Yield and antiyield genes in common bean (Phaseolus vulgaris L.)

Yarmilla Reinprecht1 | Yanzhou Qi2 | Fariba Shahmir1 | Thomas H. Smith1 | Karl Peter Pauls1

1Department of Plant Agriculture, University of Guelph, Guelph, Ontario, Canada
2Institute of Apicultural Research, Chinese Academy of Agricultural Sciences, Beijing, China

Abstract
High yield is the primary criterion for the development of new cultivars at the University of Guelph common bean breeding program. As a complex trait, yield is encoded by a number of genes contributing minor effects while also being significantly affected by environmental factors. Genes that increase yield with fixed resources have their effects by increasing input use efficiency. When suppressed, the BnMicEmUP gene has a positive effect on seed production in Arabidopsis. Preliminary work has shown that ortholog of this gene (Phvul.009G190100) exists in common bean, and its expression was negatively correlated with yield in a field test of 10 navy bean cultivars with different yield potentials. The aim of this research was to characterize the Phvul.009G190100 gene and to develop gene-based marker(s) to test for alleles associated with high yield in common bean. A database search identified a second yield-related gene (Phvul.009G202100) on the same chromosome (Pv09), which is a homolog to Phvul.009G190100. Both genes contain a DUF1118 protein domain, which has the molecular characteristics of a basic leucine zipper (bZIP) transcription factor, based on in silico analysis. Temperature switch polymerase chain reaction (PCR) markers, which were developed for both genes, were significantly associated with yield and maturity in 42 bean genotypes belonging to different market classes. The work will benefit bean breeding programs by making them more efficient in selecting high yielding cultivars, and it will directly benefit bean producers through accelerated access to new, high yielding cultivars.

KEYWORDS
antiyield gene, common bean, temperature switch PCR markers, yield

1 | INTRODUCTION
Common (dry) bean (Phaseolus vulgaris L.) is the most important legume crop used for direct human consumption. It is a good source of fiber and dietary proteins and contains little fat, sodium, or cholesterol (Duranti, 2006). Beans are also excellent sources of vitamins and minerals, with only moderate levels of calories (Hefni et al., 2010). In 2016, 30.4 million tons of beans were harvested globally with an estimated value of US$13.4 billion (available at: www.fao.org/faostat/en/#compare). Ontario beans grown on 47,600 ha (with an average yield of 2.85 tons ha⁻¹) were worth approximately $126 M in 2018 (available at: http://www.omafra.gov.on.ca/english/stats/crops/index.html).
High yield is the major objective of most plant breeding programs. The average yield of Canadian beans has increased by 59% from an average of 1.415.8 kg ha$^{-1}$ during 1961–1970 (10-year average) to 2.250.9 kg ha$^{-1}$ for period 2009–2018 (10-year average), which is a 1.6% (1971–2008) or 1.0% (1961–2018) annual increase (FAOSTAT, 2020; available at: http://www.fao.org/faostat/en/#data/QV). Generally, there is renewed interest in increasing the yield potentials of crops to meet demands from population growth that might double by 2050. In addition, there is a concern that shifts to meat-rich diets, increased biofuel production, and climate change will put additional pressure on crops to yield more with the same or decreasing inputs (Tilman et al., 2011). Studies have estimated that yields will need to increase by 2.4% annually, rather than the current average yield increase of approximately 1.2% for the major crops such as corn, soybean, wheat, or rice (Ray et al., 2013).

Common bean improvement programs have been successful in using conventional plant breeding techniques in the development and release of cultivars with greater yield potential (Kelly et al., 1998). The use of molecular markers may permit selection in the absence of selection pressure. However, they have been most effective for relatively simple traits, like disease resistance. The application of marker-assisted selection (MAS) for more complex traits, such as yield, has recently shifted to genomic selection approaches (Desta & Ortiz, 2014) that are based on genome-wide association studies (GWAS) and do not require knowledge of the underlying genes that control the trait (Atwell et al., 2010).

As a quantitative trait, yield is encoded by a number of genes, which usually have small effects on trait expression and can be significantly modified by environment. From a crop physiology perspective, a number of morphological characteristics and physiological processes contribute to yield, including flowering time, time to maturity, number of pods per plant, number of seeds per pod, and seed weight. To achieve maximum yield, all these components should be well-balanced considering that grain yield can be influenced by altering any one of them (Chapman et al., 2003). In addition, because the yield loss is usually caused by biotic and abiotic stresses (Fageria, 2002; Lynch, 2007), an improvement of resistance to these stresses should be included in breeding for high yielding cultivars. For example, the improvement of disease resistance has led to extended duration of photosynthesis in many crops but selection for high yield “has not resulted in genetic increase in photosynthesis per unit leaf area” (Richards, 2000).

Because yield determinations are labor-intensive and dependent on the environment, it would be beneficial to have a constant parameter correlated with the trait that can be used as a selection tool. Genes that increase yield with fixed resources have their effects by increasing input use efficiencies. There are some promising leads with genes associated with yield in a number of crops (Tian et al., 2017; Xu et al., 2018). Two mutants, reduced induction of nonphotochemical quenching 1 (riq1) and 2 (riq2), were identified in Arabidopsis (Yokoyama et al., 2016). They are encoded by paralogous AT1G74730 (RIQ2) and AT5G08050 (RIQ1) genes, which code for proteins of unknown function (DUF1118) and “affect the organization of light-harvesting complex II and grana stacking in Arabidopsis.” The BnMicEmUp gene, homologous to AT1G74730, was identified in a screen of upregulated genes in canola microspore cultures (Pauls et al., 2006; Shahmir & Pauls, 2021). It was shown that in Arabidopsis RNAi transgenics in which BnMicEmUp expression was suppressed, seed production was reduced (Shahmir, 2014). Bioinformatic analyses of the structure of the protein encoded by BnMicEmUp suggest that it is a leucine zipper type transcription factor localized in the chloroplast (Shahmir & Pauls, 2021). A model for a dual targeting of BnMicEmUp protein and its involvement in abscisic acid (ABA)-mediated transcriptional regulation has been proposed. Because of the negative correlation between the gene expression and yield related characteristics, this gene was characterized as an antiyield gene.

A homolog to this gene was identified in the common bean genome sequence annotation v1.0 (Phvul.009G190100) on chromosome Pv09. Preliminary results identified that the level of expression of this gene was negatively correlated with yield potentials in a field test of 10 navy bean cultivars (Qi, 2015). This observation is consistent with the suggestion that Phvul.009G190100 can also be characterized as an antiyield gene. The objectives of this research were to characterize in silico orthologs of AT1G74730 (antiyield) and its homolog, AT5G08050 genes, develop gene-based markers for these genes and test their usefulness for predicting the yield potential of beans from different market classes.

## MATERIALS AND METHODS

### 2.1 An in silico analysis of yield/antiyield genes and proteins

#### 2.1.1 Analysis of yield/antiyield gene sequences

The sequence of the Arabidopsis gene AT1G74730 (homolog of BnMicEmUp; Shahmir, 2014; Shahmir & Pauls, 2021) was used in BLAST searches for homologous genes in the genomes of common bean (P. vulgaris genome v2.1), soybean (Glycine max Wm82.a2.v1), and Arabidopsis (Arabidopsis thaliana TAIR10 [Columbia]) using the default filter criteria (BLASTN [nucleotide query to nucleotide db] with 11 word [w] length, BLOSUM62 as comparison matrix, and an expect [E] threshold of $\leq 1$ with allowed gaps) in Phytozome (Goodstein et al., 2012). The best matches were analyzed using an ab initio gene prediction and Arabidopsis/common bean/soybean specific gene-finding parameters in Fgenesh (Salamov & Solovyev, 2000). Gene structure was visualized in Gene Structure Display Server (GSDS; Hu et al., 2015).

Synten analysis was performed using the Legume Information System (LIS, Dash et al., 2016) and SoyBase (Grant et al., 2010). Expression data for antiyield genes in Arabidopsis, common bean, and soybean were retrieved from Phytozome.

Extended untranslated regions (UTRs; 1,500-bp upstream flanking sequence and 1,000 bp downstream flanking sequence, respectively) of the genes were downloaded from Phytozome and analyzed in silico for potential 5’ UTR cis-regulatory sequences using the PlantCARE...
2.1.2 Protein analyses

Multiple sequence alignments (MSAs) of yield/antiyield protein sequences were performed in Clustal Omega (Sievers et al., 2011). Final MSA outputs were produced in BoxShade 3.21 at ExPASy (Swiss Institute of Bioinformatics) and edited manually.

Conserved motif and domain scans were performed on several platforms, including Conserved Domain Database (CDD, using default parameters) at NCBI-NIH (Marchler-Bauer et al., 2015), MEME suite 5.2.0 (with searches for 20 motifs of 3–100 amino acids wide), and the MyHits at the SIB (Pagni et al., 2007).

Transmembrane helices in proteins were predicted in TMHMM Server v.2.0 at DTU Bioinformatics (Technical University of Denmark; Krogh et al., 2001) and protein fold recognition server Phyre2 (Kelley & Sternberg, 2009). Protein localization was analyzed at the DTU Bioinformatics with ChloroP 1.1 Server (Emanuelsson et al., 1999), SignalP 5.0 Server (Petersen et al., 2011), and TargetP 2.0 Server (Emanuelsson et al., 2000). Serine, threonine, or tyrosine phosphorylation sites were predicted at the DTU Bioinformatics with NetPhos 3.1 server (Blom et al., 2004).

2.2 Development of yield/antiyield gene-based TSP markers

2.2.1 Sequencing yield/antiyield genes

The genomic sequences (plus 1.5-kb upstream and 1-kb downstream) of two common bean yield/antiyield genes (Phvul.009G190100 and Phvul.009G202100) were downloaded from Phytozome. Primers, designed to overlap and cover the complete sequences, were used in polymerase chain reaction (PCR) reactions with genomic DNA from five diverse common bean genotypes, including two small seeded Mesoamerican (Lighthouse and T9905, both white [W]) and three large seeded Andean (Dynasty [dark red kidney, DRK], Yeti [white kidney, WK], and ACUG 13-C1 [cranberry, C]) beans. All primers were designed using Primer3 (Koressaar & Remm, 2007; Unterager et al., 2012) and/or Primer-BLAST (Ye et al., 2012) with the following characteristics: length 19–27 bp, melting temperature ($T_m$) 55–65°C, 40–60% GC content, 1°C maximum $T_m$ difference, one GC clamp, and maximum of three identical (poly-X) bases. Secondary structure of the primers was analyzed in IDT’s Oligo Analyzer Version 3.1 (Integrated DNA Technologies, Inc.).

Genomic DNA was isolated from young leaves from the plants grown in a controlled environment room (16-h photoperiod, 22°C) with DNeasy Plant Mini Kit (Qiagen, Canada) following manufacturer’s instructions. PCRs were performed in 20-μl volumes containing 1x PCR buffer and 3 mM MgCl₂ (both supplied with enzyme), 0.1 mM each of dNTPs (Invitrogen, Fisher Scientific, Canada), 0.6 U JumpStart Taq DNA Polymerase (Sigma, Burlington, ON, Canada), 0.25 μM each of the forward and reverse primer, and 25 ng of common bean genomic DNA, with a Bio-Rad MyCycler™ Thermal Cycler System (Bio-Rad, Mississauga, Canada). The amplification program consisted of an initial 3-min denaturation step at 94°C, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 45 s, and extension at 72°C for 1 min, with a final extension at 72°C for 10 min.

The PCR products were separated on 1% (w/v) agarose gel (with ethidium bromide added) in 1x TBE buffer at 100 V for 1–2 h. Single-band PCRs were purified (QIAquick PCR purification kit, Qiagen, Canada; Diffinity RapidTip, Sigma, Canada) and used as a template for in-house sequencing (3500 Series Genetic Analyzer [Applied Biosystems, Thermo-Fisher Scientific-CA] using a BigDye™ Terminator chemistry). Sequences were processed with BioEdit sequence alignment editor (Hall, 1999) and assembled using a CAP3 sequence assembly software (Huang & Madan, 1999). Complete sequences were aligned (Clustal Omega) and searched for single nucleotide polymorphism (SNP) differences among five genotypes in two yield/antiyield genes. Gene sequences were submitted to GenBank under accession numbers MW341245–MW341254.

2.2.2 Development of yield/antiyield gene-based TSP markers

SNPs identified in AYD1 and AYD2 gene sequences were visualized using a three-primer temperature switch PCR (TSP) technique (Hayden et al., 2009; Tabone et al., 2009). The assay consisted of a pair of a gene (locus)-specific primers (GS) and a nested allele-specific (AS) primer that is designed to assay for SNP. Forward and reverse GS primers were designed with melting temperature between 60°C and 65°C (63°C optimum), PCR product size greater than 400 bp, and located at least 100 bp from the SNP. A single AS primer (forward) was designed with the melting temperature between 43 and 48°C (45°C optimum) with 3’ end ending at the SNP. A short non-complementary stretch of 2 bp was added to the 5’-tail to increase melting temperature to 53°C.

- **AYD1-TSP (AYD1 gene, Phvul.009G190100) assay:**
  - AYD1-GS (GS) forward primer: 5'-TCGCCATAACAAAACATCAAGTGGACT-3'  
  - AYD1-GS (GS) reverse primer: 5'-CCTGTTGGTAAATATAGAGTTGGAG-3'  
  - AYD1-AS (AS) forward primer: 5'-GGAATATATATCTTTAGATAAAA GATTTTG-3'

- **AYD2-TSP (AYD2 gene, Phvul.009G202100) assay:**
  - AYD2-GS (GS) forward primer: 5'-CAGTGCAAGACGTCGGA-3'  
  - AYD2-GS (GS) reverse primer: 5'-CAGTACCTTTGGTACTCTTCA-3'  
  - AYD2-AS (AS) forward primer: 5'-ATCAACCATTGTGAAAAAT-3'.
The TSP assays were performed in a 20-μl reaction volumes containing 1× PCR buffer and 3-mM MgCl₂ (both supplied with enzyme), 0.1 mM each of dNTPs (Invitrogen), 125-ng bovine serum albumin Fraction V (Sigma-Aldrich), 0.6-U JumpStart Taq DNA Polymerase (Sigma-Aldrich), 0.05 µM each of the forward and reverse GS primers, 0.5-µM AS primer, and approximately 25 ng of common bean genomic DNA.

Amplification was done with the Bio-Rad MyCycler™ Thermal Cycler System (Bio-Rad) in 35 cycles following an initial denaturation step at 94°C for 5 min. The TSP amplification profile consisted of 15 cycles of 95°C for 30 s, 58°C for 30 s, and 72°C for 60 s (Phase 1—enrichment of GS product), followed by 5 cycles of 95°C for 10 s and 45°C for 30 s (phase 2—incorporation of AS primer into enriched GS PCRs), and 15 cycles of 95°C for 10 s, 53°C for 30 s, and 72°C for 5 s (Phase 3—competitive amplification of GS and AS PCR products). The amplification was completed with a final extension at 72°C for 10 min. The TSP PCR products were separated on a 1% ultra-pure agarose (Invitrogen) gel (with ethidium bromide added) at 100 V in 1× TBE buffer for 1–2 h. The amplicons were visualized on ChemiDoc™ XRS system with the Image Lab™ software (BioRad).

### 2.3 Testing usefulness of yield/antiyield gene-based TSP markers

The usefulness of the TSP markers was tested with 42 bean genotypes from the University of Guelph bean registration performance trials including 14 large seeded color (Andean gene pool) and 28 navy (small seeded, Mesoamerican) beans (Figure S1). Beans were evaluated for yield and a number of yield-related traits in two separate experiments at two Ontario locations (Elora and Woodstock) in 2014 (Figure S1). Beans were planted as two-row plots, 6 m long and 1.2 m wide (60-cm row-to-row spacing). Colored bean trials were harvested in Woodstock on September 26 and in Elora on October 3. Standardized cultural practices were performed in both trials as needed.

Maturity (DM) was reported as the number of days from planting to physiological maturity. Yield (YD) was expressed as seed weight measured for each plot in kg ha⁻¹ and adjusted to 18% moisture content. Harvestability (HR) was determined as standability of plants in plot at maturity using a scale of 1–5 (1. erect plants suitable for combining; 5. prostrate plants not suitable for machine harvesting). A derived trait, yield gain per day (YGD) or economic growth rate (EGR), was calculated from the yield and maturity measurements (YGD = YD/DM). Average yield gain (kg day⁻¹ ha⁻¹) describes the seed weight accumulated per growing day from a unit area for each genotype (Scully & Wallace, 1990).

### 2.4 Weather conditions during 2014 bean growing season

Daily weather data for the two Ontario locations (Elora and Woodstock) in 2014 were collected at the nearest weather stations (available at: https://climate.weather.gc.ca/; Figure S2). Environmental variation was observed between the experimental locations during the growing season (June–October). However, due to the incomplete weather data for the Woodstock site, direct comparison between the two locations was not possible for the whole bean growing seasons. As expected, Woodstock was the warmer location resulting in earlier maturity.

### 2.5 Data analysis

Analysis of variance was performed using the Proc Glimmix procedure in Statistical Analysis System (SAS) v.9.4 software (SAS Institute, 2013). For each trial, data were analyzed separately for each location and combined across locations. Genotypes and locations were considered as fixed effects, and all other effects were considered as random. The homogeneity of error variances was tested before pooling data for combined analyses using a residual analysis in SAS. The relationships among yield and yield-related traits were analyzed by correlation using SAS. The association of yield/antiyield gene-based TSP markers with yield and yield-related traits was analyzed with the Proc Reg and Proc Corr in SAS using trait’s LS mean values (after Proc Glimmix, for each location separately as well as averaged over two locations) and marker scores (AYD1: T [497 bp = 2]/G [340 bp = 1]; AYD2: A [407 bp = 2]/T [255 bp = 1]).

### RESULTS

#### 3.1 Yield/antiyield genes and proteins in common bean, soybean, and Arabidopsis

BLAST searches with Arabidopsis gene AT1G74730 (homolog to BnMicEmUP, previously associated with yield-related traits in canola [Shahmir, 2014]) against the Phytozome P. vulgaris v2.1 genome sequence database confirmed the position of the antiyield gene (AYD1, Phvu.009G190100) on chromosome Pv09 but at slightly different position (28,837,232..28,838,084, forward) compared with its placement in earlier version (v1.0) of bean genome sequence (28,180,739..28,181,741) (Qi, 2015). A search of the protein database
yielded a single homology to Pfam: 06549, a protein of unknown function (DUF1118). A homologous gene (AYD2, Phvul.009G202100) was identified on the same chromosome 1,835,947 bp apart (position 30,674,031–30,675,960, forward) (Figure 1). These genes are orthologs with the two Arabidopsis genes located on chromosomes At1 (AYD1, AT1G74730) and At5 (AYD2, AT5G08050), respectively. Four homologous genes were identified in soybean: two on chromosome Gm06 (copy 1 [higher expression in all tissues (Phytozome)]: AYD1-1, Glyma.06G190300 (position 16,709,212–16,710,732, reverse) and AYD2-1, Glyma.06G204200 (position 19,170,925–19,173,352, forward) (2,460,193 bp apart)) and two genes on chromosome Gm04 [copy 2: AYD1-2, Glyma.04G174400 (position 43,601,717–43,604,198, forward) and AYD2-2, Glyma.04G161600 (position 40,088,117–40,090,295, reverse) (3,511,422 bp apart)].

In general, the regions of three genomes containing yield/antiyield genes are syntenic (Table S1). The AYD1 gene (Phvul.009G190100) on bean chromosome Pv09 is contained in a large block (27.73–28.98 Mb) that is syntenic with recently duplicated regions of soybean chromosomes Gm06 (16.58–17.99 Mb) and Gm04 (41.26–43.90 Mb) that contain genes Glyma.06G190300 (AYD1-1) and Glyma.04g174400 (AYD1-2), respectively (LIS). The second region on bean chromosome Pv09 (30.67–30.68 Mb) containing the AYD2 gene (Phvul.009G202100) is syntenic with the region on soybean chromosomes Gm04 (39.59–40.24 Mb), which holds the AYD2 gene Glyma.04G161600 (AYD2-2). However, there is no synteny with the region of soybean chromosome Gm06, which contains the AYD2 gene Glyma.06G204200 (AYD2-1) (Table S1).

In the current annotation of the common bean genome sequence (P. vulgaris v2.1), the antiyield gene AYD1 (Phvul.009G190100) is located close to a yield quantitative traits locus (QTL) SY9.1, while its homolog, gene AYD2 (Phvul.009G202100), is located in the region of the second yield QTL SY9.2 (LIS). Similarly, four antiyield genes in the soybean genome sequence annotation Wm82.a2.v1 are located in the regions of two seed yield (SY) QTL on chromosome Gm04 and eight yield QTL on chromosome Gm06 (SoyBase; Figure 1).

**FIGURE 1** Genome locations (Mb) of yield/antiyield genes AYD1 and AYD2 in common bean, Arabidopsis, and soybean. Arabidopsis chromosomes At1 and At5 are shown as patterned (cross hatch lines) bars, common bean chromosome Pv09 as a black bar, and soybean chromosomes Gm04 and Gm06 are shown as clear bars. Yield/antiyield genes are presented in red (larger font); AYD1 genes are underlined. Genome regions associated with previously identified yield QTL (black bars) in common bean and soybean are indicated (see text for details). [Color figure can be viewed at wileyonlinelibrary.com]
3.1.2 | Gene structure

All eight yield/antiyield genes have a similar structures (Figure 2a; Table S2), consisting of two exons and one intron with sizes varying from 84p (Phvul.009G190100) to 1,399 bp (Glyma.06G204200). The only exception is the soybean gene Glyma.04G1744000 on chromosome Gm04, which has an additional intron in the 3' UTR region. In all three species, Exon 2 in the AYD1 genes is the same size (444 bp), whereas the Exon 1 in the AYD2 gene is longer in Arabidopsis and common bean (153 bp) compared with the two soybean genes (141 bp). However, the AYD2 genes are more variable. Exon 2 is longer in common bean and soybean (423 bp) compared with Arabidopsis (378 bp). The size of Exon 1 is different in all three species (Table S2). The analysis of the 5' UTR sequences (1,500-bp upstream of ATG) identified a moderate degree of divergence in these regions of the eight yield/antiyield genes (36–66% identity). The degree of identity between the 3’ UTR (1,000 bp from the stop codone) regions ranged from 35% to 52%.

Based on the transcript sequence similarity, the yield/antiyield genes in the three species separate into two groups (Figure 2b). At the level of the coding regions (CDSs), the percent identity between the two groups was 43–45%. However, sequence conservation was higher within each group and ranged from 63% (Arabidopsis and legumes) to 93% (soybean).

3.1.3 | Potential cis-regulatory elements in structurally different upstream regions of yield/antiyield genes

An in silico analysis in PlantCARE database of the promoter regions (1,500-bp upstream of ATG) of the AYD1 and AYD2 genes in Arabidopsis, common bean, and soybean genomes identified 69 potential cis-regulatory elements. In particular, the analysis of the upstream sequence of yield/antiyield genes identified core promoter elements, including CAAT and TATA boxes, as well as a large number of light responsive elements (27), elements associated with defense and stress responsiveness (7), hormonal responsiveness (9), and tissue-specific expression (4). However, the functions of a significant number (12) of predicted cis-regulatory elements are unknown (Table S3).

Only six of the regulatory elements were found in all eight genes. In addition to the two core elements (TATA and CAAT boxes), they include a G-box (light responsiveness element), a TC-rich repeat (for defense and stress responsiveness), an ABRE box (ABA response), and Unnamed_4 (unknown function). Fifty-five elements were identified in AYD1 genes, and 50 regulatory elements were predicted in 5' UTR sequences of AYD2 genes.

Forty-five promoter elements were predicted for the common bean yield/antiyield genes: 21 were common, and 12 were specific for the AYD1 (Phvul.009G190100) and AYD2 (Phvul.009G202100) genes, respectively (Table S3; Figure 3). A few elements were unique to the common bean genes. They include a MYB binding site involved in light responsiveness (MRE), which was present in both genes, as well as four cis-regulatory elements that were specific for AYD1 gene (Gap-box [light responsiveness], Box III [protein binding site], CTAG-motif [unknown], and TCCACCT [unknown]) and two elements found only in promoter region of the AYD2 gene (ATC-motif [light responsiveness] and EIRE [elicitor-responsiveness]).

3.1.4 | Yield/antiyield proteins

Based on the protein similarities, the yield/antiyield gene homologs in three species separated into two groups that corresponded with the groupings based on transcript sequence similarities. There was 26–33% amino acid identity between the groups and 65–94% identity within the groups. The protein encoded by the common bean AYD1
gene has 198 amino acids with an estimated molecular weight of 20.88 kDa, a theoretical isoelectric point (IEP, pI) of 8.80, and consists predominantly of four amino acids, namely, alanine (Ala, A, 14.1%), leucine (Leu, L, 13.1%), serine (Ser, S, 11.1%), and valine (Val, V, 11.1%). The other three members of this group have similar characteristics (Table S4).

The proteins in both groups have transit peptides at their N-termini, domains of unknown function DUF1118 (pfam 06549), which are at positions 83–197 in AYD1 and at 85–196 in AYD2 sequences, respectively (Figures 4 and S3). Both proteins contain five casein kinase II phosphorilation sites with a consensus sequence of S/TxxD/E at positions 91–131 in AYD1 and 84–124 in AYD2 sequences, respectively. Both also contain two transmembrane domains at their C-ends (in positions 140–159 and 169–191, respectively).

3.2 | Development of yield/antiyield gene-based TSP markers

The two types of yield/antiyield genes were sequenced in five common bean genotypes, including two small seeded Mesoamerican beans (white cultivars Lightning and T9905) and three large seeded Andean genotypes (Dynasty [DRK], Yeti [WK], and ACUG 13-C1 [cranberry]).

3.2.1 | AYD1 gene (Phvul.009G190100) sequence

In total, 2,807 bp of the AYD1 genes (Phvul.001G190100) in five bean genotypes were sequenced. There were no differences in the coding and intronic sequences for this gene among the analyzed genotypes, and they were identical to the Phvul.009G190100 sequence in the reference genome (G19833, Andean gene pool). The only differences that were detected were three single base substitutions in the untranslated regions. A single T to G change (at position 839 nt in the reference) was identified in 5′ UTR region of the two Mesoamerican white beans compared with the reference G19833 and other three Andean genotypes (Table 1). Two single base changes were identified in the 3′ UTR: an A to T change in the white bean T9905 (2,558 nt) and a T to C change in the cranberry bean ACUG 13-C1 (2,599 nt).

3.2.2 | AYD2 gene (Phvul.009G202100) sequence

Sections, 3,130 bp in length, of the AYD2 genes (Phvul.001G202100) were sequenced in five bean genotypes and compared with the reference genome (G19833, Andean gene pool). Fourteen single base changes were identified in the AYD2 gene sequences. Eight of the sequence differences were consistently found between Andean and Mesoamerican beans (one in 5′ UTR, one in Exon 1 and six in 3′ UTR). A single base change (Exon 1) was found in all five sequenced genotypes when compared with the reference G19833 sequence. Five additional changes were identified only in one of the two white genotypes (one in Lighthouse [5′ UTR] and four in T9905 [two in intron, one in Exon 2 and one in 5′ UTR]) and therefore not specific to the Mesoamerican gene pool (Table 1).

Three of these substitutions were identified within CDSs of the AYD2 gene. A single C to T change (2,916 nt in G19833) in Exon 2 of the AYD2 gene in the white genotype T9905 resulted in a conservative codon change (ACG to ACT) for threonine (T, Thr) at position 61 but did not change AYD2 protein in this genotype. However, two T to A changes that were observed in Exon 1 could affect the
properties of the proteins in the analyzed beans. Variation at position 1,844 nt (position in G19833) that was found in Exon 1 of the five sequenced genotypes changed the codon for serine (S, Ser [TCT] in G19833) at position 56 to the codon for threonine (T, Thr [ACT]). However, both are neutral-polar, hydrophylic amino acids. A single nucleotide difference identified at 1,713 nt (Exon 1, position 35) changed the codon (AAT) for the hydrophobic (aliphatic) amino acid isoleucine (I, Ile) at position 12 compared with the reference genotype (G19833, Andean gene pool) and three sequenced Andean genotypes to the codon (AAC) for neutral-polar (amidic) asparagine (N, Asn) in two Mesoamerican beans (Table 1 and Figure S4). These substitutions resulted in slight changes in protein properties in analyzed beans compared with the reference G19833 (Table S5).

3.2.3 Yield/antiyield gene-based TSP markers

A gel-based TSP marker (AYD1m) was developed on the basis of the single T to G substitution identified in the 5’ UTR (position −871 in G19833) of the Phvul.009G190100 (AYD1) gene sequence. The AYD1m PCR product of 497 bp was amplified from sequences with the reference (G19833) allele (T) and colored, large-seeded genotypes (Table 2), whereas the alternate allele (G), resulting in a 340-bp fragment (SNP), was observed in the navy beans (Figure 5).

The second TSP marker (AYD2m) was developed for the T to A change identified in Exon 1 (position 35 in G19833) of the AYD2 gene (Phvul.009G202100). The marker was developed for the T allele observed in the reference (G19833) genotype. The AYD2m PCR
product of 255 bp (T allele) was identified in reference (G19833) and large seeded colored genotypes, whereas the 407-bp amplicon (A allele, SNP) was present in the navy beans (Figure 5).

### 3.2.4 Validation of yield/antiyield gene-based TSP markers

As a part of Ontario bean registration trials, a set of 42 diverse bean genotypes were evaluated for yield and a number of yield-related traits in trials at two locations (Elora and Woodstock) in 2014. Significant variation was identified for all analyzed traits in navy and colored bean genotypes. For example, yield (averaged over two locations) in navy beans ranged from 2,650 (OAC Thunder) to 3,609 kg ha$^{-1}$ (Lighthouse) (Table S6) and in large seeded colored beans from 2,119 (ACUG 12-D1) to 3,111 kg ha$^{-1}$ (OAC Inferno) (Table S7).

Amplification of both yield/antiyield gene-based TSP markers was successful in most of the analyzed genotypes (except AYD1m in genotype ACUG 14-2 and AYD2m in three genotypes [HMS Medalist, ACUG 14-2, and ACUG 12-D1]). In marker AYD1m (AYD1 gene,
Phvul.009G190100, G allele (340 bp, T to G mutation) was associated with a high yield (Mesoamerican, navy beans), whereas the T allele (497 bp, wild type G19833) was linked to a low yield (Andean, large seeded colored beans). In AYD2m (AYD2 gene, Phvul.009G202100) marker, the T allele (255 bp, wild type G19833) was associated with a low yield (large seeded colored beans, Andean gene pool), and the A allele (407 bp, mutation T to A) was connected to a high yield (navy beans, Mesoamerican gene pool) (Table S8).

Both TSP markers were significantly associated with yield, maturity, YGD (derived from yield and maturity measurements), and HR (Table 3). Averaged over two locations, marker AYD1m (AYD1 gene, Phvul.009G190100) explained more variability for HR (30% vs. 28%), whereas the AYD2m marker (AYD2 gene, Phvul.009G202100) explained more variation for yield (56% vs. 38%) in a set of 42 bean genotypes. Both markers explained similar variability for maturity (34%).

Significant correlations were identified between analyzed traits as well as traits and yield/antiyield gene-based TSP markers (Table 4). Yield was negatively correlated with HR (a lower value indicates better standability) but was in positive association with maturity. There was negative correlation between two TSP markers, and they showed opposite associations with the analyzed traits. The AYD1m (AYD1 gene, Phvul.009G190100) marker was positively correlated with HR but negatively with maturity, yield, and YGD (trait derived from the maturity and yield measurements). Conversely, AYD2m (AYD2 gene, Phvul.009G202100) marker was negatively correlated with HR but was positively association with maturity, yield, and YGD (Table 4). Therefore, Phvul.009G202100, AYD2 gene can be characterized as a yield gene.

4 | DISCUSSION

4.1 | The problem being addressed

Yield is the primary criterion on which breeding lines are judged during selection in the University of Guelph dry bean breeding program and during the registration process administered by the Ontario Pulse Crop Committee. In the face of rising production costs (inputs [seed and fertilizer], equipment, labor, and land), bean producers in Ontario need high yielding cultivars in order to remain profitable. Bean
### TABLE 3  
Trait variation explained by the yield/antiyield gene-based temperature switch PCR (TSP) markers in a set of 42 diverse bean genotypes

| Trait                          | Environment | Marker R² (%) | AYD1m (Phvul.009G190100) | AYD2m (Phvul.009G202100) |
|--------------------------------|-------------|---------------|--------------------------|--------------------------|
| Harvestability (HR, Scale 1–5) | 2014        | 29.6          | 28.3                     |                          |
|                                | ERS14       | 36.0          | 45.2                     |                          |
|                                | WRS14       | 18.9          | 17.3                     |                          |
| Days to maturity (DM, days)    | 2014        | 34.2          | 34.9                     |                          |
|                                | ERS14       | 14.4          | 12.1                     |                          |
|                                | WRS14       | 45.5          | 49.7                     |                          |
| Yield (YD, kg ha⁻¹)            | 2014        | 38.2          | 56.3                     |                          |
|                                | ERS14       | 36.5          | 49.7                     |                          |
|                                | WRS14       | 30.4          | 45.9                     |                          |
| Yield gain per day (YGD = YD/DM, kg day⁻¹ ha⁻¹) | 2014 | 32.9 | 51.6 |                       |
|                                | ERS14       | 31.4          | 44.6                     |                          |
|                                | WRS14       | 23.5          | 38.6                     |                          |

Abbreviation: PCR, polymerase chain reaction.

*Environments: 2014, combined analysis over two locations in 2014; ERS14, Elora research station 2014; WRS14, Woodstock research station 2014.

**Marker R² (%), percent of trait variation explained by a marker (calculated using Proc Reg in SAS—trait’s LS mean values and marker scores).**

### TABLE 4  
Correlations between analyzed traits and yield/antiyield gene-based temperature switch PCR (TSP) markers identified in a set of 42 common bean genotypes evaluated in two locations

| Environment | Traits | DM | YD | YGD | TSP markers |
|-------------|--------|----|----|-----|-------------|
|             |        |    |    |     | AYD1m | AYD2m |
| 2014        | HR     | −0.395** | −0.374* | −0.350* | 0.544** | −0.532** |
|             | DM     | 0.633** | 0.515** | 0.693** | −0.585** | 0.591** |
|             | YD     | 0.988** | 0.743** | 0.967** | −0.618** | 0.750** |
|             | YGD    |       |     |     | −0.574** | 0.719** |
| ERS14       | HR     | −0.357* | −0.399** | −0.356* | 0.600** | −0.672** |
|             | DM     | 0.351* | 0.193d | 0.348* | −0.379* | 0.348* |
|             | YD     | 0.982** |       | 0.705** | −0.604** | 0.705** |
|             | YGD    |       |     |     | −0.560** | 0.668** |
| WRS14       | HR     | −0.285d | −0.346* | −0.339* | 0.434** | −0.415* |
|             | DM     | 0.665** | 0.548** | 0.678** | −0.684** | 0.705** |
|             | YD     | 0.989** | 0.551** | 0.678** | −0.684** | 0.705** |
|             | YGD    |       |     |     | −0.484** | 0.621** |
| AYD1m       |       | −0.941** |       |     |             |

Note: The correlation between traits and yield/anyitield gene-based TSP markers was calculated using Proc Corr in SAS—trait’s LS means (after Proc Glimmix, for each location separately as well as averaged over two locations) and marker scores for each genotype {1 or 2 [AYD1m: T allele (497 bp = 2)/G allele (340 bp = 1); AYD2m: A allele (407 bp = 2)/T allele (255 bp = 1)]}.  
*Environments: 2014, combined analysis over two locations in 2014; ERS14, Elora research station 2014; WRS14, Woodstock research station 2014.  
**TSP markers: AYD1m, temperature switch PCR (TSP) marker for the AYD1 gene, Phvul.009G190100; AYD2m, TSP marker for the AYD2 gene, Phvul.009G2020100.

**Not significant.**  
* Significant at 0.05 level.  
** Significant at 0.01 level.
improvement programs have successfully used conventional plant breeding techniques to develop and deploy dry bean cultivars with increasing seed yield potentials at a steady rate for several decades (Kelly et al., 1998). However, the development of new high yielding cultivars is challenging because of the complexity of the trait. Yield is controlled by many genes that affect a number of related traits, including maturity, seed number, and seed size. In addition, trait expression is highly influenced by environment. The use of MAS permits selection in absence of selection pressure and has become a common tool in many common bean breeding programs (Kelly et al., 2003; Miklas et al., 2006), especially for disease resistance.

Arabidopsis transformed with an antisense construct derived from the canola BnMicEmUP gene had enhanced seed production (Shahmir, 2014). Preliminary work showed that an ortholog to BnMicEmUP exists in dry bean, and the expression level of this gene (Phvul.009G190100) was negatively (but not significantly) correlated with the yield in a selected set of 10 navy beans (Qi, 2015). Therefore, both genes can be defined as antiyield genes. Because the genes that increase yield with fixed resources have their effects by increasing input use efficiencies, the purpose of this work was to characterize (in silico) antiyield gene in bean and develop gene-based markers.

4.2 Yield/antiyield genes identified on common bean chromosome Pv09

The rationale for selecting soybean (G. max Wm82.a2.v1) and Arabidopsis (A. thaliana TAIR10) genomes for comparative analyses with common bean (P. vulgaris v2.1) is that, in addition of being a model dicot plant, Arabidopsis mutants of AT1G74730, the homologous gene to BnMicEmUP, were used to characterize the function of BnMicEmUP (Shahmir, 2014; Shahmir & Pauls, 2021), and soybean is closely related to common bean, and the genomes of the two species are highly syntenic (Galeano et al., 2009; 2011; McClean et al., 2010; Reinprecht et al., 2013).

A search for a homologous sequence confirmed the location of the Phvul.009G190100 (AYD1) gene on the chromosome Pv09 but in a slightly different location compared with its placement in the common bean genome v1.0 (Schmutz et al., 2014), which was initially used to identify and characterize it as an antiyield gene (Qi, 2015). Its position would likely change with the future updates to high-quality genome assembly/annotations. The current finding that the region on bean chromosome Pv09 containing the yield/antiyield genes is syntenic with regions on soybean chromosomes Gm04 and Gm06 agrees with previous studies of syntenic relationships between common bean and soybean chromosomes (Galeano et al., 2011; McClean et al., 2010; Reinprecht et al., 2013). In particular, the relationship between AYD1 (Phvul.009g190100) and AYD2 gene (Phvul.009g202100), which are located 1.8 kb apart on chromosome Pv09 in bean, is preserved in the two syntenic regions in soybean chromosomes Gm04 an Gm06, which agrees well with the current understanding that soybean underwent a whole genome duplication event after the divergence of beans and soybeans (Schmutz et al., 2010).

Six genes (two in common bean and four in soybean) examined in the current study are located in regions to which Quantitative Trait Locus/Loci (QTL) associated with yield or yield-related traits were previously mapped. In bean, the physical locations of the yield/antiyield genes on chromosome Pv09 correspond with the locations of yield QTL previously identified on this chromosome (Blair et al., 2006; Tar'an et al., 2002). In a more recent QTL study, Mukeshimana et al. (2014) mapped two yield QTL to the same location on chromosome Pv09. In particular, the AYD1 gene Phvul.009G190100 of the current study is close to the yield QTL SY9.1 (yield reduction), whereas its homolog Phvul.009G202100 (AYD2 gene) is located near to the second yield QTL SY9.2 (yield increase), which agrees with their classification as antiyield (Phvul.009G190100) and yield (Phvul.009G202100) genes, respectively. Similarly, the four yield/antiyield genes in the soybean genome sequence (Wm82.a2.v1) are located in the regions of two SY QTL on chromosomes Gm04 (SY23-4 [Guzman et al., 2007] and SY25-1 [Palomeque et al., 2009]) and eight overlapping yield QTL on chromosomes Gm06 (SY5-1 [Renyna & Sneller, 2001], SY7-1 [Orf et al., 1999], SY19-1 and SY19-2 [Zhang et al., 2004], SY21-1 [Reinprecht et al., 2006], SY22-12 [Du et al., 2009], SY25-6 [Palomeque et al., 2009], and SY28-3 [Rossi et al., 2013]), respectively, mapped previously in different populations (SoyBase).

The high nucleotide homologies and similar gene and protein structures suggest that these genes are orthologous (share common ancestor) and belong to the same gene family characterized by a chloroplast transport peptide and a protein of unknown function (DUF1118) domain. All the genes had their highest levels of expression in green tissues (Phytozome), where most chloroplasts are found, suggesting the chloroplast as the common localization. Two yield/antiyield genes (AYD1, antiyield gene Phvul.009G190100, and AYD2, yield gene Phvul.009G202100) in common bean are paralogous. They were separated by a gene duplication event (Schmutz et al., 2014) and occupy different positions in the bean genome, both on chromosome Pv09. Although paralogs typically have the same or similar function, their function is not always conserved. Based on sequence conservation (single SNP in 5' UTR detected in this study) and higher expression in all analyzed tissues (Phytozome), Phvul.009G190100 (AYD1, antiyield gene) is the primary copy of the gene. Lacking the original selective pressure of the primary gene, Phvul.009G202100 (AYD2, yield gene) was free to mutate (eight SNPs identified in this study) and acquire new function(s). This may explain their opposing effect on yield detected in 42 bean genotypes (both Andean and Mesoamerican gene pool) that were analyzed in the current study.

The phylogenetic analysis, which grouped the four proteins encoded by the AYD1 genes, including Arabidopsis At1G74730 (RIQ2), bean Phvul.009G190100, soybean Glyma.04G174400, and Glyma.06G190300, and grouped the four AYD2 proteins, including Arabidopsis At5G08050 (RIQ1), bean Phvul.009G202100, soybean Glyma.04G161600, and Glyma.06G204200 in separate clusters, is in agreement with the previous division of the Arabidopsis RIQ genes into two clusters by Yokoyama et al. (2016). They suggested that the division corresponded to the ancient gene duplication event in
Arabidopsis and further proposed that because these genes are unique to land plants, they have a conserved function “to optimize the efficiency of light harvesting and to determine grana structure, which might play an advantageous role in survival under the severe growth conditions on land.”

The in silico modeling and bioinformatic analyses of the structure of the proteins encoded by canola BnMicEmUP genes and the homologous Arabidopsis gene (AT1G74730) indicate that they are specific leucine zipper type transcription factors called PEND proteins, first described in pea (Terasawa & Sato, 2009). Transgenic Arabidopsis plants, which overexpress AT1G74730 or BnMicEmUP, had lower SY than nontransformed control plants, and transgenics, which had reduced levels of gene expression, had increased yield (Shahmir, 2014).

The yield/antiyield genes are highly expressed in leaves (Phytozome). In all three species, the expression of the AYD1 gene was higher compared with the AYD2 gene. In beans, the expression of the AYD1 gene Phvul.009G190100 was 77% higher compared with its homologous AYD2 gene Phvul.009G202100 (Phytozome). Qi (2015) identified a negative trend between the expression of AYD1 gene Phvul.009G190100 and yield in a study with 10 navy beans evaluated over two locations. However, the correlation was not significant. That study was restricted to a small number of genotypes belonging to the same gene pool (Mesoamerican)/market class (navy). Because of its high expression in leaves and localization in chloroplasts, the previous quantitative PCR (qPCR) study was performed with leaf tissue. However, if seed were used instead of leaf tissue for the qPCRs, different results might have been obtained.

The multiple TATA and CAAT box sequences found upstream of the transcription start site in the yield/antiyield genes are core promoter elements important for accurate initiation of transcription (Kozak, 1981) that may function in promoting high levels of transcription of these genes and may also facilitate differential regulation of their activity in response to different environmental conditions (Doyle & Han, 2001). In particular, the stress responsiveness of the yield/antiyield genes is supported by the occurrence of different types of stress-related elements (ABRE elements) in the promoters of these genes. ABRE cis-acting elements are found in ABA-responsive genes and regulate gene expression in response to abiotic stress (Mundy et al., 1990; Xu et al., 1996). Shahmir and Pauls (2021) found that BnMicEmUP3 overexpression in Arabidopsis increases the sensitivity of seedlings to exogenous ABA. The environmental sensitivity of the yield/antiyield genes is further supported by the occurrence of the heat-shock cis regulatory elements (HSEs) in their promoters. Thermosensitivity of the BnMicEmUP gene homolog of Phvul.009G190100 was observed during embryogenesis in induced B. napus microspore cultures (Shahmir, 2014). In addition, cis elements identified in the promoter of the Phvul.009G190100 gene, such as the TGACG-motif and ERE, suggest that it is also regulated by jasmonic acid and ethylene. Stress-related and many defense-related genes respond to jasmonates (Creelman & Mullet, 1997) and ethylene (Müller & Munné-Bosch, 2015). Interestingly, transcription factors, such as basic leucine zipper (bZIPs) (Shinozaki et al., 2003), with domains similar to the Phvul.009G190100 protein, are involved in regulating the expression of genes with stress-related cis-acting elements in their promoters, making it possible that the yield/antiyield genes are autoregulatory. Shahmir and Pauls (2021) found that Arabidopsis transformed with 35S::BnMicEmUP had enhanced sensitivity to ABA inhibition of seed germination, whereas transgenics with anti-BnMicEmUP were less sensitive to ABA.

The current study also shows that the upstream sequence of Phvul.009G190100 gene contains many light responsive cis elements, such as the G-Box, suggesting that its expression may be regulated by light, which may be related to the chloroplast targeting of the protein. Interestingly, some regions in the promoter may bind transcription factors in a competitive manner because some of the cis-acting regulatory elements are overlapping, including Box-4 and ERE. Overall, these results suggest that it would be interesting to test the expression of yield/antiyield genes (Phvul.009G190100 and Phvul.009G202100) under a variety of stress and light regimes and to examine the role of the gene product in mediating responses to hormones and stresses. Such studies might provide interesting links between development, stress, and yield.

### 4.3 | Yield/antiyield gene-based TSP markers are associated with yield in common bean

The selection of bean genotypes from both gene pools was an appropriate approach to discriminate sequence variability of the two yield/antiyield genes. The analyses also showed that the AYD1 gene (Phvul.009G190100) was more conserved than the AYD2 gene. The only SNP that was discovered in the AYD1 gene was a single T to G substitution at the 5′ UTR (at position –847 nt from ATG in G19833) in navy beans (Mesoamerican) compared with the reference G19833 and three large seeded beans (Andean). The polymorphism is located in a light responsive Box I (TTTCAAA) element and an ERE (ATTCCAA) motif that is involved in ethylene responsiveness of genes in plants. The sequence of the AYD2 gene (Phvul.002G202100) was more variable among five genotype, with eight single nucleotide changes identified including a T to A substitution in Exon 1 that would result in the replacement of a hydrophobic amino acid (isoleucine) found in the large seeded beans and normally buried inside the protein core with a polar amino acid (asparagine) in two white (small seeded) beans. This change would likely affect protein properties.

Most of the lower yielding large seeded colored beans (Andean) had a wild-type (G19833) allele in both yield/antiyield genes. The only exception were two Andean kidney genotypes (ACUG 12-D1 and Yeti), which had mutation in AYD1 gene and wild-type allele in AYD2 gene and lower yield. This study showed that the mutation in both yield/antiyield genes (AYD1, antiyield gene Phvul.009G190100, and AYD2, yield gene Phvul.009G202100) contribute to the high yield in navy beans (Mesoamerican).

The SNPs were visualized using a gel-based TSP approach that consists of a biphasic PCR with three primers, including a set of locus-specific primers and a nested AS (SNP) primer. The original TSP assay
(Hayden et al., 2009; Tabone et al., 2009) was slightly modified to use a 1% agarose gel with added ethidium bromide. The sample DNA was dissolved in water, and standard PCR reagents were used in a three-primer system with only one AS primer. A 10-fold concentration of the AS primer compared with the GS primers was found to be optimal for the yield/antiyield gene-based TSP markers. Both markers provided accurate codominant differentiation of reference and SNP allele. Wild type (reference G19833) alleles were associated with the low yield of the large seeded colored beans (Andean gene pool), whereas the SNPs were linked to the high yield in navy beans (Mesoamerican gene pool).

The current measurements of yield in beans support many studies that show that this trait is highly variable and that absolute values fluctuate from year to year and among plots established in different locations. However, the positive correlation identified between the yields in two locations in the current study indicates that, although the genotype × environment is often significant, the overall trends and yield ranks of the lines were relatively stable.

The development of functional molecular markers for candidate genes underlining significant QTL for traits of interest can speed up and simplify breeding for traits. For example, Reinprecht et al. (2009) developed gene-based markers for mutations in two \( \omega-3 \) desaturase genes (Fad3A and Fad3B) that explained significant variation in seed linolenic acid level in soybean seed. In barley, Watt et al. (2020) developed a functional marker based on 9-bp indel at the position –84 in the 5′ UTR of HvDEP1 gene associated with the length and weight of the seed. This gene encodes a \( \gamma \)-subunit of the heterotrimeric G-protein complex and was identified previously as a candidate gene for a major QTL underlying grain length in barley.

In the current study, both yield/antiyield gene-based markers that are near yield QTL (marker AYD1m for AYD1 gene, Phvul.009G190100) or in yield QTL (marker AYD2m for AYD2 gene, Phvul.009G202100) identified previously on chromosome Pv09 (Mukeshimana et al., 2014) were significantly associated with yield in a set of 42 diverse common bean genotypes. A number of flowering and maturity QTL are also located in the same region on the chromosome Pv09 (Blair et al., 2006). In addition, a seed weight QTL (SW9.1) is located in the top arm of this chromosome (Blair et al., 2006). Therefore, the continuation of the research should include evaluation of these traits (in addition to the yield) in replicated trials with a larger set of more diverse bean genotypes.

5 CONCLUSIONS

Genes that increase yield with fixed resources have their effects by increasing input use efficiencies. We have discovered genes in bean (Phvul.009G190100 and Phvul.009G202100), which are homologous to a canola antiyield gene that are associated with yield and yield-related traits in bean. Both genes are near or in yield QTL regions identified previously in bean.

In silico analyses performed in this work suggested that Phvul.009G190100 and Phvul.009G202100 encode cbZIP transcription factors targeted to chloroplasts involved in stress responses. This study is an example of how examinations of genes with unknown function in crop plants may lead to discoveries that are agronomically relevant and ultimately to the crop improvement.

In addition, simple codominant molecular markers were developed for both yield/antiyield genes based on the TSP technique that can be used by any lab equipped for basic PCR and electrophoresis equipment.

ACKNOWLEDGMENTS

Research funding that was provided by the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA), Agriculture and Agri-Food Canada (AAFC), Ontario Bean Growers (OBG), the Natural Sciences and Engineering Research Council of Canada (NSERC), and Food from Thought from University of Guelph is highly appreciated. We acknowledge the assistance of a number of summer students with field measurements.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this study.

AUTHOR CONTRIBUTIONS

Karl Peter Pauls conceptualized this research and supervised the work. Thomas H. Smith designed and conducted field experiments. Yarmilla Reinprecht conducted lab experiments, analyzed data, and prepared the original manuscript draft. Fariba Shahmir did initial work in canola (PhD dissertation) and assisted in Yanzhou Qi’s bioinformatic analysis in bean. Yanzhou Qi did initial work in beans (MSc thesis). All authors were involved in the manuscript revision.

ETHICAL STATEMENT

This research does not contain any studies with human or animal subjects.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request.

ORCID

Yarmilla Reinprecht https://orcid.org/0000-0003-1554-7626

Karl Peter Pauls https://orcid.org/0000-0003-0636-9151

REFERENCES

Atwell, S., Huang, Y., Vilhjálmsson, B., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A. M., Hu, T. T., Jiang, R., Muliyati, N. W., Zhang, X., Amer, M. A., Baxter, I., Brachi, B., Chory, J., Dean, C., Debieu, M., Nordborg, M. (2010). Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature, 465, 627–631. https://doi.org/10.1038/nature08800

Bailey, T. L., Bodén, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. & Noble, W. S. (2009). MEME SUITE: Tools for motif discovery and searching. Nucleic Acids Research, 37, W202–W208. https://doi.org/10.1093/nar/gkp335
Blair, M. W., Iriarte, G., & Beebe, S. (2006). QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean x wild common bean (Phaseolus vulgaris L) cross. Theoretical and Applied Genetics, 112, 1149–1163. https://doi.org/10.1007/s00122-006-0217-2

Blom, N., Sichertz-Ponten, T., Gupta, R., Gammeltoft, S., & Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics, 4, 1633–1649. https://doi.org/10.1002/pmic.200300771

Chapman, C. A., Chapman, L. J., Vulinec, K., Zanne, A., & Lawes, M. J. (2004). Saturation of an intra-gene pool linkage map: Towards a unified consensus linkage map for fine mapping and synteny analysis in common bean. PLoS ONE, 11(6), e28135. https://doi.org/10.1371/journal.pone.0028135

Chavez, C. H., Fernandez, A. C., Gomez, M., & Blair, M. W. (2009). Single strand conformation polymorphism based SNP and indel markers for genetic mapping and synteny analysis of common bean (Phaseolus vulgaris L). BMC Genomics, 10, 629. https://doi.org/10.1186/1471-2164-10-629

Grant, D., Nelson, R. T., Cannon, S. B., & Shoemaker, R. C. (2010). SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Research, 38, D843–D846. https://doi.org/10.1093/nar/gkp798

Guzman, P., Diers, B., Neece, D., St. Martin, S., LeRoy, A., Grau, C., et al. (2007). QTL associated with yield in three backcross-derived populations of soybean. Crop Science, 47, 111–122. https://doi.org/10.2135/cropsci2006.01.0003

Hall, T. A. (1999). BioEdit: A user-friendly biological sequence edition editor and analysis program for windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95–98.

Hayden, M. J., Tabone, T., & Mather, D. E. (2009). Development and assessment of simple PCR markers for SNP genotyping in barley. Theoretical and Applied Genetics, 119, 939–951. https://doi.org/10.1007/s00122-009-1107-1

Hefni, M., Öhrvik, V., Tabelka, M., & Witthöft, C. (2010). Folate content in foods commonly consumed in Egypt. Food Chemistry, 121, 540–545. https://doi.org/10.1016/j.foodchem.2009.12.044

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., & Gao, G. (2015). GS3D 2.0: An upgraded gene feature visualization server. Bioinformatics, 31, 1296–1297. https://doi.org/10.1093/bioinformatics/btu817

Huang, X., & Madan, A. (1999). CAP3: A DNA sequence assembly program. Genome Research, 9, 868–877. https://doi.org/10.1101/gr.9.8.688

Kelley, L. A., & Stemberg, M. J. E. (2009). Protein structure prediction on the web: A case study using the Phyre server. Nature Protocols, 4, 363–371. https://doi.org/10.1038/nprot.2009.2

Kelly, J. D., Gepts, P., Miklas, P. N., & Coyne, D. P. (2003). Tagging and mapping of genes and QTL and molecular marker-assisted selection for traits of economic importance in bean and cowpea. Field Crops Research, 82, 13–154. https://doi.org/10.1016/S0378-4290(03)00034-0

Kelly, J. D., Kolman, J. M., & Schneider, K. (1998). Breeding for yield in dry bean (Phaseolus vulgaris L), Euphytica, 102, 343–356. https://doi.org/10.1023/A:1018392901978

Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. Bioinformatics, 23, 1289–1291. https://doi.org/10.1093/bioinformatics/btm091

Kozak, M. (1981). Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Research, 9, 5233–5252. https://doi.org/10.1093/nar/9.20.525

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. Journal of Molecular Biology, 305, 567–580. https://doi.org/10.1006/jmbi.2000.4315

Lescot, M., Déhais, P., Moreau, Y., De Moor, B., Rouzé, P., & Rombauts, S. (2002). PlantCARE: A database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research, 30, 325–327. https://doi.org/10.1093/nar/30.1.325

Lynch, J. P. (2007). Roots of the second green revolution. Australian Journal of Botany, 55, 493–512. https://doi.org/10.1071/BT06118

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., & Bryant, S. H. (2015). CDD: NCBI's conserved domain database. Nucleic Acids Research, 43, D222–D226. https://doi.org/10.1093/nar/gku1221

McClean, P. E., Mamidi, S., McConnell, M., Chikara, S., & Lee, R. (2010). Synteny mapping between common bean and soybean reveals extensive blocks of shared loci. BMC Genomics, 11, 184. https://doi.org/10.1186/1471-2164-11-184

Miklas, P. N., Kelly, J. D., Beebe, S. E., & Blair, M. W. (2006). Common bean breeding for resistance against biotic and abiotic stresses: From classical to MAS breeding. Euphytica, 147, 105–131. https://doi.org/10.1007/s10681-006-4600-5

Mukeshimana, G., Butare, L., Cregan, P. B., Blair, M. W., & Kelly, J. D. (2014). Quantitative trait loci associated with drought tolerance in common bean. Crop Science, 54, 923–938. https://doi.org/10.2135/cropsci2013.06.0427
Xu, X., Sharma, R., Tondelli, A., Russell, J., Comadran, J., Schnaithmann, F., Pillen, K., Kilian, B., Cattivelli, L., Thomas, W. T. B., & Flavell, A. J. (2018). Genome-wide association analysis of grain yield-associated traits in a pan-European barley cultivar collection. The Plant Genome, 11, 1–11. https://doi.org/10.3835/plantgenome2017.08.0073

Ye, J., Coulouris, G., Zaretsky, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics, 13, 134. https://doi.org/10.1186/1471-2105-13-134

Yokoyama, R., Yamamoto, H., Kondo, M., Takeda, S., Ifuku, K., Fukao, Y., et al. (2016). Grana-localized proteins, RIQ1 and RIQ2, affect the organization of light-harvesting complex II and grana stacking in Arabidopsis. Plant Cell, 28, 2261–2275. https://doi.org/10.1105/tpc.16.00296

Zhang, W. K., Wang, Y. J., Luo, G. Z., Zhang, J. S., He, C. Y., Wu, X. L., Gui, J. Y., & Chen, S. Y. (2004). QTL mapping of ten agronomic traits on the soybean (Glycine max L. Merr) genetic map and their association with EST markers. Theoretical and Applied Genetics, 108, 1131–1139. https://doi.org/10.1007/s00122-003-1527-2

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Reinprecht Y, Qi Y, Shahmir F, Smith TH, Pauls KP. Yield and antiyield genes in common bean (Phaseolus vulgaris L.). Legume Science. 2021;1–17. https://doi.org/10.1002/leg3.91