

# Identification of the Gene Encoding Cinnamic Acid 4-Hydroxylase and its Function to After-Cooking Darkening in Potatoes

Sandra Babin<sup>1</sup> • Karthikeyan Narayanan<sup>1</sup> • Agnes Murphy<sup>2</sup> •  
Hielke De Jong<sup>2</sup> • Gefu Wang-Pruski<sup>1\*</sup>

<sup>1</sup> Nova Scotia Agricultural College, PO Box 550, Truro, Nova Scotia, B2N 5E3, Canada

<sup>2</sup> Potato Research Centre, Agriculture and Agri-Food Canada, 850 Lincoln Road, Fredericton, New Brunswick, E3B 4Z7, Canada

Corresponding author: \* gwangpruski@nsac.ca

## ABSTRACT

After-cooking darkening (ACD) is a gray-black discoloration of the potato tuber, formed after cooking by the oxidation of an iron-chlorogenic acid complex. The trait has a negative impact on the processing quality of both fresh and processed products. Currently, the French fry industry uses preventative chemicals to control its severity. Cinnamic acid 4-hydroxylase (C4H) is an essential enzyme for the biosynthesis of chlorogenic acid and therefore has a key function in the ACD mechanism. Previously, the *c4h* gene or its expression profile had not been characterized in potatoes. In this study, the full-length *c4h* gene was cloned and sequenced from both genomic DNA and cDNA of 'Russet Burbank' tuber tissue by PCR and 5' and 3' RACE. The gene expression of *c4h* was examined by Northern hybridization in selected diploid clones known for both low and high degrees of ACD, and in cultivars used in commercial French fry production. Gene expression analyses indicated that there was no significant difference in the expression of *c4h* between the commercial cultivars 'Russet Burbank' and 'Russet Norkotah', which showed a similar degree of ACD severity in the samples produced in the season. In contrast, the level of the *c4h* transcript was significantly higher in the diploid clones with severe ACD and lower in the diploid clones with resistance to ACD. These results correlated with the levels of ACD in these tuber samples. The results suggest a significant relationship between the levels of *c4h* gene expression and the degree of ACD in potato tubers. The gene can be used to design a DNA marker for guiding breeding of ACD-resistant cultivars.

**Keywords:** diploid clones, gene cloning, gene expression, tuber discoloration

## INTRODUCTION

After-cooking darkening (ACD) is a non-enzymatic gray-black discoloration of potato tuber flesh occurring after cooking (Wang-Pruski 2007). The discoloration is due to the formation of a colorless iron-chlorogenic acid complex during the cooking process, which upon exposure to air, oxidizes to form the dark ferridichlorogenic acid (Dale and Mackay 1994; Wang-Pruski and Nowak 2004). To prevent the discoloration caused by ACD, processors in the French fry industry treat French fried potato strips with sodium acid pyrophosphate (SAPP, Na<sub>2</sub>H<sub>2</sub>P<sub>7</sub>O<sub>7</sub>), which reduces darkening by complexing the iron in the tuber. In this capacity the iron is held in a nonionizable form and cannot take part in the reaction with chlorogenic acid (Smith 1987). A rise in the number of French fry processing factories has led to an increase in SAPP usage. The phosphorus residue released from SAPP during processing has made it mandatory to eliminate SAPP from industrial wastewater. This currently involves the removal of phosphorous from wastewater through chemical precipitation, adding costs for the French fry industry (Wang-Pruski and Nowak 2004). It would, therefore, be beneficial both from economical and environmental standpoints to reduce or eliminate the use of SAPP in the processing industry (Tarn *et al.* 1992). Currently, 'Russet Burbank' and 'Shepody' are the primary cultivars used in the French fry processing industry in Canada. Both cultivars require the use of SAPP to prevent darkening. To date, no cultivars are available that possess all the traits essential for French fry processing, as well as complete resistance to ACD (Wang-Pruski 2007).

Chlorogenic acid (CGA) content in potato tubers is

correlated with the degree of ACD (Wang-Pruski 2006). Higher ACD was found in tubers with a higher concentration of CGA. Since C4H is the key enzyme for CGA biosynthesis, its gene expression will determine the content of CGA in potato tubers. CGA is not only involved in ACD, but it also has various biological roles based on its antioxidant activity, including the involvement in defense mechanisms against insects or phytopathogens, disease and fungal resistance, growth regulation, and wound response (Kühl *et al.* 1987; Yao *et al.* 1995; Friedman 1997; Griffiths and Bain 1997). In potatoes, CGA is able to provide covalent cross-links between polysaccharides and cell wall proteins; making the cell wall stronger and more resistant to invading pathogens (Yao *et al.* 1995). CGA accounts for up to 90% of the total phenolic compounds present in the potato tuber (Griffiths and Bain 1997; Lewis *et al.* 1998; Lugasi *et al.* 1999; Percival and Baird 2000). CGA is synthesized via the phenylpropanoid pathway, which has not been explored in great detail in Solanaceae family-related species. Approximately 50% of CGA is located in the potato peel and adjoining tissues.

Cinnamic acid 4-hydroxylase (C4H, EC 1.14.13.11) catalyzes the hydroxylation of *t*-cinnamic acid to form *p*-coumaric acid, during CGA biosynthesis. The C4H enzyme belongs to the CYP73 family of plant cytochrome P450 proteins. C4H enzymatic activity is induced by wounding, light, and pathogen infection in various plant species (Tanaka *et al.* 1974; Fahrendorf and Dixon 1993; Bell-Lelong *et al.* 1997; Petersen 2003). Class I and II forms of the gene encoding C4H have been sequenced in many plant species, including *Arabidopsis*, Jerusalem artichoke, red pepper, pea, alfalfa, and species of *Populus* and *Citrus*.

Class I *c4h* is the predominate form found in almost all plant species, whereas the divergent class II form has only been isolated from orange and French bean. The divergent class II *c4h* has approximately 60% sequence similarity to the class I form and differs in the N-terminus and three internal domains (Betz *et al.* 2001; Blee *et al.* 2001).

The gene expression level of *c4h* depends on the specific plant species, tissue type, as well as stress and environmental factors (Whitbred and Schuler 2000). Bell-Lelong *et al.* (1997) and Mizutani *et al.* (1997) found that in *Arabidopsis*, *c4h* was expressed in all tissues analyzed including leaves, seedlings, stems, flowers, and roots. The higher levels were found in the stems and roots, possibly because of C4H's role in the production of the monolignols required for lignification. The *c4h* gene has not been sequenced nor has its expression profile been identified in potatoes. To date, no genes in potato have been identified that relate to the control of ACD. The present investigation relates to a DNA sequence that encodes cinnamic acid 4-hydroxylase (C4H) from potato. The full length genomic DNA and cDNA of the gene for the enzyme are identified. Further, its gene function at gene expression levels of this enzyme in the potatoes has been confirmed to correlate with the degree of darkness in potatoes.

## MATERIALS AND METHODS

### Potato tuber samples

Cultivar 'Russet Burbank' was used for cloning of the *c4h* gene. Both cultivars 'Russet Burbank' and 'Russet Norkotah' potato tubers were used for *c4h* gene expression analyses. They were grown in 2002 at the Nova Scotia Agricultural College research field in Brookside, Nova Scotia. Both 'Russet Burbank' and 'Russet Norkotah' are used as major processing varieties in North America but both show higher degrees of ACD. The tubers were initially stored at 15°C for 14 d at 95% relative humidity to promote suberization. The temperature was then gradually decreased over a one-month period to a final storage temperature of 9°C. Two diploid clones, CH72.03 and 10908.06, known for low and high degrees of ACD, respectively, were obtained from the Potato Research Centre, Agriculture and Agri-Food Canada in Fredericton, New Brunswick and stored at 9°C until needed. All tubers were stored in similar conditions and tested at the same time.

In addition, a group of diploid clones were used to analyze the relationship between *c4h* gene expression and the degree of ACD using relative quantitative reverse transcription PCR (RQRT-PCR) method (Burleigh 2001). These clones were three dark clones (151, 154, 224) and three light clones (70, 167, 231) from the diploid population 13610-T; and two dark clones (52, 55) and two light clones from the diploid population 13395-B. Population 13610-T was maintained in the same manner at the Nova Scotia Agricultural College research field. Population 13395-B was maintained at the Potato Research Centre. They were selected based on their degrees of ACD from two seasons of 2002 and 2003.

### Genomic DNA and total RNA isolation

Genomic DNA was isolated from 'Russet Burbank' potato tubers using the protocol of Doyle and Doyle (1990) with minor modifications. It was used to identify the genomic DNA sequence of the *c4h* gene. Total RNA was isolated from the 'Russet Burbank' tubers according to the slightly modified method of Bachem *et al.*

**Table 1** Designed primers and their DNA sequences for cloning of *c4h* by PCR<sup>a</sup>.

Primer ID	Primer sequence (5' to 3')	Tm (°C)
AF	5'-CCCCAGGTCCAATTCCA-3'	60.25
AR	5'-TTCAGGGGATACACAAACAG-3'	59.52
BF	5'-CTGTTGTGTCATCCCCTGAA-3'	59.52
BR	5'-CCTCATTTTCTCCAGTGCT-3'	59.28
CF	5'-GGCCTTCTTGAGGGGTAC-3'	59.94
CR	5'-CCTCGTTGATCTCTCCCTTCT-3'	59.83
DF	5'-GAAGGGAGAGATCAACGAGG-3'	58.82
DR	5'-TCACAGCCTGAAGGTATGG-3'	57.16
EF	5'-CCACTGGAAGAAACCTGAAG-3'	57.36
ER	5'-TTCTGCACCAAACGTCC-3'	56.43
FF	5'-AGCATTGGAGGAAGATGAGG-3'	59.24
FR	5'-GCCAATCTACTCTCTCAGCA-3'	59.59

<sup>a</sup>The first letter refers to the primer name (A to F) and the second letter indicates the direction of the primer, forward (F) or reverse (R).

(1996). It was used for the identification of the cDNA sequence of the *c4h* gene. When gene expression was analyzed, total RNA from 'Russet Burbank', 'Russet Norkotah' and the two diploid clones, 10908.06 and CH72.03, were isolated. In preparation for RNA isolation, selected potato tubers were peeled and the cortex region removed. The remaining tuber tissue was cut into 1-cm cubes and immediately frozen in liquid nitrogen. The frozen tissue was then ground to a fine powder in liquid nitrogen and stored at -80°C until the total RNA was extracted.

### PCR primers

Primer3 primer design software (Whitehead Institute for Biomedical Research, Cambridge, MA, USA, <http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) was used to construct the primer pairs based on basic primer design principles. As there is no sequence information available for *c4h* in potato, six pairs of degenerative primers (A-F) for use in PCR were designed based on highly conserved regions between the *c4h* cDNA of red pepper (*Capsicum annuum*) and *c4h* ESTs in tomato (*Lycopersicon esculentum*) (Table 1). The primers used in 5' and 3' RACE are shown in Table 2. The primers labeled as SP are sequence-specific primers designed against sequenced regions of the *c4h* gene in potato. The universal primers, T7 (5'-TAATAC GACTCACTATAGGG-3') and SP6 (5'-GATTTAGGTGACACT ATAG-3') (New England BioLabs Inc., [www.neb.com/nebecomm/products/productS1248.asp](http://www.neb.com/nebecomm/products/productS1248.asp)) were used for sequencing the pGEM<sup>®</sup>-T plasmid (Promega Corp., [www.promega.com](http://www.promega.com)).

### Touchdown PCR

Touchdown PCR was used to amplify the *c4h* gene from genomic DNA as the primers were designed based on sequences of another species and there were suspected mismatched nucleotides between the primer and the target sequences (Sambrook and Russell 2001). The PCR reaction mixture contained the following components in a final volume of 25 µL: 50 ng of genomic DNA, 2.5 µL of 10X PCR buffer, 1 µL of 10 mM dNTPs, 1.5 U FastStart *Taq* (Roche Applied Science, PQ, Canada), and 1 µL of each primer (0.01 mM). PCR cycle parameters were as follows: denaturation for 10 min at 95°C, followed by 30 cycles of denaturing for 20 sec at 94°C, annealing at 60°C for 30 sec but after the first cycle the annealing temperature was decreased by 0.5°C every cycle; thereafter, an extension at 65°C for 1 min, following which there was

**Table 2** Sequences of primers for cloning *c4h* by 5' and 3' RACE<sup>a</sup>.

Primer ID	Primer sequence (5' to 3')	Tm (°C)
SP1	5'-TTCCTCCAGTGCTCACCATAC-3'	60.13
SP2	5'-GGTATAGAAGTGGGAAGGGACA-3'	59.35
SP3	5'-CAGGGGATGACACAACAAC-3'	58.41
SP4	5'-AGAGGAGAAGCAGCTTGAGG-3'	59.60
Oligo dT-Anchor	5'-GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTTV-3' V = A, C, or G	n/a
PCR Anchor	5'-GACCACGCGTATCGATGTCGAC-3'	67.13

<sup>a</sup>Primers SP1, SP2, SP3, and SP4 were designed based on the *c4h* gene sequence in potato. The remaining primers were supplied with the 5'/3' RACE kit. SP - sequence specific primer.

another 30 cycles of 94°C for 20 sec, 45°C for 30 sec, and 65°C for 1 min. There was then a final extension for 7 min at 72°C. The extension time during the cycle was changed from 1 min to 2 min depending on the length of the amplicon expected. If the amplified region was expected to be more than 2 kb in length, an extension time of 2 min was used.

### Reverse transcriptase-PCR (RT-PCR)

The two-step reverse transcriptase-PCR (RT-PCR) was used to amplify the *c4h* coding region from the synthesized cDNA, as described in the Eppendorf cMaster RTplusPCR kit (VWR International, PQ, Canada). The first step of the two-step method was the synthesis of the first strand cDNA from total RNA. The second step of the Eppendorf cMaster RTplusPCR kit was the PCR amplification of the first strand cDNA. Primers at the 5' and 3' ends for RT-PCR reactions were designed based on the sequenced genomic DNA (5' primer: 5'-atgatctcttctactggag-3'; 3' primer: 5'-ggtttacacaacaacaac-3').

### 5' and 3' RACE

The protocol followed is as described in the 5'/3' RACE kit (Roche Applied Science, PQ, Canada). The first strand cDNA was purified using the High Pure PCR Product Purification kit (Roche Applied Science, PQ, Canada) to remove unincorporated nucleotides and primers as per manufacturer's instructions. The addition of a homopolymeric A-tail to the 5' end of the cDNA was carried out by the enzyme terminal transferase (provided by the kit).

For rare mRNA such as for *c4h*, a second round of PCR with nested primers was required to obtain a visible PCR product. The nested sequence-specific primer (SP3) was designed 75 bp within the previously amplified region, while the reverse nested primer was complementary to the Oligo dT-anchor primer (provided with the 5'/3' RACE kit). As the concentration of the amplified dA-tailed cDNA product was unknown, the nested PCR was performed on both undiluted and diluted (1:20 in water) amplified product. The PCR conditions for amplification were identical to the conditions used in the first round of amplification.

The method of 3' RACE (Rapid Amplification of cDNA Ends) takes advantage of the naturally occurring poly(A) tail of mRNA to amplify the 3' end of a gene. The 5'/3' RACE kit (Roche Applied Science, PQ, Canada) was used to synthesize first strand cDNA and amplify the 3' end using both sequence-specific primers and supplied primers.

### DNA sequencing and data alignment

PCR products were subcloned into plasmid vector pGEM<sup>®</sup>-T (Promega Corp., WI, USA) according to the manufacturer's instructions, and subsequently stored in *E. coli* DH5 $\alpha$  strain (Inoue *et al.* 1990; Sambrook and Russell 2001). All DNA sequencing of the plasmid DNA and PCR products was performed by DNA Landmarks, Inc. (PQ, Canada). The universal primers, T7 and SP6 were used for sequencing of the pGEM<sup>®</sup>-T plasmid DNA constructs. Raw sequence chromatograms were visually edited using the software program Chromas (<http://www.technelysium.com.au>). Alignment of the sequence data was conducted using the BLAST program available from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>) (Altschul *et al.* 1990).

### Northern hybridization

Four individual samples, 'Russet Burbank', 'Russet Norkotah', a light diploid clone (CH72.03), and a dark diploid clone (10908.06), were used for this analysis. The probe used for Northern hybridization was a 550-bp cDNA fragment previously amplified using the primers AF and FR (Table 1) from the 'Russet Burbank' total RNA. The fragment was subcloned into the pGEM<sup>®</sup>-T vector as described above. The QIAfilter Plasmid Maxi kit (Qiagen, ON, Canada) was used to isolate plasmid DNA from transformed *E. coli* cells as per the manufacturer's instructions. The restriction enzyme *Pvu*II was used to release the insert from the vector and amplified using the T7 and SP6 primers using standard PCR conditions. The amplified product was purified using the Millipore

Ultrafree-DA filter kit and the probe was radioactively labeled using the DECAprime<sup>™</sup> II Random Priming DNA Labeling kit (Ambion, TX, USA). The positive control was the *c4h* probe that was transcribed to RNA (Sambrook and Russell 2001) and loaded to the denaturing gel along with the samples. Any DNA remaining in the samples was degraded by incubating for 15 min at 37°C with 20 U of DNase I. Denaturing gel electrophoresis and membrane transfer of the RNA samples (30  $\mu$ g total RNA/sample) was performed according to standard protocols (Sambrook and Russell 2001). Prehybridization of the membrane was performed in 6 mL of Ultrahyb hybridization buffer (Ambion, Inc., TX, USA) for 1 h at 42°C. The buffer was replaced with 6 mL of new Ultrahyb solution containing the denatured probe. Hybridization of the membrane was performed overnight at 42°C. Following hybridization, the membrane was gently rinsed in 5  $\times$  SSC at room temperature, then washed 2  $\times$  for 5 min each at 37°C in 5  $\times$  SSC with 0.1% SDS, and lastly in 1  $\times$  SSC with 0.1% SDS 2  $\times$  for 15 min each at 37°C. The membrane was exposed to Kodak BioMax XAR film (Fisher Scientific, ON, Canada) 5 to 10 d prior to developing. The film was developed using Kodak GBX Developer and Fixer (Fisher Scientific, ON, Canada) as per manufacturer's instructions.

Total RNAs from the four samples, 'Russet Burbank', 'Russet Norkotah', the light clone (CH72.03) and the dark clone (10908.06) were extracted. To ensure that the pattern and intensity of the hybridization signal measured for each of the four tuber samples was reproducible, the Northern hybridization experiment was replicated three times using identical conditions and the signals from the transcripts were digitally measured using the UVP Chemi Imager (UVP, USA). The intensity of the *c4h* transcript for each potato sample was measured relative to the 'Russet Burbank' signal, which was assigned a value of 1.0, using Labworks Imaging and Acquisition Software.

### Relative quantitative RT-PCR

The RQRT-PCR method was based on Burleigh (2001). Single-stranded cDNA was synthesized from the total RNA isolated from four tubers of each clone. cDNAs from each clone type (dark or light) was pooled and used as the template for relative quantitative RT-PCR. The PCR cycle number of 27 was used for all comparisons because at this cycle number, the abundance of the PCR products showed differences. The 18s rRNA primer to competitor ratio were tested and found best to be 3:7 for *c4h*. Each relative quantitative RT-PCR mix contained 1 U of Master Taq polymerase (Eppendorf, Brinkmann instruments inc., Canada), 2  $\mu$ l of the single-stranded cDNA template, 2  $\mu$ l of PCR reaction buffer (10X) containing 100 mM Tris-HCl, pH 8.3; 15 mM MgCl<sub>2</sub>; 500 mM KCl; 1% Triton X-100, 2  $\mu$ l Taqmaster (5X PCR enhancer), 25  $\mu$ M of each dNTP, 0.5  $\mu$ M of the upstream and downstream gene specific primers, 1.6  $\mu$ l of optimized 18s rRNA primer and competitor mix (Ambion Inc. TX, USA.). The final volume was made up to 20  $\mu$ l. PCR was done with initial denaturation at 95°C for 2 min, followed by optimized cycles of denaturation at 95°C for 45 sec, then annealing for 45 sec, extension at 72°C for 45 sec and a final extension at 72°C for 7 min. A 5  $\mu$ l of the amplified PCR sample was loaded on to 1.2% agarose gel and electrophoresed for 45 min at 10V cm<sup>-1</sup>.

Quantification of differential gene expression in RQRT-PCR was based on the intensities of the bands of PCR products measured using the Labworks<sup>™</sup> software (UVP Inc., Upland, CA, USA) and expressed at pixel levels as mean raw densities (MRD). MRD were measured for the two bands visualized for each PCR, representing the candidate gene-specific fragment and the 315 bp 18s rRNA internal standard (Fig. 4). The MRD readings of the candidate gene-specific bands were normalized using the 18s rRNA MRD readings. The level of gene expression was calculated using the formula given below. The RQ RT-PCR for each gene was replicated four times and the statistical analyses were performed using One-way ANOVA and the significance level was set at  $P < 0.05$ . All statistical analyses were performed using Minitab<sup>®</sup> (Minitab Software 2004).

Gene expression level =

$$\frac{\text{MRD of } c4h \text{ gene}}{\text{MRD of internal standard}} \times \text{X ratio of the competitors}$$

## Evaluation of ACD in cooked potato tubers

ACD levels of the tuber samples were measured at the same time period before the tuber samples were frozen for subsequent hybridization experiments. The level of the darkening was digitally measured based on the methods described by Wang-Pruski (2006), using a gray scale of 256 pixel units, where white has a pixel density of 255.999 and black a pixel density of 0.001. Therefore, a higher pixel density was indicative of a lighter tuber while a lower pixel density corresponded to a darker tuber. The images of tuber surfaces for each sample were captured and digitally evaluated using the UVP Chemi-Imager System and LabWorks Imaging Analysis and Acquisition Software. Four tubers for each tetraploid cultivar and two tubers for each diploid clone were evaluated due to a limited number of tubers available.

The ACD measurements for each tuber cultivar/clone were analyzed by a one-way analysis of variance using the Proc GLM (General Linear Model) procedure of SAS (Version 8, SAS Institute, NC, USA). Significance at the 5% level ( $P$ -value < 0.05) was further examined using Tukey's honestly significant difference (hsd) test ( $\alpha = 0.05$ ) to compare the mean pixel densities. Normality and constant variance were tested using the Proc Univariate procedure of SAS using the predicted and residual values. The data proved to be normal without transformation.

## RESULTS AND DISCUSSION

### The *c4h* gene in the potato genome

Analysis of the sequencing data generated from the genomic DNA and cDNA clones led to the identification of the 5'- and 3'-UTR, three exons, and two introns from the 2986 bp DNA sequence of the potato genome. The complete *c4h* gene sequence is shown in **Table 3**. The coding region of the gene is 1518 bp in length, starting at the ATG nucleotides at positions 45 to 47 and stopping at nucleotides TAA at positions 2861 to 2863. The coding region contains three exons (shown in bold) and two introns (not bold). Exon 1 is 785 bp in length (from nucleotide 45 to 829), exon 2 is 134 bp in length (from nucleotide 1363 to 1496), while exon 3 is 599 bp in length (from nucleotide 2265 to 2863). Intron 1 (between nucleotides 785 and 786) is 533 bp long and intron 2 (between nucleotides 919 and 920) is 768 bp long, respectively. The partially sequences of 5'- and 3'-UTR

measured 44 (positions 1 to 44) and 123 bp (position 2864 to 2986) in length, respectively (**Table 3**).

The C4H protein sequence is composed of 505 amino acids (not including the stop codon), as shown in **Table 4**. The typical start (ATG) and stop (TAA) codons are found at the beginning and end of the open reading frame. The splice sites for introns 1 and 2 are depicted by boxes in **Table 4**. Intron 1 is found between the second and third nucleotides of the codon (AA/G), which corresponds to the amino acid lysine (amino acid position 262). Intron 2 is positioned between the first and second nucleotides of the codon (G/CA), which codes for the amino acid alanine (amino acid position 307).

The similarity of class I *c4h* coding sequences from other plant species to the potato *c4h* coding sequence is shown in **Table 5**. The nucleotide sequence of red pepper is the most similar to potato at 91%, which was anticipated as both belong to the Solanaceae family. The sequence similarity of the remaining plant species to the potato *c4h* coding sequence was range from 82 to 67%. The *c4h* coding region in potato is 1518 bp in length, which is identical to the length of most other class I *c4h* coding regions (**Table 5**).

In the potato, the three individual exons are 785, 134, and 599 bp which together make up the complete *c4h* coding region. The lengths of the *c4h* gene have only been cloned from genomic DNA in *Arabidopsis*, pea (*Pisum sativum*), Japanese aspen (*Populus sieboldii*), and French bean (*Phaseolus vulgaris*) (Kawai *et al.* 1996; Bell-Lelong *et al.* 1997; Nedelkina *et al.* 1999; Whitbred and Schuler 2000). The first three are class I *c4h* genes, whereas the *c4h* gene from French bean is class II. **Fig. 1** depicts the positions and lengths of the exons and introns in the four plant species as well as in the potato. The similarity between the exon lengths of potato, *Arabidopsis*, pea, and Japanese aspen would suggest that there is a conserved splice position in all class I *c4h* genes. When compared, it was found that the nucleotides flanking the splice site positions for both introns were identical in all four plant species (**Fig. 1**). The intron sizes appear to relate to the complexity and size of the corresponding plant genome. Evidence to this effect is demonstrated by the pea *c4h* gene, which has one of the largest introns (1726 bp).

The C4H protein consists of 505 amino acids, excluding the stop codon (**Table 4**). A comparison of the potato C4H

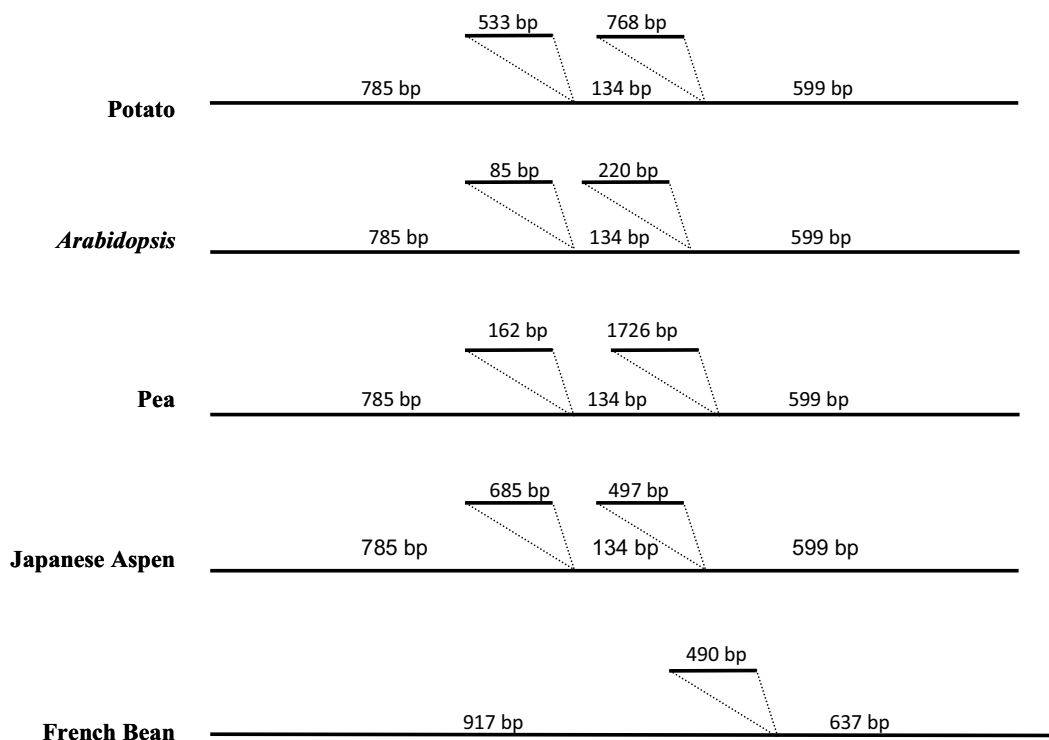


Fig. 1 Comparison of exon and intron lengths of the potato *c4h* gene to that of other plant species. Diagram not to scale.

**Table 3** DNA sequence of the *c4h* gene in the genome of 'Russet Burbank' potato cultivar. Exons are indicated in shaded areas.

1	AAACATTCTT	TTCTCAAAC	TCCCTCTGAA	AGAACTCACC	AAAAATGGAT	5'-UTR
51	<b>CTTCTCTTAC</b>	<b>TGGAGAAGAC</b>	<b>CTTAATAGGT</b>	<b>CTTTCTTTG</b>	<b>CTATTTAAT</b>	
101	<b>CGCTATTATT</b>	<b>GTCTCTAAAC</b>	<b>TTCGTTCCAA</b>	<b>GCGATTAAA</b>	<b>CTACCCCCAG</b>	
151	<b>GTCCAATTCC</b>	<b>AGTCCCAGTT</b>	<b>TTTGAAATT</b>	<b>GGCTTCAAGT</b>	<b>TGGTGATGAT</b>	
201	<b>TTGAACCATA</b>	<b>GAACCTTAC</b>	<b>TGAGTATGCT</b>	<b>AAAAAGTTG</b>	<b>GTGATGTGTT</b>	
251	<b>CTTGCTTAGA</b>	<b>ATGGGGCAA</b>	<b>GGAACCTAGT</b>	<b>TGTTGTGTCA</b>	<b>TCCCCTGAAT</b>	
301	<b>TAGCTAAAGA</b>	<b>AGTTTTACAC</b>	<b>ACACAAGGGG</b>	<b>TTGAATTTGG</b>	<b>TTCAAGAACA</b>	
351	<b>AGAAATGTTG</b>	<b>TTTTTGATAT</b>	<b>TTTTACAGG</b>	<b>AAGGGTCAAG</b>	<b>ATATGGTTTT</b>	
401	<b>TACAGTGTAT</b>	<b>GGTGACACT</b>	<b>GGAGAAAAT</b>	<b>GAGGAGGATT</b>	<b>ATGACTGTAC</b>	Exon 1
451	<b>CCTTTTTTAC</b>	<b>TAATAAGGTG</b>	<b>GTGCAGCAGT</b>	<b>ATAGAGGGGG</b>	<b>GTGGGAGTCT</b>	
501	<b>GAGGCTGCTA</b>	<b>GTGTAGTTGA</b>	<b>GGATGTGAAG</b>	<b>AAAAACCCTG</b>	<b>AATCTGCTAC</b>	
551	<b>AAATGGGATT</b>	<b>GTTTTGAGGA</b>	<b>AAAGATTGCA</b>	<b>GCTTATGATG</b>	<b>TATAATAACA</b>	
601	<b>TGTTTAGGAT</b>	<b>TATGTTTGAT</b>	<b>AGGAGATTG</b>	<b>AGAGTGAAGA</b>	<b>TGATCCCCTT</b>	
651	<b>TTTGTTAAGC</b>	<b>TTAGGGCTTT</b>	<b>GAATGGTGAG</b>	<b>AGGAGTAGAT</b>	<b>TGGCTCAGAG</b>	
701	<b>CTTTGAGTAC</b>	<b>AACATATGGT</b>	<b>ATTTTATCCC</b>	<b>TATTTTGAGG</b>	<b>CCTTCTTGA</b>	
751	<b>GAGGGTACTT</b>	<b>GAAGATTGT</b>	<b>AAGGAGGTTA</b>	<b>AGGAGAAGAG</b>	<b>GTTGAAGCTA</b>	
801	<b>TTCAAAGACT</b>	<b>ACTTTGTTGA</b>	<b>TGAAAGAAA</b>	<b>TAAGTTCAC</b>	<b>TTTTTCTTGT</b>	
851	<b>TAATCCCTTT</b>	<b>ATGCTCAATT</b>	<b>TGATCATTTG</b>	<b>TATCAGTTTT</b>	<b>ATTTATAGT</b>	
901	<b>TTAGTTTAGT</b>	<b>TGTAAGGGGT</b>	<b>GTTTGACTAA</b>	<b>ATCTTGGAAC</b>	<b>AGTATGGATC</b>	
951	<b>AATTTTGAAT</b>	<b>AGAAAAGGAA</b>	<b>GTACTAGTTG</b>	<b>ACATTCAGA</b>	<b>ATAGTAAGGG</b>	
1001	<b>TCCATTGGTT</b>	<b>AAATTTTAAA</b>	<b>AAAGGTAGTT</b>	<b>CTTGTTTTCT</b>	<b>GTTTTCAAAG</b>	
1051	<b>TGATAATGAA</b>	<b>AATTAGCGTG</b>	<b>GTGTTTGGCA</b>	<b>TATTTGGAGT</b>	<b>TGTTTTGCGA</b>	Intron 1
1101	<b>TTCTCCTGTG</b>	<b>GCAATTAGAG</b>	<b>GTTTGTGTA</b>	<b>ACGGTGGCCT</b>	<b>GTGAGAGCCT</b>	
1151	<b>AGCTTGCAGT</b>	<b>GGTAAGAGTA</b>	<b>GTGAGTGATT</b>	<b>TGGAGTAAAA</b>	<b>AAGTTAATAA</b>	
1201	<b>CTTTTTGATT</b>	<b>GATGTTTTTT</b>	<b>AAATTTTATG</b>	<b>TTGAATTCCA</b>	<b>GAATTGGCCA</b>	
1251	<b>ATAAGAATCA</b>	<b>TGTATGATTT</b>	<b>AGTGATAGTT</b>	<b>AAAGTGTCTT</b>	<b>GAGGTACTGT</b>	
1301	<b>TAGGTAGCTT</b>	<b>TTGATGGTGG</b>	<b>ACCTTGTGTT</b>	<b>TTAGTTTGTG</b>	<b>ATATTTTAT</b>	
1351	<b>TGCTTTACAC</b>	<b>AGGAAGCTTG</b>	<b>CAAATACCA</b>	<b>GAGCATGGAC</b>	<b>AGCAATGCTC</b>	
1401	<b>TAAATGTGTC</b>	<b>AATGATCAC</b>	<b>ATTCCTGAAG</b>	<b>CTCAACAGAA</b>	<b>GGGAGAGATC</b>	Exon 2
1451	<b>AACGAGGATA</b>	<b>ACGTTCTTTA</b>	<b>CATCGTTGAG</b>	<b>AACATCAATG</b>	<b>TTGCTGGTAT</b>	
1501	<b>GTTTCGAAAT</b>	<b>AACATATCTT</b>	<b>TGATTCTCTA</b>	<b>GAGTAAAATT</b>	<b>TGTTCTAGTT</b>	
1551	<b>TGGTTTAAAT</b>	<b>GATTGCATCC</b>	<b>TAGTTAGAAT</b>	<b>AAAAGTAATT</b>	<b>TATAAGTGAA</b>	
1601	<b>TGAAAATCCA</b>	<b>ATTCCAATTT</b>	<b>TGTCTATTTT</b>	<b>TCTCAAAAAGT</b>	<b>AGTAGTTGAG</b>	
1651	<b>AGTTACCAAA</b>	<b>TAAGGGGCC</b>	<b>AAGATTTAAC</b>	<b>TGTTTTTTAT</b>	<b>GTTGCCAAGG</b>	
1701	<b>ACTAGTTGGT</b>	<b>GCCTGGGCC</b>	<b>TGGGGGTAC</b>	<b>CACACACCA</b>	<b>TTTCTTGTGG</b>	
1751	<b>TAAATAAGAT</b>	<b>GTTATGTTTA</b>	<b>CATCCAAGGA</b>	<b>AGACATGTAG</b>	<b>TTTCCAAGTT</b>	
1801	<b>TGAAGGGGAA</b>	<b>ATAAGTACTA</b>	<b>TAGTAAAATG</b>	<b>AACCACATGT</b>	<b>TTCAAGTGAT</b>	
1851	<b>GGCGATGTTT</b>	<b>CTAGGCTAGG</b>	<b>TTACAAGAC</b>	<b>TTGTTAGTGA</b>	<b>CCACAATTCT</b>	Intron 2
1901	<b>TATACTACTA</b>	<b>TAAGACTTAA</b>	<b>GTCCCAAACA</b>	<b>AAGTTGGATA</b>	<b>CAATCGGGTT</b>	
1951	<b>CTATGGGTTT</b>	<b>TACTGAATTC</b>	<b>ATTGCTTTTG</b>	<b>AAGTGTGCAT</b>	<b>ACATATGAAA</b>	
2001	<b>AAGAATTTGT</b>	<b>AATGTATACA</b>	<b>TATGTAATGA</b>	<b>GATCATAAT</b>	<b>ATTTTGAAT</b>	
2051	<b>CAATAACGGG</b>	<b>TAGATCTTGG</b>	<b>AATTGCCTCT</b>	<b>TGTCCGGAAG</b>	<b>TTGTTTCATT</b>	
2101	<b>TATTGCATCG</b>	<b>CCTGTAGTA</b>	<b>AGTAATACAT</b>	<b>GAGTTTTGAT</b>	<b>ATGGTCTTAA</b>	
2151	<b>ACTTAAAG</b>	<b>TCACACATCC</b>	<b>TACCATTGAA</b>	<b>GCATGTTTTG</b>	<b>TTGTTTATAT</b>	
2201	<b>CTGTTCGTAA</b>	<b>ACTTCTTGGT</b>	<b>TAGTTGATTA</b>	<b>TTCAGCTGAT</b>	<b>ATGCTTAATT</b>	
2251	<b>ACTGTTCGTGA</b>	<b>CCAGCAATCG</b>	<b>AAACAACATT</b>	<b>GTGGTCAATT</b>	<b>GAGTGGGGTA</b>	
2301	<b>TCGCGGAACT</b>	<b>AGTCAACCAC</b>	<b>CCTCATATCC</b>	<b>AAAAGAACT</b>	<b>CCGTGATGAG</b>	
2351	<b>ATTGATACAG</b>	<b>TTCTTGGACC</b>	<b>AGGAATGCAA</b>	<b>GTGACTGAGC</b>	<b>CAGACATGCC</b>	
2401	<b>CAAGCTTCCG</b>	<b>TACCTTCAGG</b>	<b>CTGTGATCAA</b>	<b>GGAGACTCTT</b>	<b>AGACTCAGGA</b>	
2451	<b>TGGCAATTC</b>	<b>TCTTTTAGTC</b>	<b>CCACACATGA</b>	<b>ACCTTCATGA</b>	<b>TGCAAAGCTT</b>	
2501	<b>GCTGGATACG</b>	<b>ATATTCCAGC</b>	<b>TGAAAGCAA</b>	<b>ATCTTAGTTA</b>	<b>ACGCTTGGTG</b>	Exon 3
2551	<b>GCTAGCTAAC</b>	<b>AACCCGCTC</b>	<b>ACTGGAAGAA</b>	<b>ACCTGAAGAG</b>	<b>TTCAGACCTG</b>	
2601	<b>AGAGGTTCTT</b>	<b>CGAAGAGGAG</b>	<b>AAGCACGTTG</b>	<b>AGGCCAATGG</b>	<b>CAACGACTTC</b>	
2651	<b>AGATTTCTTC</b>	<b>CTTTCGGTGT</b>	<b>TGGTAGGAGG</b>	<b>AGTTGCCCGG</b>	<b>GAATTATCCT</b>	
2701	<b>TGCATTGCCA</b>	<b>ATTCTCGGCA</b>	<b>TCACTTGGG</b>	<b>ACGTTTGGTG</b>	<b>CAGAACTTTG</b>	
2751	<b>AGATGTTGCC</b>	<b>TCCTCCAGGA</b>	<b>CAGTCAAAGC</b>	<b>TCGACACCTC</b>	<b>GGAGAAAGGT</b>	
2801	<b>GGACAGTTCA</b>	<b>GTCTCCACAT</b>	<b>TTTGAAGCAT</b>	<b>TCCACCATTG</b>	<b>TGATGAAACC</b>	
2851	<b>AAGATCTTTC</b>	<b>TAA</b>	<b>ACTTTTGT</b>	<b>AATGCTATCA</b>	<b>ATTAATCATG</b>	3'-UTR
2901	<b>GTTTGTGTAA</b>	<b>ACCTTTTAA</b>	<b>TTTGACAGAA</b>	<b>AACATCTTTC</b>	<b>TTTCTTATGT</b>	
2951	<b>TTTATAAAAG</b>	<b>TCTTATTGGA</b>	<b>CTAGATTATT</b>	<b>CATTAT</b>		

amino acid sequence to C4H sequences from other plant species allows homologous regions and domains unique to the CYP73 protein family to be identified. The alignment of six C4H sequences showing high similarity to potato (red pepper, lithospermum, Madagascar periwinkle, tree cotton, wild licorice, and poplar) is shown in **Table 6**. Total four domains were found. The first domain is a hydrophobic region at the N-terminus from position 3 to 23, represented by Box A in **Table 6** (Ro *et al.* 2001). This region is responsible for membrane binding, protein stability, and is a

signal-anchor to keep the protein on the cytoplasmic side of the endoplasmic reticulum (Hotze *et al.* 1995; Nedelkina *et al.* 1999). Among the plant species shown in **Table 6**, the majority of the substitutions involve the same five amino acids, isoleucine (I), valine (V), leucine (L), phenylalanine (F), and alanine (A), which are all hydrophobic in nature. This indicates that the presence of hydrophobic amino acids is partly responsible for the function of this domain, rather than the presence of specific amino acids.

The second domain is a proline rich region which oc-

**Table 4** The *c4h* coding sequence and predicted amino acid sequence (shaded) in the potato. The two splicing sites are between AG and GC as shown in boxes.

1	ATG	GAT	CTT	CTC	TTA	CTG	GAG	AAG	ACC	TTA	ATA	GGT	CTT	TTC	TTT	GCT	ATT
1	<b>M</b>	<b>D</b>	<b>L</b>	<b>L</b>	<b>L</b>	<b>L</b>	<b>E</b>	<b>K</b>	<b>T</b>	<b>L</b>	<b>I</b>	<b>G</b>	<b>L</b>	<b>F</b>	<b>F</b>	<b>A</b>	<b>I</b>
52	TTA	ATC	GCT	ATT	ATT	GTC	TCT	AAA	CTT	CGT	TCC	AAG	CGA	TTT	AAA	CTA	CCC
18	<b>L</b>	<b>I</b>	<b>A</b>	<b>I</b>	<b>I</b>	<b>V</b>	<b>S</b>	<b>K</b>	<b>L</b>	<b>R</b>	<b>S</b>	<b>K</b>	<b>R</b>	<b>F</b>	<b>K</b>	<b>L</b>	<b>P</b>
103	CCA	GGT	CCA	ATT	CCA	GTC	CCA	GTT	TTT	GGA	AAT	TGG	CTT	CAA	GTT	GGT	GAT
35	<b>P</b>	<b>G</b>	<b>P</b>	<b>I</b>	<b>P</b>	<b>V</b>	<b>P</b>	<b>V</b>	<b>F</b>	<b>G</b>	<b>N</b>	<b>W</b>	<b>L</b>	<b>Q</b>	<b>V</b>	<b>G</b>	<b>D</b>
154	GAT	TTG	AAC	CAT	AGA	AAC	CTT	ACT	GAG	TAT	GCT	AAA	AAG	TTT	GGT	GAT	GTG
52	<b>D</b>	<b>L</b>	<b>N</b>	<b>H</b>	<b>R</b>	<b>N</b>	<b>L</b>	<b>T</b>	<b>E</b>	<b>Y</b>	<b>A</b>	<b>K</b>	<b>K</b>	<b>F</b>	<b>G</b>	<b>D</b>	<b>V</b>
205	TTC	TTG	CTT	AGA	ATG	GGG	CAA	AGG	AAC	TTA	GTT	GTT	GTG	TCA	TCC	CCT	GAA
69	<b>F</b>	<b>L</b>	<b>L</b>	<b>R</b>	<b>M</b>	<b>G</b>	<b>Q</b>	<b>R</b>	<b>N</b>	<b>L</b>	<b>V</b>	<b>V</b>	<b>V</b>	<b>S</b>	<b>S</b>	<b>P</b>	<b>E</b>
256	TTA	GCT	AAA	GAA	GTT	TTA	CAC	ACA	CAA	GGG	GTT	GAA	TTT	GGT	TCA	AGA	ACA
86	<b>L</b>	<b>A</b>	<b>K</b>	<b>E</b>	<b>V</b>	<b>L</b>	<b>H</b>	<b>T</b>	<b>Q</b>	<b>G</b>	<b>V</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>S</b>	<b>R</b>	<b>T</b>
307	AGA	AAT	GTT	GTT	TTT	GAT	ATT	TTT	ACA	GGG	AAG	GGT	CAA	GAT	ATG	GGT	TTT
103	<b>R</b>	<b>N</b>	<b>V</b>	<b>V</b>	<b>F</b>	<b>D</b>	<b>I</b>	<b>F</b>	<b>T</b>	<b>G</b>	<b>K</b>	<b>G</b>	<b>Q</b>	<b>D</b>	<b>M</b>	<b>V</b>	<b>F</b>
358	ACA	GTG	TAT	GGT	GAG	CAC	TGG	AGG	AAA	ATG	AGG	AGG	ATT	ATG	ACT	GTA	CCC
120	<b>T</b>	<b>V</b>	<b>Y</b>	<b>G</b>	<b>E</b>	<b>H</b>	<b>W</b>	<b>R</b>	<b>K</b>	<b>M</b>	<b>R</b>	<b>R</b>	<b>I</b>	<b>M</b>	<b>T</b>	<b>V</b>	<b>P</b>
409	TTT	TTT	ACT	AAT	AAG	GTG	GTG	CAG	CAG	TAT	AGA	GGG	GGG	TGG	GAG	TCT	GAG
137	<b>F</b>	<b>F</b>	<b>T</b>	<b>N</b>	<b>K</b>	<b>V</b>	<b>V</b>	<b>Q</b>	<b>Q</b>	<b>Y</b>	<b>R</b>	<b>G</b>	<b>G</b>	<b>W</b>	<b>E</b>	<b>S</b>	<b>E</b>
460	GCT	GCT	AGT	GTA	GTT	GAG	GAT	GTG	AAG	AAA	AAC	CCT	GAA	TCT	GCT	ACA	AAT
154	<b>A</b>	<b>A</b>	<b>S</b>	<b>V</b>	<b>V</b>	<b>E</b>	<b>D</b>	<b>V</b>	<b>K</b>	<b>K</b>	<b>N</b>	<b>P</b>	<b>E</b>	<b>S</b>	<b>A</b>	<b>T</b>	<b>N</b>
511	GGG	ATT	GTT	TTG	AGG	AAA	AGA	TTG	CAG	CTT	ATG	ATG	TAT	AAT	AAC	GTT	TTT
171	<b>G</b>	<b>I</b>	<b>V</b>	<b>L</b>	<b>R</b>	<b>K</b>	<b>R</b>	<b>L</b>	<b>Q</b>	<b>L</b>	<b>M</b>	<b>M</b>	<b>Y</b>	<b>N</b>	<b>N</b>	<b>M</b>	<b>F</b>
562	AGG	ATT	ATG	TTT	GAT	AGG	AGA	TTT	GAG	AGT	GAA	GAT	GAT	CCC	CTT	TTT	GTT
188	<b>R</b>	<b>I</b>	<b>M</b>	<b>F</b>	<b>D</b>	<b>R</b>	<b>R</b>	<b>F</b>	<b>E</b>	<b>S</b>	<b>E</b>	<b>D</b>	<b>D</b>	<b>P</b>	<b>L</b>	<b>F</b>	<b>V</b>
613	AAG	CTT	AGG	GCT	TTG	AAT	GGT	GAG	AGG	AGT	AGA	TTG	GCT	CAG	AGC	TTT	GAG
205	<b>K</b>	<b>L</b>	<b>R</b>	<b>A</b>	<b>L</b>	<b>N</b>	<b>G</b>	<b>E</b>	<b>R</b>	<b>S</b>	<b>R</b>	<b>L</b>	<b>A</b>	<b>Q</b>	<b>S</b>	<b>F</b>	<b>E</b>
664	TAC	AAC	TAT	GGT	GAT	TTT	ATC	CCT	ATT	TTG	AGG	CCT	TTC	TTG	AGA	GGG	TAC
222	<b>Y</b>	<b>N</b>	<b>Y</b>	<b>G</b>	<b>D</b>	<b>F</b>	<b>I</b>	<b>P</b>	<b>I</b>	<b>L</b>	<b>R</b>	<b>P</b>	<b>F</b>	<b>L</b>	<b>R</b>	<b>G</b>	<b>Y</b>
715	TTG	AAG	ATT	TGT	AAG	GAG	GTT	AAG	GAG	AAG	AGG	TTG	AAG	CTA	TTT	AAA	GAC
239	<b>L</b>	<b>K</b>	<b>I</b>	<b>C</b>	<b>K</b>	<b>E</b>	<b>V</b>	<b>K</b>	<b>E</b>	<b>K</b>	<b>R</b>	<b>L</b>	<b>K</b>	<b>L</b>	<b>F</b>	<b>K</b>	<b>D</b>
766	TAC	TTT	GTT	GAT	GAA	AGA	AAG	AG	CTT	GCA	AAT	ACC	AAG	AGC	ATG	GAC	AGC
256	<b>Y</b>	<b>F</b>	<b>V</b>	<b>D</b>	<b>E</b>	<b>R</b>	<b>K</b>	<b>K</b>	<b>L</b>	<b>A</b>	<b>N</b>	<b>T</b>	<b>K</b>	<b>S</b>	<b>M</b>	<b>D</b>	<b>S</b>
817	AAT	GCT	CTA	AAA	TGT	GCA	ATT	GAT	CAC	ATT	CTT	GAA	GCT	CAA	CAG	AAG	GGA
273	<b>N</b>	<b>A</b>	<b>L</b>	<b>K</b>	<b>C</b>	<b>A</b>	<b>I</b>	<b>D</b>	<b>H</b>	<b>I</b>	<b>L</b>	<b>E</b>	<b>A</b>	<b>Q</b>	<b>Q</b>	<b>K</b>	<b>G</b>
868	GAG	ATC	AAC	GAG	GAT	AAC	GTT	CTT	TAC	ATC	GTT	GAG	AAC	ATC	AAT	GTT	GCT
290	<b>E</b>	<b>I</b>	<b>N</b>	<b>E</b>	<b>D</b>	<b>N</b>	<b>V</b>	<b>L</b>	<b>Y</b>	<b>I</b>	<b>V</b>	<b>E</b>	<b>N</b>	<b>I</b>	<b>N</b>	<b>V</b>	<b>A</b>
919	<b>GCA</b>	ATC	GAA	ACA	ACA	TTG	TGG	TCA	ATT	GAG	TGG	GAT	ATC	GCG	GAA	CTA	GTC
307	<b>A</b>	<b>I</b>	<b>E</b>	<b>T</b>	<b>T</b>	<b>L</b>	<b>W</b>	<b>S</b>	<b>I</b>	<b>E</b>	<b>W</b>	<b>G</b>	<b>I</b>	<b>A</b>	<b>E</b>	<b>L</b>	<b>V</b>
970	AAC	CAC	CCT	CAT	ATC	CAA	AAG	AAA	CTC	CGT	GAT	GAG	ATT	GAT	ACA	GTT	CTT
324	<b>N</b>	<b>H</b>	<b>P</b>	<b>H</b>	<b>I</b>	<b>Q</b>	<b>K</b>	<b>K</b>	<b>L</b>	<b>R</b>	<b>D</b>	<b>E</b>	<b>I</b>	<b>D</b>	<b>T</b>	<b>V</b>	<b>L</b>
1021	GGA	CCA	GGA	ATG	CAA	GTG	ACT	GAG	CCA	GAC	ATG	CCC	AAG	CTT	CCG	TAC	CTT
341	<b>G</b>	<b>P</b>	<b>G</b>	<b>M</b>	<b>Q</b>	<b>V</b>	<b>T</b>	<b>E</b>	<b>P</b>	<b>D</b>	<b>M</b>	<b>P</b>	<b>K</b>	<b>L</b>	<b>P</b>	<b>Y</b>	<b>L</b>
1072	CAG	GCT	GTG	ATC	AAG	GAG	ACT	CTT	AGA	CTC	AGG	ATG	GCA	ATT	CCT	CTT	TTA
358	<b>Q</b>	<b>A</b>	<b>V</b>	<b>I</b>	<b>K</b>	<b>E</b>	<b>T</b>	<b>L</b>	<b>R</b>	<b>L</b>	<b>R</b>	<b>M</b>	<b>A</b>	<b>I</b>	<b>P</b>	<b>L</b>	<b>L</b>
1123	GTC	ACA	CAC	ATG	AAC	CTT	CAT	GAT	GCA	AAG	CTT	GCT	GGA	TAC	GAT	ATT	CCA
375	<b>V</b>	<b>P</b>	<b>H</b>	<b>M</b>	<b>N</b>	<b>L</b>	<b>H</b>	<b>D</b>	<b>A</b>	<b>K</b>	<b>L</b>	<b>A</b>	<b>G</b>	<b>Y</b>	<b>D</b>	<b>I</b>	<b>P</b>
1174	GCT	GAA	AGC	AAA	ATC	TTA	GTT	AAC	GCT	TGG	TGG	CTA	GCT	AAC	AAC	CCC	GCT
392	<b>A</b>	<b>E</b>	<b>S</b>	<b>K</b>	<b>I</b>	<b>L</b>	<b>V</b>	<b>N</b>	<b>A</b>	<b>W</b>	<b>W</b>	<b>L</b>	<b>A</b>	<b>N</b>	<b>N</b>	<b>P</b>	<b>A</b>
1225	CAC	TGG	AAG	AAA	CCT	GAA	GAG	TTC	AGA	CCT	GAG	AGG	TTC	TTC	GAA	GAG	GAG
409	<b>H</b>	<b>W</b>	<b>K</b>	<b>K</b>	<b>P</b>	<b>E</b>	<b>E</b>	<b>F</b>	<b>R</b>	<b>P</b>	<b>E</b>	<b>R</b>	<b>F</b>	<b>F</b>	<b>E</b>	<b>E</b>	<b>E</b>
1276	AAG	CAC	GTT	GAG	GCC	AAT	GGC	AAC	GAC	TTC	AGA	TTT	CTT	CCT	TTC	GGT	GTT
426	<b>K</b>	<b>H</b>	<b>V</b>	<b>E</b>	<b>A</b>	<b>N</b>	<b>G</b>	<b>N</b>	<b>D</b>	<b>F</b>	<b>R</b>	<b>F</b>	<b>L</b>	<b>P</b>	<b>F</b>	<b>G</b>	<b>V</b>
1327	GGT	AGG	AGG	AGT	TGC	CCC	GGA	ATT	ATC	CTT	GCA	TTG	CCA	ATT	CTC	GGC	ATC
443	<b>G</b>	<b>R</b>	<b>R</b>	<b>S</b>	<b>C</b>	<b>P</b>	<b>G</b>	<b>I</b>	<b>I</b>	<b>L</b>	<b>A</b>	<b>L</b>	<b>P</b>	<b>I</b>	<b>L</b>	<b>G</b>	<b>I</b>
1378	ACT	TTG	GGA	CGT	TTG	GTG	CAG	AAC	TTT	GAG	ATG	TTG	CCT	CCT	CCA	GGA	CAG
460	<b>T</b>	<b>L</b>	<b>G</b>	<b>R</b>	<b>L</b>	<b>V</b>	<b>Q</b>	<b>N</b>	<b>F</b>	<b>E</b>	<b>M</b>	<b>L</b>	<b>P</b>	<b>P</b>	<b>P</b>	<b>G</b>	<b>Q</b>
1429	TCA	AAG	CTC	GAC	ACC	TCG	GAG	AAA	GGT	GGA	CAG	TTC	AGT	CTC	CAC	ATT	TTG
477	<b>S</b>	<b>K</b>	<b>L</b>	<b>D</b>	<b>T</b>	<b>S</b>	<b>E</b>	<b>K</b>	<b>G</b>	<b>G</b>	<b>Q</b>	<b>F</b>	<b>S</b>	<b>L</b>	<b>H</b>	<b>I</b>	<b>L</b>
1480	AAG	CAT	TCC	ACC	ATT	GTG	ATG	AAA	CCA	AGA	TCT	TTC	TAA				
494	<b>K</b>	<b>H</b>	<b>S</b>	<b>T</b>	<b>I</b>	<b>V</b>	<b>M</b>	<b>K</b>	<b>P</b>	<b>R</b>	<b>S</b>	<b>F</b>	*				

curs from amino acid 34 to 41 of the C4H protein sequence (Box B in Table 6). This region is thought to be responsible for correctly orientating and folding the protein in the membrane by breaking  $\alpha$ -helix bonds (Mizutani *et al.* 1997; Koopmann *et al.* 1999). Cytochrome P450 proteins contain a conserved region that is involved in the binding and activation of dioxygen, which is necessary for oxygen incorporation into the corresponding substrates (Schalk *et al.* 1999). The consensus sequence for this region in plant P450 proteins is as follows; (A/G)(A/G)I(E/D)T. As seen in Table 6 (Box C), the sequence of the motif in the C4H protein family is AA IET and is identical in all plant species shown. One of the most important domains in the P450 family of proteins is the heme-binding domain positioned at amino acids 439 to 449 (Box D of Table 6). The consensus se-

quence for this domain in P450 proteins is PFGXGRRXC XG. In the CYP73 family, the domain (PFGVGRRSCPG) is conserved in all plant species sequenced, indicating that there is a consensus sequence specifically for C4H. The importance of this domain is that it allows the binding of the heme molecule to the enzyme which is essential for catalysis and the ability to bind carbon monoxide (Chapple 1998). The binding of the heme molecule occurs through a thiolate side chain that originates from the conserved cysteine amino acid at position 447 (Schalk *et al.* 1999). In the C4H family of proteins, the interaction of the conserved cysteine (C) and the subsequent proline (P) molecule enables the formation of a "cysteine pocket" in which the sulfur-iron bond is in the center of a hydrophobic environment (Schalk *et al.* 1999).

**Table 5** Length and similarity of plant class I *c4h* coding sequences to potato.

Species	Length (bp)	Sequence Identity (bp %)	Genbank® Acc. #
Red pepper ( <i>Capsicum annuum</i> )	1518	1379 (91)	AF212318
Lithospermum ( <i>Lithospermum erythrorhizon</i> ), <i>c4h-2</i>	1518	1249 (82)	AB055508
Poplar ( <i>Populus x generosa</i> )	1518	1200 (79)	AF302495
Japanese aspen ( <i>Populus kitakamiensis</i> )	1518	1197 (79)	D82815
Quaking aspen ( <i>Populus tremuloides</i> )	1518	1194 (79)	U47293
Lithospermum ( <i>Lithospermum erythrorhizon</i> ), <i>c4h-1</i>	1518	1182 (78)	AB055507
Grapefruit ( <i>Citrus x paradisi</i> )	1518	1159 (76)	AF378333
Chickpea ( <i>Cicer arietinum</i> )	1518	1155 (76)	AJ007449
Tree cotton ( <i>Gossypium arboreum</i> )	1518	1144 (75)	AF286648
Madagascar periwinkle ( <i>Catharanthus roseus</i> )	1518	1143 (75)	Z32563
Sweet orange ( <i>Citrus sinensis</i> ), <i>c4h-2</i> (Class I)	1560	1113 (73)	AF255014
Alfalfa ( <i>Medicago sativa</i> )	1521	1112 (73)	L11046
Zinnia ( <i>Zinnia elegans</i> )	1518	1093 (72)	U19922
Bishop's weed ( <i>Ammi majus</i> )	1521	1088 (72)	AY219918
<i>Arabidopsis thaliana</i>	1518	1083 (71)	U71080
Jerusalem artichoke ( <i>Helianthus tuberosus</i> )	1518	1073 (71)	Z17369
Wild licorice ( <i>Glycyrrhiza echinata</i> )	1518	1018 (67)	D87520

### The *c4h* gene expression analysis using Northern hybridization

The *c4h* Northern hybridization signals in the four groups of tuber samples are shown in **Fig. 2A**. The size of the *c4h* transcript was measured based on the location and known size of the transcribed *c4h* probe (710 bp) and the 25S, 18S, and 5S ribosomal RNA bands (3.8, 2.0, 0.74 kb, respectively). The size of the *c4h* signal is between 600 and 800 bp (**Fig. 2B**). The reason for such a wide spread size is due to the uneven speed of migration when the total RNA was overloaded on the agarose gel.

In order to verify the equal loading of the total RNA used for Northern hybridization, the ribosomal RNA (both the 18S rRNA and 25S rRNA bands) was quantified digitally as shown in **Fig. 2B**, prior to transferring the RNA to the membrane. The 'Russet Burbank' 18S and 25S rRNA bands were assigned a value of 1.0. The relative intensities of the remaining three samples were given a value relative to the 'Russet Burbank' RNA. The relative intensities of the *c4h* transcript were adjusted according to the proportion of RNA loaded for each sample. The mean adjusted *c4h* transcript levels of the four tuber samples are shown in **Table 7**. The dark clone showed the highest *c4h* gene expression level; the light clone showed the lowest *c4h* gene expression level, respectively; whereas the mean *c4h* transcript levels of the two tetraploid cultivars, fall in between the dark and the light clones, were not significantly different.

The potential relationship between ACD and *c4h* gene expression was analyzed by comparing the mean relative *c4h* transcript levels to their mean degree of ACD (**Fig. 3**). The mean pixel densities of 'Russet Burbank' and 'Russet Norkotah' tubers were found to be 113 and 114, respectively, which was not significantly different at  $\alpha = 0.05$ . For the light diploid clone CH72.03, the mean pixel density was higher (121 pixel units), which corresponded to a lower degree of ACD. The mean pixel density of the dark diploid clone 10908.06 was found to be significantly lower at 89 pixel units, which corresponds to a higher level of tuber darkening. As seen in **Fig. 3**, there are no significant differences in the degree of ACD or in intensity of the *c4h* transcript in 'Russet Burbank' and 'Russet Norkotah'. In contrast, the degree of darkening in the identified dark diploid sample is significantly higher when compared to the identified light diploid, and the relative intensity of the *c4h* hybridization signal is also significantly higher in the dark diploid when compared to the light diploid. These results suggest that the level of the higher *c4h* transcript produced the tubers with the higher susceptibility to ACD.

The objective of the Northern hybridization analysis was to determine if changes in *c4h* gene expression occurred in cultivars with varying degrees of ACD. The results of the Northern hybridization and the ACD evaluation

data suggest that potentially there is a relationship between *c4h* gene expression and ACD. The lack of significant differences between the tetraploid cultivars for ACD and *c4h* transcript intensity, as well as the significant differences in intensity for the dark and light diploid clones suggests that *c4h* is involved in the ACD mechanism. This evidence indicates that there is a relationship between *c4h* gene expression levels and the level of darkening in the tuber, where potato cultivars with higher *c4h* expression levels have an increased susceptibility to ACD, and *vice versa*.

### The *c4h* gene expression analysis using RQRT-PCR

The PCR products of all the RQRT-PCR were analyzed by agarose gel electrophoresis. **Fig. 4** provided an example of four gel images showing the amplified *c4h* gene products in the clones of the population 13610-T. Samples from the population 13395-B were analyzed in the same manner. After the band intensities were measured, all the data underwent statistical analyses. **Fig. 5** summarized the comparison of the *c4h* gene expression levels in both the dark and light clones of the populations 13610-T and 13395-B. The *c4h* gene showed significantly higher expression levels in the dark clones compared with that of the light clones in both populations. The difference in the population 13610-T was much greater than that of the 13395-B. This was not surprising, since the population 13610-T was bred particularly for ACD genetic mapping study using a dark and a light diploid clones as parents. As a result, the differences in the degree of ACD in the selected dark or light clones were much greater than the clones from the 13395-B population. This finding strongly supported the role of *c4h* gene expression in relationship to the degree of ACD in potato tubers.

Chlorogenic acid, a polyphenol consisting of caffeic and quinic acids, is the major phenolic secondary metabolite in potato tubers, making up to 90% of the total phenolics. Besides acting as the key determinant of the ACD severity, CGA is also associated with many plant defense mechanisms against insects or phytopathogens, disease and fungal resistance, growth regulation, and wound response (Friedman 1997). Cinnamic acid 4-hydroxylase is a key enzyme for CGA biosynthesis. This study demonstrated its role in determining the degree of ACD in potato tubers. The outcome holds significant value in developing gene markers for selecting cultivars through breeding processes. In addition, its functions in antioxidant capacity and plant defense should be further explored.

### ACKNOWLEDGEMENTS

The study was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Industrial Post-Graduate

Fellowship to S. Babin (Cantle) and the NSERC Discovery Grant to G. Wang-Pruski. The diploid potato samples were donated by Agriculture and Agri-Food Canada Potato Research Centre.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410  
 Bachem CWB, van der Hoeven RS, de Brujin SM, Vreugdenhil D, Zabeau M, Visser RGF (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. *The Plant Journal* 9, 745-753

Table 6 The identification of conserved peptide domains by the alignment of amino acid sequences most highly homologous to the potato C4H.

	BOX A	BOX B	
Potato	MDLLLEKTLIGLFFAILIAITV	SKLRSKRFKLPPGPIPVV	51
Red Pepper	MDLLLEKTLVGLFFAIVVAIV	SKLRSKRFKLPPGPIPVV	
Lithospermum	MDLLLEKALIGLFFSFIIAIV	SKLRGKFKLPPGPIPVV	
Periwinkle	MDLLLEKTLIGLFAAIVASIV	SKLRGKFKLPPGPIPVV	
Tree Cotton	MDLLEKLVLSLFFIIIFAILV	SKLRGKFKLPPGPIPVV	
Wild Licorice	MDLLLEKTLGLFTAATAITV	SKLRGKFKLPPGPIPVV	
Poplar	MDLLLEKTLGSEVAILVAIV	SKLRGKFKLPPGPIPVV	
Potato	DLNHRNLTEYAKKFGDVFLLRMGQRNLVVSSPELAKEVLHTQGVEFGSRT		102
Red Pepper	DLNHRNLTYAKKFGDIFLLRMGQRNLVVSSPELAKEVLHTQGVEFGSRT		
Lithospermum	DLNHRNLTEYAKKFGDIFLLRMGQRNLVVSSPELAKEVLHTQGVEFGSRT		
Periwinkle	DLNHRNLSDYAKKFGDIFLLRMGQRNLVVSSPELAKEVLHTQGVEFGSRT		
Tree Cotton	DLNHRNLTLAKKFGDIFLLRMGQRNLVVSSPELAKEVLHTQGVEFGSRT		
Wild Licorice	DLNHRNLTLAKKFGDIFLLRMGQRNLVVSSPELAKEVLHTQGVEFGSRT		
Poplar	DLNHRNLTLAKKFGDIFLLRMGQRNLVVSSPELAKEVLHTQGVEFGSRT		
Potato	RNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTVPFFTNKVQQYRGGWESE		153
Red Pepper	RNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTVPFFTNKVQQYRGGWESE		
Lithospermum	RNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTVPFFTNKVQQYRGGWESE		
Periwinkle	RNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTVPFFTNKVQQYRGGWESE		
Tree Cotton	RNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTVPFFTNKVQQYRGGWESE		
Wild Licorice	RNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTVPFFTNKVQQYRGGWESE		
Poplar	RNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTVPFFTNKVQQYRGGWESE		
Potato	AASVVEDVKKNPESATNGIVLRKRLQLMMYNNMFRIMFDRRFESEDDPLFV		204
Red Pepper	VASVVEDVKKNPESATNGIVLRKRLQLMMYNNMFRIMFDRRFESEDDPLFV		
Lithospermum	VESVLEDVKKIPESETVGIIVLRKRLQLMMYNNMFRIMFDRRFESEDDPLFM		
Periwinkle	AAVVEDVKKNPESATNGIVLRKRLQLMMYNNMFRIMFDRRFESEDDPLFV		
Tree Cotton	AASVVEDVKKNPEAATNGIVLRKRLQLMMYNNMFRIMFDRRFESEDDPLFV		
Wild Licorice	AASVVDVRRNPDAAGGIVLRKRLQLMMYNNMFRIMFDRRFESEDDPLFV		
Poplar	AAVVEDVKKNPEAATNGIVLRKRLQLMMYNNMFRIMFDRRFESEDDPLFN		
Potato	KLRALNGERSRLAQSFEYNYGDFIPILRPFLRGYLKICKEVKEKRLKLFKD		255
Red Pepper	KLRALNAERSRLAQSFEYNYGDFIPILRPFLRGYLKICKEVKEKRLQLFKD		
Lithospermum	KLRALNGERSRLAQSFDYNYGDFIPILRPFLRGYLKICKEVKEKRLKLFKD		
Periwinkle	KLRALNGERSRLAQSFEYNYGDFIPILRPFLRGYLKICKEVKEKRLQLFKD		
Tree Cotton	KLRALNGERSRLAQSFEYNYGDFIPILRPFLRGYLKICKEVKEKRLQLFKD		
Wild Licorice	KLRALNGERSRLAQSFEYNYGDFIPILRPFLRGYLKICKEVKEKRLKLFKD		
Poplar	KLRALNGERSRLAQSFDYNYGDFIPILRPFLRGYLKICKEVKEKRLQLFKD		
Potato	YFVDERKKLANTKSMDSNALKCAIDHILEAQKGEINEDNVLYIVENINVA		306
Red Pepper	YFVDERKKLNTKSMDSNALKCAIDHILEAQKGEINEDNVLYIVENINVA		
Lithospermum	YFVDERKKLSTKSTTTNGLKCAIDHILEAQKGEINEDNVLYIVENINVA		
Periwinkle	YFVDERKKFGSTKSMDSNALKCAIDHILEAQKGEINEDNVLYIVENINVA		
Tree Cotton	QFLEERKKLATTKRIDNALKCAIDHILDAQKGEINEDNVLYIVENINVA		
Wild Licorice	YFVDERMKLESTKSTNEGLKCAIDHILDAQKGEINEDNVLYIVENINVA		
Poplar	YFVDERKKLSTKNMSNEGLKCAIDHILDAQKGEINEDNVLYIVENINVA		
	<b>BOX C</b>		
Potato	AIETTLWSIEWGIAELVNHPIQKKLRDEIDTVLPGMQVTEPDMKLPYL		357
Red Pepper	AIETTLWSIEWGIAELVNHPIQKKLRDEIDAVLPGVQVTEPDTHKLPDL		
Lithospermum	AIETTLWSIEWGIAELVNHPIQKKLRDEIDTVLPGVQVTEPDTHKLPYL		
Periwinkle	AIETTLWSIEWGIAELVNHPIQKKLRDELETVLPGVQVTEPDYKLPYL		
Tree Cotton	AIETTLWSIEWGIAELVNHPIQKKLRNEIDTVLPGVQVTEPDTHKLPYL		
Wild Licorice	AIETTLWSIEWGIAELVNHPIQKKVRDEIDRVLPGHQVTEPDMKLPYL		
Poplar	AIETTLWSIEWGIAELVNHPIQKKLRHEIDTVLPGHQVTEPDYKLPYL		
Potato	QAVIKETLRLRMAIPLLVPHMNLHDAKLAGYDIPAESKILVNAWFLANNPA		408
Red Pepper	QAVIKETLRLRMAIPLLVPHMNLHDAKLAGYDIPAESKILVNPWFLANNPA		
Lithospermum	QAVIKETLRLRMAIPLLVPHMNLHDAKLAGYDIPAESKILVNAWFLANNPA		
Periwinkle	QAVIKETLRLRMAIPLFLPHMNLHDAKLAGYDIPAESKILVNAWFLANNPE		
Tree Cotton	QAVIKETLRLRMAIPLLVPHMNLHDAKLAGYDIPAESKILVNAWFLANNPA		
Wild Licorice	QAVIKETLRLRMAIPLLVPHMNLHDAKLAGYDIPAESKILVNAWFLANNPA		
Poplar	NAVIKETLRLRMAIPLLVPHMNLHDAKLAGYDIPAESKILVNAWFLANNPA		



Table 6 (Cont.)

		BOX D	
Potato	HWKKEEFRRPERFFEEKHEANGNDFRFL	PFVGVRRSCPGIILALPILGI	459
Red Pepper	HWKKEEFRRPERFLKEEKHVDANGNDFRFL	PFVGVRRSCPGIILALPILGI	
Lithospermum	QWKNPEEFRRPERFLEEEAKVEANGNDFRYL	PFVGVRRSCPGIILALPILGI	
Periwinkle	HWKKEEFRRPERFLEEEKVEANGNDFRYL	PFVGVRRSCPGIILALPILGI	
Tree Cotton	HWKNPEEFRRPERFFEEESKVEANGNDFRYL	PFVGVRRSCPGIILALPILGI	
Wild Licorice	NWKRPEEFRRPERFLEEEESHVEANGNDFRYL	PFVGVRRSCPGIILALPILGI	
Poplar	HWKNPEEFRRPERFLEEEAKVEANGNDFRYL	PFVGVRRSCPGIILALPILGI	
Potato	TLGRLVQNFEMLLPPPGQSKLDTSEKGGQFSLHILKHSTIVMKPRSF*		505
Red Pepper	TLGRLVQNFEMLLPPPGQSKLDTTEKGGQFSLHILKHSTIVMKPRSF*		
Lithospermum	TLGRLVQNFEMLLPPPGQSKLDTSEKGGQFSLHILKHSTIVMKPRSF*		
Periwinkle	TLGRLVQNFEMLLPPPGKSKLDTSEKGGQFSLHILKHSTIVLKPRTF*		
Tree Cotton	TLGRLVQNFEMLLPPKQSKLDTSEKGGQFSLHILKHSTIVAKPRVF*		
Wild Licorice	TLGRLVQNFEMLLPPPGQSKLDTAEKGGQFSLHILKHSTIVAKPRSF*		
Poplar	TLGRLVQNFEMLLPPPGQSKLDTAEKGGQFSLHILKHSTIVAKPRSF*		

Table 7 Mean relative intensity of *c4h* expression as detected by Northern hybridization analyses<sup>a</sup>.

Tuber sample	Number of observations	Mean relative intensity of <i>c4h</i> expression	
Dark Diploid (10908.06)	3	1.70 ± 0.53	a
Russet Burbank	3	1.12 ± 0.10	ab
Russet Norkotah	3	1.12 ± 0.11	ab
Light Diploid (CH72.03)	3	0.91 ± 0.18	b

<sup>a</sup> Means followed by the same letter are not significantly different according to Tukey's hsd test at  $\alpha = 0.05$ .

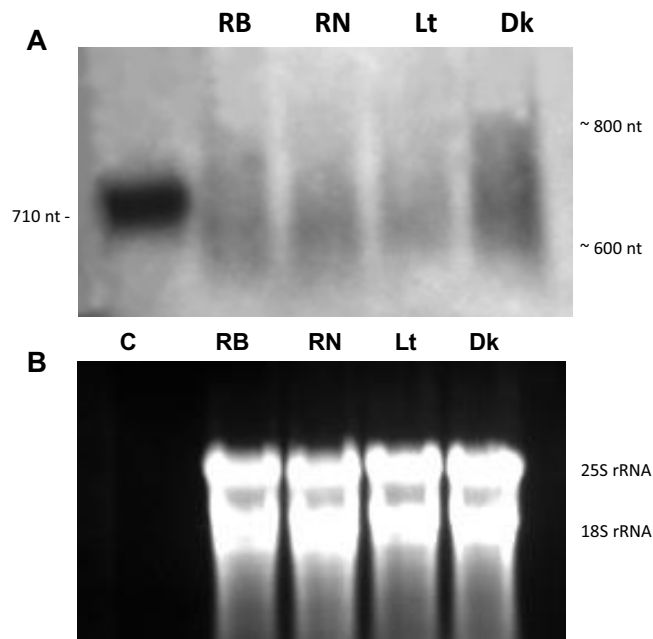


Fig. 2 Northern hybridization analysis of *c4h* expression in Russet Burbank (RB), Russet Norkotah (RN), light-ACD diploid (Lt), and dark-ACD diploid (Dk). (A) Hybridization signals from one of the three replicates using the 472 bp *c4h* probe. Lane C is the transcribed *c4h* probe used as a positive control (710 nt in length). (B) The corresponding total RNA separated by formaldehyde agarose (0.7%) gel electrophoresis.

- Bell-Lelong DA, Cusumano JC, Meyer K, Chapple C (1997) Cinnamate 4-hydroxylase expression in *Arabidopsis*. *Plant Physiology* **113**, 729-738
- Betz C, McCollum TG, Mayer RT (2001) Differential expression of two cinnamate 4-hydroxylase genes in 'Valencia' orange (*Citrus sinensis* Osbeck). *Plant Molecular Biology* **46**, 741-748
- Blee K, Choi JW, O'Connell AP, Jupe SC, Schuch W, Lewis NG, Bolwell GP (2001) Antisense and sense expression of cDNA coding for CYP73A15, a class II cinnamate 4-hydroxylase, leads to a delayed and reduced production of lignin in tobacco. *Phytochemistry* **57**, 1159-1166
- Burleigh SH (2001) Relative quantitative RT-PCR to study the expression of plant nutrient transporters in arbuscular mycorrhizas. *Plant Science* **160**, 899-904

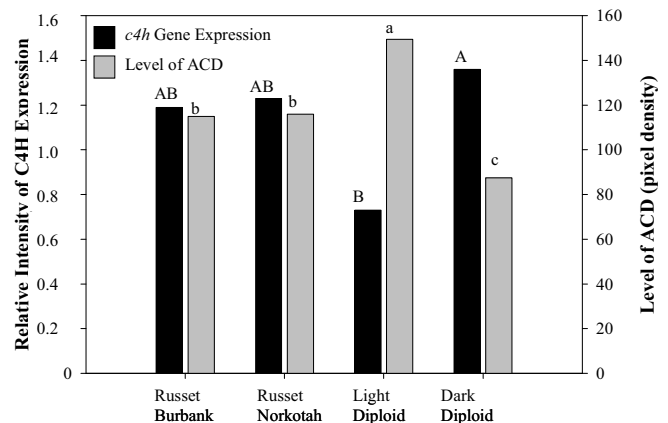


Fig. 3 Relationship between *c4h* gene expression and ACD level in potato tubers. Black bars represent the mean relative intensity of *c4h* expression (left axis) and gray bars represent the digital measurement of ACD in the potato tubers (right axis). A lower pixel density reading (right axis) represents higher ACD. Different upper case letters represent significance for the mean relative intensity of *c4h* expression, according to Tukey's hsd test at  $\alpha = 0.05$ . Different lower case letters represent significance between the ACD levels, according to Tukey's hsd test at  $\alpha = 0.05$ .

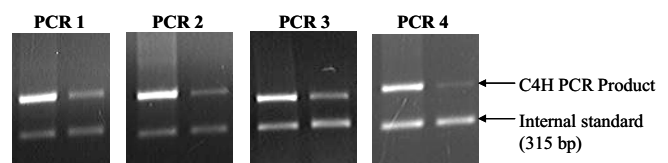
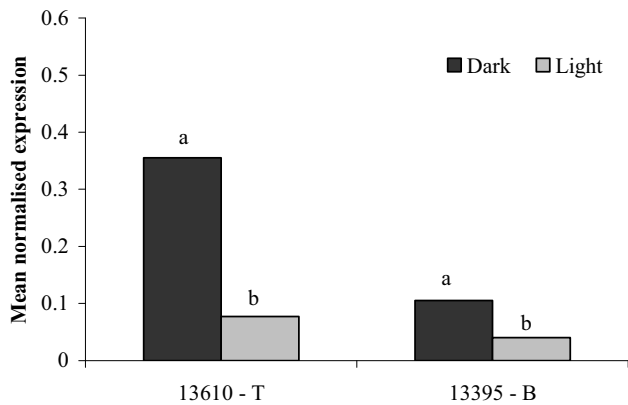


Fig. 4 RQRT-PCR analyses showing differential *c4h* gene expression in the dark and light diploid clones. The figure showed four agarose gel images, each representing one replicated PCR reaction. Each sample came from the pooled templates of three dark and three light clones in the population of 13610-T. The upper band was the amplified PCR products, the lower band of 315 bp was the amplified 18s rRNA products. The band intensity was measured as MRD levels.

- Chapple C (1998) Molecular genetic analysis of plant cytochrome P450-dependent monooxygenases. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**, 311-343
- Dale MFB, Mackay GR (1994) Inheritance of table and processing quality. In: Bradshaw JE, Mackay GR (Eds) *Potato Genetics*, CAB International Publishers, Wallingford, UK, pp 285-315
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* **1**, 13-15
- Fahrendorf T, Dixon RA (1993) Stress responses in alfalfa (*Medicago sativa* L.) XVIII: Molecular cloning and expression of the elicitor-inducible cinnamic acid 4-hydroxylase cytochrome P450. *Archives in Biochemistry and Biophysics* **305**, 509-515
- Friedman M (1997) Chemistry, biochemistry, and dietary role of potato polyphenols. *Journal of Agricultural and Food Chemistry* **45**, 1523-1540
- Griffiths DW, Bain H (1997) Photo-induced changes in the concentrations of



**Fig. 5** Statistical comparison of the *c4h* expression levels in all dark or light clones of 13610-T and 13395-B. Different letters represent significance between the dark or light clone groups, according to Tukey's hsd test at  $\alpha = 0.05$ .

individual chlorogenic acid isomers in potato (*Solanum tuberosum*) tubers and their complexation with ferric ions. *Potato Research* **40**, 307-315

- Hotze M, Schröder G, Schröder J** (1995) Cinnamate 4-hydroxylase from *Catharanthus roseus*, and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in *Escherichia coli*. *FEBS Letters* **374**, 345-350
- Inoue H, Nojima H, Okayama H** (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23-28
- Kawai S, Mori A, Shiokawa T, Kajita S, Katayama Y, Morohoshi N** (1996) Isolation and analysis of cinnamic acid 4-hydroxylase homologous genes from a hybrid aspen, *Populus kitakamiensis*. *Bioscience, Biotechnology and Biochemistry* **60**, 1586-1597
- König H, Ponta H, Herrlich P** (1998) Coupling of signal transduction to alternative pre-mRNA splicing by a composite splice regulator. *The EMBO Journal* **17**, 2904-2913
- Koopmann E, Logemann E, Hahlbrock K** (1999) Regulation and functional expression of cinnamate 4-hydroxylase from parsley. *Plant Physiology* **119**, 49-55
- Kühnl T, Koch U, Heller W, Wellmann E** (1987) Chlorogenic acid biosynthesis: Characterization of a light-induced microsomal 5-*O*-(4-coumaroyl)-D-quinic acid 3'-hydroxylase from carrot (*Daucus carota* L.) cell suspension cultures. *Archives in Biochemistry and Biophysics* **258**, 226-232
- Lewis CE, Walker JRL, Lancaster JE, Sutton KH** (1998) Determination of anthocyanins, flavonoids, and phenolic acids in potatoes. I: Colored cultivars of *Solanum tuberosum* L. *Journal of Science, Food Agriculture* **77**, 45-57
- Lugasi A, Almeida DPF, Dworschak E** (1999) Chlorogenic acid content and

antioxidant properties of potato tubers as related to nitrogen fertilization. *Acta Alimentaria* **28**, 183-195

- Mizutani M, Ohta D, Sato R** (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis* and its expression manner in planta. *Plant Physiology* **113**, 755-763
- Narayanan K** (2005) Gene expression studies in potato and their relationship to after-cooking darkening. MSc thesis, Dalhousie University, 90 pp
- Nedelkina S, Jupe SC, Blee KA, Schalk M, Werck-Reichhart D, Bolwell GP** (1999) Novel characteristics and regulation of a divergent cinnamate 4-hydroxylase (CYP73A15) from French bean: Engineering expression in yeast. *Plant Molecular Biology* **39**, 1079-1090
- Percival GC, Baird L** (2000) Influence of storage upon light-induced chlorogenic acid accumulation in potato tubers (*Solanum tuberosum* L.). *Journal of Agriculture and Food Chemistry* **48**, 2476-2482
- Petersen M** (2003) Cinnamic acid 4-hydroxylase from cell cultures of the hornwort *Anthoceros agrestis*. *Planta* **217**, 96-101
- Ro DK, Mah N, Ellis BE, Douglas CJ** (2001) Functional characterization and subcellular localization of poplar (*Populus trichocarpa* x *Populus deltoides*) cinnamate 4-hydroxylase. *Plant Physiology* **126**, 317-329
- Sambrook J, Russell DW** (2001) *Molecular Cloning: A Laboratory Manual* (3<sup>rd</sup> Edn, Vols 1-3), Cold Spring Harbor Laboratory Press, New York, USA, 18,125 pp
- Schalk M, Nedelkina S, Schoch G, Batard Y, Werck-Reichhart D** (1999) Role of unusual amino acid residues in the proximal and distal heme regions of a plant P450, CYP73A1. *Biochemistry* **38**, 6093-6103
- Smith O** (1987) Effect of cultural and environmental conditions on potatoes for processing. In: *Potato Processing* (4<sup>th</sup> Edn) Van Nostrand Reinhold Co. Inc., New York, USA, 107 pp
- Tanaka Y, Kojima M, Uritani I** (1974) Properties, development, and cellular-localization of cinnamic acid 4-hydroxylase in cut-injured sweet potato. *Plant Cell Physiology* **15**, 843-854
- Tarn TR, Tai GCC, De Jong H, Murphy AM, Seabrook JEA** (1992) Breeding potatoes for long-day, temperate climates. *Plant Breeding Reviews* **9**, 217-332
- Wang-Pruski G** (2006) Digital imaging for evaluation of potato after-cooking darkening and its comparison with other methods. *International Journal of Food Science and Technology* **41** (8), 885-891
- Wang-Pruski G** (2007) The Canon of Potato Science: 47. After-cooking darkening. *Potato Research* **50**, 403-406
- Wang-Pruski G, Nowak J** (2004) Potato after-cooking darkening. *American Journal of Potato Research* **81**, 7-16
- Whitbred JM, Schuler MA** (2000) Molecular characterization of CYP73A9 and CYP82A1 P450 genes involved in plant defense in pea. *Plant Physiology* **124**, 47-58
- Yao K, De Luca V, Brisson N** (1995) Creation of a metabolic sink for tryptophan alters the phenylpropanoid pathway and the susceptibility of potato to *Phytophthora infestans*. *Plant Cell* **7**, 1787-1799
- Yu H, Kumar PP** (2003) Post-transcriptional gene silencing in plants by RNA. *Plant Cell Reports* **22**, 167-174