MOLECULAR CLONING AND *IN SILICO* ANALYSIS OF THREE SOMATIC EMBRYOGENESIS RECEPTOR KINASE mRNA FROM DATE PALM

Imen REKIK, Amine ELLEUCH, Walid KRIAA, Noureddine DRIRA

Laboratory of plant biotechnology, Faculty of sciences of Sfax, University of Sfax. Tunisia

Rekik I., A.. Elleuch, W. Kriaa and N. Drira (2013): *Molecular cloning and in silico analysis of three somatic embryogenesis receptor kinase mRNA from date palm.* Genetika, Vol 45, No. 3, 837-853.

We report here the isolation and characterizations of three somatic embryogenesis receptor kinase (PhSERK) genes from palm date by a rapid amplification of cDNA ends (RACE) approach. PhSERKs belong to a small family of receptor kinase genes, share a conserved structure and extensive sequence homology with previously reported plant SERK genes. Sequence analysis of these genes revealed the sequence size of 11051 pb (PhSERK1), 7981 pb (PhSERK2) and 10510 pb (PhSERK3). The open reading frames of PhSERK1, PhSERK2 and PhSERK3 are 1914 pb, 1797 pb and 1719 pb respectively. PhSERKs belongs to the LRR-type cell surface RLKs, which possess a number of characteristic domains. These include an extracellular domain (EX) containing a variable number of LRR units, signal pepetide (SP) immediately followed by a single transmembrane domain (TM) and an intracellular kinase domain. The phylogenetic tree shows that the protein PhSERK1, PhSERK2 and PhSERK3 clustered within monocots SERKs proteins groups. We also predicted the secondary and tertiary with ligand binding sites structure of the protein PhSERKs.

Key words: *Phoenix dactylifera*; Phylogeny; Protein; Somatic embryogenesis receptor kinase; Structure

INTRODUCTION

The plant receptor-like kinases (RLKs) are a large group of signaling proteins in plants, and are a fundamental part of plant signal transduction. In *Arabidopsis* the RLK family contains more than 600 members, constituting 60% of kinases, including almost all of the transmembrane kinases (SHIU *et al.,* 2003). The position of RLKs in the plasma membrane, with an extracellular receptor domain and an intracellular kinase domain, makes them well suited to the task of perceiving an external signal to the cell and conducting that signal into the cell in order to elicit a response. In addition to RLKs there are a number of receptor-like proteins (RLPs). These proteins contain an extracellular domain similar to a RLK but lack the intracellular kinase

_____

Corresponding author: Imen Rekik, Faculty of sciences of Sfax, Sfax, Tunisia. Tel: 0021674676616, Fax: 0021674274437, email: imenbmc@yahoo.fr

domain (TOR *et al.,* 2009). Based on the criteria of extracellular domain structure and kinase domain phylogeny, RLKs are divided into subfamilies (SHIU *et al.,* 2003). The Somatic Embryogenesis Receptor Kinase (SERK) gene family belongs to the leucine rich repeat (LRR) subfamily of RLKs. These RLKs contain varying numbers of LRRs in their extracellular receptor domain. SERK genes belong to subgroup II (LRRII) and contain five LRR domains (SHIU *et al.,* 2003). The family has been defined according to several factors. The first is the presence of 11 exons with conserved splicing boundaries and the tendency for each exon to encode a specific protein domain. Secondly the SERK amino acid sequence contains a particular order of domains from N to C-terminal: Signal peptide (SP), leucine zipper (ZIP), 5 LRRs, a proline-rich domain (SPP), transmembrane, kinase and C-terminal domains. The SPP domain, containing the SPP motif and the C terminal domain are considered to be the characteristic domains of SERK proteins (BAUDINO *et al.,* 2001; HECHT *et al.,* 2001). Although this is largely correct for annotated SERK genes there is some divergence from the set criteria. The first SERK genes identified were linked to competence of cultured cells to form somatic embryos in carrot (*Daucus carota*), orchard grass (*Dactylis glomerata*) and *Arabidopsis thaliana* species (SCHMIDT *et al.,* 1997; Somleva *et al.,* 2000; Hecht *et al.,* 2001). Since that time SERK gene expression has been associated with somatic embryogenesis (SE) and organogenesis in numerous species (NOLAN *et al.,* 2003; ZAKIZADEH *et al.,* 2010). In *Arabidopsis* five SERK genes have been identified (HECHT *et al.,* 2001) (AtSERKs 1-5) and the gene functioning in SE is AtSERK1 (locus At1g71830). As understanding of the roles of the different members of the SERK gene family has increased, it has become apparent that these genes function in diverse signaling pathways with roles from development to defense. The *Arabidopsis* SERK gene family is subdivided into two subfamilies, generated from an ancestral gene duplication event. The first subfamily consists of AtSERKs 1 and 2 (SERK1/2) and the second subfamily, AtSERKs 3, 4 and 5 (SERK3/ 4/5) (HECHT *et al.,* 2001; HE *et al.,* 2007; ALBRECHT *et al.,* 2008).

AtSERK1 is required in conjunction with AtSERK2 for anther development and male gametophyte maturation, with double mutants lacking a failing to develop pollen (ALBRECHT *et al.,* 2005; COLCOMBET *et al.,* 2005). AtSERK1 and AtSERK3 (also called BRI1-associated kinase1 (BAK1)) function in brassinosteroid (BR) signal transduction as components of the BR receptor complex, through dimerization with brassinosteroid-insensitive 1 (BRI1) kinase (LI *et al.,* 2002; KARLOVA *et al.,* 2006), both AtSERK3 and AtSERK4 (also called BAK1-LIKE 1 (BKK1)) have been linked to programmed cell death, which can function in both developmental and pathogen defence roles (HE *et al.,* 2007; KEMMERLING *et al.,* 2007). What has emerged from studies of *Arabidopsis* SERK signaling is that these genes have a tendency to be redundant in pairs with different pairs working in different pathways. Therefore single SERK gene mutants show weak or no phenotype as a second SERK gene can complement their function. Different combinations of SERK genes act in different pathways and these combinations vary according to the pathway. For instance, AtSERK1 and 2 can complement each other in anther development, where AtSERK3 is shown not to function (ALBRECHT *et al.,* 2008). However, AtSERK1 and 3 function together in BR signaling and AtSERK3 and 4 are redundant in the programmed cell death pathway. So far a function for AtSERK5 is not known. In defence responses, AtSERK3/BAK1 functions in pathogen-associated molecular pattern (PAMP)-triggered immunity through heterodimerization with the Flagellin sensing 2 (FLS2) receptor kinase in response to binding by the bacterial PAMP, flagellin (CHINCHILLA *et al.,* 2007; HEESE *et al.,*

2007). Rice SERK, OsSERK1, shows activity in both somatic embryogenesis and fungal defence (HU *et al.,* 2005).

In this study, we isolated and characterized three new SERKs mRNA from date palm and we performed protein structure prediction of date palm. Our analyses suggested that these SERKs was orthologous to other plant SERKs, and the possible role of PhSERKs would be crucial for the explanation fungal defense of *Phoenix dactylifera.*

MATERIALS AND METHODS

Fungal material and elicitor preparation

F. oxysporum Alternaria isolate used in this study was isolated from healthy date palm tissues. Fungal cultures were routinely grown in the dark on malt extract agar (Sigma product) medium at 26±/2 ˚C. For elicitor preparation, 100 ml of sterile malt extract liquid medium in an Erlenmeyer flask was inoculated with agar plugs (5 mm) from 8 days old fungal cultures grown on malt extract agar medium. Cultures in liquid medium were incubated for 15 days at 26±/2 ˚C on a gyratory shaker at 125 rpm and then centrifuged for 15 min 4000 rpm to remove spores and mycelium. The supernatant was aseptically collected and represented the fungal elicitor preparation that was directly used for callus elicitation.

In vitro culture conditions

The immature inflorescences of *Phoenix dactylifera* cv deglet Nour were sterilized with 0.01% (w/v) mercuric chloride ($HgCl_2$) for 1 h. The explants were cultured on the callus induction media. The latter consisted of basal Murashige and Skoog (MS) salts supplemented with KH_2PO_4 (120 mg/l), myo-inositol (100 mg/l), glycine (2 mg/ l), L-glutamine (100 mg/l), adenine (25 mg/l), nicotinic acid (0.5 mg/l), HCL-pyrodoxin (0.5 mg/l), sucrose (50 g/l), and agar (8 g/l). All the cultures were maintained in darkness at 28 ˚C to obtain embryogenic or organogenic calli for 6 to 12 months. The latter were transferred to the culture room at 26-28˚C and 16 hour day-light condition. The embryogenic calli obtained were incubated in a half strength MS liquid test medium supplemented with activated charcoal (0.3 g/l) and 2, 4-D (1 mg/l) (Fki *et al.,* 2003; Kriaa *et al.,* 2012). Test medium representing the elicitation treatment consisted of MS medium where water was replaced by the fungal culture medium extract. They were then kept on a rotary shaker set at 120 rpm. The callus of date palm were collected separately and frozen immediately in liquid nitrogen and stored at -80°C for RNA extraction.

cDNA library construction (mRNA isolation, cDNA synthesis and cloning)

A *Phoenix dactylifera* deglet Nour cultivar callus were ground to a fine powder in liquid nitrogen, and homogenized in total RNA extraction buffer, according to Cathala *et al.,* (1983), except that PVP was added to the extraction buffer. The total mRNAs were selected by oligo (dT) cellulose chromatography, using the mRNA Purification Kit (QIAGEN) according to the manufacturer's specification. Double stranded cDNA was synthesized using a cDNA Synthesis Kit (QIAGEN), following the manufacturer's instructions, cDNAs were ligated to *EcoRI* adaptators, cloned into the *EcoRI* restriction site of pUC 18 and used to transform JM 105 cells. Bacteria were plated on L-Broth agarose medium containing 50 mg/ml ampicillin, 0.1 / µM IPTG and 0.1% X-Gal. fifty white colonies were randomly selected out of 1500 and screened for the presence of recombinant plasmids with inserts greater than 400 bp. Plasmid DNA from these colonies was prepared by a large scale procedure, as described by LEBRUN *et al.,* (1988).

Isolation of PhSERKs cDNA

 Sequencing and BlastX (ALTSCHUL *et al.,* 1997) of some clone reveled partial homology with a region of SERKs genes. Completion of the others domains and 5'untranslated region (UTR) of putative PhSERKs was done using the rapid amplification of 5'cDNA ends (5' RACE) kit (Invitrogen) according to the manufacturer's instructions. The first strand cDNAs was synthesized from mRNAs using gene-specific primers and gene specific nested primers (Tab. 1). In brief, the first strand cDNA synthesis was done with SuperscriptTM II RT, RNA template removed by RNase mix, purified in the SNAP columns, followed by homopolymeric tailing of cDNA and then amplification of the target cDNA. The PCR product was cloned in pGEMTeasy vector and sequenced using M13 modified primers (Tab. 1). All parts of PhSERKs cDNA are combined to obtain a full length cDNA.

Table 1. Sequence of primers used in 5' RACE PCR

PhSERK1_R 5'TCATCTCAACTTCCGTCTGAAA 3'
PhSERK1_NR 5'TGGTGTGCGCTCTTCTTTTA 3'
PhSERK2_R 5'-TCATCTTGGGCCAGACAGTTCCACGGC-3'
PhSERK2_NR 5'-CTCCTGAGAGGTTGTTGTTTGAGAGATC-3'
PhSERK3_R 5'-CAGATGGCACTAAGATAGCCGTAAAGCGGTTAAC-3'
PhSERK3_NR 5'-CTGTTTGTGTTGTACAGAACCCAATC-3'

Prediction of Putative Signal Peptide and Protein and Phylogenetic Analysis

        The signal peptide was predicted by using the "SignalP 4.0 Server" tool (http://www.cbs.dtu.dk/servicesSignalP/). The transmembrane region (TM) was predicted by RYHTHM (http://proteinformatics.charite.de/rhythm/). We predicted the secondary and tertiary structure with ligand binding sites of the proteins PhSERKs using the programs Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2).

        We have used interPro (http://www.ebi.ac.uk/interpro/) and ScanProsite (http://us.expasy.org/prosite/) to analyze the function structure of PhSERKs proteins. Detection of potential recombinant between PhSERK mRNA sequences was carried out using RDP (Recombination Detection Program) version 2.0. Sequences were aligned using CLUSTALW and the resulting alignments were visually inspected and manually edited. Thirty nine SERK protein sequences were retrieved from GenBank using BLASTP and TBLASTN with PhSERK1, PhSERK2 and PhSERK3 as query. The tree of SERK proteins was constructed with MEGA 5.1 (TAMURA *et al.,* 2011) using the Maximum Likelihood method based on the Poisson correction model. All positions containing gaps and missing data were eliminated. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed.

RESULTS AND DISCUSSION

Molecular Cloning of PhSERKs cDNA and genes prediction

In monocots, presence of the basal meristem in the leaf base makes it an ideal material for studies on somatic embryogenesis. We sequenced the fifty above-mentioned cDNA clone and nucleotide sequences were searched against GenBank databases using the BLASTX algorithm. We found that a 1kb, 1.2Kb and 1.1 pb fragments have a strong homology with protein kinase SERK1, SERK2 and SERK3 respectively. These parts corresponds to kinase domain, C terminal domain and 3 ' untranslated region of SERKs proteins. Completion of the others domains and 5'untranslated region (UTR) of putative PhSERKs was done using the rapid amplification of 5'cDNA ends using RT PCR (5' RACE) to obtain complete PhSERKs cDNAs. A full-length SERKs cDNA was isolated from a *Phoenix dactylifera* F. oxysporum induced cDNA library. The sequences were then verified by obtaining full-length cDNAs in a single RT- PCR reaction, independently, and resequencing.

The mRNA or ORF sequence lengths of PhSERK1 and PhSERK2 are 1914pb and 1797pb respectively. The 3'UTR are 384 and 376 bp and 5'UTR are 315 bp and 390 bp, with a predicted amino acid length of 637 and 598 for PhSERK1 and PhSERK2, respectively. PhSERK3 mRNA is 1734 bp in length, with 185 and 165 bp long 5' and 3'UTR, respectively, and a predicted amino acid length of 577aa (Fig. 1A; 1B; 1C). The predicted molecular mass of PhSERK1, PhSERK2 and PhSERK3 is 73, 67 and 64 kDa, respectively. Using the program RDP V 2, the recombination hypothesis was tested and the highest probabilities were found between the three PhSERKs. These probabilities were well supported by all six methods tested in RDP2. Results with GENECONV program revealed that PhSERK1 may result from a hypothetic recombination between PhSERK2 and PhSERK3 in position (135 pb- 308 pb) and (391 pb- 431 pb). Since there is not enough genomic sequences of *Phoenix dactylifera cv deglet Nour* in the NCBI website, we sought the homologue PhSERKs gene family in the Khalas variety. Furthermore, the exon–intron structure of PhtSERKs gene was deduced. We compared the PhSERK1, PhSERK2 and PhSERK3 cDNAs and whole genome shotgun sequence of Khalas variety witch reference in GenBank are ACYX02000007.1, ACYX02043306.1 and ACYX02050452.1 respectively using the BLASTN program (Al-Dous *et al.,* 2011). This comparison and gene prediction using FGENESH web site (http://linux1.softberry.com/) showed that the PhSERK1 gene (11051 pb) (Fig.1A) and PhSERK2 (7981 pb) (Fig.1B) organized into 10 introns and 11 exons. All SERK genes of monocots studied so far have conserved structures (11 exons and 10 introns) (Hu *et al.,* 2005; Jun Ma *et al.,* 2012). However the PhSERK3 (10510 pb) (FIG. 1C) gene has 10 exons and 9 introns unlike the other genes SERKS.

PhSERKs proteins prediction

As shown in figure 2 the alignment, PhSERKs belongs to the LRR-type cell surface RLKs, which possess a number of characteristic domains. These include an extracellular domain (EX) containing a variable number of LRR units and signal pepetide (SP) immediately followed by a single transmembrane domain (TM) and an intracellular kinase domain responsible for phosphorylating downstream proteins (HECHT *et al.,* 2001).

```
   1 atggcggccgtggagcgggacgcgattgtgctttggtttctatgg
     M  A  A  V  E  R  D  A  I  V  L  W  F  L  W
  46 ctgatgctggttttccaccccgtcactagggttcttgccaacgtg
     L  M  L  V  F  H  P  V  T  R  V  L  A  N  V
  91 gaaggtttatgggctgatggaggtgatgcattgcatactttacgg
     E  G  L  W  A  D  G  G  D  A  L  H  T  L  R
 136 accaacctcaatgatcccagtaatgtgctgcaaagttgggatcca
     T  N  L  N  D  P  S  N  V  L  Q  S  W  D  P
 181 actctggtcaatccatgcacatggttccatgttacatgtaacaat
     T  L  V  N  P  C  T  W  F  H  V  T  C  N  N
 226 gacaatagtgttattagagttgatcttggaaatgcacaattatct
     D  N  S  V  I  R  V  D  L  G  N  A  Q  L  S
 271 ggtacgttggtcccccagcttggccagttaaaaaaacctgcaatac
     G  T  L  V  P  Q  L  G  Q  L  K  N  L  Q  Y
 316 ctggaactttatagcaataacattagtggaacaattcctactgag
     L  E  L  Y  S  N  N  I  S  G  T  I  P  T  E
 361 ctggggagtctgacaaacctggatagcggcttaacaacaacagt
     L  G  S  L  T  N  L  D  S  R  L  N  N  N  S
 406 ctgtcgggtcaaattcctccatcattaaccaagatcaataacctc
     L  S  G  Q  I  P  P  S  L  T  K  I  N  N  L
 451 caagttctggatctgtcaacaaacaattatctggagaggttcca
     Q  V  L  D  L  S  T  N  N  Y  L  E  R  F  P
 496 tcaactggctctttttcactattcacccccatcagcgggattgct
     S  T  G  S  F  S  L  F  T  P  I  S  G  I  A
 541 ctaacactcaccatggagactttatcctgcagttttaataacaat
     L  T  L  T  M  E  T  L  S  C  S  F  N  N  N
 586 cctctattatgtggcccgagtacgggaaagccttgtccgaattct
     P  L  L  C  G  P  S  T  G  K  P  C  P  N  S
 631 cctccattttctccacctcctccatataatccaacaccaaactct
     P  P  F  S  P  P  P  P  Y  N  P  T  P  N  S
 676 tcaccaggaagtagtgcctctagcactggagcaattgctggggggg
     S  P  G  S  S  A  S  S  T  G  A  I  A  G
 721 gttgctgcaggtgctgctttgctgtttgctgctgcacctgcaattggg
     V  A  A  G  A  A  L  L  F  A  A  P  A  I  G
 766 tttgcatggtggcgtcgtcgtaaaccgcaagaacatttctttgat
     F  A  W  W  R  R  R  K  P  Q  E  H  F  F  D
 811 gtacctgctgaggaggacccctgaagttcatctggccagcttaaa
     V  P  A  E  E  D  P  E  V  H  L  G  Q  L  K
 856 aggtttctttgcgagaactgcaagttgcaactgatagttttagc
     R  F  S  L  R  E  L  Q  V  A  T  D  S  F  S
 901 aataagaatattcttggcagagggagggtttggaaaggtctataaa
     N  K  N  I  L  G  R  G  G  F  G  K  V  Y  K
 946 ggacgacttgcagatggttcattagtggcagtaaaaagactaaaa
     G  R  L  A  D  G  S  L  V  A  V  K  R  L  K
 991 gaagagcgcacaccaggtggggagcttcagtttcagacggaagtt
     E  E  R  T  P  G  G  E  L  Q  F  Q  T  E  V
1036 gagatgattagcatggctgtacatagaaacttgctccgtcttcgt
     E  M  I  S  M  A  V  H  R  N  L  L  R  L  R
1081 gggttttgcatgacgcccacagagcggctgcttgtctatccctat
     G  F  C  M  T  P  T  E  R  L  L  V  Y  P  Y
1126 atggctaatggaagtgttgcatcatgcttgagagagcggcctcca
     M  A  N  G  S  V  A  S  C  L  R  E  R  P  P
1171 tctgagccaccacttttcttggccagcttgaagaagaattgcactg
     S  E  P  P  L  L  S  W  P  A  R  R  I  A  L
1216 gggtctgcaagaggactgtcttatttgcatgatcatgtgatcca
     G  S  A  R  G  L  S  Y  L  H  D  H  C  D  P
1261 aagattattcatcgtgatgtcaaagctgcaaatattttgttggat
     K  I  I  H  R  D  V  K  A  A  N  I  L  L  D
1306 gaagagtttgaggcagttgtaggagactttggcttggccaaactt
     E  E  F  E  A  V  V  G  D  F  G  L  A  K  L
1351 atggactacaaggatacacatgtaacaactgcggtgcgtggaaca
     M  D  Y  K  D  T  H  V  T  T  A  V  R  G  T
1396 attgggcacattgctccagagtatctatccactggaaagtcctcg
     I  G  H  I  A  P  E  Y  L  S  T  G  K  S  S
1441 gagaagactgatgtttttgggtatggaatcatgcttctcgagctc
     E  K  T  D  V  F  G  Y  G  I  M  L  L  E  L
1486 attacagggcaaagagcatttgatcttgctcgtctcgcaaatgat
     I  T  G  Q  R  A  F  D  L  A  R  L  A  N  D
1531 gatgatgtgatgttgctcgactgggtaaaaggactcctgaaagat
     D  D  V  M  L  L  D  W  V  K  G  L  L  K  D
1576 agaaggctggaaacgttggtcgaccctgatctacaaaacaatttt
     R  R  L  E  T  L  V  D  P  D  L  Q  N  N  F
1621 gtagaaactgaggtggagtcactcatacaggttgccctgctctgc
     V  E  T  E  V  E  S  L  I  Q  V  A  L  L  C
1666 acacaggggctccccaatgacggccaaaaatgtcagatgtggtg
     T  Q  G  S  P  M  D  R  P  K  M  S  D  V  V
1711 agaatgctcgaaggcgatggccttgctgagatggggaggaatgg
     R  M  L  E  G  D  G  L  A  E  R  W  E  E  W
1756 cagcaagtggaggtggtacacaggagtggagatggccccacac
     Q  Q  V  E  V  V  H  Q  E  V  E  M  A  P  H
1801 cgcaactccgaatggtttatcgactccacggacaacctgcctgcg
     R  N  S  E  W  F  I  D  S  T  D  N  L  P  A
1846 gttgaattgtcaggtccaaggttgcagatccagaaaaccatctcg
     V  E  L  S  G  P  R  L  Q  I  P  E  T  I  S
1891 ctttatgcatttattggcatttga 1914
     L  Y  A  F  I  G  I  *
```

```
   1  atggagatagaagttaacctcattctggagatattgatttgcttt
      M  E  I  E  V  N  L  I  L  E  I  L  I  C  F
  46  aaattgaaatctttctctctgaggaaggttataaacacctagggcc
      K  L  K  S  F  S  L  R  K  V  I  T  P  R  A
  91  cagggttaacctcattctggagatattgatttgctttaaattgaaa
      Q  V  N  L  I  L  E  I  L  I  C  F  K  L  K
 136  tctttctctctgaggaaggttataaacacctagggcccagaagaaa
      S  F  S  L  R  K  V  I  T  P  R  A  Q  K  K
 181  tgccgtcttggtataaacttttcatgacacggaactttacagtaac
      C  R  L  G  I  N  F  H  D  R  E  L  Y  S  N
 226  aatatagtggcatgattcctagtgaccttggaaatctgacaaat
      N  I  S  G  M  I  P  S  D  L  G  N  L  T  N
 271  ttgctgtcggatgttgcatttttcataccccaactttaaaccaa
      L  L  S  D  V  A  F  F  I  P  P  L  N  Q
 316  cttgttcatgagattaaccatagtgttcaaaacgcaccggatgag
      L  V  H  E  I  N  H  S  V  Q  N  A  P  D  E
 361  cagggatttgtcaaacaacaacttatcaggagaagtgccatcaac
      Q  G  F  V  K  Q  Q  L  I  R  R  S  A  I  N
 406  aggatcctttttactattcactcccatcaggtgtttattgctctg
      R  I  L  F  T  I  H  S  H  Q  V  F  I  A  L
 451  attgaggatgtactagtgctgactttctacactcaaatctgcagt
      I  E  D  V  L  V  L  T  F  Y  T  Q  I  C  S
 496  tttgctaacaatcctctattatgtggtccgggaacaacaaaggct
      F  A  N  N  P  L  L  C  G  P  G  T  T  K  A
 541  tgtcctcgtgctcctccattatctccaccaccttcatttatttct
      C  P  R  A  P  P  L  S  P  P  P  P  F  I  S
 586  ccagcaccacctcgtctcaaggaagtagtgcctctagcactgga
      P  A  P  P  S  S  Q  G  S  S  A  S  S  T  G
 631  gcaattgctggtggagttgctgcaggtgctgctcttgctatttgct
      A  I  A  G  V  A  A  G  A  A  L  F  A  A
 676  gcacctgctatcggatttgcatggtggccccgtcgtaagccacaa
      A  P  A  I  G  F  A  W  W  R  R  K  P  Q
 721  gaacatttctttgatgtgcctgctgaagaggatcggaagttcat
      E  H  F  F  D  V  P  A  E  E  D  R  K  F  H
 766  ttgggacagctcaaaaggttttttctctgcgagaactacaagttgct
      L  G  Q  L  K  R  F  S  L  R  E  L  Q  V  A
 811  acagataatttttagcaacaagaacattttgggcgagaggtgggttt
      T  D  N  F  S  N  K  N  I  L  G  R  G  G  F
 856  ggaaaaggtctatataaacgaccgtcttgcagatggttcattagtggca
      G  K  V  Y  K  G  R  L  A  D  G  S  L  V  A
 901  gtaaagagaggctaaaacaagagcgcacaccaggcggcgagcttcaa
      V  K  R  L  K  E  E  F  P  G  G  E  L  Q
 946  tttcagacagaagttgagatgattagcatggctgtgcatcgaaac
      F  Q  T  E  V  E  M  I  S  M  A  V  H  R  N
 991  ttgcttcgacttcgtcggttttgcatgacacccaccgaacgtttg
      L  L  R  L  R  G  F  C  M  T  P  T  E  R  L
1036  cttgtatatccctatatggctaatggaagtgttgcatcatgctta
      L  V  Y  P  Y  M  A  N  G  S  Y  A  S  C  L
1081  agagagcggccaccatctgaacctccacttgattggacaactcgg
      R  E  R  P  P  S  E  P  P  L  D  W  T  T  R
1126  cgaaggattgcattgggatctgcaagggggctgtcatatctgcat
      R  R  I  A  L  G  S  A  R  G  L  H  Y  L  H
1171  gatcattgtgatccaaaaataattcatcgtgatgtcaaagctgca
      D  H  C  D  P  K  I  I  H  R  D  V  K  A  A
1216  aatatttttattggatgaagagtttgaggcagttgttggagacttt
      N  I  F  I  D  E  E  F  E  A  V  V  G  D  F
1261  ggcttggccaaacttatggattataaggatacccatgtaacaact
      G  L  A  K  L  M  D  Y  K  D  T  H  V  T  T
1306  gcggttcgtggaacaattggacacattgctccagagtacctatcc
      A  V  R  G  T  I  G  H  I  A  P  E  Y  L  S
1351  actggaaagtcctcagagaagactgatgtttttggatatggaatc
      T  G  K  S  S  E  K  T  D  V  F  G  Y  G  I
1396  atgcttttggaacttattactggtcagagggcattcgatcttgcc
      M  L  L  E  L  I  T  G  Q  R  A  F  D  L  A
1441  aggcttgcaaatgatgatgatgtaatgttgctggattgggtaaaa
      R  L  A  N  D  D  D  V  M  L  L  D  W  V  K
1486  ggacttctgaaagagaaaagctggacatgttggtagacccggat
      G  L  L  K  E  K  K  L  D  M  L  V  D  P  D
1531  cttccagaatgactatgtggaggctgaggtggagtcgctcatccaa
      L  Q  N  D  Y  V  E  A  E  V  E  S  L  I  Q
1576  gttgttctgctcatgtaccaggggctcctcctatggagcggccaag
      V  A  L  L  C  T  Q  G  S  P  V  E  R  P  K
1621  atgtcagaggtggtgagaatgctcgaaggagatggccttgctgag
      M  S  E  V  V  R  M  L  E  G  D  G  L  A  E
1666  agatgggaggaatggcagaaggttgaagtcgtgcgcttagatgta
      R  W  E  E  W  Q  K  V  E  V  V  R  L  D  V
1711  gagatggctccacataactgcaacaatgaatggatcatagactct
      E  M  A  P  H  N  C  N  N  E  W  I  I  D  S
1756  actgacaacttcatgcagtcgaattatctggcccagatga 1797
      T  D  N  F  M  Q  S  N  Y  L  A  Q  R  .
```

```
   1  atgaagcttaagttaaatgccactggcagccagcttgctgattgg
      M   K   L   K   L   N   A   T   G   S   Q   L   A   D   W
  46  aatcaaaatcaagtcaatccttgcacatggggctctgtttttttgt
      N   Q   N   Q   V   N   P   C   T   W   G   S   V   F   C
  91  gataatagcaacaatgtcatccaagtaacactgcctaacatggga
      D   N   S   N   N   V   I   Q   V   T   L   P   N   M   G
 136  ttcactggaaccttgtcagaaagcattggagtgctgaaccattta
      F   T   G   T   L   S   E   S   I   G   V   L   N   H   L
 181  aatgtattattgctgatcttgacccaattcagcttgcttgtactt
      N   V   L   L   L   I   L   T   Q   F   S   L   L   V   L
 226  gacatttgcatcaggtcattgccaaataacaatataactggtgga
      D   I   C   I   R   S   L   P   N   N   N   I   T   G   G
 271  ataccagaatcatttggaaatttatctagtctgacaaacttggat
      I   P   E   S   F   G   N   L   S   S   L   T   N   L   D
 316  ttggaaaataatctttttaagtggagaaataccagcttctcttggc
      L   E   N   N   L   L   S   G   E   I   P   A   S   L   G
 361  aggcttactaagctttcaattttgactctgaatcaaaataatctt
      R   L   T   K   L   S   I   L   T   L   N   Q   N   N   L
 406  accggaaatattcctgaatcccttttcaaacctttcaagcttaaac
      T   G   N   I   P   E   S   L   S   N   L   S   S   L   N
 451  gacatctttacaggtaacaatctaaattgtggcgcaaattttcca
      D   I   F   T   G   N   N   L   N   C   G   A   N   F   P
 496  caccttgtgcttccaatgtacccggccaaggaagatctggtaat
      H   P   C   A   S   N   V   P   G   Q   R   S   G   N
 541  acaaagattagaattgtgctcggatgtgttggaggagtaataggg
      T   K   I   R   I   V   L   G   C   V   G   V   I   G
 586  ttccttgttataggggttgcatttctgttctgcaggagtaggagg
      F   L   V   I   G   V   A   F   L   F   C   R   S   R   R
 631  aaaggttatcgaccagaaatttttgtagatgttgcaggtgaagac
      K   G   Y   R   P   E   I   F   V   D   V   A   G   E   D
 676  gatcgcagaattgcatttggacagttgaaaagatttacatggcga
      D   R   R   I   A   F   G   Q   L   K   R   F   T   W   R
 721  gaattgcaattggccacagagtggttcggtgagaaaaatattctt
      E   L   Q   L   A   T   E   W   F   G   E   K   N   I   L
 766  ggtcaggagggttttggaaaagtatataaaggtgtgcttccagat
      G   Q   G   G   F   G   K   V   Y   K   G   V   L   P   D
 811  aatacaaaaattgctgtaaaaaggttgactgattatgaaagtcct
      N   T   K   I   A   V   K   R   L   T   D   Y   E   S   P
 856  ggagggaagctgcattttttacgtgaagtcgagatgataagttta
      G   G   E   A   A   F   L   R   E   V   E   M   I   S   L
 901  gctgttcacagaaatctattaaggttgattggtttctgtacaaca
      A   V   H   R   N   L   L   F   C   T   T
 946  ccaacagaacggcttctagtttatcctttttatgcaaaacttgagt
      P   T   E   R   L   L   V   Y   P   F   M   Q   N   L   S
 991  gttgcctactgtttacgagactttaaacctgggaaccagtgtta
      V   A   Y   C   L   R   D   F   K   P   G   E   P   V   L
1036  gattggcctacgagaaaaaaaagtggctttaggtactgcacgggt
      D   W   P   T   R   K   K   V   A   L   G   T   A   R   G
1081  ttagagtaccttcacgaacattgtaacccgaaaattatacatcgt
      L   E   Y   L   H   E   H   C   N   P   K   I   I   H   R
1126  gatgttaaagctgctaatgtcttgcttgatgaggattttgaagca
      D   V   K   A   A   N   V   L   L   D   E   D   F   E   A
1171  gttgttggggactttggcttggcaaaattggttgatgtgaggagg
      V   V   G   D   F   G   L   A   K   L   V   D   V   R   R
1216  acgtctgtgacaactcaagttcgtgggactatgggccacatagca
      T   S   V   T   T   Q   V   R   G   T   M   G   H   I   A
1261  cctgaatatttgtcgactggaaggtcatctgagaaaactgatgtt
      P   E   Y   L   S   T   G   R   S   S   E   K   T   D   V
1306  tttggttatgggatcatgcttcttgaacttgttactggacaacgt
      F   G   Y   G   I   M   L   L   E   L   V   T   G   Q   R
1351  gcaatagacttttctcgcttagcagaagaagatgaagtactattg
      A   I   D   F   S   R   L   A   E   E   D   E   V   L   L
1396  cttgatcatgttaagaaactagagagggaaaagcgattggatgct
      L   D   H   V   K   K   L   E   R   E   K   R   L   D   A
1441  attgtcgaccgcaacctatgtatgaattataatggccaggaagtg
      I   V   D   R   N   L   C   M   N   Y   N   G   Q   E   V
1486  gaaataattattcaagttgctttgctttgcacccaagcttcacca
      E   I   I   I   Q   V   A   L   L   C   T   Q   A   S   P
1531  gaggatcgtcctacaatgtctgaagtagttcggatgcttgaaggg
      E   D   R   P   T   M   S   E   V   V   R   M   L   E   G
1576  gaggggctggctgagagatgggaagagtggcagcaaactgaagtt
      E   G   L   A   E   R   W   E   E   W   Q   Q   T   E   V
1621  tcacggagggaagagtctgagaggttgcaaatgagatttgattgg
      S   R   R   E   E   S   E   R   L   Q   M   R   F   D   W
1666  ggtgatgattctttgtataacccgaagccattgagctctctggt
      G   D   D   S   L   Y   N   P   E   A   I   E   L   S   G
1711  gcaagatga 1719
      A   R   *
```

Figure 1.Fig.1A showing ORF/CDS of PhSRK1 with corresponding deduced amino acid sequences. Start (ATG) and stop (TGA) codons are boxed. Fig.1B showing ORF/CDS OF PhSERK2 with corresponding deduced amino acid sequences. Start (ATG) and stop (TGA) codons are boxed. Fig.1C showing ORF/CDS OF PhSERK3 with corresponding deduced amino acid sequences. Start (ATG) and stop (TGA) codons are boxed.

Figure 2. Multiple sequences alignment of predicted amino acid sequences of SERK family protein kinases Accession numbers of SERKs included in this analysis are as follows: AcSERK1(*Ananas comosus*), AEC46975.1; CnSERK (*C. nucifera*), AAV58833.2; SpSERK1 (*Solanum peruvianum*), ABR18800.1; OsSERK1(*Oryza sativa*), AAU88198.1; MtSERK1 (*Medicago truncatula*), AAN64293.1; DcSERK (*Daucus carota*), AAB61708.1; GmSERK (*Glycine max*), ACJ64717.1; CsSERK (*Citrus sinensis*), ACP20180.1; CpSERK (*Carica papaya*), ABS32228.1; VvSERK (*Vitis vinifera*), XP_002270847.1; ZmSERK1 (*Z. mays*), NP_001105132.1; AtSERK1 (*Arabidopsis thaliana*), NP_177328.1; AtSERK2, AF384969_1; AtSERK3, AF384970_1; AtSERK4, NP_178999; NbSERK3A, (*Nicotiana benthamiana*), ADO86982.1; ClSERK (*Cyrtochilum loxense*), CBV98085.1 ; DlSERK (*Dimocarpus longan*), ADK70387.1 ; CuSERK1(*Citrus unshiu*), BAD32780.1; ZmSERK2 (*Zea mays*), NP_001105133.1 ; CpSERK1 (*Cyclamen persicum*), EB40066.1 ; CpSERK2 (*Cyclamen persicum*),ADO32950.1; TaSERK2 (*Triticum aestivum*), AEP14552.1 ; TaSERK3 (*Triticum aestivum*),AEP14553.1; MtSERK3 (*Medicago truncatula*), ADO15298.1 ; MtSERK2 (*Medicago truncatula*), ADO15291.1; AcSERK2 (*Ananas comosus*), AEZ52377.1. The different domain identified.

Surprisingly, PhSERK2 and PhSERK3 lack the signal peptide domain, which are also absent in the DcSERK and TaSERK2 and TaSERK3 (*Triticum aestivum)* proteins (Fig. 2, Fig. 3b) (SCHMIDT *et al.,* 1997) and the similarity between both proteins is 70%. At the protein level, PhSERK1, PhSERK2 and PhSERK3 showed highest identity with AcSERK1, CnSERK and TaSERK3 respectively (Baudino *et al.,* 2001; Ito *et al.,* 2005). It is clear from the amino acid sequence alignment that the PhSERKs align closely with SERK genes from other species (Fig. 2). More detailed analysis was done using Prosite for the identification of regions characteristic of SERK proteins (HOFMANN *et al.,* 1999). A hydrophobic amino acid signal peptide sequence with a possible cleavage site between positions 29 and 30 is present in PhSERK1, which is conserved in rice, maize, Arabidopsis and Medicago (Fig.3). This is followed by a leucine-zipper (LZ) sequence in PhSERK1, PhSERK2 and PhSERK3. The LZ domain are present in the AtSERK1 (Hecht *et al.,* 2001), ZmSERK1 (*Zea mays)* and OsSERK1 (*Oryza sativa*) (Baudinos *et al.,* 2001) and MtSERK (*Medicago truncatula*) (NOLAN *et al.,* 2003), but not in DcSERK (*Daucus carota*) (SCHMIDT *et al.,* 1997). The leucine-zipper sequence is followed by five LRRs in PhSERK2, as defined earlier by HECHT *et al.,* (2001) for AtSERK in *Arabidopsis thaliana.* PhSERK1 protein contains four LRR domains with glycosylation sites with delated LRR3 (Fig.2) (HECHT *et al.,* 2001). These repeats are thought to from a horseshoe-shaped cavity that promotes protein-protein interactions with small globular domains during rather diverse molecular recognition processes in animals and plants (KOBE and DEISENHOFER, 1995). The Pro-rich domain with SPP motif, located between the LRRs and transmembrane region, has been suggested to act like a hinge providing flexibility to the extracellular part of the receptor or as a region for interaction with the cell wall (HECHT *et al.,* 2001). The PhSERK3 protein shows the presence of characteristic four LRRs with delated LRR5, followed by the transmembrane domain and the C-terminal region but lacks the SPP motif. For PhSERK1 there was also a strongly preferred models of two TM helices from position 15 to 34 and 243 to 262 (Fig.3a) predicted by RHYTHM server. PhSERK2 and PhSERK3 contain one TM helices at positions 217 to 237 and 196 to 215 respectively. These results are confirmed by secondary structure prediction using the program Phyre2. The intracellular kinase domain contains 11 sub-domains of conserved amino acid sequences in all the three PhSERK proteins, as described by HANKS Et al (1988).

A putative protein kinase ATP binding signature region is present in the kinase domain at position 305 to 327 in PhSERK1, 280 –302 in PhSERK2 and 260–282 in PhSERK3. The 29-amino acid residue activation loop (A-loop) in subdomains VII and VIII, which was shown to be the active site of AtSERK1 (SHAH *et al.,* 2001), is present with 100% identity in PhSERK1, PhSERK2, and with six changes in the amino acid residues in PhSERK3 (Fig.2). Threonine at position 468, shown to be essential for AtSERK-mediated signaling and the target for phosphorylation, is highly conserved in all the three PhSERKs, as in other SERK proteins (HECHT *et al.,* 2001; NOLAN *et al.,* 2003). PhSERKs contain an active site signature of Ser/Thr protein kinases at positions 422 - 434, 397- 409 and 377-389 in subdomain VI in PhSERK1, PhSERK2 and PhSERK3, respectively, which are indicative of their function as a serine/threonine kinase. The 15 invariant amino acids found in almost all eukaryotic protein kinases (SHAH *et al.,* 2001) are present in PhSERK1, PhSERK2 and PhSERK3. Following the kinase region is a Leu-rich C-terminal domain as described by SCHMIDT *et al.,* (1997), conserved in PhSERK1 and PhSERK2 but not in PhSERK3. This domain may be involved in mediating the protein–protein interaction necessary for transmission of an intracellular phosphorylation cascade (SHAH *et al.,* 2001). There are six putative N-glycosylation sites, one in LZ, one in

LRR4, one in LRR5, two in SPP and one in domain V of the kinase region in almost PhSERK1 and PhSERK3, but in PhSERK2 there are only five N-glycosylation sites. For TaSERK1 and TaSERK2 (*Triticum aestivum),* there are N-glycosylation sites, two in LRR2, two in LRR4, one in LRR5 and one in domain V of the kinase region (BHUMICA *et al.,* 2008). In LRR5, the glycosylation is O-linked instead of N-linked in PhSERKs as shown in *Arabidopsis* and *Medicago* (SHAH *et al.,* 2001; NOLAN *et al.,* 2003). In extensions also, usually all prolines in the SPP repeat are hydroxylated and are considered to be targets for O-linked glycosylation (SCHMIDT *et al.,* 1997). PhSERK3 does not share one of the characteristic features of the SERK proteins, i.e. it lacks SPP motif but contains the C-terminal region which has also been suggested to be a SERK-specific feature (SCHMIDT *et al.,* 1997; BAUDINO *et al.,* 2001; ALBERTINI *et al.,* 2005) (Fig.3). Further works on the expression level of this gene should be done.
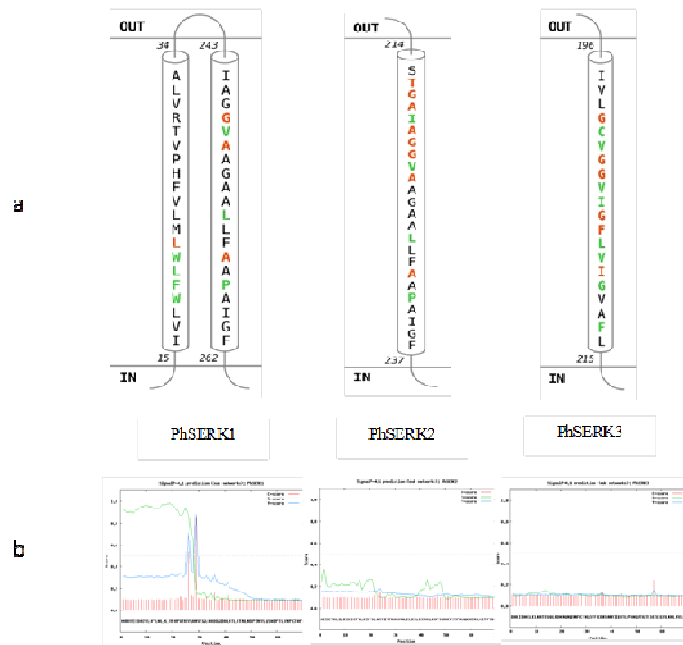
Figure 3. PhSERKs sequences analysis. a: The transmembrane region predicted by RHYTHM tool. b: Prediction of signal peptide sequence for PhSERK1, PhSERK2 and PhSERK3.

Phylogenetic relationships of the SERK proteins

To determine paralogous and orthologous relationships among the 39 SERK proteins, aligned amino acid sequences were analyzed using both distance Neighbor joining (NJ) and Maximum Likelihood method phylogenetic programs. Protein sequences of these genes were used to construct the tree. Figure 4 shows a bootstrapped Maximum Likelihood analysis, with node support estimated from 500 bootstrap replicates. This topology is also in general in agreement with traditional views regarding the phylogeny of seed plants at higher levels of

organization (CRONQUIST 1988). The phylogenetic tree of SERK protein (Fig. 4) reveals that organized into four major monophyletic groups. The SERK, SERK1 and SERK2 of the dicots are monophyletic with 30 % bootstrap support. The PhSERK3 is more closely related to TaSERK3 than the other species tested with 100% bootstrap support. The SERK 3 is joined to SERK4 of dicots except MtSERK2 at a bootstrap value of 94 %. The SERK1, SERK2 and SERK3 of the monocots are also monophyletic with 2 % bootstrap support. PhSERK2 was clustered on monophyletic group with CnSERK, AcSERK2 and AcSERK3 with 30% bootstrap. It appears that PhSERK1 is weakly united with ZmSERK1, ZmSERK2, ZmSERK3, OsSERK1, TaSERK2 and BdSERK1 with weak bootstrap support (25%).
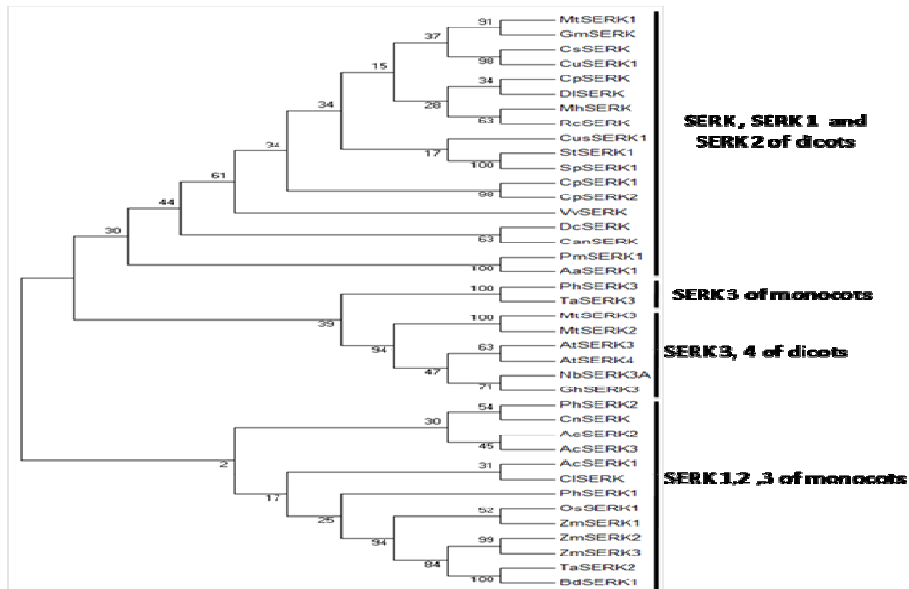
Figure 4. Molecular Phylogenetic anaylsis of SERK proteins from different plant species by Maximum Likelihood method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches.MtSERK1,2,3 (*Medicago truncatula*), GmSERK (*Glycine max*), CpSERK 1, 2 (*Cyclamen persicum*) , DlSERK (Dimocarpus longan), CsSERK *Citrus sinensis*), CuSERK1 (Citrus unshiu), VvSERK (*Vitis vinifera*), SpSERK1 (*Solanum peruvianum*), DcSERK (*Daucus carota*), AcSERK1, 2 , 3 (*Ananas comosus*), CnSERK (Cocos nucifera), OsSERK1(*Oryza sativa*), ZmSERK1,2 (*Z. mays*), SlSERK1(*Solanum lycopersicum*), TaSERK1,2,3 (Triticum aestivum) , NbSERK3A (*Nicotiana benthamiana),*AtSERK3,4 (*Arabidopsis thaliana*). StSERK1 *(Solanum tuberosum),* CanSERK *(Camellia nitidissima),* MhSERK*(Malus hupehensis),* CusSERK1*(Cucumis sativus),* BdSERK1*(Brachypodium distachyon),* PmSERK1*(Pinus massoniana),* MhSERK *(Malus hupehensis),* AaSERK1*(Araucaria angustifolia),* RcSERK *(Rosa canina),* GhSERK3*(Gossypium hirsutum),*ClSERK *(*Cyrtochilum loxense)

The PhSERK3 is closest to TaSERK3 as is evident from the phylogenetic tree while TaSERK3 was grouped with the SERK proteins which lack the SPP motif (Fig.4).

Three-dimensional structure of PhSERKs proteins

For three-dimensional structure with ligand binding sites structure of PhSERK1, 2 and 3, we used the web site that also provides a display that shows the division of family proteins into domains via 3D LigandSite program (Fig. 5), which uses an effective consensus strategy both in template selection, which combines PSI-BLAST and IMPALA and T-Coffee. The model based on template for the three SERK is proto-oncogene tyrosine-protein kinase abl1 (Nagar *et al.,* 2006).The PhSERK1 (297- 575 pb) sequence the PhSERK2 (272- 550 pb) and PhSERK3 (235- 531pb) have been modeled with 100.0% confidence by the single highest scoring template. The others residues of PhSERK sequence could be modeled at >90% confidence using multiple-templates. This region corresponds to the ATP-binding region signature, activation loop domain, and kinase domain (serine/threonine protein kinases active-site). All PhSERKs amino acids, which are involved in fixing the ligand, are shown in table 2. The results found by 3DSiteLigand program are proven using the site Prosite. The proteins PhSERK1 and PhSERK2 have the same amino acid ligand binding since both proteins have the same function in somatic embryogenesis, anther development and male gametophyte maturation (ALBRECHT *et al.,* 2005; COLCOMBET *et al.,* 2005). For protein PhSERK3, we noticed that there are differences in the amino acid ligand binding compared to PhSERK1 and PhSERK2 (Tab.2) this is normal since the function of SERK1 and SERK2 protein is the programmed cell death pathway and defense responses (CHINCHILLA *et al.,* 2007; HEESE *et al.,* 2007). The OsSERK3 and OsSERK1, showed an activity in both somatic embryogenesis and fungal defense (HU *et al.,* 2005).
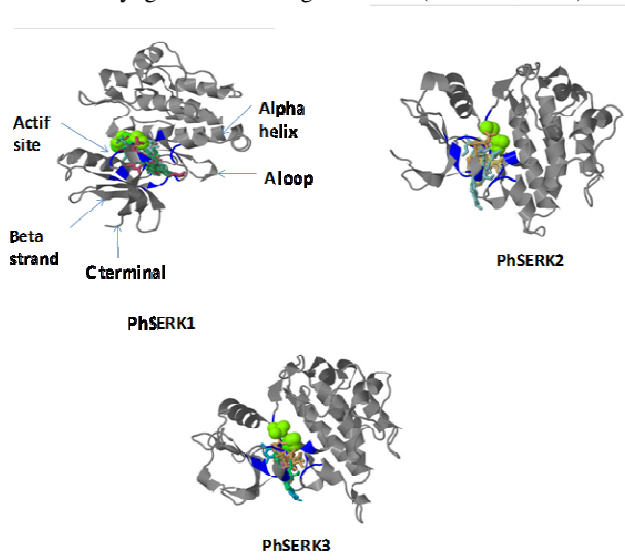

Figure 5. Three-dimensional structure with ligand binding sites of PhSERKs. The different parts of the protein are shown in the figure. PhSERKs amino acids, which are involved in fixing of the ligand, are shown in the table

In this study, we isolated PhSERK1, PhSERK2, PhSERK3 genes from an F. oxysporum induced cDNA library of palm date. Our phylogenetic and genetic evolution analysis confirmed that PhSERKs belongs to a gene family in *Phoenix dactylifera* , which leads to the conclusion that more studies must be carried out.

Tabe 2. PhSERKs amino acids which are involved in fixing of ligand

Residue	Amino acid	Residue	Amino acid	Residue	Amino acid
305	LEU	280	LEU	260	LEU
306	GLY	281	GLY	261	GLY
309	GLY	284	GLY	264	GLY
310	PHE	285	PHE	265	PHE
311	GLY	286	GLY	268	VAL
313	VAL	288	VAL	280	ALA
325	ALA	300	ALA	282	LYS
327	LYS	302	LYS	312	LEU
357	LEU	332	LEU	329	PRO
374	PRO	349	PRO	330	PHE
375	TYR	350	TYR	331	MET
376	MET	351	MET	334	LEU
379	GLY	354	GLY	335	SER
380	SER	355	SER	385	ALA
430	ALA	405	ALA	386	ASN
431	ASN	406	ASN	388	LEU
433	LEU	408	LEU	400	PHE
445	PHE	420	PHE	400	PHE

REFERENCES

ALBRECHT, C., E. RUSSINOVA, V.HECHT, E. BAAIJENS, and S. DE VRIES (2005): The Arabidopsis thaliana somatic embryogenesis receptor-like kinases1 and 2 control male sporogenesis. Plant Cell, *17***:** 3337-3349.

ALBRECHT, C., E. RUSSINOVA, B. KEMMERLING, M. KWAAITAAL, and S.C DE VRIES (2008): Arabidopsis somatic embryogenesis receptor kinase proteins serve brassinosteroid-dependent and independent signaling pathways. Plant Physiol, *148***:** 611–619.

AL-DOUS, E. K., GEORGE BINU, E. AL-MAHMOUD. MARYAM, Y. AL-JABER. MONEERA, A.JOEL MALEK (2011): De novo genome sequencing and comparative genomics of date palm (Phoenix dactylifera). Nat Biotechnol, *29*: 521–527.

ALTSCHUL, S. W.GISH, W. MILLER, E. MYERSAND, and D. LIPMAN (1997): Basic logical alignement search tool. Journal of Molecular Biology, *215*: 403-410.

BAUDINO, S., S. HANSEN, R.BRETTSCHNEIDER, V.HECHT, T.DRESSELHAUS, H. LORZ, C. DUMAS, and P. ROGOWSKY (2001): Molecular characterization of two novel maize LRR receptor-like kinases, which belong to the SERK gene family. Planta, *213*: 1–10.

BHUMICA, S., P. JITENDRA, and K. PARAMJIT (2008): Characterization of three somatic embryogenesis receptor kinase genes from wheat, Triticum aestivum. Plant Cell Rep, *27***:** 833–843.

CATHALA, J. F., B. SAVOURET, B. MENDEZ, M. KARIN, J. A. MARTIAL, and J. D. BAXTER (1983): A method for isolation of intact translationally active ribonucleic acid. DNA, *2*:329-335.

CHINCHILLA, D., C. ZIPFEL, S. ROBATZEK, B. KEMMERLING, T. NURNBERGER, J. D. G. JONES, G. FELIX, and T. BOLLER (2007): A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. Nature, *448*: 497-412.

COLCOMBET, J., D.A. BOISSON, R. ROS-PALAU, C.E VERA, and J.I SCHROEDER (2005): Arabidopsis somatic embryogenesis receptor kinases1 and 2 are essential for tapetum development and microspore maturation. Plant Cell, *17*:3350-3361.

FKI, L., MASMOUDI, R., N. DRIRA, and A. RIVAL (2003): An optimised protocol for plant regeneration from embryogenic suspension cultures of date palm, *Phoenix dactylifera* L. cv. Deglet Nour . Plant Cell Reports, *21*: 517–524.

HANKS, S. K., A.M. QUIN, and T. HIMTER (1988): The protein kinase family: Conserved features and deduced phylogeny of the catalytic domains. Science, *241*: 42-52.

HE, K., X.GOU, T.YUAN, H. LIN, T.ASAMI, S. YOSHIDA, S.D. RUSSELL, and J. LI (2007): BAK1, BKK1 regulate brassinosteroid-dependent growth and brassinosteroid independent cell-death pathways. Curr Biol, *17*:1109-1115.

HECHT, J.P., C.VIELLE, M.V. HARTOG, D.L. SCHMIDT, K.BOUTILIER, U. GROSSNICKLAUS, and S.C DE VRIES (2001): The Arabdopsis somatic embryogenesis receptor kinase 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture.Plant Physiol, *127*:803–816.

HEESE, A., D.R. HANN, S.GIMENEZ-IBANEZ, A.M.E. JONES, K.HE, J. LI, J.I. SCHROEDER, S.C PECK, and J.P RATHJEN (2007): The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. Proc Natl Acad Sci USA, *104*:12217-12222.

HOFMANN, K., P. BUCHER, L. FALQUET, and A. BAIROCH (1999): The PROSITE database: its status in 1999. Nucl Acids Res, *27***:** 215–219.

HU, H., L. XIONG, and Y. YANG (2005): Rice SERK1 gene positively regulates somatic embryoenesis of cultured cell and host defense response against fungal infection. Planta, *222***:** 107–117.

ITO, Y., K. TAKAYA, and N. KURATA (2005): Expression of SERK family receptor-like protein kinase genes in rice. Biochim Biophys Acta, *1730*: 253–258.

JUN, M., H.YEHUA, W.CHENGHOU, L. HEPING, H. ZHONGYI, and S. GUANGMING (2012): Cloning and Molecular Characterization of a SERK Gene Transcriptionally Induced during Somatic Embryogenesis in Ananas comosus cv. Shenwan. Plant Mol Biol Rep, *30*:195–203.

KARLOVA, R., S. BOEREN, E.RUSSINOVA, J.AKER, J. VERVOORT, and S. DE VRIES (2006): The Arabidopsis somatic embryogenesis receptor-like kinase1 protein complex includes brassinosteroid-insensitive1. Plant Cell, *18*: 626-638.

KEMMERLING, B. A.SCHWEDT, P.RODRIGUEZ, S.MAZZOTTA, M.FRANK, S.A. QAMAR, T. MENGISTE, S.BETSUYAKU, J.E. PARKER, C. MUSSIG (2007): The BRI1-associated kinase 1, BAK1, has a brassinolide-independent role in plant cell-death control. Curr Biol *17*: 1116-1122.

KOBE, B and J. DEISENHOFER (1995): A structural basis of the interactions between leucine-rich repeats and protein ligands. Nature, *374*:183–186.

KRIAA, W., B.SGHAIER-HAMMAMI, F. MASMOUDI-ALLOUCHE, R. BENJEMAA-MASMOUDI, and N. DRIRA (2012):The date palm (*Phoenix dactylifera* L.) micropropagation using completely mature female flowers. CR Biol, *335*: 194-2044.

LEBRUN, F. RENTIER-DELRUE, L. MERCIER (1988): A simple method for the preparation of the covalently closed circular form of plasmid DNA. Biotechniques, *6*: 3-5.

LI, J., J.WEN, K.A. LEASE, J.T. DOKE, F.E. TAX, J.C WALKER (2002): BAK1, an Arabidopsis LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. Cell, *110*: 213-222.

NAGAR, B., O.HANTSCHEL, M.SEELIGER, J.M. DAVIES, W.I. WEIS, G. SUPERTI-FURGA, J. KURIYAN (2006): Organization of the SH3-SH2 unit in active and inactive forms of the c-Abl tyrosine kinase. Mol Cell, *21*: 787–798.

NOLAN, K.E., R.R IRWANTO, and R.J ROSE (2003): Auxin up-regulates MtSERK1 expression in both Medicago truncatula root-forming and embryogenic cultures. Plant Physiol, *133***:** 218-230.

SCHMIDT, F., M.A.J. GUZZO, S.C TOONEN, and S.DE VRIES (1997): A leucine rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. Development, *124*: 2049-2062.

SHAH, K., J. VERVOORT, and S.C DE VRIES (2001): Role of threonines in the AtSERK1 activation loop in phosphorylation. J Biol Chem, *276*: 41263-41269.

SHIU, S.H and A.B BLEECKER (2003): Expansion of the Receptor-Like Kinase/Pelle Gene Family and Receptor-Like Proteins in Arabidopsis. Plant Physiol, *132*: 530-543.

SOMLEVA, M.N. E.D.L SCHMIDT, and S.C DE VRIES (2000): Embryogenic cells in Dactylis glomerata L. (Poaceae) explants identified by cell tracking and by SERK expression. Plant Cell Rep, *19*:718-726.

TOR, M. M.T LOTZE, and N. HOLTON (2009): Receptor-mediated signalling in plants: molecular patterns and programmes. J Exp Bot, *60*: 3645-3654.

ZAKIZADEH, H. B.M. STUMMANN, H. LUTKEN, and R MULLER (2010): Isolation and characterization of four somatic embryogenesis receptor-like kinase (RhSERK) genes from miniature potted rose (Rosa hybrida cv. Linda). Plant Cell Tissue Organ Cult, *101*: 331-338.

MOLEKULARNO KLONIRANJE I *in silico* ANALIZA TRI mRNK RECEPTORA KINAZE U TOKU SOMATSKE EMBRIOGENEZE KOD PALME

Imen REKIK, Amine ELLEUCH, Walid KRIAA, Noureddine DRIRA

Laboratorija za biljnu biotehnologiju, Fakultet za naukuLaboratory of plant biotechnology, Univerzitet u Sfaksu, Tunis

Izvod

Izvršeno je izolovanje i karakterizacija tri gena receptor kinaze (PhSERK) u somatskoj embriogenezi palme brzom amplifikacijom cDNK krajeva (RACE). PhSERKs pripadaju maloj familiji receptor kinaze gena koji imaju konzerviranu strukturu i ekstenzivnu homologiju sekvenci sa ranije izolovanim SERK genima. Analizom je utvrđena veličina sekvenci i to: 11051 pb (PhSERK1), 7981 pb (PhSERK2) i 10510 pb (PhSERK3). ORF kod PhSERK1, PhSERK2 i PhSERK3 su 1914 pb, 1797 pb I 1719 pb.. PhSERKs pripada LRR-tipa površine ćelije RLKs, koja ima brojne karakteristike domena. Filogenetsko stablo pokazuje da se proteini PhSERK1, PhSERK2 i PhSERK3 grupišu u klaster unutar SERKs grupe proteina.