StBEL5, a transcription factor in potato (*Solanum tuberosum* L.): sequence analysis, amplicon profile and expression pattern

A
Dissertation
Submitted in the partial fulfillment of the requirement for the award of degree of

Master of Science

In
Biotechnology

Submitted by:
NITIKA GARG
(Regd. No. 301401011)

Under the guidance of
Dr. N. Das
Professor

Department of Biotechnology
Thapar University, Patiala
July, 2016

CANDIDATE'S DECLARATION

I, hereby declare that the work which is being presented in the thesis entitled, **"StBEL5, a transcription factor in potato (*Solanum tuberosum L.*): sequence analysis, amplicon profile and expression pattern"** in the partial fulfillment of the requirement for the award of degree of Master of science in Biotechnology, Thapar University, Patiala, is an authentic record of my own research work carried out under the guidance and supervision of **Dr. N. Das**, Professor, Department of Biotechnology, Thapar University, Patiala, India. The matter embodied in this dissertation has not been submitted to any other university or institute for award of any other degree.

Nitika Garg

(Nitika Garg)

Place: Patiala
Date: July 15, 2016

CERTIFICATE

This is to certify that the dissertation entitled **"StBEL5, a transcription factor in potato (*Solanum tuberosum L.*): sequence analysis, amplicon profile and expression pattern"** submitted by **Nitika Garg** (Regd. No.301401011) in partial fulfillment of the requirement for the award of the degree of Master of science in Biotechnology, to Thapar University is a record of student's own work carried out by her under our guidance and supervision. The report has not been submitted for the award of any other degree or certificate in this or any other university or institute.

15/07/2016

(Dr. N. Das)
Supervisor
Professor
Department of Biotechnology
Thapar University
Patiala

(Dr. Dinesh Goyal)
Professor and Head
Department of Biothechnology
Thapar University
Patiala

(Dr. S. S. Bhatia)
Dean of Academic Affairs
Thapar University, Patiala

LIST OF ABBREVIATIONS

Bp	Base pairs
UDP	Uridine diphosphate
ADP	Adenosine diphosphate
INV	Invertase
kDa	Kilo Dalton
SUS	Sucrose Synthase
Nts	Nucleotides
aa	Amino acids
NCBI	National centre for Biotechnology Information
UTR	Untranslated region
kb	Kilo base pairs
cDNA	Complementary DNA
dNTP	2-deoxynucleoside-5-triphosphate
EDTA	Ethylenediamine-tetra acetic acid
TE	Tris EDTA
TFs	Trascription factors
SD	Short day
h	Hours

List of Contents

ABSTRACT

StBEL5 is a long distance travelling RNA transcript which acts as a mobile signal. It is transported from leaves to the stolon tip during the early stage of tuber formation. STBEL5, a transcription factor, is dimeric in nature ($M_{r\sim}$ 76459.1 kDa) with pI value of 6.97. It is important member of TALE superfamily which contacts with KNOTTED 1-type genes to regulate tuber formation. Considerable progress has been made on BEL5. However, with regard to the Indian potato cultivars there is no report on BEL5 till date. Sequence analysis and comparison were made using the available BEL5 sequences. Apart from sequence relatedness, conserved amino acids were found and some important protein motifs were predicted which are presented in a comprehensive manner in this report. Hydropathy character of StBEL5 was presented. Good quality genomic DNA was isolated and purified from different potato cultivars. Polymerase chain reactions (PCR) were carried out using StBEL5 cDNA-specific oligonucleotide primers and different potato genomic DNAs as template. Amplicon profile was noted and analyzed for each cultivar. RT-PCR approach was adopted to see the expression pattern of StBEL5 in different tissues of potato. This study would be useful for further research on BEL5 in potato and other *Solanaceae* family members.

1.1 About potato *(Solanum tuberosum L.)*

Potato (*Solanum tuberosum L.*) is a tuberous crop rich in carbohydrates, mainly in the predominant form of starch. It is a non-grain food crop, contributing to food and nutritional security in the world. It belongs to the family *solanaceae*, whose other constituents comprised of tomatoes, tomatillos, peppers, and egg plant. It is the modified stem which swells underground and enlarges to develop into a form of edible tuber (potato). There are nearly 4000 different varieties of potato has been reported, which differ in size, shape, color, starch content and flavor. Potatoes were first domesticated around 8000 years ago, in the highlands of Peruvian Andes-mountains in South America on the border bounded by Bolivia and Peru. In the 16th century, Spanish explorer (who discovered potatoes in South America) brought potatoes to Europe where it developed as temperate crop and with colonial expansion of European countries it got distributed throughout the world. Portuguese introduced potato in India in early 17th century and was first cultivated in Surat on West coast, and from this area it was spread to areas, which were also under Portuguese influence like Goa. Today, India is the fifth largest producer of potatoes in the world after china, Russia, USA, and Poland. Where as its commercial cultivation and consumption in large quantities, however began only from 1932 and today it is consider to be third largest staple food crop after wheat and rice.

The International Potato Center (CIP) aims at reducing poverty and managed food security on a constant basis in developing countries using scientific experimentation and analogous activities on sweet potato, potato, tuber crops and other root. Also they aimed at the modernized management of naturally obtained resources in the mountain areas and area of Andes. CIP headquarters are situated in La Molina, exterior of Lima, Peru’s capital, an irrigated coastal valley. In addition to this Center has one higher Andes experiment station in Quito, Ecuador, and a worldwide network of regional offices and collaborators. The potential of the potato crop was realized in India soon after independence and the Central Potato Research Institute (CPRI), Shimla, lay foundation in 1949 .The earliest authorization of Central Potato Research Institute (CPRI) is to volunteer basic and strategic research work which aimed at the reinforcement of varieties and renewable technologies for upgrading production and usage of potato in the country. They also work of producing disease-free product to meet the country’s demands. Potato is a cold weather crop, with average maturation time of 90-100 days and harvest time is between March-April. Its optimum yield is obtained when temperature is between 18°C to 20°C (64 to 68°F). It has been reported that tuber growth is inhibited at temperature below 10°C (50°F) and above 30°C (86°F). It occupied an area of 1.2 million hectare in our country with total population of 23.5 million tones. It is a perishable crop and the only major tuber crop that is grown in temperate regions. As a result of increased demand of this food crop because of increase in population rate, it has become important to redesign the potato crop for its improvement in the terms of quality and quantity. It is an annual, herbaceous, dicotyledonous and vegetativly propagated plant. Depending on the variety, it grows about 60 cm (24 in) high, with leaves dying back after flowering, fruiting and tuber formation. Generally, tubers with white flowers have white skin where as tuber with colored flowers such as pink, red, blue, or purple tends to have pinkish skin. The above-ground part of plant bears dark green, broad, compound leaves with oval leaflets and dies each winter and regrown in spring. Flowering starts three to four week after sprouting. Flowers are generally white, pink, or purple with yellow stamens. After flowering, a small green fruits are produce containing about 300 seeds. Different potato verities are grown from seeds, also called as “true potato seeds”, “TPS” or “botanical seed”. Below-ground part of plant continues to live even after above-ground part has died in winters. The potato plant generally lack woody stem. It could have three kinds of stems including sprouts (leafy stem), tubers and stolons (Beukema and Zaag 1979; Struik and Wiersema 1999). Tubers are the swollen portion of the underground modified stem which enlarges to develop into a potato. Energy in the form of protein and starch, also water is stored in this tuber (potato). They bear lateral bud (eyes) which grow into a new plant under favorable conditions.

The supreme species developed worldwide is *Solanum tuberosum* (a tetraploid with 48 chromosomes), which is a hybrid between the diploid species *solanum stentotomum* and diploid weed *solanum sparsipilum* with subsequent chromosome doubling. There is series of ploidy levels, ranging from four diploid species (with 2n = 2x = 24 chromosomes): *S. stenotomum*, *S. phureja*, *S. goniocalyx*, and *S. Ajanhuiri* to hexaploid (with 24 = 6x = 72 chromosomes). There are also two triploid species (with 2n = 3x = 36 chromosomes): *S. chaucha* and *S. juzepczukii*, tetraploid and one pentaploid cultivated species (with 2n = 5x = 60 chromosomes): *S. curtilobum*. The tetraploid cultivated potatoes are not diploid, so that there are four interchangeable genes at each locus. Hence, the plant is heterozygous and show polyploidy.

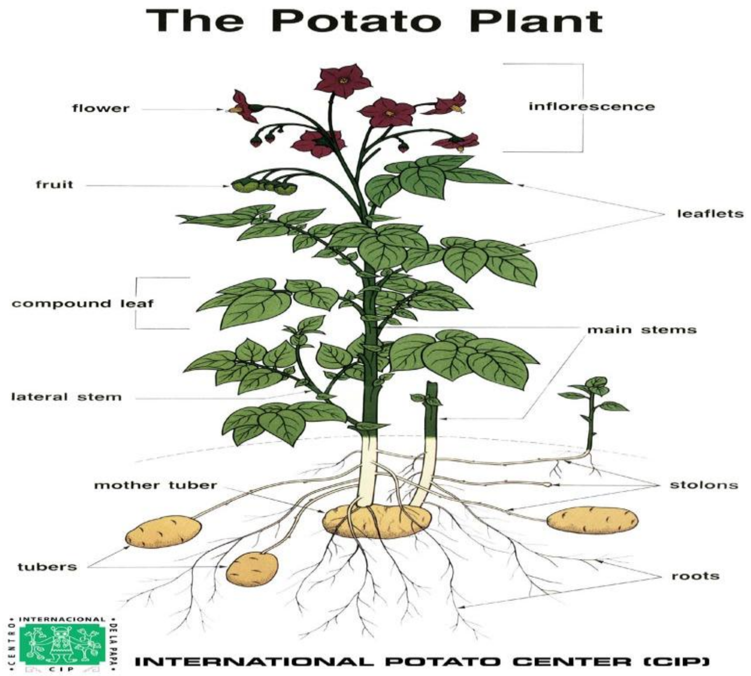


Fig. 1: Diagram showing all the parts of a potato plant (Ref: www.vomiting chiken.com)

47 potato verities have been breed in several agro-climatic regions of the India, having 28 varieties only for north Indian plains. Also, for north Indian hills and other special problem areas viz. north Bengal hills and south Indian hills Sikkim, various verities of potato have been developed. 19 out of 47 varieties possess multiplefarious resistances to different abiotic and biotic stresses. Out of these only 9 varieties are suitable for commercial purposes and processing, which are: Kufri Chipsona-1, Kufri Chipsona-2, Kufri Himsona Kufri Jyoti, Kufri Chipsona-3, Kufri Chandramukhi, Kufri Frysona, KufriLauvkar and Kufri Surya (http://cpri.ernet.in/). Currently, understanding various signals, signal transduction pathways, and particularly sugar-based signaling in potato and other plants are important research areas. All over the world, efforts are being made to produce disease-free and high-yielding potato varieties which can withstand a host of various biotic and abiotic stresses (Kooch 1996).

Nutritional aspects of potato: Potato tuber is having about 80 % is water and the rest is composed of dry matter. Parenchyma i.e. storage tissue, comprises the major part of the potato tuber. Potatoes are good dietary source rich with carbohydrate content whose predominant form is starch. Because of the resistant to enzymatic digestion in the stomach and small intestine, the portion of this starch is reaches the large intestine as in intact form and reorganize to have identical health benefits and physiological effects as fiber. It endeavor protection against colon cancer, lowers plasma cholesterol and triglyceride concentrations, improves glucose tolerance and insulin sensitivity, increases satiety, and possibly even reduces fat storage. They are also store house of vitamins and minerals such as potassium, copper, vitamin C, vitamin B, manganese, phosphate, niacin, dietary fibers and pantothenic acid. They have a exception with vitamin A and have at least about every nutrient. They are low in calories. Per 100 g of potato contains 321 KJ of energy, 17.47 g carbohydrates, 15.44 g starch, 2.2 g dietary fiber, 0.1 g fat, 2 g protein , 0.08 mg thiamine(B1), 0.03 mg Riboflavin(B2), 1.05 mg niacin(B30), 0.296 mg pantothenic acid, 0.296 mg vitamin B6, 10 ug folate B9, 19.7 mg vitamin C , 0.01 mg vitamin E, 1.9 mg vitamin K, 12 mg calcium, 0.78 mg iron, 23 mg magnesium, 0.153 mg manganese, 57 mg phosphorus, 421 mg potassium, 6 ug sodium, 0.29 mg zinc and 75g water. Potato tuber comprises of high quality dietary fiber B-group vitamins.

1.2 Tuberization: an important part of potato life cycle

Anatomy and development of the tuber: In higher plants there are two principle storage organs i.e. tubers and seed (which has cotyledonary reserves of carbohydrate, protein and lipid and endosperm). Tubers and seeds are basically different their anatomy and development. Seeds have tightly regulated developmental pathway that is initiated by fertilization and finally, ends in the programmed dehydration of the seed and whereas, tubers (somatic storage tissues) are developed by modification of somatic structures of the plant (Cutter, 1982). Out of approximately 2000 members of the genus *Solanum*, only about 160 wild and 7 cultivated species have limited ability to form tubers. Structurally, potato tubers are underground modified stems which have expanded gradually by a process of limited cell division and cell expansion (Cutter 1982). For starch and storage protein accumulation in mature tubers the biochemistry of cortical and pith cells are modified. Beneath the epidermis a periderm is formed which is plunge as the tuber develop. The tuber has indeterminate growth and may have no complex temporal or spatial patterns of gene expression involved in differentiation and starch and storage protein biosynthesis. The complex developmental process of potato in the plant is known as tuberization. Under the influence of two specific environmental conditions potato plant start tuberizing. These are short day photoperiod (SD) and cool night temperature. For the growth of potatoes such optimal conditions are only available during winter in subtropical Indo-Gangetic plains. This is the reason why the potato growing conditions in temperate countries of Europe and North America are entirely different from those in India. Tuberization is a coordinated morphophysiological process and important survival mechanism, for the tuber formation in potato plant which is highly occurring on the underground stolons (Van den Berg *et al.* 1996). This complex expanding process involves interactions between several extrinsic and intrinsic genetic, biochemical, and environmental factors. At morphological and biochemical level there are two distinct processes, associated with potato tuber formation. Stolon development and subsequent tuber induction at the stolon tip are included under morphological process. Whereas, tuber induction is characterized by both cell division and cell enlargement at cellular level (Vreugdenhil *et al.* 1999), also accompany by a change in the orientation of cell growth in the subapical region of the stolon tip (Xu *et al.* 1998b). Biochemical process includes starch synthesis (Tuberger *et al.* 2000) and the accumulations of storage protein (Taylor *et al.* 1992). Specific gene expression patterns are required to regulate both morphological and biochemical processes (Taylor *et al.* 1998; Bache metal. 2000; Verhees *et al*. 2002). Essential stages during tuberization includes, induction, initiation, and growth of an underground shoot (the stolon), and cessation of the longitudinal growth of the stolon followed by the induction, initiation and growth of a storage organ (the tuber). In leaf phytochrome B- and GA–medicated photoperiodic perception occurs followed by which RNA acts as a systemic signal in the long-distance signaling pathway to initiate tuberization in the underground stolon.

There are many environmental and hormonal factors which influence the tuber formation. Few of the important factors having profound effects are nitrogen level, temperature and light. Potatoes have many advantages with a research point of view, such as they can be easily transformed and amenable to genetic manipulation. Also they can be propagated rapidly both in tissue culture and through cuttings. Moreover, genetically they are more closely related to tomato. It has been reported that high nitrogen level inhibits tuberization; high temperature also inhibits tuberization, whereas high light and high sucrose promote tuberization. Other factors taken into consideration are that, SD promotes tuberization, many transmissible signals are involved in control of tuberization and GAs inhibit tuberization and play a role in control by photoperiod.

1.2.1 Various transcription factors involved in tuberization

Apical meristem is the origin, where a primary developmental event of plant originates (Clark, 1997; Kerstetter and Hake, 1997). Transcription factors (TFs) regulate or we can say control such types of events the molecular level. For modulating the expression of specific target gene, transcription factors TFs which are proteins in nature and act as developmental switches by binding to the DNA (or to other protein that bind to the DNA). An important family of TFs involved during regulation of developmental events in apical meristems, is *knox* (knotted-like homeo box) (Reiser *et al.,* 2000). This knotted-like homeo box (knox) gene corresponds to group of TFs known as the three-amino acid loop extension (TALE) superclass (Burglin, 1997). Because of very high level of sequence conservation in the DNA-binding region, which is designated the homeodomain, and consisting of three ὰ-helices (Kerstetter *et al.,* 1994) it becomes easier to distinguish these TFs. The third helix one is involved in DNA binding (Mann and Chan, 1996). TALE TFs contain three conserved residues of proline-tyrosine-proline located between the first and second ὰ-helix of homeodomain that has been implicated in protein interactions (Pasner *et al.,* 1999). There are two main groups of TALE homeodomain proteins in plants, KNOX and BEL types (Burglin, 1997). In specific developmental process of the shoot apical meristem (SAM), involvement of *knox* genes is supported by functional analysis and expression patterns of mutation, also it has been reported that KNOX genes are involved in regulating the level of gibberellins (GA). Gibberellins are important phytohormones involved during regulation of tuber formation. GA endogenous levels are regulated by sucrose and abscisic acid (ABA) as a result of which it is a dominant regulator in tuber formation (Xu *et al.,* 1998a). High level of GA promotes longitudinal cell expansion for stolon elongation by causing a transverse orientation of microtubules and microfibirls to the cell axis, and also prevents tuber induction (Carrera *et al.,* 1999).

Homeobox TF of potato in knox family designated as POTH1 (potato homeobox) was isolated from an early tuber cDNA library of potato. After sequence analysis it was clear that POTH1 is member of *knox* family as well as TALE superclass of homeobox, and it controls GA synthesis to regulate plant growth (Tamaoki *et al*., 1997; Kusaba *et al*., 1998b; Hay etal., 2002; Rosin *et al*., 2003a). Another gene family is BEL family (Burglin, 1997; Chan *et al*., 1998). BEL TFs have been implicated in fruits and flower development (Reiser *et al*., 1995; Dong *et al*., 2000). Subsequent studies of BEL1- like genes indicated that they are ubiquitous in the plant kingdom. Out of various BEL-like cDNAs have been identified in potato, StBEL5 is involved in tuber development by affecting hormone levels (Chen *et al*., 2003, 2004). StBEL5 is a TF which is family member of the TALE superfamily of transcription factor (Burglin 1997). To regulate numerous developmental processes BEL1-like TFs interact with KNOTTED1- type for targeting genes (Bellaoui *et al*. 2001; Muller *et al*. 2001; Smith et *al*. 2002; Smith and Hake 2003; Bhatt *et al*. 2004; Kanrar *et al*. 2006). During tuber formation in potato, various hormone levels are regulated in stolon tip by BEL1 TFs, StBEL5 and its *knox* protein partner, POTH1 (Rosin *et al.* 2003; Chen *et al*. 2003; 2004). Heterografting experiments and RNA detection method demonstrate that StBEL5 transcripts are present in phloem cells and move across a graft union to localize in stolon tips, the site of tuber induction (Benerjee *et al*. 2006). The movement of RNA to stolon tips is facilitated by a short day photoperiod, mediated by sequence tags present in the untranslated regions of the StBEL5 transcript, and correlated with enhanced tuber production (Benerjee *et al*. 2006). Thus it is noticed that mRNA of StBEL5 appears to act as a mobile signal that is delivered to the stolon tip to induce tuber formation.

In the yeast two-hybrid system the KNOTTED 1-like protein POTH1 as bait, BEL1-like proteins were identified in potato. Reduction of RNA levels for GA *20-oxidase1* in stolons and leaves, result in reduction in active GA levels and an increase in tuberization. Due to the activity of key regulatory enzymes in the GA biosynthetic pathway, the level of GA in swollen stolons decreases during tuberization(Xu *et al*., 1998a; Kloosterman *et al*., 2007). Thus, using molecular techniques we can study transcription factors involve during tuberization.

In following chapter we will be very clearly discussing about transport of different RNAs through phloem and their possible role in long distance signaling with its mechanism. Also an explanatory note is provided on tuberization, its importance and the transcription factors which regulate tuberization in potato. To make it more understandable, mechanism of TFs are also explained with evidence.

Predominant importance of phloem transport for plant growth was sighted by plant biologists by the end of 19th century. It has been suspected that phloem mediated supply of photosynthates and other organic compounds are important for plant growth. Phloem physiology becomes well established, that allows the researcher justifiably measure the translocation and higher resolution visualization of phloem tissue. Molecular biology techniques with genomics and real time microscopy regenerates phloem physiology and helps in identifying and manipulating genes encoding phloem specific proteins.

2.1 Endogenous RNA constituents of phloem, and their possible role in long distance signaling

In all living organisms diverse biological activities are coordinated, short and long distance communication networks are indispensable even in higher plants. Intra and intercellular messenger RNA transport is an important mechanism that promotes cell differentiation and plant growth (Haywood *et al*., 2002; Ding *et al*., 2003; Kim and Pai, 2009).

Several plant TF factors such as TF LEAFY and APETALA1 or the KNOTTED1- Like homeobox TF, have the qualification to move over short distance from cell to cell to plant plasmodesmata and are associated in meristem initiation and conservation (Lucas *et al*., 1995; Sessions *et al*., 2000). Recently it was proposed that mobile silencing signals from endogenous trans-acting (TA) short-interferring (si) RNA loci, might produce regional gradient of target gene expression to facilitate leaf polarization in Arabidopsis (Childwood *et al*., 2009; Dunoyer *et al*., 2010a). An astonishing observation was that, specific mRNAs can be delivered even to inaccessible plant organs through the phloem (Ruiz-Medrano *et al*., 1999), mean while many RNAs proteins and ribonucleo proteins complexes constitute in the phloem translocation stem have become fascinating candidate to serve as information transmitters (Ruiz-Medrano *et al*., 1999). Some of them have a crucial role in long distance signaling (Opark and Cruz, 2000; Lough and Lucas, 2006; Kehr, 2009; Dinant and Lemoine, 2010). Comprehensive studies focus on phloem RNA composition, which revealed different type of RNAs.

2.1.1 Import and transport mechanism of phloem- mobile RNAs

RNAs in sieve elements are thought to be imported from the adjacent companion cell through the connecting pore-plasmodesmata unit (Lough and Lucas, 2006). Some RNA-binding proteins function as chaperone and arbitrate import of RNAs from companion cells into sieve elements. RNA-specific chaperones could modify the macromolecular structure of RNAs to promote their passage through plasmodesmata and stabilize them during translocation (Yoo *et al*., 2004; Lough and Lucas, 2006; Ham *et al*., 2009). RNA stability, however, does not seem to be a primary concern within the translocation stream, since RNase activity is not encountered in sieve–tube exudates (Sasaki *et al*.,1998; Doering-Saad *et al*., 2002, Gaupels *et al*., 2008; Zhang *et al*., 2009). Depending upon the metabolic flux RNAs appear to translocate within sieve tubes from source to sink. Complex of RNP, mRNAs and several interacting proteins help in facilitating long distance transport of RNA in the phloem (Ham *et al*., 2009). Many plant biologists have discussed about protein components of long distance RNA transporting complexes, including RNA–binding proteins that facilitate trafficking from companion cells into sieve elements and large RNP complexes.

2.1.2 Full-length mRNA as a long-distance signal

In animals, mRNAs are transported interior to the cell in a tightly managed or we can say regulated process to facilitate their function. During transport and repression of translation of the mRNAs, such type of direct localization involves RNA-protein, RNA-RNA (Ferrandon *et al*., 1997) and protein-protein interactions (Elvira *et al*., 2006). Suppression of translation approves that the mobile mRNAs function only at the target sites (King *et al*., 2005). Generally, it is the untranslated regions (UTRs) that function in binding to proteins that facilitate migration of a transcript (Jansen, 2001), such migration mediate RNA stability (Derriogo *et al*., 2000; Lee and Jeong, 2006), or regulate the efficiency of translation (Gualerzi *et al*., 2003; Barreau *et al*., 2006). These cellular RNA molecules that are NA-binding proteins quickly fall prey to degradative process so they need to be protected.

Recent studies focused on phloem mRNAs of potato suggest that the tuberization signal may be a full-length mobile RNA (Banerjee *et al*., 2006). One of them is StBEL5, a member of the BEL-like family of transcription factor, which is mobile mRNAs of potato that is transported through the phloem (Muller *et al*., 2001; Chen *et al*., 2003; Smith and Hake, 2003). The studies in transgenic plant showed a positive correlation between the accumulation of StBEL5 RNA and tuber formation. Over expression of StBEL5 has consistently produced plant that exhibit overall increased vigor and enhance tuber production (Chen *et al*., 2003; Banerjee *et al*., 2006, 2009).

2.2 Role of specific transcription factors during tuber development

In the past several decade mechanism of tuberizatiozn has been the subject of considerable investigation by plant scientists. However, precise knowledge about the controlling factor involved in this growth process is not entirely clear. A tuber is a shortened thicken modified stem with leaves reduce to scale or scars transverse to the axillary bud known as eyes. Tuberization is also used to describe formation of the underground storage organ of various plants such as dahlia, Jerusalem artichoke and begonia. In potato (*Solanum tuberosum L.*), tuber formation is a complex developmental process where interactions between environmental, biochemical and genetic factors are very essential. Many important biological processes including signal transduction, carbon partitioning and meristem determination (reviewed by Ewing and Struik, 1992) are also involved. Development of potato tubers has been consider by Gregory (1965) in 3 general stages; (1) Tuber initiation, which is characterized by the development of tuber primordium and differentiation without evidence of any visible swelling; (2) Tuber enlargement, characterized by rapid visible swelling of stolon tip, cell division, also starch accumulation and (3) Tuber maturation, under this condition organ passes into the dormant period. And due to inductive conditions, under the influence of environmental factors, the transmissible signal is activated, which initiates cell division and expansion. Followed by a change is observed in the orientation of cell growth in the sub apical region of the stolon tip (Xu *et al*., 1998 a). During formation of potato tuber, all these changes are observed at morphological and biochemical level. Stolon development and subsequent tuber induction at stolon tip are involved under morphological process. Whereas, cell division and cell enlargement are observe at cellular level (Vreugdenhil *et al*., 1999). And biochemical process includes starch synthesis (Tauberge *et al*., 2000) and accumulation of storage proteins (Taylor *et al.,* 1992). These two distinct processes are regulated by specific gene expression patterns (Taylor *et al*., 1998; Bachem *et al*., 2000; Verhees *et al*., 2002).

Control over process of tuberization in potato is performed by several multiple signaling pathways such as, phytohormone signaling, RNA signaling and Calcium ions signaling (D Sarka, 2008). Phytohormone signaling involves important environmental factors such as high irradiation, low temperature and short day photoperiod which affect the potato tuberization. Gibberellins and phyB in leaves mediate the perception of these environmental factors. As a result systemic signal is produced, which is transmitted to the underground stolons to initiate tuberization (Jackson 1999). PhyB-mediated perception of gibberellic acid (GA) feedback is standardized by photoperiod responsive1 protein (PHOR1), which is a general component of GA-signaling pathways (Amador *et al*., 2001). Gibberellins and cytokinins are two most essential phytohormones involved in the regulation of potato tuber formation. GA not only just play role in controlling photoperiod, but also it is a dominant regulator in tuber formation (Xu *et al*., 1998a), also sucrose and abscisic acid (ABA) monitor its endogenous levels. GA also enhances to the cell axis by transverse orientation of microtubules and microfibrils, which is important for stolon elongation. On other hand tuber induction is prevented by high concentration of GA (Carrera *et al*., 1999). Cytokinins help in regulation of cell division, by controlling the cell cycle (Vreugdenhil 2004). Even auxin plays an important role in tuberization. Appropriate ratio of cytokinin to auxin is required for tuberization (Sergeeva *et al*., 2000). Recently it has been indicated the involvement of a family of protein, called auxin response factor6 (ARF6), a key regulator of auxin responsive gene, in tuberization (Faivre-Rampant *et al*., 2004). RNA signaling include several transcription factors, such as POTM1 (Kang and Hannapel 1995), POTH1 (Rosin *et al*., 2003), StBEL5 (Chen *et al*., 2003) and potato orthologue *Arabidopsis thaliana* CONSTANS (AtCO), independently or in concert with other regulate tuber formation in potato. Recent research firmly supported statement that "mRNA acts as a signal molecule in a long-distance signaling pathway". POTH1 and StBEL5 transcription factors of potato interact in tandem to repress the expression of *ga20oxl* (Chen *et al*., 2004), thus regulating GA synthesis and increase cytokinin level during tuberization (Rosin *et al*., 2003). During long distance signaling pathway among leaf and the stolon tip during tuber induction, StBEL5 RNA operates as a systemic signal (Banerjee *et al*., 2006). Accumulation and transport of StBEL5 RNA is promoted by SD-length and mediated by unknown RNA-binding proteins. This StBEL5 RNA is accompanied by protein chaperones to the stolon tip, where in conjugation with KNOX partner(s), it regulates transcription of various target genes involve in tuber formation. In case of calcium signaling potato needs protein dephosphorylation for transcriptional activation of some tuberizing genes (Raices *et al*., 2003). Ca^{2+} and CA-binding modulator proteins act as signal molecules for tuber induction (Jena *et al*., 1989). StCDPK1, which is Ca dependent protein kinase contains a highly conserved myristoylation site, is expressed in tuberizing stolon.

There are many transcription factors which are essential during tuberization. Out of these TALE is the important family of TFs involved in regulating the developmental events in the apical meristems is the Knox (Knotted- like homeobox) gene family (Reiser *et al*., 2000). Knox genes belong to the group of TFs known as the three amino acid loop extension (TALE) superclass (Burglin, 1997). These TFs are distinguished by a very high level of sequence conservation in the DNA- binding region, designated the homeodomain and consist of three alpha helices similar to the bacterial helix-loop-helix motifs (Kerstetter *et al*., 1994). The third helix, the recognition helix, is involved in DNA-binding (Mann and Chang, 1996). TALE TFs contain a TALE, proline- tyrosine- proline, between helices I and II in the homeodomain that has been implicated in protein interactions (Passner *et al*., 1999). These are numerous TFs from plant and animal in the TALE superclass, and the two main groups in the plants are the KNOX and BEL types (Burglin, 1997). From early tuber cDNA libraries of potato, we have isolated two groups of TFs, KNOX and BEL type that physically interact. BEL type proteins are involved in floral architecture and development in several plant species (Smith and Hake, 2003; Bhatt *et al*., 2004; Smith *et al*., 2004). An additional support for long distance movement of RNA in pumpkin was provided after the discovery of RNA-binding protein, CmPP16 (Xoconostel-Cazares *et al*., 1999). Out of the six groups of RNAs, information on the dynamics of movement is available only for StBEL5 RNA (Banerjee *et al*., 2006) which is one of the essential TF for tuberization in potato. StBEL5 is an essential member of TALE superfamily of transcription factors, (Burglin 1997). BEL1-like family of TFs is ubiquitous among plant species where they contact with KNOTTED1-types for targeting genes, as a result of which numerous developmental processes are regulated (Bellaoui *et al*., 2001; Muller *et al*., 2001; Smithet *et al*. 2002; Smith and Hake. 2003; Bhatt *et al*., 2004; Kanrar *et al*., 2006). In potato, full length mRNA of BEL1- like transcription factor, StBEL5, moves from leaf veins through the phloem to stolon tip to activate tuber formation. BEL1 transcription factor, StBEL5 and its KNOX protein partner, POTH1 regulate tuber formation by controlling hormone levels in the stolon tip (Rosin *et al*., 2003; Chen *et al*., 2003; 2004). Presence of StBEL5 transcript in phloem cells and their movement across a graft union to localize in stolon tips, to the site of tuber induction is demonstrated by using RNA detection methods and heterografting experiments (Banerjee et a., 2006). The movement of RNA is facilitated by SD photoperiod, which is mediated by sequence tags present in UTR of the StBEL5 transcript, correlated with enhanced tuber formation (Banerjee *et al*., 2006). Thus, outcome observed clarified that, the mRNA of StBEL5 appears to act as a mobile signal that transported to stolon tip to induce tuber formation.

2.2.1 Evidence for movement of BEL5 RNA through the phloem

It was being observed that light conditions help in activating the promoter of the StBEL5 gene in the veins of leaves but do not get activated in stems (Chatterjee *et al*., 2007). An abundant amount of StBEL5 RNA could be detected in stems suppoting the idea that the RNA was being transported from leaves to stems through the phloem (Banerjee *et al*., 2006). If this were the case, then BEL5 RNA would be present in cells within the stem phloem tissue. This hypothesis was confirmed by in situ hybridization of stem and stolon section with probes specific for StBEL5 and laser capture micro dissection (Banerjee *et al*., 2006). Presence of both BEL5 RNA plus POTH1 (potato homeobox 1), the KNOTTED1-type protein partner of BEL5, verified their presence in phloem cells. Moreover the presence of six other StBEL5 in phloem was confirmed by laser capture micro dissection and profiling of RNA from sieve-tube sap (Yu *et al*., 2007; Cambell *et al*., 2008). It has been suggested from the observations that the intriguing possibility that several of the BEL1 RNAs of potato participate in phloem –mediated transport to regulate development and to respond to environmental cues. Till date genes for thirteen BEL1 like TFs have been identified in potato.

Movement of StBEL5 transcript was first confirmed by grafting an overexpression line for a full-length StBEL5 construct to a wild type plant (Banerjee *et al*., 2006). In three separate plants, the StBEL5 RNA moved from the transgenic scion into the wild type stock accumulating specifically in the stolon tip. Increased in tuber production in the grafted plant and transport of transgenic StBEL5 RNA was said to be correlated (Banerjee *et al*., 2006).

2.3 Origin of the problem

Tuberization is a process of tuber development which is essential and complex durability mechanism for regulating morphological processes occurring in underground stolon tip of potato. Being a complex process it involves many important environmental, biochemical and genetic factors, at both extrinsic and intrinsic level. So, it becomes necessary to get proper knowledge on process of tuberization with regard to production of healthy tuber (large, mini and micro) in potato, where various transcription factors (TFs) play crucial roles. These TFs are developmental switches which regulate various processes within species, therefore, detailed knowledge on TFs at both biochemical level and molecular level along with their expression patterns are prerequisite to know process of tuber development. Moreover, in Indian cultivars there is no report available as of now with regard to TFs associated with tuberization process. Such types of studies are important with regards to basic and applied aspects of research in the area of potato biotechnology; it would help in improving the potato crops through transgenics. One of the objectives is to standardize a suitable and easy-to-use protocol for the isolation of DNA from potato *(Solanum tuberosum)* plants so that structure and function of the genes encoding StBEL5 TFs can be studied through PCR approach using gene-specific primers. The other objective is to know their expression patterns in different tissues of potato. A total of seven potato cultivars namely Kufri Chipsona-1 (CS-1), Kufri Chipsona-2 (CS-2), Kufri Jyoti (KJ), Kufri

Chandramukhi (KCM), Kufri Ashoka (AS), Kufri Pukhraj (PR) and Désirée (De) were chosen for molecular studies. The first six represent Indian potato cultivars and the last one is an exotic cultivar. In this study, the transcription factor StBEL5 was chosen because of its key feature of long distance transport through phloem to initiate tuber formation. It moves from leafs to stolon to regulate active tuber development by controlling various hormones. Keeping the above points in view, the following objectives were framed.

2.3.1 Objectives of the study

- Sequence analyses, multiple sequence alignment (MSA) and motif search using the available BEL5 sequences at both nucleotide and amino acid level
- Analysis of amplicon profile in the different Indian potato cultivars using their genomic DNAs and StBEL5-specific oligonucleotide primers
- RT-PCR approach for studying expression pattern of BEL5 in different potato tissues

3.1 Procurement of potato plant materials

For DNA isolation, tissue cultured potato plant were collected from the different potato cultivars such as Kufri Chipsona-1 (CS-1), Kufri Chipsona-2 (CS-2), Kufri Jyoti (KJ), Kufri Chandramukhi (KCM), Kufri Ashoka (AS), Kufri Pukhraj (PR) and Désirée (De) maintained in growth room, laboratory 4, Thapar University, Patiala. After acclimatization these plants where grown into field in the starting month of November, which is best suited for cold temperature and short day length light conditions. As potato is a cold weather crop, so optimum conditions were taken care of while cultivating above cultivars. Followed by, they were harvested after 4 months i.e. in February for further studies.

3.2 Other materials

Various enzymes used were purchased from Bangalore Genei Pvt.Ltd., Bangalore. The chemicals required were bought from Siga-aldrich India Pvt. Ltd, and Himedia Pvt. Ltd, Mumbai. Primers used were synthesized by Bangalore Genei Pvt.Ltd, Bangalore. The gel extraction Qiagen kit was purchased from Genetix.

Buffers, other chemicals and enzymes

- *Gel loading buffer (5X)*

  Sucrose - 35 % (w/v)
  EDTA - 50 mM (pH 8.0)
  Tris – 25 mM
  Bromophenol blue - 0.2 % (w/v)

- *TBE (5X) buffer*

  Tris Base - 54 g L^{-1}
  Boric acid - 28 g L^{-1}
  EDTA - 3.8 g L^{-1}
  The pH of the buffer was set at 8.0

- *TE buffer*

  Tris.HCl - 10 mM (pH 8.0)
  EDTA - 1 mM (pH 8.0)

- *Extraction buffer*
  50 mM Tris-HCL (pH 8.0)
  50 mM EDTA
  250 mM Nacl
  15 % Sucrose
- Ethanol
- Isopropanol
- 5 M potassium acetate solution
- Sodium acetate
- RNase
- Alcohol
- Chloroform

3.3 Methods

3.3.1 Isolation of genomic DNA from different potato cultivars

Procedure: Genomic DNA was isolated from the tissue cultured potato plant by the protocol described by Kumari *et al*. (2012). Plant samples were washed under tap water followed by sterile distilled water. To remove the excess water on the plant, blotting filter paper was used. Afterwards 0.7 g of plant material was weighed and fine powder was made in the presence of liquid nitrogen with the help of mortar and pestle. The fine powder was then transferred to 20 mL tube containing 5 mL of extraction buffer and 0.5 % SDS (250 µL of SDS) maintained at 65°C. Contents were mixed properly with intermittent gentle shaking and incubated at 65°C for 15 min. The solution in the tube was spun at 5500 rpm for 10 minutes using centifuge. Then 170 µL of 5.0 M potassium acetate solution was added, mixed vigorously and incubated further on ice for 20 min and solution was again centrifuged at 5500 rpm at 4°C for 15 min. Through a fine muslin cloth the supernatant was filtered and equal volume of isopropanol was added, mixed gently and incubated at -20°C for overnight. DNA was extracted by centrifugation at 12,000 rpm at 4°C for 10 min. The crude DNA pellet was washed with ice cold 70% ethanol, air dried and suspended in 50 µL of TE buffer and stored at -20°C.

3.3.2 Purification of potato genomic DNA

Materials required:

- Sterile water
- DNase-free RNase enzyme
- 3.0 M sodium acetate (pH 5.5)
- Ethanol
- TE buffer
- Extraction buffer
- 8M Licl
- DEPC-treated water
- Sodium acetate

Procedure: Further purification of DNA was done by treatment with DNase-free RNase. 400 μL of sterile water was added into eppendrof containing DNA sample. Then 3μL of DNase-free RNase enzyme was added into solution. The solution was incubated for 45 minutes at 37°C. Equal volume of phenol and chloroform (200 μL each) was added into the solution and mixed properly for 10-15 minutes by gently inversion mixing. Then the sample was centrifuged at 8000 rpm for 10 minutes. Followed by DNA was precipitated using 0.1 volume of 3.0 M sodium acetate (pH 5.5) and 2.0 volumes of ethanol. DNA was finally recovered by dissolving the pellet in 50 μL of TE buffer and stored at (-20°C). The quality and quantity of DNA was checked spectrophotometrically by measuring the A_{260}/A_{280} ratio.

3.3.3 Designing of oligonucleotide primers

The following oligonucleotide primers were designed based on the available genome sequence Corresponding to StBEL5 coding region gene in GenBank database (GenBank ID: AF406697) and primer specific StBEL5 gene in GenBank database (GenBank ID: EU200938). In case of sequence with GenBank ID: AF406697, StBEL5 coding region gene sequence of 2735 bps is comprised of the following structural features: molecular type of this sequence is mRNA means all information is provided at cDNA level. Total length of CDS region is 2067bp, starting from 148th nucleotide to 2214th nucleotide. Whereas, 5′-UTR is 147 bp long and 3′-UTR is 512 bp long starting from 2735th nucleotide to 2214th nucleotide. The whole sequence of length 2735 bp include 689 codons (including stop codon).So, the above gene sequence provides an extended complete coding sequence region. In case of sequence with GenBank ID: EU200938 include only promoter region not coding sequence of length 2219bp. Starting from 1 to 1938 it include regulatory region and mRNA joins from 1938 to 2033 (exon 1) and 2237to>2291 (exon 2). And intron length is 204 nts bp starting from 2033 to 2237.

3.3.4 Polymerase Chain Reaction (PCR)

PCR was used to amplify a specific DNA sequence in a simple, rapid and automated manner using forward and reverse primer. PCR is repeated cycling of three steps: heat denaturation of template DNA (94°C); annealing of primers to the complementary sequences in template DNA (55°C); extension of annealed primers by a thermo stable DNA polymerase (72°C).

Materials required

- Template DNA (prepared from genomic DNA of different potato cultivars)
- Primers (Resuspended with sterile TE or water)
- Buffer (usually 10X)
- Taq polymerase
- dNTPs (2mM stock)
- Sterile ddH2O

Procedure:

Table 1 Composition of reaction mixture (volume 50µL) of PCR reaction

PCR components	50µl reaction	Final concentration
DNA	Variable	<1,000ng
10X Taq reaction Buffer	5µl	1X
10µM Forward primer	1 µl	0.2 µM (0.05-1 µM)
10µM Reverse primer	1 µl	0.2 µM (0.05-1 µM)
10mM dNTPs	1 µl	200 µM
Taq polymerase	0.25 µl	1.25 units/50µl PCR
Distilled water	To make up 50 µl	

Table 2 The thermal cycling parameters of PCR

Steps	**Temperature**	**Time**
Denaturation	94°C	1 min
Annealing	55°C	2 min
Polymerization	72°C	3 min
Holding	4-10°C	Variable

Appropriate reaction setup is selected and reaction mixture is prepared by standard reagents in a tube. The master mix is prepared by aliquoting all reagents except template, there for it is added in the end. Followed by, a gentle mixing with the help of vortex is given and briefly centrifuges to collect all components to the bottom of the tube. Afterwards master mix is distributed equally to the tubes along with template to be amplified. Finally amplification parameters are set depending upon primers and template and repeated run of 30 cycles are given in a thermal cycler**.**

3.3.5 Agarose gel electrophoresis

Materials required

- Agarose (Hi media)
- 0.5 X TBE buffer
- Ethidium bromide dye (0.5 µg mL-1)
- Sterile water
- DNA samples and amplicons
- Bromophenol blue dye
- Gel Electrophoresis instrument
- UV Transilluminator
- Gel documentation system (BIO-RAD)

Procedure: Agarose gel electrophoresis was performed using standard methods (Sambrook- a laboratory manual). 1.0 % agarose gel was made in 0.5X TBE buffer and ethidium bromide dye (0.5µgmL-1) was added to it. Gel was then casted in the casting tray. The DNA samples were loaded in the wells after solidification of gel. Electrophoresis was carried out in 0.5X TBE (runningbuffer) at 2 – 5 Volt per cm till the tracking dye covered two-third of the gel length. Finally, the DNA bands were visualized under UV light.

3.3.6 Isolation of total RNA from potato tissues

Materials required

- Lithium chloride
- RNA extraction buffer
- Tris-HCL (100mM) at pH8.0
- EDTA (10mM) at pH8.0
- 1.0% SDS
- 0.2% (β-mercaptoethanol)
- RNase-free DNase
- RNase-free deionized water

Table 3 Stock solution and working solution for RNA isolation

Solution	Stock Solution	Working Solution
Tris Buffer	0.5M (pH 8.0)	100mM (pH 8.0)
LiCl	8M	100Mm
EDTA	0.5M (pH 8.0)	10mM (pH 8.0)
SDS	10%	1%
B-mercaptoethanol	0.2%	0.1ml

Procedure: Plant tissue contain high amount of polysaccharides, phenolics, nucleases and other storage material. Therefore, isolation of RNA from plant material in terms of intactness and quality is relatively difficult. For, that number of methods is reported in literature. Here we used, SDS-Phenol method described by Gilman (1987) which was used as such or with some modifications depending upon the plant material. The plant materials (0.2 to 1.0 g) were frozen and pulverized in the liquid nitrogen to a fine powder. The content were mixed in a buffer containing lithium chloride and SDS (RNA extraction buffer, 100mM tris-HCL pH8.0, 10mM EDTA pH8.0, 1.0% SDS, 0.2%

(β-mercaptoethanol) followed by direct extraction with phenol:chloroform (1:1), under ice-cold conditions, 8.0 LiCl (one third volume of the aqueous solution) was added to the supernatant and incubated for minimum 2 hours for selective precipitation of RNA. The crude RNA was further purified by RNase-free DNase treatment followed by solvent extraction and ethanol precipitation. After that RNA was dissolved in RNase-free deionized water, and kept in aliquots at -70°C for further use. The quality of the RNA samples were checked by regular and formaldehyde agarose gel electrophoresis along with RT-PCR using different potato gene specific primer.The spectrophotometric analysis using A_{260}/A_{280} ratio of the RNA samples were also measured to check the quality.

3.3.7 Reverse Transcription -Polymerase Chain Reaction (RT-PCR)

Materials required

- DEPC water
- oligo dT primer
- 10 mM DNTP
- RNase inhibitor
- Reaction buffer
- RNA sample

Procedure: First strand cDNA was synthesized using Revert Aid H Minus M-MuLV reverse transcriptase. The enzyme lacks ribonuclease H activity specific to RNA in RNA-DNA hybrids. Therefore degradation of RNA does not occur during first strand cDNA synthesis, resulting in higher yields of full-length cDNA from long templates upto 13 kb. Reverse transcription (RT) was performed using the RevertAid^{TM} H Minus first strand cDNA synthesis kit from Fermentas Life Sciences containing M-MuLV reverse transcriptase and the gene-specific reverse primer, SB5R1-2229. For each RT reaction, approx. 2.0 μg of total RNA either from tuber or leaf from Indian potato cultivars was used as template. All the steps of reverse transcription were carried out according to the manufacturer's instructions. In order to isolate full length cDNA PCR was acrried out using the individual RT product as template. 6 µl DEPC water was taken and RNA sample was added. Followed by oligo dT primer was added and mixed properly and eppendorf were kept at 65 ºC for 5 min and then at room temperature for 5 minutes. Reaction buffer (4 µl), RNase inhibitor (1 µl) were added and 10 mM DNTP was added. The tubes were incubated at 42 ºC for I hour. After initial denaturation at 94ºC for 1 min 30 s, the thermal cycling parameters during PCR were: denaturation at 94ºC for 1 min, annealing at 55ºC for 2 min; polymerization at 72ºC for 3 min for 30 cycles followed by final extension at 72ºC for 5 min.

3.3.8 Sequence analysis

NCBI tool was used to analyze the nucleotide sequence of. Open reading frame finder (ORF) available at national centre of biotechnology information using website (http//:www.ncbi.nlm.nih.gov) is used to predict the amino acid sequence. On other hand, Prot Param tool of ExPASy (expert protien analysis system) proteomics server of the swiss institute of bioinformatics (SIB;URL http//expasy.org/tools/) was used to calculate theoretical molecular weight, isoelectric point(pI) and amino acid composition of predicted protein. Prediction of hypophobic characters (Kype and Doolittle, 1982) and various secondary structures such as α helix, β helix, β turns and random coils are predicted using Prot Scale tool of ExPASy. For multiple sequence alignment, both the clustal omega (an EMBL-EBI sequence analysis tool: http:///www.ebi.ac.uk/tools) is used.

4.1 Sequence analysis, comparison and motif search in different cultivars of potato

4.1.1 Salient sequence features of *StBEL5* gene from potato

In this study, we followed 2735-bp *StBEL5* gene from potato (GenBankID: AF406697) which encodes StBEL5 protein consisting of 688 amino acids. This is basically a transcription factor involved in the process of tuber development in potato. This 2735-bp *StBEL5* gene consisting of 1 to147-bp 5′flanking region (including promoter) and from 148 to 2214 coding sequence and therefore the remaining portion i.e. from 2215 to 2735 bp refers to 3′ flanking region of this gene. Size of the ORF in *StBEL5* gene is 2067 bp (148-2214) which is equal to 689 codons including stop codon. Therefore, the StBEL5 polypeptide consists of 688 amino acids. The length of 5′-UTR is 147 bp, and 3′-UTR is 521 bp.

4.1.2 Sequence analysis of the different regions of StBEL5 by BLASTn

5′-flanking region as a query sequence: The nucleotide sequence of *StBEL5* gene was analyzed by NCBI BLAST tools. When BLASTn was done for 5′ flanking region of *StBEL5* gene, it shows largest promoter part in all *potato* gene available in database as query cover is very less. So, it shows divergence in distal part of promoter region of 5′ flanking region of *StBEL5* gene. The 147-bp 5′ flanking region of *StBEL5* gene was found to be close (100 % sequence identity) to a gene copy in

Solanum *tuberosum* BEL1-related homeotic protein5 (BEL5) mRNA (GenBank ID: NM_001287992). However sequence divergence was prominent if compared with some known *StBEL5* genes from potato, since the sequence identity values were 95% (GenBank ID: EU686384), 95% (GenBank ID: EU686380) whereas 91% (GenBank ID: XM_006361029. Query coverage ranged from 100% to 97% in the BLAST search data for these sequences.

Coding region as a query sequence: When BLASTn was done for coding region of *StBEL5* gene, it shows more homology with other gene sequences of BEL-like genes. The 2067-bp coding region was found to be identical (100% sequence identity with 100% query coverage) to a gene copy in *Solanum tuberosum* BEL1-related homeotic protein5 (BEL5) mRNA (GenBank ID: NM_001287992), 95% sequence identity with 100 % query coverage to a gene copy in *Solanum etuberosum* BEL5 protein mRNA (GenBank ID:EU686384), 95 % sequence identity with 100% query coverage to a gene copy in *Solanum paluste* BEL5 prtein (GenBank ID: EU686380), also 93 % sequence identity with 100% query coverage to a gene copy in *Solanum lycopersicum* protein 2 GenBank ID: NM_001316353).

4.1.3 Salient biochemical attributes of StBEL5 protein (GenBank Protein ID: ADN39428)

The entire 688 amino acids sequence of StBEL5 protein was analysed by using protparam tool of ExPASy resource portal under Swiss Institute of Bioinformatics (SIB) which revealed, some of the important biochemical attributes. The calculated molecular weight of the StBEL5 was found to be 76459.1 kDa with a predicted isoelectric point (pI) of 6.97. Predicted formula of StBEL5 protein is $C_{3294}H_{5210}N_{982}O_{1064}S_{27}$. Out of its total 688 amino acids, 59 were strongly basic (+) (Arg + Lys), 61 were strongly acidic (-) (Asp + Glu). The instability index (II) was computed as 52.12, which classified the protein as unstable. The amino acid composition data revealed that some of the amino acids such as Asn (8.3 %), Ile (4.4%), Lys (4.7%), Phe (3.1%), Ser (10.5%), Tyr (2.2%) and Val (4.4%) occurred more frequently as compared to their average occurance; whereas the amino acids namely Arg (3.9%), cys (0.4%), Glu (4.8%), Gly (6.4%), Leu (8.7%), Met (3.5%), Pro (4.2%) and Thr (6.1%) occurred less frequently (Doolittle 1989). The estimated half life of StBEL5 protein is 30 h (mammalian reticulocytes, in vitro), >20 h (yeast, in vivo), >10 h (Escherichia coli, in vivo) as predicted in this analysis. And Grand average of hydropathicity (GRAVY) is 0.689.

Hydropathy plot of StBEL5: The hydropathy profile was generated for StBEL5 protein using the ProtScale tool based on the Kyte- Doolittle scale

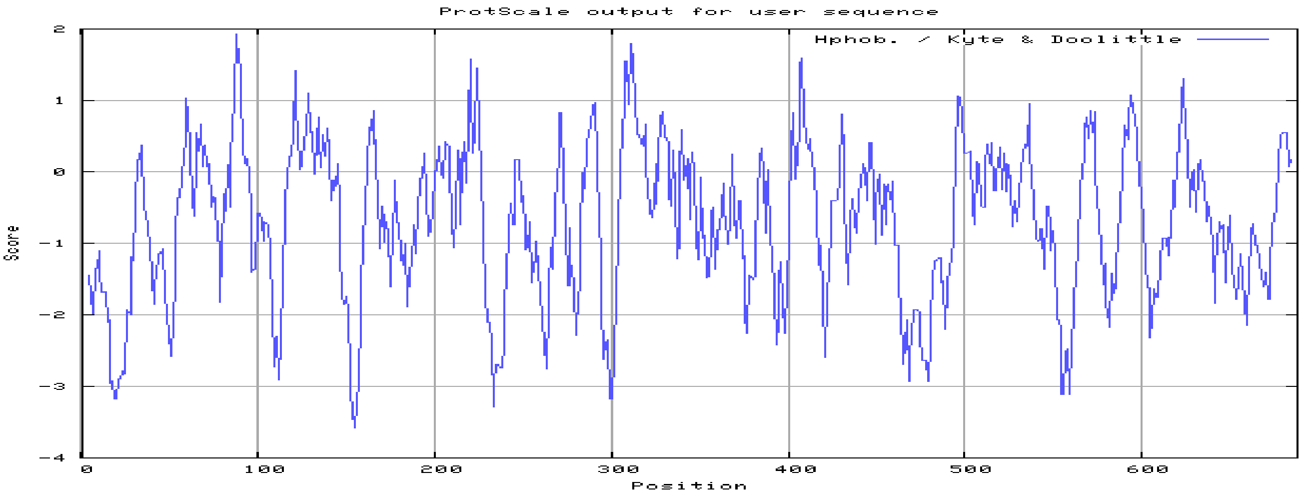


Fig. 2 Hydropathy plot of StBEL5 protein

Hydropathy plot of StBEL5 protein revealed that N terminal of amino acid sequences of StBEL5 protein is Hydrophobic i.e. at 32-34, 58-62, 65-73, 85-94, 118-131, 162-167, 200-203, 219-227, 246-248, 286-292, 305-320, 133-141, 327-332, 405-415, 495-503, 532-537, 566-573, 589-597, 620-626, 677- 684 position of amino acid sequence. Hydrophilic regions in the plots are at 5-31, 35-57, 74-84, 95-117, 142-161, 168-199, 228-245,1 249-269, 273-285, 293-304, 354-388, 390-401, 416-441, 449-494, 526-531, 538-565, 574-588, 589-619 and 634-676 positions of amino acid in StBEL5 protein.

4.1.4 BLASTp analysis of StBEL5

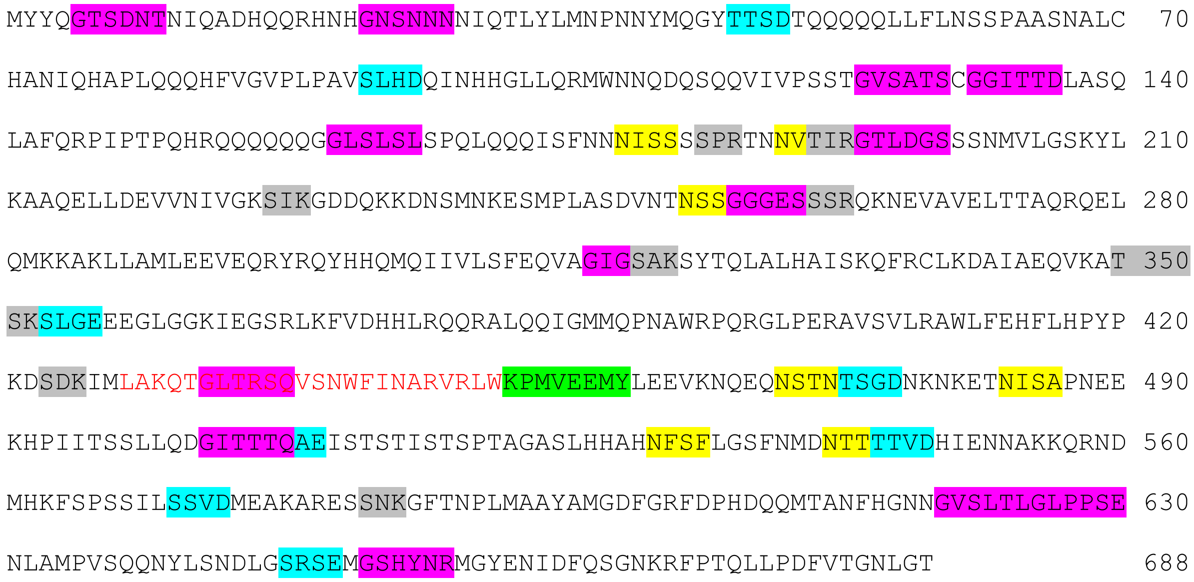
When BLASTp was done for StBEL5 protein (GenBank protein ID: AAN03621); it shows more homology with other forms of StBEL5 proteins. The 688-amino acid sequence was found to be identical (100% sequence identity with 100% query coverage) to amino acid sequence in BEL-1 like *solanum tuberosum* (GenBank protein ID: NP_001274921), 96 % sequence identity with 100 % query coverage to the amino acid sequence in homeotic protein 5 of *solanum tuberosum* (GenBank protein ID:XP_006361091), 94 % sequence identity with 100% query coverage to the amino acid sequence in *solanum palustre* gene (GenBank protein ID: ACD39464), 91 % sequence identity with 100% query coverage to the amino acid sequence in *solanum lycopersicum* (GenBank protein ID: NP_001234599). 93% sequence identity with 100% query coverage to the amino acid sequence in *solanum etuberosum* (GenBank protein ID: ACD39468). 89% sequence identity with 100% query coverage to the amino acid sequence in *solanum pennellii* (GenBank protein ID:XP_015079638). Details of some homologus BEL5 sequences are provided in Table 4.

Table 4: Details of some homologous BEL5 sequences as available in the database

Protein name	Accession No.	ORF length	Amino acids
Solanum tuberosum 1	NP_001274921	2067	688
Solanum tuberosum 2	XP_006361091	2055	684
Solanum palustre	ACD39464	2097	698
Solanum lycopersicum	NP_001234599	2100	699
Solanum etuberosum	ACD39468	2097	698
Solanum pennellii	XP_015079638	2106	701
Nicotiana syylvestris	XP_009773899	127	645

4.1.5 Significant protein motifs in StBEL5 (GenBank Protein ID: ADN39428)

The predicted amino acid sequence of StBEL5 was examined for searching the presence of motifs using the following program: (http://myhits.isbsib.ch/cgibin/motif_scan). A number of motifs were predicted as shown in Fig. 3; and the details are provided in Table 5.



```
MYYQGTSDNTNIQADHQQRHNHGNSNNNNIQTLYLMNPNNYMQGYTTSDTQQQQLLFLNSSPAASNALC   70
HANIQHAPLQQQHFVGVPLPAVSLHDQINHHGLLQRMWNNQDQSQQVIVPSSTGVSATSCGGITTDLASQ 140
LAFQRPIPTPQHRQQQQQQGGLSLSLSPQLQQQISFNNNISSSSPRTNNVTIRGTLDGSSSNMVLGSKYL 210
KAAQELLDEVVNIVGKSIKGDDQKKDNSMNKESMPLASDVNTNSSGGGESSSRQKNEVAVELTTAQRQEL 280
QMKKAKLLAMLEEVEQRYRQYHHQMQIIVLSFEQVAGIGSAKSYTQLALHAISKQFRCLKDAIAEQVKAT 350
SKSLGEEEGLGGKIEGSRLKFVDHHLRQQRALQQIGMMQPNAWRPQRGLPERAVSVLRAWLFEHFLHPYP 420
KDSDKIMLAKQTGLTRSQVSNWFINARVRLWKPMVEEMYLEEVKNQEQNSTNTSGDNKNKETNISAPNEE 490
KHPIITSSLLQDGITTTQAEISTSTISTSPTAGASLHHAHNFSFLGSFNMDNTTTTVDHIENNAKKQRND 560
MHKFSPSSILSSVDMEAKARESSNKGFTNPLMAAYAMGDFGRFDPHDQQMTANFHGNNGVSLTLGLPPSE 630
NLAMPVSQQNYLSNDLGSRSEMGSHYNRMGYENIDFQSGNKRFPTQLLPDFVTGNLGT             688
```

Fig. 3 Prediction of protein motifs in StBEL5

Table 5: Representing motif sites and the position of the motifs

Sr No.	Motif site	Motif	Motif position
1	N-glycosylation site	NISS	179-182
2	N-glycosylation site	NVTI	189-192
3	N-glycosylation site	NSSG	253-256
Ú	N-glycosylation site	NSTN	469-472
5	N-glycosylation site	NISA	483-486
6	N-glycosylation site	NFSF	531-534
7	N-glycosylation site	NTTT	542-545
8	Casein kinase II phosphorylation site.	TTST	46-49
9	Casein kinase II phosphorylation site	SLHD	93-96
10	Casein kinase IIphosphorylation site	SLGE	353-356

11	Casein kinase II phosphorylation site	TSGD	473-476
12	Casein kinase II phosphorylation site	TQAE	507-510
13	Casein kinase II phosphorylation site	TTVD	545-548
14	Casein kinase II phosphorylation site	SSVD	571-574
15	Casein kinase II phosphorylation site	SRSE	648-651
16	N-myristoylation site	GTSDNT	5-10
17	N-myristoylation site	GNSNNN	23-28
18	N-myristoylation site	GVSATS	124-129
19	N-myristoylation site	GGITTD	131-136
20	N-myristoylation site	GLSLSL	161-166
21	N-myristoylation site	GTLDGS	194-199
22	N-myristoylation site	GGGESS	256-261
23	N-myristoylation site	GIGSAK	317-322
24	N-myristoylation site	GLTRSQ	433-434
25	N-myristoylation site	GITTTQ	503-508
26	N-myristoylation site	GVSLTL	619-624
27	N-myristoylation site	GLPPSE	625-630
28	N-myristoylation site	GSHYNR	653-658
29	Protein kinase C phosphorylation site	SPR	184-186
30	Protein kinase C phosphorylation site	TIR	191-193
31	Protein kinase C phosphorylation site	SIK	227-229
32	Protein kinase C phosphorylation site	SSR	261-263
33	Protein kinase C phosphorylation site	SAK	320-322
34	Protein kinase C phosphorylation site	TSK	350-352
35	Protein kinase C phosphorylation site	SDK	423-425
36	Protein kinase C phosphorylation site	SNK	583-585
37	Tyrosine kinase phosphorylation site	KPMVEEMY	452-459
38	Homeobox domain_1	LAKQTGLTRSQV SNWFINARVRLW	428-451

In the above StBEL5 sequence mainly 38 motifs were found. Seven motifs (NISS, NVTI, NSSG, NSTN, NISA, NFSF and NTTT) were found on N- glycosylation site which are highlighted with yellow colour. From Casein kinase II phosphorylation site, eight motifs (TTST, SLHD, SLGE, TSGD, TQAE, SRSE, TTVD and SSVD) were found which are highlighted by dark green colour. Thirteen N-myristoylation sites (GTSDNT, GNSNNN, GVSATS, GGITTD, GLSLSL, GTLDGS, GGGESS, GIGSAK, GLTRSQ, GITTTQ, GVSLTL, GLPPSE and GSHYNR) were found which are pink in coloured. From Protein kinase C phosphorylation site eight motifs (SPR, TIR, SIK, SSR, SAK, TSK and SDK) were found which are highlighted by grey colour. There was also one motif (KPMVEEMY) found at tyrosin kinase phosphorylation site which was coloured green. Also one motif (LAKQTGLTRSQVSNWFINARVRLW) was found for homeobox domain which was highlighted with red colour words. Other than this one motif for HOMEOBOX_1, one motif for Asparagine-rich region, one motif for bacterial Ig-like domain 1, one motif for Glutamine-rich region and one motif for HOMEOBOX_2 were also found.

4.1.6 Multiple sequence alignment (MSA) of different StBEL5 from potato (*Solanum tuberosum*)

Multiple sequence alignment was performed for total of seven BEL5 sequence from potato (*Solanum tuberosum*) using GenBank ID: AF406697, i.e., BEL1-related homeobox protien5 (*Solanum tuberosum*) (1) with acc. No.NP_001274921, BEL1-related homeotic protein 5 isoform X1 (*Solanum tuberosum*) (2) with acc no.XP_006361091, BEL5 protein (*Solanum palustre*) with acc no.ACD39464, bell-like homeodomain protein 2 (*Solanum lycopersicum*) with acc no.NP_001234599, BEL5 protein (*Solanum etuberosum*) with acc no.ACD39468, BEL1-likehomeodomain protein 7 (*Solanum pennellii*) with acc no.XP_015079638, BEL-1like homeodomain protein 1 (*Nicotiana syylvestris*) with acc no. XP_009773899. Minor manual adjustments were also made during alignment. This study was done to examine sequence similarities and divergence, the amino acid substitutions, insertions and deletions between the seven different forms of StBEL5. Details about conserved sequences in selected sequences are given in Fig. 4.

```
NS      MYYQGTSDNIHQ-------------AADGNIQTLYLMNPNYIHQGYSNDTQQQQ-HQHQH 46
ST1     MYYQGTSDNTNIQADHQQRHNHGNS-NNNNIQTLYLMNPNNYMQGYTTSDTQQ------- 52
ST2     MYYQGTSDN-NIQADHQQHHNHGNSSNNNNIQTLYLMNPNNYMQGYTTSDTQQ------- 52
SP      MYYQGTSDN-NIQADHQ--RHHNHGNSNNNIQTLYLMNPNSYMQGYTTTDTQQ--HQNQN 55
SE      MYYQGTSDN-NIQADHQ--QHHNHGNSNNNIQTLYLMNPNSYMQGYTTTDTQQ--HQNQN 55
SL      MYYQGTSDN-NIQADHHQQQHNNLGNSNNNIQTLYLMNPNSYMQGYTTTDT--QQHLQQQ 57
SPE     MYYQGTSDN-NIQADHHQ-QHNNHGNSNHNIQTLYLMNPNSYMQRYTTTDTQRQQHLQQQ 58
        ********* .                : ***********   * *:.


NS      H-QQQ-----------------IIRHAPLQQAQQHFVGVPLPAVSLHDQ-NHAQLIHRLW 87
ST1     --QQQLLFLNSSPAASNALCHANIQHAPL—-QQQHFVGVPLPAVSLHDQINHHGLLQRMW 108
ST2     --QQQLLFLNSSPAGSNALCHANIQHAPL--QQQHFVGVPLPAVSLHDQINHHGLLQRMW 108
SP      QHQQQLLFLNSSPAGSNALGHANIPHAPL--QQQHFVGVPLPAVSLHDQINHHGLLQRMW 113
SE      QHQQQLLFLNSSPAGSNALGHANIPHAPL--QQQHFVGVPLPAVSLHDQINHHGLLQRMW 113
SL      QNQHQLLFLNSAPAGGNALSHANIQHAPL--QQQHFVGVPLPAVSLHDQINHHGLLQRMW 115
SPE     QNQQQLLFLNSAPAGGNALSHANIQHAPL--QQQHFVGVPLPAVSLHDQINHHGLSQRMW 116
          *:*                  * ****   ***************** **  * :*:*

NS      NQSGGGGGNGSHHQSQIIPSSTVVSATSCGGSTTDLASQLGFQRPIIMSPTQQ-----QQ 142
ST1     NNQD-------QSQQVIVPSSTGVSATSCGGITTDLASQLAFQRPI---PTPQHRQQ-Q 158
ST2     NNQD-------QSQQVIVPSSTGVSATSCGGITTDLASQLAFQRPI---PTPQHRQQQQ 158
SP      NNQD-------QSQQVIVPSSTVVSATSCGGTTTDLASQLAFQRPIVVSPTPQHRQQ-QQ 165
SE      NNQD-------QSQQVIVPSSTVVSATSCGGTTTDLASQLAFQRPIVVSPTPQHRQQ-QQ 165
SL      NNQD-------QSQQVIVPSSTVVSATSCGGTTTDLASQLAFQRPIVVSPTPQHRQQ-QQ 167
SPE     NNQD-------QSQQVIVPSSTVVSATSCGGTTTDLASQLAFQRPIVVSPTPQHRQQ--Q 167
        *:.        : *. *:**** ******** ********.*****   ** *      *

NS      QQGGLSLSLSPQQQ-QISFNNISAQNQV---GSNINRGLD-GISSMVLGSKYLKAAQELL 197
ST1     QQGGLSLSLSPQLQQQISFNNNISSSSPRTNNVTIRGTLDGSSSNMVLGSKYLKAAQELL 218
ST2     QQGGLSLSLSPQQQQQISFNNNISSSSPRTNNVTIRGTLDGSSSNMVLGSKYLKAAQELL 218
SP      QQGGLSLSLSPQQQQQISFNNNISSSSPRTNNVTIRGTLDGSSSNMVLGSKYLKAAQELL 225
SE      QQGGLSLSLSPQQQQQISFNNNISSSSPRTNNVTIRGTLDGSSSNMVLGSKYLKAAQELL 225
SL      QQGGLSLSLSPQQQQQISFNNNISSSSPRTNNVTIRGTMDGCSSNMILGSKYLKAAQELL 227
SPE     QQGGLSLSLSPQQQQQISFNNNISSSSPRTNNVTIRGAMYGSSSNMILGSKYLKAAQELL 227
        ************ * ******  :...      .*.  :    *.*:*************

NS      DEVVN-VGKCIKGSELEKGELMNKESILPLASDVNTSNSG----GESSSKQKNG-VAELT 251
ST1     DEVVNIVGKSIKGDDQKKDNSMNKE-SMPLASDVNTNSSG---GGESSSRQKNEVAVELT 272
ST2     DEVVNIVGKSIKGDDQKKDNS------MPLASDVNTNSSGGGGGGESSSRQKNEVAIELT 272
SP      DEVVNIVGKSIKGDDQKKENSMNKE-SMPLASDVNTNSSGG--GGESSSRQKNEVAVELT 282
SE      DEVVNIVGKSIKGDDQKKENSMNKE-SMPLASDVNTNSSGG--GGESSSRQKNEVAVELT 282
SL      DEVVNIVGKSNKGDDQKKDNSMNKE-LIPLVSDVNTNSSGGG-GGESSSRQKNEVAIELT 285
SPE     DEVVNIVGKSNKGDDQKKDNSMNKE-LMPLVSDVNTNSSGGG-GGESSSRQKNEVAVELT 285
        ***** ***. **.: :* :       :**.*****..**    *****:***  . ***

NS      TAQRQEFQMKKAKLVTMLEEVEQRYRQYHHQMQIIVSSFEQVAGIGSAKSYTQLALHAIS 311
ST1     TAQRQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVLSFEQVAGIGSAKSYTQLALHAIS 332
ST2     TAQRQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVSSFEQVAGIGSAKSYTQLALHAIS 332
SP      TAQRQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVSSFEQVAGIGSAKSYAQLALHAIS 342
SE      TAQRQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVSSFEQVAGIGSAKSYTQLALHAIS 342
SL      TAQRQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVSSFEQVAGVGSAKSYTQLALHAIS 345
SPE     TAQRQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVSSFEQVAGVGSAKSYTQLALHAIS 345
        ******:*******::******************** *******:******:********
```

```
NS            KQFRCLKDAIVGQIKATSKSLGEEDQSLGEKIEGSRLKFVDHHLRQQRALQQLGMMQPNA 371
ST1           KQFRCLKDAIAEQVKATSKSLGEE-EGLGGKIEGSRLKFVDHHLRQQRALQQIGMMQPNA 392
ST2           KQFRCLKDAIAEQVKATSKSLGEE-EGLVGKIEGSRLKFVDHHLRQQRALQQLGMMQPNA 391
SP            KQFRCLKDAIAEQVKATSKSLGEE-EGLGGKIEGSRLKFVDNHLRQQRALQQLGMMQPNA 401
SE            KQFRCLKDAIAEQVKATSKSLGEE-EGLGGKIEGSRLKFVDNHLRQQRALQQLGMMQPNA 401
SL            KQFRCLKDAISEQVKATSKSLGED-EGLGGKIEGSRLKFVDHHLRQQRALQQLGMMQPNA 404
SPE           KQFRCLKDAISEQVKATSKSLGED-EGLGGKIEGSRLKFVDHHLRQQRALQQLGMMQPNA 404
              **********  *:*********: :.*  ***********.**********:*******


NS            WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRLWK 431
ST1           WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRLWK 452
ST2           WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRLWK 451
SP            WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRLWK 461
SE            WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRLWK 461
SL            WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRLWK 464
SPE           WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRLWK 464
              ************************************************************

NS            PMVEEMYLEEMKNNQEQNSSSEENKNKNES------NKEK--MVS-KSSAPQESGTPTEI 482
ST1           PMVEEMYLEEVKNQEQNSTNTS-GDNKNKETNISAPNEEKHPIITSSLLQDGITTTQAEI 511
ST2           PMVEEMYLEEVKNQEQNSTNTS-GDNKNKE----APNEEKHPIITSSLLQDGITTTQAEI 506
SP            PMVEEMYLEEVKNQEQNS-TTS-GDNKNKETNISAPNEEKQPIITSSLLQDGTTTTQAEI 519
SE            PMVEEMYLEEVKNQEQNS-TTS-GDNKNKETNISAPNEEKQPIITSSLLQDGTTTTQAEI 519
SL            PMVEEMYLEEVKNQEQNSSNTS-GDNKNKETNISAPNEEKQPIITSSLLQDGTT--QAEI 521
SPE           PMVEEMYLEEVKNQEQNSSNTS-GDNKNKETNISAPNEEKQPIITSSLLQDGNT--QAEI 521
              **********:**::::. .:.  .***:.      *:**  ::: .      :   :**

NS            ATSTISTPPTACGSSFQAAQAHNFSFLGSLNMENNTT----------TKKPRNHDMHNNS 532
ST1           STSTISTSPTAGAS---LHHAHNFSFLGSFNMDNTT--TTVDHIENNAKKQRN-DMHKFS 565
ST2           STSTISTSPTAGAS---LHHAHNFSFLGSFNMDNTT-TTTVDHIENNAKKQRN-DMHKFS 561
SP            STSTISTSPTAGAS---LHHAHNFSFLGSFNMKNTTTTNTVDHIENNAKKPRN-DMQKFS 575
SE            STSTISTSPTAGAS---LHHAHNFSFLGSFNMENTTTTNTVDHIENNAKKPRN-DMQKFS 575
SL            STSTISTSPTAGAS---LHHAHNFSFLGSFNMENT--TTTVDHIENNAKKPRNHDMHKFS 574
SPE           STSTISTSPTAGAS---LHHAHNFSFLGSFNMENTATTTTVDHIENNAKKPRNHDMHKFS 578
              :****** *** .*     :*********:**.*.            :** ** **:: *

NS            PSNILSSVDMAETKARDHQSKKGFNPLMDAAAVYTMGEFGRFNPHEQMAATNFHGNNGVS 592
ST1           PSSILSSVDMEAK-ARESSNKGFTNPLM---AAYAMGDFGRFDPHDQQMTANFHGNNGVS 621
ST2           PSSILSSVDMEAK-ARESSNKGFTNPLM---AAYAMGDFGRFDPHDQQITANFHGNNGVS 617
SP            PSSILSLVDMEAK-ARESSNKGFTNPLM---AAYAMGDFGRFDPHDQQMTANFHGNNGVS 631
SE            PSSILSLVDMEAK-ARESSNKGFTNPLM---AAYAMGDFGRFDPHDQQMTANFHVNNGVS 631
SL            PSSILSSVEMEAK-ARESTNKGFTNPLM---AAYAMGDFGRFDPHDQQMTANFHGNNGVS 632
SPE           PSSILSSVEMEAK-ARESTNKGFTNPLM---AAYAMGDFGRFDPHDQQMTANFHGNNGVS 634
              **.*** *:*  . **:  .*   ****   *.*:**:****:**:*  ::*** *****

NS            LTLGLPPNSENLSQQSYQ------------NTIEFNRMSYENIDFQ—NKRFPSQILPDF 638
ST1           LTLGLPPSENLAMPVSQQNYLSNDLGSRSEMGSHYNRMGYENIDFQSGNKRFPTQLLPDF 681
ST2           LTLGLPPSENLAMPVSQQNYLSNDLGSRPEMGSHYNRMGYENIDFQSGNKRFPTQLLPDF 677
SP            LTLGLPPSENLAMPVNQQNYLSNELGSRPEMGSHYNRMGYENIEFQSGNKRFPTQLLPDF 691
SE            LTLGLPPSENLAMPVNQQNYLSNELGSRPEMGSHYNRMGYENIEFQSGNKRFPTRLLPDF 691
SL            LTLGLPPSENLAMPVSQQNYLSNELGSRPEIGSHYNRMGYENIDFQSGNKRFPTQLLPDF 692
SPE           LTLGLPPSENLVMPVSQQNYLSNELGSRPEIGSHYNRMGYENIDFQSGNKRFPTQLLPDF 694
              *******..:     . *               .:***.****:**  *****:::****
```

```
NS        VAGNLGT 645
ST1       VTGNLGT 688
ST2       VTGNLGT 684
SP        VTGNLGT 698
SE        VTGNLGT 698
SL        VTGNLGT 699
SPE       VTGNLGT 701
          *:*****
```

Fig. 4 Multiple sequence alignment of the BEL5 sequences from potato and other plants. Sequence alignment was based on Clustal Omega tool along with some minor manual adjustments. Dashes indicate gaps that arise during alignment. Asterisks (*) indicate the identical amino acids between all the BEL5 sequences presented here. The symbol bullet '•' denotes identical amino acid residues between the BEL family members along with semi conservative substitutions in the other members. The symbol double dot ':' denotes conservative substitution. where NS stands for *Nicotiana syylvestris*, SL for *Solanum lycopersicum*, SPe for *Solanum pennellii*, SE for *Solanum etuberosum*, ST for *Solanum* tuberosum (1) and St for *Solanum* tuberosum (2).

Above information is generated by performing multiple sequence alignment using clustal omega (MSA Tool), a new program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more protein sequences. Here, multiple sequence alignment of the deduced amino acid sequences of seven different forms of StBEL5 from the potato (*Solanum* tuberosum) cultivars (GenBank protein ID: AAN03621) was done. The star symbols denote the conserved amino acid between all the seven forms of StBEL5 proteins studied. Alignment of BEL5 protein showed some divergence with respect to the other six forms of BEL5 proteins which are highlighted with yellow colour.

After segment wise comparison between the different forms of BEL5 revealed the following observations. In the N-terminal of the all the StBEL5 forms the amino acids were nearly conserved with minor variations. In total there are 96 variant regions, including 25 semi conserved mutation highlighted with pink color, 39 conserved mutations highlighted with yellow color and only 14 non conserved mutations highlighted by green colour.

4.2 Analysis of amplicon profile in various potato cultivars using genomic DNA and *StBEL5* gene-specific primers

4.2.1 Isolation of genomic DNA from the potato cultivars

Total genomic DNA was isolated from potato cultivars namely Kufri Chipsona-1 (CS-1), Kufri Chipsona-2 (CS-2), Kufri Jyoti (KJ), Kufri Chandramukhi (KCM), Kufri Ashoka (AS), Kufri Pukhraj (PR) and Désirée (De) from the different organs of potato, collected from the field grown plant (see section 3 of

‘Materials & Methods’). The quality of genomic DNA was checked by agarose gel electrophoresis as shown in following Fig 5. And quantification data is given in Table 6.

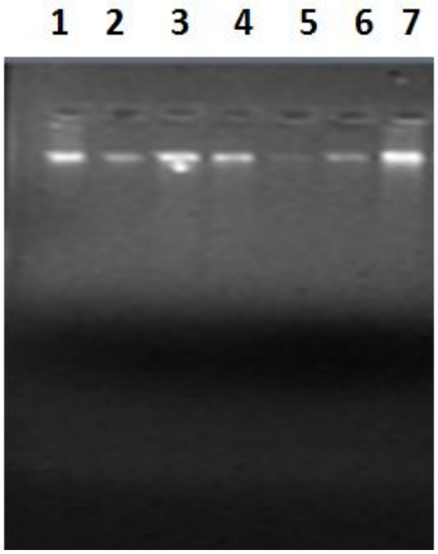


Fig. 5: Genomic DNA isolation from the potato cultivars. Agarose gel electrophoresis of genomic DNA preparation from different potato cultivars are indicated lane wise: Lane 1-CSI; Lane 2-CS2, Lane 3-KCM, Lane 4- KJ, Lane 5- AS, Lane 6-De, Lane 7-PR

Above amplicon profile shows good amount of isolated DNA from different potato cultivars. Here we can see CS1, KCM, KJ and PR indicate prominent intensity of bands, whereas, band observed in CS2, AS and De are less intense.

The quality and quantity of DNA was checked spectrophotometrically by measuring the A_{260}/A_{280} ratio respectively.

Table 6: Quantification of potato DNA using nanodrop spectrophotometer

Potato Cultivar	**Amount of Material (g)**	**Volume of DNA soln. (µL)**	**Absorb. Ratio (A_{260}/A_{280})**	**Conc. (ng/µL)**
CS1	0.7	50	1.12	690
CS2	0.7	50	1.13	880
KCM	0.7	50	1.18	900
AS	0.7	50	1.02	680
De	0.7	50	1.15	460
PR	0.7	50	1.03	740
KJ	0.7	50	1.07	780

4.2.2 Designing of oligonucleotide primers

- Oligonucleotide primers should be 10-24 nucleotides long.
- GC content should be 40-60%.
- The primer should not be self-complementary or complementary to any other primer to form primer-dimer or hair pin.
- Melting temperatures of primer pairs should not differ by more than 5ºC, so the GC Content and length must be chosen accordingly.
- The annealing temperature should be about 5°C lower than the melting temperature.
- Long run sequences of a single nucleotide should be avoided.
- Primers with significant structures are avoided.

Keeping all these parameters in mind following primers were designed using cDNA sequences of StBEL5 present in database:

- 5′-ATATAGATCAGTCTGACAAG- 3′ (SB5F1-0016; corresponding to bases 16-35 of BEL5, Acc.No. AF406697)
- 5′-TGCAGATATGTACTATCAAG-3′ (SB5F2-0041; corresponding to bases 41-60 of BEL5, Acc.No. AF406697)
- 5′-GAGACTTTCTGGTATTCATG-3′ (SB5R1-2229; complementary to the bases 2210-2229 of potato BEL5, Acc.No. AF406697)
- 5′-TGCTAATCTAATAATGATAG-3′ (SB5R2-2718; complementary to the bases 2699-2718 of potato BEL5, Acc.No. AF406697)

Details of the *StBEL5* gene-specific primers: SB5F2-0041: This primer sequence encompasses the coding region only and which appears to be conserved in a number of StBEL5 genes as reported in the database. SB5R1-2229: This sequence also appears to be conserved with many other StBEL5 gene sequences as reported from *Solanum tuberosum*.

4.2.3 PCR using *StBEL5* gene-specific primer pairs

One of the major objectives of the study is to amplify StBEL5 gene(s) preferably with the upstream regions from a few potato (*Solanum tuberosum*) accessions through PCR approach under different conditions using specific primer pairs. For this purpose, genomic DNA preparations from the following potato (*Solanum tuberosum*) accessions namely Kufri Chipsona-1 (CS-1), Kufri Chipsona-2 (CS-2), Kufri Jyoti (KJ), Kufri Chandramukhi (KCM), Kufri Ashoka (AS), Kufri Pukhraj (PR) and Désirée (De) were used during PCR. For each specific primer pair, PCR was carried out under annealing temperature 55°C. Individual primer pair-specific amplicon profiles are shown in the following sections.

Results of PCR: PCR amplification results using different potato genomic DNAs and the primer pair SB5F2-0041**,** SB5R1- 2229 are shown in Fig. 6 (Table 7). Amplification occurred only in the cases of 3 potato cultivars namely CS-2, CS1- and KJ where the size of the amplicon was around 2 kb in each case. The result suggest that *StBEL5* gene is conserved in these cultivars, however, there might be minor sequence divergence. It is likely that the above primers did not work for the remaining cultivars because of considerable sequence divergence which needs to be understood. An attempt was also made to amplify other primer pair, SB5F1-0016 with SB5R2-2718 but it did not work possibly due to significant sequence divergence and nucleotide mismatches.

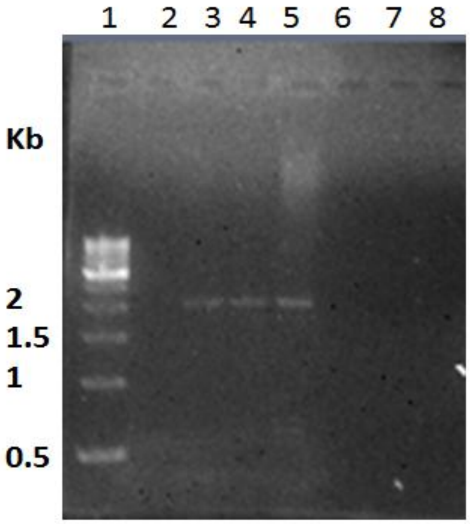


Fig. 6: PCR result using primer pair SB5F2-0041 and SB5R1-2229. PCR amplified products using gene specific primer pairs (SB5F2-0041**,** SB5R1- 2229), annealing temp. 55°C and DNA template from different potato cultivars are indicated as, *Lane* 1- 500 bp ladde*r; Lane* 2- PCR amplified bands of KCM; *Lane* 3- PCR amplified band of CS2; *Lane* 4- PCR amplified band of CS1; *Lane* 5- PCR amplified band of KJ; *Lane* 6- PCR amplified band of PR. *Lane* 7- PCR amplified band of AS; *Lane* 8- PCR amplified band of De.

Table 7: SB5F2-0041**,** SB5R1- 2229 primer pair-specific amplicons profile

Potato Accessions	**Size of amplicons (annealing temp. 55°C)**
KCM	-
CS2	~2.0 kb
CS1	~2.0 kb
KJ	~2.0 kb
PR	-
AS	-
De	-

4.3 Preliminary expression studies in different potato tissues

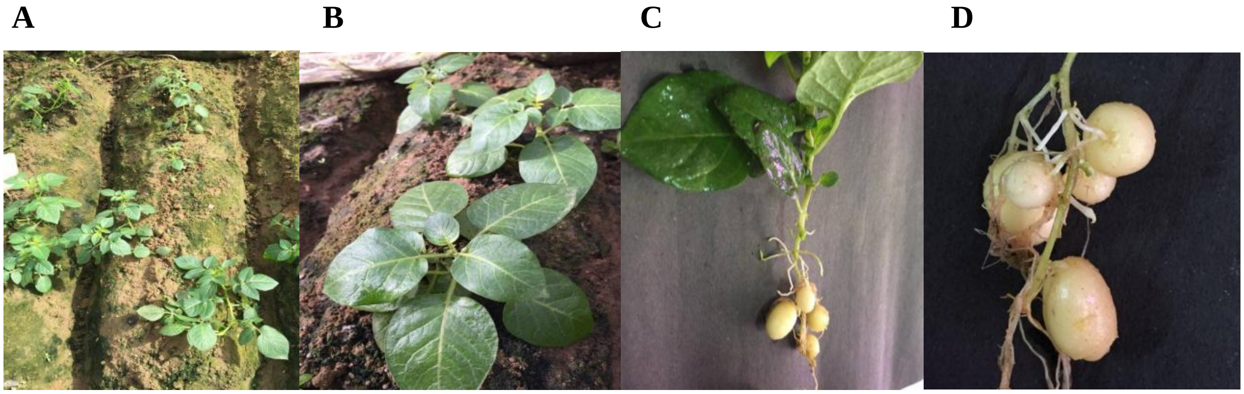


Fig. 7 A**:** potato plant grown in field, B: leaf structure of the potato plant, C: after harvesting, full plant body, D: small tubers, non tuberizing and tuberizing stolons

For RNA isolation, various tissues of potato such as leaf, smaall tuber, Non tuberization stolon and tuberizing stolon were harvested (as shown in Fig. 7) from the different cultivars of potato mainted in our Thapar University. After harvesting, Each tissue was collected seperately, manually and kept in a plastic bag under ambient conditions. Followed by, weighing their weight, tissues were separately pulverized in liquid nitrogen.

4.3.1 Isolation of RNA from potato

Total RNA was isolated from the various tissues i.e. leaf, small tuber, non tuberizing stolon and tuberising stolon samples collected from the field grown potato cultivar, Kufri Chipsona-1 (CS-1) (see section 3 of 'Materials & Methods'). The quality of RNA was checked by agarose gel electrophoresis as shown in Fig. 7. The quality and quantity of RNA was also checked spectrophotometrically by measuring the A_{260}/A_{280} ratio (Table 8).

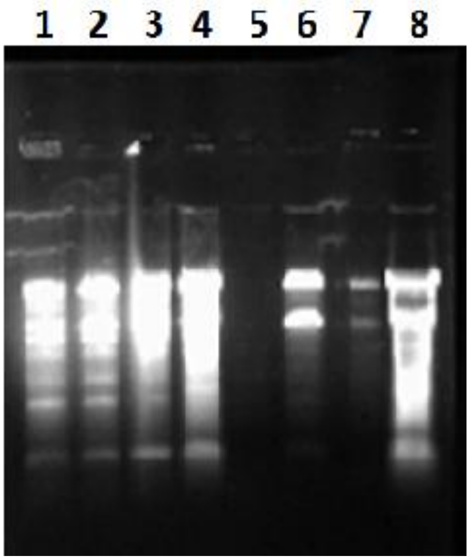


Fig. 8: Isolation of RNA using various potato tissues. RNA isolation from various potato tissues, lane wise data is given as: Lane 1- Leaf; Lane 2 tuberizing stolon, Lane 3 non tuberizing stolon, Lane 4-small tuber, Lane 5- Leaf ; Lane 6 tuberizing stolon , Lane 7 non tuberizing stolon , Lane 8 small tuber

Table 8 Quantification of RNA using nanodrop spectrophotometer

Potato Tissue	Amount of Material (g)	Volume of RNA soln. (µL)	Conc. (µg/µL)
Non tuberizing Stolon	0.7	50	580µg
Tuberizing stolon	0.7	50	690 µg
Small tuber	0.7	50	610 µg
Leaf	0.7	50	440 µg

4.3.2 Purification of RNA

Purified RNA samples from different potato tissues were checked by agarose gel electrophoresis as shown in Fig. 9. Distinct ribosomal RNA bands are seen in the gel indicating the intactness of total RNA prepration.

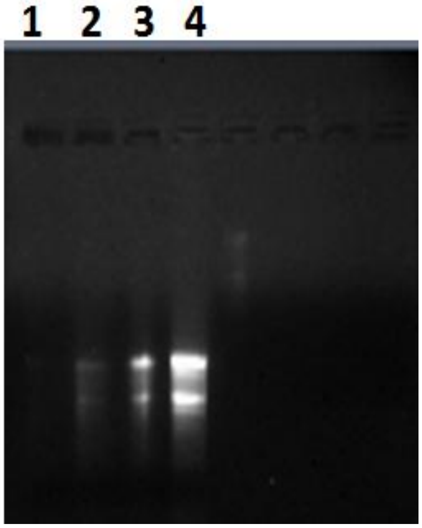


Fig. 9: Purified RNA from various potato tissues. In Fig 8 Lane1- Non tuberizing stolon, lane 2-Tuberizing stolon, lane 3-Small tuber, lane 4- Leaf

4.3.3 RT PCR using specific primer set of StBEL5 in different potato tissues

*Results of PCR***:** RT-PCR was carried out using total RNA from different potato tissues and the primer pair SB5F2-0041**,** SB5R1-2229. The RT-PCR data clearly indicated that amplification occurred only in the cases of 3 potato tissues namely tuberizing stolon, small tuber and leaf (Fig. 10). In tuberizing stolon and small tuber, the size of the amplicon was around 2 kb which was in the expected size range. Interestingly, in case of leaf, 3 amplicons were noticed out of which one common i.e. 2 kb as is found in other tissues; the remaining amplicons were of ~1.75 kb and ~0.75 kb. It is likely that more than one allelic variant of *StBEL5* genes are expressed in the leaf tissue which needs to be understood. Possibly more than one allelic variants of *StBEL5* are functional in some potato tissues. Apart from the role in tuberization the BEL5 transcription factors might have some other biological role in potato.

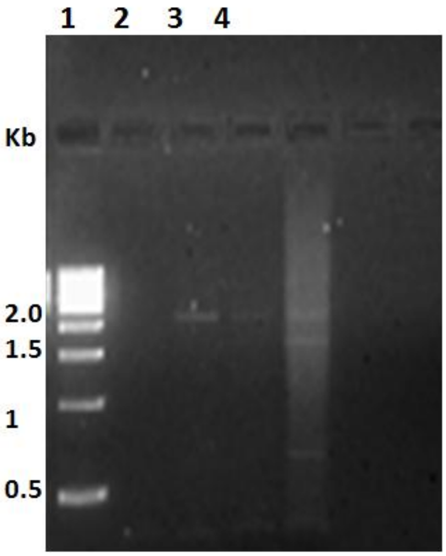


Fig. 10: RT PCR product using oligodT. RT PCR products using OligodT primer, *Lane* 1- 500 bp ladder*; Lane* 2- Non tuberising stolon ; *Lane* 3- Tuberising stolon ; *Lane* 4- Small tuber; *Lane* 5- Leaf

Table 9 RT PCR using SB5F2-0041 and SB5R1- 2229 primer pair

Potato tissues	**Size of RNA**
Non tuberising stolon	-
Tuberising stolon	~2.0 kb
Small tuber	~2.0 kb
Leaf	~2.0 kb
Leaf	~1.75kb
Leaf	~0.75kb

CONCLUSION

- StBEL5 belongs to TALE superfamily, which is considered to be a long distance travelling transcript. It is transferred from leaves to stolon to control various developmental processes to initiate tuberization. Literature survey revealed that there was no research report on these transcription factors with regards to Indian potato cultivars.
- Although there are some BEL5 sequence data available in the database at both nucleotide and amino acid level, but the salient sequence features and comparison were not highlighted in the previous research reports. Keeping this in mind, efforts are made on sequence analysis and comparison using the available BEL5 sequence at both nucleotide and amino acid level. This exercise provided a clear and comprehensive idea about the sequence relatedness between the BEL5 sequences.
- Now we know the salient sequence features, conserved amino acids and some protein motifs in the BEL5 sequences. Such type of information is not only helpful in understanding the evolutionary consequences but also important with regard to gene manipulation and protein engineering.
- In this study, the good quality genomic DNA was isolated from different commercially important Indian potato cultivars, and both quality and quantity were assessed by spectrophotometric analysis.
- The different *StBEL5* gene-specific oligonucleotide primers were used in PCR to know the allelic variation between the different cultivars. The cultivar-specific amplicons could be further studied in detail at molecular level.
- Total RNA was isolated and purified from different tissues of potato under field conditions.
- Efforts were made to study the expression patterns of StBEL5 in different tissues of potato using RT-PCR approach. Expression was noticed in non-tuberizing stolon, tuberizing stolons, small tubers and leaves.
- In conclusion this thesis work made a consolidated base for further studing BEL5 in the Indian potato cultivars.

REFERENCES

1. Amador V , Monte E , García-Martínez, JL, & Prat S (2001) Gibberellins signal nuclear import of PHOR1, a photoperiod-responsive protein with homology to Drosophila armadillo. *Cell*, *106*(3), 343-354
2. Bachem CWB, van der Hoeven, Lucker J, Oomen R, Casarini E, Jacobsen E, Visser RGF (2000) Functional genomic analysis of potato tuber life-cycle. Potato Res 43:297–312
3. Banerjee AK, Chatterjee M, Yu Y, Suh S-G, Miller WA, Hannapel DJ(2006) Dynamics of a mobile RNA of potato involved in a longdistancesignaling pathway. Plant Cell 18:3443–3457
4. Bellaoui M, Pidkowich MS, Samach A, Kushalappa K, Kohalmi SE, ModrusanZ, Crosby WL, Haughn GW (2001) The Arabidopsis BELL1 andKNOX TALE homeodomain proteins interact through a domain conserved between plants and animals. Plant Cell 13: 2455–2470
5. Bhatt AM, Etchells JP, Canales C, Lagodienko A, Dickinson H (2004)VAAMANA—a BEL1-like homeodomain protein, interacts with KNOX proteins BP and STM and regulates inXorescence stemgrowth in *Arabidopsis*. Gene 328:103–111
6. Burglin TR (1997) Analysis of TALE superclass homeobox genes (MEIS,PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved betweenplants and animals. Nucleic Acids Res 25: 4173–4180
7. Campbell BA, Hallengren J, & Hannapel, D. J. (2008). Accumulation of BEL1-like transcripts in solanaceous species. *Planta*, *228*(6), 897-906.
8. Chan RL, Gago GM, Palena CM, Gonzalez DH (1998) Homeoboxes in plantdevelopment. Biochim Biophys Acta 1442: 1–19
9. Chen H, Banerjee AK, Hannapel DJ (2004) The tandem complex of BEL and KNOX partners is required for transcriptional repressionof *ga20ox1*. Plant J 38:276–284
10. Chen H, Rosin FM, Prat S, Hannapel DJ (2003) Interacting transcription factors from the TALE superclass regulate tuber formation.Plant Physiol 132:1391–1404
11. Chen H, Rosin FM, Prat S, & Hannapel D J. (2003). Interacting transcription factors from the three-amino acid loop extension superclass regulate tuber formation. *Plant Physiology*, *132*(3), 1391-1404.
12. Clark SE (1997) Organ formation at the vegetative shoot meristem. Plant Cell 9: 1067–1076
13. Different organs of the potato plant (Solanum tuberosum L.). Plant Mol Biol 20:641–651
14. Dinant S, & Lemoine R. (2010). The phloem pathway: new issues and old debates. *Comptes rendus biologies*, *333*(4), 307-319
15. Ding B, Itaya A, Qi Y. 2003. Symplasmic protein and RNA traffic: regulatory points and regulatory factors. Curr. Opin. Plant Biol. 6:596–602

16. Doering-Saad C, Newbury HJ, Bale JS, Pritchard J. (2002). Use of aphid stylectomy and RT-PCR for the detection of transporter mRNAs in sieve elements. J. Exp. Bot. 53:631–37
17. Dong YH, Yao JL, Atkinson RG, Putterill JJ, Morris BA, Gardner RC (2000)MDH1: an apple homeobox gene belonging to the BEL1 family. Plant MolBiol 42: 623–633
18. DUNoyer P, Brosnan CA, Schott G, Wang Y, Jay F, Alioua A, Himber C, Voinnet O 2010a. An endogenous, systemic RNAi pathway in plants. EMBO J 29: 1699–1712
19. Ferrandon D, Koch I, Westhof E, & Nüsslein–Volhard, C. (1997). RNA–RNA interaction is required for the formation of specific bicoid mRNA 3′ UTR–STAUFEN ribonucleoprotein particles. *The EMBO journal*, *16*(7), 1751-1758
20. Hay A, Kaur H, Phillips A, Hedden P, Hake S, Tsiantis M (2002) TheGibberellin pathway mediates KNOTTED1-type homeobox function inplants with different body plans. Curr Biol 12: 1557–1565
21. Haywood V, Kragler F, Lucas WJ. (2002). Plasmodesmata: Pathways for protein and ribonucleoprotein signaling. Plant Cell 14:S303–25
22. Jackson SD (1999) Multiple signaling pathways control tuber induction in potato. Plant Physiol 119:1–8
23. Janssen R-P 2001. mRNA localization: message on the move. Nat. Rev. Mol. Cell Biol. 2:247–56
24. Kanrar S, Onguka O, Smith HM (2006) *Arabidopsis* inXorescencearchitecture requires the activities of KNOX-BELL homeodomainheterodimers. Planta 224:2263–2273
25. Kehr J. (2009). Long-distance transport of macromolecules through the phloem. *F1000 Biol Rep*, *1*, 31
26. Kerstetter R, Vollbrecht E, Lowe B, Veit B, Yamaguchi J, Hake S (1994)Sequence analysis and expression patterns divide Sthe maize *knotted1*-likehomeobox genes into two classes. Plant Cell 6: 1877–1887
27. Kerstetter RA, Hake S (1997) Shoot meristem formation in vegetativedevelopment. Plant Cell 9: 1001–1010Reiser L, Sanchez-Baracaldo P, Hake S (2000) Knots in the family tree:evolutionary relationships and functions of knox homeobox genes. PlantMol Biol 42: 151–166
28. Kim I, & Pai HS. (2009). Mobile macromolecules in plant development.*Journal of Plant Biology*, *52*(3), 186-192
29. Kloosterman B, Vorst O, Hall RD, Visser RGF, Bachem CW (2005) Tuber on a chip: differential gene expression during potato tuberdevelopment. Plant Biotechnol J 3:505–519
30. Kusaba S, Kano-Murakami Y, Matsuoka M, Tamaoki M, Sakamoto T,Yamaguchi I, Fukumoto M (1998b) Alteration of hormone levels in transgenic tobacco plants overexpressing the rice homeobox gene *OSH1*.Plant Physiol 116: 471–476
31. Lough TJ, & Lucas WJ. (2006). Integrative plant biology: role of phloem long-distance macromolecular trafficking. *Annu. Rev. Plant Biol, 57*, 203-232

32. Lucas WJ. 1995. Plasmodesmata: intercellular channels for macromolecular transport in plants. Curr. Opin. Cell Biol. 7:673–80
33. Mann RS, Chan SK (1996) Extra specificity from *extradenticle:* the partnershipbetween HOX and PBX/EXD homeodomain proteins. Trends Genet12: 258–262
34. Muller J, Wang Y, Franzen R, Santi L, Salamini F, Rohde W (2001) In vitro interactions between barley TALE homeodomain proteins suggest a rolefor protein-protein associations in the regulation of Knox gene function.Plant J 27: 13–23
35. Oparka KJ, Santa Cruz S. 2000. The great escape: phloem transport and unloading of macromolecules. Annu. Rev. Plant Physiol. Plant Mol. Biol. 51:323–47
36. Passner M, Ryoo HD, Shen L, Mann RS, Aggarwal AK (1999) Structure ofDNA-bound Ultrabithorax-Extradenticle homeodomain complex. Nature397: 714–719
37. Reiser L, Modrusan Z, Margossian L, Samach A, Ohad N, Haughn GW,Fischer RL (1995) The BELL1 gene encodes a homeodomain proteininvolved in pattern formation in the *Arabidopsis* ovule primordium. Cell 83: 735–742
38. Rosin FM, Hart JK, Horner HT, Davies PJ, Hannapel DJ (2003a) Overexpressionof a *knox* gene of potato alters vegetative development bydecreasing gibberellin accumulation. Plant Physiol 132: 106–117
39. Ruiz-Medrano R, Xoconostle-Cazares B, Lucas WJ. 1999. Phloem long-distance trans- ´ port of CmNACP mRNA: implications for supracellular regulation in plants. Development 126:4405–19
40. Sarkar D. (2008). The signal transduction pathways controlling in planta tuberization in potato:an emerging synthesis. *Plant Cell Reports*, *27*(1), 1-8
41. Sasaki T, Chino M, Hayashi H, Fujiwara T. 1998. Detection of several mRNA species in rice phloem sap. Plant Cell Physiol. 39:895–97
42. Sessions A, Yanofsky MF, Weigel D. 2000. Cell-cell signaling and movement by the floral transcription factors LEAFY and APETALA1. Science 289:779–82
43. Smith HM, Boschke I, Hake S (2002) Selective interaction of plant homeodomain proteins mediates high DNA-binding affinity. Proc Natl AcadSci USA 99: 9579–9584
44. Smith HMS, Hake S (2003) The interaction of two homeobox genes,*BREVIPEDICELLUS* and *PENNYWISE*, regulates internode patterningin the *Arabidopsis* in Xorescence. Plant Cell 15:1717–1727
45. Tamaoki M, Kusaba S, Kano-Murakami Y, Matsuoka M (1997) Ectopicexpression of a tobacco homeobox gene, *NTH15*, dramatically alters leafmorphology and hormone levels in transgenic tobacco. Plant Cell Physiol38: 917–927

46. Tauberger E, Fernie AR, Emmermann M, Renz A, Kossmann J, Willmitzer L, Trethewey RN (2000) Antisense inhibition of plastidial phosphoglucomutase provides compelling evidence that potato tuber amyloplasts import carbon from the cytosol in the form of glucose-6-phosphate. Plant J 23:43–53
47. Taylor MA, George LA, Ross HA, Davies HV (1998) cDNA cloning and characterization of an *a*-glucosidase gene from potato (Solanum tuberosum L.). Plant J 13:419–425
48. Taylor MA, Mad Arif SA, Kumar A, Davies HV, Scobie LA, Pearce SR, Flawell AJ (1992) Expression and sequence analysis of cDNAs induced during the early stages of tuberization in
49. van den Berg JH, Ewing EE, Plaisted RL, McMurry S, Bonierable MW (1996) QTL analysis of potato tuberization. Theor Appl Genet 93:307–316
50. Verhees J, van der Krol AR, Vreugdenhil D, van der Plas LHW (2002) Characterization of gene expression during potato tuber development in individuals and populations using the luciferase reporter system. Plant Mol Biol 50:653–665
51. Vreugdenhil D, Xu X, Jung CS, van Lammeren AAM, Ewing EE (1999) Initial anatomical changes associated with tuber formation on single-node potato (Solanum tuberosum L.) cuttings: are-evaluation. Ann Bot 84:675–680
52. Xu X, Vreugdenhil D, van Lammeren AAM (1998b) Cell division and cell enlargement during potato tuber formation. J Exp Bot 49:573–582
53. Yoo BC, Kragler F, Varkonyi-Gasic E, Haywood V, Archer-Evans S, et al. 2004. A systemic small RNA signaling system in plants. Plant Cell 16:1979–2000
54. Zhang S, Sun L, & Kragler F. (2009). The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation. *Plant physiology*, *150*(1), 378-387