REGULATION OF SEED DEVELOPMENT IN *LEGUMINOSAE*: INVESTIGATING THE ROLE OF SNF1-RELATED PROTEIN KINASE

Dissertation

Zur Erlangung des akademischen Grades
Doktor rerum naturalium (Dr. rer. nat.)

vorgelegt der
Mathematisch-Naturwissenschaftlich-Technischen Fakultät
der Martin-Luther-Universität Halle-Wittenberg

von Ruslana Radchuk

Geb. am: 22.04.1972 in Werba, Ukraine

Gutachter:
1. Prof. Dr. Ulrich Wobus, Institut für Planzengenetik und Kulturpflanzenforschung, Corrensstraße 3, 06466 Gatersleben
2. Prof. Dr. Karin Breunig, Fachbereich Biologie, Institut für Genetik, Weinbergweg 10, 06120 Halle

Halle (Saale), 10 Oktober 2005

urn:nbn:de:gbv:3-000009608
[http://nbn-resolving.de/urn/resolver.pl?urn=nbn%3Ade%3Agbv%3A3-000009608]

CONTENTS

List of abbreviations

A.tumefaciens	Agrobacterium tumefaciens
ABA	abscisic acid
ADP	adenosine diphosphate
Amp^r	ampicillin resistance
AMP	adenosine monophosphate
AMPK	AMP-activated protein kinase kinase
ATP	adenosine triphosphate
bp	base pairs
BAP	6-benzylaminopurine
BSA	bovine serum albumin
cDNA	complementary DNA
CIAP	calf intestial alkaline phosphatase
2,4D	2,4-dichlorophenoxyacetic acid
DAF	days after flowering
DAP	days after pollination
DEPC	Diethyl pyrocarbonate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetaacetic acid
EGTA	ethyleneglycol tetraacetic acid
EST	expressed sequence tag
g	gram
G	guanine
GA_3	gibberellic acid
GUS	ß-glucuronidase
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2ethansulfonic acid]
IPTG	isopropyl-β-D-thiogalactopyranoside
IBA	indole-3-butyric acid
k	kilo
Km^r	kanamicin resistance
L	litre
M	molarity
m	milli
min	minute

µ	micro
M-MLV RT	moloney murine leukemia reverse transcriptase
MOPS	N-morpholinopropanesulfonic acid
mRNA	messenger RNA
MS	Murashige and Scoog basal medium
n	nano
NAA	naphthalene acetic acid
ocs	octopine synthase
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonylfluoride
PNK	polynucleotide kinase
PPT	phosphinothricin
RNA	ribonucleic acid
rpm	rotation per minute
RT	room temperature
RT-PCR	reverse transcription-PCR
SAMS	synthetic peptide (HMRSAMSGLHLVKRR) based on the sequence around the primary phosphorylation site for AMPK
SDS	sodium dodecyl sulfate
SD	standard deviation
SEM	standard error of the mean
SNF1	sucrose-non-fermenting protein kinase
SnRK1	sucrose-non-fermenting related protein kinase
Suc	sucrose
T	thymidine
T-DNA	transferred DNA
Tris	Tris-hydroxymethylaminomethane
UV	ultraviolet
WT	wild type

1. INTRODUCTION

1.1 Embryonic stages of the plant life cycle

Seed formation is an important phase of the life cycle of higher plants. A new plant formed by sexual reproduction starts as an embryo within the developing seed, which arises from the ovule. The fertilized egg cell undergoes an unequal cell division producing two daughter cells one of which develops into a highly specialized suspensor tissue, whereas the other cell forms the embryo proper. The first division of the zygote initiates the development of the embryo.
Embryogenesis is a complex process that requires regulation of cell-specific and housekeeping genes within the embryo proper and neighboring seed tissues (e.g. endosperm) surrounding the embryo (Goldberg et al., 1994). The initiation of seed development have been studied extensively, and many regulators of these processes have been identified in *Arabidopsis* (McElver et al., 2001; Hennig et al., 2004). The regulation of these genes must be tightly coordinated and controlled in a spatial and temporal manner starting from the zygote to the mature embryo in dormant seeds.
Pea seed development has already been described in many studies (Smith, 1973; Hedley and Ambrose, 1980). Embryogenesis can be roughly divided into three main phases. In the first phase, the cell division phase, the cotyledon cells actively divide. In the second phase, maturation, the cotyledon cells expand, and reserve compounds (starch and proteins) are stored. The third phase concerns seed desiccation. At the end of the initial phase of development, the number of cells in the cotyledons is established (Smith, 1973). The transition between early and late seed development is accompanied by large-scale changes in gene expression patterns and metabolic pathways.

1.2 Hormonal and sugar signaling in seed development

ABA (abscisic acid) is an important regulator of seed development and maturation. It was shown that ABA accumulates rapidly during plant seed development or under such stress conditions as water deficiency, salt and cold treatments and further induces the expression of related genes to improve stress tolerance (Machuka et al., 1999). Studies by means of genetic and molecular technologies resulted in the identification of ABA-regulated genes and

corresponding transcription factors necessary for ABA-related signal transduction (Grill et al., 1998). Six genes necessary for the ABA-related signal transduction have been well characterized and their encoded proteins were classified as two transcription factors (VP1 of maize and ABI3 of *Arabidopsis*) (McCarty et al., 1991; Giraudat *et al*., 1992), two members of highly conserved protein phosphatase 2C family (Leung et al., 1997), one transcription regulator harboring an APETALA2 domain (ABI4) (Finkelstein et al., 1998) and one farnesyl transferase (ERA1) of *Arabidopsis* (Cutler et al., 1996). However, the complexity of ABA signaling network made it relatively difficult to isolate and analyze ABA-regulated genes (Genoud and Metraux, 1999). Functional mechanisms of ABA action, especially the related signal transduction pathways in plant cells, are not sufficiently known.

Sugars possess important hormone-like functions as primary messengers in signal transduction. In plants, sugar production through photosynthesis is a vital process, and the sugar status modulates and coordinates internal regulators and environmental cues that govern growth and development (Smeekens, 2000). Although the regulatory effect of sugars on photosynthetic activity and plant metabolism has long been recognized, the concept of sugars as central signaling molecules is relatively new. Glucose repression down-regulates the synthesis and stability of mRNAs coding for chlorophyll a/b-binding proteins, enzymes acting in starch degradation, Calvin and glyoxylate cycles. At the same time, glucose signalling induces the expression of genes encoding storage and defence proteins, as well as enzymes involved in glycolysis, nitrate assimilation, phosphate mobilization and anthocyanin biosynthesis (Smeekens and Rook, 1997). In a cross-talk with glucose signalling, cytokinins alleviate glucose repression of the photosynthetic genes and synergistically activate the expression of glucose-induced genes.

Biochemical, molecular and genetic experiments have supported a central role of sugars in the control of plant metabolism, growth and development and have revealed interactions that integrate light, stress, and hormone signaling (Roitsch, 1999; Finkelstein and Gibson, 2002) and coordinate carbon and nitrogen metabolism (Stitt and Krapp, 1999; Coruzzi and Zhou, 2001).

The coordination of signalling pathways responding to hormonal, metabolic and environmental stimuli has a central role in plant growth control. Diverse roles of Snf1-related protein kinases (SnRKs) in carbon metabolism and sugar signaling are also emerging (Halford and Hardie, 1998).

1.3 SnRK1 is the key regulator of metabolic pathways

The SnRK1(sucrose-non-fermenting-1-related protein kinase) is a plant protein kinase with a catalytic domain similar to that of SNF1 protein kinase in yeast *Sacchromyces cerevisiae* and the AMP-activated protein kinases (AMPK) in mammals.

The SNF1 family belongs to the CDPK-SnRK superfamily of serine-threonine protein kinase which include calcium-dependent protein kinases (CDPKs), CDPK-related kinases (CRKs), phosphoenolpyruvate carboxylase kinases (PPCKs), PPCK-related kinases (PEPRKs), calmodulin-dependent protein kinases (CaMKs), calcium and calmodulin-dependent protein kinases (CCaMKs) and SnRKs (Hrabak et al., 2003; Hardie, 1999, 2000). All of these kinases have a common structure, which contains in the C terminus a catalytic domain typical of eucaryotic Ser-Thr kinases. The N-terminal domains are highly variable in length and sequences among the subgroups and even between individual protein kinases.

The physiological roles of the SNF1 family are currently better defined in yeast and animals as in plants.

1.3.1 The yeast SNF1 system

Glucose repression is a general regulatory system in yeast that affects the expression of a multitude of genes. SNF1 is activated in response to low cellular glucose levels and required for the derepression of these genes (Gancedo, 1998). When the derepression mechanism is blocked, the yeast can not grow on alternative carbon sources. The SNF1 gene was originally defined in *S. cerevisiae* via mutation which can not grow on sucrose or raffinose (*snf1*-mutants) (Carlson et al., 1981) or glycerol or maltose (*cat1*-mutants) (Zimmermann et al., 1977), or ethanol-*ccr1*-mutants (Ciriacy, 1977). All of these genes turned out to be alleles of the same gene. SNF1 affects not only metabolic adaptation but also developmental processes. Mutations in the kinase complex cause defects in meiosis and sporulation (Honigberg and Lee, 1998), filamentation and invasive growth (Cullen et al., 2000), survival in stationary phase (Ashrafi et al., 1998), life span and aging (Ashrafi et al., 2000).

The SNF1 gene has been cloned and genetically mapped to a position distal to *rna3* on chromosome IV (Celenza and Carlson, 1984). The protein exists as a complex consisting of three subunits: α-(catalytic) subunit snf1, one of three β-subunits SIP1, SIP2, Gal83, and the γ-subunits snf4.

The α-subunit is generally 550-650 amino acid residues in primary sequence with the 300-amino acid serine/threonine kinase catalytic domain at the N terminus. The remaining 250-300 amino acids comprise a regulatory domain with two distinct parts. The first half of the regulatory domain contains an autoinhibitory sequence that down-regulates the kinase activity under nutrient-rich conditions (Crute et al., 1998).

The γ-subunit plays an important regulatory role. Mutation of snf4 (also known as CAT3) was detected by screening of snf1/cat1 mutants (Neigeborn and Carlson, 1984). The γ-subunit associates with the kinase complex through a constitutive independent interaction with each of the β-subunits (Jiang and Carlson, 1997). The interaction between SNF1 and SNF4 is strongly regulated by the glucose signal: SNF4 binds to the SNF1 regulatory domain in low glucose, whereas in high glucose the regulatory domain binds to the kinase domain of SNF1 itself. Finally, these interactions have been conserved from yeast to plants, indicating that homologs of the SNF1 kinase complex respond to regulatory signals by analogous mechanisms (Jiang and Carlson, 1996).

The β-subunits SIP1, SIP2 and GAL83 define a family of homologous proteins containing the ASC domain, highly conserved at the C-termini of all three proteins. This domain can mediate protein-protein interaction with the SNF1 kinase complex. The β-subunits (SIP1, SIP2, Gal83) are required for kinase function and substrate definition (Schmidt and McCartney**,** 2000). The β-subunits regulate the subcellular localization of the Snf1 kinase. Green fluorescent protein and Gal83, SIP1, or SIP2 fusion proteins show different patterns of localization to the nucleus, vacuole, and/or cytoplasm. Such independent regulation of the localization and the activity of the SNF1 kinase, combined with the distinct localization of kinases containing different β-subunits, affords versatility in regulating physiological responses (Vincent et al., 2001).

Complete and precise deletion of all three beta-subunit genes inactivates the Snf1 kinase. The Snf1-Gal83 form of the kinase is required for adherence, whereas either Snf1-Gal83 or Snf1-Sip2 is sufficient for filamentation (Vyas et al., 2003).

The yeast SNF1 can be phosphorylated at threonine 210 and activated by upstream kinases and by SNF1 itself (Celenza and Carlson, 1986). An additional regulating step in SNF1 activity is the rearrangement of the γ-subunit (McCartney and Schmidt, 2001). The two-step model for the regulation of Snf1 kinase modified by McCartney is presented in Fig. 1. The intracellular signal responsible for triggering this activation remains unknown. The AMP:ATP ratio may be important during carbon source transitions but does not appear to be the primary signal during steady-state growth (Gancedo, 1998).

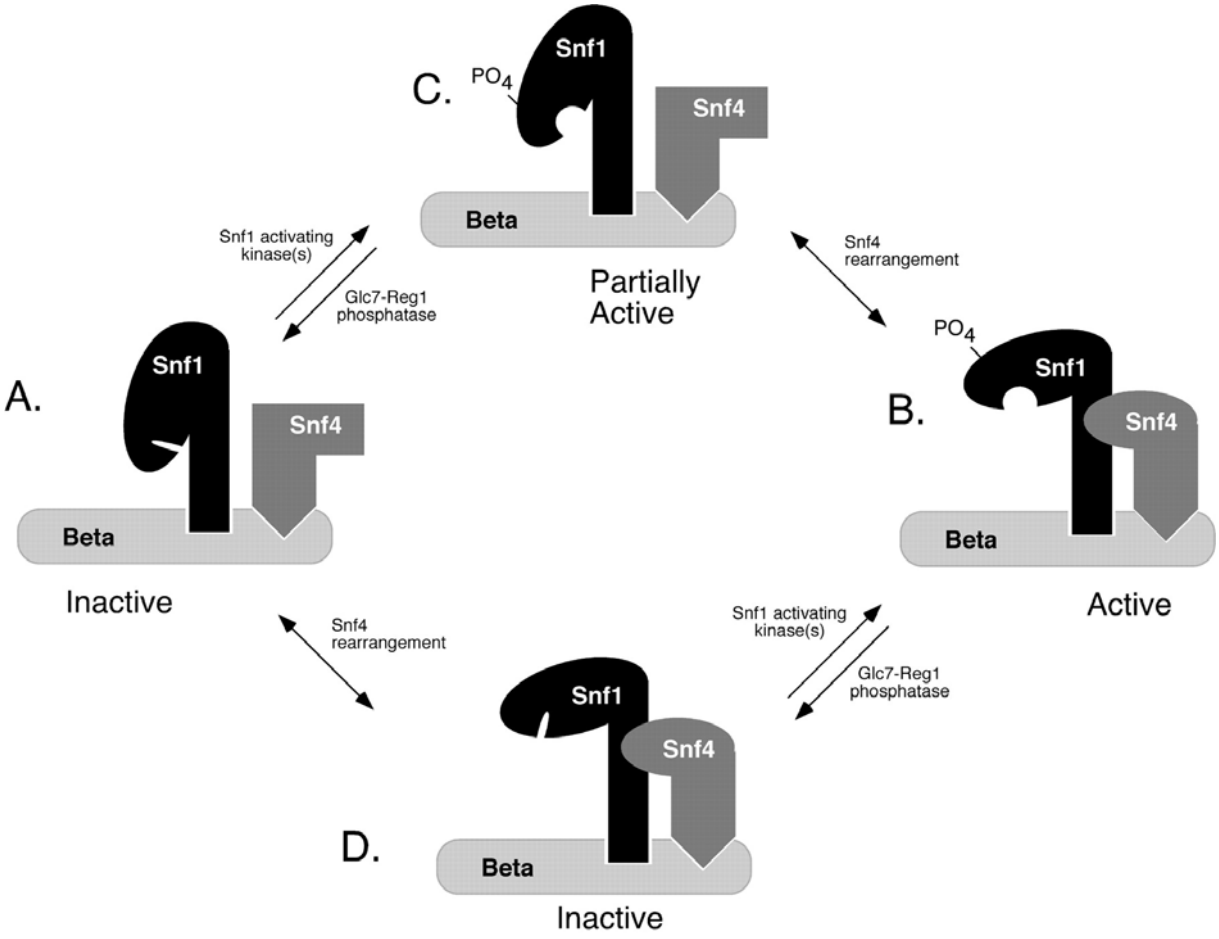


Figure 1. Two-step model for the regulation of Snf1 kinase. Proposed first by Jiang and Carlson (Jiang et al., 1996) and modified by McCartney and Schmidt et al. (2001). Under repressing conditions in the presence of glucose, the Snf1 kinase is catalytically inactive and is not phosphorylated at threonine 210 (*A*). The fully active Snf1 that is present in derepressed cells is phosphorylated on threonine 210 and has undergone a conformational rearrangement that is mediated by the Snf4 protein (*B*). Alternate intermediate forms of Snf1 are shown in *C* and *D*.

The role of SNF1 in the metabolic adaptation to glucose depletion and to growth on alternate carbon sources is understood in some details.

SNF1 regulates the transcription of metabolic genes by controlling the transcriptional repressor Mig1 and the activators Cat8 and Sip4 (Carlson, 1999).

Mig1 is a Cys_2-His_2 zinc finger protein that binds to the promoters of sucrase, galactosidase, maltase, and other glucose-repressible genes, and mutation of Mig1 or its binding sites partially relieves glucose repression. Snf1 is required for the phosphorylation of Mig1 and appears to inhibit the repression by Mig1 (De Vit et al., 1997; Treitel et al., 1998).

Mig1p repressor also can bind to a STRE (stress response element) in the *DOG2* (2-deoxyglucose-6-phosphate phosphatase) promoter. *DOG2* is induced by osmotic and oxidative stress. Snf1p kinase is likely to be involved in the signalling pathway of oxidative and/or osmotic stress in regulation of *DOG2* (Tsujimoto et al., 2000).

Sip4 is a transcriptional activator that binds to the CSRE (carbon source-responsive element) of gluconeogenic genes in *S. cerevisiae*. The Snf1 protein kinase interacts with Sip4 and regulates its phosphorylation and activator function in response to glucose limitation. (Vincent et al., 2001).

Cat8 is specifically required for derepression of CSRE -dependent genes. The biosynthetic control of CAT8 as well as transcriptional activation by Cat8p requires a functional Snf1p protein kinase (Hiesinger et al., 2001). Fig. 2 presents an example of the regulatory system responding differently to the concentration of glucose and the regulation of fructose-1,6-bisphosphatase gene.

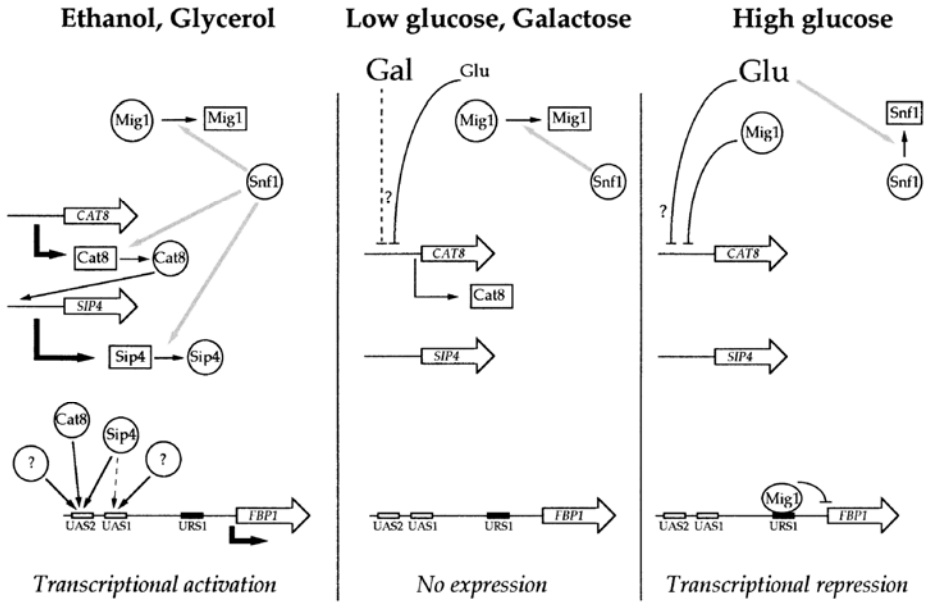


Figure 2. Proposed model for the regulation of *FBP1*(fructose-1,6-bisphosphatase gene) (Zaragoza et al., 2001). *FBP1* promoter contains two UAS (upstream activating sequence) elements called carbon-source-responsive element (CSRE) and one upstream repressing sequence (URS) able to bind the regulatory protein Mig1.

During growth on a gluconeogenic carbon source the protein kinase Snf1 is fully operative: it turns off the repressing protein Mig1 and turns on the transcriptional activators Cat8 and Sip4. Cat8 in turn is required for the transcription of *SIP4* and for activating the transcription of *FBP1* mediated by UAS1 and UAS2.

In wild-type strains growing on galactose or low glucose Snf1 is still able to inactivate Mig1. However, at low glucose *CAT8* expression is low and on galactose *CAT8* and *SIP4* are only partially derepressed. *FBP1* not is expressed under this conditions.

Snf1 is inactive at high glucose, Mig1 is in its active state and fully represses the expression of *CAT8.* The proteins that activate the transcription of *SIP4* and *FBP1* are absent and the Mig1 complex binds URS1 and represses transcription. *FBP1* is therefore completely turned off (Zaragoza et al., 2001).

Another transcriptional activator regulating by SNF1 is Adr1. Adr1 controls the expression of genes required for ethanol, glycerol, and fatty acid utilization and acts directly on the promoters of ADH2, ACS1, GUT1, CTA1 and POT1. Snf1 promotes Adr1 chromatin binding in the absence of glucose. Chromatin binding by Adr1 is not the only step in ADH2 transcription that is regulated by glucose repression. Adr1 can bind to chromatin under repressed conditions in the presence of hyperacetylated histones (Young et al., 2002).

SNF1 can regulate transcription through modification of chromatin. Snf1 activates transcription of *INO1* (inositol-1-P synthase) by phosphorylating Ser-10 of histone H3. Ser-10 phosphorylation is necessary for subsequent acetylation of Lys-14 of the same histone (Lo et al., 2001)

SNF1 protein kinase can also directly stimulate transcription by the yeast RNA polymerase II holoenzyme that may, in principle, provide a parsimonious mechanism for controlling a large array of genes. Snf1 kinase activity can stimulate transcription by the holoenzyme in a glucose-regulated manner and Snf1 interacts physically with the Srb/mediator complex of the holoenzyme (Kuchin et al., 2000).

1.3.2 The mammalian AMP-activated protein kinase system

The mammalian AMP-activated protein kinase (AMPK) was originally identified in rat liver crude extract as a regulatory kinase that ATP-dependently phosphorylates 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR), a key enzyme in cholesterol biosynthesis (Beg et al., 1978), and acetyl-CoA carboxylase (ACC), the rate-limiting enzyme of fatty acid synthesis (Carlson and Kim, 1974). Phosphorylation of these enzymes results in their inactivation. It was also shown that other metabolic enzymes are substrates of AMPK including hormone sensitive lipase, glycogen synthase and creatin kinase. AMPK has been purified from a number of species, including human, rat and pig (Carling et al., 1989; Mitchelhill et al. 1994; Sullivan et al., 1994). Like the yeast SNF1 complex, AMPK forms a heterotrimeric complex

consisting of α-(catalytic) subunit (63-kDa), β-(non-catalytic) subunit, and the γ-subunit (38 kDa) (Stapleton et al., 1994). The AMPK α-subunit is 64% identical in its catalytic core with the *S. cervisiae* Snf1p protein kinase. AMPK γ-subunit is related to the *S. cerevisiae* protein Snf4p (CAT3), whereas AMPK-β is related to the *S. cerevisiae* Sip1p/Sip2p/Gak83p family of proteins. Multiple isoforms of subunits have been identified. The hepatic AMPK catalytic subunit (α1-548 residues) is distinct from the previously cloned kinase subunit (α2- 552 residues). The isoforms have 90% amino acid sequence identity within the catalytic core but only 61% identity elsewhere. There appear to be at least three isoforms of the γ-subunit in brain with γ2 and γ3 present, distinct from the rat liver γ1 isoform. Human brain also contains multiple β-subunit isoforms distinct from the rat liver β1-isoform. Thus, a potentially large subfamily of heterotrimeric AMPKs, based on various combinations of all three AMPK subunits, may exist (Stapleton et al., 1994).

The AMPK kinase is activated by AMP (Carling et al., 1989) and by phosphorylation of an upstream protein kinase (AMPKK). The upstream and downstream kinases can form a complex. Both AMPKK and AMPK are directly activated by AMP. The molecular masses of catalytic subunits, Stokes radii, sedimentation coefficients, and the estimated native molecular masses are very similar. It seems likely that they are closely related proteins and the AMPKK builds also a heterotrimeric complex (Hawley et al., 1995). The regulation of AMPK by AMP is show in Fig. 3.

The kinase can be inactivated by the protein phosphatase-2A (PP2A) or protein phosphatase – 2C (PP2C) (Moore et al., 1991).

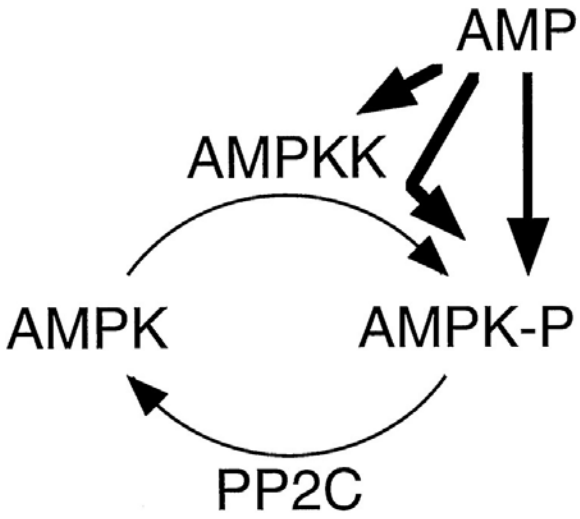


Figure 3. Proposed effects of AMP on the AMPK protein kinase cascade (Hawley et al., 1995): direct allosteric activation of the downstream protein kinase; binding to the

downstream protein kinase makes it a better substrate for the upstream protein kinase; direct activation of the upstream protein kinase.

Inactivation of AMPK by PP2C was also completely blocked by AMP (Davies et al., 1995). AMP activates the cascade by four mechanisms: (*a*) allosteric activation of AMPKK; (*b*) binding to AMPK, making it a better substrate for AMPKK; (*c*) allosteric activation of AMPK; (*d*) binding to AMPK, making it a worse substrate for PP2C. At least three of these effects are antagonised by high concentrations of ATP, so that the system is activated by high AMP, coupled with low ATP (Hawley et al., 1995).

In eucaryotic cells AMP and ATP concentrations tend to change in reciprocal directions because of the action of adenylate kinase. The reaction ($2ADP \leftrightarrow ATP+AMP$) is maintained at close equilibrium. The AMP:ATP ratio in cells varies approximately as the square of the ADP:ATP ratio. In fully energised cells, the AMP level is extremely low and AMPK is in the inactive state. If the energy status of the cells is compromised such that the ATP:ADP ratio falls, the AMP rises and the AMPK cascade switches on. It has been proposed that AMPK represents a cellular fuel gauge (Hardie et al., 1997). AMPK action initiates energy-conserving measures (such inhibition of most macromolecules biosynthesis) and mobilises the catabolism of alternative carbon sources, and therefore protects cells against environmental and nutritional stress. Activation of AMPK has been demonstrated in response to a variety of stress conditions with ATP depletion. These include treatment with arsenite, cadmium ions, heat shock, hypoxia (Corton et al. 1994) as well as exercise in skeletal muscle (Winder et al., 1997) and interruption of the blood supply in heart muscle (Kudo et al., 1996).

Targets of mammalian AMPK are HMG-CoA reductase, acetyl-CoA carboxylase, hormone-sensitive lipase, glycogen synthase, raf-1 protein kinase.

AMPK phosphorylates and inactivates a single site in a catalytically active fragment of rat liver HMG-CoA reductase, corresponding to Ser-871 (Clarke et al., 1990) and therefore inhibits sterol synthesis.

AMPK also phosphorylates three sites on rat acetyl-CoA carboxylase corresponding to Ser-79, Ser-1200, and Ser-1215 (Davies et al., 1990). Activation of AMPK by heat shock causes almost total inhibition of fatty acid synthesis in rat hepatocytes and stimulates indirectly ATP-producing catabolic pathways as fatty acid oxidation in heart muscle (Kudo et al., 1996). In rat cortical astrocytes AMPK was shown to be functionally active. Chemical induced hypoxia by sodium azide seemed to rely on the cascade: increase of the AMP/ATP ratio --> AMPK stimulation --> acetyl-CoA carboxylase inhibition --> decrease of malonyl-CoA concentration

--> carnitine palmitoyltransferase I deinhibition --> enhanced ketogenesis (Blazquez et al., 1999).

1.3.3 SnRK1 in plants

SnRK1 in plants is proposed to act as a global regulator of carbon metabolism (Halford and Hardie, 1998). Existing data support complex and distinct functions of SnRKs in plants. As in yeast, the signal transduction in plant cells by SnRK1 can be performed on transcriptional level of regulation. For example, expression of an antisense *SnRK1* construct in potato prevents transcriptional activation of a Suc-inducible Suc synthase gene, suggesting its involvement in sugar activation but not repression (Purcell et al., 1998). The *A. thaliana* pleiotropic regulatory locus (*prl1*) mutant exhibits transcriptional derepression of a variety of sugar-regulated genes and a sugar-hypersensitive growth phenotype (Nemeth et al., 1998). It has been shown that PRL1 is an inhibitor of the *A. thaliana* SnRKs (AKIN10 and AKIN11) and is regulated negatively by Glc. However, the *prl1* mutation does not seem to affect the regulation of SnRK activity by sugars (Bhalerao et al., 1999). The complex interactions of SnRK with PRL1, SKP1/ASK1 (ubiquitin ligase), and a subunit of the 26S proteasome (α4/PAD1) revealed another aspect of SnRK function in protein degradation (Farras et al., 2001).

SnRKs also play an important role in metabolism by directly phosphorylating and inactivating important enzymes and proteins. It was shown that spinach SnRK1 regulates carbon metabolism by phosphorylation and inactivates the key biosynthetic enzymes such as 3-hydroxy-3-methyl glutaryl CoA reductase, nitrate reductase (NR), and Suc phosphate synthase *in vitro* (Sugden et al., 1999b). It was also reported that SnRK1 complex can phosphorylate a spinach Suc phosphate synthase peptide. Glc-6-phosphate may act as a negative regulator of SnRK activity (Toroser et al., 2000). Interestingly, a SNF1-like gene, together with genes encoding enzymes in primary sugar metabolism (ADP-glucose pyrophosphorylase and Suc synthase), was shown to be expressed asymmetrically in tomato apical meristems, with higher expression levels in the parts destined to form leaves (Pien et al., 2001). Resent investigations showed the phosphorylation of class I heat shock protein by SnRK1 in barley (Slocombe et al., 2004).

SnRK1 is regulated transcriptionally and post-transcriptionally. Unlike mammalian AMPK, plant SnRKs does not seem to be activated directly by AMP, but is activated through

phosphorylation by an upstream protein kinase. The evidence that some plant SnRKs are activated by sugars and are involved in sugar-activated gene expression implies an opposite regulation from that of mammalian AMPKs and yeast Snf1. However, AMP appears to inhibit dephosphorylation and the concomitant inactivation of spinach SnRK activity at physiological concentrations (Sugden et al., 1999a). Therefore, the exact nature of SnRK activation in response to sugars and AMP remains unclear. The result that the tobacco SnRK NPK5 is constitutively active in yeast (Muranaka et al., 1994) underscores the differential regulation of SnRK activity in yeast and plants.

SnRK1 is implicated in a variety of distinct regulatory and developmental processes. Several SnRKs from different subfamilies *A. thaliana*, wheat, maize and rice are regulated differentially by light, temperature (Ohba et al., 2000), cytokinin (Ikeda et al., 1998), developmental stage (Takano et al., 1998) and sugars (Chikano et al., 2001). Interestingly, the ABA-induced barley SnRK, PKABA1, mediates the ABA suppression of GA-induced α-amylase gene expression in aleurone cells (Gomez-Cardenas et al., 2001). Expression of antisense SnRK1 in barley anthers causes abnormal pollen development and male sterility (Zhang et al., 2001).

Further functional analysis of SnRKs can be complicated by redundant functions and multiple interactions with regulatory proteins. Like yeast snf1 and mammalian AMPK, the plant SnRK1 works associated with different subunits. In *A. thaliana* AKINß1 and AKINß2 encode proteins similar to SIP1, SIP2 and GAL83 and AKINγ coding for a protein with similarity to SNF4 (Bouly et al., 1999). An *A. thaliana* snf4 ortholog was functionally identified by expression of an Arabidopsis cDNA library in yeast (Kleinow et al., 2000). The AKINß2, a plant ortholog of conserved Snf1/AMPK beta-subunits, has been functionally characterised and allows to form different complexes with the catalytic alpha-subunits of Arabidopsis SnRK protein kinases AKIN10 and AKIN11 *in vivo* (Ferrando et al., 2001). Another SnRK1 interaction protein SnIP1, was isolated by two-hybrid screening with barley SnRK1b, a seed-specific form of SnRK1. No homologous sequences were identified from outside the plant kingdom, but weak sequence similarity was found between the SnIP1 peptide and yeast SNF4 and its mammalian homologue AMPK (Slocombe et al., 2002). A novel class of plant SnRKs was recently identified containing only two components: a Snf1-related kinase subunit and a unique regulatory βγ-subunit, which appears to have developed by domain fusion during plant evolution (Lumbreras et al., 2001). The most complete investigation of gene expression of the SNF1-related kinase complex was performed in developing tomato seeds and germinating seeds of *Medicago truncatula* (Bradford et al., 2003; Buitink et al., 2004). cDNAs

corresponding to the kinase (LeSNF1), regulatory (LeSNF4), and localization (LeSIP1 and LeGAL83) subunits of the SnRK1 complex from tomato (*Lycopersicon esculentum* Mill.) were isolated. LeSNF4 mRNA became abundant at maximum dry weight accumulation during seed development and remained high when radicle protrusion was blocked by ABA, water stress, far-red light, or dormancy, but was low or undetectable in seeds that had completed germination or in GA-deficient seeds stimulated to germinate by GA. In contrast, LeSNF1 and LeGAL83 genes were essentially constitutively expressed in both seeds and leaves regardless of the developmental, hormonal or environmental conditions. Regulation of LeSNF4 expression by ABA and GA provides a potential link between hormonal and sugar-sensing pathways controlling seed development, dormancy, and germination. A model of possible regulation is showen in Fig. 4.

The SnRK1-complex was also investigated in germinating seeds of *M. truncatula.* Four *β*-subunits (MtAKIN*β*1-4) and three *γ* subunits (MtAKIN*βγ*, MtSNF4b and MtAKIN*γ*) were identified and characterized in seeds of *M. truncatula*. Their transcripts were found to accumulate differentially in vegetative and seed tissues and appeared to be differentially modulated during germination and the imposition of stress. MtAKIN*γ* and MtAKIN*β*3 showed identical patterns of expression upon osmotic shock, whereas transcripts of MtAKIN*γ* and MtAKIN*β*1 were strongly up-regulated upon starvation of the radicles.

Addition of glucose during the starvation process reversed this effect. MtAKIN*β*2 and MtSNF4b were specifically induced upon the re-induction of desiccation tolerance in germinated, desiccation-sensitive radicles, whereas only MtSNF4b expression was repressed by an inhibitor of ABA synthesis. Second *γ* subunit, MtAKIN*βγ*, was transiently expressed early during the induction of desiccation tolerance, and its expression could be modulated by blocking the respiratory ATP production by cyanide.

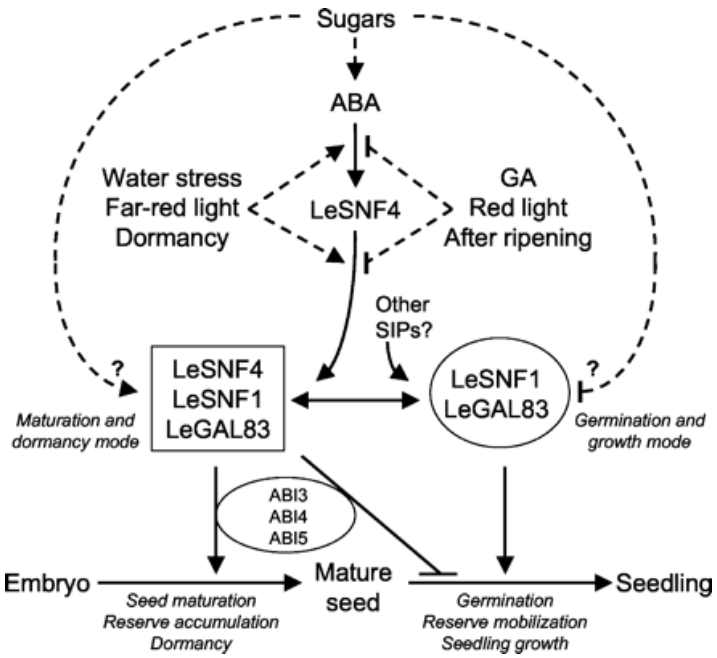


Figure 4. Model of regulation and hypothetical roles of the SnRK1 complex during seed maturation and germination proposed by Bradford et al., 2003. ABA synthesis during seed development or in response to high sugar levels during germination results in ABA accumulation, inducing LeSNF4 expression. Dashed lines both above and below LeSNF4 are shown because it is unclear where each of these factors acts in influencing expression. When LeSNF4 is present, it binds with LeSNF1/LeGAL83, potentially altering the kinase activity or its interaction with other regulatory factors or substrates, resulting in the maintenance of a "maturation/dormancy" metabolic state and inhibiting reserve mobilization. When LeSNF4 is absent, LeSNF1 may have altered activity, specificity, or interactions with other SIPs, resulting in transition to the "germination/growth" mode required for reserve mobilization and seedling growth. In either of the metabolic modes, the activity of the LeSNF1 complex may be sensitive to regulation by sugars (indicated by the broken lines) and could be involved in the regulation of ABI3, ABI4, and ABI5 and other proteins that are known to influence the transition from seed maturation and dormancy to germination and growth. (Arrows indicate promotion; bars indicate inhibition.)

1.4 Molecular genetic approaches to analyse seed development

1.4.1 Macroarry analysis

To understand the regulatory network governing metabolism in developing pea seeds, macroarray analysis of gene expression in seeds could be an effective method.

The recently developed cDNA microarray technology enabled monitoring of cell-, tissue- and developmental stage-specific gene expression profiles and simultaneous quantitative analyses of expression levels of many genes (Derisi et al., 1996; Baldwin et al., 1998; van Hal et al., 2000; Wisman and Ohlrogge, 2000). Such analyses could help to find new genes possibly involved in specific processes or signalling pathways. Today, macro- and microarray technology is an obvious choice to identify global gene expression patterns during development. The array technology was used for a general analysis of gene expression in developing *Arabidopsis* seeds. It was found that 25% and 10% of the genes tested, increased two and ten times, respectively, during *Arabidopsis* seed development. *Arabidopsis* arrays were also found to be useful for transcriptional profiling of mRNA isolated from developing oilseed rape (*Brassica napus*) seeds, and expression patterns are well correlated between the two species (Girke et al., 2000). A 1421-cDNA array was also applied for analysis of gene expression in maternal and filial tissues of barley caryopses (Sreenivasulu et al., 2004). Principle component analysis (PCA) defined distinct expression networks in the pre-storage (0, 2, and 4 days after flowering (DAF)) and early storage phase (10 and 12 DAF). During an intermediate phase (6 and 8 DAF), PCA visualizes a dramatic re-programming of the transcriptional machinery.

Gene expression in legume seeds was also investigated by array analyses. Using a 9,280 cDNA clone array, transcript patterns associated with somatic embryogenesis in soybean were identified (Thibaud-Nissen et al., 2003). The combination of reverse genetic and macroarray analyses gives a good possibility to investigate the role of unknown genes and to understand the cross talk between different signaling pathways, for instance in hormone regulated pathways. The micro- and macroarray analyses were also broadly used for characterization of transgenic plants with changed expression of certain genes, for example, transcription factors (Cluis et al., 2004) and transcriptional activators, involved in ABA signaling (Abe et al., 2003; Duque and Chua , 2003).

All these results suggest that cDNA macroarray technology can be useful to study the general expression profiles of tissue-specific or environmentally regulated genes, as well as changes in expression patterns of transgenic plants.

1.4.2 Antisense technology

The use of reverse genetics to expand our knowledge of plant molecular biology, biochemistry and physiology has been proven to be a powerful and valuable approach.
It can be traced back to the isolation of the first cDNA encoding a gene of interest, the use of the *Agrobacterium* Ti plasmid to introduce foreign DNA into plant cells (Hernalsteens *et al*., 1980) and the establishment of routine plant transformation systems (Bevan, 1984; Horsch *et al*., 1985). It became possible to repress plant genes using reverse cDNAs linked to strong promoters, with the aim of modifying metabolism. The discovery of the antisense technology in order to repress plant gene expression (van der Krol *et al*., 1988; Smith *et al*., 1988) provided the most powerful and widely-used method for investigating the roles of specific enzymes and regulatory genes in metabolism and plant development. With such molecular tools in place, plant metabolism became accessible to investigation and manipulation through genetic modification and dramatic progress was made in subsequent years (Stitt and Sonnewald, 1995), particularly in studies of legumes seeds (Weber et al., 2000) and nodules (Cordoba et al., 2003).

1.5 Research objectives

Investigation of seed development of legumes at physiological, biochemical and molecular levels are research subjects of our group. The coordination of processes in seed metabolism was intensively investigated over many years. Nevertheless, impact of single signal transduction pathways, regulation of the gene expression coding for regulatory genes involved in coordination of developing processes has to be elucidated and analysed in more details.
Thus, the aim of the research project was the molecular dissection of the regulatory network controlling gene expression and coordination of metabolic pathways during early and late embryogenesis. Special emphasis was put on the characterisation of SnRK1 as an important component of the signal transduction chain.

At the time when this work was started, there was little knowledge on the role of SnRK1 in plants. Therefore, a first goal of the present work was to identify, clone and sequence genes encoding for SnRK1 expressed in legume seeds. Based on previously obtained data, demonstrating activity of SnRK1 in *Vicia faba* embryos and highly conserved nucleotide sequences across the snf1 protein family, it was expected to find a gene, encoding SnRK1 in a *V. faba* cDNA library.

After identification of cDNA sequences, it was important to study the expression of SnRK1 genes with respect to seed development. In order to elucidate SnRK1 functions it was necessary to generate of transgenic plants carrying the cDNA sequences for VfSnRK1 in antisense orientation under a seed specific promoter with the aim to generate loss-of-function in *Pisum sativum* plants. Molecular and biochemical analysis of such transgenic plants has been performed to study the effect of SnRK1 deficiency during seed development. Macroarray analysis of transgenic seeds should clarified the components of the regulatory network including downstream effector genes as well as transcriptional regulators. The data are expected to provide an important contribution to the understanding of seed developmental processes and the function of the SnRK1 gene product.

2. MATERIAL

2.1 Organisms

Plant material

Nicotiana tabacum L. cv Gavana

Nicotiana plumbagenifolia L.

Vicia faba L.- ssp. minor cv Fribo

Vicia narbonensis L.

Pisum sativum L.- cv Erbi

Bacterial strains

Escherichia coli strains:

DH5α	recA1, endA1, gyrA96, thi-1, hsdR17, (r_K-m_K+), relA1, supE44, [u80ΔlacZΔM15, Tn10, (Tet^r)]^c (Sambrook et al., 1989)
XL1Blue	recA1, endA1, gyrA96, thi-1, hsdR17, relA1, lac[F'proAB$lacI^qZ$ΔM15, Tn10, (Tet^r)]^c (Jerpseth et al.1992)
XLOLR	D(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 thi-1 recA1 gyrA96 relA1 lac [F' proAB lacIqZDM15 Tn10 (Tet^r)]^c (Stratagene, La Jolla, USA)
XL10-Gold®^{d}	TetrD (mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB $lacI^qZ$DM15 Tn10 (Tet^r) Amy Cam^r] (Stratagene, La Jolla, USA)

A. tumefaciens strain:

EHA105 genotype C58 pTiBo542; T-region::aph, Km(S); A281; super-virulent (Hood et al., 1993)

2.2 Plasmids and genes

pBluescript SK(+)	Amp^r (Stratagene, La Jolla, USA)

pUC18	Amp^r (Yanisch-Perron et al., 1985)
pBK-CMV	Km^r (Stratagene, La Jolla, USA)
pGPTV	Km^r (Becker et al., 1992)
pUC18/VicP	vicilin B4 promoter and 3'ocs terminator cloned in pUC18 (Appendix 1)
pUC18/ USPP	USP promoter cloned in pUC18 (Appendix 1)

2.3 Primers and oligonucleotides

Sequencing primers

BK	5' ACA GGA AAC AGT CAT GAC CTT G 3'
M13-20	5'-GTA AAA CGA CGG CCA GT-3'
SNF1R	5'-GCA TCA TCA CAT TAT CCG CC-3'
SNF1U	5'-GCA CAA CAG GAC CAC CTC AGG-3'
2R_SEQUPP	5'-GGC GAG GTT TTT GAG AAG AG-3'
2R_SEQLOW	5'-CGC CAA GAG TGA ATC TAA CAC ACC-3'

RT-PCR primers

SNF3	5'-CCA YCT GCA YTT CAT GTT RTA-3'
SNF5	5'-GAT GGH CAY TTT YTS AAG AW AGY CYGG-3'

Genome walking primers

Primers for probe 1	
5END2	5'-GCG ACT ACA GAG AAA ACG G-3'
5END113	5'-CAT CAG CAA CAA AAC CGA C-3'
Primers for probe 2	
ZP12	5'-CTA TTT TAC GCG AAA TTC CAA C-3'
ZP22	5'-GCT GAT GCT GTT GTA TTC TAT TG-3'
AP1	5'-GTA ATA CGA CTC ACT ATA GGG C-3'
AP2	5'-ACT ATA GGG CAC GCG TGG T-3'
GWPI	5'-CGG TTA AGA ATC TTT ATG GCG ACT TTA TG-3'
GWPII	5'-GAC CGG TCA ATA CAT GCT CTG CAA TTT TCA-3'

GW111 5'-GAG AGA AAA CGG AAG TTG CAG TAA GAG ATGG-3'

GW221 5'-GGA GAG AGA GTG AAG AAA CCC TAA ATC GGTG-3'

Primers for promoter cloning in GUS gene construct

Prom3short SalI 5'-GAAGGTCGACCAGCGACTAGAGAGAAAACGGAAG -3'

Prom3long SalI 5'-ATCCGTCGACTTTCGAAATCTAGGGGAAAATTCA -3'

Prom5 PstI 5'-CCGGCTGCAGAAATAAATTATATAACATAATTAA -3'

EST amplifying primers

M13-21PE 5'-ACGACGTTGTAAAACGACGGCCAG –3'

MVR-26 5'-CTCACTAAAGGGAACAAAAGCTGGAG –3'.

2.4 Media

Bacterial media

LB 10g NaCl, 5 g Tryptone, 5 g Yeast extract for 1 L (pH 7.4)

TBY 5 g $MgSO_4 \cdot 7H_2O$, 5 g NaCl, 10 g Tryptone, 5 g Yeast extract for 1 L (pH 7.4)

YEB 0.5 g $MgSO_4 \cdot 7H_2O$, 5 g Beef extract, 5 g Peptone, 5 g Sucrose, 1 g Yeast extract for 1 L (pH 7.0)

SOC 0.58 g NaCl, 0.186 g KCl, 10 mM $MgCl_2$, 20 g Tryptone, 5 g Yeast extract, 20 mM Glucose for 1 L (pH 7.4)

All solidified media contain 1.5% Difco-agar.

Plant growth media

B5 1xMS micro-salts (Murashige and Scoog, 1962), 1xB5 macrosalts and vitamins (Gamborg, 1968), 3% sucrose, 0.2 mg/L kinetin, 1 mg/L 2.4D, PPT (pH 5.8)

P1 1xMS salts, 1xB5 vitamins, 3% sucrose, 2 mg/L BA, 2 mg/L NAA, 0.4 g/L Betabactyl, 4 mg/L PPT (pH 5.8)

MS4 1xMS salts, 1xMS vitamins, 2% sucrose, 4 mg/L BA, 100 mg/L myo-inositol, 0.1 mg/L IBS, 10 mg/L PPT, 0.4 g/L Betabactyl (pH 5.8)

MS1	1xMS (modified MS4) salts, 1xMS vitamins, 2% sucrose, 1 mg/L BA, 100 mg/L myo-inositol, 1 mg/L IBS, 2 mg/L GA_3, 4 mg/L PPT, 0.4 g/L Betabactyl (pH 5.8)
MS1/2	0.5 MS salts, 1xMS vitamins, 1% sucrose (pH 5.7)
LS	1xMS (modified MS4) salts, 1xB5 vitamins, 2% sucrose (pH 5.8)
T1	LS supplemented with 1 mg/L BA, 0.1 mg/L NAA, 0.5 g/L claforan, 0.1 g/L kanamycin
T2	LS supplemented with 0.1 mg/L BA, 0.25 g/L cefotaxime, 0.1 g/L kanamycin
T3	LS supplemented with 0.1 g/L cefotaxime, 0.1 g/L kanamycin

All solidified media contain 0,8 % Difco-agar.

2.5 Chemicals

Amersham, Braunschweig, Germany	[γ ^{32}P]ATP, [γ ^{33}P]ATP Nylon membrane "Hybond N+ "
BioGenes Ltd., Berlin, Germany	SAMS peptide
Difco, Detroit MI, USA	Bacto®-Agar, Bacto®-Trypton, Yeast extract, Beef extract, Peptone, Betabactyl
Duchefa, Brüssel, Belgium	Murashige-Scoog whole medium solid substance, Vitamins, NAA, Kanamycin, PPT, Kinetin 2,4D, GA_3, BA
Fluka, Buchs, Schweiz	DEPC
Gibco-BRL, Gaithersburg MD, USA	Agarose, 1kb DNA Ladder
Hoechst AG, Frankfurt a .M. Germany	Cefotaxim-natrium (Claforan),
Kodak, Rochester NY, USA	X-Ray film, Diapositive films
Merck, Darmstadt, Germany	Ethidium bromide, Benzamidine, DTT, Formamide, HEPES, Iodine, Magnesium chloride, MOPS, Potassium iodide Sodium acetate, Sodium chloride, Sodium fluoride, Sodium hydroxid, Sodium hypochloride, Sodium phosphate, Sucrose, PMSF, Polyclar AT, Tris,
Pall, Dreieich, Germany	Nylon filter Biodyne® B
Roth, Karlsruhe, Germany	Acetic acid, Acetone, Ethanol, Chloroform, Formaldehyde, Glycerol, Isoamylalkohol,

	Isopropanol, Lithium chloride, Scintillation liquid "Rotiszint eco plus", Phenol , Phosphoric acid
Roche, Basel, Schweiz	ATP, BSA, dNTPs, SDS
Schleicher & Schuell, Dassel, Germany	Blotting paper GB 002
Serva, Heidelberg, Germany	EDTA, EGTA, Brij-35®,
Sigma, Louis MO, USA	Acetocarmine, IPTG
Whatman Ltd., Maidstone, UK	Phosphocellulose P81 paper, Anopore® inorganic membrane

2.6 Enzymes and kits

Amersham, Braunschweig, Germany	Megaprime DNA Labeling Kit, Sure Cloning Kit, RediPrimeII Labeling System
Biozym Scientific, Oldendorf, Germany	Gentra RNA Isolation Kit
Clontech, Palo Alto CA, USA	Universal GenomeWalker™ Kit
Dynal, Oslo, Norway	oligo (dT)-paramagnetic beads
Gibco-BRL, Gaithersburg MD, USA	Reverse transcriptase Superscript II
Roche, Basel, Schweiz	Dnase, Klenow enzyme, Lysozym, Restriction endonucleases, RnaseA, Shrimp alkalin phosphatase, Titan One Tube RT-PCR Kit
Masherey-Nagel, Düren, Germany	PCR Product isolation Kit
Quiagen, Hilden, Germany	DNA isolation Kit, QIAquick Nucleotide Removal Kit, QIAquick PCR purification Kit, Qiagen Plasmid Kit, Mini Elute 96 VF PCR Purification Kit
Stratagene GmbH, Heidelberg, Germany	*Pfu* polymerase, pBluescript II XR cDNA Library Construction Kit, ZAP Express® cDNA Synthesis Kit
Tropix, Bedford MA, USA	GUS-Light™ Assay Kit

2.7 Laboratory tools

AGS, Heidelberg, Germany	DNA Gel-electrophoresis tanks

Biorobotics, Cambridge, UK	BioGrig robot
Du Pont, Bad Homburg, Germany	Sorvall Centrifuge RC5C
Eppendorf, Hamburg, Germany	Thermomixer 5436 and 5437, Cold centrifuge 5402, BioPhotometer
Genetix, München-Dornach, Germany	QBot colony picking robot
GFL, Burgwedel, Germany	Hybridisation oven, Water bath
Heraeus, Osterode, germany	Shaker (Vortex Genie 2 ^{TM}), Centrifuge (Biofuge 13)
PerkinElmer, New Jersey, USA	Tri-Carb 2250 liquid scintillation counter, Thermocycler9700
Quiagen, Hilden, Germany	BioRobot 9600
Raytest, Straubenhardt, Germany	Fuji BAS2000 phosphoimager, imaging plates
Stratagene, Heidelberg, Germany	UV Stratalinker® 1800, Eagle-Eye® II Still Video System
Tecan**,** Crailsheim, Germany	GENESIS Workstation re-arraying robot
Zeiss, Jena, Germany	AxioCamHDc, Axioscop

2.8 Software

Commercial software:

DNASTAR, Inc., Madison WI, USA	Lasergene® sequence analysis
Electric Genetics Reston, VA, USA	StackPACK^{TM} v 2.1.1 management system for expression variation analysis
Microsoft Corporation, USA	Excel® 2000
Imaging Research, Canada	ArrayVision^{TM} analysis of array images
Raytest, Straubenhardt, Germany	TINA 2.0 quantitative analysis of Northern images
Redasoft, Bradford; Canada	Visual Cloning 3 plasmids drawing

Academic software:

J-Express application for the analysis of gene expression data provided by microarray experiments (Dysvik and Jonassen, 2001) http://www.molmine.com/

Software R (Ihaka and Gentleman, 1996) http://www.r-project.org/

HUSAR (Heidelberg Unix Sequence Analysis Resources) (Devereux et al., 1984)
http://genome.dkfz-heidelberg.de
StatistiXL 1.5 beta statistical analysis package http://www.statistixl.com/

2.9 Databases and interactive web-programs

BLAST - Basic local alignment search tool (Altschul et al., 1990)
(http://www.ncbi.nlm.nih.gov/BLAST/)
CDART Conserved Domain Architecture Retrieval Tool (Geer et al., 2002)
(http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi)
COGs - Clusters of Orthologous Groups of proteins (Tatusov et al., 1997)
http://www.ncbi.nlm.nih.gov/COG/
CR-EST- public available online resource providing access to sequence, classification, clustering, and annotation data of crop EST projects at the IPK (Künne et al., 2005)
(http://pgrc.ipk-gatersleben.de/cr-est/)
Gene Ontology (GO) Consortium- descriptions of gene products in different databases
http://www.geneontology.org/
KEGG - Kyoto Encyclopaedia of Genes and Genomes http://www.genome.jp/kegg/
NCBI - national resource for molecular biology information http://www.ncbi.nlm.nih.gov
NRPEP - nonredundant protein database ftp://ftp.ncbi.nih.gov/blast/db/FASTA/nr.gz
NSITEP-recognition of PLANT regulatory motifs with statistics and NSITEMPR-recognition of PLANT regulatory motifs conserved in several sequences http://www.softberry.com
Pfam- Protein families database of alignments (Bateman et al., 2004)
(http://www.sanger.ac.uk/Software/Pfam/)
PHRAP- program for assembling shotgun DNA sequence data (Gordon et al., 1998)
(http://www.phrap.org)
SMART a Simple Modular Architecture Research Tool (Schultz et al., 1998)
(http://smart.embl-heidelberg.de/)

3. METHODS

3.1 Plant growth and harvesting

Tobacco plants (*Nicotiana tabacum* L. cv. Havana) were grown *in vitro* in growth rooms under a 16 h light, 8 h dark cycle at 22°C. Leaves of four-weeks-old plant were used for *A. tumefaciens* mediated gene transfer. The seeds of the F1 and the following generations of transgenic *N. tabacum* plants were germinated *in vitro* on a selective medium. Plantlets and seedlings were grown under standard greenhouse conditions (16 h light at 17°C, 8 h dark at 15°C) .

For transformation of *Pisum sativum* L. cv. Erbi, plants were grown in growth chambers under a 16 h light at 15°C, 8 h dark at 14°C cycle beyond the stage of maximum fresh weight were harvested. Thereafter immature seeds were used for *A. tumefaciens*-mediated transformation of embryonic axis (Schroeder et al., 1993) followed by cultivation of shoots *in vitro* in light chambers under 12 h light/12h dark at 18°C. Mature plants are generated by grafting shoots onto 4-5 day-old etiolated seedlings (grown under sterile conditions *in vitro*). The F1 generations of *P. sativum* plants were grown in growth chambers under a 16 h light at 15°C, 8 h dark at 14°C.

In order to study *SnRK1* expression in developing seeds and in another tissues of *Vicia faba* L., the plants were cultivated in growth chambers under 16 h light at 15°C, 8 h at 14°C dark cycle. Self-pollinated flowers were tagged at the day of anthesis (DAP 0). Only the middle seeds from each pod were collected. Seeds were harvested from 17 DAP to 35 DAP. Seed coats were separated from cotyledons; *sink*- and *source*- leaves, roots and gynoecia were collected from mature plants. Samples were immediately frozen in liquid nitrogen and stored at -80°C.

3.2 RT-PCR mediated cloning of *VfSnRK1* cDNA

Two primers SNF3 and SNF5 were designed based on conserved regions of known SnRKs (EMBL data library). 1.5 µg of total RNA from *V. faba* cotyledons at 16 DAP were reverse transcribed by using Titan One Tube RT-PCR Kit. The PCR reaction was performed as follows: 30 min at 45°C, 5 min at 94°C (hot start), 0.5 min at 50°C (annealing), 1 min at 72°C (primer extension), 0.5 min at 94°C (denaturation) for 40 cycles, and 10 min at 72°C (final

extension). To obtain full length clones of SnRK, the amplified DNA bands of ~750 bp were used for screening of a λZAPII cDNA library from *V. faba* cotyledons (Heim et al., 1993). Preparation of phage replica filters, hybridization and washing were performed according to Sambrook et al., 1989. pBK-CMV was excised from the phage DNA according to the manufacture's instruction.

3.3 Basic cloning methods and sequencing

Insert preparation: PCR products were purified using "QIAquick PCR purification Kit" or were electrophoretycally separated on agarose gel and eluted from the gel by use of "PCR Product isolation Kit" .

Purified fragments were digested by appropriate restriction enzymes. For subcloning of DNA fragments into different vectors after restriction digestion, DNA fragments were separated on agarose gels and the fragments of interest were eluted from the gels. Restriction enzymes and nucleotides were removed by "QIAquick Nucleotide Removal Kit". Purified fragments were used for the ligation reaction.

Vector preparation: Plasmid extraction and purification was done using "Qiagen Plasmid Kit" or according to the fast preparation method described by Holmes and Quigley (1981). Plasmid DNA was digested with approptiate restriction enzymes. The vector was purified from agarose gel and usually dephosphorylated by "shrimp"-phosphatase.

The standard molecular cloning methods (restriction, digestion, ligation, DNA gel electrophoresis) were performed according to Sambrook et al., 1989.

Transformation of E. coli: The transformation of *E. coli* was performed using the heat-shock procedure (Cohen et al., 1972).

DNA sequences were determined at the IPK-Gatersleben or commercially by Qiagen.

3.4 Southern and Northern hybridizations

Genomic DNA was prepared from leaves using DNA isolation Kit. For Southern hybridization, 10 μg DNA was digested with appropriate restriction enzymes, separated on 0.8% agarose gel in Tris-borate buffer (Sambrook et al. 1989), transferred onto Hybond N+ membrane using 0.4 M NaOH and hybridized with cDNA fragments radioactively labeled with [^{32}P]dCTP by RediPrimeII Labeling System in Church buffer [(250 mM sodium

phosphate, 7% SDS, 1% BSA) (Church and Gilbert, 1984)] at 65°C according to the procedure described by Sambrook et al., 1989.

Total RNA was isolated as previously described by Heim et al., 1993 and analyzed by Northern blotting. 10 μg of RNA per lane were separated on 1% formaldehyde-agarose gels and blotted on nylon membrane Hybond N+. Membranes were hybridized according to Church and Gilbert (1984) using the same probe as for Southern hybridization. Additionally a 18S rDNA fragment was used in order to measure the relative RNA amounts loaded in each slot. Relative mRNA abundance was quantified as previously described by Weber et al., (1996) by densitometry of the images by TINA evaluation software and normalized for RNA loading in each lane. The normalized signal with highest intensity was set equal to 100, and the normalized signal intensity in other lanes are shown relative to it.

3.5 Extraction and ammonium sulfate fractionation of protein

Plant material (100 mg) was ground in liquid nitrogen with 0.009 g Polyclar AT and suspended in 500 μl extraction buffer (0.25 M mannitol, 50 mM HEPES, 50 mM sodium fluoride, 1 mM EDTA, 1 mM EGTA, 1 mM benzamidine, 1 mM DTT, 0.1 mM PMSF, pH 8.2). Cell debris was removed by centrifugation at 13 000 rpm for 15 min at 4°C.

Ammonium sulphate fractionation was performed by slowly adding ammonium sulphate to the 40% saturation while stirring continuously. Samples were stirred for further 20 min at 4°C and centrifugated at 13 000 rpm for 15 min at 4°C. The precipitated protein was resuspended in 50 μl fractionation buffer (50 mM Tris-HCl, 50 mM sodium fluoride, 1 mM EDTA, 1 mM EGTA, 1 mM benzamidine, 1 mM DTT, 0.1 mM PMSF, 0.02% (v/v) Brij-35, 10% (v/v) glycerol, pH 8.2). Protein concentration was determined according to Bradford (Bradford, 1976) using a linear concentration range of bovine serum albumin as a standard.

3.6 SAMS peptide kinase activity assay

A widely used substrate for SNF1 kinase is the SAMS peptide, which is a modified version of the sequence around the AMPK target site in rat acetyl-CoA carboxylase (HMRSAMSGLHLVKRR). SAMS peptide kinase activity in *P. sativum* and *V. faba* protein extracts isolated from different tissues was measured as previously described in Davies et al., (1989).

Protein extract (5 µl) was mixed with 5 µl kinase buffer (50 mM HEPES, 50 mM sodium fluoride, 1 mM DTT, pH 7.0), 5 µl sterile water, 5 µl SAMS peptide stock solution (200 µM) and 5 µl labeled ATP stock solution (1 mM [γ ^{32}P]ATP, 1mM "cold" ATP, 25 mM magnesium chloride). The samples were incubated for 30 min at 30°C and 10 µl aliquots were spotted onto an 1 cm^2 phosphocellulose P81 paper piece. The paper pieces were washed twice in 1% (v/v) phosphoric acid for 4 min and then once in acetone and dried. The pieces of paper were transferred to a scintillation vial containing scintillation liquid "Rotiszint eco plus" and counted by liquid scintillation counter. Activity was expressed as nmol phosphate incorporated into peptide per minute per mg of an initial plant tissue.

3.7 *Agrobacterium tumefaciens* mediated transformation of *N. tabacum* and *P. sativum*

Transformation of tobacco plants

0.5 ml over night culture of *A. tumefaciens* containing the appropriate vector for plant transformation were diluted with 10 ml LS medium. From sterile *N. tabacum* leaves were cutted 0.5 cm^2 pieces and incubated in the bacterial culture in darkness for 2 days at 26°C. For callus induction the explants were washed overnight with LS medium, containing 0.5 g/L claforan and placed onto a T1 medium. After 3-4 weeks *in vitro* cultivation the explants with callus formation were transferred to T2 medium for shoot development. Regenerated seedlings were cut and placed on T3 medium until roots were formed.

Transformation of pea plants

The pea transformation procedure was modified from previously described by Schroeder et al., (1993). The immature pods with seeds at stage of two to five days beyond arising the maximum of fresh weight were harvested, sterilized in 70% (v/v) ethanol for 15 min followed by 1% (w/v) sodium hypochloride for 15 min and finally five time washed with sterile distilled water. Explants for transformation were cutted with razor blade into longitudinal segments from the embryonic axis of seeds. Then segments were fully immersed in the Agrobacterial suspension for 30-40 min. Wet segments were plated on solid B5 medium and cultivated at RT for 4 days. After co-cultivation, explants were washed five times with B5 medium and placed onto solid P1 medium for callus induction. After 15 days, the explants were transferred to MS4 medium for shoot development. Primary shoots, presumed to have arisen from pre-existing meristems, were removed and discarded. The explants were

transferred to fresh MS4 medium every 20 days for 3 or 4 passages and any developing shoots were excised and transferred to MS1 medium.

After 6 - 12 weeks of cultivation, shoots longer than 0.5 cm were grafted on 5-7 days old etiolated root stocks of *P. sativum*. For this reason, above 1.5-2 cm parts of the epicotyls were removed from the cotyledons and the remaining stacks were split with a razor-blade approx. 0.5 cm deeply. The branch for grafting was cutted off, wedge-shaped sharpened and placed between the splitted epicotyls of the seedlings. A drop of water was applied to the cutted stem to prevent its drying out. The grafted seedlings were cultivated on 1/2 MS medium. During the following days and weeks the developing shoots from the cotyledonar nodium were removed.

Grafted plants were planted into autoclaved pots, accustomed gradually to a low air humidity under transparent cover and cultivated in the growth cabinets up to complete removing of the cover.

3.8 Transient expression assay in *Nicotiana plumbagenifolia* and *Arabidopsis thaliana* protoplasts

The mesophyll protoplasts from *N. plumbagenifolia* (L.) were prepared as previously described (Takebe et al., 1968). The protoplasts from embryogenic suspension cultures of *A. thaliana* (L.) Heynh were prepared and transformed according to protocol described for tobacco protoplasts (Reidt et al., 2000). Transformation was performed without heat shock. 10 µg of the appropriate plasmid DNAs (pGUS, pSnRK_S:GUS, pSnRK_L:GUS or p35S:GUS) was used for transformation. GUS protein was assayed with GUS-Light Assay Kit. The assay conditions were also described previously (Reidt et al., 2000).

3.9 Cytological observation of pollen grains

Mature anthers were collected from transgenic plants grown in a greenhouse. Pollen grains were released on a glass slide by gently squashing of anthers with fine forceps in a drop of staining solution. Anther debris were carefully removed and cover slip was applied over the staining solution containing the pollen grains before examination under a microscope.

As staining solution was used 1% acetocarmine in 100 ml of 45% (v/v) acetic acid to check the developmental stage of the pollen or iodine-potassium iodide (1g I_2 and 2 g KI dissolved in 300 ml of distilled water) to check starch accumulation.

3.10 Generation of ESTs

3.10.1 Preparation of a pea cDNA library

375 µg of total RNA were collected from cotyledons as well as seed coat to isolate $polyA^+$ - RNA. The pBluescript II XR cDNA Library Construction Kit was used for cDNA construction following the manufacturer's protocol. The cDNA fragments have been directionally cloned into *EcoRI/XhoI* cloning sites. Transfomants were picked and kept in 384-well-microtiter-plate format. Commercial sequencing of 5538 cDNA clones was performed by the company QIAGEN and sequencing of 7680 cDNA clones was done at the IPK Gatesleben. Sequences were trimmed from the 3'-end using the DNASTAR software. After vector trimming and removal of polyA stretches, sequences shorter than 100 bases were discarded. Sequence lengths were restricted to 700 bases by cutting extra sequences from the 5' end of the sequenced cDNA.

3.10.2 EST clustering and assembly analysis

The software StackPACK v 2.1.1 was applied with default parameters for EST clustering and assembly analyses, which perform subsequent steps of masking, clustering, sequence assembly, alignment analysis and consensus partioning (Christoffels et al., 2001). The default parameters were used unless otherwise stated. The masking step is performed mainly to mask remaining vector artifacts, bacterial genomic sequences, repetitive sequences as simple repeats and plant sequences from RepBase5.02 (Jurka, 1998). The loose clustering approach, d2_cluster, was employed to identify rapidly those sequences with relative similarity by counting matching n-lenth words (n=6) and to assign a cluster ID to them. The resulting cluster could capture all transcript variants as well as "contaminating" sequences that could represent chimeric clones. Sequences that show no significant similarity to other ESTs under the defaulted parameter are defined as singleton. Subsequent alignment assemblies have been generated by using PHRAP (Gordon et al., 1998), which detects the sequence diversities and separates them as different contigs within the same clusters. Further alignment analysis is

carried out with CRAW (Chou et al., 1999; Miller et al., 1999). The longest consensus was assigned to a given contig as the primary consensus, while sub-consensus sequences, if exist, were captured and recorded within each contig as alternative consensus, which might originate from sequence variants such as alternate expression forms, chimera clones or other sequence artifacts. Finally, sequences that originated from the same cDNA clone were traced and corresponding clusters were linked (Christoffels et al., 2001). All primary consensus sequences, alternative consensus sequences and sequences of singletons were collected and nominated as TUCs (tentative unigene consensi) representing a unigene set. The StackPACK software assigned systematically distinguishable IDs (i.e. cluster ID, contig ID and consensus ID) during each step of the analyses, but kept the original EST IDs for the singleton IDs. Concatenated letter strings of "cluster ID", "contig ID" and "consensus ID" were defined as IDs for multi-sequence consensi (i.e. primary and alternative consensi). The IDs for singletons and multi-sequence consensi were adapted as TUC IDs. A flat sequence file including all pea TUCs is provided as supplemental data. An actual TUC ID for multi-sequence consensi may be very long. Since the consensus IDs are also distinctive, they were kept as identifiers for the analysis below. For example, we will refer "cl18707ct10336cn11787" as "cn11787".

3.10.3 Sequence comparison and functional classification

Annotation of individual ESTs and TUCs was performed by BLASTX2 search (Altschul et al., 1990) against the NRPEP protein database using the HUSAR software package. Searches were performed using default parameters.

Functional gene annotations were performed by finding similarities between raw EST nucleotide sequences and corresponding homologous sequence from a single database or multiple databases. The first step was the mapping of an translated EST sequence by direct sequence homology recognition in annotated protein databases. Several tools are offered to build catalogs of functional EST annotations, including tools that map ESTs to metabolic functions in KEGG (Kanehisa and Goto, 2000) or to map ESTs to functional terms in the Gene Ontology Database (Gene Ontology Consortium, 2001).

To combine the wide data spectrum of integrated databases with the advantage of a homogeneous, central entry point of homology searches, it is necessary to treat the integrated database as a virtual union of inter-linked data items. These interlinked data build so called *data linkage graphs* (DLGs). The recursive breadth-first search algorithm was applied for

extraction of DLGs (Dijkstra, 1959), which was basis for in-house developed algorithm. The algorithm was implemented in C++ and the resulting graphs are stored in a Oracle9i rDBMS.

3.10.4 UniGene set selection

The in-house developed Perl script driven on the CR-EST server (http://pgrc.ipk-gatersleben.de/cr-est/) was used for extraction of the UniGene set. The UniGene selection algorithm is based on contig assemblies, BLASTX results (E-value, score, identity percentage) as well informations about sequence length, direction and quality. All analysed EST sequences were derived from directional clones, with the result that a given sequence could be classified as originating from the 5' (forward sequence) or 3' (reverse sequence) end of the relevant cDNA. Most of our clones start at the 3'-end. In rare cases some clones casually had an another orientation. Orientation was defined based on the most probable ORF and the presence of a corresponding BLASTX hit.

In many instances the inclusion of contigs or singletons in the UniGene set was straightforward: clones for singletons that linked to no other contig or singletons were included. Contigs that linked to no other contig, including contigs that linked to singletons, were designated as isolated contigs and were also included.

Another type of genes added to the UniGene set from complex clusters with more than one consensus sequence. All complex clusters were checked up on presence of an external linking between non-redundant singletons and clusters. If no link exists, the E-value of BLASTX results was checked and the EST with the best score and at least E-value <=1E-6 was included in the UniGene set. In some cases, no BLASTX result was available or the probability of the description was doubtful. In that case, the longest sequence was selected from contig and added to the UniGene set. If link exists, it can be assessed in the context of contigs and number of links can be calculated. Sequence with an individual link was directly added to the UniGene set. Relevant to this analysis, when number of links was more than one, the longest sequence from contig was choose as preferential for the UniGene. The UniGene list was used as a basis preparation of cDNA macroarray filters.

3.11 cDNA macroarray hybridization

3.11.1 Preparation of cDNA macroarray filters

E. coli clones with cDNA fragments of interest were grown overnight in LB medium with selective antibiotic, plasmid DNA was purified via the QIAGEN Plasmid Kit and electrophoresed on an 0.8% agarose gel to confirm DNA quality.
Inserts of 4,500 cDNA clones were amplified using the M13-21PE and MVR-26 primers. The PCR products were purified and concentrated via the Mini Elute 96 VF PCR Purification Kit. PCR products showing only one DNA fragment were chosen for spotting. To get final concentration 60-100 ng, DNA was diluted in a solution containing 5 M NaCl, 1 M NaOH and 1% bromphenol blue and spotted with the MicroGrid robot onto 10x13 cm Biodyne® B nylon filter. Spotting was done from 384 well plates. 0,4 mm pins were used to spot 160 nl of DNA per spot in a 5x5 pattern.

3.11.2 RNA extraction and probe synthesis

Total RNA was extracted from tissues using the Gentra RNA Isolation Kit. The synthesis of ^{33}P-labeled cDNA and the following hybridization procedure were performed as described by Sreenivasulu et al., 2002. Poly A^+-RNA was extracted from 50 µg of total RNA using oligo (dT)-paramagnetic beads according to the manufacturer's recommendation. Poly A^+-RNA attached to the beads was directly used for the synthesis of a covalently bound first strand cDNA using reverse transcriptase Superscript II. ^{33}P-labeled second strand cDNA was obtained by a random prime reaction using the Megaprime Labeling Kit according to the instruction of the supplier, except that 10 units instead of 1 unit of Klenow polymerase were used.

3.11.3 Array hybridization

Membranes were wetted in 2xSSC and pre-hybridized for at least 3h at 65°C in Church buffer (See chapter 3.4.) containing sheared salmon sperm DNA (0.1 mg/ml). The labeled cDNA was denaturated at 95°C for 3 min and added to filter together with fresh hybridization buffer. The hybridization was carried out at 65°C for at least 14 h. After hybridization, cDNA arrays were washed three times with 40 mM sodium phosphate (pH 7.2), 1% w/v SDS, 2 mM EDTA for 20 min at 65°C, wrapped in Saran wrap and exposed to an image plate of the Fuji

BAS2000 phosphoimager for 3-6 h. Before re-use of an array, the hybridization signals were removed by successive treatment with 0.4 M NaOH for 15 min at 45°C followed by a neutralization in 0.1xSSC, 0.1% (w/v) SDS, 0.2 M Tris, pH 7.5. Successful removal of the labeled probe was controlled by an overnight exposure of the array filter.

3.11.4 Array evaluation

The image data obtained from the phosphoimager were imported into the program Array Vision for spot detection and quantification of hybridization signal intensities. After export of the determined values to a common spreadsheet program, the signals intensities were normalized. Primary statistical processing and normalization of signals in view of a background, reliability between two spots and comparison of results between repeats was performed with free software R as a statistics system and Excel program with standard statistical package. Local background was subtracted from spot intensities and signal intensity of the duplicated spots for each cDNA fragment was averaged. To allow comparison of signal intensities across experiments, the median of the logarithmically scaled intensity distribution for each experiment was set to zero (median centering of arrays, Eisen et al., 1998). Thereafter, the logarithmically scaled signal intensities of each gene was centered by its median across all experiments (median centering of genes, Eisen et al., 1998), which emphasize the differential expression of genes irrespective of absolute signal intensities. Since medial centring of genes does not yield information about signal intensity, we used the data after the first round of median centring of arrays to calculate non-logarithmic, normalized signal intensities. These normalized signal intensities were used to exclude genes with low signal intensities across all experiments.

3.11.5 Data filtering and clustering algorithmics

For further analysis of array in pea seed development, the complete dataset was reduced to cDNA fragments, which show differential expression across the experiments. Those cDNA fragments were selected based on following criteria: To exclude cDNA fragments with signals always close to the background the normalized non-logarithmic signal intensity should be exceed 3 arbitrary units (au) for at least one experiment. The ratio between the maximal and the minimal intensity of signal for a cDNA fragment across all experiments had to exceed the factor of 10. The data set filtered in this way comprised 131 cDNA fragments, which fulfil

this criteria and showing 3-fold (34% from all selected profiles) and 4-fold (66% from all selected profiles) differences in temporal expression profiles.

Cluster analysis was carried out using the program J-Express. The goal of clustering is to group together genes with similar properties. Proximities were measured by the Manhattan distance. Hierarchical clustering was preformed using the unweighted pair group method with arithmetic averages (UPGMA). The temporal expression patterns of genes were analyzed by applying the k-means clustering algorithm, which operates over a fixed number of clusters.

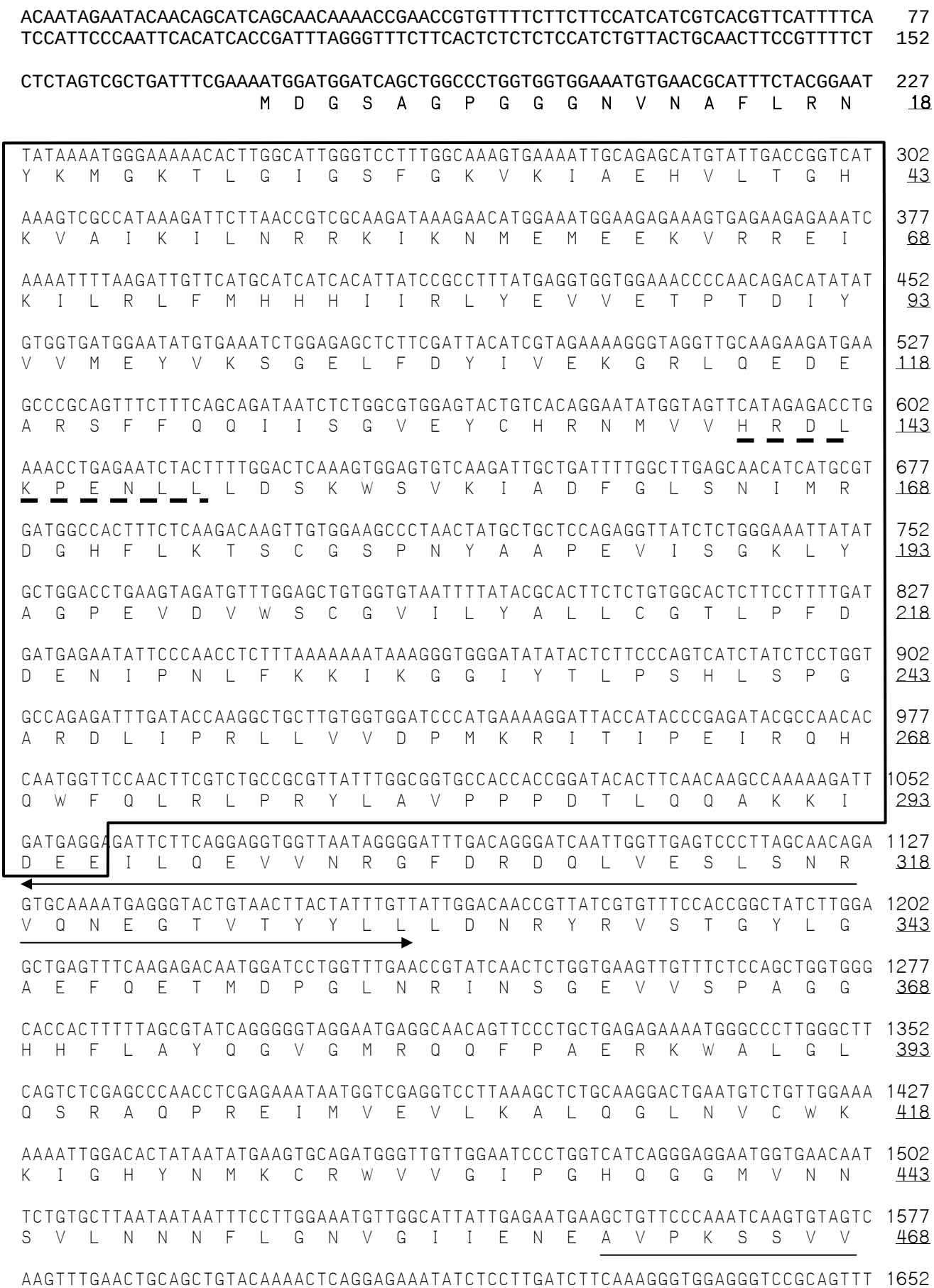
4. RESULTS

4.1. Isolation and analysis of cDNA clones encoding sucrose-non-fermenting related protein kinase (SnRK1) from *V. faba* seeds

For the identification of SnRK1 genes from *V. faba* plants the following strategy was applied. Firstly, primer pairs were designed based on the conserved region of known plant SNF1-related kinases (BLASTX search, Altschul et al., 1997). A 200-bp fragment was amplified from a *V. faba* seed cDNA library used as a template. After sequence verification, this fragment was applied as a probe to screen the same cDNA to clone a full length cDNA. 12 clones were isolated. They were characterized by DNA sequencing and three clones were identified as full-length SNF1-related clones of the same gene after comparison with other sequences in the various databases. No other SNF1-related cDNA fragments were found in this search.

The longest cDNA sequence comprised a total of 2076 nucleotides, including 174 and 372 nucleotides for 5' and 3' untranslated regions, respectively. It encodes a protein of 509 amino acids with a calculated molecular mass of 57.9 kDa. The predicted VfSnRK1 protein sequence showed homology of 84.5% to the SNF1p of *Glycine max* (Patil et al., 1999), 81.7% to the SNF1p of *L. esculentum* (Bradford et al., 2003), 81.1% to the *N. tabacum* NPK5 (Muranaka et al., 1994), 30.3% to the SNF1-like *Homo sapiens* kinase (Ruiz et al., 1994) and 41.5% to the SNF1 of *S. cerevisiae* (Celenza and Carlson, 1986).

The putative kinase catalytic domain of VfSnRK1 resides in the N-terminal part of the protein (Fig. 5, 6) and contains the 11 subdomains common to protein kinases (Hanks et al., 1988). Conserved subdomain VIII possess a Thr –175, which can be phosphorylated. Mutation of this residue to alanine in yeast results in a complete loss of *SNF1* gene function (Estruch et al., 1992). The UBA domain (Ubiquitin binding domain) is immediately adjacent to the kinase domain (position 293-332). The UBA-domain is a novel sequence motif found in several proteins having connections to ubiquitin and the ubiquitination pathway. This domain is probably a non-covalent ubiquitin binding domain consisting of a compact three helix bundle (Hofmann et al., 1996).


ACAATAGAATACAACAGCATCAGCAACAAAACCGAACCGTGTTTTCTTCTTCCATCATCGTCACGTTCATTTTCA 77
TCCATTCCCAATTCACATCACCGATTTAGGGTTTCTTCACTCTCTCTCCATCTGTTACTGCAACTTCCGTTTTCT 152
CTCTAGTCGCTGATTTCGAAAATGGATGGATCAGCTGGCCCTGGTGGTGGAAATGTGAACGCATTTCTACGGAAT 227
M D G S A G P G G G N V N A F L R N 18
TATAAAATGGGAAAAACACTTGGCATTGGGTCCTTTGGCAAAGTGAAAATTGCAGAGCATGTATTGACCGGTCAT 302
Y K M G K T L G I G S F G K V K I A E H V L T G H 43
AAAGTCGCCATAAAGATTCTTAACCGTCGCAAGATAAAGAACATGGAAATGGAAGAGAAAGTGAGAAGAGAAATC 377
K V A I K I L N R R K I K N M E M E E K V R R E I 68
AAAATTTTAAGATTGTTCATGCATCATCACATTATCCGCCTTTATGAGGTGGTGGAAACCCCAACAGACATATAT 452
K I L R L F M H H H I I R L Y E V V E T P T D I Y 93
GTGGTGATGGAATATGTGAAATCTGGAGAGCTCTTCGATTACATCGTAGAAAAGGGTAGGTTGCAAGAAGATGAA 527
V V M E Y V K S G E L F D Y I V E K G R L Q E D E 118
GCCCGCAGTTTCTTTCAGCAGATAATCTCTGGCGTGGAGTACTGTCACAGGAATATGGTAGTTCATAGAGACCTG 602
A R S F F Q Q I I S G V E Y C H R N M V V H R D L 143
AAACCTGAGAATCTACTTTTGGACTCAAAGTGGAGTGTCAAGATTGCTGATTTTGGCTTGAGCAACATCATGCGT 677
K P E N L L L D S K W S V K I A D F G L S N I M R 168
GATGGCCACTTTCTCAAGACAAGTTGTGGAAGCCCTAACTATGCTGCTCCAGAGGTTATCTCTGGGAAATTATAT 752
D G H F L K T S C G S P N Y A A P E V I S G K L Y 193
GCTGGACCTGAAGTAGATGTTTGGAGCTGTGGTGTAATTTTATACGCACTTCTCTGTGGCACTCTTCCTTTTGAT 827
A G P E V D V W S C G V I L Y A L L C G T L P F D 218
GATGAGAATATTCCCAACCTCTTTAAAAAAATAAAGGGTGGGATATATACTCTTCCCAGTCATCTATCTCCTGGT 902
D E N I P N L F K K I K G G I Y T L P S H L S P G 243
GCCAGAGATTTGATACCAAGGCTGCTTGTGGTGGATCCCATGAAAAGGATTACCATACCCGAGATACGCCAACAC 977
A R D L I P R L L V V D P M K R I T I P E I R Q H 268
CAATGGTTCCAACTTCGTCTGCCGCGTTATTTGGCGGTGCCACCACCGGATACACTTCAACAAGCCAAAAAGATT 1052
Q W F Q L R L P R Y L A V P P P D T L Q Q A K K I 293
GATGAGGAGATTCTTCAGGAGGTGGTTAATAGGGGATTTGACAGGGATCAATTGGTTGAGTCCCTTAGCAACAGA 1127
D E E I L Q E V V N R G F D R D Q L V E S L S N R 318
GTGCAAAATGAGGGTACTGTAACTTACTATTTGTTATTGGACAACCGTTATCGTGTTTCCACCGGCTATCTTGGA 1202
V Q N E G T V T Y Y L L D N R Y R V S T G Y L G 343
GCTGAGTTTCAAGAGACAATGGATCCTGGTTTGAACCGTATCAACTCTGGTGAAGTTGTTTCTCCAGCTGGTGGG 1277
A E F Q E T M D P G L N R I N S G E V V S P A G G 368
CACCACTTTTTAGCGTATCAGGGGGTAGGAATGAGGCAACAGTTCCCTGCTGAGAGAAAATGGGCCCTTGGGCTT 1352
H H F L A Y Q G V G M R Q Q F P A E R K W A L G L 393
CAGTCTCGAGCCCAACCTCGAGAAATAATGGTCGAGGTCCTTAAAGCTCTGCAAGGACTGAATGTCTGTTGGAAA 1427
Q S R A Q P R E I M V E V L K A L Q G L N V C W K 418
AAAATTGGACACTATAATATGAAGTGCAGATGGGTTGTTGGAATCCCTGGTCATCAGGGAGGAATGGTGAACAAT 1502
K I G H Y N M K C R W V V G I P G H Q G G M V N N 443
TCTGTGCTTAATAATAATTTCCTTGGAAATGTTGGCATTATTGAGAATGAAGCTGTTCCCAAATCAAGTGTAGTC 1577
S V L N N N F L G N V G I I E N E A V P K S S V V 468
AAGTTTGAACTGCAGCTGTACAAAACTCAGGAGAAATATCTCCTTGATCTTCAAAGGGTGGAGGGTCCGCAGTTT 1652

```
K  F  E  L  Q  L  Y  K  T  Q  E  K  Y  L  L  D  L  Q  R  V  E  G  P  Q  F    493

CTTTTCTTGGATCTTTGTGCTGCCTTCCTTGCGCAGCTTCGTGTTCTTTAATACAAAGATAGTACACACTCACAT 1727
L  F  L  D  L  C  A  A  F  L  A  Q  L  R  V  L  *                          509

GGAGGACCTGACCTGAGGTGGTCCTATTGTGCATGTGAATAAGATCATGTTGTACATATNATTGCTGCTCTTTCT 1802
GGGTTGACCTTATGTTTGTACTTTTAGTAATACCCTCAGTTTTTAATATCTGGGTAATCATGATTATTATGAGGC 1877
AATCCTAAATATGGTATTTGCTTCCCCTTTTCACAGAGAATTGAGTTTTTCAATCCTGTGAAACATGTGAGAATT 1952
TTATGGAAGAAAACATCCCCCAACAACAGATACAGTTATTTATGGTATTGAACCTAGTGTTATATTGACTATCTA 2027
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA                             2076
```

Figure 5. cDNA sequence of *VfSnRK1* and deduced amino acid sequence shown in single-letter code above the respective coding nucleotide sequence. Nucleotide number as well as amino acid number assignment is underlined at the right. The protein serine/threonine kinase catalytic domain is boxed. The protein kinase ATP-binding region signature is underscored with a dotted line. The ubiquitin associated domain UBA (smart00165; Hofmann et al., 1996) is indicated with arrows. The Kinase Associated domain KA1 (CDART; PF02149) is underlined. The asterisk indicates termination of the protein-coding region.

```
                                                I
  1  M D R S T G R G G G G S V D M F L R N Y K L G K T L G I G S F G  Glycine max
  1  M D G S A G P G G G - N V N A F L R D Y K M G K T L G I G S F G  Vicia faba
  1  M D G S T V Q G G S - S V E S F L R N Y K L G K T L G I G S F G  Nicotiana tabacum NPK5
  1  M D M N - - - - - - - - - - - V P N Y K L G K T L G I G S F G  Cucumis sativus
  1  M D G S G T G S R S - G V E S I L P N Y K L G R T L G I G S F G  A. thaliana KIN10
     * *                             . * * . * . * * * * * * * *

                                        II                  III
 33  K V K I A E H V R T G H K V A I K I L N R H K I K N M E M E E K  Glycine max
 32  K V K I A E H V L T G H K V A I K I L N R R K I K N M E M E E K  Vicia faba
 32  K V K I A E H T L T G H K V A V K I L N R R K I K N M E M E E K  Nicotiana tabacum NPK5
 21  K V K I A E H A L T G H K V A I K I L N R R K I K N L D M E E K  Cucumis sativus
 32  R V K I A E H A L T G H K V A I K I L N R R K I K N M E M E E K  A. thaliana KIN10
     . * * * * * *   * * * * * * . * * * * * * * * * * . . * * * *

                    IV                                  V
 65  V R R E I K I L R L F M H H H I I R L Y E V V E T P T D I Y V V  Glycine max
 64  V R R D I K I L R L F M H H H I I R L Y E V V E T P T D I Y V V  Vicia faba
 64  V R R E I K I L R L F M H P H I I R L Y E V V E T P S D I Y V V  Nicotiana tabacum NPK5
 53  V R R E I K I L R L F M H P H I I R L Y E V I E T P S D I Y V V  Cucumis sativus
 64  V R R E I K I L R L F M H P H I I R L Y E V I E T P T D I Y L V  A. thaliana KIN10
     * * * . * * * * * * * * * * * * * * * * * * * . * * * . * * *  *

                                   VIa
 97  M E Y V K S G E L F D Y I V E K G R L Q E D E A R H F F Q Q I I  Glycine max
 96  M E Y V K S G E L F D Y I V E K G R L Q E D E A R S F F Q Q I I  Vicia faba
 96  M E Y V K S G E L F D Y I V E K G R L Q E D E A R K F F Q Q I I  Nicotiana tabacum NPK5
 85  M E Y V K S G E L F D Y I V E K G R L Q E D E A R N F F Q Q I I  Cucumis sativus
 96  M E Y V N S G E L F D Y I V E K G R L Q E D E A R N F F Q Q I I  A. thaliana KIN10
     * * * *  * * * * * * * * * * * * * * * * * * * * * * * * * *

                    VIb                                 VII
129  S G V E Y C H R N M V V H R D L K P E N L L L D S K F N I K I A  Glycine max
128  S G V E Y C H R N M V V H R D L K P E N L L L D S K W S V K I A  Vicia faba
128  S G V E Y C H R N M V V H R D L K P E N L L L D S K W N V K I A  Nicotiana tabacum NPK5
117  S G V E Y C H R N M V V H R D L K P E N L L L D S K C N V K I A  Cucumis sativus
128  S G V E Y C H R N M V V H R D L K P E N L L L D S K C N V K I A  A. thaliana KIN10
     * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *    . * * *

                                  VIII
161  D F G L S N I M R D G H F L K T S C G S P N Y A A P E V I S G K  Glycine max
160  D F G L S N I M R D G H F L K T S C G S P N Y A A P E V I S G K  Vicia faba
160  D F G L S N I M R D G H F L K T S C G S P N Y A A P E V I S G K  Nicotiana tabacum NPK5
```

```
149  D F G L S N I M R D G H F L K T S C G S P N Y A A P E V I S G K  Cucumis sativus
160  D F G L S N I M R D G H F L K T S C G S P N Y A A P E V I S G K  A. thaliana KIN10
     * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

                    IX                                          X
193  L Y A G P E V D V W S C G V I L Y A L L C G T L P F D D E N I P  Glycine max
192  L Y A G P E V D V W S C G V I L Y A L L C G T L P F D D E N I P  Vicia faba
192  L Y A G P E V D V W S C G V I L Y A L L C G T L P F D D E N I P  Nicotiana tabacum NPK5
181  L Y A G P E V D V W S C G V I L Y A L L C G T L P F D D E N I P  Cucumis sativus
192  L Y A G P E V D V W S C G V I L Y A L L C G T L P F D D E N I P  A. thaliana KIN10
     * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

                                XI
225  N L F K K I K G G I Y T L P S H L S P G A R D L I P R M L V V D  Glycine max
224  N L F K K I K G G I Y T L P S H L S P G A R D L I P R L L V V D  Vicia faba
224  N L F K K I K G G M I S L P S H L S A G A R D L I P R M L I V D  Nicotiana tabacum NPK5
213  N L F K K I K G G I Y T L P S H L S S G A R E L I P S M L V V D  Cucumis sativus
224  N L F K K I K G G I Y T L P S H L S P G A R D L I P R M L V V D  A. thaliana KIN10
     * * * * * * * * *     . * * * * * *   * * * . * * *   . * . * *

257  P M K R M T I P E I R Q H P W F Q V H L P R Y L A V P P P D T L  Glycine max
256  P M K R I T I P E I R Q H Q W F Q L R L P R Y L A V P P P D T L  Vicia faba
256  P M K R M T I P E I R M H P W F Q A H L P R Y L A V P P P D T M  Nicotiana tabacum NPK5
245  P M K R I T I P E I R Q H P W F Q A H L P R Y L A V P P P D T M  Cucumis sativus
256  P M K R V T I P E I R Q H P W F Q A H L P R Y L A V P P P D T V  A. thaliana KIN10
     * * * *   * * * * * * *   *   * * *   * * * * * * * * * * * *

289  Q Q A K K I D E E I L Q E V V N M G F D R N Q L V E S L S N R I  Glycine max
288  Q Q A K K I D E E I L Q E V V N R G F D R D Q L V E S L S N R V  Vicia faba
288  Q Q A K K I D E D I L Q E V V K R G F D R N S L V A S L C N R V  Nicotiana tabacum NPK5
277  Q Q A K K I D E D I L Q E V V K M G F D R N Q L V E S L R N R I  Cucumis sativus
288  Q Q A K K I D E E I L Q E V I N M G F D R N H L I E S L R N R T  A. thaliana KIN10
     * * * * * * * * . * * * * * .     * * * * .   *     * *   * *


321  Q N E G T V T Y Y L L L D N R F R V S S G Y L G A E F Q E T M D  Glycine max
320  Q N E G T V T Y Y L L L D N R Y R V S T G Y L G A E F Q E T M D  Vicia faba
320  Q N E G T V A Y Y L L L E N Q F R A S S G Y M G A E F Q E T M E  Nicotiana tabacum NPK5
309  Q N E A T V A Y Y L L L D N R F R V S S G Y L G A E F Q E T M E  Cucumis sativus
320  Q N D G T V T Y Y L I L D N R F R A S S G Y L G A E F Q E T M E  A. thaliana KIN10
     * * .   * *   * * * . * . *   .   *   * . * * . * * * * * * * .

353  S G F N R M H S G E V A S P V V G H H S T G Y M D Y Q G V G M R  Glycine max
352  P G L N R I N S G E V V S P A G G H H - - - F L A Y Q G V G M R  Vicia faba
352  Y G Y H Q I N S S E - V L L P C W Q H L P G I M D F Q Q V G A R  Nicotiana tabacum NPK5
341  T G F N R M H P S D P T N P A V G H R L P G Y M D Y Q G M G L R  Cucumis sativus
352  G - T P R M H P A E S V A S P V S H R L P G L M E Y Q G V G L R  A. thaliana KIN10
                       .                       .     .*     *   *

385  Q Q F P V E R K W A L G L Q S R A Q P R E I M T E V L K A L Q E  Glycine max
381  Q Q F P A E R K W A L G L Q S R A Q P R E I M V E V L K A L Q G  Vicia faba
383  - Q F P V E R K W A L G L Q S R A H P R E I M T E V L K A L Q G  Nicotiana tabacum NPK5
373  A Q F P V E R K W A L G L Q S R A H P R E I M T E V L K A L R E  Cucumis sativus
383  S Q Y P V E R K W A L G L Q S R A H P R E I M T E V L K A L Q D  A. thaliana KIN10
       * .*   * * * * * * * * * * * * * *   * * * * *   * * * * * *

417  L N V C L E E D W T L Y H E C R W V A G T A G H H - E M I N N S  Glycine max
413  L N V C W K K I G H Y N M K C R W V V G I P G H Q G G M V N N S  Vicia faba
414  L N V R W K K I G P Y N M K C Q W V P G V P G H H E G M S N N S  Nicotiana tabacum NPK5
405  L N V A W K K I G H Y N M K C R W L P G I P G H H E G M I N N P  Cucumis sativus
415  L N V C W K K I G H Y N M K C R W V P - - N S S A D G M L S N S  A. thaliana KIN10
     * * *                       *   *                   *     *

448  L H S N H Y F G N D S G I I E N - E A V S K S N V V K F E V Q L  Glycine max
445  V L N N N F L G N - V G I I E N - E A V P K S S V V K F E L Q L  Vicia faba
446  I H - I Q F F G D D S T V I E N - G G V T I P N A V K F E V Q L  Nicotiana tabacum NPK5
437  V H S N H Y F G D K S T I I E N D G V V K S P N V I K F E V Q L  Cucumis sativus
445  M H D N N Y F G D E S S I I E N E A A V K S P N V V K F E I Q L  A. thaliana KIN10
                   *           * * *       *             * * *   * *
```

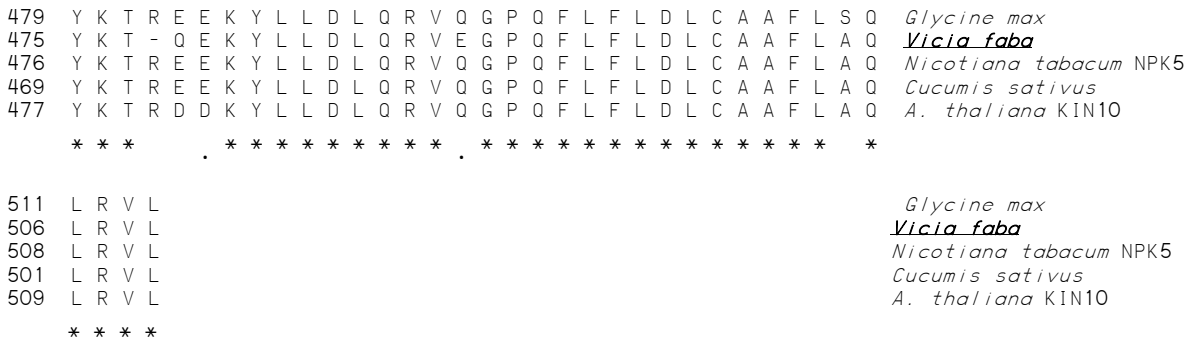


Figure 6. Amino acid sequence alignment (ClustalW method Thompson et al., 1994) of the *VfSnRK1* with other members of the SNF1 kinase family: *A. thaliana* AKIN10 (Acc. Nr Q38997 Le Guen et al., 1992), *Cucumis sativus* (Acc Nr. CAA71142 Gumpel, unpublished), *G. max* (Acc. Nr. AAD23582 Patil et al., 1999), *N. tabacum* NPK5 (Acc. Nr. A56009 Muranaka et. al, 1994). The protein kinase catalytic domains are boxed and numbered above in roman numerals. Positions with a single, fully conserved residue are indicated by asterisks; conservative substitutions are indicated by dots under the sequences.

The C-terminal putative regulatory domain of *VfSnRK1* is relatively unique. As with its yeast and animal orthologs, this regulatory domain is likely to interact with another part of the SNF1 complex. Analysis of conserved domains using CDART (Geer LY et al., 2002) indicated some conserved domains within the C-terminal part of proteins. At the end of the protein a domain exists that shows low sequence homology with kinase associated domain 1 (KA1) (position 460-509). This domain is found in the C-terminal extremity of various serine/threonine-protein kinases from fungi, plants and animals.

Thus, based on the highly conserved snf1 catalytic domain structure and comparison with known plant snf1-related protein kinases the *VfSnRK1* gene can be identified as a SnRK1 protein (Fig. 6).

The *P. sativum SnRK1* was identified from three ESTs, and one of them (PSC22B05) corresponded to a full length cDNA analogous to SnRK1 from *V. faba* encoding a protein of 57.735 kDa consists of 509 amino acid residues. The *PsSnRK1* has a domain structure similar to *VfSnRK1*. The *PsSnRK1* nucleotide sequence possesses 97.6%, and protein 98.8% identity to *VfSnRK1*. The sequences of two ESTs (PSC26L07 and PSC26F05) revealed a partial cDNA of 1113 bp and 802 bp respectively and encoded 193 and 165 from predicted 509 amino acids with the 5' upstream stop codon missing. All three ESTs belong to the same contig, are identical and encode the same cDNA.

Similarity analysis of amino acid sequences showed that *VfSnRK1* and *PsSnRK1* are most closely related to the *SnRK1* from *G. max*, StubSNF1 from *Solanum tuberosum*, NPK5 from *N. tabacum* and SNF1 from *L. esculentum.* They all together belong to the cluster of *SnRK1*s from dicot plants and are distinct from those of monocot plants (Fig. 7).

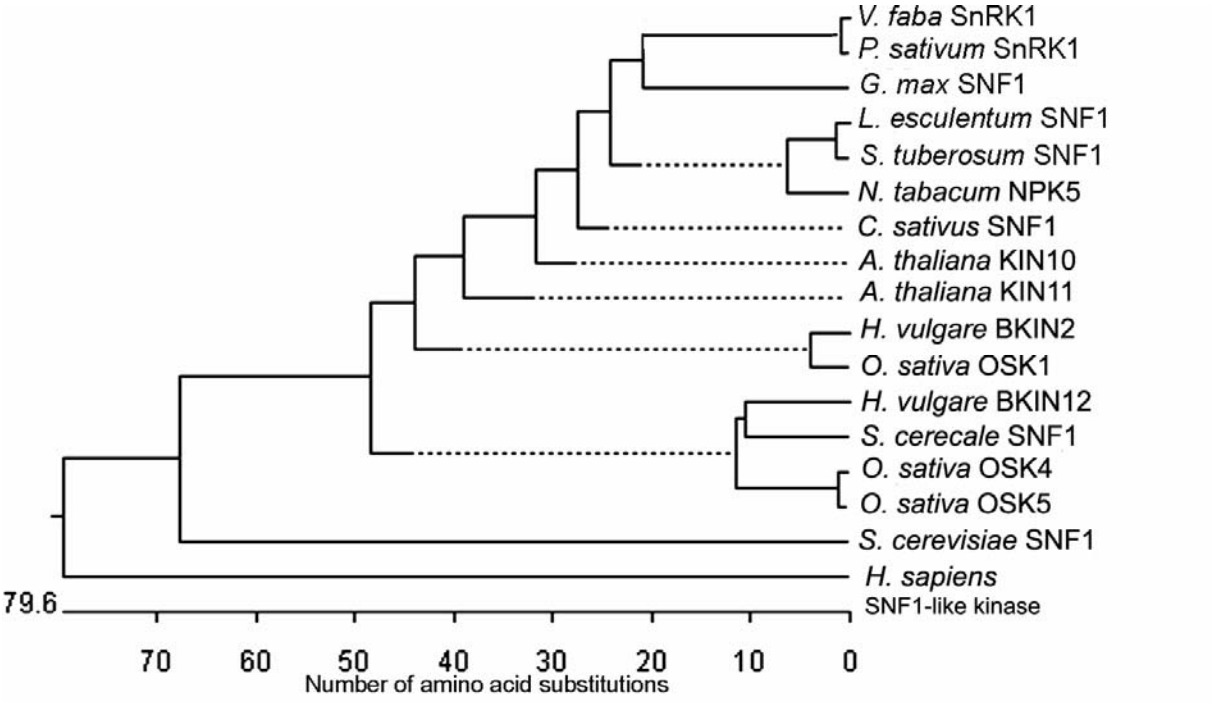


Figure 7. Phylogenetic tree showing the relationship between the *SnRK1s* sequences from higher plant as well as SNF1 from yeast and AMP-activated protein kinase from *Homo sapiens*. The tree was constructed by the neighbour-joining method using CLUSTALW (software DNAStar).

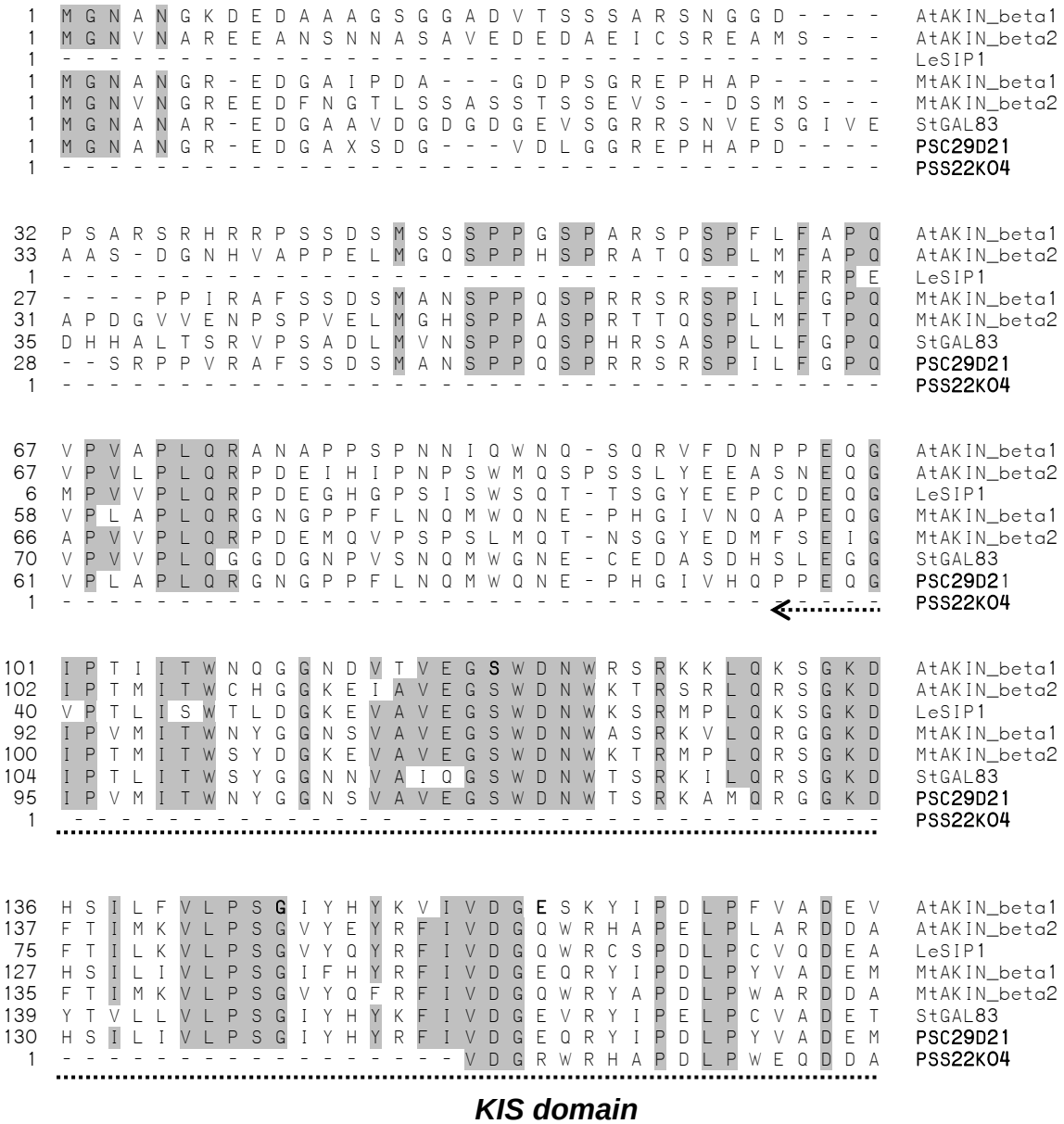
4.2. Identification of β-subunits of the SnRK1 complex in *Pisum sativum*

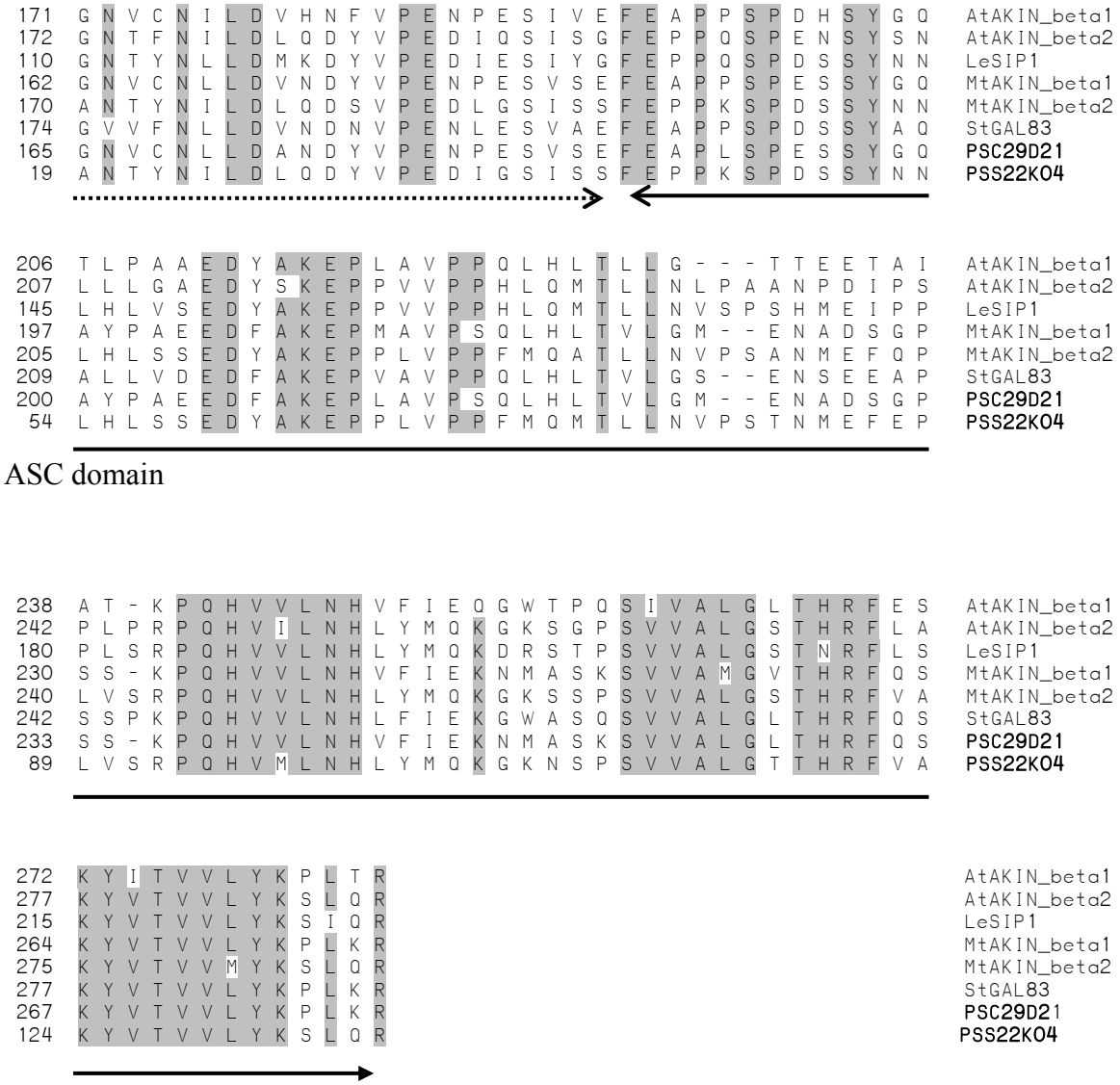
To identify β-subunit of the SnRK1 complex from *P. sativum*, the *Medicago truncatula* AKIN βγ, AKIN γ, and AKIN β1-β4 protein sequence (Acc. Nr: AY247268, AY247269, AY247271-AY247274; Buitink et al. 2004) were BLASTed against the *P. sativum* EST collection. Two ESTs encoding the full length cDNA of *PsAKINβ1* (EST ID PSC29D21) and a partial cDNA (154 from predicted 280 amino acids) with a predicted molecular weight of 30.67 kDa showed 93.1% identity AKIN β1 of *M. truncatula* (Acc. Nr. AY247271; Buitink et al., 2004), 69.3% identity with StGAL83 (Acc. Nr. AJ012215; Lakatos et al., 1999) of *S.*

tuberosum and 62.9% identity with AKIN β1 (Acc. Nr. AJ132315; Bouly et al., 1999) of *A. thaliana*.

One EST (EST ID PSS22K04) encoding a protein sequence showed similarity with the C-terminal domain of Mt*AKINβ2* (Acc. Nr. AY247272; Buitink et al., 2004). The sequence of this EST represented a partial cDNA of 770 bp with the missing 5' upstream start codon and encoded 136 from predicted 287 amino acid residues. This EST showed 89.8% identity to AKIN β2 of *M. truncatula* (Acc. Nr. AY247272; Buitink et al., 2004), 73.0% identity to LeSIP1 (Acc. Nr. AAG41995; Bradford et al., 2003) of *L. esculentum* and 73.0% identity with AKIN β2 (Acc. Nr. AJ132316; Bouly et al., 1999) of *A. thaliana*.

(A)





ASC domain

(B)

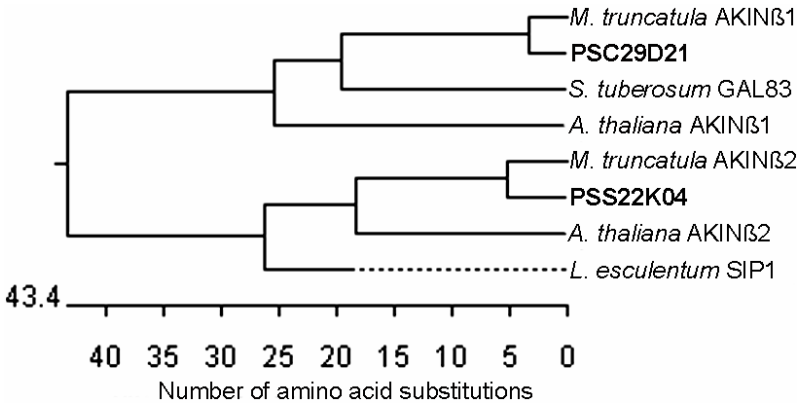


Figure 8. (A) Comparison of the deduced amino acid sequence of *P. sativum* AKIN β1 (EST ID PSC29D21) and AKIN β2 (EST ID PSS22K04) with *MtAKINβ1* (Acc. Nr. AY247271)

and MtAKIN β1 (AY247272) of *M. truncatula* (Buitink et al., 2004), *AtAKINβ1* (Acc. Nr. AJ132315) and AtAKIN β1 (Acc. Nr. AJ132316) of *A. thaliana* (Bouly et al., 1999), *LeSIP1* (Acc. Nr. AAG41995) of *L. esculentum* (Bradford et al., 2003) and *StGAL83* (Acc. Nr. AJ012215) of *S. tuberosum* (Lakatos et al., 1999). Identical amino acid residues are shaded and amino acid positions showed on the left. The conserved KIS (**ki**nase a**s**sociation) and ASC (**a**ssociation with **S**NF1 **c**omplex) domains, which in yeast bind to the catalytic subunit SNF1 and SNF4 respectively, are underlined (Jiang and Carlson, 1997).

(B) Phylogenetic tree of protein sequences represented in (A) based upon CLUSTALW analysis.

4.3 *In vivo* expression studies

4.3.1 Genomic organization of VfSnRK1

Southern blot hybridization using a *V. faba* genomic DNA was performed to determine the copy number of the *VfSnRK1* gene in the plant genome. Distinct patterns were obtained after blot hybridization at high stringency using the fragment of *VfSnRK1* cDNA as a probe. *Bam*HI digestion yielded three strongly hybridizing bands larger than expected from their restriction sites in the *VfSnRK1* cDNA (Fig.9), suggesting that the *VfSnRK1* gene has introns between predicted *Bam*HI restrictions sites. Endonuclease digestion by *Hin*dIII revealed a simple hybridization pattern with a single band. Because the *VfSnRK1* cDNA does not have a predicted *Hin*dIII restrictions site, it is suggested that the *VfSnRK1* gene has additional sites for these enzyme in its introns. Therefore, it is likely that *VfSnRK1* is present only in a single copy in the genome.

Different patterns with multiple bands were obtained by filter hybridization under low stringency conditions. These results indicate cross hybridization with other members of the SnRK family, because of high nucleotide sequence similarity (data not shown).

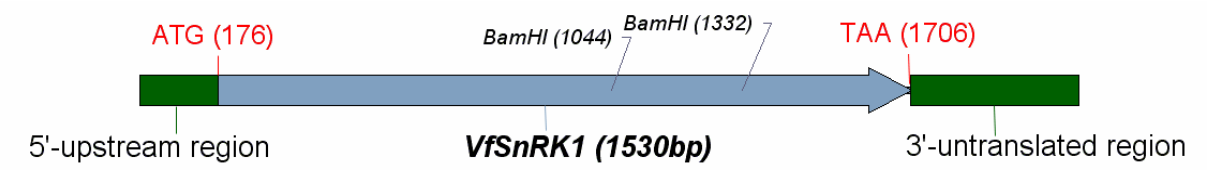


Figure 9. Schematic diagram of *VfSnRK1* cDNA. ORF (*grayed box*) is 1530 bp long. ATG and TAA are start and stop codons respectively.

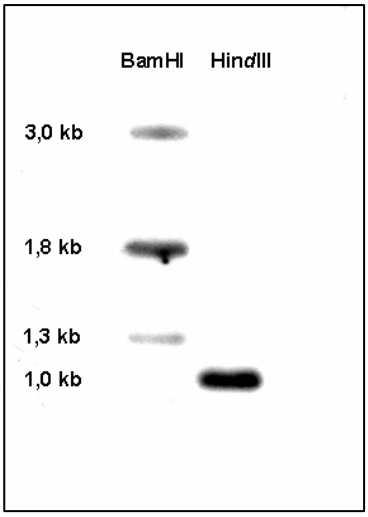


Figure 10. Southern blot analysis of *VfSnRK1* gene of *V. faba* at high stringent hybridization conditions. 10 µg of genomic DNA were digested with *Bam*HI or *Hin*dIII endonucleases, separated on 1% agarose gel and subjected to Southern blot hybridization using 200-bp ^{32}P-labeled fragment of *VfSnRK1* cDNA.

4.3.2 Tissue-specific expression and activity of SnRK1 in V. faba

Expression of *VfSnRK1* was studied in gynoecia, pods, sink- and source leaves, roots of 2 month-old plants, cotyledons and seed coat of non-germinated seeds at 16 DAP by northern-blot analysis using the SnRK1 cDNA as probe (Fig. 11). Northern blot analysis revealed *VfSnRK1* transcripts of circa 2000 bases. The transcripts were most abundant in sink-organs like roots, pods and gynoecia and hardly detectable in the seed coat .

The regulated expression of *VfSnRK1* in cotyledons of *V. faba* was determined at different stages of development. The expression level of *VfSnRK1* was highest from 19 DAP till 23 DAP but decreased after 24 DAP (Fig. 12).

VfSnRK1p kinase activity was measured using a SAMS peptide assay in 40% saturated ammonium sulfate fraction of protein extracts from developing embryos. Because the concentration of storage proteins at late stages of cotyledon development dramatically increased, it would be incorrect to express the enzyme activity on a protein basis. Therefore, the activity was calculated as nmol phosphate incorporated per minute per mg of an initial plant material. The temporal profile of enzyme activity correlated with transcripts amounts, but the peak of activity is shifted in towards two days later (Fig. 12).

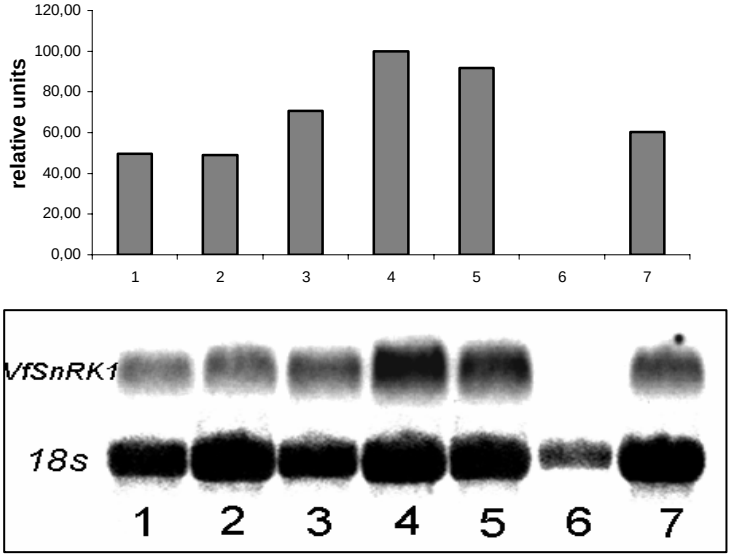


Figure 11. Expression pattern of *VfSnRK1* transcripts in different tissues of *V. faba*: 1- source leaf, 2- sink leaf, 3- pod, 4- roots, 5- gynoecia , 6- seed coat , 7- embryo (20 mg, 16 DAP). Relative mRNA abundance was quantified as described in Material and Methods, Chapter 3.4.

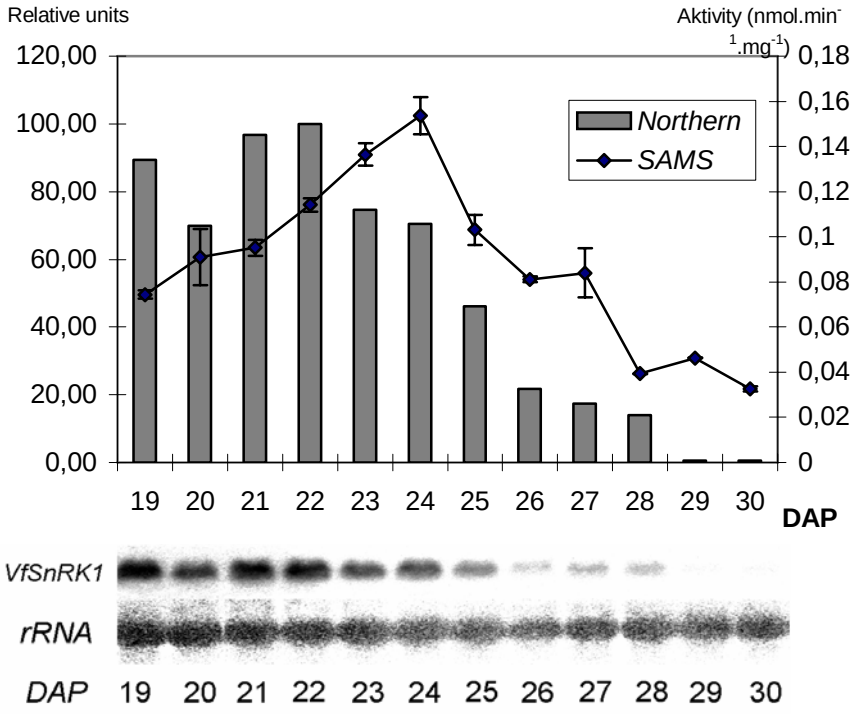


Figure 12. Comparison of *SnRK1* transcripts level and SAMS kinase activity during seed development of *V. faba*. *Bars*: the normalized signal intensity reflecting the expression level. *Line:* the kinase activity calculated as nmol phosphate incorporated per minute per mg of an initial plant material.

4.4. Isolation of the 5'-flanking sequence of the *VfSnRK1*

To better examine the structure of the *VfSnRK1* gene, a 5'-flanking region of 1.9 Kb in length was isolated by genome walking.

Immediately upstream of the translation initiation codon ATG (-9) an intron was found of 1.5 Kb length. Within the SnRK1 subfamily, few intron positions are conserved between all SnRK1 groups. SnRK1s are apparently the only one members of the CDPK-SnRK superfamily to have an intron in their 5′-untranslated region which may play a role in regulation of gene expression (Hrabak et al., 2003).

Besides the intron the following structural elements were identified in the upstream region. After excluding the intron sequence, two putative transcription initiation sites were located at -177 bp and -231 bp upstream of the ATG initiation codon. The putative TATA box is located 26 bp upstream of the second initiation codon.

A computer studies (http://www.softberry.com NSITEP-recognition of PLANT regulatory motifs with statistics and NSITEMPR-recognition of PLANT regulatory motifs conserved in several sequences) of this sequence region showed several features commonly present in DNA regulatory sequences of other genes. Short sequences with resemblance to known *cis*-acting elements were found to be located in the VfSnRK1 promoter sequence. Fig. 13 included only the ones with the highest degree of similarity to known *cis*-acting elements.

One copy similar to the sequence TTAGGGTTT was found at position –71/-63 (strand "+") (Acc. Nr. RegSite DB RSP00448). This sequence was commonly found in the promoters of the ribosomal protein (RP) L9 (*P. sativum*) genes and may have a role in regulating the coordinate production of the RP genes in plants (Moran, 2000). The same motive is also described as an inverted *telo*-box (Acc. Nr. RSP00279). It is found within the 5'- region of most plant genes encoding components of the translational apparatus, in cycling cells [e.g. PCNA (proliferating cell nuclear antigen), ribonucleotide reductase, etc.]. A number of genes encoding ribosomal proteins, also contain a *telo*-box (Manevski et *al*., 2000). At position –299/-288 (strand "-") a AATATTTTTATT motive was found, similar to those in the promoter of ribulose-1,5-bisphosphate carboxylase (*rbcS-3.6*) from pea and chlorophyll a/b binding (*cab-E*) protein from tobacco (Acc. Nr. RSP00106). This element can be recognized by AT-1 protein and is known to be regulated by light. Binding of AT-1 is modulated by phosphorylation (Cacchione et al., 1995; Datta et al., 1989).

Three copies of a conserved initiator element CATCA are located at positions –154/-149, -119/-114, -81/-76 (strand "+") upstream of ATG (Acc. Nr. RSP00512). This element is

present in both alpha-Amy1 and alpha-Amy2 genes of *Avena fatua* as well as in a number of other plant TATA-containing and TATA-less promoters (Willmott et al., 1998). The 5'-flanking region contains also two copies of CAACA-motives (positions –160/-155, -148/-143; strand "+"), which were shown to represent binding sites for RAV1 (Related to ABI3/VP1) DNA-binding protein from *A. thaliana* (Acc. Nr. RSP00308). RAV1 contains AP2 and B3-like domains. Both AP2 and B3-like domains showed autonomous DNA-binding activities specific for CAACA (Kagaya et al., 1998).

Using a method based on nested PCR, it was not possible to extend the sequence in 5'-direction further than 2.7 kb. Paired comparisons among the DNA sequence of the *VfSnRK1* promoter with sequences of plant promoter databases did not show any special pattern or similarity with any of the promoters included in the analysis (data not shown).

```
AAATAAATTATATAACATAATTAAATAATAAATTTTAATTTAATGAGATTGTAAAACTATTTT
                        -299                                -262
ACGCGAAATTCCAACTCAATAAATAAAAATATTGTTAAATTCCGCATAAACAAACAAATATAA
                             -230
AAAAAAGAACAAGATACCAAATAGCGTGGGGAATGTAAAAAATAAAAAGAGTAAAAAACGCAG
                                      -160  -154  -148
CGTTTCATTGGCGTCTTCTCGACACAATAGAATACAACAGCATCAGCAACAAAACCGAACCGT
                 -119                 -81
GTTTTCTTCTTCCATCATCGTCACGTTCATTTTCATCCATCCCCAATTCACATCACCGATTTA
GGGTTTCTTCACTCTCTCTCCATCTCTTACTGCAACTTCCGTTTTCTCTCTAGTCGCTGGTGA
GCCTTCGCGCTTCTTCCTTTTTCTCTTTTCGATTTCTACTTTTTCATGCATGATGTGTTTACT
GATTTTTTTGTTCGTGTTTTTGATCTGAAGATTCTGTTTCACCTTACGTTTGTTTGGATCTGG
ATTGTGAAAATTTTACATGAGGATGTTTTTTTGAATTGAACTGTGTTAATTGTTAACTGTGTT
TTTTGTTATCGTTAGGGTGTTTTGTAGATGTAAGATTCTCGTTTTGATTTGTTGAGGAGGATT
GGATTTCTGGGTAGTGTTGAAAAATTGAGACAAGATTGTTGTAAGCGTGGTACAAAAATTAAG
GTGGTTGTTGCAAGTGTCTTAGAAGGAATTGAGGTATTTGTATGAGGCGAGGTTTTtGAGAAG
AGACAAGTARGGKGGTGGTTAGAGTTAGAGGTGGATTTAATGTCAAAGTCAAAACCCTAGAGG
TTTTTTGGTTATTTATGCGATAAGTGCTGATAATTTGTGTCTCAAAATGTGAAGAAAGGCGCC
AACGACCTAACTTTTTGTTTGTTTTTTAACGGTCGGCTGGACTATTTTACCTTAATAGAATCA
AAATGAGAGAATTAGAACCGTGGTCATTTAACAGCTGCCAATCAAAGCACGCGTTACTTGAAC
TAGAAACGCATGCCGAGGGTCCCACATGCACACACCACAGGGAAAACAATCTGATGCATTACC
TTGGTGTGCTAGATTCACTCTTGGCGTATTTGTTGCGTCAGTTTATTGTTTATGTTAGCTCGT
```

CTATTGTCTTTTGTTTGTAGGTAGAATTTCGCACACAAACTAAATATTAGGAAAAAACGTGTA
TTTCTCGAGCTGTGCACCGAATGCTGTTAAATGAGAAGATATTTTCCATAAAAACAACGAGGA
CCTAACATGTGTCCTACATGCACACACCAAAGGGAAAACTACCTGATGCGTTACCTTGGTGTG
TTAGATTCACTCTTGGCGTTTTTGTTGCATTGGTTTCCTCGTTTCTGTTGGCTCTTCTATTGT
ATTTTATTTGTAGGTAGAATTTTTGCACACGAACTAAATATTAGGAAAAAACGTATATTTCTC
GAGCTGTGCACGTAATGCTGTTGAATGAGAAGATATTTTCCATAAAATCAACGAGGACCTCAC
GCGTGTCCTCACGCGTCTCATGAAGCGTTATAAAGGAAATATGTATATACTTCTTATTTTGCT
TTCTTTAATAGTTTCTGTTTGCATTCAATTATATACCTTGTTTAAGGTAGTCAACGGTAAAAA
TTCCAATAACTCAATTAAGTAAGTCATTAATTTTGTTTTGAAACTCGATTAGACTTGACGCTT
ATTAGAGGAAAATTATGCTGTTGACGCCAATTAAATTTGTTTAGTTGTTATAGTTCTGTTGGA
TTTTTTTTTAATTGCTTAGCATATGATCCCGTGTCTTTATTCATTTGATTTAATATCCAATGA
TGGCTTTGATCTGTTTTGCGTGCGTGTGTATCCATTTTCATATACTCTAATATGTTTGTTTCT

-1

TTTTGAATTTTCCCCTAGATTTCGAAA

Inr- initiator element

binding site for RAV1

AT-1 site

inverted telo-box

ends of the intron

Figure 13. Nucleotide sequence of *V. faba SnRK1* gene 5'-flanking sequence between –383 and –1 relative to the ATG. Elements defined as binding sites for nuclear proteins are coloured. Introns sequence is painted blue, TATA-box is redded. Possible transcription start site is shown in larger font.

To study the promoter activity, two chimeric constructs were created driving the β-glucuronidase gene under different sizes of the promoter region without intron p**SnRK_S:GUS** and with intron p**SnRK_L:GUS,** respectively (Appendix 9.2). The constructs were used for protoplasts transient expression assays. A transcriptional fusion of the *VfSnRK1* gene 5'-upstream region (promoter) with intron (long) and without intron (short) to the β-glucoronidase (GUS) reporter gene was constructed by directional cloning of a 376 bp and a 1.915 Kb promoter fragment, respectively. The fragments were PCR amplified from the genomic DNA using *Pfu* polymerase with proofreading activity using the primers Prom3short, Prom3long and Prom5 with terminal *SalI* and *PstI* cloning sites. The isolated fragments were cleaved with corresponding enzymes, purified and cloned into the binary vector pGUS digested with *SalI/PstI* restriction enzymes. Proper insertion of the different

promoter sequences into the plasmid was confirmed by DNA digestion using suitable restriction enzymes and PCR using primers designed against sequences in the VfSnRK1 promoter. The plasmid maps are presented in the Appendix.

The plasmids p**SnRK_S:GUS** and p**SnRK_L:GUS** were transformed into *A. thaliana* protoplasts and, additionally, as positive control construct **p35S:GUS** (GUS gene driven under the CaMV35S promoter) and as negative control **pGUS** (GUS gene without any promoters) (Appendix 9.1). The promoter activities were analysed by measuring the transient expression of the reporter gene at 24 hours after transformation. The transient assays demonstrated the activity of the *VfSnRK1* promoter.

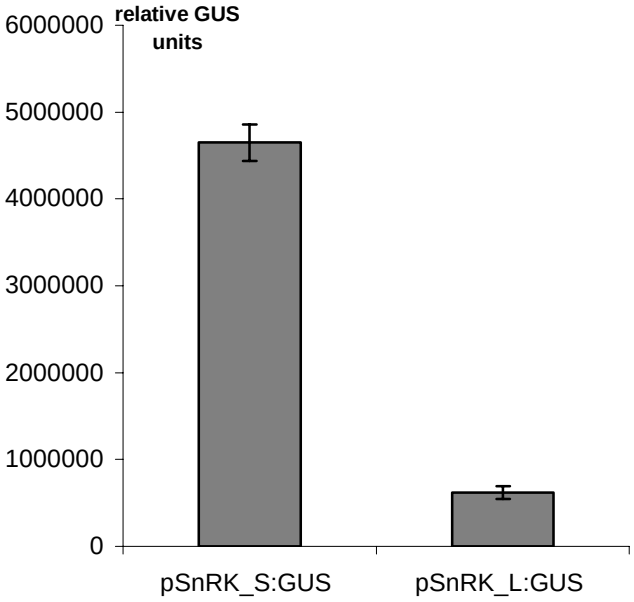


Figure 14. Transcriptional activities of the VfSnRK1 promoter fragments without intron (pSnRK_S:GUS) and with intron (pSnRK_L:GUS) in *A. thaliana* protoplasts. The columns represent the mean value (±SEM) of two repeats in two independent experiments. The groups were compared using *t*- test, P value < 0.05.

As shown in Fig. 14 the *VfSnRK1* promoter without intron shows a 7.5 fold increased activity in comparison to the *VfSnRK*1 promoter with intron. A relatively high GUS activity (33.5% of the positive control p35S:GUS) was detected (data not shown).

The transient assays clearly demonstrated the elevated levels of GUS expression depend on promoter size and presence of the intron.

4.5 Hormonal control of *VfSnRK1* promoter activity

Quantitative fluorometric GUS measurement revealed differences in GUS activity after cultivation of protoplasts in media with different hormone compositions. When protoplasts were cultivated in media containing auxin (2mg/L NAA; 1 mg/L BAP) and cytokinin (1 mg/L 2.4D) to stimulate cell division, higher GUS activity was detected. Protoplasts cultivated with cytokinin (1 mg/L BAP) leads to a smaller increase in GUS activity. In contrast, auxin alone

(2mg/L NAA; 1 mg/L 2,4D) induced GUS activity. The combination of cytokinin and ABA (1 mg/L BAP; ABA 10^{-4} M) leads to a dramatic increase of GUS activity. Low concentrations of sucrose (1 mg/L BAP; 1% sucrose) decreased the SnRK1 promoter activity again (Fig.15).

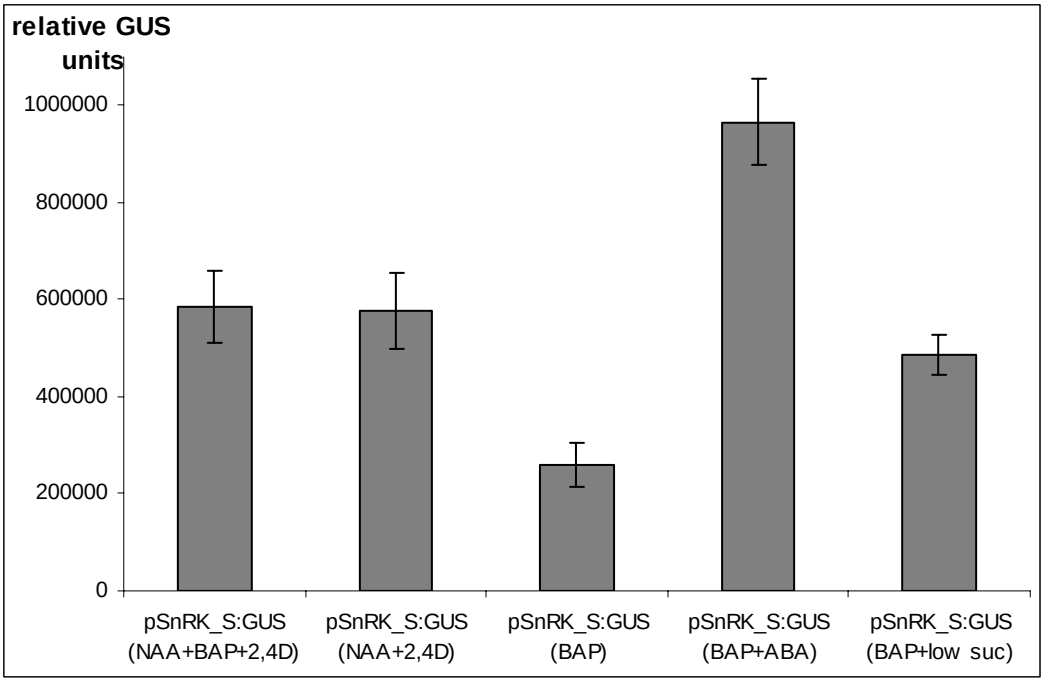


Figure 15. Analysis of GUS activity driven by the truncated VfSnRK1 promoter (pSnRK_S:GUS) in *A. thaliana* protoplasts in the transient expression assay. The protoplasts were transformed and cultivated on medium with different concentrations of hormone. The mean value (±SEM) of GUS activity represents three independent experiments. The groups were compared pairwise using one-way ANOVA test, P-value < 0.05.

Similar results were obtained with mesophyll protoplasts of *N. plumbaginifolia*. Despite of the low absolute levels of the transient activity of the promoter, the relative changes were comparable to the results obtained with protoplasts of *A. thaliana* (Fig. 16). The mesophyll protoplasts of *N. plumbaginifolia* are larger and morphological changes are more visible. The cultivation of mesophyll protoplasts on medium for 4 days with cytokinin alone leads to an arrest of cell division (Fig. 17). The protoplasts were three fold large as control protoplasts, cultivated on normal medium, and did not divide.

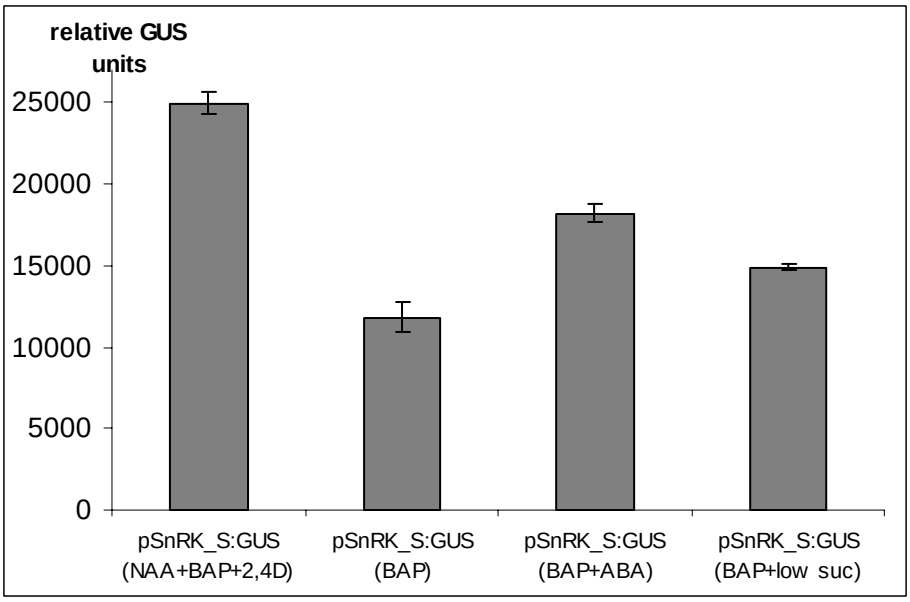


Figure 16. Measurement of GUS activity of pSnRK_S:GUS in *N. plumbaginifolia* mesophyll protoplasts using a transient expression assay. The protoplasts were cultivated in media with different concentration of hormone. The mean value (±SEM) of GUS activity represent two independent experiments. The groups were compared pairwise using one-way ANOVA test, P-value < 0.05.

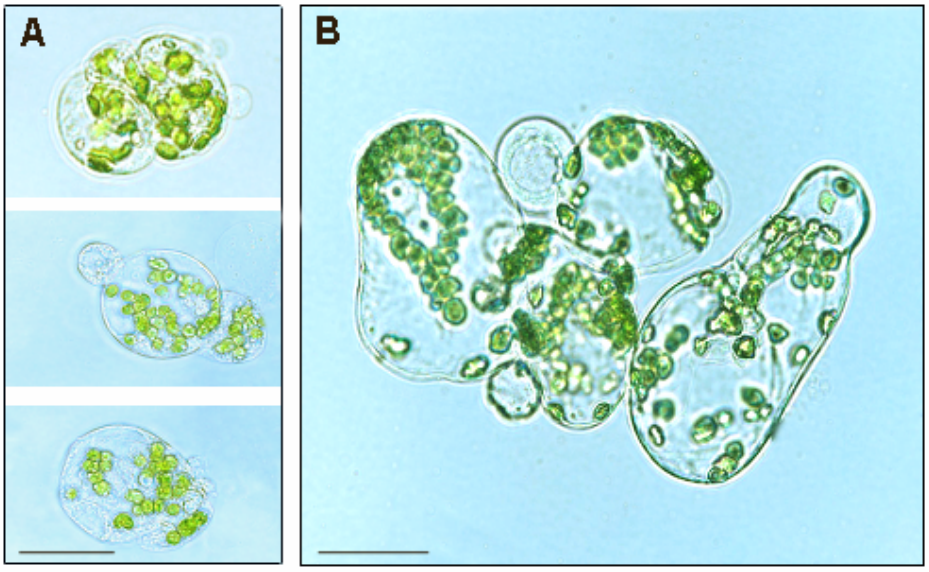


Figure 17. The *N. plumbaginifolia* mesophyll protoplasts cultivated on medium containing (A) auxin (2mg/L NAA; 1 mg/L BAP) and cytokinin (1 mg/L 2.4D) or (B) only cytokinin (1 mg/L BAP); *bar* = 50 μm.

4.6 Generation and analysis of transgenic plants

4.6.1 Vector construction for plant transformation

For VicP/VfSnRK1 antisense and VicP/VfSnRK1 sense constructs, a *SalI/SmaI* (2076 bp) full-length coding cDNA fragment was cut from the pBK-CMV vector. The DNA fragment was purified, treated with Klenow enzyme and *PNK* and inserted into the **pUC18/VicP** vector (Appendix, Chapter 9.1) cleaved with SmaI enzyme and dephosphorylated with SIP. The orientation was determined by restriction analysis and sequencing. The 3' region of octopinsynthase (*ocs)* was used to terminate all the SnRK1 genes. Sequencing was performed to select an error-free cDNA fragment. The plasmid maps of **SNF1a/s** (VfSnRK1 –cDNA (reverse) cloned in pUC18/VicP) and **SNF1s** (VfSnRK1-cDNA cloned in pUC18/VicP) are presented in the Appendix, Chapter 9.1.

To create a binary vector plasmid from pGPTV-bar (Becker et al., 1992) the GUS gene was cut out using *SacI/SmaI* endonucleases and was replaced by the *EspI/NarI* expression cassette containing VicP/VfSnRK1 antisense/3'ocs or VicP/VfSnRK sense/3'ocs using blunt-end ligation. The orientation was determined by restriction with *BamHI.* The plasmid maps of **SNF3/1** (VicP:VfSnRK1 (reverse) cloned in pGPTV) and **SNF4/16** (VicP:VfSnRK1 cloned in pGPTV) are presented in the Appendix, Chapter 9.1.

To construct an USP promoter/VfSnRK1 transcriptional fusion, a *Sal/BglII* DNA (681 bp) fragment containing the USP promoter was cut from the **pUC18/USPP** vector, purified, blunted with Klenow enzyme and treated with *PNK*. The fragment was inserted into the **pBK-CMV/SnRK1** vector (Appendix, Chapter 9.1), cleaved with *XbaI*, blunted with Klenow enzyme and dephosphorylated with SAP (for antisense orientation) or with *EcoRI*, blunted and dephosphorylated (for sense orientation). The orientation was determined by restriction with *SphI/XbaI* in case of sense cloning or with *SphI/EcoRI* in case of antisense cloning. Positive colonies were sequenced and used to create a binary vector plasmid. The *XbaI/SalI* USPP/VfSnRK1 sense (2.8 Kb) fragment and the *EcoRI/NotI* USPP/VfSnRK1 antisense (2.8 Kb) fragment were purified. Insertion of the cassettes USPP/VfSnRK1sense and USPP/VfSnRK1 antisense, were performed using *EcoRI* adaptors. The *SacI/SmaI* cleaved pGPTV-bar vector and fragments were blunted with *Pfu* polymerase and *EcoRI* adaptors were ligated with help of T4 ligase. Then the both phosphorylated by PNK inserts were ligated into the SAP treated pGPTV-bar vector. The plasmid maps of **uSNF8x** (VfSnRK1-cDNA (reverse) cloned in pBK-CMV under USP promoter) and **uSNF4e** (VfSnRK1-cDNA cloned in pBK-CMV under USP promoter) are presented in the Appendix, Chapter 9.1.

The resulting plasmids were transformed into *A. tumefaciens* strain EHA105 by the freeze-thaw method (An et al.,1988) and transferred into tobacco (*Nicotiana tabacum* cv Gavana) and pea (*Pisum sativum* L.- cv Erbi) plants.

4.6.2 Production of transgenic N. tabacum and P. sativum plants

Several aspects of the transformation procedures were examined in order to improve the transformation of pea: medium composition, pre-culture and co-cultivation conditions, concentrations of phosphinotricine for selection of transformed shoots. To determine the optimal selective conditions, the pea explants were cultivated on different concentrations of PPT (1 μg/ml, 1.5 μg/ml, 2 μg/ml, 2.5 μg/ml, 4 μg/ml). At a concentration of 4 μg/ml the explants had a reduced potential for callus generation and this concentration was chosen for further selection (Fig. 18a). However, not all regenerated shoots were transgenic on this selective medium. Therefore, the concentration of phosphinotricine in MS4 medium for shooting selection was increased to 10 μg/ml. Shoot development is shown in Figure 18b.
Since no satisfactory results were obtained with rooting of pea shoots *in vitro,* grafting was used for the further cultivation of transformants. T_0 seedlings were grafted and transferred to greenhouse conditions to obtain T_1 seeds. As an example, the efficiency of transformation for one construct ***VicSnRKas*** is shown in Table 1.

Table 1. The efficiency of pea transformation with the construct ***VicSnRKa*** in six independent experiments.

Experiment	***Number of explants, ca.***	***Formed shoots***	***Grafted shoots***	***Positive on PCR***
I	400	20	12	6
II	400	37	13	5
III	200	3	1	1
IV	400	25	16	6
V	400	24	10	4
VI	200	10	6	1
Together	2000	119	58	23

We observed similar numbers with the other constructs. The presence and integrity of the transgenes in T_0 and following plant generations were confirmed by PCR analyses. 23 individual transgenic plants from 9 independent lines containing the *VicSNFas* construct and

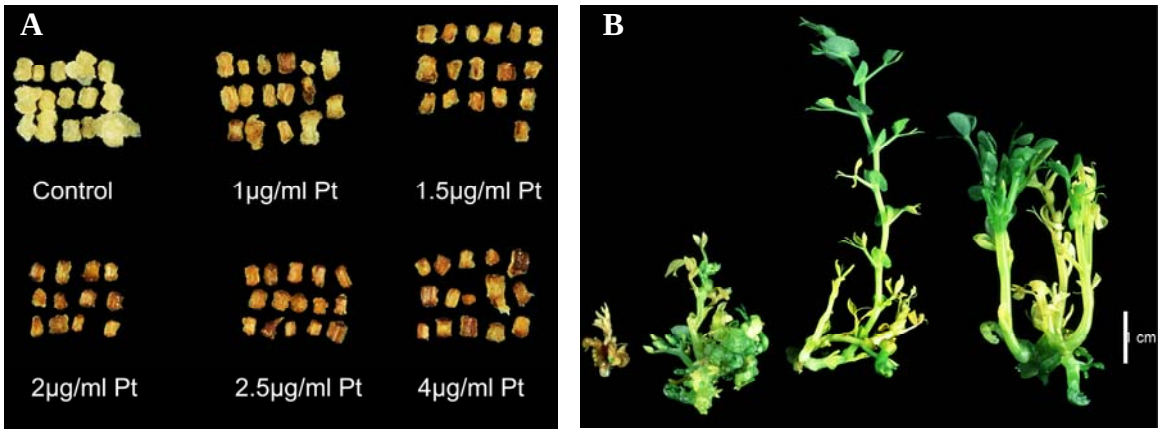


Figure 18. (A) The pea explants cultivated on different concentrations of PPT (1 µg/ml, 1.5 µg/ml, 2 µg/ml, 2.5 µg/ml, 4 µg/ml). The reduction of callus formation was observed at 4 µg/ml. (B) Transgenic shoots development after cultivation on selective medium. *bar*=1cm

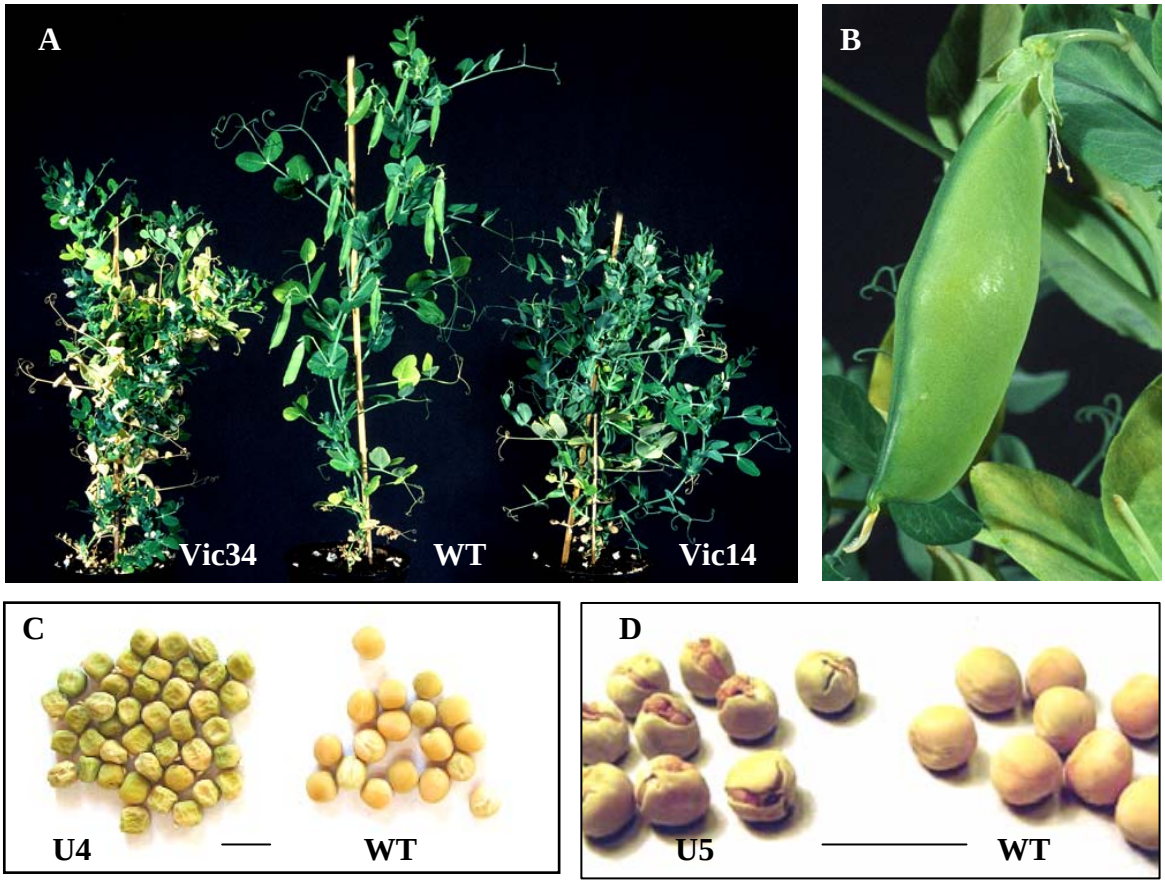


Figure 19. (A) Comparison of the phenotype of transgenic *Vic*34, *Vic*14- pea lines and WT plant (middle). (B) Pod from transgenic *Vic*34 line. (C, D) Comparison of the mature seeds from *U*4 and viviparous seeds of line *U*5 and WT; *bar*= 1cm.

25 individual plants from 3 independent lines containing the *USNFas* construct were generated.

Production of transgenic tobacco was made according to the standard protocol described in Material and Methods, Chapter 3.7. 13 lines including 23 transgenic plants were recovered.

4.6.3 Analysis of transgenic plants

Most of the transgenic plants were morphologically indistinguishable from tissue culture-derived or seed-grown plants. However, about 30% tobacco and pea transformants showed partial or complete sterility. Additionally, unusual branching effects were observed in developing shoots of some trangenic pea lines (Fig.19a). The effect was not expected taking in to account the seed specific promoters used for transformation, but could be explained by possible vicilin and USP promoter activity in meristem cells. In many cases, the mature seeds were wrinkled and smaller (Fig. 19c) and the number of seeds in a pod was reduced (Fig. 19b).

In order to study in more detail the effects of the reduced SnRK1 transcript levels in developing seeds of the transgenic plants, the following T1 transgenic lines were selected showing the phenotypical changes: the pea lines Vic*25*, Vic*34*, Vic*14* transformed with VicSNFas construct, the pea U*5* line transformed with USNFas and T_0 tobacco lines Vic *3/1/23*, Vic *3/1/36* and Vic *3/1/30* transformed with VicSNFas construct. The typical and reduction of *SnRK1* transcripts amount in seeds served as criterion of choice. Viviparous seeds were observed in line *U5* (Fig. 19d). In Vic*34*, Vic*14* lines we observed smaller fresh weight of seeds during development that can be because of both consequence of backlog in development, and/or an insufficient seed felling (Fig. 20; 21).

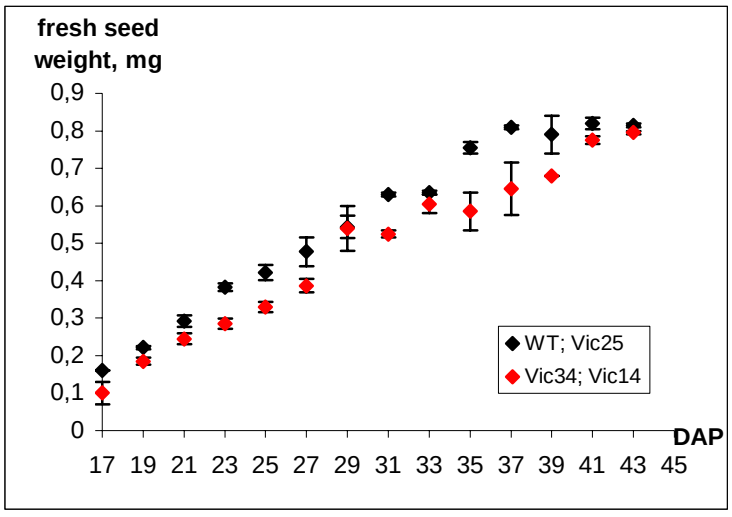


Figure 20. Fresh weight accumulation [means (±SD)] of developing pea seeds. WT –*P. sativum* cv.'Erbi', Vic*25*, Vic*34*, Vic*14*-transgenic pea lines.

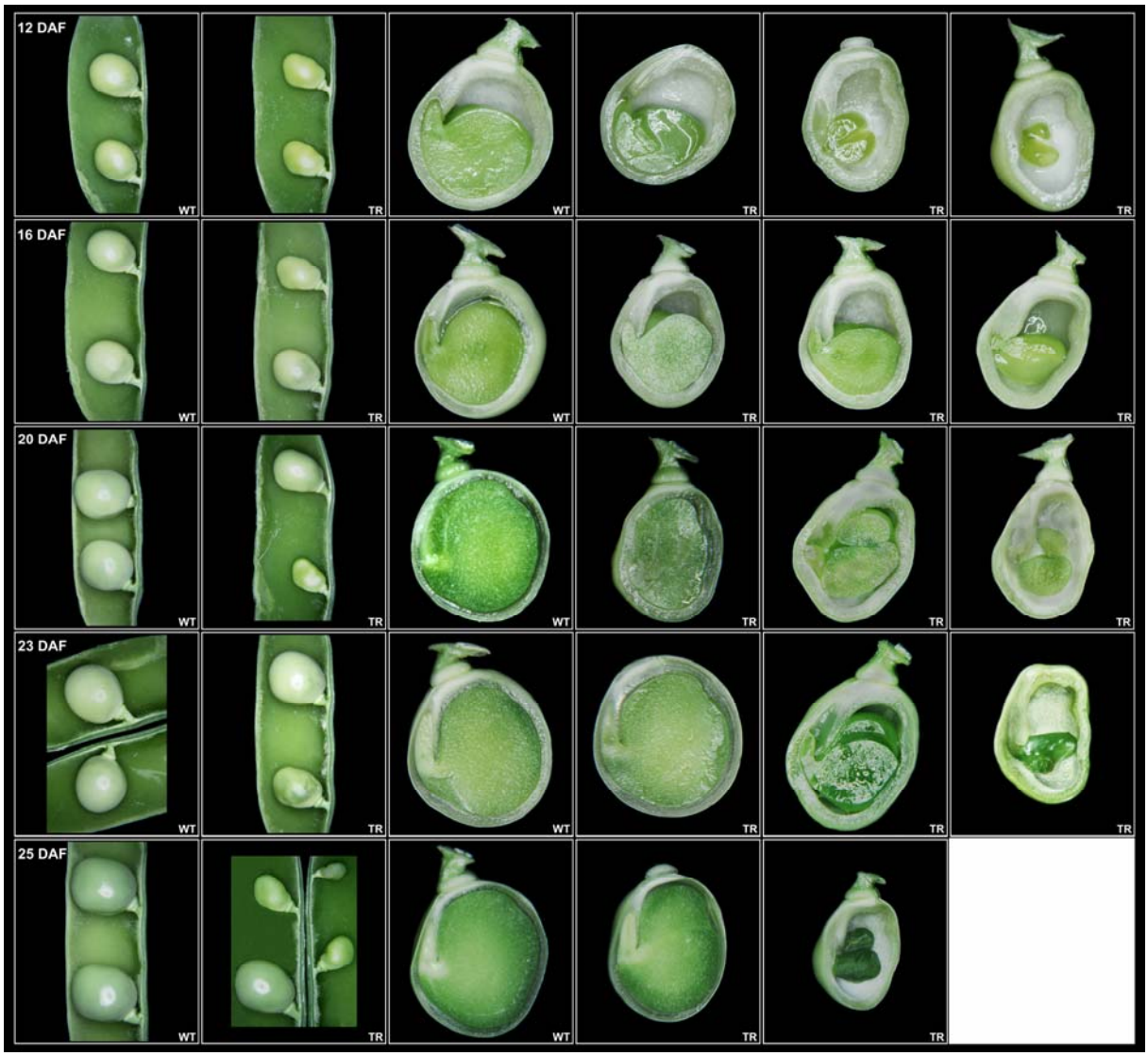


Figure 21. Seed development in WT and Vic*34* transgenic (TR) pea line. The WT pea seed is characterized by width equal to its pod cavity and normal development. The pods of Vic*34* plant contain seeds of two clearly distinguished types: with WT-like development and with strong disruption in development. WT-like seeds have smaller fresh weight and size in comparison to normal WT seeds.

The culture-derived non-transformed lines served as controls. 100% mutant phenotype in T_3 generation and PCR analysis of these plants enabled us to assume the availability of near-isogenic lines and demonstrate stable transmission of the phenotypes.

Southern blot analysis with a pea-specific cDNA fragment encoding the *PsSnRK1* revealed only one copy of the *PsSnRK1* T-DNA in the plant genome (Fig.22). The blots showed a ca. 5 kb hybridisation band, corresponding to the endogenous *PsSnRK1* and one additional band corresponding to the Vf*SnRK1* T-DNA insertion.

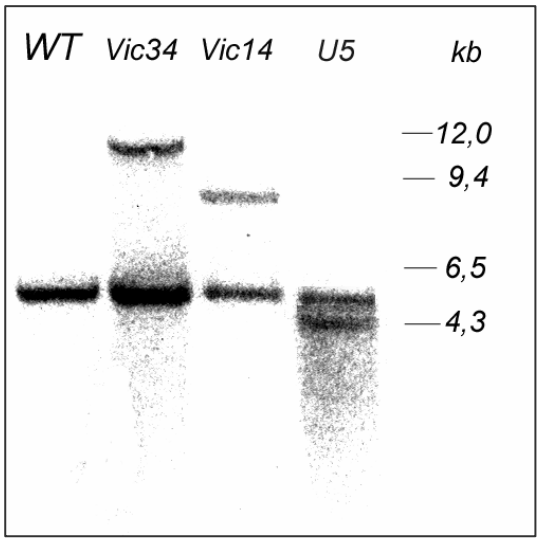


Figure 22. Southern blot analysis of the transgenic pea lines Vic*34*, Vic*14* and U*5*. 10 µg of genomic pea DNA were digested with the restriction enzyme Hin*d*III and hybridised using a ^{32}P-labeled 800-bp fragment encoding the *PsSnRK1* gene.

4.6.4 Abnormal pollen development in transgenic lines

Cytological observation of the mature pollen from sterile tobacco and pea lines indicated abnormal pollen development. In pea Vic*25*, Vic*34* lines and in T_0 tobacco Vic *3/1/28,* Vic *3/1/36* and Vic *3/1/30* plants mature anthers were found to contain two clearly distinguishable types of pollen grains. More than half of the pollen grains were round, filled with starch granules and approximately the same size as pollen from the control plants (Fig. 23). The other pollen contained little or no starch and were smaller than pollen from the control plants.

Most likely in the anthers from lines Vic*25*, Vic*34* the amount of fertile pollen was sufficient for successful fertilisation. We assume, that in these lines repression of SnRK1 occured only partial and insufficient for dramatic consequences for plant reproduction.

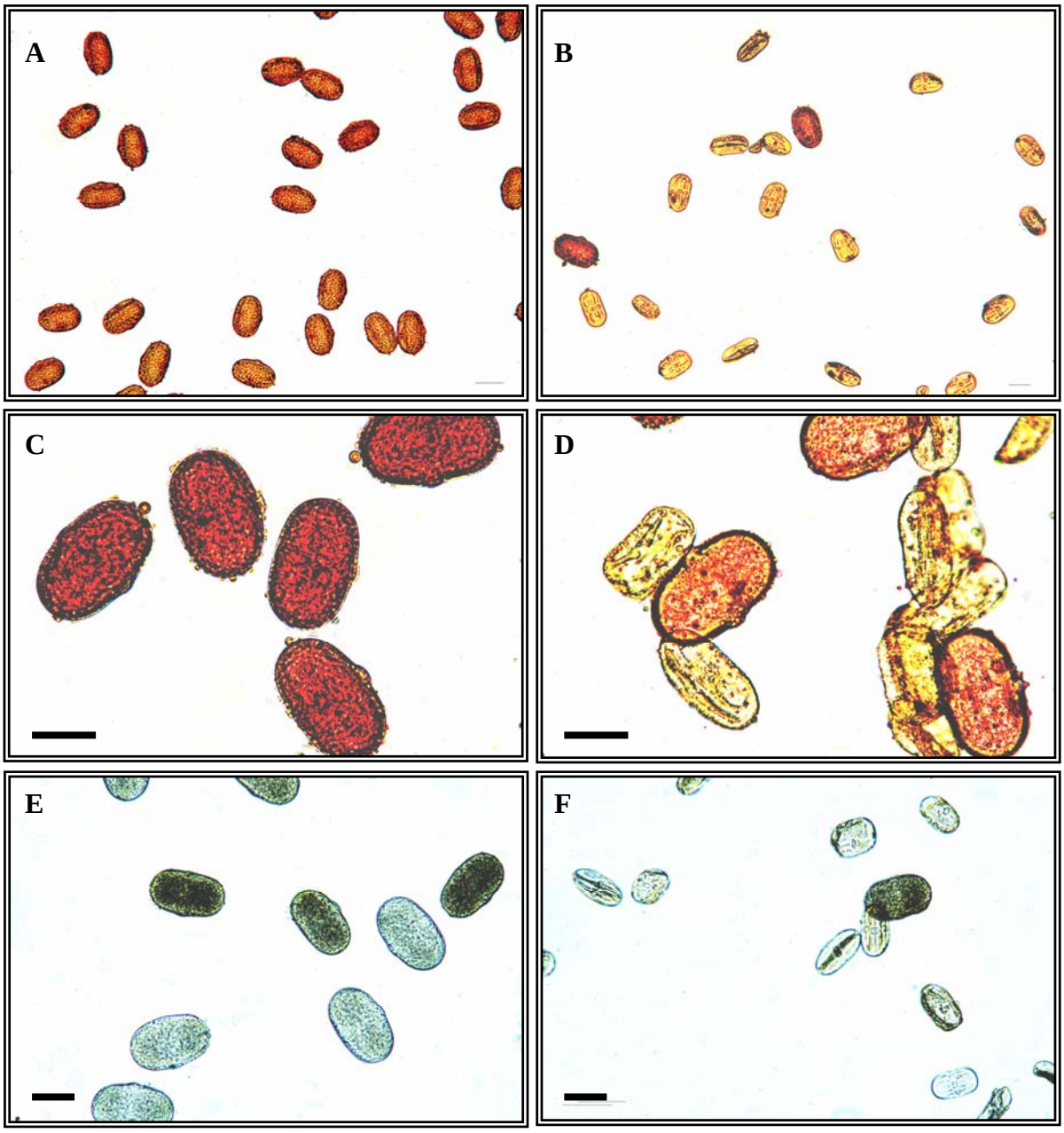


Figure 23. Pollen from WT pea plants and transgenic Vic*34* pea lines. 1% acetocarmine staining: (**A, C**)-WT pollen; (**B, D**) -Vic*34* pollen. Iodine staining: (**E**) -WT pollen (**F**) Vic*34* pollen. Staining performed as described in Material and Methods, Chapter 3.9, *bar*=25 μm.

The completely sterile T_0 tobacco and pea plants did not produce seeds and could not be analysed further. In order to prove the sterility, back-crosses were performed between fertile control tobacco plants as a pollen provider and completely sterile transgenic lines as a donor. The seeds were germinated on selective MS medium containing phosphinotricine. No transgenes were recovered. The abnormal pollen grains were not capable of fertilization.
RNA blot analyses demonstrated the reduced *SnRK1* transcript levels in anthers of sterile transgenic tobacco T_0 plants Vic *3/1/28*, Vic *3/1/36* compared to WT tobacco plants and as well as fertile line Vic *3/1/5* (Fig. 24).

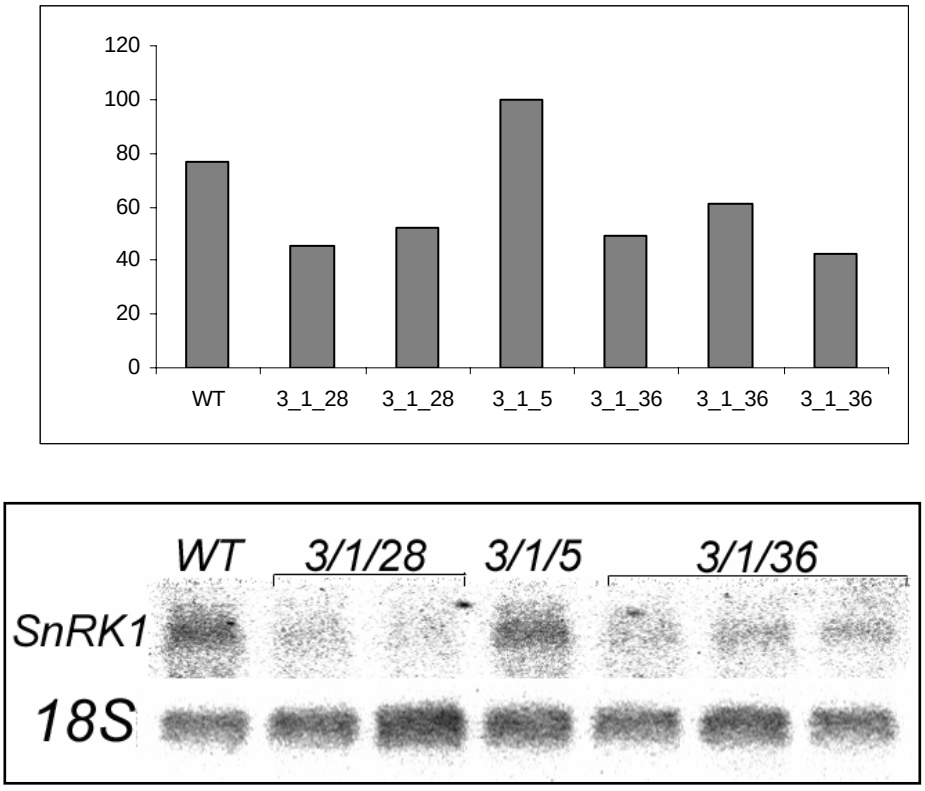


Figure 24. Antisense *SnRK1* expression in mature anthers of transgenic tobacco plants: Vic *3/1/28*, Vic *3/1/36* – were sterile plants and Vic *3/1/5*-was a fertile plant. *WT* - RNA from anthers of control non-transgenic *N. tabacum* cv. 'Havana' plant.

4.6.5 Expression analysis of transgenic pea lines

The relative level of *PsSnRK1* and *PsAKINß1* transcripts at different stages of embryo development for lines Vic*14* and Vic*34* were determined by Northern blot analysis of total RNA, using the PCR-derived probe for *PsSnRK1* and the same blots were hybridised with the probe for *PsAKINß1*. In WT the blot showed a single hybridisation band corresponding to the

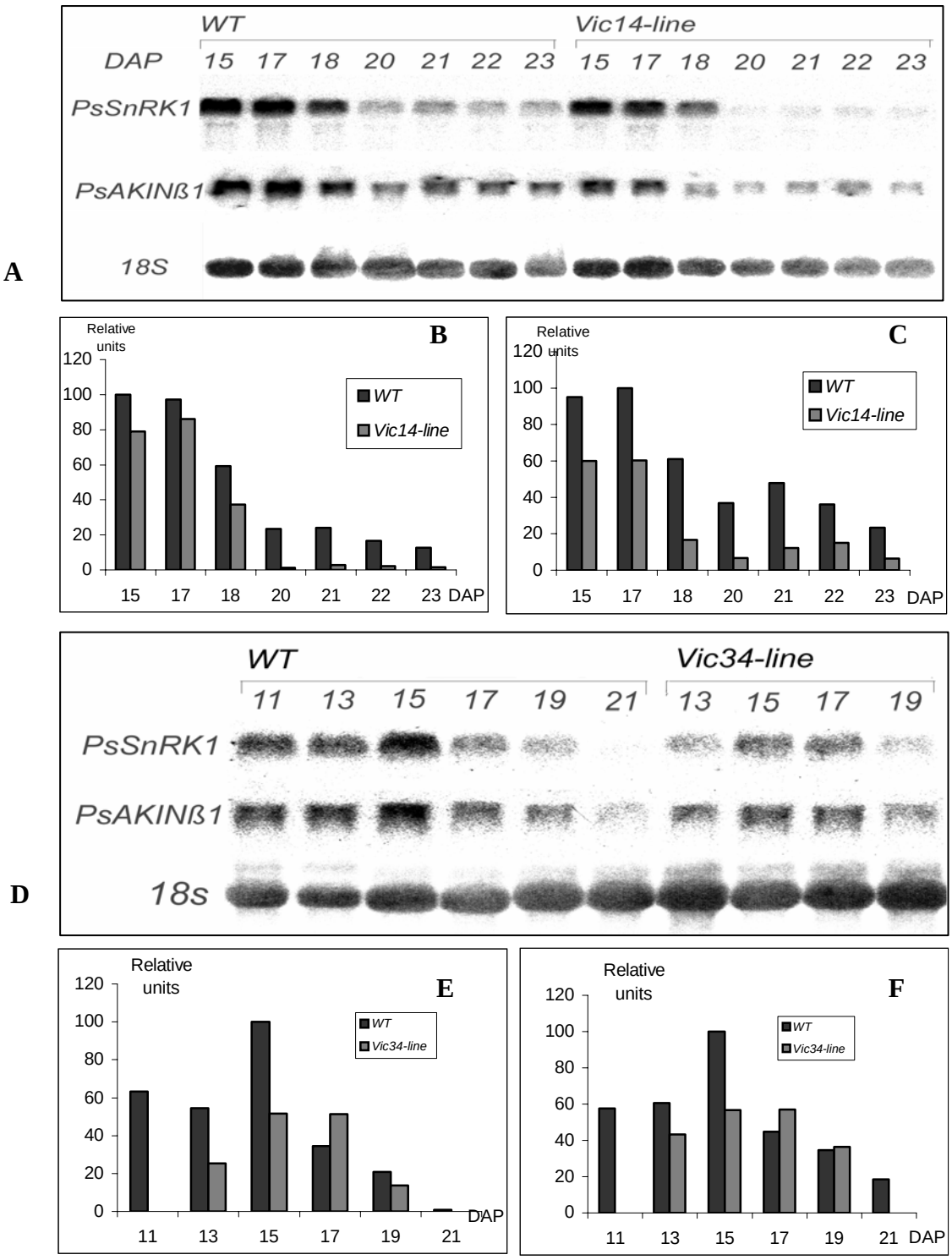


Figure 25. Expression levels of the *PsSnRK1* and *PsAKINß1* in embryos of wild type (WT) and transgenic Vic14 (**A**) and Vic34-lines (**D**). Relative mRNA abundance was quantified and normalized for RNA loading in each lane. Comparing of transcripts amount: (**B**) *PsSnRK1s* in WT and Vic14-line; (**C**) *PsAKINß1* in WT and Vic14-line; (**E**) *PsSnRK1* in WT and Vic34-line; (**F**) *PsAKINß1* in WT and Vic34-line.

endogenous sense transcript. In transgenic lines Vic*34* and Vic*14 SnRK1* transcripts amount was reduced compared to WT pea plants. Surprisingly, the transcript profiles of *PsAKINß1* were correlated with those of *PsSnRK1* in both WT and transgenic lines. It indicates that antisense reduction of *PsSnRK1* gene expression leads to a similar decrease of *PsAKINß1* mRNA(Fig. 25).

4.6.6 SnRK1 activity in transgenic pea lines

SnRK1 activity was estimated using the SAMS peptide assay in the embryos of untransformed plants and transgenic Vic34 line grown under identical environmental conditions. Five different stages of seed development (DAP 10, 16, 18, 21, 23) were chosen for analysis. The highest reduction of activity (71%) was observed in the early stage of development (DAP 16). The differences in activity were not significant at the late stage of development (DAP 23) (Fig. 26).

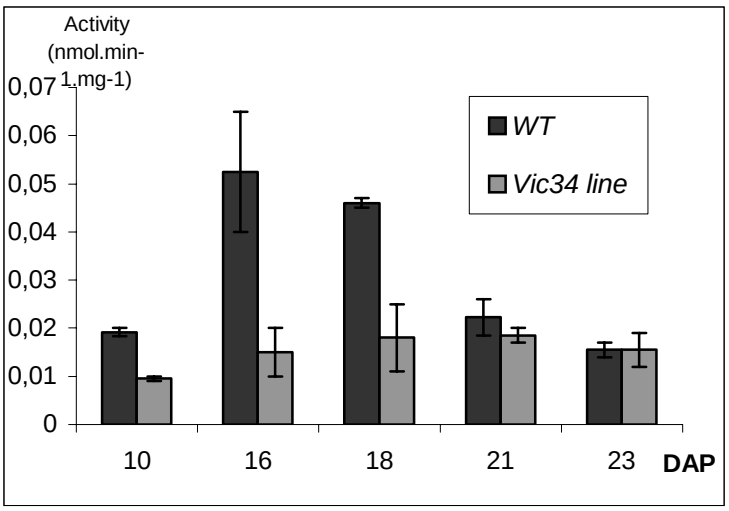


Figure 26. SAMS peptide kinase activity (nmol phosphate incorporated per minute per mg of an initial plant material) in developing embryos of wild type (WT) and transgenic Vic34-line. Means (±SD) of two independent measurements are shown for each line.

4.7. EST generation

4.7.1 Choice of developmental stages and cDNA libraries

Two cDNA libraries, one of developing cotyledons and another one from seed coat were constructed to generate ESTs. The following strategy was used to avoid the presence of high abundant ESTs corresponding to storage proteins and to expand representation of unique and non-redundant genes, which are involved in different stages during seed development. Isolated RNA from cotyledons of four different stages of *P. sativum* seed development (III-IV) (Hauxwell el al., 1990; Borisjuk et al., 1995) was mixed in equal amounts and used for the synthesis of the first-strand cDNA. RNA from seed coat from seeds of the same developmental stage was prepared in the same way. Two cDNA libraries were generated as described in Material and Methods, Chapter 3.10.1. In total 5 538 cDNA clones from the *P. sativum* cotyledon library (PSC) and 7680 cDNAs from the seed coat (PSS) library were sequenced from their 3' ends by Qiagene company and in IPK-Gatersleben.

4.7.2 Assessment of library and EST sequence quality

Based on random PCR amplification of 300 insert sequences, the average insert size of the libraries was estimated as being between 1000 and 1600 bp. As shown in Figure 27, the average sequence length is between 100 and 690 nucleotides with an average length of 620 nucleotides and a maximum of 691. The most abundant ESTs represented in the *P. sativum* EST library are shown in Table 2.

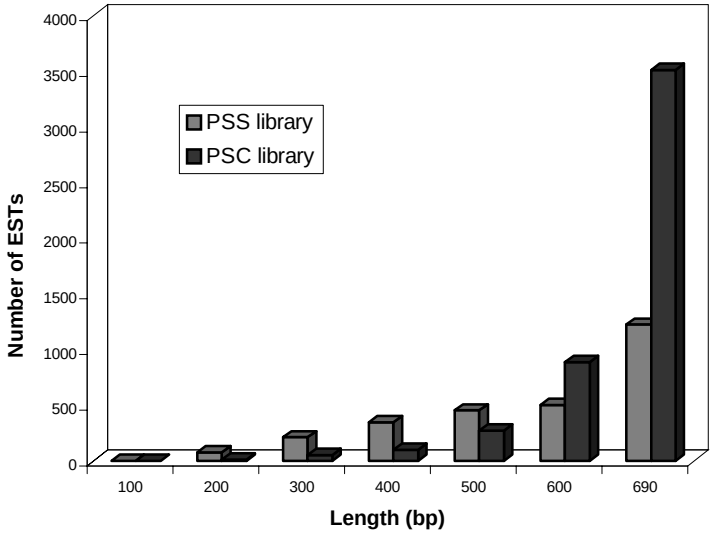


Figure 27. Distribution of lengths of the EST sequences after quality control. *Black bars*: distribution of ESTs length in the PSC library; *gray bars*: distribution of ESTs length in the PSS library. The average length of ESTs of the PSC library was 565 bp and that of the PSS library was 652 bp.

Table 2. The most abundant ESTs in the *P. sativum* EST collection

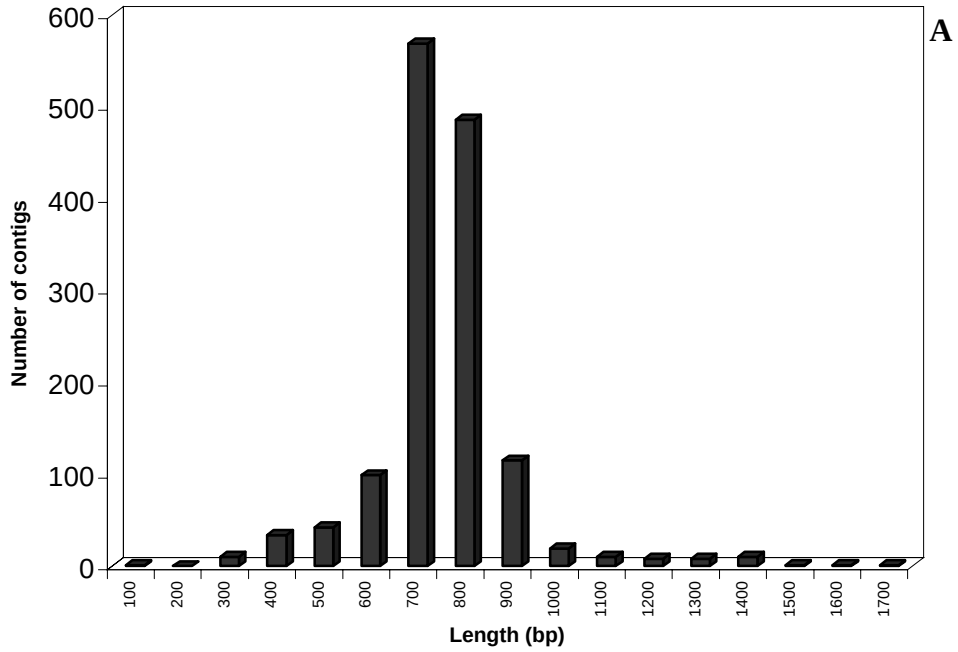
Gene name	**Number of ESTs**	*NCBI Accession number*	*Species*
Vicilin precursor	260	P13918	P. sativum
Vicilin	328	S35757	P. sativum
Convicilin precursor	32	P13915	P. sativum
Legumin A2 precursor	53	S08237	P. sativum
Legumin B LegJ precursor	49	P05692	P. sativum
Legumin B	23	S37241	V. faba
Chlorophyll a/b-binding protein	97	P27490	P. sativum
Embryonic abundant protein	101	P21747	V. faba
Photosystem II reaction center W protein	77	S53025	Spinacia oleracea
Embryonic protein BP8	71	S39475	Betula pendula
WD-40 repeat protein	69	NP_196473.1	A. thaliana
Elongation factor 1-alpha	68	Q41011	P. sativum
Peptidylprolyl isomerase	50	T50770	V. faba
Dehydration-responsive protein RD22	52	AAL26909.1	Prunus persica
Alcohol dehydrogenase 1	44	P12886	P. sativum
Unknown protein	46	At1g27150	A. thaliana
Glyceraldehyde 3-phosphate dehydrogenase	30	P34922	P. sativum
Annexin	29	CAA75308.1|	Medicago truncatula
Methionine synthase	24	AAQ08403.1	Glycine max
Histone H2A	23	P25470	P. sativum
Histone H2B like	38	CAA69025.1	A. thaliana
Histone H3	132	P02300	P. sativum
Histone H4	53	HSWT41	T. aestivum
Histone H1.b	24	S00033	P. sativum
60S ribosomal protein L2	12	CAC20221.1	Glycine max
60S ribosomal protein L3, putative	13	AAK43985.2	A. thaliana
60S ribosomal protein L6	11	NP_177546.1	A. thaliana
60S ribosomal protein L10	14	Q9M5M7	Euphorbia esula
60S ribosomal protein L10A	11	NP_197636.1	A. thaliana
60S ribosomal protein L18	14	O65729	Cicer arietinum
60S ribosomal protein L30	18	O49884	Lupinus luteus
60S ribosomal protein L44	11	Q96499	Gossypium hirsutum
40S ribosomal protein S3a	10	CAD56219.1	Cicer arietinum
40S ribosomal protein S5	14	O65731	Cicer arietinum

40S ribosomal protein S7A	27	NP_175314.1	A. thaliana
40S ribosomal protein S8	13	Q08069	Zea mays
40S ribosomal protein S9C	13	NP_198801.1	A. thaliana
40S ribosomal protein S11	11	P17093	Glycine max
40S ribosomal protein S23	12	AAF26742.1	Euphorbia esula
40S ribosomal protein S25B	48	NP_179752.1	A. thaliana
40S ribosomal protein S27A	13	S40240	Lupinus albus
60s acidic ribosomal protein P0	12	P50346	Glycine max

Sequence cleaning and quality check has been performed as described in Material and Methods, Chapter 3.10.1. A total of 8414 ESTs (4958 ESTs from cotyledons library and 3756 ESTs from seed coat library) was used for clustering .

4.7.3 EST assembly into contigs

To identify ESTs belonging to the same open reading frame (ORFs) sequences from both libraries were assembled using the StackPack software package. StackPack analysis of 8414 sequences resulted in 1082 clusters (1465 contigs) containing 6061 sequences and the 2353 singletons represent different clones. Chimeric ESTs tend to join two distinct sub-clusters. Therefore, after BLASTX alignment of the sequences large clusters were inspected manually to identity and remove chimeric clones. Using the remaining sequences, the average contig length of the remaining dataset was 748 nucleotides with its maximum at 1700 bp (Fig. 28a). Contigs with high member number encode vicilin (136 ESTs), vicilin precursor (119 ESTs) and embryonic abundant protein (101 ESTs) (Fig. 28b).



B

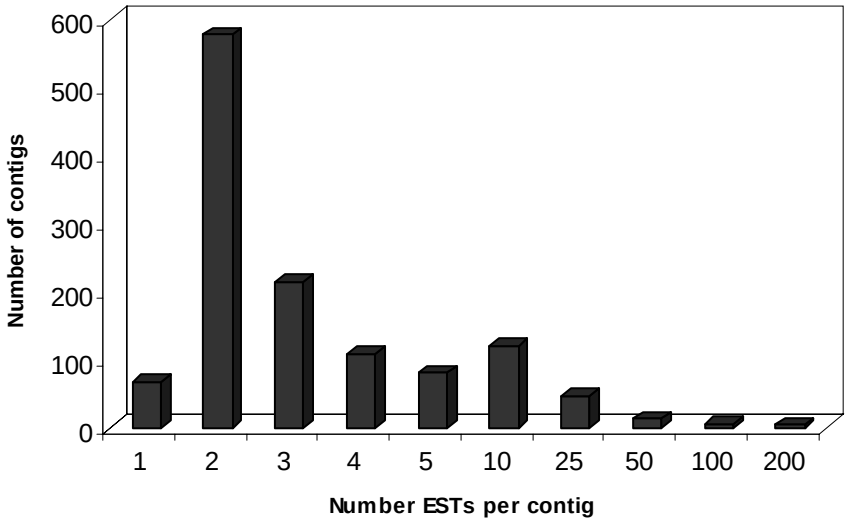


Figure 28. **(A)** Length distribution of sequences assembled within contigs. The average length of the contigs was 748 bp, corresponding to a peak between 700 and 800 bp. **(B)** Distribution of the number of ESTs assembled per contig. Most of the contigs have two ESTs. Two contigs (embryonic abundant protein and vicilin) contain more than 100 ESTs.

In summary, 3435 genes have been identified. The EST sequence of all clones along with clustering information is available at http://pgrc.ipk-gatersleben.de/est/index.php.

4.7.4 Annotation and functional classification of the ESTs

8414 ESTs were annotated with respect to gene function using results of BlastX2 comparisons with the NRPEP protein database. Information regarding length of the aligned sequence segment, score, E-value, percent of identity and other parameters were extracted from the results using a custom made Perl script (http://pgrc.ipk-gatersleben.de). The quality of the hits is illustrated in Figure 29. Generally, sequences from the PSS library give lower E-values than sequences from the PSC library, probably due to the different quality of libraries and various methods of sequencing. The percentage of hits among the sequences from both libraries is 80%, that makes 52% of sequences from PSC library and 28% of sequences from PSS library (assuming a BlastX2 E-value <1E-20 as a hit).

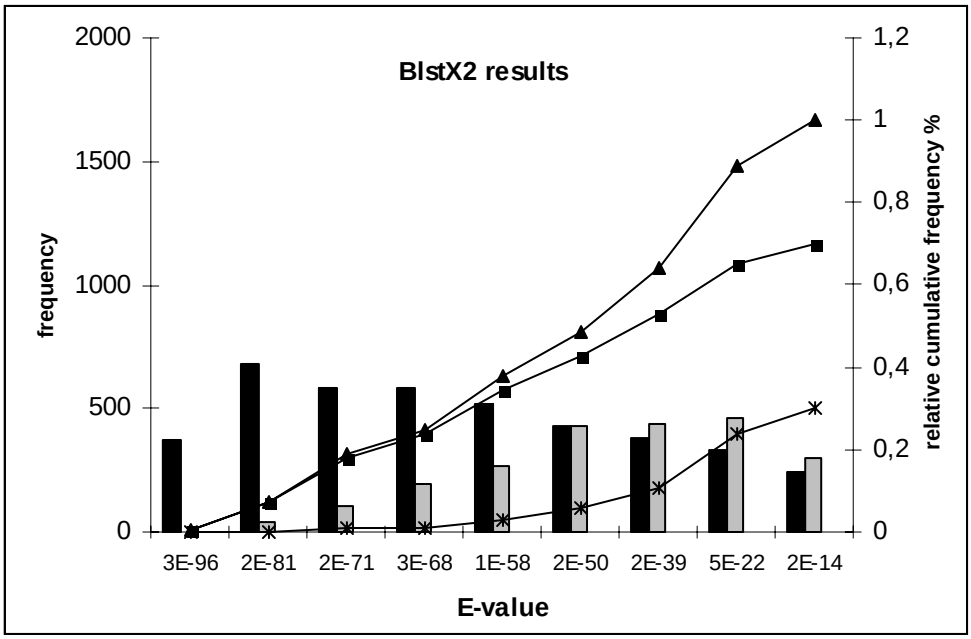


Figure 29. Quality of BlastX2 hits obtained after comparison of sequences from PSS and PSC libraries to NRPEP protein database. *Black bars*: frequency of PSC hits; *gray bars*: frequency of PSS hits; *line+triangle*: cumulative frequency of all hits; *line+square*: cumulative frequency of PSC hits; *line+cross*: cumulative frequency of PSS hits.

4.7.5 *Assignment of the P. sativum dataset to common Gene Ontology terms*

The UniGene set of clones (total 4548), which were chosen for spotting on the membrane, according to the criteria described in section Material and Methods was functionally classified. 2391 ESTs (53%) were defined as homologous (E-value <1E-20) to entries existing in the NRPEP protein database. The further procedure was carried out only with these ESTs. Functional classification according to the Gene Ontology (GO) Consortium was carried out for each EST sequence. The identified entries were recruited via extraction data linkage graphs (DLGs) from the KEGG database for functional characterization. Out of 2391 ESTs with reliable BLASTX results, 389 ESTs were annotated to possess enzymatic activity. 83 ESTs are coupled to category *'biological processes'*. The rest (1919 ESTs) were annotated manually via text search, by using the NCBI PubMed library and information about gene function available in publications. The resulting functional categories have been classified according to metabolic and biological processes. The KEGG PATHWAY database was used as a basis for classification. The results are presented in the Table 3.

Table 3. Functional categories of annotated metabolic and physiological processes applied for *P. sativum* EST annotation.

1. Metabolism	Cysteine metabolism	**2. Genetic Information Processing**
1.1 Carbohydrate Metabolism	Valine, leucine and isoleucine degradation	***2.1 Transcription***
Glycolysis / Gluconeogenesis	Valine, leucine and isoleucine biosynthesis	RNA polymerase
Citrate cycle (TCA cycle)	Lysine biosynthesis	Transcription factors
Pentose phosphate pathway	Arginine and proline metabolism	*2.2 Translation*
Pentose and glucuronate interconversions	Histidine metabolism	Ribosome
Fructose and mannose metabolism	Urea cycle and metabolism of amino groups	Translation factors
Galactose metabolism	***1.6 Metabolism of Other Amino Acids***	Aminoacyl-tRNA biosynthesis
Ascorbate and aldarate metabolism	beta-Alanine metabolism	***2.3 Folding, Sorting and Degradation***
Starch and sucrose metabolism	Selenoamino acid metabolism	Protein folding and associated processing
Aminosugars metabolism	Cyanoamino acid metabolism	Protein destination and storage
Nucleotide sugars metabolism	Glutathione metabolism	Vesicular traffic
Pyruvate metabolism	N-Glycans biosynthesis/ degradation	Ubiquitin mediated proteolysis
Glyoxylate and dicarboxylate metabolism	Keratan sulfate biosynthesis	Proteasome
Butanoate metabolism	Lipopolysaccharide biosynthesis	Storage proteins
Inositol metabolism	Glycosphingolipid metabolism	***2.4 Replication and Repair***
Inositol phosphate metabolism	***1.8 Hormon synthesis & answer***	***2.5 Transposons***
1.2 Energy Metabolism	***1.9 Metabolism of Cofact. and Vitamins***	**3. Environmental Information Processing**
Oxidative phosphorylation	Riboflavin metabolism	***3.1 Membrane Transport***
ATP synthesis	Vitamin B6 metabolism	ABC transporters
Regulation of energy metabolism	Nicotinate and nicotinamide metabolism	Ion channels
Photosynthesis Photosystem II	Pantothenate and CoA biosynthesis	Major facilitator superfamily (MFS)
Photosynthesis Photosystem I	Biotin metabolism	Lipid transport
Photosynthesis Cytochrome b6/f complex	Folate biosynthesis	Nuclear transport
Chlorophyll metabolism	One carbon pool by folate	***3.2 Signal Transduction***
Carbon fixation	***1.10 Biosynthesis of Second. Metabolites***	Nucleotid-binding protein
Reductive carboxylate cycle (CO2 fixation)	Diterpenoid biosynthesis	Membrane proteins
Sulfur metabolism	Limonene and pinene degradation	**4. Cellular Processes**
1.3 Lipid Metabolism	Stilbene, coumarine and lignin biosynthesis	***4.1 Cell cycle&chromatin remodelling***
Fatty acid biosynthesis (path 1)	Flavonoid biosynthesis	Cell cycle
Fatty acid metabolism	Erythromycin biosynthesis	Histone
Glycerolipid metabolism	Alkaloid biosynthesis I	Histone modification
Lipid storage	Alkaloid biosynthesis II	Chromatin assembling
1.4 Nucleotide Metabolism	***1.11 Biodegradation of Xenobiotics***	***4.2 Cell Growth and Division***
Purine metabolism	gamma-Hexachlorocyclohexane degradat.	Cell growth
Pyrimidine metabolism	1,2-Dichloroethane degradation	Cell structure
1.5 Amino Acid Metabolism	Tetrachloroethene degradation	Cell wall metabolism
Glutamate metabolism, Aminotransferases	Styrene degradation	Regulation of cellular processes
Alanine and aspartate metabolism	1,4-Dichlorobenzene degradation	***4.2 Cell Death /Apoptosis***
Glycine, serine and threonine metabolism	Benzoate degradation via CoA ligation	***4.3 Adaptation to atypical conditions***
Phenylalanine, tyrosine and tryptophan biosynthesis	Benzoate degradation via hydroxylation	***4.4 Disease and Defence***
Methionine metabolism		**5.Unclassified and unknown**

4.7.6 Distribution of ESTs among functional categories

Figure 30 provides a breakdown of representation by major KEGG categories and GO terms. The largest proportion of ESTs was assigned as 'Unclassified and unknown' (29.3% of total). Furthermore, four large groups emerged with 'Genetic Information Processing' (27.1%), 'Metabolism' (24%), 'Cellular Processes' (10.1%) and 'Environmental Information Processing' (9.5%) (Fig. 30a). Distribution within each individual category was as follows: the largest fraction of EST in category 'Metabolism' was associated with the function 'Carbohydrate Metabolism' (26%), followed by 'Energy Metabolism' (25%) and 'Amino Acid Metabolism' (17%). The other functions in the same category are represented a minor portion of ESTs only: 'Lipid Metabolism' (7%), 'Biosynthesis of Secondary Metabolites' (6%), 'Metabolism of Other Amino Acids' (5%), 'Biodegradation of Xenobiotics' (5%), 'Nucleotide Metabolism' (4%), 'Hormone Synthesis' (3%), 'Metabolism of Cofactors and Vitamins' (2%) (Fig. 30b).

In the category 'Genetic Information Processing' the most abundant molecular functions assigned were 'Transcription, Translation' (28%) and 'Protein Folding, Sorting, Storage and Degradation' (21%). In the same category the processes 'Replication and Repair' and 'Transposons' constituted only a small fraction (1% each) (Fig. 30c, blue parts).

The following distribution was observed across category 'Environmental Information Processing': 'Transport' occupied 5% and 'Signal transduction' 15%. Rather small amount of ESTs were assigned to the category 'Cellular Processes': 'Cell Cycle' (10%), 'Cell Growth and Division' (7%), 'Cell Death' (1%) 'Adaptation to Atypical Conditions'(1%), 'Disease and Defense' (3%) (Fig. 30c, orange and green parts of the pie chart).

4.7.7 Calibration of the experimental system

We applied high-density cDNA macro-arrays on nylon membranes to investigate level of mRNA expression at different stages of seed development. The main aim of macroarray hybridization was to compare the transcript profiles during seed development of WT and A SnRK1 antisense transgenic line of *P. sativum* to find out the genes and to identify the processes regulated by SNF1.

The first step in the analysis of macroarray data is the development of a hybridization model relating intensities of radioactive signals to mRNA abundance.

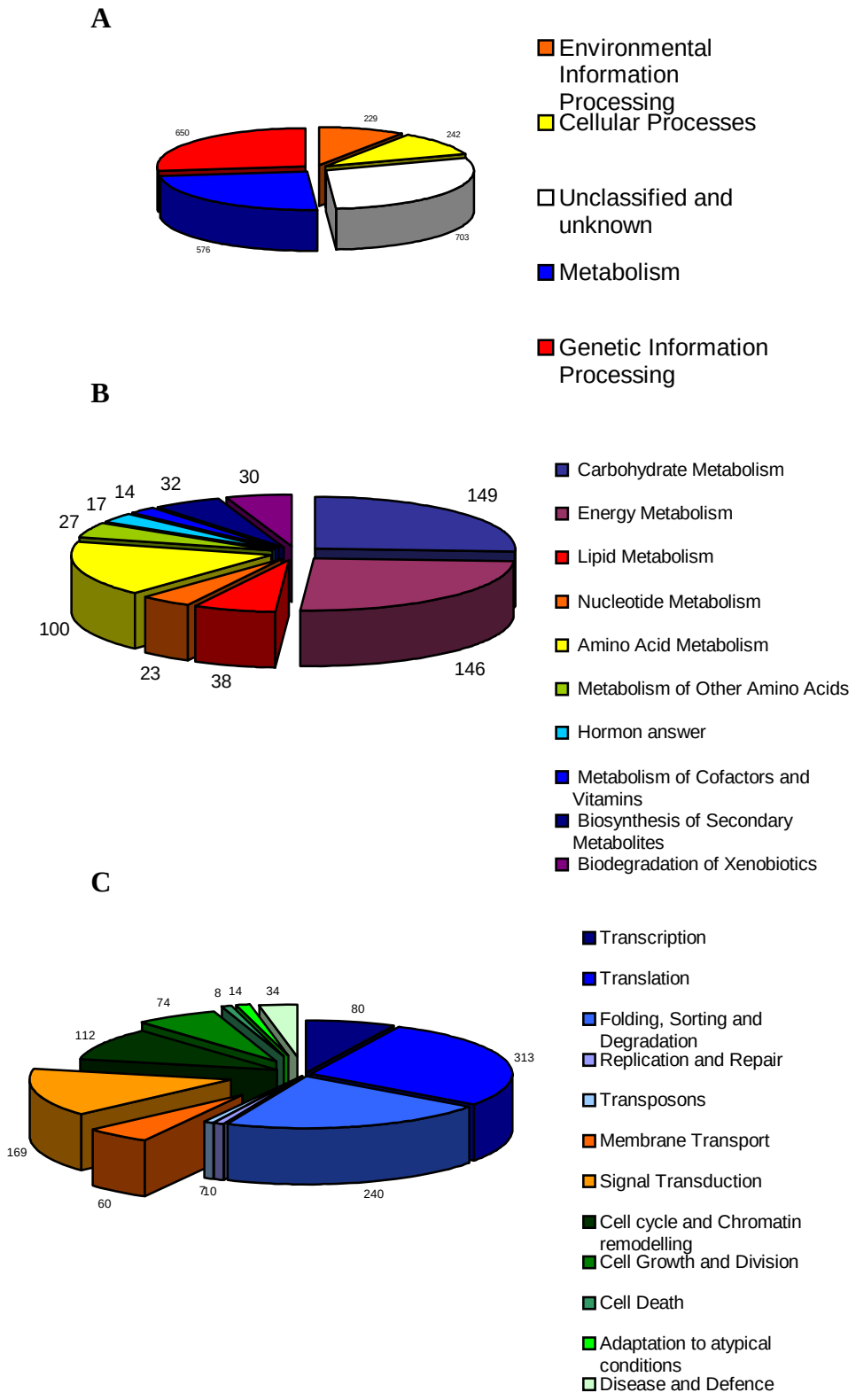


Figure 30. Annotated KEGG Orthology terms in the *P. sativum* EST libraries. **(A)** Gene ontology electronic annotation of UniGenes set from *P. sativum*. **(B)** Annotation associated with 'Metabolism' identified in the *P. sativum* EST sequence dataset. **(C)** Annotation associated with 'Genetic Information Processing' (*blue*), 'Environmental Information Processing' (*orange*) and 'Cellular Processes' (*green*).

For selection of optimal DNA concentration for spotting, a test- filter with 129 double-spotted clones in three DNA dilutions was prepared with final DNA concentration of 120-200 ng/µl, 60-100 ng/µl and 30-60 ng/µl. The test filter was hybridized with a ^{33}P-labeled average probe. The hybridization signals were detected using a phosphoimager. Resulting images were processed with a ArrayVision^{TM} software package for spot detection, and data files were exported to a standard spreadsheet program Excel. log_2(ratio) of measured signals was used to represent the expression levels.

As shown in the Figure 31, the signal saturated more quickly at concentration of 120-200 ng/µl. Moreover, at this concentration the some cases intensity of hybridization signals was lower than at higher dilution. At low DNA concentration (30-60 ng/µl) the weak signals disappeared and strong signals did not reach the threshold of sensitivity.

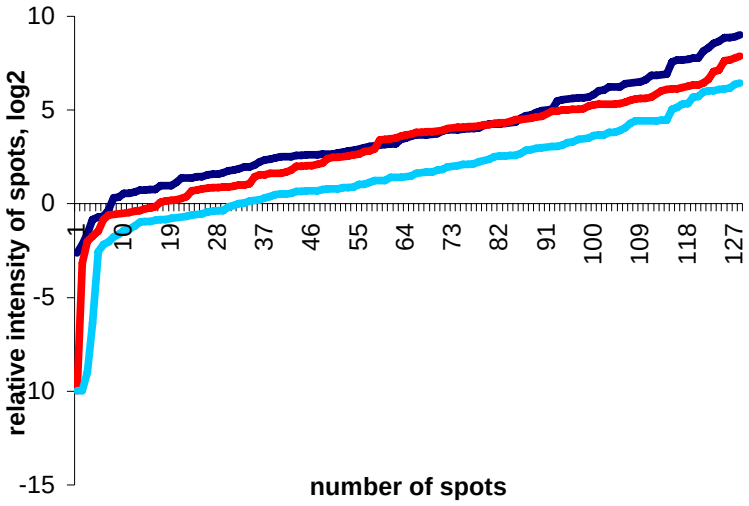


Figure 31 *log2* values of hybridization signals after intensity-dependent normalization. *Deep blue*: intensity of array signals at concentration 120-200 ng/µl DNA per spot; *blue*: intensity of array signals at concentration 60-100 ng/µl DNA on spot; *light blue*: intensity of array signals at concentration 30-60 ng/µl DNA on spot.

Therefore, the optimal concentration of DNA for spotting was determined to be 60-100 ng/µl. At this concentration, the weak signals were still measurable, and the strong signals were not saturated. The reproducibility between the two spots double was sufficient and the Pearson correlation factor 0.93 was within the limits of reliability (data not shown).

In order to produce an array, inserts of 4548 cDNA clones, representing 4018 unique clones and some additional internal control cDNAs were amplified with the vector specific M13-21PE and MVR-26 primers. The PCR-fragments were purified, visualized on agarose gel and

spotted in duplicate onto nylon membranes with a BioGrid robot. mRNA samples, isolated from WT pea embryos of six different stages of development (11, 13, 15, 17, 19 and 21 DAP) and from embryos of the transgenic line *Vic34* in four different stages of development (13, 15, 17 and 19 DAP), were labelled and hybridised onto the DNA array filter. The developmental stages correspond to the end of pre-storage phase and beginning of storage phase.

A comparison of the hybridization results obtained after labeling of two independent RNA probe preparations and hybridization with two different membranes is shown in a scatter plot (Fig. 32). Almost all signals show a less than threefold deviation from the diagonal. However, it is also evident that the accuracy decreased with lower signal intensities. Therefore, signals below background in all experiments were excluded from further evaluation.

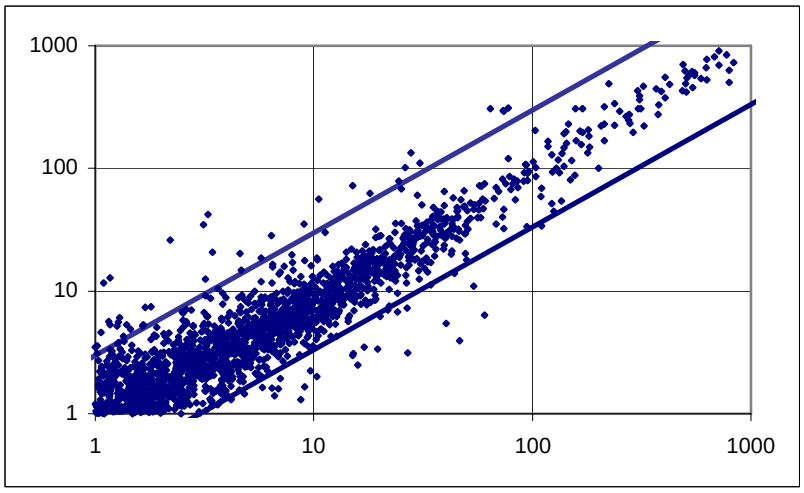


Figure 32. Comparison of gene expression in two independent experiments: embryo 19 DAP *vs* embryo 19 DAP. Dashed lines indicate threefold up- as well as down-regulation. On X- and Y-axe shown the measured spot intensity.

The reliability and validity of the array experiments were checked by comparison of the expression patterns of internal controls of genes presented more than twice on the array but derived from different cDNA clones of the same genes. They were in a case of high abundant genes. All clones corresponding to the same gene showed a similar trend of expression and their expression pattern can be interpreted with considerable accuracy.

Statistical processing and normalization was performed as described in Material and Methods, Chapter 3.11.4.

Pearson correlation changes in limits should be between 0.9-0.99. 131 genes showing 3-fold (34% from all selected profiles) and 4-fold (66% from all selected profiles) differences in temporal expression profiles were selected for the analysis. To allow comparison of signal intensities across experiments the median of the *log* scaled intensity distribution for each experiment was set to zero (Eisen et al., 1998).

4.8. Gene expression changes during seed development. Differences between WT and transgenic *SnRK1* antisense line

The temporal expression patterns were composed of different clusters with similar profiles by applying the k-means clustering algorithm. Cluster analysis visualised two clearly distinct groups of genes. Genes which are higher expressed in WT than in transgenic seeds during the beginning of the transition phase (DAP 13-15) belong to cluster A (Figure 33A). Cluster B consists of genes, which are down-regulated in WT seeds during the storage phase(DAP 17-19), show a prolonged expression in transgenic seeds (Fig. 33B).

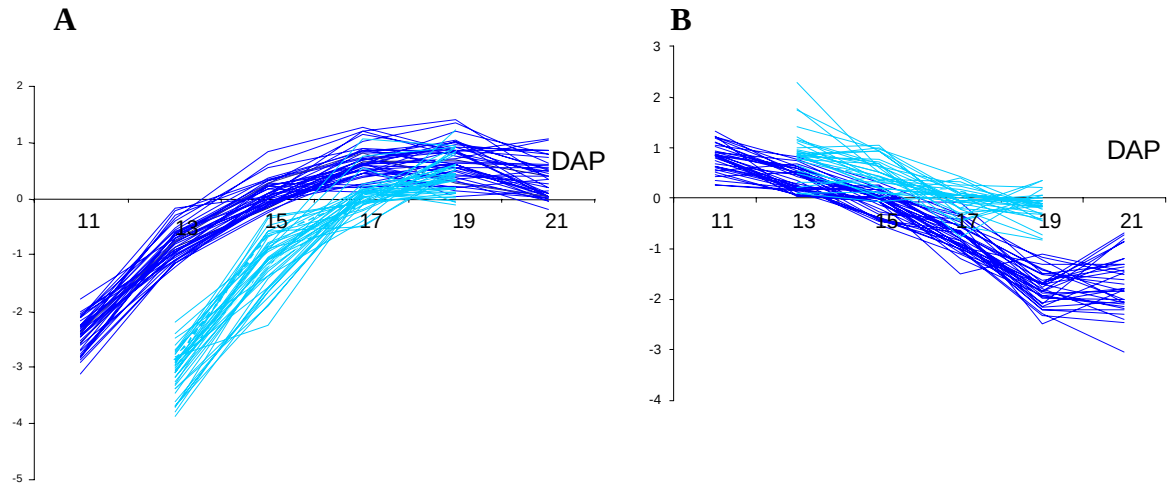


Figure 33. Temporal patterns of gene expression in developing embryo (11, 13, 15, 17, 19 and 21 DAP) of WT (*blue lines)* and transgenic (*light-blue lines)* SnRK1 antisense Vic34 line (13, 15, 17 and 19 DAP) of *P. sativum*. **Cluster A** represent a group of genes, which are generally up-regulated during the transition phase, but showing a later induction in transgenic seeds. **Cluster B** represent a group of genes, which is down-regulated during the transition phase, but shows prolonged expression in transgenic seeds.

All genes in clusters A and B were annotated according to the classification described above. The full list of genes is presented in Appendix, Chapter 9.1.

Cluster A consists of 57 genes from which 49 were unique. 12 are unannotated or show low E-value. From 45 annotated genes, 6 encode unknown proteins, 4 are unclassified and 35 can be functionally categorized into 13 functional classes.

Genes associated with carbohydrate and energy metabolism, like alcohol dehydrogenase, gluconolactonase, malate dehydrogenase and phosphoribulokinase as well as different glycosyltransferases, are present in this cluster. The most widely presented genes are connected with amino acid synthesis, protein biosynthesis destination and storage, such as acetohydroxy acid isomeroreductase, different regulators of transcription, one ribosomal protein, aminoacid-tRNA synthetase, embryonic abundant protein, vicilin and 26S proteasome regulatory subunit.

Some genes, regulating cell cycle and cell growth, like translationally controlled tumour protein, histone H3, DEAD box RNA helicase and cellulose synthase, also belong to this cluster. To the same cluster belong additional genes annotated to the function 'Signal transduction' (proton pump interactor, ATP binding protein, protein phosphatase 2C and small G-protein) and some unclassified genes (pirin, lectin and cold-inducible protein)

Cluster B contains74 genes, them 67 of were unique. A large group of genes (24) from this cluster is unannotated, and from the rest 50 genes 7 were with unknown function and 43 genes represent 10 functional classes. The most widely presented genes were annotated to the function 'Genetic information processing': 7 different ribosomal proteins, poly(A)-binding protein, RNA-binding protein and peptidyl isomerase. A set of genes encodie enzymes of carbohydrate and energy metabolism (isocitrate dehydrogenase, ATP synthase, transketolase, fructose-bisphosphate aldolase, cytochrome c oxidase) and are connected to photosynthesis (plastocyanin and chlorophyll a/b-binding protein) belong to this cluster as well. Furthermore, several genes encoding components of cell cycle (all histones except histone H3) and genes assigned to cell division (annexin and tubulin) are represented in this cluster. An analogous picture emerged for genes encoding transporters (aquaporin water channel, dicarboxylate/tricarboxylate carrier, adenine nucleotide translocator, sugar transporter, acyl-CoA binding protein and HEM protein) and for genes required for signal transduction and defence (ADP-ribosilation factor, calmodulin, putative protein kinase and polygalacturonase inhibitor).

Distinct groups of co-regulated genes and genes with very similar expression profiles were manually extracted from the clusters and investigated in detail for the presence of possible interaction between processes and interconnected regulation and are discussed in the next chapter.

5. DISCUSSION

5.1 Cloning and characterisation of *V. faba* and *P. sativum* SnRK1 and β subunits of SnRK1 complex in *P. sativum*

In this study, the full-length cDNA fragment corresponding to *SnRK1* gene termed *VfSnRK1* was isolated from a cDNA library from *V. faba* developing cotyledons, whereas the *P. sativum SnRK1* was identified as a full-length cDNA from a *P. sativum* EST collection.

The gene encoding the SNF1 protein kinase was identified first from *S. cerevisiae* via mutation which did not allow *snf1*-mutants grow on sucrose or raffinose (Carlson et al., 1981). SNF1 is activated in response to low cellular glucose levels and required for the derepression of genes repressed by glucose (Gancedo, 1998). Generation of *snf1*-deficient yeast mutants provided a useful way for the isolation and characterisation of *snf1* genes from different organisms by functional complementation of the mutation. Thus, the *LeSNF1* from *L. esculentum* complemented a *snf1* mutant, restoring both growth on Suc and derepression of invertase and indicating that it encodes a functional kinase able to respond to Glc repression signals in yeast (Bradford et al., 2003).

Both genes from *V. faba* as well as from *P. sativum* exhibit high predicted amino acid sequence homology to each other (98.8 %), as it was expected from their close taxonomic relationships. 84.5% to the snf1 gene of *G. max* (Patil et al., 1999), 81.7% to the snf1 gene of *L. esculentum* (Bradford et al., 2003), 81,1% to the NPK5 of *N. tabacum* (Muranaka et al., 1994), 30.3% to the SNF1-like kinase of *H. sapiens* (Ruiz et al., 1994) and 41.5% to the SNF1 of *S. cerevisiae* (Celenza and Carlson, 1986).

The protein consists of a putative kinase catalytic domain in the N-terminal part of both proteins and contains 11 subdomains common to protein kinases (Hanks et al., 1988). The Thr can be phosphorylated in conserved subdomain VIII. Mutation of this residue to alanine results in a complete loss of *SNF1* gene function in yeast (Estruch et al., 1992). The UBA domain (Ubiquitin binding domain) is immediately adjacent to the kinase domain and is responsible for non-covalent binding to ubiquitin (Hofmann et al., 1996). Within the putative regulatory C-terminal part there is a domain that shows low sequence homology to kinase associated domain 1 (KA1).

Southern blot analyses revealed *VfSnRK1* and *PsSnRK1* as single copy genes. Apparently, there is only one copy of the SnRK1 gene in legumes. This confirms results received by Patil

et al. (1999). Also, by cloning the SNF1 homologue from *G. max* only one gene of *SnRK1* was found. However, by hybridization under low stringency different patterns with multiple bands were obtained. This indicates the possible presence of other members of the SnRK family.

The yeast SNF1 kinase is a heterotrimeric complex composed of a catalytic subunit α that interacts with two other subunits: the SNF4 (γ) and one of the β subunits (SIP1, SIP2 or GAL83) (Jiang and Carlson, 1997). As in yeast, plants also possess several isoforms of the β and γ subunits. Four isoforms of the β subunits and three γ subunits were identified in *M. truncatula* seeds (Buitink et al., 2004). Based on similarity of amino acid sequences, the *P. sativum* EST collection was screened for the possible presence of β and γ subunits of the *SnRK1* complex. A full-length cDNA corresponding to homologues of *MtAKIN β1* (93.1% identity) and a partial cDNA similar to *MtAKIN β2* (89.8% identity) were isolated from the *P. sativum* EST collection. Both proteins consist of conserved KIS (kinase association) and ASC (association with SNF1 complex) domains, which bind to the subunit SNF1 and SNF4 respectively in yeast.

5.2 The *VfSnRK1*, *PsSnRK1* and *PsAKIN β1* subunits have similar expression profiles during seed development and *V. faba* SnRK1 kinase activity correlates with gene expression

Northern blot analyses revealed that transcripts of *VfSnRK1* are most abundant in roots, pods and gynoecia, less abundant in leaves, and hardly detectable in seed coat. Since organs with high *SnRK1* expression are assigned as sink-organs, a connection between *SnRK1* signal transduction and sink-activity is indicated.

The expression pattern of *VfSnRK1* and *PsSnRK1* during seed development is similar. Our interest is concentrated on two phases of seed development: organ formation and organ maturation. *SnRK1* expression levels were highest at the early stages of seed development and decreased fluently during seed storage phase.

Surprisingly, the transcript profiles of *PsAKINß1* were correlated with those of *PsSnRK1* in WT (Fig. 25). This indicates a similar mechanism of the regulation of *PsSnRK1* and *PsAKINß1* gene expression. The investigation of *MtAKINß1* expression during seed germination in *M. truncatula* showed increasing transcript amounts after osmotic treatment and starvation (Buitink et al., 2004). Also in *Arabidopsis* leaves, the *AKINß* transcripts accumulated in the dark, a condition that affects carbon availability (Bouly, 1999). It is known

that sugar concentration decreases during legume seed development (Weber et al., 1996). Therefore, expression of *PsAKINß1* is either independent of sugar concentration, or the regulation of transcription in *sink*-organs has another mechanism as in *source*-organs. High similarity between the nucleotide sequences *PsAKINß1* and *PsAKINß2* and the absence of the 5'-prime part of the *PsAKINß2* gene makes a possible cross hybridisation by northern blot analysis. Therefore, an individual expression analysis of *PsAKINß2* was not carried out.
The temporal profiles of SnRK1 enzyme activity during seed development correlate with transcript amounts, but the maximum of activity shifted towards two days later. Perhaps, there is an additional mechanism of the regulation of *SnRK1* activity. For example, the yeast snf1 can be phosphorylated on threonine 210 and than activated with upstream kinases and with SNF1 itself (Celenza et al.,1986). Indeed, *SnRK1* and *AKINß1* genes are probably involved in regulation of processes of early embryogenesis **.**

5.3 Changes in pea seeds caused by antisense-repression of SnRK1

Antisense technology genetics is a suitable tool to investigate the possible functions of SnRK1 during pea seed development. This work shows that the antisense repression of the *P. sativum* gene coding for *SnRK1* leads to morphological and or developmental modifications in seeds. The *VfSnRK1* antisense construct was expressed in tobacco and pea seeds under the vicilin- or USP-promoter, which active during storage phase.
Since the production of transgenic pea plants is still not a routine process, experiments with tobacco plants, a more simple model, were carried out in parallel with pea plant transformations to check the possible effect. Taking in to account the high nucleotide identity and the homology of regulatory mechanisms between yeast and plants (Buitink et al., 2004), and also taking into account the USP- and vicilin- promoter activities in tobacco seeds (Baumlein et al., 1991; Jiang et al., 1996) the occurrence of a phenotype could be expected. We found that about 30% tobacco and pea plants showed either partial or complete sterility. Cytological observation of the pollen indicated abnormal pollen development: pollen grains were shaped, contained small amount or no starch and were smaller than pollen from the control plants. Similar results have been found after antisense expression of *SnRK1* in barley plants that were controlled by seed-specific Glu-1D-1 promoter (Zhang et al., 2000). Pollen from antisense SnRK1 barley plants contained small amount of starch and ovule development was also affected resulting in sterility. We did not expect such dramatic effects by using seed-

specific promoters. However, new and more detailed research revealed activity of a pea and *V. faba* seed specific vicilin- and USP- promoters in pollen (Zakharov et al., 2004). The T_0 complete sterile tobacco and pea lines did not produce seeds and could not be analysed further. Northern blot analysis of RNA isolated from mature anthers of transgenic tobacco lines revealed an obvious reduction of SnRK1 transcripts in sterile lines.

Further research on the phenotype focused on transgenic pea lines which were not sterile, but showed a seed growth alterations. Pods of transgenic pea lines contain two clearly distinguishable types of seeds. After fertilisation all seeds started to develop, but during different time intervals, a part of the seeds was arrested in growth and died. However, the defects in embryo development were already visible at very early stages. Growth of seeds, which continue to develop, lagged behind in comparison to wild type seeds.

In many cases mature seeds were wrinkled and smaller in size. The number of seeds in a mature pod was reduced. This could be explained as insufficient seed feeding. The seeds from one line showed premature germination (vivipary), that could indicate the defective transfer of a hormonal signal during seed maturation (Finkelstein et al., 2002). Northern blot analysis of RNA from these seeds revealed reduced SnRK1 transcripts levels compared to WT pea seeds. Additioaly, the specific SnRK1activity of has also been reduced.

To summarize, we obtained pea lines with up to 50% reduced SnRK1 transcript amount and with up to 70% repressed of SnRK1 kinase activity. Application of the method of macroarray hybridisation allowed us to estimate in details changes in gene expression, resulting by repression of SnRK1 kinase.

5.4 Global changes in transcripts profiles during seed development in a transgenic *SnRK1*-antisense pea lines

The "*omic*" experimental approaches have great value to get a global, comprehensive knowledge of the molecular events of regulators processes during seed development and provide information for a deeper understanding of different aspects of gene regulation.

In addition, the established EST collection provides an opportunity for a fast search and easy cloning of new genes of interest.

5.4.1 High abundant ESTs from a P. sativum seed EST library (quantitative estimation)

A total of 8414 good quality ESTs, which were assembled in 1082 clusters (1465 contigs) with 6061 sequences and 2353 singletons were derived from two cDNA libraries of developing pea cotyledons and seed coat. The search of known gene families *in silico* offers an approach for discovering new genes. Determining the validity of using *in silico* expression data as a true reflection of *in vivo* transcript abundance is extremely important. A rigorous statistical test developed by Audic and Claverie (1997) delineates more precisely and extends the limits within which *in silico* expression data can be used confidently. To demonstrate the differential expression of a gene, its EST assembly must be composed of more than four ESTs to be considered as having greater than a basal level of expression. Recent publications demonstrate the efficiency of this technique to find new genes (Ewing et al., 1999; Mekhedov et al., 2000). Using the different already known cDNA and protein sequences in plants as bait for a BLASTN or a TBLASTN search in the plant genome databases, the representation of high abundance genes in *P. sativum* EST libraries was discovered. The purpose of those investigations was a quantitative definition of frequency of high abundant and rare genes, as well as a definition of genes involved in developments.

The stages of seed development for construction the *P. sativum* EST library have been chosen in that way, to cover two main developmental processes: early embryogenesis and start of storage phase. Thus, widely represented genes in the collection belong to cell division, regulation, synthesis and maturation of storage proteins. Some genes are involved in embryonic photosynthesis or other genes are associated with cell expansion and desiccation. Embryogenesis starts with a morphogenesis phase during which the embryo passes through several distinct stages (globular, heart and torpedo) during which all embryo structures are formed. At the end of embryogenesis, cell division arrests and the seed differentiates and accumulates the storage components (Raz et al., 2001). The abundance of several proteins decreased in the transition from cell division and differentiation. This was the case for annexins, and there is strong evidence that annexins are involved in cell division. For example, they accumulate during the cell cycle and peak at the end of mitosis in tobacco (*N. tabacum*) cells (Proust et al., 1999). Also, it is possible to attribute gene with high mRNA abundance which similar to gene from *Cicer arietinum* named *Cap28*. High expression of this gene was detected in early stages of seed development in pea. The direct function of *Cap28* is unknown, but the transcription profiles allow to assume that the gene is associated in development with organs undergoing cell division and cell elongation and that it may be involved in one of these processes (Romo et al., 2002).

After cessation of cell division, seed storage compounds are synthesized. The most abundant ESTs, which are mainly associated for protein storage in legume seeds, were identified to encode vicilins and legumins. The high abundance of ESTs corresponding to chlorophyll *a*/*b*-binding protein and photosystem II reaction center protein reflects the photosynthesis activity to produce both energy-rich compounds and oxygen (Borisjuk et al., 2003).

The processes of intensive cell division, expansion and storage are coupled with active protein synthesis. Different histones and ribosomal proteins are highly abundant during early embryogenesis for example genes involved in cell cycle or protein biosynthesis. Some transcripts, whose abundance varied during seed filling, corresponded to enzymes involved in methionine (Met) biosynthesis. Among the essential amino acids synthesized in plants, Met is a fundamental metabolite because it functions not only as a building block for protein, but also as the precursor of Ado-Met, the primary donor of methyl-groups and the precursor of polyamines and the plant ripening hormone ethylene (Ravanel et al., 1998). 24 ESTs corresponding to the enzyme Met synthase catalyzing the last step of the *de novo* biosynthetic pathway of Met are present in the library.

Furthermore, three genes with unknown functions are highly abundant in pea seed. One of them was defined as embryonic abundant protein similar to USP (unknown seed protein) from *V. faba*, which reflects the most abundant mRNA present in cotyledons during early seed development. However, the corresponding protein does not accumulate in cotyledons (Baumlein et al., 1991). This protein contains a characteristic BURP domain (Acc. Nr. Pfam03181.10), which was also found at the C-terminus of several different plant proteins. The function of the domain is unknown (Hattori et al., 1998). A second gene containing a BURP domain is widely represented in the *P. sativum* EST library and is possibly connected to seed maturation. About 52 ESTs correspond to this gene, encoding the dehydration-responsive protein similar to RD22 from *Prunus persica*.

Another gene is encoding a protein with weak similarity in its C-terminal part to the embryonic protein BP8 from *Betula pendula*. However, there are no other homologues in protein and domain databases, and there is only one similar DNA sequence from *M. truncatula* without any connection to possible function.

The list of high abundant ESTs corresponds to the highly expressed proteins identified by proteomics analysis in *M. truncatula* seeds (Gallardo et al., 2003).

5.4.2 Putative functional distribution of ESTs in classes (qualitative estimation)

4548 EST clones, containing 4018 unique clones' cDNA fragments were used for preparation of array filters. To obtain qualitative information about ESTs, the BLASTX results were annotated according to the description. The aim of protein-level annotation is to assign functions to the products of the genome. Databases of protein sequences and functional domains and motifs are powerful resources for this kind of annotation. Nevertheless, about half of the predicted proteins in a new genome sequence tend to have no obvious function. (Stein, 2001). Therefore, it is necessary to remember, that annotation often delivers only a theoretical probability for a function. Understanding the function of genes and their products in the context of cellular physiology is the goal of process-level annotation. Finally, the reliability of the conclusions obviously depends on the quality of the available gene annotation.

The annotation was performed at three levels: automatic annotation, manual annotation and search for the logic connection among components.

The automatic annotation includes analysis of some parameters determining the reliability of annotation, like a score and E-value. At this level it was already possible to define genes, unique for species as well as genes with still no identified functions. From a total of 4548 ESTs clones from *P. sativum* seed EST library about 18% seem to be specific for *Pisum.* Since there are no proteins similar to the translation products were presented in the public databases at the time of analysis. From the remaining 82 % only little more than half part could be securely assigned. Another half is named "low-homology". This class would include novel proteins as well as proteins with known functions. Reason for the low BLASTX score could be low quality of some sequences. The comparison of E-values of BLASTX hits and average lengths of ESTs between cotyledons and seed coat EST library showed that ESTs from seed coat library are generally shorter and accordingly the BLASTX E-values are lower. Anyway, a large part of all ESTs could not be annotated. A similar observation was made for ESTs from *Arabidopsis* seeds (White et al., 2000).

ESTs with secure annotation (2391 clones) were taken for a second approach called "functional classification", which means classification according to metabolic and biological processes. A part of the functional classification can also be performed automatically *via* extraction of data linkage graphs (DLGs) from the functionally categorized KEGG and GO databases. 16% from 2391 securely assigned ESTs could be annotated to encode enzymes and 3% ESTs were associated to the biological processes. The automatic classification facilitates

the annotation process, but requires an additional manual correction. Some trivial errors generated by KEGG arose because of the non-plant pathways imports. Moreover, only a minor part of the genes can be categorized in this way because links to KEGG or GO databases in NCBI accessions gene entries are missed. Thus, the basic part of the functional classification was carried out manually.

For enzymes which are known as involved in more than one pathway retained only one. The reason for it was the simplification of primary classification, though it is definitely not an optimum choice and it misrepresents the reality. It should be taken into account in conclusions.

The ESTs were annotated according to four large functional groups: 'Genetic Information Processing', 'Metabolism', 'Cellular Processes' and 'Environmental Information Processing'. The majority of ESTs belongs to categories 'Genetic Information Processing' and 'Metabolism'.

High EST abundance were assigned to molecular functions such as 'Transcription, Translation' and 'Protein Folding, Sorting, Storage and Degradation' in the category 'Genetic Information Processing' indicating the beginning of storage processes in developing seeds.

The distribution within the category 'Metabolism' showed that the largest fraction of EST was associated with the function 'Carbohydrate Metabolism', followed by 'Energy Metabolism' and 'Amino Acid Metabolism', whereas other functions of the same category ('Lipid Metabolism', 'Biosynthesis of Secondary Metabolites', 'Metabolism of Other Amino Acids', 'Biodegradation of Xenobiotics', 'Nucleotide Metabolism', 'Hormone Synthesis' and 'Metabolism of Cofactors and Vitamins') represented only a small portion. This reflects the high activity of metabolic processes, which are characteristic for the period of early embryogenesis.

A low representation of ESTs was observed across the category 'Environmental Information Processing', including the functions 'Transport' and 'Signal transduction'. Regulatory genes are not as highly expressed as storage protein genes or genes that are essential for general cell metabolism. It confirms that the observed abundance of ESTs is in agreement with common knowledge about the biology of plant cells and of developing seeds in particular (White et al., 2000).

Additional information are necessary for a deeper understanding of the results received by hybridization about regulation of expression and possible connections between pathways. Such information was derived from literature for all selected genes.

5.5 Differential gene expression in response to SnRK1 down-regulation in pea seeds

Macroarray analysis was used to identify differentially expressed genes in WT and SnRK1 antisense transgenic pea seeds. Signal intensities (reflecting relative mRNA levels) of 131 differentially expressed cDNA fragments were subjected to cluster analysis. Two major groups of cDNA fragments, showing distinct expression patterns during pea seed development were created by k-mean clustering. The first cluster of EST profiles represents a group of genes, which are up-regulated in WT seeds during the transition phase. Transgenic seeds showed retarded induction of these genes. 57 profiles belong to this cluster. 74 profiles build a second cluster, which represented groups of genes, which are down-regulated in WT during the transition phase but showed prolonged expression activity in transgenic seeds. The macroarray analysis provides a good opportunity to compare simultaneously the large number of genes and allows to estimate complex changes and to find possible connections to regulatory processes.

However, macroarray analysis allows to estimate the transcripts amount of genes, but does not provide any information about its activity and about factors entailing changes of the transcription levels. A main problem arising by handling of the results is to recognize primary and secondary consequences, caused by repression of the *SnRK1* gene. A comparison of the results received by macroarray hybridisation and the information about the experimental confirmation of SnRK1p functions from various sources, recreates results in a more consistent picture of the consequences caused by repression of this gene.

SnRK1 is involved in the regulation of various cellular processes. Depending on recognition signals, SnRK1p is capable to interact with different subunits of the SnRK1-complex and possible recognises different substrates (Bouly et a., 1999). SnRK1 can regulate cellular processes at least on three different levels. Firstly, SnRK1 regulates gene transcription on the level of chromatin remodelling by regulation of histone H3 (Lo et al., 2001). Secondly, regulation can be performed directly on the transcriptional level, by regulating of transcription factors activity. SnRK1 regulates transcription through direct interactions with the transcription apparatus (Kuchin et al., 2000). Thirdly, the regulation can be carried out on the posttranslational level by modulating the activity of various enzymes: 3-hydroxy-3-methyl glutaryl CoA reductase, nitrate reductase (NR) and Suc phosphate synthase *in vitro* (Sugden et al., 1999b), as well as participating in ubiquitin-mediated proteolysis.

The main part of the above mentioned ways of regulation has been elucidated in yeast. Similar mechanisms are expected in plants.

5.5.1 SnRK1 regulates the transition from cell division to the protein storage

5.5.1.1 SnRK1 is involved in cell cycle regulation probably by mediation of brassinosteroid signal transduction.

Glucose repression plays a major role in the regulation of carbon metabolism in higher plants as well as in other organisms (Sheen, 1990; Stitt and Sonnewald , 1995). Feeding of plants with glucose or sucrose (which is converted to glucose and fructose) results in either transcriptional or post-transcriptional down-regulation (or both) of genes involved in chlorophyll biosynthesis, Calvin cycle, gluconeogenesis, starch degradation and glyoxylate cycle, but leads to the activation of genes for glycolysis, defence responses, nitrate and phosphate metabolism, biosynthesis of anthocyanin pigments and storage proteins (Koch, 1996; Jang and Sheen, 1997).

It is impossible to explain the effects obtained after repression of SnRK1 without taking into account the connections between pathways and metabolite signalling in the cells. Characterization of the SnRK1 antisense repression indicated the existence of an intensive cross talk between glucose and cytokinin signalling at the early stage of seed development. Some genes, which were down-regulated in WT during the transition phase, showed prolonged expression activity in the transgenic line. Among them, there are genes encoding plastocyanin and chlorophyll a/b binding protein, which are known to be transcriptionally regulated on a sugar dependent manner.

The transcript levels of chlorophyll a/b binding protein (*CAB1*) and plastocyanin (*PC*) are regulated in *Arabidopsis* by the hexokinase-dependent pathway (Sheen et al., 1999). Expression analysis of *Arabidopsis PC* showed its strong repression in the presence of 7% glucose in wild-type plants. Light-grown seedlings show a similar transient and sucrose-sensitive increase in PC mRNA levels (Dijkwel et al., 1996). The expression of both genes seems to be regulated by transcription factor ABI4. Efforts are currently being directed to confirm whether ABI4 can directly bind to the *CAB1* and the *PC* promoters under high glucose conditions. In addition, the predicted amino acid sequence of ABI4 contains a serine/threonine-rich domain, which is a possible target for protein kinases including the SnRK1 family (Soderman et al., 2000). This reflects the ABA regulation of these genes. The coordination between ABA signalling and SnRK1 signal transduction will be further discussed in Chapter 5.5.1.4.

It is interesting that the SnRK1 promoter region contains the AATATTTTTATT motive at the position –299/-288 upstream of the start codon, similar to the promoter of the chlorophyll a/b binding (*cab-E*) protein from tobacco. This motive can be recognized by the AT-1 protein and is known to be regulated by light. Binding of AT-1 is modulated by phosphorylation (Cacchione et al., 1995; Datta et al., 1989). This can explain the possible coordinated expression of SnRK1 and chlorophyll a/b binding protein.

More details come from an earlier description of the light and hormonal dependent gene regulation of the genes in the prl1 *Arabidopsis* mutant (Salchert et al., 1998). The mutation affects potential regulatory functions in glucose signalling and was isolated by screening for plants showing hypersensitive or insensitive growth responses to glucose or sucrose. SnRKs are implicated in the pleiotropic regulation of metabolic, hormonal and stress responses through their interaction with PRL1. Progress was made to understand the signal transduction cascade by analyzing the function of the SNF1 kinase complex and the regulatory PRL1 (pleiotropic regulatory locus 1) protein. It has been discussed, that the observed changes in gene transcription can be explained by functional changes in the SnRK1/PRL1 complex.

In the *Arabidopsis PRL1* mutant the sensitivity for glucose is increased as a result of transcriptional derepression of glucose-regulated genes, among them the *CAB* gene.

PRL1 (an α-importin-binding nuclear WD protein) is considered to be a subunit of the SnRK1 complex, and the activity of SnRK1 is regulated by PRL1 (PRL1 operates as SnRK1 inhibitor).

In addition, the SnRK1/PRL1 complex is involved in cell cycle regulation by modulation of brassinosteroid signalling at the transcriptional and post-translation level.

On the one hand, the SnRK1/PRL1 complex regulates the expression of CPD (constitutive photomorphogenesis dwarf) gene, controlling the rate limiting step in brassinosteroid biosynthesis. Abnormal activation of the SnRK1 kinase in the *prl1* mutant correlates with cell elongation defects and down-regulation of the CPD gene in *prl1*-mutants.

On the other hand, it is known that SnRK1 phosphorylates and inactivates HMG-CoA reductase catalysing the first committed step in steroid biosynthesis. It is known that brassinosteroid induces cell elongation and cell division.

Finally, PRL1 modulates the binding of the SCF complexes and of SnRK1 to the 26S proteasome (Farras et al., 2001). This is a very important part of the mechanism of cell cycle regulation. It is described in more detail later (Chapter 5.5.1.3).

5.5.1.2 Regulation of cell cycle genes by SnRK1 at the level of chromatin remodelling

Different histones which are involved in the regulation of the cell cycle, chromatin remodeling and cell growth and division, showed changed transcriptional levels in transgenic pea seeds.

The histone H3 mRNA was down-regulated in transgenic seeds. It is known, that expression of this histone is clearly linked to DNA replication. A similar histone H3 gene was cloned from alfalfa (*M. sativa*) and described as a replication-dependent histone H3.1 variant protein (Robertson et al., 1997). Besides, it was found in yeast, that the Snf1-a histone kinase can modify the histone H3 serine-10 and lysine-14 together with the acetyltransferase Gcn5 (Lo et al ., 2001). The yeast transcriptional activator Gcn4 recruits Gcn5 to selective promoters, thereby creating local domains of histone H3 hyperacetylation and subsequent transcriptional activation (Kuo et al., 2000). The possible connection between Gcn4 and SnRK1 will be discussed in Chapter 5.5.1.5.

Histone H3 was the only one of all histones down-regulated in the transgenic pea seeds. Another 11 clones encoding different histones showed the expression profiles similar to histone H3 during seed development and were up-regulated in transgenic seeds at early stages of development. One clone encoding H1.41 histone had been previously described from *P. sativum* . In the C-terminal domain, there are four lysine-rich repeating motifs of SXK(T/S)PXKKXK which may be involved in chromatin condensation and decondensation. Southern blot analysis of nuclear DNA shows that PsH1-41 belongs to a multigene family (Woo et al., 1995).

Two clones encode different H2A histones. One clone encodes a protein which is similar to a H2B histone-like protein from *A. thaliana* and is also similar to a reproductive organ-preferential H2B histone from *Capsicum annuum* which is highly expressed in floral buds. CaH2B was shown to be induced by wounding, treatment by methyl jasmonate (MJ) (Kim et al., 1998). Seven clones encode the same protein which is similar to histone H4 from *T. aestivum* .

5.5.1.3 SnRK1 involved in ubiquitin-dependent proteolysis and cytokinin signal transduction

The detailed characterization of signalling intermediates has demonstrated that various branches of the signalling network are spatially separated and take place in different cellular compartments. In addition, proteasome-mediated degradation of signalling intermediates most probably acts in concert with subcellular partitioning events as an integrated checkpoint.

As already mentioned in Chapter 5.1, the *PsSnRK1* contains the UBA domain (Ubiquitin binding domain) responsible for non-covalent binding to ubiquitin. It was also described that SKP1/ASK1, a conserved SCF (Skp1-cullin-F-box) ubiquitin ligase subunit, which suppresses the *skp1-4* mitotic defect in yeast, interacts with the PRL1-binding C-terminal domains of SnRKs. SKP1/ASK1 is co-immunoprecipitated with a cullin SCF subunit (AtCUL1) and SnRK kinase. SKP1/ASK1, cullin and proteasomal alpha-subunits show nuclear co-localization in differentiated *Arabidopsis* cells and are observed in association with mitotic spindles and phragmoplasts during cell division (Farras et al., 2001).

It seems that reduction of SnRK1 activity in our transgenic plants caused the repression of ubiquitin-dependent proteolysis of proteins, involved in the promotion of cell cycle.

It is interesting that the expression of one clone defined as component of the ubiquitin-dependent protein catabolism, namely 26S proteasome regulatory subunit RPN11, was activated later in the transgenic SnRK1 antisense pea line compared to WT. RPN11 is a metalloprotein that possibly plays a role in deubiquination. Proteasomes in the RPN11 yeast mutant assemble normally but fail to either deubiquitinate or degrade ubiquitinated substrate *in vitro* (Verma et al., 2002). The late activation of RPN11 transcription can be either a direct consequence of SnRK1 repression or a secondary reaction to the prolonged phase of cell division in our transgenic pea seeds.

Except for that, two clones encoding different chains of tubulin, needed for microtubule organization and participating in the cell cycle, showed also prolonged activation of transcription in the transgenic line.

It is known, that both auxin (Gray and Estelle, 2001) and cytokinin (Smalle et al., 2002) signaling is mediated by regulation of protein degradation, and SnRK1s were detected in association with purified 26S proteosomes and with the Skpl-Cullin-F-box (SCF)-proteosomal complex co-localized with tubulin during mitosis.

Probably, the SnRK1 kinase modulates the function of the cell cycle progression apparatus, including cyclins and cyclin-dependent kinases. In *S. cerevisiae* the cyclin-dependent protein kinase, and the Snf1p protein kinase Pho85p act antagonistically in controlling glycogen

accumulation (Huang et al., 1998). Moreover, the Pho85 can phosphorylate and prepare the substrate for the ubiquitination complex SCF^{CDC4}, as it was described for the yeast transcription factor Gcn4, which promotes the expression of amino acid and purine biosynthesis genes (Shemer et al., 2002). The degradation of the Gcn4 by $Cdc34/SCF^{CDC4}$ is cell cycle regulated.

Thus, SnRK1 may, in fact, take part in the control of cell cycle by sugar and hormone (brassinosteriode, auxine, cytokinine) signal transduction, by regulating gene transcription and cell cycle dependent protein degradation. Interruption of signal transduction on the SnRK1- level entails the prolongation of the phase of early embryogenesis. It is clear, that SnRK1 is involved in a complicated signaling network and discussed explanations need more direct experimental support.

Additional evidence about the role of SnRK1 in cell cycle regulation was received by studying the SnRK1s 5'upstream region. Transient expression of the chimeric beta-glucuronidase (GUS) gene under control of the VfSnRK1 promoter was examined in cultivated protoplasts of *N. tabacum*. The promoter activity was highest at the appropriate concentrations of auxin and cytokinin, when cells were competent to divide. Cell cultivation in the presence of cytokinin or auxin alone was insufficient to induce cell division. But cultivation in the presence of only auxin or without any hormones did not repress the activity. In contrast, clear repression of GUS activity was shown by cultivation of protoplasts in a medium supplemented with cytokinin only. This arrests the cell cycle in G1/S phase. These data imply that SnRK1 may be involved in the regulation of the cell cycle either directly or indirectly. Specific up-regulation of GUS activity in response to abscisic acid (ABA) treatment and low sugar concentration was observed, indicating a critical role of the regulation of SnRK1 by phytohormones and sugar in legumes seed development.

5.5.1.4 Repression of SnRK1 causes the interruption of ABA signals

The transition between seed development and germination is connected to changes in gene expression patterns and metabolic pathways. An important regulator of seed maturation is abscisic acid (ABA). The genomic approaches with respect to ABA signaling in seeds allows to identify the transcription factors and proteins involved in the interactive network regulating late embryogenesis and seed development.

The macroarray analysis of gene expression profiles in SnRK1-antisense plants identified changes in transcription patterns of genes, known to be ABA regulated or involved in ABA signal transduction. Generally these genes showed a similar type of change: induction of transcription begins much later in the transgenic seeds. However, later it reaches a similar level than the WT plants. An exception from this pattern are CAB and PC gene, which show opposite profiles, but transcription of these genes also seems to be ABA regulated.

The clone encoding ABA insensitive transcription factor 3 (*ABI3*) is similar to previously described ones from *P. sativum* (unpublished). In *A. thaliana,* this seed-specific transcription factor plays a regulatory function during seed maturation (Mönke et al., 2004). The ABI3 protein is an important regulator of *Arabidopsis* embryo development (Koornneef et al., 1998) and works both as an activator of embryo maturation and as a repressor of germination (Kurup et al., 2000). An essential target of this regulator is the highly conserved RY motif [DNA motif CATGCA(TG)], present in many other seed-specific promoters of dicot and monocots including the promoters of the *V. faba* legumin and USP genes (Baumlein et al., 1986, Baumlein et al., 1991). The two clones encoding the USP gene were also found to be among the genes with changed expression patterns. Thus, it is quite possible, that these changes are caused by ABI3.

It is interesting that the SnRK1s 5'-flanking region contains two copies of a CAACA-motif, known as a binding site for RAV1 (Related to ABI3/VP1) DNA-binding protein from *A. thaliana.* RAV1 contains also AP2 and B3-like domains. Both AP2 and B3-like domains of RAV1 were demonstrated to possess autonomous DNA-binding activities specific for CAACA (Kagaya et al., 1998).

The clone encoding a serine/threonine phosphatase [(similar to protein phosphatase 2C (PP2C) from *A. thaliana*)], showed a significant reduction of transcription as well. Biochemical and molecular genetic studies of *Arabidopsis* identified PP2C enzymes as key players in plant signal transduction processes. For instance, the ABI1/ABI2 PP2Cs are central components in ABA signal transduction (Rodriguez, 1998). There is evidence that AMPK (mammals' homologue of yeast SNF1 and plants SnRK1) can be inactivated by PP2C (Moore et al., 1991).

As already mentioned, the expression of at least two proteins, chlorophyll a/b binding protein (*CAB1*) and plastocyanin (*PC*), seems to be also ABA dependent. In transgenic seeds both genes have prolonged transcription activities. Expression of both of these proteins is regulated by the transcription factor ABI4, which in turn is SnRK1 regulated (Soderman et al., 2000).

The change in transcription level of alcohol dehydrogenase 1 (ADH1) in the transgenic pea seeds can also be a consequence of modulated ABA signalling. Three clones encoding the same alcohol dehydrogenase 1 from *P. sativum* (Llewellyn et al., 1987) showed a similar temporal profile during seed development. In transgenic seeds the induction of their activity started much later, however, then reached the level of WT plants. In plants, ADH1 responds to sugars (Koch et al., 2000). Other stresses can also affect *Adh1* expression (Dolferus et al., 1994). The transcription factor ABI4 from maize regulates the level of transcription of *Adh1* gene in response to glucose- and ABA content (Xiping Niu et al., 2002). In yeast, the global transcriptional activator swi/snf regulates the transcription of both *Adh1* and *Adh2* genes through modification of chromatin, and Snf1 co-regulates the swi/snf occupancy (Geng et al., 2004). However, there is no sufficient confirmation that the same mechanism also exists in plants.

It's interesting that expression profiles of vicilin, one of the main storage proteins, are also shifted towards reduction of transcript amount, similar to ABI3 and USP mRNAs. Previously, it was shown, that the vicilin promoter responds to ABA in developing transgenic tobacco seed (Jiang, et al., 1996).

There may be the following explanation for how the SnRK1 can transfer the ABA signal. Investigation of SnRK1 in tomato seeds showed, that the regulatory subunit of the SnRK1 complex, LeSNF4, was induced in response to ABA or dehydration, but e was not ABA dependent (Bradford et al., 2003). There is evidence about the presence of ABA-responsive elements in the promoter region of the *LeSNF4* gene. Our study of the SnRK1s 5'upstream region in the transient expression system on tobacco protoplasts showed a clear response to ABA. The VfSnRK1 promoter is repressed in the presence of cytokinin alone in the cultivation medium, however, it can be activated by supplementation with ABA and low.

Thus, when SNF4 is expressed during seed maturation (probably in response to ABA), it can bind to SNF1/SIP (or other subunits) and alter the kinase activity to promote gene expression involved in the accumulation of storage reserves. For example, SNF4 might modulate SNF1 activity or substrate specificity to express metabolic pathways that result in the reduction in Glc and accumulation of Suc and oligosaccharides. The antisense repression of SnRK1 entails the interruption the ABA signal exactly at the moment of the occurrence of the seed maturations phase.

5.5.1.5 SnRK1 is involved in regulation of protein synthesis

Some genes that are involved in amino acid synthesis, protein biosynthesis and storage showed similar profiles. Therefore, they may be coordinately regulated.

The enzymes, which catalyse the charging of tRNAs with corresponding amino acid, for instance lysyl-tRNA synthetase and methionyl-tRNA synthetase, showed later induction of expression levels in transgenic seeds. The transcriptional regulation of amino acid tRNA synthetases in yeast represents a very complex mechanism. During amino acid starvation, the accumulation of uncharged tRNAs result in the activation of protein kinase GCN2. The target for GCN2 is the factor initiation of translation eIF2α. Phosphorylation causes inhibition of the eIF2 complex, leading to a general reduction of protein synthesis. This general reduction of protein synthesis is directly linked to increased translation of the transcriptional activator GCN4. cDNA microarray analysis of yeast cells cultivated under amino acid starvation conditions showed altered regulation of transcription of 539 genes, including amino acid biosynthetic pathway genes (Halford et al., 2004). In amino acid-starved cells, when Gcn4p is highly induced, ribosomal proteins (RP) were strongly repressed, the *KRS1/GCD5, ILS1,* and *MES1* genes, encoding aminoacyl-tRNA synthetases for Lys, Ile, and Met, respectively, were induced (Natarajan et al., 2001).

The opposite situation was observed in transgenic pea seeds. At least three explanations are possible: the GCN4p is inactive, the amino acid level is increased, or the signal transduction to GCN4 is interrupted.

Eight different clones which represent different ribosomal proteins were up-regulated showing similar transcriptional profiles in transgenic seeds. The 60S ribosomal protein L9, previously cloned from GA-treated dwarf peas was the most interesting of them. In yeast cells, the ribosomal protein L9 as a product of *GRC5* gene, when presented in multiple copies, increased the expression of transcription factor GCN4 (Nika et al., 1997). One possible explanation of this consequence is the 60S L9 is expressed for possible compensation of GCN4 activity. The induction of ribosomal protein genes might enhance the translation process. The presence of GCN4 in plants is under debate (Halford et al., 2004).

There are no strong evidences confirming the direct interaction of SnRK1 and GCN4 (or even GCN2), but it is possible, that it could be regulated indirectly, or both genes are involved in common signal transduction pathways. There could also be an additional link between sugar and amino acid metabolism.

5.5.1.6 SnRK1 activity changes entail global changes in sugar metabolism pathways

A high number of genes, involved in sugar metabolism, showed changes in their temporal expression profiles. Among them are gluconolactonase, cytosolic malate dehydrogenase 1, glycosyltransferase, and cellulose synthase, which were up-regulated in transgenic pea seeds at the early stage of development. To the group of genes, which showed prolonged expression activity in transgenic lines, belong genes involved in carbohydrate metabolism, such as isocitrate dehydrogenase 1, fructose-bisphosphate aldolase, and transketolase 1. The full list of genes is given in section "Results". It reflects the consequences of SnRK1 repression. However, there is no experimental evidence on how these genes are regulated by SnRK1.

Based on the reported results, the following general conclusions about the role SnRK1 in pea seed development are possible:

1) SnRK1 is involved in cell cycle regulation probably by mediation of brassinosteroide signal transduction.
2) SnRK1 regulates the transcription of many cell cycle genes at the level of chromatin remodelling.
3) SnRK1 is involved in ubiquitin-dependent proteolysis and cytokinin signal transduction.
4) Repression of SnRK1 causes interruption of ABA signals.
4) SnRK1 is involved in the regulation of protein synthesis.

Most of the conclusions are based on correlative data and need further experimental evidence.

6. Abstract

SnRK1 (**s**ucrose-**n**on-fermenting-1-**r**elated) protein kinases are involved in the regulation of plant metabolism controlling both gene expression and phosphorylation. The aim of the study was to investigate the role of SnRK1 in pea seed development.

1. A full length cDNA was isolated by screening a *V. faba* seed cDNA library with an RT-PCR fragment as a probe. The fragment was amplified by primers, designed from conserved region of known plant SNF1-related kinases, using the *V. faba* seed cDNA library as a template. The resulted sequence was designated *VfSnRK1* and encodes a protein of 509 amino acid residues. The full length *SnRK1* of *P. sativum* was identified as an EST in a pea EST library and encodes a protein of 509 amino acid residues as well. The *PsSnRK1* has a domain structure similar to *VfSnRK1*. The *PsSnRK1* nucleotide sequence possesses 97.6%, and protein 98.8% identity to *VfSnRK1*.

Southern blot hybridization using a *V. faba* genomic DNA revealed a single copy in the genome.

Temporal and spatial expression patterns were analyzed by Northern blot hybridization in different tissues of *V. faba* plants. The transcripts were most abundant in sink-organs like roots, pods and gynoecia and hardly detectable in the seed coat. During seed development the expression levels of *VfSnRK1* were highest from 19 DAP to 23 DAP and decreased after 24 DAP. VfSnRK1p kinase activity was measured using a SAMS peptide assay in 40% saturated ammonium sulfate fraction of protein extracts from developing embryos. The temporal profile of enzyme activity correlated with transcripts amounts, but the peak of activity is shifted towards two days later.

2. To better examine the structure of the *VfSnRK1* gene, a 5'-flanking region of 1.9 Kb in length was isolated by genome walking. Immediately upstream of the translation initiation codon an intron was found of 1.5 Kb length. To study the promoter activity, two chimeric constructs were created driving the β-glucuronidase gene under different sizes of the promoter region without and with intron. The transient assays clearly demonstrated that elevated levels of GUS expression depend on the presence of the intron and showed that this sequence is a functional promoter. Moreover, promoter activity was down-regulated in the presence of cytokinin alone. These data imply that SnRK1 may be involved in the regulation of the cell cycle either directly or indirectly. Specific up-regulation of GUS activity in response to abscisic acid (ABA) treatment and low sugar concentration was observed, indicating a critical role of the regulation of SnRK1 by phytohormones and sugar in legumes seed development.

3. To study the effect of SnRK1 deficiency, transgenic pea and tobacco plants were generated carrying a gene for VfSnRK1 in antisense orientation under control of vicilin promoter. Selected transgenic lines were characterized with decreased levels of *PsSnRK1* mRNA and reduced up to 71% SAMS peptide activity. Antisense inhibition of SnRK1 resulted in reduced seeds fresh weight, defect of pollen development.

4. To dissect the SnRK1-antisense phenotype at the molecular level, a search for genes with differential expression patterns in transgenic plant *versus* wild type seeds has been performed using cDNA macroarray analysis. For this approach two cDNA libraries from pea embryos and seed coats were constructed. A total of 8414 good quality ESTs, which were assembled in 1082 clusters (1465 contigs) with 6061 sequences and 2353 singletons were derived, annotated and characterized. 4548 EST clones, containing 4018 unique clones' cDNA fragments were used for preparation of array filters. Radioactive labeled cDNA probes were prepared from RNA isolated from embryo of developing seeds of wild type (11-21 DAP) and transgenic plant (13-19 DAP), which correspond to the transition phase of seed development, and hybridized to cDNA macroarrays. Signal intensities (reflecting relative mRNA levels) of 131 differentially expressed cDNA fragments were subjected to cluster analysis. Two major groups of cDNA fragments, showing distinct expression patterns during pea seed development were created by k-mean clustering. The first cluster of EST profiles represents a group of genes, which are up-regulated in WT seeds during the transition phase. Transgenic seeds showed retarded induction of these genes. 57 profiles belong to this cluster. 74 profiles build a second cluster, which represented groups of genes, which are down-regulated in WT during the transition phase but showed prolonged expression activity in transgenic seeds.

The transition between early and late seed development in pea is accompanied by large-scale changes in gene expression patterns and metabolic pathways. The macroarray analysis provides a good opportunity to compare simultaneously the large number of genes and allows to estimate complex changes and to find possible connections to regulatory processes. A following general conclusions about the role SnRK1 in pea seed development is possible: SnRK1 is involved in cell cycle regulation probably by mediation of brassinosteroide signal transduction and regulation of transcription of many cell cycle genes at the level of chromatin remodelling. SnRK1 is involved in ubiquitin-dependent proteolysis and cytokinin signal transduction and regulation of protein synthesis. Repression of SnRK1 causes interruption of ABA signals.

A comparison of the results received by macroarray hybridisation and the information about the experimental confirmation of SnRK1p functions from various publications, recreates results in a more consistent picture of the consequences caused by repression of this gene.

7. Zusammenfassung

SnRK1 Protein Kinasen (sucrose-non-fermenting-1-related) sind an die Regulation des pflanzlichen Metabolismus beteiligt, mittels Steuerung von Genwirkung und Phosphorylierung. Das Ziel dieser Arbeit war, die Rolle von SnRK1 in der Erbsensamenentwicklung zu untersuchen.

1. Eine Voll-Längen -cDNA wurde gefunden, indem eine *V. faba*-DNA-Bibliothek mit einem RT-PCR Fragment als Sonde durchmustert wurde. Das Fragment wurde durch Primer amplifiziert, die von bekannten pflanzlichen SNF1-ähnnliche Kinasen abgeleitet wurden. Die *V. faba*-cDNA-Bibliothek diente als Schablone. Die resultierende DNA-Sequenz wurde als *VfSnRK1* bezeichnet und kodiert ein Protein von 509 Aminosäuresten. Die Voll-Längen cDNA von SnRK1 aus *P. sativum* wurde als EST aus einer Erbsen-EST-Bibliothek isoliert und kodiert ebenfalls ein Protein von 509 Aminosäureüberresten. Das *PsSnRK1* hat eine Domänenstruktur, die der von *VfSnRK1* ähnlich ist. Die Nukleotidsequenz von *PsSnRK1* ist zu 97,6% identisch, die Identität der abgeleiteten Proteine beträgt 98,8%.

Die Southern-Blot-Hybridisierung von *V.faba* genomischer DNS, zeigt eine einzige Kopie im Genom auf.

Zeitliche und räumliche Expressionsmuster wurden durch Northern- Blot-Hybridisierung in unterschiedlichen Geweben von *V. faba* analysiert. Transkriptspiegel waren hoch in *sink*-Organen wie Wurzeln, Hülsen und Fruchtknoten und kaum nachweisbar in der Samenschale.

Während der Samenentwicklung waren die Expressionsspiegel von *VfSnRK1* zwischen 19 TNB (Tage nach Befruchtung) und 23 DAP am höchsten und nach 24 DAP verringert. VfSnRK1p-Kinaseaktivität wurde mit einem SAMS-Peptidassay in einer zu 40% gesättigtem Ammoniumsulfat Fraktion der Proteinextrakte von sich entwickelnden Embryos gemessen. Das zeitliche Profil der Enzymaktivität bezogen auf Transkriptenmengen war ähnlich, aber die Spitze der Aktivität war zwei Tage nach hinten verschoben.

2. Um die Struktur des *VfSnRK1* Gens zu überprüfen, wurde die 5'-flankierende Region von 1,9 Kb Länge mittels „*Genom walking*“ kloniert. Unmittelbar stromaufwärts vom Initiationskodon der Translation wurde ein Intron von 1,5-Kb-Länge gefunden. Um die Promoter Aktivität zu prüfen, wurde zwei chimäre Konstrukte des β-glucuronidase-Gens hinter den Promoterregionen mit und ohne Intron gemacht. Die transiente Expression in Protoplasten zeigte, dass GUS- Expressionshöhe vom Vorhandensein des Intron abhängt und bewies, dass diese Sequenz einen funktionellen Promoter darstellt. Die Promoteraktivität wurde durch Cytokinine herunterreguliert.

Diese Daten zeigen, dass SnRK1 an die Regelung des Zellenzyklus entweder direkt oder indirekt beteiligt ist.

Eine spezifische Hochregulation der GUS-Aktivität wurde durch Behandlung mit Abscisinsäure (ABS) und niedrigen Zuckerkonzentration beobachtet. Das zeigt die kritische Rolle der Regulation von SnRK1 durch Phytohormone und Zucker in der Hülsenfruchtsamenentwicklung.

3. Um den Effekt des Mangels an SnRK1 zu untersuchen, wurden transgene Erbsen- und Tabakpflanzen mit einem Gen für VfSnRK1 in antisense Orientierung unter Steuerung des Vicilin Promoter erzeugt. Ausgewählte transgene Linien mit verringerten Gehalt an *PsSnRK1*-mRNA und bis zu 71 % verringerter SAMS Peptidaktivität wurden weiter charakterisiert. Antisense Reduktion von *SnRK1* bewirkte verringerte Samenfrischgewichte und Defekte in der Pollenentwicklung.

4. Um den SnRK1-antisense-Phänotypus auf dem molekularen Niveau zu untersuchen, wurde eine differentielle Genexpressionsanalyse der transgenen und der Wildtyp-Samen mittels macroarray-Analyse durchgeführt. Dafür wurden zwei cDNA Bibliotheken aus Erbsenembryos und Samenschalen konstruiert. Eine Gesamtmenge von 8414 ESTs, in 1082 Clustern (1465 Contigs) und mit 6061 Sequensen wurden charakterisiert. Davon wurden 2353 Singletons abgeleitet, annotiert und charakterisiert. 4548 EST Klone, darunter 4018 Einzel-cDNA-Fragmente, wurden für Arrayfilter verwendet.

Radioaktiv markierte cDNA-Proben wurden aus RNA von Embryonen sich entwickelnder Samen von Wildtyp (11-21 DAP) und transgenen Pflanzen (13-19 DAP, Übergangsphase der Samenentwicklung) vorbereitet und auf die DNA Macroarrayfilter hybridisiert. Die Signalintensität (entspricht dem relativen mRNA Expressionsniveau) von 131 unterschiedlich exprimierten DNA-Fragmenten wurde einer Clusteranalyse unterzogen.

Zwei Hauptgruppen von cDNAs, die eindeutige veränderte Expressionsmuster während der Erbsensamenentwicklung zeigen, wurden durch *k-mean clustering* identifiziert.

Das erste Cluster von EST-Profilen stellt eine Gruppe Genen dar, die in den WT-Samen während der Übergangsphase hochreguliert sind. Transgene Samen zeigten verzögerte Induktion dieser Gene, 57 Profile gehören diesem Cluster an. 74 Profile bilden ein zweites Cluster, das von Genen repräsentiert wird, die in WT Samen während der Übergangsphase herunterreguliert sind, aber in den transgenen Samen eine verlängerte Expressionsaktivität hatten.

Der Übergang zwischen früher und später Samenentwicklung der Erbse ist von den großräumigen Änderungen der Genexpressionsmuster sowie Stoffwechselwegen verbunden.

Die Macroarray-Analyse ist eine gute Methode, große Mengen von Genen gleichzeitig in ihrer Expression zu vergleichen und kann helfen, komplexe Änderungen abzuschätzen und mögliche Aufschlüsse über Regulationsprozesse zu finden.

Die folgende allgemeine Zusammenfassunge über die Rolle *SnRK1* in der Erbsensamenentwicklung ist möglich: SnRK1 spielt eine Rolle in Zellenzyklusregulation vermutlich durch Vermittlung von Brassinosteroidsignalen sowie durch Regulation der Transkription vieler Zellenzyklusgene auf Ebene des *chromatin remodelling*. SnRK1 steuert weiterhin ubiquitin-abhängige Proteolyse, Cytokininsignaltransduktion sowie Regulation der Proteinbiosynthese. Unterdrückung von *SnRK1*-Aktivität unterbricht den ABA-Signaltransduktionsweg. Ein Vergleich der Resultate, der Macroarray Hybridisierungen mit publizierten Informationen über die Rolle von SnRK1p ergibt damit ein konsistenteres Bild der Rolle von SnRK1 während der Samenentwicklung.

8. REFERENCES

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* **15**: 63-78

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25-29

Ashrafi K, Farazi TA, Gordon JI (1998) A role for *Saccharomyces cerevisiae* fatty acid activation protein 4 in regulating protein N-myristoylation during entry into stationary phase. *J Biol Chem* **273**: 25864-25874

Ashrafi K, Lin SS, Manchester JK, Gordon JI (2003) Sip2p and its partner snf1p kinase affect aging in *S. cerevisiae*. *J Biol Chem* **278**: 13390-13397

Altschul SF, Gish W, Miller W, Myers E, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410

Altschul SF, Madden TL, Schäffer AA, Zhang Z, Miller W und Lipman DJ (1997) Gapped BLAST and PSI-BLAST:a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402

An G, Ebert PR, Mitra A, Ha SB (1988) Plant Molecular Biology Manual 1-19, Martinus Nijhoff, Dordrecht, Netherlands

Audic S, Claverie JM (1997) The significance of digital gene expression profiles. *Genome Res*. **7**:986-995

Baldwin D, Crane V, Rice D (1999) A comparison of gel-based, nylon filter and microarray techniques to detect differential RNA expression in plants. *Curr Opin Plant Biol* **2**: 96-103

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR (2004) The Pfam protein families database. *Nucleic Acids Res* **32** Database issue: D138-41

Baumlein H, Boerjan W, Nagy I, Bassuner R, Van Montagu M, Inze D, Wobus U (1991) A novel seed protein gene from *Vicia faba* is developmentally regulated in transgenic tobacco and Arabidopsis plants. *Mol Gen Genet* **25**: 459-67

Baumlein H, Wobus U, Pustell J, Kafatos FC (1986) The legumin gene family: structure of a B type gene of *Vicia faba* and a possible legumin gene specific regulatory element. *Nucleic Acids Res* **14**: 2707-2720

Becker D, Kemper E, Schell J, Masterson R (1992) New plant binary vectors with selectable markers located proximal to the left T-DNA border. *Plant Mol Biol* **20**: 1195-1197

Beg ZH, Stonik JA, Brewer HB Jr (1978) 3-Hydroxy-3-methylglutaryl coenzyme A reductase: regulation of enzymatic activity by phosphorylation and dephosphorylation. *Proc Natl Acad Sci U S A* **75**: 3678-3682.

Bevan MW (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Research* **12**: 8711–8722

Bhalerao RP, Salchert K, Bako L, Okresz L, Szabados L, Muranaka T, Machida Y, Schell J, Koncz C (1999) Regulatory interaction of PRL1 WD protein with *Arabidopsis* SNF1-like protein kinases. *Proc Natl Acad Sci U S A* **96**: 5322-5327

Blazquez C, Woods A, de Ceballos ML, Carling D, Guzman M (1999) The AMP-activated protein kinase is involved in the regulation of ketone body production by astrocytes. *J Neurochem* **73**: 1674-82

Borisjuk L, Rolletschek H, Radchuk R, Weschke W, Wobus U, Weber H (2004) Seed development and differentiation: a role for metabolic regulation. *Plant Biol* **6**: 375-86

Borisjuk L, Weber H, Panitz R, Manteufel R, Wobus U (1995) Embryigenesis of *V. faba*.: Histodifferentiation in Relation to Starch and Storage Protein Synthesis. J.*Plant Physiol* **147**: 203-18

Borisjuk L, Walenta S, Rolletschek H, Mueller-Klieser W, Wobus U, Weber H (2002) Spatial analysis of plant metabolism: Sucrose imaging within *Vicia faba* cotyledons reveals specific developmental patterns. *The Plant Journal* **29**: 521-530

Bouly JP, Gissot L, Lessard P, Kreis M, Thomas M (1999) *Arabidopsis thaliana* proteins related to the yeast SIP and SNF4 interact with AKINalpha1, an SNF1-like protein kinase. *Plant J* **18**: 541-550

Bradford KJ, Downie AB, Gee OH, AlvaradoV, Yang H. and Dahal P (2003) Abscisic acid and gibberellin differentially regulate expression of genes of the SNF1-related kinase complex in tomato seeds. *Plant Physiol* **132**: 1560-1576

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem* **72**: 248-254

Buitink J, Thomas M, Gissot L, Leprince O (2004) Starvation, osmotic stress and desiccation tolerance lead to expression of different genes of the regulatory β and γ subunits of the SnRK1 complex in germinating seeds of *Medicago truncatula*. *Plant, Cell & Environment* **27**: 55-67

Cacchione S, Cerone MA, De Santis P, Savino M (1995) Superstructural features of the upstream regulatory regions of two pea rbcS genes and nucleosomes positioning: theoretical prediction and experimental evaluation. *Biophys Chem* **53**: 267-281

Carling D, Clarke PR, Zammit VA, Hardie DG (1989) Purification and characterization of the AMP-activated protein kinase. Copurification of acetyl-CoA carboxylase kinase and 3-hydroxy-3-methylglutaryl-CoA reductase kinase activities. *Eur J Biochem* **86**: 129-136

Carlson CA, Kim KH (1974) Regulation of hepatic acetyl coenzyme A carboxylase by phosphorylation and dephosphorylation. *Arch Biochem Biophys* **164**: 478-489

Carlson M (1999) Glucose repression in yeast. *Curr Opin Microbiol* **2**: 202-207

Carlson M, Osmond BC, Botstein D (1981) Mutants of yeast defective in sucrose utilization. *Genetics* **98**: 25-40

Celenza JL, Carlson M (1984) Cloning and genetic mapping of SNF1, a gene required for expression of glucose-repressible genes in *Saccharomyces cerevisiae*. *Mol Cell Biol* **4**: 49-53

Celenza JL, Carlson M (1986) A yeast gene that is essential for release from glucose repression encodes a protein kinase. *Science* **233**: 1175-1180

Chikano H, Ogawa M, Ikeda Y, Koizumi N, Kusano T, Sano H (2001) Two novel genes encoding SNF-1 related protein kinases from *Arabidopsis thaliana*: differential accumulation of AtSR1 and AtSR2 transcripts in response to cytokinins and sugars, and phosphorylation of sucrose synthase by AtSR2. *Mol Gen Genet* **264**: 674-681

Chou A, Burke J (1999) CRAWview: for viewing splicing variation, gene families, and polymorphism in clusters of ESTs and full-length sequences. *Bioinformatics* **15**: 376-381

Christoffels A, van Gelder A, Greyling G, Miller R, Hide T, Hide W (2001) STACK: Sequence Tag Alignment and Consensus Knowledgebase. *Nucleic Acids Res* **29**: 234-238

Church GM, Gilbert W (1984) Genomic sequencing. *Proc Natl Acad Sci USA* **81**: 1991-1995

Ciriacy M (1977) Isolation and characterization of yeast mutants defective in intermediary carbon metabolism and in carbon catabolite derepression. *Mol Gen Genet* **154**: 213-220

Clarke PR, Hardie DG (1990) Regulation of HMG-CoA reductase: identification of the site phosphorylated by the AMP-activated protein kinase in vitro and in intact rat liver. *EMBO J* **9**: 2439-2446

Cluis CP, Mouchel CF, Hardtke CS (2004) The Arabidopsis transcription factor HY5 integrates light and hormone signaling pathways. *Plant J* **38**:332-347

Cohen SN, Chang AC, Hsu L (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc Natl Acad Sci U S A* **69**: 2110-2114

Cordoba E, Shishkova S, Vance CP, Hernandez G (2003) Antisense inhibition of NADH glutamate synthase impairs carbon/nitrogen assimilation in nodules of alfalfa (*Medicago sativa* L.). *Plant J* **33**: 1037-1049

Corton JM, Gillespie JG, Hawley SA, Hardie DG (1995) 5-aminoimidazole-4-carboxamide ribonucleoside. A specific method for activating AMP-activated protein kinase in intact cells? *Eur J Biochem* **229**: 558-565

Coruzzi GM, Zhou L (2001) Carbon and nitrogen sensing and signalling in plants: emerging 'matrix effects'. *Curr Opin Plant Biol* **4**: 247-253

Crute BE, Seefeld K, Gamble J, Kemp BE, Witters LA (1998) Functional domains of the alpha1 catalytic subunit of the AMP-activated protein kinase. *J Biol Chem* **273**: 35347-35354

Cullen PJ, Sprague GF Jr (2000) Glucose depletion causes haploid invasive growth in yeast. *Proc Natl Acad Sci U S A* **97**: 13619-1362

Cutler S, Ghassemian M, Bonetta D, Cooney S and McCourt P (1996) A protein farnesyl transferase involved in abscisic acid signal transduction in *Arabidopsis*. *Science* **273**: 1239-1241

Datta N, Cashmore AR (1989) Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. *Plant Cell* **11**: 1069-1077

Davies SP, Carling D, Hardie DG (1989) Tissue distribution of the AMP-activate protein kinase, and lack of activation by cyclic AMP-dependent protein kinase, studies using a specific and sensitive peptide assay. *Eur J Biochem* **186**: 123-128

Davies SP, Helps NR, Cohen PT, Hardie DG (1995) 5'-AMP inhibits dephosphorylation, as well as promoting phosphorylation, of the AMP-activated protein kinase. Studies using bacterially expressed human protein phosphatase-2C alpha and native bovine protein phosphatase-2AC. *FEBS Lett* **377**: 421-425

De Vit MJ, Waddle JA, Johnston M (1997) Regulated nuclear translocation of the Mig1 glucose repressor. *Mol Biol Cell* **8**: 1603-1618

Derisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM (1996) Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nature Genetics* **14**: 457-460

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* **12**: 387-395

Dijkstra WE (1959) A note o two problems in connection with graphs. *Numerische Mathematik* **1**: 269-271

Dijkwel PP, Kock P, Bezemer R, Weisbeek PJ, Smeekens S (1996) Sucrose Represses the Developmentally Controlled Transient Activation of the Plastocyanin Gene in Arabidopsis thaliana Seedlings. *Plant Physiol* **110**: 455-463

Dolferus R, Jacobs M, Peacock WJ, Dennis ES (1994) Differential interactions of promoter elements in stress responses of the Arabidopsis Adh gene. *Plant Physiol* **105**: 1075-1087

Duque P, Chua NH (2003) IMB1, a bromodomain protein induced during seed imbibition, regulates ABA- and phyA-mediated responses of germination in Arabidopsis. *Plant J* **35**:787-799

Dysvik B, Jonassen I (2001) J-Express: Exploring Gene Expression Data using Java. *Bioinformatics* **17**: 369-370

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95**: 14863-14868

Estruch F, Treitel MA, Yang X, Carlson M (1992) N-terminal mutations modulate yeast SNF1 protein kinase function. *Genetic* **132**: 639-650

Ewing RM, Ben Kahla A, Poirot O, Lopez F, Audic S, Claverie JM (1999) Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Res* **9**:950-959.

Farras R, Ferrando A, Jasik J, Kleinow T, Okresz L, Tiburcio A, Salchert K, del Pozo C, Schell J, Koncz C (2001) SKP1-SnRK protein kinase interactions mediate proteasomal binding of a plant SCF ubiquitin ligase. *EMBO J* **20**: 2742-2756

Ferrando A, Koncz-Kalman Z, Farras R, Tiburcio A, Schell J, Koncz C (2001) Detection of *in vivo* protein interactions between Snf1-related kinase subunits with intron-tagged epitope-labelling in plants cells. *Nucleic Acids Res* **29**: 3685-3693

Finkelstein RR, Wang ML, Lynch TJ, Rao S and Goodman HM (1998) The *Arabidopsis* abscisic acid response locus *ABI4* encodes an APETALA2 domain protein. *The Plant Cell* **10**: 1043-1054

Finkelstein RR, Gibson SI (2002) ABA and sugar interactions regulating development: cross-talk or voices in a crowd? *Curr Opin Plant Biol* **5**: 26-32

Gallardo K, Le Signor C, Vandekerckhove J, Thompson RD, Burstin J (2003) Proteomics of Medicago truncatula seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiol* **133**: 664-682

Gamborg OL, Miller RA, Ojima K (1968) Nutrient requirements of suspension cultures of soybean root cells. *Exptl Cell Res* **50**: 151-158

Gancedo JM (1998) Yeast carbon catabolite repression. *Microbiol Mol Biol Rev* 62: 334-361

Geer LY, Domrachev M, Lipman DJ, Bryant SH (2002) CDART: protein homology by domain architecture. *Genome Res* **12**: 1619-1623

Gene Ontology Consortium (2001) Creating the gene ontology resource: design and implementation. *Genome Res* **11**: 1425-1433

Geng F, Laurent BC (2004) Roles of SWI/SNF and HATs throughout the dynamic transcription of a yeast glucose-repressible gene. *EMBO J* **23**: 127-137

Genoud T, Metraux JP (1999) Crosstalk in plant cell signaling: structure and function of the genetic network. *Trends Plant Sci* **4**: 503-507

Giraudat J, Hauge B, Valon C, Smalle J, Parcy F, Goodman H (1992) Isolation of the *Arabidopsis ABI*3 gene by positional cloning. *The Plant Cell* **4**:1251-1261

Girke T, Todd J, Ruuska S, White J, Benning C, Ohlrogge J (2000) Microarray analysis of developing Arabidopsis seeds. *Plant Physiol* **124**: 1570-1581

Goldberg RB, de Paiva G, Yadegari R Plant embryogenesis: zygote to seed. *Science* 1994 **266**: 605-614

Gomez-Cardenas A, Zentella R, Walker-Simmons MK, Ho THD (2001) Gibberellin/abscisic acid antagonism in barley aleurone cells: Site of action of the protein kinase PKABA1 in relation to gibberellin signaling molecules. *Plant Cell* **13**: 667–679

Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. *Genome Res* **8**: 195-202

Gray WM, Estelle I (2000) Function of the ubiquitin-proteasome pathway in auxin response. *Trends Biochem Sci* 25: 133-1338

Grill E, Himmelbach A (1998) ABA signal transduction. *Curr Opin Plant Biol* 1:412-418

Halford NG, Hardie DG (1998) SNF1-related protein kinases: global regulators of carbon metabolism in plants? *Plant Mol Biol* **37**: 735-748

Halford NG, Hey S, Jhurreea D, Laurie S, McKibbin RS, Zhang Y, Paul MJ (2004) Highly conserved protein kinases involved in the regulation of carbon and amino acid metabolism. *J Exp Bot* **55**: 35-42

Hardie DG (1999) PLANT PROTEIN SERINE/THREONINE KINASES: Classification and functions. *Annu Rev Plant Physiol Plant Mol Biol* **50**: 97-131

Hardie DG (2000) Plant protein-serine/threonine kinases: classification into subfamilies and overview of function. *In* M Kreis, JC Walker, eds,. *Advances in Botanical Research* Incorporating Advances in Plant Pathology, Plant Protein Kinases, Academic Press **32**: 1-44

Hardie DG, Carling D (1997) The AMP-activated protein kinase--fuel gauge of the mammalian cell? *Eur J Biochem* **246**: 259-273

Hauxwell AJ, Corke FM, Hedley CL, Wang TL (1990) Storage protein gene expression is localised to regions lacking mitotic activity in developing pea embryos. An analysis of seed development in Pisum sativum XIV. *Development* **110**: 283-289

Hawley SA, Selbert MA, Goldstein EG, Edelman AM, Carling D, Hardie DG (1995) 5'-AMP activates the AMP-activated protein kinase cascade, and Ca2+/calmodulin activates the calmodulin-dependent protein kinase I cascade, via three independent mechanisms. *J Biol Chem* **270**: 27186-271891

Hedley CL, Ambrose MJ (1980) An analysis of seed development in *Pisum sativum* L. *Annals of Botany* **46**: 89–105

Heim U, Weber H, Baumlein H, Wobus U (1993) A sucrose-synthase gene of *Vicia faba* L.: expression pattern in developing seeds in relation to starch synthesis and metabolic regulation. *Planta* **191**: 394-401

Hennig L, Gruissem W, Grossniklaus U, Kohler C (2004) Transcriptional programs of early reproductive stages in *Arabidopsis*. *Plant Physiol* **135**: 1765-1775

Hernalsteens JP, Van Vliet F, De Beuckeleer M, Depicker A, Engler G (1980) The *Agrobacterium tumefaciens* Ti plasmid as a host vector system for introducing foreign DNA in plant cells. *Nature* **287**: 754–656

Hiesinger M, Roth S, Meissner E, Schuller HJ (2001) Contribution of Cat8 and Sip4 to the transcriptional activation of yeast gluconeogenic genes by carbon source-responsive elements. *Curr Genet* **39**: 68-76

Hofmann K, Bucher P (1996) The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway. *Trends Biochem Sci* **5**: 172-173

Holmes DS, Quigley M (1981) A rapid boiling method for the preparation of bacterial plasmids. *Anal Biochem* **114**: 193-197

Honigberg SM, Lee RH (1998) Snf1 kinase connects nutritional pathways controlling meiosis in Saccharomyces cerevisiae. *Mol Cell Biol* **8**: 4548-4555

Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. *Trans Res* **2**: 208–218

Horsch RB, Try JE, Hoffmann NL, Wallroth M, Eichholtz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. *Science* **227**: 1229–1233

Hrabak EM, CWM Chan, M Gribskov, JF Harper, JH Choi, N Halford, J Kudla, S Luan, HG Nimmo, M.R Sussman, M Thomas, K Walker-Simmons, J-K. Zhu, Harmon AC (2003) The *Arabidopsis* CDPK-SnRK superfamily of protein kinases. *Plant Physiol* **132**: 666-680

Huang D, Farkas I, Roach PJ (1996) Pho85p, a cyclin-dependent protein kinase, and the Snf1p protein kinase act antagonistically to control glycogen accumulation in *Saccharomyces cerevisiae*. *Mol Cell Biol* **16**: 4357-4365

Ihaka R, Gentleman R (1996) R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**: 299-314

Ikeda Y, Koizumi N, Kusano T, Sano H (1999) Sucrose and cytokinin modulation of WPK4, a gene encoding a SNF1-related protein kinase from wheat. *Plant Physiol* **121**: 813-820

Jang JC, Sheen J (1997) Sugar sensing in higher plants. *Trends Plant Sci* **2:** 208-214

Jerpseth B, Greener A, Short JM, Viola J, Kretz PL (1992) XL-1 Blue MRF ' *E . coli* cells: McrA , McrB , McrF , Mrr , HsdR derivative of XL-1 Blue cells. *Stratagies* **5**: 81-88

Jiang L., Abrams S.R., and Kermode A.R (1996) Vicilin and napin storage protein gene promoters are responsive to ABA in developing transgenic tobacco seed, but lose sensitivity following premature desiccation. *Plant Physiol* **110**: 1135-1144

Jiang R, Carlson M (1996) Glucose regulates protein interactions within the yeast SNF1 protein kinase complex. *Genes Dev* **10**: 3105-3115

Jiang R, Carlson M (1997) The Snf1 protein kinase and its activating subunit, Snf4, interact with distinct domains of the Sip1/Sip2/Gal83 component in the kinase complex. *Mol Cell Biol* 17: 2099-2106

Jürgens G, Mayer U (1994) Arabidopsis. *In* JBL Bard, ed, *Embryos, Color Atlas of Development*. Wolfe Publishing, London: 7-21

Jurka J (1998) Repeats in genomic DNA: mining and meaning. *Curr Opin Struct Biol* **8**: 333-337

Kagaya Y, Ohmiya K, Hattori T (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res* **27**: 470-478

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**: 27-30

Kim SA, Kwak HJ, Park MC, Kim SR (1998) Induction of reproductive organ-preferential histone genes by wounding or methyl jasmonate. *Mol Cells* **8**: 669-677

Kleinow T, Bhalerao R, Breuer F, Umeda M, Salchert K, Koncz C (2000) Functional identification of an arabidopsis *snf4* ortholog by screening for heterologous multicopy suppressors of snf4 deficiency in yeast. *Plant J* **23**: 115-122

Koch KE (1996) Carbohydrate-modulated gene expression in plants. *Ann Rev Plant Physiol Plant Mol Biol* **46**: 509-540

Koch KE, Ying Z, Wu Y, Avigne WT (2000) Multiple paths of sugar-sensing and a sugar/oxygen overlap for genes of sucrose and ethanol metabolism. *J Exp Bot* **51**: 417-427

Koornneef M, Jorna, ML, Derswan D, Karssen CM (1982) The isolation of abscisic acid (ABA) deficient mutants by selection of induced revertants in non-germinating gibberellin sensitive lines of *Arabidopsis thaliana* (L.) Heynh . *Theor. Appl. Genet*. **61**: 385-393

Kuchin S, Treich I, Carlson M (2000) A regulatory shortcut between the Snf1 protein kinase and RNA polymerase II holoenzyme. *Proc Natl Acad Sci USA* **97**: 7916-7920

Kudo N, Gillespie JG, Kung L, Witters LA, Schulz R, Clanachan AS, Lopaschuk GD (1996) Characterization of 5'AMP-activated protein kinase activity in the heart and its role in

inhibiting acetyl-CoA carboxylase during reperfusion following ischemia. *Biochim Biophys Acta* **1301**: 67-75.

Kuo MH, vom Baur E, Struhl K, Allis CD (2000) Gcn4 activator targets Gcn5 histone acetyltransferase to specific promoters independently of transcription. *Mol Cell* **6**: 1309-1320

Kurup S, Jones HD, Holdsworth MJ (2000) Interactions of the developmental regulator ABI3 with proteins identified from developing Arabidopsis seeds. *Plant J* **21**: 143-155

Künne C, Lange M, Funke T, Miene H, Thiel T, Grosse I, Scholz U (2005) CR-EST: a resource for crop ESTs. *Nucl Acid Research* **33**: 619-621

Lakatos L, Klein M, Hofgen R, Banfalvi Z (1999) Potato StubSNF1 interacts with StubGAL83: a plant protein kinase complex with yeast and mammalian counterparts. *Plant J* **17**: 569-574

Le Guen L, Thomas M, Bianchi M, Halford NG, Kreis M (1992) Structure and expression of a gene from *Arabidopsis thaliana* encoding a protein related to SNF1 protein kinase. *Gene* **120**: 249-254

Leung J, Merlot S and Giraudat J (1997) The *Arabidopsis* ABSCISIC ACID-INSENSITIVE2 (*ABI*2) and *ABI1* genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. *The Plant Cell* **9**: 759-771

Llewellyn DJ, Finnegan EJ, Ellis JG, Dennis ES, Peacock WJ (1987) Structure and expression of an alcohol dehydrogenase 1 gene from *Pisum sativum* (cv. "Greenfeast"). *J Mol Biol* **195**: 115-23

Lo WS, Duggan L, Tolga NC, Emre, Belotserkovskya R, Lane WS, Shiekhattar R, Berger SL (2001) Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science* **293**: 1142-1146

Lumbreras V, Alba MM, Kleinow T, Koncz C, Pages M (2001) Domain fusion between SNF1-related kinase subunits during plant evolution. *EMBO Rep* **2**: 55-60

McCarty DR, Hattori T, Carson CB, Vasil V, Lazar M, Vasil I (1991) The *Viviparous-1* developmental gene of maize encodes a novel transcriptional activator. *Cell* **66**: 895-905

McCartney RR, Schmidt MC (2001) Regulation of Snf1 kinase. Activation requires phosphorylation of threonine 210 by an upstream kinase as well as a distinct step mediated by the Snf4 subunit. *J Biol Chem* **76**: 36460-36466

Machuka J, Bashiardes S, Ruben E, Spooner K, Cuming A, Knight C, Cove D (1999) Sequence analysis of expressed sequence tags from an ABA-treated cDNA library identifies stress response genes in the moss physcomitrella patens. *Plant Cell Physiol* **40**: 378-387

Manevski A, Bertoni G, Bardet C, Tremousaygue D, Lescure B (2000) In synergy with various cis-acting elements, plant insterstitial telomere motifs regulate gene expression in Arabidopsis root meristems. *FEBS Lett* **483**: 43-46

McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA, Tossberg J, Nickle T, Levin JZ, Law M, Meinke D, Patton D (2001) Insertional mutagenesis of genes required for seed development in Arabidopsis thaliana. *Genetics* **159**: 1751-1763

Mekhedov S, de Ilarduya OM, Ohlrogge J (2000) Toward a functional catalog of the plant genome. A survey of genes for lipid biosynthesis. *Plant Physiol* **122**:389-402

Miller C, Gurd J, Brass A (1999) A RAPID algorithm for sequence database comparisons: application to the identification of vector contamination in the EMBL databases. *Bioinformatics* **15**: 111-121

Mitchelhill KI, Stapleton D, Gao G, House C, Michell B, Katsis F, Witters LA, Kemp BE (1994) Mammalian AMP-activated protein kinase shares structural and functional homology with the catalytic domain of yeast Snf1 protein kinase. *J Biol Chem* **269**: 2361-2364

Monke G, Altschmied L, Tewes A, Reidt W, Mock HP, Baumlein H, Conrad U (2004) Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA. *Planta* 219: 158-166

Moore F, Weekes J, Hardie DG (1991) Evidence that AMP triggers phosphorylation as well as direct allosteric activation of rat liver AMP-activated protein kinase. A sensitive mechanism to protect the cell against ATP depletion. Eur J Biochem. **199**: 691-697

Moran DL (2000) Characterization of the structure and expression of a highly conserved ribosomal protein gene, L9, from pea. *Gene* **253**: 19-29

Muranaka T, Banno H and Machida Y (1994) Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces cerevisiae SNF1 that constitutively activates expression of the glucose-repressible SUC2 gene for a secreted invertase of S. cerevisiae. *Mol Cell Biol* **14**: 2958-2965

Murashige T, Skoog F (1962) A revised médium for rapid growth and bioassays with tobacco tissue cultures. *Physiol Plant* **15**: 473-497

Natarajan K, Meyer MR, Jackson BM, Slade D, Roberts C, Hinnebusch AG, Marton MJ (2001) Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. *Mol Cell Biol* **21**: 4347-4368

Neigeborn L, Carlson M (1984) Genes affecting the regulation of SUC2 gene expression by glucose repression in Saccharomyces cerevisiae. Genetics **108**: 845-858

Nemeth K, Salchert K, Putnoky P, Bhalerao R, Koncz-Kalman Z, Stankovic-Stangeland B, Bako L, Mathur J, Okresz L, Stabel S, Geigenberger P, Stitt M, Redei GP, Schell J, Koncz C (1998) Pleiotropic control of glucose and hormone responses by PRL1, a nuclear WD protein, in Arabidopsis. *Genes Dev* **12**: 3059-3073

Nika J, Erickson FL, Hannig EM (1997) Ribosomal protein L9 is the product of GRC5, a homolog of the putative tumor suppressor QM in *S. cerevisiae*. *Yeast* **13**: 1155-1166

Ohba H, Steward N, Kawasaki S, Berberich T, Ikeda Y, Koizumi N, Kusano T, Sano H (2000) Diverse response of rice and maize genes encoding homologs of WPK4, an SNF1-related protein kinase from wheat, to light, nutrients, low temperature and cytokinins. *Mol Gen Genet* **263**: 359-366

Patil S, Penheiter, Chollet R, Sarath G (1999) Cloning of a full length SNF1-related protein-Ser/Thr kinase cDNA from soybean root nodules (Accession No. AF28443) *Plant Physiol* **119**: 1568

Pien S, Wyrzykowska J, Fleming AJ (2001) Novel marker genes for early leaf development indicate spatial regulation of carbohydrate metabolism within the apical meristem. *Plant J* **25**: 663-674

Proust J, Houlne G, Schantz ML, Shen WH, Schantz R (1999) Regulation of biosynthesis and cellular localization of Sp32 annexins in tobacco BY2 cells. *Plant Mol Biol* **39**:361-372

Purcell PC, Smith AM, Halford NG (1998) Antisense expression of a sucrose non-fermenting-1-related protein kinase sequence in potato results in decreased expression of sucrose synthase in tubers and loss of sucrose-inducibility of sucrose synthase transcripts in leaves. *Plant J* **14**: 195-202

Ravanel S, Gakiere B, Job D, Douce R (1998) The specific features of methionine biosynthesis and metabolism in plants. *Proc Natl Acad Sci U S A* **95**:7805-7812

Raz V, Bergervoet JH, Koornneef M (2001) Sequential steps for developmental arrest in Arabidopsis seeds. *Development* **128**:243-252

Reidt W, Wohlfarth T, Ellerstrom M, Czihal A, Tewes A, Ezcurra I, Rask L, Baumlein H (2000) Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. *Plant J* **21**: 401-408

Robertson AJ, Kapros T, Waterborg JH (1997) A cell cycle-regulated histone H3 gene of alfalfa with an atypical promoter structure. *DNA Seq* **7**: 209-216

Rodriguez PL (1998) Protein phosphatase 2C (PP2C) function in higher plants. *Plant Mol Biol* **38**: 919-927

Roitsch T (1999) Source-sink regulation by sugar and stress. *Curr Opin Plant Biol* **2**: 198-206

Romo S, Dopico B, Labrador E (2002) The expression of a new *Cicer arietinum* cDNA, encoding a glutamic acid-rich protein, is related to development. *J. Plant Physiol* **159**:1375-1381

Ruiz JC, Conlon FL, Robertson EJ (1994) Identification of novel protein kinases expressed in the myocardium of the developing mouse heart. *Mech Dev* **48**:153-164

Salchert K, Bhalerao R, Koncz-Kalman Z, Koncz C (1998) Control of cell elongation and stress responses by steroid hormones and carbon catabolic repression in plants. *Philos Trans R Soc Lond B Biol Sci* **353**: 1517-1520

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a laboratory manual. Cold Spring Harbour, NY: Cold Spring Harbour Laboratory Press

Schmidt MC, McCartney RR (2000) beta-subunits of Snf1 kinase are required for kinase function and substrate definition. *EMBO J* **19**: 4936-4943

Schroeder HE, Schotz AH, Wardley-Richardson T, Spencer D, Higgins T (1993) Transformation and Regeneration of Two Cultivars of Pea (*Pisum sativum* L.). *Plant Physiol* **101**: 751-757

Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **9511**:5857-5864

Sheen J (1990) Metabolic repression of transcription in higher plants. *Plant Cell* **2**: 1027-1038

Shemer R, Meimoun A, Holtzman T, Kornitzer D (2002) Regulation of the transcription factor Gcn4 by Pho85 cyclin PCL5. *Mol Cell Biol* **22**: 5395-5404

Slocombe SP, Laurie S, Bertini L, Beaudoin F, Dickinson JR and Halford NG (2002) Identification of SnIP1, a novel protein that interacts with SNFI- related protein kinase (SnRK1), *Plant Molecular Biology* **49**: 31-44

Slocombe SP, Beaudoin F, Donaghy PG, Hardie DG, Dickinson JR, Halford NG (2004) SNF1-related protein kinase (SnRK1) phosphorylates class I heat shock protein. *Plant Physiol Biochem* **42**: 111-116.

Smalle J, Kurepa J, Yang P, Babiychuk E, Kushnir S, Durski A, Vierstra RD (2002) Cytokinin growth responses in Arabidopsis involve the 26S proteasome subunit RPN12. *Plant Cell* **14**: 17-32.

Smeekens S (2000) Sugar-induced signal transduction in plants. *Ann Rev Plant Physiol Plant Mol Biol* **51**: 49-81

Smeekens S, Rook F (1997) Sugar Sensing and Sugar-Mediated Signal Transduction in Plants. *Plant Physiol* **115**: 7-13

Smith CJ, Watson CF, Ray J, Bird CR, Morris PC, Schuch W, Grierson D (1988) Antisense RNA inhibition of polygalacturonase genes in transgenic tomatos. *Nature* **334**: 724–726.

Smith DL (1973) Nucleic acid, protein, and starch synthesis in developing cotyledons of *Pisum arvense*. *Annals of Botany* **37**: 795–804

Soderman EM, Brocard IM, Lynch TJ, Finkelstein RR (2000) Regulation and function of the Arabidopsis ABA-insensitive4 gene in seed and abscisic acid response signaling networks. *Plant Physiol* **124**: 1752-1765

Sreenivasulu N, Altschmied L, Panitz R, Hahnel U, Michalek W, Weschke W, Wobus U (2002) Identification of genes specifically expressed in maternal and filial tissues of barley caryopses: a cDNA array analysis. *Mol Genet Genomics* **266**: 758-767

Sreenivasulu N, Altschmied L, Radchuk V, Gubatz S, Wobus U, Weschke W (2004) Transcript profiles and deduced changes of metabolic pathways in maternal and filial tissues of developing barley grains. *Plant J.* **37**:539-553.

Stapleton D, Gao G, Michell BJ, Widmer J, Mitchelhill K, Teh T, House CM, Witters LA, Kemp BE (1994) Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are homologs of proteins that interact with yeast Snf1 protein kinase. *J Biol Chem* **269**: 29343-29346

Stein L (2001) Genome annotation: from sequence to biology. *Nat Rev Genet* **2**: 493-503

Stitt M, Krapp A (1999) The interaction between elevated carbon dioxide and nitrogen nutrition: The physiological and molecular background. *Plant Cell Environ* **22**: 583–621

Stitt M, and Sonnewald U (1995) Regulation of Metabolism in Transgenic. *Plants Ann Rev of Plant Physiol and Plant Mol Biol* **46**: 341-368

Sugden C, Crawford RM, Halford NG, Hardie DG (1999a) Regulation of spinach SNF1-related (SnRK1) kinases by protein kinases and phosphatases is associated with phosphorylation of the T loop and is regulated by 5'-AMP. *Plant J* **19**: 433-439

Sugden C, Donaghy PG, Halford NG, Hardie DG (1999b) Two SNF1-related protein kinases from spinach leaf phosphorylate and inactivate 3-hydroxy-3-methylglutaryl-coenzyme A reductase, nitrate reductase, and sucrose phosphate synthase in vitro. *Plant Physiol* **120**: 257-274

Sullivan JE, Carey F, Carling D, Beri RK (1994) Characterisation of 5'-AMP-activated protein kinase in human liver using specific peptide substrates and the effects of 5'-AMP analogues on enzyme activity. *Biochem Biophys Res Commun* **200**: 1551-1556

Takano M, Kajiya-Kanegae H, Funatsuki H, Kikuchi S (1998) Rice has two distinct classes of protein kinase genes related to SNF1 of *Saccharomyces cerevisiae*, which are differently regulated in early seed development. *Mol Gen Genet* **260**: 388-394

Takebe I, Otsuki Y, Aoki S (1968) Isolation of tobacco mesophyll cells in intact and active state. *Plant and Cell Physiol* **9**: 115-124

Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. *Science* **278**: 631-637

Thibaud-Nissen F, Shealy RT, Khanna A, Vodkin LO (2003) Clustering of microarray data reveals transcript patterns associated with somatic embryogenesis in soybean. *Plant* Physiol **132**: 118-136

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680

Toroser D, Plaut Z, Huber SC (2000) Regulation of a plant SNF1-related protein kinase by glucose-6-phosphate. *Plant Physiol* **123**: 403-412

Treitel MA, Kuchin S, Carlson M (1998) Snf1 protein kinase regulates phosphorylation of the Mig1 repressor in *Saccharomyces cerevisiae*. *Mol Cell Biol* **18**: 6273-6280

Tsujimoto Y, Izawa S, Inoue Y (2000) Cooperative regulation of DOG2, encoding 2-deoxyglucose-6-phosphate phosphatase, by Snf1 kinase and the high-osmolarity glycerol-mitogen-activated protein kinase cascade in stress responses of *Saccharomyces cerevisiae*. *J Bacteriol* **182**: 5121-5126

van der Krol A, Lenting PE, Veenstra J, van der Meer IM, Koes RE, Gerats AGM, Mol JNM, Stuitje AR (1988) An anti-sense chalcone synthase gene in transgenic plants inhibits flower pigmentation. *Nature* **333**: 866–869

van Hal NL, Vorst O, van Houwelingen AM, Kok EJ, Peijnenburg A, Aharoni A, van Tunen AJ, Keijer J (2000) The application of DNA microarrays in gene expression analysis. *J Biotech* **78**: 271-80

Verma R, Aravind L, Oania R, McDonald WH, Yates III JR, Koonin EV, Deshaies RJ (2002) Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26*S* proteasome. *Science* **298**: 18 611-614

Vincent O, Townley R, Kuchin S, Carlson M (2001) Subcellular localization of the Snf1 kinase is regulated by specific beta subunits and a novel glucose signalling mechanism. *Genes Dev* **15**: 1104-1114

Vyas VK, Kuchin S, Berkey CD, Carlson M (2003) Snf1 kinases with different beta-subunit isoforms play distinct roles in regulating haploid invasive growth. *Mol Cell Biol* **23**: 1341-1348

Weber H, Borisjuk L, Wobus U (1996) Controlling seed development and seed size in *Vicia faba*: a role for seed coat-associated invertases and carbohydrate state. *Plant J* **10**: 823-834

Weber H, Rolletschek H, Heim U, Golombek S, Gubatz S, Wobus U (2000) Antisense-inhibition of ADP-glucose pyrophosphorylase in developing seeds of Vicia narbonensis moderately decreases starch but increases protein content and affects seed maturation. *Plant J* **24**: 33-43

White J, Todd J, Newman T, Focks N, Girke T, Martinez de Ilarduya Ol, Jaworski J, Ohlrogge J, Benning C (2000) A new set of *Arabidopsis* expressed sequence tags from developming seeds. The metabolic pathway from carbohydraes to seed oil. *Plant Phys* **124**: 1582-1594

Willmott RL, Rushton PJ, Hooley R, Lazarus CM (1998) DNase1 footprints suggest the involvement of at least three types of transcription factors in the regulation of alpha-Amy2/A by gibberellin. *Plant Mol Biol* **38**: 817-825

Winder WW, Wilson HA, Hardie DG, Rasmussen BB, Hutber CA, Call GB, Clayton RD, Conley LM, Yoon S, Zhou B (1997) Phosphorylation of rat muscle acetyl-CoA carboxylase by AMP-activated protein kinase and protein kinase A. *J Appl Physiol* **82**: 219-225.

Wisman E, Ohlrogge J (2000) *Arabidopsis* microarray service facilities. *Plant Physiol* **124**: 1468-1471

Woo HH, Brigham LA, Hawes MC (1995) Molecular cloning and expression of mRNAs encoding H1 histone and an H1 histone-like sequences in root tips of pea (*Pisum sativum* L.). *Plant Mol Biol* **28**: 1143-1147

Xiping Niu, Tim Helentjaris and Nicholas J. Bate (2002) Maize ABI4 Binds Coupling Element1 in Abscisic Acid and Sugar Response Genes. *The Plant Cell* **14**: 2565-2575

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119

Young ET, Kacherovsky N, Van Riper K (2002) Snf1 protein kinase regulates Adr1 binding to chromatin but not transcription activation. *J Biol Chem* **277**: 38095-38103

Zakharov A, Giersberg M, Hosein F, Melzer M, Muntz K, Saalbach I (2004) Specific promoters direct gene expression in non-seed tissue. *J Exp Bot* **55**:1463-1471

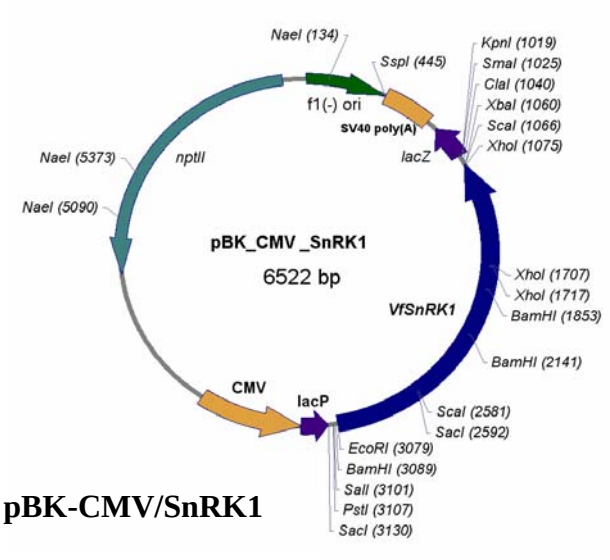
Zaragoza O, Vincent O, Gancedo JM (2001) Regulatory elements in the FBP1 promoter respond differently to glucose-dependent signals in *Saccharomyces cerevisiae*. *Biochem J* **359**: 193-201

Zhang Y, Barcelo P, Shewry PR, Jones HD, Davey, Lazzeri PA, Halford NG (2000) Antisense expression of a SnRK1 protein kinase sequence in transgenic barley (*Hordeum vulgare* L): evidence of an important role for SnRK protein kinase in carbohydrate metabolism during pollen development. In: 8th Intern. Barley Genetics Symp. (Ed. S. Logue) Adelaide Munning Printers Pty **3**: 134-136

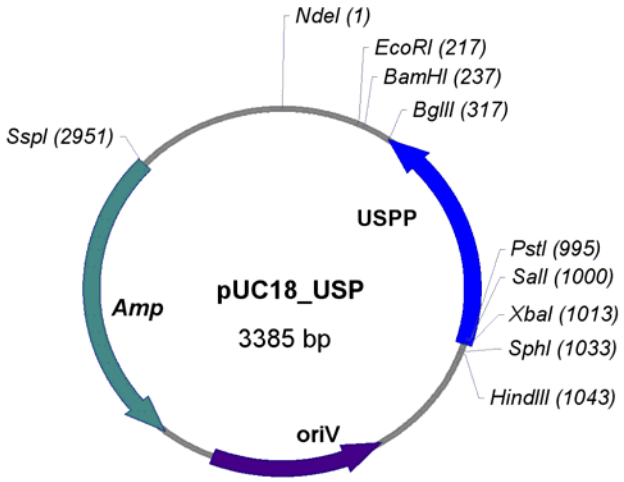
Zimmermann FK, Kaufmann I, Rasenberger H, Haubetamann P (1977) Genetics of carbon catabolite repression in Saccharomycess cerevisiae: genes involved in the derepression process. *Mol Gen Genet* **151**: 95-103

9. APPENDIX

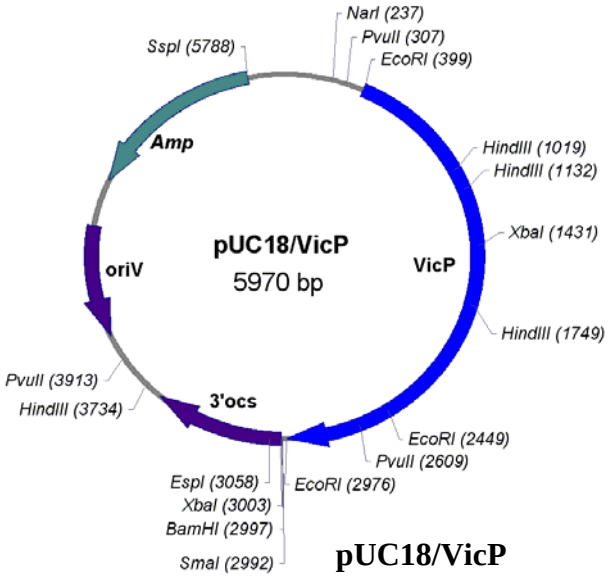
9.1 Cloning maps



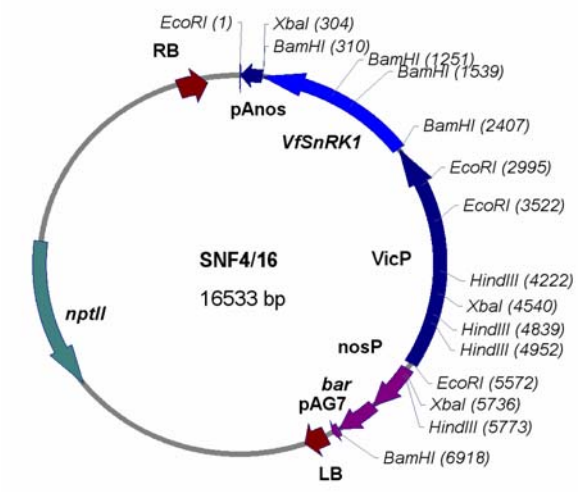
pBK-CMV/SnRK1



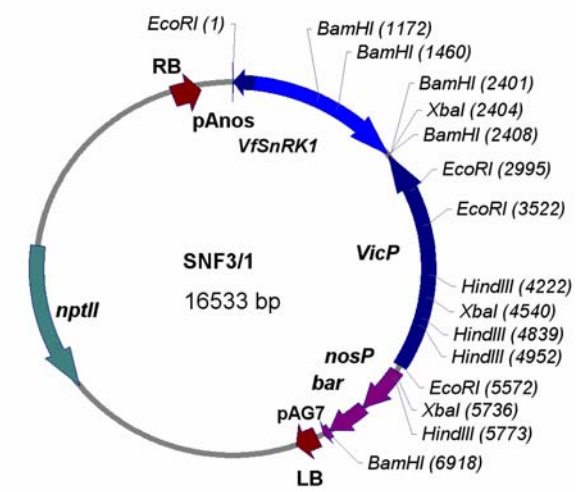
pUC18/USPP



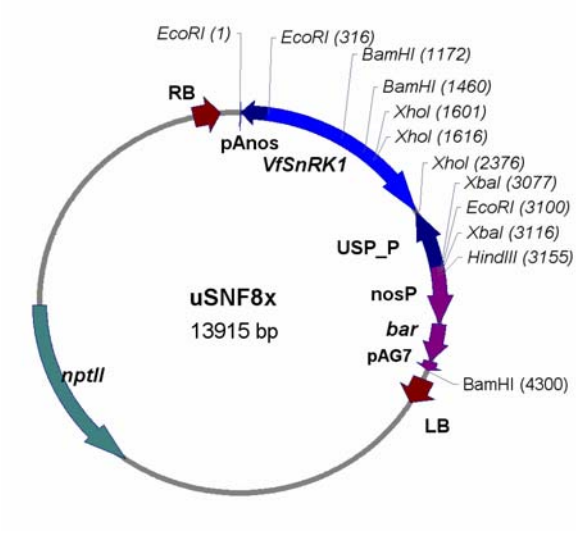
pUC18/VicP



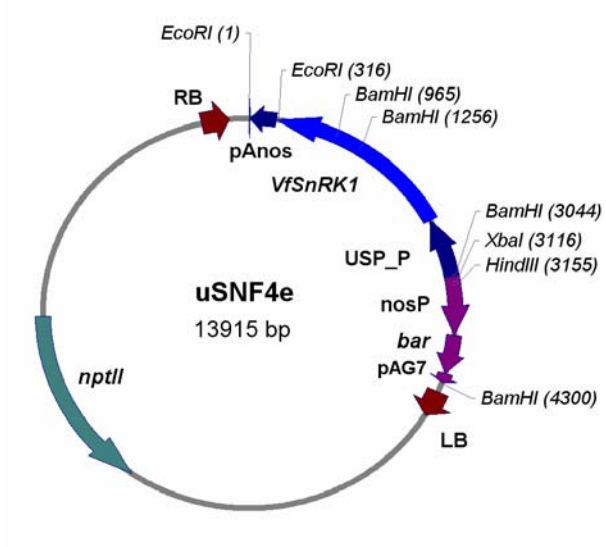
SNF4/16



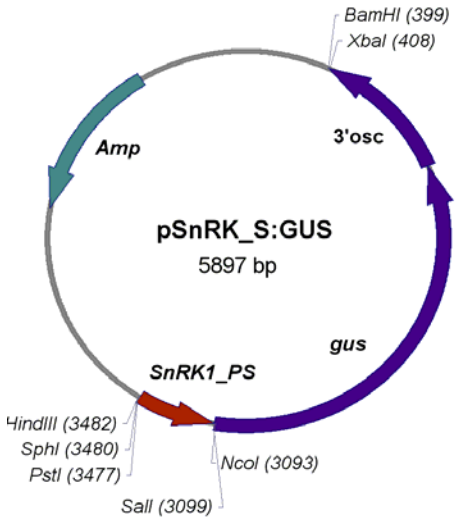
SNF3/1



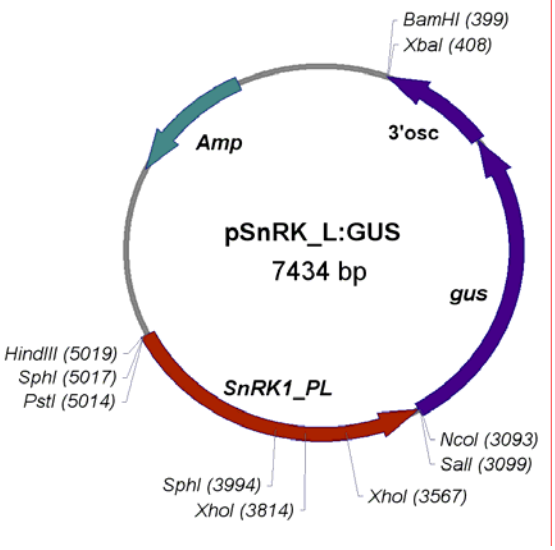
uSNF8x



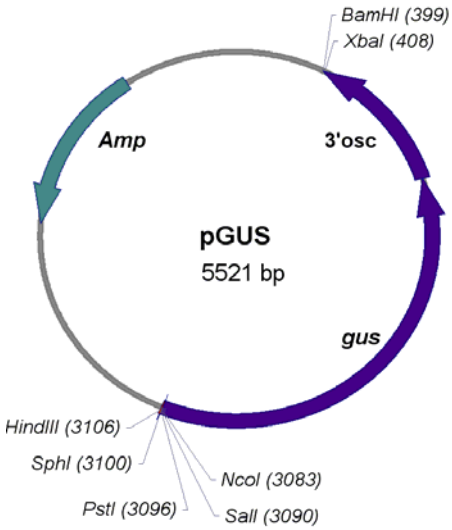
uSNF4e



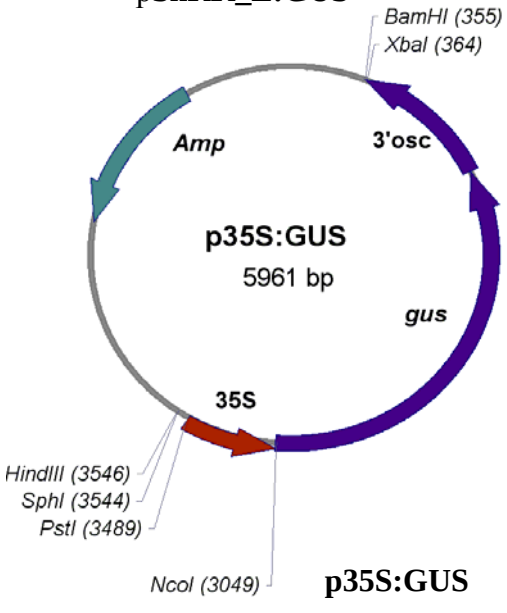
p**SnRK_S:GUS**



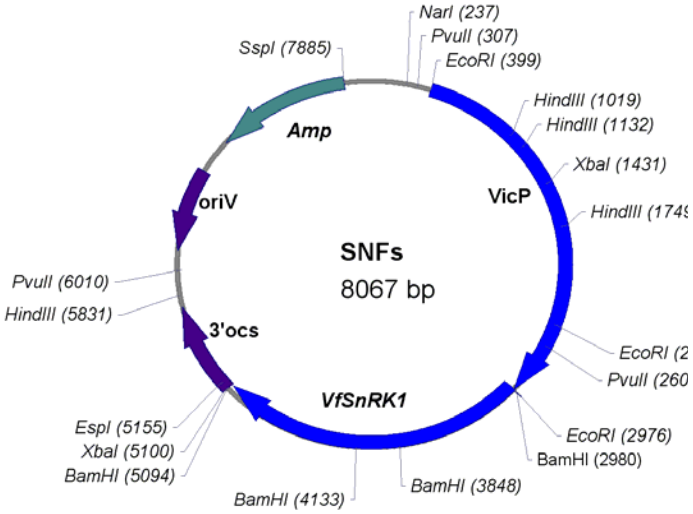
p**SnRK_L:GUS**



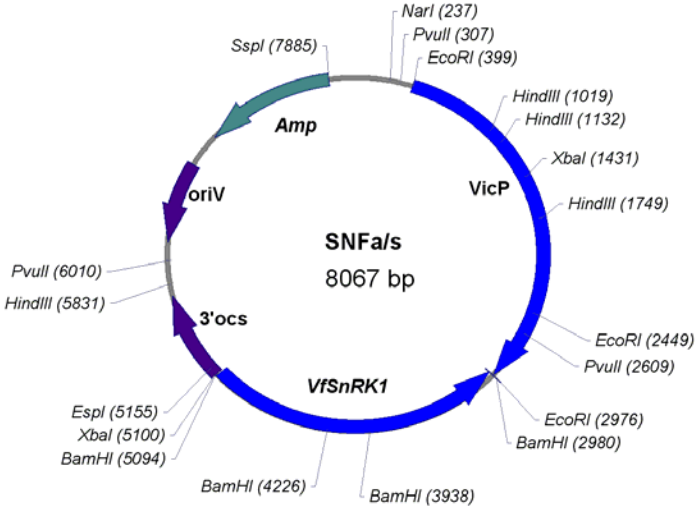
pGUS



p35S:GUS



SNFs



SNFas

VfSnRK 1- *V. faba* SNF1-related protein kinase gene, **CMV** - promoter for eukaryotic expression, **SV40 poly(A)** - polyadenylation sequence provides signal required for termination of eukaryotic expression, **lacP** - IPTG inducible promoter for prokaryotic expression, ***lacZ*** - the coding sequence of *lacZ* gene provides α-complementation for blue-white colour selection of recombinant plasmides, ***nptII*** - neomycin- and kanamycin-resistance gene for prokaryotic selection, **f1(-) ori** – origin of replication, **VicP** - vicilin promoter for expression in plant, **USP_P** - USP (unknown seed protein) promoter for expression in plant, **pAnoc** - polyadenylation sequence from nopaline synthase, ***nosP*** - promoter from nopaline synthase gene, ***bar*** - selective gene of phosphinothricine resistance, **pAG7** - polyadenylation sequence from agropine synthase gene 7, ***nptII*** - neomycin- and kanamycin-resistance gene for prokaryotic selection, **LB** - left border, **RB** - right border, ***gus*** - glucuronidase reporter gene, **3'ocs** - polyadenylation sequence from octopine synthase, Amp – ampicilline resistance gene for prokaryotic expression, **SnRK1_PS** – 376 bp 5'-flanking region of *VfSnRK1*, **SnRK1_PL** - 1,915 Kb 5'-flanking region of *VfSnRK1*.

9.1 List of ESTs

Table 1. The ESTs from cluster A

EST_ID		E-value of BLASTX result	Accesion number	*Species*	Description	EC number
1. Metabolism						
1.1 Carbohydrate Metabolism						
PSC22H11u	Glycolysis / Gluconeogenesis,	3,00E-84	P12886	*P. sativum*	Alcohol dehydrogenase	1.1.1.1
PSS07H15u	Glycolysis / Gluconeogenesis,	9,00E-79	P12886	*P. sativum*	Alcohol dehydrogenase	1.1.1.1
PSS09H10u	Glycolysis / Gluconeogenesis,	4,00E-21	P12886	*P. sativum*	Alcohol dehydrogenase	1.1.1.1
PSC33C10u	Pentose phosphate pathway	6,00E-62	NP_565392.1	*A. thaliana*	Gluconolactonase	3.1.1.17
1.2 Energy Metabolism						
PSS07D09u	Carbon fixation	1,00E-48	AAO15575.1	*Lupinus albus*	Malate dehydrogenase	1.1.1.37
PSC30A07u	Carbon fixation	2,00E-73	E96835	*A. thaliana*	Phosphoribulokinase	2.7.1.19
1.6 Metabolism of Other Amino Acids						
PSC30K22u	Glycosphingolipid metabolism	4,00E-57	AAK73021.1	*Gossypium arboreum*	Ceramide glucosyltransferase	
1.9 Metabolism of Cofactors and Vitamins						
PSC25P16u	Pantothenate and CoA biosynthesis	4,00E-63	CAB61890.1	*P. sativum*	Acetohydroxy acid isomeroreductase	1.1.1.86
1.10 Biosynthesis of Secondary Metabolites						
PSC26L10u	Flavonoid biosynthesis	3,00E-35	BAB60720.1	*N. tabacum*	Glucosyltransferase	2.4.1.91
PSC27D17u	Flavonoid biosynthesis	8,00E-37	NP_181994.1	*A. thaliana*	Glucosyltransferase	2.4.1.91
2. Genetic Information Processing						
2.1 Transcription						
PSC25E19u	Unclassified transcriptional regulators	6,00E-59	A84668	*A. thaliana*	Argonaute (AGO1)-like protein	
PSC27J23u	Unclassified transcriptional regulators	6,00E-46	T00420	*A. thaliana*	Small nuclear ribonucleoprotein D2	
PSC27N04u	Unclassified transcriptional regulators	2,00E-57	NP_188725.1	*A. thaliana*	Heterogeneous nuclear ribonucleoprotein	
PSC30D03u	Transcription factors	2,00E-87	BAC10553.1	*P. sativum*	ABA insensitive 3	
2.2 Translation						
PSC27H08u	Aminoacyl-tRNA biosynthesis	2,00E-53	CAC12821.1	*N. tabacum*	Lysyl-tRNA synthetase	6.1.1.6
PSC29N17u	Aminoacyl-tRNA biosynthesis	2,00E-52	NP_565938.1	*A. thaliana*	Methionyl-tRNA synthetase	6.1.1.10
PSC31E12u	Ribosome Ortholog	1,00E-58	O49884	*Lupinus luteus*	60s ribosomal protein L30	

2.3 Folding, Sorting and Degradation					
PSC33E02u	Other proteins (folding, sorting)	5,00E-39	P21747	*V. faba*	Embryonic abundant protein precursor (clone pUSP92)
PSC34D04u	Other proteins (folding, sorting)	4,00E-91	P21745	*V. faba*	Embryonic abundant protein precursor (clone USP Vf30.1)
PSC21D24u	Storage proteins	8,00E-73	P13918	*P. sativum*	Vicilin precursor
PSC33J21u	Storage proteins	e-101	P13918	*P. sativum*	Vicilin precursor
PSC21C18u	Storage proteins	2,00E-68	S35757	*P. sativum*	Vicilin, 47K
PSC24I14u	Storage proteins	8,00E-60	S35757	*P. sativum*	Vicilin, 47K
PSC26D22u	Storage proteins	3,00E-76	S35757	*P. sativum*	Vicilin, 47K
PSC33I08u	Storage proteins	2,00E-58	S35757	*P. sativum*	Vicilin, 47K
PSC25B20u	Proteasome	4,00E-86	NP_197745.1	*A. thaliana*	26S proteasome regulatory subunit (RPN11)
2.4 Replication and Repair					
PSC23A09u		3,00E-24	AAF35250.1	*Zea mays*	Mismatch binding protein Mus3
3. Environmental Information Processing					
3.1 Membrane Transport					
PSC35H04u	Major facilitator superfamily	2,00E-44	T49900	*A. thaliana*	2-oxoglutarate/malate translocator
3.2 Signal Transduction					
PSC33F05u	Nucleotid-binding protein	2,00E-22	NP_193689.1	*A. thaliana*	ATP binding protein
PSC23G17u		3,00E-44	NP_680572.1	*A. thaliana*	Protein phosphatase 2C (PP2C)
PSC24B23u		3,00E-38	AAM18133.1|	*M. truncatula*	Small G-protein ROP3
4. Cellular Processes					
4.1 Cell cycle&chromatin remodelling					
PSC25K01u	Cell cycle	7,00E-78	P50906	*P. sativum*	Translationally controlled tumor protein homolog
PSC30H24u	Histone	3,00E-68	P02300	*M. sativa*	Histone H3
4.2 Cell Growth and Division					
PSC32E08u	Cell growth	1,00E-50	AAN74636.1	*P. sativum*	DEAD box RNA helicase
PSC30L06u	Cell growth	2,00E-71	BAC57282.1	*O. sativa*	Cellulose synthase-4 2.4.1.-
5.Unclassified and unknown					
PSC22P01u		2,00E-44	D96542	*A. thaliana*	Pirin
PSC24L17u		3,00E-99	P02867	*P. sativum*	Lectin precursor
PSC25H03u		5,00E-28	NP_195729.1	*A. thaliana*	Lectin
PSS05J22u		3,00E-25	T09593	*M. sativa*	Cold-inducible protein
PSC21N06u		6,00E-44	NP_190500.1	*A. thaliana*	Putative protein
PSC28P08u		1,00E-34	NP_197742.1	*A. thaliana*	Putative protein
PSC30F17u		2,00E-40	AAM65866.1	*A. thaliana*	Unknown protein
PSC30O15u		2,00E-72	BAC43629.1	*A. thaliana*	Unknown protein
PSC33M14u		5,00E-20	NP_195939.1	*A. thaliana*	Putative protein
PSC33M23u		4,00E-35	BAB89541.1	*O. sativa*	P0435B05.20

6.Low E-value						
PSC27H24u		1,00E-17	AAF91388.1	*Myxococcus xanthus*	SocE	
PSC23A07u		3,00E-17	F86346	*A. thaliana*	Hypothetical protein F24J8.1	
PSC28C14u		3,00E-15	AAO22663.1	*A. thaliana*	Unknown protein	
PSC30A23u		3,00E-12	BAA97212.1	*A. thaliana*	Unknown protein	
PSC21M18u		4,00E-09	NP_175819.1|	*A. thaliana*	Hypothetical protein	
PSC34M18u		1,00E-03	NP_177792.1	*A. thaliana*	Unknown protein	
PSC32C17u		3,10E-02	AAN71038.1	*Drosophila melanogaster*	AT07931p	
PSC25A11u		3,40E+00	Q01870	*Penicillium chrysogenum*	wetA protein	
PSC27O16u		8,30E+00	AAL08852.1	*Cowdria ruminantium*	NADH dehydrogenase	
PSC21B24u	NONE					
PSC24A12u	NONE					
PSS21O02u	NONE					

Table 2. The ESTs from cluster B

EST_ID	**Pathway/ Structural component**	**E-value of BLASTX results**	**Accession number**	**Species**	**Description**	**EC number**
1. Metabolism						
1.1 Carbohydrate Metabolism						
PSC29O14u	Citrate cycle (TCA cycle)	1,00E-50	AAS49171	*P. sativum*	Isocitrate dehydrogenase 1	1.1.1.42
1.2 Energy Metabolism						
PSS21F16u	ATP synthesis	2,00E-32	P49087	*Zea mays*	Vacuolar ATP synthase catalytic subunit A	3.6.3.14
PSS12B15u	Carbon fixation	2,00E-33	T09541	*Capsicum annuum*	Transketolase 1	2.2.1.1
PSS08G11u	Carbon fixation	3,00E-64	P46257	*P. sativum*	Fructose-bisphosphate aldolase, cytoplasmic	4.1.2.13
PSS21J10u	Photosynthesis	1,00E-57	P16002	*P. sativum*	Plastocyanin, chloroplast precursor	
PSS22N10u	Photosynthesis	6,00E-99	S33775	*P. sativum*	Chlorophyll a/b-binding protein	
2. Genetic Information Processing						
2.1 Transcription						
PSS20B07u	Unclassified transcriptional regulators	2,00E-39	AAF63202.1	*Cucumis sativus*	Poly(A)-binding protein	
2.2 Translation						
PSC31I22u	Ribosome Ortholog	3,00E-68	NP_172256.1	*A. thaliana*	40S Ribosomal protein S15	

PSC24P15u	Ribosome Ortholog	4,00E-66	CAA09042.1	*Cicer arietinum*	40S ribosomal protein S6	
PSC32G17u	Ribosome Ortholog	3,00E-68	AAM65510.1	*A. thaliana*	60S ribosomal protein L4-B	
PSS14C23u	Ribosome Ortholog	3,00E-57	P49210	*Oryza sativa*	60S ribosomal protein L9	
PSC26O24u	Ribosome Ortholog	1,00E-77	P30707	*P. sativum*	60S ribosomal protein L9 (Gibberellin-regulated protein GA)	
PSC30C08u	Ribosome Ortholog	1,00E-80	AAK27801.1	*O. sativa*	60S ribosomal protein L21	
PSS15B03u	Ribosome Ortholog	1,00E-27	BAB86520.1	*O. sativa*	60S ribosomal protein L28-like	
PSS19C01u	Ribosome Ortholog	1,00E-37	AAS47512	*G. max*	60S ribosomal protein L37e	
2.3 Folding, Sorting and Degradation						
PSS20E21u	Protein folding and associated processing	1,00E-46	Q41651	*V. faba*	Peptidyl-prolyl cis-trans isomerase, chloroplast precursor	5.2.1.8
3. Environmental Information Processing						
3.1 Membrane Transport						
PSS16N05u	Ion channels	2,00E-33	CAD56222.1	*Cicer arietinum*	Aquoporin-like water channel protein	
PSC31H07u	Major facilitator superfamily	2,00E-48	CAC84547.1	*N. tabacum*	Dicarboxylate/ tricarboxylate carrier	
PSC26D12u	Major facilitator superfamily	8,00E-82	CAA05979.1	*Lupinus albus*	Adenine nucleotide translocator	
PSC29E14u	Other transporters	4,00E-36	Q39315	*Brassica napus*	Acyl-CoA binding protein	
PSC32K17u	Major facilitator superfamily	8,00E-54	NP_186959.1	*A. thaliana*	Sugar transporter family protein	
PSS21O14u	Other transporters	9,00E-21	NP_181056.1	*A. thaliana*	HEM protein	
3.2 Signal Transduction						
PSS05G23u	Nucleotid-binding protein	1,00E-41	NP_191788.1	*A. thaliana*	ADP-ribosylation factor-like	
PSS20B03u		8,00E-70	P17928	*Medicago sativa*	Calmodulin	
PSS11B24u		3,00E-32	BAB21240.1	*O. sativa*	Putative protein kinase	
4. Cellular Processes						
4.1 Cell cycle&chromatin remodelling						
PSC35H05u	Histone	7,00E-28	S59560	*P. sativum*	Histone H1.41	
PSS23F19u	Histone	3,00E-33	P25470	*P. sativum*	Histone H2A	
PSC24F23u	Histone	4,00E-52	O65759	*Cicer arietinum*	Histone H2A	
PSC32K09u	Histone	5,00E-44	CAA69025.1	*A. thaliana*	Histone H2B	
PSC26O10u	Histone	2,00E-41	AAA34292.1	*T. aestivum*	Histone H4	
PSC23H18u	Histone	4,00E-36	AAA34292.1	*T. aestivum*	Histone H4	

PSS12A07u	Histone	3,00E-36	AAA34292.1	*T. aestivum*	Histone H4	
PSC32E09u	Histone	5,00E-39	AAA34292.1	*T. aestivum*	Histone H4	
PSC24G06u	Histone	8,00E-39	AAA34292.1	*T. aestivum*	Histone H4	
PSC32F09u	Histone	1,00E-45	AAA34292.1	*T. aestivum*	Histone H4	
PSC24J03u	Histone	8,00E-39	AAA34292.1	*T. aestivum*	Histone H4	
4.2 Cell Growth and Division						
PSS20O14u	Cell structure	8,00E-31	S56674	*Fragaria x ananassa*	Annexin homolog RJ4	
PSC31D15u	Cell structure	2,00E-72	CAA75308.1	*M. truncatula*	Annexin	
PSC33N04u	Cell structure	6,00E-80	P29500	*P. sativum*	Tubulin beta-1 chain	
PSS13M16u	Cell structure	2,00E-61	P33623	*Anemia phyllitidis*	Tubulin alpha-1 chain	
4.4 Disease and Defence						
PSC22A22u		5,00E-52	Q05091	*Pyrus communis*	Polygalacturonase inhibitor precursor	
5.Unclassified and unknown						
PSC25G05u		7,00E-59	NP_564562.1	*A. thaliana*	Expressed protein	
PSS15O08u		3,00E-72	NP_758764.1	*Erwinia amylovora*	Hypothetical protein	
PSS10H19u		8,00E-33	NP_181711.1	*A. thaliana*	Unknown protein	
PSS11B14u		2,00E-20	NP_567799.1	*A. thaliana*	Expressed protein	
PSC31P19u		2,00E-33	AAO19365.1	*O. sativa*	Unknown protein	
PSC21L17u		4,00E-33	NP_567210.1	*A. thaliana*	Expressed protein	
PSC24A01u		1,00E-28	T05120	*A. thaliana*	Hypothetical protein F7H19.70	
6.Low E-value						
PSC26I07u		1,00E-19	CAB01913.1	*Sesbania rostrata*	Histone H4	
PSS18J18u		4,00E-19	AAM63485.1	*A. thaliana*	Cytochrome c oxidase subunit	
PSS24N01u		2,00E-14	CAA70550.1	*synthetic construct*		
PSC27B07u		2,00E-11	T08942	*A. thaliana*	Proton pump interactor	
PSS21L10u		9,00E-07	NP_176449.1	*A. thaliana*	Short-chain dehydrogenase/reductase	
PSS12F15u		1,00E-04	CAA70111.1	*A. thaliana*	HSC70-G7 protein	
PSC73K07u		2,00E-04	T09709	*Gossypium hirsutum*	ADP,ATP carrier	
PSS07F17u		2,00E-04	BAB10508.1	*A. thaliana*	Unknown protein	
PSS17E19u		1,30E-01	AAF02153.1	*A. thaliana*	Unknown protein	
PSS21E13u		1,50E+00	P24499	*Trypanosoma brucei*	ATP synthase A chain	
PSS11P15u		5,10E+00	NP_741630.1	*Caenorhabditis elegans*	G-protein coupled receptor	
PSS13O18u	NONE					

PSS16H21u	NONE					
PSS19H08u	NONE					
PSS20L02u	NONE					
PSS18C21u	NONE					
PSS13H17u	NONE					
PSC05C05u	NONE					
PSS14B04u	NONE					
PSC12A13u	NONE					
PSC05L03u	NONE					
PSS24E08u	NONE					
PSS14L11u	NONE					
PSC05M01u	NONE					
PSC21M12u	NONE					
PSS21K24u	NONE					

Acknowledgements

This work was done at the Institute of Plant Genetics and Crop Plant Research (IPK) in Gatersleben. I would like to thank all the co-workers of the Institute, especially from the "Gene Expression" and "Bioinformatics" groups for the scientifically stimulation environment that has promoted this work.

I would like to express my gratitude to Dr. Hans Weber for his careful supervision, helpful and stimulating discussions. I am also grateful to him for his help during writing and correcting this thesis.

I am thankful to Prof. Ulrich Wobus, director of the IPK Gatersleben and head of the Gene Expression group, for giving me the opportunity to work in his research group and for critical reading of the manuscript.

I am also thankful to Dr. Wolodymyr Radchuk, Dr. Renate Manteuffel, Dr. Winfriede Weschke, Dr. Hardy Rolletschek and Dr. Ljudmilla Borisjuk for the fascination discussions and interesting suggestion on various topics of the developmental biology of crop plants and sharing new experimental methods with me.

I am particularly grateful to Dr. Uwe Scholz, Dr. Lothar Altschmied and Dr. Usr Hähnel for introducing me to bioinformatic tool and to handling of robots, also for the performance of annotation of ESTs and clustering analysis.

I wish to thank Mrs. Elsa Fessel, Mrs. Angela Stegmann, Mrs. Kathrin Blaschek and Mrs. Sabina Skiebe for excellent technical assistance, Mrs. Angela Schwarz for her assistance in the transformation of pea and Dr. Annegret Tewes for protoplasts transformation. I thankful to other colleagues from our group, for the nice and stimulating atmosphere in the group, for scientific debates and for their help during experimental work.

I would like to thank to Mrs. Susanne König for DNA sequencing, Mrs. Heike Ernst and Mrs. Schäfer for the photographic work, Mrs. Ursula Tiemann and Mrs. Karin Lipfert for the fast and qualified help with the graphical presentation of the results.

Finally, particularly thanks to my husband and children for the endless patience and belief in me.

Curriculum vitae

Name	Ruslana I. Radchuk
Address	An der alten Mühle 18 06466 Gatersleben Germany
Date of birth	April 22, 1972
Place of birth	Werba, Ukraine
Nationality	Ukrainian, citizen of Ukraine
Marital status	Married, have two children
Languages	Ukrainian (native), Russian (native), German (fluently), English (fluently)

Education and employment

1979-1989	Secondary school, Kozelets, Chernigiv Region, Ukraine
1989-1994	Study at the Department of Genetics, Faculty of Biology of National Taras Shevchenko University of Kyiv. The diploma thesis: "*Agrobacterium tumefaciens*-mediated transformation of tobacco plants with phosphinotricine resistance *bar*-gene“. The practical work was done at the Institute of Plant Physiology and Genetic, National Academy of Sciences of Ukraine.
1994-1996	Scientific researcher at the Department of Biotechnology, Institute of Horticulture, National Academy of Agrarian Sciences of Ukraine, Kyiv. Research interest: studying the biochemical and molecular aspects of plant disease resistance.
1998	Guest scientist at the Federal Centre for Breeding Research on Cultivated Plants (BAZ), Quedlinburg, Germany. Research interest: genetic transformation of oil-seed rape plants, expression of antibiotic nisin in transgenic rape.
January, 1999-present	Post-graduate student at the Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany. The Ph.D. thesis : ” Regulation of seed development in *Leguminosae*: investigating the role of SNF1-related protein kinase”. The work was performed in the Department “Gene Expression” under supervision of Dr. H. Weber.

Papers published

Borisjuk L, Rolletschek H, **Radchuk R**, Weschke W, Wobus U, Weber H (2004) Seed development and differentiation: a role for metabolic regulation. (Review) *Plant Biol*. **6**:375-386

Rolletschek H., Borisjuk L., **Radchuk R**., Miranda M., Heim U., Wobus U., Weber H (2004) Seed-specific expression of a bacterial phosphoenolpyruvate carboxylase in *Vicia narbonensis* increases protein content and improves carbon economy. *Plant Biotech.J*. **2**: 211-219

Weschke W, Panitz R, Gubatz S, Wang Q, **Radchuk R**, Weber H, Wobus U (2003) The role of invertases and hexose transporters in controlling sugar ratios in maternal and filial tissues of barley caryopses during early development. *Plant J* **33**: 395-411

Radchuk VV, Klocke E, **Radchuk RI**, Neumann M, Blume YaB (2000) Production of transgenic rape plants (*Brassica napus* L.) using Agrobacterium tumefaciens. *Genetika* **36**:932-941

Abstract published

Radchuk R, Radchuk V, Weschke W, Wobus U, Weber H (2004) Transcript profiling of pea (*Pisum sativum*) seed development using expressed sequence tag-based (EST) cDNA arrays. *5th European Conference on Grain Legumes*, June 7-11, Dijon, France (page no. 257)

Radchuk R, Radchuk V, Wobus U, Weber H (2004) A hormon-regulated signal transduction pathway in developing pea seeds changed by transgene expression-a cDNA macroarray analysis. *Mini Symposium on Expression Data Analysis*, September 02-03, IPK-Gatersleben, Germany (page no. 11)

Rolletschek H, Borisjuk L, **Radchuk R**, Miranda M, Heim U, Wobus U, Weber H (2004) Seed-specific expression of a bacterial phosphoenolpyruvate carboxylase in *Vicia narbonensis* increases protein content and improves carbon economy. *Tagung "Molekularbiologie der Pflanzen"* , March 09 – 12, Dabringhausen, Germany

Radchuk V, Gubatz S, Borisjuk L, **Radchuk R**, Menteuffel R, Steinbiss HH, Panitz W, Wobus U, Weschke W (2004) A Small cysteine-rich protein expressed in maternal tissues is essential for seed development in barley (*Hordeum vulgare*). *3th Plant Genomics European Meetings*, September 22-25, Lyon, France (page no. 193)

Radchuk V, **Radchuk R**, Tewes A, Wobus U, Weschke W (2004) Characterization of two genes of cobalamine-independent methionine synthase from barley (*Hordeum vulgare*). *IIIth Conf. All-Russ. Genetics and Breeding Soc. "Genetics in XXI century: state of the art and perspectives of development"* June 6-12, Moscow, Russia, V.1 (page no. 260)

Radchuk V, Gubatz S, **Radchuk R**, Panitz W, Manteuffel R, Weschke W, Wobus U (2003) Zelluläre Desintegrationsprozesse an der maternal-filialen Kontaktzone des Gerstenkorns und die mögliche Rolle Cystein-reicher Peptide. *Tagung "Molekularbiologie der Pflanzen"*, February 25 – 28 Dabringhausen, Germany

Radchuk R, Heim U, Wobus U, Weber H (2002) Antisense-repression of SnRK1 causes abnormal pollen development and male sterility in transgenic tobacco and pea plants. *Botanikertagung*, September 22-27, Freiburg , Germany (page no. 198)

Weber H, Borisjuk L, Heim U, Miranda M, Rolletschek H, **Radchuk R**, Wobus U (2002) Functional characterisation of seed-specific sucrose transport in legumes. *Botanikertagung*, September 22-27, Freiburg , Germany (page no. 121)

Radchuk V, **Radchuk R**, Gubatz S, Sreenivasulu N, Wobus U, Weschke W (2002) Cloning and characterisation of a gene spacifically expressed in the nuclear projection of the developing barley grain. *Botanikertagung*, September 22-27, Freiburg , Germany (page no. 27)

Erklärung

gemäss §28 Abs. 2 Promotionsordnung 03.02.2004

Hiermit erkläre ich , dass ich mich mit der vorliegenden wissenschaftlichen Arbeit erstmals um die Erlangung des Doktorgrades bewerbe, die Arbeit selbständig und ohne fremde Hilfe verfasst habe, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Gatersleben, April 2005 Ruslana Radchuk