	18	39	49
<b>Secondary metabolism</b>	77	2	57	33	34	41
<b>Hormone metabolism</b>	74	1	69	40	75	57
<b>DNA/Chromatin</b>	57	3	43	19	56	58
<b>Redox</b>	50	2	20	9	24	30
<b>Development</b>	45	0	26	14	37	33
<b>Cell wall</b>	41	3	62	24	27	53
<b>Nucleotide metabolism</b>	34	1	15	3	26	18
<b>Minor CHO metabolism</b>	33	2	15	11	15	16
<b>Photosystem</b>	32	0	35	19	26	67
<b>Mitochondrial/ATP synthesis</b>	29	3	9	12	9	11
<b>Major CHO metabolism</b>	24	1	15	7	21	15
<b>Tetrapyrrole synthesis</b>	12	2	6	4	9	13
<b>Glycolysis</b>	10	1	9	8	10	10
<b>Not assigned</b>	2457	289	1386	928	1414	1560

note that these comprise different sets of genes, both cis- and trans-acting, which may have evolved in a tissue dependent manner. For example a considerable number of genes associated to stress response, show leaf or tuber specific QTLs likely to reflect the different environmental influences that leaves and tubers need to respond to and the differences in the responses themselves.

The formation of tubers requires the expression of many genes controlling the initiation of underground stems (stolons) that under favourable conditions are able to develop tubers acting as sink organs. The genome-wide RNAseq data [25] demonstrates that there are few genes that are exclusively expressed in tubers, which would imply that the tuber-specific development and response to the environment have evolved through modulation of expression and thus subfunctionalization of genes. In Figure 3, a schematic overview of important

metabolic routes and associated genes is provided on the basis of presence of either tuber (red) or leaf (blue) specific eQTLs. Thus, any gene active within a metabolic pathway that is under genetic control (eQTL) is represented as either a red box (tuber) or leaf (blue box). Genes active in metabolic routes which are not under genetic regulation and thus do not produce eQTLs, are not represented. In this manner, tissue dependent variation in gene expression based on a genetic factor can be easily visualized and this provides a useful resource in the identification of key regulatory genes in the respective tissues. In both tubers and leaves, we found tissue-dependent differential expression of genes associated with major carbohydrate metabolism (CHO) (Additional file 2 and Additional file 3). In tubers, eQTLs for genes involved in the sucrose degradation and starch biosynthesis were overrepresented whereas for the leaf expression dataset, there appears to



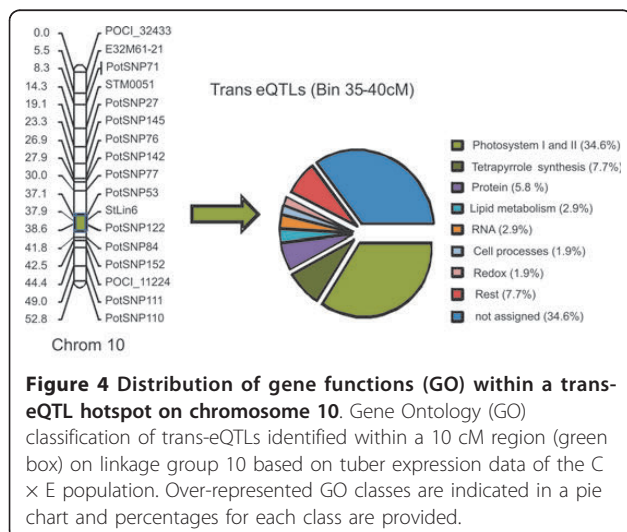
be a stronger bias towards differential regulation of starch degradation. In both tissues similar gene functions are expressed, but these are encoded by different transcripts or are differentially regulated, as reflected by the different positions of eQTLs for these genes in the two tissues. Here, we find that differentiation in the regulation of carbohydrate metabolism across tissues is strongly associated to their function, i.e. the tuber as a carbohydrate storage organ (starch synthesis) and the transient starch produced and subsequently degraded at night in the leaves. Using the MapMan visualization tool, relevant metabolic routes can be analysed in greater detail including secondary metabolites, glycolysis, transport genes or transcription factors, providing insight into tissue dependent differential regulation and whether the underlying genetic component is either cis- or trans-acting. This sort of tissue dependency not only allows the identification of genes active in the respective tissues with different mechanisms of transcriptional control (eQTLs), but also provides insight in organ dependent regulation which is the basis for tissue specific network reconstructions.

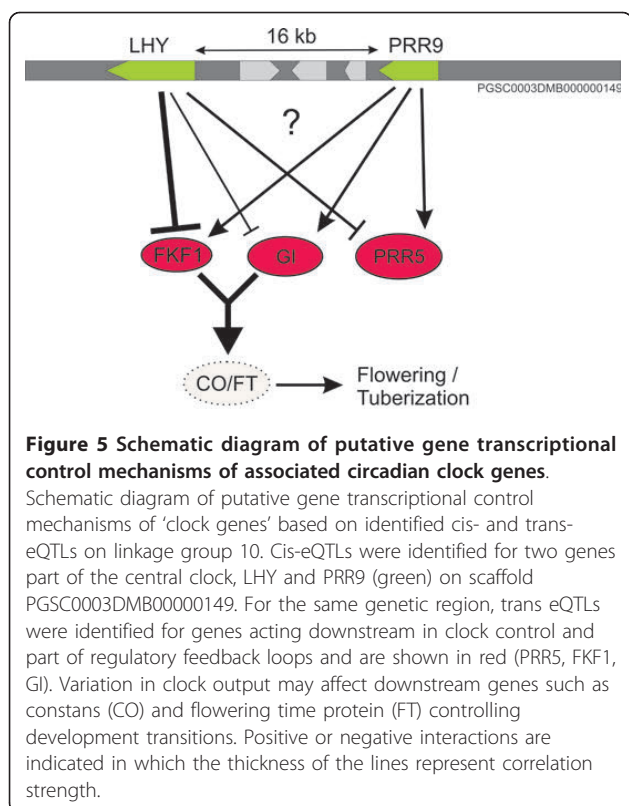
#### Identification of trans-regulatory networks

Co-localization of trans-acting eQTLs may indicate a similar mechanism of transcriptional control. We initially looked for over-representation of GO categories in those regions in the genome with a high number of trans-acting eQTLs from either leaf or tuber. On chromosome 10 (35-40 cM), we found a disproportionately large number of tuber-specific eQTLs (42.3%) associated with photosystem I, II, and tetrapyrrole metabolism (Figure 4 and Table 2). Greening of potato tubers after exposure to light is unacceptable for consumption and processing industry, as it is often associated with high

levels of glycoalkaloid content. Edwards et al. [33], however, did not find a direct metabolic link between Chlorophyll (Chl) and glycoalkaloid biosynthesis after exposure to light. Other studies have shown genotype dependence of the tuber response to light and the ratio of glycoalkaloid and chlorophyll synthesis, as well as of physiological state and intrinsic values [34,35]. Transcriptome-wide RNAseq data show that many photosynthesis-associated genes are expressed at relatively high levels in the tuber peel [25]. During harvesting and storage, tubers may be differentially exposed to short periods of light. This may induce expression of photosynthesis-related genes, but cannot account for the genetic mapping of these expression differences. An increase in the level of expression of genes related to Chl biosynthesis may therefore reflect a heritable factor influencing sensitivity to light exposure. Interestingly, we identified functionally similar genes associated to light reactions that are differentially expressed in the leaves of the C × E genotypes (Figure 3; Light Reactions *top right corner*).

Although, over-representation of functional GO categories within eQTL hotspots can reveal regulatory networks, information on the regulation of a single or few genes can be highly relevant as well. Differential expression of genes involved in maintaining circadian rhythms and clock outputs were found to be partially regulated in the leaf by a region at the top of chromosome 10 (0-10 cM). In the genomic region spanning the QTL interval, cis-eQTLs for genes/array features with high homology to *LHY* (late elongated hypocotyl; MICRO.14662.C1) and *PRR9* (Pseudo Response Regulator 9; PotatoF1061.scf) were identified. Both genes code for integral components of the central clock [36-39]. Interestingly, trans-acting eQTLs for genes with strong homology to members of the central clock; *GIGANTEA* (MICRO.7284.C1), *FKF1* (MICRO.15722.C1) and Pseudo Response Regulator 5 (*PRR5*; MICRO.7036.C1) proteins, all mapped to the same 10 cM genomic region. The best-associated genetic marker is *POCI\_32433*, which maps on genome scaffold PGSC0003DMB000000149 [25] and contains both the *LHY* and *PRR9* genes, suggesting that *LHY* and/or *PRR9* may be involved in regulation of *GIGANTEA*, *FKF1*, *PRR5* genes. *LHY* expression is negatively correlated with downstream effectors *GI*, *FKF1* and *PRR5*, while *PRR9* expression is positively correlated (Figure 5). The circadian clock is involved in many aspects of plant growth and development and is associated with transition from vegetative growth to flowering and adaption to environmental and seasonal changes [reviewed in [40]]. The identification of natural genetic variation in the regulation of central clock genes may prove valuable in the unravelling of complex traits such as plant maturity, timing of





flowering and yield characteristics. For potato, little is known about clock control in relation to plant growth, but timing of tuberization appears to be under direct control of the circadian clock through transcriptional regulation of *CONSTANS* (*CO*) [41] controlling the synthesis and transport of a potato flowering time ortholog associated with tuber induction [42]. In *Arabidopsis*, the central clock includes two MYB genes, *CCA1* and *LHY* [37,43]. In the potato genome, only one ortholog (*LHY*) can be identified possibly indicating a different mechanism of clock control. Similar to *Arabidopsis*, expression patterns of *GI* and *FKF1* in potato follow a diurnal rhythm with peak expression in the afternoon, while *LHY* expression peaks at dawn in agreement with the observed negative correlation (data not shown). In *Arabidopsis*, *GI* and *FKF1* form a complex that controls *CDF1* stability [44]. *CDF1* in turn controls *CO* expression depending on time of day [44]. Variation in the timing of gene expression of *GI* and *FKF1* within a population could therefore cause a phase shift or a change in phase amplitude effecting downstream effectors such as *CO* and *FT* orthologs. Here, we propose a tentative role of *LHY* or *PRR9* in controlling the clock output by targeting expression of *GI*, *FKF1* and *PRR5* (Figure 5). Analogous to the possible difference of physiological state of the tubers at harvest as inferred from our data, the time point of sampling on

the day itself may be crucial for the detection of diurnally controlled expression differences. Additional sequencing of *LHY* and *PRR9* alleles in tetraploids can potentially reveal trait associations related to different clock control.

The strategy of deductive reasoning as used in the above examples requires a priori knowledge of the pathways and the potential interactors. For a model species such as *Arabidopsis* there is a wealth of scientific evidence on gene function, expression patterns and protein interactions that can be integrated in genetical genomics studies. Still, many gene functions are not yet resolved. For the majority of other plant species the annotation of biological functions of genes is often based on sequence similarity alone, which can generate various assumptions lacking scientific support. Genetical genomics studies can provide evidence supporting proposed functions of genes through newly found interactions and co-regulation of genes. This approach will lead to better understanding of transcriptional regulation, and may lead to identification of key regulatory genes underlying trait QTLs.

#### Heterozygosity based QTLs (false cis-eQTLs)

Based on sequence homology of the representative unigenes to the recently published potato genome sequence, physical map positions can be attributed to a large portion of the features present on the array. However, due to the heterozygous nature of potato, the sequences of genotypes are often relatively divergent from the published genome sequence, which makes distinguishing between allelic variation, different gene family members or multi-copy genes distributed throughout the genome not straightforward. SNP frequency estimations in potato range from 1 SNP per 29 bp, found in a single phase comparison targeting 6.6 Mb of sequence [25], to 1 SNP every 87 bp between pairs of randomly selected sequences [45]. The potato unigene set used for microarray oligo design [46] brings forth similar issues, as ESTs used in the assembly originate from various tetraploid cultivars containing allelic variation. Hence, unigenes may contain allelic variation or represent different alleles of the same gene. Consequently, designed oligonucleotides based on these sequences can potentially act as allelic discriminants depending on the genotype used for array hybridizations. Hughes et al. [47] showed that even a single nucleotide mismatch can reduce signal strength up to 50% depending on the position of the mismatch. To assess allele hybridization specificity within the C × E population, we hybridized genomic DNA of both parental lines on the POCI microarray array (Methods). For 94% of the potato oligonucleotides on the array, hybridization ratios between the C and E parent could be



obtained of which only 1921 features showed signal strength variation larger than 33% between both parents (Additional file 4). Of the identified cis-eQTLs, only 391 and 385 features in tuber and leaf data respectively, result in both a cis-eQTL and a differential genomic hybridization ratio ( $> 33\%$ ). These cis-eQTLs are likely to be false eQTLs since variation in observed signal strength may not reflect actual expression differences. Although the percentage of potential false eQTLs (4.7%) appears to be low considering SNP frequency in potato, the actual percentage may be higher due to several factors. Firstly, the  $C \times E$  population is the result of a backcross, giving rise to a common allele present in both parental lines and for this common allele no specificity can be discriminated when hybridizing parental genomic DNA. Within the population however, any oligonucleotide with unique specificity for the common allele (a) will segregate in an expected 1:2:1 (aa:a:-) ratio, resulting in differential signals that can be interpreted as expression differences. Secondly, in cases for which gene expression is relatively stable across members of the population, hybridization specificity among alleles can result in the identification of eQTLs that are based on much smaller differences than can be statistically inferred based on genomic hybridization signals. Thirdly, oligonucleotides for the array were designed based on coding sequences and when these oligonucleotides span intron-exon splicing junctions, hybridization with genomic DNA may be hindered. The use of high coverage RNAseq data would bypass most of these issues as SNP calling would allow allele distinction [48], while read counts will enable comparison of expression levels of alleles and the detection of eQTLs.

## Conclusions

In this paper we describe the identification of expression QTLs (eQTLs) in two different tissues of one of the most important food crops in the world; potato. With the completed potato genome sequence, population-wide differential gene expression can be queried to differentiate between cis- and trans-acting eQTLs. Overlapping QTLs present in both tissues are predominantly cis-acting while for tissue specific QTLs, the percentage of trans-acting QTLs increases. The type of genetic material screened and sampling strategy applied in genetical genomics studies has a high impact on the output of a genetical genomics study. Interesting regulatory networks have been identified on chromosome 10 associated to the photosystem and circadian clock control. Identification of key regulatory genes and networks unique to either source or sink tissues not only gives us insight in tissue dependent gene subfunctionalization but will also greatly enhance the identification of the causative polymorphism (s) underlying important trait QTLs.

## Methods

### Plant material

Ninety-six individuals, including the parental clones, of a diploid backcross population ( $C \times E$ ) were used in this study. This population is derived from an original cross between potato clones C (USW533.7) and E (77.2102.37) and is described in detail in [32]. Tuber samples originate from a field experiment, grown in repeats during the normal potato-growing season in the Netherlands (April-September). Mature tubers were collected from three plants and representative samples were mechanically peeled and immediately frozen in liquid nitrogen before being ground into a fine powder and stored at  $-80^{\circ}\text{C}$ . Tuber total RNA was extracted from the 96 samples using the hot phenol method described previously [49], DNaseI treated and purified using Qiagen RNeasy columns (Qiagen). Tubers of the  $C \times E$  population were planted in soil filled pots in the greenhouse and grown until stolon emergence after which fully open but still growing young leaves were collected from two replicates and immediately frozen in liquid nitrogen. Leaf RNA was isolated using KingFisher Flex system and the MagMAX<sup>TM</sup>-96 total RNA isolation kit according to the manufacturer's instruction and DNaseI treated prior to labelling.

### Microarray hybridizations and data processing

All tuber samples were labelled with both Cy3 and Cy5-dye using the Low RNA Input Linear Amplification Kit, PLUS, Two-colour (Agilent technologies) according to the manufacturer's protocol starting with 2  $\mu\text{g}$  of purified total RNA. Leaf samples were labelled with Cy3-dye following the same protocol as the tuber samples but starting amounts were 200 ng of purified total RNA. Hybridization and washing was performed according to the Agilent's two-colour hybridization protocol with the following change: 1  $\mu\text{g}$  of labelled Cy5 and/or Cy3 cRNA was used as input in the hybridization mixture. Slides were scanned on the Agilent DNA Microarray Scanner and data extracted using the feature extraction software package (v9.1.3.1) using a standard two-colour protocol. Genes which show consistent low expression were removed and data sets were independently normalized using the quantile normalization procedure (mean) available in Genstat<sup>®</sup> 11.1. As tuber samples were measured twice, only genes with a Pearson correlation coefficient higher than 0.8 across the 94 individuals between the Cy3 and Cy5 datasets were included resulting in 19,956 features. For calculation of genome distribution in tubers the Cy3 data was used. All raw and normalized data files are available as Additional files 5, 6, 7, 8, 9, 10, 11, 12, 13,&14 including annotation and genome mapping information. All normalized expression data for leaf and tuber samples have been deposited in ArrayExpress (E-MTAB-808, E-MTAB-701).

Genomic DNA hybridizations were performed after labelling 800 ng of genomic DNA from both parents (C and E) in duplo, using the BioPrime<sup>®</sup> labelling kit (Invitrogen) with a modified dNTPmix (1.2 mM of dATP, dGTP, dTTP, 0.6 mM dCTP and 5  $\mu$ l dCTP-Cy3 or dCTP-cy5 Agilent technologies) and incubated for 16 h at 37°. Labelled samples were purified using PureLink<sup>™</sup> PCR purification System (Invitrogen) and heat fragmented for 30 min resulting in fragments ranging from ~150 to 600 bp. 1  $\mu$ g of labelled samples (cy3 and cy5) were hybridized at 65° for 17 h in a standard swop-dye experiment using the independently labelled samples. Washing and scanning of the slides was carried out as described for the gene expression experiments. Feature normalization and ratio calculations were carried out using default method available in Agilent Feature extraction software package (v9.1.3.1). Features showing consistent differential expression (> 33%) between both parents are presented in Additional file 4.

#### Genetic map and linkage to the potato genome

The genetic map used in all QTL studies was generated using mapping software Joinmap 4.0<sup>®</sup> [50] and is a modified version of an earlier C  $\times$  E genetic map [51], including all sequence based SNP markers. Additional markers were obtained from the tuber expression data set, in cases where both cy3 and cy5 hybridization signals could be unambiguously scored as present or absent. Marker names originating from tuber expression data have POCI as a prefix. All sequence based markers present in the map are linked to EST contigs or EST singletons and these sequences were blasted against the potato genome scaffolds (v3.4). Segregating markers with unique and significant genome scaffold hits (> 90% homology) could be subsequently linked to the physical genome map (Additional file 15). The C  $\times$  E generated genetic map was validated aligning the sequence based markers along the draft scaffold/pseudo-molecules available from the PGSC website [25]. The potato oligo (60-mer) microarray (POCI) used in the experiments, contains 42,034 features based on a potato unigene set [46]. To allow discrimination between cis and trans-eQTLs all unigenes were blasted against the genome scaffolds sequences, predicted Coding sequences (CDS) and predicted gene regions (including 5' and 3'-UTR's). Features with a unique and significant hit were assigned to genome scaffolds for which the majority has chromosome information [25] and results are presented in Additional file 16. Identified QTLs on the same linkage group as their physical map position are identified as cis-acting while QTLs on different linkage groups are defined as trans-acting. Features on the array for which no physical map position could be assigned are classified as unknown.

#### QTL analysis

We have used the main module of the R/qtl package, and optimized it for high-throughput QTL mapping in outcrossing species such as potato. The R-script checks the ratio of missing values before automatically converting JoinMap segregation scores to R/qtl scores after which the QTL mapping is performed [52]. The genotypic scores and genetic map order used in this study are available in Additional file 15 and 17. The program relies on R/qtl for the QTL mapping where QTL information is extracted from the summary.scanone method and the explained variance of the QTL is computed for each QTL using the *makeqtl* and the *fitqtl* functions. The QTL interval is computed using LOD interval method. The default analysis performs a QTL mapping using the Haley-Knott regression with a mapping step size of 5 cM. The LOD threshold used to determine the QTLs is calculated using Li & Ji algorithm [53]. The number of simulation replicates to perform for the sim.geno function is default to 16 and the n.draws parameter allows further adjustment. For tuber and leaf expression data QTL analysis was run using the LOD interval method with default step size and significance LOD threshold of 4.35.

#### Gene ontology and eQTL visualization

Gene ontology (GO) assignment for all the potato micro-array features has been described previously (Kloosterman et al. 2008). GO information was downloaded through Agilent E-array published designs (<https://earray.chem.agilent.com/earray/>). The number of array features having eQTLs (cis or trans) were counted per GO group and totals are listed. To reduce the number of classes, GO identifiers targeting the same metabolic routes were in some cases merged (Table 2). GO classification of the POCI microarray has been previously linked to the expression data visualization tool MapMan (Kloosterman et al. 2008). MapMan is a user-driven tool that normally displays large datasets (e.g. gene expression data arrays) onto diagrams of metabolic pathways or other processes (<http://mapman.gabipd.org/web/guest/mapman>). Instead of visualizing gene expression data we display the presence of tissue specific eQTLs (eQTL present in leaf = blue box; eQTL present in tuber = red box). In this manner a quick overview can be obtained for each pre-designed metabolic route and genes can be identified that are differentially expressed within a segregating population (Figure 3). Additional information on gene name and function can be obtained for each box when running the MapMan software (publicly available).

## Additional material

**Additional file 1: Frequency distribution of eQTLs over the different potato chromosomes.**

**Additional file 2: Tissue specific and Tissue dependent eQTLs of tuber and leaf tissues belonging to the major carbohydrate metabolism groups.**

**Additional file 3: Schematic overview of tuber and leaf eQTL specificity for major CHO metabolism genes.**

**Additional file 4: Hybridization specificity of genomic DNA from C and E parental clones.**

**Additional file 5: Tuber Cy3 and Cy5 raw data + genome info partI.**

**Additional file 6: Tuber Cy3 and Cy5 raw data + genome info partII.**

**Additional file 7: Tuber Cy3 and Cy5 raw data + genome info partIII.**

**Additional file 8: Tuber Cy3 and Cy5 qnormalized data + genome info partI.**

**Additional file 9: Tuber Cy3 and Cy5 qnormalized data + genome info partII.**

**Additional file 10: Tuber Cy3 and Cy5 qnormalized data + genome info partIII.**

**Additional file 11: Leaf Cy3 Leaf raw data + genome info partI.**

**Additional file 12: Leaf Cy3 Leaf raw data + genome info partII.**

**Additional file 13: Leaf Cy3 qnormalized + genome info partI.**

**Additional file 14: Leaf Cy3 qnormalized + genome info partII.**

**Additional file 15: Marker to Genome.**

**Additional file 16: 4x44k 015425 POCl annotation.**

**Additional file 17: Genotypic scores CxE population.**

## Abbreviations

SNP: Single nucleotide polymorphism; QTL: Quantitative trait loci; LOD: Logarithm of odds

## Acknowledgements

We would like to thank Robert Rap for his help in developing the R/qtl package for high throughput QTL analysis in potato. The work presented was carried out and funded by the EU-SOL project (PL 016214-2 EU-SOL) and the research program of the Centre of BioSystems Genomics (CBSG), which is part of The Netherlands Genomics Initiative/Netherlands Organization for Scientific Research.

## Author details

<sup>1</sup>Wageningen UR Plant Breeding, Wageningen University and Research Center, PO Box 386, 6700 AJ Wageningen, the Netherlands. <sup>2</sup>Graduate School Experimental Plant Sciences, Wageningen, The Netherlands. <sup>3</sup>Centre for BioSystems Genomics, P.O. Box 98, 6700 AA Wageningen, The Netherlands. <sup>4</sup>KeyGene N.V., P.O. Box 216, 6700 AE Wageningen, The Netherlands.

## Authors' contributions

BK designed the tuber expression study, performed QTL analysis, data interpretation and wrote the manuscript. AMA designed and performed the leaf expression study and carried out hybridizations together with MO. PC optimized the R/qtl program and performed QTL and statistical analysis. GL designed the leaf expression study and helped draft the manuscript. RGFV and CWBB helped design the tuber expression experiment and help draft the manuscript. All authors read and approved the final manuscript.

Received: 14 July 2011 Accepted: 7 February 2012

Published: 7 February 2012

## References

1. Jansen RC, Nap JP: Genetical genomics: the added value from segregation. *Trends Genet* 2001, **17**(7):388-391.
2. West MA, Kim K, Kliebenstein DJ, van Leeuwen H, Michelmore RW, Doerge RW, St Clair DA: Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. *Genetics* 2007, **175**(3):1441-1450.
3. Potokina E, Druka A, Luo Z, Wise R, Waugh R, Kearsley M: Gene expression quantitative trait locus analysis of 16 000 barley genes reveals a complex pattern of genome-wide transcriptional regulation. *Plant J* 2008, **53**(1):90-101.
4. Keurentjes JJ, Fu J, Terpstra IR, Garcia JM, van den Ackerveken G, Snoek LB, Peeters AJ, Vreugdenhil D, Koornneef M, Jansen RC: Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *Proc Natl Acad Sci USA* 2007, **104**(5):1708-1713.
5. Drost DR, Benedict CI, Berg A, Novaes E, Novaes CR, Yu Q, Dervinis C, Maia JM, Yap J, Miles B, et al: Diversification in the genetic architecture of gene expression and transcriptional networks in organ differentiation of Populus. *Proc Natl Acad Sci USA* 2010, **107**(18):8492-8497.
6. Jordan MC, Somers DJ, Banks TW: Identifying regions of the wheat genome controlling seed development by mapping expression quantitative trait loci. *Plant Biotechnol J* 2007, **5**(3):442-453.
7. Kirst M, Basten CJ, Myburg AA, Zeng ZB, Sederoff RR: Genetic architecture of transcript-level variation in differentiating xylem of a eucalyptus hybrid. *Genetics* 2005, **169**(4):2295-2303.
8. DeCook R, Lall S, Nettleton D, Howell SH: Genetic regulation of gene expression during shoot development in Arabidopsis. *Genetics* 2006, **172**(2):1155-1164.
9. Fu J, Swertz MA, Keurentjes JJ, Jansen RC: MetaNetwork: a computational protocol for the genetic study of metabolic networks. *Nat Protoc* 2007, **2**(3):685-694.
10. Keurentjes JJ: Genetical metabolomics: closing in on phenotypes. *Curr Opin Plant Biol* 2009, **12**(2):223-230.
11. Morreel K, Goeminne G, Storme V, Sterck L, Ralph J, Coppieters W, Breyne P, Steenackers M, Georges M, Messens E, et al: Genetical metabolomics of flavonoid biosynthesis in Populus: a case study. *Plant J* 2006, **47**(2):224-237.
12. Acharjee A, Kloosterman B, de Vos RCH, Werij JS, Bachem CWB, Visser RGF, Maliepaard C: Data integration and network reconstruction with ~ omics data using Random Forest regression in potato. *Analytica Chimica Acta*, corrected proof.
13. Breitling R, Li Y, Tesson BM, Fu J, Wu C, Wiltshire T, Gerrits A, Bystrykh LV, de Haan G, Su AI, et al: Genetical genomics: spotlight on QTL hotspots. *PLoS Genet* 2008, **4**(10):e1000232.
14. Terpstra IR, Snoek LB, Keurentjes JJ, Peeters AJ, van den Ackerveken G: Regulatory network identification by genetical genomics: signaling downstream of the Arabidopsis receptor-like kinase ERECTA. *Plant Physiol* 2010, **154**(3):1067-1078.
15. Wu C, Delano DL, Mitro N, Su SV, Janes J, McClurg P, Batalov S, Welch GL, Zhang J, Orth AP, et al: Gene set enrichment in eQTL data identifies novel annotations and pathway regulators. *PLoS Genet* 2008, **4**(5):e1000070.
16. Kliebenstein DJ, West MA, van Leeuwen H, Loudet O, Doerge RW, St Clair DA: Identification of QTLs controlling gene expression networks defined a priori. *BMC Bioinformatics* 2006, **7**:308.
17. Druka A, Druka I, Centeno AG, Li H, Sun Z, Thomas WT, Bonar N, Steffenson BJ, Ullrich SE, Kleinhofs A, et al: Towards systems genetic analyses in barley: Integration of phenotypic, expression and genotype data into GeneNetwork. *BMC Genet* 2008, **9**:73.
18. Kliebenstein D: Quantitative genomics: analyzing intraspecific variation using global gene expression polymorphisms or eQTLs. *Annu Rev Plant Biol* 2009, **60**:93-114.
19. Gerrits A, Li Y, Tesson BM, Bystrykh LV, Weersing E, Ausema A, Dontje B, Wang X, Breitling R, Jansen RC, et al: Expression quantitative trait loci are highly sensitive to cellular differentiation state. *PLoS Genet* 2009, **5**(10):e1000692.
20. Potokina E, Druka A, Luo Z, Moscou M, Wise R, Waugh R, Kearsley M: Tissue-dependent limited pleiotropy affects gene expression in barley. *Plant J* 2008, **56**(2):287-296.

21. Sergeeva LI, Vonk J, Keurentjes JJ, van der Plas LH, Koornneef M, Vreugdenhil D: **Histochemical analysis reveals organ-specific quantitative trait loci for enzyme activities in Arabidopsis.** *Plant Physiol* 2004, **134**(1):237-245.
22. Zahn LM, Leebens-Mack JH, Arrington JM, Hu Y, Landherr LL, de Pamphilis CW, Becker A, Theissen G, Ma H: **Conservation and divergence in the AGAMOUS subfamily of MADS-box genes: evidence of independent sub- and neofunctionalization events.** *Evol Dev* 2006, **8**(1):30-45.
23. Lynch M, Force A: **The probability of duplicate gene preservation by subfunctionalization.** *Genetics* 2000, **154**(1):459-473.
24. Li L, Paulo MJ, Strahwald J, Lubeck J, Hofferbert HR, Tacke E, Junghans H, Wunder J, Draffehn A, van Eeuwijk F, et al: **Natural DNA variation at candidate loci is associated with potato chip color, tuber starch content, yield and starch yield.** *Theor Appl Genet* 2008, **116**(8):1167-1181.
25. Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, et al: **Genome sequence and analysis of the tuber crop potato.** *Nature* 2011, **475**(7355):189-195.
26. Keiner R, Kaiser H, Nakajima K, Hashimoto T, Drager B: **Molecular cloning, expression and characterization of tropinone reductase II, an enzyme of the SDR family in Solanum tuberosum (L.).** *Plant Mol Biol* 2002, **48**(3):299-308.
27. Wang P, Dawson JA, Keller MP, Yandell BS, Thornberry NA, Zhang BB, Wang IM, Schadt EE, Attie AD, Kendziorski C: **A model selection approach for expression quantitative trait loci (eQTL) mapping.** *Genetics* 2011, **187**(2):611-621.
28. Grieve IC, Dickens NJ, Pravenec M, Kren V, Hubner N, Cook SA, Aitman TJ, Petretto E, Mangion J: **Genome-wide co-expression analysis in multiple tissues.** *PLoS One* 2008, **3**(12):e4033.
29. Hughes KA, Ayroles JF, Reedy MM, Drnevich JM, Rowe KC, Ruedi EA, Caceres CE, Paige KN: **Segregating variation in the transcriptome: cis regulation and additivity of effects.** *Genetics* 2006, **173**(3):1347-1355.
30. Brem RB, Kruglyak L: **The landscape of genetic complexity across 5,700 gene expression traits in yeast.** *Proc Natl Acad Sci USA* 2005, **102**(5):1572-1577.
31. Petretto E, Mangion J, Dickens NJ, Cook SA, Kumaran MK, Lu H, Fischer J, Maatz H, Kren V, Pravenec M, et al: **Heritability and tissue specificity of expression quantitative trait loci.** *PLoS Genet* 2006, **2**(10):e172.
32. Celis-Gamboa C, Struijk PC, Jacobsen E, Visser RGF: **Temporal dynamics of tuber formation and related processes in a crossing population of potato (Solanum tuberosum).** *Ann Appl Biol* 2003, **143**(2):175-186.
33. Edwards EJ, Saint RE, Cobb AH: **Is there a link between greening and light-enhanced glycoalkaloid accumulation in potato (Solanum tuberosum L) tubers?** *J Sci Food Agric* 1998, **76**(3):327-333.
34. Griffiths DW, Dale MFB, Bain H: **The effect of cultivar, maturity and storage on photo-induced changes in the total glycoalkaloid and chlorophyll contents of potato (Solanum tuberosum).** *Plant Science* 1994, **98**(1):103-109.
35. Percival GC: **The influence of light upon glycoalkaloid and chlorophyll accumulation in potato tubers (Solanum tuberosum L.).** *Plant Science* 1999, **145**(2):99-107.
36. McClung CR: **A modern circadian clock in the common angiosperm ancestor of monocots and eudicots.** *BMC Biol* 2010, **8**:55.
37. Schaffer R, Ramsay N, Samach A, Corden S, Putterill J, Carre IA, Coupland G: **The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering.** *Cell* 1998, **93**(7):1219-1229.
38. Nakamichi N, Kiba T, Henriques R, Mizuno T, Chua NH, Sakakibara H: **PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock.** *Plant Cell* 2010, **22**(3):594-605.
39. Nakamichi N, Kita M, Ito S, Yamashino T, Mizuno T: **PSEUDO-RESPONSE REGULATORS, PRR9, PRR7 and PRR5, together play essential roles close to the circadian clock of Arabidopsis thaliana.** *Plant Cell Physiol* 2005, **46**(5):686-698.
40. Imaizumi T: **Arabidopsis circadian clock and photoperiodism: time to think about location.** *Curr Opin Plant Biol* 2010, **13**(1):83-89.
41. Martinez-Garcia JF, Virgos-Soler A, Prat S: **Control of photoperiod-regulated tuberization in potato by the Arabidopsis flowering-time gene CONSTANS.** *Proc Natl Acad Sci USA* 2002, **99**(23):15211-15216.
42. Navarro C, Abelenda JA, Cruz-Oro E, Cuellar CA, Tamaki S, Silva J, Shimamoto K, Prat S: **Control of flowering and storage organ formation in potato by FLOWERING LOCUS T.** *Nature* 2011, **478**(7367):119-122.
43. Wang ZY, Tobin EM: **Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression.** *Cell* 1998, **93**(7):1207-1217.
44. Sawa M, Nusinow DA, Kay SA, Imaizumi T: **FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis.** *Science* 2007, **318**(5848):261-265.
45. Simko I, Haynes KG, Jones RW: **Assessment of linkage disequilibrium in potato genome with single nucleotide polymorphism markers.** *Genetics* 2006, **173**(4):2237-2245.
46. Kloosterman B, De Koeyer D, Griffiths R, Flinn B, Steuernagel B, Scholz U, Sonnewald S, Sonnewald U, Bryan GJ, Prat S, et al: **Genes driving potato tuber initiation and growth: identification based on transcriptional changes using the POCI array.** *Funct Integr Genomics* 2008, **8**(4):329-340.
47. Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR, et al: **Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer.** *Nat Biotechnol* 2001, **19**(4):342-347.
48. Turro E, Su SY, Goncalves A, Coin LJ, Richardson S, Lewin A: **Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads.** *Genome Biol* 2011, **12**(2):R13.
49. Bachem CW, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RG: **Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development.** *Plant J* 1996, **9**(5):745-753.
50. Van Ooijen JW: *JoinMap<sup>®</sup> 4, Software for the Calculation of Genetic Linkage Maps in Experimental Populations* Wageningen: Kyazma B.V.; 2006.
51. Anithakumari AM, Tang J, van Eck HJ, Visser RG, Leunissen JA, Vosman B, van der Linden CG: **A pipeline for high throughput detection and mapping of SNPs from EST databases.** *Mol Breed* 2010, **26**(1):65-75.
52. Broman KW, Wu H, Sen S, Churchill GA: **R/qtl: QTL mapping in experimental crosses.** *Bioinformatics* 2003, **19**(7):889-890.
53. Li J, Ji L: **Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix.** *Heredity* 2005, **95**(3):221-227.

doi:10.1186/1471-2229-12-17

**Cite this article as:** Kloosterman et al.: Organ specificity and transcriptional control of metabolic routes revealed by expression QTL profiling of source-sink tissues in a segregating potato population. *BMC Plant Biology* 2012 **12**:17.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

