

RESEARCH ARTICLE

Open Access

The characterisation of *AOP2*: a gene associated with the biosynthesis of aliphatic alkenyl glucosinolates in *Arabidopsis thaliana*

Calida S Neal^{1,2}, Dale P Fredericks^{1,3}, Cara A Griffiths¹, Alan D Neale^{1*}

Abstract

Background: Glucosinolates, a group of nitrogen and sulfur containing compounds associated with plant-insect interactions, are produced by a number of important *Brassicaceae* crop species. In *Arabidopsis* the *AOP2* gene plays a role in the secondary modification of aliphatic (methionine-derived) glucosinolates, namely the conversion of methylsulfinylalkyl glucosinolates to form alkenyl glucosinolates, and also influences aliphatic glucosinolate accumulation.

Results: This study characterises the primary structural variation in the coding sequences of the *AOP2* gene and identifies three different *AOP2* alleles based on polymorphisms in exon two. To help determine the regulatory mechanisms mediating *AOP2* expression amongst accessions, *AOP2* 5' regulatory regions were also examined however no major differences were identified. Expression of the *AOP2* gene was found to be most abundant in leaf and stem tissue and was also found to be light dependent, with a number of light regulatory elements identified in the promoter region of the gene. In addition, a study was undertaken to demonstrate that the *Arabidopsis* *AOP2* gene product is functional *in planta*. The over-expression of a functional *AOP2* allele was found to successfully convert the precursor methylsulfinyl alkyl glucosinolate into the alkenyl form.

Conclusions: The expression of the *AOP2* gene has been found to be influenced by light and is most highly expressed in the photosynthetic parts of the *Arabidopsis* plant. The level of *AOP2* transcript decreases rapidly in the absence of light. *AOP2* exists as at least three alleles in different *Arabidopsis* accessions and we have demonstrated that one of these, *AOP2-2*, is functionally able to convert methylsulfinyl glucosinolates into the alkenyl form. The demonstration of the *in planta* functionality of the *Arabidopsis* *AOP2* gene is an important step in determining the feasibility of engineering glucosinolate profiles in food plants.

Background

Glucosinolates are a group of secondary metabolites found exclusively in the Plant Kingdom, predominantly within the *Brassicaceae* family. Glucosinolates are substrates for endogenous thioglucosidases called myrosinases. Following tissue disruption, glucosinolates are hydrolysed by myrosinases to produce compounds with a range of biological activities. These activities include the protection of the plant against pathogens and herbivores as well as plant recognition by specialist predators. Glucosinolates are also responsible for the flavour and

anti-carcinogenic properties of many *Brassica* vegetables [1-3]. Recent interest has focussed on the regulation of glucosinolate synthesis and the potential for manipulating glucosinolate profiles in food crops [4].

The biosynthesis of glucosinolates occurs in three main stages. Initially the precursor amino acids, such as phenylalanine and methionine may be elongated by the addition of one or several methylene groups. During the second stage the modified precursor amino acids are converted into glucosinolates (glucone formation), and finally secondary modification of the glucosinolate structure takes place [1,2,5,6]. Modifications of the basic glucosinolate molecule create an enormous variety of glucosinolate structures. Secondary modification such as oxidation, hydroxylation, methoxylation, desaturation

* Correspondence: alan.neale@sci.monash.edu.au

¹School of Biological Sciences, Monash University, Melbourne, Victoria, 3800, Australia

Full list of author information is available at the end of the article

and glycosylation primarily occurs on the methionine side chain and occasionally on the glucose moiety [3]. The side chain modification of the glucosinolate molecule is of particular importance as the structure of the side chain largely determines the nature of the products formed following glucosinolate hydrolysis by myrosinases [3,6].

The major steps involved in 4C aliphatic glucosinolate side chain modification are outlined in Figure 1. Following the formation of methylthioalkyl glucosinolate from methionine, an initial oxidation occurs, catalyzed by the product of the *GS-OX* locus to form methylsulfinylalkyl glucosinolate. Further modification of the side chain is performed by the protein products of the *GS-ALK* and *GS-OHP* loci [7]. In *Arabidopsis* the *GS-ALK* and *GS-OHP* loci map to a region near the top of chromosome four, termed the *GS-AOP* locus [8,9] and fine scale mapping has led to the identification of the three genes, *AOP1*, *AOP2* and *AOP3*, that encode 2-oxoglutarate-dependent dioxygenases [10,11]. The *GS-ALK* (*AOP2*)

locus protein product is responsible for the formation of alkenyl glucosinolates whilst the *GS-OHP* (*AOP3*) locus protein product is thought to be responsible for the formation of hydroxyalkyl glucosinolates [7]. The alkenyl glucosinolate can also be further hydroxylated through the action of the *GS-OH* locus protein product [7,10,11].

The *AOP1* gene is believed to be the ancestral gene that gave rise to *AOP2* and *AOP3* through a series of gene duplication events. Currently the function of *AOP1* is unknown [11]. Heterologous expression in *Escherichia coli* has shown that *AOP2* catalyses the conversion of 3-methylsulfinylpropyl- and 4-methylsulfinylbutyl glucosinolates to the corresponding alkenyl glucosinolates, 2-propenyl and 3-butenyl, respectively. The heterologous expression of *AOP3* displayed only a weak conversion of 3-methylsulfinylpropyl glucosinolate to 3-hydroxypropyl glucosinolate [11]. However, in the accessions examined a perfect correlation was reported between the expression of a functional *AOP2* and *AOP3* and the accumulation of alkenyl and hydroxyalkyl glucosinolates,

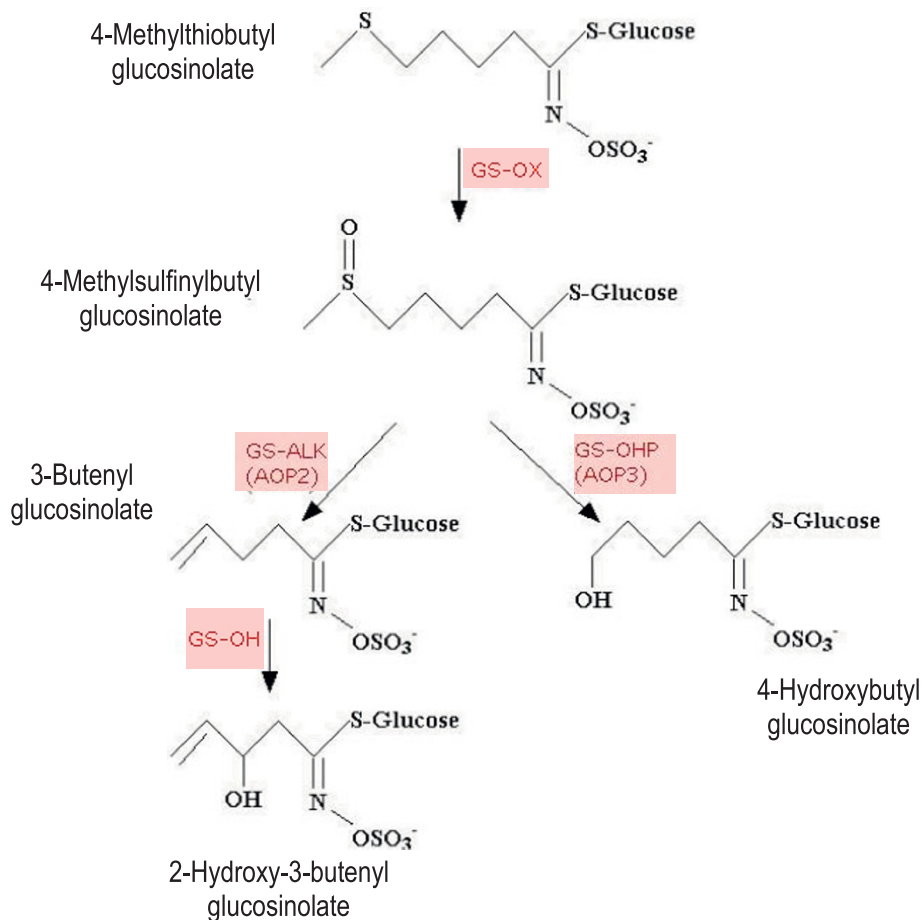


Figure 1 The modification of the 4-methylthiobutyl glucosinolate side chain [7,9-11]. Arrows show the steps for 4C aliphatic glucosinolate modification with the locus responsible for each biochemical reaction shown in red text. The structures and chemical names of intermediates and end products are provided.

respectively [11]. *Arabidopsis* displays differential *AOP* leaf expression whereby a particular accession expresses either *AOP2* or *AOP3* or neither, but not both, and in some *Arabidopsis* accessions the absence of both functional enzymes leads to the accumulation of the precursor methylsulfinylalkyl glucosinolate [11].

The *AOP2* homologue (*BoGSL-ALK*) from collard (*Brassica oleracea*) has been cloned and characterised and it was found that the gene product is able to catalyse the conversion of methylsulfinylalkyl glucosinolate to the alkenyl form *in planta* [12]. Broccoli, on the other hand, contains a non-functional allele of *BoGSL-ALK* and hence accumulates the 4-methylsulfinylbutyl glucosinolate, glucoraphanin [12]. The accumulation of glucoraphanin is of particular importance as the hydrolysis of glucoraphanin by myrosinase leads to the production of the isothiocyanate, sulforaphane, which is a major inducer of phase II detoxification enzymes that are important in chemoprotection from carcinogens [13].

It is thought that natural variation at the *AOP* locus, in particular the differential expression of the *AOP2* and *AOP3* genes, accounts for a large amount of the variation in the glucosinolate profiles and the levels of glucosinolate accumulation found between different *Arabidopsis* accessions [11,14,15]. A mechanism for the differential leaf expression, whereby a particular accession transcribes either *AOP2* or *AOP3* but not both, has recently been reported to be due a complete inversion of the *AOP2* and *AOP3* structural genes in some accessions, causing the *AOP3* gene to be expressed from the *AOP2* promoter [15]. The important role that *AOP2* plays as part of a feedback loop which regulates the transcription of other aliphatic glucosinolate biosynthetic genes has also been reported, however, the mechanisms underlying this regulation have not yet been elucidated [16]. Due to the key role that the *AOP2* structural gene and promoter plays in the aliphatic glucosinolate biosynthetic pathway, the *AOP2* coding sequence and 5' regulatory region has been characterised in a range of accessions in this report.

The accumulation of glucosinolates has been well studied in *Arabidopsis*. An extensive survey of the glucosinolates accumulated in the *Arabidopsis* accession Col-0 demonstrated that the composition and concentration of glucosinolates varies both developmentally and spatially. It was found that dormant and germinating seeds contain the highest glucosinolate concentration and diversity, followed by inflorescences, siliques, leaves and roots [17]. This pattern of accumulation has also been found in other glucosinolate producing species [18] and correlates with the current theories on the optimal distribution of defence substances which state that the maximum level of protection is provided to the plants'

reproductive tissues [19]. In this study the spatial pattern of *AOP2* transcript accumulation was examined.

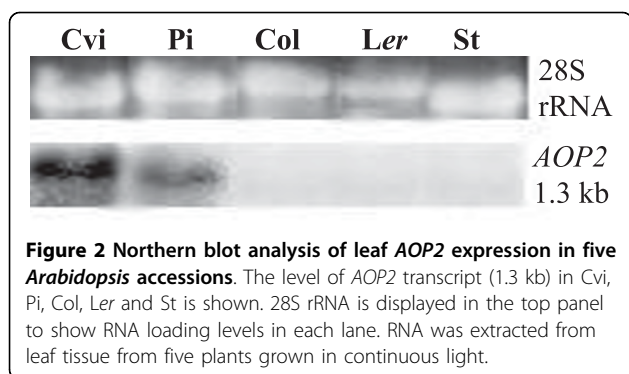
The accumulation of glucosinolates in various plants species is also influenced by many environmental factors [20-22]. Climatic and seasonal conditions have long been known to influence glucosinolate concentration in various crop species [23-25]. Glucosinolate levels have also been found to vary throughout the day, with a two-fold change in total glucosinolate concentration being recorded in *B. oleracea* during one diurnal cycle [26,27]. It was found that the concentration of 3-indolylmethyl glucosinolate was reduced in *Brassica napus* seedlings when kept in darkness for 24 hours [28]. It has also been found that watercress (*Nasturtium officinale* R. Br.) accumulates higher levels of the glucosinolate, gluconasturtiin, when grown under long day conditions [29]. Currently little is known about changes in glucosinolate accumulation in response to environmental factors in *Arabidopsis*. In particular, little is known about the environmental factors that activate or repress glucosinolate biosynthetic genes [30]. Therefore a study was initiated to investigate the level of transcript accumulation of the *AOP2* gene in response to light.

As many plant species within the *Brassicaceae* are important crop plants, the ability to alter glucosinolate profiles *in planta* is potentially valuable from an agricultural perspective. Therefore the current study also aimed to demonstrate that the *Arabidopsis AOP2* gene product is functional *in planta*. This was done through the introduction of the *AOP2* gene from the *Arabidopsis* Pitztal accession (*AOP2_{pi}*) into the Columbia accession. It was shown that Columbia plants accumulate methylsulfinylalkyl glucosinolates and it was hypothesised that the introduction of this gene would result in the conversion of the precursor methylsulfinylalkyl glucosinolate to the alkenyl glucosinolate form.

Results

AOP2 gene expression in the leaves of five *Arabidopsis* accessions

The accumulation of the *AOP2* transcript in the leaves of five *Arabidopsis* accessions: Col, Cvi, Ler, Pi and St, was investigated. Northern blot analysis indicates that of these accessions, Cvi and Pi exhibit high levels of *AOP2* gene expression, whereas Ler and St display low or undetectable levels of *AOP2* gene expression (shown in Figure 2). Figure 2 also indicates that leaf expression of *AOP2* in Col is not detectable by northern analysis. Expression of a non-functional *AOP2* transcript in the Col accession has previously been reported [11]. Our inability to detect the *AOP2* message in the leaves of Col suggests that the gene may be transcribed at a low level and/or the transcript may be unstable.



Analysis of the *AOP2* structural gene

The *AOP2* structural gene was sequenced in Col and the two *AOP2*-expressing accessions (Cvi and Pi) and two *AOP3*-expressing accessions (Ler and St). The sequenced region encompassed all introns and exons between the translation initiating ATG codon and the stop codon. Intron/exon boundaries were predicted from the Cvi sequences using the Genscan program found at (<http://genes.mit.edu/GENSCAN.html>) and confirmed by sequencing RT-PCR products. The *AOP2* gene was predicted to contain three exons at positions 1-368, 588-1257, and 1647-1907 bp (Figure 3).

The *AOP2* gene sequence exhibits complete identity between the five accessions across introns and exons with the exception of three regions in exon two where polymorphisms were found to occur (shown in Figure 3). The *AOP2* allele present in the Cvi and St accessions was named allele 1 (*AOP2-1*), the Pi allele which contains two polymorphisms (18 bp and 9 bp deletions) was named allele 2 (*AOP2-2*), whereas the Col and Ler alleles that contain these two polymorphisms as well as an additional 5 bp frame shift deletion was named allele 3 (*AOP2-3*).

The 18 bp and 9 bp polymorphisms cause a nine amino acid difference between the predicted protein product encoded by the *AOP2-1* and *AOP2-2* alleles (432 and 423 amino acids, respectively). If the *AOP2-3* transcript containing the 5 bp deletion was translated, the frame shift would result in a truncated protein of 172 amino acids being produced, which is in agreement with the Col *AOP2* gene product (*AOP2-3*) having been reported as non-functional [11].

Exon 2 from *AOP2* in twelve additional accessions was amplified and sequenced (data not shown). Six of these accessions (Di-I, Ni, Sf1, Rsch-4, Sf2, Su-O) generated sequence characteristic of the *AOP2-1* allele, whereas the accessions C24, Can-O, Cond, Sha and Sorbo generated sequence characteristic of the *AOP2-2* allele, whilst La-O appears to harbour the *AOP2-3* allele. Northern analyses show that the six accessions analysed containing the *AOP2-2* allele accumulate *AOP2* transcript in

the leaves, whereas leaf *AOP2* expression in the three accessions containing *AOP2-3* could not be detected. The presence of the *AOP2-1* allele does not correlate either positively or negatively with *AOP2* expression as four of the *AOP2-1* containing accessions accumulate leaf *AOP2* transcripts and four accumulate leaf *AOP3* transcripts (data not shown).

AOP2 spatial expression profile

An analysis of the *AOP2* transcript levels in various organs (leaf, stem, floral tissue, roots and dormant seeds) of Pi and Ler plants grown in CL was conducted (Figure 4a and 4b). Leaf tissue was collected at the four leaf stage of vegetative growth from Ler and the six leaf stage from Pi. The stem, root and floral tissues were collected from both types of plant approximately four days after the plants had formed floral bolts. In Pi the highest level of the *AOP2* transcript was detected in the leaf tissue and the transcript was also found to accumulate in stem tissue and to a lesser extent in the floral organs. The root tissue and dormant seeds displayed little or no detectable *AOP2* transcript. No *AOP3* expression was detected in Pi plants whereas Ler accumulated *AOP3* transcript in the same spatial pattern as that exhibited by *AOP2* in Pi (data not shown). As expected, no *AOP2* expression was found in any part of the Ler accession plant. This indicates that, at least in the plant organs of the accessions tested, the accession-specific leaf expression of the *AOP2* and *AOP3* genes extends to the other tissues.

A comparison of the *AOP2* 5' regulatory region amongst accessions

Two kb of the 5' regulatory region upstream of the *AOP2* gene was sequenced with the intention of identifying any promoter differences between *AOP2*-expressing accessions (Cvi and Pi) and Col which could help explain the different expression profiles observed amongst these accessions. The *AOP2* non-expressing accessions (Ler and St) were also included in the analyses. Interestingly, the Col and Pi 5' regulatory region sequences were found to be identical to each other (Figure 5), notwithstanding the substantial difference in *AOP2* transcript accumulation levels in these two accessions. This study has detected no *AOP2* expression in the Col accession using either northern analysis (Figure 2) or RT-PCR, although it has previously been reported that a low level of *AOP2* transcript (*AOP2-3*) is detectable in Col by RT-PCR [11]. This suggests that the *AOP2-3* transcript, which encodes a non-functional protein, may be degraded rapidly.

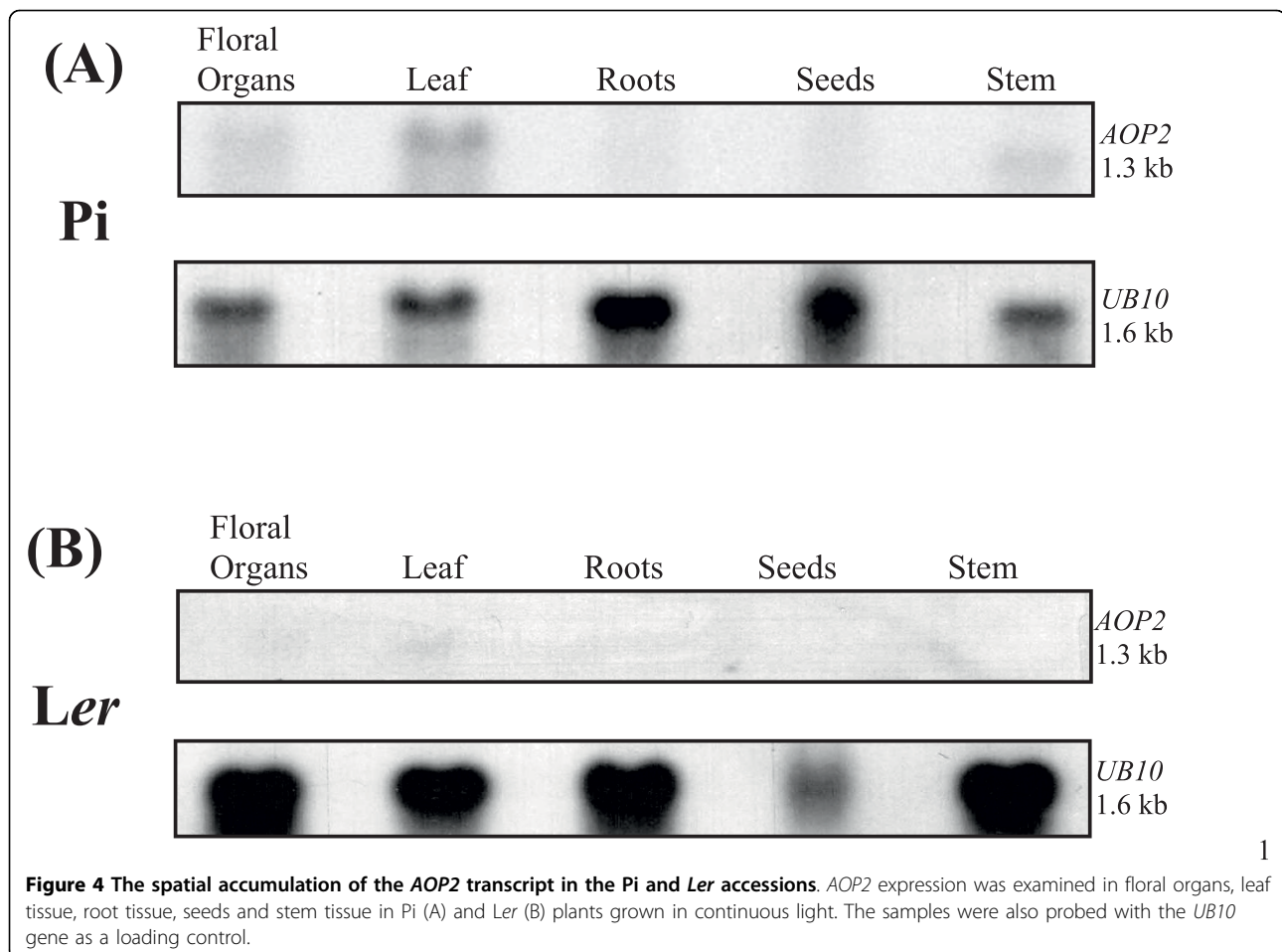
Figure 5 shows that the Ler and St sequences were identical to each other and the Cvi 5' regulatory sequence also shows a high level of similarity to the Ler/St sequence with only some minor differences being found.

```

1      11      21      31      41      51      61      71      81      91
ATGGGTTCATGCAGTCTTCAGCTTCCACTCATCAATCTTGCGGCAAAAACCCTAGAACCAGCAAGCTCAAAGTGGGCTGAAGTGAGGAGTGATGTCCGTA
101    111    121    131    141    151    161    171    181    191
AAGCTCTCGAAGACTTCGGTGTTTTCGAAGCTTCATATGATAAAGTGTCATTGGAGCTTCAGGAATCAATTATGAAAACCATGGAAGAGCTTTTTCGCATT
201    211    221    231    241    251    261    271    281    291
GCCAGTTGAGCACAAACAACGAACGTGTGCCCCAAACCTTACGTTGGATATTTGAATCATAAATATCTCTCCGAGAGTTTGGGGATAGTAATGCTAAT
301    311    321    331    341    351    361    371    381    391
ATTTTGGAGACATCACGAATTTACTCACAACTATGGCTCATGGTGACGGTAACGAGAATATCAGGcgaaaactcatatgtttttactatgtttag
401    411    421    431    441    451    461    471    481    491
taaaaatcatgtaatctttttctctaaaagtaaaatgttaattcactctcaaatcgtagaaccaatgaactgaagacaaaaaaaaaattatc
501    511    521    531    541    551    561    571    581    591
tgaaaacaaattttctgcatattcttaattgtatttctgtttttcatatgaatcctcactttgtttttgtgaaacagTAAAACGATCCAG
601    611    621    631    641    651    661    671    681    691
TGTTTGCGGAGAAGTAGTGGAAATAGATGTGATGGTGAGAAGATGGTATGGAGAGCTTTGGAATAGAAAATACATTGTATGACCACTCTAAAGTCTA
701    711    721    731    741    751    761    771    781    791
CGG....CGTATGCGACTGATGAAGTATATTGCACCACCTGAGGGTGATGCTAATACTACTGTTGACGATAATGTGATCTTCTTGCTAAAGCTTAATAT Col
CGG....CGTATGCGACTGATGAAGTATATTGCACCACCTGAGGGTGATGCTAATACTACTGTTGACGATAATGTGATCTTCTTGCTAAAGCTTAATAT Ler
CGGAATACCGTATGCGACTGATGAAGTATATTGCACCACCTGAGGGTGATGCTAATACTACTGTTGACGATAATGTGATCTTCTTGCTAAAGCTTAATAT St
CGGAATACCGTATGCGACTGATGAAGTATATTGCACCACCTGAGGGTGATGCTAATACTACTGTTGACGATAATGTGATCTTCTTGCTAAAGCTTAATAT Pi
CGGAATACCGTATGCGACTGATGAAGTATATTGCACCACCTGAGGGTGATGCTAATACTACTGTTGACGATAATGTGATCTTCTTGCTAAAGCTTAATAT Cvi
801    811    821    831    841    851    861    871    881    891
TGTTGGTGTTGAACCATATGTTGGTTAAAGTTATGCCGATATAGTGATGATGTTATGCTAATGCTAGTGTTATGCTGGTGTGGTGCTAATGT
901    911    921    931    941    951    961    971    981    991
AACGCTGATACCGGTGTTATGATAATCT.....TAATGTTGATGCTAATGTTGCCGTCGGTGGCGTTAATGCTAATACTGATC Col
AACGCTGATACCGGTGTTATGATAATCT.....TAATGTTGATGCTAATGTTGCCGTCGGTGGCGTTAATGCTAATACTGATC Ler
AACGCTGATACCGGTGTTATGATAATCTAAATGTTGATGCTGAAGCTAATGTGATGCTAATGTTGCCGTCGGTGGCGTTAATGCTAATACTGATC St
AACGCTGATACCGGTGTTATGATAATCT.....TAATGTTGATGCTAATGTTGCCGTCGGTGGCGTTAATGCTAATACTGATC Pi
AACGCTGATACCGGTGTTATGATAATCTAAATGTTGATGCTGAAGCTAATGTGATGCTAATGTTGCCGTCGGTGGCGTTAATGCTAATACTGATC Cvi
1001   1011   1021   1031   1041   1051   1061   1071   1081   1091
TTGGTGTGGTGTTATGCAATTCTAATGTGCTGTGAATGCTAAAAACTG.....GTGGTGATGATGTTGAGGCTAATGATGATGATAGGAAAA Col
TTGGTGTGGTGTTATGCAATTCTAATGTGCTGTGAATGCTAAAAACTG.....GTGGTGATGATGTTGAGGCTAATGATGATGATAGGAAAA Ler
TTGGTGTGGTGTTATGCAATTCTAATGTGCTGTGAATGCTAAAAACTG.....GTGGTGATGATGTTGAGGCTAATGATGATGATAGGAAAA St
TTGGTGTGGTGTTATGCAATTCTAATGTGCTGTGAATGCTAAAAACTG.....GTGGTGATGATGTTGAGGCTAATGATGATGATAGGAAAA Pi
TTGGTGTGGTGTTATGCAATTCTAATGTGCTGTGAATGCTAAAAACTG.....GTGGTGATGATGTTGAGGCTAATGATGATGATAGGAAAA Cvi
1101   1111   1121   1131   1141   1151   1161   1171   1181   1191
GAAGTTGGTTTTACTTTGCATACTGATAAAAACCTTTTCACAGTACTTTTTCAACATGAAATTGAAGTTTGGAGGTAAGCAAAAGATGAAGTGG
1201   1211   1221   1231   1241   1251   1261   1271   1281   1291
ATCAGAGTGAACCATCCCGAATACTTTCATTGTTATTGCCGAGATCCCTATGTgtaagttttttttcattctgttttttactttctatggcacaaaa
1301   1311   1321   1331   1341   1351   1361   1371   1381   1391
tattaacctttattcatctctcaaggttcatactagctagttcggtcatgtggcaaacagtctcgaaaagtcgagctaagggtctttggccttttacaa
1401   1411   1421   1431   1441   1451   1461   1471   1481   1491
aaattacatatgtgcatgatctatactatttgattagaggcatatataaactcaaaatttaaataaataatatagcaaaccatttctctagcaat
1501   1511   1521   1531   1541   1551   1561   1571   1581   1591
cactttaatgctttaactcatacaaaatctcaatctctatgatttatttgctctatcactttcaatgttggtcgatgcactactatgtttttctc
1601   1611   1621   1631   1641   1651   1661   1671   1681   1691
gtacataatctctttgttcataatcattctttggtgatgatgCTCTTATGAATGGTAGAATCAGGGCTCCGTATCACGAGTAAGAGTGACCGAG
1701   1711   1721   1731   1741   1751   1761   1771   1781   1791
AAAAAGAGGACAAGATACAGCAGCAATTTTCAGTGTCAAAACCGAGCTATGTCATAGAGGCACCAAAGAAGCTTGTGGACGAGAGCATCCAGCTC
1801   1811   1821   1831   1841   1851   1861   1871   1881   1891
TCTTCAGACCTTTTGATTACCGAGCTTGTTACATCTATCACTCAGAAGCTGGTCGCAAAATTCAATATACTCTTCAAGCTTATTGTGCCGTCTCCGA
1901
AGCATAA

```

Figure 3 The AOP2 structural gene sequence. The sequence of the AOP2 structural gene in the Col, Ler, St, Pi and Cvi accessions is shown. The sequences from the five accessions are identical where the sequence is presented as a single line. Exons are shown in upper case letters and introns are shown in lower case letters. The sequence from the five accessions is shown separately in the regions where the three polymorphisms identified in exon 2 (5 bp, 18 bp and 9 bp) are located. Each sequencing reaction was performed three times on different PCR products to ensure accuracy of the sequence.



The Col/*Pi* sequence differs substantially from these other three 5' regulatory regions as it contains a 37 bp deletion and 45 bp and 3 bp insertions, as well as a number of single base pair changes throughout the sequence. The comparison between the *Cvi* and *Pi* sequences suggests that the major deletion/insertion differences in their 5' regulatory regions is not pivotal in determining the expression profile of *AOP2* since both these alleles are expressed at comparatively similar levels (Figure 2).

It has previously been reported that the *Ler* accession contains a complete inversion of the *AOP2* and *AOP3* structural genes that may have resulted in a promoter exchange. This would suggest that this regulatory region may be driving *AOP3* leaf expression in *Ler* [15]. Gene-specific primers were used to investigate the physical juxtaposition of the *AOP2/3* regulatory regions and the *AOP2/3* structural genes in *Ler*. The primers were designed to amplify a fragment extending around 300 bp upstream and 300 bp downstream from the initiating ATG codon. Both Col and *Pi* were included as controls. Amplifications using mixed primers were included in an attempt to detect any *AOP* promoter-CDS linkage

anomalies. The mixed *AOP2/3* primer combinations generated no fragments amongst these accessions (Figure 6). The *AOP3* promoter-*AOP3* CDS primer combination generated the expected fragments in the three accessions, whereas the *AOP2* combination generated a fragment in Col and *Pi* but not *Ler* despite repeated attempts. Sequencing of the *AOP3* *Ler* PCR fragment (data not shown) indicates that in *Ler* the proximal 289 nucleotides of *AOP3* regulatory region is linked to the *AOP3* CDS (extending 375 nucleotides into the *AOP3* structural gene). The sequence from the overlapping regulatory region fragments extend this linkage 2 kb upstream indicating that an inversion has not occurred within 2 kb upstream of the *AOP3* structural gene in the *Ler* accession used in this study.

The comparison between the regulatory regions upstream of the *AOP2* structural gene in the *AOP2*-expressing (*Cvi/Pi*) and *AOP2* non-expressing (*Ler/St*) accessions found only a small number of exclusive differences, comprised of single and double base pair substitutions and insertion/deletions, and one 4 bp deletion. The majority of these differences are clustered around

1 11 21 31 41 51 61 71 81
 AATACACGTTTCTCAGTTTGGTCTGAATCGGCAACTCGAAAAGCTCTTCCATGGCTTCAACAGAGTTCTTAAGCTTTACAGACAATA Co/P/L/S/
 AATACACGTTTCTCAGTTTGGTCTGAATCGGCAACTCGAAAAGCTCTTCCATGGCTTCAACAGAGTTCTTAAGCTTTACAGACAATA Cv

91 101 111 121 131 141 151 161 171
 TATGTCAAAGAAGCTTCGAAAACACCGAGTCTCTAGA.....AGCTTCCTGGTTT Co/P
 TATGTCAAAGAAGCTTCGAAAACACCGTAGTCTCTAGAGCTTTACGGACATCAGTTTCTACTTCATCCCACTTTGAGCTTCTCGTT L/S/Cv

181 191 201 211 221 231 241 251 261
 TAGGGTTGGTTAGAGAAATCGATAACCGGGTCCGGGAGCTGAAAAGATATAGAACGAGAGAAGAGAAATCTGAATCCATTATTTTGT Co/P
 TAGGGTTGGTTAGAGAAATCGATAACCGGGTCCGGGAGCTGAAAAGATATAGAACGAGAGAAGAGAAATCTGAATCCATTATTTTGT L/S/Cv

271 281 291 301 311 321 331 341 351
 TCTGAGAATTTGAGGTTTTTTTTGCTTTATGATTTGGCCTAATAATTAAGAATTTTGTATGGAAGAATTCGTAATTAGAAG Co/P/L/S/Cv

361 371 381 391 401 411 421 431 441
 TTAGTATATGCGTTACCGATGCTACGCTGCTTATATGCGTAGATGTTTATACATTAAATAAATATGCAATATTCGCGTTGATAAAGA Co/P/Cv
 TTAGTATATGCGTTACCGATGCTACGCTGCTTATATGCGTAGATGTTTATACATTAAATAAATATGCAATATTCGCGTTGATAAAGA L/S

451 461 471 481 491 501 511 521 531
 TTTATCTCTTTATCCAAATAAGAGTTAAAATATCAAAGAGTTAAGAGATAA.....AAAAAAGTTTGAATTCAGGAAAAGATAAAT Co/P
 GTTATCTCTTTATCCAAATAAGAGTTAAAATATCAAAGAGTTAAGAGATAA.....AAAAAAGTTTGAATTCAGGAAAAGATAAAT L/S
 GTTATCTCTTTATCCAAATAAGAGTTAAAATATCAAAGAGTTAAGAGATAA.....AAAAAAGTTTGAATTCAGGAAAAGATAAAT Cv

541 551 561 571 581 591 601 611 621
 ATATATAACCGTGTGATCATCTTAAAATAAATGTCTACATTTTACTTACCTAGCTAACCACTGTTACAACCCCATGTTGATAATGTCT Co/P/Cv
 ATATATAACCGTGTGATCATCTTAAAATAAATGTCTACATTTTACTTACCTAGCTAACCACTGTTACAACCCCATGTTGATAATGTCT L/S

631 641 651 661 671 681 691 701 711
 CTTGTATTACGGCATATGTTAAATTTAATAAACAATGAATAATTTCTTAAATTACATAATACAAGAGAA.....GAAAACTAATAAATGAC Co/P/L/S/Cv

721 731 741 751 761 771 781 791 801
 TGACTAT.AGACAAAATTTGTTAACTTTAAACAATGATTTTCAATTTTCCTTTTCCTTCTAGTTGAGATACCTCTTCAATTTCTCTG Co/P
 TGACTATAGAACAAAATTTGTTAACTTTAAACAATGATTTTCAATTTTCCTTTTCCTTCTAGTTGAGATACCTCTTCAATTTCTCTG L/S
 TGACTATAGAACAAAATTTGTTAACTTTAAACAATGATTTTCAATTTTCCTTTTCCTTCTAGTTGAGATACCTCTTCAATTTCTCTG Cv

811 821 831 841 851 861 871 881 891
 CAGGTATGTTATGTAATAATATGTTTAGACTTTTA.....TTAACAAAAGAAAAGAAAAGAACAGACATGAGCC...TGGGCTAAAAA Co/P
 CAGGTATGTTATGTAATAATATGTTTAGACTTTTA.....TTAACAAAAGAAAAGAAAAGAACAGACATGAGCC...TGGGCTAAAAA L/S
 CAGGTATGTTATGTAATAATATGTTTAGACTTTTA.....TTAACAAAAGAAAAGAAAAGAACAGACATGAGCC...TGGGCTAAAAA Cv

901 911 921 931 941 951 961 971 981
 GCAACATTAAGAATTCGGGTCAGTTCTAATATATAGATCAGCTCCTTTATGCGATCAGTTTAAAGCCGATCACTGTTCTTACAT Co/P/L/S/Cv

991 1001 1011 1021 1031 1041 1051 1061 1071
 TCCTAAGTCGGCCTTCTTCCATTTTCCACACGACATGAATTTGAGAAATGAAGCAAAACGCCAAAAGACATTTTATCTCTAATTTGTC Co/P/Cv
 TCCTAAGTCGGCCTTCTTCCACACGACATGAATTTGAGAAATGAAGCAAAACGCCAAAAGACATTTTATCTCTAATTTGTC L/S

1081 1091 1101 1111 1121 1131 1141 1151 1161
 TCATTTTATGATTTTTTGGGCTATTCAA.....AATATATATAGCAAAACGTACAAGTCAAAAACCTGATCGATCGTCTCGTATTCTGTTAT Co/P
 TCATTTTATGATTTTTTGGGCTATTCAA.....AATATATATAGCAAAACGTACAAGTCAAAAACCTGATCGATCGTCTCGTATTCTGTTAT L/S/Cv

1171 1181 1191 1201 1211 1221 1231 1241 1251
 ACATTTGTCATAA.....AAAACCTTAAACAATACCTAACTGATTAATCATCTTCCACCTGCAATCTTATTA.....TACAGAGACGACGTGTATGC Co/P
 ACATTTGTCATAA.....AAAACCTTAAACAATACCTAACTGATTAATCATCTTCCACCTGCAATCTTATTA.....TACAGAGACGACGTGTATGC L/S/Cv

1261 1271 1281 1291 1301 1311 1321 1331 1341
 AGCAAGTAAACAAAAGCTCTGATCCAAAGTAATGGCCAAATGATACAGTTGGATCCAAACCGATAAAGTAAAAAAGATTCTATCTTTA Co/P
 .GCAAGTAAACAAAAGCTCTGATCCAAAGTAATGGCCAAATGATACAGTTGGATCCAAACCGATAAAGTAAAAAAGATTCTATCTTTA L/S/Cv

1351 1361 1371 1381 1391 1401 1411 1421 1431
 AAATGGGACCTCTCTCC.....TTTTCCTCTCATTATTGCTATGTTCAAAACGTACAATCATCGGCTCGCTCGGAAATTTAAATCATATACAA Co/P/L/S/Cv

1441 1451 1461 1471 1481 1491 1501 1511 1521
 TATGCTTACTAAGTATTATTCACGTATAAGTAACAATATGCAACTAAAGTTTGGGCGACGAATTACTTACAATAAAAAAATTTTGTAA Co/P/L/S/Cv

1531 1541 1551 1561 1571 1581 1591 1601 1611
 ACTCAAGTCCAAATGTTAGTAAATAAATTCACATTTTATCTCAAAAAAAGAGAAAAGATAAAGATAAAGATAAAGATAAAGATAAAG Co/P
 ACTCAAGTCCAAATGTTAGTAAATAAATTCACATTTTATCTCAAAAAAAGAGAAAAGATAAAGATAAAGATAAAGATAAAGATAAAG L/S/Cv

1621 1631 1641 1651 1661 1671 1681 1691 1701
 TTAACCTTAAGTTATAACTCGTTGAGTTGTTCTTTCATCTGTACAACAATAATGCACACATCGATGGATTGATATTTTCATGATGAA Co/P/L/S1
 TTAACCTTAAGTTATAACTCGTTGAGTTGTTCTTTCATCTGTACAACAATAATGCACACATCGATGGATTGATATTTTCATGATGAA Cv

1711 1721 1731 1741 1751 1761 1771 1781 1791
 AATAAACGATTACCACCTTTCTTTTTTTGATATAAATGTGAACTTTAAATGATATACCACTTCCCTCTCTGTCTATTATTA Co/P/L/S/Cv

1801 1811 1821 1831 1841 1851 1861 1871 1881
 TTAAGCATGTGTTGCTTAAATTAAGCAAAGCGCAAAAAAAGCTTAAATACAATCACTTGTGAATAATTTCTTATAAATGGGGACC Co/P/L/S/Cv

1891 1901 1911 1921 1931 1941 1951 1961 1971
 CTTCACTATTCTTACTCACACAGAAGGAAAAATCTAGAGCTAGCAAAGTAAAAACAATTAATATAACAGAAAGCTAAAGGTAATTT Co/P
 CTTCACTATTCTTACTCACACAGAAGGAAAAATCTAGAGCTAGCAAAGTAAAAACAATTAATATAACAGAAAGCTAAAGGTAATTT L/S/Cv

1981 1991 2001 2011 2021 2031 2041 2051 2061
 TCTTATGCGTTTCAATGTTTTTTTCTTATAAATATTGCTACTACTCTGCTAACATCTGCAACCTTTTTCGATTACACCAAAAAGG Co/P/L/S/Cv

2071
AAAGAAGAATG Co/P/L/S/Cv

Figure 5 The AOP2 5' regulatory region sequence. The sequence of the AOP2 5' regulatory region in the Col, Ler, St, Pi and Cvi accessions is shown. The accession to which the sequence pertains is indicated at the end of sequence; Co = Col, P = Pi, L = Ler, S = St and Cv = Cvi. The sequence changes between the accessions are highlighted in gray. Each sequencing reaction was performed three times on different PCR products to ensure accuracy of the sequence. Putative light and diurnal response elements found in the AOP2 5' regulatory region are highlighted. Yellow; GATA/I box (GATAA) and GT box (GRWAAW where R = A/G and W = A/T); Blue; CBS (AAAAAATCT), ME (CCACAC) and TBX (AAACCCT); Green; FORC^A/PBX element (T/ATGGGC/ATGGGCC). The translation initiating ATG codon at the end of the sequence is in bold.

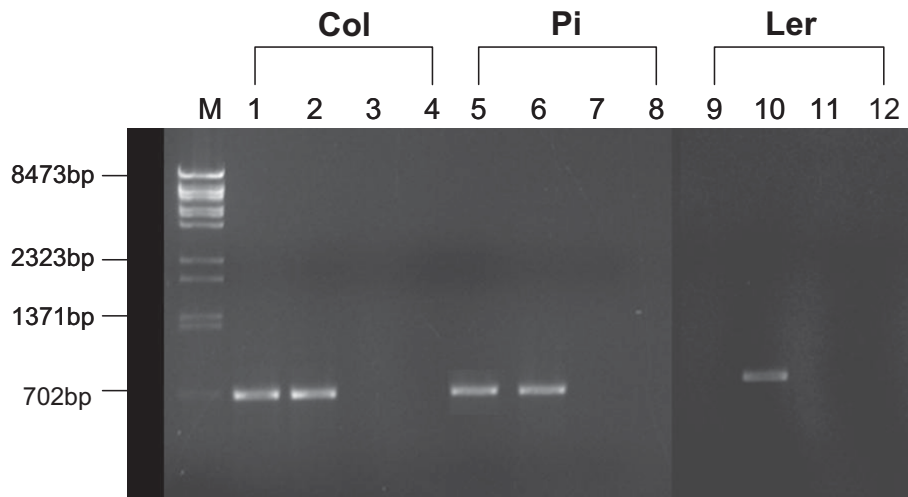


Figure 6 Gel electrophoresis of PCR-amplified fragments encompassing *AOP2* and *AOP3* promoter regions and the structural genes. PCR reactions, using genomic DNA from the Col, Pi and Ler as template, were performed with *AOP2* and *AOP3* gene-specific primers designed to hybridize at flanking positions in promoter regions and CDS regions approximately 300 bp distal from the ATG translation initiation codon. Lanes 1,5 & 9; products obtained from *AOP2* promoter and CDS primers, Lanes 2, 6 & 10; products obtained from *AOP3* promoter and CDS primers, Lanes 3, 7 & 11; products obtained from *AOP2* promoter and *AOP3* CDS primers, Lanes 4, 8 & 12; products obtained from *AOP3* promoter and *AOP2* CDS primers. Lane M contains λ /BstII size markers.

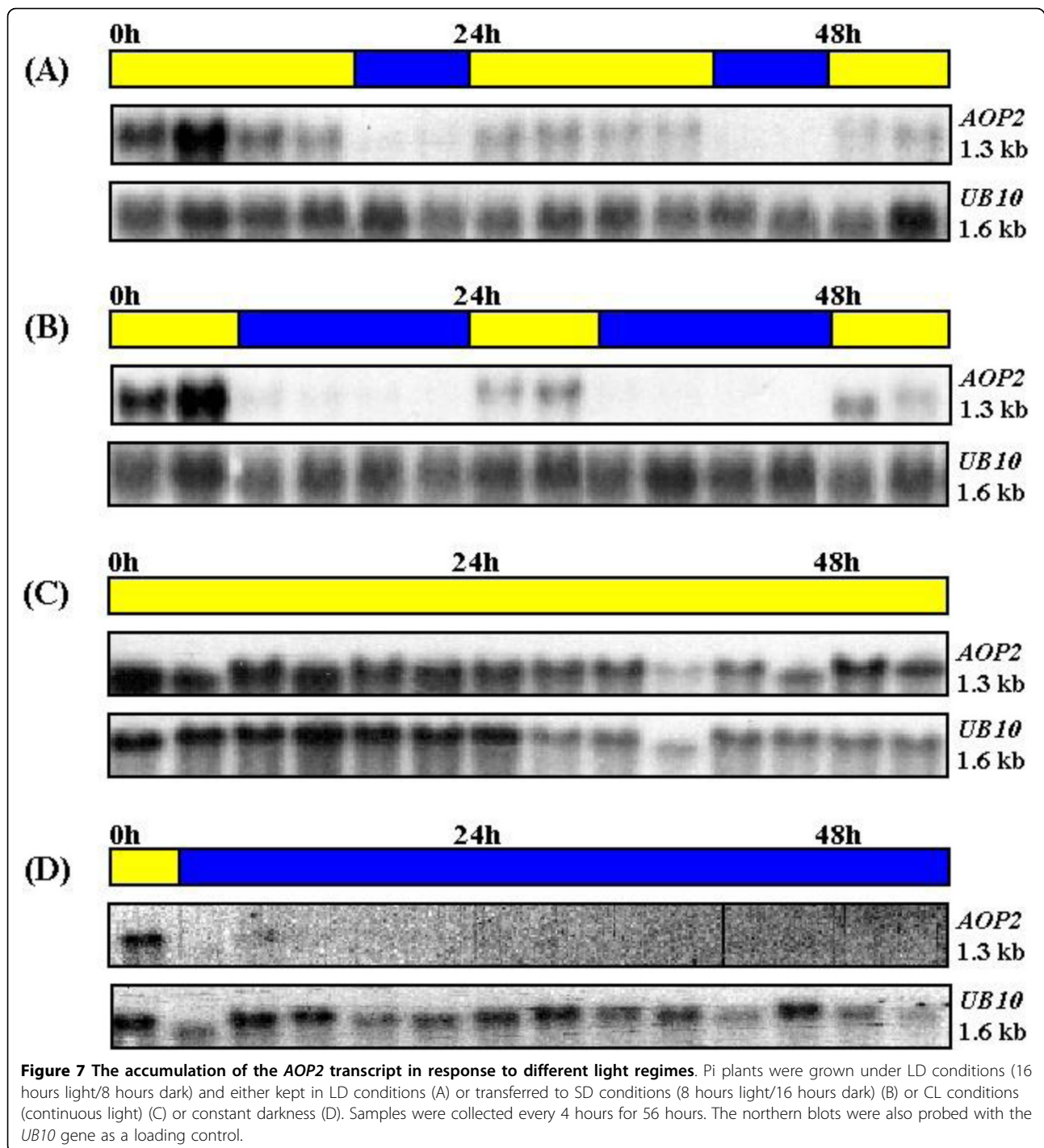
nucleotides 1184 to 1232 upstream of the ATG (Figure 5; position 847-897). No potential *cis*-acting elements (see below) were identified in this region and it is not clear that these minor changes would explain the substantial difference in the leaf level of *AOP2* transcript observed amongst these accessions.

***AOP2* expression is regulated by light**

Several light-responsive elements were identified in the *AOP2* 5' regulatory region (Figure 5) including GT-boxes and GATA/I-boxes [31-33]. The *AOP2* promoter also contains a CBS element that is found in the promoters of many clock-regulated genes that have a peak of expression around mid-day [34,35] although Harmer & Kay [36] suggest that CBS does not act alone to specify circadian phase. One ME and one TBX element [37] are also found in the *AOP2* promoter. These elements, along with flanking sequences are thought to specify the time of day phase and level of transcriptional activity through interactions with photocycles, thermocycles and the circadian clock [37]. Interestingly, the *AOP2* promoter contains a number of *FORC*^A elements which occur in *Arabidopsis* genes that are up-regulated in response to oomycete and fungal infection [38]. Similar elements (GATA, CBS, TBX & *FORC*^A) are found in the *Ler AOP3* promoter (data not shown). The positions of the elements in the *AOP2* promoter sequence are highlighted in Figure 5. Sequence polymorphisms between the *AOP2*-expressing accessions (Cvi/Pi) and non-expressing accessions (*Ler*/St) lead to the disruption of

only one putative GT box response element in the non-expressing accessions at position 767 (Figure 5).

An examination of the microarray data located at "The Diurnal Project" (<http://diurnal.cgrb.oregonstate.edu/>) shows *AOP2* exhibiting a very low level of expression in Col that peaks around midday when plants are grown in LD conditions at constant temperature [39]. However, this microarray-based *AOP2* expression pattern in Col is very similar in both amplitude and phase to that depicted for *AOP2* in *Ler*, which is generally accepted as being an *AOP2* non-expressor. This contrast with the *AOP3* expression pattern in *Ler* which shows a substantially higher diurnal expression with a maximum at ZT 12 under LD conditions. Since neither of the accessions utilized in "The Diurnal Project" express substantial levels of *AOP2* transcript, the accumulation of the *AOP2* transcript in Pi plants under various light conditions at a constant temperature was investigated. The Pi plants, which express *AOP2* but not *AOP3*, were initially grown under long day (LD) conditions (16 hours light: 8 hours dark) until they had reached the six-leaf stage. The plants were then transferred to either short day (SD) conditions (8 hours light: 16 hours dark), continuous light (CL) or remained under LD conditions. Samples were collected at 4-hour intervals and the transcript level of the *AOP2* gene was subsequently analysed by northern blot (Figure 7). The *AOP2* transcript was detectable in plants exposed to light, however it was found that when the plants entered a dark period the expression of *AOP2* decreased to an almost undetectable



level within four hours in both LD (Figure 7A) and SD (Figure 7B) conditions. When plants again entered a light period the transcript level returned to a detectable level within a few hours. Plants that were transferred from LD conditions to CL showed relatively constant expression of *AOP2* (Figure 7C) whereas no transcript was detectable at all in plants transferred to continuous darkness. Hence no evidence of *AOP2* transcript

oscillations in the absence of environmental stimuli was obtained.

Constitutive expression of the *AOP2-2* allele in the Col background alters the glucosinolate profile

Transgenic *Arabidopsis* lines were created in the Col accession. Col expresses low levels of *AOP2* (Figure 2) and no *AOP3* transcript. Absence or low levels of *AOP2*

or AOP3 expression should result in accumulation of methylsulfinylalkyl glucosinolate [11] in the Col accession, whereas the introduction of a functional AOP2 gene would be expected to convert methylsulfinylalkyl glucosinolate to the alkenyl glucosinolate form. The AOP2_{pi} gene (AOP2-2) from the Pi accession driven by the CaMV35S promoter was used to create a 35S:AOP2_{pi} line in the Col accession.

Ten independent transformants were obtained. Following confirmation of the presence of the construct by the amplification of the expected 699 bp PCR fragment, the transcript levels from the AOP2 gene within those lines was analysed. Northern blot analysis (Figure 8) shows the presence of AOP2 transcript in all the appropriate transgenic lines. RNA from an AOP2 expressing accession (Cvi) was included as a positive control (lane 1). The inclusion of wild-type Col RNA on the northern blot shows that the low expression of AOP2 was not detectable by northern analysis in untransformed Col plants. A degree of variability in the transcript levels of the AOP2 gene is observed in the transformed lines.

HPLC analysis of the transgenic lines was conducted to determine whether the introduction of AOP2 into the Col background alters the glucosinolate profile. Five T₃ generation homozygous AOP2 lines were obtained and used for HPLC analysis. The selected lines were 35S:AOP2_{pi}: 3, 4, 8, 9 and 10 from Figure 8. Figure 9(A) shows the typical HPLC result obtained for wild-type Col. The identities of the peaks were confirmed using standards and peaks 1 and 2 were found to correspond to the two precursor glucosinolates present in wild-type Col, 3-methylsulfinylpropyl (4.7 mins) and 4-methylsulfinylbutyl glucosinolates (9.9 mins), respectively. The HPLC analysis of Col accession was repeated several times with similar results and corresponds with the profile reported in the literature [11,12].

Figure 9(B) shows the typical HPLC results obtained for the 35S:AOP2_{pi} lines 3, 4, 8, 9 and 10. Evidence for

the expression of active AOP2 in the transgenic lines was provided by the appearance of two new peaks: 2-propenyl (12.2 mins; peak 3) and 3-butenyl glucosinolate (20.6 mins; peak 4), which were not observed in the wild-type Col control. Line 3 shows a high degree of conversion of the 4-methylsulfinylbutyl glucosinolate precursor to the 3-butenyl glucosinolate. Line 8 showed a similar result, lines 9 and 10 showed slightly less conversion and line 4 showed the least amount of conversion. Lines 3 and 8 also showed almost complete conversion of the 3-methylsulfinylpropyl glucosinolate precursor into 2-propenyl glucosinolate.

Discussion

Much of the genetic and phenotypic diversity evident in natural *Arabidopsis* accessions occurs as a result of drastic mutations, such as deletions and nucleotide changes resulting in the introduction of stop codons. In some cases these changes occur in several accessions which suggest they may be adaptive [40]. In the present study, three AOP2 alleles were identified based on the presence or absence of polymorphisms, which were all located in the second exon. When compared with AOP2-1 (the reference allele), the allele, termed AOP2-3, was found to contain one 18 bp and one 9 bp deletion polymorphism, in addition to the 5 bp deletion polymorphism previously identified [11]. The Col and *Ler* accessions were found to possess the AOP2-3 allele and AOP2 transcript was not detected in the leaves of either of these accessions. If the AOP2-3 allele is transcribed, the 5 bp mutation would cause the introduction of a series of stop codons in the transcript, which may have a negative effect on mRNA stability [41]. It should be noted that the *Ler* accession used in the previous study [11] was found not to possess the 5 bp polymorphism in exon 2, whereas the *Ler* used in the current study was found to possess the AOP2-3 allele containing the 5 bp deletion. The expression of AOP2 in the leaves was not detected

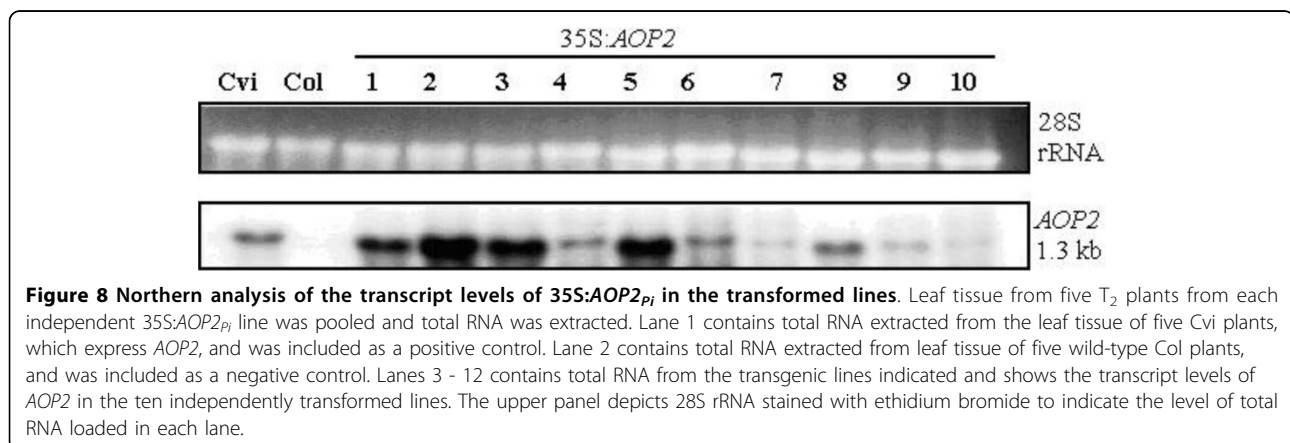


Figure 8 Northern analysis of the transcript levels of 35S:AOP2_{pi} in the transformed lines. Leaf tissue from five T₂ plants from each independent 35S:AOP2_{pi} line was pooled and total RNA was extracted. Lane 1 contains total RNA extracted from the leaf tissue of five Cvi plants, which express AOP2, and was included as a positive control. Lane 2 contains total RNA extracted from leaf tissue of five wild-type Col plants, and was included as a negative control. Lanes 3 - 12 contains total RNA from the transgenic lines indicated and shows the transcript levels of AOP2 in the ten independently transformed lines. The upper panel depicts 28S rRNA stained with ethidium bromide to indicate the level of total RNA loaded in each lane.

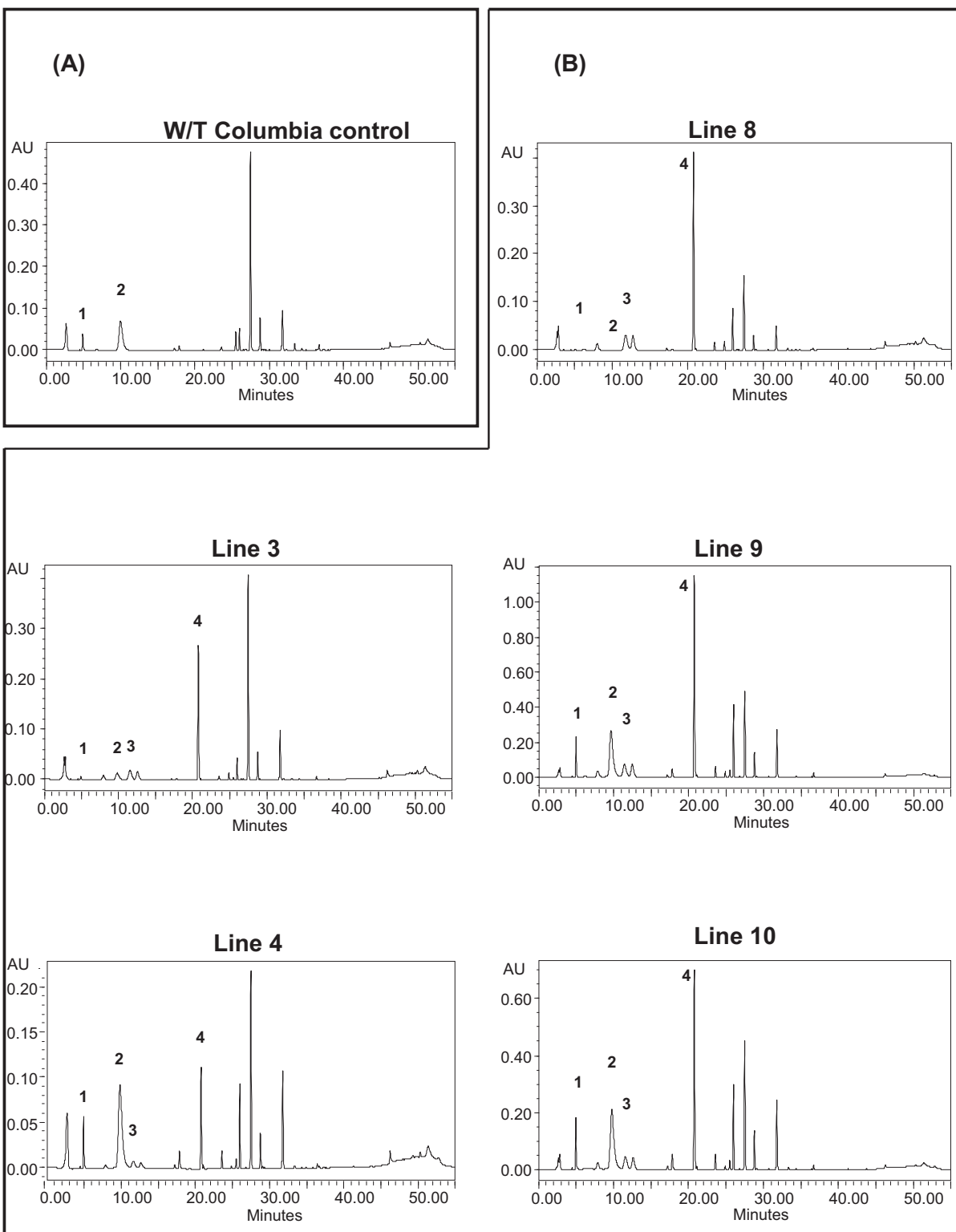


Figure 9 The typical HPLC profile obtained for wild-type Col (A) and the five *35S:AOP2* transgenic lines (3, 4, 8, 9 and 10) (B). **Panel A:** The typical HPLC profile of wild-type Col is shown. Peak 1 corresponds to the 3-methylsulfinylpropyl glucosinolate, and peak 2 corresponds to 4-methylsulfinylbutyl glucosinolate. **Panel B:** The HPLC profiles of transgenic lines 3, 4, 8, 8 and 10 are shown. Peak 1 corresponds to 3-methylsulfinylpropyl glucosinolate; peak 2 corresponds to 4-methylsulfinylbutyl glucosinolate, peak 3 corresponds to 2-propenyl glucosinolate peak 4 corresponds to 3-butenyl glucosinolate.

in the *Ler* accession used in either study [11]. In addition, the Pi-0 accession used in the previous study was found to express *AOP3* instead of *AOP2* [11], whereas the Pi accession used in the current study displays *AOP2* leaf expression exclusively. These discrepancies are most likely due to the use of different Pi and *Ler* strains. The Pi and *Ler* used in the present study are laboratory strains that have been selected over a number of years.

The *AOP2-2* allele that was found to be transcribed in our Pi accessions does not contain the 5 bp deletion but does contain the two in-frame 18 bp and 9 bp deletion polymorphisms. This would result in a predicted protein product nine amino acids shorter than *AOP2 1*. These polymorphisms occur in a region of the *AOP2* protein that contains several asparagine, valine and alanine rich repeats. Our results, in conjunction with previously reported results [11], indicate that both *AOP2-1* and *AOP2-2* encode functional protein products.

The *AOP2-1* allele was present in both Cvi and St but only expressed in Cvi and not the St accession. St expresses *AOP3* in the leaves (data not shown). As the St *AOP2-1* allele does not contain the 5 bp deletion, mechanisms other than mRNA instability arising from a frame shift mutation are likely to be responsible for the lack of *AOP2* expression in St.

The *AOP2* 5' regulatory region was examined for differences between the *AOP2*-expressing (Cvi/Pi) and *AOP2* non-expressing (*Ler*/St) accessions. No sequence differences in the *AOP2* 5' regulatory regions differentiating the *AOP2*-expressing and *AOP2* non-expressing accessions were detected within 1 kb of the ATG initiation codon amongst any of these accessions and only a few minor differences were found within the 1 kb analysed further upstream. In addition, none of the major primary sequence differences found in the *AOP2*-expressing accessions correlate with the level of *AOP2* transcripts. Hence it is likely that primary sequence differences found in the *AOP2* promoter regions do not account for the variation in *AOP2* transcript expression levels amongst the accessions analysed, although reporter gene experiments are needed to confirm this.

Differential *AOP2/3* expression in the leaves of *Arabidopsis* accessions has been reported as being associated with an *AOP2/3* structural gene inversion and *AOP2/3* promoter swap [15]. In the *Ler* accession, which expresses *AOP3* in the leaves, such an inversion would be expected to bring the *Ler AOP3* CDS under the control of the *AOP2* regulatory region. Our results suggest that an inversion has not occurred within the 2 kb upstream of *AOP3* in our *Ler* accession or within the 2 kb upstream of the *AOP2* gene in those accessions analysed that express *AOP2*. It should be noted that our results do not preclude an inversion/s having occurred

outside the regions analysed which may bring the *AOP* regulatory regions under flanking epigenetic influences that regulate the spatial expression of the *AOP* genes. Our inability to obtain a product using the *AOP2*-specific primers on *Ler* DNA may indicate a rearrangement has occurred in the *AOP2* region that affects primer hybridization or amplification. Faint larger bands in the *AOP2* primer *Ler* reaction as well as some bands originating from the use of mixed *AOP2/3* primer combinations on other *Arabidopsis* accessions occurred on occasion but these results were difficult to reproduce. Nevertheless the results raise the possibility that a variety of rearrangements may have occurred in the *AOP* region amongst *Arabidopsis* accessions. This warrants further investigation given the implications for glucosinolate production and plant-insect interactions. Other regulatory mechanism differences associated with transcription factor activity or differential mRNA stability may also need to be investigated in order to fully understand the regulation of *AOP* expression.

The spatial expression analysis showed high levels of *AOP2* transcript were present in the vegetative leaves of Pi. Although there is not necessarily a direct correlation between the site of gene expression and metabolite accumulation, this correlates with reports that aliphatic glucosinolates, in particular C₃ and C₄ alkenyl, hydroxyalkyl and methylsulfinylalkyl glucosinolates, accumulate to high levels in *Arabidopsis* vegetative leaf tissue [17,42].

It has been reported that *Arabidopsis* roots accumulate a similar glucosinolate profile to that of rosette leaves, except that they contain a greater proportion of indole glucosinolate [17]. This spatial and temporal analysis of glucosinolates was conducted in the Col-0 *Arabidopsis* accession, which contains non-functional alleles of *AOP2* and *AOP3* and thus, does not accumulate alkenyl or hydroxyalkyl glucosinolates [17]. In the current study the *AOP2* transcript level was analysed in the Pi and *Ler* accessions and was found to be at low to undetectable levels in the roots and we have been unable to detect root *AOP3* in either accession. Therefore it is possible that the lack of a substantial *AOP2* transcript level would correlate with a lack of alkenyl glucosinolates in *Arabidopsis* roots in general. It is also interesting to note that substantial *AOP2* expression is only observed in photosynthetic regions of the plant. Thus, it would be interesting to determine whether *AOP2* transcript is detectable in light grown roots, and if so, whether the change in gene expression would result in the expected alteration in glucosinolate composition. A study has recently been conducted using *Arabidopsis* roots to analyse phenylpropanoid accumulation. It was reported that monolignol glucoside and flavonoid accumulation are induced by light in *Arabidopsis* roots [43].

The response to light was found to require proteins involved in photomorphogenesis and was repressed by proteins involved in skotomorphogenesis. The light-induced accumulation of monolignol glucosides and flavonoids in root tissue was also found to involve changes in the level and cell specificity of phenylpropanoid gene expression [43].

As the levels of many secondary metabolites, such as alkaloids, phenylpropanoids and glucosinolates have been reported to be influenced by light [26,27,44,45] the effect of light on the level of *AOP2* transcript in the leaves was analysed in the Pi accession throughout a diurnal cycle and it was found that the levels of the transcript varied dramatically. High levels of transcript are observed in light periods, however during dark periods the transcript was found to decrease to an almost undetectable level (Figure 7). This suggests there is a light requirement for *AOP2* transcript accumulation. This is supported by our spatial analysis. Multiple GT- and GATA/I-boxes were identified in the *AOP2* 5' regulatory region (Figure 5) and these may be important in the light-mediated regulation of the gene. The *AOP2* 5' regulatory region was found to have three main regions containing five, three and two light responsive elements. A high degree of degeneracy has been reported for the GT-boxes [31,33]. Therefore, as there may be additional motifs within the outlined regions that contain GT-boxes that were not detected by the bioinformatics packages, functional analysis will need to be conducted to determine which regions are critical for the light-responsive regulation of *AOP2*. In *Ler*, the *AOP3* transcript also accumulates preferentially in light grown plants (data not shown) although the *AOP3* transcript does not drop as dramatically in the dark as does the *AOP2* transcripts in Pi plants.

The *FORC^A* elements found in the *AOP2* promoter are activators of oomycete- and fungal-pathogen response genes under continuous light, but acts as a repressor under long days and prolonged darkness ([38]. The same element has also been identified as a circadian and diurnal response element, PBX that occurs in many genes associated with protein synthesis [37]. Under hot day/cold night conditions the PBX element confers midnight expression, whereas under photocycles alone PBX confers midday expression. Evrard *et al.* [38] have proposed that these elements, which are conserved throughout the plant kingdom, participate in co-ordinating gene expression associated with photosynthesis, protein production and defence. If *AOP2* is primarily regulated by these elements, the imposition of thermocycles may phase shift the transcript accumulation cycles. A comparison of the diurnal microarray data for *AOP2* and *AOP3* reveals that imposition of thermocycles under long days changes the peak transcript

accumulation from around midday (*AOP2*; ZT = 8, *AOP3*; ZT = 12) to midnight (ZT = 24) suggesting that the *FORC^A* elements may play a role in regulating *AOP* expression.

The activity of the *Arabidopsis AOP2* gene has previously been investigated by the heterologous expression of the *AOP2* protein in *E. coli* and it was reported that the conversion of the precursor methylsulfinylalkyl glucosinolate to the alkenyl glucosinolate forms is controlled by the *AOP2* gene [11]. To demonstrate the activity of *AOP2 in planta*, ten independent lines over-expressing the *AOP2* gene from the Pi accession were created. PCR analysis, designed to amplify a junction fragment spanning the 35S promoter, confirmed that all of the lines contain the appropriate constructs. The northern blot analysis of the ten lines showed expression of the appropriate transgene in all of the lines. The variability in transgene transcript levels in this analysis may be explained by position effects [46], gene copy number [47] and the varying ratios of heterozygous and homozygous T₂ plants included in the pooled samples analysed.

The glucosinolate profiles in the five 35S:*AOP2*_{Pi} over-expressing T₃ lines analysed by HPLC showed that the *AOP2-2* allele catalysed the conversion of the two precursor methylsulfinylalkyl glucosinolates to 2-propenyl and 3-butenyl glucosinolates. In the majority of lines tested, the extent of conversion to 3-butenyl glucosinolate was substantial and provides the first demonstration of manipulation of the glucosinolate pathway utilising the activity of *Arabidopsis AOP2 in planta*. A previous study has also shown *in planta* modification of the glucosinolate profile in the Col *Arabidopsis* accession by the introduction of the *Brassica oleracea AOP2* homologue (*BoGLUCOSINOLATE-ALK*). It was found that *BoGLUCOSINOLATE-ALK* is able to convert more than 80% of the 4-methylsulfinylbutyl glucosinolate precursor into 3-butenyl glucosinolate, and the 3-methylsulfinylpropyl glucosinolate precursor into 2-propenyl glucosinolate [12].

Conclusions

An examination of the *AOP2* gene across several *Arabidopsis* accessions has revealed that the structural gene encoding *AOP2* has three alleles. While one of the *AOP2* alleles (*AOP2-3*) encodes a truncated protein which is unlikely to be functional, the other alleles (*AOP2-1* and *AOP2-2*) have now been demonstrated to be functional *in planta*. No major differences were identified in the sequence of the *AOP2* 5' regulatory region between the accessions. *AOP2* expression in leaf tissue is light regulated and numerous putative light response elements are found within the 5' regulatory region of the *AOP2* gene. Furthermore, *AOP2* transcript

was found to accumulate predominantly in photosynthetic regions of the plant, with little or no transcript present in root and seeds.

Methods

Plant material

The *Arabidopsis* accessions used in this study were either laboratory stocks: Columbia (Col), Landsberg *erecta* (Ler), Niederenz (Ni); Pitztal (Pi); San feliu 2 (Sf2); Stockholm (St) or obtained from the *Arabidopsis* Biological Resource Center (ABRC): C24 (CS906); Canary Islands (Can-0) (CS1064); Cape Verdi Islands (Cvi) (CS8580); Conlara (Cond) (CS6175); Dijon-I (Di-1) (CS1108); Landsberg (La-0) (CS1298); Rschew (Rsch-4) (CS1494); San feliu-1 (Sf1) (CS6855); Shakhara (Sha) (CS6180); Sorbo (CS931); Southport (Su-0) (CS1540).

Northern blot analysis

The *AOP2* cDNA (1.3 kb) was digested with the restriction enzymes *EcoRI* and *HindIII* to yield a 300 bp fragment that was used as a probe. The coding sequence of the *Arabidopsis UBIQUITIN10 (UB10)* gene was used to quantify gene expression. A *UB10* probe was obtained by amplification of the *UB10* gene from genomic *Arabidopsis* DNA using the primers (shown in 5'-3' orientation) GGAAAGACGATTACTCTTGAGG and GTCTTAACGAAAATCTGCATACC.

Plant growth conditions

Plants were grown in continuous light under gro-lux fluorescent lighting tubes (Sylvania), unless otherwise stated. For light regulation studies Pi plants were grown in LD conditions (16 hours light/8 hours dark) under fluorescent light in glasshouse (GH) conditions until the 6-leaf stage. The plants were then either left in LD conditions or transferred to SD (8 hours light/16 hours dark) or CL conditions (24 hours light) or constant darkness. The temperature was maintained at around 23°C under the different light conditions.

Amplification and analysis of the *AOP2* structural gene and 5' regulatory region

All primers were designed based on the published genome sequence of the Columbia *Arabidopsis* accession [48]. The structural *AOP2* gene was amplified using Pfu polymerase (Promega) using two sets of primers. The first amplification was performed using genomic *Arabidopsis* DNA as template with the following flanking primers (shown in 5'-3' orientation): *AOP2*: TATTGCTACTATCTGCTAACATCTGCAAC and TTCGTA-TACCAC ATGGTTTTTGATACTTG.

The second amplification was performed using 1 µl of the original reaction as template and the following nested primers (shown in 5'-3' orientation): *AOP2*:

ATAGAGCTCTAGGATCCGATTACACCAAAAAGG-GAAAGAAGAATG and TGAGAGCTCGATCTAGA GAAAGTCGGAATCAAGCGCACACATTG.

The 5' ends of the nested primer pair incorporate the restriction enzyme sites *BamHI* and *XbaI*, respectively. These sites were subsequently used to sub-clone the 1.9 kb fragment into the pBluescript and pART7 vectors. Internal sequencing of the *AOP2* gene was performed using the following primers (shown in 5'-3' orientation): *AOP2*: GAGCTTCAGGAATCAATTATGAAAA, GAC-TAGTATGAACCTTGAAGAATGAATAAA and CGGTAACGAGAATATCAGGTG.

The *AOP2* 5' regulatory region was amplified from *Arabidopsis* genomic DNA in two parts (distal and proximal) using Taq polymerase (Promega). Each fragment was amplified three times to identify random PCR errors. The primers are shown in 5'-3' orientation.

AOP2 (distal): AATACACGTTTCTCAGTTTGGTC TG and TTAGAGGAATAAAAATGTCTTTTGGC

AOP2 (proximal): GCCAAAAGACATTTTATTCTCT TAA and TCTTCTTTCCCTTTTGGTGTAATC.

Fragments were either sequenced directly as PCR products or were TA sub-cloned into pBluescript and sequenced using the T3 and T7 primers (Promega).

The following bioinformatics packages were utilised in order to identify consensus sequences of potential *cis*-acting elements in the 5' regulatory region of the *AOP2* gene: **PlantCARE** <http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>[49]; **PLACE** <http://www.dna.affrc.go.jp/PLACE/>[50,51] and **AGRIS** <http://arabidopsis.med.ohio-state.edu/>[52].

Amplification of fragments linking the *AOP2* and *AOP3* structural genes and 5' regulatory regions

The following *AOP2* and *AOP3* gene-specific primers were designed to hybridize at flanking positions in promoter regions and CDS regions approximately 300 bp distal from the ATG translation initiation codon. The primers are shown in 5'-3' orientation.

AOP2 promoter; ATATAAAATGTGAACTTTA ATAG

AOP2 CDS; TAAACAATAGTAAAAACATATGAG

AOP3 promoter; TATAGAGGTTTATTTTACTGT ATC

AOP3 CDS; TTTTGGTAACTTCAAAAGTAGAG

Preparation of transgenic plants

The *AOP2_{Pi}* gene was amplified using Pi genomic DNA as template and ligated into the *BamHI* and *XbaI* sites of the pART7 vector downstream of the CaMV35S promoter [53]. Following sequence analysis, the pART7: *AOP2_{Pi}* construct was digested with *NotI* and the fragment containing the CaMV35S promoter region linked to the *AOP2* gene was ligated into the binary vector

pMLBART [53] and subsequently transformed into *Agrobacterium tumefaciens* strain AGL1. The binary vector pMLBART contains an herbicide (BASTA) resistance gene to aid in the selection of transformed *Arabidopsis* lines. The floral dip method adapted from [54] was used to transform wild-type Col plants. The seeds were collected from these plants and sown in soil. When the seedlings reached the 4-leaf stage they were sprayed twice (three days apart) with the herbicide (0.1% (v/v) BASTA (active ingredient glufosinate-ammonium)/0.1% (v/v) Silwet L-77 mixture). The resulting BASTA resistant seedlings (T_1 generation) were allowed to self and the T_2 seedlings were sprayed with BASTA. Tissue from resistant seedlings was further characterised by PCR and northern blot analysis. Seed was also collected from individual BASTA resistant plants and placed under selection to identify homozygous transformed lines. The T_3 generation homozygous plants were subsequently used in HPLC analysis.

PCR analysis of transgenic plants

Primers were designed to detect the presence of the 35S:*AOP2_{Pi}* construct in transformed *Agrobacterium* and plants. The forward primer (5'-AATGTCAAAGATA-CAGTCTCAGAAG-3') was designed to hybridise with the 35S promoter and the reverse primers (5'-ATGTTCTCCAAAATATTAGCATTAC-3') was designed to hybridise to the *AOP2* coding region. The expected fragment size is 699 bp for the *AOP2* construct.

Glucosinolate extraction from plant tissue

Glucosinolates were extracted from T_3 plants according to the following protocol (A. Leong., personal communication). Plant tissue was freeze-dried and then ground to powder and 150 mg of the powder was then suspended in 5 ml of 70% methanol. If appropriate, glucosinolate standards were added, and the mixture was incubated at 70°C for 20 minutes. The samples were then centrifuged at 3000 rpm for 5 minutes and the supernatant added to a 1 ml column of DEAE-Sephadex A25. The column was then washed twice with deionised water and twice with 1 ml of 0.02 M sodium acetate. Sulfatase solution (75 μ l) (prepared as described by [55]) was then added to the column and left to stand overnight. Desulfonated glucosinolates were eluted in 1 ml aliquots of deionised water and analysed by HPLC.

HPLC analysis of glucosinolates

Glucosinolate samples were analyzed using a Waters (USA) 600E HPLC with a Waters 486 absorbance detector (229 nm). Samples (100 μ l) were separated using a Phenomenex Gemini C18 column (150 \times 4.6 mm i.d., 5 μ m particle size) operated at 1 ml/min at room temperature using the following separation gradient adapted from [56]. Solvent A: H₂O; Solvent B: MeCN: 1.5 - 5%

(v/v) B (6 min), 5 - 7% (v/v) B (2 min), 7 - 21% (v/v) B (10 min), 21-29% (v/v) B (5 min), 29 - 57% (v/v) B (14 min), followed by a cleaning cycle: 57 - 93% (v/v) B (2 min), 5 min hold, 93 - 1.5% (v/v) B (3 min), 6 min hold. Glucosinolate peaks were identified by spiking the samples with the appropriate standards. 2-propenyl glucosinolate was purchased from Sigma-Aldrich. 3-methylsulfinylpropyl glucosinolate, 4-methylsulfinylbutyl glucosinolate and 3-butenyl glucosinolate were kindly supplied by Michael Reichelt (Max Planck Institute for Chemical Ecology, Jena, Germany). The HPLC peaks were analyzed using the Millennium32 software.

Acknowledgements

We thank Dr Michael Reichelt (Max Planck Institute for Chemical Ecology, Jena) for supplying glucosinolate HPLC standards and Dr Audrey Leong (Department of Primary Industry, Victoria) for technical advice. Dale Fredericks and Calida Neal were supported by an Australian Postgraduate Award and Cara Griffiths by a Faculty of Science Dean's Postgraduate Research Scholarship.

Author details

¹School of Biological Sciences, Monash University, Melbourne, Victoria, 3800, Australia. ²Dept of Gastroenterology/Renal Unit, School of Medicine, Flinders University, Adelaide, South Australia, 5042, Australia. ³ARC Special Research Centre for Green Chemistry, Monash University, Melbourne, Victoria, 3800, Australia.

Authors' contributions

DF carried out the gene isolation and initial gene characterization and regulation studies. CN undertook further gene characterization, created the transgenic plants and performed the HPLC analysis. CG performed gene characterization and assisted with the manuscript preparation. AN initiated the experimental program and gave technical and intellectual guidance. All authors contributed to data analysis and interpretation. CN and AN drafted the manuscript. All authors read and approved the final manuscript.

Received: 18 September 2009 Accepted: 11 August 2010

Published: 11 August 2010

References

1. Halkier BA: **Glucosinolates**. *Naturally Occurring Glycosides: Chemistry, Distribution and Biological Properties* New York, John Wiley & sons Ltd/Jan R 1999, 193-223.
2. Mithen R: **Glucosinolates-biochemistry, genetics and biological activity**. *Plant Growth Regul* 2001, **34**:91-103.
3. Mikkelsen MD, Petersen BL, Olsen CE, Halkier BA: **Biosynthesis and metabolic engineering of glucosinolates**. *Amino Acids* 2002, **22**:279-295.
4. Grubb CD, Abel S: **Glucosinolate metabolism and its control**. *Trends Plant Sci* 2006, **11**:89-100.
5. Halkier BA, Du L: **The biosynthesis of glucosinolates**. *Trends Plant Sci* 1997, **2**:425-431.
6. Wittstock U, Halkier BA: **Glucosinolate research in the *Arabidopsis* era**. *Trends Plant Sci* 2002, **7**:263-270.
7. Giamoustaris A, Mithen R: **Genetics of aliphatic glucosinolates IV. Side chain modification in *Brassica oleracea***. *Theor Appl Genet* 1996, **93**:1006-1010.
8. Parkin I, Magrath R, Keith D, Sharpe A, Mithen R, Lydiate D: **Genetics of aliphatic glucosinolates. II. Hydroxylation of alkenyl glucosinolates in *Brassica napus***. *Heredity* 1994, **72**:594-598.
9. Mithen R, Clarke J, Lister C, Dean C: **Genetics of aliphatic glucosinolates. III. Side chain structure of aliphatic glucosinolates in *Arabidopsis thaliana***. *Heredity* 1995, **74**:210-215.
10. Hall C, McCallum D, Prescott A, Mithen R: **Biochemical genetics of glucosinolate modification in *Arabidopsis* and *Brassica***. *Theor Appl Genet* 2001, **102**:369-374.

11. Kliebenstein DJ, Lambrix VM, Reichelt M, Gershenzon J, Mitchell-Olds T: **Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in *Arabidopsis*.** *Plant Cell* 2001, **13**:681-693.
12. Li G, Quiros CF: **In planta side-chain glucosinolate modification in *Arabidopsis* by introduction of dioxygenase Brassica homolog BoGSL-ALK.** *Theor Appl Genet* 2003, **106**:1116-1121.
13. Fahey JW, Zhang Y, Talalay P: **Broccoli sprouts: an exceptionally rich source of inducers of enzymes that protect against chemical carcinogens.** *Proc Natl Acad Sci USA* 1997, **94**:10367-10372.
14. Wilson D: **Early-flowering mutants of a late-flowering accession of *Arabidopsis thaliana*.** *PhD thesis* Monash University, Melbourne 2001.
15. Chan EK, Rowe HC, Kliebenstein DJ: **Understanding the evolution of defense metabolites in *Arabidopsis thaliana* using Genome-Wide Association Mapping.** *Genetics in press*.
16. Wentzell AM, Rowe HC, Hansen BG, Ticconi C, Halkier BA, Kliebenstein DJ: **Linking metabolic QTLs with network and cis-eQTLs controlling biosynthetic pathways.** *PLoS Genetics* 2007, **3**:e162.
17. Brown PD, Tokuhisa JG, Reichelt M, Gershenzon J: **Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*.** *Phytochemistry* 2003, **62**:471-481.
18. Fahey JW, Zalcmann AT, Talalay P: **The chemical diversity and distribution of glucosinolates and isothiocyanates among plants.** *Phytochemistry* 2001, **56**:5-51.
19. Zangerl AR, Bazzaz FA: **Plant Resistance to herbivores and pathogens.** Chicago, University of Chicago Press 1993.
20. Josefsson E: **Content of *p*-hydroxybenzylglucosinolate in seed meals of *Sinapis alba* as affected by heredity, environment and seed part.** *J Sci Food Agri* 1970, **21**:94-97.
21. Kondra ZP, Downey RK: **Glucosinolate content of rapeseed (*Brassica napus* L and *B. campestris* L.) meal as influenced by pod position on the plant.** *Crop Sci* 1970, **10**:54-56.
22. Yan X, Chen S: **Regulation of plant glucosinolate metabolism.** *Planta* 2007, **226**:1343-1352.
23. Bible B, Chong C: **Correlations of temperature and rainfall with thiocyanate ion content in roots of radishes grown on two soil types.** *Hortic Sci* 1975, **10**:484-485.
24. Neil LJ, Bible B: **Thiocyanate ion (SCN-) content of hypocotyl-root region of *Raphanus sativus* as affected by environment.** *J Sci Food Agric* 1972, **23**:1379-1382.
25. Rosa EAS, Heaney RK, Portas CAM, Fenwick GR: **Changes in glucosinolate concentrations in *Brassica* crops (*B oleracea* and *B napus*) throughout growing seasons.** *J Sci Food Agric* 1996, **71**:237-244.
26. Rosa EAS, Heaney RK, Rego FC, Fenwick GR: **The variation of glucosinolate concentration during a single day in young plants of *Brassica oleracea* var *acephala* and *capitata*.** *J Sci Food Agric* 1994, **66**:457-463.
27. Rosa EAS: **Daily variation in glucosinolate concentrations in the leaves and roots of cabbage seedlings in two constant temperature regimes.** *J Sci Food Agric* 1997, **73**:364-368.
28. Bodnaryk RP: **Effects of wounding on glucosinolates in the cotyledons of oilseed rape and mustard.** *Phytochemistry* 1992, **31**:2671-2677.
29. Engelen-Eigles G, Holden G, Cohen JD, Gardner G: **The effect of temperature, photoperiod, and light quality on gluconasturiin concentration in watercress (*Nasturtium officinale* R. Br.).** *J Agr Food Chem* 2006, **54**:328-334.
30. Mikkelsen MD, Petersen BL, Glawischmig E, Jensen AB, Andreasson E, Halkier BA: **Modulation of CYP79 genes and glucosinolate profiles in *Arabidopsis* by defense signaling pathways.** *Plant Physiol* 2003, **131**:298-308.
31. Guilfoyle TJ: **The structure of plant gene promoters.** *Genetic engineering* New York, Plenum Press/Setlow JK 1997, 15-47.
32. Gourrierc JL, Li Y-F, Zhou D-X: **Transcriptional activation by *Arabidopsis* GT-1 may be through interaction with TFIIA-TBP-TATA complex.** *Plant J* 1999, **18**:663-668.
33. Zhou DX: **Regulatory mechanism of plant gene transcription by GT-elements and GT-factors.** *Trends Plant Sci* 1999, **4**:210-214.
34. Wang Z-Y, Kenigsbuch D, Sun L, Harel E, Ong MS, Tobin EM: **A Myb-related transcription factor is involved in the phytochrome regulation of an *Arabidopsis*.** *Lhcb gene* 1997, **9**:491-507.
35. Michael TP, McClung CR: **Phase-specific circadian clock regulatory elements in *Arabidopsis*.** 2002, **130**:627-638.
36. Harmer SL, Kay SA: **Positive and negative factors confer phase-specific circadian regulation of transcription in *Arabidopsis*.** *Plant Cell* 2005, **17**:1926-1940.
37. Michael TP, Mockler TC, Breton G, McEntee C, Byer A, Trout JD, Hazen SP, Shen R, Priest HD, Sullivan CM, Givan SA, Yanovsky M, Hong F, Kay SA, Chory J: **Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules.** *PLoS Genetics* 2008, **4**(2):e14.
38. Evrard A, Ndatimana T, Eulgem T: **FORCA, a promoter element that responds to crosstalk between defense and light signaling.** *BMC Plant Biology* 2009, **9**:2.
39. Mockler TC, Michael TP, Priest HD, Shen R, Sullivan CM, Givan SA, McEntee C, Kay SA, Chory J: **THE DIURNAL PROJECT: Diurnal and Circadian Expression Profiling, Model-Based Pattern Matching and Promoter Analysis.** *Cold Spring Harb Symp Quant Biol* 2007, **72**:353-363.
40. Weigel D, Mott R: **The 1001 genomes Project for *Arabidopsis thaliana*.** *Genome Biol* 2009, **10**:107.
41. Maderazo AB, Belk JP, He F, Jacobson A: **Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution.** *Mol Cell Biol* 2003, **23**:842-851.
42. Kliebenstein DJ, Gershenzon J, Mitchell-Olds T: **Comparative Quantitative Trait Loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seed.** *Genetics* 2001, **159**:359-370.
43. Hemm MR, Rider SD, Ogas J, Murry DJ, Chapple C: **Light induces phenylpropanoid metabolism in *Arabidopsis* roots.** *Plant J* 2004, **38**:765-778.
44. Mancinelli AL: **Light-dependent anthocyanin synthesis - a model system for the study of plant photomorphogenesis.** *Bot Rev* 1985, **51**:107-157.
45. Vazquez-Flota FA, Luca VD: **Jasmonate modulates development- and light-regulated alkaloid biosynthesis in *Catharanthus roseus*.** *Phytochemistry* 1998, **49**:395-402.
46. Matzke AJM, Matzke MA: **Position effects and epigenetic silencing of plant transgenes.** *Curr Opin Plant Biol* 1998, **1**:142-148.
47. Gelvin SB: **The introduction and expression of transgenes in plants.** *Curr Opin Biotech* 1998, **9**:227-232.
48. **Arabidopsis Genome Initiative: Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*.** *Nature* 2000, **408**:796-815.
49. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Peer Yvd, Rouzé P, Rombauts S: **PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences.** *Nucleic Acids Res* 2002, **30**:325-327.
50. Higa K, Ugawa Y, Iwamoto M, Korenaga T: **Plant cis-acting regulatory DNA elements (PLACE) database: 1999.** *Nucleic Acids Res* 1999, **27**:297-300.
51. Prestridge DS: **SIGNALSCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements.** *CABIOS* 1991, **7**:203-206.
52. Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Molina C, Kurtz M, Grotewold E: **AGRIS: Arabidopsis Gene Regulatory Information Server, an information resource of Arabidopsis cis-regulatory elements and transcription factors.** *BMC Bioinformatics* 2003, **4**:25.
53. Gleaves AP: **A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome.** *Plant Mol Biol* 1992, **20**:1203-1207.
54. Clough SJ, Bent AF: **Floral dip: simplified methods for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*.** *Plant J* 1998, **16**:735-743.
55. Graser G, Schneider B, Oldham NJ, Gershenzon J: **The methionine chain elongation pathway in the biosynthesis of glucosinolates in *Eruca sativa* (Brassicaceae).** *Arch Biochem Biophys* 2000, **378**:411-419.
56. Reichelt M, Brown PD, Schneider B, Oldham NJ, Stauber E, Tokuhisa J, Kliebenstein DJ, Mitchell-Olds T, Gershenzon J: **Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*.** *Phytochemistry* 2002, **59**:663-671.

doi:10.1186/1471-2229-10-170

Cite this article as: Neal et al.: The characterisation of AOP2: a gene associated with the biosynthesis of aliphatic alkenyl glucosinolates in *Arabidopsis thaliana*. *BMC Plant Biology* 2010 **10**:170.