# University of São Paulo
# "Luiz de Queiroz" College of Agriculture

## Protein-protein interaction between a candidate effector of *S. scitamineum* and a transcription factor of *A. thaliana* to study the smut-sugarcane pathosystem

### Jéssica Fernanda Mendes

Dissertation presented to obtain the degree of Master in Science. Area: Genetics and Plant Breeding

Piracicaba
2022

**Jéssica Fernanda Mendes**
**Biologist**

**Protein-protein interaction between a candidate effector of *S. scitamineum* and a transcription factor of *A. thaliana* to study the smut-sugarcane pathosystem**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Profª Drª **CLAUDIA BARROS MONTEIRO-VITORELLO**

Dissertation presented to obtain the degree of Master in Science. Area: Genetics and Plant Breeding

**Piracicaba**
**2022**

2

**ACKNOWLEDGMENTS**

First I would like to acknowledge the University of São Paulo and "Luiz de Queiroz" campus (ESALQ/USP) for the opportunity of being a Master student. I also thank the Department of Genetics and Plant Breeding for the support, especially all the professors, colleagues and staff.

I thank the funding institutions for financial support and scholarships, the São Paulo Research Foundation (FAPESP), the National Council for Scientific and Technological Development (CNPq) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES).

With much appreciation I want to thank my advisor Dra. Claudia Barros Monteiro Vitorello for all the professional, personal and emotional support. I am so grateful for all the learning, advice, friendship and patience you always give. Science and the academic career became even more attractive after joining her research group.

I would also like to thank Dr. Fabio Tebaldi Silveira Nogueira for all the collaboration, for donating biological material and for allowing the use of equipment in his lab. With much gratitude I want to thank his students, Leticia Frizzo and Flaviani Pierdonà, for all the assistance through the experiments for this work.

To all lab members of Genomics Group: Deepak Sehgal, Elaine Vidotto, Gustavo Crestana, Gustavo Lima, Hugo Rody, Joyce Ferreti, Jonathan Macedo, Juliana Calderan, Laina Oliveira, Luiz Quecine, Marcella Ferreira, Mariana Braga, Mariana Marrafon, Mauricio Jampani, Pedro Vilanova, Renato Bombardelli, Tiarla Souto, Thiago Maia and Clesivan Pereira. I am very thankful for working with people like you guys. Our work environment is very healthy and friendly. Thank you for all the help and emotional support through this journey.

I thank the board of examiners for enthusiastically accepting this invitation and contributing to this work and to my professional career.

I also thank my family, especially my parents Silvana Oliveira and Gerson Mendes, and my stepfather Donisete Oliveira. Their support and encouragement were essential for me to follow this path. I also much appreciate my fiance's support, Marcelo Vieira, for helping me with difficult and stressful moments during this period of my life. I thank all my friends Maristela, Ramana, Vitória, Danielle and Delaney for always encouraging and telling me positive stories.

4

# SUMMARY

# RESUMO

**Interação proteína-proteína entre um candidato a efetor de *S. scitamineum* e um fator de transcrição de *A. thaliana* para estudar o patossistema da cana-carvão**

A cana-de-açúcar é uma cultura de grande valor econômico. A partir da cana-de-açúcar é possível gerar etanol, bioplásticos, hidrocarbonetos e biogás. O material genético da cana-de-açúcar é extremamente complexo, poliplóide e com muitos elementos transponíveis. Embora existam variedades resistentes de cana-de-açúcar, essa cultura pode sofrer perdas econômicas significativas devido a doenças. A doença do carvão da cana-de-açúcar é causada pelo fungo biotrófico *Sporisorium scitamineum* e apresenta como principal sintoma o desenvolvimento de um chicote formado por teliósporos do fungo e tecidos da planta. Após a infecção, a cana-de-açúcar passa por uma extensa alteração transcricional. O patógeno utiliza efetores para auxiliar na penetração e colonização nos tecidos do hospedeiro. Este estudo teve como objetivo validar uma interação proteína-proteína, entre o candidato a efetor (CE) g5159 previamente descrito pelo nosso grupo de pesquisa, e um fator de transcrição de *Arabidopsis thaliana*. O primeiro capítulo consistiu em uma revisão bibliográfica sobre os conceitos principais discutidos ao longo da dissertação. O SPL13 faz parte da família dos fatores de transcrição do tipo SQUAMOSA-PROMOTER BINDING PROTEIN-like e está envolvido nas fases de transição da planta, como por exemplo para as fases vegetativa e reprodutiva. A interação entre CE g5159 e AtSPL13 foi mostrada anteriormente por meio de um ensaio Yeast-Two-Hybrid em nosso grupo, e agora usamos o ensaio de complementação de fluorescência bimolecular para validar essa interação proteína-proteína. Usamos o SPL13 de *Arabidopsis thaliana* e atualmente estamos trabalhando na amplificação por PCR de SPL13 das variedades de cana-de-açúcar RB925354, SP80-3280 e IAC66-6. A *Arabidopsis thaliana* é amplamente utilizada como planta modelo devido ao sequenciamento completo de seu genoma e facilidade de manipulação da planta. No segundo capítulo, mostramos os resultados do ensaio BiFC, indicando que ocorreu a interação CE g5159 e AtSPL13. Nossos resultados também nos levaram a especular que o CE g5159 pode ser capaz de reter a localização subcelular do AtSPL13 do núcleo. No entanto, é necessário investigações futuras. No último capítulo, inferimos relações filogenéticas entre ortólogos de SPLs de cana-de-açúcar e outras espécies. Os ortólogos do SPL13 de *A. thaliana* e da cana-de-açúcar agruparam no mesmo clado e seu alinhamento apresentou regiões conservadas entre o SPL13 de *A. thaliana* e as cultivares de cana-de-açúcar. Além disso, utilizamos dados de transcriptoma publicados para analisar os padrões de expressão de SPLs em cana-de-açúcar infectada. Assim, este é o primeiro estudo a mostrar uma interação proteína-proteína entre um efetor candidato de *S. scitamineum* e um fator de transcrição envolvido no crescimento vegetativo e reprodutivo da planta. Nossos resultados possibilitam pesquisas futuras em relação aos mecanismos moleculares usados pelo fungo para melhor compreender o patossistema e avançar em programas de melhoramento.

Palavras-chave: Interação planta-patógeno, Interação proteína-proteína, Ensaio de BiFC, Squamosa-promoter binding protein-like, SPL13, Cana-de-açúcar, Arabidopsis

**ABSTRACT**

**Protein-protein interaction between a candidate effector of *S. scitamineum* and a transcription factor of *A. thaliana* to study the smut-sugarcane pathosystem**

Sugarcane is a crop of great economic value. It is possible to generate ethanol, bioplastics, hydrocarbons and biogas from the production of sugarcane. Sugarcane's genetic material is very complex, being polyploid and with several transposable elements. Although there are resistant varieties of sugarcane, this crop can undergo significant economic losses because of diseases. The sugarcane smut disease is caused by the biotrophic fungus *Sporisorium scitamineum,* and its main symptom is the development of a whip formed by teliospores of the fungus and plant's tissues. After infection, sugarcane goes through an extensive transcriptional change. The pathogen uses effectors to assist penetration and colonization into the host's tissues. This study aimed to validate an interaction between the candidate effector g5159 previously described from our research group and a transcription factor SPL13. The first chapter consists in a bibliographic review of main concepts discussed in the dissertation. The tSPL13 is part of the SQUAMOSA-PROMOTER BINDING PROTEIN-like transcription factor family, and it is involved in the transition phases of the plant, such as to the vegetative and reproductive phases. The CE g5159 and AtSPL13 interaction was previously shown through a Yeast-Two-Hybrid assay in our group and now we used bimolecular fluorescence complementation (BiFC) assay to validade this protein-protein interaction. We used SPL13 from *Arabidopsis thaliana* and we are currently working on SPL13 PCR-amplification of RB925354, SP80-3280 and IAC66-6 varieties of sugarcane. *Arabidopsis thaliana* is extensively used as a model plant because of its complete genome sequencing and facility to manipulate the plant. In the second chapter we showed our BiFC assay results, indicating that the CE g5159 and AtSPL13 interaction occurred. We also speculate that CE g5159 may retain AtSPL13 subcellular location from the nucleus. However, future research is necessary to better comprehend the mechanism. In the last chapter we inferred phylogenetic relationships among sugarcane SPLs orthologs and other flowering species. Orthologs of SPL13 from *A. thaliana* and sugarcane clustered in the same clade and their alignment showed conserved regions between SPL13 from *A. thaliana* and sugarcane cultivars. We also used published transcriptome data to uncover the expression patterns of SPLs in infected sugarcane. This is the first study to show a protein-protein interaction between a candidate effector from *S. scitamineum* and a transcription factor involved in the vegetative and reproductive growth of the plant. Our results allow future investigation within molecular mechanisms used by the fungus to better understand the pathosystem and to advance in breeding programs.

Keywords: Plant-pathogen interaction, Protein-protein interaction, BiFC Assay, Squamosa-promoter binding protein-like, SPL13, Yeast-Two-Hybrid, Sugarcane, Arabidopsis

# PREFACE

This dissertation is a requirement for a degree in Master in Agronomy in the program of Genetics and Plant Breeding at University of São Paulo, "Luiz de Queiroz" campus, in Piracicaba, São Paulo, Brazil.

The research described here was guided by Professor Dra. Claudia Barros Monteiro-Vitorello between August 2019 and December 2021. The dissertation is divided into three chapters.

The first chapter is a bibliographic review on some subjects that will be included in second and third chapter, such as the sugarcane genome and the smut disease that affects this crop, the host manipulation by effectors of pathogens, *Arabidopsis thaliana* as a model plant to study the smut-sugarcane pathosystem, the study and techniques of protein-protein interactions and the transcription factor family of SQUAMOSA-PROMOTER BINDING protein.

The second chapter shows the results of a bimolecular fluorescence complementation (BiFC) assay performed in order to validate a protein-protein interaction between a candidate effector (CE) g5159 and a transcription factor SPL13.

The third and last chapter consists of a phylogenetic analysis of sugarcane SPLs orthologs and other flowering species. Published data of transcriptome analysis were also shown to uncover the expression patterns of sugarcane infected by *Sporisorium scitamineum*.

## CHAPTER 1: GENERAL INTRODUCTION

1. **Sugarcane**

Sugarcane is grown in over 100 countries globally, and it is the sixth crop in economic value (FAO, 2019; ISO, 2021). Regarding the production of sucrose, sugarcane is responsible for more than 80% of sugar production, and it is the second largest biofuel crop in the world (Feng et al., 2021). In addition, second-generation ethanol, biogas, biomethane, bioplastic, electricity, and hydrocarbons are other products obtained from sugarcane (Sugarcane.org, 2021). Energy production from sugarcane and its advancement is promising because it can reduce costs, cogenerate electricity, and increase income. Beyond that, the production is estimated to increase over the years (Heinrichs et al., 2017; Cursi et al., 2021).

Sugarcane is part of the Poaceae family, Saccharinae sub-tribe, and *Saccharum* genus. The Saccharum genus consists of various species, including *S. spontaneum, S. robustum, S. officinarum, S. barberi, S. sinensi,* and *S. edule,* which originated modern sugarcane varieties. However, the currently cultivated hybrid sugarcane consists mainly of *Saccharum officinarum* and *Saccharum spontaneum* genomes (D'Hont, 2005; Piperidis et al., 2010; Piperidis and D'Hont, 2020). These two subspecies' contributions in modern varieties resulted in increased sugar accumulation and disease resistance (Amalraj et al., 2005; D'Hont et al., 1998; Screenivasan et al., 1987). The subgenome *S. officinarum* is octoploid, while *S. spontaneum* shows a variation in its polyploidy level. Also, these species may differ regarding their transposable elements content and chromosomal rearrangements. The genetic material is unevenly distributed in the hybrid genome, resulting in a complex genome, having a combination of aneuploidy and polyploidy, resulting in 100 to 130 chromosomes (Thirugnanasambandam et al., 2018).

The sequencing and analysis of the sugarcane hybrid genome remain a challenge for the scientific community because of its complex nature (highly polyploidy and aneuploidy), the abundance of transposons and repetitive elements, which occupy about 50% of the entire genome (de Setta et al., 2014). As a result, sugarcane genomics uses several methodologies and techniques to decipher, for instance, how different alleles behave regarding gene expression (Souza et al., 2019; Sforça et al., 2019; Garsmeur et al., 2018; Thirugnanasambandam et al., 2018;  Vilela et al., 2017; Cardoso-Silva et al., 2014). Although there are cultivars of sugarcane resistant to certain diseases, some can still cause significant economic losses. These diseases can be caused by bacteria, such as leaf scald disease and ratoon stunting disease. Sugarcane can also undergo diseases caused by viruses, for instance,

mosaic disease and yellow leaf syndrome. Moreover, many diseases are caused by fungi, such as brown stripe disease, yellow spot disease, rust disease, and smut disease (Huang et al., 2018).

## 2. Pathosystem: Smut Disease in Sugarcane

The smut disease is caused by fungi of the Ustilaginales order, which mainly affects plants of the grass family that are economically important, generating losses in the productivity and quality of cultivars (Horst, 2013). There are 1,500 species known as smut fungi, representing one of the largest groups of phytopathogens. They are characterized by the production of black teliospores, usually in the host's floral organs, influencing the reproduction of plants infected (Zuo et al., 2019). The cause of sugarcane smut disease is the biotrophic fungus *Sporisorium scitamineum (Syd.) [Piepenbring et al. (2002) (Syn: Ustilago scitaminea H. and P. Sydow),* first found in 1877, in the Natal region of South Africa (Huang et al., 2018; Monteiro-Vitorello et al., 2018; Bhuiyan et al., 2021). The first description of sugarcane smut in Brazil was in 1946, detected in the State of São Paulo, infecting cultivars POJ21 and POJ36 and soon spreading to other sugarcane areas (Bergamin-Filho et al., 1987; Rago, Casagrande and Massola-Junior, 2009). However, after the intense smut epidemic in 1980 affecting the cultivar NA56-79, the incorporation of genetic resistance in commercial cultivars, the treatment of setts for planting, and the destruction of sick plants (roguing) kept the disease under control (Monteiro-Vitorello et al., 2018).

*Sporisorium scitamineum* penetrates the host through the sprouting buds infecting meristematic tissues. It grows systemically, forming teliospores in the peripheral tissues of the whip structure (Sundar et al., 2012; Marques et al., 2016). The pathogen has three phases in its life cycle: 1. haploid yeast-like cells; 2. dikaryotic mycelium; and 3. diploid teliospores (Fig 1). A *sine qua non*-condition for infection is the fusion of sexually compatible haploid cells, which allows the development of the dikaryotic hyphae. The fusion of compatible cells occurs in a pheromone-receptor recognition system controlled by two loci, *a* and *b* (Bakkeren, Kämper, and Schirawski 2008; Vollmeister et al. 2012). These two loci *a* and *b* in the *S. scitamineum* genome are physically linked on chromosome 2, with a distance of 59 kbp (Taniguti et al., 2015). Variations in the mating-type loci structure imply differences in sexual behavior, differentiating tetra and bipolar mating systems. Bipolar mating systems locus (when the two loci are physically linked), meiosis produces two types of mating reactions and 50% of compatible interactions between sporidia, allowing for an increased rate of self-

fertilization and, consequently, greater homozygosity among members of a population (Bakkeren and Kronstad, 1994; Raudaskoski and Kothe, 2010).

Although hosts of smut-causing fungi present different symptoms, sporogenesis is a fundamental characteristic of the pathogen`s signals. The disease's main symptom in sugarcane is the development of a dark structure in the form of a whip. The structure originates in the primary meristems of the lateral shoots and stem´s apex, consisting of fungal and plant tissues. Moreover, the whip can present different morphologies (twisted and closed), with different lengths and structures (long, short, and multiple whips) (Sundar et al., 2012). Other symptoms such as increase in tillering, poor cane formation, and the presence of galls on leaves and stems may be present as well (Fig 2) (Sundar et al., 2012; Monteiro-Vitorello et al., 2018).

Sugarcane undergoes an extensive transcriptional change as early as six hours after the pathogen infection. In resistant genotypes, the initial phase of the disease is characterized by a robust oxidative burst and alterations on the antioxidant systems to delay or impair fungal proliferation (Peters et al., 2016). Furthermore, lignin and phenolic compounds accumulated in the initial stages of infection, followed by forming a protective barrier composed of lignin, cellulose, and arabinoxylan (Marques et al., 2018). On the other hand, in susceptible plants during whip development, the changes occur in carbon partitioning, tissue lignification via a bifunctional PTAL (phenylalanine/tyrosine ammonia-lyase) overexpression, breakdown of starch, and hormonal changes, in which auxin is the most influenced (Schaker et al., 2016). Also, whip development and the transition from the vegetative to reproductive stages of the plant development share some of the molecular mechanisms involving the MADS-box family of transcriptional factors and flowering inducers, such as APETALA1 (AP1) and 3 (AP3), COL6, AGAMOUS, VIN3, Target of Eat 2 (TOE2) and Flowering Locus (FT) (Schaker et al., 2016).

In addition, a genomic approach combined with the transcriptomic data produced for smut-resistant and -susceptible inoculated with *S. scitamineum* was used to define the set of resistance analog genes (RGAs) expressed in smut-resistant plants (Rody et al., 2019). The data revealed, for instance, that chromosomes 2 and 5 of the ancient genotype (*S. spontaneum*) grouped most of the RGAs responsive to smut, which were not present in the modern genotypes of sugarcane. Later, Rody et al. (2021) used this data to build a sugarcane biological network based on *A. thaliana*. They defined molecular events varying in strongly

connected modules that reached the primary metabolism connecting the sugarcane immune system.

The complete chromosome-scale genome of a Brazilian strain of *S. scitamineum* was sequenced (Taniguti et al., 2015), unveiling 438 specific genes that may be associated with adaptation to their hosts. Regarding the fungal mechanisms to provoke disease, the genome sequence suggested that the pathogen can overcome plant defense using detoxification of defense molecules (proteases inhibitors, pisatin demethylase, benzoate 4-monooxygenase, chorismate mutase, superoxide dismutase, catalase) and to grow and colonize plant tissues using an arsenal of nutrient uptake-related proteins and cell wall degrading enzymes. Many enzymes that degrade cell walls are encoded by the *S. scitamineum* genome, such as enzymes that degrade lignin, pectin, hemicellulose, cellulose, and cuticles. In addition, approximately 13% of the genes were expressed only *in planta* compared to *in vitro* growth. The function of these genes indicated a possible association with the detoxification of plant defense molecules (Taniguti et al., 2015). Also, *S. scitamineum* activates a mechanism to resist peroxides and other reactive oxygen species produced by the host when infecting resistant sugarcane plants (Peters et al., 2020). In total, *S. scitamineum* predicted secretome comprised 305 proteins, including 54 carbohydrate-active enzymes (CAZymes) and 70 effector candidates (Taniguti et al., 2015).

Although smut fungi of different species and genera share 85% of gene content, they have specific genes related to host-specialization. Benevenuto et al. (2018) analyzed the genome of different species of Ustilaginaceae smut fungi and found effectors, transposases, and reverse transcriptases to be often species-specific. The EuKaryotic Orthologous Groups (KOG) (Tatusov et al., 2003) detected terms related to recombination, replication and repair overrepresented in *U. hordei* and *U. esculenta*, while *Melanopsichum pennsylvanicum* was overrepresented by RNA processing and modification terms. In the *S. scitamineum*, chromatin structure and dynamics were overrepresented. For effectors, smut fungi have a set of conserved genes and some specific to each species. For instance, the leaf-specific effector candidates (um06223 and um12217) were found only in *Ustilago maydis*.
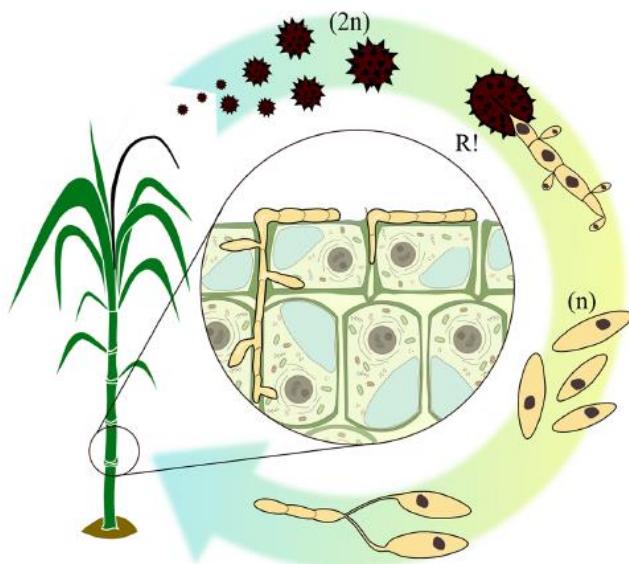
**Figure 1.** Development stages in the life cycle of *Sporisorium scitamineum.* (2n) diploid teliospores, (n) haploid yeast cells after meiosis (R!); fusion of the hypha (Taniguti et al., 2015).
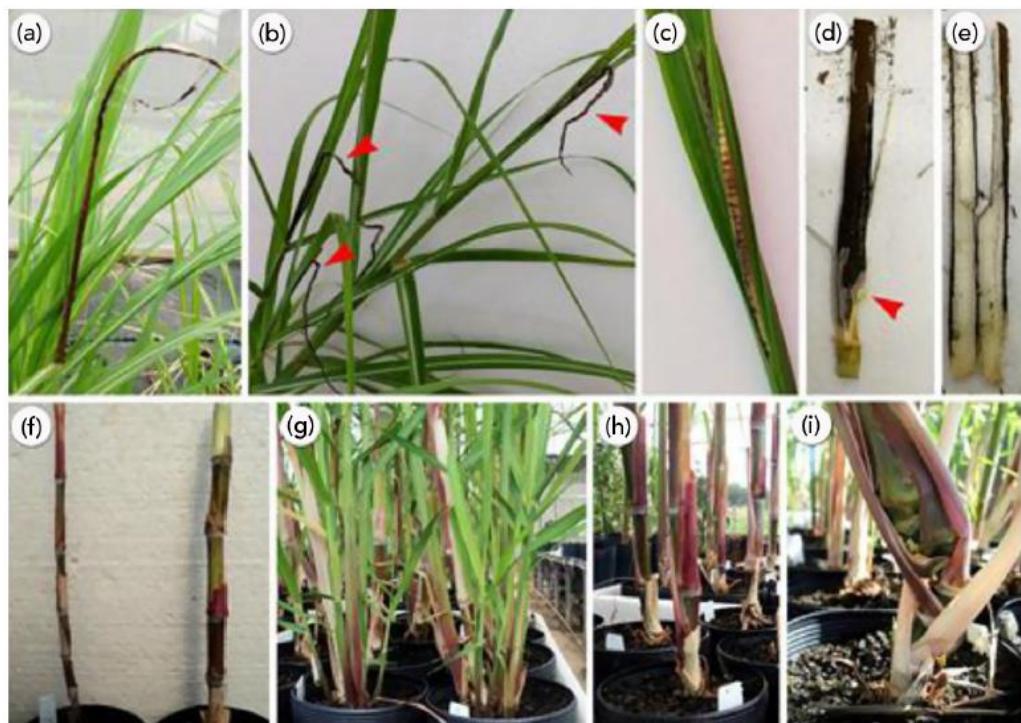


**Figure 2.** Symptoms of smut disease in sugarcane. (a) Whip structure, derived from the apical meristem of cane; (b) three whips emerging from the side reeds of the cane (red arrows); (c) tumor-like vesicle developed over a central leaf vein; (d) basal increase of a whip (red arrow); (e) longitudinal section of the whip shown in (d); (f) inoculated plants (left) and control (right) showing differences in stem diameter after the emission of whips (120 DAI); (g) tillering of infected plants; h) healthy single stem plants; (i) gall formation at the base of the sugarcane stem (Monteiro-Vitorello et al., 2018).

### 3. Host manipulation by effectors

Host manipulation can be achieved by several mechanisms used by biotrophic fungus, resulting in a complex scenario characterized by many molecules, signaling components, and molecular networks. Recently, one of the components used by the sugarcane smut to defeat host defenses, was identified as a small secreted protein translocated from the fungus to the plant cell, the first functional study of a *S. scitamineum* effector protein (Jaswal et al., 2020; Ling et al., 2021). In smuts, effectors can contribute to the various steps of plant infection and colonization stages, because of their biotrophic behavior, including epidermal penetration and intracellular growth (Depotter et al., 2021). The genes encoding these effector proteins are constantly under selection pressure from the plant's host immune system resulting in the classical coevolution model of 'arms race' described by Jones et al., 2006 (Zuo et al., 2019; Selin et al., 2016). Recently, not only did the interacting pair effector-R protein be predicted to undergo fast evolution but also other interacting proteins of smut pathogens and their hosts functioning to determine compatible or incompatible reactions (Beckerson et al., 2019).

Distinct effectors can use unique mechanisms in different cell localization of the plant, and time during disease progression to overcome the host's defense barriers (Jaswal et al. 2020; Toruño et al. 2016; van der Linde and Göhre, 2021). For instance, the studies of effectors explained functions such as enhancing pathogen virulence, cell death hijacking, inhibiting RNAi silencing pathways, and suppressing defense-related genes (Jaswal et al., 2020; Tabassum et al., 2021). Assays using a heterologous system and studies establishing effectors overexpression resulted in HR (Hypersensitive Response)-induced cell death, ROS (Reactive Oxygen Species) accumulation, suppression of PCD (Program Cell Death) pathway, and leakage of essential electrolytes (Ramachandran et al., 2016).

The definition and knowledge about effectors are constantly in progress. However, effectors can generally be classified into two classes depending on their action site: cytoplasmic effectors and apoplastic effectors (Sperschneider et al., 2018). Apoplastic effectors can protect the pathogen from hydrolytic enzymes of the host or be perceived by membrane receptors to induce the immune response. They are frequently smaller and have a higher cysteine content than cytoplasmic effectors (Jaswal et al., 2020). Cytoplasmic effectors are secreted and translocated to the host cell cytoplasm, often modulating a signal transduction cascade, leading to the expression of genes associated with immunity (Kim et al., 2020). The plant immune system uses then two distinct strategies to perceive pathogen attacks: 1) the PAMP-Triggered Immunity (PTI) with the involvement of apoplastic effectors,

and 2) the Effector-Triggered Immunity (ETI), where effectors are translocated inside the plant cell.

Generally, there are some criteria to predict a protein as an effector. An important one is the presence of a signal peptide at the N-terminal of the protein. All effectors must be secreted. In addition to this feature, others are considered for *in silico* identification, such as the absence of transmembrane domain, small size, high mRNA expression during infection cycle, high cysteine content, no conserved PFAM domain (Protein Families Database), repeat motifs, species or family specificity, and high scores in the predictors effector P and apoplast P (Sperschneider et al., 2018; Jaswal et al., 2020). An accepted checklist of features defined by Saunders et al. (2012) provided to predict fungal effectors: i) presence of a secretion signal; ii) genes that have their expression induced *in planta*; iii) proteins sharing some similarity with secreted proteins expressed in haustorium (HESP) or with avirulence genes (Avr); iv) genes encoding for proteins harboring known effector motifs or a nuclear localization signal (NLS); v) small cysteine-rich proteins (≤135aa, ≥5%); vi) proteins with internal repeats; vii) encoded by genes flanking long intergenic regions; viii) proteins without a PFAM domain.

For *S. scitamineum*, combining genomic approaches and the use of model plants, Teixeira-Silva et al. (2020) and Maia et al. (in press) investigated the ability of candidate effectors (CEs) to suppress sugarcane immune responses and determined the subcellular location of these CEs. The authors described that four out of six CEs could induce MTI or ETI and defined their location as nuclear, cytoplasmic, or cell wall-associated.

In this study, we analyzed one of the effectors described by the aforementioned authors. The CE g5159 is a suppressor of MTI when delivered into *Nicotiana benthamiana* cells via the type-three secretion system of *Pseudomonas fluorescens* EtHAn. The effector was cloned into the pEDV6 vector and expressed by *P. fluorescens* EtHAn for translocation into the plant cells and analyzed in different co-infiltration experiments with *Pseudomonas syringae* pv. tomato DC3000. Also, CE g5159 is a suppressor of ETI based on AvrB-induced experiments (Maia et al., in press). The subcellular localization of a protein fusion g5159 + GFP using agro-transformation, transient expression, and *N. benthamiana* leaves revealed an attachment to the plant cell wall. In the same experiment, a pull-down assay followed by mass spectrometry for protein identification, revealed that g5159 had a few potential interactors in *N. benthamiana*, in which the most relevant was an endochitinase (Teixeira-Silva et al., 2020).

**4.** ***Arabidopsis thaliana*** **as a model plant to study the smut-sugarcane pathosystem**

Studies with various pathogens of valued agronomic crops use model plants, mainly *Arabidopsis thaliana*, to reveal host factors affecting the outcome of diseases. *A. thaliana* is widely used as an experimental model because of their accessible genetic manipulation to obtain mutants, and the expression of genes of interest. Besides, they have a short life-cycle, present an extensive database of mutants and germplasm, and the complete genome sequence is available (Meinke, 1998; Andargie; Li, 2016; Arabidopsis Genome Initiative, 2000). Various authors used *Arabidopsis* as a model plant to study pathosystems involving fungi, for instance, *Cladosporium fulvum, Magnaporthe oryzae, Blumeria gramininis, and Sclerotinia sclerotiorum* (van Esse et al., 2008; Schmidt et al., 2014, Park et al., 2009; Barbacci et al., 2020; Coleman et al., 2020; Quin et al., 2020).

 Considering smuts, using model plants can overcome the experimental limitations of pathosystems such as smut-sugarcane, due to the genetic complexity of their hosts (León-Ramírez et al., 2004). The most common symptoms described in alternative hosts for smut were the growth of fungal mycelium on stems and leaves, increased root number in monocots or development of adventitious roots in dicots, chlorosis, increased anthocyanins, necrosis, and stunting. For *Ustilago maydis*, the fungus inoculated in *Arabidopsis* seedlings produced a white mycelium on the leaf surface and the invasion of tissues, penetration through stomata, increased anthocyanin formation, development of chlorosis, increased formation of secondary roots, induction of malformations in the leaves and petioles, induction of tissue necrosis, and stunting (Méndez-Morán et al., 2005; Martínez-Soto et al.; 2013). However, the fungus did not produce teliospores. Transcriptomic data produced by Martínez-Soto et al. (2013) showed that the genetic machinery used in *Arabidopsis*-smut infection is potentially similar to that used in maize infection.

The only smut fungi (*Thecaphora thlaspeos)* adapted to Brassicaceae hosts had its genome sequenced and the transcriptomic data made available, revealing that the fungus balances its virulence during biotrophic growth, allowing long-lived and systemic infection of *A. thaliana* (Frantzeskakis et al., 2017*).* Like many other smuts, *T. thlaspeos* produced spores in siliques, replacing seeds in their natural host *Arabis hirsuta*. However, the authors did not observe sporulation in siliques of *A. thaliana.* The genome of *T. thlaspeos* shares various characteristics with other smut genomes, including genome size and the coding of core candidate effectors (Courville et al., 2019). For three candidates, the authors functionally demonstrated effector activity. However, the molecular mechanisms underlying fungal sporulation remain elusive in hosts- and non-hosts plants.

Furthermore, sequences and functional information from *Arabidopsis* suggested to be adequate to model a sugarcane biological network to predict links between the primary metabolism and perception of the smut pathogen (Rody et al., 2021). In our group, an ongoing study using *Arabidopsis* as a host for *S. scitamineum* showed that the fungus forms appressorium, besides using stomata also to penetrate plant tissues (Silva, 2019).

## 5. Studies of protein-protein interaction in pathosystems

The plant-pathogen interplay, the battle-front of the initial contact and pathogen penetration of plant tissues, involves interaction among various proteins. These interactions support several cellular functions in pathogenicity, impacting plant growth and development. The multiple methods established to analyzed protein-protein interactions have different technical characteristics. For instance, they are used *in vitro* or *in vivo* or involve enzymatic or fluorescent markers (Miura et al., 2018; Xing et al., 2016). Some examples are Yeast-Two-Hybrid Assay (Y2H) (Fields and Song, 1989), Pull-Down Assay (Gerace et al., 2015), Forster Resonance Energy Transfer (FRET) (Sun et al., 2013), and Bimolecular Fluorescence Complementation Assay (BiFC) (Ghosh et al., 2000; Kerppola, 2008). Some methods are relatively inexpensive and allows high-throughput assays, such as the Y2H system, with the activity of B-galactosidase (Fields and Song, 1989). Some others are time-consuming, experimentally tricky, and more expensive, for instance, the FRET. Yet, some methods might require higher training, it can include complex data processing, and it might miss plant-specific post-translational protein modification or infer incorrect subcellular localization (Walter et al., 2004). Therefore, each method has pros and cons regarding its use (Miura, 2018).

Our group aimed to identify host interactors of candidate effectors (CE) of *S. scitamineum* using a normalized prey cDNA expression library obtained from *Arabidopsis*, a yeast two-hybrid cDNA library from *Arabidopsis* (Matiolli and Melloto, 2018) to detect *Arabidopsis* proteins interacting with these effectors. This library employed cDNA made from *Arabidopsis* leaves exposed to several stressors and untreated leaves, which increased transcripts representation. The library was built from a pool of mRNA extracted from plants challenged with plant and human bacterial pathogens, the flg22 elicitor, the phytotoxin coronatine, and several hormones associated with environmental stress responses. One of the results showed that a transcription factor called *Squamosa-promoter binding protein 13* (SPL13), associated with flowering pathways, interacts with a CE g5159 (unpublished data),

previously detected by Teixeira-Silva et al. (2020). In order to validate the result and select sugarcane candidate sequences orthologous to SPL13, we conducted this study. We chose BiFC according to Silva et al. (2018) to validate the result obtained with the Y2H system.

The BiFC technique was first described using *E. coli* (Ghosh et al., 2000). The authors showed that the green fluorescent protein (GFP) could be divided into the C and N terminal fragments. The individual parts of the protein were not fluorescent when fragmented. However, if the fragments interacted, they could fold together and restore fluorescence (Fig 3). In plants, Walter et al. (2004) described several BiFC vectors compatible with plants using a genetic mutant of GFP, the YFP (yellow fluorescent protein) marker (Ohad et al., 2007).

The study showed that BiFC could be a reliable and valuable tool for studying protein-protein interactions with some advantages, such as 1) the possibility of visualizing interactions in the plant cell; 2) the observed signals can be strong, concluding that BIFC is sensitive and that it can detect interactions under low protein expression; 3) the fluorescence signals generated in the BIFC can be quantified; 4) it is possible to detect the subcellular region where the interaction occurs, and 5) BIFC can investigate the structural basis for the formation of the protein complex.

Nonetheless, BiFC also has some significant drawbacks. The most relevant is that the fluorescent protein halves are prone to self-assembly independent of a protein-protein interaction event, the so-called self-assembly (Horstman et al., 2014; Kudla and Bock, 2016). Therefore, to avoid experimental problems leading to false conclusions, it is advisable to include internal controls. Besides the standard negative controls, it is appropriate to use mutated versions of the proteins under study that cannot bind to their interacting partner. Also, although a single positive combination can provide evidence of protein-protein interaction, the experiment should include all the possible combinations of the N- and C-terminal of the protein of interest and the halves of the fluorescent protein. This precaution is necessary because the interaction ability may be lost to spatial restrictions of the protein complex's three-dimensional structure, which can impair the reconstruction of the FP (Horstman et al., 2014).
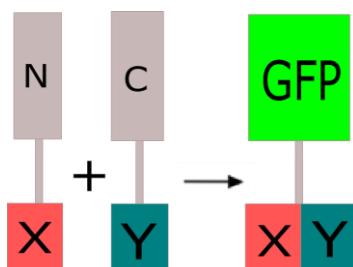
**Figure 3.** Scheme of interaction of proteins X and Y by the BIFC assay. N and C are the respective fragments of the N- and C-terminal of the fluorescence protein

Various reports used BiFC to confirm the interaction between an effector and a plant host protein for pathosystems involving fungi and bacteria. One of the results for a bacterial pathogen unveiled that an effector recognition is converted to defense activation, using a series of BiFC experiments performed in *N. benthamiana* leaves. The authors proved physical interaction among receptors, defense regulators, and effectors (Un Huh et al., 2017). For the sugarcane-smut pathosystem, the most recent study revealed the functional role of the first described effector of *S. scitamineum*. The effector SsPele1 showed a strong co-expression with sugarcane Plant Elicitor Peptide Receptor1 (ScPEPR1), which encodes a receptor-like kinase for the perception of plant elicitor peptide1 (ScPep1). The authors confirmed physical interaction between SsPele1 and ScPEPR1 using the BiFC strategy (Ling et al., 2021).

## 6. Squamosa-promoter binding proteins (SPL) transcription factor family

The SQUAMOSA-PROMOTER BINDING PROTEINS (SPL) constitute a family of plant-specific transcription factors that control several characteristics in plant growth, such as vegetative phase change, flowering time, leaf, branching and initiation rate, gibberellin biosynthesis and signaling, anthocyanin biosynthesis, plant response to stress, and somatic embryogenesis (Preston and Hileman, 2013; Li et al., 2019). *SPL* genes were first identified in *Antirrhinum majus* as regulators of early flower development in the floral meristem (Klein et al., 1996). They share a conserved SBP domain, with 75 to 79 amino acid residues with two $Zn2+$-binding sites: Cys-Cys-His-Cys and a putative conserved nuclear localization signal (NLS) (Birkenbihl et al., 2005; Preston and Hileman, 2013) (Figure 4). SPLs bind specifically to the *cis*-element TNCGTACAA in the promoter region of their target genes (Cardon et al., 1997; Cardon et al., 1999; Birkenbihl et al., 2005).
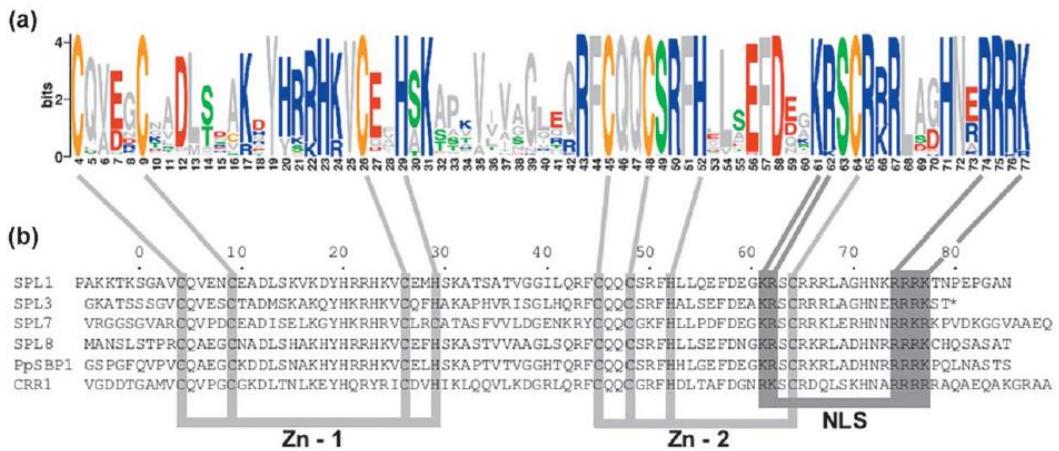
**Figure 4.** Sequence conservation in the SBP-domain described in Birkenbihl et al. (2005). (a) SPL domain showing sequence conservation as a stack of letters using the program WebLogo (WebLogo.berkeley.edu). Cysteine residues are in yellow, basic amino acid residues involved in DNA binding and nuclear localization are in blue, acidic amino acid residues are in red, and threonine and serine are in green. (b) Amino acid sequence alignment of SBP-domains used in DNA-binding: SPL1, SPL3, SPL7, SPL8 from *Arabidopsis*, PpSBP1 from Physcomitrella and CRR1 from Chlamydomonas. The conserved basic amino acids of the NLS are shaded dark grey, while the two Zn-coordinating structures (Zn-1 and Zn-2) with the involved cysteine and histidine residues are in light grey.

Some authors sorted *SPL* genes according to the sequence, size, and gene structure in short and long groups (Birkenbihl et al., 2005). In Arabidopsis, the groups also vary in expression pattern, the largest genes (*SPL1, SPL7, SPL12, SPL14* and *SPL16)* are expressed constitutively, while the mid-sized and small genes are upregulated mainly in flower development. Only the smaller genes are targeted by the microRNAs miR156 and miR157 (exception *SPL8*), which regulate their expression (Schwab et al., 2005). The number of introns within the *SPL* family ranges from 1-10, but phylogenetic analysis showed conservation of exon-intron structures between genes grouped (Preston and Hileman, 2013).

The genomes of plants, so far studied, have a variable number of genes encoding a different family SPL member. For instance, Arabidopsis has 16 (Cardon et al., 1999), maize has 31 (Mao et al., 2016) and wheat has 56 genes (Zhu et al., 2020). Recently, two groups independently defined a set of *SPL* genes in the genome of *S. spontaneum* AP85-441 without taking alleles into account (Zhang et al., 2018). Feng et al., (2021) described 17 genes, whereas Liu et al., (2021) described 30 *SPLs*. The results were different because of the strategies used. Feng et al., (2021) used the set of 19 genes described for sorghum as queries to search the *S. spontaneum* genome. On the other hand, Liu et al., (2021) used the HMM (Hidden Markov Model) profile of the SBP domain as a query to search the genome sequence. Both used the chromosome assembly to infer the correct position of each gene. However, names and classification varied between the two analyses. Feng and collaborators

clustered *SPLs* into six subfamilies (clades I-VI) using 72 *SPLs* from four species, including 17 from Arabidopsis, 19 from rice, 19 from sorghum and 17 from *S. spontaneum*. All the SPLs identified by Feng et al., (2021) have similar features as those described by Birkenbihl et al., (2005) in *Arabidopsis,* including the signal for nuclear localization (NBL) and were highly similar to the sorghum proteins (Figure 5). The sequences varied from 199 to 967 amino acids, according to the long and short groups described above.
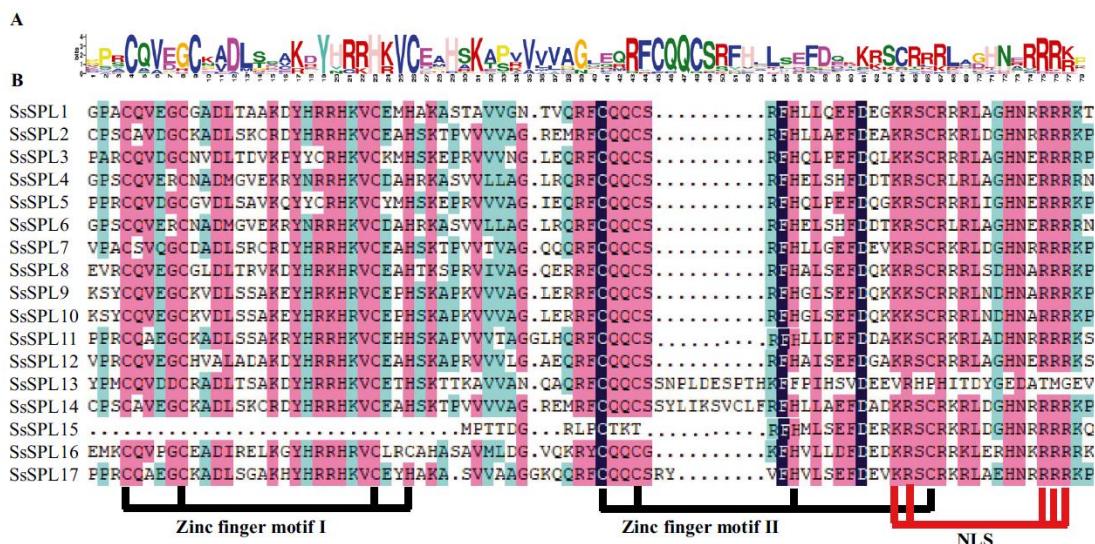


**Figure 5.** Sequence conservation in the SBP-domain described in Feng et al. (2021). (a) SBP domain showing sequence conservation as a stack of letters as logo. Cysteine residues are in blue, basic amino acid residues involved in DNA binding and nuclear localization are in red, acidic amino acid residues are in purple, and serine are in green. (b) Multiple alignment performed by DNAMAN showing protein sequence similarities among SPLs of *S. spontaneum*. Two zinc finger binding sites and nuclear localization signals (NLSs) are indicated. The conserved basic amino acids of the NLS are pink, while the two Zn-coordinating structures (Zn-1 and Zn-2) with the involved cysteine and histidine are in light blue.

Different from Feng et al., (2021), Liu and collaborators (2021), described two out of 30 SPLs as membrane associated (SPL4 and SPL12) (Figure 6). The sequences varied in size from 189 to 1064 amino acids. The authors defined 8 clusters in a phylogenetic analysis using *Arabidopsis*, *Vitis vinifera*, *Ananas comosus*, *Sorghum bicolor* and *Oryza sativa* SPL sequences. Interestingly, both groups (Liu et al., 2021; Feng et al., 2021) analyzed the expression profiles of SBP genes using RNA-seq data of various tissues and development stages revealing a variant pattern of expression for *S. spontaneum* SPLs. They also analyzed the mi156 target and expression pattern of this set of SPLs. Considering the naming of SPLs there is no correlation between the work of the two groups.
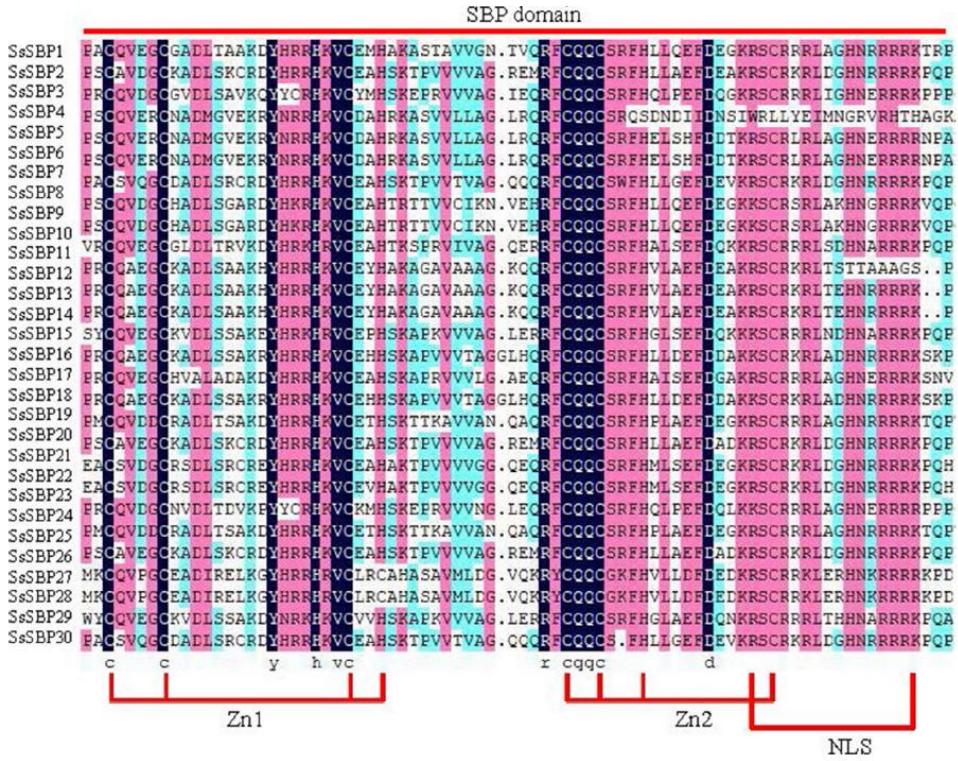
**Figure 6.** Sequence conservation in the SBP-domain described in Liu et al. (2021). Multiple alignment showing protein sequence similarities among SPLs of *S. spontaneum*. Two zinc finger binding sites and nuclear localization signals (NLSs) are indicated. The conserved basic amino acids of the NLS are pink, while the two Zn-coordinating structures (Zn-1 and Zn-2) with the involved cysteine and histidine are in dark blue.

Functional analyses of different *SPL* genes of various plant species have uncovered their roles in plant growth and development, including leaf development, plant architecture, vegetative-to-reproductive phase change and flowering, inflorescence branching, organ size, fruit development, grain yield and quality (reviewed in Cao et al., 2019). However, in sugarcane, besides the general description of SPLs described above, there is little knowledge about the SPL family.

## 7. Objectives

Our hypothesis considers that the candidate effector g5159 of *S. scitamineum* physically interacts with a protein of the SPL family to impair the progression of the transition from vegetative to reproductive stage in resistant genotypes. The first step to accomplish that is the primary objective of this study: to validate a previously identified interaction in a Y2H experiment of SPL13 protein of the model plant *Arabidopsis thaliana* with the candidate effector 5159. We also aimed to perform *in silico* studies to characterize the SPL13 in sugarcane and find orthologs of SPLs for future validation in BiFC experiments and functional studies.

**7.1 Specific Objectives**

- Select, amplify and clone in *Escherichia coli* the transcription factor *SPL13* gene of *Arabidopsis thaliana* and the candidate effector g5159 of *Sporisorium scitamineum*;

- Obtain the BiFC vectors with AtSPL13 and CE g5159 and appropriated controls;

- Perform the BiFC assay to detect fluorescence emission in *Nicotiana benthamiana* leaves using transient expression of genes cloned in *Agrobacterium tumefaciens*;

- Infer phylogenetic relationships among sugarcane SPLs orthologs and other flowering species in search of the sugarcane SPL13 ortholog;

- Uncover the expression patterns of smut-inoculated sugarcane SPLs based on previously obtained RNAseq data.

**References**

Ali, M., Alia, K., Atif, R., et al. Genome-Wide Identification and Comparative Analysis of SQUAMOSA PROMOTER BINDING PROTEINS (SBP) Transcription Factor Family in *Gossypium raimondii* and *Arabidopsis thaliana*. **Pak. J. Bot** 49(3), p. 1113-1126, 2017.

Andargie, M., Li, Jianxiong. *Arabidopsis thaliana*: A Model Host Plant to Study Plant–Pathogen Interaction Using Rice False Smut Isolates of *Ustilaginoidea virens*. **Frontiers In Plant Science**, v. 7, p.1-10, 23 fev. 2016.

Amalraj, V., Alfonse, B.N. On the Taxonomy of the Members of 'Saccharum Complex'. **Genetic Resources And Crop Evolution**, v. 53, n. 1, p.35-41, 10 ago. 2005.

Arabidopsis Genome Initiative. "Analysis of the Genome Sequence of the Flowering Plant *Arabidopsis Thaliana*." **Nature** 408: 796. 2000.

Bakkeren, G., Kämper, J., Schirawski, J. Sex in smut fungi: structure function and evolution of mating-type complexes. **Fungal Genet Biol** 45: S15-S21. 2008. 10.1016/j.fgb.2008.04.005

Bakkeren, G., Kronstad, J.W. Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. **Proc. Natl. Acad. Sci. USA** 91, p. 7085–7089. 1994.

Barbacci, A., Navaud, O., Mbengue, M., Barascud, M., Godiard, L., Khafif, M., Lacaze, A., Raffaele, S. Rapid identification of an Arabidopsis NLR gene as a candidate conferring susceptibility to *Sclerotinia sclerotiorum* using time-resolved automated phenotyping. **The Plant Journal**, v. 103, n. 2, p. 903-917, 21 abr. 2020. http://dx.doi.org/10.1111/tpj.14747.

Beckerson, WC., Rodríguez de la Vega, RC., Hartmann, FE., Duhamel, M., Giraud, T., Perlin, MH. Cause and Effectors: Whole-Genome Comparisons Reveal Shared but Rapidly Evolving Effector Sets among Host-Specific Plant-Castrating Fungi. **mBio**. 5, 10(6):e02391-19. 2019. doi: 10.1128/mBio.02391-19. PMID: 31690676; PMCID: PMC6831777.

Benevenuto, J., Teixeira-Silva, N., Kuramae, E., Croll, D., Monteiro-Vitorello, C. Comparative Genomics of Smut Pathogens: insights from orphans and positively selected genes into host specialization. **Frontiers In Microbiology**, v. 9, p. 1-17, 6 abr. 2018. http://dx.doi.org/10.3389/fmicb.2018.00660.

Bergamin-Filho, A., Amorim, L., Cardoso, C.O.N., Sanguino, A., Irvine, J.E., Silva, W.M. Carvão da cana-de-açúcar e sua epidemiologia. São Paulo: Copersucar, 1987. 23p. (Boletim Técnico – Edição especial)

Bhuiyan, S., Magarey, R., Mcneil, M., Aitken, K. Sugarcane Smut, Caused by *Sporisorium scitamineum*, a Major Disease of Sugarcane: a contemporary review. **Phytopathology**, p. 1-13, 19 nov. 2021. http://dx.doi.org/10.1094/phyto-05-21-0221-rvw.

Birkenbihl, R.P., Jach, G., Saedler, H., Huijser, P. Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. **J Mol Biol.** 352(3), p. 585–596. 2005.

Cardoso-Silva, C.B., Costa, E.A., Mancini, M.C., Balsalobre, T.W.A., Canesin, L.E.C., Pinto, L.R., Carneiro, M.S., Garcia, A.A.F., de Souza, A.P., Vicentini, R. De novo assembly and transcriptome analysis of contrasting sugarcane varieties. **PloS One,** *9*(2), p. 1-10. 2014.

Cardon, G., Hohmann, S., Klein, J., Nettesheim, K., Saedler, H., Huijser, P. Molecular characterisation of the Arabidopsis SBP-box genes. **Gene** 237, p. 91–104. 1999.

Cardon GH, Höhmann S, Nettesheim K, Saedler H, Huijser P. Functional analysis of the *Arabidopsis thaliana* SBP-box gene SPL3: a novel gene involved in the floral transition. **Plant J** 12, p. 367–377. 1997.

Cao R, Guo L, Ma M, Zhang W, Liu X and Zhao H. Identification and Functional Characterization of Squamosa Promoter Binding Protein-Like Gene TaSPL16 in Wheat (*Triticum aestivum L.*). **Front. Plant Sci.** 10:212. 2019. doi: 10.3389/fpls.2019.00212

Coleman, A., Maroschek, J., Raasch, L., Takken, F.L.W., Ranf, S., Huckelhoven, R. The Arabidopsis leucine-rich repeat receptor-like kinase MIK2 is a crucial component of early immune responses to a fungal-derived elicitor. **New Phytologist**, v. 229, n. 6, p. 3453-3466, 30 dez. 2020. http://dx.doi.org/10.1111/nph.17122.

Courville, K., Frantzeskakis, L., Gul, S., Haeger, N., et al. Smut infection of perennial hosts: the genome and the transcriptome of the Brassicaceae smut fungus Thecaphorathlaspeos reveal functionally conserved and novel effectors. **New Phytologist** 222. p. 1474-1492. 2019. doi: 10.1111/nph.15692.

Depotter, J.R.L., Zuo, W., Hansen, M., Zhang, B., Xu, M., Doehlemann, G. Effectors with Different Gears: Divergence of Ustilago maydis Effector Genes Is Associated with Their Temporal Expression Pattern during Plant Infection. **J. Fungi** 7, 16. 2021. https://doi.org/10.3390/jof7010016

D'Hont, A., Ison, D., Alix, K., Roux, C., Glaszmann, J.C. Determination of basic chromosome numbers in the genus Saccharum by physical mapping of ribosomal RNA genes. **Genome**, v. 41, n. 2, p.221-225, abr. 1998.

D'Hont, A. Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. **Cytogenetic and Genome Research**, 109(1-3), p. 27–33. 2005 doi:10.1159/000082378

de Setta, N., Monteiro-Vitorello, C.B., Metcalfe, C.J. *et al.* Building the sugarcane genome for biotechnology and identifying evolutionary trends. **BMC Genomics** 15, 540 (2014). https://doi.org/10.1186/1471-2164-15-540.

FAO.http://www.fao.org/fileadmin/templates/est/COMM_MARKETS_MONITORING/Sugar/Documents/sugar_assessment_food_outlook_may_2019.pdf. Visited 25/11/2021

Feng, X., Wang, Y., Zhang, N. *et al.* Systematic Identification, Evolution and Expression Analysis of the *SPL* Gene Family in Sugarcane (*Saccharum spontaneum*). **Tropical Plant Biol.** 14, p. 313–328. 2021. https://doi.org/10.1007/s12042-021-09293-4.

Fields, S., Song, O. A novel genetic system to detect protein–protein interactions. **Nature**, v. 340, n. 6230, p.245-246, jul. 1989.

Frantzeskakis, L., Courville, KJ., Plücker, L., Kellner, R., Kruse, J., Brachmann, A., Feldbrügge, M., Göhre V. The Plant-Dependent Life Cycle of Thecaphora thlaspeos: A Smut Fungus Adapted to Brassicaceae. **Mol Plant Microbe Interact.** 30(4). p. 271-282. 2017. doi: 10.1094/MPMI-08-16-0164-R. Epub 2017 Apr 19. PMID: 28421861.

Garsmeur, O., Droc, G., Antonise, R et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. **Nature Communications**, v. 9, n. 1, p.1-10, 6 jul. 2018.

Gao, L., Kelliher, T., Nguyen, L., Walbot, V. *Ustilago maydis* reprograms cell proliferation in maize anthers. **The Plant Journal**, v. 75, n. 6, p. 903-914, 8 ago. 2013. http://dx.doi.org/10.1111/tpj.12270.

Gerace, E., Moazed, D. Affinity pull-down of proteins using antiFLAG M2 agarose beads. **Methods Enzymol**, 559, 99-110. 2015.

Ghosh, I., Hamilton, A.D., Regan, L. Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. **Journal Of The American Chemical Society**, v. 122, n. 23, p.5658-5659, jun. 2000.

Horst RK. Smuts. In: Horst RK (Ed.) Westcott's Plant Disease Handbook. Springer Netherlands, The Netherlands. pp. 373-379. 2013.

Horstman, A., Tonaco, IA., Boutilier, K., Immink, RG. A cautionary note on the use of split-YFP/BiFC in plant protein-protein interaction studies. **Int J Mol Sci** 30, 15(6). p. 9628-43. 2014. doi: 10.3390/ijms15069628. PMID: 24886811; PMCID: PMC4100113.

Huh, SU., Cevik, V., Ding, P., Duxbury, Z., Ma, Y., Tomlinson L, et al. Protein-protein interactions in the RPS4/RRS1 immune receptor complex. **PLoS Pathog** 13(5): e1006376. 2017 https:// doi.org/10.1371/journal.ppat.1006376

Huang, Y., Li, W., Zhang, R., Wang, X. Diagnosis and Control of Sugarcane Important Diseases. **Color Illustration Of Diagnosis And Control For Modern Sugarcane Diseases, Pests, And Weeds**, p. 1-103, 2018. http://dx.doi.org/10.1007/978-981-13-1319-6_1.

ISO. International Sugar Organization. Disponível em: https://www.isosugar.org/sugarsector/sugar. Acesso em: 06 dez. 2021.

Jaswal, R., Kiran, K., Rajarammohan, S., Dubey, H., Singh, P., Sharma, Y., Deshmukh, R., Sonah, H., Gupta, N., Sharma, T.R. Effector Biology of Biotrophic Plant Fungal Pathogens: current advances and future prospects. **Microbiological Research**, v. 241, p. 1-27, dez. 2020. http://dx.doi.org/10.1016/j.micres.2020.126567.

Jones, J., Dangl, J. The plant immune system. **Nature** 444, p. 323-329. 2006.

Kerppola, TK. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. **Annu Rev Biophys**. 37, p. 465-87. 2008. doi: 10.1146/annurev.biophys.37.032807.125842. PMID: 18573091; PMCID: PMC2829326.

Kim, S., Kim, C.Y., Park, S.Y. *et al.* Two nuclear effectors of the rice blast fungus modulate host immunity via transcriptional reprogramming. **Nat Commun** 11, 5845, p. 1-11. 2020. https://doi.org/10.1038/s41467-020-19624-w

Klein, J., Saedler, H., Huijser, P. A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum* majus floral meristem identity gene SQUAMOSA. **Mol Gen Genet, 250,** p. 7–16. 1996.

Kudla, J., Bock, R. Lighting the Way to Protein-Protein Interactions: Recommendations on Best Practices for Bimolecular Fluorescence Complementation Analyses. **Plant Cell** 28(5):1002-8. 2016. doi: 10.1105/tpc.16.00043. Epub 2016 Apr 20. PMID: 27099259; PMCID: PMC4904677.

León-Ramírez, C., Cabrera-Ponce, J., et al. Infection of alternative host plant species by Ustilago maydis. **New Phhytologist** 164, p. 337-346. 2004.

Li, X., Lin, E., Huang, H., Niu, M., Tong, Z., Zhang, J. Molecular Characterization of SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Gene Family in *Betula luminifera*. **Frontiers In Plant Science**, v. 9, n. 608, p. 1-17, 4 maio 2018. http://dx.doi.org/10.3389/fpls.2018.00608.

Ling, H., Fu, X., Huang, N., et al. A sugarcane smut fungus effector simulates the host endogenous elicitor peptide to suppress plant immunity. **New Phytologist**, p. 1-15. 2021. doi: 10.1111/nph.17835.

Liu, Y., Aslam, M., Yao, LA. et al. Genomic analysis of SBP gene family in *Saccharum spontaneum* reveals their association with vegetative and reproductive development. **BMC Genomics** 22, 767, p. 1-16. 2021. https://doi.org/10.1186/s12864-021-08090-3

Maia, T., Rody, H., Bombardelli, R., et al. A bacterial type three secretion-based delivery system for functional characterization of *Sporisorium scitamineum* plant immune suppressing effector proteins. **Phytopathology**. 2021. (in press).

Marques, J. P. R., Appezzato-da-Gloria, B., Piepenbring, M., et al. Sugarcane smut: shedding light on the development of the whip-shaped sorus. **Annals Of Botany**, p.1-13, 27 ago. 2016.

Martínez-Soto, D., Robledo-Briones, A., et al. Transcriptomic analysis of *Ustilago maydis* infecting Arabidopsis reveals important aspects of the fungus pathogenic mechanisms. **Plant Signaling & Behavior**, v. 8, n. 8, p.1-13, ago. 2013.

Matiolli, C., Melloto, M. A Comprehensive Arabidopsis Yeast Two-Hybrid Library for Protein-Protein Interaction Studies: A Resource to the Plant Research Community. **Molec. Plant-Microbe Interactions**, v. 31, n. 9, p. 899-902, jul. 2018.

Mao, H., Yu, L., Li, Z., et al. Genome-wide analysis of the SPL family transcription factors and their responses to abiotic stresses in maize. **Plant Gene** 6, p. 1-12. 2016.

Meinke, D. W. *Arabidopsis thaliana*: A Model Plant for Genome Analysis. **Science**, v. 282, n. 5389, p. 662-682, 23 out. 1998.

Méndez-Morán, L., Reynaga-Pena, C.G., Springer, P.S. and Ruiz-Herrera, J. *Ustilago maydis* Infection of the Non-Natural Host *Arabidopsis thaliana*. **Phytopathology**, 95, p. 480-488. 2005.

Miura, K. An Overview of Current Methods to Confirm Protein- Protein Interactions. **Protein & Protein Letters** 25, p. 728-733. 2018.

Monteiro-Vitorello, C.B., et al. Progress in understanding fungal diseases affecting sugarcane: smut. In: ROTT, Philippe. Achieving sustainable cultivation of sugarcane. Gainesville: Burleigh Dodds, p. 1-25. 2018.

Monteiro-Vitorello, C.B. Exploring the molecular cross-talk between the biotrophic pathogen *Sporisorium scitamineum* and its host sugarcane: The use of Arabidopsis as a model system. 2017. Dados não publicados.

Nir, O., Keren, S., Shaul, Y., The Analysis of Protein-Protein Interactions in Plants by Bimolecular Fluorescence Complementation, **Plant Physiology,** 145 (4), p. 1090–1099, 2007. https://doi.org/10.1104/pp.107.107284

Park, J.Y., Jin, J., Lee, Y.W., Kang, S. and Lee, Y.H. Rice blast fungus (*Magnaporthe oryzae*) infects Arabidopsis via a mechanism distinct from that required for the infection of rice. **Plant Physiology**, *149*(1), p.474-486. 2009.

Peters, L. P., Carvalho, G., Vilhena, M. B., Creste, S., Azevedo, R., Monteiro-Vitorello, C. B. Functional analysis of oxidative burst in sugarcane smut-resistant and -susceptible genotypes. **Planta**, v. 245, n. 4, p. 749-764, 21 dez. 2016. http://dx.doi.org/10.1007/s00425-016-2642-z.

Peters, L.P., Teixeira-Silva, N., Bini, A., et al. Differential responses of genes and enzymes associated with ROS protective responses in the sugarcane smut fungus. **Fungal Biology** 124, p. 1039-1051. 2020.

Piperidis, N., D'hont, A. Sugarcane genome architecture decrypted with chromosome-specific oligo probes. **The Plant Journal**, p. 2039-2051, jun. 2020.

Piperidis, G., Piperidis, N., D'Hont, A. Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. **Mol Genet Genomics** 284, p. 65–73. 2010. https://doi.org/10.1007/s00438-010-0546-3

Preston, J.C., Hileman, L.C. Functional Evolution in the Plant SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) Gene Family. **Frontiers In Plant Science**, v. 4, n. 80, p. 1-13, 2013. http://dx.doi.org/10.3389/fpls.2013.00080.

Rago, A., Casagrande, M., Massola, J., Nelson, S. Variabilidade patogênica de *Ustilago scitaminea* no estado de São Paulo. Summa Phytopathologica [online]. 2009, v. 35, n. 2 [Acessado 8 Dezembro 2021] , pp. 93-97. Disponível em: <https://doi.org/10.1590/S0100-54052009000200002>. Epub 04 Ago 2009. ISSN 1980-5454. https://doi.org/10.1590/S0100-54052009000200002.

Ramachandran, S.R., Yin, C., Kud, J., Tanaka, K., Mahoney, A.K., Xiao, F., Hulbert, S.H. Effectors from wheat rust fungi suppress multiple plant defense responses. **Phytopathology** 107, p. 75-83. 2016.

Raudaskoski, M., Kothe, E. Basidiomycete mating type genes and pheromone signaling. **Eukaryot Cell.** 9(6), p. 847-59. 2010. doi: 10.1128/EC.00319-09.

Rody, HVS., Bombardelli, RGH., Creste, S., Camargo, LEA., Van Sluys, MA., Monteiro-Vitorello, CB. Genome survey of resistance gene analogs in sugarcane: genomic features and differential expression of the innate immune system from a smut-resistant genotype. **BMC Genomics**. 20(1):809. 2019. doi: 10.1186/s12864-019-6207-y. PMID: 31694536; PMCID: PMC6836459.

Rody, HVS., Camargo, LEA., Creste, S., Van Sluys, M-A., Rieseberg, LH., Monteiro-Vitorello, CB. Arabidopsis-Based Dual-Layered Biological Network Analysis Elucidates Fully Modulated Pathways Related to Sugarcane Resistance on Biotrophic Pathogen Infection. **Front. Plant Sci.** 12:707904. p. 1-12. 2021. doi: 10.3389/fpls.2021.707904

Saunders, DGO., Win, J., Cano, LM., Szabo, LJ., Kamoun, S., et al. Using Hierarchical Clustering of Secreted Protein Families to Classify and Rank Candidate Effectors of Rust Fungi. **PLoS ONE** 7(1): e29847. 2012. doi:10.1371/journal.pone.0029847

Schwab, R., Palatnik, J.F., Riester, M., Schommer, C., Schmid, M., Weigel, D. Specific effects of microRNAs on the plant transcriptome. **Dev. Cell 8,** p. 517–527. 2005.

Screenivasan, T. V., Ahloowalia, B. S., Heinz, D. J. Sugarcane improvement through breeding. **Cytogenetics**, v. 5, p. 211–253. 1987.

Selin, C., de Kievit, TR., Belmonte, MF., Fernando WGD. Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. **Front. Microbiol.** 7:600. 2016. doi: 10.3389/fmicb.2016.00600

Silva, M. *Arabidopsis thaliana* as a model plant for the study of plant-pathogen interaction between *Sporisorium scitamineum* and sugarcane. 2019. 51 f. Dissertação (Mestrado) - Curso de Genética e Melhoramento de Plantas, Universidade de São Paulo, Piracicaba, 2019.

Silva, G., Silva, E., Correa, J., Vicente, M., Jiang, N., Notini, M., Junior, A., Jesus, F., Castilho, P., Carrera, E. Tomato floral induction and flower development are orchestrated by the interplay between gibberellin and two unrelated micro RNA -controlled modules. **New Phytologist**, v. 221, n. 3, p. 1328-1344, 19 out. 2018. http://dx.doi.org/10.1111/nph.15492.

Sforça D.A., Vautrin S., Cardoso-Silva C.B., et al. Gene Duplication in the Sugarcane Genome: A Case Study of Allele Interactions and Evolutionary Patterns in Two Genic Regions. **Front. Plant Sci**. 2019. 10:553. doi: 10.3389/fpls.2019.00553

Souza, G.M., et al. Assembly of the 373k gene space of the polyploid sugarcane genome reveals reservoirs of functional diversity in the world's leading biomass crop. **Gigascience**, v. 8, n. 12, p.1-18, 29 nov. 2019.

Sugarcane.org (Org.). **Sugarcane Products.** 2017. Disponível em: <https://sugarcane.org/#>. Acesso em: 04 dez. 2021.

Sun, Y., Rombola, C., Jyothikumar, V., Periasamy A. Férster resonance energy transfer microscopy and spectroscopy for localizing protein-protein interactions in living cells. **Cytometry A**, 83, p. 780-793. 2013.

Sundar A.R,. Barnabas E.L., Malathi P., Viswanathan, R. A mini-review on smut disease of sugarcane caused by *Sporisorium scitamineum*. In: Mworia J, (Ed.) Botany. Rijeka, Croatia: InTech; 2012, p. 109–128. 2012.

Schmidt, S.M., Kuhn, H., Micali, C., Liller, C., Kwaaitaal, M., Panstruga, R. Interaction of a *B lumeria graminis f. sp. hordei* effector candidate with a barley ARF-GAP suggests that host vesicle trafficking is a fungal pathogenicity target. **Molecular Plant Pathology**, 15(6), p. 535-549. 2014.

Sperschneider, J., Dodds, P.N., Gardiner, D.M., Singh, K.B., Taylor, J.M. Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. **Molecular plant pathology** 19, p. 2094-2110. 2018.

Taniguti, L.M., Schaker, P.D.C., Benevenuto, J., Peters, L.P., Carvalho, G., Palhares A., et al. Complete Genome Sequence of *Sporisorium scitamineum* and Biotrophic Interaction Transcriptome with Sugarcane. **PLoS ONE** 10(6), p. 1-31. 2015. doi:10.1371/journal.pone.0129318.

Tatusov, R.L., Fedorova, N.D., Jackson, J.D. *et al.* The COG database: an updated version includes eukaryotes. **BMC Bioinformatics** 4, 41. 2003. https://doi.org/10.1186/1471-2105-4-41

Teixeira-Silva, N., Schaker, P., Rody, H., Maia, Thiago., Garner, M., et al. Leaping into the Unknown World of *Sporisorium scitamineum* Candidate Effectors. **Journal Of Fungi**, v. 6, n. 4, p. 1-26, 4 dez. 2020. http://dx.doi.org/10.3390/jof6040339.

Thirugnanasambandam, P., Hoang, N., Henry, R. The Challenge of Analyzing the Sugarcane Genome. **Frontiers In Plant Science**, v. 9, p.1-18, 14 maio 2018.

Toruño, TY., Stergiopoulos, I., Coaker, G. Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners. **Annu Rev Phytopathol** 4 (54), p. 419-41. 2016. doi: 10.1146/annurev-phyto-080615-100204.

Van Esse, H.P., van't Klooster, J.W., Bolton, M.D., Yadeta, K.A., Van Baarlen, P., Boeren, S., Vervoort, J., de Wit, P.J., Thomma, B.P. The *Cladosporium fulvum* virulence protein Avr2 inhibits host proteases required for basal defense. **The Plant Cell,** 20(7), p. 1948-1963. 2008.

Van der Linde, K., Göhre, V. How Do Smut Fungi Use Plant Signals to Spatiotemporally Orientate on and In Planta?. **J. Fungi** 7, 107. 2021. https://doi.org/10.3390/ jof7020107

Vilela, M., Bem, l., Van Sluys, M et al. Analysis of Three Sugarcane Homo/Homeologous Regions Suggests Independent Polyploidization Events of *Saccharum officinarum* and *Saccharum spontaneum*, **Genome Biology and Evolution,** 9 (2), p. 266–278. 2017. https://doi.org/10.1093/gbe/evw293.

Vollmeister E., Schipper K., Baumann S., Haag C., Pohlmann T., Stock J., Feldbrügge M. Fungal development of the plant pathogen *Ustilago maydis*. **FEMS Microbiol Rev**. 36(1), p. 59-77. 2012. doi: 10.1111/j.1574-6976.2011.00296.x.

Walter, M., et al. Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. **The Plant Journal**, v. 40, n. 3, p.428-438, 14 set. 2004.

Xing, S., et al. Techniques for the analysis of protein-protein interactions in vivo. **Plant Physiology**, p.727-758, 25 abr. 2016.

Zhang, J., Zhang, X., Tang, H., Zhang, Q., Hua, X., Ma, X., Zhu, F., Jones, T., Zhu, X., Bowers, J et al. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum L.* **Nat Genet** 50, p. 1565–1573. 2018.

Zhu, T., Liu, Y., Ma, L., et al. Genome-wide identification, phylogeny and expression analysis of the *SPL* gene family in wheat. **BMC Plant Biol** 20, 420. 2020. https://doi.org/10.1186/s12870-020-02576-0

Zuo, W., Ökmen, B., Depotter, J., Ebert, M., Redkar, A., Villamil, J., Doehlemann, G. Molecular Interactions Between Smut Fungi and Their Host Plants. **Annual Review Of Phytopathology**, v. 57, n. 1, p. 411-430, 25 ago. 2019. http://dx.doi.org/10.1146/annurev-phyto-082718-100139.

**CHAPTER 2: SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE 13 (SPL13) INTERACTS WITH A CANDIDATE EFFECTOR OF *Sporisorium scitamineum***

**Abstract**

*Sporisorium scitamineum* (Ssc) (Syd.) M. Piepenbr., M. Stoll &amp; Oberw., 2002 is a biotrophic fungal pathogen of sugarcane causing the smut disease, which is disseminated worldwide in crop fields. Sugarcane is one of the most prominent crops for biofuel production worldwide. However, fungal diseases affecting production cause significant economic losses. Previously, predicted candidate effectors (CEs) were analyzed to identify host targets in experiments using a comprehensive Arabidopsis yeast two-hybrid Library for protein-protein interaction studies and a co-immunoprecipitation (CoIP) assay followed by mass spectrometry resulting in the identification of an SPL13 transcription factor (Squamosa-Promoter Binding Protein-Like). Therefore, this study aimed to validate the interaction CE g5159-SPL13 using bimolecular fluorescence complementation assays (BiFC). Here, we present the cloning methods involving sequence adaptation and a combination of Golden Gate and Gateway strategies as the first steps to investigate the interaction CE g5159-SPL13 using bimolecular fluorescence complementation assays (BiFC). The amplicons containing the CE g5159 (with no signal peptide) and SPL13 coding regions were cloned using the Golden Gate assembly method and recombined into the BiFC expression vectors pSITE-nYFP and pSITE-cYFP. The CE g5159 sequence was adapted without losing the reading frame to introduce a base substitution necessary to proceed with cloning. Every cloning step was followed by insert sequencing before moving forward. The introduction of both vectors with the appropriate inserts into *Agrobacterium tumefaciens* GV3101, infiltration in *Nicotiana benthamiana* leaves, expression of both proteins determined their subcellular location and confirmed the interaction when analyzed using confocal microscopy. Effector biology is a significant field of investigation in fungal genetics and host-pathogen interactions, and implementing strategies to determine host interactors will significantly improve functional studies.

**Keywords:** Protein-protein interaction; bimolecular fluorescence complementation assay (BiFC); AtSPL13; CE g5159

## 1. Introduction

Sugarcane production significantly impacts the Brazilian economy, acting as the largest producer and exporter of ethanol and sugar globally. Brazil produced 642.7 million tons of sugarcane, 34 billion liters of ethanol and 29.8 million tons of sugar in 2019-2020. In July 2019, our country exported approximately 1.82 million tons of sugar, and 209.6 million liters of ethanol, resulting in significant economic value (Conab, 2019). However, many diseases can affect sugarcane productivity, resulting in substantial losses in sugar content and cane yield.

Sugarcane smut is a disease caused by the biotrophic fungus *Sporisorium scitamineum*, which is known to develop a whip-like structure formed from the apex of plants and composed of fungal and plant tissues. The whip is associated with fungal reproduction

and teliospores production. Teliospores can be easily spread by wind and rain, resulting in the rapid progress of the disease, which can cause losses from 30% to 100% in productivity (Que et al., 2014). Smut fungi present a biotrophic lifestyle to complete their life cycle. Sugarcane infection is only possible after fusing two mating-compatible sporidial cells forming dikaryotic hyphae. Hyphal growth induces the formation of a specialized structure (appressorium) at the tip to penetrate host tissues.

The plant-pathogen interaction has evolved, establishing a highly regulated cell-to-cell communication. While plants recognize pathogens by activating signaling cascades and hormone signaling for defense responses, phytopathogens manipulate hormonal and cellular signaling to invade hosts' cells. Signaling cascades include reactive oxygen species, Ca+ pathways and mitogen-activated protein kinase (Tabassum et al., 2021). *S. scitamineum* induces a robust oxidative burst, cell wall modification, and defense hormone signaling in sugarcane smut resistant plants (Peters et al., 2016; Schaker et al., 2017; Marques et al., 2018; Vicente et al., 2021). Some of the transcripts in this first response undergo alternative splicing not detected in plants susceptible to smut (Bedre et al., 2019). In addition, the fungus modulates the host's meristematic functions toward differentiation of reproductive organs during whip emission in susceptible plants (Schaker et al., 2016). Recently, our group unraveled that in resistant plants, the modulation of the meristematic functions occurs via upregulation of negative regulators of flowering transition (unpublished data).

Effector biology is a recent and powerful line of investigation to understand pathosystems and direct plant breeding programs (Depotter et al., 2021; Hogenhout et al., 2009; Yuan et al., 2021). Regarding *S. scitamineum* recently had described a putative effector proved to be involved in plant immune defense (Ling et al., 2021). The authors showed co-expression and physical interaction of the effector *SsPele1* with the sugarcane *PLANT ELICITOR PEPTIDE RECEPTOR1* (*ScPEPR1*), which encodes a receptor-like kinase suppression of ScPEPR1-mediated immune responses. Other *S. scitamineum* candidate effectors have been examined considering their subcellular location within host cells (Teixeira-Silva et al., 2020) and ability to suppress MTI (Pattern-Triggered Immunity) and ETI (Effector-Triggered Immunity) plant immune responses (Maia et al., 2022).

Our group used a heterologous Y2H system to identify potential interactors of *S. scitamineum* CE. This experiment revealed a protein-protein interaction between CE g5159 and an SPL13 (Squamosa Promoter Binding Protein-Like) of *Arabidopsis thaliana. A. thaliana* is a model plant used frequently to set the basis of plant-pathogen interactions to overcome some of the issues regarding complex genomes (León-Ramírez et al., 2004). The

CE 5159, in experiments using a transient expression after *Agrobacterium* transformation of *Nicotiana benthamiana* leaves, localized associated to the cell wall (Teixeira-Silva et al., 2020). Some interactors of *N. benthamiana* were identified in co-immunoprecipitation assays followed by protein identification using liquid chromatography-tandem mass spectrometry (CoIP/LC-MS/MS). The most promising candidate was a plant endochitinase A with a high identification score. The CE g5159 suppressed MTI and ETI in classical experiments using co-infiltration experiments with *Pseudomonas syringae* pv. *tomato* DC3000 (PTI) and the AvrB-induced suppressors of effector-triggered immunity (ETI), respectively (Maia et al., 2022). The gene encoding CE g5159 has paralogs organized in tandem localized in a region of chromosome 15 syntenic to *S. reilianum*, the maize smut pathogen. However, CE g5159 is exclusive of the *S. scitamineum* genome (Maia et al., 2022). The protein sequence of CE g5159 has the N-terminal signal peptide followed by a disordered region (Figure 1a).

The *A. thaliana SPLs* play a role in the transition of juvenile to adult and the vegetative to reproductive phases (Fornara and George Coupland, 2009). *SPL13* of *A. thaliana* also regulates the transition from vegetative to reproductive stages in other plant species (Gao et al., 2018). A single mutation inducing loss of function of SPL13 function showed a significant effect on shoot development of *A. thaliana* (Xu et al., 2016). The miRNAs, miR156 and miR157, target transcripts of some SPL members. For instance, miR156 is an essential regulator of vegetative transition in *A. thaliana* and other flowering plants. At the beginning of the plant's development, miR156 is expressed at a high level, and its expression decreases as the shoot develops. This reduction follows the induction of *SPL13* expression (He et al., 2018). SPL13 amino acid sequence shares all the features of the family members, including the disorder region, the SBP-box with two Zn2+-binding sites: Cys-Cys-His-Cys and a putative conserved nuclear localization signal (NLS) (Birkenbihl et al., 2005) (Figure 1b).

To validate the interaction between AtSPL13-g5159, we performed an additional protein-protein interaction technique. We showed through the BiFC method that the interaction occurs in the cytoplasm of *N. benthamiana* leaves. Although these are preliminary results and more experiments should be performed with the sugarcane SPL13 ortholog, it is tempting to speculate that the effector prevents the role of SPL13, impairing the transition from the vegetative to reproductive stages, leading to sugarcane smut resistance.
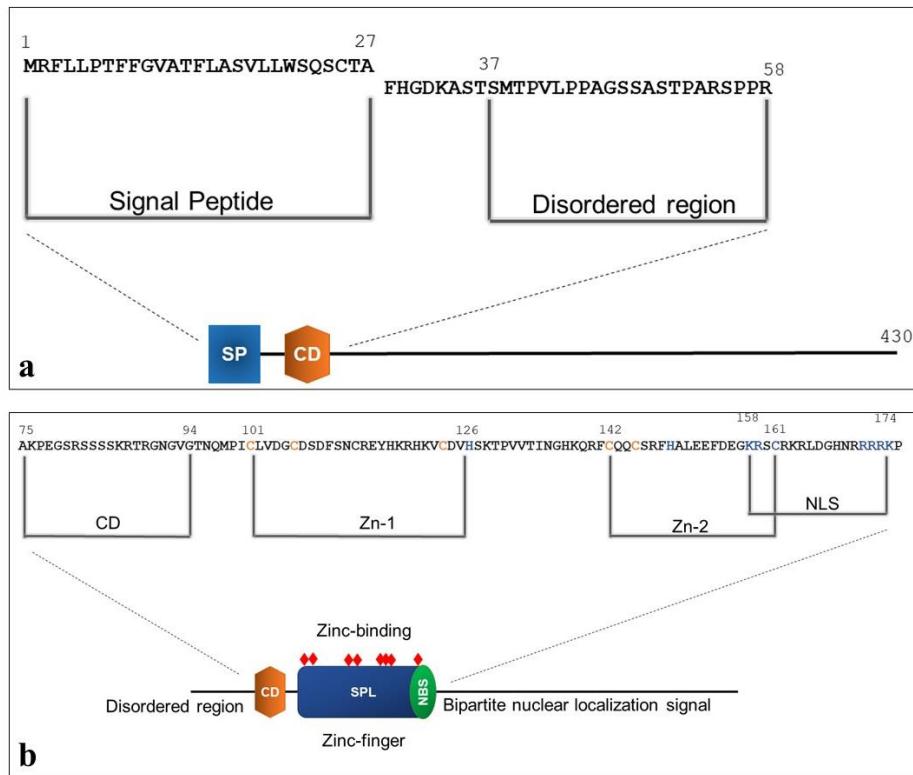
**Figure 1.** Protein structure of a) candidate effector g5159 of *S. scitamineum* and b) transcription factor SPL13 of *A. thaliana.*

## 2. Material and Methods

### 2.1 Biological material and Experimental Condition

For RNA extraction, we used 30 days-old *Arabidopsis thaliana* Columbia (Col-0). Seeds of *A. thaliana* were chemically treated (70% alcohol immersion for 5 minutes; 2% sodium hypochlorite solution for 20 minutes under agitation; and 4 washes in sterilized ultrapure water). The seeds were maintained in the dark at 4°C for four days to break the dormancy. Then, they were cultivated in petri dishes with MS media (Murashige & Skoog, 1962) in a concentration of 4,31 g.L-1, 1,5% of sucrose and 1% of agar. Petri dishes were maintained in the dark for 4 days at 10°C and then in a photoperiod of 16 hours of light and 8 hours of dark, at 22°C for 26 days.

Teliospores of *Sporisorium scitamineum* Ssc39 were used to inoculate germinated buds of sugarcane, and 48 hours after inoculation, samples were collected for total RNA extraction (Teixeira-Silva et al., 2020). The RNA was used for cDNA synthesis, from where the CE g5159 was amplified (primers are listed in Table 1). For cloning methods, we used strains of *e.coli* DH5α, TOP10 and DB3.1. For Agroinfiltration and the BiFC experiment, we

used three-four weeks old tobacco (*Nicotiana benthamiana*) plants and GV3101 strain of *Agrobacterium tumefaciens.*

## 2.2 RNA Extraction and cDNA Synthesis

A volume of 70 mg of macerated and frozen *A. thaliana* samples were separated for total RNA extraction. We used Trizol® (Ambion by life Technologies™) according to the protocol. Then, we used the purification kit PureLink™ RNA Mini Kit (Invitrogen™). The samples were treated with DNase (Sigma Aldrich™), and the quality and quantity of the extraction was analyzed by electrophoresis gel and spectrophotometer BioDrop (Harvard Bioscience™).

For cDNA synthesis, 4µg of total RNA and Oligo(dT) primers were used according to the kit GoScript™ Reverse Transcription System (Promega) instructions.

## 2.3 *In silico* sequence analysis and primer design

We used TAIR (https://www.arabidopsis.org/) and Uniprot (https://www.uniprot.org/) databases to analyze *the Arabidopsis thaliana SPL13* gene. The gene is located in chromosome 5 of the genome sequence, where it is duplicated in tandem: *SPL13A* (Locus: AT5G50570) and *SPL13B* (Locus: AT5G50670) (Gandikota, 2007). *SPL13A* is 2639pb long, while *SPL13B* is 2499 bp (Figure 2), and they are located 30kb apart. The transcripts and protein sequences are the same.

For cloning procedures, we first designed the primers using the Primer3 program (Rozen & Skaletsky, 1999) and primer-blast considering a GC content of 40% and a TM of 60%. We also used the NetPrimer program (http://www.premierbiosoft.com/netprimer/) to analyze primer quality regarding inappropriate parameters, such as hairpin formation (up to -3), self-dimers (up to -6) and cross dimer (up to -5). Next, we added vector tags to the primers for cloning procedures using Benchling software.
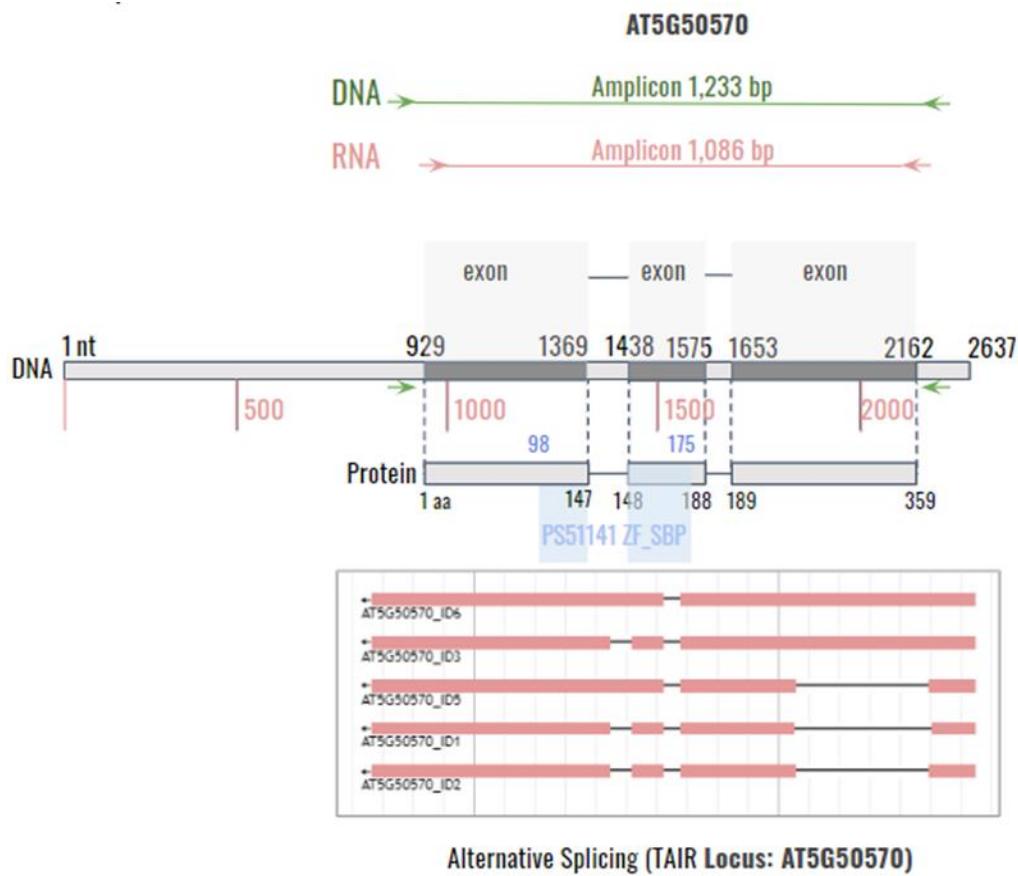
**Figure 2.** Sequence of the gene, the protein and alternative splicing of SPL13A. Exons are represented in gray and zinc-fingers domains are represented in blue. The numbers represent the nucleotide and amino acid positions.

## 2.4 Cloning

The transcription factor SPL13 of *A. thaliana* was amplified through a touchdown PCR with a pair of primers containing Golden Gate and Gateway cloning tags. Thermocycle conditions were set according to Korbie and Mattick., (2008), in which the annealing temperature decreased 1ºC every cycle. Every cycle started at 95ºC for 2 minutes, followed by 95ºC for 20 seconds. Extension temperatures were set at 65ºC in every cycle. We set the first annealing temperature at 66ºC for 20 seconds and performed one cycle decreasing the annealing temperature until 58ºC. Then, we performed 34 cycles with an annealing temperature of 54ºC. The candidate effector g5159 of *S. scitamineum* was amplified without the signal peptide and in two separate fragments to induce a silent mutation in order to remove an additional restriction site in the gene. It was necessary to use two pairs of primers to induce this base substitution (Table 1). Amplifications were performed using Phusion 2x High Fidelity enzyme (Thermo Scientific) according to the manufacturer's instructions. PCR

product sizes were verified by gel electrophoresis and purified with the gel extraction method using the Wizard SV Gel and PCR Clean-Up System kits (PROMEGA).

Amplicons were recombined with the PUPD2 vector (Varquez-Vilar et al., 2017) using ESP3I restriction enzyme (Thermo Scientific) and T4 ligase (NEB). Next, PUPD2 vectors containing, separately, the coding region, gateway AttL1 site and AttL2 site were recombined with the Golden Gate acceptor vector pICSL86900_OD vector (a gift from Mr Mark Youles, Addgene plasmid # 86178; http://n2t.net/addgene:86178; RRID: Addgene_86178) using BSAI/ECO3I restriction enzyme (NEB) and T4 ligase. Finally, the coding regions with AttLs sites were recombined into BiFC expression vectors pSITE-nYFP and pSITE-cYFP (Martin et al., 2009) using Gateway® LR Clonase® II Enzyme mix (Sigma-Aldrich). Recombinant vectors were transformed into chemically competent DH5α cells. The vectors were purified from single colonies grown overnight (37°C; 240 rpm) with the PureYield™ Plasmid Miniprep System (Promega). All constructs were verified for the correct insertions by sequencing. Vectors Map generated on www.benchling.com can be found in the supplementary material (Supplementary Figure 1).

**Table 1.** Primers used to clone SPL13 from *Arabidopsis thaliana* and candidate effector g5159

| PRIMERS | SEQUENCE 5'- 3' |
| --- | --- |
| At_SPL13_F | GGCGTCTCACTCGAATGGACTGGAATTTCAAACTTAGC |
| At_SPL13_R | GGCGTCTCACTCAAAGCACTTCTTCTTCTACTACTCCC |
| 5159_F | GGCGTCTCACTCGAATGTTCCACGGGGACAAAGC |
| 5159_R_mut | GGCGTCTCAGTCCATGTTTCTTGAGAGCCTGGA |
| 5159_F_mut | GGCGTCTCAGGACGCTTCCGcCTCTACATCACCT |
| 5159_R | GGCGTCTCACTCAAAGCTTAGACGTGGATGTGTCCTGC |

## 2.5 Agroinfiltration and Confocal Microscopy

Expression vectors were transformed into GV3101 *A. tumefaciens* strain by electroporation. Constructs were verified by sequencing. Colonies grown overnight in LB medium were pelleted and resuspended in an infiltration buffer (10 mM MgCl2, 10 mM MES, 200 μM acetosyringone). Cells were infiltrated in 3–4-week-old *Nicotiana benthamiana* leaves at OD600 = 0.2. All constructs were co-infiltrated with p19 post-transcriptional inhibition suppressor and a nuclear marker. Protein-protein interaction was analyzed 2 days

after infiltration through images captured in a confocal microscope (Nikon C2+ ). For positive controls, we used two constructs for BiFC with tomato genes, *PROCERA/DELLA* and *LANCEOLATE* (Silva et al., 2018). For negative controls, we infiltrated one final construct with empty BiFC vectors carrying C and N terminals of YFP separately.

# 3. Results

## 3.1 Positive controls

For positive controls of the experiment, we used BiFC expression vectors containing *PROCERA/DELLA* and *LANCEOLATE* (LA) coding regions from Tomato (*Solanum lycopersicum*), which were previously shown to interact in the nucleus of *N. benthamiana* by Silva et al., (2018). A vector containing the reporter gene *mcherry* was used as a nuclear marker (Silva et al., 2015). Strong YFP fluorescence was observed with the positive control combinations in the nucleus of *N. benthamiana* cells (Figure 3).
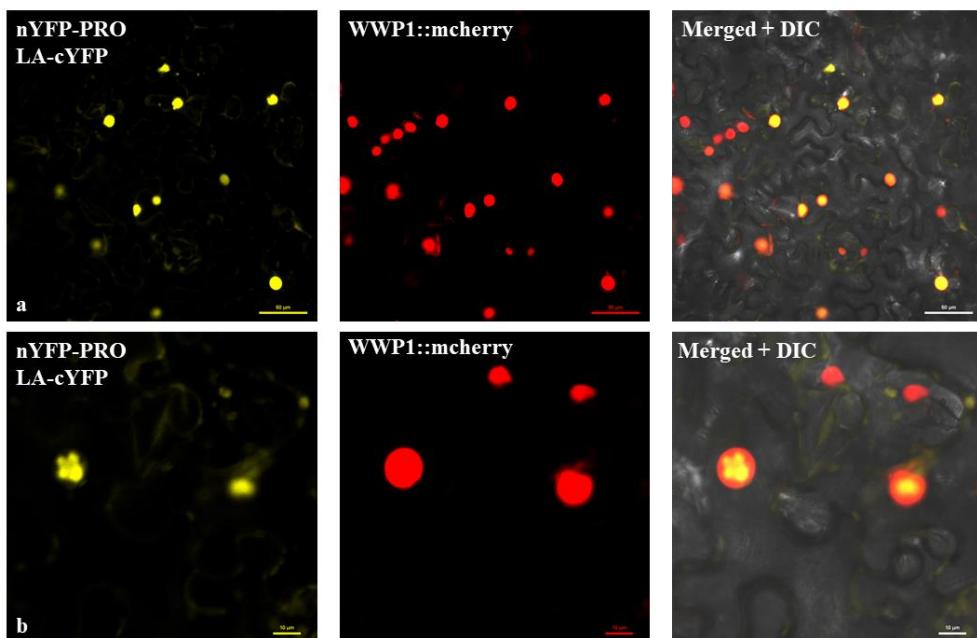


**Figure 3**. Bimolecular fluorescence complementation experiment using *Nicotiana benthamiana* leaves infiltrated with *A. tumefaciens*. The combination of nYFP-PRO and LA-cYFP was used as a positive control. Mcherry was used as a nuclear marker. The image of differential interference contrast (DIC) with the merged image is also exhibited. a) Bars, 50 µm and b) Bars, 10 µm

## 3.2 Negative controls

Negative controls were used to analyze background fluorescence of empty vectors carrying C and N terminus of YFP. The selection of negative controls was based on recent studies that performed BiFC assays (Li et al., 2019; Silva et al., 2018; Yang et al., 2008).

Thus, the combinations of nYFP-5159 with cYFP, nYFP with 5159-cYFP, nYFP-AtSPL13 with cYFP and nYFP with AtSPL13-cYFP were used as negative controls. Weak or no fluorescence signal was observed. The reporter gene *mcherry* was used as a nuclear marker (Figure 2).
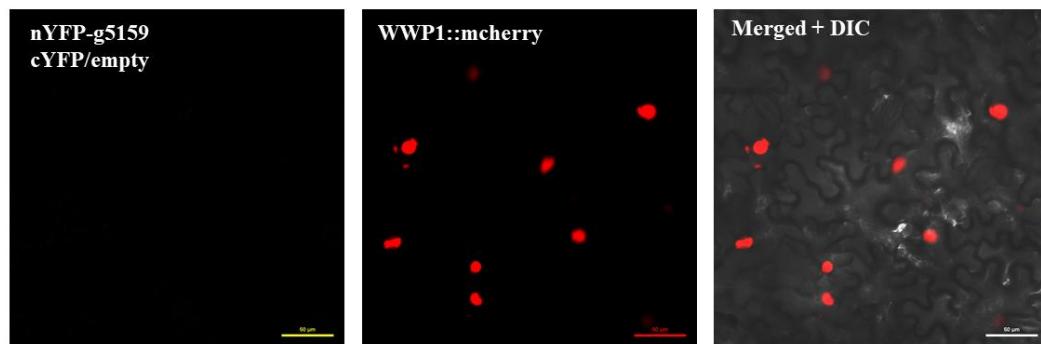


**Figure 4.** Bimolecular fluorescence complementation experiment using *Nicotiana benthamiana* leaves infiltrated with *A. tumefaciens*. The combination of nYFP-g5159 with an empty vector carrying cYFP did not show fluorescence background. Mcherry was used as a nuclear marker. The image of differential interference contrast (DIC) with the merged image is also exhibited. Bars, 50 µm.

## 3.3 Candidate effector g5159 and SPL13 showed interaction

Strong YFP fluorescence was detected in the cytoplasm for the treatment with cells infiltrated with nYFP-g5159 and AtSPL13-cYFP expression vectors (Figure 5). Treatment with cells infiltrated with nYFP-AtSPL13 and g5159-cYFP did not show any YFP fluorescence. However, this may happen because of topological constraints (Miller et al., 2015). This result indicates that the candidate effector g5159 of *S. scitamineum* physically interacts with the transcription factor SPL13 of *A. thaliana* in the cytoplasm of *N. bethamiana*, corroborating with previous Y2H assay performed by our group (unpublished results).
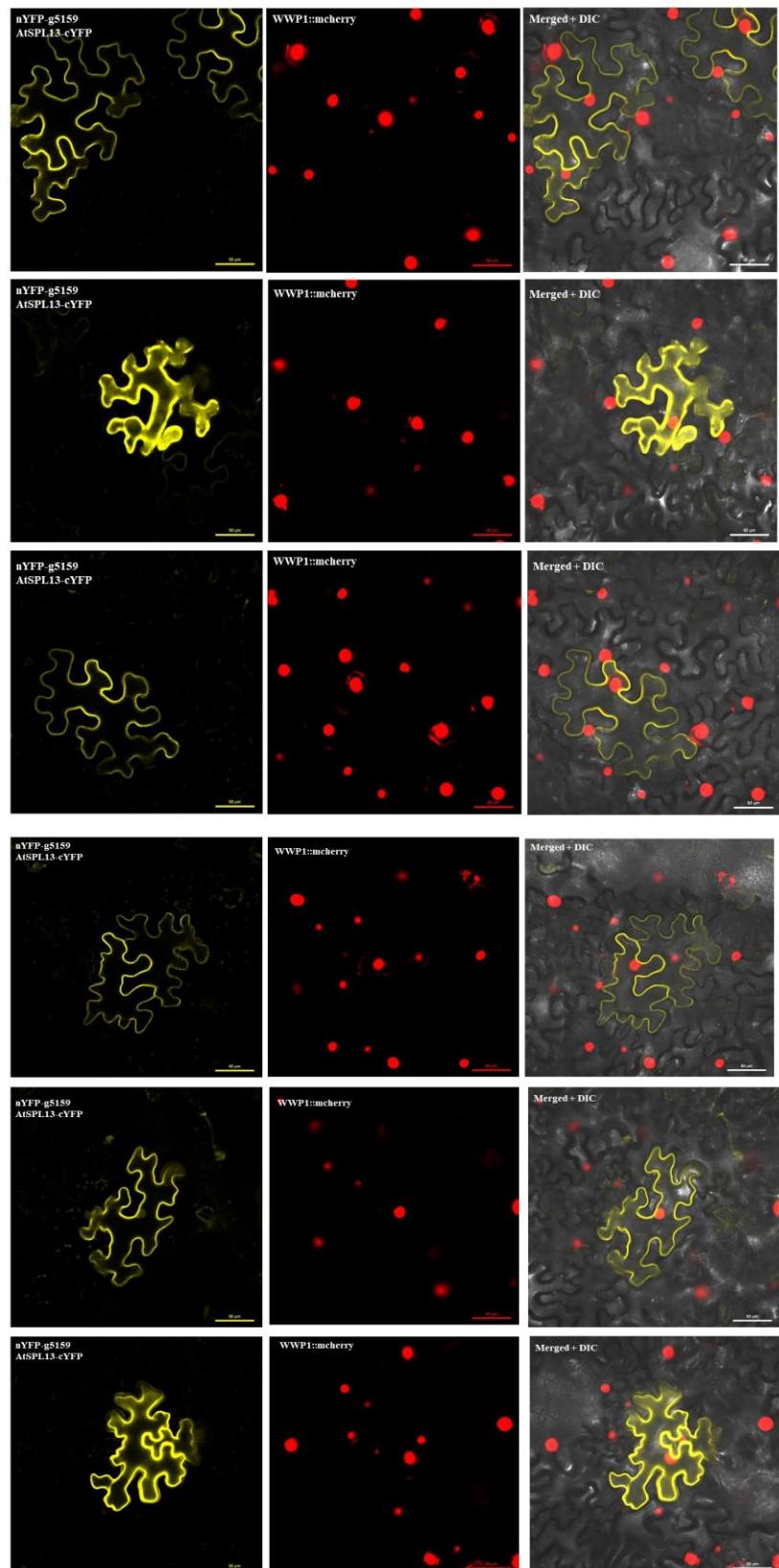
**Figure 5.** Bimolecular fluorescence complementation experiment using *Nicotiana benthamiana* leaves infiltrated with *A. tumefaciens.* The combination of nYFP-g5159 with cYFP-AtSPL13 showed strong fluorescence signals. Mcherry was used as a nuclear marker. The image of differential interference contrast (DIC) with the merged image is also exhibited. Bars, 50 µm.

**4. Discussion**

Our images from the BiFC assay suggested that the proteins of interest, the candidate effector (CE) g5159 and the transcription factor SPL13 from *A. thaliana*, interacted in the cytoplasm. These results confirmed the interaction between g5159 and SPL13 identified in Y2H assay (unpublished results).

A recent study detected the CE g5159 of *Sporisorium scitamineum* to play an important role inside the pathosystem. The CE g5159 effector tagged in both C and N terminus of the fluorescence protein GFP was found in only one specific compartment, the cell wall. The authors also found a plant endochitinase A to be a strong candidate interactor of CE g5159. Furthermore, the authors tested the influence of this protein regarding the suppression of basal defenses responses. They found that CE g5159 increased Pst (*P. syringae)* disease symptoms with the production of chlorotic sites and water-soaked lesions in leaves of *N. benthamiana.* The candidate effector also showed to suppress AvrB-induced ETI responses in tobacco, suggesting that the fungus produces effectors to inactivate hypersensitive responses cell death as a strategy to sabotage the host immune system (Teixeira-Silva et al., 2020).

Usually, we expected to observe transcription factors localized in the nucleus to induce gene expression. However, transcription factors are initially in the cytoplasm. Only after receiving a signal from the cell membrane signal transduction, transcription factors are activated and then translocated from the cytoplasm into the nucleus, where they interact with the corresponding DNA *cis*-acting regulatory element (Liu et al., 2018). Indeed, SPL13-CE g5159 interaction observed in the cytoplasm agrees with our hypothesis that the effector holds SPL13 translocation to the nucleus, somehow interfering in smut resistant plants with the transition from vegetative to reproductive stages. In sugarcane, resistant plants upregulated negative regulators of floral transition (vegetative-reproductive stage) such as *PIE1* (Uniprot ID: Q7X9V2), *ESD7* (F4HW04), *BRM* (Q6EVK6), *CLF* (P93831), *TEM2* (P82280), *UBP12* (Q9FPT1) and *SYD* (F4IHS2) (work in progress). This same effect is not observed in susceptible plants that induce reproductive transition earlier, and only later, after whip emission, repression of flowering genes is observed.

The biotrophic fungus, *Puccinia striiformis f. sp. tritici* (Pst), secretes an effector, PstGSRE1, which disrupts the nuclear accumulation of a zinc-finger-type LSD1 transcription factor (TaLOL2) of wheat that is involved in host resistance related to ROS signaling. In

addition, the study suggested that this effector may modify the subcellular localization of this transcription factor (Qi et al., 2019).

It is not uncommon to have smut pathogens interfering with flowering pathways. *For example, Villosiclava virens*, the cause of false smut disease in rice, infects flowers and converts grains into a false smut ball (Fan et al., 2016). Although the molecular interaction within this pathosystem is unsettled, Fan et al. (2015) showed that flowering-related genes are suppressed in infected hosts, suggesting that the fungus can maintain its colonization by inhibiting flowering in rice. Similar results described by Schmitz et al. (2018) studying *Ustilago maydis*, the corn smut pathogen, demonstrated downregulation of transcription factors involved in plant development, and upregulation of the ones expressed in flowers. The authors suggested that the pathogen inhibits the transition to the adult phase and induces the flowering pathway (Schmitz et al., 2018).

Both SPL13 and CE g5159 present disordered domains. These domains are regions that do not fold into secondary and tertiary structures spontaneously (Covarrubias et al., 2017). Instead, they "adjust" their folding to the interactor. This feature confers certain flexibility resulting in variable folds and many possible interactors. A study detected over 50% of common proteins targeted by effectors showing disorder domains (Ceulemans et al., 2021).

As mentioned before, the interaction between smut and sugarcane revealed that the pathogen modulates the host's meristematic function in early infection (Shacker et al., 2017). Considering that *SPL13* is involved in the phase transition, we also investigated the interactors of AtSPL13 described on STRING database through text mining (https://string-db.org/) are flowering repressors, such *as Sclafmutze (SMZ), Schnarchzapfen (SNZ), TOE2, TOE3*, and *AP2 (*Zheng et al., 2019). Recently, our group revealed that resistant plants 48 hai presented highly upregulated flowering repressors compared to control samples and susceptible plants (ongoing work).

Therefore, it is tempting to speculate that CE g5159 interferes with SPL13 migration to the nucleus to modulate the transition from the vegetative to the reproductive phase earlier in resistant plants and later in susceptible ones to allow the whip development. We will investigate this hypothesis further.

## 5. Conclusion and future perspectives

The mechanisms in which *Sporisorium scitamineum* manipulates and reprogram host's transcriptional profiles are unclear. Recently, candidate effectors were selected as strong

players regarding smut resistance in sugarcane. The CE g5159 especially showed the most contrasting results (Teixeira-Silva et al., 2020). Here, we show that CE g5159 interacts with SPL13, which is involved in plant's development and transition to vegetative and reproductive phases. Therefore, SPL13 is also an excellent candidate to explore how flowering and whip development pathways are associated. This study provides evidence and outlines future investigation on how the effector g5159 cooperates towards differentiation of meristematic functions in sugarcane.

Experiments of colocalization of SPL13 and CE g5159 are being developed in our lab to detect if AtSPL13 subcellular localization varies when there is no interaction with CE g5159. Furthermore, attempts to PCR-amplify SPL13 transcription factor from sugarcane varieties (RB925345, SP80-3280 and IAC66-6) are also being performed to develop a new BiFC assay with sugarcane proteins. Repetition of the BiFC experiment with our constructs will be performed one more time.

## References

Bedre, R., et al. Genome-wide alternative splicing landscapes modulated by biotrophic sugarcane smut pathogen. **Scientific Reports**, v. 9, n. 1, p.1-12, 20 jun. 2019.

Benchling [Biology Software]. 2020. Retrieved from https://benchling.com.

Birkenbihl, R.P., Jach, G., Saedler, H., Huijser, P. Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. **J Mol Biol.** 352(3), p. 585–596. 2005.

Ceulemans, E., Ibrahim, H., Coninck, B., Goossens, A. Pathogen Effectors: Exploiting the Promiscuity of Plant Signaling Hubs. **Trends In Plant Science**, v. 26, n. 8, p. 780-795, ago. 2021.

Conab. "Cana-de-Açúcar - 4° Levantamento Da Safra 2018/19 - Abril/2019.

Covarrubias, A.A., Cuevas-Velazquez, C.L., Romero-Pérez, P.S. et al. Structural disorder in plant proteins: where plasticity meets sessility. **Cell. Mol. Life Sci**. 74, p. 3119–3147. 2017. https://doi.org/10.1007/s00018-017-2557-2

Cui, L., Zheng, F., Wang, J., Zhang, C., Xiao, F., Ye, J., Li, C., Ye, Z., Zhang, J. MiR156a-targeted SBP-Box transcription factor SlSPL13 regulates inflorescence morphogenesis

by directly activating SFT in tomato. **Plant Biotechnology Journal**, n. 18, p. 1670-1682, 2020.

Cursi, D.E., Hoffmann, H.P., Barbosa, G.V.S. *et al.* History and Current Status of Sugarcane Breeding, Germplasm Development and Molecular Genetics in Brazil. **Sugar Tech**. 2021.

Depotter, J.R.L., Zuo, W., Hansen, M., Zhang, B., Xu, M., Doehlemann, G. Effectors with Different Gears: Divergence of Ustilago maydis Effector Genes Is Associated with Their Temporal Expression Pattern during Plant Infection. **J. Fungi** 7, 16. 2021. https://doi.org/10.3390/jof7010016

Fan, J., Guo, X.Y., Li, L., Huang, F., Sun, W.X., Li, Y., Huang, Y.Y., Xu, Y.J., Shi, J., Lei, Y., Zheng, A.P. and Wang, W.M. Infection of Ustilaginoidea virens intercepts rice seed formation but activates grain-filling-related genes. **J. Integr. Plant Biol.** 57, p. 577–590. 2015.

Fan, J., Yang, J., Wang, Y., Li, G., Li, Y., Huang, F., Wang, W. Current understanding on Villosiclava virens, a unique flower-infecting fungus causing rice false smut disease. **Mol. Plant Pathology**, v. 9, n. 17, p. 1321-1331, 2016.

Fornara, F., & Coupland, G. Plant Phase Transitions Make a SPLash. **Cell,** 138(4), p. 625–627. 2009. doi:10.1016/j.cell.2009.08.011

Gandikota, M., Birkenbihl, R., Hohmann, S., Cardon, G., Saedler, H., Huijser, P. The miRNA156/157 recognition element in the 3′ UTR of the Arabidopsis SBP-box gene SPL3 prevents early flowering by translation inhibition in seedlings. **The Plant journal**: for cell and molecular biology, v. 49, p. 683-93, 2007.

Gao, R., Gruber, M.Y., Amyot, L. *et al.* SPL13 regulates shoot branching and flowering time in *Medicago sativa* . **Plant Mol Biol** 96, p. 119–133. 2018. https://doi.org/10.1007/s11103-017-0683-8

He, J., Xu, M., WIillmann, M., Mccormick, K., Hu, T., Yang, L., Starker, C., Voytas, D., Meyers, B., Poethig, R. Threshold-dependent repression of SPL gene expression by miR156/miR157 controls vegetative phase change in Arabidopsis thaliana. **Plos Genetics**, v. 14, n. 4, p. 1-28, 19 abr. 2018. http://dx.doi.org/10.1371/journal.pgen.1007337.

Heinrichs, R., Otto, R., Magalhães, A., Meirelles, G. Importance of Sugarcane in Brazilian and World Bioeconomy. **Knowledge-Driven Developments In The Bioeconomy**, p. 205-217, 2017. http://dx.doi.org/10.1007/978-3-319-58374-7_11.

Hogenhout, S. A., Van der Hoorn, R. A. L., Terauchi, R., & Kamoun, S. Emerging Concepts in Effector Biology of Plant-Associated Organisms. **Molecular Plant-Microbe Interactions**, 22(2), p. 115–122. 2009. doi:10.1094/mpmi-22-2-0115

Korbie, D. J., Mattick, J. S. Touchdown PCR for increased specificity and sensitivity in PCR amplification. **Nature Protocols**, 3(9), p. 1452–1456. 2008. doi:10.1038/nprot.2008.133

León-Ramírez, C., Cabrera-Ponce, J., et al. Infection of alternative host plant species by *Ustilago maydis*. **New Phhytologist** 164, p. 337-346. 2004.

Li, B., Zhao, Y., Wang, S., Zhang, X., Wang, Y., Shen, Y., Yuan, Z. Genome-wide identification, gene cloning, subcellular location and expression analysis of SPL gene family in *P. granatum L.* **BMC Plant Biology.** p. 1-19, 2021.

Li, D., Zhang, Ha., Mou, M., Chen, Y., Xiang, S., Chen, L., Yu, D. Arabidopsis Class II TCP Transcription Factors Integrate with the FT–FD Module to Control Flowering. **Plant Physiology**, v. 181, n. 1, p. 97-111, 24 jun. 2019. http://dx.doi.org/10.1104/pp.19.00252.

Ling, H., Fu, X., Huang, N., et al. A sugarcane smut fungus effector simulates the host endogenous elicitor peptide to suppress plant immunity. **New Phytologist**, p. 1-15. 2021. doi: 10.1111/nph.17835.

Liu, Y., Aslam, M., Yao, LA. et al. Genomic analysis of *SBP* gene family in *Saccharum spontaneum* reveals their association with vegetative and reproductive development. **BMC Genomics** 22, 767. 2021. https://doi.org/10.1186/s12864-021-08090-3

Liu, Y., Li, P., Fan, L. et al. The nuclear transportation routes of membrane-bound transcription factors. **Cell Commun Signal** 16, 12. p. 1-9. 2018. https://doi.org/10.1186/s12964-018-0224-3

Maia, T., Rody, H., Bombardelli, R., et al. A bacterial type three secretion-based delivery system for functional characterization of *Sporisorium scitamineum* plant immune suppressing effector proteins. **Phytopathology**. 2021. (in press).

Marques, JPR., Hoy, JW., Appezzato-da-Glória, B., Viveros, AFG., Vieira, MLC., Baisakh, N. Sugarcane Cell Wall-Associated Defense Responses to Infection by *Sporisorium scitamineum*. **Front. Plant Sci.** 9:698. p. 1-14. 2018. doi: 10.3389/fpls.2018.00698

Martin, K., Kopperud, K., Chakrabarty, R., Banerjee, R., Brooks, R., Goodin, MM. Transient expression in *Nicotiana benthamiana* fluorescent marker lines provides enhanced definition of protein localization, movement and interactions in planta. **Plant J** 59(1), p. 150-162. 2009. doi: 10.1111/j.1365-313X.2009.03850.x.

Miller, K. E., Kim, Y., Huh, W.-K., & Park, H.-O. Bimolecular Fluorescence Complementation (BiFC) Analysis: Advances and Recent Applications for Genome-Wide Interaction Studies. **Journal of Molecular Biology**, 427(11), p. 2039–2055. 2015. doi:10.1016/j.jmb.2015.03.005

Murashige, T., Skoog, F. A revised medium for rapid growth and bioassays with tabacco tissue cultures. **Physiologia Plantarum**, v.15, p.473-97, 1962.

Peters, L. P., Carvalho, G., Vilhena, M. B., Creste, S., Azevedo, R., Monteiro-Vitorello, C. B. Functional analysis of oxidative burst in sugarcane smut-resistant and -susceptible genotypes. **Planta**, v. 245, n. 4, p. 749-764, 21 dez. 2016. http://dx.doi.org/10.1007/s00425-016-2642-z.

Qi T., Guo J., Liu P., He F., Wan C., Islam M.A., Tyler B.M., Kang Z., and Guo J. Stripe Rust Effector PstGSRE1 Disrupts Nuclear Localization of ROS-Promoting Transcription Factor TaLOL2 to Defeat ROS-Induced Defense in Wheat. **Mol. Plant**. 12, p. 1624–1638. 2019.

Que., et al. Genome sequencing of *Sporisorium scitamineum* provides insights into the pathogenic mechanisms of sugarcane smut. **BMC Genomics** 15:996. 2014.

Rozen, S. & Skaletsky H. Primer3 on the WWW for general users and for biologist programmers BT- Bioinformatics Methods and Protocols." In eds. Stephen Misener and Stephen A Krawetz. Totowa, NJ: Humana Press, 365–86. 1999. https://doi.org/10.1385/1-59259-192-2:365.

Schaker, P.D.C., Palhares, A.C., Taniguti, L.M., Peters, L.P., Creste, S., Aitken, K.S., Van Sluys, M., Kitajima, J.P., Vieira, M.L.C., Monteiro-Vitorello, C.B. RNAseq Transcriptional Profiling following Whip Development in Sugarcane Smut Disease. **Plos One**, v. 11, n. 9, p. 1-21, 1 set. 2016. http://dx.doi.org/10.1371/journal.pone.0162237.

Schaker, PDC., Peters, LP., Catald,i TR., Labate, CA., Caldana, C., Monteiro-Vitorello, CB. Metabolome Dynamics of Smutted Sugarcane Reveals Mechanisms Involved in Disease Progression and Whip Emission. **Front. Plant Sci.** 8:882. p. 1-17. 2017. doi: 10.3389/fpls.2017.00882

Schmitz, L., McCotter, S., Kretschmer, M., et al. Transcripts and tumors: regulatory and metabolic programming during biotrophic phytopathogenesis. **F1000Research** 7(F1000 Faculty Rev):1812 ( 2018. https://doi.org/10.12688/f1000research.16404.1

Silva, P. A., Silva, J. C., Caetano, H. D., Machado, J. P., Mendes, G. C., Reis, P. A., Brustolini, O. J., Dal-Bianco, M. and Fontes, E. P. Comprehensive analysis of the endoplasmic reticulum stress response in the soybean genome: conserved and plant-specific features. **BMC Genomics**. 16, 783. 2015.

Silva, G., Silva, E., Correa, J., Vicente, M., Jiang, N., Notini, M., Junior, A., Jesus, F., Castilho, P., Carrera, E. Tomato floral induction and flower development are orchestrated by the interplay between gibberellin and two unrelated micro RNA -controlled modules. **New Phytologist**, v. 221, n. 3, p. 1328-1344, 19 out. 2018. http://dx.doi.org/10.1111/nph.15492.

Tabassum, N., Blilou, I. Cell-to-Cell Communication During Plant-Pathogen Interaction. **Molecular Plant-Microbe Interactions**, p. 1-42, 19 out. 2021.

Teixeira-Silva, N., Schaker, P., Rody, H., Maia, Thiago., Garner, M., et al. Leaping into the Unknown World of *Sporisorium scitamineum* Candidate Effectors. **Journal Of Fungi**, v. 6, n. 4, p. 1-26, 4 dez. 2020. http://dx.doi.org/10.3390/jof6040339.

Vazquez-Vilar, M., Quijano-Rubio, A., Fernandez-del-Carmen, A., Sarrion-Perdigones, A., Ochoa-Fernandez, R., Ziarsolo, P., Blanca, J., Granell, A., Orzaez, D. GB3.0: a platform for plant bio-design that connects functional DNA elements with associated biological data. **Nucleic Acids Res**. 45, p. 2196-2209. 2017. https://doi.org/10.1093/nar/gkw1326

Vicente, C., Legaz, M.-E., Sánchez-Elordi, E. Physiological Basis of Smut Infectivity in the Early Stages of Sugar Cane Colonization. **J. Fungi** 7, 44. 2021. p. 1-18. https://doi.org/10.3390 /jof7010044

Xu, M., Hu, T., Zhao, J., Park, M., Earley, K., Wu, G., Yang, L., Poethig, R. Developmental Functions of miR156-Regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Genes in *Arabidopsis thaliana*. **Plos Genetics**, v. 12, n. 8, p. 1-29, 19 ago. 2016. http://dx.doi.org/10.1371/journal.pgen.1006263.

Yang, Y., Nicolas, M., Zhang, J., Yu, H., Guo, D., Yuan, R., et al. The TIE1 transcriptional repressor controls shoot branching by directly repressing BRANCHED1 in Arabidopsis. **PLoS Genet** 14(3): e1007296. 2008. https://doi.org/10.1371/ journal.pgen.1007296

Yuan, M., Jiang, Z., Bi, G. et al. Pattern-recognition receptors are required for NLR-mediated plant immunity. **Nature** 592, p. 105–109. 2021. https://doi.org/10.1038/s41586-021-03316-6

Zheng, J., Ma, Y., Zhang, M., Lyu, M., Yuan, Y., Wu, Binghua. Expression Pattern of FT/TFL1 and miR156-Targeted SPL Genes Associated with Developmental Stages in *Dendrobium catenatum*. **International Journal Of Molecular Sciences**, p. 1-14, jun. 2019.

# CHAPTER 3: CHARACTERIZATION OF SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE FAMILY

**Abstract**

The Squamosa-Promoter Binding protein-like gene family (*SPLs*) is involved in many aspects of plant growth and development, such as phase transition, flowering and also activation of the immune system. Recently, two independent studies successfully defined SPL genes of *Saccharum spontaneum* (17 in first and 30 in second). Most of the SPLs follow a gene organization presenting a disorder region, the SPL-box with two $Zn^2+$-binding sites and a putative conserved nuclear localization signal (NLS). Our objective was to perform *in silico* analysis to identify orthologs of SPLs in modern genotypes of sugarcane using sequences of *Arabidopsis*, tomato, sorghum, soy, *Zea mays* and *S. spontaneum*. We also analyzed previous RNAseq data to investigate the expression of *SPLs* genes. Our Maximum Likelihood (ML) phylogeny recovered well supported branches with bootstrap values above 75. Most of the paralogs clustered into subgroups. The multiple protein sequence alignment showed a similar region between the orthologs (290-390 amino acids), being a conserved region. The SPLs with a similar expression pattern among the treatments were grouped closely. Intermediate sugarcane genotype showed two differentially expressed SPLs after whip emission, while resistant and susceptible cultivars did not show differentially expressed SPLs. This difference may be due to the fact that the functionally related genes of *A. thaliana* were recovered in the same groups. Our sugarcane ortholog of *SPL13* clustered in the same group of *SPL13* of *A. thaliana*. In previous studies, orthologs of *SPL13* of *S. spontaneum* also clustered in the same group of *AtSPL13*. whip development, and a time-course study of gene expression from shortly after inoculation until whip development may reveal different patterns in *SPL13* expression.

**Keywords:** *Saccharum*, SPL, Phylogeny, Gene expression

## 1.Introduction

The development and growth of plants involve a precise and complex mechanism controlled by molecular signaling components that activate transcription factors in a specific order and time. Transcription factors (TF) interact with gene regulatory sequences inducing their expression. The genome of *Arabidopsis thaliana* contains over 25 families of transcription factors, representing more than 5% of all genes, of which 45% are from families specific to plants (Riechmann et al., 2000). Among plant-specific TFs, the Squamosa-Promoter Binding protein-like gene family (*SPLs*) are critical regulatory components of various roles in plant growth and development. Guo et al. (2008), using a genome-wide analysis of TF containing SPL-box, revealed that this family evolved from a common ancestor of green plants and underwent duplication events followed by sequence diversification in each lineage, including exon-intron loss processes. In addition, Zhang et al.

(2018) identified divergence of miR156/529 binding sites following the duplication event, and they suggested that this event might have contributed to the functional diversity of *SPL* genes.

The first definition of function related to an SPL TF described in *Antirrhinum majus* was to regulate the expression of MADS-box genes in early flower development (Klein et al., 1996). Since then, various reports have described functions in growth and development (Chen et al., 2010), including their role associated with stress tolerance in plants (Cui et al., 2014) and the activation of the immune defense against a pathogen (Padmanabhan et al., 2013). For instance, the *SPL6* from *Arabidopsis* functions in resistance against the bacterial pathogen *Pseudomonas syringae* expressing the AvrRps4 effector and positively modulates defense gene expression. Also, the authors showed an association of the TIR-NB-LRR (TNL) immune receptor N of *N. benthamiana* with SPL6 to activate the expression of defense-related genes against the TMV (Tobacco mosaic virus) in the presence of the defense-eliciting TMV-p50-U1 effector (Padmanabhan et al., 2013). In *Arabidopsis,* Xu et al. (2016) divided the SPLs according to function as follows: 1) *SPL2, SPL9, SPL10, SPL11, SPL13*, and *SPL15* as contributors to both the juvenile-to-adult vegetative transition and the vegetative-to-reproductive transition; 2) *SPL3, SPL4* and *SPL5* as promoters of the floral meristem identity transition; and 3) *SPL6* may be important for certain physiological processes. Previous work of our group that performed a Y2H assay (unpublished data) showed that the SPL13 interacts with the *Sporisorium scitamineum* candidate effector g5159 (Teixeira-Silva et al., 2020).

Genome-wide analysis of SPLs defined clusters of related sequences usually forming 6 to 8 groups. For instance, six groups were formed in the maize B73 genome (total of 31 *SPLs* genes) (Mao et al., 2016), eight groups in wheat (Zhu et al., 2020), eight groups in cotton (Cai et al., 2017), eight groups in *Jatropha curcas* (Yu et al, 2020), eight groups in *Brassica juncea* (Gao et al., 2019), and six groups in *Fragaria vesca* (Xiong et al., 2018). Recently, Feng et al., (2021) defined six clades for *Saccharum spontaneum* (Ss) using 17 SPLs in addition to the TFs of *Arabidopsis*, rice, and sorghum genes. Another group expanded the analysis to define 30 SPLs in *S. spontaneum*, detecting seven clades using the same group of references (Liu et al., 2021). According to the studies, SsSPLs showed a closer relationship with sorghum SPLs (85% to 96.7%) than with *Arabidopsis* and rice. However, the gene structure of *SsSPLs* was more diverse than sorghum and rice (Feng et al., 2021). The authors also suggested that *SsSPLs* originated from two common ancestors and went through the gain or loss of exons during the evolutionary process. Two gene pairs of SsSPL experienced segmental duplication: *SsSPL2/SsSPL14* and *SsSPL13/SsSPL13*, and another two went through tandem duplication: *SsSPL4/SsSPL6* and *SsSPL9/SsSPL10*. In *Arabidopsis,*

*SPL13* is also detected as a pair of tandemly organized genes named *SPL13A* and *B*. They also suggested that the duplication events contributed to SPL development in *S. spontaneum* (Feng et al., 2021). Most of the SPLs described followed the same gene organization for family members having a disorder region, the SPL-box with two Zn2+-binding sites: Cys-Cys-His-Cys and a putative conserved nuclear localization signal (NLS) (Birkenbihl et al., 2005). Only two SPLs described by Liu et al., (2021), SPL4 and 12, did not present the conserved NLS. Correlations among designating SPLs in various species are not straightforward. For example, there is no consensus in naming described by Liu et al., (2021) and Feng et al. (2021) for SPLs in *S. spontaneum*.

Although the two recent aforementioned attempts on elucidating SPLs within *S. spontaneum* genome, commercial sugarcane cultivars are interspecific hybrids of *S. spontaneum* and *S. officinarum* that may carry different collections of SPL orthologs. Therefore, to define SPL orthologs in modern genotypes of sugarcane (Souza et al., 2019; Cardoso-Silva et al., 2014, Schaker et al., 2016) and identify orthologs of the SPL13, we set up a series of in silico analysis using as references the SPL sequences of *Arabidopsis*, tomato, sorghum, soy, *Zea mays*, and *S. spontaneum.* We also examined previously obtained RNAseq data to investigate the expression of the *SPLs* genes.

## 2. Material and Methods

### 2.1 Orthologs and paralogs survey of SPLs

Sixteen SPL sequences were obtained for *Arabidopsis thaliana* from Preston and Hileman (2013). All protein sequences were then used as SPL queries during BLASTp searches (Supplementary Table 1) against the predicted proteins in the genomes of *Sorghum bicolor* (Sobic), *Zea mays* (GRMZ), *Oryza sativa* (LOC), *Glycine max* (Glyma), *Solanum lycopersicum* (Solyc), all the four obtained from Phytozome v12. They were all compiled in an excel sheet and in a blast format .txt dataset. In addition, also composing the BLAST database were the protein sequences from *Saccharum spontaneum* (Sspon) clone AP85-441 (Zhang et al., 2018), from the sugarcane cultivar SP80-3280 (evm) (Souza et al., 2019), and a set of 88,488 Open Reading Frames (ORFs) obtained from a collection of sugarcane *de novo* transcripts assembled by two independent experiments (comp and gg), comprising transcripts from six sugarcane cultivars (Cardoso-Silva et al., 2014, Schaker et al., 2016). These sequences were previously included in our in-house BLAST database. BLASTp searches used a cutoff of $e^{-05}$ and a minimum of 40% of identity and 80% of query coverage. Sequences that

passed BLASTp cutoffs (N=182) were declared as SPL orthologs and finally merged to form the SPLdataset (Supplementary File 1).

## 2.2 SPL Phylogeny

First, multiple amino acid alignment of SPLdataset was performed using the MAFFT server v7 (Katoh et al., 2019) with the parameters of "E-INS-i" (Very slow; recommended for <200 sequences with multiple conserved domains and long gaps; 2 iterative cycles only) and "Leave gappy regions" (Not recommended for >~1,000 sequences) set. Then, Maximum Likelihood (ML) phylogenetic relationships among amino acid sequences of SPLdataset were inferred using the IQ-Tree software v2.1.3 (Nguyen et al., 2015). The best-fit model of molecular evolution for each tree was selected by the ModelFinder software (Kalyaanamoorthy et al., 2017) with the parameter -m MFP in IQ-Tree, and choosing the model that minimizes the Bayesian information criterion (BIC) score. BIC selected the model Q. mammal+F+R6. In addition, the ultrafast bootstrap approximation (UFBoot) (Hoang et al., 2018) parameter was set as -B 1000 to specify the number of bootstraps replicates, alongside the hill-climbing nearest neighbor interchange (NNI) search with parameter -bnni to reduce the risk of overestimating branch supports. The consensus trees were visualized and edited with iTOL software v6. Finally, Consensus disorder and SBP domains predicted by InterproScan v5.53-87.0 (Jones et al., 2014) were plotted alongside sequences in the tree.

## 2.3 Sugarcane SPL expression

The expression profiles were previously obtained by Rody et al. (2019). However, only eight sugarcane sequences comprising orthologs of the six major SPL clades were selected for this study. The data described in Rody et al. (2019) composed: 1) RB925354 (intermediate resistance) 200 days after infection (dai), immediately after whip emission, obtained from Taniguti et al. (2015); 2) SP80-3280 (resistant) and IAC66-6 (susceptible) 48 hours after inoculation and respective controls of mock-inoculated plants. Here we only used the SPL data from the complete data analysis obtained using TMM (Trimmed Mean of M-values) normalized CPM (counts per million) values, as well as Log2FoldChange (inoculated/control).

## 3. Results
## 3.1 ML Phylogeny of SPLs

Our ML phylogeny provided a tree with most of the well supported branches as indicated by the bootstrap values above 75 (Figure1). Different SPL proteins were divided into orthologs and paralogs from *Arabidopsis thaliana, Sorghum bicolor, Zea mays, Oryza sativa, Glycine max, Solanum lycopersicum, Saccharum spontaneum* clone AP85-441, and a set of *de novo* transcripts from different sugarcane cultivars. SPL proteins tend to form groups.

All SPL proteins were recovered into 9 major groups, with the majority presenting a high nodal bootstrap support. Furthermore, most of the paralogs, including the ones from different varieties of sugarcane, were clustered into subgroups. Genes that encode a protein length between 100 and 500 amino acids were clustered in the same groups, separately from the protein sequences with approximately 1000 amino acids.

The paralog of SPL13 from *A. thaliana* and sugarcane were clustered within the same group. Nested within a major clade harboring two *A.* thaliana SPL13 (AT5G50570, AT5G50670) were sequences from *S. spontaneum* (N=4), sugarcane cultivars comp (N=1) and evms (N=4). Closely related to the subclade of *A. thaliana* SPL13 were the sequences from *G. max* (N=4) forming a separated subclade. Bootstrap also indicated two sequences of *Z. mays* forming a separated subclade.

*S. lycopersicum* had two sequences within the major clade of SPL13. One sequence grouped closer to *A. thaliana* sequences, whereas the other grouped with sugarcane cultivars evms and *S. spontaneum* sequences. *S. bicolor* also had two SPL13 sequences. One grouped in a separate subclade with sugarcane evms (N=2) and comp (N=1) sequences, while the other was found closely related to sequences of *O. sativa* (N=1) and *Z. mays* (N=2).
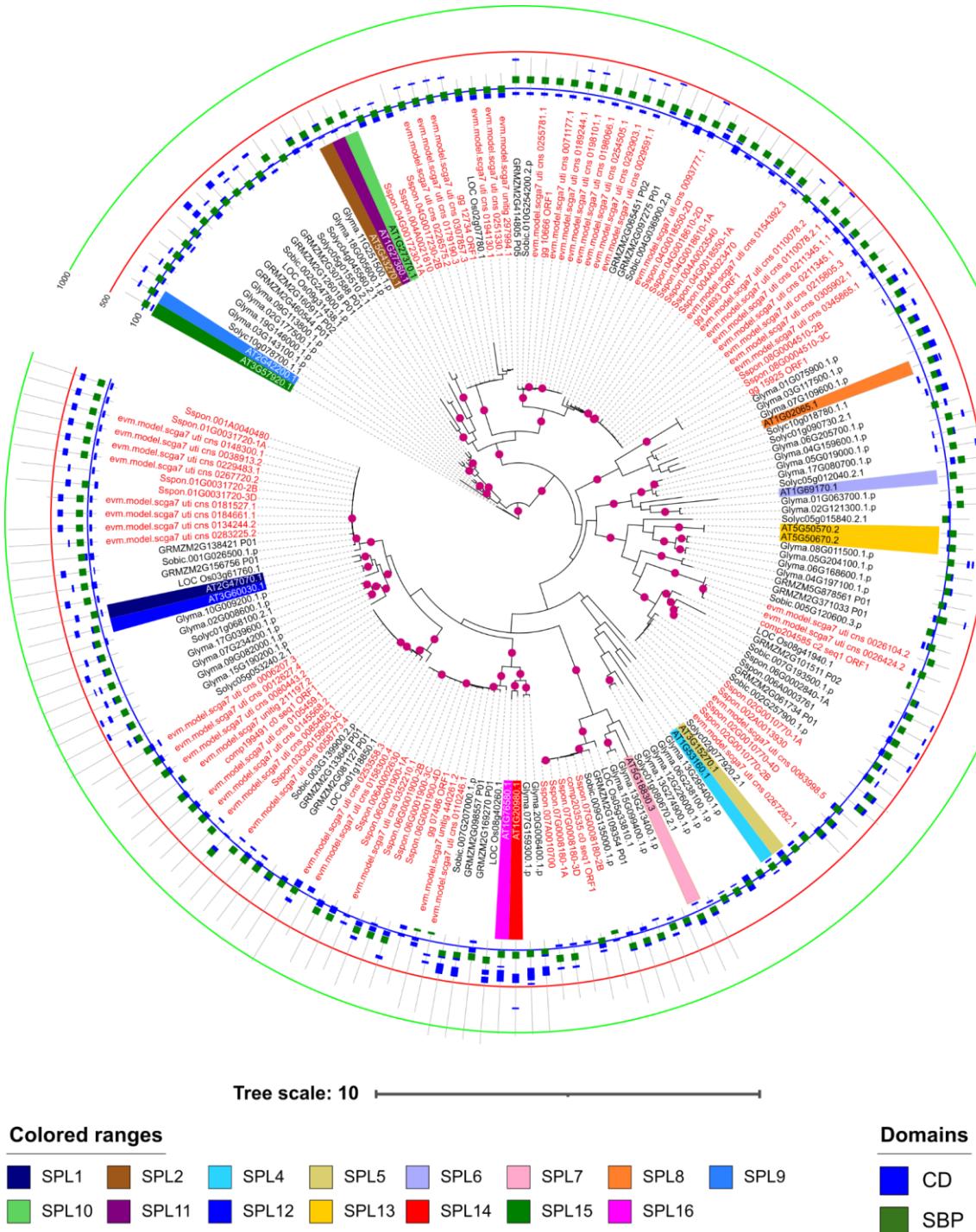
**Figure 1.** Unrooted Maximum Likelihood Phylogenetic consensus trees as calculated by the IQ-Tree software showing phylogenetic relationships among the multiple amino acid sequence alignment of 182 sequences in SPLdataset (see methods). Branch lengths are drawn to scale. Colored ranges represent *A. thaliana* SPLs as obtained from the TAIR11 database (https://www.arabidopsis.org/, accessed on 29/11/2021). Nodal support values are given as local bootstraps above the branches indicated by dark pink circles if equal to 100. Conserved disorder (CD) and SBP (Squamosa-Promoter Binding Protein) domains are drawn alongside leaves in the tree. Outermost circles indicate length of sequences.

The multiple protein sequence alignment of *Arabidopsis thaliana*, *sorghum bicolor, saccharum spontaneum* AP85-441 and sugarcane cultivars showed that it was possible to

align many regions between the orthologs. The region between 290-390 amino acids is similar in the majority of orthologs, being a conserved region (Figure 2).
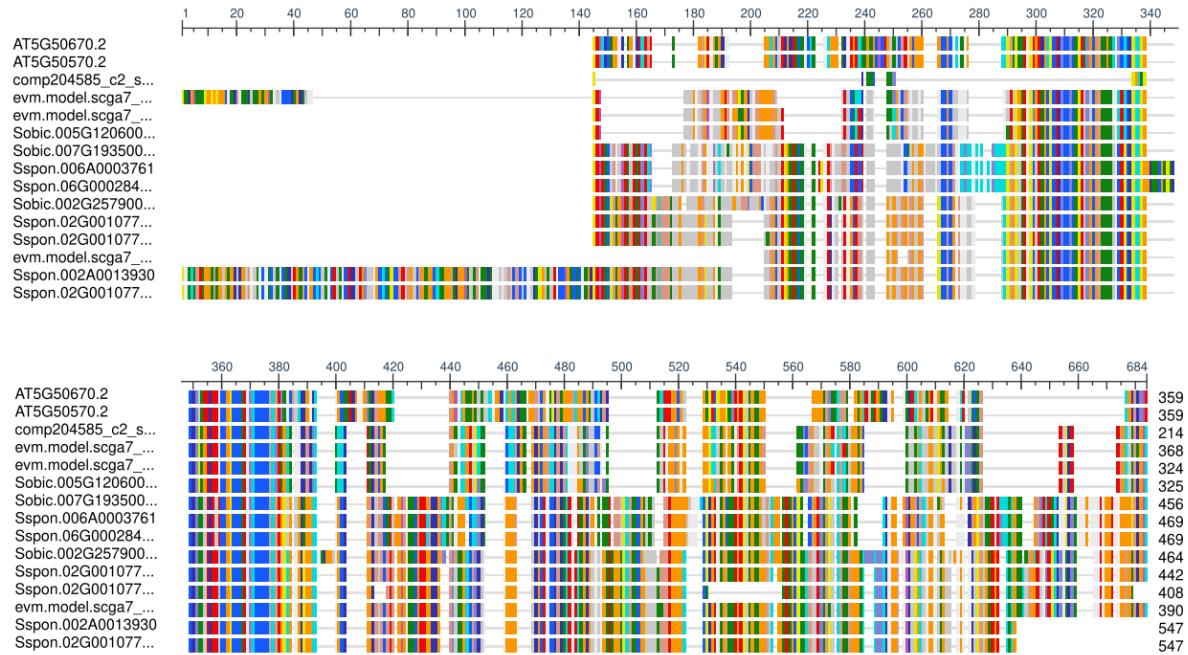


**Figure 2.** Multiple sequence alignment (MSA) of SPL13 ortholog protein sequences among *Arabidopsis thaliana*, *Sorghum bicolor*, *Saccharum spontaneum* AP85-441, and sugarcane cultivars (comp and evms). MSA was generated with the MAFFT web server. Amino acids are depicted as Rasmol Amino Acid colors.

In order to compare the sequences described by Feng et al., (2021) and Lui et al., (2021), we used the SGD: Saccharum Genome Database (http://sugarcane.zhangjisenlab.cn/sgd/html/index.html). BlastP analysis using the sugarcane sequences as queries provided the correspondent SPL (best hit) according to Liu et al., (2021) nomenclature and phylogenetic analysis (Table 1). Thirty-one sequences of *S. spontaneum* were used to build our analysis. They were recovered in BlastP queries. Our set included sequences used by the two groups.

**Table 1.** SPLs of sugarcane modern genotypes orthologs of *S. spontaneum* described in Liu et al., (2021).

| Sugarcane Modern Genotypes | *Arabidopsis* SPL | Liu et al. (2021) SPL | clade | % identity |
|---|---|---|---|---|
| evm.model.scga7_uti_cns_0181527.1 | AtSPL5 | SsSBP1 | II | 99% |
| evm.model.scga7_uti_cns_0110078.2.1 | AtSPL11 | SsSBP11 | V | 98% |
| evm.model.scga7_uti_cns_0105459.1 | AtSPL12 | SsSBP8 | IV | 91% |
| evm.model.scga7_uti_cns_0283225.2 | AtSPL5 | SsSBP1 | II | 99% |
| evm.model.scga7_uti_cns_0058773.4 | AtSPL16 | SsSBP8 | IV | 91% |
| evm.model.scga7_uti_cns_0229483.1 | AtSPL16 | SsSBP1 | II | 99% |
| evm.model.scga7_uti_cns_0148300.1 | AtSPL16 | SsSBP1 | II | 99% |
| evm.model.scga7_uti_cns_0215805.3 | AtSPL11 | SsSBP11 | V | 94% |
| evm.model.scga7_uti_cns_0085485.1 | AtSPL12 | SsSBP8 | IV | 91% |
| evm.model.scga7_uti_cns_0267720.2 | AtSPL16 | SsSBP1 | II | 99% |
| evm.model.scga7_uti_cns_0006207.3 | AtSPL16 | SsSBP8 | IV | 91% |
| evm.model.scga7_uti_cns_0071177.1 | AtSPL5 | SsSBP29 | V | 91% |
| evm.model.scga7_uti_cns_0345865.1 | AtSPL11 | SsSBP11 | V | 98% |
| evm.model.scga7_uti_cns_0026104.2 | AtSPL13 | SsSBP21 | VII | 95% |
| evm.model.scga7_uti_cns_0184661.1 | AtSPL5 | SsSBP1 | II | 98% |
| evm.model.scga7_uti_cns_0093777.1 | AtSPL11 | SsSBP12 | V | 95% |
| evm.model.scga7_uti_cns_0352210.1 | AtSPL14 | SsSBP20 | II | 95% |
| evm.model.scga7_uti_cns_0145568.2 | AtSPL12 | SsSBP8 | IV | 91% |
| evm.model.scga7_uti_cns_0110078.2 | AtSPL11 | SsSBP12 | V | 95% |
| evm.model.scga7_uti_cns_0029591.1 | AtSPL5 | SsSBP29 | V | 90% |
| evm.model.scga7_uti_cns_0134244.2 | AtSPL5 | SsSBP1 | II | 99% |
| evm.model.scga7_uti_cns_0255781.1 | AtSPL5 | SsSBP29 | V | 87% |
| evm.model.scga7_uti_cns_0179190.3 | AtSPL5 | SsSBP10 | V | 99% |
| evm.model.scga7_uti_cns_0226575.2 | AtSPL5 | SsSBP10 | V | 99% |
| evm.model.scga7_uti_cns_0254505.1 | AtSPL5 | SsSBP29 | V | 91% |
| evm.model.scga7_uti_cns_0267282.1 | AtSPL5 | SsSBP18 | VII | 87% |
| evm.model.scga7_uti_cns_0154392.3 | SPL11 | SsSBP11 | V | 98% |
| evm.model.scga7_uti_cns_0292903.1 | SPL5 | SsSBP29 | V | 90% |
| evm.model.scga7_uti_cns_0012627.4 | SPL12 | SsSBP8 | IV | 84% |
| evm.model.scga7_uti_cns_0080443.2 | SPL12 | SsSBP8 | IV | 91% |
| evm.model.scga7_uti_cns_0307857.3 | SPL5 | SsSBP10 | V | 97% |
| evm.model.scga7_uti_cns_0253558.3 | SPL12 | SsSBP20 | II | 87% |
| evm.model.scga7_uti_cns_0211345.1 | SPL11 | SsSBP12 | V | 95% |
| evm.model.scga7_uti_cns_0194130.1 | SPL5 | SsSBP10 | V | 99% |
| evm.model.scga7_uti_cns_0305902.1 | SPL11 | SsSBP11 | V | 98% |
| evm.model.scga7_unitig_211197.2 | SPL1 | SsSBP9 | IV | 97% |
| evm.model.scga7_uti_cns_0038913.2 | SPL16 | SsSBP1 | II | 99% |
| evm.model.scga7_uti_cns_0211345.1.1 | SPL11 | SsSBP12 | V | 95% |
| evm.model.scga7_uti_cns_0158300.4 | SPL1 | SsSBP20 | II | 87% |
| evm.model.scga7_uti_cns_0063998.5 | SPL13 | SsSBP2 | VII | 98% |
| evm.model.scga7_uti_cns_0198101.1 | SPL5 | SsSBP29 | V | 89% |
| evm.model.scga7_unitig_297994.1 | SPL5 | SsSBP10 | V | 99% |
| evm.model.scga7_uti_cns_0198066.1 | SPL5 | SsSBP29 | V | 91% |
| evm.model.scga7_uti_cns_0026424.2 | SPL13 | SsSBP23 | VII | 95% |
| evm.model.scga7_uti_cns_0189244.1 | SPL5 | SsSBP29 | V | 89% |
| evm.model.scga7_uti_cns_0251330.1 | SPL5 | SsSBP10 | V | 99% |
| evm.model.scga7_unitig_440261.2 | SPL12 | SsSBP20 | II | 87% |

## 3.2 Sugarcane SPL expression

The SPLs that showed a similar pattern of expression among the treatments were grouped closely (Figure 3). Intermediate smut-resistant cultivar RB925345 sugarcane showed two differentially expressed (DE) SPLs after whip emission (white stars in Figure 3: gg_04693 SPL11 and gg_12734 SPL5), while resistant (SP80-3280) and susceptible (IAC66-6) cultivars did not show DE SPLs.
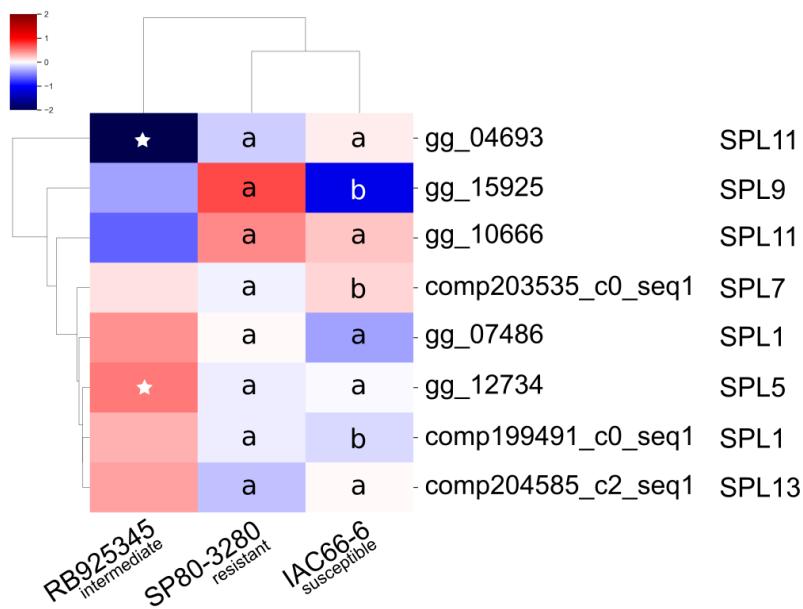


**Figure 3.** Clustered heatmap showing the expression profiles of eight sugarcane sequences comprising orthologs of six major SPL clades uncovered by ML Phylogeny. Differential expression profiles are shown as Log2FoldChange (inoculated/control) for transcriptomes of three sugarcane genotypes (RB925345, SP80-3280, and IAC66-6) with contrasting degrees of resistance to smut. Stars indicate significant differential expression at $p < 0.05$. In addition, different letters inside maps show ANOVA statistically significant differences at $p < 0.05$ between the means of three TMM normalized CPM values of replicates as calculated by EdgeR software.

Although RNAseq data did not uncover DE SPLs in the smut- resistant and susceptible sugarcane cultivars, ANOVA showed the two genotypes as having expression differences in terms of CPM values of inoculated treatments for the three genes of gg_15925 SPL9, comp203535_c0_seq1 SPL7, and comp199491_c0_seq1 SPL1. Therefore, our analyzed data suggested divergent expression patterns between cultivars. For example, the observed expression for gg_15925 SPL9 that was up-regulated in smut-resistant sugarcane (red color) whereas down-regulated (blue color) in the smut-susceptible.

## 4. Discussion

Our phylogeny clade was divided into different major groups formed by distinct species, and this divergence among groups can indicate the diversity of functions that SPLs proteins hold among different species. The distinct *A. thaliana* SPL sequences we used in this study from Preston and Hileman (2013) were recovered in the same groups, as shown in their results. As of today, we know that these genes might be functionally related. For instance, *AtSPL9* and *AtSPL15*, which were in the same clade, are involved in shoot development, having overlapping functions (Xu et al., 2016; Schwarz et al., 2008). Also, another clade formed by *AtSPL2, AtSPL10*, and *AtSPL11* contributed to the developmental transition in conjunction with *AtSPL9* and *AtSPL15*, although these last two play dominant roles. Moreover, *AtSPL4* and *AtSPL5*, recovered together, promote meristem identity transition (Xu et al., 2016). Finally, *AtSPL1* and *AtSPL12*, also clustered, share substantial sequence similarity, including intron positions (Cardon et al., 1999). Indeed, *AtSPL1* and *AtSPL12* are a pair of duplicated SBP-box genes. Another pair of duplicated genes that were also forming a clade is *AtSPL14* and *AtSPL16* (Guo et al., 2008) (Figure 1).

We defined the ortholog of *AtSPL13* in the set of sugarcane proteins (evm.model.scga7_uti_cns_0026104.2) and in the *S. spontaneum* dataset provided by Liu et al., (2021). In their study, both *AtSPL13* and *SsSBP21* clustered in the same group. We checked and SsSBP21 is the ortholog of *AtSPL13*. Similarly, our analysis showed comparable results because our orthologs of *SPL13* in sugarcane (evm.model.scga7_uti_cns_0026104.2 and evm.model.scga7_uti_cns_0063998.5) were closely related to *SPL13* of *A. thaliana*. Unfortunately, sequences used by Feng et al., (2021) were not available for searching. However, for certain the *SPL13* described by Feng is not an ortholog of *AtSPL1*3, since the authors described the sequence as not having the canonical NLS.

Moreover, since our protein sequence alignment showed a conserved region between the orthologs (in 290 and 390 amino acids), it is possible to infer that the SPL of interest of this study (SPL13) is present in sugarcane. Future functional experiments can be performed in order to investigate how the SPL13 acts in infected sugarcane (Figure 2).

Although there are no differently expressed genes in resistant and susceptible varieties, their expression patterns differ *in SPL9, SPL7* and *SPL1* (marked by different letters a and b in Figure 3). These differences are given between the group of infected SP and the group of infected IAC, inoculated 48 hours after inoculation. We speculate that the differences observed for the RB925345 genotype are related to whip emission, since Schaker et al. (2017)

described a series of molecular events involving vegetative to reproductive stages transition. Whip development occurs instead of any flower structure from the apex of sugarcane plants. Liu et al., (2021) studied the spatiotemporal expression patterns of SsSBP genes using RNA-seq data of different organs and tissues. In the case of *SsSBP21 (AtSPL13)*, the data sustained low expression levels in leaf development and female reproductive organs of *S. spontaneum*.

Although RNAseq data did not uncover differential expression of *SPL13* 48 hours after inoculation of *S. scitamineum* in mock-inoculated plants, we did not abandon the hypothesis that *SPL13* may influence disease progression. The more sensitive analysis of RT-qPCR might help understand further the expression profiles of *SPL13* in sugarcane. A time-course of gene expression evaluation from shortly after inoculation until whip development may reveal a specific moment where *SLP13* may be relevant.

## References

Cai, C., Guo, W., & Zhang, B. Genome-wide identification and characterization of SPL transcription     factor family and their evolution and expression profiling analysis in cotton. **Scientific Reports,** 8(1). p. 1-13. 2018. doi:10.1038/s41598-017-18673-4

Cardon, Guillermo; Hohmann, Susanne; Klein, Joachim; Nettesheim, Klaus; Saedler, Heinz; Huijser, Peter. Molecular characterisation of the Arabidopsis SBP-box genes. **Gene**, Cologne, v. 237, p. 91-107, jul. 1999.

Cardoso-Silva, C.B., Costa, E.A., Mancini, M.C., et al. De Novo Assembly and Transcriptome Analysis of Contrasting Sugarcane Varieties. **Plos One**, v. 9, n. 2, p. 1-10, 11 fev. 2014. http://dx.doi.org/10.1371/journal.pone.0088462.

Chen, X., Zhang, Z., Liu, D., Zhang, K., Li, A., et al. SQUAMOSA promoter binding protein-like transcription factors: star players for plant growth and development. **J Integr Plant Biol** 52, p. 946–951. 2010.

Cui, L., Shan, J., Shi, M., Gao, J. P., Lin, H. X. The miR156-SPL9-DFR pathway coordinates the relationship between development and abiotic stress tolerance in plants. **Plant J**. 80, p. 1108–1117. 2014.

Feng, X., Wang, Y., Zhang, N. et al. Systematic Identification, Evolution and Expression Analysis of the SPL Gene Family in Sugarcane (Saccharum spontaneum). **Tropical Plant Biol**. 14, 313–328. 2021. https://doi.org/10.1007/s12042-021-09293-4.

Gao, J., Peng, H., Chen, F., Liu, Y., Chen, B., Li, W. Genome-wide identification and characterization, phylogenetic comparison and expression profiles of SPL transcription factor family in B. juncea (Cruciferae). **PLoS ONE** 14(11). p. 1-17. 2019. https://doi.org/10.1371/journal. pone.0224704

Guo, A.-Y., Zhu, Q.-H., Gu, X., Ge, S., Yang, J., & Luo, J. Genome-wide identification and evolutionary analysis of the plant specific SBP-box transcription factor family. **Gene**, 418(1-2), p. 1–8. 2008. doi:10.1016/j.gene.2008.03.016

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., and Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. **Mol. Biol. Evol.** 35, p. 518–522. 2018. doi:10.1093/molbev/msx281.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., et al. InterProScan 5: genome-scale protein function classification. **Bioinformatics** 30, p. 1236–1240. 2014. doi:10.1093/bioinformatics/btu031.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., and Jermiin, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. **Nat. Methods** 14, p. 587–589. 2017. doi:10.1038/nmeth.4285.

Katoh, R., Yamada. (Briefings in Bioinformatics 20:1160-1166) Access the recommendation on F1000MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. 2019.

Klein, J., Saedler, H., Huijser, P. A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA. **Mol Gen Genet**. 250(1), p. 7–16. 1996. https://doi. org/10.1007/BF02191820.

Liu, Y., Aslam, M., Yao, LA. et al. Genomic analysis of *SBP* gene family in *Saccharum spontaneum* reveals their association with vegetative and reproductive development. **BMC Genomics** 22, 767. 2021. https://doi.org/10.1186/s12864-021-08090-3

Mao, H., Yu, L., Li, Z., Yan, Y., Han, R., Liu, H., Ma, M. Genome-wide analysis of the SPL family transcription factors and their responses to abiotic stresses in maize. **Plant Gene** 6: 12, p. 1-12. 2016.

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. **Mol. Biol. Evol.** 32, 268–274. 2015. doi:10.1093/molbev/msu300.

Padmanabhan, MS., Ma, S., Burch-Smith, TM., Czymmek, K., Huijser, P., et al. Novel Positive Regulatory Role for the SPL6 Transcription Factor in the NTIR-NB-LRR Receptor-Mediated Plant Innate Immunity. PLoS Pathog 9(3): e1003235. 2013. doi:10.1371/journal.ppat.1003235

Preston, J., Hileman, L. Functional Evolution in the Plant SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) Gene Family. **Frontiers In Plant Science**, v. 4, n. 80, p. 1-13, 2013. http://dx.doi.org/10.3389/fpls.2013.00080.

Riechmann, J.L. and O.J. Ratcliffe. A genomic perspective on plant transcription factors. Curr. **Opinion in Plant Biol**, 3. p. 423-434. 2000.

Rody, HVS., Bombardelli, RGH., Creste, S., Camargo, LEA., Van Sluys, MA., Monteiro-Vitorello, CB. Genome survey of resistance gene analogs in sugarcane: genomic features and differential expression of the innate immune system from a smut-resistant genotype. **BMC Genomics**. 20(1):809. 2019. doi: 10.1186/s12864-019-6207-y. PMID: 31694536; PMCID: PMC6836459.

Schaker, P.D.C., Palhares, A.C., Taniguti, L.M., Peters, L.P., Creste, S., Aitken, K.S., Van Sluys, M., Kitajima, J.P., Vieira, M.L.C., Monteiro-Vitorello, C.B. RNAseq Transcriptional Profiling following Whip Development in Sugarcane Smut Disease. **Plos One**, v. 11, n. 9, p. 1-21, 1 set. 2016. http://dx.doi.org/10.1371/journal.pone.0162237.

Schwarz, S., Grande, AV. Bujdoso, N., Saedler, H., Huijser, P. The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis. **Plant Mol Biol** 67, p. 183-195. 2008.

Souza, G., Van Sluys, M., Lembke, C., Lee, H., Margarido, G., Hotta, C., Gaiarsa, J., Diniz, A., Oliveira, M., Ferreira, S. Assembly of the 373k gene space of the polyploid sugarcane genome reveals reservoirs of functional diversity in the world's leading biomass crop. **Gigascience**, v. 8, n. 12, p. 1-18, 29 nov. 2019. http://dx.doi.org/10.1093/gigascience/giz129.

Teixeira-Silva, N., Schaker, P., Rody, H., Maia, Thiago., Garner, M., et al. Leaping into the Unknown World of *Sporisorium scitamineum* Candidate Effectors. **Journal Of Fungi**, v. 6, n. 4, p. 1-26, 4 dez. 2020. http://dx.doi.org/10.3390/jof6040339.

Xiong, J., Zheng, D., Zhu, H.-Y., Chen, J.-Q., Na, R., & Cheng, M. Z.-M. Genome-wide identification and expression analysis of SPL gene family in woodland strawberry *Fragaria vesca.* **Genome.** 2018. doi:10.1139/gen-2018-0014

Xu, M., Hu, T., Zhao, J., Park, MY., Earley, KW., Wu, G., et al. Developmental functions of miR156-regulated SQUAMOSA PROMOTER BINDING PROTEINLIKE (SPL) genes in *Arabidopsis thaliana*. **PLoS Genet.** 12(8):e1006263, p. 1-29. 2016.

Yu, N., Yang, J.-C., Yin, G.-T., Li, R.-S., & Zou, W.-T. Genome-wide characterization of the SPL gene family involved in the age development of *Jatropha curcas*. **BMC Genomics**, 21(1). p. 1-14. 2020. doi:10.1186/s12864-020-06776-8

Zhang, S., Ling, L. Diversification of SQUAMOSA promoter binding protein-like (SPL) genes by changes of miR156/529 binding sites in land plants. **Plant Gene**, v. 14, p. 55-63, jun. 2018. http://dx.doi.org/10.1016/j.plgene.2018.04.007.

Zhu, T., Liu, Y., Ma, L., Wang, X., Zhang, D., Han, Y. Ma, L. Genome-wide identification, phylogeny and expression analysis of the SPL gene family in wheat. **BMC Plant Biology**, 20(1), p. 1-14. 2020. doi:10.1186/s12870-020-02576-0

# SUPPLEMENTARY MATERIAL

**Supplementary Figure 1.** Maps of vectors used in cloning procedures



Copy of pUPD2 (2690 bp)
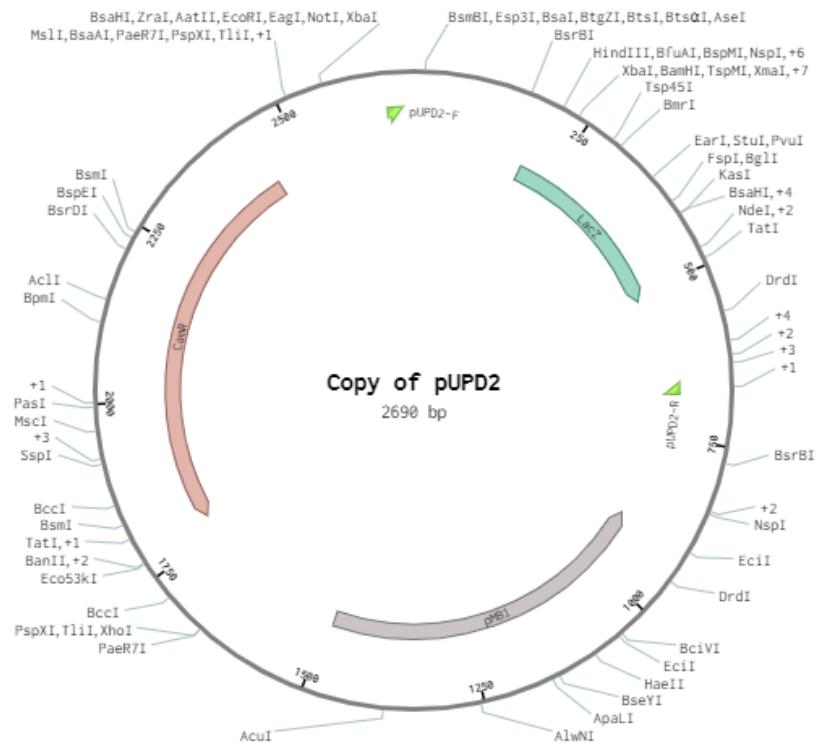
## Copy of pICSL86900_OD (7384 bp)
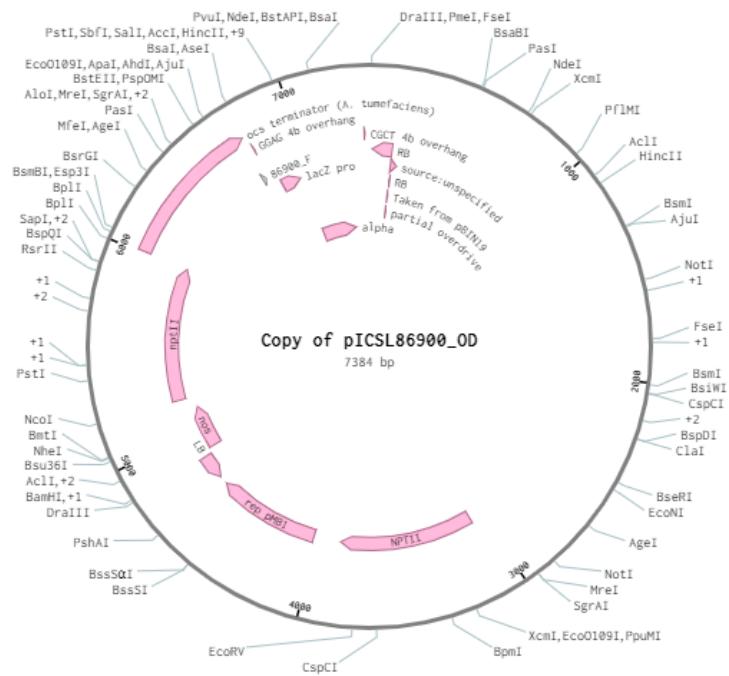


## pSITE-BiFC-cEYFP-C1 (5470 bp)

pSITE-BiFC-nEYFP-C1 (5791 bp)
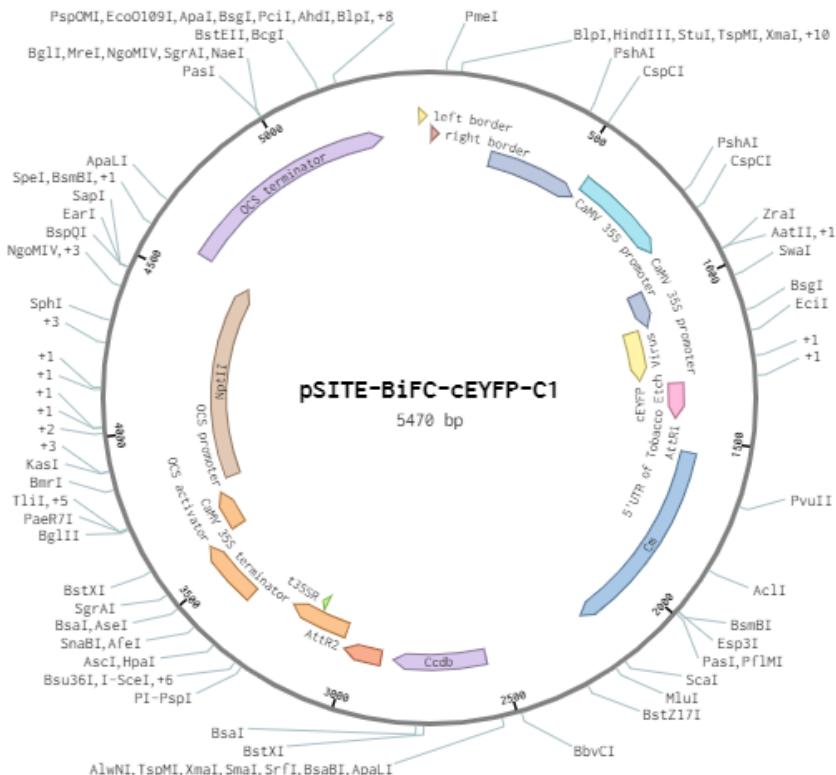


pSITE-BiFC-nEYFP-C1
5791 bp
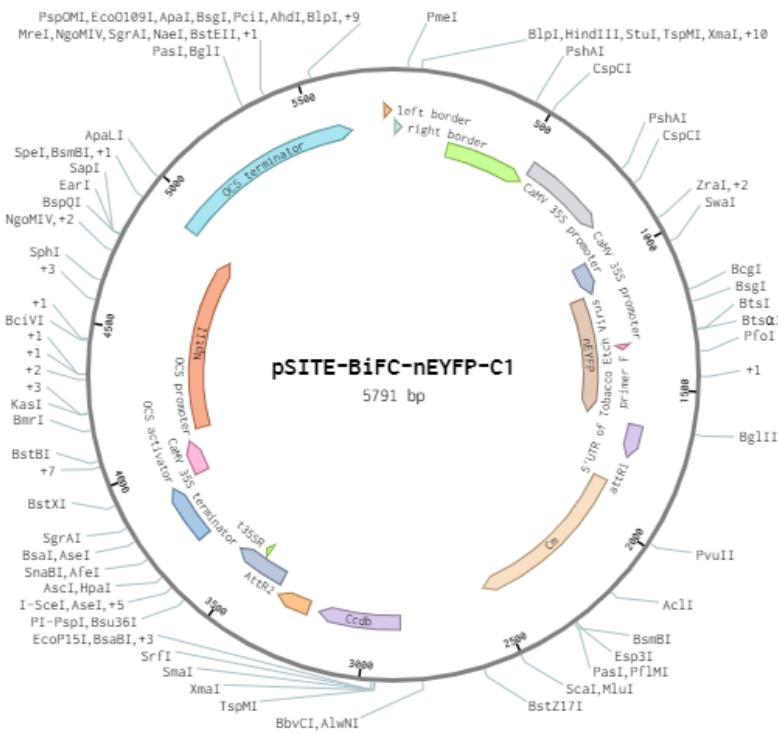
**Supplementary Table 1.** Blastp searches for *Arabidopsis thaliana* orthologs of SPLs with Genbank accession, percentage of identity, alignment length, mismatch, gap open, Q. start, Q. end, S. start, S. end, E value, bitscore and query coverage.

| Supplementary Table 1. Result of BLASTp searches for *Arabidopsis thaliana* orthologs of SPLs. | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Genbank accession** | **Query** | **Subject** | **Percent identity** | **Alignment length** | **Mismatch** | **Gap Open** | **Q. start** | **Q. end** | **S. start** | **S. end** | **Evalue** | **Query coverage** |
| NM_180137 | Atspl1 | AT2G47070.1 | 100 | 881 | 0 | 0 | 1 | 881 | 1 | 881 | 0 | 1834 | 100 |
| NM_001084134.1 | Atspl11 | AT1G27360.1 | 100 | 393 | 0 | 0 | 1 | 393 | 1 | 393 | 0 | 815 | 100 |
| NM_115866 | Atspl12 | AT3G60030.1 | 100 | 927 | 0 | 0 | 1 | 927 | 1 | 927 | 0 | 1930 | 100 |
| NM_124445.3 | Atspl13 | AT5G50570.2 | 100 | 359 | 0 | 0 | 1 | 359 | 1 | 359 | 0 | 736 | 100 |
| NM_124445.3 | Atspl13 | AT5G50670.2 | 100 | 359 | 0 | 0 | 1 | 359 | 1 | 359 | 0 | 736 | 100 |
| NM_101951 | Atspl14 | AT1G20980.1 | 100 | 1035 | 0 | 0 | 1 | 1035 | 1 | 1035 | 0 | 2148 | 100 |
| NM_115654.2 | Atspl15 | AT3G57920.1 | 100 | 354 | 0 | 0 | 1 | 354 | 1 | 354 | 0 | 739 | 100 |
| NM_112390.4 | Atspl5 | AT3G15270.1 | 100 | 181 | 0 | 0 | 1 | 181 | 1 | 181 | 6.22E-132 | 370 | 100 |
| NM_001203404.1 | Atspl7 | AT5G18830.3 | 100 | 818 | 0 | 0 | 1 | 818 | 1 | 818 | 0 | 1692 | 100 |
| NM_202009.2 | Atspl8 | AT1G02065.1 | 100 | 235 | 0 | 0 | 1 | 235 | 1 | 235 | 1.1E-175 | 489 | 96 |
| NM_129782.2 | Atspl9 | AT2G42200.1 | 100 | 375 | 0 | 0 | 1 | 375 | 1 | 375 | 0 | 774 | 100 |
| NM_106308 | Atspl16 | AT1G76580.1 | 99,899 | 986 | 1 | 0 | 3 | 988 | 35 | 1020 | 0 | 2041 | 99 |
| AJ011644.1 | Atspl6 | AT1G69170.1 | 99,419 | 172 | 1 | 0 | 1 | 172 | 1 | 172 | 6.05E-122 | 354 | 82 |
| NM_001084134.1 | Atspl11 | AT1G27370.1 | 77,922 | 385 | 76 | 3 | 1 | 377 | 1 | 384 | 0 | 562 | 96 |
| NM_115866 | Atspl12 | Glyma.07G159300.1.p | 77,5 | 80 | 18 | 0 | 127 | 206 | 103 | 182 | 5.31E-33 | 142 | 69 |
| NM_180137 | Atspl1 | GRMZM2G169270_P01 | 75,641 | 78 | 19 | 0 | 105 | 182 | 164 | 241 | 1.02E-31 | 138 | 71 |
| NM_106308 | Atspl16 | Sspon.06G0001900-3C | 75,51 | 49 | 12 | 0 | 79 | 127 | 177 | 225 | 6.03E-17 | 90,5 | 67 |
| NM_180137 | Atspl1 | Glyma.07G159300.1.p | 75 | 84 | 21 | 0 | 100 | 183 | 97 | 180 | 3.07E-33 | 142 | 72 |
| NM_106308 | Atspl16 | Glyma.02G008600.1.p | 73,864 | 88 | 23 | 0 | 74 | 161 | 148 | 235 | 1.81E-34 | 147 | 62 |
| NM_115866 | Atspl12 | GRMZM2G169270_P01 | 73,171 | 82 | 22 | 0 | 127 | 208 | 165 | 246 | 3.69E-32 | 139 | 69 |
| NM_106308 | Atspl16 | Sspon.006A0002630 | 73,077 | 52 | 14 | 0 | 76 | 127 | 164 | 215 | 9.89E-18 | 93,2 | 68 |
| NM_106308 | Atspl16 | Sspon.06G0001900-1A | 73,077 | 52 | 14 | 0 | 76 | 127 | 164 | 215 | 9.9E-18 | 93,2 | 68 |
| NM_106308 | Atspl16 | AT1G20980.1 | 72,079 | 1010 | 241 | 20 | 6 | 988 | 40 | 1035 | 0 | 1375 | 99 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0253558.3 | 71,951 | 82 | 23 | 0 | 127 | 208 | 91 | 172 | 1.56E-31 | 137 | 70 |
| NM_115866 | Atspl12 | Sspon.006A0002630 | 71,739 | 46 | 13 | 0 | 127 | 172 | 170 | 215 | 3.91E-13 | 78,2 | 66 |
| NM_115866 | Atspl12 | Sspon.06G0001900-1A | 71,739 | 46 | 13 | 0 | 127 | 172 | 170 | 215 | 3.91E-13 | 78,2 | 66 |
| NM_115866 | Atspl12 | Sspon.06G0001900-2B | 71,739 | 46 | 13 | 0 | 127 | 172 | 180 | 225 | 2.95E-13 | 78,6 | 65 |
| NM_115866 | Atspl12 | evm.model.scga7_unitig_440261.2 | 71,429 | 35 | 10 | 0 | 174 | 208 | 1 | 35 | 7.23E-08 | 60,8 | 65 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0110246.1 | 71,429 | 35 | 10 | 0 | 174 | 208 | 1 | 35 | 7.29E-08 | 60,8 | 65 |
| NM_101951 | Atspl14 | AT1G76580.1 | 71,211 | 1049 | 259 | 22 | 1 | 1035 | 1 | 1020 | 0 | 1426 | 100 |
| NM_115866 | Atspl12 | LOC_Os08g40260.1 | 70 | 90 | 27 | 0 | 127 | 216 | 187 | 276 | 1.41E-33 | 144 | 73 |
| NM_180137 | Atspl1 | AT3G60030.1 | 69,321 | 942 | 213 | 25 | 1 | 881 | 1 | 927 | 0 | 1179 | 100 |
| NM_115866 | Atspl12 | AT2G47070.1 | 69,321 | 942 | 213 | 25 | 1 | 927 | 1 | 881 | 0 | 1191 | 100 |
| NM_115866 | Atspl12 | Seita.6G208600.1.p | 68,889 | 90 | 28 | 0 | 127 | 216 | 187 | 276 | 1.15E-32 | 141 | 69 |
| NM_180137 | Atspl1 | Sspon.06G0001900-2B | 68,085 | 47 | 15 | 0 | 105 | 151 | 179 | 225 | 4.79E-13 | 77,8 | 68 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_115866 | Atspl12 | Sobic.007G207000.1.p | 66,667 | 90 | 30 | 0 | 127 | 216 | 178 | 267 | 1.72E-31 | 137 | 70 |
| NM_101951 | Atspl14 | Solyc01g068100.2.1 | 65,487 | 113 | 31 | 2 | 119 | 231 | 145 | 249 | 2.58E-36 | 153 | 72 |
| NM_112390.4 | Atspl5 | AT1G53160.1 | 65,054 | 186 | 48 | 6 | 1 | 181 | 1 | 174 | 4.4E-64 | 198 | 100 |
| NM_106308 | Atspl16 | Sspon.01G0031720-3D | 65 | 100 | 35 | 0 | 74 | 173 | 97 | 196 | 1.24E-36 | 153 | 63 |
| NM_115866 | Atspl12 | Glyma.20G006400.1.p | 64,602 | 113 | 40 | 0 | 94 | 206 | 61 | 173 | 5.97E-37 | 154 | 72 |
| NM_180137 | Atspl1 | Glyma.20G006400.1.p | 63,889 | 108 | 32 | 1 | 83 | 183 | 64 | 171 | 3.39E-35 | 149 | 74 |
| NM_106308 | Atspl16 | LOC_Os03g61760.1 | 63,265 | 98 | 36 | 0 | 74 | 171 | 144 | 241 | 4.76E-33 | 142 | 63 |
| NM_106308 | Atspl16 | GRMZM2G156756_P01 | 61,947 | 113 | 37 | 1 | 59 | 171 | 138 | 244 | 2.3E-37 | 156 | 65 |
| NM_101951 | Atspl14 | Glyma.02G008600.1.p | 61,719 | 128 | 47 | 2 | 77 | 203 | 112 | 238 | 7.3E-38 | 158 | 64 |
| NM_101951 | Atspl14 | Glyma.10G009200.1.p | 61,719 | 128 | 47 | 2 | 77 | 203 | 115 | 241 | 5.22E-38 | 158 | 66 |
| NM_106308 | Atspl16 | Sobic.003G139900.2.p | 61,364 | 88 | 34 | 0 | 73 | 160 | 122 | 209 | 1.4E-27 | 124 | 62 |
| NM_180137 | Atspl1 | Ha412HOChr16g0773431 | 59,829 | 117 | 42 | 1 | 105 | 216 | 161 | 277 | 1.44E-35 | 150 | 75 |
| NM_180137 | Atspl1 | HanPSC8Chr16g0717331 | 59,829 | 117 | 42 | 1 | 105 | 216 | 161 | 277 | 1.46E-35 | 150 | 75 |
| NM_180137 | Atspl1 | HanXRQChr16g0748361 | 59,829 | 117 | 42 | 1 | 105 | 216 | 161 | 277 | 1.44E-35 | 150 | 75 |
| NM_106308 | Atspl16 | GRMZM2G138421_P01 | 59,483 | 116 | 41 | 1 | 56 | 171 | 128 | 237 | 2.92E-35 | 149 | 65 |
| NM_112390.4 | Atspl5 | Seita.9G022800.1.p | 59,13 | 115 | 42 | 3 | 28 | 140 | 117 | 228 | 6.66E-32 | 124 | 62 |
| NM_101951 | Atspl14 | Sobic.003G139900.2.p | 58,947 | 95 | 38 | 1 | 109 | 203 | 120 | 213 | 8.95E-28 | 125 | 61 |
| NM_115866 | Atspl12 | AT1G20980.1 | 58,594 | 128 | 41 | 3 | 111 | 230 | 108 | 231 | 1.52E-35 | 150 | 72 |
| NM_101951 | Atspl14 | AT3G60030.1 | 58,594 | 128 | 41 | 3 | 108 | 231 | 111 | 230 | 2.2E-35 | 150 | 64 |
| NM_106308 | Atspl16 | Glyma.10G009200.1.p | 58,268 | 127 | 51 | 1 | 35 | 161 | 114 | 238 | 1.07E-34 | 147 | 68 |
| NM_112390.4 | Atspl5 | Sspon.001A0040480 | 57,798 | 109 | 43 | 1 | 39 | 147 | 131 | 236 | 5.41E-31 | 122 | 60 |
| NM_112390.4 | Atspl5 | Sspon.01G0031720-1A | 57,798 | 109 | 43 | 1 | 39 | 147 | 131 | 236 | 5.41E-31 | 122 | 60 |
| NM_101951 | Atspl14 | Sspon.01G0031720-3D | 57,6 | 125 | 50 | 2 | 81 | 203 | 64 | 187 | 1.06E-35 | 150 | 68 |
| NM_106308 | Atspl16 | LOC_Os01g18850.1 | 57,292 | 96 | 35 | 1 | 65 | 160 | 96 | 185 | 9.56E-25 | 115 | 62 |
| NM_112390.4 | Atspl5 | Glyma.13G295400.1.p | 56,667 | 120 | 44 | 3 | 69 | 181 | 40 | 158 | 5.87E-33 | 118 | 62 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0134244.2 | 56,637 | 113 | 46 | 1 | 29 | 141 | 122 | 231 | 2.1E-31 | 123 | 62 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0181527.1 | 56,637 | 113 | 46 | 1 | 29 | 141 | 122 | 231 | 1.02E-31 | 123 | 62 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0184661.1 | 56,637 | 113 | 46 | 1 | 29 | 141 | 122 | 231 | 2.12E-32 | 123 | 62 |
| NM_112390.4 | Atspl5 | Glyma.12G226000.1.p | 56,41 | 117 | 50 | 1 | 35 | 150 | 38 | 154 | 8.43E-34 | 121 | 64 |
| NM_112390.4 | Atspl5 | Sobic.001G026500.1.p | 55,752 | 113 | 47 | 1 | 29 | 141 | 121 | 230 | 6.26E-31 | 121 | 62 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0038913.2 | 55,725 | 131 | 50 | 2 | 59 | 189 | 135 | 257 | 2.09E-36 | 153 | 66 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0148300.1 | 55,725 | 131 | 50 | 2 | 59 | 189 | 135 | 257 | 2.44E-36 | 153 | 66 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0229483.1 | 55,725 | 131 | 50 | 2 | 59 | 189 | 135 | 257 | 1.88E-36 | 153 | 66 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0267720.2 | 55,725 | 131 | 50 | 2 | 59 | 189 | 135 | 257 | 1.81E-36 | 153 | 66 |
| NM_106308 | Atspl16 | Sspon.001A0040480 | 55,725 | 131 | 50 | 2 | 59 | 189 | 135 | 257 | 5.71E-36 | 151 | 66 |
| NM_106308 | Atspl16 | Sspon.01G0031720-1A | 55,725 | 131 | 50 | 2 | 59 | 189 | 135 | 257 | 5.72E-36 | 151 | 66 |
| NM_101951 | Atspl14 | GRMZM2G081127_P01 | 55,238 | 105 | 42 | 2 | 99 | 203 | 94 | 193 | 1.83E-27 | 124 | 62 |
| NM_115866 | Atspl12 | GRMZM2G098557_P01 | 54,955 | 111 | 46 | 1 | 127 | 233 | 170 | 280 | 5.48E-31 | 135 | 72 |
| NM_112390.4 | Atspl5 | Glyma.17G039600.1.p | 54,955 | 111 | 47 | 1 | 30 | 140 | 131 | 238 | 2.77E-29 | 117 | 61 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_112390.4 | Atspl5 | Glyma.06G238100.1.p | 54,918 | 122 | 51 | 2 | 63 | 181 | 65 | 185 | 1.23E-37 | 131 | 66 |
| NM_112390.4 | Atspl5 | GRMZM2G156756_P01 | 54,918 | 122 | 45 | 3 | 26 | 147 | 128 | 239 | 8.05E-31 | 121 | 67 |
| NM_112390.4 | Atspl5 | Glyma.09G082000.1.p | 54,867 | 113 | 38 | 3 | 30 | 141 | 126 | 226 | 1.05E-29 | 118 | 62 |
| NM_180137 | Atspl1 | Glyma.02G008600.1.p | 54,839 | 1023 | 307 | 28 | 1 | 881 | 1 | 1010 | 0 | 972 | 100 |
| NM_106308 | Atspl16 | Glyma.07G234200.1.p | 54,667 | 150 | 58 | 2 | 78 | 227 | 158 | 297 | 8.97E-37 | 154 | 68 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0283225.2 | 54,622 | 119 | 51 | 1 | 29 | 147 | 121 | 236 | 1.02E-31 | 124 | 66 |
| NM_115866 | Atspl12 | Glyma.02G008600.1.p | 54,589 | 1035 | 337 | 28 | 1 | 927 | 1 | 1010 | 0 | 971 | 100 |
| NM_112390.4 | Atspl5 | GRMZM2G138421_P01 | 54,128 | 109 | 47 | 1 | 33 | 141 | 121 | 226 | 4.84E-28 | 113 | 60 |
| NM_112390.4 | Atspl5 | HanPSC8Chr17g0784411 | 54,128 | 109 | 50 | 0 | 38 | 146 | 34 | 142 | 1.56E-34 | 122 | 60 |
| NM_112390.4 | Atspl5 | HanXRQChr17g0816731 | 54,128 | 109 | 50 | 0 | 38 | 146 | 34 | 142 | 1.56E-34 | 122 | 60 |
| NM_112390.4 | Atspl5 | Glyma.15G190200.1.p | 53,982 | 113 | 39 | 3 | 30 | 141 | 126 | 226 | 3E-29 | 117 | 62 |
| NM_106308 | Atspl16 | GRMZM2G081127_P01 | 53,271 | 107 | 40 | 1 | 52 | 158 | 91 | 187 | 9.37E-27 | 122 | 64 |
| NM_112390.4 | Atspl5 | Ha412HOChr17g0842881 | 53,211 | 109 | 51 | 0 | 38 | 146 | 34 | 142 | 6.59E-33 | 118 | 60 |
| NM_180137 | Atspl1 | Glyma.10G009200.1.p | 53,061 | 1029 | 305 | 29 | 1 | 881 | 1 | 999 | 0 | 951 | 100 |
| NM_115866 | Atspl12 | Glyma.10G009200.1.p | 52,697 | 1038 | 341 | 26 | 1 | 927 | 1 | 999 | 0 | 961 | 100 |
| NM_180137 | Atspl1 | Solyc01g068100.2.1 | 52,685 | 987 | 332 | 22 | 18 | 881 | 20 | 994 | 0 | 879 | 98 |
| NM_112390.4 | Atspl5 | GRMZM2G126018_P01 | 52,542 | 118 | 49 | 3 | 63 | 176 | 55 | 169 | 3.62E-30 | 117 | 63 |
| NM_115866 | Atspl12 | Solyc01g068100.2.1 | 52,51 | 996 | 363 | 25 | 20 | 927 | 21 | 994 | 0 | 886 | 98 |
| NM_112390.4 | Atspl5 | Lsat_1_v5_gn_1_128400.1 | 52,381 | 126 | 53 | 2 | 12 | 136 | 1 | 120 | 1.7E-35 | 124 | 69 |
| NM_112390.4 | Atspl5 | Lsat_1_v5_gn_1_128440.1 | 52,381 | 126 | 53 | 2 | 12 | 136 | 1 | 120 | 1.15E-36 | 127 | 69 |
| NM_112390.4 | Atspl5 | Glyma.04G159600.1.p | 52,212 | 113 | 46 | 3 | 63 | 175 | 184 | 288 | 8.36E-28 | 112 | 62 |
| NM_101951 | Atspl14 | Glyma.20G006400.1.p | 52,158 | 1066 | 422 | 24 | 1 | 1035 | 1 | 1009 | 0 | 947 | 100 |
| NM_112390.4 | Atspl5 | Solyc02g077920.2.1 | 51,825 | 137 | 53 | 5 | 15 | 147 | 9 | 136 | 2.47E-37 | 129 | 73 |
| NM_101951 | Atspl14 | Glyma.07G159300.1.p | 51,747 | 1059 | 430 | 26 | 11 | 1035 | 8 | 1019 | 0 | 922 | 99 |
| NM_112390.4 | Atspl5 | Glyma.17G080700.1.p | 51,695 | 118 | 49 | 3 | 63 | 180 | 167 | 276 | 1.37E-29 | 117 | 65 |
| NM_124445.3 | Atspl13 | Ha412HOChr08g0364421 | 51,613 | 217 | 84 | 5 | 1 | 215 | 1 | 198 | 1.89E-62 | 207 | 60 |
| NM_124445.3 | Atspl13 | HanPSC8Chr08g0341841 | 51,613 | 217 | 84 | 5 | 1 | 215 | 1 | 198 | 1.89E-62 | 207 | 60 |
| NM_124445.3 | Atspl13 | HanXRQChr08g0354031 | 51,613 | 217 | 84 | 5 | 1 | 215 | 1 | 198 | 9.53E-63 | 208 | 60 |
| NM_124445.3 | Atspl13 | Ha412HOChr05g0237991 | 51,364 | 220 | 83 | 10 | 1 | 215 | 1 | 201 | 1.44E-59 | 200 | 60 |
| NM_124445.3 | Atspl13 | HanPSC8Chr05g0219871 | 51,364 | 220 | 85 | 9 | 1 | 215 | 1 | 203 | 5.66E-59 | 198 | 60 |
| NM_124445.3 | Atspl13 | HanXRQChr05g0227861 | 51,364 | 220 | 85 | 9 | 1 | 215 | 1 | 203 | 5.66E-59 | 198 | 60 |
| NM_112390.4 | Atspl5 | Glyma.13G274900.1.p | 51,316 | 152 | 67 | 4 | 35 | 181 | 38 | 187 | 1.12E-37 | 131 | 81 |
| NM_112390.4 | Atspl5 | Seita.4G270400.1.p | 50,926 | 108 | 51 | 1 | 63 | 170 | 160 | 265 | 1.76E-28 | 113 | 60 |
| NM_112390.4 | Atspl5 | Ha412HOChr10g0445241 | 50,877 | 114 | 56 | 0 | 23 | 136 | 37 | 150 | 3.49E-32 | 121 | 63 |
| NM_112390.4 | Atspl5 | HanXRQChr10g0432891 | 50,877 | 114 | 56 | 0 | 23 | 136 | 37 | 150 | 3.49E-32 | 121 | 63 |
| NM_112390.4 | Atspl5 | Glyma.05G019000.1.p | 50,847 | 118 | 50 | 3 | 63 | 180 | 172 | 281 | 2.17E-28 | 114 | 65 |
| NM_129782.2 | Atspl9 | AT3G57920.1 | 50,815 | 368 | 117 | 16 | 32 | 375 | 27 | 354 | 6.51E-99 | 301 | 92 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0267282.1 | 50,806 | 124 | 53 | 4 | 48 | 168 | 24 | 142 | 5.88E-31 | 114 | 67 |
| NM_106308 | Atspl16 | Seita.9G022800.1.p | 50,658 | 152 | 73 | 1 | 38 | 189 | 107 | 256 | 7.31E-37 | 154 | 69 |
| NM_129782.2 | Atspl9 | Glyma.02G177500.1.p | 50,622 | 241 | 96 | 8 | 31 | 255 | 27 | 260 | 2.05E-55 | 189 | 60 |

| NM_112390.4 | Atspl5 | Glyma.06G205700.1.p | 50,442 | 113 | 46 | 3 | 63 | 175 | 70 | 172 | 3.4E-26 | 106 | 62 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_115654.2 | Atspl15 | AT2G42200.1 | 50,272 | 368 | 119 | 16 | 27 | 354 | 32 | 375 | 3.87E-93 | 286 | 93 |
| NM_180137 | Atspl1 | Sspon.006A0002630 | 50 | 82 | 36 | 1 | 70 | 151 | 139 | 215 | 7.22E-14 | 80,1 | 72 |
| NM_180137 | Atspl1 | Sspon.06G0001900-1A | 50 | 82 | 36 | 1 | 70 | 151 | 139 | 215 | 7.23E-14 | 80,1 | 72 |
| NM_112390.4 | Atspl5 | LOC_Os02g07780.1 | 50 | 110 | 55 | 0 | 63 | 172 | 68 | 177 | 1.76E-28 | 109 | 61 |
| NM_112390.4 | Atspl5 | Sobic.010G254200.2.p | 50 | 108 | 52 | 1 | 63 | 170 | 160 | 265 | 1.22E-27 | 110 | 60 |
| NM_112390.4 | Atspl5 | Sspon.01G0031720-3D | 50 | 138 | 62 | 3 | 29 | 162 | 74 | 208 | 9.11E-31 | 121 | 74 |
| NM_124445.3 | Atspl13 | Lsat_1_v5_gn_5_12360.1 | 49,772 | 219 | 85 | 7 | 1 | 215 | 1 | 198 | 1.23E-57 | 196 | 60 |
| NM_106308 | Atspl16 | Glyma.20G006400.1.p | 49,66 | 1029 | 428 | 24 | 5 | 988 | 26 | 1009 | 0 | 875 | 99 |
| NM_101951 | Atspl14 | Lsat_1_v5_gn_9_48121.1 | 49,439 | 1070 | 428 | 25 | 1 | 1003 | 1 | 1024 | 0 | 902 | 97 |
| NM_001203404.1 | Atspl7 | Glyma.13G213400.1.p | 49,246 | 729 | 347 | 15 | 78 | 794 | 53 | 770 | 0 | 654 | 88 |
| NM_115866 | Atspl12 | Ha412HOChr05g0226701 | 49,239 | 985 | 403 | 25 | 1 | 927 | 1 | 946 | 0 | 857 | 100 |
| NM_115866 | Atspl12 | HanPSC8Chr05g0211201 | 49,239 | 985 | 403 | 25 | 1 | 927 | 1 | 946 | 0 | 857 | 100 |
| NM_202009.2 | Atspl8 | Lsat_1_v5_gn_4_44160.1 | 49,206 | 252 | 85 | 9 | 1 | 235 | 1 | 226 | 3.07E-58 | 189 | 96 |
| NM_202009.2 | Atspl8 | Lsat_1_v5_gn_7_11081.1 | 49,143 | 175 | 63 | 6 | 83 | 235 | 41 | 211 | 1.93E-42 | 148 | 62 |
| NM_112390.4 | Atspl5 | Solyc04g045560.2.1 | 49,107 | 112 | 51 | 3 | 63 | 173 | 68 | 174 | 1.47E-27 | 109 | 61 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0189244.1 | 49,074 | 108 | 53 | 1 | 63 | 170 | 160 | 265 | 4.24E-26 | 106 | 60 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0198101.1 | 49,074 | 108 | 53 | 1 | 63 | 170 | 160 | 265 | 4.49E-26 | 106 | 60 |
| NM_106308 | Atspl16 | Glyma.07G159300.1.p | 49,033 | 1034 | 437 | 24 | 5 | 988 | 26 | 1019 | 0 | 863 | 99 |
| NM_180137 | Atspl1 | Ha412HOChr05g0226701 | 48,921 | 973 | 378 | 33 | 1 | 881 | 1 | 946 | 0 | 825 | 100 |
| NM_180137 | Atspl1 | HanPSC8Chr05g0211201 | 48,921 | 973 | 378 | 33 | 1 | 881 | 1 | 946 | 0 | 825 | 100 |
| AJ011644.1 | Atspl6 | Glyma.01G063700.1.p | 48,889 | 135 | 61 | 2 | 46 | 172 | 95 | 229 | 1.32E-31 | 123 | 60 |
| AJ011644.1 | Atspl6 | Glyma.02G121300.1.p | 48,889 | 135 | 61 | 2 | 46 | 172 | 95 | 229 | 9.97E-32 | 123 | 60 |
| NM_115866 | Atspl12 | HanXRQChr05g0218781 | 48,882 | 984 | 408 | 24 | 1 | 927 | 1 | 946 | 0 | 857 | 100 |
| NM_180137 | Atspl1 | HanXRQChr05g0218781 | 48,66 | 970 | 385 | 32 | 1 | 881 | 1 | 946 | 0 | 823 | 100 |
| NM_112390.4 | Atspl5 | Ha412HOChr17g0813761 | 48,649 | 111 | 52 | 1 | 26 | 136 | 106 | 211 | 1.25E-26 | 107 | 61 |
| NM_112390.4 | Atspl5 | HanPSC8Chr17g0756041 | 48,649 | 111 | 52 | 1 | 26 | 136 | 106 | 211 | 1.25E-26 | 107 | 61 |
| NM_112390.4 | Atspl5 | HanXRQChr17g0787951 | 48,649 | 111 | 52 | 1 | 26 | 136 | 106 | 211 | 1.12E-26 | 107 | 61 |
| NM_202009.2 | Atspl8 | HanPSC8Chr17g0785731 | 48,606 | 251 | 78 | 11 | 1 | 235 | 1 | 216 | 2.26E-59 | 192 | 96 |
| NM_112390.4 | Atspl5 | Lsat_1_v5_gn_1_128420.1 | 48,529 | 136 | 63 | 2 | 12 | 146 | 1 | 130 | 1.33E-35 | 124 | 75 |
| NM_106308 | Atspl16 | Solyc01g068100.2.1 | 48,421 | 190 | 70 | 6 | 82 | 261 | 146 | 317 | 1.82E-35 | 150 | 71 |
| NM_180137 | Atspl1 | Lsat_1_v5_gn_6_117840.1 | 48,384 | 959 | 390 | 27 | 1 | 881 | 1 | 932 | 0 | 812 | 100 |
| NM_115866 | Atspl12 | Lsat_1_v5_gn_1_39821.1 | 48,323 | 954 | 415 | 23 | 20 | 927 | 26 | 947 | 0 | 827 | 98 |
| NM_202009.2 | Atspl8 | Ha412HOChr17g0844341 | 48,207 | 251 | 80 | 10 | 1 | 235 | 1 | 217 | 2.16E-59 | 192 | 96 |
| NM_202009.2 | Atspl8 | HanXRQChr17g0818221 | 48,207 | 251 | 80 | 10 | 1 | 235 | 1 | 217 | 2.16E-59 | 192 | 96 |
| NM_180137 | Atspl1 | Lsat_1_v5_gn_1_39821.1 | 48,156 | 976 | 382 | 29 | 1 | 881 | 1 | 947 | 0 | 821 | 100 |
| NM_101951 | Atspl14 | Sspon.06G0001900-4D | 47,952 | 415 | 190 | 8 | 414 | 804 | 330 | 742 | 4.39E-109 | 360 | 76 |
| NM_115866 | Atspl12 | Glyma.09G082000.1.p | 47,855 | 1049 | 406 | 29 | 1 | 926 | 1 | 1031 | 0 | 840 | 99 |
| NM_180137 | Atspl1 | Solyc05g053240.2.1 | 47,83 | 1014 | 367 | 35 | 10 | 877 | 2 | 999 | 0 | 783 | 99 |

| NM_129782.2 | Atspl9 | Lsat_1_v5_gn_9_28021.1 | 47,826 | 253 | 82 | 10 | 31 | 255 | 18 | 248 | 7.22E-55 | 187 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_106308 | Atspl16 | Lsat_1_v5_gn_9_48121.1 | 47,761 | 1005 | 432 | 26 | 5 | 957 | 62 | 1025 | 0 | 807 | 96 |
| NM_115866 | Atspl12 | Glyma.15G190200.1.p | 47,706 | 1046 | 412 | 26 | 1 | 926 | 1 | 1031 | 0 | 841 | 99 |
| NM_001203404.1 | Atspl7 | Glyma.15G099400.1.p | 47,672 | 816 | 361 | 24 | 8 | 794 | 7 | 785 | 0 | 659 | 96 |
| NM_112390.4 | Atspl5 | Lsat_1_v5_gn_3_31000.1 | 47,581 | 124 | 53 | 3 | 54 | 168 | 117 | 237 | 7.52E-27 | 107 | 64 |
| NM_101951 | Atspl14 | LOC_Os08g40260.1 | 47,551 | 980 | 430 | 20 | 117 | 1035 | 184 | 1140 | 0 | 809 | 89 |
| NM_115866 | Atspl12 | Lsat_1_v5_gn_6_117840.1 | 47,513 | 985 | 406 | 21 | 1 | 927 | 1 | 932 | 0 | 828 | 100 |
| NM_180137 | Atspl1 | Glyma.15G190200.1.p | 47,409 | 1042 | 375 | 29 | 1 | 880 | 1 | 1031 | 0 | 831 | 99 |
| NM_112390.4 | Atspl5 | GRMZM2G414805_P05 | 47,368 | 114 | 59 | 1 | 29 | 142 | 127 | 239 | 1.93E-27 | 110 | 63 |
| NM_115866 | Atspl12 | Solyc05g053240.2.1 | 47,321 | 1008 | 401 | 28 | 19 | 922 | 17 | 998 | 0 | 795 | 98 |
| NM_202009.2 | Atspl8 | Solyc10g018780.1.1 | 47,28 | 239 | 96 | 8 | 1 | 235 | 1 | 213 | 1.12E-56 | 186 | 96 |
| NM_202009.2 | Atspl8 | Glyma.03G117500.1.p | 47,266 | 256 | 102 | 6 | 1 | 235 | 1 | 244 | 7.06E-56 | 185 | 96 |
| NM_101951 | Atspl14 | Sspon.06G0001900-3C | 47,043 | 372 | 181 | 6 | 408 | 764 | 217 | 587 | 1.61E-97 | 329 | 76 |
| NM_180137 | Atspl1 | Glyma.09G082000.1.p | 46,98 | 1043 | 378 | 29 | 1 | 880 | 1 | 1031 | 0 | 834 | 99 |
| NM_112390.4 | Atspl5 | Lsat_1_v5_gn_9_28021.1 | 46,875 | 128 | 66 | 1 | 11 | 138 | 26 | 151 | 4.15E-33 | 124 | 71 |
| NM_101951 | Atspl14 | HanPSC8Chr16g0717331 | 46,861 | 1099 | 460 | 26 | 1 | 1035 | 1 | 1039 | 0 | 883 | 100 |
| NM_101951 | Atspl14 | Ha412HOChr16g0773431 | 46,77 | 1099 | 461 | 26 | 1 | 1035 | 1 | 1039 | 0 | 881 | 100 |
| NM_101951 | Atspl14 | HanXRQChr16g0748361 | 46,77 | 1099 | 461 | 26 | 1 | 1035 | 1 | 1039 | 0 | 881 | 100 |
| NM_101951 | Atspl14 | gg_07486_ORF1 | 46,755 | 678 | 324 | 11 | 388 | 1035 | 1 | 671 | 0 | 591 | 63 |
| NM_115866 | Atspl12 | Glyma.17G039600.1.p | 46,667 | 1020 | 390 | 24 | 19 | 926 | 27 | 1004 | 0 | 793 | 98 |
| NM_115866 | Atspl12 | Glyma.07G234200.1.p | 46,609 | 1032 | 407 | 29 | 19 | 926 | 27 | 1038 | 0 | 797 | 98 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0253558.3 | 46,507 | 959 | 459 | 18 | 117 | 1035 | 88 | 1032 | 0 | 804 | 89 |
| NM_001203404.1 | Atspl7 | Solyc01g080670.2.1 | 46,465 | 792 | 363 | 19 | 20 | 794 | 6 | 753 | 0 | 634 | 95 |
| NM_106308 | Atspl16 | Sspon.06G0001900-3C | 46,277 | 376 | 176 | 7 | 370 | 724 | 217 | 587 | 1.07E-95 | 323 | 67 |
| NM_180137 | Atspl1 | Glyma.07G234200.1.p | 46,198 | 1052 | 380 | 28 | 1 | 880 | 1 | 1038 | 0 | 821 | 99 |
| NM_106308 | Atspl16 | Glyma.17G039600.1.p | 46,078 | 204 | 82 | 5 | 78 | 281 | 157 | 332 | 1.24E-36 | 154 | 74 |
| NM_202009.2 | Atspl8 | Glyma.07G109600.1.p | 46,008 | 263 | 97 | 7 | 1 | 235 | 1 | 246 | 7.77E-55 | 182 | 96 |
| NM_112390.4 | Atspl5 | Sspon.04G0017230-2B | 45,985 | 137 | 70 | 1 | 16 | 148 | 9 | 145 | 2.21E-29 | 114 | 73 |
| NM_101951 | Atspl14 | Sspon.06G0001900-2B | 45,872 | 654 | 326 | 9 | 408 | 1035 | 217 | 868 | 0 | 572 | 76 |
| NM_106308 | Atspl16 | HanPSC8Chr16g0717331 | 45,813 | 1039 | 445 | 24 | 6 | 988 | 63 | 1039 | 0 | 801 | 99 |
| NM_124445.3 | Atspl13 | Solyc05g015840.2.1 | 45,798 | 238 | 95 | 8 | 101 | 315 | 20 | 246 | 1.04E-51 | 177 | 60 |
| NM_106308 | Atspl16 | Ha412HOChr16g0773431 | 45,717 | 1039 | 446 | 24 | 6 | 988 | 63 | 1039 | 0 | 801 | 99 |
| NM_106308 | Atspl16 | HanXRQChr16g0748361 | 45,717 | 1039 | 446 | 24 | 6 | 988 | 63 | 1039 | 0 | 801 | 99 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0110246.1 | 45,595 | 908 | 441 | 18 | 167 | 1035 | 1 | 894 | 0 | 720 | 84 |
| NM_180137 | Atspl1 | Glyma.17G039600.1.p | 45,577 | 1040 | 370 | 24 | 1 | 880 | 1 | 1004 | 0 | 796 | 99 |
| NM_106308 | Atspl16 | gg_07486_ORF1 | 45,575 | 678 | 323 | 14 | 350 | 988 | 1 | 671 | 0 | 568 | 65 |
| NM_180137 | Atspl1 | AT1G76580.1 | 45,556 | 180 | 50 | 3 | 24 | 183 | 40 | 191 | 8.23E-35 | 147 | 82 |
| NM_106308 | Atspl16 | AT2G47070.1 | 45,556 | 180 | 50 | 3 | 8 | 159 | 24 | 183 | 6.15E-36 | 151 | 71 |
| NM_106308 | Atspl16 | Sspon.06G0001900-4D | 45,519 | 424 | 197 | 8 | 370 | 764 | 324 | 742 | 3.23E-104 | 346 | 76 |
| NM_202009.2 | Atspl8 | Ha412HOChr16g0785211 | 45,418 | 251 | 84 | 10 | 1 | 235 | 1 | 214 | 1.35E-49 | 167 | 96 |

| NM_202009.2 | Atspl8 | HanPSC8Chr16g0728221 | 45,418 | 251 | 84 | 10 | 1 | 235 | 1 | 214 | 1.35E-49 | 167 | 96 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_202009.2 | Atspl8 | HanXRQChr16g0759931 | 45,418 | 251 | 84 | 10 | 1 | 235 | 1 | 214 | 1.35E-49 | 167 | 96 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0179190.3 | 45,255 | 137 | 71 | 1 | 16 | 148 | 9 | 145 | 4.51E-29 | 113 | 73 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0226575.2 | 45,255 | 137 | 71 | 1 | 16 | 148 | 9 | 145 | 1.08E-28 | 112 | 73 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0307857.3 | 45,255 | 137 | 71 | 1 | 16 | 148 | 9 | 145 | 4.92E-29 | 113 | 73 |
| NM_112390.4 | Atspl5 | gg_12734_ORF1 | 45,255 | 137 | 71 | 1 | 16 | 148 | 9 | 145 | 4.48E-29 | 113 | 73 |
| NM_101951 | Atspl14 | evm.model.scga7_unitig_440261.2 | 44,983 | 907 | 449 | 16 | 167 | 1035 | 1 | 895 | 0 | 716 | 84 |
| NM_124445.3 | Atspl13 | AC233751.1_FGP002 | 44,954 | 109 | 47 | 3 | 38 | 146 | 59 | 154 | 1.32E-18 | 91,3 | 87 |
| NM_112390.4 | Atspl5 | Ha412HOChr09g0399221 | 44,853 | 136 | 57 | 4 | 24 | 146 | 29 | 159 | 8.85E-28 | 109 | 68 |
| NM_112390.4 | Atspl5 | HanPSC8Chr09g0373801 | 44,853 | 136 | 57 | 4 | 24 | 146 | 29 | 159 | 8.85E-28 | 109 | 68 |
| NM_112390.4 | Atspl5 | HanXRQChr09g0387921 | 44,853 | 136 | 57 | 4 | 24 | 146 | 29 | 159 | 8.85E-28 | 109 | 68 |
| NM_112390.4 | Atspl5 | Lsat_1_v5_gn_5_81161.1 | 44,828 | 145 | 73 | 4 | 17 | 157 | 26 | 167 | 1.23E-27 | 108 | 78 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0352210.1 | 44,823 | 763 | 382 | 13 | 304 | 1035 | 1 | 755 | 0 | 613 | 71 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0253558.3 | 44,549 | 954 | 476 | 18 | 79 | 988 | 88 | 1032 | 0 | 773 | 92 |
| NM_106308 | Atspl16 | Sspon.06G0001900-2B | 44,529 | 658 | 320 | 13 | 370 | 988 | 217 | 868 | 2.27E-180 | 549 | 75 |
| NM_112390.4 | Atspl5 | evm.model.scga7_unitig_297994.1 | 44,526 | 137 | 72 | 1 | 16 | 148 | 9 | 145 | 1.19E-29 | 113 | 73 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0194130.1 | 44,526 | 137 | 72 | 1 | 16 | 148 | 9 | 145 | 1.02E-29 | 113 | 73 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0251330.1 | 44,526 | 137 | 72 | 1 | 16 | 148 | 9 | 145 | 1.29E-29 | 113 | 73 |
| NM_112390.4 | Atspl5 | Sspon.004A0021810 | 44,526 | 137 | 72 | 1 | 16 | 148 | 9 | 145 | 1.32E-28 | 110 | 73 |
| NM_112390.4 | Atspl5 | Sspon.04G0017230-1A | 44,526 | 137 | 72 | 1 | 16 | 148 | 9 | 145 | 1.35E-28 | 110 | 73 |
| NM_202009.2 | Atspl8 | Solyc01g090730.2.1 | 44,49 | 245 | 94 | 9 | 1 | 235 | 1 | 213 | 4.88E-52 | 174 | 96 |
| NM_112390.4 | Atspl5 | GRMZM2G133646_P01 | 44,444 | 117 | 57 | 2 | 51 | 161 | 117 | 231 | 3.36E-24 | 102 | 61 |
| NM_106308 | Atspl16 | GRMZM2G098557_P01 | 44,377 | 987 | 482 | 21 | 57 | 988 | 138 | 1112 | 0 | 756 | 94 |
| NM_129782.2 | Atspl9 | Glyma.19G146000.1.p | 44,267 | 375 | 149 | 15 | 31 | 375 | 15 | 359 | 2.65E-70 | 228 | 92 |
| NM_106308 | Atspl16 | Seita.6G208600.1.p | 44,236 | 1015 | 486 | 19 | 41 | 987 | 115 | 1117 | 0 | 789 | 96 |
| NM_124445.3 | Atspl13 | GRMZM2G371033_P01 | 44,196 | 224 | 102 | 7 | 86 | 303 | 1 | 207 | 1.37E-45 | 161 | 61 |
| NM_106308 | Atspl16 | Solyc05g053240.2.1 | 44,162 | 197 | 89 | 4 | 65 | 261 | 133 | 308 | 1.94E-35 | 150 | 71 |
| NM_115866 | Atspl12 | Seita.9G022800.1.p | 44,142 | 956 | 464 | 26 | 22 | 925 | 29 | 966 | 0 | 681 | 98 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0012627.4 | 44,141 | 256 | 113 | 10 | 694 | 925 | 443 | 692 | 3.58E-43 | 172 | 80 |
| NM_101951 | Atspl14 | LOC_Os03g61760.1 | 44,131 | 213 | 68 | 7 | 39 | 203 | 25 | 234 | 6.36E-36 | 152 | 78 |
| NM_129782.2 | Atspl9 | Glyma.03G143100.1.p | 44,094 | 381 | 149 | 16 | 31 | 375 | 15 | 367 | 1.18E-70 | 229 | 92 |
| NM_106308 | Atspl16 | LOC_Os08g40260.1 | 44,066 | 969 | 471 | 20 | 79 | 988 | 184 | 1140 | 0 | 761 | 92 |
| NM_115866 | Atspl12 | Sobic.001G026500.1.p | 44,062 | 960 | 462 | 27 | 22 | 926 | 29 | 968 | 0 | 663 | 98 |
| NM_115866 | Atspl12 | GRMZM2G156756_P01 | 43,88 | 964 | 459 | 25 | 22 | 925 | 29 | 970 | 0 | 664 | 98 |
| NM_001203404.1 | Atspl7 | Lsat_1_v5_gn_6_14240.1 | 43,869 | 791 | 362 | 19 | 29 | 794 | 31 | 764 | 0 | 616 | 94 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0352210.1 | 43,848 | 764 | 381 | 16 | 264 | 988 | 1 | 755 | 0 | 593 | 73 |
| NM_180137 | Atspl1 | Seita.9G022800.1.p | 43,823 | 947 | 435 | 23 | 21 | 879 | 29 | 966 | 0 | 670 | 98 |
| NM_124445.3 | Atspl13 | Glyma.05G204100.1.p | 43,672 | 403 | 170 | 17 | 1 | 359 | 1 | 390 | 4.53E-78 | 248 | 100 |
| NM_112390.4 | Atspl5 | AT5G50570.2 | 43,651 | 126 | 67 | 3 | 52 | 175 | 92 | 215 | 4.98E-27 | 108 | 69 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_112390.4 | Atspl5 | AT5G50670.2 | 43,651 | 126 | 67 | 3 | 52 | 175 | 92 | 215 | 4.98E-27 | 108 | 69 |
| NM_180137 | Atspl1 | AT1G20980.1 | 43,627 | 204 | 90 | 4 | 20 | 208 | 38 | 231 | 9.94E-37 | 154 | 83 |
| NM_101951 | Atspl14 | Seita.6G208600.1.p | 43,612 | 1135 | 516 | 26 | 3 | 1034 | 4 | 1117 | 0 | 840 | 99 |
| NM_101951 | Atspl14 | GRMZM2G169270_P01 | 43,59 | 1131 | 512 | 26 | 3 | 1035 | 4 | 1106 | 0 | 821 | 99 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0283225.2 | 43,587 | 959 | 463 | 28 | 22 | 925 | 29 | 964 | 0 | 657 | 98 |
| NM_101951 | Atspl14 | Sobic.007G207000.1.p | 43,578 | 1129 | 528 | 23 | 3 | 1035 | 4 | 1119 | 0 | 830 | 99 |
| NM_101951 | Atspl14 | Sspon.006A0002630 | 43,578 | 654 | 316 | 10 | 408 | 1035 | 207 | 833 | 9.93E-170 | 522 | 76 |
| NM_101951 | Atspl14 | Sspon.06G0001900-1A | 43,578 | 654 | 316 | 10 | 408 | 1035 | 207 | 833 | 1.02E-169 | 522 | 76 |
| NM_001084134.1 | Atspl11 | Glyma.11G251500.1.p | 43,537 | 294 | 138 | 9 | 1 | 274 | 1 | 286 | 1.84E-51 | 181 | 70 |
| NM_129782.2 | Atspl9 | GRMZM2G307588_P01 | 43,509 | 285 | 115 | 11 | 18 | 269 | 3 | 274 | 4.54E-47 | 167 | 67 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0134244.2 | 43,496 | 961 | 463 | 29 | 22 | 925 | 29 | 966 | 0 | 659 | 98 |
| NM_124445.3 | Atspl13 | Sobic.005G120600.3.p | 43,421 | 228 | 105 | 6 | 101 | 323 | 78 | 286 | 2.45E-40 | 148 | 62 |
| NM_129782.2 | Atspl9 | Seita.6G205500.1.p | 43,382 | 272 | 94 | 11 | 31 | 255 | 26 | 284 | 8.11E-45 | 162 | 60 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0229483.1 | 43,375 | 966 | 456 | 28 | 22 | 925 | 29 | 965 | 0 | 651 | 98 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0267720.2 | 43,375 | 966 | 457 | 27 | 22 | 925 | 29 | 966 | 0 | 653 | 98 |
| NM_115866 | Atspl12 | Sspon.001A0040480 | 43,375 | 966 | 456 | 29 | 22 | 925 | 29 | 965 | 0 | 648 | 98 |
| NM_115866 | Atspl12 | Sspon.01G0031720-1A | 43,375 | 966 | 456 | 29 | 22 | 925 | 29 | 965 | 0 | 648 | 98 |
| NM_124445.3 | Atspl13 | evm.model.scga7_uti_cns_0026104.2 | 43,231 | 229 | 103 | 8 | 101 | 323 | 122 | 329 | 2.41E-42 | 154 | 62 |
| NM_101951 | Atspl14 | GRMZM2G156756_P01 | 43,204 | 206 | 71 | 4 | 43 | 203 | 33 | 237 | 3.75E-38 | 159 | 77 |
| NM_101951 | Atspl14 | Sobic.001G026500.1.p | 43,204 | 206 | 68 | 5 | 43 | 203 | 33 | 234 | 2.29E-38 | 159 | 77 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0229483.1 | 43,203 | 949 | 437 | 24 | 21 | 879 | 29 | 965 | 0 | 645 | 98 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0038913.2 | 43,168 | 966 | 458 | 28 | 22 | 925 | 29 | 965 | 0 | 647 | 98 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0148300.1 | 43,168 | 966 | 458 | 28 | 22 | 925 | 29 | 965 | 0 | 646 | 98 |
| NM_101951 | Atspl14 | AT2G47070.1 | 43,137 | 204 | 91 | 4 | 38 | 231 | 20 | 208 | 1.75E-36 | 153 | 70 |
| NM_101951 | Atspl14 | GRMZM2G138421_P01 | 43,137 | 204 | 74 | 5 | 43 | 205 | 30 | 232 | 1.61E-35 | 150 | 78 |
| NM_180137 | Atspl1 | Sspon.001A0040480 | 43,098 | 949 | 438 | 24 | 21 | 879 | 29 | 965 | 0 | 644 | 98 |
| NM_180137 | Atspl1 | Sspon.01G0031720-1A | 43,098 | 949 | 438 | 24 | 21 | 879 | 29 | 965 | 0 | 644 | 98 |
| NM_115866 | Atspl12 | GRMZM2G138421_P01 | 43,079 | 968 | 475 | 28 | 11 | 925 | 15 | 959 | 0 | 654 | 99 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0110246.1 | 43,079 | 903 | 462 | 18 | 129 | 988 | 1 | 894 | 0 | 684 | 87 |
| NM_180137 | Atspl1 | Sobic.001G026500.1.p | 43,022 | 953 | 437 | 26 | 21 | 880 | 29 | 968 | 0 | 655 | 98 |
| NM_115866 | Atspl12 | Sspon.01G0031720-3D | 43,01 | 937 | 438 | 28 | 40 | 925 | 5 | 896 | 0 | 610 | 96 |
| NM_180137 | Atspl1 | Sspon.01G0031720-3D | 42,998 | 914 | 408 | 23 | 52 | 879 | 10 | 896 | 0 | 617 | 94 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0148300.1 | 42,993 | 949 | 439 | 24 | 21 | 879 | 29 | 965 | 0 | 642 | 98 |
| NM_001203404.1 | Atspl7 | HanXRQChr08g0342871 | 42,984 | 791 | 374 | 17 | 19 | 794 | 14 | 742 | 0 | 592 | 95 |
| NM_180137 | Atspl1 | GRMZM2G138421_P01 | 42,978 | 947 | 441 | 22 | 20 | 879 | 25 | 959 | 0 | 651 | 98 |
| NM_001203404.1 | Atspl7 | Ha412HOChr08g0353151 | 42,966 | 789 | 377 | 17 | 19 | 794 | 16 | 744 | 0 | 598 | 95 |
| NM_001203404.1 | Atspl7 | HanPSC8Chr08g0331191 | 42,966 | 789 | 377 | 17 | 19 | 794 | 16 | 744 | 0 | 598 | 95 |
| NM_106308 | Atspl16 | GRMZM2G169270_P01 | 42,951 | 1071 | 509 | 25 | 8 | 988 | 48 | 1106 | 0 | 772 | 99 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0267720.2 | 42,932 | 948 | 442 | 22 | 21 | 879 | 29 | 966 | 0 | 645 | 98 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0283225.2 | 42,932 | 948 | 440 | 23 | 21 | 879 | 29 | 964 | 0 | 653 | 98 |

| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0134244.2 | 42,927 | 205 | 71 | 5 | 43 | 203 | 33 | 235 | 3.82E-37 | 155 | 71 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_106308 | Atspl16 | evm.model.scga7_unitig_440261.2 | 42,92 | 904 | 463 | 18 | 129 | 988 | 1 | 895 | 0 | 683 | 87 |
| NM_101951 | Atspl14 | GRMZM2G098557_P01 | 42,895 | 1126 | 533 | 22 | 3 | 1035 | 4 | 1112 | 0 | 815 | 99 |
| NM_001203404.1 | Atspl7 | Ha412HOChr07g0294661 | 42,894 | 781 | 373 | 17 | 29 | 793 | 21 | 744 | 0 | 600 | 94 |
| NM_001203404.1 | Atspl7 | HanPSC8Chr07g0276901 | 42,894 | 781 | 373 | 17 | 29 | 793 | 21 | 744 | 0 | 600 | 94 |
| NM_001203404.1 | Atspl7 | HanXRQChr07g0286041 | 42,894 | 781 | 373 | 17 | 29 | 793 | 21 | 744 | 0 | 600 | 94 |
| NM_180137 | Atspl1 | GRMZM2G156756_P01 | 42,857 | 980 | 436 | 28 | 8 | 879 | 7 | 970 | 0 | 652 | 99 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0134244.2 | 42,842 | 950 | 440 | 24 | 21 | 879 | 29 | 966 | 0 | 656 | 98 |
| NM_115866 | Atspl12 | LOC_Os03g61760.1 | 42,842 | 957 | 478 | 24 | 21 | 925 | 28 | 967 | 0 | 682 | 98 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0038913.2 | 42,677 | 949 | 442 | 25 | 21 | 879 | 29 | 965 | 0 | 643 | 98 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0038913.2 | 42,647 | 204 | 72 | 5 | 43 | 203 | 33 | 234 | 2.8E-36 | 152 | 71 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0148300.1 | 42,647 | 204 | 72 | 5 | 43 | 203 | 33 | 234 | 2.93E-36 | 152 | 71 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0229483.1 | 42,647 | 204 | 72 | 5 | 43 | 203 | 33 | 234 | 1.9E-36 | 153 | 71 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0267720.2 | 42,647 | 204 | 72 | 5 | 43 | 203 | 33 | 234 | 1.82E-36 | 153 | 71 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0283225.2 | 42,647 | 204 | 72 | 5 | 43 | 203 | 33 | 234 | 9.96E-37 | 154 | 71 |
| NM_101951 | Atspl14 | Sspon.001A0040480 | 42,647 | 204 | 72 | 5 | 43 | 203 | 33 | 234 | 2.8E-36 | 152 | 71 |
| NM_101951 | Atspl14 | Sspon.01G0031720-1A | 42,647 | 204 | 72 | 5 | 43 | 203 | 33 | 234 | 2.8E-36 | 152 | 71 |
| NM_112390.4 | Atspl5 | Sobic.007G193500.1.p | 42,636 | 129 | 56 | 3 | 63 | 174 | 130 | 257 | 2.25E-24 | 102 | 62 |
| NM_180137 | Atspl1 | LOC_Os03g61760.1 | 42,549 | 973 | 446 | 26 | 8 | 879 | 7 | 967 | 0 | 677 | 99 |
| NM_129782.2 | Atspl9 | Glyma.09G113800.1.p | 42,513 | 374 | 159 | 17 | 31 | 375 | 27 | 373 | 4.59E-65 | 214 | 92 |
| NM_115866 | Atspl12 | Sspon.03G0015860-3C | 42,507 | 367 | 165 | 14 | 594 | 925 | 510 | 865 | 8.57E-67 | 244 | 72 |
| NM_106308 | Atspl16 | Sobic.007G207000.1.p | 42,41 | 1087 | 501 | 22 | 8 | 988 | 52 | 1119 | 0 | 778 | 99 |
| NM_124445.3 | Atspl13 | Glyma.08G011500.1.p | 42,402 | 408 | 174 | 19 | 1 | 359 | 1 | 396 | 8.84E-73 | 234 | 100 |
| NM_106308 | Atspl16 | Sspon.006A0002630 | 42,401 | 658 | 309 | 14 | 370 | 988 | 207 | 833 | 3.02E-163 | 504 | 68 |
| NM_106308 | Atspl16 | Sspon.06G0001900-1A | 42,401 | 658 | 309 | 14 | 370 | 988 | 207 | 833 | 3.11E-163 | 504 | 68 |
| NM_124445.3 | Atspl13 | Lsat_1_v5_gn_4_184960.1 | 42,387 | 243 | 107 | 5 | 79 | 311 | 46 | 265 | 1.79E-45 | 161 | 65 |
| NM_180137 | Atspl1 | Seita.6G208600.1.p | 42,381 | 210 | 82 | 6 | 105 | 298 | 186 | 372 | 1.97E-32 | 140 | 85 |
| NM_112390.4 | Atspl5 | Lsat_1_v5_gn_5_12360.1 | 42,282 | 149 | 77 | 2 | 32 | 172 | 48 | 195 | 5.22E-31 | 119 | 78 |
| NM_124445.3 | Atspl13 | GRMZM2G061734_P01 | 42,248 | 258 | 100 | 11 | 87 | 311 | 95 | 336 | 8.24E-47 | 168 | 63 |
| NM_106308 | Atspl16 | Lsat_1_v5_gn_1_39821.1 | 42,105 | 228 | 82 | 7 | 8 | 189 | 31 | 254 | 1.06E-35 | 150 | 73 |
| NM_101951 | Atspl14 | Solyc05g053240.2.1 | 41,985 | 262 | 94 | 6 | 22 | 234 | 4 | 256 | 1.47E-43 | 176 | 71 |
| NM_115866 | Atspl12 | Lsat_1_v5_gn_9_48121.1 | 41,85 | 227 | 104 | 6 | 26 | 240 | 66 | 276 | 1.3E-37 | 157 | 77 |
| NM_001084134.1 | Atspl11 | Ha412HOChr17g0855561 | 41,722 | 302 | 135 | 11 | 114 | 393 | 91 | 373 | 8.34E-47 | 169 | 71 |
| NM_001084134.1 | Atspl11 | HanPSC8Chr17g0795991 | 41,722 | 302 | 135 | 11 | 114 | 393 | 91 | 373 | 9.35E-47 | 168 | 71 |
| NM_001084134.1 | Atspl11 | HanXRQChr17g0828971 | 41,722 | 302 | 135 | 11 | 114 | 393 | 91 | 373 | 9.55E-47 | 168 | 71 |
| NM_129782.2 | Atspl9 | Seita.2G254300.1.p | 41,637 | 281 | 123 | 11 | 31 | 283 | 19 | 286 | 5.76E-46 | 165 | 67 |
| NM_106308 | Atspl16 | AT3G60030.1 | 41,629 | 221 | 77 | 3 | 8 | 189 | 25 | 232 | 2.04E-38 | 159 | 71 |
| NM_124445.3 | Atspl13 | Glyma.04G197100.1.p | 41,623 | 382 | 151 | 17 | 32 | 359 | 8 | 371 | 2.35E-77 | 246 | 91 |
| NM_129782.2 | Atspl9 | LOC_Os09g31438.1 | 41,549 | 284 | 112 | 11 | 31 | 272 | 21 | 292 | 3.88E-44 | 160 | 65 |

| NM_112390.4 | Atspl5 | comp203535_c0_seq1_ORF1 | 41,538 | 130 | 54 | 3 | 63 | 170 | 55 | 184 | 2.2E-21 | 94,4 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_112390.4 | Atspl5 | Sspon.007A0010700 | 41,538 | 130 | 54 | 3 | 63 | 170 | 176 | 305 | 2.46E-21 | 94,4 | 60 |
| NM_112390.4 | Atspl5 | Sspon.07G0008180-1A | 41,538 | 130 | 54 | 3 | 63 | 170 | 176 | 305 | 2.46E-21 | 94,4 | 60 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0012627.4 | 41,502 | 253 | 121 | 7 | 651 | 879 | 443 | 692 | 1.02E-40 | 164 | 77 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0029591.1 | 41,497 | 147 | 83 | 2 | 24 | 170 | 122 | 265 | 1.42E-28 | 113 | 81 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0071177.1 | 41,497 | 147 | 83 | 2 | 24 | 170 | 122 | 265 | 1.55E-28 | 113 | 81 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0198066.1 | 41,497 | 147 | 83 | 2 | 24 | 170 | 122 | 265 | 1.63E-28 | 113 | 81 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0254505.1 | 41,497 | 147 | 83 | 2 | 24 | 170 | 122 | 265 | 1.63E-28 | 113 | 81 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0255781.1 | 41,497 | 147 | 83 | 2 | 24 | 170 | 121 | 264 | 5.29E-29 | 114 | 81 |
| NM_112390.4 | Atspl5 | evm.model.scga7_uti_cns_0292903.1 | 41,497 | 147 | 83 | 2 | 24 | 170 | 122 | 265 | 1.45E-28 | 113 | 81 |
| NM_112390.4 | Atspl5 | gg_10666_ORF1 | 41,497 | 147 | 83 | 2 | 24 | 170 | 122 | 265 | 2.86E-30 | 115 | 81 |
| NM_001084134.1 | Atspl11 | Solyc05g015510.2.1 | 41,481 | 270 | 142 | 8 | 114 | 371 | 117 | 382 | 3.88E-46 | 168 | 66 |
| NM_001084134.1 | Atspl11 | Ha412HOChr17g0813761 | 41,418 | 268 | 118 | 7 | 113 | 371 | 84 | 321 | 4.85E-48 | 171 | 66 |
| NM_001084134.1 | Atspl11 | HanPSC8Chr17g0756041 | 41,418 | 268 | 118 | 7 | 113 | 371 | 84 | 321 | 4.85E-48 | 171 | 66 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0085485.1 | 41,406 | 384 | 181 | 14 | 577 | 925 | 423 | 797 | 9.76E-69 | 248 | 93 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0105459.1 | 41,406 | 384 | 181 | 14 | 577 | 925 | 406 | 780 | 5.96E-69 | 248 | 93 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0145568.2 | 41,406 | 384 | 181 | 14 | 577 | 925 | 406 | 780 | 6.83E-69 | 248 | 93 |
| NM_001084134.1 | Atspl11 | Seita.1G091900.1.p | 41,392 | 273 | 141 | 6 | 116 | 371 | 201 | 471 | 1.32E-44 | 166 | 65 |
| NM_112390.4 | Atspl5 | Ha412HOChr08g0353151 | 41,322 | 121 | 60 | 1 | 25 | 145 | 105 | 214 | 6.43E-21 | 93,2 | 67 |
| NM_112390.4 | Atspl5 | HanPSC8Chr08g0331191 | 41,322 | 121 | 60 | 1 | 25 | 145 | 105 | 214 | 6.43E-21 | 93,2 | 67 |
| NM_101951 | Atspl14 | Lsat_1_v5_gn_1_39821.1 | 41,176 | 238 | 97 | 5 | 18 | 213 | 5 | 241 | 8.54E-44 | 176 | 73 |
| NM_180137 | Atspl1 | LOC_Os08g40260.1 | 41,15 | 226 | 94 | 7 | 105 | 314 | 186 | 388 | 1.42E-33 | 144 | 86 |
| NM_180137 | Atspl1 | Lsat_1_v5_gn_9_48121.1 | 41,148 | 209 | 106 | 4 | 25 | 216 | 66 | 274 | 6.66E-37 | 154 | 81 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0080443.2 | 41,146 | 384 | 182 | 14 | 577 | 925 | 405 | 779 | 1.98E-68 | 247 | 93 |
| NM_112390.4 | Atspl5 | Ha412HOChr10g0468341 | 41,139 | 158 | 83 | 3 | 24 | 176 | 31 | 183 | 3.73E-27 | 107 | 85 |
| NM_112390.4 | Atspl5 | HanPSC8Chr10g0439961 | 41,139 | 158 | 83 | 3 | 24 | 176 | 31 | 183 | 3.73E-27 | 107 | 85 |
| NM_112390.4 | Atspl5 | HanXRQChr10g0455511 | 41,139 | 158 | 83 | 3 | 24 | 176 | 31 | 183 | 3.73E-27 | 107 | 85 |
| NM_001084134.1 | Atspl11 | gg_04693_ORF1 | 41,053 | 285 | 140 | 9 | 116 | 378 | 2 | 280 | 2.17E-43 | 158 | 67 |
| NM_115654.2 | Atspl15 | GRMZM2G307588_P01 | 41,045 | 268 | 100 | 9 | 10 | 227 | 2 | 261 | 7.52E-42 | 153 | 62 |
| NM_001084134.1 | Atspl11 | HanXRQChr17g0787951 | 40,977 | 266 | 122 | 6 | 113 | 371 | 84 | 321 | 1.07E-48 | 173 | 66 |
| NM_124445.3 | Atspl13 | AT1G69170.1 | 40,948 | 232 | 113 | 8 | 85 | 312 | 109 | 320 | 7.68E-35 | 135 | 64 |
| NM_001084134.1 | Atspl11 | Sspon.004A0023540 | 40,84 | 262 | 129 | 5 | 1 | 250 | 6 | 253 | 2.82E-46 | 168 | 64 |
| NM_001084134.1 | Atspl11 | Sspon.04G0018610-1A | 40,84 | 262 | 129 | 5 | 1 | 250 | 6 | 253 | 2.82E-46 | 168 | 64 |
| NM_001084134.1 | Atspl11 | AT5G43270.1 | 40,838 | 382 | 179 | 10 | 1 | 371 | 1 | 346 | 1.02E-60 | 205 | 94 |
| NM_180137 | Atspl1 | Sobic.007G207000.1.p | 40,773 | 233 | 92 | 8 | 105 | 318 | 177 | 382 | 1.48E-32 | 140 | 87 |
| NM_112390.4 | Atspl5 | Sobic.009G135000.1.p | 40,769 | 130 | 55 | 3 | 63 | 170 | 186 | 315 | 7.25E-21 | 92,8 | 60 |
| NM_129782.2 | Atspl9 | GRMZM2G126018_P01 | 40,625 | 288 | 126 | 8 | 31 | 282 | 14 | 292 | 1.39E-45 | 164 | 67 |
| NM_112390.4 | Atspl5 | HanXRQChr08g0342871 | 40,496 | 121 | 61 | 1 | 25 | 145 | 103 | 212 | 1.42E-20 | 92 | 67 |
| NM_129782.2 | Atspl9 | HanPSC8Chr10g0417881 | 40,327 | 367 | 146 | 15 | 31 | 375 | 20 | 335 | 2.11E-57 | 193 | 92 |
| NM_124445.3 | Atspl13 | Sspon.002A0013930 | 40,239 | 251 | 97 | 7 | 87 | 304 | 244 | 474 | 1.43E-43 | 161 | 61 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_124445.3 | Atspl13 | Sspon.02G0010770-1A | 40,239 | 251 | 97 | 7 | 87 | 304 | 244 | 474 | 1.43E-43 | 161 | 61 |
| NM_124445.3 | Atspl13 | Sspon.02G0010770-2B | 40,239 | 251 | 97 | 7 | 87 | 304 | 100 | 330 | 5.01E-44 | 160 | 61 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0058773.4 | 40,217 | 184 | 75 | 3 | 8 | 160 | 29 | 208 | 6.02E-29 | 129 | 68 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0080443.2 | 40,209 | 383 | 186 | 10 | 536 | 879 | 401 | 779 | 2.23E-71 | 254 | 90 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0085485.1 | 40,209 | 383 | 186 | 10 | 536 | 879 | 419 | 797 | 2.51E-71 | 255 | 91 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0305902.1 | 40,206 | 291 | 139 | 9 | 1 | 271 | 6 | 281 | 1.13E-47 | 169 | 69 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0345865.1 | 40,206 | 291 | 139 | 9 | 1 | 271 | 6 | 281 | 1.13E-47 | 169 | 69 |
| NM_106308 | Atspl16 | Sspon.06G0001900-2B | 40,116 | 172 | 51 | 5 | 8 | 127 | 54 | 225 | 6.3E-18 | 93,6 | 75 |
| NM_115654.2 | Atspl15 | Glyma.19G146000.1.p | 40,054 | 372 | 153 | 16 | 26 | 354 | 15 | 359 | 4.08E-53 | 182 | 93 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0105459.1 | 39,948 | 383 | 187 | 10 | 536 | 879 | 402 | 780 | 2.52E-70 | 252 | 90 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0145568.2 | 39,948 | 383 | 187 | 10 | 536 | 879 | 402 | 780 | 2.65E-70 | 251 | 90 |
| NM_101951 | Atspl14 | LOC_Os01g18850.1 | 39,894 | 188 | 91 | 4 | 42 | 211 | 11 | 194 | 1.42E-28 | 128 | 68 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0080443.2 | 39,891 | 183 | 76 | 3 | 8 | 160 | 11 | 189 | 8.07E-29 | 128 | 60 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0085485.1 | 39,891 | 183 | 76 | 3 | 8 | 160 | 29 | 207 | 1.17E-28 | 127 | 60 |
| NM_129782.2 | Atspl9 | Ha412HOChr10g0445241 | 39,891 | 366 | 154 | 14 | 31 | 375 | 20 | 340 | 1.68E-59 | 199 | 92 |
| NM_129782.2 | Atspl9 | HanXRQChr10g0432891 | 39,891 | 366 | 154 | 14 | 31 | 375 | 20 | 340 | 1.68E-59 | 199 | 92 |
| NM_001084134.1 | Atspl11 | GRMZM2G097275_P01 | 39,792 | 289 | 129 | 9 | 110 | 370 | 106 | 377 | 1.84E-47 | 172 | 66 |
| NM_115654.2 | Atspl15 | Sobic.002G247800.1.p | 39,781 | 274 | 96 | 11 | 12 | 227 | 3 | 265 | 3.56E-40 | 149 | 61 |
| NM_202009.2 | Atspl8 | Glyma.01G075900.1.p | 39,768 | 259 | 92 | 8 | 1 | 235 | 1 | 219 | 3.99E-40 | 143 | 96 |
| NM_106308 | Atspl16 | comp199491_c0_seq1_ORF1 | 39,674 | 184 | 76 | 3 | 8 | 160 | 29 | 208 | 1.67E-28 | 127 | 68 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0006207.3 | 39,674 | 184 | 76 | 3 | 8 | 160 | 29 | 208 | 1.33E-28 | 128 | 68 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0105459.1 | 39,674 | 184 | 76 | 3 | 8 | 160 | 11 | 190 | 1.15E-28 | 127 | 60 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0145568.2 | 39,674 | 184 | 76 | 3 | 8 | 160 | 11 | 190 | 1.98E-28 | 127 | 60 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0253558.3 | 39,655 | 232 | 96 | 7 | 105 | 318 | 90 | 295 | 1.35E-32 | 140 | 87 |
| NM_124445.3 | Atspl13 | Glyma.06G168600.1.p | 39,627 | 429 | 174 | 20 | 1 | 359 | 1 | 414 | 1.76E-68 | 224 | 100 |
| NM_180137 | Atspl1 | evm.model.scga7_unitig_211197.2 | 39,589 | 778 | 349 | 19 | 152 | 879 | 27 | 733 | 2.4E-157 | 482 | 83 |
| NM_001084134.1 | Atspl11 | Glyma.18G005600.1.p | 39,583 | 384 | 191 | 14 | 1 | 371 | 1 | 356 | 5.22E-49 | 175 | 94 |
| NM_101951 | Atspl14 | Seita.5G030800.1.p | 39,572 | 187 | 83 | 5 | 42 | 203 | 32 | 213 | 2.3E-26 | 120 | 67 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0058773.4 | 39,459 | 185 | 88 | 4 | 42 | 203 | 29 | 212 | 2.36E-29 | 130 | 67 |
| AJ011644.1 | Atspl6 | HanPSC8Chr08g0308961 | 39,437 | 142 | 73 | 3 | 41 | 175 | 21 | 156 | 2.3E-21 | 93,6 | 64 |
| NM_124445.3 | Atspl13 | evm.model.scga7_uti_cns_0026424.2 | 39,326 | 267 | 134 | 9 | 101 | 359 | 78 | 324 | 8E-43 | 154 | 72 |
| NM_001084134.1 | Atspl11 | Sspon.04G0018610-2D | 39,31 | 290 | 143 | 7 | 1 | 271 | 6 | 281 | 1.71E-43 | 158 | 69 |
| NM_112390.4 | Atspl5 | Sspon.006A0003761 | 39,286 | 140 | 56 | 4 | 63 | 174 | 136 | 274 | 2.34E-21 | 94 | 62 |
| NM_112390.4 | Atspl5 | Sspon.06G0002840-1A | 39,286 | 140 | 56 | 4 | 63 | 174 | 136 | 274 | 2.34E-21 | 94 | 62 |
| NM_115866 | Atspl12 | Ha412HOChr16g0773431 | 39,224 | 232 | 106 | 5 | 26 | 240 | 66 | 279 | 8.29E-37 | 154 | 80 |
| NM_115866 | Atspl12 | HanPSC8Chr16g0717331 | 39,224 | 232 | 106 | 5 | 26 | 240 | 66 | 279 | 8.01E-37 | 154 | 80 |
| NM_115866 | Atspl12 | HanXRQChr16g0748361 | 39,224 | 232 | 106 | 5 | 26 | 240 | 66 | 279 | 8.29E-37 | 154 | 80 |
| NM_180137 | Atspl1 | GRMZM2G098557_P01 | 39,207 | 227 | 104 | 8 | 105 | 318 | 169 | 374 | 1.67E-31 | 137 | 87 |
| NM_115866 | Atspl12 | GRMZM2G081127_P01 | 39,15 | 871 | 401 | 25 | 92 | 925 | 81 | 859 | 1.83E-172 | 526 | 90 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0255781.1 | 39,114 | 271 | 149 | 6 | 116 | 371 | 101 | 370 | 2.42E-44 | 162 | 65 |
| NM_101951 | Atspl14 | Glyma.07G234200.1.p | 39,098 | 266 | 101 | 7 | 19 | 230 | 7 | 265 | 7.85E-37 | 155 | 72 |
| NM_115654.2 | Atspl15 | Glyma.03G143100.1.p | 39,075 | 389 | 141 | 20 | 26 | 354 | 15 | 367 | 1.03E-48 | 171 | 93 |
| NM_124445.3 | Atspl13 | Sobic.002G257900.1.p | 39,016 | 305 | 131 | 11 | 36 | 304 | 62 | 347 | 1.66E-44 | 162 | 75 |
| NM_202009.2 | Atspl8 | Seita.3G022100.1.p | 39 | 200 | 88 | 6 | 39 | 235 | 43 | 211 | 1.82E-23 | 101 | 80 |
| NM_106308 | Atspl16 | Lsat_1_v5_gn_6_117840.1 | 38,989 | 277 | 103 | 9 | 8 | 222 | 29 | 301 | 5.69E-37 | 154 | 76 |
| NM_101951 | Atspl14 | GRMZM2G133646_P01 | 38,974 | 195 | 98 | 5 | 42 | 216 | 29 | 222 | 9.52E-29 | 128 | 69 |
| NM_115866 | Atspl12 | evm.model.scga7_unitig_211197.2 | 38,938 | 791 | 361 | 23 | 173 | 925 | 27 | 733 | 2.48E-149 | 462 | 81 |
| NM_101951 | Atspl14 | comp199491_c0_seq1_ORF1 | 38,919 | 185 | 89 | 4 | 42 | 203 | 29 | 212 | 5.69E-29 | 129 | 67 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0006207.3 | 38,919 | 185 | 89 | 4 | 42 | 203 | 29 | 212 | 1.05E-28 | 128 | 67 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0105459.1 | 38,919 | 185 | 89 | 4 | 42 | 203 | 11 | 194 | 4.29E-29 | 129 | 61 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0145568.2 | 38,919 | 185 | 89 | 4 | 42 | 203 | 11 | 194 | 7.61E-29 | 128 | 61 |
| NM_106308 | Atspl16 | Seita.5G030800.1.p | 38,889 | 180 | 77 | 3 | 8 | 158 | 32 | 207 | 1.93E-27 | 124 | 68 |
| NM_180137 | Atspl1 | comp199491_c0_seq1_ORF1 | 38,747 | 942 | 433 | 24 | 9 | 879 | 14 | 882 | 0 | 573 | 99 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0058773.4 | 38,747 | 942 | 433 | 24 | 9 | 879 | 14 | 882 | 0 | 575 | 99 |
| NM_106308 | Atspl16 | Sobic.001G026500.1.p | 38,696 | 230 | 87 | 3 | 9 | 189 | 33 | 257 | 6.93E-38 | 157 | 72 |
| NM_180137 | Atspl1 | GRMZM2G081127_P01 | 38,673 | 874 | 399 | 22 | 63 | 879 | 66 | 859 | 1.14E-178 | 541 | 93 |
| NM_180137 | Atspl1 | GRMZM2G133646_P01 | 38,662 | 957 | 431 | 26 | 1 | 879 | 1 | 879 | 0 | 580 | 99 |
| NM_001203404.1 | Atspl7 | comp203535_c0_seq1_ORF1 | 38,659 | 701 | 383 | 13 | 137 | 803 | 54 | 741 | 7.64E-147 | 452 | 82 |
| NM_124445.3 | Atspl13 | Seita.5G432500.1.p | 38,628 | 277 | 131 | 13 | 86 | 350 | 83 | 332 | 2.27E-41 | 152 | 74 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0058773.4 | 38,582 | 959 | 442 | 29 | 17 | 925 | 21 | 882 | 0 | 560 | 98 |
| NM_001084134.1 | Atspl11 | gg_10666_ORF1 | 38,571 | 280 | 139 | 7 | 1 | 271 | 6 | 261 | 8.49E-47 | 164 | 69 |
| NM_129782.2 | Atspl9 | Solyc10g078700.1.1 | 38,564 | 376 | 171 | 14 | 34 | 375 | 23 | 372 | 1.94E-57 | 194 | 91 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0006207.3 | 38,535 | 942 | 435 | 24 | 9 | 879 | 14 | 882 | 0 | 570 | 99 |
| NM_115866 | Atspl12 | comp199491_c0_seq1_ORF1 | 38,269 | 959 | 445 | 29 | 17 | 925 | 21 | 882 | 0 | 553 | 98 |
| NM_180137 | Atspl1 | Seita.5G030800.1.p | 38,235 | 952 | 426 | 23 | 6 | 879 | 12 | 879 | 0 | 573 | 99 |
| NM_101951 | Atspl14 | Glyma.09G082000.1.p | 38,224 | 259 | 101 | 7 | 19 | 231 | 7 | 252 | 2.33E-37 | 156 | 72 |
| NM_180137 | Atspl1 | Sspon.01G0031720-2B | 38,212 | 895 | 419 | 23 | 21 | 848 | 29 | 856 | 2.37E-162 | 499 | 94 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0006207.3 | 38,205 | 958 | 447 | 27 | 17 | 925 | 21 | 882 | 0 | 552 | 98 |
| NM_124445.3 | Atspl13 | Seita.2G266500.1.p | 38,17 | 317 | 133 | 12 | 36 | 311 | 53 | 347 | 9.83E-42 | 155 | 77 |
| NM_001084134.1 | Atspl11 | Sspon.04G0018550-2D | 38,158 | 304 | 152 | 8 | 1 | 291 | 67 | 347 | 1.13E-42 | 157 | 74 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0134244.2 | 38,095 | 231 | 88 | 3 | 9 | 189 | 33 | 258 | 2.16E-37 | 156 | 72 |
| NM_180137 | Atspl1 | Sspon.03G0015860-3C | 38,073 | 436 | 223 | 12 | 484 | 879 | 437 | 865 | 2.58E-68 | 248 | 79 |
| NM_124445.3 | Atspl13 | AT1G27360.1 | 38,009 | 221 | 126 | 5 | 98 | 315 | 172 | 384 | 3.93E-30 | 122 | 61 |
| NM_001084134.1 | Atspl11 | GRMZM2G414805_P05 | 37,993 | 279 | 139 | 7 | 1 | 271 | 6 | 258 | 1.73E-44 | 163 | 69 |
| NM_115654.2 | Atspl15 | gg_15925_ORF1 | 37,885 | 227 | 103 | 5 | 26 | 239 | 49 | 250 | 4.01E-30 | 121 | 60 |
| NM_001084134.1 | Atspl11 | Lsat_1_v5_gn_3_31000.1 | 37,838 | 370 | 146 | 15 | 5 | 371 | 1 | 289 | 3.84E-48 | 171 | 93 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0283225.2 | 37,826 | 230 | 89 | 3 | 9 | 189 | 33 | 257 | 5.32E-37 | 155 | 72 |
| NM_101951 | Atspl14 | HanXRQChr05g0218781 | 37,809 | 283 | 126 | 7 | 19 | 264 | 7 | 276 | 1.07E-39 | 164 | 78 |
| NM_106308 | Atspl16 | Glyma.09G082000.1.p | 37,778 | 270 | 103 | 7 | 8 | 227 | 29 | 283 | 3.08E-37 | 155 | 81 |

| NM_106308 | Atspl16 | Glyma.15G190200.1.p | 37,778 | 270 | 103 | 7 | 8 | 227 | 29 | 283 | 2.14E-37 | 156 | 75 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_124445.3 | Atspl13 | LOC_Os08g41940.1 | 37,736 | 265 | 99 | 7 | 86 | 305 | 102 | 345 | 2.37E-42 | 156 | 61 |
| NM_180137 | Atspl1 | LOC_Os01g18850.1 | 37,728 | 933 | 429 | 23 | 20 | 879 | 7 | 860 | 0 | 560 | 98 |
| NM_001203404.1 | Atspl7 | Seita.3G246000.1.p | 37,676 | 783 | 399 | 15 | 89 | 803 | 98 | 859 | 5.38E-160 | 491 | 87 |
| NM_001084134.1 | Atspl11 | GRMZM2G065451_P02 | 37,631 | 287 | 134 | 9 | 1 | 272 | 8 | 264 | 3.52E-39 | 149 | 69 |
| NM_129782.2 | Atspl9 | Sspon.004A0023540 | 37,551 | 245 | 129 | 4 | 32 | 255 | 131 | 372 | 9.73E-37 | 142 | 60 |
| NM_129782.2 | Atspl9 | Sspon.04G0018610-1A | 37,551 | 245 | 129 | 4 | 32 | 255 | 131 | 372 | 9.73E-37 | 142 | 60 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0154392.3 | 37,5 | 288 | 151 | 6 | 1 | 271 | 1 | 276 | 5.19E-43 | 158 | 69 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0105459.1 | 37,474 | 475 | 209 | 12 | 16 | 462 | 3 | 417 | 3.43E-79 | 276 | 90 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0145568.2 | 37,474 | 475 | 209 | 12 | 16 | 462 | 3 | 417 | 5.14E-79 | 276 | 90 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0012627.4 | 37,421 | 473 | 211 | 12 | 16 | 462 | 3 | 416 | 1.52E-78 | 272 | 77 |
| NM_115654.2 | Atspl15 | Glyma.09G113800.1.p | 37,366 | 372 | 165 | 14 | 26 | 354 | 27 | 373 | 2.1E-48 | 171 | 93 |
| NM_115866 | Atspl12 | Seita.5G030800.1.p | 37,288 | 944 | 459 | 25 | 22 | 925 | 29 | 879 | 0 | 550 | 98 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0029591.1 | 37,248 | 298 | 153 | 8 | 1 | 288 | 6 | 279 | 6.02E-47 | 169 | 73 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0071177.1 | 37,248 | 298 | 153 | 8 | 1 | 288 | 6 | 279 | 5.77E-47 | 169 | 73 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0198066.1 | 37,248 | 298 | 153 | 8 | 1 | 288 | 6 | 279 | 7.11E-47 | 169 | 73 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0254505.1 | 37,248 | 298 | 153 | 8 | 1 | 288 | 6 | 279 | 7.11E-47 | 169 | 73 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0292903.1 | 37,248 | 298 | 153 | 8 | 1 | 288 | 6 | 279 | 5.26E-47 | 169 | 73 |
| NM_115654.2 | Atspl15 | Sspon.08G0004510-3C | 37,22 | 223 | 103 | 4 | 26 | 236 | 142 | 339 | 1.04E-29 | 122 | 60 |
| NM_001203404.1 | Atspl7 | Sobic.009G135000.1.p | 37,194 | 777 | 413 | 12 | 89 | 803 | 101 | 864 | 5.86E-156 | 480 | 87 |
| NM_124445.3 | Atspl13 | Seita.6G223300.1.p | 37,184 | 277 | 134 | 10 | 83 | 327 | 96 | 364 | 6.02E-42 | 155 | 68 |
| NM_124445.3 | Atspl13 | Glyma.09G113800.1.p | 37,175 | 269 | 120 | 8 | 44 | 299 | 28 | 260 | 6.81E-37 | 140 | 71 |
| NM_001203404.1 | Atspl7 | LOC_Os05g33810.1 | 37,17 | 834 | 436 | 19 | 32 | 797 | 23 | 836 | 4.55E-165 | 503 | 94 |
| NM_106308 | Atspl16 | Sspon.06G0001900-3C | 37,143 | 280 | 150 | 9 | 727 | 988 | 501 | 772 | 3.01E-41 | 167 | 67 |
| NM_115866 | Atspl12 | Sobic.003G139900.2.p | 37,041 | 953 | 465 | 26 | 17 | 925 | 21 | 882 | 1.68E-174 | 533 | 98 |
| NM_101951 | Atspl14 | Lsat_1_v5_gn_6_117840.1 | 37,037 | 297 | 134 | 8 | 19 | 264 | 7 | 301 | 2.41E-41 | 169 | 77 |
| NM_115866 | Atspl12 | Ha412HOChr16g0773431 | 37,017 | 543 | 302 | 14 | 397 | 926 | 523 | 1038 | 8.57E-98 | 334 | 80 |
| NM_115866 | Atspl12 | HanPSC8Chr16g0717331 | 37,017 | 543 | 302 | 14 | 397 | 926 | 523 | 1038 | 5.96E-97 | 332 | 80 |
| NM_115866 | Atspl12 | HanXRQChr16g0748361 | 37,017 | 543 | 302 | 14 | 397 | 926 | 523 | 1038 | 8.57E-98 | 334 | 80 |
| NM_115866 | Atspl12 | LOC_Os01g18850.1 | 36,975 | 952 | 455 | 28 | 21 | 925 | 7 | 860 | 0 | 553 | 98 |
| NM_115866 | Atspl12 | Sspon.01G0031720-2B | 36,968 | 917 | 445 | 27 | 22 | 894 | 29 | 856 | 7.48E-152 | 473 | 94 |
| NM_101951 | Atspl14 | Sspon.06G0001900-2B | 36,937 | 222 | 81 | 6 | 3 | 165 | 4 | 225 | 1.31E-25 | 118 | 76 |
| NM_124445.3 | Atspl13 | Sspon.02G0010770-4D | 36,928 | 306 | 156 | 10 | 36 | 334 | 53 | 328 | 6.15E-44 | 160 | 83 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0189244.1 | 36,913 | 298 | 154 | 8 | 1 | 288 | 6 | 279 | 3.91E-45 | 165 | 73 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0198101.1 | 36,913 | 298 | 154 | 8 | 1 | 288 | 6 | 279 | 3.52E-45 | 165 | 73 |
| NM_001084134.1 | Atspl11 | Sobic.010G254200.2.p | 36,913 | 298 | 154 | 7 | 1 | 288 | 6 | 279 | 3.64E-46 | 167 | 73 |
| NM_106308 | Atspl16 | GRMZM2G133646_P01 | 36,893 | 206 | 104 | 5 | 8 | 189 | 29 | 232 | 1.8E-29 | 130 | 71 |
| NM_115654.2 | Atspl15 | Sspon.08G0004510-2B | 36,889 | 225 | 103 | 5 | 26 | 236 | 138 | 337 | 4.75E-28 | 117 | 60 |
| NM_115654.2 | Atspl15 | GRMZM2G160917_P02 | 36,713 | 286 | 103 | 11 | 11 | 227 | 2 | 278 | 1.85E-29 | 120 | 61 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AJ011644.1 | Atspl6 | Lsat_1_v5_gn_1_14540.1 | 36,634 | 202 | 94 | 6 | 1 | 172 | 1 | 198 | 5.43E-25 | 101 | 82 |
| NM_129782.2 | Atspl9 | AT5G43270.1 | 36,626 | 243 | 119 | 7 | 31 | 255 | 119 | 344 | 6.51E-37 | 142 | 60 |
| NM_101951 | Atspl14 | Sspon.06G0001900-4D | 36,596 | 470 | 196 | 15 | 3 | 397 | 4 | 446 | 2.29E-59 | 222 | 76 |
| NM_129782.2 | Atspl9 | AT1G69170.1 | 36,585 | 287 | 138 | 8 | 40 | 308 | 90 | 350 | 1.1E-34 | 135 | 72 |
| NM_124445.3 | Atspl13 | AT5G43270.1 | 36,486 | 222 | 103 | 7 | 97 | 311 | 165 | 355 | 5.2E-29 | 119 | 60 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0085485.1 | 36,458 | 480 | 220 | 11 | 9 | 462 | 14 | 434 | 2.12E-76 | 269 | 91 |
| NM_001203404.1 | Atspl7 | Sspon.07G0008180-2B | 36,418 | 670 | 375 | 14 | 170 | 803 | 1 | 655 | 1.11E-122 | 387 | 78 |
| NM_129782.2 | Atspl9 | GRMZM2G460544_P01 | 36,416 | 346 | 155 | 14 | 72 | 375 | 86 | 408 | 3.62E-36 | 139 | 81 |
| NM_101951 | Atspl14 | Sspon.06G0001900-3C | 36,404 | 228 | 74 | 6 | 3 | 165 | 4 | 225 | 2.28E-25 | 117 | 76 |
| NM_001203404.1 | Atspl7 | GRMZM2G109354_P01 | 36,364 | 847 | 450 | 17 | 32 | 803 | 18 | 850 | 6.3E-164 | 500 | 94 |
| NM_106308 | Atspl16 | HanXRQChr05g0218781 | 36,286 | 350 | 148 | 14 | 8 | 305 | 29 | 355 | 1.85E-38 | 159 | 86 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0080443.2 | 36,25 | 480 | 207 | 10 | 16 | 462 | 3 | 416 | 6.7E-78 | 273 | 90 |
| NM_001084134.1 | Atspl11 | Sobic.004G036900.2.p | 36,211 | 417 | 205 | 16 | 1 | 378 | 6 | 400 | 2.61E-47 | 171 | 96 |
| NM_115654.2 | Atspl15 | Ha412HOChr10g0445241 | 36,141 | 368 | 149 | 14 | 26 | 354 | 20 | 340 | 5.57E-47 | 166 | 93 |
| NM_115654.2 | Atspl15 | HanXRQChr10g0432891 | 36,141 | 368 | 149 | 14 | 26 | 354 | 20 | 340 | 5.57E-47 | 166 | 93 |
| NM_124445.3 | Atspl13 | Sobic.007G193500.1.p | 36,131 | 274 | 107 | 7 | 99 | 327 | 128 | 378 | 1.76E-41 | 154 | 64 |
| NM_124445.3 | Atspl13 | evm.model.scga7_uti_cns_0063998.5 | 36,066 | 305 | 135 | 7 | 33 | 304 | 1 | 278 | 1.85E-45 | 163 | 76 |
| NM_180137 | Atspl1 | Ha412HOChr16g0773431 | 36,063 | 574 | 308 | 14 | 332 | 880 | 499 | 1038 | 2.26E-96 | 328 | 75 |
| NM_180137 | Atspl1 | HanPSC8Chr16g0717331 | 36,063 | 574 | 308 | 14 | 332 | 880 | 499 | 1038 | 6.74E-96 | 327 | 75 |
| NM_180137 | Atspl1 | HanXRQChr16g0748361 | 36,063 | 574 | 308 | 14 | 332 | 880 | 499 | 1038 | 2.26E-96 | 328 | 75 |
| NM_115654.2 | Atspl15 | HanPSC8Chr10g0417881 | 36,043 | 369 | 149 | 14 | 23 | 354 | 17 | 335 | 3.31E-46 | 164 | 94 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0110078.2 | 35,904 | 415 | 209 | 15 | 1 | 378 | 6 | 400 | 1.7E-43 | 161 | 96 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0110078.2.1 | 35,904 | 415 | 209 | 15 | 1 | 378 | 6 | 400 | 1.7E-43 | 161 | 96 |
| NM_106308 | Atspl16 | Sspon.06G0001900-4D | 35,872 | 407 | 192 | 12 | 8 | 359 | 54 | 446 | 1.31E-49 | 192 | 76 |
| NM_001203404.1 | Atspl7 | Sspon.07G0008180-3D | 35,754 | 537 | 318 | 7 | 282 | 803 | 92 | 616 | 2.47E-100 | 327 | 64 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0211345.1 | 35,663 | 415 | 210 | 15 | 1 | 378 | 6 | 400 | 3.48E-43 | 160 | 96 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0211345.1.1 | 35,663 | 415 | 210 | 15 | 1 | 378 | 6 | 400 | 3.48E-43 | 160 | 96 |
| NM_124445.3 | Atspl13 | LOC_Os09g31438.1 | 35,565 | 239 | 123 | 8 | 101 | 319 | 74 | 301 | 1.08E-24 | 107 | 61 |
| NM_115654.2 | Atspl15 | Solyc10g078700.1.1 | 35,526 | 380 | 163 | 17 | 28 | 354 | 22 | 372 | 7.35E-49 | 172 | 92 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0080443.2 | 35,509 | 521 | 242 | 14 | 17 | 527 | 3 | 439 | 4.18E-76 | 268 | 93 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0105459.1 | 35,429 | 525 | 238 | 15 | 17 | 527 | 3 | 440 | 8.45E-77 | 270 | 93 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0145568.2 | 35,429 | 525 | 238 | 15 | 17 | 527 | 3 | 440 | 1.79E-76 | 270 | 93 |
| NM_115866 | Atspl12 | GRMZM2G133646_P01 | 35,368 | 950 | 472 | 30 | 22 | 925 | 26 | 879 | 2.79E-153 | 477 | 98 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0215805.3 | 35,322 | 419 | 207 | 16 | 1 | 378 | 6 | 401 | 2.93E-41 | 155 | 96 |
| NM_101951 | Atspl14 | Sspon.006A0002630 | 35,321 | 218 | 80 | 6 | 3 | 165 | 4 | 215 | 6.37E-25 | 116 | 76 |
| NM_101951 | Atspl14 | Sspon.06G0001900-1A | 35,321 | 218 | 80 | 6 | 3 | 165 | 4 | 215 | 6.38E-25 | 116 | 76 |
| NM_124445.3 | Atspl13 | Sspon.006A0003761 | 35,216 | 301 | 116 | 11 | 99 | 344 | 134 | 410 | 1.76E-37 | 144 | 69 |
| NM_124445.3 | Atspl13 | Sspon.06G0002840-1A | 35,216 | 301 | 116 | 11 | 99 | 344 | 134 | 410 | 1.76E-37 | 144 | 69 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0012627.4 | 35,115 | 524 | 240 | 13 | 17 | 527 | 3 | 439 | 1.06E-76 | 268 | 80 |
| NM_101951 | Atspl14 | Ha412HOChr05g0226701 | 35,077 | 325 | 158 | 8 | 19 | 306 | 7 | 315 | 1.03E-39 | 164 | 82 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_101951 | Atspl14 | HanPSC8Chr05g0211201 | 35,077 | 325 | 158 | 8 | 19 | 306 | 7 | 315 | 1.03E-39 | 164 | 82 |
| NM_180137 | Atspl1 | Sspon.06G0001900-4D | 35,058 | 599 | 307 | 15 | 105 | 655 | 178 | 742 | 2.05E-91 | 309 | 63 |
| NM_001084134.1 | Atspl11 | Sspon.004A0023470 | 35,024 | 414 | 214 | 13 | 1 | 378 | 6 | 400 | 2.03E-43 | 161 | 96 |
| NM_001084134.1 | Atspl11 | Sspon.04G0018550-1A | 35,024 | 414 | 214 | 13 | 1 | 378 | 6 | 400 | 2.03E-43 | 161 | 96 |
| NM_115866 | Atspl12 | Lsat_1_v5_gn_9_48121.1 | 34,915 | 527 | 289 | 17 | 397 | 895 | 525 | 1025 | 4.42E-75 | 273 | 77 |
| NM_001084134.1 | Atspl11 | evm.model.scga7_uti_cns_0093777.1 | 34,856 | 416 | 211 | 12 | 1 | 378 | 6 | 399 | 6.29E-41 | 154 | 96 |
| NM_124445.3 | Atspl13 | Solyc05g012040.2.1 | 34,828 | 290 | 149 | 9 | 31 | 307 | 166 | 428 | 9.26E-35 | 137 | 77 |
| NM_180137 | Atspl1 | Sobic.007G207000.1.p | 34,746 | 590 | 332 | 15 | 330 | 880 | 543 | 1118 | 9.27E-95 | 326 | 87 |
| NM_180137 | Atspl1 | AT1G20980.1 | 34,722 | 576 | 322 | 16 | 337 | 880 | 481 | 1034 | 1.77E-82 | 290 | 83 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0085485.1 | 34,667 | 525 | 241 | 15 | 17 | 527 | 21 | 457 | 5.4E-73 | 260 | 93 |
| NM_115866 | Atspl12 | Seita.6G208600.1.p | 34,642 | 586 | 331 | 13 | 374 | 926 | 551 | 1117 | 6.13E-93 | 322 | 69 |
| NM_124445.3 | Atspl13 | Lsat_1_v5_gn_4_113241.1 | 34,601 | 263 | 122 | 7 | 82 | 340 | 96 | 312 | 2.47E-32 | 127 | 72 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0110246.1 | 34,583 | 587 | 336 | 13 | 330 | 880 | 319 | 893 | 8.72E-96 | 324 | 81 |
| NM_180137 | Atspl1 | evm.model.scga7_unitig_440261.2 | 34,576 | 590 | 333 | 14 | 330 | 880 | 319 | 894 | 5.56E-96 | 325 | 81 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0158300.4 | 34,576 | 590 | 333 | 14 | 330 | 880 | 25 | 600 | 7.11E-99 | 324 | 63 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0253558.3 | 34,576 | 590 | 333 | 14 | 330 | 880 | 456 | 1031 | 2.56E-95 | 326 | 87 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0352210.1 | 34,576 | 590 | 333 | 14 | 330 | 880 | 179 | 754 | 3.28E-97 | 325 | 63 |
| NM_180137 | Atspl1 | gg_07486_ORF1 | 34,576 | 590 | 333 | 14 | 330 | 880 | 95 | 670 | 1.2E-97 | 323 | 63 |
| NM_180137 | Atspl1 | GRMZM2G169270_P01 | 34,576 | 590 | 333 | 14 | 330 | 880 | 530 | 1105 | 3.22E-94 | 324 | 71 |
| NM_115866 | Atspl12 | Sobic.007G207000.1.p | 34,558 | 599 | 343 | 14 | 364 | 926 | 533 | 1118 | 3.03E-94 | 325 | 70 |
| NM_180137 | Atspl1 | GRMZM2G098557_P01 | 34,518 | 591 | 333 | 14 | 330 | 880 | 535 | 1111 | 9.98E-95 | 325 | 87 |
| NM_180137 | Atspl1 | LOC_Os08g40260.1 | 34,506 | 597 | 324 | 17 | 330 | 880 | 564 | 1139 | 1.34E-93 | 323 | 86 |
| NM_115866 | Atspl12 | GRMZM2G169270_P01 | 34,477 | 583 | 341 | 12 | 374 | 926 | 534 | 1105 | 1.66E-92 | 320 | 69 |
| NM_115866 | Atspl12 | Glyma.20G006400.1.p | 34,459 | 592 | 325 | 16 | 370 | 926 | 445 | 1008 | 2.61E-81 | 287 | 72 |
| NM_101951 | Atspl14 | Seita.9G022800.1.p | 34,448 | 299 | 169 | 8 | 94 | 377 | 125 | 411 | 1.11E-31 | 138 | 89 |
| NM_115866 | Atspl12 | LOC_Os08g40260.1 | 34,416 | 616 | 336 | 16 | 344 | 926 | 559 | 1139 | 2.84E-94 | 326 | 73 |
| NM_180137 | Atspl1 | Sspon.06G0001900-2B | 34,354 | 588 | 333 | 14 | 330 | 878 | 292 | 865 | 2.8E-95 | 322 | 68 |
| NM_115866 | Atspl12 | Glyma.07G159300.1.p | 34,343 | 594 | 323 | 17 | 370 | 926 | 455 | 1018 | 8.15E-80 | 283 | 69 |
| NM_124445.3 | Atspl13 | Glyma.05G019000.1.p | 34,323 | 303 | 127 | 10 | 42 | 295 | 126 | 405 | 4.48E-33 | 132 | 71 |
| NM_115866 | Atspl12 | Sspon.03G0015860-3C | 34,286 | 350 | 186 | 11 | 17 | 353 | 199 | 517 | 5.12E-41 | 167 | 72 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0110246.1 | 34,281 | 598 | 345 | 15 | 364 | 926 | 309 | 893 | 1.46E-93 | 319 | 65 |
| NM_180137 | Atspl1 | Seita.6G208600.1.p | 34,237 | 590 | 330 | 14 | 330 | 880 | 547 | 1117 | 9.91E-95 | 325 | 85 |
| NM_115866 | Atspl12 | evm.model.scga7_unitig_440261.2 | 34,224 | 599 | 345 | 15 | 364 | 926 | 309 | 894 | 3.72E-93 | 318 | 65 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0158300.4 | 34,224 | 599 | 345 | 15 | 364 | 926 | 15 | 600 | 5.36E-96 | 318 | 61 |
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0352210.1 | 34,224 | 599 | 345 | 15 | 364 | 926 | 169 | 754 | 2.48E-94 | 318 | 61 |
| NM_115866 | Atspl12 | gg_07486_ORF1 | 34,224 | 599 | 345 | 15 | 364 | 926 | 85 | 670 | 4.48E-95 | 317 | 61 |
| NM_129782.2 | Atspl9 | evm.model.scga7_uti_cns_0026104.2 | 34,221 | 263 | 132 | 8 | 74 | 316 | 122 | 363 | 3.6E-27 | 114 | 65 |
| NM_115866 | Atspl12 | GRMZM2G098557_P01 | 34,171 | 597 | 349 | 13 | 364 | 926 | 525 | 1111 | 1.17E-92 | 321 | 72 |
| NM_115866 | Atspl12 | Sspon.06G0001900-4D | 34,095 | 613 | 322 | 17 | 127 | 702 | 179 | 746 | 4.44E-89 | 304 | 62 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_115866 | Atspl12 | evm.model.scga7_uti_cns_0253558.3 | 34,057 | 599 | 346 | 14 | 364 | 926 | 446 | 1031 | 2.73E-92 | 318 | 70 |
| NM_180137 | Atspl1 | Sspon.03G0015860-3C | 34,043 | 329 | 180 | 10 | 9 | 309 | 192 | 511 | 6.46E-39 | 160 | 79 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0229483.1 | 34,043 | 611 | 359 | 11 | 458 | 1035 | 368 | 967 | 2.51E-83 | 294 | 71 |
| NM_180137 | Atspl1 | Glyma.20G006400.1.p | 34,028 | 576 | 339 | 15 | 332 | 880 | 447 | 1008 | 1.48E-81 | 287 | 74 |
| NM_115866 | Atspl12 | Sspon.06G0001900-2B | 34,003 | 597 | 345 | 15 | 364 | 924 | 282 | 865 | 2.59E-92 | 315 | 65 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0267720.2 | 33,987 | 612 | 359 | 11 | 458 | 1035 | 368 | 968 | 5.18E-83 | 293 | 71 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0038913.2 | 33,879 | 611 | 360 | 11 | 458 | 1035 | 368 | 967 | 2.56E-83 | 294 | 71 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0148300.1 | 33,879 | 611 | 360 | 11 | 458 | 1035 | 368 | 967 | 3.69E-83 | 294 | 71 |
| NM_101951 | Atspl14 | Sspon.001A0040480 | 33,879 | 611 | 360 | 11 | 458 | 1035 | 368 | 967 | 9.06E-83 | 293 | 71 |
| NM_101951 | Atspl14 | Sspon.01G0031720-1A | 33,879 | 611 | 360 | 11 | 458 | 1035 | 368 | 967 | 9.2E-83 | 293 | 71 |
| NM_101951 | Atspl14 | Sspon.01G0031720-3D | 33,879 | 611 | 360 | 11 | 458 | 1035 | 299 | 898 | 9.84E-84 | 294 | 68 |
| NM_101951 | Atspl14 | AT2G47070.1 | 33,874 | 555 | 315 | 15 | 501 | 1034 | 357 | 880 | 1.08E-84 | 296 | 70 |
| NM_101951 | Atspl14 | Lsat_1_v5_gn_6_117840.1 | 33,857 | 573 | 332 | 15 | 481 | 1034 | 387 | 931 | 2.48E-80 | 285 | 77 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0080443.2 | 33,846 | 130 | 74 | 3 | 466 | 584 | 311 | 439 | 2.52E-15 | 85,1 | 60 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0105459.1 | 33,846 | 130 | 74 | 3 | 466 | 584 | 312 | 440 | 2.57E-15 | 85,1 | 60 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0145568.2 | 33,846 | 130 | 74 | 3 | 466 | 584 | 312 | 440 | 2.61E-15 | 85,1 | 60 |
| NM_101951 | Atspl14 | Glyma.15G190200.1.p | 33,832 | 334 | 151 | 9 | 19 | 302 | 7 | 320 | 4.23E-37 | 155 | 79 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0229483.1 | 33,816 | 553 | 329 | 13 | 463 | 988 | 425 | 967 | 3.99E-72 | 262 | 66 |
| NM_106308 | Atspl16 | Sspon.001A0040480 | 33,816 | 553 | 329 | 13 | 463 | 988 | 425 | 967 | 8.71E-73 | 264 | 66 |
| NM_106308 | Atspl16 | Sspon.01G0031720-1A | 33,816 | 553 | 329 | 13 | 463 | 988 | 425 | 967 | 8.81E-73 | 264 | 66 |
| NM_180137 | Atspl1 | Glyma.07G159300.1.p | 33,788 | 586 | 340 | 16 | 328 | 880 | 448 | 1018 | 1.99E-82 | 290 | 72 |
| NM_106308 | Atspl16 | AT3G60030.1 | 33,748 | 563 | 298 | 19 | 466 | 987 | 398 | 926 | 6.09E-76 | 272 | 71 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0134244.2 | 33,715 | 611 | 361 | 11 | 458 | 1035 | 369 | 968 | 1.7E-82 | 292 | 71 |
| NM_101951 | Atspl14 | Sspon.06G0001900-3C | 33,651 | 315 | 181 | 9 | 737 | 1035 | 470 | 772 | 6.07E-38 | 157 | 76 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0038913.2 | 33,635 | 553 | 330 | 13 | 463 | 988 | 425 | 967 | 3.81E-72 | 262 | 66 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0148300.1 | 33,635 | 553 | 330 | 13 | 463 | 988 | 425 | 967 | 7.76E-72 | 261 | 66 |
| NM_106308 | Atspl16 | Sspon.01G0031720-3D | 33,635 | 553 | 330 | 13 | 463 | 988 | 356 | 898 | 3.68E-72 | 261 | 63 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0267720.2 | 33,574 | 554 | 330 | 12 | 463 | 988 | 425 | 968 | 2.06E-72 | 263 | 66 |
| NM_101951 | Atspl14 | Glyma.09G082000.1.p | 33,564 | 578 | 326 | 20 | 500 | 1035 | 471 | 1032 | 4.7E-78 | 280 | 72 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0283225.2 | 33,552 | 611 | 362 | 11 | 458 | 1035 | 367 | 966 | 1.54E-81 | 289 | 71 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0134244.2 | 33,454 | 553 | 331 | 13 | 463 | 988 | 426 | 968 | 1.55E-70 | 257 | 72 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0283225.2 | 33,454 | 553 | 331 | 13 | 463 | 988 | 424 | 966 | 1.49E-71 | 260 | 72 |
| NM_180137 | Atspl1 | AT1G76580.1 | 33,448 | 583 | 336 | 15 | 320 | 880 | 467 | 1019 | 6.93E-79 | 280 | 82 |
| NM_101951 | Atspl14 | GRMZM2G138421_P01 | 33,383 | 674 | 391 | 17 | 394 | 1035 | 314 | 961 | 1.48E-83 | 295 | 78 |
| NM_101951 | Atspl14 | Glyma.15G190200.1.p | 33,333 | 579 | 326 | 18 | 500 | 1035 | 471 | 1032 | 6.59E-78 | 280 | 79 |
| NM_101951 | Atspl14 | Glyma.07G234200.1.p | 33,276 | 577 | 325 | 17 | 500 | 1035 | 482 | 1039 | 8.03E-80 | 286 | 72 |
| NM_106308 | Atspl16 | GRMZM2G156756_P01 | 33,273 | 553 | 332 | 11 | 463 | 988 | 430 | 972 | 8.38E-72 | 261 | 65 |
| NM_180137 | Atspl1 | Sobic.003G139900.2.p | 33,263 | 947 | 478 | 22 | 9 | 879 | 14 | 882 | 4.2E-148 | 462 | 99 |
| NM_106308 | Atspl16 | GRMZM2G138421_P01 | 33,212 | 551 | 331 | 11 | 463 | 988 | 423 | 961 | 1.85E-71 | 259 | 65 |
| NM_106308 | Atspl16 | Ha412HOChr05g0226701 | 33,178 | 428 | 186 | 15 | 8 | 361 | 29 | 430 | 2.26E-38 | 159 | 91 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_106308 | Atspl16 | HanPSC8Chr05g0211201 | 33,178 | 428 | 186 | 15 | 8 | 361 | 29 | 430 | 2.26E-38 | 159 | 91 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0105459.1 | 33,094 | 139 | 93 | 0 | 500 | 638 | 311 | 449 | 1.76E-16 | 89 | 61 |
| NM_101951 | Atspl14 | evm.model.scga7_uti_cns_0145568.2 | 33,094 | 139 | 93 | 0 | 500 | 638 | 311 | 449 | 1.66E-16 | 89 | 61 |
| NM_106308 | Atspl16 | evm.model.scga7_uti_cns_0085485.1 | 33,077 | 130 | 75 | 3 | 466 | 584 | 329 | 457 | 2.34E-14 | 82 | 60 |
| NM_101951 | Atspl14 | Lsat_1_v5_gn_1_39821.1 | 32,993 | 588 | 336 | 18 | 471 | 1034 | 393 | 946 | 1.26E-80 | 286 | 73 |
| NM_115866 | Atspl12 | AT1G20980.1 | 32,931 | 580 | 322 | 14 | 383 | 926 | 486 | 1034 | 8E-77 | 275 | 72 |
| NM_106308 | Atspl16 | Sobic.001G026500.1.p | 32,911 | 553 | 334 | 13 | 463 | 988 | 427 | 969 | 1.09E-71 | 260 | 72 |
| NM_101951 | Atspl14 | Ha412HOChr05g0226701 | 32,879 | 587 | 342 | 15 | 474 | 1034 | 385 | 945 | 1.6E-85 | 300 | 82 |
| NM_101951 | Atspl14 | HanPSC8Chr05g0211201 | 32,879 | 587 | 342 | 15 | 474 | 1034 | 385 | 945 | 1.6E-85 | 300 | 82 |
| NM_101951 | Atspl14 | Glyma.02G008600.1.p | 32,862 | 566 | 337 | 15 | 500 | 1034 | 456 | 1009 | 4.67E-83 | 294 | 64 |
| NM_106308 | Atspl16 | Seita.9G022800.1.p | 32,796 | 558 | 328 | 14 | 463 | 988 | 426 | 968 | 3.06E-72 | 262 | 69 |
| NM_180137 | Atspl1 | evm.model.scga7_unitig_440261.2 | 32,778 | 180 | 85 | 7 | 153 | 318 | 1 | 158 | 4.84E-09 | 64,7 | 81 |
| NM_180137 | Atspl1 | evm.model.scga7_uti_cns_0110246.1 | 32,778 | 180 | 85 | 7 | 153 | 318 | 1 | 158 | 6.3E-09 | 64,3 | 81 |
| NM_124445.3 | Atspl13 | comp204585_c2_seq1_ORF1 | 32,743 | 226 | 124 | 9 | 142 | 359 | 9 | 214 | 1.9E-15 | 78,2 | 61 |
| NM_106308 | Atspl16 | AT2G47070.1 | 32,601 | 592 | 329 | 16 | 435 | 987 | 320 | 880 | 1.4E-80 | 284 | 71 |
| NM_101951 | Atspl14 | HanXRQChr05g0218781 | 32,538 | 587 | 344 | 15 | 474 | 1034 | 385 | 945 | 1.31E-84 | 298 | 78 |
| NM_106308 | Atspl16 | Solyc01g068100.2.1 | 32,513 | 569 | 324 | 18 | 467 | 987 | 437 | 993 | 1.25E-72 | 263 | 71 |
| NM_101951 | Atspl14 | GRMZM2G156756_P01 | 32,493 | 674 | 399 | 14 | 395 | 1035 | 322 | 972 | 5.39E-83 | 293 | 77 |
| NM_124445.3 | Atspl13 | GRMZM2G101511_P02 | 32,443 | 262 | 118 | 13 | 123 | 342 | 162 | 406 | 1.45E-15 | 81,6 | 61 |
| NM_101951 | Atspl14 | AT3G60030.1 | 32,379 | 559 | 323 | 14 | 501 | 1034 | 398 | 926 | 1.05E-78 | 281 | 64 |
| NM_106308 | Atspl16 | Lsat_1_v5_gn_1_39821.1 | 32,327 | 563 | 334 | 15 | 451 | 987 | 405 | 946 | 2.77E-77 | 276 | 73 |
| NM_106308 | Atspl16 | Glyma.15G190200.1.p | 32,112 | 573 | 327 | 17 | 466 | 988 | 472 | 1032 | 2.56E-73 | 266 | 75 |
| NM_124445.3 | Atspl13 | GRMZM5G878561_P01 | 32,071 | 396 | 157 | 15 | 1 | 329 | 1 | 351 | 1.78E-37 | 143 | 92 |
| NM_180137 | Atspl1 | Lsat_1_v5_gn_9_48121.1 | 31,978 | 541 | 329 | 13 | 332 | 849 | 501 | 1025 | 2.15E-71 | 261 | 81 |
| NM_106308 | Atspl16 | Glyma.07G234200.1.p | 31,93 | 570 | 328 | 13 | 466 | 988 | 483 | 1039 | 1.03E-73 | 267 | 68 |
| NM_101951 | Atspl14 | Glyma.10G009200.1.p | 31,904 | 583 | 344 | 14 | 481 | 1034 | 440 | 998 | 6.19E-81 | 288 | 66 |
| NM_101951 | Atspl14 | Sobic.001G026500.1.p | 31,899 | 674 | 403 | 13 | 395 | 1035 | 319 | 969 | 1.64E-82 | 292 | 77 |
| NM_101951 | Atspl14 | Solyc05g053240.2.1 | 31,777 | 557 | 327 | 16 | 500 | 1019 | 446 | 986 | 8.87E-76 | 274 | 71 |
| NM_180137 | Atspl1 | Sspon.006A0002630 | 31,695 | 590 | 325 | 15 | 330 | 880 | 282 | 832 | 6.14E-79 | 276 | 72 |
| NM_180137 | Atspl1 | Sspon.06G0001900-1A | 31,695 | 590 | 325 | 15 | 330 | 880 | 282 | 832 | 6.24E-79 | 276 | 72 |
| NM_106308 | Atspl16 | LOC_Os03g61760.1 | 31,646 | 553 | 336 | 12 | 463 | 988 | 432 | 969 | 1.3E-77 | 277 | 63 |
| NM_115866 | Atspl12 | Sspon.006A0002630 | 31,553 | 599 | 336 | 16 | 364 | 926 | 272 | 832 | 9.82E-77 | 271 | 66 |
| NM_115866 | Atspl12 | Sspon.06G0001900-1A | 31,553 | 599 | 336 | 16 | 364 | 926 | 272 | 832 | 9.98E-77 | 271 | 66 |
| NM_106308 | Atspl16 | Ha412HOChr05g0226701 | 31,478 | 575 | 354 | 15 | 439 | 987 | 385 | 945 | 5.72E-78 | 278 | 91 |
| NM_106308 | Atspl16 | HanPSC8Chr05g0211201 | 31,478 | 575 | 354 | 15 | 439 | 987 | 385 | 945 | 5.72E-78 | 278 | 91 |
| NM_115866 | Atspl12 | AT1G76580.1 | 31,407 | 1073 | 472 | 36 | 25 | 926 | 40 | 1019 | 2.41E-117 | 386 | 97 |
| NM_106308 | Atspl16 | HanXRQChr05g0218781 | 31,359 | 574 | 356 | 14 | 439 | 987 | 385 | 945 | 3.27E-78 | 279 | 86 |
| NM_106308 | Atspl16 | Glyma.17G039600.1.p | 31,273 | 550 | 324 | 12 | 466 | 988 | 483 | 1005 | 2.63E-69 | 254 | 74 |
| NM_101951 | Atspl14 | Seita.5G030800.1.p | 31,162 | 568 | 340 | 16 | 500 | 1033 | 329 | 879 | 2.41E-72 | 261 | 67 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NM_101951 | Atspl14 | Seita.9G022800.1.p | 31,111 | 675 | 409 | 13 | 394 | 1035 | 317 | 968 | 6.9E-82 | 290 | 89 |
| NM_106308 | Atspl16 | Lsat_1_v5_gn_6_117840.1 | 31,065 | 573 | 332 | 18 | 448 | 987 | 389 | 931 | 1.52E-71 | 259 | 76 |
| NM_101951 | Atspl14 | LOC_Os03g61760.1 | 31,055 | 673 | 407 | 16 | 394 | 1035 | 323 | 969 | 3.07E-84 | 297 | 78 |
| NM_106308 | Atspl16 | Glyma.02G008600.1.p | 31,041 | 567 | 332 | 16 | 466 | 987 | 457 | 1009 | 4.72E-77 | 276 | 62 |
| NM_106308 | Atspl16 | Glyma.09G082000.1.p | 30,912 | 647 | 356 | 19 | 408 | 988 | 411 | 1032 | 4.87E-76 | 274 | 81 |
| AJ011644.1 | Atspl6 | Glyma.10G009200.1.p | 30,688 | 189 | 100 | 4 | 1 | 172 | 33 | 207 | 2.97E-18 | 86,7 | 82 |
| NM_124445.3 | Atspl13 | AC233751.1_FGP002 | 30,488 | 246 | 113 | 10 | 147 | 350 | 308 | 537 | 3.19E-13 | 74,7 | 87 |
| NM_106308 | Atspl16 | Seita.5G030800.1.p | 30,427 | 562 | 338 | 17 | 466 | 986 | 330 | 879 | 4.45E-67 | 246 | 68 |
| NM_106308 | Atspl16 | Glyma.10G009200.1.p | 30,293 | 581 | 340 | 13 | 448 | 987 | 442 | 998 | 2.84E-75 | 271 | 68 |
| NM_106308 | Atspl16 | Solyc05g053240.2.1 | 30,055 | 549 | 333 | 12 | 466 | 972 | 447 | 986 | 6.11E-73 | 265 | 71 |

**Supplementary File 1**. Sequences that passed BLASTp cutoffs (N=182), declared as SPL orthologs and merged to form the SPL dataset

>AT2G42200.1
MEMGSNSGPGHGPGQAESGGSSTESSSFSGGLMFGQKIYFEDGGGGSGSSSSGGRSNRRVRGGGSGQSGQIPRCQVEGCGMDL
TNAKGYYSRHRVCGVHSKTPKVTVAGIEQRFCQQCSRFHQLPEFDLEKRSCRRRLAGHNERRRKPQPASLSVLASRYGRIAPS
LYENGDAGMNGSFLGNQEIGWPSSRTLDTRVMRRPVSSPSWQINPMNVFSQGSVGGGGTSFSSPEIMDTKLESYKGIGDSNCA
LSLLSNPHQPHDNNNNNNNNNNNNNNNTWRASSGFGPMTVTMAQPPPAPSQHQYLNPPWVFKDNDNDMSPVLNLGRYTEPDNCQ
ISSGTAMGEFELSDHHHQSRRQYMEDENTRAYDSSSHHTNWSL*

>AT2G47070.1
MEARIDEGGEAQQFYGSVGKRSVEWDLNDWKWDGDLFLATQTTRGRQFFPLGNSSNSSSSCSDEGNDKKRRAVAIQGDTNGAL
TLNLNGESDGLFPAKKTKSGAVCQVENCEADLSKVKDYHRRHKVCEMHSKATSATVGGILQRFCQQCSRFHLLQEFDEGKRSC
RRRLAGHNKRRRKTNPEPGANGNPSDDHSSNYLLITLLKILSNMHNHTGDQDLMSHLLKSLVSHAGEQLGKNLVELLLQGGGS
QGSLNIGNSALLGIEQAPQEELKQFSARQDGTATENRSEKQVKMNDFDLNDIYIDSDDTDVERSPPPTNPATSSLDYPSWIHQ
SSPPQTSRNSDSASDQSPSSSSEDAQMRTGRIVFKLFGKEPNEFPIVLRGQILDWLSHSPTDMESYIRPGCIVLTIYLRQAET
AWEELSDDLGFSLGKLLDLSDDPLWTTGWIYVRVQNQLAFVYNGQVVVDTSLSLKSRDYSHIISVKPLAIAATEKAQFTVKGM
NLRQRGTRLLCSVEGKYLIQETTHDSTTREDDDFKDNSEIVECVNFSCDMPILSGRGFMEIEDQGLSSSFFPFLVVEDDDVCS
EIRILETTLEFTGTDSAKQAMDFIHEIGWLLHRSKLGESDPNPGVFPLIRFQWLIEFSMDREWCAVIRKLLNMFFDGAVGEFS
SSSNATLSELCLLHRAVRKNSKPMVEMLLRYIPKQQRNSLFRPDAAGPAGLTPLHIAAGKDGSEDVLDALTEDPAMVGIEAWK
TCRDSTGFTPEDYARLRGHFSYIHLIQRKINKKSTTEDHVVVNIPVSFSDREQKEPKSGPMASALEITQIPCKLCDHKLVYGT
TRRSVAYRPAMLSMVAIAAVCVCVALLFKSCPEVLYVFQPFRWELLDYGTS*

>AT1G53160.1
MEGKRSQGQGYMKKKSYLVEEDMETDTDEEEEVGRDRVRGSRGSINRGGSLRLCQVDRCTADMKEAKLYHRRHKVCEVHAKAS
SVFLSGLNQRFCQQCSRFHDLQEFDEAKRSCRRRLAGHNERRRKSSGESTYGEGSGRRGINGQVVMQNQERSRVEMTLPMPNS
SFKRPQIR*

>AT1G27370.1
MDCNMVSSFPWDWENLIMSNQSKTENEKKQQSTEWEFEKGEGIESIVPDFLGFEKVSSGSATSFWHTAVSKSSQSTSINSSSP
EDKRCNLASQSSPGDSSSNIDFLQVKPSTALEVPIASAESDLCLKLGKRTYSEEFWGRNNNDLSAVSMNLLTPSVVARKKTKS
CGQSMQVPRCQIDGCELDLSSSKDYHRKHRVCETHSKCPKVVVSGLERRFCQQCSRFHAVSEFDEKKRSCRKRLSHHNARRRK
PQGVFPLNSERVFDRRQHTSMLWNGLSLNTRSEEKYTWGTTYETKPTQMESGFTLSFQRGNGSEDQLFTGSTLSFSAFQTSGG
FSAGKSNIQLPDKGVGECSGGLHESHDFYSALSLLSTTSDSQGIKHTPVAEPPPIFGTFPSHFI*

>AT1G76580.1
MGEVGPQVAAPISVHHPMGKKRDLYVQPHWLVSSQLPKDDWQMNRWKWDGQRFEAIELQGESLQLSNKKGLDLNLPCGFNDVE
GTPVDLTRPSKKVRSGSPGSGGGGGGNYPKCQVDNCKEDLSIAKDYHRRHKVCEVHSKATKALVGKQMQRFCQQCSRFHLLSE
FDEGKRSCRRRLDGHNRRRRKTQPDAITSQVVALENRDNTSNNTNMDVMALLTALVCAQGRNEATTNGSPGVVPQREQLLQILN
KIKALPLPMNLTSKLNNIGILARKNPEQPSPMNPQNSMNGASSPSTMDLLAALSASLGSSAPEAIAFLSQGGFGNKESNDRTK
LTSSDHSATTSLEKKTLEFPSFGGGERTSSTNHSPSQYSDSRGQDTRSSLSLQLFTSSPEEESRPKVASSTKYYSSASSNPVE
DRSPSSSPVMQELFPLHTSPETRRYNNYKDTSTSPRTSCLPLELFGASNRGATANPNYNVLRHQSGYASSGSDYSPPSLNSNA
QERTGKISFKLFEKDPSQLPNTLRTEIFRWLSSFPSDMESFIRPGCVILSVYVAMSASAWEQLEENLLQRVRSLVQDSEFWSN
SRFLVNAGRQLASHKHGRIRLSKSWRTLNLPELITVSPLAVVAGEETALIVRGRNLTNDGMRLRCAHMGNYASMEVTGREHRL
TKVDELNVSSFQVQSASSVSLGRCFIELENGLRGDNFPLIIANATICKELNRLEEEFHPKDVIEEQIQNLDRPRSREEVLCFL
NELGWLFQRKWTSDIHGEPDFSLPRFKFLLVCSVERDYCSLIRTVLDMMVERNLGKDGLLNKESLDMLADIQLLNRAIKRRNT
KMAETLIHYSVNPSTRNFIFLPSIAGPGDITPLHLAASTSSSDDMIDALTNDPQEIGLSCWNTLVDATGQTPFSYAAMRDNHS
YNTLVARKLADKRNGQISLNIENGIDQIGLSKRLSSELKRSCNTCASVALKYQRKVSGSRRLFPTPIIHSMLAVATVCVCVCV
FMHAFPMVRQGSHFSWGGLDYGSI
*

>AT1G27360.1
MDCNMVSSSQWDWEHLIMSNPSRTEDDSKQLPTEWEIEKGEGIESIVPHFSGLERVSSGSATSFWHTAVSKSSQSTSINSSSP
EAKRCKLASESSPGDSCSNIDFVQVKAPTALEVSVASAESDLCLKLGKRTYSEEYWGRNNNEISAVSMKLLTPSVVAGKSKLC
GQSMPVPRCQIDGCELDLSSAKGYHRKHKVCEKHSKCPKVSVSGLERRFCQQCSRFHAVSEFDEKKRSCRKRLSHHNARRRKP
QGVFSMNPERVYDRRQHTNMLWNGVSLNARSEEMYEWGNNTYDTKPRQTEKSFTLSFQRGNGSEDQLVASSSRMFSTSQTSGG
FPAGKSKFQLHGEDVGEYSGVLHESQDIHRALSLLSTSSDPLAQPHVQPFSLLCSYDVVPK*

>AT1G02065.1

MLDYEWDNPSSIVLSGDERNPDSDPTRSSFSFFDPISHYNNDHRHITISPPLLSSFSNQQQQHHLTLYGQTNSNNQFLHHHHH
HHSLYGSTTTTTPYGASDPIYHPHSSAPPASLFSYDQTGPGSGSGSSYNFLIPKTEVDFTSNRIGLNLGGRTYFSAADDDFVS
RLYRRSRPGESGMANSLSTPRCQAEGCNADLSHAKHYHRRHKVCEFHSKASTVVAAGLSQRFCQQCSRFHLLSEFDNGKRSCR
KRLADHNRRRRKCHQSASATQDTGTGKTTPKSPNDSGVKASSSPSSNAPPTISLECFRQRQFQTTASSSTSASSSSNSMFFSS
G*


>AT1G20980.1
MDEVGAQVAAPMFIHQSLGRKRDLYYPMSNRLVQSQPQRRDEWNSKMWDWDSRRFEAKPVDVEVQEFDLTLRNRSGEERGLDL
NLGSGLTAVEETTTTTQNVRPNKKVRSGSPGGNYPMCQVDNCTEDLSHAKDYHRRHKVCEVHSKATKALVGKQMQRFCQQCSR
FHLLSEFDEGKRSCRRRLAGHNRRRRKTTQPEEVASGVVVPGNHDTTNNTANANMDLMALLTALACAQGKNAVKPPVGSPAVP
DREQLLQILNKINALPLPMDLVSKLNNIGSLARKNMDHPTVNPQNDMNGASPSTMDLLAVLSTTLGSSSPDALAILSQGGFGN
KDSEKTKLSSYENGVTTNLEKRTFGFSSVGGERSSSSNQSPSQDSDSRGQDTRSSLSLQLFTSSPEDESRPTVASSRKYYSSA
SSNPVEDRSPSSSPVMQELFPLQASPETMRSKNHKNSSPRTGCLPLELFGASNRGAADPNFKGFGQQSGYASSGSDYSPPSLN
SDAQDRTGKIVFKLLDKDPSQLPGTLRSEIYNWLSNIPSEMESYIRPGCVVLSVYVAMSPAAWEQLEQKLLQRLGVLLQNSPS
DFWRNARFIVNTGRQLASHKNGKVRCSKSWRTWNSPELISVSPVAVVAGEETSLVVRGRSLTNDGISIRCTHMGSYMAMEVTR
AVCRQTIFDELNVNSFKVQNVHPGFLGRCFIEVENGFRGDSFPLIIANASICKELNRLGEEFHPKSQDMTEEQAQSSNRGPTS
REEVLCFLNELGWLFQKNQTSELREQSDFSLAR
FKFLLVCSVERDYCALIRTLLLDMLVERNLVNDELNREALDMLAEIQLLNRAVKRKSTKMVELLIHYLVNPLTLSSSRKFVFLP
NITGPGGITPLHLAACTSGSDDMIDLLTNDPQEIGLSSWNTLRDATGQTPYSYAAIRNNHNYNSLVARKLADKRNKQVSLNIE
HEVVDQTGLSKRLSLEMNKSSSSCASCATVALKYQRRVSGSQRLFPTPIIHSMLAVATVCVCVCVFMHAFPIVRQGSHFSWGG
LDYGSI*


>AT1G69170.1
MDSWSYGRSVFMSNETLLPCDTFAKNRRFEQRLSNNDDVLISDMAGNSNGFSAVSITKVVPEEEDEENISSSSKFSSQELNRI
DFKLRSFLDLGNDDDDTSSRGFALPSKKSRASNLCSQNPLCQVYGCSKDLSSSKDYHKRHRVCEAHSKTSVVIVNGLEQRFCQ
QCSRFHFLSEFDDGKRSCRRRLAGHNERRRKPAFYFLPGKRHKLLRTSQDVVGNKFLENSSLVLPESFPGSLLYRVIDEDDHR
TSRLVSFKDEPTCSMFPTNEQNSSRTYESKPAIYSTEVSSIWDLHETAASRSTRALSLLSAQSQQHLSKFPNTTFSITQPNQN
LNHSSSTDYHQMEQPLWIDPGKTNSAGSSSCKGKGTSTVDLLQLSSHLQRIEQQRNYTGDVKQEYNELYFPGS*


>AT3G60030.1
MEARIEGEVEGHSLEYGFSGKRSVEWDLNDWKWNGDLFVATQLNHGSSNSSSTCSDEGNVEIMERRRIEMEKKKKRRAVTVVA
MEEDNLKDDDAHRLTLNLGGNNIEGNGVKKTKLGGGIPSRAICCQVDNCGADLSKVKDYHRRHKVCEIHSKATTALVGGIMQR
FCQQCSRFHVLEEFDEGKRSCRRRLAGHNKRRRKANPDTIGNGTSMSDDQTSNYMLITLLKILSNIHSNQSDQTGDQDLLSHL
LKSLVSQAGEHIGRNLVGLLQGGGGGLQASQNIGNLSALLSLEQAPREDIKHHSVSETPWQEVYANSAQERVAPDRSEKQVKVN
DFDLNDIYIDSDDTTDIERSSPPPTNPATSSLDYHQDSRQSSPPQTSRRNSDSASDQSPSSSSGDAQSRTDRIVFKLFGKEPN
DFPVALRGQILNWLAHTPTDMESYIRPGCIVLTIYLRQDEASWEELCCDLSFSLRRLLDLSDDPLWTDGWLYLRVQNQLAFAF
NGQVVLDTSLPLRSHDYSQIITVRPLAVTKKAQFTVKGINLRRPGTRLLCTVEGTHLVQEATQGGMEERDDLKENNEIDFVNF
SCEMPIASGRGFMEIEDQGGLSSSFFPFIVSEDEDICSEIRRLESTLEFTGTDSAMQAMDFIHEIGWLLHRSELKSRLAASDH
NPEDLFSLIRFKFLIEFSMDREWCCVMKKLLNILFEEGTVDPSPDAALSELCLLHRAVRKNSKPMVEMLLRFSPKKKNQTLAG
LFRPDAAGPGGLTPLHIAAGKDGSEDVLDALTEDPGMTGIQAWKNSRDNTGFTPEDYARLRGHFSYIHLVQRKLSRKPIAKEH
VVVNIPESFNIEHKQEKRSPMDSSSLEITQINQCKLCDHKRVFVTTHHKSVAYRPAMLSMVAIAAVCVCVALLFKSCPEVLYV
FQPFRWELLEYGTS*


>AT3G57920.1
MELLMCSGQAESGGSSSTESSSLSGGLRFGQKIYFEDGSGSRSKNRVNTVRKSSTTARCQVEGCRMDLSNVKAYYSRHKVCCI
HSKSSKVIVSGLHQRFCQQCSRFHQLSEFDLEKRSCRRRLACHNERRRKPQPTTALFTSHYSRIAPSLYGNPNAAMIKSVLGD
PTAWSTARSVMQRPGPWQINPVRETHPHMNVLSHGSSFTTCPEMINNNSTDSSCALSLLSNSYPIHQQQLQTPTNTWRPSSG
FDSMISFSDKVTMAQPPPISTHQPPISTHQQYLSQTWEVIAGEKSNSHYMSPVSQISEPADFQISNGTTMGGFELYLHQQVLK
QYMEPENTRAYDSSPQHFNWSL*


>AT3G15270.1
MEGQRTQRRGYLKDKATVSNLVEEEMENGMDGEEEDGGDEDKRKKVMERVRGPSTDRVPSRLCQVDRCTVNLTEAKQYYRRHR
VCEVHAKASAATVAGVRQRFCQQCSRFHELPEFDEAKRSCRRRLAGHNERRRKISGDSFGEGSGRRGFSGQLIQTQERNRVDR
KLPMTNSSFKRPQIR*


>AT5G43270.1
MECNAKPPFQWELENLISFGTSTAEVPRKLKPMEWEIDGFDCTSLYSSSFAYAGSSGSDIAHAFSKSSKSTSISSSSAEVRTH
NFTSETGESLPGEFAKGIDTSPSLELSFGSGDPVLGLKLGKRTYFEDFWEVENAKGLGLPVTLASSSVSPVKKSKSIPQRLQT
PHCQVEGCNLDLSSAKDYHRKHRICENHSKFPKVVVSGVERRFCQQCSRFHCLSEFDEKKRSCRRRLSDHNARRRKPNPGRTY

DGKPQVDFVWNRFALIHPRSEEKFIWPSSKHVPSRVLMPQPAKTEISDTEHNRFGLLDPKTKTARAELFSKEKVTISSHMGAS
QDLDGALSLLSNSTTWVSSSDQPRRFTLDHHPSSNLQPVAHRSAAQLNSVSGYWQPDPPAVEGPTALHRNGVGQFNENYFSLN
QFYN*

>AT5G18830.3
MSSLSQSPPPPEMDIQPPALVNDDPSTYSSALWDWGDLLDFAADERLLVDQIHFPPVLSPPLPPLIPTQTPAESELDPSPEES
GSGSDRVRKRDPRLICSNFIEGMLPCSCPELDQKLEDAELPKKKRVRGGSGVARCQVPDCEADISELKGYHKRHRVCLRCATA
SFVVLDGENKRYCQQCGKFHLLPDFDEGKRSCRRKLERHNNRRKRKPVDKGGVAAEQQQVLSQNDNSVIDVEDGKDITCSSDQ
RAEEEPSLIFEDRHITTQGSVPFTRSINADNFVSVTGSGEAQPDEGMNDTKFERSPSNGDNKSAYSTVCPTGRISFKLYDWNP
AEFPRRLRHQIFQWLANMPVELEGYIRPGCTILTVFIAMPEIMWAKLSKDPVAYLDEFILKPGKMLFGRGSMTVYLNNMIFRL
IKGGTTLKRVDVKLESPKLQFVYPTCFEAGKPIELVVCGQNLLQPKCRFLVSFSGKYLPHNYSVVPAPDQDGKRSCNNKFYKI
NIVNSDPSLFGPAFVEVENESGLSNFIPLIIGDAAVCSEMKLIEQKFNATLFPEGQEVTACSSLTCCCRDFGERQSTFSGLLL
DIAWSVKVPSAERTEQPVNRCQIKRYNRVLNYLIQNNSASILGNVLHNLETLVKKMEPDSLVHCTCDCDVRLLHENMDLASDI
HRKHQSPIESKVNPPSSGCCCVSSQKDIPSRILNFNKDPEAGLDCKERIQADCSPDSGGKETDPLLNKEVVMNVNDIGDWPRK
SCIKTHSALAFRSRQTMFLIATFAVCFAVCAVLYHPNKVTQLAVAIRMRLCSTELSHNSFSESFPSERDRI*

>AT5G50670.2
MDWNFKLSSGYLSGFDQEPDLSPMDGSISFGGSSQSKADFSFDLKLGRNIGNSSSVFGDTEQVISLSKWKDSALAKPEGSRSS
SSKRTRGNGVGTNQMPICLVDGCDSDFSNCREYHKRHKVCDVHSKTPVVTINGHKQRFCQQCSRFHALEEFDEGKRSCRKRLD
GHNRRRRKPQPEHIGRPANFFTGFQGSKLLEFSGGSHVFPTTSVLNPSWGNSLVSVAVAANGSSYGQSQSYVVGSSPAKTGIM
FPISSSPNSTRSIAKQFPFLQEEESSRTASLCERMTSCIHDSDCALSLLSSSSSSVPHLLQPPLSLSQEAVETVFYGSGLFEN
ASAVSDGSVISGNEAVRLPQTFPFHWE*

>AT5G50570.2
MDWNFKLSSGYLSGFDQEPDLSPMDGSISFGGSSQSKADFSFDLKLGRNIGNSSSVFGDTEQVISLSKWKDSALAKPEGSRSS
SSKRTRGNGVGTNQMPICLVDGCDSDFSNCREYHKRHKVCDVHSKTPVVTINGHKQRFCQQCSRFHALEEFDEGKRSCRKRLD
GHNRRRRKPQPEHIGRPANFFTGFQGSKLLEFSGGSHVFPTTSVLNPSWGNSLVSVAVAANGSSYGQSQSYVVGSSPAKTGIM
FPISSSPNSTRSIAKQFPFLQEEESSRTASLCERMTSCIHDSDCALSLLSSSSSSVPHLLQPPLSLSQEAVETVFYGSGLFEN
ASAVSDGSVISGNEAVRLPQTFPFHWE*

>comp199491_c0_seq1_ORF1
MEAGFLWKESGDQASAARMTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPVED
GSADCSNSAVADGDNNHVVSVQRGRSSEEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTRTTVVCIKNV
EHRFCQQCSRFHLLQEFDEGKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSTLFLLLKQLSGLEPGSSSEQINHPNS
LVNILKNLAAIASTQSYQDILKNANSNLILSNAGNNAANGSIVHEQTIRSTPIGRESLAEEPPVKRRVQDFDLNDACIEEAES
RTDKIVFKLFGKEPKDFPVDLREQIQDWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISLSNDGLWRK
GWLYARVQDCLTLSCNGSLMFASPWQPVIGDKLQRLCVTPIAVACSSSVKFSVQGFNIVQPTTKLLCVFGEKYLIQEETQLLL
EDSTMQQGPQCLTFSCSFPCTSGRGFIEVEDYDQSSLSVPFVVTDNDVCSEIRTLEHGLNLISFDETSERIDDLMAYRSRSLH
FLHEIGWLLQRSHMRATSEQPQYCPDRFPVARFRWLLSFAIDHEWCAVVTKLLNTMFQGDIDLDVPSPIEFALGENLLLTAVR
KCSKPLVECLLRYTATNYAPVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGF
TPEEYARQRGHVSYIQMVQDKIDRRVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAF
FHPRTNRFLSNRPAMLSLVAIAAVCVCVGLLMKSPPQVACMRPFLWDHIRWGPN

>gg_15925_ORF1
MAPPPLPLPLPMPLQMPGMAGMPGDHVYPALGMVKREGGADGAGAGRIGLNLGRRTYFSPGDMLAVDRLLMRSRLGGVFGLGF
GGVHHQPPRCQAEGCKADLSGAKHYHRRHKVCEYHAKASVVAAGGKQQRFCQQCSRFHVLSEFDEVKRSCRKRLAEHNRRRRK
PATTASSKDAASPPANKKPNGGSITSSYSTDNKNLSTAKSTMSSNTSSVISCLDQAGNKQQLARPTLTLGASQDKDHHQQQLS
TMLQVQAASGHHQEQHFITSLQVHNNGGSGIGGGNNNNILSCSSVCSSGALPSANGEVSDQNTTTTNTGNGGSSNNMHNLFEV
DFM

>gg_10666_ORF1
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTLSSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVESHVN
NTGTNGRVDDSGNSRSSMIAFSQGEPLISLKLGKRAYFENVCSGQDAKVSAPSDVTSTATVVKKTKVSQQNAKNWYCQVEGCK
VDLSSAKDYNRKHKVCVVHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLTHHNARRRKPQADTISFNSARLSTM
FYDTSQRSNLFFSQPLYGQVRSTAV

>gg_07486_ORF1
LSLQLFGSTEEDIPPKIDSLNKYLSSESSNPLDERSPSSSPPITRKFFPIHSVDEEVRHPHITDYGEDATMGEVSTSQAWCAP
PLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQDRTGRIIFKLFGKEPSTIPGNLRDDIVNWLKHSPTEMEG

YIRPGCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRTDSQLVSYKAGMTRLSKSWRTWNTPELTFVSP
IAVVGGRKTSLVLKGRNLSIPGTQIHCTSIGKYISKEVLCSAYPGTIYDDSGVETFDLPGQPDLILGRCFIEVENRFRGNSFP
VIVASSSVCLELRNLEVELEDSQVLDVSSDGQIHDSPQSKTRNQVLHFLNELGWLFQRASACTLSSRSDVSDLDLIQFSTARF
KYLLLFSNERDWCSLTKTLLEILAKRSLASEELSKETMEMLAEIHLLNRAVKRKSRSMVHLLVQFVVICPDNSKVYPFLPNLP
GPGGLTPLHLAASIENAEDIVDALTDDPQQIGLTYWQSVLDDDGQSPETYAKLRNHNSYNELVAQKLVDMKNNQVTITVNGDE
IHMDQLGNVGDHKKSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPYIHSMLAIAAVCVCVCVFMRALLRINSGKSFKWE
RLDFGTI

>comp203535_c0_seq1_ORF1
MVAAAEVEDVATEFLAGARKKTKTAARRGKAAAAGAAAGAAGGGAVRAAVAEMKCQVPGCEADIRELKGYHRRHRVCLRCAHA
SAVMLDGVQKRYCQQCGKFHVLLDFDEDKRSCRRKLERHNKRRRRKPDSKGTLDKEIDDQLDLSADVSGDGELREENMEGTTS
EMLETVLSNKVLDRETPVGSEDVLSSPTCTQPGLQNDQSKSLQNDQSKSVVTFAASVEACIGAKQESVKLANSPMHDTKSAYS
SSCPTGRISFKLYDWNPAEFPRRLRHQIFEWLASMPVELEGYIRPGCTILTVFIAMPQHMWDKLSDDAADLLRNLVNSPNSLL
LGKGAFFIHVNNMLFQVLKDGATLMSTRLDVQAPRIDYVHPTWFEAGKPVDLILYGSSLDQPNFRSLLSFDGDYLKHDCYRLT
SHDTFDHIKNGDLIPDSQHEIFRIKITQSRPDIHGPAFVEVENIFGLSNFVPILFGSKQLCSELERIQDALCGSYSKNNVLGE
LLSASSDPHERRKLHSSVMSGFLIDIGWLIRNPTPDEFKNVLSSTNIQRWICILKFLIQNDFINVLEIIVKSMDSIMGSEVLS
NLERGRLEDHVTAFLGYVSHARNIVDRRANHDEKTQMDTGGISVNSPNQPSLGASLPLASENTDIGGDNKLNSADEEETMPLV
TRDVSRRQCCQPDMTAGWLKPSLIVSYPGGATRMRLVMTVVVAAVLCFTACLVLFHPHGVGVLAAPVKRYLSSDSAS

>gg_04693_ORF1
VIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKKAKAAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCE
PHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARRRKPQPEALSFGSSRLSTMFYDTRQQTSLLFGQGPY
GQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHISSQQVWNNFMPHGVHQDFDGFTALKGTSAKVLDQGVEASAAVSNP
NGAPDLQHALSLLSNSSAGAGTNQQPTAQLHHHHAGLSTLVASSSSVMQASSPGLWQDGGGAALGHHAHAHAHTRFQALDATG
SGAAIATTAQELLQLPRPPSSYDSSSHYDLMR

>comp204585_c2_seq1_ORF1
MFVLKVVHCSQICRFHMLSEFDEGKRSCRKRLDGHNRRRRKPQHDLTNLGGFFPYHQVNQFEVYPRTIPTVGQNSDAMQLVGR
QQPFSISFSRTPNQFPFPQGGGSALHAARRGLLAEGSSHTGSSTCNNALSGTLGPECALSLLSPLLHRPSAAGIPTDGQPQVA
SSLSRIAAVSQAATTAMTIAFAAGVGHHVFVPDAMFEDPSQALPFSWQ

>gg_12734_ORF1
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPTRRGGNEREAKRARGEDGGRGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFDDRRQISFVWNKAPV
SHVRPFTSPWDSSSDLKLPHAKETREVSTKVGTITGQVHLDKSHMSSAIPTLSHGKDEPWPMKGPDMSITASKFDGAPDLQHA
LSLLSVGACGLPDPVHQTSCVIQFTGASENSSDLHLSHGGNSGPASCADGQHIATQPQSQLVRFTMDTSNNVYEPSFFGVNQI
N

>Glyma.10G009200.1.p
MEARFGAEAYHFCGVGASSDLRGVGKRSSEWDLNDWRWDGDLFIASRLNPVVPADGVGVGRQFFPTGSGTGILVAGGPSNSSS
TSEEVDPRDPKGNKEGDKKRRVIVLEDDGLNEEGGTLSLKLGGHASAVVDREVASWDGTNGKKSRVSGSTSNRAVCQVEDCSA
DLSKAKDYHRRHKVCEMHSKASRALVGNAMQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNKRRRKTNHEAVPNGSSLNDDQT
SSYLLISLLKILSNMHSDRSDQTTDQDLLTHILRSLASQNGEQGSKNIANLLREPENLLREDGSSRKSEMMSTLFSNGSQGSP
TDTRQHETVSIAKMQQQVMHAHDARAADQQITSSIKPSMSNSPPAYSEARDSTAGQIKMNNFDLNDIYIDSDDGMEDLERLPV
STNLVTSSLDYPWAQQDSHQSSPPQTSGNSDSASAQSPSSFSGEAQSRTDRIVFKLFGKEPNDFPLVLRAQILDWLSHSPTDM
ESYIRPGCIVLTIYLRQAEALWEELCYDLTSSLNRLLDVSDDTFWRNGWVHIRVQHQMAFIFNGQVVIDTSLPFRSNNYSKIL
TVSPIAAPASKRAQFSVKGVNLIRPATRLMCALEGKYLVCEDAHMSMDQSSKEPDELQCVQFSCSVPVMNGRGFIEEDVCSEI
CTLEPLLELSETDPDIEGTGKIKAKNQAMDFIHEMGWLLHRSQLKLRMVQLNSSEDLFPLKRFKWLIEFSMDHDWCAAVRKLL
NLLLDGTVNTGDHPSLYLALSEMGLLHKAVRRNSKQLVECLLRYVPENISDKLGPEDKALVDGENQTFLFRPDVVGPAGLTPL
HIAAGKDGSEDVLDALTNDPCMVGIEAWKNARDSTGSTPEDYARLRGHYAYIHLVQKKINKRQGAAHVVVEIPSNTTESNTNE
KQNELSTTFEIGKAEVIRGQGHCKLCDKRISCRTAVGRSLVYRPAMLSMVAIAAVCVCVALLFKSSPEVICMFRPFRWENLDF
GTS*

>Glyma.17G080700.1.p
MESWSYIPEEKECVSNEAPSPPNTLGRSKNSFLGWELKTPCGFSTDMLGLGHHSIENQGFEELGFPEMLGKHMSDDLVASVPT
RKVDGNRQQQQHHSDRITGTVVDAPSGFSQRGDSNSLIDLKLGRFAEHGDFTDPAFSKVLSSSESSTPPKRVRTSGLHSQTAY
CQVYGCNKDLSSCKDYHKRHKVCEVHSKTAIVIVNGIEQRFCQQCSRFHLLAEFDDGKRSCRKRLAGHNERRRKPQMGIHPGK
SGTLLQPCGDSRFQGSMLTSSFIRPEMLPSGVMCSEKYGMSDFWRPIKAEHGTGFRNLSSMAATNGHPLPRSLFSSYNGKQFP
FLHENSSTSTTGSIFCESNSHYPSALGAQNSGLRPLFQDTVGNEDFSVFDTSPNVQGLSGISDSCCALSLLSSQSQNSSSQSS

GIPLAHSLVIPSSHSHHYNMSQVSDKIGISSQTSSSRVSDSFPPELNHADGSHLSPVLISGNNDIINFDMADGIFQGSDYLNV
KDRLSDEDGATIDLLQLSSQLQRVEHERQSLQVKQENDSSCTLRIT*

>Glyma.17G039600.1.p
MEAEFGAKNQYLYGPVVSGMKKAVVGNGKRSLEWDLNDWRWDGDLFTAQPLNSVPSDCRGRQFSPPHPEIPAKNANPSTTNLS
YSVFISGEGKRELEKRRGVIGEGEGEGLNYEAGSLSLNLAGQGYPLMVEEEEKSGKKTKVIGTTTASTTTASNRAVCQVQDC
RADLSNAKDYHRRHKVCDIHSKATMALVGNVMQRFCQQCSRFHVLQEFDEGKRSCRRRLAGHNRRRRKTHPDVSVVNEGSLND
QRGSNYLLMSLLRILTNLHSNGSDHTRNQDILSHLLKNLASLAGPNNGGRLAPLLEESKDLVNAGTPGAAQDKPNLNSNAPEA
SRPSSFIKTDNVIITQDPTMPVAQYETPADGRTQKCIASGDGGVGNSKPPSRPLFSNVCQPRDSLPSQLTTAETKVGRGNLNN
IDLNNGYNDIQNTVENHKKPYPPVASVMGSIDRASWLQRDSLKSSPPQTSRNSDSTSIQSPSSSSGEAQSRTDRIVFKLFGKD
PSDFPLLIRSQILNWLSHSPTEMESYIRPGCIILTIYLRLEKSAWEELYCNLGSSLRKLLAESNDSFWRTGWVYARVQHSVAF
LYNGQVVLDVPLHLKSPQHCRISCIKPLAVPASASAQFIVKGFNLFQSSTRLLCALEGKYLVHASCHGLIGGADAPIQHLSFS
CHIPNVTGRGFIEVEDHGLSSCSFPFIVAEQEVCSEICKLENVIEAAETTDDIQIKKQQMEEMTRALDFLQEMGWLLHRSHVK
VKLGPMAPFRDLFQFNRFAWLVDFSMDHGWCAVMNKLLDIIFEGGVDAGEHASIELALLNMGLLHTAVKRNCRPMVELLLRFV
PVKTSDGADSEVKQVDEASDRFLFRPDTVGPAGLTPLHVAASMSGLEDVLDALTDDPRMLGIEAWKSSRDSTGLTPNDYACLR
GYYSYIQLVQKKTNKKGERQHVVDITGIVVDSYTTQKSNGHRTSMVYWPVMLTMVTIAAVCVCVALLFKSSPRVYYVFQPFN
WESLEYGAM*

>Glyma.09G113800.1.p
MSSGSVTKAPPPSSSSSPNSSSESLDGLKFGQKIYFEDVSVAATTTTQGAYKTSRVVSSPSSASKKGRGGSLQPPRCQVEGCK
VDLSGAKAYYSRHKVCTMHSKFPTVIVAGLEQRFCQQCSRFHLLSEFDEGKRSCRRRLAGHNERRRKPPPSSTLASHYGRLST
PIFDNSGRAGGFLMEFASYPKVTLRNTLPTQRSAEPALGNHAATLTWQGSSDTPSEFFLQESGSGGGTSFLGSKHPLLESYTGV
TDSNCALSLLSRQTWGSRNTAPATSVELNNNLLNFNGTSMTQFSAASSQTAAIHQLPNATSWLYLKGVGSGDCSPEGVPDLGL
GQISPPLDSHLHGELNVSQQGRRHYMDLGESRACESSHWSL*

>Glyma.09G082000.1.p
MESQLEGKNQYLYGPVVPEMKSVGKRSLEWDLNDWKWDGDLFTARQLNSVPSDCRSRELFPTDPEILATGGASNSLSSAYDDA
NLGEGKRELEKRRRGVNDDGGVEMNDGAGSLNLNLGGQVYPIMEGEEKSGKKTKMTASTSSRAVCQVEDCRADLSNAKDYHRR
HKVCDMHSKATQALVGNVMQRFCQQCSRFHVLQEFDEGKRSCRRRLAGHNKRRRKTLPDATVVNGGSLNEEKGSSYLLMSLLR
ILSNMHSNGSDNMRNQDVLSHLLRNLASLAGTINGRNIVSLLEGSQGLVKAGTSGAAQNVPNTNSDGPEPSRPFDSFIKMDDG
LIHRDPPESMVQRETTPANDMAKKCIASGSDGVGSLKSPSVPQSSNVLLSRDGLPPQSVAAQTTVGRIGLSNIDLNNVYDDVQ
DYVENTRNSCPPLPSGNGSLDHPLWIQCDSLKSSPPQTSRNSDSTSTQSPSSSSGEAQSRTDRIVFKLFGKAPNDFPHALRSQ
ILNWLSHSPTEIESYIRPGCIILTIYLRLENSAWEELCYNLESSLRKLAAPNDSFWRTGWIYTRVQHSVAFLYNGQVVLDAPL
RLKSPQNCQILCVKPLAVSASSSAQFVVKGFNFLLSNTRLLCALEGKYLVQDSCYDLIDSADAVNGHQELQHLSFSCHVPNVT
GRGFIEVEDNGLSSCSFPFIVAEQEICLEICTLDNVIEAAEMADDNQIKTNLMEEKTQALYFIQEMGWLLHRSRVKVRLGPMA
PVQDRFHFNRFIWLVGFSMDHDWCAVMKKLLNIIFEGTVDTGDHASVELALLEMGLLHKAVKRNCRPMVEILLKFVPVKASDG
GDSNEKQVNKSPDRFIFRPDTVGPVGLTPLHVAASMHGSENVLDALTDDPGMVGTEAWKSAQDATGLTPYDYASMRGYYSYIQ
LVQSKTSNTCKSQHVLDIPGTLVDSNTKQKQSDRHRSSKVSSLQTEKIETTAMPRRCGLCQQKLAYGGMRRALVYRPAMLSMV
AIAAVCVCVALLFKSSPKVYYVFQPFSWESLEYGSI*

>Glyma.04G159600.1.p
MESWSYVAEEKGYVSSEALSPPNTLERSKGSCFGWELKTPCSFSNDMLGLGQQNVEDQGFGELGFPGMLGKQLCGDDPNPIGA
VQSRKVDDVVVDCRITSPSPVVAAPNGFSSREDFNSKLSNALVDSNGGNSLIDLKLGWFADHGDAMVPTFSKGAPILSSSELS
TPPKRVRASGAHSQTAYCLVYGCNKDLSGCKDYHKRHKVCEVHSKTSKVIVNGIEQRFCQQCSRFHLLAEFDDGKRSCRKRLA
GHNERRRKPQVGIHSGKAGRLLQPYADSRFQETMLTSSFICQEILPSGVLSSEKYGASDMWRPIKAEDSTGFTFRPLSSIPIT
NGHPQSRSLFPSYSGKQLPFLHENGATSATGNIFSENNNQYNPPSLGGQSSGSQSLFQDSSLGREDFNVFDTESAVQGLSGIS
DSGCALSLLSSQSQNSSGQSSGIPIAHPMVIPSSHSHYNMSRVSEKMIGISSQASSGVSDRFPPEFNPVDGSHLSPILISDNN
EIINFEMANGIFQGSDFVNVKDRLSCEDSLTIDLLQLSSQLQRVENQRLSLQVKQENDSSCSLRIT*

>Glyma.04G197100.1.p
MEEQNRFGVYRMKGEFSVDLKLGHHVGNSGTESALANKSKDAAAAGVSKMASSPSGSSKRARAINSTSLTVACLVDGCNSDLS
NCRDYHRRHKVCELHSKTPEVTIGGFKQRFCQQCSRFHSLEQFDERKRSCRKRLDGHNRRRRKPQPEPLSRPGSFLSNYQGTQ
LLPFSSPHVYPSTAIVSPAWSASLVTSSADARLHNHNQQQHHHHQQDLFLGSSHKEGKQVAFLHNNHPSTAAALSSQTPPPFP
SSSVRHTLLRTDPLSETCGGARSKMFCESLTTTTTNNNSVHDTTPCALSLLSSSQTHTPGNGLNQMVQPPPSMSLMQPLGLSL
QDNNSLESVDHCSSMYSLASDHGSQCNEAAPQLFPFQWE*

>Glyma.07G109600.1.p
MLDYEWGNPSTIMLTGNEDSSAAGSDQTHRQIFDHYASHTLIPDHFLHAPMHSTATTAAATTVDFTHHHHFNPQPHHHNHMNHH
HHHHAFFDPRAFHGASNTSYPPPPSMLSLDPLPHVNAVGSGPGHGPGPGPGPGGLLLVPKSEEVGRPMDFVGSRIGLNLGGRT

```
YFSSSEDDFVSRLYRRSRPVEPGSTISSNSPRCQAEGCNADLSQAKHYHRRHKVCEFHSKAATVIAAGLTQRFCQQCSRFHLL
SEFDNGKRSCRKRLADHNRRRRKTQQPTQEIQKSQPSLDNVARSPPDSGAQSSSSVTVAVSPPDYFRQRPYQSTNNPSSISSS
LFFSSG*
```

>Glyma.07G234200.1.p
```
MEAEFGGKNQYLYGPVVSGMKKAVVGNGKRSLEWDLNDWRWDGDLFTAQPLNSVPSDCRGCQFFPPHPEIPAKNANPSTTNLS
SSVFILGEGKRELEKRRRDVIAEGEGEGLNDEGGSLSLNLGGQGYPLMLEEEEKSGKKTKVIGTNTTTTTTTTSNRAVCQVQD
CRADLSNAKDYHRRHKVCDVHSKATMALVGNVMQRFCQQCSRFHVLQEFDEGKRSCRRRLAGHNRRRRKTHPDVSVVNEGSLN
DQRDSNYLLMSLLRILTNLHSNGSDHTRNQDILSHLLKNLASLAGPNNGGRLAPLLEESKGLVNAGTHGADHDKPNLNSNAPE
ASRPSSSIKTDNGIIAQDPPMSVVQYETPANGMTQKCIASGDGVGNLKPPSGPLLSNVCEPRDSVPSQLTTAETKVGRGNLNN
IDLNNVYNDIQNTVENHKKPYPPVASGMGFIDHASWLQCDSLKSSPPQTSRNSDSTSTQSPSSSSGEAQSRTDRIVFKLFGKD
PSDFPLLLRSQILNWLSRSPTEIESYIRPGCIILTIYLRLEKSAWEELYCNLGSSLRKLLAASNDSFWRTGWVYARVQHAVAF
LYNGQVVLDVPLRLKSPQHCMISCINPLAVPASASAQFIVKGFNLSQSSTRLHCALEGKYLVHASCHDLIGGADAPIQHLSFS
CQIPSVTGRGFIEVEDHGLSSCSFPFIVAEQEVCSEICKLENVIEEAETTDDIQIKNQHMEEKTQALDFLQEMGWLLHRSHVK
FKLGSMAPFHDLFQFNRFAWLVDFSMDHGWCAVMKKLLDIIFEGGVDAGEHASIELALLNMGLLHRAVKRNCRPMVELLLRFV
PVKTSDGADSEMKQVAEAPDRFLFRPDTVGPAGLTPLHVAASMSGSENVLDALTNDPRMVGIEAWKSARDSTGLTPNDHACLR
GYYSYIQLVQNKTNKKGERQHLVDIPGTVVDSNTTQKQSDGNRTCRVPSLKTEKIETTAMPRQCRACQQKVAYGGMKTAMVYR
PVMLSMVTIAVVCVCVALLFKSSPRVYYVFQPFNWESLEYGAM*
```

>Glyma.07G159300.1.p
```
MEKVAPPPILMHRKRDLSYDIVSAGPNESWRWEWDSVRFAGKPPPPPPLSPNNDVVFEAESVVPPLQLKLGGTTRVNNNNINI
NVSNKRVRSGSPGTASYPMCQVDNCREDLSKAKDYHRRHKVCEAHSKASKALLANQMQRFCQQCSRFHPLSEFDEGKRSCRRR
LAGHNRRRRKTQPEDVTSATPAPAAAANLEIFNLLTAIAGASQGKFEEKRSQVSDREQLVQILNKIPLPADLATKLLDAGSGN
VNGKKDHVQLQTPSSSYQCHESHDLLNHTPAAPLTMDLLAVLSTTLSGGSAPDSSASPSQNRSCSSDGGSADQTRQQQFFSVG
GERSSSSSQSPVEDSDCQEVRVNLPLQLFSSSPEDDSLPKLASSRKYFSSDSSNPAEERSPSSSPPIVEMQFDLQDGARGLKP
ESISSGRGVNANKEASQSHSSNISLDLFKGSNNWIQQPSSLQSVPFQAGYTSSGSDHSPPSLNSDAQDRTGRIMFKLFDKHPS
HFPGTLRAQIYNWLSNRPSDMESYIRPGCVVLSIYASMSSADWERLEENFLQHVHSLIQNSDSDFWRNGRFLVHSGSRLVSHK
DGKIRICKPWRTWKSPELISVSPLAIVSGQETSISLKGRNLSTLGTKIHCTGTGSYASAEVIGSAHSGVMYDKIKLSGFKVQD
VSPGVLGRCFIEVENGFKGNSFPVIIADETICKELRPLESEFDEEEKICDAISEEHEHHFGRPRSREEALHFLNELGWLFQRE
RFSYVHEVPCYSLDRFKFVLIFAVERNCCMLIKTLLDVLVGKHLQGEWLSTGSVEMLNAIQLLNRAVKGKYVGMVDLLIHYSI
PSKNGTSRKYVFPPNLEGPGGITPLHLAACTSGSESVVDSLTSDPQEIGLKCWESLVDANGQSPHAYAMMRNNDSYNALVARK
LADRQRGEISVTIANAIEQQSLRVELKQKQSYLVKRGQSSCAKCANAEIRYNRRVPGSHGLLHRPFIYSMLAVAAVCVCVCVF
FRGRPFVGSVAPFSWENLDYGTM*
```

>Glyma.15G099400.1.p
```
MEPQSLPPPPTMDDAQDPSSIWDFSYLLDFNLDDQDGIFTSSNSLPLPLNDDDDAAAPEISQNDRVRKRDPRMTCSNFLAGRV
PCACPELDAKLEDEGLPVKKRARTARASASARCQVPGCEADISELKGYHRRHRVCLRCANAATVMLNDEAKRYCQQCGKFHVL
SDFDEGKRSCRRKLERHNTRRRRKPPADSGGAAHSELLQPVAQNEEHNYHVEAGKDCSNLSTDVNDVGVSPDHEDEPVSIPSL
APDHEDEPVSIPSLAPEAQNINSDSVVSLAVSGEMRVNSGNTSNSPSYCNNKNAYSSVCQTGRISFKLYDWNPAEFPRRLRHQ
IFQWLANMPVELEGYIRPGCTILTVFIAMPNIMWITLLKDSLEYVHDFVAPGKMLSGRGTALVHLNDMIFRVMKDGTSVTKVE
VNMLAPRLHYVHPTYFEAGKPMEFVACGSNLLQPKFRLLVSFSGKYLKCEYCVPSPHSWTEDNISCAFDNQLYKIYVPHTEES
LFGPAFIEVENESGLSNFIPVLIGDKEICTEMKTLQQKLDVSLLSKQFQSASGGSICSSCETFALSHTSSSDFLVDIAWLLKD
TTSENFDRVMTASQIQRYCHLLDFLICNDSTIILGKILPNLIILTESVESNNVVINRTSDVDITQLLNHIHNASNVICQKHQQ
SRSIIVHSEMEGFKLAQGCSQDNMLSVAINSQGILSRTDEKWGVLKSPTSNDKTERIPLLKRDIIMSVEELPERYGRRCLGRG
FLTSRPAIFVIVSVAVCLGVCVAVLHPGRVSELAVSVRRCLFNY*
```

>Glyma.15G190200.1.p
```
MEARLEGKNQYLYGPVVPEVKSVGKRTLEWDLNDWKWDGDLFTARQLNSVPSDCRSRELFPADPEILVTGDASNNLSSAYDDV
NLGEGKRELEKRRRGVIDEGGVEMNDGAGSLNLNLGGQVYPIMEGEEKSGKKTKLTASTSSRAVCQVEDCRADLSNVKDYHRR
HKVCDMHSKATQALVGNVMQRFCQQCSRFHVLQEFDEGKRSCRRRLAGHNKRRRKTHPDATVVNGGSLNEEKGSSYLLMSLLR
ILSNMHSNGSDNMRNQDVLSHLLRNLASLTGTINGRNIVSLLEGSQDLVKAGTSGAAQNVPNTNSNGPEPSRPLYSSIKMDDG
LIHRDPPESLVQCETTPANDMAKECIASGNDEVGSLKSPSVPLSTNVLLSMDSLPPQSIAAQTTVGRIGLSNIDLNNVYDDVQ
DYVENTRNCRPPLPSGNGSLDHPLLVQCDSLKSSPPQTSRNSDSTSTQSPSSSSGEAQSRTDRIVFKLFGKAPNDFPHALRSQ
ILNWLSHSPTEIESYIRPGCIMLTIYLRLENSAWEELCYNLGPSLRKLAASNDCFWRTGWIYTRVQHSVAFLYNGQVVLDAPL
RLKSPQSCQILCVKPLAVSASSCAQFVLKGFNFLLSNSRLLCALEGKYLVQDNCYDLIDSVDAANGHHELQHLRFSCHVPNVT
GRGFIEVEDNGLSSCSFPFIVAEQEICSEICKLENVIEAAETADDIQIKTKLMEEKTQALYFIQEMGWLLHRSRVKVRLGPVA
PVQDNFHFNRFMWLVGFSMDHDWCAVMKKLLNIVFEGTVDTGDHASVELALLEMGLLHKAVKRNFRPMVELLLKFVPVKASDG
GDSNEKQINKSPDRFLFRPDTVGPARLTPLHVAASMHGSENVLDALTDDPGMVGSEAWKSAQDATGLTPYDYASLRGYYSYIQ
LVQRKTSNTCKNQQHVLDIPGNLVDSNTKQKQSDGHRSSKVLSLQTEKIETTAMRHCGLCQQKLVYGGMRRALVFRPAMLSMV
```

AIAAVCVCVALLFKSSPKVYYVFQPFSWESLEYGSI*

>Glyma.13G213400.1.p
MEPLSLPPLPVMDDAEDPSSIWDFSYLLDFDLDDQDIIINNNNSLPLPLNDHPAPSEIPQNDRVRKRDPRLTCSNFLAGRVPC
ACPELDAILEDEGLPGKKRARTARASASARCQVPACEVDISELKGYHRRHRVCLRCANAATVMLNGEAKRYCQQCGKFHVLSD
FDEGKRSCRRKLERHNTRRRRKPTADSGGATHSELLQPVAENEENNYDVEAGKDCSNLSTDINDVGVSLDLEDEPAPIPSSAP
EAQNINSDSVVSLAVSGETRVNSGNTSNSPSYCDNKSAYSSMCQTGRISFKLYDWNPAEFPRRLRHQIFQWLASMPVELEGYI
RPGCTILTIFIAMPNIMWINLLKDPLEYVHDIVAPGKMLSGRGTALVHLNDMIFRVMKDGTSVTNVKVNMHAPKLHYVHPTYF
EAGKPMEFVACGSNLLQPKFRLLVSFSGKYLKCEYCVPSPHSWTEDNISCAFDNQLYKIYVPHTEESLFGPAFIEVENESGLS
NFIPVLIGDKKICTEMKTLQQKLDVSLLSKQFRSASGGSICSSCETFALSHTSSSDLLVDIAWLLKDTTSENFDRVMTASQIQ
RYCHLLDFLICNDSTIILGKILPNLIILTESMKSNVVINRTSDVDIMQLLNHIHNARNAVYQKHPKSESIIVHSEMKGFILAQ
GCSQDNKLSVAINSQGIQSRADEKWGVLKVKSPTSNEKMERIPLLKRDIIMSMEELPERYGRRCLGRGFLTFRPAIFVLVSVA
VCLGVCVAVLHPGRVSELAVSVRRCLFNY*

>Glyma.13G274900.1.p
MEIKSGERRQVSIEKVRRDNTVEVEDEEEEDESVVGGEGGLAEDERKKIGGGSGRRGSSGVSPPSCQAEICGADLTVAKRYHR
RHKVCELHSKAPSVMVAGLRQRFCQQCSRFHELAEFDEAKRSCRRRLARHNERRRKSNAETCNEGCSGSKGQHPAETHCRNTD
DWGRIQMNIAKNSGYKFFHIR*

>Glyma.13G295400.1.p
MVVRSDRLKRDVADCGDNNQNGYVEQEKKIKGVHICGVVNADLHEAKQYHRRHRVCEYHVKAQVVLVDEVRQRFCQQCSRFHE
LAEFDDTKRSCRSSLAGHNERRRKNSDQSQAEGSSRNKGTGHPQLKDITCGQADERGRIQITIHENAAYKHFQIR*
>Glyma.02G008600.1.p
MEAKFGAEAYHFYGVGASSDLRGVGKRSSEWDLNDWRWDGDLFIASRLNPVPADGVGVGQQFFPIGSGIPVAGGPSNSSSTSE
EVDPRDPKANKEGDKKRRVIVLEDDGLNEEGGTLSLKLGGHASAVVDREVGSWDGTNGKKSRVSGSTSNRAVCQVEDCSADLS
KAKDYHRRHKVCEMHSKASRALVGNAMQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNKRRRKTNHEAVPNGSSLNDDQTSSY
LLISLLKILSNMHSDRSDQTTDQDLLTHILRSLASQNGEQGGKNIANLLREPENLLREDGSSRKSEMMSTLFSNGSQGSPSNI
RQHETVSMAKMQQQVMHAHDAGASDQQITSSIKPSMSNSPPAYSEARDSTAGQIKMNNFDLNDIYIDSDDGMEDLERLPVSTN
LVTSSLDYPWAQQDSHQSSPPQTSGNSDSASAQSPSSSSGEAQSRTDRIVFKLFGKEPNDFPLVLRAQILDWLSHSPTDMESY
IRPGCIVLTIYLRQAEALWEELCYDLTSSLNRLLDVSDDTFWRNGWVHIRVQHQMAFIFNGQVVIDTSLPFRSNNYSKILTVS
PIAVPASKRAQFSVKGVNLIRPATRLMCALEGKYLVCEDDHMSMDQCSKEPDELQCVQFSCSVPVMNGRGFIEIEDQGLSSSF
FPFIVVEEDVCSEICTLEPLLELSETDPDIEGTGKIKAKNQAMDFIHEMGWLLHRSQLKLRMVSSVDLFPLKRFKWLIEFSMD
HDWCAAVRKLLNLLFDGTVNTGDHPSLYLALSEMGLLHKAVRRNSKHLVELLLRYVPENISDKLGPEEKALVDGENQTFLFRP
DVDGTAGLTPLHIAAGKDGSEDVLDALTNDPCMVGIEAWKNARDSTGSTPEDYARLRGHYAYIHLVQKKINKKQGAAHVVVEI
PSNMTENNTNKKQNELSTIFEIGKPEVRRGQGHCKLCDNRISCRTAVGRSMVYRPAMLSMVAIAAVCVCVALLFKSSPEVICM
FRPFRWENLDFGTS*

>Glyma.02G121300.1.p
MESWSYVPEEKGYLFSDEMDFSLDAFMRSRKALVEWDNKSNFVEKDGFNSDREVVRSMEFVDLGFPDLLQKSFHGSQPLETSS
YELDSNNSSKRGNSSPHVIALDSSMGEEESDSKHLSSLVESKTHDSSLIDLKLGRLADCKGASSDKVAKEGFTLTSIHPTTLW
KRARTSSLPAQAPVCQVYGCNMDLSSSKDYHKRHKVCDAHSKTAKVIVNGIEQRFCQQCSRFHLLAEFDDGKRSCRRRLAGHN
ERRRKPQFDYVTGKQHKILQSYQGTKYLGSSLQNKPQFPFQDIFQSSILFPGKHAQFSQSGRVKLEEDSIYGSQLAAPVTLDQ
ESSSCALSLLSDQSQYPSCHAAGNPLASPQVFSGIRIPDRDSRVSKKPLRISPADKLAPNESFPCDTISKEVIKNRSAATFSS
VGHALEVHNGDDICQPSELFSIKHSLSSEHAATVDLFQLSSHLQRVEQQRNSVLVKWENEDYCFPTV*

>Glyma.02G177500.1.p
MGSASVTKAPPSSSSSPNSSSTESLDGLKFGQKIYFEEVSVAAAAATTTQGGNKTRGGVSASKKGRGGSVQAPRCQVAGCKVD
LSGAKAYYSRHKVCAMHSKSPTVTVAGLEQRFCQQCSRFHLLSEFDQGKRSCRKRLAGHNERRRKPPPSSVLTSRYGRLSLPI
FDNNGRDDGFLMEFASYPKLTLRNTLPTQRSTEPAPGNHAATLTWQGGSDTPSEFFLQGSGSSGGTSFLSSKHPLVESYTGVT
DSNCALSLLSSQTWGSRNTAPATSFELNNNLFNFNGTSMTQFAVASLQAAAIHQLPNATSWLYLKGVDSGNCSPERVPNHGLA
QISPPLHSHVHSELDVSQQGRRHYMI*

>Glyma.19G146000.1.p
MASDAKLSSSESLNGLKFGQKIYFEDVGLATPATSLTSSSSSSSAATVTSSSSSRKGRGGSVQPAQPPRCQVEGCKVDLSDAK
AYYSRHKVCGMHSKSPSVIVAGLQQRFCQQCSRFHQLPEFDQGKRSCRRRLAGHNERRRKPPTSSLLTSRYARLSSSAFDNSG
RGSNFLMELTSYPKLSLRNSLPTPRSSELAPGNQTSTLSWNGNSETSSDLFLQGSVGGTSFASPGHPPGESYIGVTDTSCALS
LLSNQTWGSRNTAPSLGLSNMINFNGTPLTQLAASSHGASIHQLPNTSWFFKGIDSGNCSPEVVPDLGLGQISQPLNSQLHGE
LDLSQQGRRHYMDLEQSRAYESAHWSL*

>Glyma.03G117500.1.p
MLDYEWGNPSTIMLSGNEDSSAAGSDQTHRQIFDHYASHTLIPDHFLHAPTHSTATTAAAATVDFTHHHHFTPQPHHHNQNHM
NHHHHHAFFDPRAFHGASNTSYPPPHSMLSLDPLPHVNAVGPGHGPGSGPGGLLLVPKSEDVGRPMDFVGSRIGLNLGGRTYF
SSSEDDFVSRLYRRSRPVEPGSTASSNSPRCQAEGCNADLSQAKHYHRRHKVCEFHSKAATVIAAGLTQRFCQQCSRFHLLSE
FDNGKRSCRKRLADHNRRRRKTQQPTQEIQKSQSSLENVARSPPDSGAQSSSSVTVAVSPPDYFRQRPYQSTNPSSTTSSSLF
FSSG*

>Glyma.03G143100.1.p
MASDAKLSSSESLNGLKFGQKIYFEDVGFATPATSLTSSATTATTVTSSSSSSSRKGRGGSVQPAQPPRCQVEGCEVDLSNAK
TYYSRHKVCGMHSKSPTVIVAGLQQRFCQQCSSELILAAPRFHQLPEFDQGKRSCRRRLAGHNERRRKPPPSSLLTSRYARLS
SSAFDHSGRGGNFLMELASYPKLSLRNSLPTPRSSELAPGNQTSTLSWHGNSETSSDLFLQGSVGGTSFAGPGHPPGESYTGV
TTSSCALSLLSNQTWGSRNPAPSLGLNNMVNFHGTPMTQPAASSHGASIHQLPNTPWCFKGIDSGNCSPEGVPDLGLGQISQP
LSSQLHGELDLSQQGRRHYMDLEQSRAYESAHWSL*

>Glyma.20G006400.1.p
MDQVAPPPILMHRKRDLSYAVVSPAPNPSWSWDWDSVRFAGKPPPPLSSPNDDVVFEESVAPPLQLNLGGRTNNSNSNKRVRS
GSPGTSSYPMCQVDNCREDLSKAKDYHRRHKVCEAHSKASKALLANQMQRFCQQCSRFHPLSEFDEGKRSCRRRLAGHNRRRR
KTQPEDVTSATPAPAAAANLEIFDLLTAIAGASQGKFEEKRSQVPVREQLVQILNRIPLPADLATKLLDAGSGNVNGKKDQVQ
LQTPSSYQRHESHDQLNHTPAAPLTMDLLAVLSTTLSGGSAPDASASPSQNHSCNSDGGSADQTRQQQFFSVGGERSSSSSRS
PVEDSDCQEDVRVNLPLQLFSSSPEDDSLPKLASSRKYFSSDSSNPAEERSPSSSPVVEMLFDLQGGARGLKPESISSGREVI
ANKEASQSHSSNISLDLFKGSNNRIQQPSSLQSVPFQAGYTSSGSDHSPPSLNSDAQDRTGRIMFKLFDKHPSHFPGTLRAQI
YNWLSNRPSDMESYIRPGCVVLSIYASMSSADWEKLEENFLQHVHSLIQNSDSDFWRNGRFLVHSGSQFVSHKDGKIRICKPW
RTWKSPELISVSPLAIVSGHETSISLKGRNLSTPGTKIHCTGTGSYASAEVIGSAYSGVMYDKIKLSAFKVQDVSHGVLGRCF
IEVENGFKGNSFPVIIADETICKELRPLESEFDEEEKICDAISEEHEHHFGRPRSREEALHFLNELGWLFQRERFSYVHEVPY
YSLDRFKFVLTFAVERNCCMLVKTLLDVLVGKHLQGEWLSTGSVEMLNAIQLLNRAVKGKYVGMVDLLIHYSIPSKNGTSRKY
VFPPNLEGPGGITPLHLAAGTSGSESVVDSLTSDPQEIGLKCWESLVDANGQTPHAYAMMRNNDSYNALVAHKLADRRRGEIS
VTIENAIEQQSLRVELKEKQSNLVKRGQSSCAKCANAEFRFNRRVPGSHGLLHRPFIYSMLAVAAVCVCVCVFFRGRPFVGSV
APFSWENLDYGTM*

>Glyma.05G204100.1.p
MDWNLKALSWDLSEVDQANLPNMETMEGSSRYGMYRTKGEFSVDLKLGQVGNSGTESVLTKSKDAVGVSKMTSSSTSGSSKRA
RALSNGTQTVSCLVDGCHADLSNCRDYHRRHKVCEVHSKTAQVSIGGQKQRFCQQCSRFHSLEEFDEGKRSCRKRLDGHNRRR
RKPQPESLTRSGSFLSNYQGTQLLPFSSSHEYPSTTVVNPTWGGVLTTSGDVRLHGHNQHHHQVHLADKQDLFLGSSPAGYKE
GKQLAFMQGDHTLNNRSPHLPGASVGQAQMFLRTSLYSESDGLRCKMFCDSLTSSIQDNNTSRALSLLSSPPQPLSPGNGLNL
MVNSHSPLMQPLGLSLHDNSLGSVDPVLGPNDSSSMYNIGSNGSQGNEAPPLFPFQWE*

>Glyma.05G019000.1.p
MESWSYVPEEKECVSNEALSPPNTLGRSKNSFLGWELKTPCGFSNDMLGLGHHSIENQGFEELGFPEMLGKHMSDDLIGSVPT
RKVNGNRQQQHSDRITGTVVDAPSGFSQRGDSNSKLSNSNSLIDLKLGRFADHGDFTDRAFSKVLSSSESSTPPKRVRASGLH
SQTAYCQVYGCNKDLSSCKDYHKRHKVCEVHSKTAIVIVNGIEQRFCQQCSRFHLLAEFDDGKRSCRKRLAGHNERRRKPQMG
IHPGKSGRLLQPCGDSRFQGSMLTSSFICPEMLPSGVTCSEKYGMSDFWRPIKAEHGTGFRNLSSMAATNGHPPSRSLFPSYN
GNQFPFLHENSATSTTGSIFCESNSPYPSALGAQNSGLRPLFQDTVGNENFNVFDTSSTVQGLSGISDSCCALSLLSSQSQNS
SSQSSGIPLAHSLVIPSSHSHNYNMSQVSEKIGISSQTSSSRVSDSFPSELNHADGSHLSPVLISDNNCIVNFDMADGIFQAS
DYLNVKDHLSGEDGATIDLLQLSSQLQRVEHERQSLQVKQENDSSCTLRIT*

>Glyma.11G251500.1.p
MECNAKSLGQWDWEHLFFFNTKAAENPRLQPTDCSSEADREINVGMLHPSGGSACSGSELIHASSSRSSKSASNNSSSNGDSK
ASILTLEGSHDDSTGKKELCNGEPIETSPAAVPSPVSGEPLLTLKLGKRLYFEDVCTGSDSKKPSSFDIPISSGKKHKTSSQT
LQHPSCQVEGCGLDLSSAKDYHRKHRVCEAHSKSPKVVITGLERRFCQQCSRFHALSEFDDKKRSCRRRLSDHNARRRKPQPD
SIQLNPSALSSSPYDGRQIVSPFAFSRTATNLAWQDMHPQTKEFLPKPAKAFNKIPSIVSMLSDDSSGLLTSRGIATKSTVPD
VAGMEDPIRLSDPNGTQDVNRALSLLSTNSSSSWGAYETKSLSLEHSTRATHSITQHHLPLPSSAEYWHTDHQQHSSMCISFT
DCDSSNRFQDFQLLSSPYESPFPCNQLD*

>Glyma.08G011500.1.p
MDWNLKAPSWDLSEVDQANLPNMETMEGSSRYGMYRTKGEFSVDLKLGQVGNSGRESSVLTKSKDGGAVGVSKMISSSSSGSA
KRARSLSNGTQTVSCLVDGCHSDLSNCRDYHRRHKVCEVHSKTAQVTIGGQKQRFCQQCSRFHSLEEFDEGKRSCRKRLDGHN
RRRRKPQPEVVLTRSASFLSNYHQGTQLLSFSSSHVYPSTTVVNPTLGGVVTSGDVRLHGQNHHHHQMHLVDKQDLFLGSSPT
GYKEGKHLAFMQGDHTLNNQSPHLPGASVGPAQMFLRTSPYSESGGLRCKMFCDSLTGSAHGNTSRALSLLSSPPQPHSPGNG
LNQMVNPHSPLMQPLGLSLHDNSLGSVDPVLGPNGSDHSSSMHHIGSNGSQVNEAPPLFPFQWE*

>Glyma.01G075900.1.p
MLDYEWGNPSTIILAGNEGGTTGKPDQTRRQIFDHNFTMLSNNNILQGVSHQYHHHFNPPQPHQHHAFFDPRAFQSAFTTSFS
PPPRSMLSLNSIPPLATPHCGLGFLLMHKPDDVMNFEGSAHRLGHGPGLGLNLGGRTYFDPFNDNTVNRLYRTKTVESVTTLS
SNLPRCQAEGCNADLSDAKHYHRRHKVCEYHSKAATVVAAGLTQRFCQQCNRFHLLSEFDNGKRSCRKRLADHNRRRKTQQI
AQEIHKSQHSLDNVARSLPESGAQSSSSSGKVAVRVTP*

>Glyma.01G063700.1.p
MESWSYVPEEKGYLFSDEMEFSLDAFMRSRKALVEWDNKSNFVERDGFNSDREVVKSMEFVDLGFPDLLQKSFHGSLPLETSS
YELDSNNSSKRGNSSTHVIALDSSMGEEESNSKHLSSLVESKTQDSSLIDLKLGRLADCKGASSDKVEKEGFTLTSIHPTTLS
KRARTSSLPAQAPVCQVYGCNMDLSSSKDYHKRHKVCDAHSKTAKVIVNGIEQRFCQQCSRFHLLVEFDDGKRSCRRRLAGHN
ERRRKPQFDYMTGKQHKILQSYQAGTKYLGSSLQNKPMFPFQDIFQSSILFPGKHAQISQSGRVKLEEDSIYSSQLTAPITLD
QESSSCALSLLSDQSQYLSHHAACNPLTSHRVFSGICIPNRDGQISETPLRISSADKHAPNESFPCNTTSKEVIKNRSAATFS
NTGHALRVHHGDDICQPSELLDVKHHFSSEHAATVDLFQLSSHLQRVEQQRNSVLVKWENEDYCFPTV*

>Glyma.12G226000.1.p
MEIKSRERIQVSIEKVRRDSTVEVEDEEEEDESVVGGEGGLAEDEKKKGGGGSGRRGSSGVSPPSCQAEMCGADLTVAKRYHR
RHKVCELHSKAPSVMVAGLRQRFCQQCSRFQELAEFDEAKRSCRRRLARHNERRRKSNAETCNEGCSGSKGQHPTESHCRNAD
VWGRIQMNIAKNSGYNSFHIR*

>Glyma.06G168600.1.p
MDWNLKAPSWDLVDHVDKATTLQQNIGTKEEQNRFGVYRMKGEFSVDLKLGHHHVGNSGTESALANKSKDAAAASGAAASAGV
SKMASSPSGSSKRARTISSTSLTVSCLVDGCNSDLSNCRDYHRRHKVCELHSKTPEVTIAGLKQRFCQQCSRFHSLEQFDERK
RSCRKRLDGHNRRRRKPQPEPLSRPGSFLSNYQGTQLLPFSGPHVYASTAMVSPDWSGALVTSSADARLHNHNHHHHQQDLFL
GSSHKEGKQVAFLHSNHHPSTLSNQNPPPFPSPSICHTLLRTRPLSETCGGGARSKMFCESLTTTTTASNNNNNSLHDTIPCA
LSLLSSSQTHTPGGNGLNQMVQQPPPPSMSLMMQPLGLSLQDNNNNSLESVDHCSSIYSLASDHHGSHCNEAAPQLFPFQWE*

>Glyma.06G205700.1.p
EDFNSKLSNSLGDSNGGNSLIDLKLGRFDDHGDAMVPTFSKGAPILSLLESSTPPKRVRALGAHSQTAYCQVYGCNKDLSGCK
DYHKGHKVCEVHSKTYKVIVNGIEQRFCQQCSRFHLLAEFDDGKRSCRKCLAGHNERRRKPQVGIHSGKAGRLLQLYPDLSKP
EDNTSFTFRPLSSIPITNGHPQSRSLFPSYSEKQVPLLHKNGATSATGNVFSENNNQYHPPSLGGQSSGSQSLFQDTSLGSED
FNVFNTESAVQGLSQSSGIPMARPMVIPSSHSHYSMNQVSEKMMGISSQASNKVFDRFPPEFNPVDGRHLRPILISDNNEIFN
FEMADGIFQGSDFVNVKDHLSCEDGLTFDSFQLSSPLQRVSNRGSSCR*

>Glyma.06G238100.1.p
MERERALFERVRKNSTVEEEVENEVEEEDESGVSITEEEKKKGVGSSGGRRGSSGGSGGVSPPSCQAERCGADLTDAKRYHRR
HKVCEFHSKAPVVVVAGLRQRFCQQCSRFHDLAEFDESKRSCRRRLAGHNERRRKSNPEASNEGSTKGHHPKETTHCRLANER
ERIQMNLPGSSGYKSFNIR*

>Glyma.18G005600.1.p
MEWNAKSLGQWDWEHLFFNTKAAENPRLQPTDCSSEADPEINVGLLYPSGGSACSGSELIHASSSRSSKSASNNSSSNGDSRE
SILTLEGSQDDSTGKKELCNGEPVETSPAAVPSPVSGEPLLTLKLGKRLYFEDVCTGSDSKKPSSFEVPILSGKKYKTSSQTV
QHPSCQVEGCGLDLSSAKDYHRKHRVCEAHSKSPKVVIAGLERRFCQQCSRFHALSEFDDKKRSCRQRLSDHNARRRKTQPDS
IQLNPSALSSSPYDGRQIMSPFAFSRTATNLAWQDMHSSKHPQTKEFLPKPAKAFNKIPSIVSMLSDDSSGLLTSRGIATKSI
VPGMEVPNGTQDVNRALSLLSTNSSWCPYETKSLSLEHSTRTTHSITQHRLPLPSSEYWHTDQQNSSMCISFSDCDSSNRFQD
FQLLSSPYESGFPCNQLD*

>LOC_Os08g41940.1
MEWDLKMPPAASWELADELENSGGGGVPAAVSSSSAAVGGGVNAGGGGRQECSVDLKLGGGLGEFGGGGAQPRVAVAGEPAKGK
GPAAAATGAAAAASSAPAKRPRGAAAAGQQQCPSCAVDGCKEDLSKCRDYHRRHKVCEAHSKTPLVVVSGREMRFCQQCSRFH
LLQEFDEAKRSCRKRLDGHNRRRRKPQPDPMNSASYLASQQGARFSPFATPRPEASWTGMIKTEESPYYTHHQIPLGISSRQQ
HFVGSTSDGGRRFPFLQEGEISFGTGAGAGGVPMDQAAAAAAASVCQPLLKTVAPPPPHGGGGSGGGKMFSDGGLTQVLDSD
CALSLLSAPANSTAIDVGGGRVVVQPTEHIPMAQPLISGLQFGGGGGGSSAWFAARPHHQAATGAAATAVVVSTAGFSCPVVES
EQLNTVLSSNDNEMNYNGMFHVGGEGSSDGTSSSLPFSWQ*

>LOC_Os08g40260.1
MQREVGPQVAPPMFLHQIQPLPPHATAAKKRGNPWPAAAVAAAEAKGGGNWNPRMWDWDSRALTAKPSSDALRVNAGLSHHQQ
QQQQSPPAAAKAAEALRQGGGGSGGLNLQLGLREDAATPMDVSPAATTVSSSPSPPASSAPAQEPVVRPSKRVRSGSPGSASG
GGGGGGGGGNSGGGGGSYPMCQVDDCRADLTNAKDYHRRHKVCEIHGKTTKALVGNQMQRFCQQCSRFHPLSEFDEGKRSCRR

RLAGHNRRRRKTQPTDVASQLLLPGNQENAANRTQDIVNLITVIARLQGSNVGKLPSIPPIPDKDNLVQIISKINSINNGNSA
SKSPPSEAVDLNASHSQQQDSVQRTTNGFEKQTNGLDKQTNGFDKQADGFDKQAVPSTMDLLAVLSTALATSNPDSNTSQSQG
SSDSSGNNKSKSQSTEPANVVNSHEKSIRVFSATRKNDALERSPEMYKQPDQETPPYLSLRLFGSTEEDVPCKMDTANKYLSS
ESSNPLDERSPSSSPPVTHKFFPIRSVDEDARIADYGEDIATVEVSTSRAWRAPPLELFKDSERPIENGSPPNPAYQSCYTST
SCSDHSPSTSNSDGQDRTGRIIFKLFGKEPSTIPGNLRGEIVNWLKHSPNEMEGYIRPGCLVLSMYLSMPAIAWDELEENLLQ
RVNTLVQGSDLDFWRKGRFLVRTDAQLVSYKDGATRLSKSWRTWNTPELTFVSPIAVVGGRKTSLILKGRNLTIPGTQIHCTS
TGKYISKEVLCSAYPGTIYDDSGVETFDLPGEPHLILGRYFIEVENRFGRNSFPVIIANSSVCQELRSLEAELEGSQFVDGSS
DDQAHDARRLKPKDEVLHFLNELGWLFQKAAASTSAEKSDSSGLDLMYFSTARFRYLLLFSSERDWCSLTKTLLEILAKRSLA
SDELSQETLEMLSEIHLLNRAVKRKSSHMARLLVQFVVVCPDDSKLYPFLPNVAGPGGLTPLHLAASIEDAVDIVDALTDDPQ
QIGLSCWHSALDDDGQSPETYAKLRNNNAYNELVAQKLVDRKNNQVTIMVGKEEIHMDQSGNVGEKNKSAIQALQIRSCNQCA
ILDAGLLRRPMHSRGLLARPYIHSMLAIAAVCVCVCVFMRALLRFNSGRSFKWERLDFGTI*

>LOC_Os01g18850.1
MSSGLKKKGLEWDLNDWRWDSNLFLATPSNASPSKCSRRELGRAEGEIDFGVVDKRRRVSPEDDDGEECINAATTNGDDGQIS
GQRGRSSEDEMPRQGTCSSSGPCCQVDGCTVNLSSARDYNKRHKVCEVHTKSGVVRIKNVEHRFCQQCSRFHFLQEFDEGKKS
CRSRLAQHNRRRRKVQVQAGVDVNSLHENHSLSNTLLLLLKQLSGLDSSGPSEQINGPNYLTNLVKNLAALAGTQRNQDMLKN
ANSAAIASHTGNYVAKGNSLHDSRPHIPVGTESTAEEPTVERRVQNFDLNDAYVEGDENRTDKIVFKLFGKEPNDFPSDLRAQ
ILSWLSNCPSDIESYIRPGCIILTIYMRLPNWMWDKLAADPAHWIQKLISLSTDTLWRTGWMYARVQDYLTLSCNGNLMLASP
WQPAIGNKHQILFITPIAVACSSTANFSVKGLNIAQPTTKLLCIFGGKYLIQEATEKLLDDTKMQRGPQCLTFSCSFPSTSGR
GFIEVEDLDQSSLSFPFVVAEEDVCSEIRTLEHLLNLVSFDDTLVEKNDLLASRDRALNFLHEFGWFLQRSHIRATSETPKDC
TEGFPAARFRWLLSFAVDREFCAVIKKLLDTLFQGGVDLDVQSTVEFVLKQDLVFVAVNKRSKPLIDFLLTYTTSSAPMDGTE
SAAPAQFLFTPDIAGPSDITPLHIAATYSDTAGVLDALTDDPQQLGIKAWKNARDATGLTPEDYARKRGHESYIEMVQNKIDS
RLPKAHVSVTISSTTSTTDFTEKHASQSKTTDQTAFDVEKGQQISTKPPLSCRQCLPELAYRHHLNRFLSTRPAVLSLVAIAA
VCVCVGLIMQGPPHIGGMRGPFRWNSLRSGPK*

>LOC_Os03g61760.1
MEAARVGAQSRHLYGGGLGEPDMDRRDKRLFGWDLNDWRWDSDRFVATPVPAAEASGLALNSSPSSSEEAGAASVRNVNARGD
SDKRKRVVVIDDDDVEDDELVENGGGSLSLRIGGDAVAHGAGVGGGADEEDRNGKKIRVQGGSPSGPACQVEGCTADLTGVRD
YHRRHKVCEMHAKATTAVVGNTVQRFCQQCSRFHPLQEFDEGKRSCRRRLAGHNRRRRKTRPEVAVGGSAFTEDKISSYLLLG
LLGVCANLNADNAEHLRGQELISGLLRNLGAVAKSLDPKELCKLLEACQSMQDGSNAGTSETANALVNTAVAEAAGPSNSKMP
FVNGDQCGLASSSVVPVQSKSPTVATPDPPACKFKDFDLNDTYGGMEGFEDGYEGSPTPAFKTTDSPNCPSWMHQDSTQSPPQ
TSGNSDSTSAQSLSSSNGDAQCRTDKIVFKLFEKVPSDLPPVLRSQILGWLSSSPTDIESYIRPGCIILTVYLRLVESAWKEL
SDNMSSYLDKLLNSSTGNFWASGLVFVMVRHQIAFMHNGQLMLDRPLANSAHHYCKILCVRPIAAPFSTKVNFRVEGLNLVSD
SSRLICSFEGSCIFQEDTDNIVDDVEHDDIEYLNFCCPLPSSRGRGFVEVEDGGFSNGFFPFIIAEQDICSEVCELESIFESS
SHEQADDDNARNQALEFLNELGWLLHRANIISKQDKVPLASFNIWRFRNLGIFAMEREWCAVTKLLLDFLFTGLVDIGSQSPE
EVVLSENLLHAAVRMKSAQMVRFLLGYKPNESLKRTAETFLFRPDAQGPSKFTPLHIAAATDDAEDVLDALTNDPGLVGINTW
RNARDGAGFTPEDYARQRGNDAYLNMVEKKINKHLGKGHVVLGVPSSIHPVITDGVKPGEVSLEIGMTVPPPAPSCNACSRQA
LMYPNSTARTFLYRPAMLTVMGIAVICVCVGLLLHTCPKVYAAPTFRWELLERGPM*

>LOC_Os02g07780.1
MDWMPPPKPTSPRSPPLLWDWADAAVPGSSSGEVSAAAAAAAAHPGRRRKEKRGRAEEGGGGGGGEVRCQVEGCGVELVGVKDY
HRKHRVCEAHSKFPRVVVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLYDHNARRRKPQTDVFSYASARPPSSLLFDDNRQI
SFVWNKAPLSHVRPFAISPWESSSEVGTTDGHIYLDKSHISKSLPAFNTDIDELLPMKDFSAATIWMFFGVLFIRTAQESCMS
FH*

>LOC_Os09g31438.1
MATGGSGGGGGGGGGDDVHGLKFGKKIYFEQDAAASASAAAVESSSTSSGGGGKKGKGVAAAAAPPPPLPPRCQVEGCGVDL
SGVKPYYCRHKVCYMHAKEPIVVVAGLEQRFCQQCSRCSVHMVRFHQLPEFDQEKKSCRRRLAGHNERRRKPTPGPLSSRYGR
LAASFHEEPGRSRSFVVDFSYPRVPSSVRDAWPAIQPSDRMSGSIQWQGGHELHPHRSAVAGYSDHHAFSSHGGSAAGAPMLH
HPAFELTSGGCLAGVATDSSCALSLLSTQPWDTTQSTSSHNRSPPMSSTASAFGGGNNPVSPSVMASNYMAASPGWNSSSRGH
DGARNVHLPPPHGVVLNEVPPGSVHHGHFSGELELALQGGAPSNRPEAEHGSGSGAFSHSTNAMNWSL*

>LOC_Os05g33810.1
MDAPGGGGGGGGGGGVDAGEPVWDWGNLLDFAVHDDDSLVLPWGDDSIGIEADPAEAALLPPAPSPQPAEAEAEAAGPASLP
SSMQAEGSKRRVRKRDPRLVCPNYLAGRVPCACPEIDEMAAALEVEDVATELLAGARKKPKGAGRGSGAAVGGSGGGASRGTP
AEMKCQVPGCEADIRELKGYHRRHRVCLRCAHAAAVMLDGVQKRYCQQCGKFHILLDFDEDKRSCRRKLERHNKRRRRKPDSK
GILEKDIDDQLDFSADGSGDGELREENIDVTTSETLETVLSNKVLDRETPVGSDDVLSSPTCAQPSLQIDQSKSLVTFAASVE
ACLGTKQENTKLTNSPVHDTKSTYSSSCPTGRVSFKLYDWNPAEFPRRLRHQIFEWLSSMPVELEGYIRPGCTILTVFVAMPQ
HMWDKLSEDTGNLVKSLVNAPNSLLLGKGAFFIHVNNMIFQVLKDGATLTSTRLEVQSPRIHYVHPSWFEAGKPIDLILCGSS

```
LDQPKFRSLVSFDGLYLKHDCRRILSHETFDCIGSGEHILDSQHEIFRINITTSKLDTHGPAFVEVENMFGLSNFVPILVGSK
HLCSELEQIHDALCGSSDISSDPCELRGLRQTAMLGFLIDIGWLIRKPSIDEFQNLLSLANIQRWICMMKFLIQNDFINVLEI
IVNSLDNIIGSELLSNLEKGRLENHVTEFLGYVSEARNIVDNRPKYDKQRQVDTRWAGDYAPNQPKLGISVPLAESTGTSGEH
DLHSTNAASGEEENMPLVTKALPHRQCCHPETSARWLNAASIGAFPGGAMRMRLATTVVIGAVVCFAACVVLFHPHRVGVLAA
PVKRYLSRNYSS*
```

>Sobic.010G254200.2.p
```
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTLTSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVQRHVK
NTGTNGRVDDSGNSPSSMIAFSQGEPLISLKLGKRAYFENVCGGQEAKVSAPSDVTSAVTVVKKTKVSQQNAKNWYCQVEGCK
VDLSSAKDYNRKHKVCVAHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLTHHNARRRKPQADAISFNSSRLSSM
FYDTSHRTNLFFSQPLYGQVRSNAVSPWDNSGGFKFMETKHLSVQPTKTVGLDELHFSTPQISTSVVAHTGHHHNFDRFMPFK
GTNTKVLNQVSGVEASATASNSNNGAPEPGRALSLLSDGSWGSNSAVIQQPPSHADAGALPPLATVAVSNTAARHPLDPSLGR
FWQDDHPPLDGPAPQIQELAHLRTW*
```

>Sobic.003G139900.2.p
```
MEASFLWKESGDQASAARMTGGNKNKSLNWDLNDWRWDANLFLATPAAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPEE
DGSAGCSNSAVADGDNNHVVSVQRARSSEEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTRTTVVCIKN
VEHRFCQQCSRFHLLQEFDEGKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSALFLLLKQLSGLESGSSSEQINHPN
SLVNILKNLTAIASTQSYQDILKNANSNSIPSNAANNAANGSIMHEQTIRSTPIGRESLAEEPPVKRRVQDFDLNDACIEEAE
SRTDKIVFKLFGKEPKDFPVDLREQIQNWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISLSNDGFWR
KGWLYARVQDCLTLSCNGSLMFASPWQPVIGDKHQRLCVTPIAVACSSSVKFSVQGFNIVQPTTKLLCVFGENYLIQEETQML
LEDSTMQQGPQCLTFSCSFPCTSGRGFIEVEDYDQSSLSVPFVVTDNDVCSEIRTLEHGLNLISFDETPERIDDLMAYRSRSL
HFLHEIGWLLQRSHVRATSEQPQYCPDRFPVARFKWLLSFAIDHEWCAVVRKLLNTMFQGDIDLDVPSPVEFALGENLLLTAV
KKCSKPLVECLLRYTTTNYAPVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATG
FTPEEYARQRGHVSYIQMVQDKIDRRMSRDHVLVTIPTTIDAVGKHASGLKSADQITFGVEKQLNINETSSCRQCVQQVQQFA
FQPRTNRFLSNRPAMLSLVAIAAVCVCVGLLMKSPPQVACMRPFLWDNVRWGPN*
```

>Sobic.009G135000.1.p
```
MDVPDSGGGAPDAGEPIWDWGNLLDFVVQDDDSLVLPWDDAAAGIAAAADPTEAATAPLLPASPPLPQPVEVEPEPEPEPEPE
PGPVLPPPPLRVQGIGRRVRKRDPRLVCPNYLAGRVPCACPEVDEMVAAAEVEDVATEFLAGARKKTKTAAAAARRSKAAAAG
AAPGAAGGGAARAAAAEMKCQVPGCEADIRELKGYHRRHRVCLRCAHAAAVMLDGVQKRYCQQCGKFHILLDFDEDKRSCRRK
LERHNKRRRRKPDSKGTLDKEIDEQLDLSADVSGDGELREENMEGTTSEMLETVLSNKVLDRETPVGSEDVLSSPTCTQPSLQ
NDQSKSVVTFAASVEACIGAKQESVKLVNSPMHDTKSAYSSSCPTGRISFKLYDWNPAEFPRRLRHQIFEWLASMPVELEGYI
RPGCTILTVFIAMPQHMWDKLSDDAADLLRNLVNSPNSLLLGKGAFFIHVNNMLFQVLKDGATLMSTRLDVQAPRIDYVHPTW
FEAGKPGDLILYGSSLDQPNFRSLLSFDGDYLKHDCYRLTSHDTFDRVENGDLIPDSQHEIFRINITQSRPDIHGPAFVEVEN
IFGLSNFVPILFGSKQLCSELERIQDALCGSYSKNNVLGELLSASSDPHERRKLHSSVMSGFLIDIGWLIRKTTPDEFKNVLS
STNIQRWIHILKFLIQNDFINVLEIIVKSMDTIIGSEILSNLERGRLEDHVTTFLGYVSHARNIVERRANHDEKTQIETGGII
VNSPNQPSLGASLPHASENTDIGGDNKLNSADEEETMPLVTRDVSHRHCCQPDMTARWLKPSLIVTYPGGATRMRLVTTVVVA
AVLCFTACLVLFHPHGVGVLAAPVKRYLSSDSAS*
```

>Sobic.005G120600.3.p
```
MDWAAASGAPSWGAAATDPGPTMLSFAGPSSSTTPAADAEARLQDFAAGLAQSARPAGAAGAGRRSRAAGGGAGVEACSVDGC
RSDLSRCREYHRRHKVCEAHSKTPVVVVAGQEQRFCQQCSRFHMLSEFDEGKRSCRKRLDGHNRRRRKPQHDLTNLGGFFPYH
QVNQFEVYPRTIPTVGQNSDAMHLVGRQQPFSISFSRTPNQFPFPQGGGSTLHAARPGHLLAEGSSHTGSSTCSNPLSGTLGP
ECALSLLSSSLHHPSAAGIPTAGQPQVASSLSRIAAVSQAATTAMATAFAAGVGHHVFVPDAMFEDPSQALPFSWQ*
```

>Sobic.004G036900.2.p
```
MGSFGMDWNQKASVLWDWENLPPIGANMSDNPRMTAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSASTDS
SPKAKGNSMELNFAPAKAPDKNTDSGKVDDARTSPSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPLASA
KKAKSAPNTQKSYCQVEGCKVDLSSAKEYHRKHKVCEPHSKASKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNA
RRRKPQPEAFSFGSSRLSTMFYDTRQQTNLPFGQGPYGQMRSCANSSWDNPVGGFKFAETKAPWLRAAGVDGLHLSSQQVWNN
FMPHGVHQDFDGFTAFKETSAKVLDQGAEASAVVSNPDGAPDLQRALSLLSNSSAGAGTNQQPTAQLHHHHAGLSTLAASSSS
VMQASSPGLWQDSSGTALGHHAHARFQALDAPGSGAAIATAQELLQLPRSPSLYDSSSSHYDLMR*
```

>Sobic.007G207000.1.p
```
MQREVGPQVAPPLFLHQIQPMPPHAAAAKKRGHPWPAAAGAAVAPAEAAAGNWNPRLWDWDSRALTARPSSDALRLAGGQQPQ
PSAKAAEAQRQGTAGGNGALNLQLGLREDAATPMDASPTAPAASSSPSPPPASASAGQEPVVRPSKRVRSGSPGSAGGGSGGG
GAANGGAGYPMCQVDDCRADLTSAKDYHRRHKVCETHSKTTKAVVANQAQRFCQQCSRFHPLAEFDEGKRSCRRRLAGHNRRR
RKTQPTDVASQLLLPGNQENAANRTQDIVNLITVIARLQGSNVGKVPSIPPIPDKQNLVEIISKINSLNNTTSAAKSPPSEVV
```

DLNASQEQQEQRQDSVEKTTNGIDKQTVPSTMDLLGVFSTGLATSTPETNTSQSQGSSDSSGNNKSKSHSTEPATVVNSHDKP
TRDFPAAGFMRSNSTHESRPHIYKQIEQETRPYLSLQLFGSTEADIPPKMDSVNKYLSSESSNPLDERSPSSSPPITRKFFPI
HSVDEEVRHPHITDYGEDATMGEVSTSQAWCAPPLDLFKDSERPVENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQDRTGRI
IFKLFGKEPSTIPGNLRDDIVNWLKHSPTEMEGYIRPGCIVLSMYLSMAAIAWDELEENLLQRVNSLVQSSDMDFWRKGRFLV
RTDSQLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVVGGRKTSLILKGRNLSIPGTQIHCTSIGKYISKEVLCSAYPGTIYDD
SGVETFDLPGQPDLILGRCFIEVENRFRGNSFPVIVASSSVCQELRNLEVELEDSQVLDVPSDGQIHDSRQSKTRVQVLHFLN
ELGWLFQRASACTLSTRSDMSDLDLIKFSTARFKYLLLFSSERDWCSLTKTLLDILAKRSLVSKELSKETMEMLAEIHLLNRA
VKRKSRSMVHLLVQFVVICPDNSKVYPFLPNLPGPGGLTPLHLAASIENAEDIVDALTDDPQQIGLTCWQSVLDEDGQSPETY
AKLRNHNSYNELVAQKLVDVKNNQVTITVNGDEICMDQLGNVGDHKQSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPY
IHSMLAIAAVCVCVCVFMRALLRINSGKSFKWERLDFGTI*

>Sobic.007G193500.1.p
MDWDLKAPGAWDLAELEHEQHAAATAAAAGPSGGHANAGVAGTGTGCRPPPGAGGAPAECSVDLKLGGLGEPGAAPPARRERE
AAAAAAAAAGKVASASSGAAAGAKRPRPGGQHQQGQGQHQQQQCPSCAVEGCKADLSKCRDYHRRHKVCEAHSKTPVVVVAGR
EMRFCQQCSRFHLLAEFDADKRSCRKRLDGHNRRRRKPQPDTMASASFITSQQGTRFSPFAPPRLEASWPGVIKTEESPYYTH
QIPLGSSSNRQQHFVGSTTPAYAKEGRRFPFLQEGEISFATGVVLEPPACQPLLKTVAVAPPESSSSSSSAGGGNKMFSDGLTR
VLDSECALSLLSAPANSSGIDVSRMVRPTEHVPMAQPVVSGLQFGSSSWFSPRPQASTGVTAGSGFPSCPAVVEGEQQQLNTV
LSPHDNEVNYGGMFHVGGGSGGGGEGSSDGGTSSSMPFSWQ*

>Sobic.001G026500.1.p
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAQAANGSGLNSSPSSSEEAEEDVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGSLSLRIGGTAVGVEAMENSDVNEDERNGKKIRVQGGGSSGPACVEGCGADLTAAKD
YHRRHKVCEMHAKASTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAVGGTASIEDKVSNYLLLS
LIGICANLNSDNVQHSNSQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANVAEEAAGPSNSKLPFVNGS
HCGQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNS
DSTSTQSLSSSNGDTQCRTDKIVFKLFDKVPGDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMS
LHLDKLLSSSTDDFWASGLVFVMVRRHLAFMLNGQIMLDRPLAPSSHDYCKILCVKPVAAPYSATINFRVEGFNLLSTSSRLI
CSFEGRCIFQEDTDFIAENAECEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKDVCSEVSELESIFESSSNE
HANVNDDARDQALEFINELGWLLHRANRMSKEDETDTSLAALNMWRFRNLGVFAMEREWCAVIKMLLDFLFIGLVDVGSRSPE
DVVLSENLLHAAVRRKSVNMVRFLLRYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGISAW
SNARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGHVVLGVPSSICSVITDGVKPGDVSLEICRPMSAAVPSCLLCSRQA
RVYPNSTTKTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWESLERGAM*

>Sobic.002G247800.1.p
MATGGSGSDDVHGLTFGKKIYFEQDAAVAAAGGSGSGSATAAAGGRKGKGVATGGGRPASASAASATQPPRCQVDGCGVDLSA
VKQYYCRHKVCYMHSKEPRVVVAGIEQRFCQQCSRFHQLPEFDQGKRSCRRRLIGHNERRRKPPPGPLTSRYGRLAASFQEPG
RFRSFLLDFSYPRVPSSVRDAWQGVQPGGDRMLGTIQWHGNQEPHPHRSAAAGYGNHAYNCHGGLVAGASMLPSAFELPPGGC
VAGVTTDSSCALSLLSTQPWDTTPQSGSHNRSPAMSAAASAFEGTPVAPSVMASSYTASSTWTASRDPADGARNAQQHHDDAL
HLVHPGPVHHGHFSGELELALQGSGPPPPPNMPHVDHGSSSGTFSHSTTNAMNWSL*

>Sobic.002G257900.1.p
MDWDLKMPVSWDLPDLEHDAAMPQPPPIAATAAAASPAAASGIAAAATAAPPHAAATRGAPSRAECSVDLKLGGLGEFGAVAD
GTKEPAAAATATAAPSAPSASPMKRPRSGPGGAAGAQCPSCAVDGCKADLSKCRDYHRRHKVCEAHSKTPVVVVAGREMRFCQ
QCSRFHLLAEFDEAKRSCRKRLDGHNRRRRKPQPDAMNSGSFMTSQQGLFSSAGTRFSSFPAPRPEPSWSGVIKSEDSSYYTP
HPHPVLSNRPHFAGGSTSSAYSKEGRRFPFLQDGDQVSFSAGAGTLEVSTVCQPLLKTAAAAAPPPESSSSNKMFSDGLTPVL
DSDCALSLLSSPANSSSVDVSRMVVQPAAEHIPMAQPLVPSSSLQHHQHHQFGSSPGGWFACSQAGSSAVSGAGTGTGGFAC
PASVESEQLNTVLVPSNDAGHEMNYHGIFHVGGEGSSDGTSPSLPFSWQ*

>Solyc01g090730.2.1
MLDYEWENQSSFMFTSDQNTQEENDQNNNQFLDPQAHFPHITSYQHPQQQQQQQLQNPHFPQFQATPNNTHLISSMYDPRAYG
VPYTQTHDTSMLSLQPTGGFMVVPKSEPQFGGGIHDFSSSSSRIGLNLGGRTYFASSSEDDFVNRLYRRTRAVDAGSVNSPKC
QAEGCNADLTHAKHYHRRHKVCEFHSKASTVIAAGLTQRFCQQCSRFHVLSEFDNGKRSCRKRLADHNRRRRKNIQQENNKKQ
PQPSSKSPAESGAQSSTVTVAISPPRIPVDGFRQRTYQQVTTSSSFPMGD*

>Solyc01g068100.2.1
MEASVGERFYHMGGTDLRGLGKRSLEWDLTDWKWDGDLFIATPLQQNPSNYQSRQFFPVETGNLASSNSSSSCSDEVNHGMEQ
QRRELEKRRRVIVVDEDDSGPLSLKLGGQGEPAADAGREMSNWDGAAGKRTKLAAPAAARAVCQVDDCGTDLSKAKDYHRRHK
VCEMHSKASRALVGNVMQRFCQQCSRFHALQEFDEGKRSCRRRLAGHNKRRRKTQSETVANNNSLNDGQTSGYSLMSLLKILS
NMHSNGANHTEDQDLLSHLLRSLASQGPTNGDKSLSGLLQESSNLLNNRSILRNPEIASLISNGSQAPPRPKERQFTNSAAEM

```
PQKRLEDARTASSQSPGILFPIQSNSQAYTPGRESTTGRSKLIDFDLNDAYVDSDDCGDDIDRSPVPECPSWLQQDSHQSSPP
QTSGNSDSASAQSPSSSSGDNQNRTDRIVFKLFGKGPSDFPFVVRAQILDWLSHSPTEIESYIRPGCVVLTIYLRLPESAWEE
LSYDLSSSLSRLLDVHGGDSFWTKGWIYIRVQNQIAFVCDGQVLLDMSLPCVSNDGSTLLSVRPIAVPVSDRVQFLVKGYNLT
KPSTRLLCSLEGNYLDPEADNEVEEQVAGGDKDDKLQSLNFTCSIPAVGGRGFIEVEDHGVSNSFFPFIIAEEDVCSEIRMLE
SDLELTSLDYVKGQTNNIEARNQAMDFIHELGWLLHRNNLRARLEHFGPNAVLHPLKRFKWLVEFSVDHEWCAVVKKLLNILL
DGTVGGDDSSLKYALTEMGLLHKAVRRNSRPLVELLLTYTPTNVADDLCSEYQSLVGVGGQFLFRPDCVGPGGLTPLHIAAGI
DGYEDVLDALTDDPGKVAIEAWKNTRDSTGFTPEDYARLRGHYSYIHLVQRKISKKANSGHIVVDIPRVPSVVENSNQKDEVC
ATTSLEISMTERKAIPRPCRLCDRKLAYGSRSRSLLYRPAMFSMVAMAAVCVCVALLFRGSPEVLYIFRPFRWEMVDFGTS*
```

>Solyc01g080670.2.1
```
MISSLLSGDDPAASSNFDWSDLLDFDLHEQLNISFDDPLHQEQQPETEFVAPVIPSSEDSPHSQDTDAGRIRKRDPRMACSNF
LAGRIPCACPELDEKMEEEEMAGIGPGKKRARTVRASAGAGARCQVPDCEADISELKGYHKRHRVCLRCANATSVVLDGHSKR
YCQQCGKFHILSDFDEGKRSCRRKLERHNNRRRRKATDTSKTSAEKESQQLTTADDVSGDDDIVKDNTCMGSQLGEKEILLES
EGHVPICSTQGIQNNHSDSFTASGETQVDAEKENYKNSHSPSYYDNKSALSSVCPTGRISFKLYDWNPAEFPRRLRHQIFQWL
ASMPVELEGYIRPGCTILTVFVAMPTFKWGKLLEDPAAHLYELIASPGNMLRGRGSFLIYLNNMVFRVTKGENSVVKVKLKGP
APKLMSIYPTCFEAGKPMEFFACGSNLMQPRFRFLVSFGGRYLGNDINVVPSDCKYEGDSSSTEHQLLKIHVPRTEADLFGPA
FVEVENESGLSNFIPILIAEKDICAEMKEIQRKFCSGGSECTAVCSPCEASTSRKSEFSEFMLDVAWLLREPSSENVQILASV
QMQRFNYLLNILMESQSTIILERVLSYFENIVKRNMLAGITDADMTLFQKNILEKNILLKERLHLKEYFAGDSGQIMQELPNL
QDTAVPHKHNIEFGPTYWELTSRVPLLDAELPLRVKEQQSGKSCGFLVRKTLLTSRTLVFVISGFALCLGLCATFLHPRKVGD
IAMTIRRCLFDKT*
```

>Solyc04g045560.2.1
```
MFGSGFDSPPVGVYEKNSNCTVEEGCTSSAAPMMYGLGKRNNSSVELPLPSLSKKTKSSTHNPLKPHCQVDSCNLDLSSAKQY
HRKHRVCHIHSKCPKVIILGRHRRFCQQCSRFHSLSDFDENKRSCRTRLSDHNARRRKPQFNSTSTRFSSLFYDNKQSMNLVF
NNAPLVHSKPTADSTWETSEVSKFTLTRGLTLKPYKADSFNGQSLVGGFKLSGTTEMQTNASNLYSVSNDTKAEVFSQGVKET
TFSLNMGVAPELPRALSLLSNENEFISIGHSKHVNQIIIPEQVTHAIPQSLPFESSERWQAEQHLFNPHIHTLVANGHSGGSF
QEIQLSKDPFDDSYMSMD*
```

>Solyc05g053240.2.1
```
MEARNFHGPVVSNMEVSGKKSREWDSNDWVWDGDRFTAEPLNSLPSDCRSKQLFPIGSEIPETATGIFNGFSSGAGELTLGND
KGRKELEKRRRTIVIDDDDEQNGEAGSLNLKLGEQLYPVMEEEVEKWEGKNGKKTKISGVSSNRAVCQVQDCRADLSSAKDYH
RRHKVCEVHSKAAKALVGNVMQRFCQQCSRFHVLEEFDEGKRSCRRRLAGHNKRRRKTHPENVANGASVNDEGGSNYLLISLL
RILANVQFNSSDQTKDQDLLSHLLRNLASLAGAANERNASGLLPAPSDLQNPGTSMEAPKEDSLRPNANCLTIPASEVKEKRM
DRGTSDAERGISQNLCALRPETLCCRKESLPINANASVTTSAPLKLNIDLNNIYDDSQGGIQKLQNSDVFVNPGAASSGCPLW
ISHDPHKSSSTRTSLNSGSTSSLSPSSSSGEAQSRTDRIVFKLFGKDPGEIPTGLRKQVLDWLSHSPTDIESYIRPGCIILTI
YLRMDKPIWEELYSDLNSSLRKLLNASAGSFWRTGWVYSRVKDRVAFLFNGQVVLDTPLPSHRSCGISIIKPIAVCASERVQF
LVKGFNLSRPTTRFLCAMEGKYLVQGNCTDVMVGADSCMDYNEIQSLSFPCTVPNATGRGFIEIEDHGLSSNFFPFIVAEKDV
CSEIRTLESIIEAAKMDDGFLRGTEEFQARDQALDFLHELGWLLHRCHLKFRVGSGASLNLFPFQRFHRLIDFSIDHDWCAVV
KKLLDVFFNGVVDVGQQSSLDIPLQEVGILHRAVRRKCRSMIDVLLKYRHHGAFDKSGLQTQQDDRGYLFRPDTVGPGGLTPL
HVVASLAGYENILDALIDDPGEVGIEAWKSARDSTGLTPNDYACLRGHYSYVHMVQKKINQKPGDGHVVLDIPGSLLDSNLKQ
KLSDGHRSVKVTSFQTEKSLGKPIHRQCKQCKQKLSYGNSGTSLVYKPAMLSMVAIAAICVCVALLFKSSPEVLYSFRPFRWE
LLNTTWIGLATGIVVVDASDLCPIMLCKISKLLCILTICTANCLHLWF*
```

>Solyc05g012040.2.1
```
MEPLSYALEGRGLIFPDNVELSVDTVSRNRSIVKEWNLNPFCDVDKSICGFSQEVTENTEFLGSGIADILKKSAASNPCPGVL
SGEMSDGCGKMLSSTSMFNSFEPIPGEVGLGAMFNNSATKSNNPMSSLIDLKLEELTDHGELRTNQSSKESSILSSPESSLAA
GKAQTKSSHSRFSISQVHDQKEEESTTVYREVKSNQSLKRNSALTSVNSSLQGKRLRTTNFHSEIPVCQVHGCNKDLSSSKDY
HKRHKVCNEHSKTAIVIVNGIEQRFCQQCSRFHLLAEFDEGKRSCRKRLAGHNERRRKPQFDTHWGSRFLDMTSQRRVPFLFP
EIFPGSFFYQENYEDNNNSKHPKLEHKPFGISQLAISVKNEQFPAKSIQHQYGMRKQDPSKVHTGGTTLSIQEFSKGQNSSCA
LSLLSAHSQNLLHNSTDVSPTLRWTVEPNHHVYVKGRDHNLEKSPRVSVVKSLTPTELYSSDVAEQDVVVQVPDCEAVSFGIQ
RDHGHDQRSNSINSKNCLSLEGGPTMDLVQLSSHLQRVEQQKNSVQVKQENDIFCSFTST*
```

>Solyc05g015510.2.1
```
MEWNVKWDWGNLVMFGSKASESPKELQLTDWGVEEEGELDGGSFNLSSGGSGTGGYGSDLKCGSSIKSSISASTDSSPKDGFK
VSNFAFETFNASPEDPSKKLESSKAEVSRNSPPMEASVGSVEPVIGLKIGKRTCETFGGGSSAKVSSFPPNPASSAAATKKAK
SSTQNAPIPHCQVEGCNLDLSSAKEYYRKHRVCDSHSKSPKVIVAGVARRFCQQCSRFHSVSEFDDKKRSCRRRLSDHNARRR
KPHQETIQFNSARLSSLFYDSRQPMNLVLNEAQLIHSRAAANATWESTQDSKFSITREFTPKPERTGSTNGKSLLERNQFSRA
VGAHSNASSLLLPSKGTTAEVFNRGAKESMFNMVTGPEFPRALSLLSTNSWGSSEPESVSLNHANQTSMPEQMMQAIPPHMSS
QYWQAGQHSSDPRYHTLAAANSNSGGSFQEMGVFKAPFDTDFYLNALN*
```

>Solyc05g015840.2.1
MESSSSSSKRAKAPGNIAHCLVDGCNADLSECREYHRRHKVCEVHSKTAKVTIAGRDQRFCQQCSRFHSLVEFDDGKRSCRKR
LDGHNRRRRKPQPDSMAKNSGILFGQQGTKLLSFSSQQIFPSAVVSSAWAGVVKTDSDMVLYNNQSHMNGMDSQNSFPDSSGH
SYKGGSQFQFMQGSDRSLTEAPLFEHPTSAAGISSSGQKIFSSGLNDIVDSDRALSLLSSAPAVTREIGLSHMVQQPASIPRS
QSHGLQYDGLSHFPFAQDFNSKPQDSHVSNSSSPLHFHDMLQNAQDESSTIPASQQTLAFMWD*

>Solyc10g078700.1.1
MELGSVSSSGNSSSSDSLNGLKFGKKIYFGNAGVGVQVKNGCGSSPVNGDGNLPPAPATTKRGRGGLVQGGHPPRCQVEGCQA
DLSDAKAYYSRHKVCGMHSKSPTVVVAGLEQRFCQQCSRFHQLTEFDQGKRSCRRRLACHNERRRKPPSGSLFSTHYGNLSSS
IFENNSSRSGSFLVDFSSHQNVNESSWPNTRASEQGWDHQSSGKFLQRPWLNNSENAASELVLQGSATRTSYHGVPSGDYFPG
VSDSSGALSLLSNRSWGSRNRPPSLGVNSQVHIDGVHTIQPSGSHGAPTNHFSSPSLSFKGNEASSSSHEMPPDLGLGQMLQA
SDNPYCGELGMAQHGDGRQYMELDQSKGYHPSVQNVHWTL*

>Solyc10g018780.1.1
MLDYNWGEQSSVMLSNDEQNQENEQNREFFDPYGPNTQTYTENPYPHQHFPNIQHQNHQFFHTQNNPTTLYDPRSYETCYTPH
PHPQSSMLQQNSGFMVVPKSEPNQFVGNGSGIDFGSSSSSRIGLNLGGRTYFSSSEDHDFVNRLYRRSRAVETAGSIMNTPRC
QAEGCNADLTHAKHYHRRHKVCEFHSKASTVIAAGITQRFCQQCSRFHLLSEFDNGKRSCRKRLADHNRRRRKNQQANLENSN
KDTTTNSPSDTLARSPADSGAHCSSVTVAISPPRISLDVFRQRP*

>Solyc02g077920.2.1
METNKWEGKRSITEAEKEEDEHGSVEEDSKRKRVLTLSGRKLVGEGSAHPSCQVDQCTADMADAKPYHRRHKVCEFHSKSPIV
LISGLQKRFCQQCSRFHLLAEFDDAKRSCRRRLAGHNERRRKITYDSHGENLG*

>Sspon.001A0040480
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGLLSLRVGGSAVGVGVMEKSDVNEDERNGKKIRVQGGNSSGPACQVEGCGADLTAAKD
YHRRHKVCEMHAKASTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAIGGTGSIEDKVSNYLLLS
LIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNGC
GQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSDS
TSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSLH
LDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATIDFRVKGFNLLGTSSRLICS
FEGRCVFQEDTGSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKDVCSEVSELESIFESSSSEHA
NVNDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFAMERDWCAVIKMLLDFLFIGLVDVGSRSAEEV
VLSENLLHAAVRRKSVNMVRFLLKYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGISAWSN
ARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICSVITDGVKPGDVSLEICRPMSASVPSCLLCSRQARV
YPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM

>Sspon.002A0013930
MGLNIPEITTIPLTASVSYAYTAPGALGKGLKALGKVFAESLPRGSRQRAIGKEALGKEFFAESHAAGSRQRLCRRPEAPSAK
PTVTSRNVLTTPLFSSWLAVPGILFLGGAGAPARPRARGAYLCTALADGKRWSLRLRGVSFMDWDLKMPVSWDLPDLEHDAAV
PQPPPIAATAAAAAASPMAASGIAAAAAAPSRAECSVDLKLGGLGEFGAADGTKEPAAAATATAAPSAPSASPMKRPRSGPGG
AAGAQCPSCAVDGCKADLSKCRDYHRRHKVCEAHSKTPVVVVAGREMRFCQQCSRFHLLAEFDEAKRSCRKRLDGHNRRRRKP
QPDTMNSGSFMTSQQGTRFSSFPAPRPEPSWSGVIKSEDSSYYTPHAVLSNRPHFAGGSTSSAYSKEGRRFPFLQDGDQISFS
AGAGTLEVSTVCQPLLKTTAAAALPPPESSSSNKMFSDGLTPVLDSDCALSLLSSPANSSVDVSRMVVQPAAEHIPMAQPLVP
SSSLQQHHQFGSSPGWFACSQAGSSGAVSAAGTGGFACPASVESEQLNT

>Sspon.004A0021810
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPARRGGNEREAKRARGEDGGGGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFGPDMSITASKFDGAP
DLQRALSLLSVGACGLPDPVHQTSCVIQFTGASENSSDLHLSHGGNSGPASCADGQHIATQPQSQLVRFTMDTSNNVYEPSFF
GVNQIN

>Sspon.004A0023470
MGSFGMDWNQKASVLWDWELPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSALTDSSP
KAKGNSMELNFAPAKAPDKNTDSGKADDARTSPLSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKK
AKAAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARR
RKPQPEALSFGSSRLSTMFYSSNTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQKVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQGIETSAAVSNPNGAPDLQHALSLLSNSSAGAGTNQQPIAQLHHHHAGLSTLVASSSS

AMQASSPGLWQDGGGAALGHHAHAHAHTRFQALDATGSGAAIATTAQELLQLPRPPSLYDSSSSHYDLMR

>Sspon.004A0023540
MGSFGMDWNQKASVLWDWELPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSALTDSSP
KAKGNSMELNFAPAKAPDKNTDSGKADDARTSPLSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKK
AKVAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARR
RKPQPEALSFGSSRLSTMFYGSSPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQKVWNNFMPHGVHQDFDG
FTALKGTSAKVLDQGLEASAAVSNPNGAPDLQHALSLLSNSSAGAATNQQPTAQLHHHHAGLSTLVASSSSAMQASSPGLWQD
GGGTALGHHAHAHAHTRFQALDATGSGAAIATTAQELLQLPRPPSSYDSSLSHYDLMR

>Sspon.006A0002630
MQREVGPQVAPPLFLHQIQPMPPHAAAAAKKRGHPWPAAGAAVAPTEAAAGNWNPRLWDWDSRALTARPSSDALRLAGGQQQL
PAEAQRQGAGGSGALNLQLGCDAATPMDASPTAPVASSSPSSPPASAAAGQEPVVRPTSACGPDRRGALEEDCGRWGCQRGAS
YPMCQVDDCRADLTSAKDYHRRHKVCETHSKTTKAVVANQAQRFCQQCSSNPLDERSPSSSPPITHKFFPIHSVDEEVRHPHI
TDYGEDATMGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQIVNWLKHSPTEMEGYIRP
GCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRTDSQLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVV
GGRKTSLVLKGRNLSIPGTQIHCTSIGKYISKEVLCSAYPGTIYDDSGVETFDLPGQPDLILGRCFIEVENRFRGNSFPVIVA
SSSVCLELRNLEVELEDSQVLDVSSDGQIHDSPQSKTRNQVLHFLNELGWLFQRASACTLSSRSDVSDLDLIQFSTARFKYLL
LFSNERDWCSLTKTLLEILAKRSLASEELSKETMEMLAEIHLLNRAVKRKSRSMVHLLVQFVVICPDNSKVYPFLPNLPGPGG
LTPLHLAASIENAEDIVDALTDDPQQIGLTCWQSVLDDDGQSPETYAKLRNHNSYNELVAQKLVDMKNNQVTITVNGDEIRMD
QLGNVGDHKKSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPYIHSMLAIAAVCVCVCVFMRALLRINSGKSFKWERLDF
AAIVM

>Sspon.006A0003761
MDWDLKAPGAWDLAELEHDQHAAATAGAAAGPSGGHANAGAAGTGTGCRPPPGAGAAPAECSVDLKLGGLGECGEPGAAPPAR
RERDAAAAAAAAGKAASASPGAAAGAKRPRPGGQHQGQGQGQGHQQQQCPSCAVEGCKADLSKCRDYHRRHKVCEAHSKTPV
VVVAGREMRFCQQCSSYLIKSVCLFRFHLLAEFDADKRSCRKRLDGHNRRRRKPQPDTMASASFITSQQGTRFSPFAPPRLEA
TSWPGVIKTEENPYYTHQIPLGSSSRQQHFVGSTTPAYAKEGRRFPFLQEGEISFATGVVLEPPACQPLLKTVAVAPPESSSS
SAGGGKMFSDGLTRVLDSECALSLLSAPANSSGIDVSRMVRPTEHVPMAQPVVSGLQFGSSSWFSPRPQASTGATTGGGFPSC
PAVVEGEHQQLNTVLSPHDNEVNYGGMFHVGGGSGGGEGSSDGGTSSSMPFSWQ

>Sspon.007A0010700
MDAPGSGGGGAVDAGEPIWDWGNLLDFVVQDDDSLAFPWDDAAGIAAANPTEAATAALLPASPPLPQPVEPEPEPEPGPVLPP
PPLRVQGIGRRVRKRDPRLVCPNYLAGRVPCACPEVDEMVAAAEVEDVATEFLAGARKKTKTAARRGKAAAAGAAAGAAGGGA
VRAAVAEMKCQVPGCEADIRELKGYHRRHRVCLRCAHASAVMLDGVQKRYCQQCGKFHVLLDFDEDKRSCRRKLERHNKRRRR
KPDSKGTLDKEIDDQLDLSADVSGDGELREENMEGTTSEMLETVLSNKVLDRETPVGSEDVLSSPTCTQPGLQNDQSKSLQND
QSKSVVTFAASVEACIGAKQESVKLANSPMHDTKSAYSSSNFLDAYDTRSLLSFDGDYLKHDCYRLTSHDTFDHVENGDLIPD
SQHEIFRIKITQSRPNIHGPAFVEMYPKNREGSMGSTSSSPLPSASCGGMVENIFGLSNFVPILFGSKQLCSELERIQDALCG
SYSKNNVLGELLGASSDPHERRKLHSSVMSGFLIDIGWLIRNPTPDEFKNVLSSTNIQRWICILKFLIQNDFINVLEIIVKSM
DSIMGSEVLSNLERGRLEDHVTAFLGYVSHARNIVDRRGNHDEKTQIETGGISVNSPNQPSLGASLPLASERNWGDELGNGDC
P
>GRMZM2G061734_P01
MDWDLKMPVSWDLPDLEHDAMPPPPPVSAASTAAASGIAAAAAAPSSATAPSRAECSVDLKLGGLGEFGAADGTATKEPAAAT
AAPSASPMKRPRLGPGGAGGAQCPSCAVDGCKADLSKCRDYHRRHKVCEAHSKTPVVVVAGREMRFCQQCSRFHLLLEFDEAK
RSCRKRLDGHNRRRRKPQPDTMNSGSFMTSQQGLFSSAGTRFSSFPAPRPEPSWSGVIKSEDSSYYTHHPVLSNRPHVAGTST
SPAYSKEGRRFPFLQDGDQVSFSASGAGTLEVSTVCQPLLKTTAVAPPPPESSKMLAPVLDSDCALSLLSSPANSSSVDVSR
MVQPAERIPMAQPLVPGLQHHQFGGSPVPDWFAGSGAVPAAGTGGFACPHSVESEQFNTVLVPSSDGGHEMNYHGIFHVGGEG
SSDGTSPSLPFSWQ*

>GRMZM2G126018_P01
MATGGGGGSDDFHGLTFGQKIYFEQDAGAAGGSATGGRKGKGAAASASAAQPPRCRVDGCGVDLSAVKQYYCRHKVCYMHSKE
PRVFVAGIEQRFCQQCSRFHQLPEFDQGKRSCRRRLIGHNERRRKPPPGPLTSRYGPLAASLVQEPGRFRSFLLDFSYPRVPS
SVRDDAWQWQGMQHGGDRMLLGTASAIQWHGNQEPHPPHRSAAAGYGNHAYNNCHGGGGLVAAAAGASMLSSAFEPPPGGGGC
VAGATAADSSCALSLLSTQPWDTPTQSGSSHHNRSPAMSAAAGAFEVAPSSVVASSYTASSSWSGSRDPADGARKNAQRHDDG
LHLVHPPGPVHHGHFSGELELALQGSGPPHVDHGGGTCFSHSDTNAMNWSL*

>GRMZM2G109354_P01
MDAPDSGGGGSADAGEPIWDWGNLLDFVVQDDDSLILPWDDAAGIAASDPTEAATALLPASPPLPQSVEEEPEPEPEPGPVLP
PPPLRVQGIGRRVRKRDPRLVCPNYLAGIVPCACPEVDEMVAAAEVEDVATEFLAGARKKTKTAGRRGKAEAAGTAGGTVRAA

VAEMKCQVPGCEADIRELKGYHRRHRVCLRCAHAATVMLDGVQKRYCQQCGKFHVLLDFDEDKRSCRRKLERHNKRRRRKPDS
KGTLDKEIDEPLDMPADVSGSDELREENMEGITSEILETVLSNKVMDRETPVGSEDVLSSPTCTQPCLQNDQSKSVVTFAASV
EACIGAKQESIKLANSPMHDTKSAYSSSCPTGRISFKLYDWNPAEFPRRLRHQIFEWLGSMPVELEGYIRPGCIILTVFIAMP
QHMWDKLSDDAADLLRNLVNSPNSLLLGKGAFFIHVNNMLFQVLKDETTLMSTRLDIQAPRIDYVHPTWFEAGKPVNLILYGS
SLDQPNFRSLLSFGGDYLKHDCYRLPSHDTFDGFESGDFIPDSQHEIFRIHITQSRPDIYGPAFVEVENMFGLSNFVPILFGS
KQLCSELERIHDALCGSYSENNVLGELLSASSDPHEHRKLCSSVMSGFLIDIGWLIRKPTPDEFKSVLSSTNIQRWICMLKFL
IQNDFINVLEIIVKSMDSIMGSEVLSNLERGRLEDHVTAFLGYVSHARNIADRRANHDEKTQIETRGIIVTSPNQPNLGASVP
FASENTDFGGDNKLNSTDEEEIMPLVTRDVSHRHCCQPDTTVRWLKPSLIVTYPGGATRMRLVTTVAVAAVLCFTACLVLFHP
HGVGVLAAPVKRYLSSDSAS*

>GRMZM2G081127_P01
MKKRRWWFECEYAGGEGSASPGYFAPRSTASPSGFRLCSAAPASRIRGEIDFGVVVDKRRRLSPEEDGSGGCSNSAVADGDNS
HVVCVQRGRNTEETRPRKGANSSTTPSCQVDGCQADLSGARDYHKRHKVCEAHTRTTVVCINNVEHRFCQQCSRFHLLHEFDD
GKKSCRSRLAQHNGRRRKVQPQPAVNGNSMNEDQSLSSTLFLLLKQLSGLESGSSSEQINHPNSLVNLLKNLAAIASTHAYQD
VLKNATSISSNDGNNAANGSIMHEQTIRSIPVRRESLAEEPAVKRRVQDFDLNDSCIEEAESRTDKIVFKLFGKEPKDFPVDL
REQILNWLSHYPTDMESYIRPGCVILTIYLHLPNWMWDEFNDDPASWIENLISLSNDGFWRTGWLYARVQDCLTLSCNGSLMF
ASPWQPVIGDKHQRLCVTPIAVDCSSSVKFSVKGFNIVQPTTKLLCVFDEKYLIQEETQMLLEDSTMQQGPQCLTFSCSFPCT
SGRGFIEIEDYDQSSLSVPFVVTDKDVCSEIRMLEHGLDLVSFDQTSKRIDDLMIYRSRALHFLHEIGWLLQRSHVRATSEQR
QYCPDRFPVARFRWLLSFAVDQEWCAVLRKLLNTMFQGDIDVLSPIEFALGENLLHTAVKKRSKPLVEFLLRYTTTNIAPVGG
GDGAPVQFLFTPAMTELSNITPLHIAATISDAIGVLDALTDDPQQLGIKAWKNARDATGFTPEEYAAKRGNISYIQMVQDKID
RRVTRAHVSVTIPSTIDTVGKHGSRMKPADQITFGVEKKQLSINQTLSCRQCVQQAQQLAFHPRTNRFLSNRTAMLSLVSIAA
VCVCVGLIMKSLPQVGCMKPFLWDNIRWGPN*

>GRMZM2G371033_P01
KRPRPGHAQQAAPACSVQGCDADLSQCRDYHRRHKVCVAHSKAPFVTVAGQQQRFCQQCSRFHLLEEFDEVKRSCRKRLDGHN
RRRRKPQPDPLSPAGLFGDHHHGGVTRFTSYPHRFTTSIAKPEWGAVVKTETDAFQDQYYPAVHLNGAGAGALFHDKDRKHLP
FLTNHNNGCQPFTVTTTASSESSSKHSNGNCALSLLSDNPTPAQAQTATIPTAQPLVATASQYGRLPGGDDDDVSLAGMSYMR
VGDSRQASNLTTSSPGPATAAATQQHLQYHHGYYYHVSGGEQGNNSPDGAAIQALPFSSW*

>GRMZM2G138421_P01
MEAARFGTQSSHLYGGGLDLNRREKRVFGWDLNDWSWDSQRFVATPVQQQAANGSGLNSSPSSSEEAVEEVARNGGFRGDTDK
RKRVIDDDETEDQCTVLHGGRSLSLKIGGAAGGVGAMENSDVNEDNERNGKKVRVQGGSSSGPACQVEDCGADLAAEKDYHRR
HKVCEKHAKASTAVVGNTVQRFCQQCSRFHLLHEFDEGRRSCRRRLAGHNRRRRKTRPDTAIGGTASIEDKVSNYLLLSLIGI
CANMNSDNAQNANGQELLSTLLKNLGSVAKSLEPKELCKFLEAYQSLQNGLNAGTSGIANGAEEAAGPSNSKLPSVNDSHYGQ
ASSSVVPVQSKATTVVTPEPASCKLKDFDLNDTCIDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSDSTS
TQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPSDIESHISPGCIILTVYVRLVESAWQELSENMSLLLG
KLLNSSTDDFWASGLVFVMVRHHLAFMLNGQIMLDRPLAPSSHHYCKILCVRPVAAPYSATINFRVEGFNLLSTSSRLICSFE
GHCIFHEDTDSGAENVEYKDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKDVCAEVSELESIFESSSNEHANA
RDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFAMEREWCAVIKMLLDFLFIGLVDVGSRSAEEMVLSENL
LHAAVRRKSVNMVRFLLRYKPNKNSKGTAQAYLFRPDALGPSKITPLHIAAATSDAEDVLDALTHDPGLIGISAWSNARDETG
FTPEDYARQRGNGAYLNLVQKKIDKHLGEGHVVLGVPSSICPVITDGVKPGDVSLEICRAMSASVPGCLLCSRQARMHPSSTA
RTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*

>GRMZM2G097275_P01
MGSFGMDWNQKASVLWDWENLPPVAVGASGSENPRMAAAPQALHSACGTISSSSEMGYGSSKSSVSASADDSSPKAKGKSMEL
NFAPDKVPDKNTDSGKVDDARTSPSSPVIAISSAEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPPASPKKTKVTPNK
KSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQKKKSCRRRLNDHNARRRKPQPEAL
SFGSSRLSAMFYDTRQHTSLLFGQGPYGQMRSCASSSWDNPGGGFKFAETKGPWLRAGGVDGLHQPSQQVWNNFMPHGVHQDF
DGFTAFKGTNAKVLHQGVEASAVASNQNGAPQPDLHRALSLLSNSWAGGGTGTGTDQHPTAQLRHAALSSSLAASSVSVAQAS
PAGLWQDGGAALGHHDYAHAFQAINTPGGGGAAAPIAVGQELLQLPRPPSLYDDGSSSRYDLMR*

>GRMZM2G414805_P05
MGSFGMNWNQKDPMVWDWEHLVPSVSNAVTRHGSANSSGGTLTSNSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVERHVK
NTGTNGRVDDSGNSPSSMIAFNQGEPLISLKLGKRAYFENACGGQDAKVSAASDVTSAASVVKKTKVSQQNAKNWYCQVEGCK
VDLSSAKDYNRKHKVCVVHSKATKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRLMHHNARRRKPQADTISFNSSTMFYD
TRQRTNLFFSQPLYGQVRSNAGSSWDNLGGLKFMETKHPPVHPTKTASPDELHFSALQITSAAAHTGHHHDLDGFMAFKGTST
KVLNQGTSPAWAAASSSNNGGPEGGRALSLLSDGSWGSSSAVIQQPTSHADAGALLPPLATVAVSNAAAAAGHPLDPSPGRFW
PQDDHPPLVDGPATQIPELAHLRIW*

>GRMZM2G307588_P01
MATGGGSSRSDDVRGLKFGKKIYFEQDGGSGSGAGAVGGRKGKGVATGGARPASAASAAQPPRCQVDGCGVDLSAVKQYYCRH
KVCNMHSKEPRVFVAGIEQRFCQQCSRFHQLHEFDQGKRSCRRRLIGHNERRRKPPPGPLTSRYGRLAASLQEPGRFRSFLLD
FSYPRVPSSVRDAWPGIQHGGDRMLGTVQWHGHQEPPHPHRSAAAGYGNHAAYNCHGGLVAGGAPMLSSAAFELPPGGCVAGV
AADSSCALSLLSTQPWDTTSHDHRSPAMPAAGAFDGTPVAPSVMASSYAASSAWTGSRDPAADGARNAQRLDDALHLVHPGSA
AVHFSGELELALQGSGGPPHLPRVDHGGSGGGTFNHSTTSAMNWSL*

>GRMZM2G169270_P01
MQREVGPQVAPPLFLHQIQPMPPHAAAAAKKRGHPWPAAGAAEAGAGNWNPRLWDWDSRALTARPSSDALRLAGGQPQPAAEA
QRQGAGALNLQLGLQEDSTTPMDASPTAPAAAAAASPPASAAAGQEPVVRPSKRVRSGSPGGAGGGPGGGGTANGGASYPMCQ
VDDCRADLTSAKDYHRRHKVCETHSKTTKAVVGNQAQRFCQQCSRFHPLAEFDEGKRSCRRRLAGHNRRRRKSQPTDVASQLL
LPVNQENAANRTQDIVNLITVIARLQGSNVGKVPSIPPIPDKQNLVEIISKINSLNNATSPAKSPSPEVVVLNTSQEQREQGH
DSVDKTTNGIDKQTVPSTMDLLGVFSTGFATSTPVTNTSQSQGSSDSSGNNKSKSHSTEPATVVNSHDISTQDFPAAGFMRSN
STQESRPHIYKQTEHETRPYLSLQLFGSSEEDIPPKMDSLNKYLSSESSNPLDERSPSSSPPITRKFFPIHSVDEEVRHPHIT
DFGEDATMGEVSTSQAWCAPPLDLFKDLERPLENGSPPNPGYQSCYVSTSCSDHSPSTSNSDGQDRTGRIIFKLFGKEPSTIP
GNLRDDIVNWLKHSPTEMEGYIRPGCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRTGSKLVSYKAGM
TRLSKSWRTWNTPELTFVSPIAVVGGQKISLILKGRNLSIPGTQIHCTSTGKYISKEVLCSAYPGTIYDDSGVETFDLPGQPD
FILGRCFIEVENRFRGNSFPVIVASSSVCQELRSLEVELEDSQVLDVSSDGQIHDCRQSKTRVQVLHFLNELGWLFQRASACT
LSTRPDVSDLDLTQFSITRFKYLLLFSSERDWCSLTKTLLDILAKRSLVSEELSKETMEMLAEIHLLNRAVKRKSRRMVHLLV
QFVVLCLDNSKVYPFLPNFPGPGGLTPLHLAASIENAEDIVDALTDDPQQVGLTCWQSALDEDGQSPETYAKLRNHNSYNELV
AQKLVDMKNSQVTITVNGDEIHMDQLGNVDDRKKSGVQALQIRSCSQCAILESGVLRQPMRSRGLLARPYIHSMLAIAAVCVC
VCVFMRALLRINSGKSFKWERLDYGTI*

>GRMZM2G156756_P01
MEAARFGARSSHLYGSGLGELDLNRRESRVFGWDLNEWSWDSQRFTATPVPVEAANGSGLNSSPSSSEEAGEEMAKNGGLGGD
TDKRKRAVVIDDDDDDETEDQDTVVNGGGGSLSLIIGGSAAGAGAMENGDVNEDERNGKKIRVQGGGSSGPACQVEGCGADLTA
AKDYHRRHKVCEMHAKATTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNKRRRKTRPDTASCGTASIEDKISNYL
LLSLIGICANLNSDNVQHSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANATEEAAGPSNSKLPFV
NGSHCGHASSSVVPVQSKATIVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTS
GNSDSTSTQSLSSSNGDAQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTIYLRLVESAWQELSE
NMSLHLDKLLSSSTDNFWASGLVFVMVRRRLAFMLNGQIMLDRPLAPSSHHYCKILCVKPVAAPYSATINFRVEGSNLLSTSS
RLICSFEGRCIFQEDTDSVAENDEYEDRAIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKDICFEVSELESIFESS
SIEHADANDNAREQALEFLNELGWLLHRANRMSKENVTDTSVATFSMWRFRNLGVFAMEREWCAVIKMLLDFLFIGLVDVGSR
SPEEVVLSENLLHAAVRRKSVNMVRFLLRYRPNKNSKGTAQTYLFRPDALGPSMITPLHIAAATSDAEDVLDVLTDDPGLIGI
SAWSNARDGTGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGRVVLGVPSSICSVITDGVKPGDVSLEICMPMSASVPGCLLCS
RKARVYPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*

>GRMZM2G160917_P02
MESGGGGDDQLHGLKFGKKIYFEDAAGSSSGSSSGGGSAPAPPATQQPSPPAASPRAPAGGGRRGRAAAGGAGPSTAPAPARC
QVDGCNVDLTDVKPAYYCRHKVCKMHSKEPRVLVNGLEQRFCQQCSRFHQLPEFDQLKKSCRKRLAGHNERRRRPPPGPLASR
YGRHAASLGEEPGRLRSFMLDFSYPRVSSAMRGGFPAVRPGGERVPGGIQWQAGLDPRHHQGAVAGYGAHYGSEGGSSSSARP
PVFPGPGPELPPGGCLAGVPADSSCALSLLSTQPWDAAHSHSHSHAAPTAGFDGGSPVAPSLMAASSYIAPSPWTETDSWGHEGG
RSVPQLLPPDDVPLGEVHSGSSSHHGQFSGELELALQGNRPAPGSAAPPAPRNNQGSAGTFDQAGNTMDWSL*

>GRMZM5G878561_P01
MDWDAKMLPAWDLGTVVGPSSSGGGALDLKLGAPTSSRAVLTAAAGAAPALPSANPPPPPPPPSSSAPAKRPRPGLARQAVPA
CSVQGCDADLSRCRDYHRRHKVCEAHSKTPVVTVAGQQQRFCQQCSSVKPSASSRCMRCAPVLSGACPYNATRQVQWNWTSVQ
FRFHLLGEFDEVKRSCRKRLDGHNRRRRKPQPDPLNPAGLFGNHHGVTRFTSYPQLFAASMEQPKWSVVKTEADVFQDQYYPA
VHLNGAGGGAIGSHFHGKEQKHFPFLTSHSNGGGGGGDVATAAALGCQPFTITTTASSESDSKQSNGNCALSLLSDNPTSAQTA
MIPTTAQSLAAAMQYGGRLPGSVSVSGGGGGDVSLTGMSYMRVGDSRQASILTASSPGHAAVASPVPATTAAATPPPPPQLLQ
YHHGYGYYHVSGGEQGNNPDGAAVHALPFSSW*
>GRMZM2G133646_P01
MEASFLRKDGGDQASAATMSGEHRNKGLDWDLNDWRWDANLFLATPAAIVPSECSSRKLSRAQGEIDFGVVVDKRRRLSPEED
GSAGCSNSGVADGDNNHVVPAQRGRSSEEETPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTRTTVVCIKNVEHR
FCQQCSRFQLLQEFDEGKKSCRSRLAQHNGRRRKVQPQTVVNGNSMSENQSLSNTSILLLKQLSRLESGSSSEPINHPNSLVN
LLKNLAAIASTQSYQDILKNANSNSILSNAGNNAANGSTVHEQTIRSIPIRGESLAEEPPVKRRVHDFDLNDACFEEVESRTD
KIVFKLFGKEPKDFPVDLREQIQHWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISISNDGLWRKGWL
YARVQDCLTLSYNGSLMFASPWQLVIGDKHQRLCVTPIAVACSSSVKFSVKGFNIVQPTTKLLCVFGDKYLIQEETQMLLEDS
TVQQGPQCLTFSCSFSCTSGRGFIEVEDYDQSSLCVPFVVADKDVCSEIRMLEHGLNLISFDETSERIDDLMAYRNRSLHFLH

EIGWLLQRSHVRATSEQPQYCPDRFPVERFRWLLSFAVDQEWCAVLRKLLNTMFQGDIDLDVPSPIEFALGENVLLTAVKKRS
KPLVEFLLRYTATDYAPVGSGAGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPRQLGIKAWKNARDATGFTPE
EYARKRGHISYIQLVQDKIDRRVSRAHVSVTIPSTIDTAGKHASRLKPADQITFGVEKKQRGMNQTLSCRQCVQQVQQLAFHS
RTNRFLSNRPAMLSLVAIAAVCVCVGLIMKSLPQVGGMRPFLWDKIRWGPN*


>GRMZM2G101511_P02
MKNLHLVRSFPPVATVPELLARLFPFSRPPPALAAAAAHCSQAGLAFVSFTFSDALMRINICCSPIVCSRFSPLPRTRELLRL
LFRAVSCFHALWRRRDGMPPPPPPPRYSAIGSSSSRRCLLPGDSSPPPQMKRRHGLACRVERDRGVLPALATINHLMFRCRVVS
CTATVSLQNVFCSSMKRFHLLAEFDADKRSCRKRLDGHNRRRKPQPDTMASASFIASQQGTRFSPFAHPRLEASWPPGVMKT
EESPYHITHQIPLGSSSSSRQQHFVALGAATPAYAKEGRRFPFLQEGEISFATGVVLEPPAAAPACQPLLRTGAPSESSGAGG
SKMFSDQGLARVLDSDCALSLLSAPANSSGIDVSRMVRPTEHVPMAQQPVVPGLQFGSASWFPRPQASTGGSFVPSCPAAVEG
EQQLNAVLGPNDSEVSMNYGGMFHVGGGSGGGEGSSDGGTSSSMPFSWQ*


>GRMZM2G098557_P01
MQREVGPQVAPPLFIHHQIQPMPPHAAAAVKKRGQPCPAAAAAPAEAAAGNWNPRLWDWDSRALTARPSADALRLAGGQPQPQ
PQQAAEVHRQGAGGSGALKLQLGPREGSTAPTDASPTAPAASLSPSPPASGQDLVVRPSKRVRSGSPGSAGGGGGGAANGGAG
YPMCQVDECRADLTGAKDYHRRHKVCETHSKTIRAVVANQAQRFCQQCSRFHPLAEFDEGKRSCRRRLAGHNRRRRKTQPADV
SSQLLLPGNQENAANRTQDIVNLITVIAHLHGSSVGKVPSIPPIPDKQNLVEIISKINSFNNMTSADKSPPSEVVDLDALQEQ
QVQRQDSVGKTTNGIDKQTVPSTMDLLGVFPTGLATSTPETNTSQSQGSSDSSGNNKSKSHSTEPVTVVNSHDKSTRDFSAAG
FMRSNSTHESQPHIYKQTEQETRPYLSLQLFGSTEEDFPPKMDSVNKYLSSESSNPLDERSPSSSPPITRKFFPIHSVDEEVR
HPHITDYGEDATMGEVSTNQAWLAPPLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQDRTGRIIFKLFGKE
PSTIPGNLRDDIVNWLKHSPTEMEGYIRPGCLVLSMYLLMPGIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRTNSQLVS
YKAGMTRLSKSWRTWNTPELTLVSPIAVVGGQKTSLILKGRNLSIPGTQIHCTSIGKYISKEVLCSAYPGTIYDDSGVETFDL
PGQPDLILGRCFVEVENRFRGNSFPVIVASSSVCQELRNLEVEFEDSQVLDVSSDGQIHDSRQPKTSVQVLHFLNELGWLFQR
ASACTSSTRSDVSDLDLIRFSTARFRYLLLFCSERDWCSLTKTLLDILAKRSLASEELSKETMEMLAEIHLLNRAVKRKSRNM
VHLLVKFVVICPDNSKVYPFLPNLPGPGGLTPLHLAASIENAEDIVDALTDDPQQTGVTCWQTVLDDDGQSPETYAKLRNHNS
YNELVAQKLVDMKNNQVTVRVNGDGIRADRLGNDVGDRKRSGVQALQIRSCSQCAILESGVLMQPVRSRGFLARPYIHSMLAI
AAVCVCVCVFMRALLRINSGKSFKWERLDYGTI*


>GRMZM2G065451_P02
MGSFGMDMDWNQKASVLLWDWENLPPAAANGSESHRTTAAAPQFAGLEATGHEPAPSSVSSSIHSLPSAKGNKNKNNMELISF
APAKAPDKDTGSVLSSGEPVVLGLKLGKRTYFEDGCGLGQSGKSSAALGTASAATPPGPAKKAKAAAAAPTAQQQKSYCQVEG
CRTDLSSAKDYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQKKKSCRRRLNDHNARRRKPQPEALPFGSSRLS
AMFYDTRRQASLLFGEGPYGQMRSCASSWWDSPGAGGGGFKFAETRGAPWLKPARAAADGLLPLSGQQQQQVWNDSIPHVAHQ
EDFDGFTVTAFKGISPKTLDHQGAGVEACAAVSSPDGAPDLQRALSLLSNSPAGGGGTNTNHPPPPPAAELHHHRAAPSSCLA
AGSVSVLQEASSPGLWQDGGGSAALGHHAHARLQALDAAMATAQELLQLPRPPPSYGGSSSHYDLMR*


>GRMZM2G460544_P01
MEAGGASGGDDGDGQLHGLKFGKKIYFEDAGGSSSSGGSGSASATAPPATQQPSPPAASPRAAPGGGRRGRAAAGGAGPSLPP
APPRCLVDGCNADLTDAKTYYCRHKVCEMHSKEPRVVVNGLELRFCQQCSRFHNLAEFDQQKKSCRKRLAGHNERRRRPPPGP
LASRYGRLAAPLGEEPGGRFRSFMLDFSYPRVPSAMRDGFPAVRPGGERVPGSNIQWQAASLDPPPPHHQSAAAGYGAHSYGS
PGSSSSSRPPVFPAGPELPPGGGCLAGVPTDSSCALSLLSTQPWDAAGHSAGHGHGHAASLPATAGFDGNPVAMASSYIAPPS
PWTDSRAHEGGRRNVPQLPPDVPLGDVHSGPGTHHGQFSGELELALQGNRPAPAPRNGQGSAGAFDQAGSTTEWSL*


>Sspon.01G0031720-1A
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGLLSLRVGGSAVGVGVMEKSDVNEDERNGKKIRVQGGNSSGPACQVEGCGADLTAAKD
YHRRHKVCEMHAKASTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAIGGTGSIEDKVSNYLLLS
LIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNGC
GQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSDS
TSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSLH
LDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATIDFRVKGFNLLGTSSRLICS
FEGRCVFQEDTGSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKDVCSEVSELESIFESSSSEHA
NVNDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFAMERDWCAVIKMLLDFLFIGLVDVGSRSAEEV
VLSENLLHAAVRRKSVNMVRFLLKYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGISAWSN
ARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICSVITDGVKPGDVSLEICRPMSASVPSCLLCSRQARV
YPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*


>Sspon.01G0031720-2B

MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGLLSLRVGGSAVGVGVMEKSDVNEDERTERRSGCKEATQDYHRRHKVCEMHAKASTAV
VGNTVRGSASNVVADNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPF
VNGCGQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQYSTQSPPQTSG
NSDSTSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSEN
MSLHLDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATINFRVKGFNLLSTSSR
LICSFEGRCIFQEDTDSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPYIIAEKDVCSEVSELESIFESSS
SEHANVNDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFAMERDWCAVIKMLLDFLFIGLVDVGSRS
AEEVVLSENLLHAAVRKKSVNMVRFLLRYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGIS
AWSNARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICSVITDGVKPGDVSLEICRPMSASVPSCLLCSR
QARVYPNSTSRTFLYRPAMLTVMGVAV

>Sspon.01G0031720-3D
MPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGDTDKRKRVVVIDDDETEDQDTVVNGGGLLSLRVGGSAVGVGAMEKSDV
NEDERNGKKIRVQGGNSSGPACQVEGCGADLTAAKDYHRRHKVCEMHAKASTAMVGNTVQRFCQQCSRFHLLQEFDEGKRSCR
RRLAGHNRRRRKTRPDTAIGGTASIEDKNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEA
AGPSNSKLPFVNGCGQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQD
STQSPPQTSGNSDSTSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLV
ESAWQELSENMSLHLDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRSLAPSSHHYCKILCVRPVAAPYSATINFRVK
GFNLLSTSSRLICSFEGRCIFQEDTDSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPYIIAEKDVCSEVS
ELESIFESSSSEHANVNDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFAMERDWCAVIKMLLDFLF
IGLVDVGSRSAEEVVLSENLLHAAVRKKSVNMVRFLLRYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVL
TDDPGLIGISAWSNARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICPVITDGVKPGDVSLEICRPMSA
SVPSCLLCSRQARVYPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM

>Sspon.02G0010770-1A
MGLNIPEITTIPLTASVSYAYTAPGALGKGLKALGKVFAESLPRGSRQRAIGKEALGKEFFAESHAAGSRQRLCRRPEAPSAK
PTVTSRNVLTTPLFSSWLAVPGILFLGGAGAPARPRARGAYLCTALADGKRWSLRLRGVSFMDWDLKMPVSWDLPDLEHDAAV
PQPPPIAATAAAAAASPMAASGIAAAAAAPSRAECSVDLKLGGLGEFGAADGTKEPAAAATATAAPSAPSASPMKRPRSGPGG
AAGAQCPSCAVDGCKADLSKCRDYHRRHKVCEAHSKTPVVVVAGREMRFCQQCSRFHLLAEFDEAKRSCRKRLDGHNRRRRKP
QPDTMNSGSFMTSQQGTRFSSFPAPRPEPSWSGVIKSEDSSYYTPHAVLSNRPHFAGGSTSSAYSKEGRRFPFLQDGDQISFS
AGAGTLEVSTVCQPLLKTTAAAALPPPESSSSNKMFSDGLTPVLDSDCALSLLSSPANSSVDVSRMVVQPAAEHIPMAQPLVP
SSSLQQHHQFGSSPGWFACSQAGSSGAVSAAGTGGFACPASVESEQLNT

>Sspon.02G0010770-2B
MDWDLKMPVSWDLPDLEHDAAVPQPPPIAATAAAAAASPMAASGIAAAAAAPSRAECSVDLKLGGLGEFGAADGTKEPAAAAT
ATPAPSAPSASPMKRPRSGPGGAAGAQCPSCAVDGCKADLSKCRDYHRRHKVCEAHSKTPVVVVAGREMRFCQQCSRFHLLAE
FDEAKRSCRKRLDGHNRRRRKPQPDTMNSGSFMTSQQGTRFSSFPAPRPEPSWSGVIKSEDSSYYTPHAVLSNRPHFAGGSTS
SAYSKEGRRFPFLQDGDQISFSAGAGTLEVSTVCQPLLKTTAAAALPPPESSSSNKMFSDGLTPVLDSDCALSLLSSPANSSV
DVSRMVVQPAAEHIPMAQPLVPSSSLQQHHQFGSSPGWFACSQAGSSGAVSAAGTGGFACPASVESEQLNTVLQAPSSDAGHE
MNYHGIFHVGGEGSTDGTSPSLPFSWQ*

>Sspon.02G0010770-4D
MDWDLKMPVSWDLPDLEHDAAVPQPPPIAATAAAAAASPMAASGIAAAAAVPSRAECSVDLKLGGLGEFGAADGTKEPAAAAT
ATPAPSAPSASPMKRPRSGPGGAAGAQCPSCAVDGCEADLSKCRDYHRRHKVCEAHSKTPVVVVAGREMRFCQQCSRFHLLAE
FDEAKRSCRKRLDGHNRRRRKPQPDTMNSGSFMTSQQGTRFSSFPEPSWSGVIKSEDSSYYTPHAVLSNRPHFAGGSTSSAYS
KEGRRFPFLQDGDQISFSAGAGTLEVSTVCQPLLKTTAAAALPPPESSSSNKIVDVSRMVVQPAAEHIPMAQPLVPSSSLQQH
HQFGSSPGWFACSQAGSSGAVSAAGTGGFACPASVESEQLNTVLQAPSSDAGHEMNYHGIFHVGGEGSTDGTSPSL

>Sspon.03G0015860-3C
MTVQSHSVFRLDESQGPPEPEEVEKKRPAAQPRSWGRPPSERVVTETRESPRALRQQVTLLFQVPTGNRTLGRGVAVRAPDKR
QGILRPPFHREESPSASVPPLQDPGDENHGIVLAAGAELRYFLVLVLFISFRICPNPRDLWILVIRANNRRINSPQKSSKKKN
KRGDRVEGKAAEMEAGFLWKESGDQASAARMTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVV
VDKRRRLSLVEDGSADCSNSAVADGDNNHVVSVQRGRSSEEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEA
HTRTTVVCIKNVEHRFCQQCSRFHLLQEFDEGKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSTLFLLLKQLSGLES
GSSSEQINHPNSLVNILKNLAAIASTQSYQDILKNVNSNSILSNAGNNAASGSIVHEQTIRSTPIGRESLAEEPPVKRRVQDF
DLNDPCIEEAEVEDYDQSSLSVPFVVTDNDVCSEIRTLEHGLNLISFDETSERIDDLMAYRSRSLHFLHEIGWLLQRSHMRAT
SEQPQYYPDRFPVARFRWLLSFAIDHEWCAVVTKLLNTMFQGDIDLDVPSPIEFALGENLLLTAVRKCSKPLVECLLRYTATN
YAPVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGFTPEEYARQRGHVSYIQM

VQDKIDRRVSRDHVSVTIPTMIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAFFHPRTNRFLSNRPAMLS
LVAIAAVCVCVGLLMKSPPQVGCMRPFLWDHIRWGPN


>Sspon.04G0017230-1A
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPARRGGNEREAKRARGEDGGGGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFGPDMSITASKFDGAP
DLQRALSLLSVGACGLPDPVHQTSCVIQFTGASENSSDLHLSHGGNSGPASCADGQHIATQPQSQLVRFTMDTSNNVYEPSFF
GVNQIN*


>Sspon.04G0018550-1A
MGSFGMDWNQKASVLWDWELPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSALTDSSP
KAKGNSMELNFAPAKAPDKNTDSGKADDARTSPLSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKK
AKAAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARR
RKPQPEALSFGSSRLSTMFYSSNTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQKVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQGIETSAAVSNPNGAPDLQHALSLLSNSSAGAGTNQQPIAQLHHHHAGLSTLVASSSS
AMQASSPGLWQDGGGAALGHHAHAHAHTRFQALDATGSGAAIATTAQELLQLPRPPSLYDSSSSHYDLMR


>Sspon.04G0018610-1A
MGSFGMDWNQKASVLWDWELPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSALTDSSP
KAKGNSMELNFAPAKAPDKNTDSGKADDARTSPLSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKK
AKVAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARR
RKPQPEALSFGSSRLSTMFYGSSPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQKVWNNFMPHGVHQDFDG
FTALKGTSAKVLDQGLEASAAVSNPNGAPDLQHALSLLSNSSAGAATNQQPTAQLHHHHAGLSTLVASSSSAMQASSPGLWQD
GGGTALGHHAHAHAHTRFQALDATGSGAAIATTAQELLQLPRPPSSYDSSLSHYDLMR


>Sspon.04G0017230-2B
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPARRGGNEREAKRVRGEDGGRGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFDDRRQISFVWNKAPV
SHVRPFTSPWDSSSDLKLPHAKETREVSTKVGTITGQVHLDKSHMSSAIPTLSHGKDEPWPMKGPDMSITASKFDGAPDLQRA
LSLLSVGACGLPDPVHQTSCVIQFTGASENSSDLHLSHGGNSGPASCADGQHIATQPQSQLVRFTMDTSNNVYEPSFFGNSSD
LHLSHGGNSGPASCADGQHIATQPQSQLVRFTMDTSNNVYEPSFFGVNQIN*


>Sspon.04G0018610-2D
MGSFGMDWNQKASVLWDWELPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSALTDSSP
KAKGNSMELNFAPAKAPDKNTDSGKADDAKTSPSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKK
AKAAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARR
RKPQPEALSFGSSRLSTMFYSSNTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQKVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQ


>Sspon.04G0018550-2D
MRRRRTARAPWLSSVLMRIERQVVPFSGEYFGLLLHLDWRTCGGSCFGRQVGAAAQILDPGMGSFGMDWNQKASVLWDWELPI
GANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSALTDSSPKAKGNSMELNFAPAKAPDKNTD
SGKADDAKTSPSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKKAKAAPNTQKSYCQVEGCKVDLS
SAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARRRKPQPESLSFGSSRLSTMFYSS
NTRQQKSLLFGQDPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQKVWNNFMPHGVHQDFDGFTALKGTSAK
VLDQ


>Sspon.06G0001900-1A
MQREVGPQVAPPLFLHQIQPMPPHAAAAAKKRGHPWPAAGAAVAPTEAAAGNWNPRLWDWDSRALTARPSSDALRLAGGQQQL
PAEAQRQGAGGSGALNLQLGCDAATPMDASPTAPVASSSPSSPPASAAAGQEPVVRPTSACGPDRRGALEEDCGRWGCQRGAS
YPMCQVDDCRADLTSAKDYHRRHKVCETHSKTTKAVVANQAQRFCQQCSSNPLDERSPSSSPPITHKFFPIHSVDEEVRHPHI
TDYGEDATMGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQIVNWLKHSPTEMEGYIRP
GCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRTDSQLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVV
GGRKTSLVLKGRNLSIPGTQIHCTSIGKYISKEVLCSAYPGTIYDDSGVETFDLPGQPDLILGRCFIEVENRFRGNSFPVIVA
SSSVCLELRNLEVELEDSQVLDVSSDGQIHDSPQSKTRNQVLHFLNELGWLFQRASACTLSSRSDVSDLDLIQFSTARFKYLL
LFSNERDWCSLTKTLLEILAKRSLASEELSKETMEMLAEIHLLNRAVKRKSRSMVHLLVQFVVICPDNSKVYPFLPNLPGPGG
LTPLHLAASIENAEDIVDALTDDPQQIGLTCWQSVLDDDGQSPETYAKLRNHNSYNELVAQKLVDMKNNQVTITVNGDEIRMD
QLGNVGDHKKSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPYIHSMLAIAAVCVCVCVFMRALLRINSGKSFKWERLDF
AAIVM*

>Sspon.06G0002840-1A
MDWDLKAPGAWDLAELEHDQHAAATAGAAAGPSGGHANAGAAGTGTGCRPPPGAGAAPAECSVDLKLGGLGECGEPGAAPPAR
RERDAAAAAAAAGKAASASPGAAAGAKRPRPGGQHQQGQGQGQGHQQQQCPSCAVEGCKADLSKCRDYHRRHKVCEAHSKTPV
VVVAGREMRFCQQCSSYLIKSVCLFRFHLLAEFDADKRSCRKRLDGHNRRRRKPQPDTMASASFITSQQGTRFSPFAPPRLEA
TSWPGVIKTEENPYYTHQIPLGSSSRQQHFVGSTTPAYAKEGRRFPFLQEGEISFATGVVLEPPACQPLLKTVAVAPPESSSS
SAGGGKMFSDGLTRVLDSECALSLLSAPANSSGIDVSRMVRPTEHVPMAQPVVSGLQFGSSSWFSPRPQASTGATTGGGFPSC
PAVVEGEHQQLNTVLSPHDNEVNYGGMFHVGGGSGGGEGSSDGGTSSSMPFSWQ

>Sspon.06G0001900-2B
MQREVGPQVAPPLFLHQIQPMPPHAAAAAAAKKRGHPWPAAGAAVAPTEAAAGNWNPRLWDWDSRALTARPSSDALRLAGGQQ
QLPASMAAEAQRQGAGGSGALNLQLGLREDAATPMDAGPTAPVASSSPSSPPTSAAAGQEPVVRPSKRVRSGSPGSAGGGSAG
GGGAANGGASYPMCQVDDCRADLTSAKDYHRRHKVCETHSKTTKAVVANQAQRFCQQCSSNPLDERSPSSSPPITHKFFPIHS
VDEEVRHPHITDYGEDATMGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQDRTGRIIF
KLFGKEPSTIPGNLRDDIVNWLKHSPTEMEGYIRPGCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRT
DSQLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVVGGRKTSLVLKGRNLSIPGTQIHCTSIGKYISKEVLCSAYPGTIYDDSG
VETFDLPGQPDLILGRCFIEVENRFRGNSFPVIVASSSVCLELRNLEVELEDSQVLDVSSDGQIHDSPQSKTRNQVLHFLNEL
GWLFQRASSCTLSSRSDVSDLDLIQFSIARFKYLLLFSNERDWCSLTKTLLEILAKRSLASEELSKETMEMLAEIHLLNRAVK
RKSRSMVQLLVQFVVICPDNSKVYPFLPNLPGPGGLTPLHLAASIENAEDIVDALTDDPQQIGLTCWQSVLDDDGQSPETYAK
LRNHNSYNELVAQKLVDMKNNQVTITVNGDEIHMDQLGNVGDHKKSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPYIH
SMLAIAAVCVCVCVFMRALLRINSGKSFKWERLDFVAIVM*

>Sspon.06G0001900-3C
MQREVGPQVAPPLFLHQIQPMPPHAAAAAAAKKRGHPWPAAGAAVAPTEAAAGNWNPRLWDWDSRALTARPSSDALRLAGGQQ
QQQPASMAAEAQRQGAGGSGALNLQLGLREDAATPMDASPTAPVASSSPSSPPTSAAAGQEPVVRPSKRVRSGSPGSAGGGSA
GGGAANGGASYPMCQVDDCRADLTSAKDYHRRHKVCETHSKTTKAVVANQAQRFCQQCSSNPLDERSPSSSPPITHKFFPIHS
VDEEVRHPHITDYGEDATMGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQDRTGRIIF
KLFGKEPSTIPGNLRDDIVNWLKHSPTEMEGYIRPGCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRT
DSQLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVVGGRKTSLVLKGRNLSIPGTQIHCTSIGKYISKEVLCSAYPGTIYDDSG
VETFDLPGQPDLILGRCFIEVENRFRGNSFPVIVASSSVCLELRNLEVELEDSQVLDVSSDGQIHDSPQSKTRNQVLHFLNEL
GWLFQRASACLPLPSKLAWPGGLTPLHLAASIENAEDIVDALTDDPQQIGLTCWQSVLDDDGQSPETYAKLRNHNSYNELVAQ
KLVDMKNNQVTITVNGDEIHMDQLGNVGDHKKSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPYIHSMLAIAAVCVCVC
VFMRALLRINSGKSFKWERLDFGTI*

>Sspon.06G0001900-4D
MQREVGPQVAPPLFLHQIQPMPPHAAAAAAAKKRGHPWPAAGAAVAPTEAAAGNWNPRLWDWDSRALTARPSSDALRLAGGQQ
QQPAFMAAEAQRQGAGGSGALNLQLGLREDAATPMDASPTAPVASSLPSSPPTSAAAGQEPVVRPSKRVRSGSPGSAGGGSAG
GGAANGGASYPMCQVDDCRADLTSAKDYHRRHKVCETHSKTTKAVVANQAQRFCQQCSRFHPLAEFDEGKRSCRRRLAGHNRR
RRKTQPTDVASQLLLPVNQENAANRTQDIVNLITVIARLQGSNVGKVPSIPPIPDKQNLVEIISKINSLNNTTPAAKSPPSES
SNPLDERSPSSSPPITHKFFPIHSVDEEVRHPHITDYGEDATMGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSCYAS
TSCSDHSPSTSNSDGQDRTGRIIFKLFGKEPSTIPGNLRDDIVNWLKHSPTEMEGYIRPGCLVLSMYLSMPAIAWDELEENLL
QRVNSLVQSSDLDFWRKGRFLVRTDSQLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVVGGRKTSLVLKGRNLSIPGTQIHCT
SIGKYISKEVLCSAYPGTIYDDSGVETFDLPGQPDLILGRCFIEVENRFRGNSFPVIVASSSVCLELRNLEVELEDSQVLDVS
SDGQIHDSPRSKTRNQVLHFLNELGWLFQRASACTLSSRSDVSDLDLIQFSTARFKYLLLFSNERDWCSLTKTLWKFLPREAW
PARNYQRKLWRCWLRFIF*

>Sspon.07G0008180-1A
MDAPGSGGGGAVDAGEPIWDWGNLLDFVVQDDDSLAFPWDDAAGIAAANPTEAATAALLPASPPLPQPVEPEPEPEPGPVLPP
PPLRVQGIGRRVRKRDPRLVCPNYLAGRVPCACPEVDEMVAAAEVEDVATEFLAGARKKTKTAARRGKAAAAGAAAGAAGGGA
VRAAVAEMKCQVPGCEADIRELKGYHRRHRVCLRCAHASAVMLDGVQKRYCQQCGKFHVLLDFDEDKRSCRRKLERHNKRRRR
KPDSKGTLDKEIDDQLDLSADVSGDGELREENMEGTTSEMLETVLSNKVLDRETPVGSEDVLSSPTCTQPGLQNDQSKSLQND
QSKSVVTFAASVEACIGAKQESVKLANSPMHDTKSAYSSSNFLDAYDTRSLLSFDGDYLKHDCYRLTSHDTFDHVENGDLIPD
SQHEIFRIKITQSRPNIHGPAFVEMYPKNREGSMGSTSSSPLPSASCGGMVENIFGLSNFVPILFGSKQLCSELERIQDALCG
SYSKNNVLGELLGASSDPHERRKLHSSVMSGFLIDIGWLIRNPTPDEFKNVLSSTNIQRWICILKFLIQNDFINVLEIIVKSM
DSIMGSEVLSNLERGRLEDHVTAFLGYVSHARNIVDRRGNHDEKTQIETGGISVNSPNQPSLGASLPLASERNWGDELGNGDC
P*

>Sspon.07G0008180-2B
MLDGVQKRYCQQCGKFHVLLDFDEDKRSCRRKLERHNKRRRRKPDSKVTLDKEIDDQLDLSADVSGDGELREENMEGTTSEML

ETVLSNKVLDRETPVGSEDVLSSPTCTQPGLQNDQSKSLQNDQSKSVVTFAASVEACIGAKQESVKLANSPMHDTKSAYSSSC
PTGRISFKLYDWNPAEFPRRLRHQIFEWLASMPVELEGYIRPGCTILTVFIAMPQHMWDKLSDDAADLLRNLVHSPNSLLLGK
GAFFIHVNNMLFQVLKDGATLMSTRLDVQAPRIDYVYPTWFEAGKPVDLILYGSSLDQPNFRSLLSFDGDYLKHDCYRLTSHD
TFDHVENGDLIPDSQHEIFRIKITQSRPDIHGPAFVEVENIFGLSNFVPILFGSKQLCSELERIQDALCGSYSKNNVLRELLS
ASSDPHERRKLHSSVMSGFLIDIGWLIRNPTPDEFKNVLSSTNIQRWICILKFLIQNDFINVLEIIVKSMDSIMGSEVLSNLE
RGRLEDHVTAFLGYVSHARNIVDRRANHDEKTQIDTGGISVNSPNQPSLGASLPLASENTDIGGDNKLNSADEEETMPLVTRD
VSRRQCCQPDMTAGWLKPSLIVTYPGGATRMRLVTTVVVAAVLCFTACLVLFHPHGVGVLAAPVKRYLSSDSAS*

>Sspon.07G0008180-3D
MVGAALQSTAVAIVSDRVGGGGRSGVAGSSGGFENMEGTTSEMLETVLSNKVLDRETPVGSEDVLSSPTCTQPGLQNDQSKSL
QNDQSKSVVTFAASVEACIGAKQESVKLANSPMHDTKSAYSSSCPTGRISFKLYDWNPAEFPRRLRHQIFEWLASMPVELEGY
IRPGCTILTVFIAMPQHMWDKLSDDAADLLRNLVNSPNSLLLGKGAFFIHVNNMLFQVLKDGATLMSTRLDVQAPRIDYVHPT
WFEAGKPVDLILYGSSLDQPNFRSLLSFDGDYLKHDCYRLTSHDTFDHVENGDLIPDSQHEIFRIKITQSRPNIHGPAFVEVE
NIFGLSNFVPILFGSKQLCSELERIQDALCGSYSKNNVLGELLSASSDPHERRKLHSSVMSGFLIDIGWLIRNPTPDEFKNVL
SSTNIQRWICILKFLIQNDFINVLEIIVKSMDSIMGSEVLSNLERGRLEDHVTAFLGYVSHARNIVDRRGNHDEKTQIETGGI
SVNSPNQPSLGALLPLASENTYIGDDNKLNSADEEETMPLVTRDVSHRQCCQPDMTARWLKPSLIVTYPGGATRMRLVMTVVV
AAVLCFTACLVLFHPHGVGVLAAPVKRYLSSDSAS*

>Sspon.08G0004510-2B
MMSGRLNVTTSSSPGDDFPFAAMQQHQPPPPPPPPYVGFEHGVAGGGGGQRGAGGVGIQQHHLYDGLDFAAALQFQQEAPHHH
HQLLTLPSSLGPMAPPPLPLQMPGMAGMPGDHVYPALGMVKREGGADGAAAGRIGLNLGRRTYFSPGDMLAVDRLLMRSRLGG
VFGLGFGGAHHQPPRCQAEGCKADLSGAKHYHRRHKVCEYHAKASVVAAGGKQQRFCQQCSRYVFHVLSEFDEVKRSCRKRLA
EHNRRRRKPATTASSKDAASPPANKKPNGGSITSSYSTDNKNLSTAKSTMSSNTSSVISCLDQAGNKQQLARPTLTLGASQDK
DHQQQQLSTMLQVQAAGGHHQEQHFITSLQVHNNGGGGGGNNILSCSSVCSSGALPSANGEVSDQNTTTTNTGNGGSSNNMHNL
FEVDF

>Sspon.08G0004510-3C
MMSGRLNVTTSSSPGDDFPFAPMQQHQPPPPPPPPYVGFEHGVAGGGGGQRGAGGVGIQQHHLYDGLDFAAALQFQQEAPHHH
HQLLTLPSSLGPMAPPPLPLPMPLQMPGMAGMPGDHVYPALGMVKREGGADGAAAGRIGLNLGRRTYFSPGDMLAVDRLLMRS
RLGGVFGLGFGGAHHQPPRCQAEGCKADLSGAKHYHRRHKVCEYHAKASVVAAGGKQQRFCQQCSRFHVLSEFDEVKRSCRKR
LAEHNRRRRKPATTASSKDAASPPANKKPNGGSITSSYSTDNKNLSTAKSTMSSNTSSVISCLDQAGNKQQLARPTLTLGASQ
DKDHQQQQLSTMLQVQAAGGHHQEQHFITSLQVHNNGGSGIGGGNNILSCSSVCSSGALPSANGEVSDQNTTTTNTGNGGSS
NNMHNLFEVDFM*

>evm.model.scga7_uti_cns_0181527.1
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDDETEDQDTVVNGGGLLSLRVGGSAVGVGAMEKSDVNEDERNGKKIRVQGGSSSGPACQVEGCGADLTAAK
DYHRRHKVCEMHAKASTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAVGGTASIEDKVSNYLLL
SLIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNG
CGQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSD
STSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSL
HLDKLLSSSTDDFWASGLEDTDSVAENAEYEDRDIECLSFCCSVPGRRGRGFIE

>evm.model.scga7_uti_cns_0110078.2.1
MGSFGMDWNQKASVLWDWENLLPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSASTDS
SPKAKGNSMELNFAPAKAPDKNTDSGKADDARTSLSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASA
KKAKAAPNTQKSYCQVEGCNVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNA
RRRRKPQPEALSFGSSRLSTMFYDTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQQVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQGVEASAAVSNPNGAPDLQHALSLLSNSSAGAGTNQQPTAQLHHHHAGLSTLVASSSS
VMQASSPGLWQDGGGAALGHHAHAHAHTRFQALDATGSGAAIATTAQELLQLPRPPSSYDSSSSHYDLMR*

>evm.model.scga7_uti_cns_0105459.1
MTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPVEDGSADCSNSAVADGDNNHV
VSVQRGRSSEEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTRTTVVCIKNVEHRFCQQCSRFHLLQEFD
EGKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSTLFLLLKQLSGLEPGSSSEQINHPNSLVNILKNLAAIASTQSYQ
DILKNANSNLILSNAGNNAANGSIVHEQTIRSTPIGRESLAEEPPVKRRVQDFDLNDACIEEAESRTDKIVFKLFGKEPKDFP
VDLREQIQDWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISLSNDGLWRKGWLYARVQDCLTLSCNGS
LMFASPWQPVEDYDQSSLSVPFVVTDNDVCSEIRMLEHGLNLISFDETSERIDDLMAYRSRSLHFLHEIGWLLQRSHMRATSE
QPQYCPDRFSVARFRWLLSFAIDHEWCAVVTKLLNTMFQGDIDLDVPSPIEFALGENLLLTAVRKCSKPLVECLLRYTATNYA

PVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGFTPEEYARQRGHVSYIQMVQ
DKIDRRVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAFFHPRTNRFLSNRPAMLSLV
AIAAVCVCVGLLMKSPPQVACMRPFLWDHIRWGPN*


>evm.model.scga7_uti_cns_0283225.2
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVVTPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGSLSLRVGGSAVGVGAMEKSDVNEDERNGKKIRVQGGSSSGPACQVEGCGADLTAAKD
YHRRHKVCEMHAKASTAMVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAIDGTASIEDKVSNYLLLS
LIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAHQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNGC
GQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSDST
STQSLSSSNGDTQCRTDKIVFKLFNKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSLHL
DKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATINFRVKGFNLLSTSSRLICSF
EGRCIFQEDTDSVAENAEFEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKDVCSEVSELESIFESSSSEHAN
VNDNARDQALEFLNELGWLLHRANRMSKHDETDTSLAPFNMWRFRNLGVFAMERDWCAVIKMLLDFLFIGLVDVGSRSAEEVV
LSENLLHAAVRKKSVNMVRFLLRYKPNKNSKGTAQTYLFRPDALGPSMIIPLHIAAATSDAEDVLDVLTDDPGLIGISAWSNA
RDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICSVITDGVKPGDVSLEICRPMSASVPSCRLCSRQARVY
PNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*


>evm.model.scga7_uti_cns_0058773.4
MEAGFLWKESGDQASAARMTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPVED
GSADCSNSAVADGDNNHVVSVQRGRSSEEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTKTTVVCIKNV
EHRFCQQCSRFHLLQEFDEGKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSTLFLLLKQLSGLEPGSSSEQINHPNS
LVNILKNLAAIASTQSYQDILKNANSNSILSNAGNNAANGSIVHEQTIRSTPIGRESLAEEPTVKRRVQDFDLNDACIEEAES
RTDKIVFKLFGKEPKDFPVDLREQIQNWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISLSNDGLWRK
GWLYARVQDCLTLSCNGSLMFASPWQPVIGDKLQRLCVTPIAVACSSSVKFSVQGFNIVQPTTKLLCVFGEKYLIQEETQMLL
EDSTMQQGPQCLTFSCSFPCTSGRGFIEVEDYDQSSLSVPFVVTDNDVCSEIRTLEHGLNLISFDETSERIDDLMAYRSRSLH
FLHEIGWLLQRSHMRATSEQPQYYPDRFPVARFRWLLSFAIDHEWCAVVTKLLNIMFQGDIDLDVPSPIEFALGENLLLTAVR
KCSKPLVECLLRYTATNYAAVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGF
TPEEYARQRGHVSYIQMVQDKIDRRVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAF
FHPRTNRFLSNRPAMLSLVAIAAVCVCVGLLMKSPPQVGCMRPFLWDNIRWGPN*


>evm.model.scga7_uti_cns_0229483.1
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGLLSLRVGGSAVGVGVMEKSDVNEDERNGKKIRVQGGNSSGPACQVEGCGADLTAAKD
YHRRHKVCEMHAKASTAMVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAIGGTASIEDKVSNYLLLS
LIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNGC
GQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSDS
TSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSLH
LDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATINFRVKGFNLLGTSSRLICS
FEGRCVFQEDTDSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKDVCSEVSELESIFESSSSEHA
NVNDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFSMERDWCAVIKMLLDFLFIGLVDVGSRSAEEV
VLSENLLHAAVRRKSVNMVRFLLKYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGISAWSN
ARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICSVITDGVKPGDVSLEICRPMSASVPSCLLCSRQARV
YPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*


>evm.model.scga7_uti_cns_0148300.1
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGLLSLTVGGSAVGVGVMEKSDVNEDERNGKKIRVQGGNSSGPACQVEGCGADLTAAKD
YHRRHKVCEMHAKASTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAIGGTASIEDKVSNYLLLS
LIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNGC
GQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSDS
TSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSLH
LDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATINFRVKGFNLLGTSSRLICS
FEGRCVFQEDTDSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKAVCSEVSELESIFESSSSEHA
NVNDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFAMERDWCAVIKMLLDFLFIGLVDVGSRSAEEV
VLSENLLHAAVRRKSVNMVRFLLKYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGISAWSN
ARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICSVITDGVKPGDVSLEICRPMSASVPSCLLCSRQARV
YPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*

>evm.model.scga7_uti_cns_0215805.3
MGSFGMDWNQKASVLWDWENLLPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSASTDS
SPKAKGNSMELNFAPAKAPDKNTDSGKADDARTSLSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASA
KKAKAAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNA
RRRKPQPEALSFGSSRQSTMFYGNTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQQVWN
NFMPNGVHQDFDGFTALKGTSAKVLDQGVEASAAVSNPNGAPDLQHALSLLSNSSAGAGTNQQPTAQLHHHHAGLSTLVASSS
SVMQASSPRLWQDGSGAALGHHAHAHAHAHARFQALDATGSGAAIATTAQELLQLPRPPSSYDSSSSHYDLMR*

>evm.model.scga7_uti_cns_0085485.1
MEAGFLWKESGDQASAARMTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPVED
GSADCSNSAVADGDNNHVVSVQRGRSSEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTRTTVVCIKNVE
HRFCQQCSRFHLLQEFDEGKKSCRSRLAKHNGRRRKVQPQAAVGNGNSMNENQSLSSTLFLLLKQLSGLESGSSSEQINHPNSL
VNILKNLAAIASTQSYQDILKNVNSNSILSNAGNNAASGSIVHEQTIRSTPIGRESLAEEPPVKRRVQDFDLNDPCIEEAESR
TDKIVFKLFGKEPKDFPVDLREQIQDWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDEVLDDPASWIEKLISLSNDGLWRKG
WLYARVQDCLTLSCNGSLMFASPWQPVEDYDQSSLSVPFVVTDNDVCSEIRTLEHGLNLISFDETSERIDDLMAYRSRSLHFL
HEIGWLLQRSHMRATSEQPQYYPDRFPVARFRWLLSFAIDHEWCAVVTKLLNIMFQGDIDLDVPSPIEFALGENLLLTAVRKC
SKPLVECLLRYTATNYAPVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGFTP
EEYARQRGHVSYIQMVQDKIDRRVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAFFH
PRTNRFLSNRPAMLSLVAIAAVCVCVGLLMKSPPQVGCMRPFLWDHIRWGPN*

>evm.model.scga7_uti_cns_0267720.2
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGLLSLRVGGSAVGVGVMEKSDVNEDERNGKKIRVQGGNSSGPACQVEGCGADLTAAKD
YHRRHKVCEMHAKASTAMVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAIGGTASIEDKVSNYLLLS
LIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNGC
GQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSDS
TSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSLH
LDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATINFRVKGFNLLGTSSRLICS
FEGRCVFQEDTDSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKDVCSEVSELESIFESSSSEHA
NVNDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFSMERDWCAVIKMLLDFLFIGLVDVGSRSAEEV
VLSENLLHAAVRRKSVNMVRFLLKYKPNKNSNGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGISAWSN
ARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICSVIITDGVKPGDVSLEICRPMSASVPSCLLCSRQAR
VYPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*

>evm.model.scga7_uti_cns_0006207.3
MEAGFLWKESGDQASAARMTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPVEN
GSADCSNSAVADGDNNHVVSVQRGRSSEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTRTTVVCIKNV
EHRFCQQCSRFHLLQEFDEGKKSCRSRLAKHNGRRRKVQPQAAVGNGNSMNENQSLSSTLFLLLKQLSGLEPGSSSEQINHPNS
LVNILKNLAAIASTQSYQDILKNANSNLILSNAGNNAANGSIVHEQTIRSTPIGRESLAEEPPVKRRVQDFDLNDACIEEAES
RTDKIVFKLFGKEPKDFPVDLREQIQDWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISLSNDGLWRK
GWLYARVQDCLTLSCNGSLMFASPWQPVIGDKLQRLCVTPIAVACSSSVKFSVQGFNIVQPTTKLLCVFGEKYLIQEETQLLL
EDSTMQQGPQCLTFSCSFPCTSGRGFIEVEDYDQSSLSVPFVVTDNDVCSEIRTLEHGLNLISFDETSERIDDLMAYRSRSLH
FLHEIGWLLQRSHMRATSEQPQYCPDRFLVTRFRWLLSFAIDHEWCAVVTKLLNTMFQGDIDLDVPSPIEFALGENLLLTAVR
KCSKPLVECLLRYTATNYAPVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGF
TPEEYARQRGHVSYIQMVQDKIDRRVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAF
FHPRTNRFLSNRPAMLSLVAIAAVCVCVGLLMKSPPQVACMRPFLWDHIRWGPN*

>evm.model.scga7_uti_cns_0071177.1
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTLTSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVERHVK
NTGTNGRVDGSGNSPSSMIAFSPGEPLISLKLGKRAYFENVCGGQDAKVSAPSDVTSAATVVKKTKVSQQNAKNWYCQVEGCK
VDLSSAKDYNRKHKVCVVHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLTHHNARRRKPQADTISFNSSRLSTM
FYDTSQRTNLFFSQPLYGQVRSTAVPSWDNLGGFKFMETKHLSVQPTKIVGLDELQFSTPQISTSVVAHTGHHNNFDGFMPFK
GTNTKVLNQGVEASAATSNSNNGAPEPGRVLSLLSDGSWGSSSAVIQQTTSHADAGALPPLATVAVSNAAAGHPLDPSPGRLW
QCDPPPLDGTAQIQELAHLRTW*

>evm.model.scga7_uti_cns_0345865.1
MGSFGMDWNQKASVLWDWENLLPIGTNGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSASTDS
SPKAKGNSMELNFAPAKAPDKNTDSGKADDARTSLSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASA
KKAKVAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNA

```
RRRKPQPEALSFGSSRLSTMFYDTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQQVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQ
```

>evm.model.scga7_uti_cns_0026104.2
```
MLFSWLVVCSCSCSSRGIFPGIVFAVLTISWLGAGRRRFLSFLAGGMDWAAASGAPSWGAAATDPGPTMLSFAGPSSSTSPAA
EARLQNFAAGLVHSARPAGAAGAGRRSRAAGGGAGAEACSVDGCRSDLSRCREYHRRHKVCEAHAKTPVVVVGGQEQRFCQQC
SRFHMLSEFDEGKRSCRKRLDGHNRRRRKPQHDLTNLGGFFPYHQVNQFEVYPRTIPTAGQNSDAMQLVGRQQPFSISFSRTP
NQFPFPQGGGSALHAARPGLLAEGSSHTGSSTCNNALSGTLGPECALSLLSPLLHRPSAAGIPTNGQPEVASSLSRIAAVSQA
ATTAMTIAFAAGVGHHVFVPDAMFEDPSQALPLSWQ*
```

>evm.model.scga7_uti_cns_0184661.1
```
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDADETEDQDTVVNGGGLLSLRVGGSAVGVGAMEKSDVNEDERNGKKIRVQGGSSSGPACQVEGCGADLTAAK
DYHRRHKVCEMHAKASTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAVGGTASIEDKVSNYLLL
SLIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNG
CGQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSD
STSTQSLSSSNGDTQ
```

>evm.model.scga7_uti_cns_0093777.1
```
MGSFGMDWNQKASVLWDWELPIGANGSENPRIAAAPQGEAKFAGLEATRHEAVHSSCGAFSSSSEMGYGSSKSSVSASTDSSP
KAKGNSMELNFAPAKAPDKNTDSGKADDARTSPSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKK
AKAAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARR
RKPQPEALSFGSSRLSTMFYSNTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRASGVDGFHLSSQKVWNNF
MPHSVHQDFDGFTALKGTSAKVLDQGVETSAAVSNPNGAPDLQHALSLLSNSSAGAGTNQQPTAQLHHHHAGLSTLVASSSPA
MQASSPGLWQDGGGAALGHHAHTHAHTRFQALDATGSGAAIATTAQELLQLPRPPSSYDSSSSHYDLMR*
```

>evm.model.scga7_uti_cns_0352210.1
```
MDLLGVFSTGLATSTPETNTSQSQGSSDSSGNKSKSHSTEPAIVVNSHDKPTRDFPAAGFMRSNNTYESRPHIYKQTEQETRP
YLSLQLFGSTEEDIPPKIDSLNKYLSSESSNPLDERSPSSSPPITHKFFPIHSVDEEVRHPHITDYGEDATMGEVSTSQAWCA
PPLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQDRTGRIIFKLFGKEPSTIPGNLRDDIVNWLKHSPTEME
GYIRPGCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRTDSQLVSYKAGMTRLSKSWRTWNTPELTFVS
PIAVVGGRKTSLVLKGRNLSIPGTQIHCTSIGKYISKEVLCSAYPGTIYDDSGVETFDLPGQPDLILGRCFIEVENRFRGNSF
PVIVATSSVCLELRNLEVELEDSQVLDVSSDGQIHDSPQSKTRNQVLHFLNELGWLFQRASACTLSSRSDVSDLDLIQFSTAR
FKYLLLFSNERDWCSLTKTLLEILAKRSLASEELSKETMEMLAEIHLLNRAVKRKSRSMVQLLVQFVVICPDNSKVYPFLPNL
PGPGGLTPLHLAASIENAEDIVDALTDDPQQIGLTCWQSVLDDDGQSPETYAKLRDHNSYNELVAQKLVDMKNNQVTITVNGD
EIHMDQLGNVGDHKKSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPYIHSMLAIAAVCVCVCVFMRALLRINSGKSFKW
ERLDFGTI*
```

>evm.model.scga7_uti_cns_0145568.2
```
MTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPVEDGSADCSNSAVADGDNNHV
VSVQRGRSSEEEEEERPRKGASSSTTPSCQVDCCHADLSGARDYHKRHKVCEAHTRTTVVCIKNVEHRFCQQCSRFHLLQEFD
EGKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSTLFLLLKQLSGLEPGSSSEQINHPNSLVNILKNLAAIASTQSYQ
DILKNANSNLILSNAGNNAANGSIVHEQTIRSTPIGRESLAEEPPVKRRVQDFDLNDACIEEAESRTDKIVFKLFGKEPKDFP
VDLREQIQDWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISLSNDGLWRKGWLYARVQDCLTLSCNGS
LMFASPWQPVEDYDQSSLSVPFVVTDNDVCSEIRMLEHGLNLISFDETSERIDDLMAYRSRSLHFLHEIGWLLQRSHMRATSE
QPQYCPDRFSVARFRWLLSFAIDHEWCAVVTKLLNTMFQGDIDLDVPSPIEFALGENLLLTAVRKCSKPLVECLLRYTATNYA
PVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGFTPEEYARQRGHVSYIQMVQ
DKIDRRVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAFFHPRTNRFLSNRPAMLSLV
AIAAVCVCVGLLMKSPPQVACMRPFLWDHIRWGPN*
```

>evm.model.scga7_uti_cns_0110078.2
```
MGSFGMDWNQKASVLWDWENLLPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSASTDS
SPKAKGNSMELNFAPAKAPDKNTDSGKADDARTSLSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASA
KKAKAAPNTQKSYCQVEGCNVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNA
RRRKPQPEALSFGSSRLSTMFYDTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQQVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQGVEASAAVSNPNGAPDLQHALSLLSNSSAGAGTNQQPTAQLHHHHAGLSTLVASSSS
VMQASSPGLWQDGGGAALGHHAHAHAHTRFQALDATGSGAAIATTAQELLQLPRPPSSYDSSSSHYDLMR*
```

>evm.model.scga7_uti_cns_0029591.1

```
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTLTSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVERHVK
NTGTNGRVDGSGNSPSSMIAFSPGEPLISLKLGKRAYFENVCGGQDAKVSAPSDVTSAATVVKKTKVSQQNAKNWYCQVEGCK
VDLSSAKDYNRKHKVCVVHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLTHHNARRRKPQADTISFNSSRLSTM
FYDTSQRTNLFFSQPLYGQVRSTAVPSWDNLGGFKFMETKHLSVHPTKTVGLDELQFSTPQISTSVVAHTGHHNNFDGFMPFK
GTNTKVLNQGVEASVATSNSNNGAPEPRRALSLLSDGSWGSSAVIQQTTSHADAGALPPLATVAVSNAAAGHPLDPSPGRLWQ
CDPPPLDGTAQIQELAHLQTW*
```

>evm.model.scga7_uti_cns_0134244.2
```
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDDETEDQDTVVNGGGLLSLRVGGSAVGVGAMEKSDVNEDERNGKKIRVQGGSSSGPACQVEGCGADLTAAK
DYHRRHKVCEMHAKASTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAVGGTASIEDKVSNYLLL
SLIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNG
CGQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSD
STSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSL
HLDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATINFRVKGFNLLSTSSRLIC
SFEGRCIFQEDTDSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPYIIAEKDVCSEVSELESIFESSSSEH
ANANDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFAMERDWCAVIKMLLDFLFIGLVDVGSRSAEE
VVLSENLLHAAVRKKSVNMVRFLLRYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGISAWS
NARDETGFTPEDYARQRGNDAYLNLVQKKIDKHVGKGQVVLGVPSSICSVITDGVKPGDVSLEICRPMSASVPSCLLCSRQAR
VYPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*
```

>evm.model.scga7_uti_cns_0255781.1
```
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTLSSSSELGHGSSKSSISASIDSPSGVGNSEFNFAAVESHVNN
TGTNGRVDDSGNSRSSMIAFSQGEPLISLKLGKRAYFENVCSGQDAKVSAPSDVTSTATVVKKTKVSQQNAKNWYCQVEGCKV
DLSSAKDYNRKHKVCVVHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLTHHNARRRKPQADTISFNSARLSTMF
YDTSQRSNLFFSQPLYGQVRSSAVPSWDNLGGFKFMETKHLSVQPTKTVGLDELQFSTPQISTSVVAHTGHHHNFDGFMPFKG
TNTKVLNQVSGVEASATASNSNGAPEVGRALSLLSDGSWGSSSAIIQRTTSHADAAGALPPLATVAVSNNAAGHPLDPSLGRL
WQCDPPPPEGTAQIQELAHLRTW*
```

>evm.model.scga7_uti_cns_0179190.3
```
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPTRRGGNEREAKRARGEDGGRGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFDDRRQISFVWNKAPV
SHVRPFTSPWDSSSDLKLPHAKETREVSTKVGTITGQVHLDKSHMSSAIPTLSHGKDEPWPMKGPDMSITASKFDGAPDLQRA
LSLLSVGACGLPDPVHQTSCVIQFTGASENSSDLHLSHGGNSGPASCADGQHIATQPQSQLVRFTMDTSNNVYEPSFFGVNQI
N*
```

>evm.model.scga7_uti_cns_0226575.2
```
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPTRRGGNEREAKRARGEDGGRGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKYPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFDDRRQISFVWNKAPV
SHVRPFTSPWDSSSDLKLPHAKETREVSTKVGTITGQVHLDKSHMSSAIPTLSHGKDEPWPMKGPDMSITASKFDGAPDLQRA
LSLLSVGACGLPDPVHQTSCVIQFTGASENSSDLHLSHGGNSGPASCADGQHIATQPQSQLVRFTMDTSNNVYEPSFFGVNQI
N*
```

>evm.model.scga7_uti_cns_0254505.1
```
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTLTSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVERHVK
NTGTNGRVDGSGNSPSSMIAFSPGEPLISLKLGKRAYFENVCGGQDAKVSAPSDVTSAATVVKKTKVSQQNAKNWYCQVEGCK
VDLSSAKDYNRKHKVCVVHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLTHHNARRRKPQADTISFNSSRLSTM
FYDTSQRTNLFFSQPLYGQVRSTAVPSWDNLGGFKFMETKHLSVHPTKTVGLDELQFSTPQISTSVVAHTGHHNNFDGFMPFK
GTNTKVLNQGVEASVATSNSNNGAPEPGRALSLLSDGSWGSSAVIQQTTSHADAGALPPLATVAVSNAAAGHPLDPSPGRLW
QCDPPPLDGTAQIQELAHLRTW*
```

>evm.model.scga7_uti_cns_0267282.1
```
VHGNGGAAAGGGASGTAAEGKRKEKAAAPAVPRCQVEGCHVALADAKDYHRRHKVCEAHSKAPRVVVLGAEQRFCQQCSRFHA
ISEFDGAKRSCRRRLAGHNERRRKSNASEAMARGVAHPHGVTVGKFPVTVERLTVGKKIQSGLLFKCKSYGTVDSGFVFIPKH
PV*
```

>evm.model.scga7_uti_cns_0154392.3
```
MDWNQKASVLWDWENLLPIGANGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSASTDSLPKAK
```

GNSMELNFAPAKAPDKNTDSGKADDARTSLSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASAKKAKA
APNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNARRRKP
QPEALSFGSSRLSTMFYDTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHISSQQVWNNFMPHG
VHQDFDGFTALKGTSAKVLDQVSSETCPLWLCLVEKIWVKSYCSTFRCYLVKFVQPWIN*


>evm.model.scga7_uti_cns_0292903.1
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTLTSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVERHVK
NTGTNGRVDGSGNSPSSMIAFSPGEPLISLKLGKRAYFENVCGGQDAKVSAPSDVTSAATVVKKTKVSQQNAKNWYCQVEGCK
VDLSSAKDYNRKHKVCVVHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLTHHNARRRKPQADTISFNSSRLSTM
FYDTSQRTNLFFSQPLYGQVRSTAVPSWDNLGGFKFMETKHLSVHPMKTVGLDELQFSTPQISTSVVAHTGHHNNFDGFMPFK
GTNTKVLNQGVEASVATSNSNNGAPEPRRALSLLSDGSWGSSAVIQQTTSHADAGALPPLATVAVSNAAAGHPLDPSPGRLWQ
CDPPPLDGTAQIQELAHLQTW*


>evm.model.scga7_uti_cns_0012627.4
MTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPVENGSADCSNSAVAGGDNNHV
VSVQRGRSSEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTRTTVVCIKNVEHRFCQQCSRFHLLQEFDE
GKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSTLFLLLKQLSGLEPGSSSEQINHPNSLVNILKNLAAIASTQSYQD
ILKNANSNLILSNAGNNAANGSIVHEQTIRSTPIGRESLAEEPPVKRRVQDFDLNDACIEEAESRTDKIVFKLFGKEPKDFPV
DLREQIQDWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISLSNDGLWRKGWLYARVQDCLTLSCNGSL
MFASPWQPVEDYDQSSLSVPFVVTDNDLLNTMFQGDIDLDVPSPIEFALGENLLLTAVRKCSKPLVECLLRYTATNYAPVGSG
DGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGFTPEEYARQRGHVSYIQMVQDKIDR
RVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAFFHPRTNRFLSNRPAMLSLVAIAAV
CVCVGLLMKSPPQVACMRPFLWDHIRWGPN*


>evm.model.scga7_uti_cns_0080443.2
MTGGNRNKSLDWDLNDWRWDTNLFLATPAAAAPSECSIRELSRAQGEIDFGVVVDKRRRLSPVENGSADCSNSAVADGDNNRV
VSVQRGRSSEEEEERPRKGASSSTTPSCQVDGCHADLSGARDYHKRHKVCEAHTRTTVVCIKNVEHRFCQQCSRFHLLQEFDE
GKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSTLFLLLKQLSGLEPGSSSEQINHPNSLVNILKNLAAIASTQSYQD
ILKNANSNLILSNAGNNAANGSIVHEQTIRSTPIGRESLAEEPPVKRRVQDFDLNDACIEEAESRTDKIVFKLFGKEPKDFPV
DLREQIQDWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPASWIEKLISLSNDGLWRKGWLYARVQDCLTLSCNGSL
MFASPWQPVEDYDQSSLSVPFVVTDNDVCSEIRTLEHGLNLISFDETSERIDDLMAYRSRSLHFLHEIGWLLQRSHMRATSEQ
PQYCPDRFPVTRFRWLLSFAIDHEWCAVVTKLLNTMFQGDIDLDVPSPIEFALGENLLLTAVRKCSKPLVECLLRYTATNYAP
VGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDPQQLGIKAWKNARDATGFTPEEYARQRGHVSYIQMVQD
KIDRRVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSINQTSSCRQCVQQVQQLAFFHPRTNRFLSNRPAMLSLVA
IAAVCVCVGLLMKSPPQVACMRPFLWDHIRWGPN*


>evm.model.scga7_uti_cns_0307857.3
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPARRGGNEREAKRARGEDGGRGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFDDRRQISFVWNKAPV
SHVRPFTFPWDSSSDLKLPHAKETREVSTKVGTITGQVYLDKSHMSSAIPTLSHGKDEPWPMKGPDMSITASKFDGAPDLQHA
LSLLSVGACGLPDPVHQTSCVIQFTGASENSSDLHLSHGGNSGPASCADGQHIATQPQSQLVRFTMDTSNNVYEPSFFGVNQI
KPLDHLDK*


>evm.model.scga7_uti_cns_0253558.3
MAAEAQRQGAGGSGALNLQLGLREDAATSMDASPTAPVASSSPSSPPTSAAAGQEPVVRPTKRVRSGSPGSAGGGSAGGGAAN
GGASYPMCQVDDCRADLTSAKDYHRRHKVCETHSKTTKAVVANQAQRFCQQCSRFHPLAEFDEGKRSCRRRLAGHNRRRRKTQ
PTDVASQLLLPVNQENAANRTQDIVNLITVIARLQGSNVGKVPSIPPIPDKQNLVEIISKINSLNNTTPAAKSPPSEVVDLNA
SQEQQEQRQDSVDKTTNVIDKQTVPSTMDLLGVFSTGLATSTPETNTSQSQGSSDSSGNNKSKSHSTEPAIVVNSHDKPTRDF
PAAGFMRSNSTYESRPHIYKQTEQETRPYLSLQLFGSTEEDIPPKIDSLNKYLSSESSNPLDERSPSSSPPITHKFFPIHSVD
EEVRHPHITDYGEDATMGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQDRTGRIIFKL
FGKEPSTIPGNLRDDIVNWLKHSPTEMEGYIRPGCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRTDS
QLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVVGGRKTSLVLKGRNLSIPGTQIHCTSIGKYISKEILCSAYPGTIYDDSGVE
TFDLPGQPDLILGRCFIEVENRFRGNSFPVIVASSSVCLELRNLEVELEDSQVLDVSSDGQIHDSPQSKTRNQVLHFLNELGW
LFQRASACTLSSRSDVSDLDLIQFSTARFKYLLLFSNERNWCSLTKTLLEILAKRSLASEELSKETMEMLAEIHLLNRAVKRK
SRSMVHLLVQFVVICPDNSKVYPFLPNLPGPGGLTPLHLAASIENAEDIVDALTDDPQQIGLTCWQSVLDDDGQSPETYAKLR
NHNSYNELVAQKLVDMKNNQVTITVNGDEIHMDQLGNVGDHKKSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPYIHSM
LAIAAVCVCVCVFMRALLRINSGKSFKWERLDFGTI*


>evm.model.scga7_uti_cns_0211345.1

```
MGSFGMDWNQKASVLWDWENLLPIGANGSENPRMAAAPQGEAKFAGLQATRHESVHSSCGTFSSSSEMGYGSSKSSVSVSTDS
SPKAKGNSMELNFAPAKAPDKNTDSGKADDARTSLSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASA
KKAKAAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNA
RRRKPQPEALSFGSSRLSTMFYDTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQQVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQGVEASAAVSNPNGAPDLQHALSLLSNSSAGAGTNQQPTAQLHHHHAGLSTLVASSSS
VIPASSPGLWQDGGGAALGHHAHAHAHTRFQALDATGSGAAIATTAQELLQLPRPPSSYDSSSSHYDLMR*
```

>evm.model.scga7_uti_cns_0194130.1
```
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPTRRGGNEREAKRARGEDGGRGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFDDRRQISFVWNKAPV
SHVRPFTSPWDSSSDLKLPHAKETREVSTKVGTITGQVHLDKSHMSSAIPTLSHGKDEPWPMKGLIFQTLFLPAAEMCSRKKA
AICFVNWFTIFVDKGL*
```

>evm.model.scga7_uti_cns_0305902.1
```
MGSFGMDWNQKASVLWDWENLLPIGTNGSENPRMAAAPQGEAKFAGLEATRHESVHSSCGTFSSSSEMGYGSSKSSVSASTDS
SPKAKGNSMELNFAPAKAPDKNTDSGKADDARTSLSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASA
KKAKVAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNA
RRRKPQPEALSFGSSRLSTMFYDTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQQVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQ
```

>evm.model.scga7_unitig_211197.2
```
MEGVVEVMVSVDRYLAEESNDSVCHDRFHLLQEFDEGKKSCRSRLAKHNGRRRKVQPQAAVNGNSMNENQSLSSTLFLLLKQL
SGLEPGSSSEQINHPNSLVNILKNLAAIASTQSYQDILKNANSNLILSNAGNNAANGSIVHEQTIRSTPIGRESLAEEPPVKR
RVQDFDLNDACIEEAESRTDKIVFKLFGKEPKDFPVDLREQIQDWLSHYPTDMESYIRPGCVILTIYLRLPNWMWDELDDDPA
SWIEKLISLSNDGLWRKGWLYARVQDCLTLSCNGSLMFASPWQPVIGDKLQRLCVTPIAVACSSSVKFSVQGFNIVQPTTKLL
CVFGEKYLIQEETQLLLEDSTMQQGPQCLTFSCSFPCTSGRGFIEVEDYDQSSLSVPFVVTDNDVCSEIRTLEHGLNLISFDE
TSERIDDLMAYRSRSLHFLHEIGWLLQRSHMRATSEQPQYCPDRFPVARFRWLLSFAIDHEWCAVVTKLLNTMFQGDIDLDVP
SPIEFALGENLLLTAVRKCSKPLVECLLRYTATNYAPVGSGDGAPVQFLFTPAMTGLSNITPLHIAATISDATGVLDALTDDP
QQLGIKAWKNARDATGFTPEEYARQRGHVSYIQMVQDKIDRRVSRDHVSVTIPTTIDTVGKHTSGLKSADQITFGVEKQLSIN
QTSSCRQCVQQVQQLAFFHPRTNRFLSNRPAMLSLVAIAAVCVCVGLLMKSPPQVACMRPFLWDHIRWGPN*
```

>evm.model.scga7_uti_cns_0038913.2
```
MEAARFGTQSSHLYGGGLGELDLNRRENRVFGWDLNDWSWDSQRFVATPVPAEAANGSGLNSSPSTSEEAEEEVARNGGLRGD
TDKRKRVVVIDDDETEDQDTVVNGGGLLSLTVGGSAVGVGVMEKSDVNEDERNGKKIRVQGGSSSGPACQVEGCGADLTAAKD
YHRRHKVCEMHAKASTAVVGNTVQRFCQQCSRFHLLQEFDEGKRSCRRRLAGHNRRRRKTRPDTAIGGTASIEDKVSNYLLLS
LIGICANLNSDNVQNSNGQELLSTLLKNLGSVAKSLEPKELCKLLEAYQSLQNGSNAGTSGTANAAEEAAGPSNSKLPFVNGC
GQASSSVVPVQSKATMVVTPEPASCKLKDFDLNDTCNDMEGFEDGQEGSPTPAFKTADSPNCASWMQQDSTQSPPQTSGNSDS
TSTQSLSSSNGDTQCRTDKIVFKLFDKVPSDLPPVLRSQILGWLSSSPTDIESHIRPGCIILTVYLRLVESAWQELSENMSLH
LDKLLSSSTDDFWASGLVFVMVRRRLAFMLNGHIMLDRPLAPSSHHYCKILCVRPVAAPYSATINFRVKGFNLLGTSSRLICS
FEGRCVFQEDTDSVAENAEYEDRDIECLSFCCSVPGPRGRGFIEVEDSGFSNGFFPFIIAEKAVCSEVSELESIFESSSSEDA
NVNDNARDQALEFLNELGWLLHRANRMSKEDETDTSLAPFNMWRFRNLGVFAMERDWCAVIKMLLDFLFIGLVDVGSRSAEEV
VLSENLLHAAVRRKSVNMVRFLLKYKPNKNSKGTAQTYLFRPDALGPSTITPLHIAAATSDAEDVLDVLTDDPGLIGISAWSN
ARDETGFTPEDYARQRGNDAYLNLVQKKIDKHLGKGQVVLGVPSSICSVITDGVKPGDVSLEICRPMSASVPSCLLCSRQARV
YPNSTSRTFLYRPAMLTVMGVAVVCVCVGILLHTFPRVYAAPTFRWELLERGPM*
```

>evm.model.scga7_uti_cns_0211345.1.1
```
MGSFGMDWNQKASVLWDWENLLPIGANGSENPRMAAAPQGEAKFAGLQATRHESVHSSCGTFSSSSEMGYGSSKSSVSVSTDS
SPKAKGNSMELNFAPAKAPDKNTDSGKADDARTSLSSVIAISSGEPVLSLKLGKRTYFEDVCGGQSVKSLPSDTSAVTPPASA
KKAKAAPNTQKSYCQVEGCKVDLSSAKEYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLSEFDQKKKSCRRRLNDHNA
RRRKPQPEALSFGSSRLSTMFYDTRQQTSLLFGQGPYGQMRSCASSSWDNPVGGFKFEETKAPWLRAAGVDGFHLSSQQVWNN
FMPHGVHQDFDGFTALKGTSAKVLDQGVEASAAVSNPNGAPDLQHALSLLSNSSAGAGTNQQPTAQLHHHHAGLSTLVASSSS
VIPASSPGLWQDGGGAALGHHAHAHAHTRFQALDATGSGAAIATTAQELLQLPRPPSSYDSSSSHYDLMR*
```

>evm.model.scga7_uti_cns_0158300.4
```
MGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSCYASTSCSDHSPSTSNSDGQDRTGRIIFKLFGKEPSTIPGNLRDDI
VNWLKHSPTEMEGYIRPGCLVLSMYLSMPAIAWDELEENLLQRVNSLVQSSDLDFWRKGRFLVRTDSQLVSYKAGMTRLSKSW
RTWNTPELTFVSPIAVVGGRKTSLVLKGRNLSIPGTQIHCTSIGKYISKEILCSAYPGTIYDDSGVETFDLPGQPDLILGRCF
IEVENRFRGNSFPVIVASSSVCLELRNLEVELEDSQVLDVSSDGQIHDSPQSKTRNQVLHFLNELGWLFQRASACTLSSRSDV
SDLDLIQFSTARFKYLLLFSNERDWCSLTKTLLEILAKRSLASEELSKETMEMLAEIHLLNRAVKRKSRSMVHLLVQFVVICP
```

DNSKVYPFLPNLPGPGGLTPLHLAASIENAEDIVDALTDDPQQIGLTCWQSVLDDDGQSPETYAKLRNHNSYNELVAQKLVDM
KNNQVTITVNGDEIHMDQLGNVGDHKKSGVQALQIRSCSQCAILESGVLRQPVRSRGLLARPYIHSMLAIAAVCVCVCVFMRA
LLRINSGKSFKWERLDFGTI*

>evm.model.scga7_uti_cns_0063998.5
AAPSRAECSVDLKLGGLGEFGAADGTKEPAAAATATPAPSASPMKRPRSGPGGAAGAQCPSCAVDGCKADLSKCRDYHRRHKV
CEAHSKTPVVVVAGREMRFCQQCSRFHLLAEFDEAKRSCRKRLDGHNRRRRKPQPDTMNSGSFMTSQQGTRFSSFPAPRPEPS
WSGVIKSEDTSYYTPHAVLSNRPHFAGGSTSSAYSKEGRRFPFLQDGDQISFSAGAGTLEVSTVCQPLLKTTAAAALLPPESS
SSNKMFSDGLTPVLDSDCALSLLSSPANSSVDVSRMVVQPAAEHIPVAQPLVPSSSLQQHHQFGSSPGWFACSQAGSSGAVSA
AGTGGFACPASVESEQLNTVLQAPSSDAGHEMNYHGIFHVGGEGSSDGTSPSLPFSWQ*

>evm.model.scga7_uti_cns_0198101.1
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTFTSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVERHVK
NTGTNGRVDGSGNSPSSMIAFSPGEPLISLKLGKRAYFENVCGGQDAKVSAPSDVTSAATVAKKTKVSQQNAKNWYCQVQGCK
VDLSSAKDYNRKHKVCVVHSKAPKVVAAGLERRFCQQCSRFHGLAEFDQNKRSCRRCLTHHNARRRKPQADTISFNSSRLSTM
FYDTSQRTNLFFSQPLYGQVRSTAVPSWDNLGGFKFMETKRLSVQPTKTVGLDELQLSTPQISTSVVAHTGHHNNFDGFMPFK
GTNTKVLNQGVEASVSTSNSNNGAPEPGRALSLLSDGSWGSSSAVIQQTTSHADAGASPPLATVAVSNAAAGHPLDPSPGRLW
QCDPPPLDGTAQIQELAHLPCHLEGLGTSG*

>evm.model.scga7_unitig_297994.1
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPARRGGNEREAKRARGEDGGRGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFDDRRQISFVWNKAPV
SHVRPFTFPWDSSSDLKLPHAKETREVSTKVGTITGQVYLDKSHMSSAIPTLSHGKDEPWPMKGLIFQTLILPAAEMCSRKKL
LFALLTGSPYLLTKGCKESLV*

>evm.model.scga7_uti_cns_0198066.1
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTLTSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVERHVK
NTGTNGRVDGSGNSPSSMIAFSPGEPLISLKLGKRAYFENVCGGQDAKVSAPSDVTSAATVVKKTKVSQQNAKNWYCQVEGCK
VDLSSAKDYNRKHKVCVVHSKAPKVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLTHHNARRRKPQADTISFNSSRLSTM
FYDTSQRTNLFFSQPLYGQVRSTAVPSWDNLGGFKFMETKHLSVHPTKTVGLDELQFSTPQISTSVVAHTGHHNNFDGFMPFK
GTNTKVLNQGVEASVATSNSNNGAPEPGRALSLLSDGSWGSSSAVIQQTTSHADAGALPPLATVAVSNAAAGHPLDPSPGRLW
QCDPPPLDGTAQIQELAHLRTW*

>evm.model.scga7_uti_cns_0026424.2
MDWAAASGAPSWGAAATDPGTTILSFAGPSSSTSPTADAEARLQDFAAGLVQSARPAGAAGAGRRSRAAGGGAGAEACSVDGC
HSDLSRCREYHRRHKVCEAHAKTPVVVVGGQEQRFCQQCSRFHMLSEFDEGKRSCRKRLDGHNRRRRKPQHDLTNLGGFFPYH
QVNQFEVYPRTIPTVGQNSDAMQLVGRQQPFSISFSRTPNQFPFPQGGGSALHAARRGLLAEGSSHTGSSTCNNALSGTLGPE
CALSLLSPLLHRPSAAGIPTDGQPQVASSLSRIAAVSQAATTAMTIAFAAGVGHHVFVPDAMFEDPSQALPFSWQ*

>evm.model.scga7_uti_cns_0189244.1
MGSFGMNWNQKNSMVWDWEHLVPSVSNAVTRHGSANSSGGTFTSSSELGHGSSKSSISASIDSPSGVGNSLEFNFAAVERHVK
NTGTNGRVDGSGNSPSSMIAFSPGEPLISLKLGKRAYFENVCGGQDAKVSAPSDVTSAATVAKKTKVSQQNAKNWYCQVQGCK
VDLSSAKDYNRKHKVCVVHSKAPKVVAAGLERRFCQQCSRFHGLAEFDQNKRSCRRCLTHHNARRRKPQADTISFNSSRLSTM
FYDTSQRTNLFFSQPLYGQVRSTAVPSWDNLGGFKFMETKRLSVQPTKTVGLDELQLSTPQISTSVVAHTGHHNNFDGFMPFK
GTNTKVLNQGVEASVSTSNSNNGAPEPGRALSLLSDGSWGSSSAVIQQTTSHADAGASPPLATVAVSNAAAGHPLDPSPGRLW
QCDPPPLDGTAQIQELAHLPCHLEGLGTS

>evm.model.scga7_uti_cns_0251330.1
MEWTAPKPASSPPLLWDWGDHAATGSGSSSGDAPTRRGGNEREAKRARGEDGGRGAEVRCQVEGCGLDLTRVKDYHRKHRVCE
AHTKSPRVIVAGQERRFCQQCSRFHALSEFDQKKRSCRRRLSDHNARRRKPQPDAFTFASARLPSSLFDDRRQISFVWNKAPV
SHVRPFTSPWDSSSDLKLPHAKETREVSTKVGTITGQVHLDKSHMSSAIPTLSHGKDEPWPVKGLIFQTLFLPAAEMCSRKKL
LFALLTGSPYLLTKGCKESLV*

>evm.model.scga7_unitig_440261.2
FHPLAEFDEGKRSCRRRLAGHNRRRRKTQPTDVASQLLLPVNQENAANRTQDIVNLITVIARLQGSNVGKVPSIPPIPDKQNL
VEIISKINSLNNTTPAAKSPASEVADLNASQEQQEQRQDSVEKTTNVIDKQTVPSTMDLLGVFSTGLATSTPETNTSQSQGSS
DSSGNNKSKSHSTEPAIVVNSHDKPTRDFPAAGFMRSNSAYESRPHIYKQTEQETRPYLSLQLFGSTEEDIPPKIDSLNKYLS
SESSNPLDERSPSSSPPITRKFFPIHSVDEEVRHPHITDYGEDATMGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSC
YASTSCSDHSPSTSNSDGQDRTGRIIFKLFGKEPSTIPGNLRDDIVNWLKHSPTEMEGYIRPGCLVLSMYLSMPAIAWDELEE

NLLQRVNSLVQSSDLDFWRKGRFLVRTDSQLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVVGGRKTSLVLKGRNLSIPGTQI
HCTSIGKYISKEVLCSAYPGTIYDDSGVETFDLPGQPDLILGRCFIEVENRFRGNSFPVIVASSSVCLELRNLEVELEDSQVL
DVSSDGQIHDSPQSKTRNQVLHFLNELGWLFQRASACTLSSRSDVSDLDLIQFSTARFKYLLLFSNERDWCSLTKTLLEILAK
RSLASEELSKETMEMLAEIHLLNRAVKRKSRSMVHLLVQFVVICPDNSKVYPFLPNLPGPGGLTPLHLAASIENAEDIVDALT
DDPQQIGLTYWQSVLDDDGQSPETYAKLRNHNSYNELVAQKLVDMKNNQVTITVNGDEIHMDQLGNVGDHKKSGVQALQIRSC
SQCAILESGVLRQPVRSRGLLARPYIHSMLAIAAVCVCVCVFMRALLRINSGKSFKWERLDFGTI

>evm.model.scga7_uti_cns_0110246.1
FHPLAEFDEGKRSCRRRLAGHNRRRRKTQPTDVASQLLLPVNQENAANRTQDIVNLITVIARLQGSNVGKVPSIPPIPDKQNL
VEIISKINSLNNTTPAAKSPASEVVDLNASQEQQEQRQDSVEKTTNVIDKQTVPSTMDLLGVFSTGLATSTPETNTSQSQGSS
DSSGNNKSKSHSIEPAIVVNSHDKPTRDFPAAGFMRSNSAYESRPHIYKETEQETRPYLSLQLFGSTEEDIPPKIDSLNKYLS
SESSNPLDERSPSSSPPITRKFFPIHSVDEEVRHPHITDYGEDATMGEVSTSQAWCAPPLDLFKDSERPIENGSPPNPGYQSC
YASTSCSDHSPSTSNSDGQDRTGRIIFKLFGKEPSTIPGNLRDDIVNWLKHSPTEMEGYIRPGCLVLSMYLSMPAIAWDELEE
NLLQRVNSLVQSSDLDFWRKGRFLVRTDSQLVSYKAGMTRLSKSWRTWNTPELTFVSPIAVVGGRKTSLVLKGRNLSIPGTQI
HCTSIGKYISKEVLCSAYPGTIYDDSGVETFDLPGQPDLILGRCFIEVENRFRGNSFPVIVASSSVCLELRNLEVELEDSQVL
DVSSDGQIHDSPRSKTRNQVLHFLNELGWLFQRASACTLSSRSDVSDLDLIQFSTARFKYLLLFSNERDWCSLTKTLLEILAK
RSLASEELSKETMEMLAEIHLLNRAVKRKSRSMVHLLVQFVVICADNSKVYPFLPNLPGPGGLTPLHLAASIENAEDIVDALT
DDPQQIGLTYWQSVLDDDGQSPETYAKLRNHNSYNELVAQKLVDMKNNQVTITVNGEIHMDQLGNVGDHKKSGVQALQIRSCS
QCAILESGVLRQPVRSRGLLARPYIHSMLAIAAVCVCVCVFMRALLRINSGKSFKWERLDFGTI