Chapter 4

Molecular Analysis of Transgenic Plants

4.1 Introduction

Plant transformation has developed into an essential experimental means to understand the organization and regulation of eukaryotic genes and also to investigate various aspects of plant biology such as physiology, genetics and developmental biology among others. The accuracy and simplicity of the approach, in addition to the potential of resolution of fine characteristics at the molecular level, led to its acceptance to address several problems. It is well established that both epigenetic and heritable variation (somaclonal variation) could be generated via tissue culture [332]. *Agrobacterium tumefaciens* mediated plant transformation is the most widely exploited method for introducing foreign genes into plant genome. A high level of perfect transgenic loci with absolute conservation of the host genome could be developed via this method [333]. The molecular mechanisms involved in the integration of transfer DNA (T-DNA) are still not well characterized; nevertheless integration is considered to occur by illegitimate recombination [334]. The arrangement of transgenic loci is determined by the genomic factors and does not depend on the method used for transgene introduction [335]. The complication of the integration mechanisms paves the way to transgenic loci consisting of multiple copies of the transgene [336, 337]. The other major factor contributing to the disparity is related to the transformation process. None of the methods of plant transformation presently in use (excluding plastid transformation) provides any mechanism for introducing the transgene at a specific locus. In addition, the technique does not provide any control over the copy number of the gene being introduced. Consequently, the foreign DNA gets inserted at random positions in the host chromosome(s). Therefore, independent transgenic lines having the same cassette of transgene sequences may act differently depending on their location in the host genome (position effect) [332]. If the introduction of transgene is into the heterochromatic region of host genome where its expression is subdued, it may experience negative position effect as there is little probability of improving such transgenic lines as even the selection of transformed cell requires expression of the transformed gene (*i.e.* scorable or selectable marker gene). A different scenario could be an extremely elevated level of expression due to the presence of an enhancer element in the vicinity of introduced gene. Transgene lodged in the subtelomeric region may experience positive position effect since gene(s) in that region are known to be highly expressed [338].

Therefore, the study of transgene expression and its consequences are of vital importance. Transgene expression is influenced by many factors, in particular, the site of integration within the plant genome, gene silencing, and the promoter attached with it. While some of these factors can be bypassed to some extent in the experimental design, it is still pertinent to correlate the phenotypic differences between the transgenic and control plants with transgene expression. To investigate it thoroughly the ultimate effect of transgene(s) on the expression of other genes and metabolites must also be studied. The polymerase chain reaction is an important technique in transgenic research as it could screen putative transgenic plants at an early developmental stage, when plant material is a limiting factor for other confirmatory analyses. High sensitivity, specificity, less time consuming and low costs are few advantages of PCR. In this method a precisely defined fragment, unique for transgene is amplified *in-vitro* by elongation of transgene specific primers and visualized in agarose gel after electrophoresis on the basis of size of the amplified fragment. Amplified product can be further confirmed by sequencing or southern blotting. However, in most cases further information is required concerning the number of copies of the inserted transgene in the transgenic, its expression both at mRNA and protein levels and its impact on other cellular activities [332]. Transgene expression can be evaluated at the RNA, protein and bio-analytical levels. Northern blot, dot blot, microarray and qRT-PCR are some fundamental and well established methods to examine the expression of transgene at transcript level. The transgene which is finally expressed as protein can be confirmed by western blot analysis or bio-analytical assay. Western blot is performed with a protein specific antibody, which is visualized by biochemical reaction. In addition, the effect of transgene introduction can also be visualized via a comparative proteomics approach since it can reveal the influence of alien gene and can be used even to distinguish the landraces, populations, varieties and even species [121, 149].

As described, genetic transformation of sweet potato was carried out with three different constructs and the regenerated plants were subjected to the kanamycin selection. Therefore, knowledge about the successful transgenic events is imperative for downstream analysis.

4.2 Materials and methods

4.2.1 Plant growth and maintenance

Both the transgenic and wild type was grown in parallel in the greenhouse in identical conditions with one plant per pot (diameter, 30 cm) to eliminate the environmental or developmental influence(s), if any. Each line was grown at least in four pots filled with a mix of clay loam and vermiculite (3:1) and was watered sufficiently. Mature tubers were harvested and eight identical sized mature tubers (two from each pot) were pooled to normalize the growth and developmental effects. Identical sized mature tubers were collected from both the wild type and the overexpressed lines, peeled off and cut into slices. Samples of slices were pooled to normalize the growth and developmental effects, quickly frozen in liquid nitrogen, and stored at –80°C until use.

4.2.2 Molecular analysis of putative transgenic lines

To check the integrity of the transferred gene, PCR analysis of genomic DNA from the putative transformants was carried out using gene specific primers. Integrity of the selectable marker gene *nptII* was assessed at the genome level by PCR analysis using genomic DNA isolated from the putative transformants and the wild-type plants as described earlier [339]. The presence of transgene was confirmed by PCR using gene specific primers (*nptII*F 5’ -ATGATTGAACAAGATGGATTGCACGCAGG -3’ and *nptII*R 5’- GAAGAA CTCGTCAAGAAGGCGATA -3’) for *nptII*, which delimits 0.8 kb fragment from the *nptII* coding region. For *AmA1* gene specific primers (*AmA1*F 5’-CACCATGGCGGGATTACCAGTG-3’ and *AmA1*R 5’-CAAGGAAGAACCCTCTTGTTTCC-3’) were used which delimits the 1.02 kb fragment. PCR analysis was performed in 20 µl reaction mix containing 1X PCR buffer, 2 mM $MgCl_2$, 200 µM dNTP mix, 0.5 µM of each primer and 0.5 units of Taq polymerase. The cycling conditions employed were 3 min at 94°C for denaturation and 30 amplification cycles using 94°C for 1 min, 58°C for 45 sec annealing and 72°C for 45 sec followed by 5 min extension at 72°C. The products were electrophoresed in 0.8% agarose gels.

4.2.3 Molecular analysis of the putative transgenic lines at transcript level

Transcript accumulation of transgene was assessed by northern blot analysis as well as by real time quantitative PCR (qRT-PCR).

4.2.3.1 Northern blot analysis

Northern blotting analysis is a conventional technique for analysis of the size and steady-state level of a specific RNA in a composite sample. Briefly, the RNA is size-fractionated by gel electrophoresis and transferred by blotting onto a membrane to which the RNA is covalently bound. Then, the membrane is analyzed by hybridization to one or more specific probes that are labelled for subsequent detection. For the analysis of the transcript accumulation pattern, Northern blot analysis was performed. Briefly, total RNA was isolated by TriPure Isolation Reagent (Roche Diagnostics) following manufacturer's recommendation. With the use of formaldehyde as a denaturant, 10 µg of total RNA was resuspended in RNA loading buffer, denatured at 65°C for 10 min, quenched in ice for 1 min and further subjected to gel electrophoresis on 1.2% formaldehyde denaturing agarose gel according to [324]. After electrophoresis, the gel was rinsed thrice in DEPC treated water and after rinsing the gel was neutralized in 2X SSC. Ethidium bromide staining under UV light was used to ascertain equal gel loading and efficient transfer to nylon membrane. *AmA1* amplified fragment was used as probe. The membranes were hybridized and exposed to Kodak X-ray film and autoradiographed. The detailed methodology is described earlier in section 3.2.4.

4.2.3.2 qRT-PCR analysis

Total RNA was isolated by as described above and cDNAS were prepared either by using SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen) or by 3'-RACE kit (Invitrogen). For the standardization of endogenous control to normalize the expression data gene specific primers of *Tubulin* (*TubRT*F 5'- AGGACCCTTGTGTTTGGTGTTAA- 3' and *TubRTR* 5'-CCCACTCATCGTTGCAGAAA-3'), *GAPDH* (*GAPDHRT*F 5'-AAGAAAACAAAAGCACGGCACTA-3' and *GAPDHRT*R 5'-AAGTGGAAAAAGGATTCGGTGTAT-3') and *Actin* (*Actin*F 5'-CTCCCCTAATGAGTGTGATGTGAT-3' and *Actin*R 5'-GAGCCCCATGAGAACATTACCA-3') were used. In addition, to evaluate the transcript

abundance, *GUS* and *AmA1* gene specific primes were used. The *GUS*-specific primers (*GUS*F 5'-TGGTAATTACCGACGAAAACGGC-3' and *GUS*R 5'-ACGCGTGGTTACAGTCTTGCG-3') and *AmA1*-specific primers (*AmA1RT*F 5'-GGGAATGATCCTCGCGAAA-3' and *AmA1RT*R 5'- AAAATCATGCACATCCGACCTA-3'; *AmA1UTRRT*F 5'- GAGATAATAGAATTGGGATCCAACAAC-3' and *AmA1UTRRT*R 5'-CCAAAGAGACGACTTACAACGTTTT-3') were used respectively to assess the transgene expression in the putative transgenic lines. All the primers were designed using the Primer Express Software v3.0.1. The qRT-PCR was performed, in two biological and three technical replicates, by an ABI PRISM 7700 Sequence Detection System (Applied Biosystems) using SYBR green dye. Mean of the Ct values for target and endogenous control was considered for calculating the relative quantitation (RQ) value using comparative Ct ($2^{-\Delta\Delta Ct}$) method.

4.2.4 Analysis of copy number of transgene

qRT-PCR was used to determine the copy number of the *AmA1* gene per genome in the constitutive overexpression (COE) and tuber specific overexpression (TOE) lines. The method of quantification used a standard curve for a series of copies of the *AmA1* gene. Plasmid pSB8 containing a single copy of *AmA1* gene per plasmid was prepared and the copy number of the *AmA1*gene was calculated per molecular weight. Alternatively, to generate a standard curve for the endogenous gene *actin* and *AmA1*, genomic DNA from transgenic line was used as described earlier [340, 341]. Standard curves were calibrated using six concentrations 50, 5.0, 0.5, 0.05, 0.005 and 0.0025 ng /reaction of the genomic DNA of the wild type and transgenic lines. These standard curves were used for the relative quantitation of the *actin* and *AmA1*. A non template control (NTC) was also prepared as a negative control for the data analysis.

4.2.5 Molecular analysis of putative transgenic lines at protein level

4.2.5.1 Quantitative enzyme assay of *GUS* (β-glucuronidase, EC 3.2.1.31)

GUS (β-glucuronidase) activity was determined using the fluorometric method described earlier with few modifications [342]. The tissues were ground and homogenized with extraction buffer (50 mM sodium phosphate, pH 7.0, 10 mM beta-mercaptoethanol, 10 mM EDTA, 0.5 mM PMSF, 0.1% sodium lauryl sarcosine and 0.1% Triton X-100). The homogenates were then centrifuged (10000 g for 5 min at 4°C) and the supernatants were further used for the assay. The

reaction mixture in triplicate, consisted of 50 mM sodium phosphate, pH 7.0, 1 mM 4-methylumbelliferyl-β-Dglucuronide (MUG) and the tissue extract was incubated at 37°C for appropriate time intervals (5-60 min). The reaction was terminated by the addition of 0.2 M Na_2CO_3. Fluorescence was measured at 455 nm using a Varian Cary Eclipse Fluorescence Spectrophotometer (Agilent) set at an excitation wavelength of 365 nm.

4.2.5.2 Immunoblot analysis of putative transgenic lines

4.2.5.2.1 Protein extraction and SDS- PAGE

Protein extraction was carried out by the method described by **Hoffman et al. (1988**) [343]. Proteins were extracted from all the transgenic lines and from the wild type. Proteins were extracted from leaves as well as from the tuber (specifically for tuber specific lines). Leaves and mature tubers were ground to powder in liquid nitrogen and transferred to an open-mouthed 50 ml tube. Immediately, the leaves and tuber powder was homogenized in respective homogenizing buffers [50 mM Tris-HCl (pH 6.8), 2 mM EDTA, 20% glycerol, 5 mM DTT and 2 mM PMSF] and [50 mM Tris-HCl (pH 8.2), 2 mM EDTA, 20% glycerol, 5 mM DTT and 2 mM PMSF]. The samples were kept for 1 hr on gyrorotatory shaker at 4°C. The soluble proteins were recovered as supernatant by centrifugation at 10,000 x *g* for 10 min at 4°C. The supernatant, so obtained, was the soluble protein fraction. The concentration of protein extract was determined by Bradford assay (Bio-Rad) and 50 µg of protein was precipitated in 10% TCA overnight at 4°C and subjected to SDS-PAGE on 12.5% (w/v) acrylamide gels [344]. The 12.5% gels were prepared and protein samples were loaded after denaturation with addition of 1/6 volume of 6X Laemmli buffer (50 mM Tris-HCl pH 6.8, 100 mM DTT, 2% SDS, 0.1% bromophenol blue, 10% glycerol) and boiling for 5-10 min. Initially the proteins were run at low voltage so that the proteins were stacked and when the proteins entered the separating gel, voltage was increased. When the run was over, one of the gels used for loading control was stained with Coomassie brilliant blue R-250 (CBB-0.2%, 50% methanol, 10% acetic acid and destained by the destaining solution (40% methanol and 10% acetic acid) and images were digitized with a FluorS equipped with a 12-bit camera (Bio-Rad, CA). Another gel was subjected to immunoblotting.

4.2.5.2.2 Immunobltting

Immunoblot analysis was done by resolving the protein extracts from wild-type and the transgenic plants on a uniform 12.5% SDS-PAGE and then electrotransferred onto nitrocellulose membrane (Amersham Biosciences, Bucks, U.K.) at 150 mA for 3 h. The membrane was blocked with 5% (w/v) nonfat milk in TTBS buffer (0.1 M Tris, pH 7.9, 0.15 M NaCl, 0.1% Tween 20). The resolved proteins were probed with the primary polyclonal antibody, raised in rabbit against an antigenic peptide of AmA1. Immunodetection was performed by incubation of the membrane-bound proteins with alkaline phosphatase and with horseradish peroxidase (HRP) conjugated anti rabbit IgG secondary antibody. Antibodies were diluted to varying ratios (1:1000-1:5000) in Tris-buffered saline (TBS). The signals were detected using NBT/BCIP (nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate) method and the HRP conjugated cross reactivity signal was detected as per manufacturer's instruction.

4.2.5.2.3 Densitometric quantitation

The intensity of the bands corresponding to AmA1 was quantified using Quantity One 1-D Analysis Software (Bio-Rad).

4.2.5.3 Comparative proteomics

4.2.5.3.1 Protein isolation and 2-DE

Protein was isolated from the mature tubers by the method of **Jiang et al. (2012)** with few modifications [89]. The tubers were peeled, sliced and grounded to fine powder with liquid nitrogen and suspended in acetone containing 10% trichloroacetic acid and 0.07% ß-mercaptoethanol, and precipitated overnight at -20°C. The precipitates were recovered by centrifugation at 10,000 x *g* at 4°C for 10 min and washed twice with 0.07% ß-mercaptoethanol in acetone and then dried. The dried pellets were resuspended in the resuspension buffer (100 mM Tris-HCl buffer, pH 8.0, 10 mM EDTA, and 30% sucrose, 2% CHAPS, 2% SDS, 10 mM PMSF and 2% ß-mercaptoethanol) and further extracted with an equal-volume of Tris-saturated phenol. Proteins were precipitated overnight at -20°C by the addition of 0.1 M ammonium acetate in methanol, 4 times the volume of the sample. The pellet was washed 2-3 times with 0.1 M ammonium acetate in methanol and then with 0.07% ß-mercaptoethanol in acetone. The

protein pellets were solubilized using IEF sample buffer (8 M urea, 2 M thiourea and 4% w/v CHAPS) and the concentration of protein was determined using 2-D Quant kit (GE Healthcare). Aliquots of 250 µg protein were diluted with 2D rehydration buffer (8 M urea, 2 M thiourea, 4% w/v CHAPS, 20 mM DTT, 0.5% v/v pharmalyte (pH 4–7) and 0.05% w/v bromophenol blue) followed by rehydration of the IPG strips (13 cm; pH 4–7 and pH 3-10) with 250 µl of the solution. Electrofocusing was performed using the IPGphor system (GE Healthcare) at 20°C for 32 000 VhT. The focused strips were subjected to reduction with 1% w/v DTT in 10 ml of equilibration buffer (6 M urea, 50 mM Tris-HCl (pH 8.8), 30% v/v glycerol and 2% w/v SDS), followed by alkylation with 2.5% w/v iodoacetamide in the same buffer. The strips were then loaded on 12.5% polyacrylamide gels for SDS-PAGE. To reduce gel-to-gel variation, each protein preparation was analysed through at least three parallel 2-D gels, representing three technical replicates. The electrophoresed proteins were stained with Silver Stain Plus Kit (Bio-Rad).

4.2.5.3.2 Image acquisition and data analysis

The gel images were scanned using the Fluor-S MultiImager system (Bio-Rad), and processed and analyzed with PDQuest gel analysis software version 7.2.0 (Bio-Rad). The pI and experimental molecular mass were calculated from the scanned images using standard molecular mass marker proteins. To compare spots across gels, a match set representing a standard image of three replicates, representing three biological replicates, was created for each sample. Each spot on the standard gel was quantified by several criteria for consistancy in size and shape for all the replicate gels and being within the linear range of detection. The spots detected by the software program were further verified manually to eliminate any possible artifacts, such as gel background or streaks. In addition to ascertain the quantification for the scores, the PDQuest software was used to assign quality scores to each gel spot. The spots with a quality score less than 30 were eliminated from further analysis. The high-quality spot quantities were used to calculate the mean value for a given spot, and the value was used as the spot quantity on the standard gel. The correlation coefficient, representing the association between the spot intensities on replicates, was maintained at a minimum of 0.8 between gel images. The spot densities on the standard gel were normalized against the total density in the gel image. To facilitate the comparison of the standard gels for each of the samples, the spot volumes were further

normalized using three unaltered protein spots across all the gels to avoid experimental variations in 2-DE gels.

4.3 Results

4.3.1 Confirmation of the *GUS* putative transformants

As stated above, 25 and 15 regenerated lines survived the kanamycin selection from cv. SP-6 and SP-17, respectively in the *GUS* transformed lines. All the *GUS* transformants were checked for successful integration of the transgene. While 18 out of 25 regenerated plants in cv. SP-6 were found to be positive, 9 out of 15 regenerated plants were positive in cv. SP-17. We observed no detectable morphological difference between the wild-type and the transgenic plants. Successful integration of the transgene was examined by PCR analysis at genome level in 12 putative transformants (6 from cv. SP-6 and 6 from cv. SP-17) using *nptII* gene specific primers. The presence of 0.8 kb amplicons in the putative transformants confirmed the successful integration of the transgene (**Fig. 4.1A and B**).

4.3.2 Transcript accumulation and enzyme assay of *GUS* transformants

Out of these 12 transgenic plants, 4 (2 from both the genotypes) were further selected to assess the transcript accumulation and relative *GUS* expression. The transcript accumulation of *GUS* showed 3- to 12-fold expression compared to the wild-type. Accumulation pattern also revealed more *GUS* transcript in the transformants of cv. SP-6 when compared with that of SP-17 transformants (**Fig. 4.2 A**). The *GUS* activity of each transgenic plant was significantly higher than that of wild-type and the untransformed plant. In addition, there was a positive correlation between the transcript accumulation and enzyme activity of *GUS* transformants (**Fig. 4.2 B**).

4.3.3 Molecular analysis of putative *AmA1* transformants by PCR

To check the integrity of the transferred gene, PCR was carried out with primers designed from the coding sequence of the AmA1 gene. Successful integration of transgene was examined at genome level in all the kanamycin selected putative transformants. Out of 18 regenerated constitutive overexpression (COE) lines, 13 were found to be true transformants as revealed by the amplification of 1.02 kb fragments which was otherwise absent in wild type (cv. SP-6).

Additionally, out of 15 regenerations from the tuber specific promoters, all the lines were found to be positive (**Fig. 4.3 A and B**).

4.3.4 Analysis of transgene expression

4.3.4.1 Immunoblot analysis

All the PCR positive 13 COE lines and 15 TOE lines were selected for further investigation at protein level by immunoblot analysis. Immunodetection of AmA1 in tubers of the constitutive lines showed a 35-kDa band in all of the transgenic events but at varying levels, suggesting the expression of AmA1 protein in the alien environment (**Fig. 4.4 A)**. Based on the protein expression profile six constitutive transgenic lines were selected for further analysis. The band intensity corresponding to AmA1 was quantified by densitometry analysis in all the COE lines. The transgenic lines were categorized into three groups, *i.e.* low expression, moderate expression and high expression (**Fig. 4.4 B and Table 4.1**). A cut of value in the range of 5500 to <6000 INT/mm^2 was considered for moderate expression and accordingly for the low and high expression group. Out of 13 transgenic lines 7 were categorized as high expression lines whereas 2 and 4 lines were categorized as moderate and low expression lines, respectively. For further analysis 4 lines were considered from high expression group and one each from the moderate and low expression groups.

In the TOE lines, all the 15 transgenic events showed the expression of AmA1 as a 35-kDa band corresponding to the transgene product which was absent in the wild type. To reduce the transformants to transformants variation due to difference in the background of the blots, band intensity was calculated in each blots and based on the values six lines were selected for further analysis (**Fig. 4.5)**.

4.3.4.2 Evaluation of the transcript abundance by northern blot analysis

Since expression of most of the seed proteins is temporally and spatially regulated, to study the developmentally regulated expression of *AmAl* transgene in sweet potato as well as the stability of expression, the transcript abundance were analyzed. The best independent transgenic events from both the category were selected for this analysis. Total RNA was isolated from the developing tubers as well as from the mature tubers of both the COE and TOE lines. The

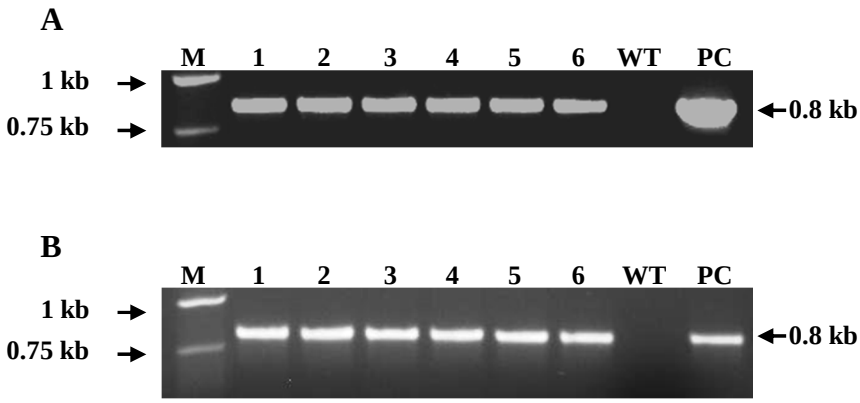


Fig. 4.1. Confirmation of transgene integration. Successful integration of the transgene was confirmed by PCR analysis in 6 putative transformants, each from cv. SP-6 (**A**) and SP- 17 (**B**), respectively. An amplicon of 0.8 kb confirmed the successful transgene integration in putative transformants which was absent in wild type (WT) plants. Expression plasmid pBI121 was used as positive control (PC).

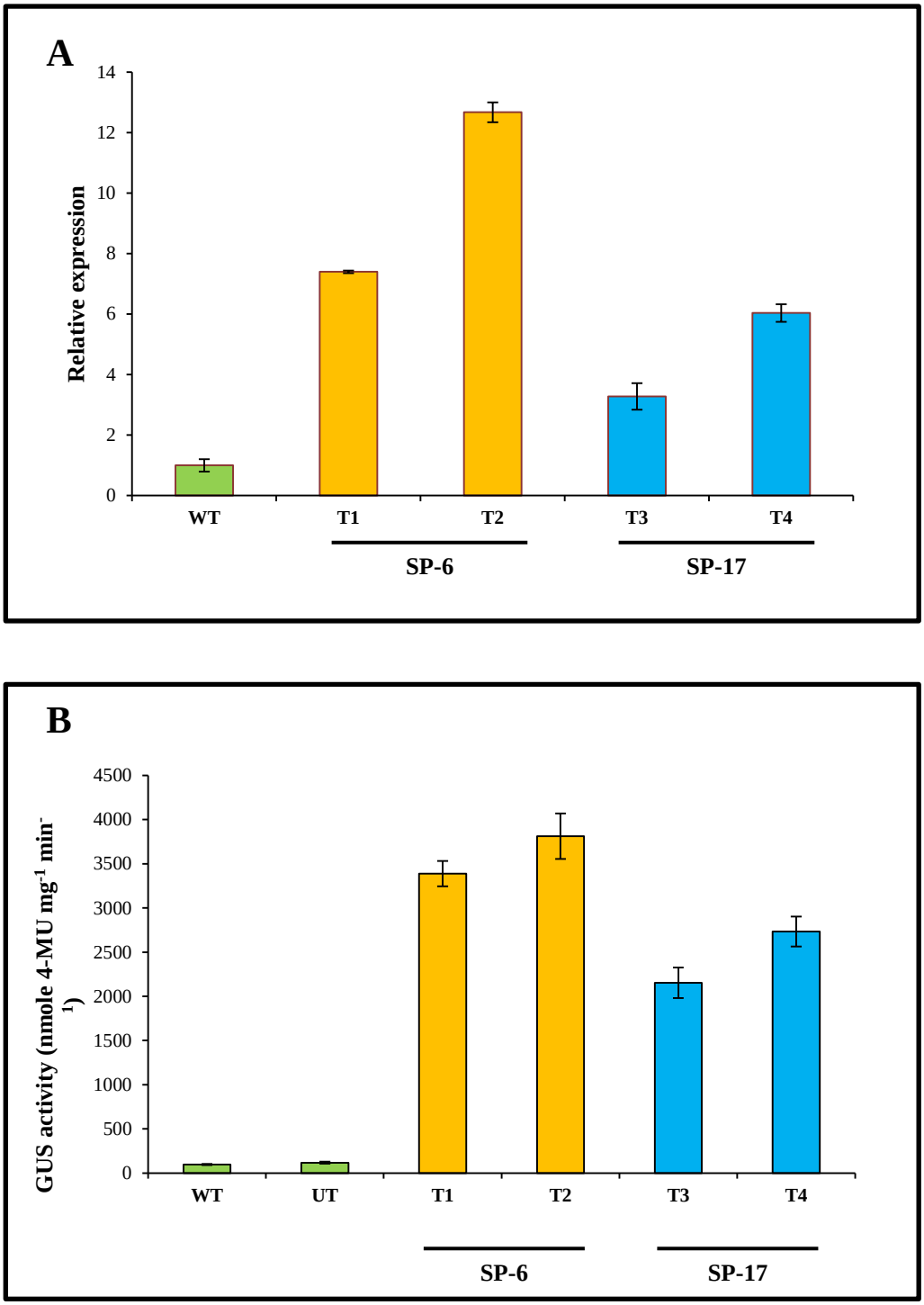


Fig. 4.2. Transcript accumulation and enzymatic activity of *GUS*. The transcript accumulation of *GUS* showed a 3 to 12-fold increase in expression than that of wild type (WT) plant (**A**). Higher *GUS* activity was observed in the transgenic plants than WT and UT (untransformed plant) (**B**). Each analysis was performed in 4 transgenic plants (2 from each of the genotypes).

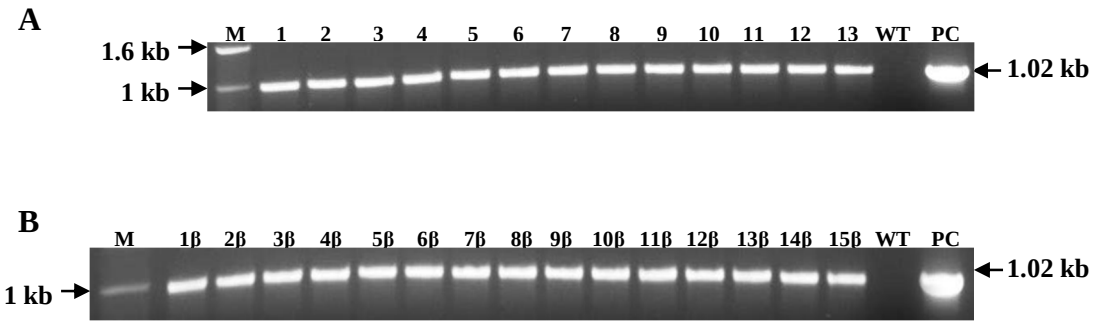
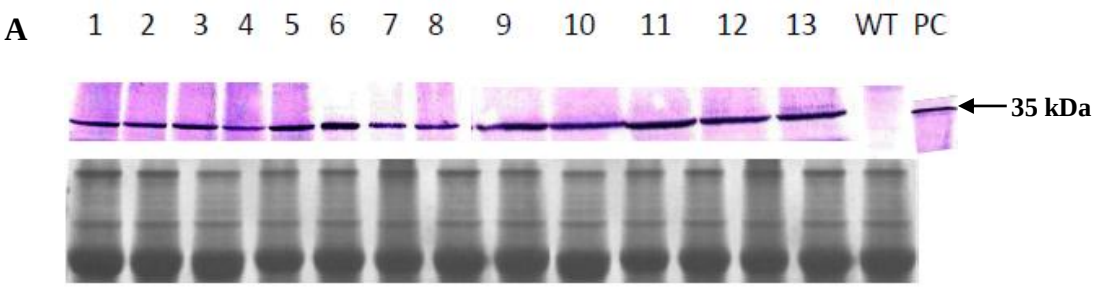


Fig. 4.3. Confirmation of transgene integration in AmA1 transformants. Successful integration of the transgene *AmA1* was confirmed by PCR analysis in 13 putative transformants from pSB8 (COE lines) (**A**) and in 15 putative transformants from pSB8β (TOE lines) (**B**). An amplicon of 1.02 kb confirmed the successful transgene integration in putative transformants which was absent in wild type (WT) plants. Expression plasmid pSB8 was used as positive control (PC). M represents the marker lane.



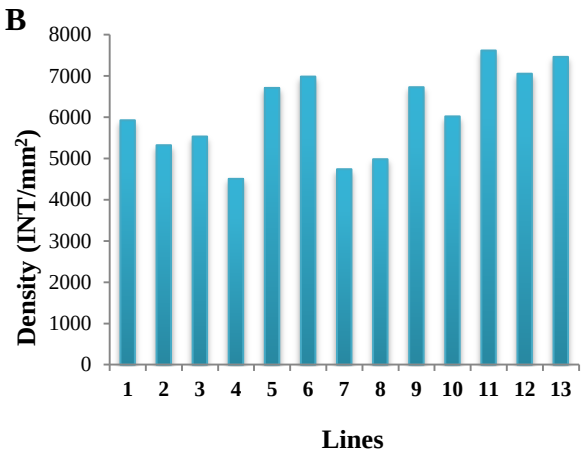


Table 4.1

High expression	Moderate expression	Low expression
5	**1**	2
6	3	4
9		**7**
10		8
11		
12		
13		

Cutoff value
High expression ≥ 6000 INT/mm²
Moderate expression ≥ 5500- <6000 INT/mm²
Low expression < 5500 INT/mm²

Fig. 4.4. Immunodetection of AmA1 expression in COE lines. Transgene expression was confirmed at protein level in all the PCR positive 13 constitutive (COE) lines. Immunodetection of AmA1 in tubers of the constitutive lines showed a 35 kDa band in all of the transgenic events. CBB stained gel as loading control is shown in the lower panel (**A**). Densitometry analysis was performed in all the constitutive lines (**B**) and categorized into three groups, *i.e.* low expression, moderate expression and high expression based on a cutoff value 5500 INT/mm² (Table 4.1). Transgenic lines selected for further analysis are shown in red colors.

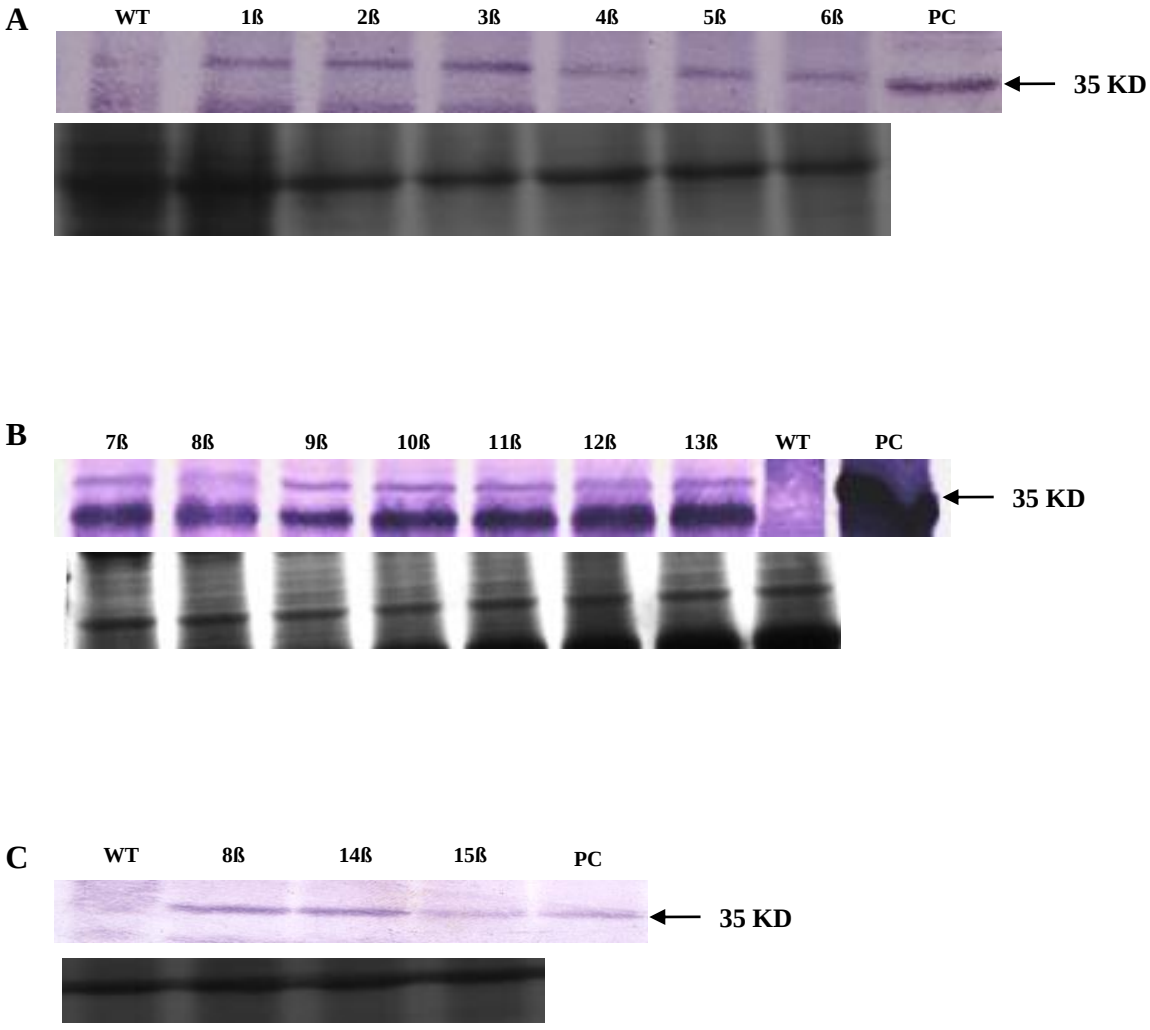


Fig. 4.5. Immunodetection of AmA1 expression in TOE lines. Transgene expression was confirmed in all the PCR positive 15 tuber specific (TOE) lines (**A-C**). CBB stained gel as loading control is shown in the lower panel of each blot.

Northern blot probed with *AmAl* revealed its steady expression up to the maturity albeit with slight reduction in the mature tuber (**Fig. 4.6 A**). This perhaps is due to its nature as a seed storage protein as they act as a sink to act as a biological reserve. Intriguingly, the expression of *AmAl* in TOE line was lower than that of COE line at both the developmental stages as revealed by the quantification of the autoradiogram (**Fig. 4.6 B**). Additionally, the quantification of the autoradiogram also revealed that the transcript of *AmAl* is more stable in the COE lines even upto the maturity whereas it reduced more in TOE lines.

4.3.5 Evaluation of transcript abundance by qRT-PCR

4.3.5.1 Standardization of endogenous control

The identification of stable reference genes is crucial for normalizing the levels of target mRNA for the accurate quantification of mRNA transcripts using qRT-PCR. For the standardization of endogenous control three housekeeping genes *tubulin, glyceraldehyde 3-phosphate dehydrogenase* (*GAPDH*) and *actin* were used. Initially a housekeeping gene *tubulin* was used as the endogenous control based on the reports by **Kim et al. (2008)** for normalization [345]. It was not found to be suitable as an ideal endogenous control for accurate normalization as its Ct value varies across the lines including wild type as revealed by its dissociation/melt curve (**Fig. 4.7 A-D**). To negate the possibility of pipetting error, Ct value as well as dissociation/melt curve of target gene *AmA1* were also analyzed which revealed no detectable errors (**Fig. 4.7 E and F**). *GAPDH* showed more or less a similar pattern and its transcript level was also not constant across the samples (**Fig. 4.8 A and B**). However, expression level of *actin* was relatively constant and exhibited the most stable expression status for all samples (**Fig. 4.9 A-C**). and *actin* was used an ideal internal reference gene.

4.3.5.2 Transcript accumulation of *AmA1*

Transcript level of *AmA1* was then analyzed in six selected COE and TOE lines taking *actin* as endogenous reference gene. Transcript analysis showed a higher level of transcript accumulation of the transgene in most of the transgenic lines when compared to the wild-type. However, no steady and reproducible pattern of expression was observed in the transgenic lines selected (**Fig. 4.10 A and B)**. It may perhaps be due to some nonspecific binding of primers with some endogenous seed proteins/albumin since the transgene *AmA1* is a seed albumin. Therefore,

to increase the specificity of the primers, new set of primers were used taking the partial sequence of cDNA as well as the 102 bp of 3' UTR (**Fig. 4.11)**. The significance of 102 bp of 3' UTR has already been discussed in **Section 2.5**. Furthermore, a new set of primers (*AmA1UTRRT*F 5'- GAGATAATAGAATTGGGATCCAACAAC-3' and *AmA1UTRRT*R 5'-CCAAAGAGACGACTTACAACGTTTT-3') was used and a consistent pattern of transcript level across the transgenic lines was obtained. The transgene expression level of transgenic lines revealed more than 3- to 12-fold higher transcript abundance in different COE lines (**Fig. 4.12 A)**. However, in TOE lines, 5- to 9-fold higher expression of transgene was observed which was in concordance with the northern blot analysis (**Fig. 4.12 B)**. To check it further, RNA was isolated from the shoot as well as from the tubers of COE lines and cDNA was synthesized. In addition, RNA was also isolated from tubers of TOE. The expression level of *AmA1* in the shoots and tubers of COE lines was compared with the expression level of *AmA1* in TOE by qRT-PCR. This analysis revealed an overall prominent level of transgene expression in tubers as compared to shoots in constitutive lines. Nonetheless, TOE had lower expression in comparison to the expression in the COE lines vis-à-vis the transgene expression specifically in the tuber. However, the transgene expression was higher in the TOE line when compared to the aerial portion of COE lines (**Fig. 4.13 A)**. Furthermore, leaky expression was checked by qRT-PCR itself in the TOE lines and for that cDNA was synthesized from the RNA isolated from aerial portion and the sample was thus named ß' as well as from the tubers (named as ß). The expression analysis revealed a comparable level of expression in ß' with that of wild type (**Fig. 4.13 B)**. To check the disparity of transgene expression at protein level between the COE and TOE lines, immunoblot analysis was performed with two most promising,one each from COE and TOE lines. Which were used in northern and qRT-PCR analysis. The immunoblot analysis revealed 1.5–fold higher expression of transgene in COE line when compared to TOE line. (**Fig. 4.13 C**).

4.3.5.3 Determination of transgene copy number

Quantification of the copy number of the *AmA1* overexpressing lines by absolute quantification in qRT-PCR revealed a low copy number of the transgene. The results of this quantification revealed the presence of a single copy of the transgene in most of the transgenic events with very few having two to three copies as shown in **Table 4.2**.

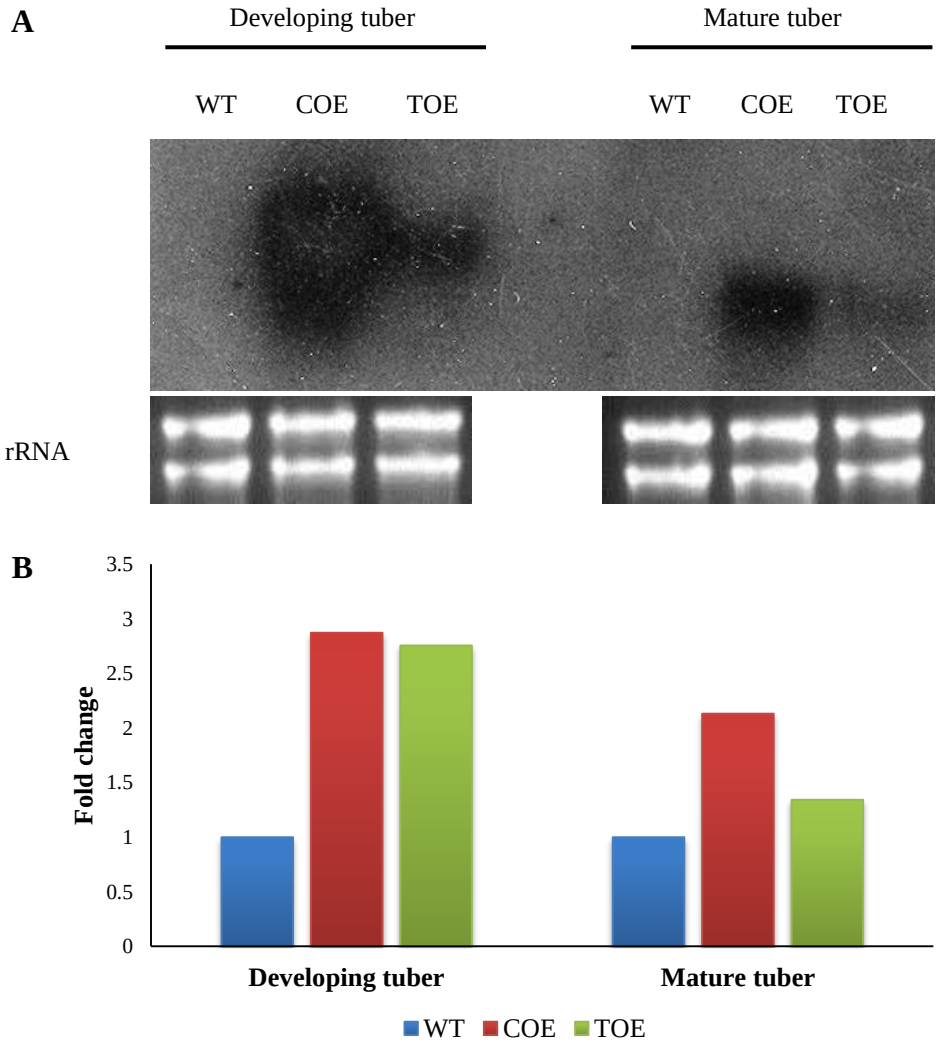


Fig. 4.6. Evaluation of the transcript accumulation by northern blot. Total RNA was isolated from the developing tubers as well as from the mature tubers of two independent lines from both the categories (COE and TOE) and transcript level was analyzed (**A**). Ethidium bromide-stained RNA served as loading control. Further quantification of the autoradiogram by Quantity One Analysis Software was performed (**B**).

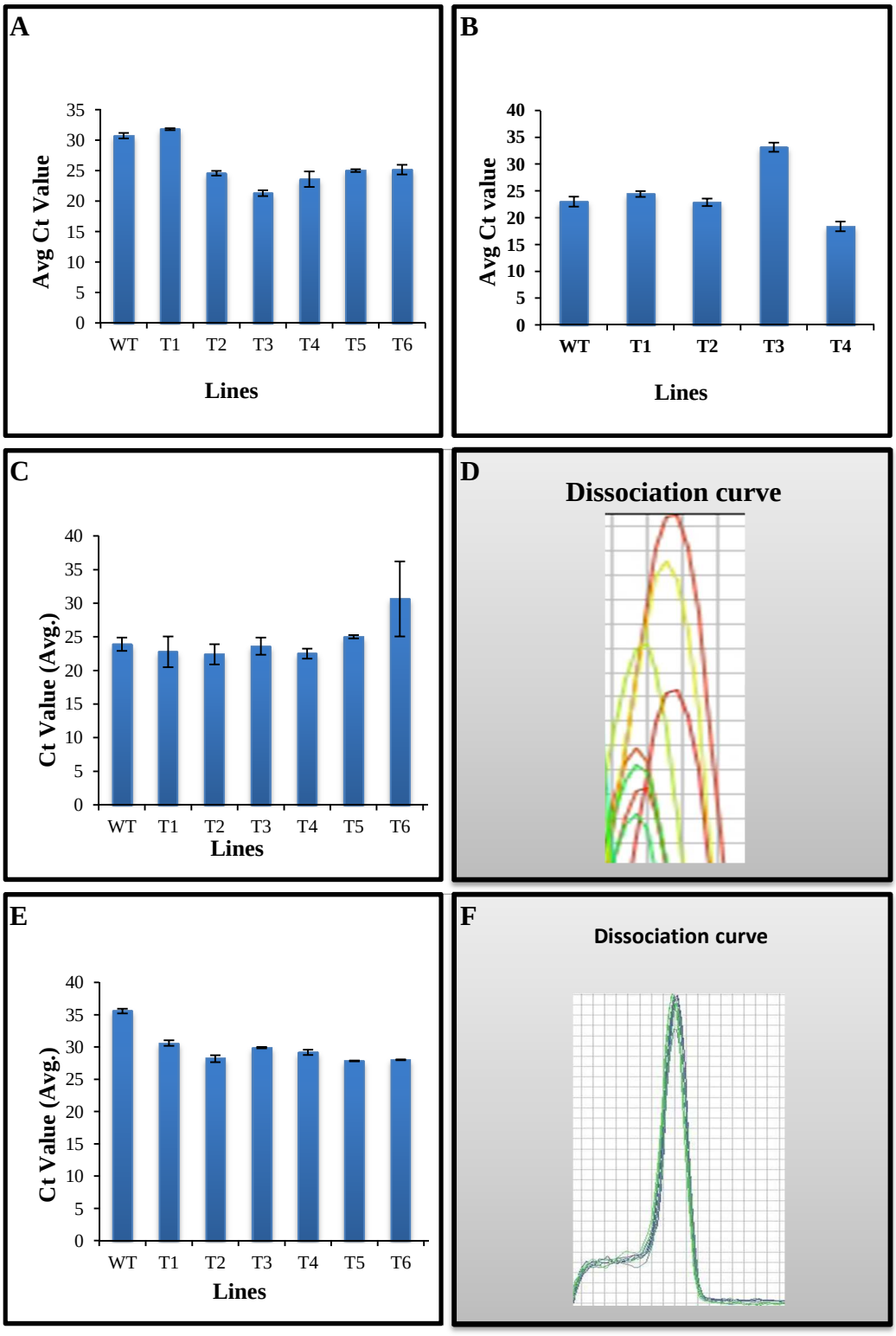


Fig. 4.7. The identification of stable reference genes by qRT-PCR. *Tubulin* was analyzed as endogenous reference gene for accurate normalization. Histogram represents the Ct value of tubulin in wild type as well as in transgenic lines (**A-C**), A constant expression level was not observed in case of tubulin as revealed by dissociation curve (**D**), To negate the possibility of pipetting error, Ct value of target gene *AmA1* and its dissociation curve were also analyzed (**E** & **F**).

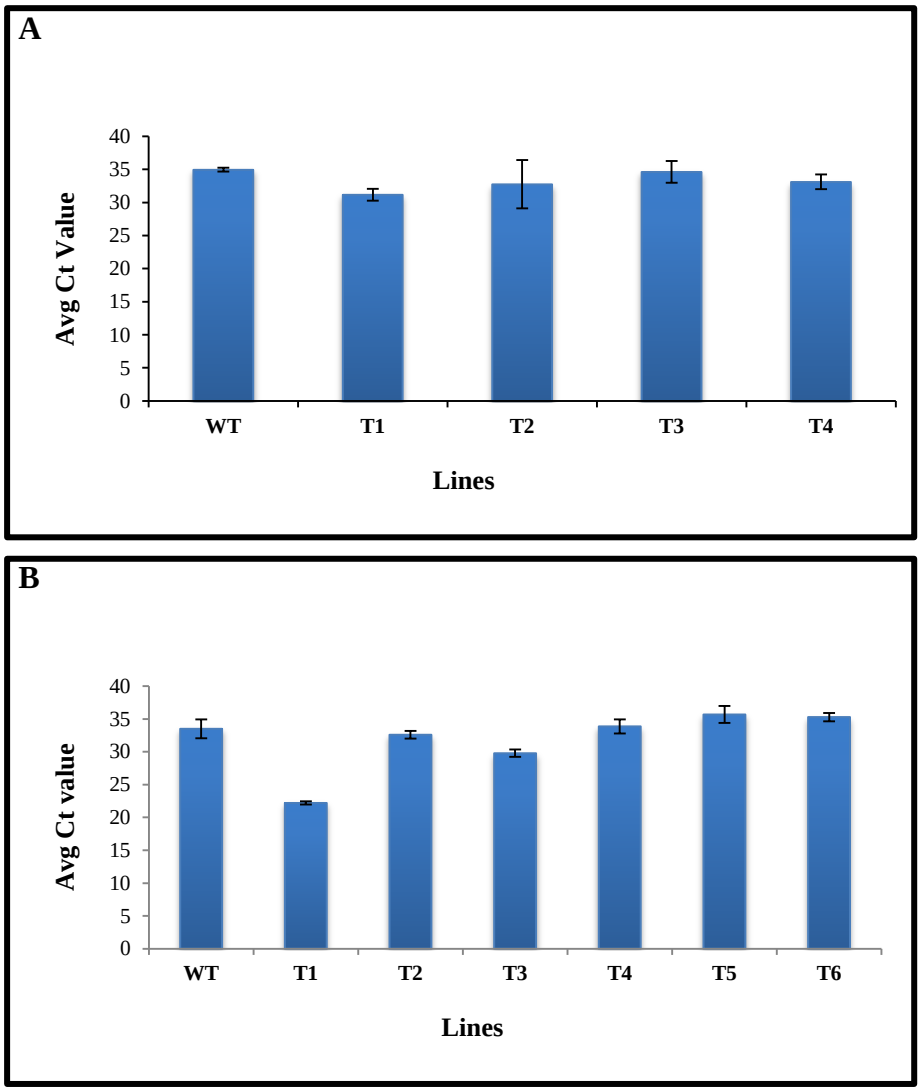
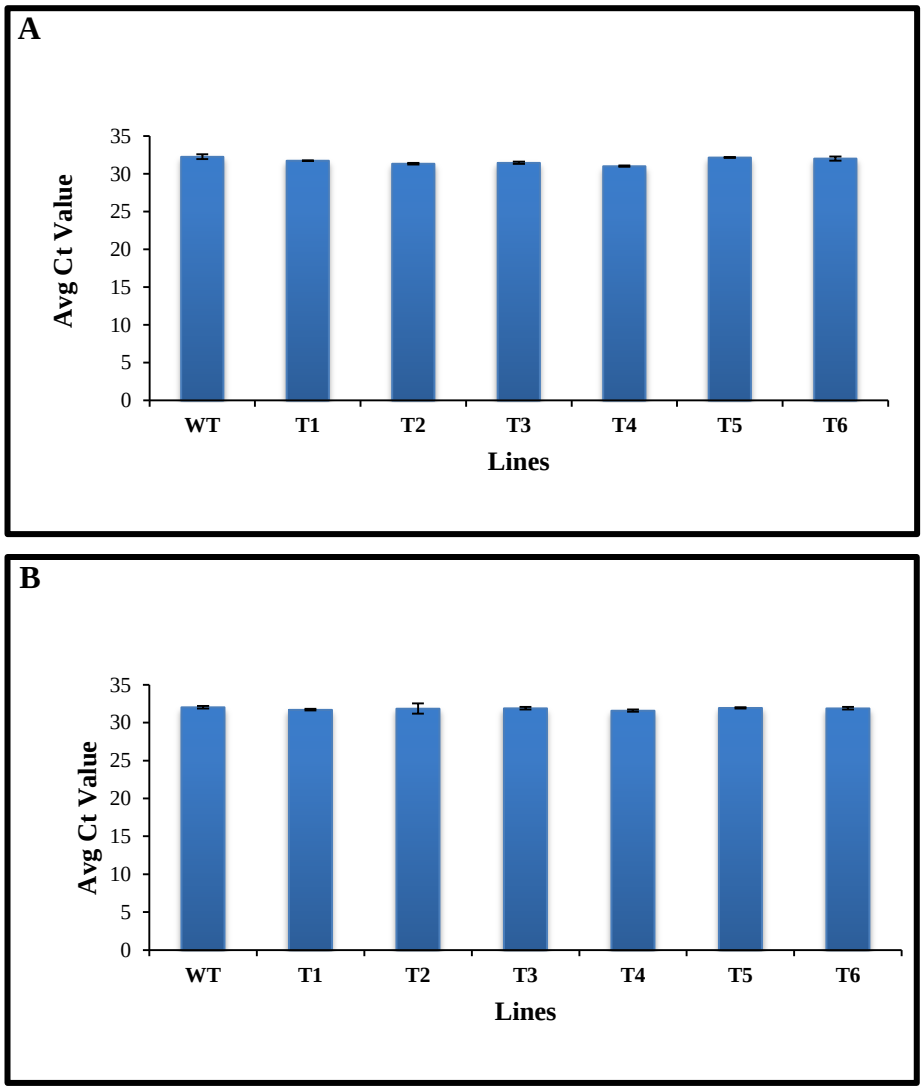


Fig. 4.8. Analysis of expression pattern of *GAPDH*. Expression pattern of GAPDH was analyzed to assess its suitability as endogenous reference gene for normalization of qRT-PCR expression (**A & B**).



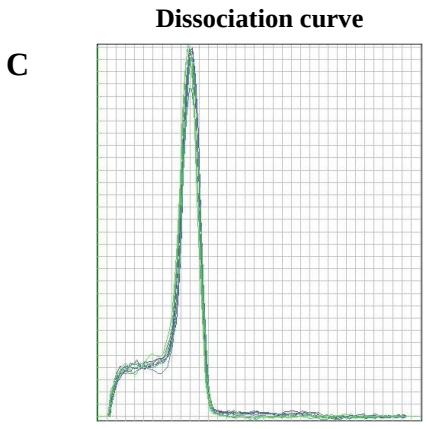


Fig. 4.9. Analysis of expression pattern of *actin*. A steady level of expression was observed in case of *actin* (**A & B**) as revealed by dissociation curve (**C**).

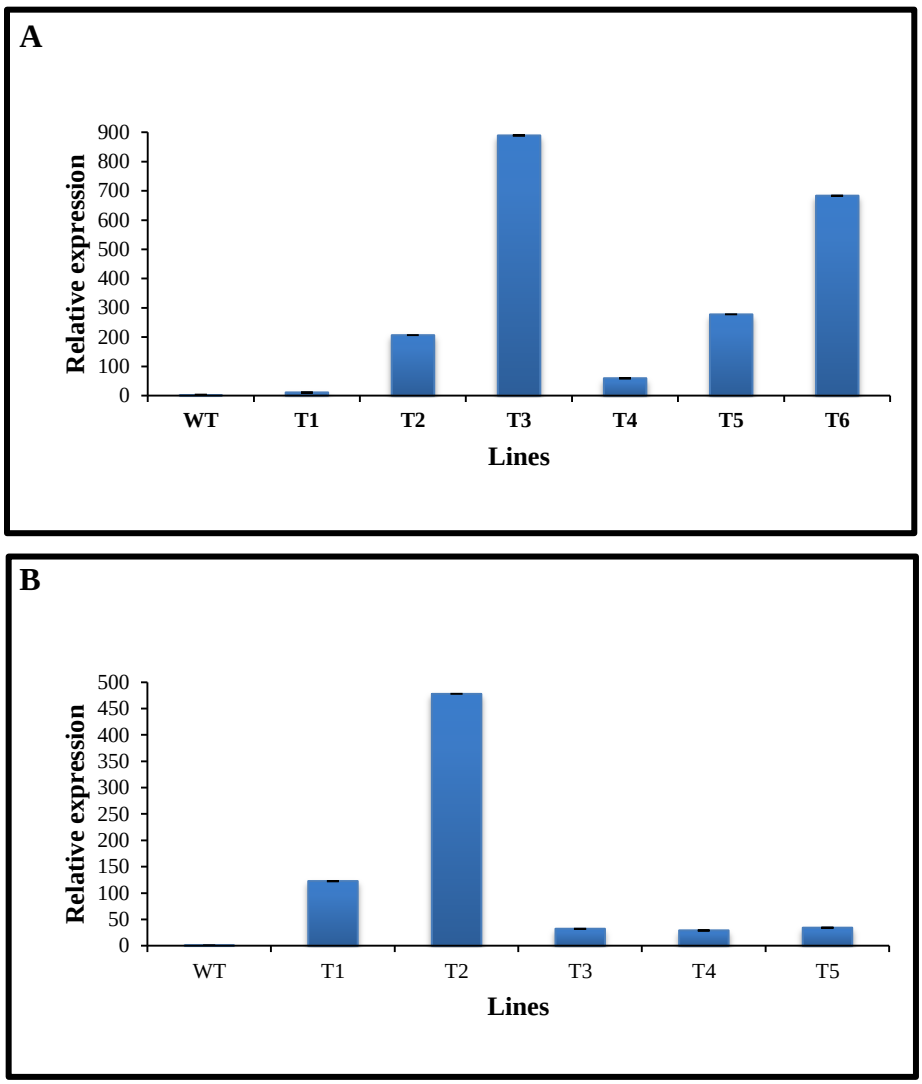


Fig. 4.10 Transcript accumulation of *AmA1*. Transcript accumulation of AmA1 was analyzed taking *actin* as endogenous reference gene in selected lines. A higher level of transcript accumulation of the transgene showed a variable pattern of expression (**A &B**).

***Amaranthus hypochondriacus* seed protein AmA1 gene, complete cds and 3' UTR**

ATGGCGGGATTACCAGTGATTATGTGCCTAAAATCAAATAACAACCAGGAGTACTTAAGATATCAAAG
TGATAATATTCAACAATATGGTCTTCTTCAATTTTCAGCTGATAAGATTTTAGATCCATTAGCTCAAT
TTGAAGTCGAACCTTCCAAGACTTATGATGGTCTTGTTCACATCAAATCTCGCTACACTAACAAATAT
TTGGTTAGGTGGTCTCCCAATCATTATTGGATTACAGCATCAGCCAATGAACCAGATGAAAATAAAAG
CAATTGGGCATGCACATTATTCAAACCACTTTACGTAGAAGAAGGTAACATGAAAAAGGTTCGACTTT
TGCACGTCCAATTAGGTCATTATACAGAAAATTATACCGTTGGTGGGTCCTTCGTATCATACTTATTT
GCCGAATCAAGTCAAATTGATACCGGCTCTAAAGACGTATTCCATGTCATAGATTGGAAATCAATCTT
TCAATTTCCCAAAACATATGTCACATTTAAAGGAAATAATGGAAAATATTTAGGGGTTATCACAATTA
ATCAACTTCCATGTCTACAATTTGGGTATGATAATCTTAATGATCCAAAGGTGGCTCATCAAATGTTT
GTCACTTCTAATGGTACTATTTGCATTAAATCCAATTATATGAACAAGTTTTGGAGACTCTCTACGGA
TAATTGGATATTAGTCGATGGGAATGATCCTCGCGAAACTAATGAAGCTGCTGCGTTGTTTAGGTCGG
ATGTGCATGATTTTAATGTGATTTCGCTTTTGAACATGCAAAAAACTTGGTTTATTAAGAGATTTACG
AGTGGTAAGCCTGAGTTTATAAATTGTATGAATGCAGCTACTCAAATTGTTGATGAAACTGCTATTTT
AGAGATAATAGAATTGGGATCCAACAACTAATATATTGGATTGCTTTTAAGATTCAAATTAAAGTCTA
GTTGTTAATGTAAGGAATAAAACGTTGTAAGTCGTCTCTTTGGAAACAAGAGGGTTCTTCCTTG

Blue- New amplicon
Green- Old amplicon
Underlined- 3' UTR
Red and blue- Seq considered for new primers

New primers for AmA1 real time PCR from cDNA and 3'UTR

AmA1 RTUTRFP
GAGATAATAGAATTGGGATCCAACAAC

AmA1 RTUTRRP
CCAAAGAGACGACTTACAACGTTTT

Fig. 4.11. Gene specific primers with increased specificity. AmA1 cDNA sequence with 102 bp of 3' UTR is shown in the figure and new primer set was made taking cDNA as well as 3' UTR region (shown in red and blue colors).

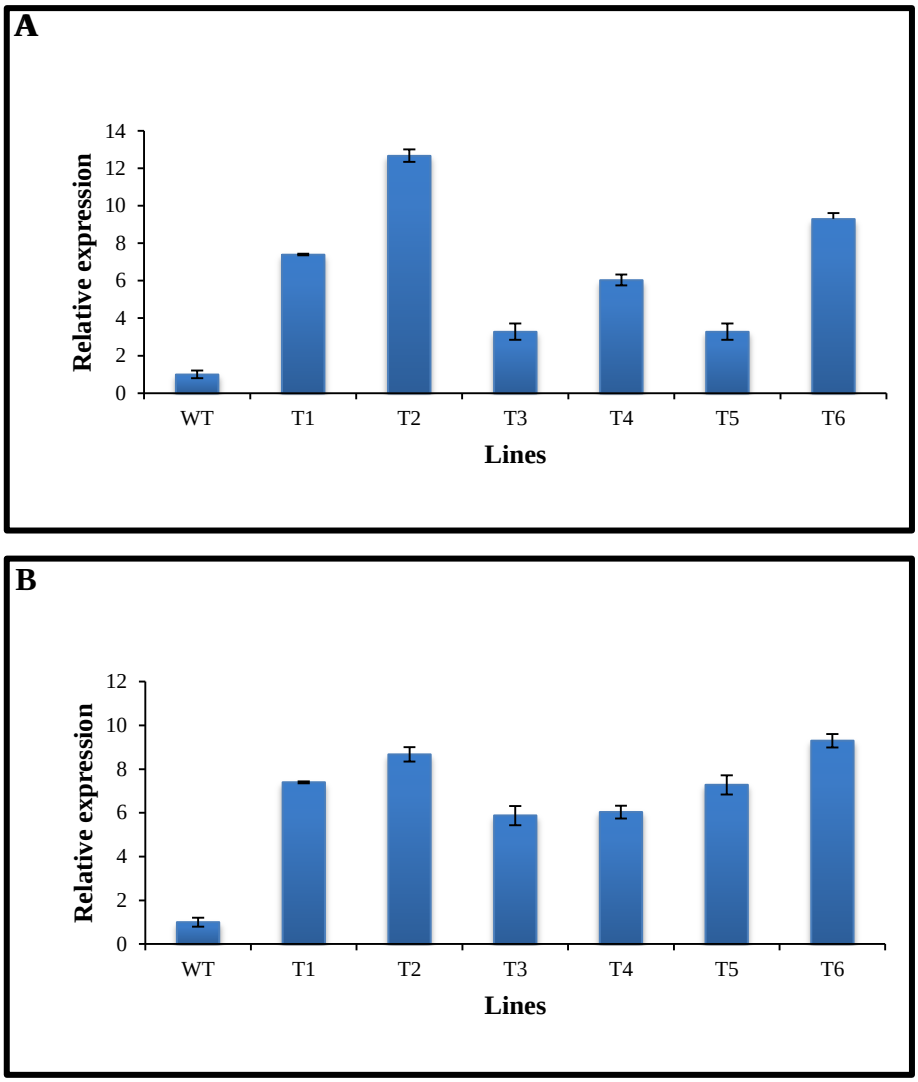


Fig. 4.12. Transcript accumulation transgenic lines. The expression level of transgene, AmA1 was analyzed by qRT-PCR in COE lines (**A**) and in TOE lines (**B**).

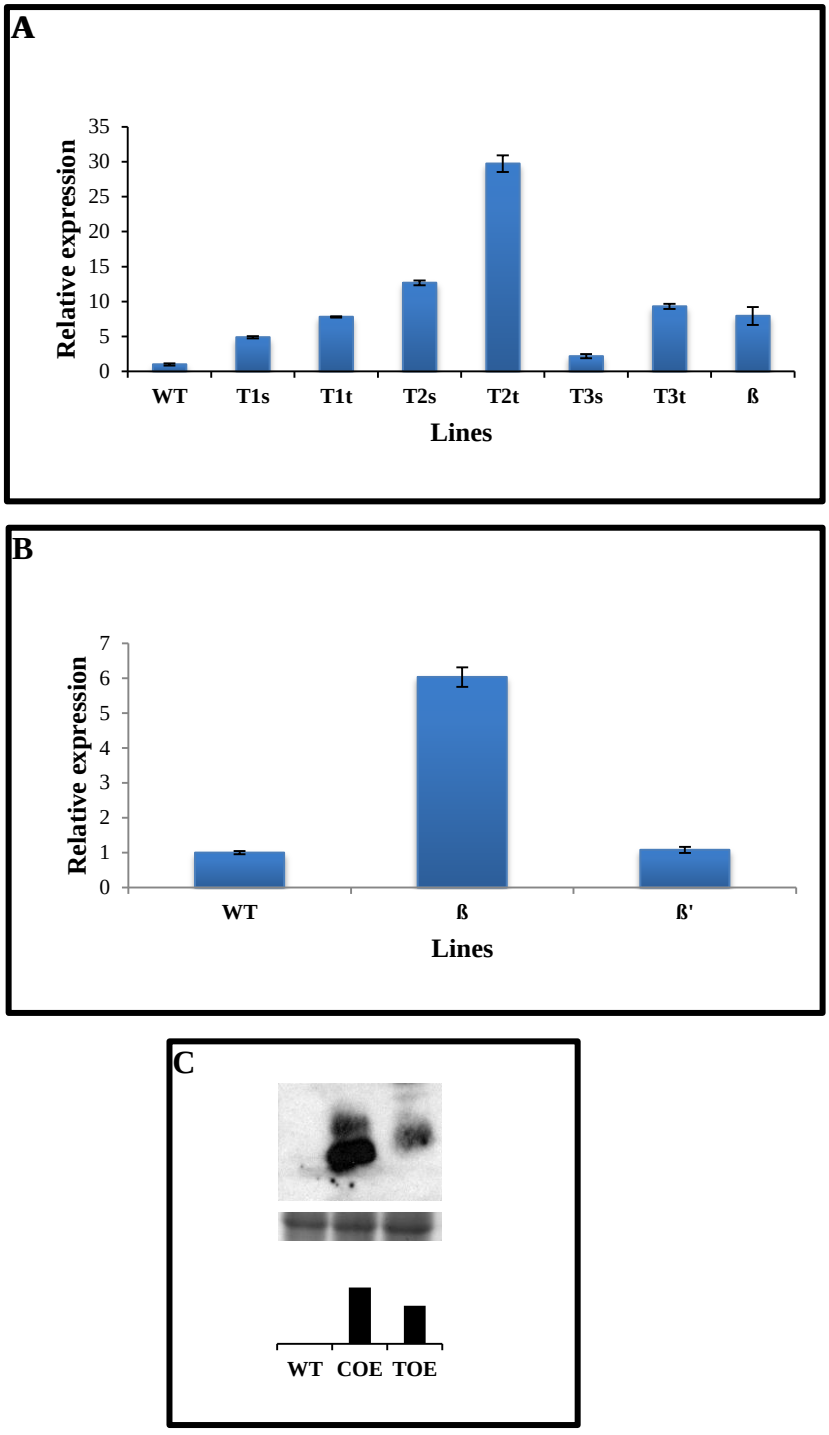


Fig. 4.13. Comparative analysis of transgene expression in COE and TOE lines. The expression level of *AmA1* in the shoots and tubers of COE lines was compared with the expression level of *AmA1* in tuber of tuber specific line by qRT-PCR (**A**) and further leaky expression was checked in aerial portion the tuber specific lines (**B**). ß and ß' represents tuber and aerial portion, respectively. Expression was also compared by immunodetection by AmA1 specific antibody and densitometry was done by Quantity One Analysis Software (**C**).

Table 4.2: Transgene copy number in different transgenic lines

COE Lines	Copy no.	TOE Lines	Copy no.
T1	1	1ß	1
T2	1	2ß	1
T3	1	3ß	1
T4	1	4ß	1
T5	1	5ß	1
T6	2	6ß	1
T7	2	7ß	3
T8	1	8ß	1
T9	2	9ß	1
T10	1	10ß	1
T11	3	11ß	2
T12	1	12ß	1
T13	1	13ß	1
		14ß	1
		15ß	1

4.3.6 Comparative proteomics

To evaluate the increase in protein content due to the transformation of AmA1, proteins were extracted simultaneously from same mass of mature tuber, and an equal volume (250 µl) of protein was separated by 2-DE as described in Materials and Methods. Another approach was to investigate the differential regulation of proteins across the wild type and transgenic line. As a primary step, a reproducible 2-DE method was established by taking two different pH ranges of IPG strips *i.e.* 4-7 and 3-10. The comparative analysis of protein spots in 2-DE gels revealed that most of the proteins are resolved in a range of pH 4-7 (**Fig. 4.14**). Therefore, further 2-DE analysis was performed only with the IPG strips of pH 4-7.

To check transgene effect(s) vis-à-vis to the increase in total protein content at molecular level, proteins were isolated simultaneously from wild type and transgenic tubers taking equal weight in equal volume of buffer. To examine the effect of AmA1 towards the increase in protein content, the changes in the tuber proteome of transgenic sweet potato was monitored primarily using 1-D electrophoresis (1-DE). The proteins were extracted simultaneously from same mass of mature tuber from wild type and transgenic lines, and equal sample volumes (100 µl) were loaded in 1-D gels. The analysis showed no change in overall protein profile in both the transgenic lines. However, change in total contents of proteins was evident in both the COE and TOE lines than that of wild type (**Fig. 4.15 A**). Since 2-DE is one of the most proficient and potent methods to study intricate details of gene expression at the level of proteins and the separation and resolution is its advantages over 1-DE, a similar analysis was performed also by 2-DE. Three parallel 2-D gels, representing three technical replicates with equal volume from each wild-type and transgenic sample were run and then computationally combined into a representative standard gel, the first-level match set. Analysis of 2-D gels revealed an increase in total number of spots (491) in transgenic in comparison to the wild type (467) (**Fig. 4.15 B**).

To investigate the transgene mediated change in tuber proteome both the wild type (WT) and COE line was subjected to comparative proteomic analysis by 2-DE. The first-level match set was produced as described earlier and the replicates had a correlation co-efficient of variation above 0.8 as displayed by scatter plots (**Fig. 4.16)**. Further, the gels showed more than 90% high quality protein spots suggesting high reproducibility among the replicates (**Table 4.3**). Furthermore, a second level match set was then developed, which allowed comparison of the

standard gels from both the wild type and COE line and a second normalization was done with a set of three unaltered spots identified across the gels. The filtered spot quantities from the standard gels were assembled into a data matrix of high-quality spots from all the representative gels for further analysis. A comprehensive 2-DE analysis revealed several contrasting trends, as well as the commonalities in the proteomes of both the wild type and transgenic lines under investigation (**Fig. 4.17 A and B**). Whereas 303 protein spots were found to be common to both the WT and COE, 102 spots were exclusive to WT and 111 to COE (**Fig. 4.17 C**). The comparative proteomic analysis of WT and COE unraveled several common as well as transgene-specific expressions of proteins suggesting that subtle changes in the genome might lead to distinct proteome.

4.4 Discussion

The most important objective of this work was to evaluate the successful transgenic events at molecular level. Initially, an efficient regeneration and transformation system was established in sweet potato by *A. tumefaciens* mediated plant transformation method using *GUS* as a reporter gene. The kanamycin selected putative transformants were examined for successful integration and expression of transgene at genome, transcript and protein level. The expression of various traits in the transgenic plants can be severely impeded by the individual characteristics and unadaptability of the alien genes and proteins to the particular subcellular environment of the new host [346]. Since sweet potato is a tuber crop, TOE lines were generated using pSB8ß construct in which transgene *AmA1* is under the control of ß-amylase promoter to facilitate tuber specific expression. The COE lines were generated using pSB8 construct in which *AmA1* is under the control constitutive CaMV-35S promoter. PCR as well as expression analysis revealed the successful integration and expression of transgene. Transformation efficiency was measured as percentage of confirmed transgenic plants out of total number of plants regenerated. PCR analysis revealed 72% and 60% transformation efficiency in cv. SP-6 and SP-17 respectively which was quite higher than the earlier reports of 20% and 30.8% [248, 250]. The transcript accumulation and enzyme activity were also higher in cv. SP-6. Therefore, cv. SP-6 was selected for further investigations. In order to improve the nutritional status of sweet potato in terms of protein, the *AmA1* was introduced in such a way that expression would be influenced in both constitutive and tuber-specific manner.

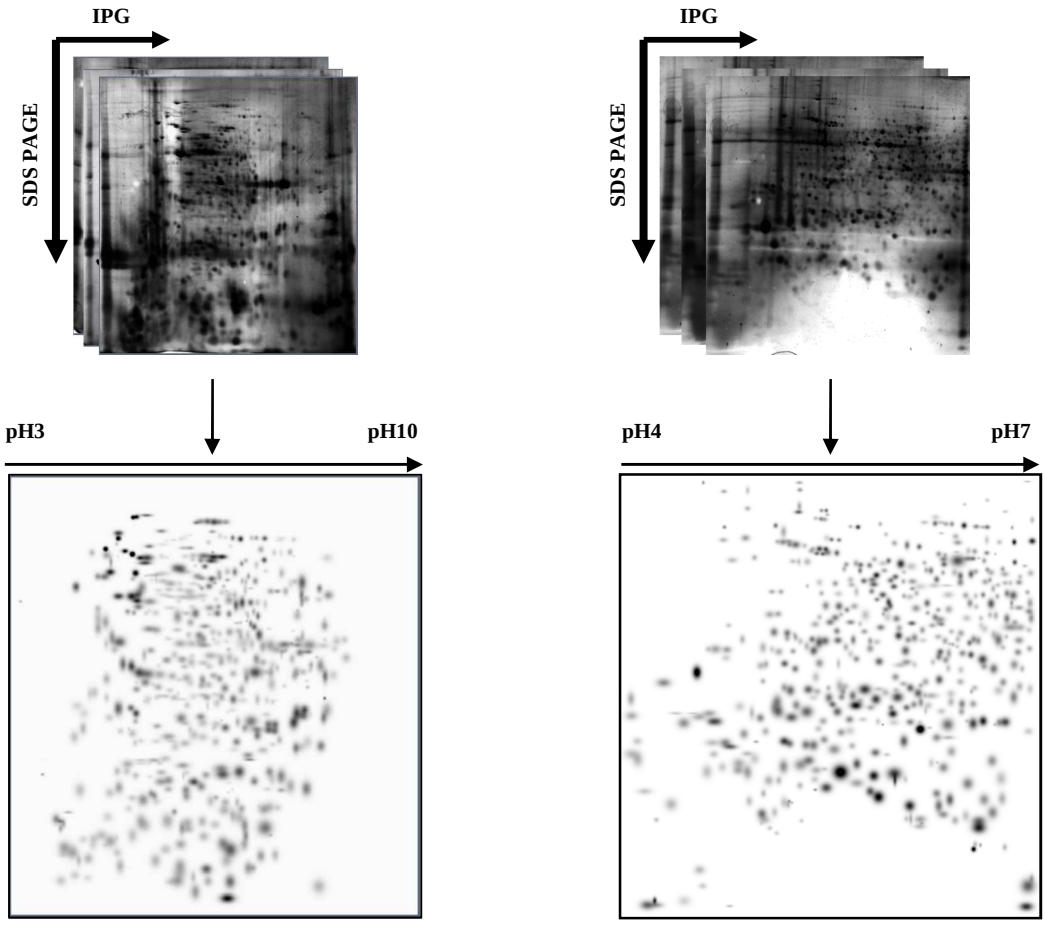


Fig. 4.14. Identification of suitable pH range. Suitable pH range was determined for 2-DE analysis taking two different pH ranges of IPG strips *i.e.* 4-7 and 3-10.

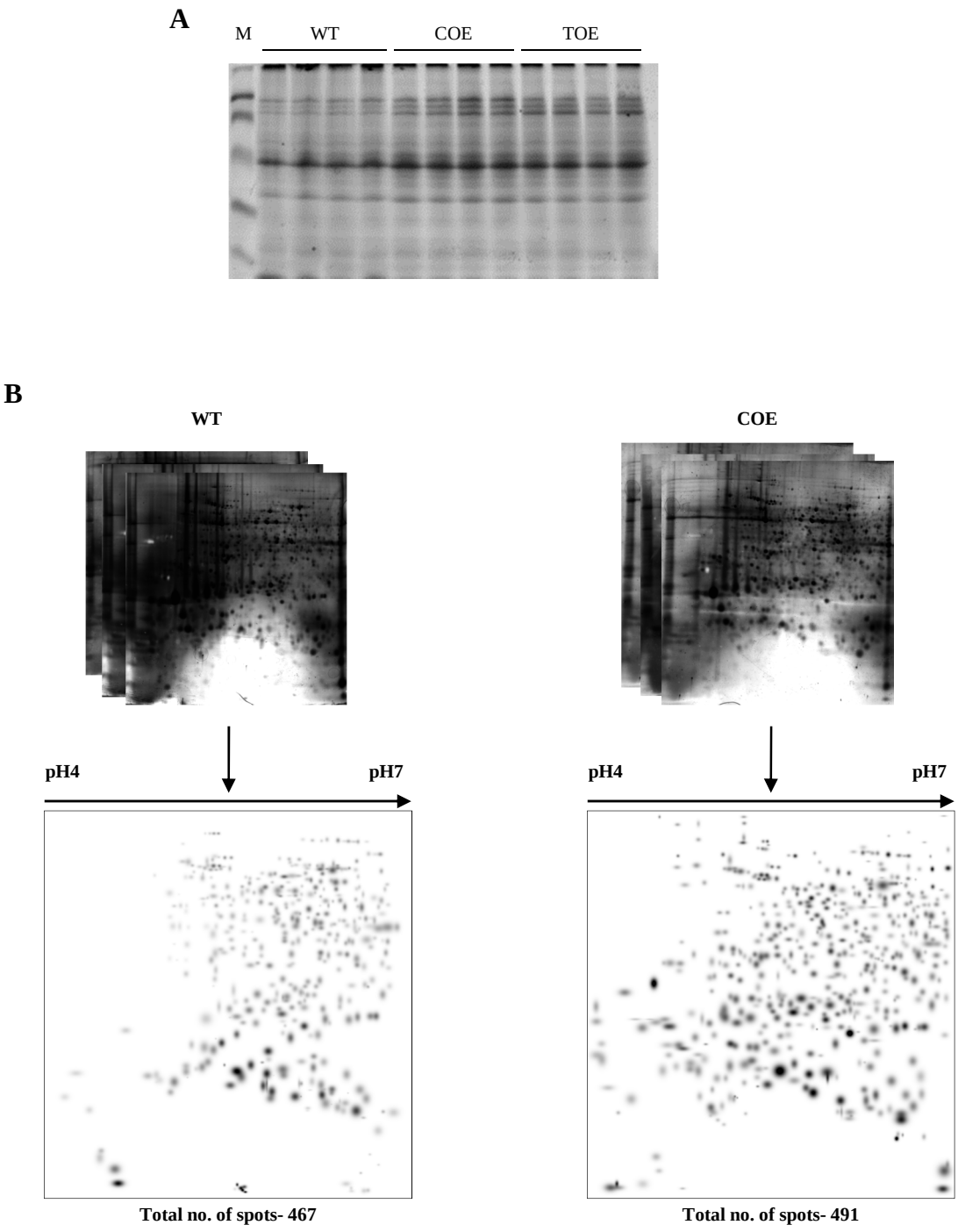


Fig. 4.15. Comparative proteomic analysis. Proteomic analysis of the mature potato tubers of wild type (WT) and the transgenic line (COE and TOE) was performed. Proteins were extracted from same mass of mature tuber, and an equal volume of 100 µl and 250 µl of protein was separated by 1-DE (**A**) and 2-DE (**B**), respectively.

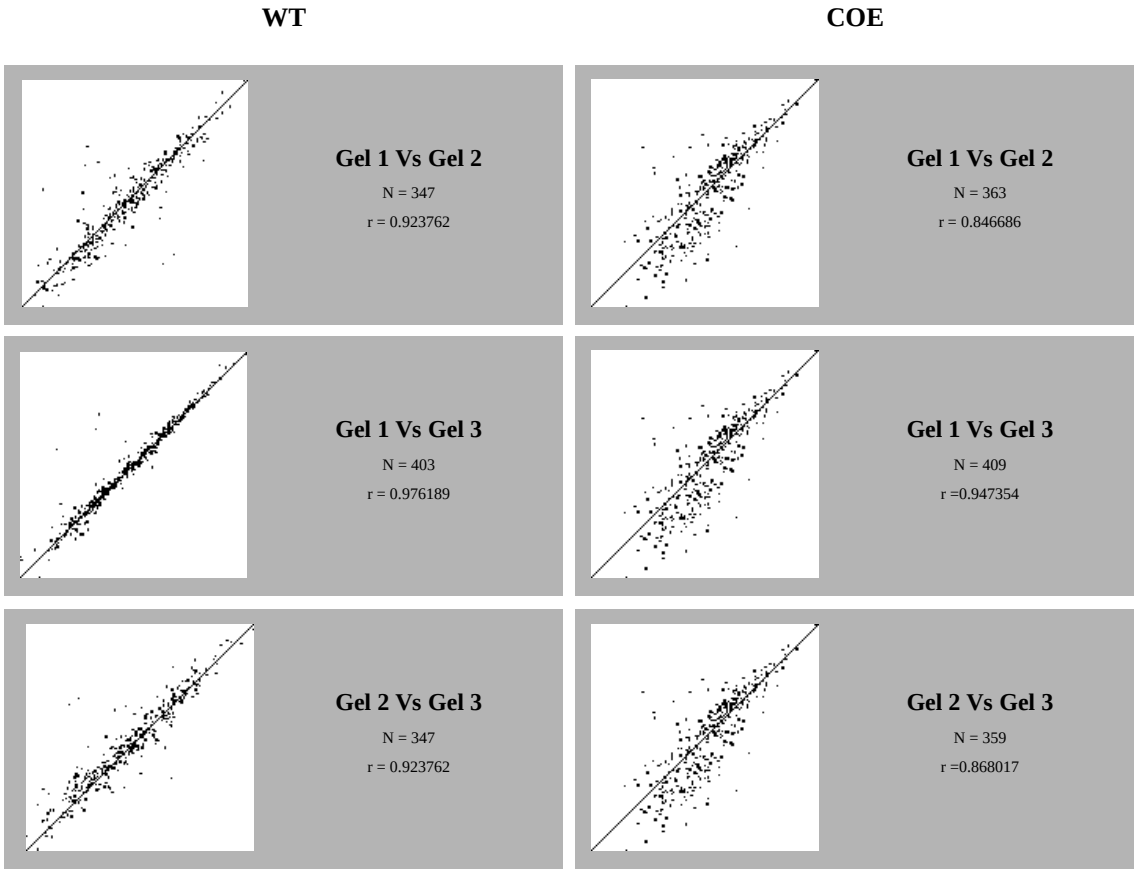


Table 4.3 Reproducibility 2-Dimensional

Lines	Average no of spots	High quality spots	Reproducibility (%)
WT	376	367	97.61
COE	384	370	96.35

Fig. 4.16. Reproducibility of 2-Dimensional gels. Scatter plots displaying a correlation coefficient of variation above 0.8 between the three replicates in wild type (WT) and transgenic line (COE). Percentage of reproducibility was determined by high quality spots out of the total spots. High quality score > 30 was assigned by PDQuest software

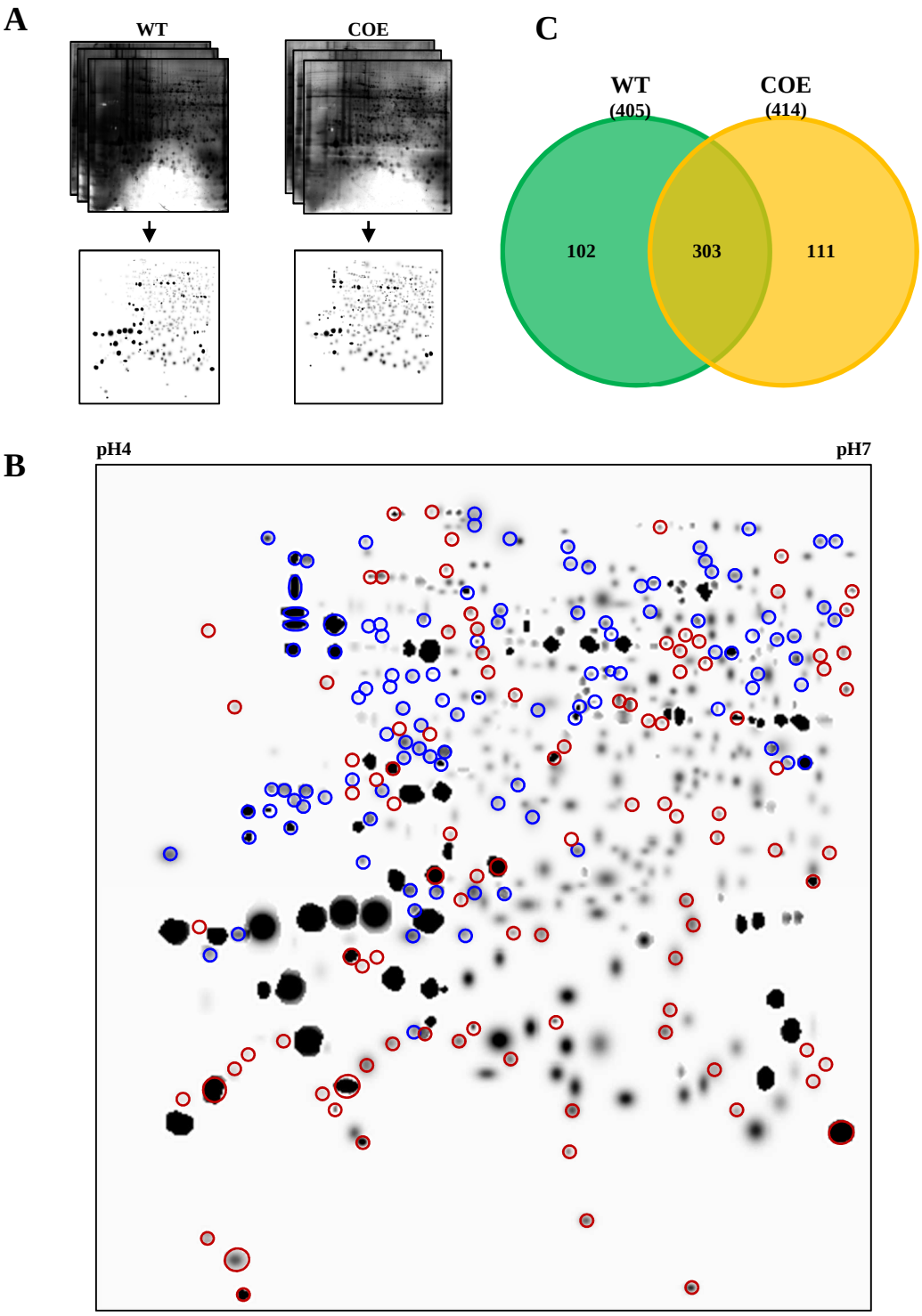


Fig. 4.17. Comparative proteomic analysis of differential tuber proteome. The proteomes displaying three replicate gels were computationally integrated into the "standard gel" (**A**). The differential proteome was developed from the "standard gels". The exclusive protein spots are shown with red and blue circles for wild type (WT) and transgenic (COE) lines, respectively (**B**). The Venn diagram shows the common and exclusive proteins in WT and COE tubers (**C**). The areas in the diagram are not proportional to the number of proteins in the groups.

Successful integration of transgene in the host genome is not enough criteria to guarantee the expression of the desired traits. Therefore, it is pertinent to check the stable and steady expression of transgene in the alien environment. Northern blot, dot blot, microarray and qRT-PCR are few methods which can deduce the transcript accumulation pattern of a specific gene including the transgene itself. Northern blot analysis and reverse-transcription polymerase chain reaction (RT-PCR)-based techniques are more commonly used to analyze transgene expression. Quantification of an individual gene in qRT-PCR takes place during the exponential phase of the amplification; therefore the sensitivity of this method is very high [347, 348]. qRT-PCR is an extensively used technique for gene expression analysis owing to its sensitivity, reproducibility and dynamic quantification range allowing the detection of both low-abundance mRNAs and slight variations in gene expression. Consequently, qRT-PCR has become the ideal approach for the substantiation of high-throughput analysis such as microarray and the quantitation of gene expression [349, 350]. Based on normalization using a steadily expressed reference gene it is employed to quantify relative levels of gene(s) expression [349, 351]. Accurate normalization is indispensable for reliable qRT-PCR outcomes otherwise the results obtained may perhaps lead to erroneous conclusions and could have significant impact [352, 353]. Despite the importance of systemic validation of reference genes, this is still underutilized in plant sciences [354]. This implies that, prior to their use in qRT-PCR normalization, prospective reference genes must be thoroughly evaluated for their stability under the applied experimental conditions. Gene expression analyses for a broad range of experimental conditions have relied on the use of conventional housekeeping genes such as *actin*, *tubulin*, *GAPDH*, elongation factor-1α (EF1α), and 18S rRNA, for normalization of the data. Nonetheless, in several cases, the transcripts expressed from these genes are unstable, such that variations in the respective expression levels can lead to a misinterpretation of the outcomes. In this study, reference gene was standardized for the accurate qRT-PCR normalization considering three housekeeping genes *Tubulin*, *GAPDH* and *Actin*. Out of these three, *Actin* was found to be the ideal reference gene for the accurate normalization of the qRT-PCR analysis. In a similar analysis in sweet potato by **Park et al. (2012)**, the stability and expression of ten isolated candidate reference genes of sweet potato was analyzed using two algorithms, geNorm and NormFinder. The samples representing four sweet potato cultivars were subjected to four different stress conditions, *i.e.*, cold, drought, salt and oxidative stress. The findings revealed that, for sweet potato, discrete reference genes or

combinations thereof should be preferred for further use in data normalization subject to the experimental conditions and the particular cultivar. In general, the genes *ADP-ribosylation factor* (*ARF*), *ubiquitin extension protein* (*UBI*), *cytochrome c oxidase subunit Vc* (*COX*), *glyceraldehyde-3- phosphate dehydrogenase* (*GAPDH*) and *ribosomal protein L* (*RPL*) were validated as the most suitable reference gene set for every cultivar across the samples examined. Intriguingly, the genes *actin* and *tubulin*, even though extensively used, were not found to be most suitable reference genes [350]. In this study *actin* was found to be the suitable reference gene at least in cv. SP-6 and SP-17; while, *GAPDH* was not found to be an ideal reference gene. Findings with respect to the *tubulin* further corroborated this investigation (**Fig. 4.5-4.7)**. A similar study has been done with several candidate reference genes across a range of tissues, organs, developmental stages, and stress conditions in different plant species, such as Arabidopsis, tobacco, rice, tomato, potato, poplar, melon and many others [355-360].

Further analysis of transcript level of the transgene revealed a range of >3- to 12-fold higher expression in different COE and TOE lines (**Fig. 4.10 A and B)**. To study the developmentally regulated expression of transgene, *AmAl* in sweet potato as well as the stability of its expression, a comparative analysis was conducted to check the transcript level in developing and mature tubers from the COE and TOE lines. Both independent lines were selected on the basis of high expression of transgene. The transcript accumulation pattern by northern analysis revealed higher transcript abundance in COE line than that of the TOE line. Additionally, the expression of *AmA1* was invariably more stable in COE line. This was further correlated by qRT-PCR analysis which revealed no leaky expression in the aerial portion of TOE line as well (**Fig. 4.4 and 4.11**). Leaky expression of transgene of tissue specific or marker gene(s) have been reported earlier by several studies [361, 362].

Photosynthetic activity of the leaf canopy (source), the ability of the plant to translocate the photo-assimilates to the tuberous root (sink) and the capacity of the tuberous root to accommodate or capture assimilates are the deciding factors for total dry matter of the tuberous roots of sweet potato [363]. Elevating the source and the sink activities and its maintenance for extended periods pertaining to higher economic yield is one of the crucial aspects for agronomic improvements. Source strength and sink capability consist of two key components, namely the size and activities of the source and the sink [364]. Tuberization is a complex process involving

various metabolic cues that too include massive accumulation of starch and proteins. Tuber as a storage organ acts as a sink and competes for the available photo-assimilates [61, 62]. Indeed, transgenic crops with changes in source and/or sink capacities were produced which were either impaired [365] or superior in yield [366], demonstrating the feasibility and putative impact of such approaches on yield increase. Nevertheless, findings from this study suggested the lack of leaky expression in the TOE lines. However, the expression of transgene in COE line was higher which was in concordance with immunoblot analysis (**Fig. 4.13 B**). Except few lines the transcript accumulation pattern of COE and TOE lines further substantiated the immunoblot analysis *i.e.* the expression of AmA1 at protein level (**Fig. 4.12 and 4.13 B**). Since during the tuberization process the entire photo-assimilates are directed towards the tuber (sink) and COE lines having an additional advantage of AmA1 expression even in the source tissues. Therefore, cumulative expression of AmA1 in source as well as in sink tissue may perhaps be the reason behind the higher expression of transgene in COE lines in comparison to the TOE line. Additionally, seed storage proteins have a crucial role in storage tissue therefore source directed expression of AmA1 in tubers is but natural. Even after maturity, there is a constant expression of AmA1 in source tissue that may be the reason behind its steady expression in COE lines. Otherwise also AmA1 is reported to be localized in cytoplasm unlike many seed storage proteins which usually localize in protein storage bodies. Even at protein level TOE lines evidently were found to maintain the expression of AmA1 as revealed by immunodetection of transgene (**Fig. 4.13 B**). The delayed breakdown of this protein could also be due to its cytoplasmic localization. Proteolytic enzymes are not released into the cytosol until the final stages of cellular disorganization and an analogous mechanism may regulate AmA1 expression as well [127].

Genetic transformation of plants has become an extensively used expertise that serves several objectives in plant biotechnology. However, considerable disparity in transgene expression is frequently observed within populations of transformants transformed with even the same transgene construct and often vary >100-fold with respect to transgene expression levels. These inter-transformant differences of transgene expression delimits proper evaluation of transgenes and might be most undesirable when high-throughput transgene screening is expected. The common plant transformation strategy now is to produce a sufficiently high number of transformants to find some of them with the desired level of expression. Different factors are thought to contribute towards the variation in transgene expression including the

transgene copy number, RNA silencing, transgene insertion site and the involvement of certain regulatory sequences to drive transgene expression [367, 368]. With the genetic transformation methods currently used in plants, it is neither feasible to introduce a defined number of transgene into the genome nor possible to target efficiently the alien DNA to specific positions in the genome. Consequently, repeat arrangements and truncated and/or rearranged transgene copies are often observed in transgenic lines. Therefore, independent transgenic lines vary with respect to number, arrangement, and position of transgene copies they harbor in the genome [367]. Nonetheless, *A. tumefaciens* mediated genetic transformation is a random process and usually results in lower transgene copy numbers than direct transformation methods [369]. The presence of multiple transgene copies show lower or unstable transgenic expression and occasionally implicated in transgene mediated gene silencing. While multiple copies of the transgene are helpful for overexpression experiments, single copy transformation events are preferred for the majority applications as they are constant over several generations of subsequent breeding [370, 371]. Transgene copy number can be positively or negatively associated with transgene expression as plants with larger number of transgene copies resulted in a lower level of transgene expression, unstable expression or even gene silencing. On the other hand, the insertion of only one or two copies tends to result in higher levels of expression [372, 373]. Southern blot analysis is the traditional method generally used for determining the copy number of a transgene. However, this method is laborious and time-consuming, and requires large quantity of genomic DNA and use of hazardous radioisotope. Moreover copy number detection using all these techniques does not reflect the rearranged transgene copies lacking relevant restriction site(s) and in cases of concatemers [340]. For determination of transgene mRNA, Northern blot analysis and reverse transcription followed by the polymerase chain reaction (PCR) are often used, but the results obtained with these methods are usually only qualitative or semi-quantitative. qRT-PCR for copy number detection is an effective, sensitive technique and consumes less time than the southern blotting [371]. In present study, majority of the transformants from both the categories have single copy of transgene insertion. However, few COE as well as TOE lines harbor more than one copy of transgene but the maximum transgene copy number was found to be three. Moreover, both the positive and negative correlation was observed with the transgene copy no and its expression.

Comparative proteome profiling has remained an attractive tool for unambiguous comparison between cultivars of a particular crop species; even to categorize single mutations with multiple effects, although the outcome is affected by genotypic divergence, organs and tissues, developmental stages besides, different environmental cues. The proteomics studies revealed the influence and efficacy of such approach to distinguish the landraces, populations, varieties and even species [148, 149]. Proteomic analyses of plant organs or tissues have been used to monitor developmental changes in seed and tuber, environmental stress responses, and transgene mediated changes in transgenic crops as well [121, 148, 149, 289, 295, 299, 374].

In order to perceive transgene mediated changes in tuber proteome, a comparative proteomic analysis of the wild type (WT) and one promising COE line was performed. The comparative proteomic analysis of the WT and COE lines unraveled several conserved as well as unique expression of proteins, suggesting that subtle changes in the genome might lead to distinct proteome (**Fig. 4.17**). The relative distribution of seed proteins is principally genetically determined, besides the variability caused by nutrient modulation and environmental influences. Storage tissues have intrinsic compositional plasticity concomitantly from the alteration of the source-sink relationship; this may be disturbed by the accumulation of alien proteins as an alternative sink protein [375, 376]. Despite the fact that there are many studies concerning the genetic control and spatial and temporal regulation of seed storage proteins, our current knowledge about their exact functional role and physiological relevance remains unknown. Only recently focus has been given to study crop plants overexpressing storage protein [376, 377]; however, the modus operandi of cellular network and physiological consequences toward sensing a storage protein has not been elucidated. Overexpression of extrinsic or underexpression of any intrinsic protein directs the cellular system for an overall proteome rebalancing. Reduced or increased levels of these proteins is compensated by the expression of some other proteins leading to the rebalancing of the nitrogen sink to maintain the metabolic cues at more or less a constant level. Moreover, evolutionary background of proteome rebalancing is also reported earlier [378]. Seed storage proteins are known to serve as the sink to regulate the movement of photosynthate into developing organs [379]. It is conceivable that AmA1 as a storage protein might act as a sink protein in transgenics, thereby regulating the movement of metabolites, including the amino acids, into the developing tuber where they are fixed into newly synthesized proteins and consequently enhance the level of essential amino acids. In addition, it has been

revealed that AmA1 regulates the biology of nutrition, and its ectopic expression in storage organs such as tuber and seed helps increase protein and amino acid accumulation [112, 113, 288]. This perhaps also is the reason for the increased no of protein spots in COE proteome. When protein isolated from equal weight of tissues from WT and COE line, COE line showed a marked increase in protein spots (**Fig. 4.15 B**). It has also been established that AmA1 plays an important physiological role during organ development and is crucial for the homoblastic growth in potato tubers, the gradual transition cascade from stolon to mature tuber. In a recent study, it has suggested that proteome rebalancing due to AmA1 expression might lead to nutrient enhancement and increased yield. It is thus conceivable that AmA1 might play a crucial role during seed germination and seedling growth as a nutrient source and growth-promoting substance [113, 121]. The comparative proteomic approach in AmA1 overexpressed potato tuber also revealed its role in tuberization. AmA1-regulated functional protein network and its combinatorial effects cause the protein enhancement and determine the organ development, tuber in particular. The introduction of the AmA1 gene by means of a constitutive and tuber specific promoter may lead to numerous changes within the plant proteome that may be related to many pathways or general expression variation of individual proteins [121]. Therefore, in-depth study of the uniquely expressed as well as the conserved proteins tubers will shed new insight on AmA1 regulated proteome rebalancing in sweet potato.