Ana Teresa Talhinhas Barata Alhinho

The role of NEP-TC in the somatic embryogenesis of tamarillo (*Solanum betaceum* Cav.)

Tese de Mestrado em Biodiversidade e Biotecnologia Vegetal

2016

· U C ·

Universidade de Coimbra

DEPARTAMENTO DE CIÊNCIAS DA VIDA

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

The role of NEP-TC in the somatic embryogenesis of tamarillo (*Solanum betaceum*)

Dissertação apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biotecnologia Vegetal, realizada sob a orientação científica da Doutora Sandra Isabel Correia (Centro de Ecologia Funcional - Universidade de Coimbra) e da Professora Doutora Paula Cristina Veríssimo Pires (Faculdade de Ciências da Universidade de Coimbra)

Ana Teresa Talhinhas Barata Alhinho

2016

The picture on the cover is a mosaic by Nathan Hopkins

Agradecimentos

Em primeiro lugar, gostaria de agradecer às minhas orientadoras. À Doutora Sandra Correia, pela sua orientação sempre disponível, pelos seus ensinamentos e auxílio no laboratório, pela revisão de todo este trabalho e sobretudo por toda a amizade e incentivo. À Professora Doutora Paula Veríssimo, a quem estou verdadeiramente grata pela orientação e integração no Laboratório de Biotecnologia Molecular, pela sua disponibilidade e atenção no delineamento experimental e revisão do trabalho e pela confiança que depositou em mim.

Ao Professor Doutor Jorge Canhoto, por me ter confiado este projeto e pela atenta revisão de todo o trabalho.

À Doutora Rosário Faro, pela cedência das células competentes necessárias à expressão da proteína. À Doutora Mónica Zuzarte, pela ajuda na obtenção dos cortes histológicos, assim como à Doutora Luísa Cortes, do CNC, pelo auxílio na obtenção e análise das imagens de microscopia confocal.

À Doutora Aileen Justies, da NanoTemper Technologies, pela oportunidade de realizar os ensaios de termoforese e pela sua disponibilidade na análise dos resultados.

À Inês Mota e ao Nuno Beltrão, pela preciosa ajuda na expressão da proteína nos momentos em que estava ocupada com outras vertentes do trabalho.

Um agradecimento muito especial à Professora Doutora Alexandra Rosa da Costa, da Universidade de Évora, que muito me motivou para o estudo das plantas, e cujos ensinamentos e constante incentivo foram determinantes na escolha do meu percurso académico.

Aos meus colegas de mestrado, em especial à Diana Graça, pelas tardes passadas no Sr. Borges a discutir resultados da tese (e o quão detestamos o Jar-Jar Binks), e sobretudo por me ter feito sentir que não estava sozinha nesta demanda. May the Force be with you (eu disse-te que isto não era uma tese completa sem referências a Star Wars).

Aos meus amigos eborenses, em particular à Inês Amaro, minha parceira de tantos dias passados na biblioteca de Évora a trabalhar nas respetivas teses. Um dia vamos estar no rooftop de um edifício desenhado por ti e rir de todos os momentos em que nos ocorreu desistir. Ao Pedro Pereira, por ter sido uma "bridge over troubled water". Por me ter apoiado incondicionalmente e ter sempre uma palavra de conforto e incentivo nos momentos de dúvida. À Patrícia Serrano e ao Luís Caldeira, meus "afilhados" da Universidade de Évora, sempre disponíveis para me ouvirem e motivarem a fazer melhor, e a quem sempre tentei nunca desiludir.

Ao "Mestre André" Caeiro, meu companheiro nesta aventura (e em tantas outras). Obrigada pelo auxílio no laboratório e por teres sido o meu revisor não-oficial de todo este trabalho. Esta era a parte em que eu escrevia aqui algo emotivo, mas como não gostamos dessas coisas vou pôr algo que só tu compreendes: "we're gonna need a bigger boat".

À minha família, sobretudo aos meus pais, Ana e Silvino, e à minha avó Bia, por todo o apoio ao longo do meu percurso académico.

Index

Index of figures

Index of tables

Index of Abbreviations

2,4-D – 2,4-Diclorophenoxiacetic acid

ACC – 1-aminocyclopropane-1-carboxylic acid

ATP – Adenosine triphosphate

CUC – *CUP-SHAPED COTYLEDON*

dcSAM – Decarboxylized S-adenosil-L-methionine

DNA – Desoxyribonucleic acid

EC – Embryogenic callus

EDTA – Ethylenediaminetetraacetic acid

ELISA – Enzyme-Linked Immunosorbent Assay

EMSA – Electrophoretic mobility shift assay

FAD – Flavin adenine dinucleotide

GST – Glutatione S-transferase

IDA – Iminodiacetic acid

IMAC – Immobilized metal ion affinity chromatography

IPTG – Isopropyl β-D-1-thiogalactopyranoside

IR – Infrared

LC-MS – Liquid chromatography–mass spectrometry

LEC – Leaf embryogenic callus

MAT – Methionine adenosyltransferase

MST – Micro Scale Termophoresis

NAA – Naphtaleneacetic acid

NAD – Nicotinamide and adenine dinucleotide

NEC – Non-embryogenic callus

NEP-TC – Non-embryogenic protein of tamarillo callus

NTA – Nitriloacetic acid

ori – Replication origin

PCR – Polymerase chain reaction

ppGpp – Guanosine pentaphosfate

RFM – Rossmann-fold

RID – ROOT INITIATION DEFECTIVE

RNA – Ribonucleic acid

SAH – S-adenosil-homocysteine

SAHH – S-adenosil-homocysteine hidrolase

SAM – S-adenosil-L-methionine

SE – Somatic embryogenesis

SERK – SOMATIC EMBRYOGENESIS RECEPTOR- LIKE KINASE

TEMED - Tetramethylethylenediamine

TIR – Translation initiation region

UV – Ultraviolet

Abstract

Tamarillo (*Solanum betaceum* Cav.) is a small solanaceous tree known for its edible and nutritious fruits. This species is extremely useful for understanding the process of induction of somatic embryogenesis (SE), an important tool for plant cloning and *in vitro* propagation. During this process, both embryogenic (EC) and non-embryogenic (NEC) *calli* arise from the same explant, in a medium with plant growth regulators (picloram or 2,4-dichlorophenoxyacetic acid).

NEP-TC (26.5 kDa) is a protein identified as being consistently present in tamarillo's non-embryogenic *calli*, having been associated with the SpoU SAM-dependent RNA methyltransferase family, by bioinformatic comparison with homologous sequences found in the model plant *Arabidopsis thaliana*.

The present study aims to elucidate the role of NEP-TC on tamarillo somatic embryogenesis, as well as perform an extensive characterization of the protein, by activity tests to evaluate the methyltransferase activity of recombinant NEP-TC expressed in *Escherichia coli* and immunohistochemical studies for the localization of NEP-TC on both embryogenic and non-embryogenic *calli*. Bioinformatic studies were also performed to determine the structure and phylogeny of this protein.

After successful protein expression on a heterologous system, specific activity results show that NEP-TC has affinity towards RNA, while showing no activity rates towards DNA, thus confirming previous bioinformatic studies performed. Analyzing the protein's affinity towards ribosomal RNA (rRNA), it is clear that the protein has affinity towards this substrate. When it comes to immunolocalization, NEP-TC shows a differential expression, with a consistent presence in peripheral areas of non-embryogenic tissues and a less distinct manifestation in embryogenic tissues and cells in

the process of differentiation. Bioinformatic analysis shows that this protein belongs to a gene family with only one gene that has been suffering alterations over time originating the 38 genes present in the genome of this tree.

Key words: bioinformatics; embryogenic competence; immunohistochemistry; *in vitro* culture; NEP-TC; recombinant protein expression; RNA methyltransferases; somatic embryogenesis

Resumo

O tamarilho (*Solanum betaceum* Cav.) é uma árvore de pequeno porte da família das solanáceas, conhecido pelos seus frutos comestíveis e nutritivos. O estudo desta espécie é bastante útil na compreensão do processo de indução de embriogénese somática (ES), uma importante ferramenta para clonagem de plantas e sua propagação *in vitro*. Durante este processo, tanto calos embriogénicos (CE) como não-embriogénicos (CNE) surgem no mesmo explante, num meio de cultura com reguladores de crescimento vegetal (picloram ou ácido 2,4-diclorofenoxiacetico).

A NEP-TC é uma proteína expressa de forma consistente em calos não-embriogénicos de tamarilho, tendo sido associada à família SpoU de metiltransferases de RNA dependentes de SAM por comparação bioinformática com sequencias homólogas presentes na planta modelo *Arabidopsis thaliana.*

O presente estudo tem como objetivo a elucidação do papel da NEP-TC na embriogénese somática de tamarilho, mas também fazer uma caracterização extensiva desta proteína, através da realização de ensaios de atividade específica para avaliação da ação como metiltransferase da NEP-TC recombinante expressa em *Escherichia coli*, assim como estudos de imunohistoquímica para determinação dos locais de expressão da proteína em tecidos embriogénicos e não-embriogénicos. Foram também realizados ensaios de bioinformática para determinação da filogenia e estrutura desta proteína.

Após expressão da proteína num sistema heterólogo, os resultados de atividade mostram que a NEP-TC apresenta de facto afinidade para com RNA, não apresentando taxas de atividade em relação a DNA, confirmando desta forma estudos de bioinformática realizados previamente. Analisando a afinidade da proteína relativamente a RNA ribossomal (rRNA), esta apresenta afinidade para com este substrato. Quanto à

imunolocalização, a NEP-TC mostra expressão diferencial, com uma presença consistente em tecidos não embriogénicos, e uma manifestação menos distinta em tecidos embriogénicos e tecidos em diferenciação. A análise bioinformática mostra que esta proteína pertence a uma família de genes com apenas um gene que sofreu alterações ao longo do tempo, originando os 38 genes presentes na árvore.

Palavras-chave: bioinformática; competência embriogénica; cultura *in vitro*; embriogénese somática; expressão de proteínas recombinantes; imunohistoquímica; NEP-TC; metiltransferases de RNA

1. Introduction

1.1. General introduction

The induction of somatic embryogenesis (SE) was first described by Stewart *et al*. (1958) and Reinert (1959) on carrot (*Daucus carota* L.), and is currently a biotechnological tool of utmost importance, not only as a model of regeneration but also as a system for fundamental studies of physiology and biochemistry of plant development (Quiroz-Figueroa *et al.*, 2006). However, the molecular mechanisms behind this complex process are still widely unknown, and most of the current SE protocols are developed based on trial and error tests, and there is still a large number of species that are recalcitrant to this process (Jiménez, 2005; Takac *et al*., 2011). Woody plants exhibit a higher level of recalcitrance in relation to other species, being particularly difficult to study due to its long life cycle (Canhoto, 2010).

Tamarillo is a woody plant of the Solanaceae family, whose somatic embryogenesis was first described by Guimarães *et al*. (1988). Since then, the somatic embryogenesis of this plant has been widely studied, and optimized protocols have been established for the development of somatic embryos and further regeneration of adult plants (Correia and Canhoto, 2012).

Somatic embryogenesis is a multifactorial process, and may be influenced by a variety of physiological, genetic and exogenous parameters (stress and plant growth regulators) (Feher, 2015). To determine the role of these factors, studies have been carried out in tamarillo: proteomic studies (Correia *et al*., 2012) indicate the relevance of several families of proteins in the process. Among these, NEP-TC (*Non Embryogenic Protein of Tamarillo Callus*) stands out, being consistently expressed in non-embryogenic *calli*, but not in embryogenic *calli*, and with varying expression levels in induced tissues. By sequence bioinformatics analysis, it was possible to determine that this protein is a putative RNA methyltransferase (Correia, 2011). Recent studies in

Arabidospis thaliana (Ohbayashi *et al*., 2011; Shinohara *et al.*, 2014) indicate that methyltransferases can have a decisive role in the regulation of cellular stress during the induction of somatic embryogenesis.

However, the role of this protein has not yet been fully unraveled, and it is considered that it may constitute an important tool for a better understanding of the mechanisms of differentiation between different types of embryogenic *calli* during somatic embryogenesis induction. Thus, studies are required to perform a biochemical and physiological characterization of this protein role on this morphogenic process.

1.2.Somatic embryogenesis

Somatic embryogenesis may be described as the process from which differentiated somatic cells (haploid or diploid) cultured on a define medium originate structures that resemble zygotic embryos (Jiménez, 2001). The embryos resulting from this process have similar characteristics to zygotic embryos, such as the bipolar organization and the absence of vascular contact with parental tissues, as well as the stages of development - globular, heart, torpedo and cotyledonary (Jiménez, 2001; Canhoto, 2010).

This totipotency was first described in carrot (Stewart *et al*., 1958; Reinert, 1959), quickly becoming the target of studies which confirmed this process as an important means of plant regeneration from plant cell culture systems, as well as a good model for the study of biochemical, physiological and morphogenical processes that occur during the plant embryogenesis (Zimmerman, 1993; Jiménez, 2005; Quiroz-Figueroa *et al.*, 2006).

The induction of SE is a multi-factorial event, and the determination of physical and chemical factors involved in the whole reprogramming of the cell for acquisition of embryogenic competence is difficult (Karami *et al.*, 2009). It is well established that growth regulators and stresses are two key factors in mediating the signaling cascade that leads to reprogramming of gene expression, followed by successive divisions that allow either the development of embryogenic *calli* and the subsequent formation of somatic embryos or the direct formation of embryos (Feher, 2008). In fact, some authors suggest that the great variability of induction conditions may be due to the interaction of growth regulator/stress factor with the levels of endogenous auxins of the explant (Feher, 2015).

Auxins play a key role in SE, since the responses of cells in culture to these regulators constitute a key event in the cell adaptation (in genetic, metabolic and physiological levels), leading to the acquisition of embryogenic competence by somatic cells, and being also relevant to the proliferation and maintenance of pro-embryogenic masses (Feher *et al*., 2003). 2,4-dichlorophenoxyacetic acid (2,4-D) is particularly efficient in this process: many embryogenic *in vitro* systems depend on the use of exogenous 2,4-D as a SE inducer (Pasternak *et al*., 2002). The development of a cell with embryogenic capacity is largely related to the methylation of DNA in the presence of 2,4-D, by changes in the chromatin structure (under the influence of this synthetic auxin) that lead to genomic reprogramming of somatic cells (Leljak-Levanic *et al*., 2004; Feher, 2015). In fact, some authors suggest that 2,4-D acts primarily as a stress inducer, and its role as growth regulator is merely secondary (Zavattieri *et al*., 2010).

In vitro culture conditions include a high number of stress factors to the plant cell (oxidative, temperature and osmotic stress, hypoxia, UV-radiation and physical/chemical treatments, among others), and the stress associated with the

induction of SE can result in a generalized response which is translated into the reorganization of chromatin: this, it is believed, causes an activation of the embryogenic program (Feher, 2005). This activation of successive differentiations and dedifferentiations is characterized by the occurrence of organized patterns of DNA methylation and acetylation of histones (Leljak-Levanic *et al*., 2004).

DNA methylation is a remarkable process, since it involves a covalent modification of the cell's genetic material, consequently changing its expression (Karami and Saidi, 2010). In *Arabidopsis*, it has already been shown that the action of the methyltransferase-1 influences gene expression during embryogenesis induction: thus, the changes in chromatin structure for DNA methylation in the presence of 2,4-D lead to genome reprogramming in somatic cells, and genes needed for the acquisition of embryogenic competence are expressed (Xiao *et al*., 2006). In studies conducted in *Cucurbita pepo*, the highest rate of DNA methylation occurred during early stages of development of embryos in the medium containing 2,4-D, while the occurrence of this phenomenon decreased during the maturation of embryos in medium without auxins (Leljak-Levanic *et al*., 2004). In epidermal cells of carrot, DNA (cytosine-5)-methyltransferase (*MET*) genes were expressed after induction of SE in the presence of 2,4-D, and the formation of embryogenic masses was suppressed after addition of 5-azacitidine, an inhibitor of DNA methylation (Yamamoto *et al.*, 2005).

Figure 1 shows the factors that may influence acquisition of embryogenic competence, as well as the changes that occur in gene and protein expression during this developmental process.

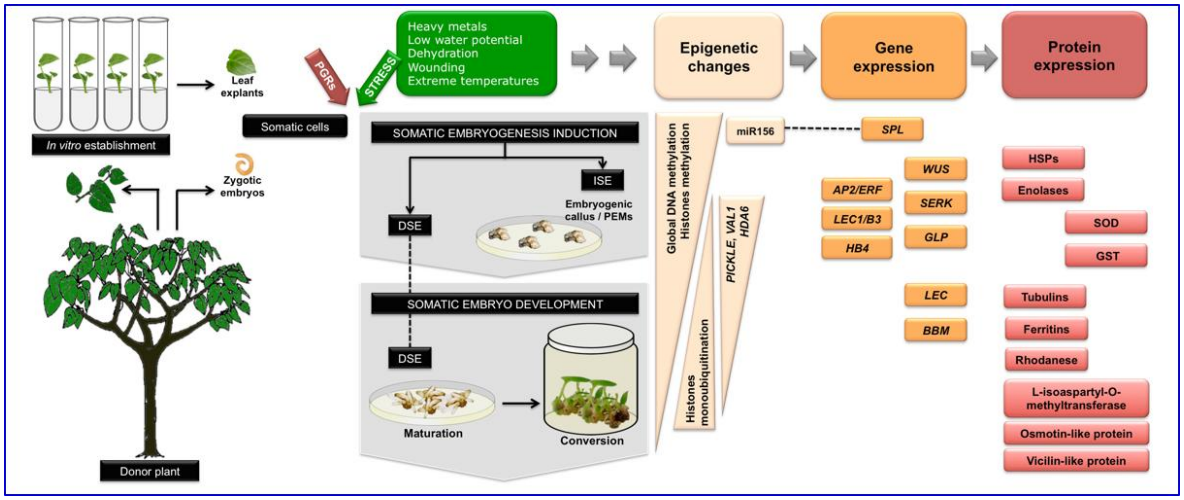

Figure 1 -Representation of embryogenic competence acquisition (Adapted from Correia *et al*., 2016)

Since the early stages of SE are characterized by the activation of genes related to stress, some authors consider the hypothesis that this phenomenon is an extreme response to adverse conditions by the cells in culture (Zavattieri *et al*., 2010). Among the stress -related genes identified as expressed during the induction of SE, it is given relevance to those responsible for the production of glutathione-s-transferase (*GST* genes): products of these genes were detected in abundance during induction of SE, as well as in somatic embryos, and their expression can be induced by exogenous factors like pathogen attack, tissue injury and presence of auxins (Galland *et al*., 2007). These genes undergo regulation during SE induced by auxins and act as regulators of genes expressed in response to these growth regulators (Thibaud-Niessen *et al.*, 2003). The *MtSK1* gene was also identified as being overexpressed during the induction of embryogenesis in *Medicago truncatula*: in this case, the stress factors were wounds applied to the explants and it was, therefore, concluded that the expression of this gene

is independent of the levels of hormones present in embryogenic cultures (Rosa and Nolan, 2006).

More recently, developments in proteomic analysis enabled the identification and characterization of proteins to become a relevant area in the study of plants: there has been a growing interest in the role of proteins involved in the induction of SE, as these allow for a better understanding of the changes at the cellular level during this phenomenon (Takac *et al*., 2011). Most research has been based on comparative studies between embryogenic and non-embryogenic *calli*, focusing also on the identification of protein markers for the different stages of development of the embryos (Correia *et al*., 2016). Analyzing the protein profiles, there are differences in metabolism and stress-related proteins during this process (Zhang *et al.*, 2009).

Despite all this knowledge, SE process is far from being completely understood, as many aspects of molecular cell reprogramming and totipotent acquisition by the cell are not yet elucidated, remaining one of the least understood areas of this phenomenon (Yang and Zhang, 2010). Therefore it is critical to use appropriate model systems that can contribute to a better understanding of the molecular network involved in this process. Tamarillo is a promising model for such kind of approaches, since the SE induction process for this plant has been widely studied and optimized (Correia and Canhoto, 2012).

1.3.Tamarillo

Tamarillo is a small tree from South America, particularly from the region of the Andes (Peru, Chile and Ecuador), having been introduced later in Central and Western India, southern Europe, Australia, New Zealand and the Portuguese Islands (Prohens

and Nuez, 2001). The main producers of tamarillo are the United States of America, Ecuador, Colombia, Australia and New Zealand, with an annual production of approximately 2000 tons (Prohens and Nuez, 2001; Correia and Canhoto, 2012).

The commercial interest of this plant is due to its edible fruits that can be eaten raw, incorporated into recipes or used in the confection of jellies (due to their high content in pectin), drinks and other processed foods (Bohs, 1989; Duke and DuCellier, 1993). Recent studies (Hurtado *et al.*, 2009; Osorio *et al.*, 2012; Birth *et al.*, 2013; Atiqah, 2014) indicate that these fruits have a high nutritional value, showing significant concentrations of vitamins, minerals and bioactive components such as anthocyanins, carotenoids and flavonoids (Mertz *et al.*, 2009, 2010). In addition, compounds present in these fruits have high antioxidant activity, and therefore have important preventive and therapeutic properties of neurodegenerative diseases (Hassan and Bakar, 2013). Thus, it is considered that the fruits of tamarillo are an important resource whose beneficial properties on human health should be explored (Osorio *et al.*, 2012).

Figure 2 - *Solanum betaceum*. (A) Tamarillo tree from the Botanical Garden of the University of Coimbra. (B) Flowers. (C) Fruit. (Correia and Canhoto, 2012)

Tamarillo is a fast growing tree, reaching dimensions between 1 and 5 meters. The leaves are evergreen and with a heart-shaped structure at the base, reaching lengths between 10 to 30 cm, presenting a musky odor. The flowers, usually grouped, and developing from axyllary meristems, blossoming in the spring and remaining in bloom until summer (Morton, 2002; Prohens and Nuez, 2001). The fruits (which may appear isolated or in associations of 3 to 12 units) are oval shaped, with dimensions between 5 to 10 cm long and 3 to 5 cm wide, reaching their maturity between October and April (Correia, 2011; Correia and Canhoto, 2012). The epicarp can present different colors ranging from yellow to dark purple, while the color of the pulp varies between red and yellow (Bohs, 1989; Morton, 2002). It is an autogamic plant, usually requiring factors such as wind and insects for pollination to occur (Prohens and Nuez, 2001).

Tamarillo propagation can be achieved through seeds, cuttings or grafting. When propagated via seeds, germination is usually effective, although when originating from plantations with a large number of cultivars cross-pollination can occur (contributing to a segregation of features) making the method unuseful if the aim is the propagation of selected genotypes (Pringle and Murray, 1991; Prohens and Nuez, 2001; Correia and Canhoto, 2012). This fact, coupled with the likelihood of impairment of the plant's health, makes the conventional propagation methods less adequate for improvement of tamarillo cultivars. Thus, biotechnological methodologies such as *in vitro* cloning and genetic transformation emerge as important alternatives to traditional methods of propagation (Correia and Canhoto, 2012).

There are many methodologies applied to obtain tamarillo plants *in vitro* (Correia and Canhoto, 2012). The first methodology described was micropropagation from axillary shoots (Cohen and Elliot, 1979), followed by regeneration by organogenesis from leaf explants (Obando *et al*., 1992) and from protoplasts

(Guimarães *et al*., 1996). Somatic embryogenesis (SE) is a methodology widely described for this plant (Correia, 2011; Correia and Canhoto, 2012; Correia *et al*., 2012). The pioneering studies related to SE were performed in mature zygotic embryos, hypocotyls, cotyledons, as well as protoplasts (Guimarães *et al*., 1988; 1996).

1.3.1. Tamarillo somatic embryogenesis

The first reports of somatic embryogenesis in tamarillo were described by Guimarães *et al*. (1988). In this protocol, the embryogenesis was induced by the synthetic auxin 1-naphthaleneacetic acid (NAA), occurring zygotic embryo differentiation in somatic embryos, in a process that comprises only one phase. In another protocol (Lopes *et al.*, 2000), the auxin used it is 2,4-D or picloram, if the explants used correspond to zygotic embryos or foliar explants, respectively. These explants produce embryogenic *calli* that can be kept in sub-culture in a medium supplemented with auxins, and somatic embryos develop when the *calli* are transferred to a medium containing gibberellic acid. This process involves two phases, and its main advantage is the fact that the embryogenic *calli* can be maintained in culture for extended periods. These protocols are summarized in figure 3.

However, the time that elapses until the embryos develop is long and *calli* may lose their embryogenic capacity over time (Canhoto, 2010). It is important to clarify that in any of the situations it was possible to verify that the rate of embryogenesis was more significant when adding high levels of sucrose to the culture medium (Canhoto *et al*., 2005).

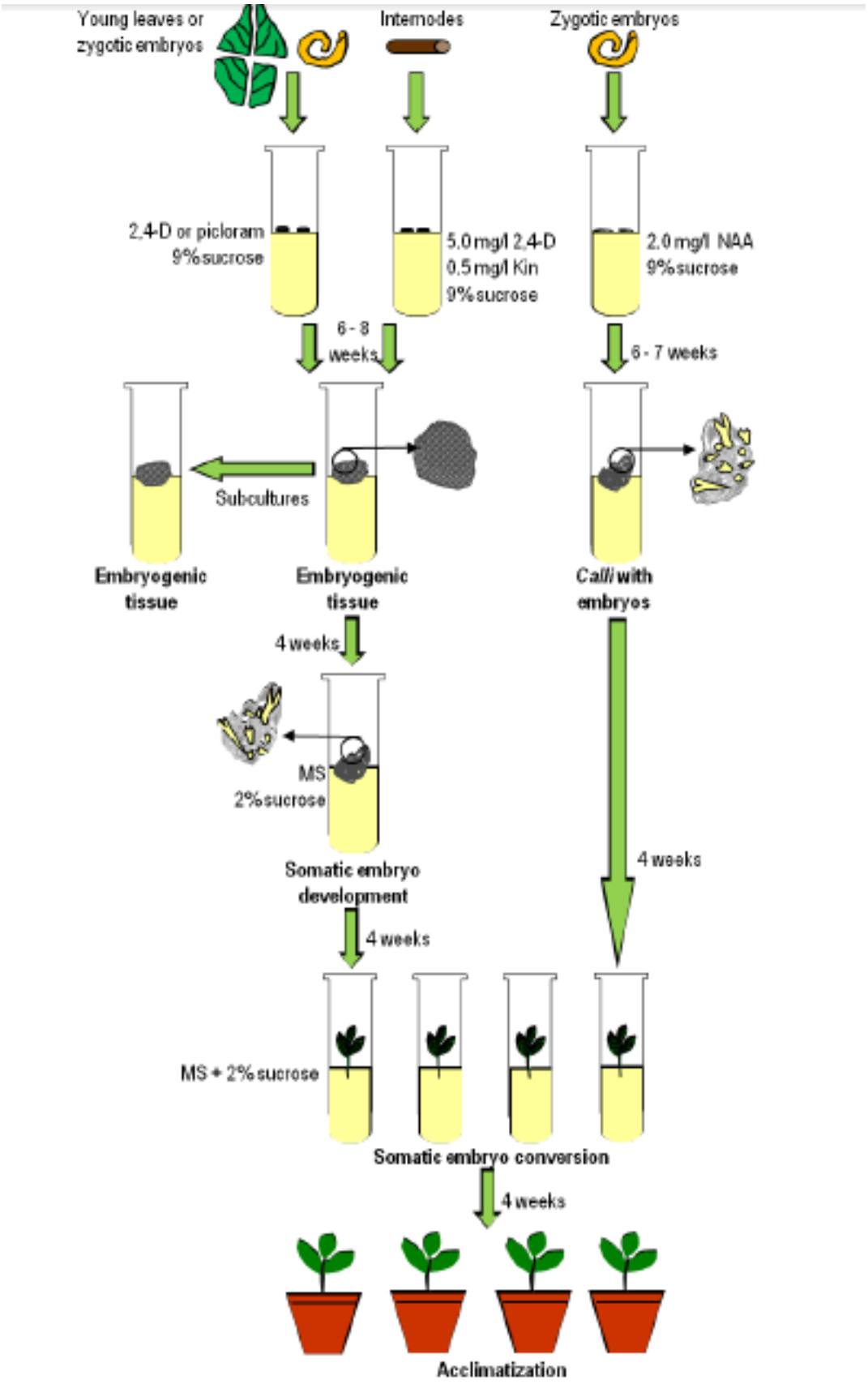

Figure 3 – Schematic representation of SE induction and plant regeneration in tamarillo. Adapted from Canhoto *et al*. (2005)

1.4.*Non-Embryogenic Protein of Tamarillo Callus* (NEP-TC)

As already mentioned, somatic embryogenesis is an important tool for large scale propagation of plants, constituting an important model system for the study of embryogenic development (Quiroz-Figueroa, 2006). Recently, molecular approaches have been used for detection of embryogenic markers, and several genes related to cell differentiation, morphogenesis and drought tolerance have been identified (Ikeda *et al.*, 2006; Yang and Zhang, 2010), as well as those related to the induction of somatic embryogenesis and acquisition of embryogenic competence, like *LEC* (*LEAFY COTYLEDON*) (Gaj *et al*., 2005) and *SERK* (*SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE*) (Aker *et al.*, 2006). Ikeda and co-workers (2006), isolated and analyzed mutants with embryogenic deficiencies (derived from the model plant *Arabidopsis thaliana*). However, most of the genes that cause this mutation were related to maintenance-related phenomena, such as division and cellular differentiation, response to plant hormones and other processes, and few specific genes for embryogenesis have been identified.

In the particular case of tamarillo, comparative studies conducted allowed the analysis of molecular patterns present in embryogenic (EC) and non-embryogenic *calli* (NEC), aiming to identify protein markers associated with the two types of tissue (Ferreira *et al.*, 1998; Faro *et al*., 2003). Studies with liquid chromatography coupled to mass spectrometry (LC-MS), followed by computer analysis in data bases (Correia *et al*., 2012), revealed that proteins produced by explants are grouped into three distinct categories: metabolism, cellular defense/virulence and protein synthesis, as seen in figure 4.

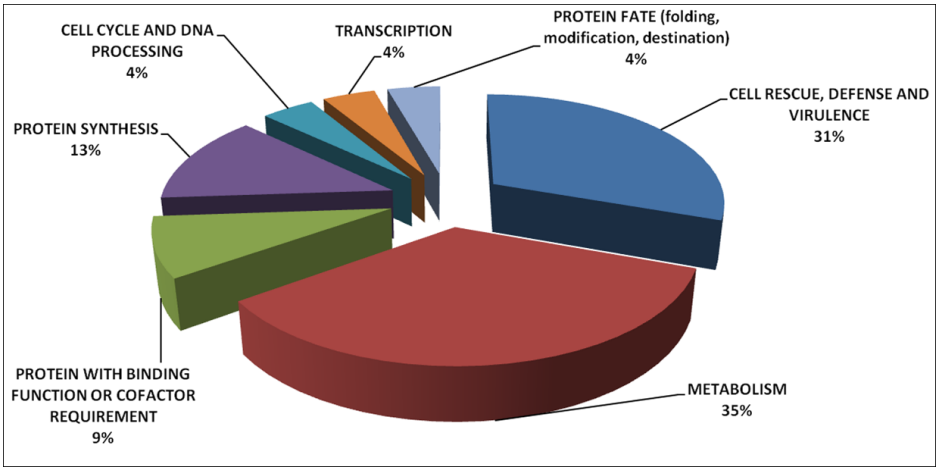

Figure 4- Distribution of proteins identified during the induction of SE in tamarillo (Correia, 2011)

Given that there is no information regarding tamarillo's genome, proteins expressed during its SE have been identified based on databases available to other species, including *Arabidopsis thaliana*, *Solanum lycopersicum* (tomato) and potato *Solanum tuberosum* (potato). However, this method shows a disadvantage, not allowing the identification of proteins specific to this species, due to the absence of similar genes in the databases (Correia *et al*., 2012). Previous to these studies, the protein NEP-TC (26.5 kDa; AFI57511) was identified, and on which interest befell, because it is expressed consistently in non-embryogenic *calli* obtained from different explants, such as zygotic embryos and hypocotyls, suggesting that it can be a good molecular marker for the absence of somatic embryogenesis competence in tamarillo (Ferreira *et al.*, 1988). The sequence of this protein is shown in figure 5.

```
        10         20         30         40         50         60
MENEKKLESF VLVHNIAKRH NVGTLARSAT AFGVSEMILV GRRDFNAFGS HGSTSHVRFR

        70         80         90        100        110        120
HFHSLADAKT FLKERDCDIC GVEITENAVA INEHPFKRST AFLLGNEGTG LSTKECEICD

       130        140        150        160        170        180
FFVYIPQYGC GTASLNVTVA ASIVLHQFGV WAGFSERTRE GNKFIVAERP FKQAKKNYCM

       190        200        210        220
ETSESVAEER RLKKENLSNG FFEDTGKEES PSNLLDTLFD D
```

Figure 5- NEP-TC's sequence, as inferred from cDNA (Correia, 2011)

The comparison of the protein's sequence with those available in databases revealed that it presents homology with proteins of the SpoU methylase family, the second-largest group of RNA methyltransferases: the S-adenosil-L-methionine (SAM) dependent methyltransferases (Correia, 2011).

1.5. RNA methyltransferases

It is the structure of the RNA molecule that allows its inherent ability to represent various functions in biological systems, this competence being increased by the occurrence of directed post-transcriptional modifications catalyzed by specific enzymes (Machnicka *et al.*, 2013). These changes occur in the three main RNA fractions (tRNA, rRNA and mRNA), as well as in snRNA, constituting one of the most conserved properties of the RNA molecule throughout the ages (Sing *et al.*, 2010).

Approximately one-third of the more than 100 modifications in the RNA molecule correspond to the addition of methyl groups (Motorin and Helm, 2011), however the available data on this phenomenon are not explicit as to its purpose, or the molecular basis of its function, especially because the detection and quantification of

methylation of nucleotides can be technically demanding, as well as the fact that RNA methylation may present varied purposes (Kelner *et al*., 2010).

The transfer reactions from the methyl group donor to the RNA nucleotides are catalyzed by a large number of RNA methyltransferases, divided into four superfamilies (Czerwoniec *et al.*, 2009). The largest number of methyltransferases known to date corresponds to the Rossmann-fold superfamily (RFM), while the other superfamilies are called SPOUT, radical S-adenosil-L-methionine (SAM) and FAD/NAD binding proteins. All methyltransferases, except rare exceptions, use SAM as a donor of methyl groups (Motorin and Helm, 2011).

Since NEP-TC putatively belongs to the family of SAM-dependent methyltransferases (Correia, 2011), focus will be given upon their action mechanism.

SAM (also called AdoMet) is an important metabolic intermediate in all forms of life, and one of its most studied functions is its role in the transfer of methyl groups to different substrates, such as RNA and ions, whose methylation is an important signal of the regulation of maturation and interaction with other macromolecules (Kozbial and Mushegian, 2005). All parts of the SAM molecule are used for a series of biochemical reactions, having been identified about 15 families of proteins that bind to this molecule, with additional functions that include methylation of phospholipids and small molecules, and also the synthesis of polyamines and formation of radicals (Loenen, 2006). It is estimated that approximately 95% of this intermediary is used for methylation, with the remaining 3-5% used for production of decarboxylated SAM (dcSAM) (Zaheer and Clarkson, 2004)

This molecule is a conjugate of methionine and ATP's adenosine group, formed in a reaction catalyzed by the enzyme methionine adenosil transferase (MAT), also known as SAM synthetase (Loenen, 2006). The transfer of methyl groups promoted by

SAM to other macromolecules produces a molecule of S-adenosil-homocysteine (SAH) as a by-product. SAH is the substrate of the enzyme methionine synthase, which uses a derivative of folate (MTHF) as methyl groups donor for regeneration of methionine (Struck *et al*., 2012). The SAH molecule was identified as a competitive inhibitor of transmethylation reactions, and should be removed in order for these reactions to proceed. The only way known to catabolize SAH in eukaryotes is mediated by SAH hydrolase (SAHHA), which catalyzes a reversible reaction whose equilibrium depends on the synthesis of SAH: when this molecule suffers hydrolysis, the resulting molecules of adenosine and homocysteine must be continuously removed (Moffatt and Weretilnyk, 2001). This mechanism is summarized in figure 6.

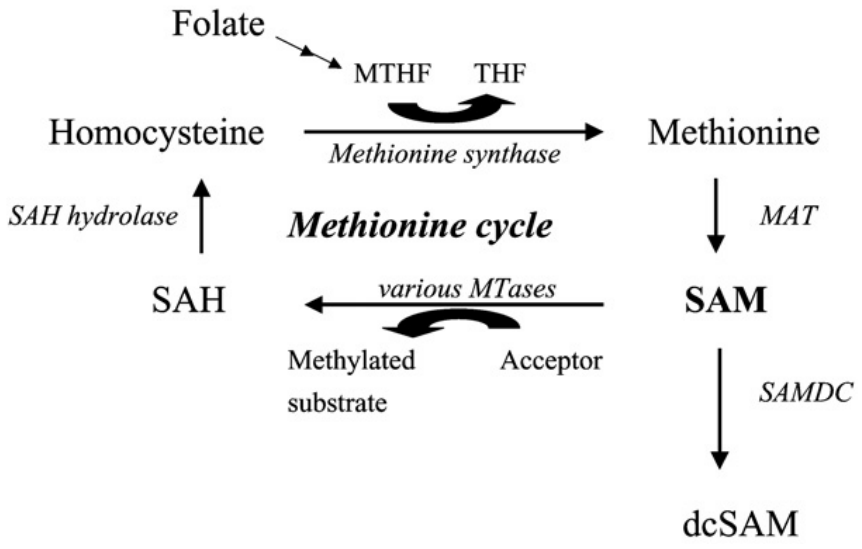

Figure 6- Methionine cycle. (Loenen, 2006)

Transmethylation reactions are associated with all aspects of plant metabolism and development, each modification catalyzed by a specific methyltransferase (Moffatt and Weretilnyk, 2001). Transmethylation activity of these enzymes has been detected in

fractions of the Golgi apparatus of several plants (tomato, tobacco and flaxseed), although its place of production corresponds to the cytosol (Liu *et al*., 2015).

One of the most important roles of these enzymes is assigned to spermidine synthase, an enzyme involved in the polyamines pathway (putrescine, spermidine and spermine), small positively charged molecules that bind to DNA, RNA and proteins, affecting the DNA frameshift and modulating signal transduction (Wallace and Fraser, 2004). The decarboxylation of SAM produces dcSAM, and this molecule's methionine group is used as substrate for spermidine synthetase to convert putrescine in spermidine. A second molecule of dcSAM is used by spermine synthase for conversion of spermidine into spermine (Loenen, 2006).

Another phenomenon in which methyl substrates have a preponderant role corresponds to ethylene synthesis: a gaseous phytohormone consisting of two carbon atoms and four hydrogen atoms (C_2H_4), produced in virtually all plant tissues and cells (Bradford, 2008). This plant hormone affects a variety of processes during the plant life cycle, including seed germination, fruit ripening, senescence, abscission, gravitropism and response to stress (Wang *et al*., 2013). The biosynthetic pathway of ethylene only requires two steps: the conversion of SAM into ACC (1-aminocyclopropane-1-carboxylic acid), in which 5-methyladenosine is released, and the conversion of ACC into ethylene (Bradford, 2008). 5-methyladenosine is used to regenerate methionine, which in turn is used in the production of SAM, restarting the cycle of production of ethylene, or participating in another type of reactions as donor of methyl groups (Sauter *et al*., 2013).

Other processes include the methylation of plant sterols located in cell membranes, and the loss of this activity can be associated with defects in plant development (Diener *et al*., 2000). Phenolic compounds of cell walls (which include

flavonoids and anthocyanins) are produced by the action of specific methyltransferases: in the case of the flavonoids, the methylation influences its solubility, sub-cellular partitioning and anti-microbial activity (Moffatt and Weretilnyk, 2001). Secondary compounds (such as caffeine, nicotine, myo-inositol and biotin) can be methylated in response to stress, growth stimuli and other metabolic signals (Shoji *et al*., 2000). Other secondary compounds susceptible to transformation by SAM dependent methyltransferases include salicylates and dihydrojasmonates, that act as growth regulators. In the particular case of jasmonic acid (which is responsible for intracellular signaling in plants), the overexpression of the carboxylmethyltransferase reveals that the methylated form of this compound is also an important intracellular signal, in addition to its role as inter-plant signaling in response to herbivory and pathogens: the occurrence of methylation enhances their action (Seo *et al*., 2001).

However, there are no reports of enzymes related to the inhibition of the induction process of somatic embryogenesis, and molecular mechanisms from which the expression of the *NEP-TC* gene may interfere with this process are not known.

1.6. Methylation of RNAs

1.6.1. Transfer RNA

Transfer RNA (tRNA) corresponds to the type of RNA subject to a greater number of modifications, being the reference model for the study of these phenomena (Motorin and Helm, 2011). The modification of this molecule is one of the most efficient and versatile mechanisms known to date, and can be divided into two types, according to their predominant role: mRNA decoding or structural and metabolic

stabilization (Motorin and Helm, 2010). To date, more than 90 modified nucleosides were identified in the tRNA molecule, and most of the changes include methylation reactions (Machinicka *et al*., 2013). Most of the transmethylation reactions of tRNA molecules are held by SAM-dependent methyltransferases: reduction of SAM's content as it is consumed as a substrate can be translated into incomplete modifications in the tRNA molecule. The precursors of SAM are ATP and methionine, which may answer the question of why methionine is used as the initiation codon for protein synthesis (Hori, 2014). To date, several tRNA methyltransferases have been identified as members of the superfamily SPOUT (to which the SpoU methyltransferases group belongs, including proteins such as NEP-TC), based on crystallographic structure (Shao *et al*., 2013).

tRNA methyltransferases modify a particular nucleoside, in a specific position of the molecule: this raises the question of the specificity of these enzymes for its substrate. In general, the methyltransferases recognize the structure surrounding the target location, which is often the tRNA L structure: consequently, the recognition by the enzyme involves several steps of induced fit binding (Hori, 2014).

Nucleoside modifications of tRNA have been extensively studied in bacteria and yeast, and most of their mechanisms of action have been identified. However, the occurrence of these modifications in plants is rarely documented (Chen *et al*., 2010). In bacteria and yeast, it is believed that modifications of the tRNA molecule act as biological sensors, since these phenomena occur according to changes in growth conditions (Johanson and Byström, 2005). tRNA modification may vary between the methylation of just a nucleoside or a pathway involving multiple protein complexes: for example, in *Saccharomyces cerevisiae* at least 25 proteins are involved in modifying only a nucleoside (mcm5s2U) (Huang *et al.*, 2008). In *Escherischia coli*, modification

of nucleosides is not essential to the viability of the cells, however the absence of certain methyltransferases in the bacteria's metabolism can lead to cell death: in the particular case of the nucleoside m5U54, the absence of its modification will only promote cell growth deficiencies. However, mutants that do not produce the necessary methyltransferase that promotes modification of this nucleoside present no cell viability (Urbonavicius *et al.*, 2007).

Growing conditions can affect changes of the tRNA molecule both qualitative and quantitatively, so the modification of tRNA nucleosides has been considered a regulatory mechanism, on which this molecule acts as a biological sensor (Chen *et al*., 2010). In amino acid deprivation conditions, the bacterial tRNA binds to the ribosome, signaling the relA gene expression, responsible for production of guanosine pentaphosphate (ppGpp). This molecule will bind to the RNA polymerase active site, ceasing the transcription of tRNA and rRNA and favoring amino acid biosynthetic pathways (Traxler *et al*, 2008). Consequently, the production and accumulation of this molecule is accompanied by global changes in gene expression and cell physiology, that are believed to improve survival in extreme conditions: this phenomenon is called *stringent response*, having already been observed in different bacteria (Haiser *et al.*, 2008). For example, strains of *E. coli* grown in media with reduced content of amino acids and iron suffer a reduction in translation efficiency and increased occurrence of errors during this process, affecting cellular metabolism (Sorensen, 2001). The same applies to the bacteria *Streptomyces coelicolor*, that in conditions of amino acid deprivation undergoes morphological changes (cleavage and degradation) induced by changes in the tRNA molecule (Haiser *et al.*, 2008).

In the particular case of plants, studies in tobacco showed that the abundance and variety of methylated nucleosides is higher in intact plants than in plants that

exhibit tumors (Jones and Scott, 1981). These modifications also differ according to the level of maturity of the plant and/or transport of the nucleosides to chloroplasts and mitochondria (Knoop, 2011). It is well known that plants undergo major environmental changes during their life cycle, which raises the following question: the changes in the tRNA molecule occur as a consequence of the environmental stimuli or are related to the stages of plant development? (Chen *et al.*, 2010). Recent advances in the methodology for the analysis of plant genomes allow a better understanding of the mechanisms that regulate RNA methylation phenomena, however the molecular siginificance of these mechanisms remains unknown (Saze *et al*., 2012).

1.6.2. Messenger RNA

The methylation of messenger RNA (mRNA) is the type of methylation better understood in eukaryotes, since the reaction leads to the development of the *cap* structure, and the importance of this process for the initiation of translation has been extensively studied (Motorin and Helm, 2010). The *cap* is a guanine nucleotide that connects with the mRNA via a 5'-5' triphosphate binding. Guanosine is methylated at position 7 after the *cap*'s addition, by action of a methyltransferase (Cowling, 2010).

The process of the *cap*'s addition starts at one unchanged end of the mRNA molecule, ending in a triphosphate group (formed by a nucleotide associated with three phosphate groups, which bind to the 5' carbon): this process starts before the transcription's end, as the pre-mRNA is synthesized (Shatkin and Manley, 2000). Firstly, one of the terminal phosphate groups is removed by action of the enzyme RNA triophosphatase, and a GTP molecule is added to the newly formed biphosphate, by action of the GTP guanylyltransferase. In this process there is loss of a pyrophosphate group from GTP and a 5'-5' triphosphate binding is promoted (Kapp and Lorsch, 2004).

Then, the nitrogen present in guanine's position 7 suffers methylation from the mRNA guanine N-7 methyltransferase (which uses SAM as substrate), creating the *cap* structure. This structure can undergo additional modifications, some of them promoted by other methyltransferases (Fetcher and Brownlee, 2005).

The 7-methylguanosine *cap* plays an important role in the mRNA translation, and it is also involved in this molecule's stabilization against attack by exonucleases, promoting transcription, splicing, polyadenylation and nuclear export (Cowling, 2010).

1.6.3. Ribosomal RNA

Ribosomal RNA (rRNA) of every living organism contains a high number of methylated nucleotides. The conservation of these methylated residues is not universal, but most of the places where methylation occurs are present in the same region of the rRNA: this way, an equivalency can be established between species (Motorin and Helm, 2011). The enzymatic mechanism in which the methylated residue of rRNA is created can occur in two different pathways. In the first case, a single enzyme has specificity to recognize the substrate for subsequent methylation: this strategy is often used for base methylation processes in all organisms and to 2'-O type methylation in bacteria. A second mechanism is based on the use of a sub-unit of a methyltransferase that is associated with a specific RNA with ability to promote base pairing with RNA substrate: the specificity of RNA substrate recognition is the responsibility of the guide-RNA and the enzymatic catalysis is due to a methyltransferase. This strategy is used in eukaryotic organisms and archaea to form 2'-O residues on rRNA (Decatur and Fournier, 2003).

The methylation of rRNA is an important process as it allows a stabilization of its loops (consequently allowing a correct folding of the RNA molecule), and it also has an important role in the translation of mRNA: rRNA methylated residues connect with the tRNA, stabilizing the codon-anticodon specificity interaction required for the occurrence of translation (Blanchard and Puglisi, 2001; Baillieu *et al.*, 2009).

Although rRNA is a quite stable molecule, it is known that under certain stressful physiological conditions it can be extensively degraded (Deutscher, 2003). In *S. cerevisiae*, methylation of rRNA in random locations can severely compromise cell growth: if the production of peptidyl transferases (involved in the formation of peptide bonds during the translation) is affected, the accumulation of ribosomes, as well as their activity, are compromised: poorly directed methylation can reduce between 70 to 100% the rate of translation by decreased production of ribosomal subunits, caused by a slow processing of pre-rRNA and a quick *turn over* of rRNA, as well as impairment of the activity of ribosomes that do not undergo degradation (Liu *et al.*, 2008).

In plants, studies carried out in tobacco allowed the detection of an increase in transcription of rRNA during the formation of *calli* (induced by auxins and cytokinins), with increased levels of rRNA hypermethylation two weeks after induction, reversed when the plant undergoes regeneration: this hypermethylation does not occur randomly, affecting specific families of genes. These data suggest that pluripotency and cellular proliferation occurring during the induction of SE are associated with an improved expression of rRNA and its hypermethylation (Koukalova *et al.*, 2005). Similarly, studies by Ohbayashi *et al.* (2011) in *A. thaliana* revealed an accumulation of rRNA precursors during induction of SE in hypocotyls, which allowed to infer that there is an active production of rRNA during this process. In agreement with these results, mutants for *ROOT INITIATION DEFECTIVE2* gene (*RID2*), responsible for the expression of a

methyltransferase, do not produce *calli* and present an abnormal accumulation of various intermediates of pre-rRNA. This gene is expressed in meristematic tissues after incubation in induction media, indicating that their activity is linked to the proliferative capacity of cells (Ohbayashi *et al*., 2011). After regenerating *A. thaliana* plants from *calli*, Shinohara *et al*. (2014) determined that pre-meristematic aggregates present a strong expression of *CUP-SHAPED COTYLEDON* genes (*CUC*), which regulate the establishment and maintenance of the apical meristem of the stem. *ROOT INITIATION DEFECTIVE3* gene (*RID3*) regulates negatively the expression of the gene *CUC1* and the restriction of cell division in pre-meristematic cells. *In silico* analysis allowed to verify that the *RID3* is orthologous to *IPI3* gene, responsible for mediation of the rRNA processing in *S. cereviseae*. In mutants (*rid3*) for this gene, rRNA precursors accumulate in the cells at high levels: by comparison, it is possible to verify that *RID3* is as relevant for rRNA processing as *IPI3*. The mutants *rid3* and *rid2*, the latter already studied by Ohbayashi *et al*. (2011), were compared. It is known that both do not produce an RNA methyltransferase, and both show deficiencies in their development, like the *pointed leaf phenotype* (common in mutants that suffered ribosomal alterations), having been also verified a disturbance in the proliferation of pre-meristematic cell masses and in the expression of genes CUC. This indicates that the changes in rRNA (with emphasis on the role of methyltransferases) can be involved in gene expression and regulation of the development of pre-meristems (Shinohara *et al.*, 2014).

In *A. thaliana*, the study of root hair patterns is an important system for the study of the molecular basis of cell fate (Schiefelbein *et al.*, 2009). Studies by Wieckowski and collaborators (2012) using mutants of this plant for the *DIM1A* gene (responsible for the expression of a rRNA dimethylase involved in rRNA post-transcriptional processing), expressed in fast-growing regions, showed that gene expression during the

development of the root epidermis is dependent on a good rendering of the rRNA: the mutant plants don't perform the post-transcription processing of 18S rRNA, displaying abnormal development of root meristem and leaves. These results show the importance of the regulation of rRNA modifications and the importance of methylation in plant growth and development (Wieckowski and Schiefelbein, 2012).

The SAM-dependent rRNA methyltransferases mediate the methylation of two adenosines in the 3' terminal of rRNA's ribosomal subunits of bacteria, archaea and eukaryotes (Richter *et al.*, 2010). However, there is limited information concerning the methylation of plant mitochondrial rRNA (Wieckowski and Schiefelbein, 2012). Studies conducted by Richter *et al*. (2010) in *A. thaliana* aimed to verify if the plant's mitochondrial rRNA presents methylation residues, and how methyltransferases act as transcription factor. The identification of the rRNA methyltransferase Dim1B has shown that it is involved in the methylation of the two adenines in the 18S mitochondrial rRNA. In addition, the comparison of the orthologous genes related to Dim1B in other Angiosperms suggests that this methylation is a common phenomenon in higher plants, though its functional importance in this process is still unknown (Richter *et al.*, 2010).

1.7. Expression and purification of recombinant proteins

Due to its easy handling, plasmids are the preferred vectors for cloning genes or specific fragments of DNA. Plasmids used in cloning derive from naturally occurring ones that were manipulated by genetic engineering in order to incorporate desirable features (Videira, 2001). The elements that compose expression plasmids include origin of replication (*ori*), an antibiotic resistance marker, transcription promoters, translation

initiation regions (TIR), as well as terminator sequences of transcription and translation (Sorensen and Mortensen, 2005).

The promoter is a central element that affects the strength and duration of the transcription process and, consequently, the protein yield (Sorensen and Mortensen, 2005). T7 (from phage T7) is one of the most important promoters, being an element of the pET system (Novagen), the most widely used system for expression of recombinant proteins in *E. coli* (Correa and Oppezzo, 2011). First described by Studier and Moffat (1986), this system is based on T7 RNA polymerase for induction of recombinant protein production. This polymerase only transcribes genes under the control of the promoter T7, reaching a transcription rate eight times higher than that obtained by polymerases of the bacterium, which translates into a high protein yield (Lost *et al.*, 1992). The pET expression mechanism is shown in figure 7.

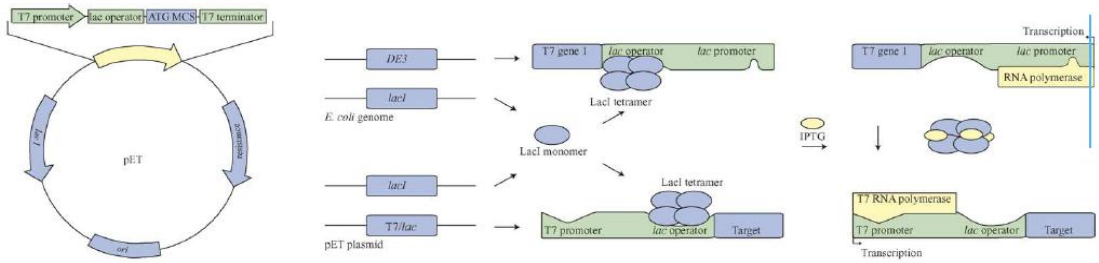

Figure 7- pET expression system mechanism (Sorensen and Mortensen, 2005). Expression with the pET system requires a host strain that allows the production of RNA polymerase T7 under the control of the *lacUV5* promoter (derived from the *lac* operon), induced in the presence of IPTG. The *lacUV5* promoter is controlled only by the *lac* operon repressor LacI, that allows the induction with IPTG, even in the presence of glucose. The addition of IPTG to the medium ends the repression caused by LacI, resulting in T7 polymerase expression and consequent gene transcription and recombinant protein production.

There are several hosts used for recombinant protein expression, however the main organism selected for this process is *Escherischia coli*, due its easy genetic manipulation *in vitro*, rapid growth in inexpensive substrates, easy maintenance in culture, well known genetic traits and high availability of compatible cloning vectors (Gopar and Kumal, 2013).

Different strains of *E. coli* have been used as hosts for cloning and expression vectors for specific sequences. The strains used currently derive from clonal lines established 20 years ago, designed to receive, maintain and express genes of non-native DNA (Chart *et al.*, 2010). Strains used for protein expression must present deficiencies in the production of proteases and contain the necessary genetic elements of expression: the strain BL-21 is the most widely used for this purpose (Sorensen and Mortensen, 2005). It is a strain with the ability to grow rapidly in minimal medium, non-pathogenic and insensitive to variations in glucose content (thus reducing the production of acetate) and the metabolic stress caused by the production of a large amount of recombinant protein (Phue *et al.*, 2008). Additionally, it has a copy of the gene for T7 polymerase, allowing easy integration of the pET system (Sorensen and Mortensen, 2005).

Due to its physical-chemical properties, the proteins do not allow high yield screenings. Consequently, affinity tags have become an indispensable tool in proteomics (Waugh, 2005). The use of affinity tags on recombinant protein expression allows for easy modification of the protein of interest, enabling its easy identification, production and isolation (Young *et al.*, 2012). Among the various available tags, the hexahistidine tag (His-tag) is the preferred for protein purification of high yield (Waugh, 2005).

Recombinant proteins with His-tag, but also phosphorylated proteins and antibodies, can be purified by Immobilized-metal affinity chromatography (IMAC), a

separation technique based on protein's differential affinity with metal ions (Porekar and Menarth, 2001; Cheung *et al*., 2012). Associated with capillary electrophoresis, IMAC allows the simultaneous achievement of end-product analysis and efficient purification (Cheung *et al*., 2012).

In IMAC, the adsorption of proteins is based on the interaction between a metal ion and the electron-donor immobilized on the surface of the protein (Cheung *et al*., 2012). This chromatography uses a matrix to which covalent metal-chelating groups are linked: the most used correspond to nitrilotriacetic acid (NTA) and iminodiacetic acid (IDA) (Block *et al.*, 2009). When metal ions are added to the matrix, there is the formation of a complex that interacts with the compound to purify: for this to be possible, the ions must have free coordination sites for connection to the solvent/solute (Ueda *et al*., 2003). Since this interaction has a reversible nature, it can be used for adsorption followed by disruption in non-denaturing conditions (Chaga, 2001). This mechanism is summarized in figure 8.

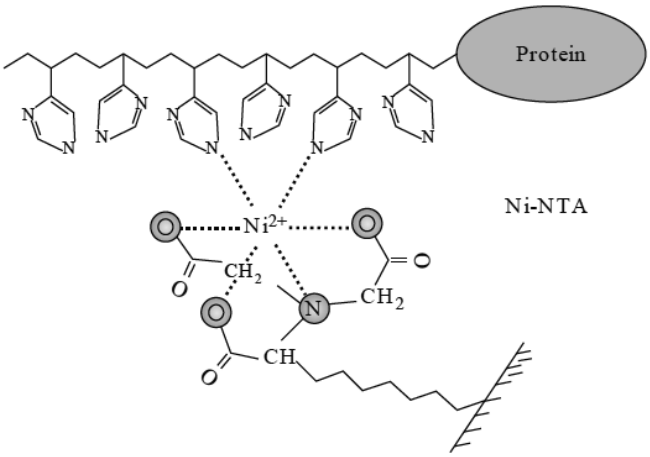

Figure 8- His-tag and IMAC ligands interaction. (Block *et al*., 2009)

The choice of the metallic ion immobilized on IMAC ligands depends on the technique's application: trivalent cations (Al^{3+}, Ga^{3+} and Fe^{3+}) or tetravalent (Zr^{4+}) are preferred for phosphoproteins and phosphopeptides, while divalent cations (Cu^{2+}, Ni^{2+} and Zn^{2+}) are used for the purification of His-tagged proteins (Block *et al.*, 2009). Elution of the desired protein is obtained by protonation, linking or by extracting the metal ion with a stronger chelating agent such as EDTA. The use of elution buffers with low pH is also a recurrent method, although it presents some limitations: some proteins are sensitive to low pH, so the change of ligand (using imidazole) in neutral pH is more favorable (Menarth and Porekar, 2012).

In the particular case of proteins with His-tag, the separation is allowed due to the interaction of the histidine residue with the chelator immobilized in the matrix (Arnau *et al*., 2006). Through successive associations/dissociations the molecule glides over the matrix surface, dissociating itself from the metal ion, as seen in figure 9 (Knecht *et al.*, 2009; You *et al.*, 2014).

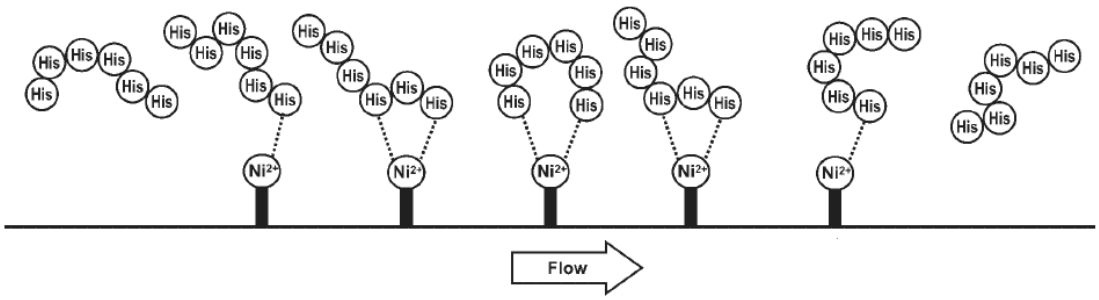

Figure 9 – Association/dissociation of the recombinant protein from the IMAC matrix. (Knecht *et al*., 2009).

The advantages of this methodology include the fact that it allows high ligand stability, high protein loading capacity, full recovery of the ligand after chromatography, it does not interfere with the structure of the protein, as well as its reduced cost and the fact that the matrix can be used several times without any loss of its capacity (Arnold, 1991; Porath, 1992; Ueda *et al.*, 2003).

1.8. Enzymatic activity analysis

1.8.1. Spectrophotometric assays

Spectrophotometric measurement is the most popular analytical tool in the field of analysis of a variety of compounds in simple as well as in complex mixtures, being a well-established fact that it is less expensive, follows a simple procedure, and provides a high accuracy and reproducibility from a small number of samples (Bhawani *et al*., 2015). The study of methyltransferase activity and its regulation through this method continues to be an important and promising area of research. Thus, there is currently a demand for the ability to screen large compound libraries to identify and optimize small molecules that modulate methyltransferase activity (Graves *et al*., 2013).

In the present study, the assay performed relies on the production of SAH in methylation reactions carried by SAM-dependent methyltransferases. As seen on figure 10, the removal of the methyl group from SAM generates SAH, which is rapidly converted to S-ribosylhomocysteine and adenine by adenosylhomocysteine nucleosidase. The resulting adenine is converted to hypoxantine, which is converted to urate and hydrogen peroxide. The rate of production of hydrogen peroxide is measured with a colorimetric assay by an increase in absorbance at 510 nm.

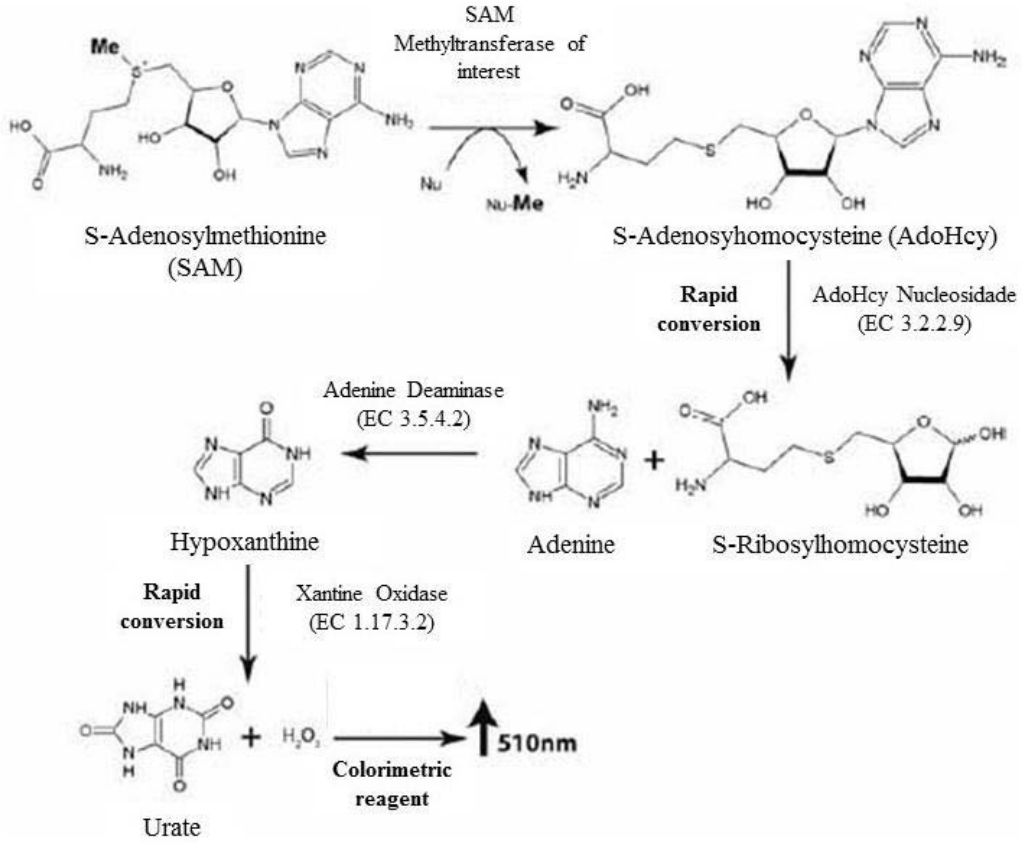

Figure 10 – Schematic representation of the reaction sequence and methyltransferase activity detection by Calbiochem's CBA096 SAM Methyltransferase Assay kit.

1.8.2. Thermophoresis

Thermophoresis is a phenomenon observed in mixtures of mobile particles that exhibit different responses in relation to a temperature gradient (Duhr and Braun, 2006). *Micro scale thermophoresis* (MST) is a technique that allows the quantification of interactions between biomolecules by thermophoretical detection of changes in conformation and dimensions of a molecule by induction of a binding event (Jerabek-Willemsen *et al*., 2011). The detailed protocol of this technique is summarized in figure 11.

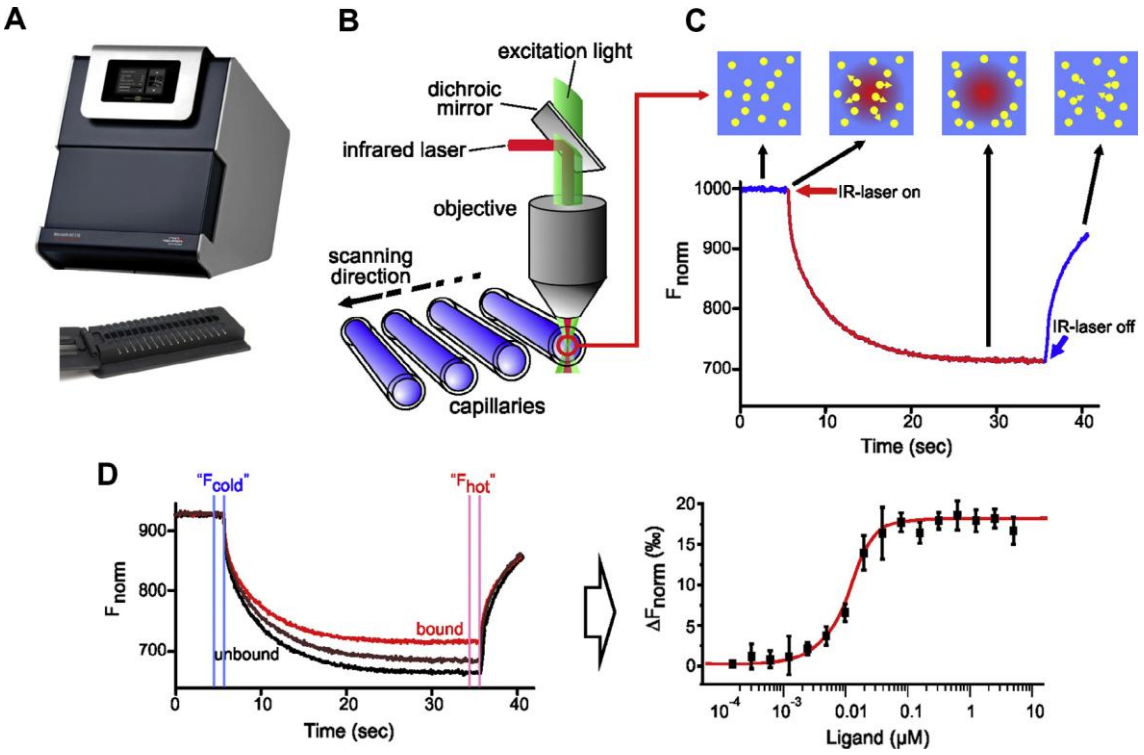

Figure 11 - Schematic representation of the Micro scale thermophoresis (MST) protocol (Jerabek-Williamsen *et al.*, 2014). The experimental procedure of MST is performed with the Monolith NT. 115 equipment (A), in which thermophoresis is induced and detected in glass capillaries containing a suspension with fluorescent molecules. An infrared (IR) laser with a 1480 nm emission wavelength is focused through a lens into the capillaries, producing a microscopic temperature gradient (B). When the temperature variation is induced, the fluorophores in the suspension are excited and the emitted fluorescence is sensed by the same objective that detects the IR laser. This mechanism allows a follow-up of the depletion or accumulation of thermophoresis-dependent fluorophores in temperature gradient induced by IR (C, D). The themophoretical movement is detected via fluorescence of an element linked to the protein: a fluorescent marker, a fusion protein or inherent fluorescence of the analyzed protein

For derivation of the binding constants by MST, multiple capillaries with constant concentrations of fluorescent molecules and increasing concentrations of ligands are analyzed sequentially and the thermophoresis is detected. Changes in thermophoresis of fluorescent molecules can also be used to calculate the equilibrium constants of binding. The advantages of this technique include its easy application, low sample consumption, as well as the fact that it allows the quantification of the interactions without factors like the weight and dimensions of the biomolecules being a limitation: this gives the technique a wider range of applicability, as well as flexibility in experimental planning, enhanced by the fact that any type of buffer and complex bioliquid (like plasma and cell lysates) may be used (Seidel *et al*., 2013). Additionally, this approach can be used for quantification of enzyme activity and changes in proteins and nucleic acids (Wienken *et al*., 2011).

1.9. Objectives

SE has been extensively described for tamarillo, and there are several protocols established and optimized for the induction of this phenomenon, as well as regeneration of adult plants from embryogenic *calli*.

As already described, SE is a complex biological process influenced by many factors, from which stress is one of the most important ones. Even though several protocols have been optimized and established for the induction of this process, the mechanisms behind the explants' response to environmental conditions and further reprogramming of the cell and embryogenic competence acquisition are still widely unknown.

NEP-TC is a protein that was identified by bioinformatics studies as a putative RNA methyltransferase involved in the process of SE induction, being consistently

present in non-embryogenic *calli*. Since the role of this protein has not been fully unravelled and recent studies specify that methyltransferases can have a significant role in the regulation of cellular stress throughout SE induction, studies are required to perform a biochemical and physiological characterization of this protein.

Thus, the aim of this work is to characterize this protein through different approaches: i) bioinformatic studies to determine its phylogeny and structure; ii) immunohystochemical analysis to determine its expression on EC and NEC, as well as proembryogenic masses; iii) enzyme activity assays using recombinant NEP-TC in order to determine the correct substrate for this protein, as well as perform a kinetic characterization of this protein.

The results of these parameters can jointly contribute for a deeper biochemical and physiological characterization of NEP-TC, thus allowing a deeper understanding of the process of somatic embryogenesis.

2. Material and Methods

2.1. Bioinformatics analysis of NEP-TC

According to the information avaliable at UniProt (http://www.uniprot.org/), NEP-TC has a conserved domain between the positions 10 and 146 (Pfam PF00588) that identifies this protein as a SAM-dependent methyltransferase belonging to the SpoU rRNA methylase family, by a 76% homology with the orthologous sequences in the model plant *Arabidopsis thaliana* (Correia, 2011).

The 3D structure of the protein was obtained with the information available at UniProt, using the ModBase: Database of Comparative Protein Structure Models tool (https://modbase.compbio.ucsf.edu).

Phylogenetic analysis of NEP-TC was performed using the software SeaView Version 4 (Gouy *et al*., 2010). Sequences for other plants were obtained from Plaza dicots 3.0 (Proost *et al*., 2015) and aligned with the one corresponding to NEP-TC's, and sequences for pseudogenes were eliminated. The phylogenetic tree PhyML was obtained with the software's default settings and a 1000 bootstrap, which says how well the node of interest is supported in the model used to generate a phylogenetic tree.

2.2. NEP-TC expression analysis in tamarillo somatic embryogenesis

2.2.1. Plant material and culture conditions

In this work one genotype (TV310) of a red tamarillo was used, being previously established from seed *in vitro* germination, followed by micropropagation, trough shoot tip culture, in Murashige and Skoog (1962) (MS) medium, containing 3% sucrose, 0.2 mg/L of benzylaminopurine (BAP) and 0.6% (w/v) agar. The shoots (1 cm long) were subcultured once a month using the same medium and kept in a growth chamber at 25

ºC, in a 16h light / 8h dark photoperiod. Leaf segments used in the induction of SE were excised from developed shoots after 4-6 weeks.

SE induction in leaf explants was achieved in MS medium with 9% (w/v) sucrose, 5 mg/L of picloram and 0.25% (w/v) of Phytagel (Sigma). The pH of all the media used was adjusted between 5.6 and 5.8 and the media were autoclaved at 121 ºC for 20 minutes. The induction phase was carried out in a growth chamber at 24 ºC ± 1 during approximately 12 weeks. Samples were collected for subsequent analysis at several time-points of the process (0, 4, 8 and 12 weeks).

The embryogenic and non-embryogenic masses resulting from the induction process were carefully isolated and subcultured *in vitro* in the same conditions.

2.2.2. Protein extracts from embryogenic and non-embryogenic *calli*

The protein extracts were obtained from *calli* according to the method proposed by Zhang *et al*. (2009). The plant material used for such corresponded to 1 g of non-embryogenic callus and 500 mg of embryogenic callus, which were macerated and homogenized in 4 ml of acetone solution with 10% (v/v) TCA and 0.2% (w/v) DTT. After incubation overnight at -20 ºC, the macerated tissues were subjected to centrifugation during 30 minutes, 13200 rpm at 4 ºC. The pellet was rinsed with an acetone solution with 0.2% (w/v) DTT, followed by a 30 minutes incubation at -20 °C. Thereafter, there was a new centrifugation, under the same conditions as the previous one. The resulting supernatant was discarded and the *pellet* was dried in a flux chamber during 1 hour, being resuspended twice in an IEF solution, by addition of 250 µL and 500 µL to the non-embryogenic and embryogenic *calli* pellet, respectively. In order to promote the protein's releasement, the samples were exposed to ultrasounds for 15

minutes, followed by a period of agitation on a rotating incubator, for 2 h. After this period, the supernatant was collected and stored at -20 °C for further analysis.

2.2.3. Anti NEP-TC antibody purification

An antiserum containing polyclonal antibodies raised against the peptide antigen IPQYGCGTASLN, corresponding to residues 123-134 of an exposed region of NEP-TC (Genosphere Biotechnologies, Paris, France) was purified by affinity chromatography. To this end, 1.5 ml of Affigel matrix-15 (Bio-Rad) were washed with 10 volumes of 0.1 M MOPS pH 7.5, followed by the addition of the NEP-TC peptide and incubation at 4 ± 1 ºC overnight on a roller. After incubation, and in order to block the matrix, it was incubated during 1 hour at 4 ± 1 °C with ethanolamine 1.0 M, pH 8, on a roller.

When the matrix was ready, the purification ensued. The matrix was washed to remove ethanolamine residues through successive washes with an excess of 10 mM Tris-HCl pH 7.5. Then, 5 ml of antiserum were added, and incubated at 4 ± 1 °C on a roller during 1 h. After the collection of the filtrate, the column was washed with 20 volumes of 10 mM Tris-HCl pH 7.5, followed by 10 volumes of 10 mM Tris-HCl pH 7.5, supplemented with 500 mM NaCl. The purpose of this last washing was the change of the column's ionization, so that the antibody's elution is facilitated. This was done using 10 volumes of glycine 100 mM pH 2.8, since glycine competes with the antibody. Then, the matrix was washed with an excess of 10 mM Tris-HCl pH 7.5 and the antibodies were eluted with 5 volumes of $MgCl_2$, 5.4 M pH 6.2, to guarantee that antibodies that remain attached to the matrix are removed. Thus, the matrix can be reused for new purifications.

In order to concentrate the antibodies, the elutions performed with glycine were dialyzed in 1 L of Tris-HCl, 50 mM, 2.0 M NaCl, pH 8.2. Dialysis was held overnight at 4 ± 1 °C. After dialysis, a second concentration was carried out by placing the dialysis tube on *aquacide* (Calbiochem) during 1 hour.

2.2.4. Protein separation and immunoblot analysis

Protein fractions were denatured and subjected to a SDS-PAGE electrophoresis according to the method described by Laemmli (1970). The samples were incubated with denaturant solution (125 mM Tris-HCl, pH 6.8; 100 mM glycine; 40% (v/v) glycerol; 4% (w/v) SDS; 200 mM of DDT 0.001% (w/v) bromophenol blue) during 15 minutes at 95 ºC and loaded on a 12.5% polyacrylamide gel (1.5 mM Tris-HCl, pH 8.8; 40% acrylamide-bisacrylamide (29:1); 20% SDS (v/v): 10% ammonium persulfate (AMPS); 3 µL of tetramethylethylenediamine (TEMED); miliQ water) , placed on a Mini-Protean Tetra System (Biorad) and subjected to a 120 V potential for 90 minutes at room temperature.

After electrophoresis, the gels were stained for 15 minutes using a 0.25% solution (w/v) of Coomassie Brilliant Blue R250 in 45% (v/v) methanol and 5% (v/v) acetic acid. The discoloration was achieved by successive washes in a bleaching solution containing 25% (v/v) methanol and 5% (v/v) acetic acid. After revelation the gels were preserved in water.

Non-stained gels were transferred to a Immobilon PVDF membrane (Millipore) (activated previously in pure methanol) by placing it in contact with the gel in a tank Mini-Protean Tetra (Biorad) containing transfer buffer (1.44% (w/v) glycine, 0.3%

(w/v) Tris, 20% (v/v) methanol, 0.2% (v/v) SDS 20%, miliQ water), having been applied a potential of 100 V for 90 minutes at room temperature.

At the end of the transference, the membrane was placed in a blot holder (Millipore) and subjected to several washing steps in a SNAP i.d. ™ equipment (Millipore) with a TBS-T solution, followed by a TBS-T blocking solution with 0.25% (w/v) milk. The membrane was then incubated overnight with the primary antibody at 4 ºC. After the first incubation, the membrane was washed three times with TBS-T, under the same conditions described previously. This was followed by a new incubation with secondary antibodies during 1 h at 4 ºC. After incubation, the membrane was again washed three times with TBS-T.

The blotting revelation was done by incubating the membranes with ECF from the Amersham kit (GE Healthcare) for 5 minutes, followed by revelation in a VersaDoc 3000 (BioRad) equipment, using the software QuantityOne, version 4.5.2.

When the membranes were incubated twice with new primary and secondary antibodies, a membrane stripping protocol was applied, by placing them in water for 5 minutes, followed by 5 minutes in a solution of 0.2 M NaOH and in water again.

2.2.5. Samples fixation and cryostat sectioning

The localization of NEP-TC expression on embryogenic tissues was based on the method described by Casson *et al.* (2005), modified for highly hydrated material. The process started with the immersion of the material in a fixing solution of ethanol and acetic acid (3:1, v/v) during an incubation period of 48 h at 4 ºC, with agitation, followed by dehydration by immersion in different solutions with increasing concentrations of sucrose. Firstly, the embryogenic material was placed in cell strainers

on a multiwell plate with a solution of PBS pH 7.4 with 10% sucrose, having been subjected to a vacuum during 15 minutes and then incubated for 48 h at 4 ºC, with agitation. This procedure was repeated with PBS solutions with 15% and 34% sucrose, the latter being supplemented with 0.001% of safranin. After dehydration, the material was frozen in molds using O.C.T (TissueTek ®), and these were immersed in liquid nitrogen. The molds were preserved at -80°C for further analysis.

The material sectioning was carried out using a Leica CM3050S cryostat, and the sections' thickness corresponded to 14 µm. The histological sections were fixed on microscope slides covered with poly-L-lysine (Thermo Scientific).

2.2.6. Immunolocalization

Immunohistochemical analysis of the material was based on the method described by Sauer *et al.* (2006). The tissue sections were marked using a hydrophobic marker (PAP pen), and a permeabilization of the tissues ensued by incubation at 37 °C for 30 minutes, with a 2% (w/v) driselase solution. After incubation, the sections were washed with PBS pH 7.4 solution during 5 minutes, for removal of the permeabilization solution. An incubation with 75 µL of blocking solution (10% BSA in PBS pH 7.4) followed, for 1 hour at room temperature. The sections were subsequently incubated with a 1:10 solution of primary antibodies (75 µL), overnight on a wet chamber. The antibodies dilution was performed in a blocking solution with 1% BSA (w/v) in PBS buffer, pH 7.4. A control group was established, in which the tissues were not incubated with antibodies, in order to help further image analysis with Fiji software: since the tissue revealed itself autofluorescent, the values of integrated density determined for the

control without antibody labelling were subtracted to the ones obtained for the labelled tissues.

After the first incubation, the sections were washed again with PBS pH 7.4, followed by incubation with secondary antibody AlexaFluor 633 (Thermo Fisher), diluted in blocking solution 1% BSA (w/v) in PBS buffer, pH 7.4, during 1 hour in the wet chamber. After incubation, the material was covered with DAKO fluorescent mounting medium and cover slips.

The observations were performed in a confocal Zeiss LSM510 META model microscope, with 621 nm excitation and 639 nm emission.

Image analysis was performed using ZEN software (Zeiss), Version 2.1, as well as Fiji Image J (Schneider *et al*., 2012).

2.3. Enzymatic characterization of recombinant NEP-TC

2.3.1. Recombinant NEP-TC expression in *Escherischia coli*

Preparation of the culture medium

1L of LB medium (Sigma-Aldrich) was prepared for *E. coli* growth, in a final concentration of 25 g/L. 500 mL were supplemented with 30 mg/mL kanamycin and 7.5 g of agar were added, for distribution in Petri dishes. The remaining 500 mL were stored for additional growth.

Transformation of the Escherischia coli cells

The plasmid encoding the NEP-TC cDNA sequence (produced by GeneCust) was suspended in 10 mL of mQ water. 2 µL of this suspension were added to an aliquot

(2 mL) of *E. coli* competent cells (BL-21 strain), which was incubated on ice for 30 minutes. After incubation, a thermal shock was applied at 42 ºC for 1 minute, followed by a new incubation on ice for 2 minutes. 1 mL of LB medium was added to the aliquot, and cell growth was promoted for 1 hour at 37 ± 1 °C, with continuous shaking. This was followed by the spreading of cells on Petri dishes containing LB medium supplemented with kanamycin. The transformed cells' growth occurred overnight at 37 ±1 °C.

This step also had the goal of obtaining a cell bank with the capacity to express the protein of interest whenever necessary. To this end, the transformed cells were placed in 5 mL of LB medium and grew overnight, with agitation, at 37 ± 1° C. After growth, the cells were divided into 2 mL aliquots and frozen at -80° C after the addition of a cryopreserving agent (glycerol).

For NEP-TC sequence confirmation, the plasmid was also transformed into competent *E. coli* cells BL-21 and the isolated clones were used for plasmid extraction and DNA sequencing, using the High Pure Plasmid Isolation kit (Roche).

Induction of NEP-TC expression

The expression medium was prepared by dissolution of 12.5 g of LB medium in 500 mL of distilled water. This was later autoclaved at 121 ºC for 20 minutes. At the same time, transformed cell colonies were placed in 5 mL of LB medium supplemented with 100 µL of kanamycin solution (30 mg/mL), on a growth chamber with agitation at 37 ± 1 ºC. After cell growth, 4 mL of cell suspension were added to the expression medium, and cell growth proceeded for a period of 5 h, at 37 ± 1° C with agitation. When the ideal optical density was achieved (0.850 nm), protein expression was

induced by addition of 1 mL of IPTG (100 mg/mL) to the cell suspension, in a final concentration of 1 mM.

The protein expression occurred overnight at 30 ± 1 °C with agitation.

2.3.2. Cell lysis and recombinant NEP-TC purification

Considering that NEP-TC is an intracellular protein, it was necessary to resort to cell lysis to allow its purification. Thus, the expression culture was centrifuged at high speed and low temperatures (5000 g, 4 ºC) for cell sedimentation. In order to re-suspend the pellet obtained by centrifugation, a first lysis buffer was prepared (2 M urea, 0.05 M Tris and 0.3 M NaCl) and the pH was adjusted to 8. The re-suspension of cells was performed with 4 mL of the referred buffer, to which were added 45 mL of a second lysis buffer (0.1% Triton X-100, 1 mM DTT), supplemented with 100 µL of DNAse, as well as 100 µL magnesium.

Purification of the protein from the lysate was carried out by ion-metal affinity chromatography (IMAC). The lysate was incubated with the matrix for 1 hour at 4 ± 1 °C. After incubation, the matrix was subjected to successive centrifugations with elution buffers, with increasing concentrations of Imidazole (20 mM, 50 mM, 200 mM and 500 mM) for removal of the purified protein. Afterwards, a molecular exclusion chromatography was performed using a Superdex 200-PG column 160 x 15 cm (GE Healthcare Life Sciences), utilizing a renaturation buffer (0.25 M Tris-HCl, 0.15 M NaCl, pH 8.5). Additionally, the protein fractions collected during the molecular exclusion chromatography were concentrated using a centricon filter unit. The final protein concentration was determined with the Bradford method (Bradford, 1976).

2.3.3. Protein profile confirmation

In order to confirm if the protein profile obtained by expression and purification corresponds to the intended, a SDS-PAGE electrophoresis was performed, in accordance with the method described above (section 2.2.4). A Western blot was also performed, in accordance with the method described in the same section.

2.4. Substrates isolation

2.4.1. Total DNA and RNA extraction

Total DNA extraction from non-embryogenic *calli* was performed using the NucleoSpin Plant II kit (Macherey-Nagel), while the extraction of total RNA from the same tissues was performed using the kit NucleoSpin RNA Plant (Macherey-Nagel). The quantification of these substrates was performed in a NanoDrop spectrophotometer.

Total RNA quality was verified by electrophoresis analysis. An agarose gel was prepared (1% agarose, 1x TBE buffer in DEPC treated water, Midori Green staining solution), the samples were mixed with a 5x loading buffer and loaded into the gel. A potential of 100 V was applied during 45 minutes, at room temperature. The gel images were obtained with a GelDoc equipment (BioRad).

2.4.2. rRNA purification

The rRNA fraction was extracted from the agarose gel using the High Pure PCR Product Purification kit (Roche). The elution's RNA concentration was determined in a NanoDrop spectrophotometer.

An additional electrophoretical analysis of rRNA was performed with Experion RNA StdSens Analysis kit (BioRad).

2.4.3. Purification of ribosomes

Ribosome purification was performed according to the method described by Rivera *et al*. (2015) with slight modifications. 13 g of non-embryogenic *calli* were frozen in liquid nitrogen and macerated to a fine powder, and resuspended in 2 volumes of plant extraction buffer (50 mM Tris-HCl pH 9, 30 mM $MgCl_2$, 400 mM KCl, 17% sucrose). The suspension was homogenised by vortexing. The homogenate was then passed through two layers of cheese cloth, followed by a new filtration with eight layers. The crude lysate was layered over a sucrose cushion buffer (20mM Tris-HCl pH 7.6, 5mM $MgCl_2$, 50mM NH4Cl, 60% sucrose) using a pipette and centrifuged at 10 ºC for 3h. The resulting pellet was suspended with resuspension buffer (50mM KCl, 20mM Tris-HCl ph 7.6, 5 mM $MgCl_2$), in a final volume of 160 µL, and stored at -80 ºC for further analysis.

In order to determine the ribosome concentration in the suspension, the 260/280 nm ratio was determined in a Jenway 7305 spectrometer. According to the literature (Boulter and Parthier, 1982), 15 units of absorbance at 260 nm (A_{260}) correspond to 1 mg/mL of plant ribosomes, which have an average molecular weight of 492 µg/µL, corresponding to 0.492 mM. Thus, the quantity of ribosomes in the 10 µL used in the enzyme activity assays was calculated according to this relation.

2.5. Enzyme activity assays

Enzymatic characterization of NEP-CT was performed using the CBA096 SAM Methyltransferase Assay (Calbiochem-Merck Millipore) kit, which is based on the measurement of the absorbance resulting from the production of SAH in SAM-

dependent methylation reactions. The final product of this reaction (H_2O_2) is measured by interaction with the colorimetric agent. The specific activity was then determined on the basis of the ratio between the speed of the reaction and the concentration of protein. The absorbance was measured at 510 nm in a Spectra Max Plus 384 spectrophotometer.

All the assays were performed in duplicate in a micro well plate, with two wells for each sample. NEP-TC was firstly used in different concentrations ranging from 136.5 to 714 µg/µL, while the substrates, DNA and RNA, had concentrations corresponding to 53 µM and 57 µM, respectively. In further assays, NEP-TC had a concentration of 714 µg/µL and the substrates, corresponding to rRNA and ribosomes, had a final concentration of 0.011mM and 0.027 mM, respectively.

2.6. Thermophoresis

The thermophoretical activity of the protein was determined through the use of the Monolith NT His-Tag Labeling RED-tris-NTA kit (NanoTemper), using total RNA as substrate. Recombinant NEP-TC was used in a concentration of 200 nM, while the total RNA sample had an initial concentration of 1 mM, being sequentially diluted in a 1:2 proportion, reaching a final concentration of 100 µM. The substrate dilutions, mixed with the protein, were loaded into the capillaries, and the results were obtained with the Monolith NT. 115 equipment (NanoTemper).

3. Results

3.1. Bioinformatics of NEP-TC

The purpose of this analysis was to determine the phylogeny of NEP-TC, a protein expressed in non-embryogenic tissues of tamarillo, by comparing its sequence with other known sequences from proteins expressed in other plant and algae species. Thus, a bioinformatic analysis was performed using the SeaView software, Version 4. To this purpose, sequences from the HOM03D005634 gene family were searched in the software Plaza Dicots 3.0 and inserted on a Fasta file with NEP-TC's sequence. The sequences were aligned using SeaView and a phylogenetic tree was obtained (Figure 12).

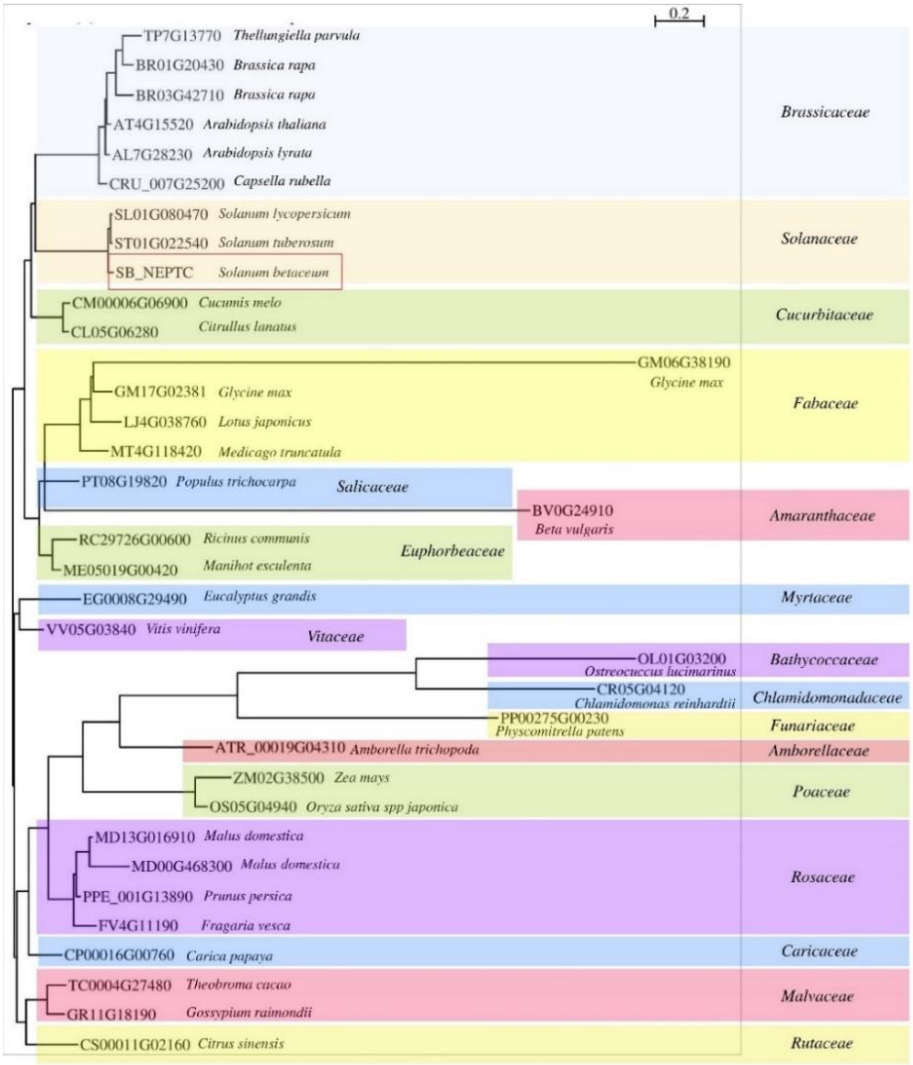

Figure 12 – Phylogeny of NEP-TC. The units of branch length are nucleotide substitutions per site.

As seen on figure 12, genes similar to the one that encodes NEP-TC are present in sixteen plant families, as well as two algae families, comprising 31 different species in total.

The 3D structure of the protein, shown on figure 13, was obtained with the information available at UniProt, using the Protein Model Portal.

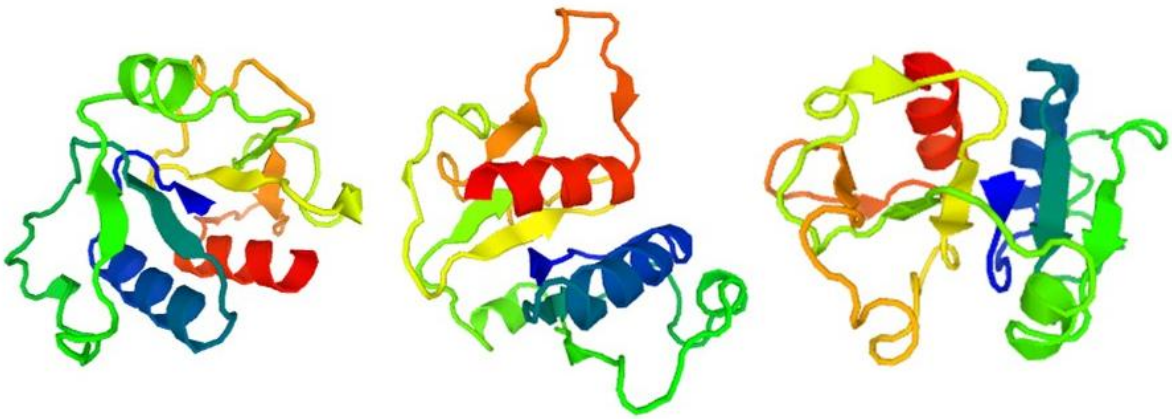

Figure 13 – Predicted three-dimensional structure of NEP-TC, shown from different angles. The blue areas correspond to hydrophobic regions and green areas correspond to polar regions. Red areas correspond to positively charged aminoacids. Orange areas represent glycine residues while yellow areas represent proline residues. The model was created in 31-03-2014 and it's available at www.proteinmodelportal.com

According to the information at the Protein Model Portal, NEP-TC has a globular structure, characterized by the presence of an alpha/beta knot, as well as a tRNA (guanine-N1-)-methyltransferase N-terminal, characteristics that predict that this protein is a tRNA/rRNA methyltransferase of the SpoU family. Some unintegrated signatures indicate that this protein is similar to OBP33pep-like proteins, rRNA methyltransferases found in *A. thaliana*, with a PF00588 domain of the SpoU rRNA methylase family.

3.2. NEP-TC expression analysis in tamarillo somatic embryogenesis

3.2.1. *In vitro* establishment of plant material

In order to obtain embryogenic and non-embryogenic *calli* of tamarillo, as well as tissues in the process of differentiation, leaf explants were placed on a MS medium with 9% sucrose and 5 mg/L of the synthetic auxin picloram to induce SE.

Before sectioning of the plant material and incubation with the purified antibodies, the samples were individually observed to determine the percentage of explants forming embryogenic callus. These results are summarized in table 1.

Table 1 – Results of the somatic embryogenesis induction protocol on tamarillo leaf segments. Results are presented as mean ± standart deviation

Number of explants	55.71 ± 19.83
Contamination (%)	28.11 ± 25.09
Non-responsive (%)	11.79 ± 21.77
Responsive (%)	17.29 ± 17.29
Induction (%)	11.00 ± 4.37

The induction were performed with approximately 55 explants per assay, from which 17.3% responded to the treatment. From the responsive explants only 11.00% produced embryogenic *calli*.

Representative images of the tissues in development, as well as the embryogenic and non-embryogenic material used for the immunolocalization are pictured below, in figures 14 and 15.

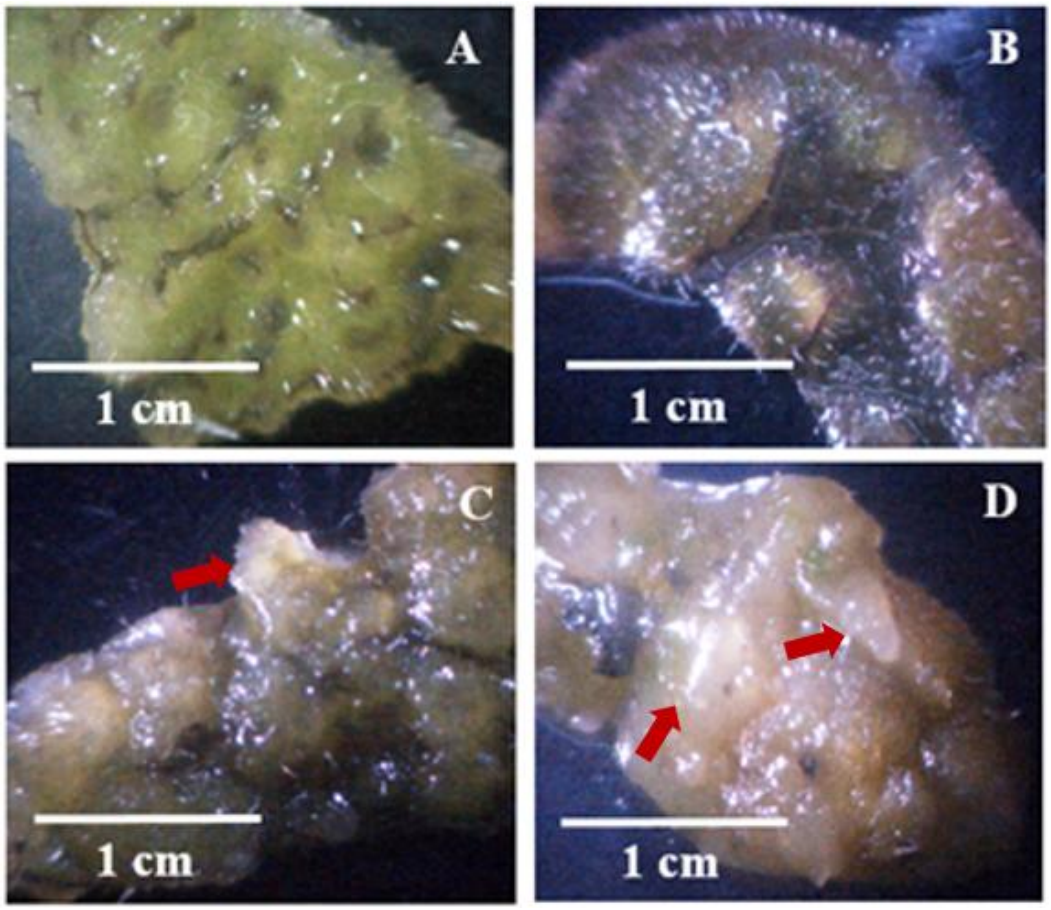

Figure 14 – Time course induction of SE. (A) Tamarillo leaf after SE induction (from a plant growing *in vitro* for 2 months). (B) Explant 6 weeks after induction (C) Explant 8 weeks after induction, with development of embryogenic areas. (D) Explant 10 weeks after induction. The arrows indicate dedifferentiated regions with possible induction of embryogenic cells.

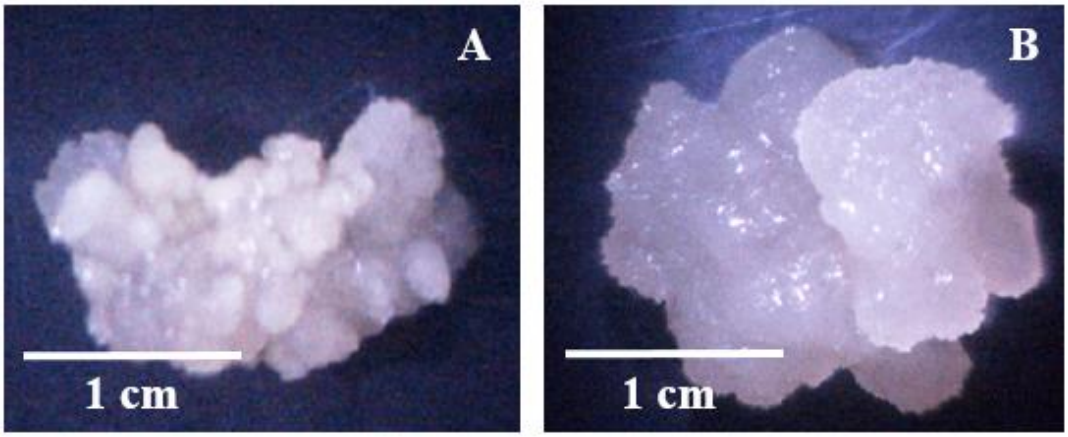

Figure 15 – Isolated and subcultured *calli* induced from tamarillo leaf explants. (A) Embryogenic callus (B) Non-embryogenic callus.

Figure 14 shows proembryogenic tissues throughout the process of somatic embryogenesis induction. Wounded leaves start to dedifferentiate after the 4th week of induction and *calli* appear in the explants 6-8 weeks after induction (Fig. 14B). Around the 10th week (Fig. 14D) areas of embryogenic cell formation can be distinguished amongst the more friable tissues. These embryogenic areas were isolated from the surrounding tissues and subcultured in TP medium, forming compact globular cell masses (Fig. 15A), while the non-embryogenic tissues cultured in the same medium are friable, translucent and fast growing (Fig. 15B).

3.2.2. Western blotting analysis

In order to confirm affinity of the purified antibodies towards NEP-TC, a Western blot was carried out. Analysing figure 16, it can be seen that the antibodies show an affinity towards recombinant NEP-TC, with consistent blotting in the 25 kDa strip. Thus, the purified anti-NEP-TC antibodies could be used to determine the expression of NEP-TC in plant tissues, as will be shown further.

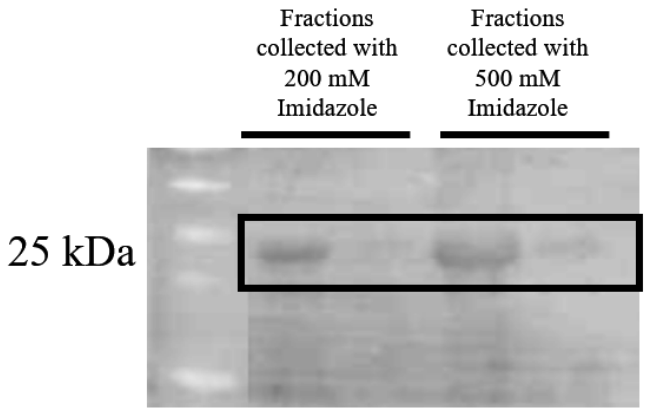

Figure 16 – Blotting of NEP-TC by specific antibodies purified from rabbit anti-NEP-TC serum. The boxes comprise NEP-TC fractions from affinity chromatography.

For western blotting analysis, SDS-PAGE profiles were obtained and analysed (Figure 17). The blotted samples chemiluminescence was determined with Fiji software, and the results are summarized in figure 18.

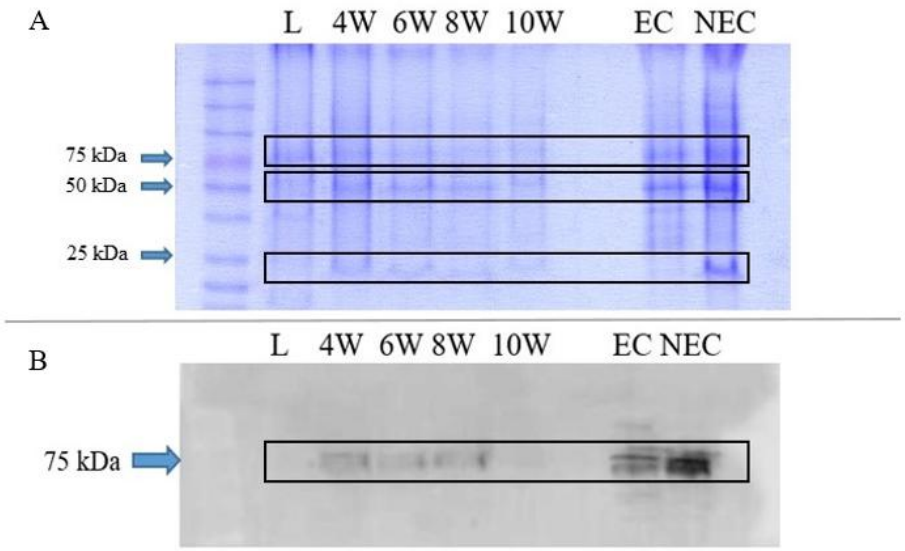

Figure 17 – SDS-PAGE (A) and blotting of NEP-TC (B) from protein extracts. L represents extracts from tamarillo leaves, while 4W-10W correspond to the selected induction points in the time course (4 weeks till 10 weeks). EC and NEC correspond to the extracts from embryogenic and non-embryogenic *calli*. 20µL of each sample were loaded in the wells.

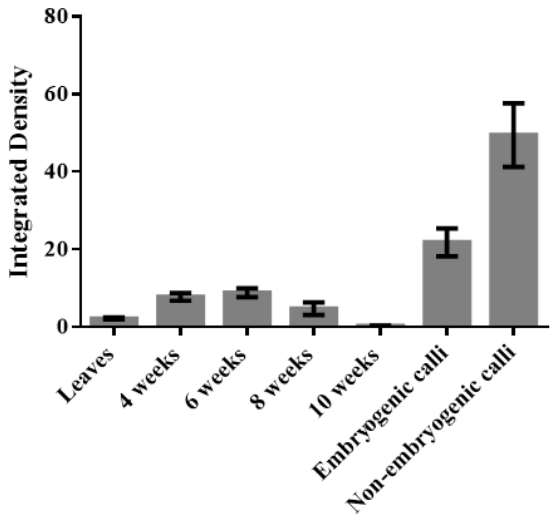

Figure 18 – Integrated density results for the Western blot assay.

Analysing the SDS-page results in figure 17, three different bands, corresponding to 25, 50 and 75 kDa, appear more intensively stained, having a higher expression in the extracts corresponding to EC and NEC than in the time course extracts. There is no detectable presence of the 25 kDa band in the leaves' extracts. The band is present in all extracts corresponding to the induction points of the time course, with its expression decreasing significantly after the 8th week. These results are corroborated by the Western blot analysis, in which NEP-TC was labelled in extracts corresponding to the 4th, 6th and 8th weeks after induction, while there's no significant labelling on the 10th week (Fig. 17B). Nonetheless, there is intense labelling in extracts from EC and NEC, with a striking high level of fluorescence in the extracts from NEC. The integrated density results (obtained by multiplying the average chemiluminescence intensity with the analysable area of the image) confirm the differences seen on blot labelling.

One noteworthy detail is the fact that in the Western blot the protein was labelled in the 75 kDa strip, showing no labelling in the 50 kDa or 25 kDa strip.

3.2.3. Immunolocalization

In order to determine the localization of NEP-TC expression in the embryogenic and non-embryogenic tissues of tamarillo, as well as in the dedifferentiated tissues, sections of explants and developed *calli* were obtained and incubated with specific antibodies for this protein. The images obtained by confocal microscopy for the antibody labelling of NEP-TC are shown below, on Figures 19 and 20.

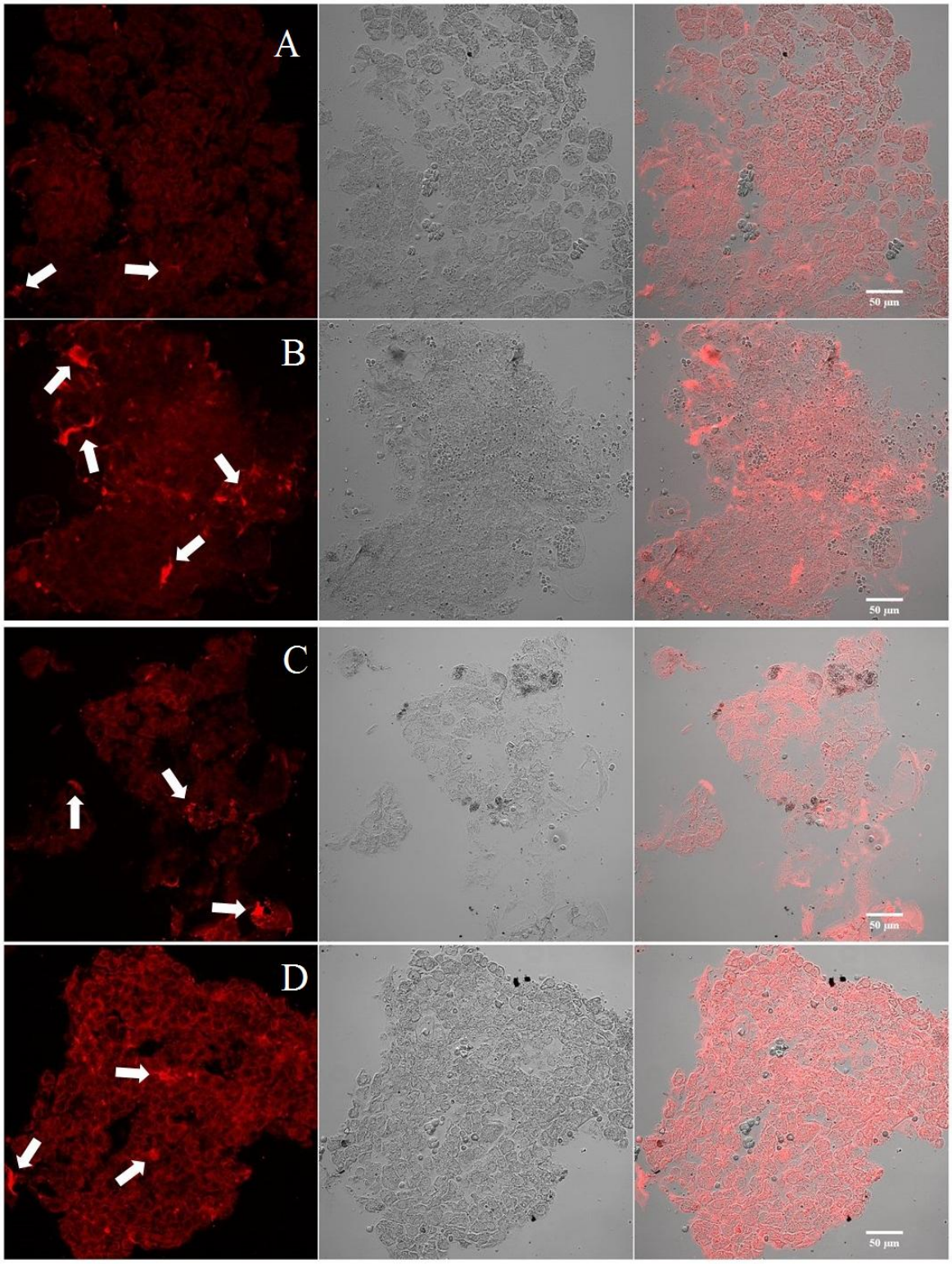

Figure 19 – Immunolocalization of NEP-TC during the timecourse of SE induction in leaf explants. The arrows indicate the places where the protein has been labelled (A) Explant after 4 weeks of induction. (B) Explant after 6 weeks of induction. (C) Explant after 8 weeks of

induction. (D) Explant after ten weeks of induction. The images are composed in three elements: the first image corresponds to the fluorescence, while the second corresponds to the transmission image. The third image is the merge of the two.

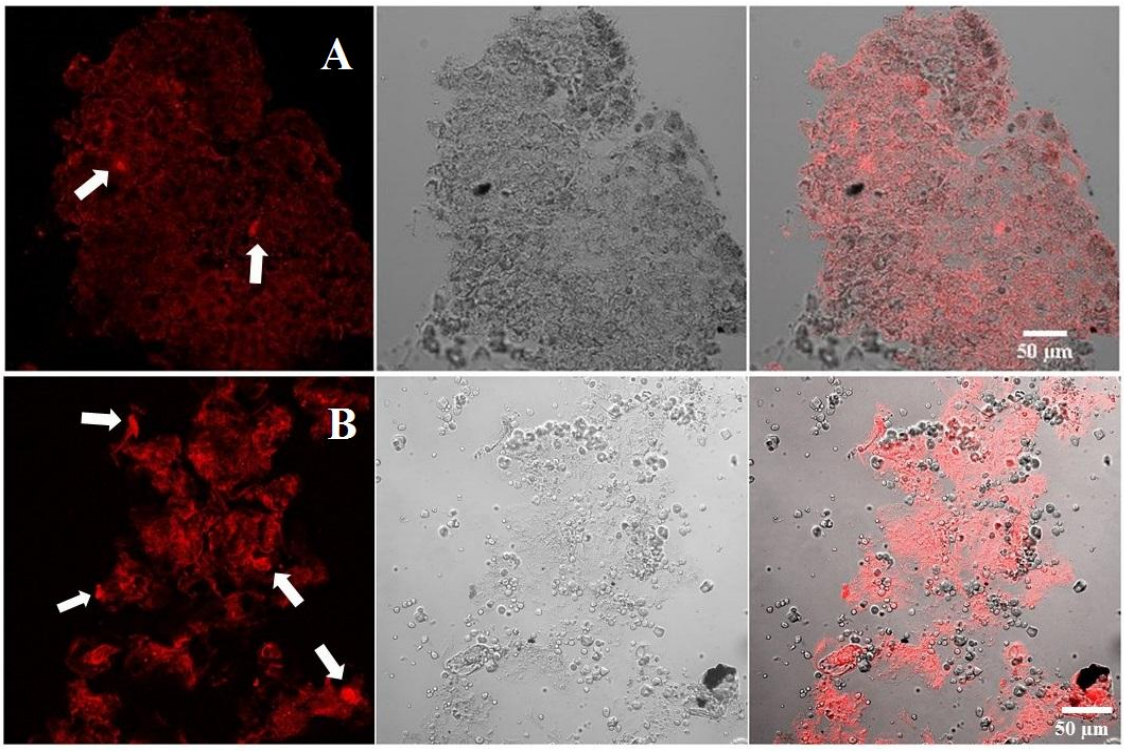

Figure 20 – Immunolocalization of NEP-TC in EC and NEC. The arrows indicate the places where the protein has been labelled. (A) Embryogenic callus. (B) Non-embryogenic callus. The images are composed in three elements: the first image corresponds to the fluorescence, while the second corresponds to the transmission image. The third is the merge of the two.

Analysing the images corresponding to the proembryogenic tissues, it is clear that four weeks after the induction, NEP-TC appears in some distinct labelled spots, that increase in frequency along the induction phase (Figure 19A), reaching its peak in the 6th week. The obvious exception to this increasing tendency is the 8th week, where the labelling is reduced. However, it increases consistently after, in the 10th week of development. When it comes to the embryogenic tissues (Figure 20), the non-embryogenic *calli* present significant labelling in the whole tissues' extension, whilst in

embryogenic *calli* the presence of NEP-TC is not frequent, or even absent, thus showing that the immunolocalization results are consistent with those from the Western blot.

The values of integrated density obtained for the several tissues are summarized in figure 21.

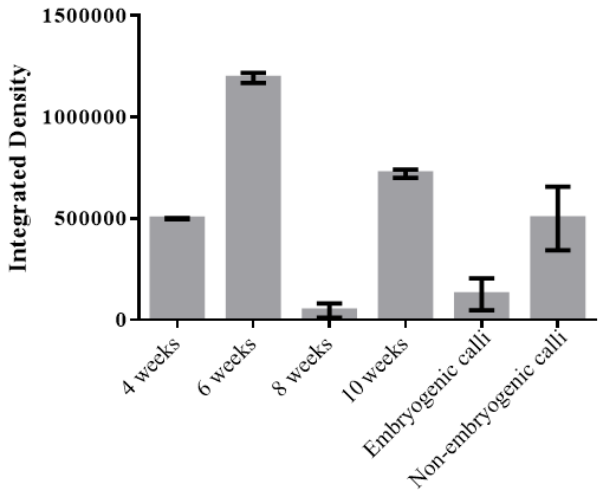

Figure 21 – Integrated density results for the immunolocalization assays.

A closer look at the tissues that presented labelling with the antibody allowed to observe NEP-TC's expression in the peripheral area of the cells (Figure 22).

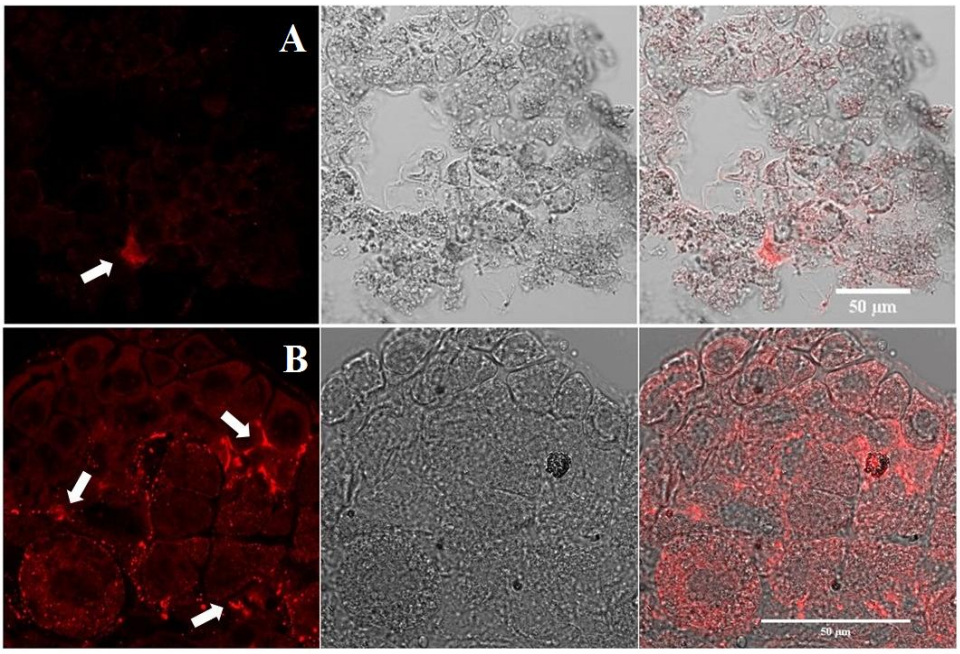

Figure 22 – NEP-TC labelling in embryogenic and proembryogenic tissues. The arrows indicate the places where the protein has been labelled. (A) Embryogenic callus. (B) Explant after 10 weeks of induction. The expression is evident in peripheral areas of the tissues. The images are composed in three elements: the first image corresponds to the fluorescence, while the second corresponds to the transmission image. The third image is the merge of the two.

3.3. Enzymatic Characterization of recombinant NEP-TC

3.3.1. Expression of recombinant NEP-TC in *Escherischia coli*

In order to perform enzymatic assays to determine NEP-TC's specific activity, the recombinant protein was expressed on a heterologous system. Thus, *E. coli* cells were transformed with a plasmid encoding the NEP-TC gene, the protein's expression was induced with IPTG and, after cell lysis, several chromatographies and a dialysis were performed in order to isolate it.

The plasmid sequencing results were aligned with NEP-TC's cDNA sequence using the online tool MultAlin (http://multalin.toulouse.inra.fr/multalin/), and it was possible to determine that NEP-TC's sequence was successfully inserted into the expression plasmid.

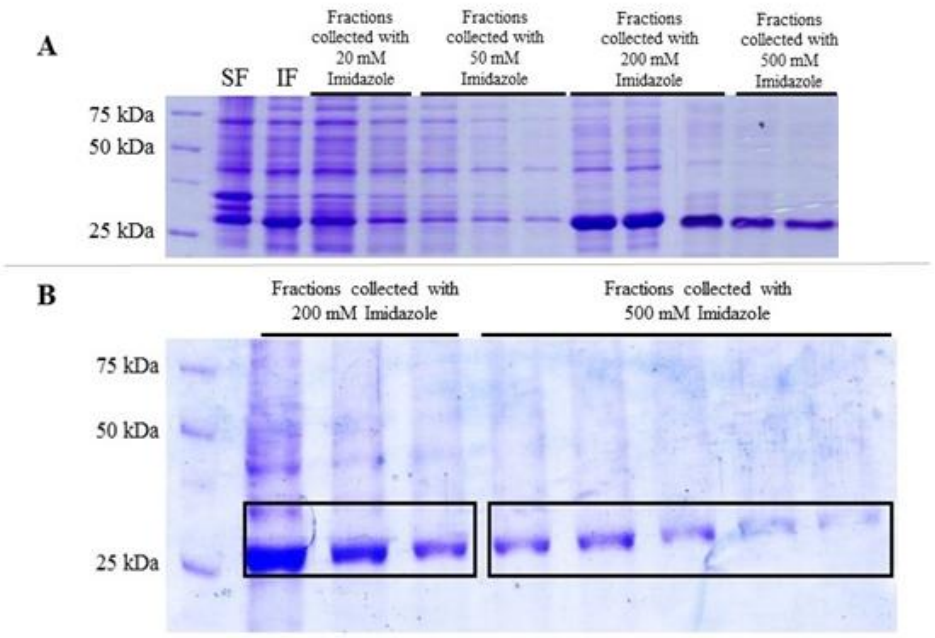

Figure 23 – NEP-TC fractions collected from the IMAC column. (23A) "SF" and "IF" correspond to the proteins' soluble and insoluble fractions. (23B) Detail of the fractions collected with 200 and 500 mM Imidazole. The boxes comprise the strips that correspond to approximately 25 kDa. 20µL of each sample were loaded in the wells.

Analysing the results shown on Figure 23, it is possible to see the fractions collected from the ion-metal affinity chromatography. As seen on Figure 23A, the purified protein matches the 25 kDa strips, and NEP-TC molecular weight is 26.5 kDa. It is also possible to see some protein aggregation in the first purified fractions, in the 50 kDa and 75 kDa strips. It is also clear that the protein has a higher concentration in its soluble form.

The protein fractions collected with 200 mM and 500 mM of imidazole (Fig. 23B) show a higher level of purity, and therefore were selected to be used in further assays, being pooled and concentrated. A molecular exclusion chromatography was performed in order to determine the protein's aggregation, as well as to exchange the

buffer in which the protein is suspended, because the original buffer is not appropriate for the enzymatic assays.

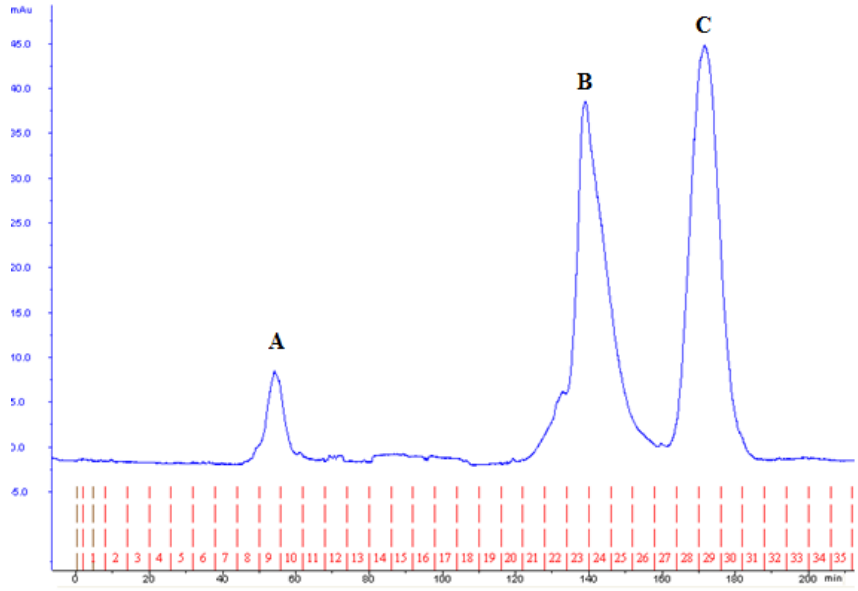

Figure 24 – Gel filtration results. The first peak (A) corresponds to NEP-TC in its aggregated form, while the second peak (B) corresponds to the protein on its purified form. The third peak (C) corresponds to Imidazole.

The peak fractions were collected and analysed by Bradford assays. While peaks A and B revealed the protein's presence, peak C had lower concentration values. Thus, it can be inferred that peak C corresponds to Imidazole.

The dialysed pool fractions were analysed by SDS-Page, and the results are shown in Figure 25. The dialysis was necessary in order to evaluate the protein's stability and to guarantee the buffer exchange for the kinetic activity assays.

The protein concentration was determined in the end of this process by Bradford assays, with an average concentration of 136.50 ± 29.50 µg/µL, which corresponds to 0.1365 mg/mL.

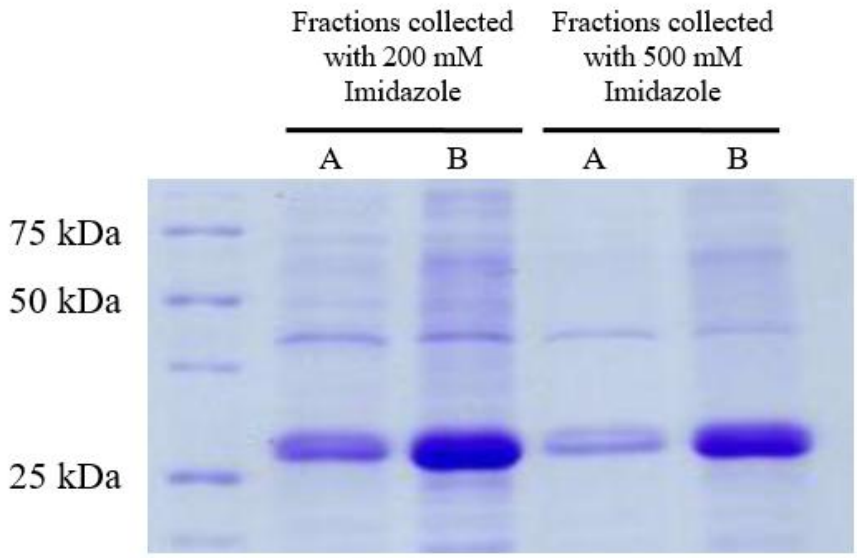

Figure 25 – SDS-page results of NEP-TC pooled fractions dialysis. A corresponds to fractions dialysed against a buffer without urea, while B comprises fractions dialysed against a buffer supplemented with 0.2 M of urea. 20µL of each sample were loaded in the wells.

3.3.2. Substrate obtainment and analysis

In order to validate the RNA and ribosome extraction from NEC, and also to extract the band corresponding to rRNA, an electrophoresis was performed. The RNA concentration in the samples corresponds to 57 µM, whilst the ribosome concentration is 2.7 mM, as shown on table 2.

The electrophoresis results are shown in Figure 26. The migration of the total RNA sample allowed the identification of the 18S and 28S bands, corresponding to the rRNA fraction, which was to be isolated. The ribosome samples did not migrate, but were stained by the colouring agent, thus confirming the presence of these organelles.

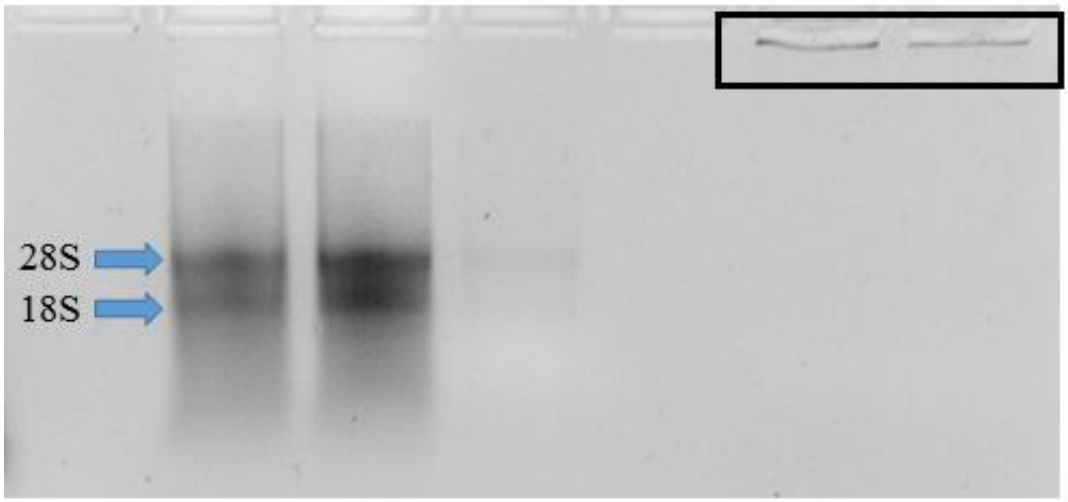

Figure 26 – Total RNA and ribosome electrophoresis results. The bands corresponding to the ribosome suspension are within the box.

The rRNA band (between 18S and 28S) was removed physically and submitted to a treatment in order to remove the agarose. After isolation, quantification in a NanoDrop spectrophotometer was performed. The results are displayed in table 2.

Table 2 – Quantification of total RNA, rRNA and ribosomes

Substrate	**Concentration (µg/mL)**	**$A_{260/280}$**
Total RNA	384.00 ± 6.680	2.107 ± 0.003
rRNA	3.1 ± 0.000	1.771 ± 0.030
Ribosomes	27.00	1.312

As seen on table 2, the analysis revealed an rRNA concentration of 3.1 µg rRNA/mL, which corresponds to 0.011 mM rRNA. Since this concentration is extremely low, this fraction cannot be analysed by regular electrophoresis. Therefore, a virtual electrophoresis was performed, using Experion RNA StdSens Analysis Kit (BioRad). Results are shown in Figure 27.

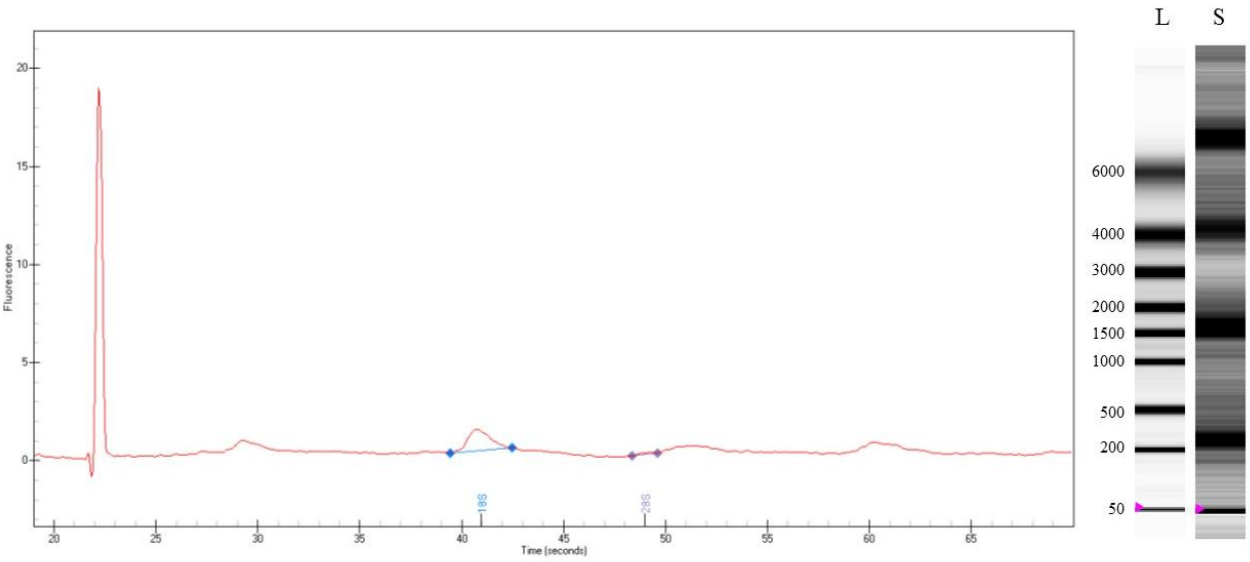

Figure 27 - Experion RNA StdSens Analysis Kit results. The ladder is represented with "L", while thevirtual gel corresponding to the rRNa sample is represented with "S".

Analysing the graph in Figure 27, it can be seen that the sample's concentration is extremely low. Nonetheless, there was detection of rRNA in the 18S and 28S strips, which correspond to the peaks. Examining the virtual gel, is it possible to conclude that rRNA didn't maintaing its integrity throughout the extraction process, so the kit only detected a collection of fragments corresponding to 18S and 28S rRNA. Although the rRNA is not whole, it can be used for further enzymatic assays.

3.3.3. Enzymatic activity

Since there are no data concerning the specificity of NEP-TC beside the ones obtained by homology of sequences with *A. thaliana*, some preliminary assays were

needed to determine if the protein shows indeed affinity towards RNA or if it also has the ability to methylate DNA. Quantification of the recombinant protein by Bradford assays shows that it has an average concentration of approximately 625 µg/µL. This purified protein was used in Calbiochem's SAM Methyltransferase Assay kit to determine its affinity towards different substrates. The concentrations of DNA and RNA used in the assays correspond to 53 µM and 57 µM, respectively. Since the bioinformatics data show that this protein might be a rRNA methyltransferase, enzymatic assays were performed with both ribosomes isolated from non-embryogenic *calli* and rRNA extracted from total RNA isolated from the same tissue (as shown in figure 26). The concentration of rRNA corresponds to 0.011 mM, whereas the ribosome concentration corresponds to 0.027 mM. The specific activity values for these substrates are summarized in table 3.

Table 3 – NEP-TC's specific activity according to different substrates. The results for DNA weren't obtained in triplicate.

Substrate	**Specific activity (nmol/min/ mg protein)**
RNA	0.0323 ± 0.0245
DNA	0 - 0.0228
Ribosomes	0.0393 ± 0.0285
rRNA	0.337 ± 0.152

According to the results displayed on table 3, it is possible to infer that NEP-TC has higher specific activity values towards total RNA, with an average value of 0.0323

nmol/min/mg protein. However, the specific activity towards DNA has a value of 0 nmol/min/mg protein, with tendency to be negative. This preliminary screening allowed the elimination of DNA as a potential substrate, and the subsequent assays were performed in order to determine which fraction of the total RNA this protein methylates.

After defining that the protein does in fact methylate rRNA, several concentrations of this substrate were tested. The results are summarized in Figure 28.

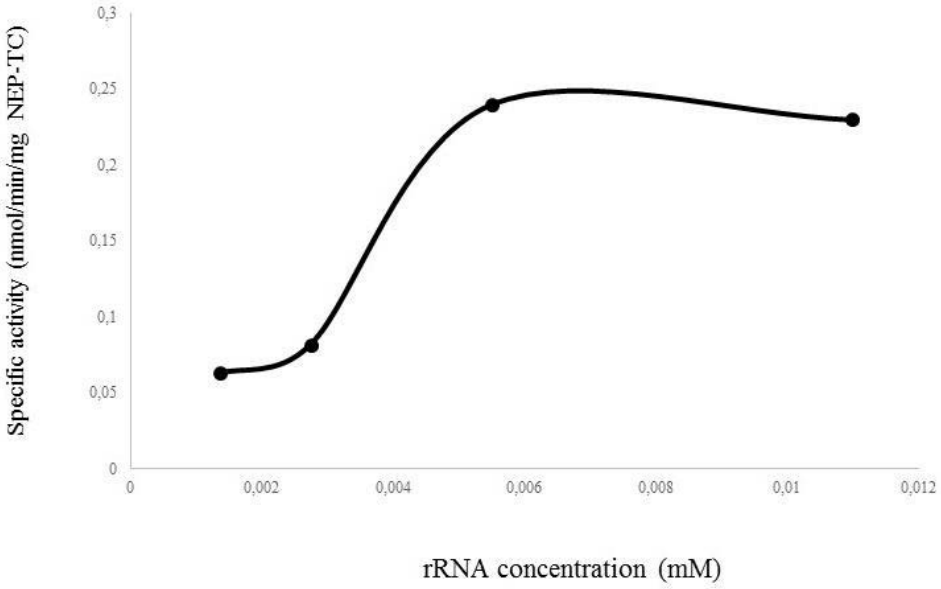

Figure 28 – Relation between specific activity values and rRNA concentration

Evaluating the graph, it is possible to determine that for the highest concentration, corresponding to 0.011 mM rRNA, the reaction reaches a saturation plateau, while for concentrations between 0.0014 mM and 0.0055 mM the enzyme is in log phase. Due to the insufficiency of assays regarding NEP-TC's specific activity, the determination of the proteins' kinetic parameters was not possible.

3.3.4. Thermophoresis

Thermophoresis is a technique based on temperature gradients that allows the determination of the protein's affinity towards a given substrate. A sample of total RNA from NEC was used, in an initial concentration of 1 mM. After a short incubation of the labelled protein with different concentrations of RNA (100 µM- 3 nM), the samples were loaded into capillaries and MST analysis was performed. The results are shown in Figure 29.

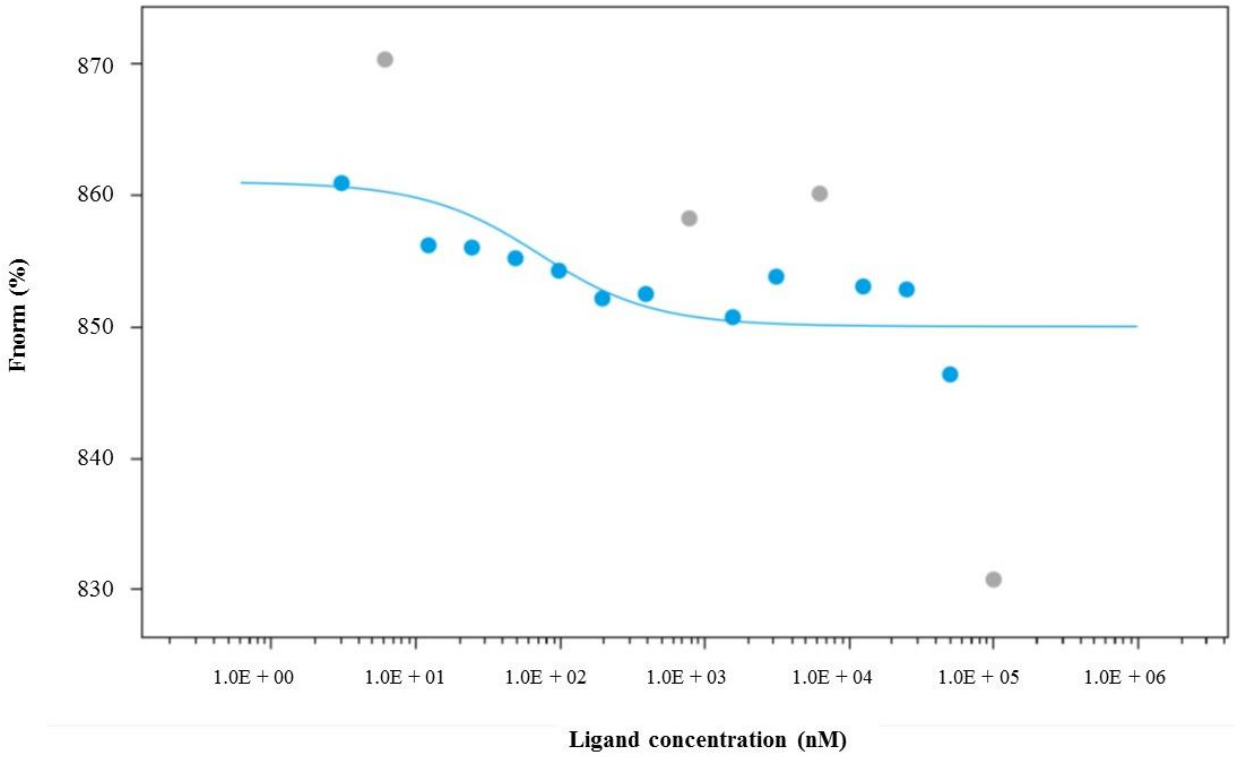

Figure 29 – Thermophoretic interaction between NEP-TC and RNA.

Analysing the graph, it is clear that NEP-TC shows affinity towards RNA, since the average fluorescence results are high, ranging from 850 to 860. A dissociation

constant (Kd) in the nM range could be assumed for this interaction. There is, although, some irregular binding of the protein to the substrate.

The protein was also incubated with Tafa, a control ligand provided in the kit, and the results are shown in Figure 30.

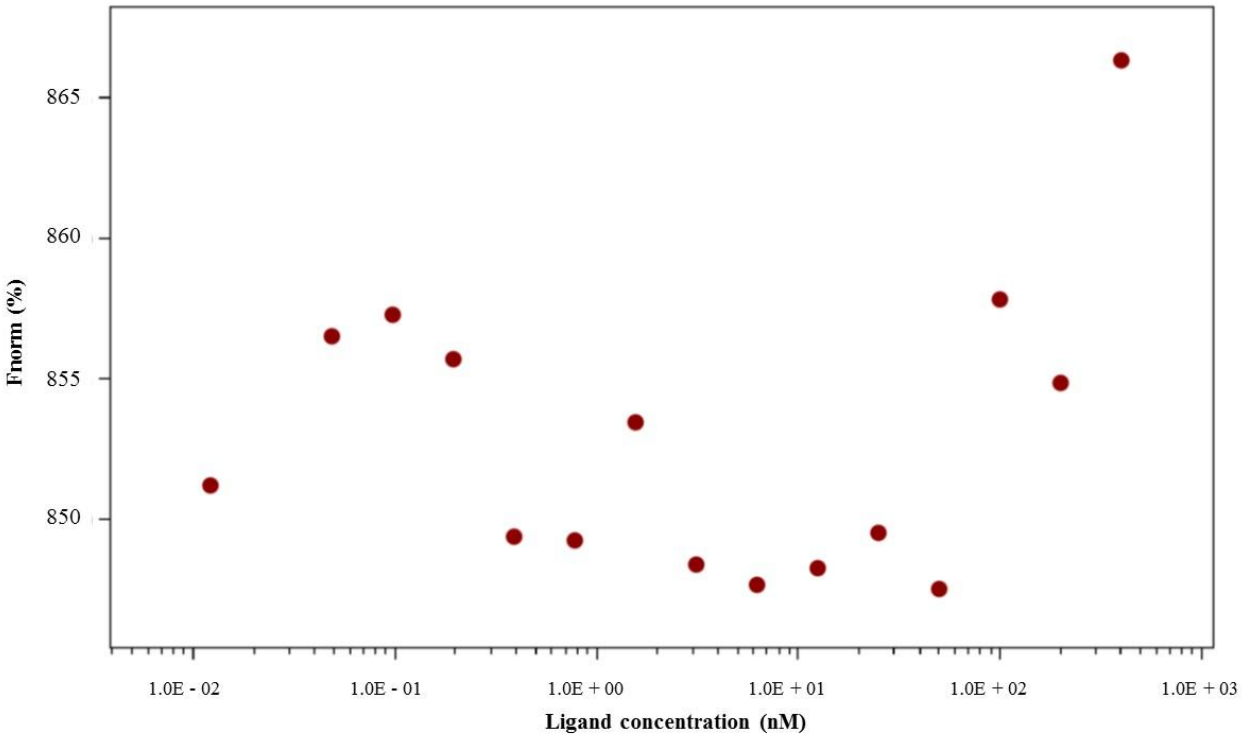

Figure 30 - Thermophoretic interaction between NEP-TC and Tafa.

The determination of the Kd was not possible, since the substrate concentrations were too high. Also, the protein binding is extremely uneven, and the cause of this irregularity will be discussed further.

4. Discussion

4.1. Bioinformatics

One of the most frequent types of RNA modifications is SAM-dependent methylation, and although the function of this modification remains largely unknown, it is conceivable that it can alter the three-dimensional structure of RNA and its interaction with ligands, a phenomenon that has been hard to demonstrate experimentally (Bauerle *et al.*, 2015). Therefore, bioinformatics studies concerning a comprehensive sequence analysis and a systematic comparison of protein structures are fundamental to infer their phylogenetical history and identify important differences and similarities between them, as well as their function (Feder *et al.*, 2003).

Regarding NEP-TC, the only available data regarding bioinformatics studies of this protein were presented by Correia (2011), in which the protein was identified as a putative SAM-dependent tRNA/rRNA methyltransferase, belonging to the SpoU family (the second-largest group of RNA methyltransferases), by sequence homology with *A. thaliana*. A three-dimensional modulation of the protein was also performed in order to determine the most appropriate sequences for polyclonal antibody production. Nonetheless, further studies were required in order to determine the phylogeny of this protein, as well as a more extensive analysis of its structure.

Therefore, phylogenetic analysis of the protein was carried out by alignment of sequences from the HOM03D005634 gene family with NEP-TC's gene sequence. According to protein databases (Dicots Plaza 3.0), the HOM03D005634 gene family comprises 38 genes expressed in 31 species, and the main molecular function of these genes is related to RNA methyltransferase activity.

Analysing a phylogenetic tree, the horizontal dimension gives the amount of genetic change that has occurred, and the branches represent evolutionary lineages over

time: the longer the branch, the larger the amount of change in the nucleotide sequences. Furthermore, the tree's internal nodes represent hypothetical ancestors of the descendant clades of that node.

Scrutinising *NEP-TC*'s phylogenetic tree, all the species present belong to the *Viridiplantae* clade, which comprises some of the smallest and largest eukaryotes. Most phylogenetic analysis of *Viridiplantae* have reported two subclades: *Chlorophyta*, that contains most of the green algae, and *Streptophyta*, that contains the land plants (Ruhfel *et al*., 2014). In *NEP-TC*'s tree, both land plants and algae can be found. It is clear that all the genes have a common ancestor, which means that the HOM03D005634 gene family to which they all belong has only one gene that has been suffering alterations over time, originating the 38 genes shown in the tree.

Focusing on the Solanaceae family, to which tamarillo belongs, there is polytomy, meaning that NEP-TC has the same common ancestor as SL01G080470, present in tomato (*Solanum lycopersicum*), and ST01G022540, present in potato (*Solanum tuberosum*). Both these genes encode tRNA/rRNA methyltransferases. These are the two genes to which NEP-TC is more closely related. Not surprisingly, the genes that are least related to NEP-TC are the ones expressed in algae, since they're the only organisms in the tree that are not angiosperms. When it comes to plants, the least relatable genes are expressed in earthmoss (*Physcomitrella patens*), soy (*Glycine max*) and beetroot (*Beta vulgaris*). These results show some insight into the evolutionary history of genes that encode these unique methyltransferases in plants, and might be useful in the determination of NEP-TC's role in tamarillo's physiology.

Regarding the protein's structure, it was already described in the results that it has an alpha/beta knot characteristic of some SAM-dependent methyltransferases. This structure is extremely rare in protein structure, as it poses a folding problem (Lim *et al*.,

2003). Although the detailed role of this structure is not yet understood, its functional importance is without doubt: the knotted region encompasses the conserved motifs predicted to be SAM's binding site, playing an important role in dimer formation (Basturea and Deutscher, 2007). Comparing NEP-TC with other rRNA methyltransferases described in literature, Michel *et al*. (2002) characterized an rRNA methyltransferase expressed in *E. coli* (RImB), with a size similar to NEP-TC's (29 kDa), that also has a globular shape and the characteristic alpha-beta-knot. Furthermore, this protein has also the tendency to form dimers when the NaCl concentration is inferior to 500 mM. RsmH, a rRNA methyltransferase characterized by Wei and co-workers (2012) is a globular protein with a molecular weight of 35 kDa, and also tends to form dimers when in solution, which indicates that *in vivo* it is active in its dimeric architecture. These data reinforce the idea that NEP-TC is in fact an rRNA methyltransferase, for its behaviour and structure are similar to others described in the literature for these enzymes.

When it comes to the protein's function, according to CATH: Protein Structure Classification Database, several proteins similar to NEP-TC show methyltransferase activity towards tRNA and rRNA, as well as ribosome binding and protein homodimerization functions. Additionally, this protein's knotted structure might be relevant in the organism's response to stress. This topic will be discussed further.

4.2. Immunohistochemistry

In recent years, the advent of high-throughput methods for expression analysis allowed the collection of a wealth of information about plant gene expression (Paciorek *et al.*, 2006). However, in order to elucidate protein function, interaction and activity,

global data do not suffice: to this purpose, cellular and subcellular resolution are also needed (Sauer *et al*., 2006). Although direct visualization involving recombinant DNA techniques (such as green fluorescent protein) permitted the collection of a large amount of information regarding plant protein localization, they present one drawback: the recombinant proteins may not reflect in an accurate way the localization of the wild-type equivalent because their molecular properties are altered (Paciorek *et al.*, 2006). Immunolocalization with specific antibodies for the protein of interest is a complementary technique, which avoids the risk of inducing side effects by a fusion protein, such as misexpression, mistargeting, altered stability, or toxicity (Sauer and Friml, 2010).

Previous immunoblot studies concerning NEP-TC (Correia, 2011) proved that this proteins' expression is not exclusive of NEC, being also expressed in initial explants and earlier induction stages. This expression decreases when embryogenic areas start to develop in induced explants. Immunolocalization assays performed with tamarillo embryogenic tissue sections (Casimiro, 2014) supported these evidence, and were made in order to determine the subcellular localization of this protein. This purpose was not fully accomplished, because the polyclonal antibodies used were not purified, and reacted with non-biological components. It was only possible to determine that it has a cytoplasmatic expression, probably near the nucleus. Therefore, in the present study the immunohistochemical analysis was performed using anti-NEP-TC specific antibodies purified from the serum.

As described in the results, NEP-TC presents observable expression in early stages of SE induction. Moreover, there's a decrease of the protein's expression in the 8^{th} week of development, when embryogenic tissues start to arise in the explant. These results are in accordance with previous data (Correia, 2011), showing only one

difference: after the 8th week, the protein's expression increases. This phenomenon might be justified by the fact that the proembryogenic masses in the 10th week of development presented a considerable area of non-embryogenic tissue, where the protein has its main expression. Concerning NEC and EC, image analysis shows that EC has a considerably lower fluorescence when compared to NEC, and protein labelling is almost absent. When it comes to NEC, it displays a consistent expression of the protein throughout the samples' extension, thus confirming all the assumptions made in previous studies that NEP-TC is mainly expressed in these tissues.

However, the integrated density results for NEC are lower or quite similar to some presented for the induced tissues. NEC are mucilaginous and friable tissues, with a high water content (Canhoto, 2010), and have to be submitted to rigorous dehydration conditions in order to be prepared for preservation in O.C.T and further sectioning with a cryostat microtome. Despite the optimized dehydration process, some of NEC's inner cells may have retained a high water content, what might have contributed for their eventual ripping and detachment from the microscope slides during the incubation process. This fact may have contributed for a decrease in the analysable area of the NEC tissues, which is smaller to the one of the induced and embryogenic masses that easily retained their integrity. Since the integrated density results are obtained based on the average fluorescence multiplied by the area of the explant, a reduced value for the latter might explain the lower results obtained for NEC.

When it comes to the validation of these results by blotting of the proteins, comparison of both integrated fluorescence graphs shows that the results are consistent in both Western blot and immunolocalization assays. In both techniques, NEP-TC's expression is higher in the 4th and 6th weeks after induction, with a decrease in the 8th week. There's one difference in the results for the 10th week: in the Western blot assays,

there is no detectable labelling of the protein, while in the immunolocalization assays it increases to higher levels after the 8th week. In both assays the protein's expression is higher in NEC than in EC.

These immunolocalization and blot results are in accordance with what was previously observed for this protein (Correia, 2011), except for the differences shown in the 10th week after induction. As referred before, several weeks after induction both embryogenic and non-embryogenic *calli* start to arise from the same explant (Lopes *et al*., 2000). As shown in the results, it is clear that in the 10th week the tissue reaches a high level of heterogeneity. It is possible that the explant used to extract the proteins that were used in the blotting assays showed a higher area of embryogenic tissue, and since NEP-TC's expression is lower in EC, the results were inferior when compared to those of immunolocalization.

The SDS-page results reveal that there is a significant presence of protein in the extracts, in the 25 kDa, 50 kDa and 75 kDa strips. Surprisingly, when it comes to blotting, there was no labelling of NEP-TC with specific antibodies in the 25 kDa or 50 kDa strips, with consistent labelling only in the 75 kDa strip. So, there is no correlation with NEP-TC's molecular weight (26.5 kDa), and the possible explanation for this is that in the extracts NEP-TC has a higher concentration as a trimer. Thus, it can be assumed that in natural conditions this protein presents an oligomeric conformation. These results are consistent with literature reports (Tkaczuc *et al*., 2007; Motorin and Helm, 2011), for it was predicted that methyltransferases from the SpoU family are usually active in their multimeric form.

In terms of subcellular localization, a close analysis of the labelling in the augmented sections reveals that NEP-TC has a peripheral expression in the cells, being

mostly present in the cytosol. These results are in accordance with what was previously predicted and partially observed for this protein (Correia, 2011; Casimiro, 2014).

Data available in the literature concerning the subcellular localization of RNA methyltrasferases in plants is scarce or even absent. Studies performed with *A. thaliana* (Krupkova *et al*., 2007; Held *et al*., 2011) and *Nicotiana tabacum* (Liu *et al*., 2015) revealed that SAM-dependent methyltransferases have their catalytic domain in the cytosolic side of the Golgi apparatus, and are extremely mobile in the plant cell's cytosol. Still, none of these enzymes is a RNA methyltrasferase, so the results obtained for NEP-TC are, as far as we are concerned, pioneer. There are, however, some data available regarding the subcellular localization of RNA methyltransferases in yeast. In recent years, several RNA methyltransferases have been identified in *S. cerevisiae*: Rcm1 and Nop2, responsible for the methylation of m^5C residues of rRNA, are expressed in the nucleolus (Sharma *et al*., 2013; Bourgeois *et al*., 2015), while Btm5 and Btm6, responsible for the methylation of m^3U of 25S RNA are expressed in the nucleolus and the cytoplasm, respectively (Sharma *et al*., 2014). BUD23, that methylates rRNA and is required for efficient nuclear export of its subunits, is also expressed in the nucleolus (White *et al*., 2008). Establishing the places of expression of these enzymes in yeast will not only help identifying the function of their homologs in higher organisms, but will also enable understanding the role of these modifications in ribosome function and architecture (Sharma *et al*., 2013).

Furthermore, there is extensive information about the subcellular localization of these enzymes on protein databases. According to CATH: Protein Structure Classification Database, several proteins similar to NEP-TC that act as rRNA methyltransferases are expressed in the cytosol of both yeast and bacteria. Thus, it is possible to establish a correlation between NEP-TC and these proteins, once NEP-TC is

mainly expressed in the cytosol, as was observed in the immunolocalization assays. There is as well one non-characterized putative RNA methyltransferase expressed in chloroplasts of *A. thaliana*. Further studies regarding NEP-TC should include an extensive characterization of the protein's location in the cell, and not only the tissues in general.

An interesting finding, although outside the scope of this work, is that the *calli* observed in confocal microscopy present autofluorescence, once the control groups (that were not incubated with antibodies) presented fluorescence in all spectra.

4.3. Enzymatic characterization of recombinant NEP-TC

Extracts enriched by a specific protein are rarely easily obtained from natural host cells. Hence, recombinant protein production is frequently the sole applicable procedure (Sorensen and Mortensen, 2005). As the biotechnology industry has rapidly expanded in recent years, the expression of a spectrum of recombinant proteins in different systems has been a major feature and challenge (Rosano and Ceccarelli, 2014). The predominant expression system for production of heterologous proteins is the *Escherichia coli*, whose genetics is better understood than that of any other microorganism (Demain and Vaishnav, 2009). Its genome can be easily and precisely modified, allowing rapid protein expression and high yields: *E. coli* can accumulate recombinant proteins up to 80% of its dry weight, and survives a variety of environmental conditions (Demain and Vaishnav, 2009; Gopal and Kumar, 2013).

In order to perform assays to determine NEP-TC's specific activity and kinetics, a large quantity of the protein was required. Since extracts from EC and NEC only allow the obtainment of a low protein yield (Correia, 2011) and the recombinant protein

presents a His-tag that allows an easier purification, the expression of recombinant NEP-TC on a heterologous system was the preferred approach. Thus, an expression vector with NEP-TC's gene was cloned on a BL-21 strain of *E. coli*, its expression was induced, and the purified protein was obtained after several isolation steps.

Analysing the purified protein fractions, it is clear that recombinant NEP-TC's expression and purification was successful, because it matches the 25 kDa strips. However, it is clear that there is some protein aggregation in the first purified fractions, for it is possible to see some faint bands in the 50 kDa and 75 kDa strips, and this phenomenon occurs because NEP-TC tends to aggregate when outside the cytosol. This occurrence is supported by the available literature, because RNA methyltransferases from the SpoU family are reported to be functional in their multimeric form (Motorin and Helm, 2011).

Considering the final protein concentration of the expression process, it was already mentioned in the results that quantification by Bradford assays revealed a final concentration of 0.1365 mg/mL in the concentrated fractions, which had 15 mL. Therefore, it is possible to infer that 500 mL of expression culture yielded approximately 2.048 mg of protein, which corresponds to 4.1 mg of protein per litre of expression culture. Even though several recombinant rRNA methyltransferases have been expressed in heterologous systems, the data concerning the protein yield of the whole process are usually not mentioned in the papers' results. Agarwalla and co-workers (2002) reported a successful expression of a recombinant rRNA methyltransferase in *E. coli* with a final concentration of 2.5 mg of purified protein per litre of expression culture. Comparing these reference values with the ones obtained for NEP-TC, it is clear that the whole optimized process used for the expression of this protein is more advantageous, as it permitted the obtainment of a considerably higher

concentration of purified protein. Considering this process' high yield, the present methodology used for NEP-TC's expression can continue to be used for further characterization of this enzyme.

Concerning NEP-TC's specificity, the only available data were reported by Correia (2011): firstly, Faro and co-workers (2003) isolated a cDNA corresponding to the gene that encodes NEP-TC from non-embriogenic *calli* and sequenced it. Later, the isolated sequence was compared by bioinformatic approaches with sequences present the plant model *A. thaliana*, showing a high degree of similarity (76%) with tRNA/rRNA methyltransferases from the SpoU family (Correia, 2011). Thus, practical approaches were adopted in this work with the purpose of proving that NEP-TC is in fact a RNA methyltransferase.

Methyltransferase activity assays with the recombinant protein were firstly performed with RNA and DNA extracted from NEC, in order to eliminate DNA as a potential substrate. The results show that NEP-TC has specific activity towards RNA, while showing negative activity values when the substrate is DNA. These results confirmed that this protein is a RNA methyltransferase. Nonetheless, evidence were still needed about which fraction of the total RNA this protein methylates. Therefore, based on the predictions made by bioinformatics analysis, rRNA was the selected substrate. rRNA and ribosomes were extracted from NEC and used as a substrate in a specific activity assay. Both substrates were methylated by the recombinant protein, with rRNA showing consistently higher values. There are already characterized RNA methyltransferases in *E. coli* (that the bioinformatic assays determined as similar to NEP-TC) that present a similar behavior: YbeA is an rRNA methyltransferase active as a homodimer that docks into the ribosomal subunits, being specific for both substrates and functioning as a quality marker for the translational initiation (Purta *et al*., 2008).

Moreover, the specific activity assays allowed the determination of the correct rRNA concentrations that must be used when performing kinetic assays: as seen on the results, concentrations above 0.011 mM cause a quick saturation of the reaction, while for concentrations between 0.0014 mM and 0.0055 mM the reaction reaches its ideal rate. Nonetheless, even lower concentrations of the substrate should be used, once they allow a broader understanding of the protein's activity range.

The protein specificity results were confirmed by a thermophoretical assay, in which the protein showed affinity towards a sample of total RNA. Nonetheless, there was some irregular binding of the protein to the substrate, since there are some points that do not follow the fluorescence curve. This phenomenon might be justified by the fact that the substrate sample was heterogeneous: the protein shows affinity towards one of RNA's fractions, and was binding to each of them separately until binding to the right one, which was assumed to be rRNA. Therefore, this procedure must be optimized and repeated with lower concentrations of the right substrate to reach the unbound plateau. It was only possible to determine that NEP-TC has a K_D in the nM range. Furthermore, the results obtained with the control ligand Tafa indicate that the dimerization/oligomerization of NEP-TC might influence the results, so measures to ensure that the protein is in its monomeric form should be taken into account.

Even though it was not possible to perform an extensive characterization of the protein (determination of its K_M, K_D , K_{cat} and Vmax), the enzymatic assays performed allowed the identification of the correct substrate for this enzyme, and specific activity results with different rRNA helped to determine the precise concentrations to be used in further kinetic studies. This findings will also help further thermophoresis assays, for the correct substrate and its concentrations will be considered, as well as the protein's structure.

Regarding this proteins' function, although several rRNA methyltransferases have been identified in plants, there is virtually no information about the physiological role of these enzymes. However, in bacteria and yeast, they are frequently expressed as a response to abiotic stress, like nutrient restriction or exposure to antibiotics (Liu *et al.*, 2008; Sharma *et al*., 2013).

Abiotic stress responses are important for sessile organisms such as plants because they cannot survive without coping with the changes in the environment. The term "abiotic stress" comprises stresses caused by complex environmental conditions, such as UV, high/low temperatures, osmotic changes, heavy metals and hypoxia (Hirayama and Shinozak, 2010). These stress situations trigger signals that promote alterations in the plant's physiology and growth, to ensure survival in hostile environments. While little is known about the primary receptors that sense these stresses, it is known that the plant's response to these factors is accompanied by several downstream signalling cascades that ultimately promote massive changes in the transcriptome (Zeller *et al*., 2009). Recent studies have shown that biotic and abiotic stresses cause long-term regulation of gene expression, mostly thanks to epigenetic gene regulatory mechanisms, such as chromatin remodelling through histone modification or DNA methylation (Kim *et al*., 2015).

When it comes to RNA methylation, the information is scarce. Studies on *A. thaliana* show that a tRNA methyltrasferase production is enhanced during the course of the plant's infection with the bacterial pathogen *Pseudomonas syringae*, as a response to this stress factor (Ramirez *et al*., 2015). Also in this plant, an uncharacterized RNA methyltransferase has been identified as expressed when the plants are submitted to UV stress, and loss of function of this enzyme causes

hypersensitivity: thus, it is involved in the plant's adaptation to extreme conditions (Piofczyk *et al*., 2015).

The induction of SE is a process which conditions comprise several stress factors to the plant cell. NEP-TC is an RNA methyltransferase and, as mentioned before, these proteins are usually expressed when the cells undergo extreme conditions, and are part of the transcriptome reprograming of the cell as a response to stress. Since the induction of SE in tamarillo is a process in which the explants are exposed to osmotic stress, wounding and contact with picloram (an herbicide), it can be hypothesised that this protein might have influence in the cell's response to stress conditions and further acquisition of embryogenic competence.

Recently, comparative proteomic studies have been performed in order to determine which classes of proteins are expressed in the induction of SE and the differences between EC and NEC's proteomes. Souza-Reis and co-workers (2016) identified several stress-related proteins expressed in the induction of SE in sugarcane (*Saccharum spp*.), including peroxidases (hydrogen peroxide has been described as a SE inducer, triggering a signalling cascade and promoting the expression of stress-related genes), heat-shock proteins, glutathione-s-trasferases (that have antioxidant and detoxifying properties) and 14-3-3 proteins. This last group of proteins has the ability to bind to a multitude of functionally diverse signalling proteins, being involved in the plant's systemic response to both biotic and abiotic stress (Denison *et al*. 2011). Also in sugarcane, Heringer and co-workers (2015) compared the proteomic profiles of both EC and NEC, identifying 14 exclusive proteins in EC, most of them highly related to stress: the expression of these proteins suggests that stress experienced by the explants may be responsible for the physiological modulation of the cells, thereby allowing the acquisition of embryogenic competence. NEC presented 10 exclusive proteins, related

to protein degradation, which is indicative of callus with low metabolic activity and, consequently, a low level of cell differentiation that possibly prevent the development of somatic embryos.

In maize (*Zea mays*), Varhanikova and co-workers (2014) identified stress related antioxidant enzymes in EC, like ascorbate peroxidase, responsible for maintaining the intracellular levels of H_2O_2. Also, they identified the presence of lipoxygenases, enzymes that are related to the plant's response to biotic/abiotic stress, as well as wounding (Yan *et al*., 2013). Therefore, it can be concluded that these proteins are relevant for cellular reprogramming in stress conditions.

In oil palm (*Elaeis guineensis*), a proteomic identification of differentially expressed proteins during the acquisition of somatic embryogenesis, showed the presence three main groups during the induction phase: proteins related to cellular proliferation, energy production and stress - the last group being the predominant one (Silva *et al*., 2014). These results show that the knowledge of the expression pattern of different stress-associated proteins during *in vitro* embryogenesis is important since stress itself has a crucial role as a trigger of the somatic embryogenesis process.

The assumption that NEP-TC might be related to stress is supported by this protein's own structure: it has an alpha-beta-knot, ubiquitous in globular homopolymers, and conserved throughout evolution, suggesting that they play a crucial role in the proteins' enzymatic activity and binding (Virnau *et al*. 2007). Furthermore, some literature data suggest that a potential functional role of this unusual knotted topology is to increase resistance against degradation (Virnau *et al*., 2007), increase thermal and mechanical stability (Taylor and Lin, 2003; King *et al*., 2007; Sulkowska *et al*., 2009) or even alter enzymatic activity (Alam *et al*., 2002) when the proteins are submitted to extreme conditions.

5. Conclusions and Future Perspectives

Somatic embryogenesis is a relevant biotechnological tool for *in vitro* culture and improvement of plants, but also an important model for fundamental studies of physiology and biochemistry of plant embryo development.

The advance of high-throughput protein analysis allowed the identification and characterization of proteins expressed in plant tissues, and interest has been focused on proteins involved in the induction of SE, and the potential role that these enzymes might play in the reprograming of the cells' onthogenic program. The present work was a step forward in the characterization of one of this protein: NEP-TC (*Non-embryogenic protein of tamarillo callus*).

Previous studies made in order to characterize this protein involved three main approaches. Firstly, bioinformatic studies were performed and revealed that it is a putative tRNA/rRNA methyltransferase, for its sequence presents high identity with one that encodes a tRNA/rRNA methyltransferase in *Arabidopsis thaliana.* Immunohistochemistry approaches were used to determine the protein's subcellular localization in embryogenic and non-embryogenic *calli*. This task was not fully accomplished, but it was determined that NEP-TC probably has a cytoplasmic expression near the nucleus. Finally, enzymatic activity assays were performed to confirm NEP-TC's role as a RNA methyltransferase. The results were not conclusive, since the protein used was from extracts and it didn't have a high level of purity. Nonetheless, the reduced concentration of protein allowed the quantification of some methyltransferase activity towards total RNA extracts.

Therefore, the aim of this study was to improve the methodologies used in NEP-TC's characterization.

Regarding the bioinformatics approaches, the first factor studied was NEP-TC's phylogeny, in order to understand in which plant species other genes similar to this one are expressed. This gene belongs to a gene family that has been suffering speciation over time, and its orthologous are expressed in 29 species of land plants and two green algae species. These results allow some insight into the evolutionary history of plant methyltransferases, which might help the unveiling of these enzymes' function in the organism.

Also, the protein's structure was determined and analysed using online protein databases. It was determined that NEP-TC has a globular shape and the alpha-beta knot characteristic of RNA methyltransferases. Furthermore, literature data indicate that this knot is essential in the protein's survival in extreme conditions, thus providing one explanation to the fact that NEP-TC is expressed when somatic embryogenesis is induced. Nonetheless, structural and functional studies can be optimized. As was discussed, the recombinant protein expression system allowed the production of a high quantity of purified protein. Therefore, one approach that can be used to determine experimentally the protein's structure is X-ray crystallography.

When it comes to immunohistochemistry, not only NEC and EC were used, but also explants with different levels of dedifferentiation, in order to determine the variances in the protein's expression throughout the induction phase. The revelation of the protein's labelling in tissues using specific antibodies showed that NEP-TC's expression is indeed higher in NEC than in EC, and it is also expressed in the first weeks of proembryogenic development, decreasing after the 8^{th} week. These results complemented data presented in previous studies, that only contemplated EC and NEC and immunoblot analysis performed with extracts. Another potentially important finding of this work, although outside the original objectives, is that the *calli* present

autofluorescence. The identification of the factors that potentiate this phenomenon could prove to be interesting.

The third part of this project consisted on the expression of highly purified recombinant NEP-TC for methyltransferase activity assays. The expression was extremely successful, for it allowed the production of high concentrations of recombinant protein. Thus, this work provides an optimized expression system for recombinant NEP-TC that confers high yields, and can be used in further assays that require high amounts of protein (like X-ray crystallography, already mentioned).

The methyltransferase activity results performed with RNA and its fractions confirmed that this protein is in fact an RNA methyltransferase, with specificity to rRNA. As far as it is concerned, these results are pioneer, because there are few rRNA methyltrasferases identified in plants.

The optimization of the conditions for the methyltransferase activity assays will allow an easier determination of NEP-TC's kinetic parameters, essential for an even more extensive characterization of this protein. Additionally, further termophoretic studies should be optimized and conducted in order to determine other aspects of the protein's kinetics, using the correct substrate and taking into account the protein's conformation. The compilation of these results might provide some insight about the role of NEP-TC in the somatic embryogenesis process.

Overall, the results presented here contribute to a better understanding of some biochemical aspects underlying somatic embryogenesis of tamarillo. Even though the data here presented do not entirely describe the full role and importance of NEP-TC in this process, they confirm some data available in the literature and lay methodological foundations for future studies regarding this phenomenon.

6. References

Agarwalla, S., Kealey, J., Santi, D., Stroud, R. (2002) "Characterization of the 23 S Ribosomal RNA m5U1939 methyltransferase from *Escherichia coli"*. The Journal of Biological Chemistry 277: 8835-8840.

Alam, M., Yamada, T., Carlsson, U., Ikai, A. (2002) "The importance of being knotted: effects of the C-terminal knot structure on enzymatic and mechanical properties of bovine carbonic anhydrase". FEBS Letters 519: 35-40.

Arnold, F. (1991) "Metal-affinity separations: a new dimension in protein processing". Biotechnology 9: 151-156.

Atiquah, N., Maisarah, A., Ahmah, R. (2014) "Comparison of antioxidant properties of tamarillo (*Cyphomandra betacea*), cherry tomato (*Solanum lycopersicum* var. *cerasiform*) and tomato (*Lyopersicon esulentum*)". International Food Research Journal 21: 2355-2362.

Baillieu, A., Fabret, C., Liang, X., Piekna, D., Fournier, M., Rousset, J. (2009) "Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy". Nucleic Acids Research 37: 7665-7677.

Basturea, G., Deutscher, M. (2007) "Substrate specificity and properties of the *Escherichia coli* 16S rRNA methyltransferase, RsmE". RNA 13: 1969-1976.

Bauerle, M., Schwalm, E., Booker, S. (2015) "Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation". The Journal of Biological Chemistry 290: 3995-4002.

References

Bhawani, S., Fong, S., Ibrahim, M. (2015) “Spectrophotometric Analysis of Caffeine”. International Journal of Analytical Chemistry doi: 10.1155/2015/170239 (published online).

Blanchard, S., Puglisi, J. (2001) “Solution structure of the A loop of 23S ribosomal RNA”. Proceedings of the National Academy of Sciences 98: 3720-3725.

Bohs, L. (1989) “Ethnobotany of the genus *Cyphomandra* (Solanaceae)”. Economic Botany 43:2 143-163.

Boulter, D., Parthier, B. (1982) “Nucleic acids and Proteins in Plants – Structure, biochemistry and physiology of proteins” *in* Encyclopedia of Plant Physiology ISBN 978-3-642-68239-1.

Bourgeois, G., Ney, M., Gaspar, I., Aigueperse, C., Schaefer, M., Kellner, S., Helm, M., Motorin, Y. (2015) “Eukaryotic rRNA modification by yeast 5-methylcytosine methyltransferases and human proliferation-associated antigen p120”. PLoS ONE 10 (7).

Bradford, K. (2008) “Shang Fa Yang: Pioneer in plant ethylene biochemistry”. Plant Science 175: 2-7.

Bradford, M. (1976) “Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding”. Analytical Biochemistry 72: 248-254.

Canhoto J. M. (2010) "Biotecnologia Vegetal, da clonagem de plantas à transformação genética", Imprensa da Universidade de Coimbra

Canhoto J.M., Lopes, M.L., Cruz, G.S., (2005) "Protocol for somatic embryogenesis: tamarillo (*Cyphomandra betacea* (Cav.) Sendt.)". In: Jain, S.M., Gupta, P.K. (Eds.), Protocol for Somatic Embryogenesis in Woody Plants. Springer, Dordrecht, pp. 379–389.

Cantara, W., Crain, P., Rozenski, J., McCloskey, J., Harris, K., Zhangs, X., Vendeix, F., Fabris, D., Agris, P. (2011) "The RNA modification database, RNAMDB: 2011 update". Nucleic Acids Research, 39:195-201.

Casimiro, B. (2014) "Análise do envolvimento da proteína NEP-TC na embriogénese somática de tamarilho (*Cyphomandra betacea*)". Masters Thesis, Universidade de Coimbra, Portugal.

Casson, S., Spencer, M., Walker, K., Lindsey, K. (2005) "Laser capture microdissection for the analysis of gene expression during embryogenesis of *Arabidopsis*". The Plant Journal 42: 111-123.

Chaga, G. (2001) "Twenty-five years of immobilized metal ion affinity chromatography: past, present and future". Journal of Biochemical and Biophysical Methods 49: 313-334.

Chart, H., Smith, H., Ragione, R., Woodward, M. (2000) "An investigation into the pathogenic properties of *Escherichia coli* strains BLR, BL21, DH5a and EQ1". Journal of Applied Microbiology, 89: 1048-1058.

Cohen, D., Elliot, D. (1979) "Micropropagation methods for blueberries and tamarillos". Combined Proceedings. International Plant Propagators Society 29: 177-179.

References

Correa, A., Opezzo, P. (2011) “Tuning different expression parameters to achieve soluble recombinant proteins in *E. coli*: advantages of highthroughput screening”. Biotechnology Journal 6: 715-730.

Correia S. I, (2011) “Somatic embryogenesis in *Cyphomandra betacea* (Cav.) Sendt (tamarillo): optimization and molecular analysis” PhD Thesis, Universidade de Coimbra, Portugal.

Correia S. I., Canhoto J. M. (2012) “Biotechnology of tamarillo (*Cyphomandra betacea*): From in vitro cloning to genetic transformation” Scientia Horticulturae, 148: 161–168.

Correia S., Vinhas, R., Manadas, B., Lourenço, A. S., Veríssimo, P., Canhoto, J. M. (2012) “Comparative proteomic analysis of auxin-Induced ambryogenic and nonembryogenic tissues of the solanaceous tree *Cyphomandra betacea* (tamarillo)” Journal of Proteome Research, 11:1666−1675.

Correia, S., Alves, A., Veríssimo, P., Canhoto, J. (2016) “Somatic Embryogenesis in Braod-Leaf Woody Plants: What We Can Learn from Proteomics”. *In*: Germana, M. and Lambardi, M. (Eds.) *In vitro* Embryogenesis of Higher Plants. Springer. Science+Business Media, New York, 1359:117-130.

Correia, S., Pinto, G., Canhoto, J. (2016) “Molecular biology of somatic embryogenesis in hardwoods”. *In*: Yill-Sung Park, Jan M Bonga, Heung-Kyu Moon (Eds.) Vegetative Propagation of Forest Trees. National Institute of Forest Science (NIFoS). Seoul, Korea. pp 97-122.

Correia, S.I., Cunha, A.E., Salgueiro, L., Canhoto, J.M., (2012) “Somatic embryogenesis in tamarillo (*Cyphomandra betacea*): approaches to increase efficiency

of embryo formation and plant development". Plant Cell Tissue and Organ Culture, 109: 143–152.

Cowling, V. (2010) "Regulation of mRNA cap methylation". Biochemical Journal 425: 295-302.

Czerwoniec, A., Kasprzak, J.M, Kaminska, K.H., Rother, K., Purta, E., Bujnicki, J.M (2011) "RNA nucleothyde methylation". In: Grosjean H (ed.) "DNA and RNA Modification Enzymes: Comparative Structure, Mechanism, Functions, Cellular Interactions and Evolution", 289–302.

Decatur, W., Fournier, M. (2003) "RNA-guided nucleotide modification of ribosomal and other RNAs". Journal of Biological Chemistry 278: 695-698.

Denison, F., Paul, A., Zupanksa, A., Frejl, R. (2011) "14-3-3 proteins in plant physiology". Seminars in Cell and Developmental Biology 22: 720-727.

Diener, A., Li, H., Zhou, W., Whoriskey, J., Nes, W., Fink, G. (2000) "Sterol metyltransferase 1 controls the level of cholesterol in plants". Plant Cell 12: 853-870.

Duhr, S., Braun, D. (2006) "Why molecules move along a temperature gradient". Proceedings of the National Academy of Science 103: 19678–19682.e0133321.

Faro, R., Lopes, M., Canhoto, J., Cheung, A., Cruz, S. (2003) "Identification and molecular characterization of a non-embryogenic *calli* protein in *Cyphomandra betacea*". 7th International Botanical Meeting – Plant Cell Biology. Lisbon, Portugal.

Feder, M., Pas, J., Wyrwicz, L., Bujnicki, J. (2003) "Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2'-O-methyltransferases". Gene 302: 129-138.

Feher, A. (2005) “Why somatic plant cells start to form embryos?”. Plant Cell Monographs 2: 85-101.

Feher, A. (2008) “The initiation phase of somatic embryogenesis: what we know and what we don’t”. Acta Biologica Szegediensis 52: 53-56.

Feher, A. (2015) “Somatic embryogenesis – stress induced remodeling of plant cell fate”. Biochimica et Biophysica Acta 1849: 385-402.

Fetcher, P., Brownlee, G. (2005) “Recognition of mRNA cap structures by viral and cellular proteins”. Journal of General Virology 86: 1239-1249.

Galland, R., Blervacq, A., Blassiau, C., Smagghe, B., Decottignies, J., Hilbert, J. (2007) “Glutathione-S-Transferase is detected during somatic embryogenesis in chicory”. Plant Signaling and Behaviour 2: 343-348.

Gouy, M., Guindon, S., Gascuel, O. (2010) “SeaView version 4 : a multiplatform graphical user interface for sequence alignment and phylogenetic tree building”. Molecular Biology and Evolution 27: 221-224

Gopal, G., Kumar, A. (2013) “Strategies for the Production of Recombinant Protein in *Escherichia coli*”. Protein Journal, 32: 419-425.

Graves, T., Zhang, Y., Scott, J. (2013) “A universal competitive fluorescence polarization activity assay for *S*-adenosylmethionine utilizing methyltransferases”. Analytical Biochemistry 373: 296-306.

Grossman, T., Kawasaki, E., Punreddy, S., Osburne, M. (1998) “Spontaneous cAMP-dependent repression of gene expression in stationary phase plays a role in recombinant expression instability”. Gene 209: 95-103.

Guimarães M.L., Tomé, M.C., Cruz, G.S., (1996) "*Cyphomandra betacea* (Cav.) Sendt. (tamarillo)" In: Bajaj, Y.P.S. (Ed.), "Biotechnology in Agriculture and Forestry" vol. 35, Trees IV. Springer Verlag, Berlin, pp. 120–137.

Guimarães, M.L., Cruz, G.S., Montezuma-de-Carvalho, J.M., (1988) "Somatic embryogenesis and plant regeneration in *Cyphomandra betacea* (Cav.) Sendt" Plant Cell Tissue and Organ Culture, 15: 161–167.

Haiser, H., Karginov, F., Hannon, G., Elliot, M. (2008) "Developmentally regulated cleavage of tRNAs in the bacterium *Streptomyces coelicolor*". Nucleic Acids Research 36: 732-741.

Hassan, S., Bakar, M. (2013) "Antioxidative and Anticholinesterase Activity of *Cyphomandra betacea* Fruit". The Scientific World Journal Volume 2013.

Held, M., Be, E., Zemelis, S., Withers, S., Wilkerson, C., Brandizzi, F. (2011) "CGR3: A Golgi-Localized protein influencing homogalacturonan methylesterification". Molecular Plant 4: 834-844.

Heringer, A., Barroso, T., Macedo, A., Santa-Catarina, C., Souza, G., Floh E. (2015) "Label-Free quantitative proteomics of embryogenic and non-Embryogenic callus during sugarcane somatic embryogenesis". PLoS ONE 10(6):e0127803. doi:10.1371/journal.pone.0127803.

Hirayama, T., Shinozaki, K. (2010) "Research on plant abiotic stress responses in the post-genome era: past, present and future". The Plant Journal 61: 1041-1052.

Hori, H. (2014) "Methylated nucleosides in tRNA and tRNA methyltrasferases". Frontiers in Genetics (published online).

Huang, B., Lu, J., Byström, A. (2008) "A genome-wide screen identifies genes required for formation of the wobble nucleoside 5-methoxycarbonylmethyl-2-thiouridine in *Saccharomyces cerevisiae*". RNA 14: 2183-2194.

Hurtado, N., Morales, A., Miret, M., Gilete, M., Heredia, F. (2009) "Colour, pH stability and antioxidant activity of anthocyanin rutinosides isolated from tamarillo fruit (*Solanum betaceum* Cav.)". Food Chemistry 117: 88-93.

Ikeda, M., Umehara, M., Kamada, H. (2006) "Embryogenesis-related genes: its expression and roles during somatic and zygotic embryogenesis in carrot and *Arabidopsis"*. Plant Biotechnology, 23: 153–161.

Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H., Duhr, S., Baaske, P., Breitsprecher, D. (2014) "MicroScale Thermophoresis: Interaction analysis and beyond". Journal of Molecular Structure 1077: 101-113.

Jerabek-Willemsen, M., Wienken, C., Braun, D., Baaske, P., Duhr, S. (2011) "Molecular Interaction Studies Using Microscale Thermophoresis". Assay and Drug Development Technologies 9: 342-353.

Jiménez, V.M. (2001) "Regulation of *in vitro* somatic embryogenesis with emphasis on to the role of endogenous hormones". Revista Brasileira de Fisiologia Vegetal 13: 196–223.

Jiménez, V.M. (2005) "Involvement of plant hormones and plant growth regulators on *in vitro* somatic embryogenesis". Plant Growth Regulation 47: 91–110.

Johanson, M., Byström, A. (2005) “Transfer RNA modifications and modifying enzymes in Saccharomyces cerevisiae”. Fine Tuning of RNA Functions by Modification and Editing 12: 87-120.

Jones, L., Scott, T. (1981) “Transfer ribonucleic acid modification and its relationship to tumorous and nontumorous plant growth”. Plant Physiology 67:535-538.

Kapp, L., Lorsch, J. (2004) “The molecular mechanics of eukaryotic translation”. Annual Review of Biochemistry 73: 657-704.

Kellner, S., Burhenne, J., Helm, M. (2010) “Detection of RNA modifications”. RNA Biology, 7: 237-247.

Kim, J., Sasaki, T., Ueda, M., Sako, K., Seki, M. (2015) “Chromatin changes in response to drought, salinity, heat, and cold stresses in plants”. Frontiers in Plant Science 6: 1-12.

King, N., Yeates, E., Yeates, T. (2007) “Identification of rare slipknots in proteins and their implications for stability and folding”. Journal of Molecular Biology 373: 153-166.

Knoop, V. (2011) “When you can’t trust the DNA: RNA editing changes transcript sequences”. Cellular and Molecular Life Sciences 68: 567-586.

Koukalova, B., Fojtova, M., Lim, K., Fulnecek, J., Leitch, A., Kovarik, A. (2005) “Dedifferentiation of tobacco cells is associated with ribosomal RNA gene hypomethylation, increased transcription, and chromatin alterations”. Plant Physiology 139: 275–286.

Kozbial, P., Meshegian, A. (2005) “Natural history of S-adenosylmethionine-bindig proteins”. BMC Structural Biology 5: 19-45.

Krupkova, E., Immerzeel, P., Pauly, M., Schmulling, T. (2007) “The TUMOROUS SHOOT DEVELOPMENT2 gene of *Arabidopsis* encoding a putative methyltransferase is required for cell adhesion and co-ordinated plant development”. The Plant Journal 50: 735-750.

Leljak-Levanic, D., Mihaljevic, S., Bauer, N., Jelaska, S. (2004) “Changes in DNA methylation during somatic embryogenesis in *Cucurbita pepo L.*”. Plant Cell Reports 23: 120-127.

Lim, K., Zhang, H., Tempczyk, A., Krajewski, W., Bonander, N., Toedj, J., Howard, A., Eisenstein, E., Herzberg, O. (2003) “Structure of the YibK methyltransferase from *Haemophilus influenzae* (HI0766): a cofactor bound at a site formed by a knot”. Proteins: Structure, Function and Bioinformatics 51 (1): 56-67.

Liu, B., Liang, X., Prybylska, D., Liu, Q., Fournier, M. (2008) “Mis-targeted methylation in rRNA can severely impair ribosome synthesis and activity”. RNA Biology 5: 249-254.

Liu, J., Hayashi, K., Matsuoka, K. (2015) “Membrane topology of Golgi-localized probable S-adenosylmethionine-dependent methyltransferase in tobacco (*Nicotiana tabacum*) BY-2 cells”. Bioscience, Biotechnology and Biochemistry 79: 2007-2013.

Loenen, W. (2006) “S-Adenosylmethionine: jack of all trades and master of everything?”. Biochemical Society Transactions, 34: 330-335.

Lopes M., Ferreira M., Carloto J., Cruz, G. Canhoto J. (2000) “Somatic embryogenesis induction in tamarillo (*Cyphomandra betacea*)”. In: Jain, S., Gupta, P., Newton, R. (Eds.), “Somatic Embryogenesis in Woody Plants”, vol.6. Kluwer Academic Publishers, Dordrecht, 433–455.

Lost, I., Guillerez, J., Dreyfus, M. (1992) “Bacteriophage T7 RNA polymerase travels far ahead of ribosomes in vivo”. Journal of Bacteriology 174: 619-622.

Machinicka, M., Milanowska, K., Oglou, O., Purta, E., Kurkowska, M., Olchowik, A., Januszewski, W., Kalinowski, S. Horkawickz, S., Rother, K., Helm, M., Bujnicki, J., Grosiean, H. (2013) “MODOMICS: a database of RNA modification pathways – 2013 update”. Nucleic Acids Research, 41.

Meadows, L.R. (2002) “Growing tamarillo relatives”. In: The New Zealand Home Garden. www.naturalhub.com.

Merali, S., Clarkson, A. (2004) “S-adenosylmethionine and *Pneumocystis*”. FEMS Microbiology Letters 237: 179-186.

Mertz, C., Brat, P., Veyrat, C., Gunata, Z. (2010) “Characterization and thermal lability of carotenoids and vitamin C of tamarillo fruit (*Solanum betaceum* Cav.)”. Food Chemistry 119: 653-659.

Mertz, C., Gancel, A., Gunata, Z., Alter,P., Mayer, C., Vaillant, F., Perez, A., Ruales, J., Brat, P. (2009) “Phenolic compounds, carotenoids and antioxidant activity of three tropical fruits”. Journal of Food Composition and Analysis 22: 381-387.

Moffatt, B., Weretilnyk, E. (2001) “Sustaining S-adenosyl-L-methionine-dependent methyltransferase activity in plant cells”. Physiologia Plantarum 113: 435-442.

Motorin, Y., Helm, M. (2010) “tRNA stabilization by modified nucleotides”. Biochemistry 49: 4934-4944.

Motorin, Y., Helm, M. (2011) “RNA nucleotide methylation”. WIREs RNA. DOI: 10.1002/wrna.79.

Nascimento, G., Hamm, L., Baggio, C., Werner, M., Iacomini, M., Cordeiro, L. (2013) "Structure of a galactoarabinoglucuronoxylan from tamarillo (*Solanum betaceum*), a tropical exotic fruit, and its biological activity". Food Chemistry 141: 510-516.

Obando, M., Jordan, M. (2001) "Regenerative responses of *Cyphomandra betacea* (Cav.) Sendt. (tamarillo) cultivated *in vitro"*. Acta Horticulturae 560: 429-432.

Ohbayashi, I., Konishi, M., Ebine, K., Sugiyama, M. (2011) "Genetic identification of *Arabidopsis* RID2 as an essential factor involved in pre-rRNA processing". The Plant Journal 67: 49–60.

Osorio, C., Hurtado, N., Dawid, C., Hofmann, T., Mira, F., Morales, A. (2012) "Chemical characterisation of anthocyanins in tamarillo (*Solanum betaceum* Cav.) and Andes berry (*Rubus glaucus* Benth.) fruits". Food Chemistry 132: 1915-1921.

Paciorek, T., Sauer, M., Balla, J., Wisnievska, J., Friml, J. (2006) "Immunocytochemical technique for protein localization in sections of plant tissues". Nature Protocols 1: 104-107.

Pasternak, T., Prinsen, E., Ayaydin, F., Miskolczi, P., Potters, G., Asard, H., Onckelen, H., Dudits, D., Feher, A. (2002) "The role of auxin, pH and stress in the activation of embryogenic cell division in leaf protoplast derived cells of alfalfa". Plant Physiology 129: 1807-1819.

Phue, J., Lee, S., Trinh, L., Shiloach, J. (2008) "Modified *Escherichia coli* B (BL21), a Superior Producer of Plasmid DNA Compared With *Escherichia coli* K (DH5a)". Biotechnology and Bioengeneering, 101:4.

Piofczyk, T., Jeena, G., Pecinka, A. (2015) "*Arabidopsis thaliana* natural variation reveals connections between UV radiation stress and plant pathogen-like defense responses". Plant Physiology and Biochemistry 93: 34-43.

Porath, J. (1992) "Immobilized metal ion affinity chromatography". Protein Expression and Purifications 3: 263-281.

Pringle, G., Murray, B. (1991) "Reproductive biology of the tamarillo, *Cyphomandra betacea* (Cav.) Sendt. (Solanaceae), and some wild relatives". New Zealand Journal of Crop and Horticultural Science 19: 263-273.

Prohens J., Nuez F. (2001) "The Tamarillo (*Cyphomandra betacea*)". Small Fruits Review, 1: 43-68.

Proost, S., Van Bel, M., Vaneecchoutte, D., Van de Peer, Y., Inze, D., Mueller-Roeber, B., Vandepoele, K. (2015) "PLAZA 3.0: an access point for plant comparative genomics". Nucleic Acids Research 43 (database issue)

Purta, E., Kaminska, K., Kasprzak, J., Bujnicki, J., Douthwaite, S. (2008) "YbeA is the m3C methyltransferase RlmH that targets nucleotide 1915 in 23S rRNA". RNA 14: 2234-2244.

Quiroz-Figueroa, F.R., Rojas-Herrera R., Galaz-Avalos, R.M., Loyola-Vargas, V.M. (2006) "Embryo production through somatic embryogenesis can be used to study cell differentiation in plants". Plant Cell Tissue and Organ Culture, 86: 285–301.

Ramirez, V. Gonzales, B., Lopez, A., Castello, M., Gil, M. Zheng, B., Chen, P. Vera, P. (2015) "Loss of a conserved tRNA anticodon modification perturbs plant immunity". PLOS Genetics.

Reinert, J. (1959) “Über die kontrolle der morphogenese und die induktion von adventivembryonen in gewebekulturen aus karotten”. Planta 53: 318–333.

Richter, U., Ku, K., Okada, S., Brennicke, A., Weihe, A., Borner, T. (2010) “A mitochondrial rRNA dimethyladenosine methyltransferase in *Arabidopsis*”. The Plant Journal 61: 558-569.

Rivera, M., Maguire, B., Lake, J. (2015) “Isolation of ribosomes and polysomes”. Cold Spring Harbor Protocols 3: 293-299.

Rose, R., Nolan, K. (2006) “Genetic regulation of somatic embryogenesis with particular reference to *Arabidopsis thaliana* and *Medicago truncatula*”. In Vitro Cellular and Developmental Biology 42: 473-481.

Ruhfel, B., Gitzendanner, M., Soltis, P., Soltis, D., Burgeligh, G. (2014) “From algae to angiosperms – inferring the phylogeny of green plants (*Viridiplantae*) from 360 plastid genomes”. BMC Evolutionary Biology 14: 23-50.

Sauer, M., Friml, J. (2010) “Immunolocalization of proteins in plants”. Methods in Molecular Biology 655: 253-263.

Sauer, M., Paciorek, T., Benkova, E., Friml, J. (2006) “Immunocytochemical techniques for whole-mount *in situ* protein localization in plants”. Nature Protocols 1: 98-103.

Saze, H., Kazuo, T., Kanno, T., Nishimura, T. (2012) “DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation”. Plant and Cell Physiology 53: 766-784.

Schiefelbein, J., Kwak, H., Wieckowski, Y., Barron, C., Bruex, A. (2009) "The gene regulatory network for root epidermal cell-type pattern formation in *Arabidopsis*". Journal of Experimental Botany 60: 1515–1521.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W. (2012) "NIH Image to ImageJ: 25 years of image analysis". Nature Methods 9: 671-675.

Seidel, S., Dijkman, P., Lea, W., Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J., Srinivasan, P., Baaske, P. (2013) "Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions". Methods 59: 201-315.

Seo, H., Song, J., Cheong, J., Lee, Y., Lee, J., Hwang, I., Choy, Y. (2001) "Jasmonic acid carboxyl methyltransferase: a key enzyme for jasmonate-regulated plant responses". Proceedings of the National Academy of Sciences USA 98: 4788-4793.

Shao, Z., Yan, W., Peng, J., Zou, Y., Li, F. (2013) "Crystalstructure of tRNA methyltransferase Trm10: insight on the catalytic mechanism and recognition of tRNA substract". Nucleic Acids Research 42: 509-525.

Sharma, S., Yang, J., Duttman, S., Watzinger, P., Kotter, P., Entian, K. (2014) "Identification of novel methyltransferases, Bmt5 and Bmt6, responsible for the m3U methylations of 25S rRNA in *Saccharomyces cerevisiae*". Nucleic Acids Research 42: 3246–3260.

Sharma, S., Yang, J., Watzinger, P., Kotter, P., Entian, K. (2013) "Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively". Nucleic Acids Research 41: 9062-9076.

Shatkin, A., Manley, J. (2000) “The ends of the affair: capping and polyadenylation”. Nature Structure and Molecular Biology 7: 838-842.

Shinohara, N., Ohbayashi, I., Sugiyama, M. (2014) “Involvement of rRNA biosynthesis in the regulation of *CUC1* gene expression and pre-meristematic cell mound formation during shoot regeneration”. Frontiers in Plant Science 5: 159.

Shoji, T., Yamada, Y., Hashimoto, T. (2000) “Jasmonate induction of putrescine N-methyltransferase genes in the root of *Nicotiniana sylvestris*”. Plant Cell Physiology 41: 831-839.

Silva, R., Carmo, L., Luis, Z., Silva, L., Pereira, J., Mehta, A. (2014) “Proteomic identification of differentially expressed proteins during the acquisition of somatic embryogenesis in oil palm (*Elaeis guineensis* Jacq.). Journal of Proteomics 104: 112-127.

Sorensen, H., Mortensen, K. (2005) “Advanced genetic strategies for recombinant protein expression in *Escherichia coli”.* Journal of Biotechnology, 115: 113-128.

Sorensen, M. (2001) “Charging levels of four tRNA species in *Escherichia coli* $Rel^{+}$and Rel^{-} strains during amino acid starvation: a simple model for the effect of ppGpp on translational accuracy”. Journal of Molecular Biology 307: 785-798.

Souza-Reis, R., Vale, E., Heringer, A., Catarina, C., Silveira, V. (2016) “Putrescine induces somatic embryo development and proteomic changes in embryogenic callus of sugarcane”. Journal of Proteomics 130: 170-179.

Steward, F.C. (1958) “Growth and development of cultivated cells. III Interpretations of the growth from free cell to carrot plant”. American Journal of Botany 45: 709-713.

Studier, F., Moffat, B. (1986) “Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes”. Journal of Molecular Biology, 189: 113-130.

Sulkowska, J., Sulkowski, P., Szymczak, P., Cieplak, M. (2009) “Stabilizing effects of knots on proteins”. Proceedings of the National Academy of Sciences 105: 19714-19719.

Takac, T., Pechan, T., Samaj, J. (2011) “Differential proteomics of plant development”. Journal of Proteomics 47: 577-588.

Taylor, W., Lin, K. (2003) “Protein knots - a tangled problem”. Nature 421: 25.

Thibaud-Nissen, F., Shealy, R., Khanna, A., Vodkin, L. (2003) “Clustering of microarray data reveals transcript patterns associated with somatic embryogenesis in soybean”. Plant Physiology 132: 118–136.

Tkaczuk, K., Horkawicz, S., Purta, E., Bujnicki, J. (2007) “Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases”. BMC Informatics 8: 73 – 105.

Traxler, M., Summers, S., Nguyen, H., Zacharia, V. Smith, J., Conway, T. (2008) “The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*”. Molecular Microbiology 68: 1128-1148.

Urbonavicius, J., Jager, G., Bjork, G. (2008) “Amino acid residues of the *Escherichia coli* tRNA(m5U54) methyltransferase (TrmA) critical for stability, covalent binding of tRNA and enzymatic activity”. Nucleic Acids Research 35: 3293-3305.

Varhanikova, M., Uvackova, L., Skultety, L., Pretova, A., Obert, B., Hajduch; M. (2014) "Comparative quantitative proteomic analysis of embryogenic and non-embryogenic *calli* in maize suggests the role of oxylipins in plant totipotency". Journal of Proteomics 104: 57-65.

Videira, A. (2001) "Vectores de clonagem" *in* "Engenharia Genética – Princípios e aplicações" LIDEL Edições Técnicas ISBN 9727571638.

Virnau, P., Mirny, L., Kardar, M. (2007) "Intricate knots in proteins: function and evolution". PLOS Computational Biology 2(9): e122.

Wallace, H., Fraser, A. (2004) "Inhibitors of polyamine metabolism". Amino Acids 26: 353-365.

Wang, F., Xiankui, C., Sun, Y., Dong, C. (2013) "Ethylene signaling and regulation in plant growth and stress responses". Plant Cell Reports 32: 1099-1109.

Waugh, D. (2005) "Making the most of affinity tags". Trends in Biotechnology 23: 316-320.

Wei, Y., Zhang, H., Gao, Z., Wang, W., Shtykova, E., Xu, J., Liu, Q., Dong, Y. (2012) "Crystal and solution structures of methyltransferase RsmH provide basis for methylation of C1402 in 16S rRNA". Journal of Structural Biology 179: 29-40

White, J., Li, Z., Saradana, Z., Bujnicki, J., Marcotte, E., Johnson, A. (2008) "Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits". Molecular Cell Biology 28: 3151-3161.

Wieckowski, Y., Schiefelbein, J. (2012) "Nuclear ribosome biogenesis mediated by the DIM1A rRNA dimethylase is required for organized root growth and epidermal patterning in *Arabidopsis*". The Plant Cell 24: 2839-2856.

Wienken, C., Baaske, P., Duhr, S., Braun, D. (2011) "Thermophoretic melting curves quantify the conformation and stability of RNA and DNA". Nucleic Acids Research 10: 1093.

Xiao, W., Custard, K., Brown, R., Lemmon, B., Harada, J., Goldberg, R., Fischer, R. (2006) "DNA methylation is critical for *Arabidopsis* embryogenesis and seed viability". The Plant Cell 18: 805-814.

Yamamoto, N., Kobayashi, H., Togashi, T., Mori, Y., Kikuchi, K., Kuriyama, K., Tokuji, Y. (2005) "Formation of embryogenic cell clumps from carrot epidermal cells is supressed by 5-azacytidine, a DNA methylation inhibitor". Journal of Plant Physiology 162: 47-54.

Yan, L., Zhai, Q., Wei, J., Li, S., Wang, B. (2013) "Role of Tomato Lipoxygenase D in Wound-Induced Jasmonate Biosynthesis and Plant Immunity to Insect Herbivores". PLoS Genetics 9(12): e1003964. doi:10.1371/journal.pgen.1003964.

Yang X., Zhang X. (2010) "Regulation of somatic embryogenesis in higher plants". Critical Reviews in Plant Sciences, 29: 36-57.

Zavattieri, M., Frederico, A., Lima, M., Sabino, R., Arnholdt-Schmitt, B. (2010) "Induction of somatic embryogenesis as an exemple of stress-related plant reactions". Electronic Journal of Biotechnology DOI: 10.2225/vol13-issue1-fulltext-4.

Zeller, G., Henz, S., Widmer, C., Saschenberg, T. Ratsch, G., Weigel, D., Laubinger, S. (2009) “Stress-induced changes in the *Arabidopsis thaliana* transcriptome analyzed using whole-genome tiling arrays”. The Plant Journal 58: 1068-1082.

Zhang, J., Mab, H., Chen, S., Ji, M., Perl, A., Kovacs, L., Chen, S. (2009) “Stress response proteins differential expression in embryogenic and non-embryogenic callus of *Vitis vinifera* L. cv. Cabernet Sauvignon – A proteomic approach”. Plant Science 177: 103-113.

Zimmerman, J.L. (1993) “Somatic embryogenesis: a model for early development in higher plants”. The Plant Cell 5: 1411-1423.