β-glucuronidases in plants

Ph.D. Thesis of
Matteo Ronchini

Supervisor: prof. Paolo Gerola

Academic year 2013/2014
CONTENTS

Abstract

General introduction ............................... 1

Chapter 1 .................................................. 16
Artefacts in histochemical localization of LAT52-GUS
construct in Nicotiana alata pollen tubes growing through style

Chapter 2 .................................................. 29
β-glucuronidases activity in pollen

Chapter 3 .................................................. 38
GUS genes expression in Arabidopsis thaliana and Nicotiana
  tabacum

Chapter 4 .................................................. 53
Identification of GUS genes in N. tabacum and phylogenetic
  analysis of GH79 GUSs

Materials and method ................................ 72

General conclusions and perspective ............... 83

Appendix
ABSTRACT

In the last decade, the interest for the presence and role of endogenous GUS in plants increased, being previously limited to the use of GUS from E. coli as a reporter gene. Three different GUS genes have been identified in A. thaliana and different roles have been suggested, mainly associated to plant cell wall remodelling and to the regulation of the presence in the active form of molecules with regulative functions.

This thesis investigated several aspects: A) artefacts in histochemical GUS detection, B) role of GUS in pollen tube germination and growth, C) GUS expression in N. tabacum and A. thaliana, D) identification of GUS genes in N. tabacum and their phylogenetic analysis in angiosperms.

A) Histochemical detection of E.coli GUS activity in transformed plants can be impaired by the presence of GUS inhibitors and by the solubility of an intermediate reaction product formed when X-glu is used as substrate. The expression of LAT52 has been revised.

B) The use of saccharolattone, a GUS inhibitor, suggests that GUS is involved in pollen tube germination and growth.

C) GUS expression was observed in all organs of N. tabacum and A. thaliana.

Expression studies, in collaboration with researchers of Calabria University, have been performed in Arabidopsis, by in situ mRNA hybridization: GUS3 is specifically expressed in border like cells and probably it is involved in their detachment from root tip; GUS1 and GUS2 are expressed in the root cup meristem, resulting perhaps involved in regulation of the mitotic cycle.

D) In N. tabacum GUS2 was completely sequenced and one GUS1 and two GUS3 have been partially sequenced. By bioinformatics analysis, GUS genes have been identified in other angiosperms species and a phylogenetic analysis have been conducted.
**GENERAL INTRODUCTION**

**β-glucuronidases: classification and roles**

Glycosil hydrolases (GHs) are enzymes which hydrolyse the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. GHs have been subdivided into 132 families, which are continuously updated in the Carbohydrate Active EnZymes (CAZy) database (Cantarel et al 2009).

In plants, GHs are mainly involved in cell wall polysaccharides metabolism; other functions are remodulation of glycans, mobilization of energy reserves, defense, symbiosis, signaling, secondary plant metabolism and metabolism of glycolipids (Minic 2008).

β-glucuronidases are GHs which catalyze the hydrolysis of O-linked glucuronic acid. These enzymes are both exo- or endo-acting, depending upon whether they act at the (usually non-reducing) end or in the middle, respectively, of a saccharide chain.

β-glucuronidases have been identified in all living organisms and, according to their aminoacidic sequences, have been classified in three GH families (GH1, GH2 and GH79) (Cantarel et al 2009).

GH1 comprehend klotho enzymes, an extracellular mammal cell membrane enzyme that hydrolyzes steroid β-glucuronides (Tohyama et al 2004).

The GH2 family contains a large number of exo-acting GUSs from mammals, bacteria and fungi but none from plants (Arul et al 2008). In bacteria, GH2 GUS is involved in carbon source mobilization (Wenzl et al 2005).

Presence of GUS activity in floral intestinal bacteria can cause enterohepatic circulation of toxic, carcinogenic compounds and thus promote tumor formation at different sites, including the large bowel (Arimochi et al 1999, McBain & Macfarlane 1998, Wallace et al 2010).
β-glucuronidase (uidA), widely used as reporter gene in plants (Jefferson 1989, Jefferson et al 1986), is included in the GH2 family. The GH79 family contains both exo-acting (β-glucuronidase) and endo-acting (heparanase) hydrolases (Michikawa et al 2012). To this family belong vertebrate heparanases, which cleave heparan sulphate side chains thus inducing structural alterations of the extracellular matrix (Fux et al 2009), fungi enzymes that hydrolyze beta-glucuronosyl residues of arabinogalactan proteins (AGPs) (Konishi et al 2008), bacterial β-glucuronidases (Michikawa et al 2012) and all the known plant β-glucuronidases (Arul et al 2008, Eudes et al 2008, Morimoto et al 1995, Woo et al 2007).


In 2006 Sudan et al. demonstrated the presence of GUS activity in several plants, including the main model plants (Arabidopsis thaliana, Oryza sativa, Nicotiana tabacum, Zea mays), and proposed a role of endogenous GUS in cell (Sudan et al 2006). sGUS from Scutellaria baicalensis was the first GUS gene sequenced in plants (Sasaki et al 2000) and it was observed to be homologous to heparanase and therefore included in the GH79 family.
Recently, a genome analysis revealed the presence of three GUS genes from the GH79 family in *Arabidopsis thaliana*: AtGUS1, AtGUS2 and AtGUS3 (Woo et al. 2007). Other GUS genes can be identified in the genome of other plants species, all homologous to sGUS and belonging to the GH79 family.

Regarding the role, GUS in plants is thought to be involved both in the regulation of the presence of physiologically active signal molecules, mainly associated to the cleavage of flavonoid glucuronides and in changes in cell wall composition, associated to processes like cell growth and cell-cell adhesion.

It has been demonstrated that the flavonoid luteolin accumulates in the apoplastic space of mesophyll cells of rye seedling leaves as luteolin 7-O-diglucuronide-4’-O-glucuronide. GUS present in the apoplastic space hydrolyzes it to luteolin 7-O-diglucuronide (Anhalt & Weissenböck 1992, Shulz & Weissenbock 1987), probably playing a role in hydrogen peroxide and auxin levels regulation.

In *Scutellaria baicalensis* the activity of apoplastic GUS releases the flavonoid baicalein, which is involved in two different processes: hydrogen peroxide levels regulation in the oxidative burst induced by plant defense mechanism (Morimoto et al. 1998) and the induction of apoptosis following mechanical damage (Hirunuma et al. 2011).

The involvement of GUS in stress responses has been observed also in *Hordeum vulgare*, where the inoculation with the fungus *Blumeria graminis* causes up regulation of the GUS gene (Eckey et al. 2004).

The role of GUS in regulative processes has to be considered in relation to UDP-glucuronyltransferase (UGT). The two enzymes work in tandem, reversibly glycosylating molecules, like flavonoids, with regulative function.

This UGT-GUS tandem function seems to operate in the regulation of cell division in the root cap meristem. It has been proposed that PsUGT1 activate cell division by glycosylating, and therefore removing, a mitosis inhibitor, probably a flavonoid (Woo et al. 1999, Woo et al. 2005). Such component is vice versa released by the action of GUS, with consequent cell cycle inhibition (Wen et al. 2004, Woo et al. 2005).
Concerning changes in cell wall composition, enzymes codified by the genes *AtGUS1* (At5g07830), *AtGUS2* (At5g07830) and *AtGUS3* (At5g34940) were found within the *Arabidopsis thaliana* cell wall proteome (Bayer et al 2006, Minic et al 2007).

GUS enzymes acting on glucuronic acid (GlcA) residue could be involved in two processes in the cell wall: cell wall plasticity and intercellular adhesion.

The involvement of GUS activity in the degradation of arabinogalactan proteins (AGPs) present in the cell wall was first hypothesized in a study conducted on seeds extracts of *Raphanus sativus* (Sekimata et al 1989).

In another work tissue-specific localization and partial purification of *AtGUS2*, a GUS active isolated from *Arabidopsis thaliana* (Eudes et al 2008) have been reported. In this study it has been demonstrated that over-expression of *AtGUS2* could modify GlcA content in the polysaccharidic side chains of AGPs and influence the content of other sugars such as galactose, arabinose and xylose.

GlcA residues present at the non-reducing ends of the side chains of most AGPs potentially prevent the access of glycosyl hydrolases to sugar residues located within these chains.

Transgenic plants in which *AtGUS2* activity was suppressed display defects compared to wild-type plants. In particular, dark-grown hypocotyls seedlings showed a decrease of elongation compared to the wild type. The authors have supposed that deglycosylation of AGPs have a role in cell growth and, in particular, could contribute to hypocotyl growth (Eudes et al 2008). *AtGUS2* would be one of the enzymes involved in this process.

GlcA residues have also been demonstrated to be important in intercellular organization and attachment in plant meristem (Iwai et al 2002).

In *Nicotiana plumbaginifolia* it was observed that the loss of one unit of glucuronic acid in the pectin molecule, due to inhibition of glucuronyltransferase, induced drastic morphological abnormalities like crumbled shoots (Zhong et al 2005).
These findings evidence the importance of glucuronic acid residues in cell wall and are consistent with the consequences of inhibition of GUS activity. Inhibition of GUS activity leads to reduced elongation of stem, root and root hairs in *N. tabacum* (Sudan et al 2006). Inhibition of root hairs elongation is particular interesting because of their apical growth mechanism very similar to pollen tube growth. This consideration is supported by the existence of a common pathways involved in polarized cell-tip expansion in pollen tubes and root hairs (Hafidh et al 2012). The only GUS gene (*GUS2*) considered in the study was not expressed in *Nicotiana tabacum* pollen. However, the presence of GUS genes also in pollen has been reported in the literature (Honys & Twell 2004, Pina et al 2005). Thus, the genes that codify GUS enzymes in *Nicotiana* pollen remain undetected.

**Importance of investigating genic expression in pollen**

In this work particular attention is dedicated to the study of β-glucuronidase gene expression and its role in pollen. The importance of studying genic expression in pollen is related to gametophytic selection and to the comprehension of the mechanism of pollen growth. Gametophytic selection has been widely studied starting from the end of nineteenth century (Buchholz 1922, Mulcahy 1986) and its importance as a plant breeding tool has been widely reviewed by Hormaza and Herrero (Hormaza & Herrero 1996). When a diploid organism, heterozygous at N loci, undergoes meiosis, it has the potential of producing $2^N$ different haploid genotypes. In the case of pollen, this great variety is often associated with an equally impressive overabundance of individuals. There are a vast array of pollen genotypes competing for a relative small number of ovules. Such competition holds the potential for an extremely intense selection. Selective pressures could be applied and act at different stages of pollen development and reproductive processes.
These pressures produced effects on progenies. In recent studies it was observed that *Phalaenopsis* seedlings derived from pollination at high temperature were more vigorous under warm growing conditions and those derived from cold pollination were more vigorous under cold growing conditions (Chang et al 2010). The main limitation of gametophytic selection in plant breeding is that it can only be used for traits that are expressed in gametophyte. Hence the importance of knowing gametophyte gene expression.

Male gametophytic gene expression has been investigated in several studies and reviewed by Twell and coworkers (Twell et al 2006). Comparing the results of the different published works it is possible to note the remarkable overlap of gametophytic and sporophytic gene expression. In addition, the number of male specific gametophyte expressed genes is gradually decreasing as new sporophytic datasets emerge, especially those from more specialized tissues and individual cells. These genes are characterized by very high expression signals, highlighting their importance and their potential as targets for functional analysis.

Male gametophyte gene expression can be divided into two major phases, early and late (Mascarenhas 1990). Early genes become active after meiosis is completed, that is at the tetrad stage. Late genes become active after microspore mitosis. Late genes make a significantly greater contribution to the group of highly expressed pollen-specific genes while early genes expression program is much more similar to that of the sporophyte.

Given that the number of putative male gametophyte specific genes gradually decrease, a new expression profile must be considered in addition to early and late genes. This comprises non-specific enhanced pollen-expressed genes. Such pollen-enhanced genes are defined as genes with maximum male gametophytic expression at least five times higher than the maximum expression in the sporophyte.
From a functional point of view, in *Arabidopsis thaliana* microspore the most expressed genes are those involved in protein synthesis while among gametophytic specific genes the most expressed are cell wall and transport genes. In mature pollen, genes involved in cell wall synthesis and metabolism comprise more than 19% of highly expressed specific mRNAs. In pollen tube growth *in vivo* up regulated genes, compared to dry pollen, are those codifying for proteins with function of transporter, antiporter, symporter and calcium ion binding (Qin et al. 2009). The relevance of genes involved in cell wall synthesis emerge from these studies. Among them there are genes codifying for glycolsil hydrolases, whose role in pollen has been recently reviewed (Mollet et al. 2013). However, in this review the presence of β-glucoronidases hasn’t been reported, thus demonstrating that the role and the importance of these enzymes has to be better elucidated and confirmed also under this aspect.

**Nicotiana and Arabidopsis as model plants for genetic studies**

*Nicotiana tabacum*, *Nicotiana alata* and *Arabidopsis thaliana* are the three species selected in order to investigate the presence and role of GUS and the genes that codify the related enzymes. *Nicotiana* species are widely used in studies regarding interspecific hybridization, inheritance and gene transfer because of the large number of widely varying species in the genus and some advantaging characteristics (Lewis 2011). Many species of the genus, like *Nicotiana tabacum*, present sizable flowers that can be manipulated easily and often a large amount of seed can be obtain from single pollination. *Nicotiana alata* presents the advantage of producing great quantity of pollen not dispersed from the anther and so easy to be collected for further studies. In *Nicotiana* several species are characterized by self-incompatibility (McClure 2009), and are thus suitable to studies of this phenomenon and the interspecific crosses between self-incompatible (SI) and self-compatible (SC) species (Murfett et al. 1996).
N. alata (SI species) pollen can be used to successfully pollinate N. tabacum (SC species), but not vice versa (Murfett et al 1996), generating a completely sterile hybrid (Nikova et al 1999)

Since the beginning of the 20th century, N. tabacum is believed to have originated by hybridisation of N. sylvestris (S-genome component) with a species in the Tomentosae section (T-genome component) of Nicotiana (Gerstel 1960, Goodspeed & Clausen 1928).

The T-genome component has generated a long discussion: N. tomentosiformis, N. otophora, or an introgression hybrid of the two were considered as candidates ancestors of N. tabacum (Murad et al 2002).

Genomic in situ hybridization (Chase et al 2003, Lim et al 2000) indicated that N. tomentosiformis is the most probably N. tabacum ancestor and comparison of chloroplast genome confirmed N. sylvestris as “female” ancestor (Yukawa et al 2006).

Thus Nicotiana tabacum is actually considered to be a natural allotetraploid originated within the last 200000 year by the cross of N. sylvestris and N. tomentosiformis (Clarkson et al 2005, Murad et al 2002).

Thanks to genomic in situ hybridization it is possible to identify chromosomes derived from each progenitor and hybrid chromosomes are present as well (Lim et al 2007, Lim et al 2004). Some translocation between T-genome and S-genome appear ubiquitous in all N. tabacum cultivars, and probably fixed, whereas others are specific to particular cultivars.

N. tabacum shows a genome downsizing with respect to the sum of its parent genomes (Renny-Byfield et al 2011). This reduction seems to be due to the loss of repeats derived from the T-genome component of N. tabacum.

Partial sequencing of N. tabacum genome (Tobacco Genome Initiative) and recent publication of N. sylvestris and N. tomentosiformis genomes (Sierro et al 2013) will allow further investigations regarding the evolution of the three genomes and make easier to sequence specific genes in N. tabacum and thereafter to investigate their expression.
Although *Nicotiana alata* genomic resources available are very scarce, this species, with respect to *Nicotiana tabacum*, has the advantage to present a diploid genome. 

*Arabidopsis thaliana* is a flowering diploid plant of the family of *Cruciferae* widely used as model because present a short life cycle, it is easy to cultivate, its genome (the smallest among angiosperm) has been sequenced and annotated (AGI 2000) and a large number of mutants are available (Meinke et al 1998, Rédei 1975).

**Topics of this thesis**

This thesis touches one or more topics about β-glucuronidases. Chapter 1 regards GUS as reporter gene in plants. In particular experiments are described and discussed related to histochemical localization of *E. coli* GUS activity in pollen tubes growth *in vivo*, *semi-in vivo* and *in vitro*. 

Chapter 2 concerns the study of GUS activity and its role in pollen of *N. alata* and *N. tabacum*. In chapter 3 the study of both GUS activity and the expression pattern of GUS genes in different plant species are reported. 

Gene expression in root apex of *Arabidopsis thaliana* was investigated thanks to a recently developed whole mount multi probe *in situ* hybridization technique, in addition to traditional approach. In *Arabidopsis thaliana*, the expression of UGT genes was also investigated, given that they possibly work in tandem with GUS genes.

Chapter 4 regards the identification and phylogenetic analysis of GUS genes in plants and in particular in *Solanaceae*.


Gerstel DU. 1960. Segregation in new allopolyploidd of Nicotiana. I. Comparison of 6x (N. tabacam) x tomentosiformis and 6x (N. tabacam x otophora). Genetics 45: 1723-34


Yukawa M, Tsuzuki T, Sugiura M. 2006. The chloroplast genome of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*: complete sequencing confirms that the *Nicotiana sylvestris* progenitor is the maternal genome donor of *Nicotiana tabacum*. *Mol Genet Genomics* 275: 367-73

CHAPTER 1

Artefacts in histochemical localization of LAT52-GUS construct in Nicotiana alata pollen tubes growing through style

INTRODUCTION

GUS from E. coli (uidA) is the most widely used reporter gene in plants. This is due to the advantages that it offers with respect to the other reporter systems: the stability of the enzyme, the simplicity and sensitivity of the assays, the variety of available substrates and the absence of toxic effects on plants (Jefferson 1987, Jefferson 1989, Jefferson et al 1986). Several efforts have been spent to overcome artefacts due to the presence of endogenous GUS activity. However, the difference in pH optimum (neutral for E. coli GUS and acidic for plant GUSs) (Alwen et al 1992, Sudan et al 2006) makes negligible the interference due to endogenous GUS, restricting it to a background in long histochemical treatments.

Constructs of uidA associated to promoters of genes specifically expressed in pollen have been used not only to investigate gene expression since microsporogenesis up to ovule fecundation, but also as pollen markers, allowing to recognize specific pollens or pollen tubes. In particular, LAT52 promoter-uidA construct has been used to observe pollen tubes behaviour in vivo (Gerola et al 2000, Johnson et al 2004). LAT52 was identified for the first time in Solanum lycopersicum as a pollen specific gene expressed late in the anther, after meiosis, during microsporogenesis (Twell et al 1989). This pattern of expression has been observed also in Nicotiana tabacum and Arabidopsis thaliana (Eady et al 1994, Twell et al 1990).
It has been demonstrated that LAT52 is involved in pollen germination and polarized tube growth (Kaothien et al 2005, Muschietti et al 1994, Zhang et al 2008) and encodes a small cysteine-rich protein which interacts with a pollen specific receptor protein kinase LePRK2 (Johnson & Preuss 2003, Tang et al 2002).

Gerola and co-workers (2000) used pollen from Nicotiana alata plants transformed with LAT-52 promoter-uidA construct (GUS-pollen) to investigate the so called pollen mentor effect in gametophytic self-incompatibility. They observed that GUS activity in pistils pollinated by GUS-pollen was histochemically observable in the stigma, the higher part of the style and in the ovary, while was practically undetectable in the lower part of the style. According to these results it was suggested that the LAT52 promoter activity is differentially regulated during pollen tube growth along the pistil (Gerola et al 2000).

An analogous GUS staining pattern was observed in our laboratory in N. tabacum pistils pollinated with N. alata GUS-pollen. A high variability was however observed in the histochemical detection in the uppermost part of the style (Gerola, personal communication). Moreover, enzymatic assays in extracts from pistils pollinated with GUS-pollen did not reveal any significant difference between GUS activity in the higher and lower part of the style, suggesting the potential presence of staining artefacts (Pilotto et al 2002, Pisoni et al 2004).

Two kinds of artefacts could interfere with histochemical detection of GUS activity: diffusion of the reaction product (Mascarenhas & Hamilton 1992, Stomp 1992) and the presence of inhibitors of GUS activity (Bahieldin et al 2005, Fior & Gerola 2009, Ramadan et al 2011, Thomasset et al 1996). In particular, the presence of a low molecular weight inhibitor of GUS activity was observed in N. tabacum stylar extract (Fior & Gerola 2009).

However, there is no indication in literature on artefacts on histochemical GUS detection due to the presence of GUS-inhibitors, which have been reported to interfere with the in vitro measured enzymatic activity.

In this work we further investigated the LAT52 promoter activity in vitro, in semi-vivo and in vivo pollen tube growth.
We observed that it is not differentially regulated and that artefacts are responsible of the lack of GUS detection in pollen tubes in the lower part of the style.
Only pollen from *N. alata* plants transformed with *LAT52* promoter – *uidA* construct was used in this work.

**RESULTS**

In order to solve the controversy between the histochemical analysis (low or no GUS activity in pollen tubes present in the lower part of the style) and the enzymatic data (similar activity in the pollen tubes present in the upper and lower part of the style), GUS activity was histochemically analysed in semi-*in vivo* grown pollen tubes.

Pistils were collected 24 hours after pollination (when the pollen tubes reach about the middle of the style), cut off at the lower part of the style and immersed in pollen tubes growth medium. After few hours pollen tubes grow out from the style, elongating into the medium. GUS activity, histochemically detected by “X-Glu reaction”, was clearly evident in the protruding pollen tubes (fig. 1.1), in agreement with the enzymatic results (data not shown) and in contrast with what observed in *in vivo* grown pollen tubes (Gerola et al 2000). However, as only the apical region of pollen tubes was protruding from the style, we decided to verify the presence of GUS enzyme and activity also in the old region of the pollen tube.

![Image](image.png)

**Fig.1.1** Histochemical GUS activity staining of *N. alata* GUS-pollen tube grown semi-*in vivo* in *N. tabacum* pistil.

Pistil 24 h after pollination was cut and immersed vertically in growth medium. Protruding pollen tubes after 48 h were stained with X-glu. Blue precipitate is present in almost 50% of pollen tubes given that pollen was obtained from *N. alata* plants hemizigous for the *LAT52* promoter. Note that staining is detectable mainly in pollen tube tip (an example is indicated by arrow). Image was acquired in transmitted light.
Pollen tubes were then *in vitro* grown for 24 hours and the presence of GUS was analysed by both immunohistochemical and histochemical detection. For immunohistochemical analysis, pollen tubes were treated with rabbit anti *E.Coli GUS* antibody, followed by treatment with a mouse anti-rabbit secondary antibody linked to the fluorochrome Alexa Fluor® 488. As it can be seen in fig. 1.2, pollen tubes show the apical region separated by the old one by a callose plug and the presence of GUS, revealed by Alexa fluorescence, is observable all along the pollen tube, with higher intensity in the tip with respect to the old region. This difference in immunohistochemical staining is probably due to the fact that the pollen tube tip is rich in cytoplasmic content, where the GUS is localized, while the old one is amply vacuolated and the cytoplasm is limited to a thin layer under the plasma membrane.

![Fig. 1.2 Immunohistochemical localization of GUS enzyme in *N. alata* pollen after 24 h of *in vitro* growth. Image was acquired with epifluorescence microscope.](image)

Partially different results were obtained by the histochemical analysis (fig. 1.3). In fact, X-Glu treatment revealed GUS activity in pollen grains and in pollen tubes tips, while the staining was faint or absent in the vacuolated old region. Since the first product of X-Glu hydrolysis is a soluble component (Stomp 1992), which only after dimerization and oxidation gives rise to the blue precipitate, we verified the existence of leakage of the reaction product.
To avoid product dispersion in the reaction medium, pollen tubes were stained by X-Glu after growth in 2% agarose: blue precipitate was clearly evident in the agar surrounding the pollen grains and the pollen tubes (fig. 1.4).

Fig. 1.3 Histochemical staining for GUS activity in *N. alata* pollen tube grown *in vitro*. Pollen tubes after 3 h of growth were stained with aniline blue (A) and X-glu (B). Arrows indicate callose plugs. Pollen tubes after 24 h of growth were stained with X-glu (C). Images were acquired in transmitted light. Detailed explanation is reported in text.

Fig. 1.4 X-glu staining dispersion *in vitro*. (A) *N. alata* GUS-pollen grains on *N. tabacum* style stained with X-glu. Note precipitate dispersed around pollen grains. (B) *N. alata* GUS-pollen tubes stained with X-glu after 24 h of growth. Arrow indicated precipitate dispersed around pollen tube. Images were acquired in transmitted light.
GUS activity was then analysed in *in vivo* grown pollen tubes. Pollinated pistils were collected, dissected in different ways and stained with X-Glu treatment. The best staining evident at the stereo-microscope was observed when the pistil was dissected by removal of the epidermis, with minimal damage of the styal transmitting tissue (fig.1.5). However, analysis at the transmitting light microscope revealed that the blue precipitate was present inside the pollen tubes tip and on the transmitting tissue cells, indicating a leakage from the pollen tubes (fig. 1.6).

![Image](image1.png)

**Fig. 1.5** Histochemical staining for GUS activity in *N. alata* GUS-pollen tube grown *in vivo* in *N. tabacum* pistil. Style 48 h after pollination was dissected removing epidermis, than was stained with X-glu. Image was acquired in transmitted light.

![Image](image2.png)

**Fig. 1.6** X-glu staining dispersion *in vivo*. Pollen tubes grown through style stained with X-glu (A) and aniline blue (B). Aniline blue allowed a precise identification of pollen tubes while X-glu diffused also in tissue around pollen tubes (see zones indicate by arrows). Note that X-glu staining isn’t detectable in all pollen tube length. Images were acquired in transmitted light (A) and epifluorescence (B)
In order to reduce leakage of the reaction product we tried to use ELF® 97-β-D-glucuronide (ELF® 97) as substrate of the histochemical reaction, a substrate developed for GUS activity detection in polyacrylamide gels after electrophoretic separation (Zhou et al 1996). In fact no soluble intermediate is formed upon hydrolysis and the reaction gives rise directly to a fluorescent precipitate.

On the contrary of what observed when using X-Glu, ELF® 97 histochemical staining of *in vitro* grown pollen tubes (fig. 1.7) revealed the presence of GUS activity all along the pollen tube, although, as expected by the immunohistochemical detection, in lower amount in the old part of the pollen tube with respect to the tip. Old part of pollen tube was clearly evidenced by fluorescein diacetate vital staining (fig. 1.8).

---

**Fig. 1.7** Histochemical staining for GUS activity in *N. alata* GUS-pollen grown *in vitro*. Pollen tubes after 24 h of growth were stained with ELF® 97. Detailed explanation is reported in text. Image was acquired with epifluorescence microscope.

**Fig. 1.8** Fluorescein diacetate vital staining of *N. alata* pollen tube grown *in vitro* for 24h. Image was acquired with epifluorescence microscope.
The use of ELF® 97 as substrate for histochemical detection of GUS activity turned out to be better than X-Glu also in semi-in vivo (fig. 1.9) and in vivo pollen tubes (1.10). In pistils collected 24 hours after hyper-pollination, GUS activity was much more evident in the upper and lower part of the style than in the central portion. This staining pattern is due to the fact that in *N. tabacum* the transmitting tissue become thinner in the upper part of the style and therefore in hyper-pollinated pistils pollen tube tips, with high GUS activity, are present in two region of the style: in the lower part of the style, where pollen tubes arrived during their growth, and in the upper part, where pollen tubes “in excess” were stopped by the taper of the transmitting tissue.

Mainly old, with low GUS activity, regions of the pollen tubes that grew to the bottom of the style are present in the central portion, where aniline blue staining is prevailing.

Fig. 1.9 Histochemical staining for GUS activity in *N. alata* GUS-pollen grown semi-in vivo. Pistils 24 hours after pollination were cut 2 cm from stigma and vertically immersed in growth medium. Protruding pollen tubes were stained with aniline blue (A) and ELF® 97 (B). Images were acquired with epifluorescence microscope.
Fig. 1.9 Histochemical staining for GUS activity in *N. alata* GUS-pollen grown in vivo. Pistil 36 hours after pollination was collected and epidermis removed. Then style was stained with ELF® 97. Detailed explanation is reported in text. Image was acquired with epifluorescence microscope.

**DISCUSSION**

The pollen specific promoter *LAT52* was supposed to be regulated during pollen growth through the pistil: active in the stigma and the ovary and inactive along the style (Gerola et al 2000). This hypothesis was based on histochemical assay of GUS activity in pistils pollinated with pollen from *N. alata* plants transformed with a construct formed by the *LAT52* promoter associated to E. coli GUS as reporter gene (GUS pollen). However similar GUS activity was assayed fluorimetrically in the lower and higher part of pollinated styles (Pilotto et al 2002) and presence of GUS inhibitor was reported in *Nicotiana* styles (Fior & Gerola 2009, Pisoni et al 2004). These results questioned the reliability of the differential X-glu histochemical staining observed in the pollinated style, although the observation that the GUS inhibitor is uniformly distributed along the style (Pisoni et al 2004) does not support the hypothesis that GUS activity in pollen tubes is inhibited only in the lower part of the style.

In any case, these observations, together with a high variability in GUS staining in the pistils pollinated with GUS pollen, led us to verify if *LAT52* promoter is effectively regulated during pollen tube growth. Semi-*in vivo* experiments demonstrated that GUS is present and is histochemically detectable in the tips of pollen tubes grown through two thirds of the style, i.e. penetrated in that stylar region where GUS activity is usually low or not histochemically detectable.
Moreover, an immunohistochemical approach demonstrated that GUS is present both in the tip and in the old part of *in vitro* grown pollen tubes, clearly in higher amount in the tip, which is rich in cytoplasmic content, with respect to the old part, characterized by vacuolization and reduced cytoplasm.

However, in the histochemical detection of GUS activity by X-Glu staining the reaction product (a blue precipitate) was observable in pollen grains and in the tip of pollen tubes, but it was practically absent in the old part. In addition dispersion of precipitate was observed outside pollen grains and pollen tubes, both in vitro and in vivo. In particular, it has been observed that the method used to dissect pollinated pistils before X-Glu staining influences the histochemical detection of GUS activity. When the style is completely dissected and the pollen tubes, free of transmitting tissue cells, are exposed to the medium, the few blue precipitate is observable, localized inside pollen tubes tips. Instead when the pistil is gently dissected by removing the epidermis, abundant blue precipitate is observable in the style. However further inspection revealed that, over the reaction product present inside the pollen tube tips, a large amount of the precipitate was associated to the transmitting tissue cells, outside the pollen tubes.

From these observation we can conclude that *LAT52* promoter is not differentially regulated during pollen tube growth along the style. More factors can influence the histochemical detection of GUS activity in pollinated pistils. In fact the first product of X-Glu hydrolysis is a soluble indoxyl which dimerizes and precipitates only in the presence of particular redox condition (Stomp 1992). The plasma membrane in the old part of pollen tubes has lost its semi permeability and the cytoplasm, where GUS is localized, is reduced to a thin layer under the plasma membrane. Absence of fluorescein diacetate vital staining in old pollen tube parts confirmed this consideration. In these conditions the indoxyl, product of X-Glu hydrolysis, diffuse outside the pollen tube and precipitate on the surrounding tissue, like in the style dissected by epidermis removal, or is dispersed in the medium, like in the completely dissected style.
In the pollen tube tip the membrane remains semi permeable for longer time, moreover GUS is present in higher amount and the cytoplasm fill the tip, so that the indoxyl produced by X-Glu hydrolysis in the centre of the tip precipitates before reaching the plasma membrane and diffusing out of the pollen tube. Blue precipitate is therefore observable in the pollen tube tips present along the style, although during longer reaction times the plasma membrane can lose semi-permeability, with diffusion of the indoxyl and appearance of blue precipitate no more strictly associated to the presence of GUS.

It has also be considered that the transmitting tissue is constituted by two cylindrical portions that, starting under the stigma surface, converge and fuse in the centre of the style. That is, under the stigma there are “two” transmitting tissues which fuse constituting a central transmitting tissue with the shape of an eight which progressively change to a thinner central cylinder (Bell & Hicks 1976). Therefore the space available for pollen tube growth decreases in the higher part of the style and, when the pistil is hyper-pollinated, pollen tubes that germinate later enter the transmitting tissue but cannot go ahead along the style. In hyper-pollinated pistils, the higher portion of style is therefore rich in pollen tubes tips and, therefore, in GUS activity, while, on the opposite, few pollen tubes tips are present in less pollinated pistils.

Moreover the procedure used to dissect the pistil might both damage the integrity of plasma membrane with dispersion of X-glu staining and also change the exposure of GUS to the GUS inhibitor present in the style. All these factors influence histochemical GUS detection in different extent, causing the observed high variability.

Removal of the artefacts associated to diffusion of the reaction product was obtained by substituting X-Glu with ELF® 97 as substrate. In fact no soluble intermediate is produced by ELF® 97 hydrolysis and the fluorescent precipitate which is produced is also visible inside the old part of the pollen tubes.


Muschiatti J, Dircks L, Vancanneyt G, McCormick S. 1994. LAT52 protein is essential for tomato pollen development: pollen expressing antisense
LAT52 RNA hydrates and germinates abnormally and cannot achieve fertilization. *Plant J* 6: 321-38


CHAPTER 2

B-glucuronidases activity in pollen

Parts of the contents of this chapter was reported in:

INTRODUCTION

Pollen is the male gametophyte of flowering plants and it is an independent organism respect to the plant. Once germinated on stigma, pollen has the role to carry the two sperm cells to the ovule allowing the double fertilization process and seed setting. These processes requires a massive cell wall deposition to promote fast pollen tube elongation and a tight control of the cell wall remodeling to modify its mechanical properties.

In cell wall metabolism are involved a great number of enzymatic families included glycoside hydrolases (GHs) (Minic & Jouanin 2006). According a recent review in pollen tube have been identified the following GH families: 3, 9, 10, 17, 28, 35, 43 and 51 (Mollet et al 2013), GH79 family wasn’t considered.

However in past presence of β-glucuronidase activity was observed in pollen, but the enzymes involved wasn’t characterized. Expression of GUS genes in pollen has been discussed in the chapter 3.

First observations regarding presence of GUS activity in pollen dated to 1965 (Gorska-Brylass 1965). Similar results were obtained several decades later using pollen of Portulaca grandiflora (Sood 1980).

Presence of endogenous β-glucuronidase activity was histochemically detected using in male gametophyte of Nicotiana tabacum (Plegt & Bino 1989). GUS activity couldn’t be demonstrated at premeiotic and meiotic stages in anthers.
During the later stage of development blue precipitation, due to X-glu staining, was observed in tapetal and sporogenous cells. Mature bicellular pollen contained high level of GUS, in tricellular pollen such presence couldn’t be demonstrated. This difference was explained with a different regulation of protein-synthetizing apparatus in bicellular and tricellular pollen. Similar results have been reported some years later (Alwen et al 1992). In particular was observed GUS activity only in mature *N. tabacum* pollen. Regarding GUS role in pollen has been suggested that GUS activity may be correlated with the presence of glucuronic acid-rich arabinogalactan proteins (AGPs) in style, contributing to the hydrolysis of the acid carbohydrates (Plegt & Bino 1989). AGPs is also present in cell wall of pollen grains and tubes (Nguema-Ona et al 2012). AGPs have different roles in plants: cell-cell recognition in plant reproduction, cell adhesion, signal transduction and nutrient resources (Cheung & Wu 1999, Nguema-Ona et al 2012, Schultz et al 2000, Zhang et al 2011). All functions important in pollen development and in its growth along the style. It has recently been observed that overexpression or suppression of *AtGUS2* gene in *Arabidospis thaliana* induced an alteration in composition of AGPs (Eudes et al 2008), supporting the initial hypothesis of Plegt and Bino (1989). Pollen tubes show an apical growth modality such as root hairs. In these least cells has been observed a reduced growth due to by inhibition of GUS activity (Sudan et al 2006), suggesting to verify the presence of a similar effect in pollen tubes. Starting from the previous considerations the activity and role of GUS enzymes have been investigated in pollen of *N. alata* and *N. tabacum*.

**RESULTS**

**Effects of inhibitors on GUS endogenous activity**

The use of inhibitors is one of the approaches used to investigate biological role of enzymes.
Contrasting results have been reported in literature on efficiency of traditional GUS inhibitors on endogenous GUS activity in plants (Muhitch 1998). We tested three different components known to inhibit GUS activity of *E. coli* (saccharolactone, glucuronic acid and galacturonic acid). The inhibitory efficacy was tested on GUS activity histochemically detected in pollen grains. As it can be seen in figure (fig. 2.1) saccharolactone is the most efficient GUS inhibitor: at 1 mM concentration inhibition is already evident and at 10 mM no GUS staining is anymore observable. Thus saccharolactone has been therefore used for investigating the role of endogenous GUS in pollen tubes germination and growth.

Fig. 2.1 - glu histochemical staining of wild type *N. alata* pollen grains in presence different concentration of GUS inhibitors: control (on the left); saccharolactone: 1 mM (A), 4 mM (B), 10 mM (C); glucuronic acid: 1 mM (D), 4 mM (E), 10 mM (F); galacturonic acid: 1 mM (G), 4 mM (H), 10 mM (I). Saccharolactone showed the highest inhibitory effect followed by glucuronic acid and galcturonic acid.
Effects of saccharolactone on pollen tubes germination and growth

The effect of SL on pollen tube germination and growth was tested in different conditions: *in vitro*, *in vivo* and semi-*in vivo*.

*In vitro*

Pollen grains have been incubated in the growth medium in the presence or absence of SL at different concentrations. As it can be seen in fig. 2.2, significant reduction of germination was observed only in the presence of 30 mM SL, while an effect on pollen tube growth was observable at 5 mM SL and in the presence of 30 mM SL pollen tube elongation was practically completely inhibited.

Fig. 2.2 *N. alata* pollen *in vitro* germinated in presence of 30 mM (B) saccharolactone after 5 h resulted almost completely scrambled respect to control (A). Images were acquired in transmitting light.

*In vitro* inhibition *N. alata* wild type pollen germination (C) and elongation (D) in presence of different concentration of saccharolactone. Pollen tube elongation (D) resulted inhibited at lower concentration with respect to germination (C). Pollen grains resulted almost all germinated after 1 h (C).
**In semi- v**ivo
Pollinated pistils were cut and immersed in growth medium in the presence or absence of 20 mM saccharolactone.
As it can be seen in fig the presence of SL strongly inhibits the elongation of pollen tubes protruding from the style (fig. 2.3).
No effect was observed when, instead of SL, 20 mM mannitole was added to the growing medium (data not shown).

![Fig. 2.3 Semi- in vivo inhibition of N. alata wild type pollen grown in N. tabacum pistils.](image)
Pollen was allowed to elongate for 48 h in growth medium in presence 20 mM saccharolactone (A) and in its absence (B). pollen tubes were evidenced using aniline blue and images were acquired in epifluorescece.

**In v**ivo
A new technique has been set up to investigate the effects of SL on in vivo grown pollen tubes. *N. tabacum* pistils were collected two hours after pollination with *N. alata* wild type pollen. Without damaging the transmitting tissue, the epidermis of the style was then were immersed in pollen growth medium in the presence or absence of 20 mM saccharolactone.
After 24 hour the style was dissected and the pollen tubes were stained by aniline blue. As it can be seen in fig. 2.4, perforation of epidermis and infiltration of growth medium within the style did not alter pollen tubes elongation along the pistil while, when saccharolactone was present in the growth medium, pollen tubes were not able to grow across the perforated stylar region.
Fig. 2.4 *in vivo* inhibition of *N. alata* wild type pollen grown in *N. tabacum* pistils. The epidermis of the style region at 2 cm below the stigma was then perforated and the pistils the whole style was immersed in medium in absence (A) or presence (B) of 20 mM sacccharolactone. After 24 hour the style was dissecting and the pollen tubes were evidenced by aniline blue. Images were acquired in epifluorescece and then assembled to obtain an overall picture.

**DISCUSSION**

Histochemical analysis in both *N. alata* and *N. tabacum* pollen demonstrated that endogenous GUS activity is only partially inhibited by 1 mM SL and that one order of magnitude higher concentration of SL is required to observe strong inhibition.

Pollen tube germination was less sensitive to the presence of SL with respect to pollen tube growth. In fact 30 mM SL was required to observe a significant effect on pollen tube germination, while pollen tube elongation was sensible to the presence of 5 mM SL and an almost complete inhibition was observable in the presence of 20 mM SL.
Probably, the role of GUS in weakening the cell wall at the germinative pore level, allowing the pollen tube to protrude under the turgor pressure, is less important than the role played in the cell wall remodelling required for pollen tube elongation.

It is also particularly interesting that the SL concentration required to inhibit pollen tube growth is analogous to that reported in literature as inhibitor of root hairs elongation (Sudan et al 2006), process characterized by an apical growth mechanism similar to that present in pollen tubes.

In a more general view, all the processes where GUS activity is required for plant cell wall remodelling, such as root hair elongation and cell elongation, SL behave as inhibitor at 10-20mM concentrations (Schoenbeck et al 2007, Sudan et al 2006). These results are in agreement with observation made on pollen tubes.

To understand the role of GUS in plant cell wall remodelling it has to be considered that, although glucuronic acid is not a main component of plant cell wall, it is present, bound by glycosidic bond, at the end of polysaccharidic chains in emicelluloses (mainly xylans), AGPs and pectins (Eudes et al 2008). Its removal by GUS in wall remodelling is therefore essential to allow the action of the other glycosil hydrolases.

Drastic inhibition of pollen tube growth by SL has been observed also in semi-in vivo and in vivo experiments.

In this last case GUS might play another role in addition to cell wall remodelling. In fact it has been demonstrated that GUS activity is essential for the access of glycosil hydrolases to the polysaccaridic side chain of AGPs (Eudes et al 2008) and AGPs present in the transmitting tissue are deglycosilated by pollen tubes, which probably use them as nutrient resource (Cheung 1995, Wang et al 1993).

The inhibitory effect of SL on pollen tube growth in vivo could be therefore explained by the alteration of two important processes: cell wall remodelling and mobilization of nutrient resources.


CHAPTER 3

GUS genes expression in Arabidopsis thaliana and Nicotiana tabacum

Parts of the contents of this chapter was reported in:

INTRODUCTION


In 2006 histochemical analysis indicated that GUS activity associated to the young portion of the organs, where it seems to be involved in cell elongation, is higher with respect to that observed in the adult tissues (Sudan et al 2006). However, the presence of GUS activity in the apical region of roots is controversial. In fact, on the basis of comparison of plants grown in sterile or non-sterile conditions, Eudes et al. (2008) hypothesized that the histochemical GUS activity observed in roots is due to microflora associated enzyme.

Expression pattern of GUS genes has been also investigated in different organs of A. thaliana (Woo et al 2007). A detailed expression pattern could be also obtained from data made available on Arabidopsis eFP Browser (Winter et al 2007).

In our work we were particularly interested in investigating GUS genes expression in the root and in the pollen.
Concerning the root, we were interested to clarify the controversial data on GUS expression (Eudes et al 2008, Sudan et al 2006), concentrating particularly on border like cells and on the root cap meristem in A. thaliana. During root growth, root cap turnover and programmed cell separation from the cap periphery result in the delivery of detached cell populations into the rhizosphere. The root results enclosed by large populations of detached somatic cells, that are termed root border cells or root border-like cells depending on the pattern of their release and organization (Vicré et al 2005). Border cells are defined as cells that detach from the root cap as individual cells and form small aggregates, whereas border-like cells (present in Arabidopsis) are released as blocks or sheets of cells that remain attached to each other. Homogalcturonans and arabinogalactan-protein seem to play a role in the attachment and organization of the border like cells (Durand et al 2009). GUS enzymes seems to be involved in the deglycosilation of these molecules (Eudes et al 2008). Border cells and border like cells play a key role in controlling root interaction with living microbes of the rhizosphere. As their separation from root tip proceeds, the cells synthesize and secrete a hydrated mucilage that contains polysaccharides, secondary metabolites, antimicrobial proteins and extracellular DNA (exDNA). This exDNA-based matrix seems to function in root defence in a way similar to that of recently characterized neutrophil extracellular traps (NETs) in mammalian cells (Driouich et al 2013). The number of border cells is regulated by environmental stimuli (Hawes et al 2000) and they are formed by mitosis by a group of meristematic cells, distinct from the root apical meristem, named root cap meristem. Mitosis in the cap meristem is suppressed indefinitely (presumably at the G2 phase of the cell cycle) once a species- specific number of border cells accumulates on the cap periphery (Wen et al 2008, Wen et al 2004). Removal of extant border like cells reactivates mitotic division in the root cup meristem (Brigham et al 1998). It has been observed that UDP-glucuronyltransferase (UGT) gene expression is correlated with cell division in the root cup meristem.
Moreover, when *uidA* was expressed under the control of the promoter of a *UDP-glucuronyltransferase* (*PsUGT1*) gene, GUS expression was lethal in all the tested species (pea, alfalfa and *A.thaliana*) and when root tips were incubated in the presence of saccharolactone (SL), an inhibitor of GUS activity, an increase in border cells production was observed, indicating an increase in the mitotic activity of root cap meristem (Wen et al 2004). These data support the hypothesis that *PsUGT1* and GUS operate in tandem in the root cap meristem, reversibly glycosylating a molecule, probably a flavonoid, involved in cell cycle regulation (Woo et al 1999, Woo et al. 2005). In this thesis *GUS* expression in root tip of *A. thaliana* has been analysed by using a recently developed whole mount multi probe *in situ* hybridization technique (Bruno et al 2011).

The interest to investigate GUS gene expression in pollen originates by the observations that GUS activity is histochemically detectable in pollen grains and that GUS plays a role in pollen tube germination and growth (see chapter 2). Little information on the argument are available in literature. Analysis of pollen transcriptome reported in literature revealed the expression of *AtGUS1* and *AtGUS2* (Honys & Twell 2004, Pina et al 2005). However, no indication of GUS expression in pollen grains or germinated pollen tubes was obtained from the analysis of a recently published list of genes expressed in pollen (Hafidh et al 2012).
RESULTS

Histochemical localization of GUS activity

GUS activity was observed by histochemical staining at different stages (4, 7, 10 and 17 day after germination) in N. tabacum, N. alata plants grown in sterile conditions (fig.3.1 and 3.2).

Activity was detectable both in differentiated and apical zones of roots at all the considered stages. However localization and intensity of staining change during plant development (fig. 3.2). Staining appeared more intense in secondary respect to primary root and, usually, in long roots it was mainly limited to the apical part (fig 3.1 and 3.2). Intense staining was observed in root zone attached to the rest of the seeds (fig. 3.1B).

In cotyledons activity was detectable in vasculature, trichomes and apex (fig.3.1B and C).

In S. lycopersicum plants grown in sterile conditions GUS activity was detectable at different stages (at 7 and 10 days) in the same root region, cotyledons and leaves as in Nicotiana (data not shown). Activity was detectable also in the remains of the seed (data not shown).

In A. thaliana GUS activity was detected in different part of the plant: inflorescence, shoot, siliques, leaves and roots (figs.3.3 and 3.4).

In inflorescence GUS activity was detected in filament, anther, apical meristem, pollen grains and pollen tubes growing in the stigma. It was also present in the flower pedicel (fig. 3.3A) and, in siliques, in the placenta and along the fusion line between the two carpels, where the siliqua open (fig. 3.3D).

In the root, GUS activity is present both in the tip and in the differentiated region, where it is more evident. (fig. 3.4).

A change in staining pattern is observable in correspondence of the transition zone between shoot and root: the whole shoot appears stained while root shows the presence of activity only in the central part. This can be explained by the change in stele organization between root and shoot.
As a control that the observed staining was due to endogenous enzymatic activity, it was verified that it was not detectable in the presence of 10 mM saccharolactone or at pH 7 (data not shown).

Fig. 3.1 *N. tabacum* seedlings grown in continuous light stained with X-glu. (A) 17-day seedlings. Staining is detectable in cotyledon and in roots. More the roots are elongated more the activity was limited to the apical part. (B) Root particular of a 7-day seedling. Intense staining is localized in the root zone attached to the rest of the seeds. (B) Cotyledon particular of a 17-day seedling. Trichomes (indicated by arrows) appears stained. Image was acquired in transmitted light.
Fig. 3.2 *N. alata* seedlings stained with X-glu: 4-day (A) and 7-day (B). Staining is clearly detectable in roots. Image was acquired in transmitted light.

Fig. 3.3 *Arabidopsis thaliana* inflorescence and siliques stained with X-glu. (A) Complete Inflorescence. (B) Particular represents a stigma pollinated and an anther. Stained pollen grains are visible in the anther. The intense stained zone in stigma is associated to pollen tubes descending toward ovules. (C) Bottom part of a flower. Abscission zone results stained. (D) Particular of a siliques with an ovule. Staining is visible in the placenta and along the carpel fusion line. Image was acquired in transmitted light.
Expression pattern of \textit{GUS} and \textit{UGT} genes in \textit{A. thaliana}

Specific primers for semiquantitative PCR analysis were designed on the basis of the GUS cDNA sequences reported in literature (see Materials and Methods).

As it can be seen in fig. 3.4, \textit{AtGUS} genes were expressed in almost all the plant organs: in flowers the highest expression levels were observed (fig. 3.4). Regarding shoot apex subsequently verifications allow to determine that all \textit{AtGUS} are expressed in this organ. Among the three genes, \textit{GUS2} seemed to be the most expressed.

We analysed also \textit{UGT85A} gene expression. Their expression resulted more differentiated with respect to that of \textit{AtGUS} genes. \textit{UGT85A1}, \textit{A2} and \textit{A4} were expressed in almost all plant organs. \textit{UGT85A3} and \textit{UGT85A7} resulted the less expressed genes. Regarding shoot apex subsequently verifications allow to determine that all \textit{AtUGT85A} genes, with the exception of \textit{AtUGT85A7}, are expressed in this organ.
These results were in agreement with expression pattern reported in Arabidopsis eFP Browser (Winter et al. 2007).

GUS gene expression was also analysed, in collaboration with University of Calabria, in the root tip at tissue level by whole mount multi probe hybridization technique, recently developed for plant samples (fig. 3.6). While AtGUS1 and AtGUS2 showed a similar expression pattern (epidermis, cortical zone, cap and quiescent centre), AtGUS3 expression was limited to the border like cells.

AtGUS1 and AtGUS2 expression was also detected in secondary root apex (data not shown).

By enhancing the probe signal for AtGUS1 and AtGUS2 mRNA using tyramide amplification system, the observed expression pattern was confirmed (data not shown). These experiments are only at preliminary level.

**Expression pattern of GUS genes in *N. tabacum* and *N. alata***

Specific primers for semiquantitative PCR analysis were designed on the basis of predicted GUS sequences of *N. tabacum* (see Materials and Methods and chapter 4). Primer specificity was verified by sequencing the product of amplification and by comparing the result with GUS gene sequences from *A. thaliana* and *S. lycopersicum*. This allowed to assign the specific primers to the different GUS genes.

Expression of GUS1, GUS2 and GUS3 genes was detected in all vegetative (stem, leaves, root) and reproductive (ovary, pistil, flower) organs considered (fig. 3.5). In the leaves, all genes were expressed both at young and senescing states. Preliminary experiments showed that also GUS3B was expressed (see chapter 4 for GUS genes classification), but these results have to be confirmed.

In *N. tabacum* GUS3 gene was strongly expressed in mature pollen and in pollen tubes at 4 and 24 h after germination (fig. 3.5). At a minor extent, also GUS2 was expressed in pollen, while GUS1 expression was hardly detectable.

Preliminary experiments in *N. alata* showed similar results (data not shown).
Fig. 3.4 Expression pattern of GUS and UGT genes in *A. thaliana*. Different plant parts were considered: cotyledons, shoot apex, young leaves, expanded leaves, caulinar leaves, roots, flowers and whole plants (seedlings). Actin2 was used as internal control.
Fig. 3.5 Expression pattern of GUS genes in different organs of *Nicotiana tabacum*: young leaves (Y.L.), old leaves (O.L.), pistils without ovary (Pi), ovary (Ov), flowers without anther and pistils (Fl), stem (St), roots (Ro), pollen grains (Po), pollen germinated for 4 hours (G.P. 4h), pollen germinated for 24 h (G.P. 24 h.). G= DNA genomic, A= actin, B= GUS, C= GUS2, D= GUS3. Ladder range is from 100 up to 1000 bp.
Fig. 3.6 Multiprobe *in situ* hybridization in *Arabidopsis thaliana* seedling root tips. Images were acquired on a Leica SP2 confocal microscope with a 40X oil immersion objective.

(2) A - Red: *AtGUS2* Digoxigenin riboprobe, sheep anti-DIG and AF555 donkey anti-sheep.

B - Green: *AtGUS1* Biotin riboprobe, mouse anti-BIO and AF488 donkey anti-mouse

C - Blue: *AtGUS3* Fluorescein riboprobe rabbit anti-FITC and AF647 chicken antirabbit.

D - Merge A, B, C. Scale bar 75 µm.

(2) Control performed using sense riboprobe: A - *AtGUS1*, B - *AtGUS2* and C - *AtGUS3*. Note the background fluorescence in all samples, enhanced by high sensibility of observation conditions used. Scale bar 50 µm.
DISCUSSION

GUS activity distribution was histochemically investigated in *N. tabacum, N. alata, S. lycopersicum* and *A. thaliana.*

As previously observed (Sudan et al 2006), GUS activity, confirmed by the expression pattern of the *GUS* genes obtained by PCR analysis, has been detected in all the vegetative and reproductive organs of all tested species, associated to the vascular tissue and to both young and senescent regions. Concerning the conflicting results reported in literature (Eudes et al 2008, Sudan et al 2006), we observed GUS activity in the root tip in sterile grown plants, confirming the results obtained by Sudan and co-workers (2006), in contrast to what reported by Eudes and co-workers (2008), who suggested that GUS activity in the root was due to the microflora.

Particularly interesting results have been obtained by the *in situ* hybridization technique. *GUS3* gene resulted expressed only in the border-like cells, in that region where they detach from the root tip. These results can be related to the observation reported in literature that glucuronyl transferase activity is required for cell-cell adhesion. The presence of glucuronic acid is probably required for the attachment of border like cells to the root tip and its removal by GUS is one of the steps for their detachment. *GUS3* expression seems to be regulated in function of such role.

*GUS1* and *GUS2* are instead expressed in the cortex and in root cup meristem, not in the root tip meristem. Expression in the root cortex might be related to cell wall remodelling in the cell elongation process. The expression in the root cup meristem might instead be related to the inhibition of the mitotic cycle reported to be onset in the root cap meristem when a certain number of border cells are formed (Brigham et al 1998). In fact, it has been hypothesized (Woo et al 2005) that a component, probably a flavonoid, inhibitor of mitotic cycle, is sequestered and released by respectively UDP-glucuronyltransferase (UGT) and GUS, which, by their tandem coordinated action, regulate the mitotic cycle. Experiments are planned to investigate the expression of the different *UGT* genes in the root tip of *Arabidopsis.*
To this aim will be useful the tyramide method, that we have seen can be applied to the whole mount multi probe hybridization protocol (data not shown), with a signal amplification which can allow the detection of low expressed genes.

particularly interesting is also the observation that GUS activity is present in abscission zones, like in the pedicel of the flower, in the placenta and along the fusion line of silique carpels. This finding confirms the role of GUS in the processes of cell wall remodelling, certainly important in the abscission process, and also support the hypothesis of a role of GUS in cell detachment, suggested for border like cells.

GUS activity in *Arabidopsis* was also detected in pollen grains and in pollen tubes growing in the stigma, confirming the observations in *N. tabacum* (see chapter 2). Expression pattern analysis indicated that *GUS3* is more expressed in pollen and pollen tubes with respect to *GUS1* and *GUS2* genes. However quantitative PCR is required to confirm the data.


CHAPTER 4

Identification of GUS genes in N. tabacum and phylogenetic analysis of GH79 GUSs

INTRODUCTION

The first GUS gene sequenced in plants was from Scutellaria baicalensis and it showed high homology to heparanases of family-79 glycosil hydrolase (Sasaki et al 2000). Thereafter, three different GUS genes have been identified in A. thaliana and named: AtGUS1, AtGUS2 and AtGUS3 (Woo et al 2007). Also these genes resulted to be to belong to heparanase of GH-79 family: AtGUS1 code for heparanase 2, AtGUS2 code for heparanase and AtGUS3 code for heparanase 3. Other plant GUS sequences can only be derived by genomic bank analysis.

In the Sol Genomic Network we found the sequences of GUS genes from Solanum lycopersicon that were useful for identification of GUS genes in N. tabacum. Genomic sequences of N. sylvestris and N. tomentosiformis, progenitors of N. tabacum, are only recently available (Sierro et al 2013). According to a bioinformatics analysis, it seems that GUSs of GH-2 family are missing in plants, where are present only GUSs from GH-79 family (Arul et al 2008).

On the basis of the N. tabacum GUS genes sequences that we determined and of the other GUS sequences reported in literature (Sasaki et al 2000, Woo et al 2007) or obtained by bioinformatics genome bank analysis, we derived an interesting phylogenetic analysis. Phylogenetic relationship of GH79 GUSs from plants, fungi and vertebrates was present in the literature, but it was not supported by an adequate statistical analysis (Konishi et al 2008).
RESULTS AND DISCUSSION

Sequence analysis of GUSs genes in A. thaliana, S. lycopersicum, N. sylvestris and N. tomentosiformis

*A. thaliana* β-glucuronidases

*AtGUS1* (AT5G61250), *AtGUS2* (AT5G07830) and *AtGSU3* (AT5G34940), identified in *A. thaliana*, contain 8 introns (fig. 4.1), a number different from what previously reported (Woo et al 2007).

In addition to the canonical TATA-box in the [-39, -26] region they contain other regulative elements, such as TATA-variant sequences (TATAΔ-PLMs), and, except *AtGUS2*, also TC-elements (tab. 4.1).

These last elements are a class of novel regulatory motifs that seems to be involved in the complex modulation of gene expression in plants (Bernard et al 2010).

![Fig. 4.1 Structure of A. thaliana GUS genes.](image)

<table>
<thead>
<tr>
<th>Gene</th>
<th>TATA-box PLM</th>
<th>TATAΔ-PLMs</th>
<th>TC[-39,-26]-PLMs</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>AtGUS1</em></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><em>AtGUS2</em></td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><em>AtGUS3</em></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Tab. 4.1 Presence of regulatory elements involved transcription in *A. thaliana* GUS gene promoter sequences.
S. lycopersicum β-glucuronidases

The *S. lycopersicum* genome (T.G.C. 2012) has been recently made available on Solanum Genomics Network and GenBank (NCBI) and it was adopted as reference species for further studies in *Solanaceae.* It was possible to retrieve in this database five GUS genes: two GUS1 (*SlGUS1A* and *SlGUS1B*), one GUS2 (*SlGUS2*) and two GUS3 (*SlGUS3A* and *SlGUS3B*) (accession numbers are reported in tab. 4.2). Unlike *A. thaliana*, in *S. lycopersicum* GUS1 and GUS3 genes are present in duplicate. *SlGUS1A* maps on chromosome 5, *SlGUS1B* on chromosome 6, *SlGUS2* on chromosome 3, *SlGUS3A* on chromosome 7 and *SlGUS3B* on chromosome 10. All genes contain 8 introns (fig. 4.2). The positions and dimensions of introns are different also in the duplicated genes. Exons, although separated by introns with different length in the different genes, show a similar dimension in all genes.

Regarding regulatory elements, all the genes do not present canonical TATA-box in the [-39,-26] region but rather TATA-variant sequences and TC-elements (tab. 4.2).

![Fig. 4.2 Structure of S. lycopersicum GUS genes.](image-url)
Starting from *S. lycopersicum* GUS sequences, it was possible to retrieve five GUS genes in *N. tomentosiformis* and *N. sylvestris* genomes (tab. 4.3).

<table>
<thead>
<tr>
<th>Gene</th>
<th>GenBank accession number</th>
<th>Chromosome</th>
<th>TATA-box PLM</th>
<th>TATA∆-PLMs</th>
<th>TC[-39,-26]-PLMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIGUS1A</td>
<td>XM_004239417</td>
<td>5</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>SIGUS1B</td>
<td>XM_004241543</td>
<td>6</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>SIGUS2</td>
<td>XM_004235932</td>
<td>3</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>SIGUS3A</td>
<td>XM_004242684</td>
<td>7</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>SIGUS3B</td>
<td>XM_004249554</td>
<td>10</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Tab. 4.2 Information regarding *S. lycopersicum* GUS genes: accession number, chromosome on which map each gene and presence of transcriptional regulatory elements

*N. tabacum* progenitors β-glucuronidases

Starting from *S. lycopersicum* GUS sequences, it was possible to retrieve five GUS genes in *N. tomentosiformis* and *N. sylvestris* genomes (tab. 4.3).

<table>
<thead>
<tr>
<th>Species</th>
<th>Genomic GenBank accession number</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>N. tomentosiformis</em></td>
<td>ASAG01046754</td>
<td>NtomGUS1A</td>
</tr>
<tr>
<td><em>N. tomentosiformis</em></td>
<td>ASAG01017701</td>
<td>NtomGUS1B</td>
</tr>
<tr>
<td><em>N. tomentosiformis</em></td>
<td>ASAG01011971</td>
<td>NtomGUS2</td>
</tr>
<tr>
<td><em>N. tomentosiformis</em></td>
<td>ASAG01070395</td>
<td>NtomGUS3A</td>
</tr>
<tr>
<td><em>N. tomentosiformis</em></td>
<td>ASAG01008598</td>
<td>NtomGUS3B</td>
</tr>
<tr>
<td><em>N. sylvestris</em></td>
<td>ASAF01000528</td>
<td>NsGUS1A</td>
</tr>
<tr>
<td><em>N. sylvestris</em></td>
<td>ASAF01014810</td>
<td>NsGUS1B</td>
</tr>
<tr>
<td><em>N. sylvestris</em></td>
<td>ASAF01019790</td>
<td>NsGUS2</td>
</tr>
<tr>
<td><em>N. sylvestris</em></td>
<td>ASAF01048865</td>
<td>NsGUS3A</td>
</tr>
<tr>
<td><em>N. sylvestris</em></td>
<td>ASAF01087681</td>
<td>NsGUS3B</td>
</tr>
</tbody>
</table>

Tab. 4.3 Accession numbers of the *N. tomentosiformis* and *N. sylvestris* genomic sequences in which GUS genes were identified.
All the identified genes show the same numbers of introns (8) present in *S. lycopersicum* (fig. 4.3). *NtomGUS1B* differently from other genes presents a very long intron (ca 3 kb).

Exons, although separated by introns with different length in the different genes, show a similar dimension in all genes.

Traditional TATA-box result absent in all genes, while other regulatory elements are present (tab4.4).

![Gene Structure](image)

**Fig. 4.3 Structure of *N. sylvestris* and *N. tomentosiformis* GUS genes.**

<table>
<thead>
<tr>
<th>Gene</th>
<th>TATA-box PLM</th>
<th>TATAΔ-PLMs</th>
<th>TC[-39,-26]-PLMs</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>NsGUS1A</em></td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td><em>NsGUS1B</em></td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td><em>NsGUS2</em></td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td><em>NsGUS3A</em></td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td><em>NsGUS3B</em></td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td><em>NtomGUS1A</em></td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td><em>NtomGUS1B</em></td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td><em>NtomGUS2</em></td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td><em>NtomGUS3A</em></td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td><em>NtomGUS3B</em></td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

**Tab. 4.4 Transcriptional regulatory elements in *N. sylvestris* and *N. tomentosiformis* GUS gene sequences.**
Identification and sequencing of GUS genes in *N. tabacum*

Likely partial sequences of *GUS1*, *GUS2* and *GUS3* of *N. tabacum* were determined by bioinformatics analysis of EST database (NCBI). The same specific primers utilized to investigate *GUS* expression (chapter 3) or expressly designed primers were used for PCR amplification of cDNA. Obtained amplification bands were sequenced and the cDNA sequence was completely (*GUS2*) or partially (*GUS1* and *GUS3*) determined (see Materials and Methods).

Only one EST was determined for *GUS1* sequence: SGN-U493366 (Solanum Genomics Network). *GUS1* partial sequence resulted homologue to *SlGUS1B* and showed strong homology to *N. tomentosiformis GUS1B*. The presence of *GUS1B* derived from *N. sylvestris* has not demonstrated yet. Nothing is yet known regarding the presence of *GUS1A* in *N. tabacum* genome.

To predict *GUS2* sequence, the following EST sequences were aligned: FG139192, EB452249, EB102900, EB680524, AM846742, FG152993, AM846742 and EB102900. A complete coding sequence of *GUS2*, with partial 5’ and 3’ UTR, was obtained and it displayed strong homology with the *N. tomentosiformis GUS2* gene. The sequence was published on GenBank (Accession Number: KF148025). The presence of *GUS2* derived from *N. sylvestris* has not been observed.

To predict *GUS3* sequence, the following EST sequences were aligned: SGN-E824103 and SGN-E1088185 (Solanum Genomics Network).

Specific primers were designed on the basis of the reconstructed sequence and only one band was obtained by PCR amplification of cDNA. However, sequence analysis revealed that more than one gene was amplified (fig.4.4). Genomic DNA amplification with the same primers gave two distinct bands (fig. 3.5). Sequence analysis demonstrated that *GUS3B* genes from both progenitors are present in *N. tabacum*.

Regarding *GUS3A*, preliminary experiments indicate its presence in *N. tabacum*, but it has to be confirmed by sequence analysis. At the moment, it is not known if *N. tabacum* genome contains *GUS3A* genes of both progenitors.
Fig. 4.4 Part of electropherogram of GUS3 cDNA sequence of *N. tabacum*. Double peak are in correspondence of the difference in sequence between *N. tomentosiformis* and *N. sylvestris* GUS3 genes.

**Comparison of GUS sequence analysis in the different species**

In *A. thaliana* three GUS genes have been identified (Woo et al 2007), while five GUS genes are found in the *S. lycopersicum, N. sylvestris* and *N. tomentosiformis* genomes.

*N. tabacum*, being an allotetraploid species originated from *N. sylvestris* and *N. tomentosiformis*, could potential have ten GUS genes. Actually, experimental evidence support the presence of at least five GUS genes and only one (*NtGUS2*) have been completely sequenced in its coding part. Thus it has to be demonstrated if the remaining genes are effectively present or if they have been lost during genome shrinkage process (see general introduction).

The analysis of gene structures shows that in each considered species GUS genes maintain the same number of introns and the exons show similar dimension.

Regarding identification of transcription regulatory elements, determined by the approach proposed by Bernard and co-workers (Bernard et al 2010), canonical TATA-box were present in *A. thaliana* GUSs but not in *Solanaceae* GUSs.
TATA-variant sequences, which in plants are generally observed in gene with a broad expression pattern, are present in all GUSs of the species considered, while the presence of TC-elements is not so general.

**Protein sorting, motif analysis and post translational modifications**

β-glucuronidase aminoacidic sequences, deducted from the cDNA nucleotide sequence, were analysed by bioinformatics methods (see Materials and Methods) to find the presence of functional domain architecture. GUS proteins of *A. thaliana* and *Scutellaria baicalensis* were also included in the analysis.

Aminoacids motifs related to heparanase (β-glucuronidase), GH79 family and (trans)glycosidases superfamily were present. The glutamic acid, identified in position 212 of *Scutellaria baicalensis* GUS as characteristic of the active site of glycoside hydrolases (Sasaki et al 2000), is conserved in all examined GUSs. Also the nucleophile residue, characteristic of the active site of heparanase, results conserved.

Another common characteristic is the presence of an N-terminal signal peptide (fig.4.4B). For few of them, like SlGUS1B, there was a consensus between all the sorting prediction programs used.

However, with the exception of WolfPsort program, there was enough consensus for the sorting of the different GUSs, mainly to the plasma membrane or to the cell wall/extracellular space, a localization consistent with the identification of β-glucuronidase among the cell wall proteins in *Arabidopsis thaliana* and *Scutellaria baicalensis* (Bayer et al 2006, Irshad et al 2008, Minic et al 2007).

The presence of extensive post-translational modifications has been suggested, on the basis of bioinformatics predictions, for *A. thaliana* β-glucuronidases (Eudes et al 2008, Woo et al 2007). By ScanProsite program we verified the existence of possible sites for post-translational modifications in the deduced GUS sequences. The analysis showed the presence of multiple putative phosphorylation sites and of several N-glycosylation sites.
Fig. 4.5 Some of the conserved motifs obtained by protein sequences analysis with MEME Suite. (A) Heparanase mofis (consensus of all proteins considered) contains Glu (E) residue of active site (indicated by arrow). (B) N-terminal signal peptide: consensus motif in NsGUS2, NtomGUS2, NtGUS2, NtomGUS1B, NsGUS1B, SIGUS1B, SIGUS2; however is present also in other protein considered (see appendix). Detailed explanation is reported in text.

<table>
<thead>
<tr>
<th>Protein</th>
<th>WP</th>
<th>Ps</th>
<th>ESL</th>
<th>CELLO</th>
<th>TP</th>
<th>PS</th>
<th>SiP</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIGUS1A</td>
<td>Mit</td>
<td>Ext</td>
<td>Ext</td>
<td>Lys, Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>SIGUS1B</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>SIGUS2</td>
<td>Vac</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>SIGUS3A</td>
<td>PIM</td>
<td>PIM</td>
<td>Ext</td>
<td>PLM, Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>SIGUS3B</td>
<td>Chl</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext, PLM</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NsGUS1A</td>
<td>Nuc, End, Cyt</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NsGUS1B</td>
<td>PIM</td>
<td>PIM</td>
<td>Ext</td>
<td>Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NsGUS2</td>
<td>Vac</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext, Lys</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NsGUS3A</td>
<td>Vac</td>
<td>PIM</td>
<td>Ext</td>
<td>PIM</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NsGUS3B</td>
<td>PIM</td>
<td>PIM</td>
<td>Ext</td>
<td>Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NtomGUS1A</td>
<td>Vac</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NtomGUS1B</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext, Lys</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NtomGUS2</td>
<td>Vac</td>
<td>Ext</td>
<td>Ext</td>
<td>Lys</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NtomGUS3A</td>
<td>PIM</td>
<td>PIM</td>
<td>Ext</td>
<td>PIM</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NtomGUS3B</td>
<td>Chl</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
<tr>
<td>NtGUS2</td>
<td>Ext</td>
<td>Ext</td>
<td>Ext</td>
<td>Lys</td>
<td>SP</td>
<td>SP</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Tab. 4.5 Prediction sorting of GUS proteins of *Solanum lycopersicum, Nicotiana tomentosiformis, Nicotiana sylvestris* and *N. tabacum*.

Abbreviations: chloroplast (Chl), cytoplasm (Cyt), endoplasmic reticulum (EnR), Extracellular (Ext), mitochondrion (Mit), nucleus (Nuc), plasma membrane (PLM), secretory pathway (SP) and vacuole (Vac). WP: WolfPsort, Ps:Psort, TP: TargetP, PS: PrediSi, SiP: SignalP. SignalP predicts the presence of the signal peptide. References relative to software used in analysis are reported in materials and method.
Tab. 4.6 Numbers of post-translational modifications in the deduced protein
determined using ScanProsite. Abreviations: casein kinase II (CK2), N-glycosiation
sites (N-Glyc), protein kinase C(PKC), tyrosin kinase (TyrK), cAMP-kinase and
gGM-kinase (cAMP-cGMP)

<table>
<thead>
<tr>
<th>Protein</th>
<th>Serin protease</th>
<th>CK2</th>
<th>N-Glyc</th>
<th>PKC</th>
<th>TyrK</th>
<th>cAMP-cGMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>NtomGUS1A</td>
<td>1</td>
<td>13</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NtomGUS1B</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NtomGUS2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>NtomGUS3A</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>10</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NtomGUS3B</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NsGUS1A</td>
<td>1</td>
<td>13</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>NsGUS1B</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>NsGUS2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NsGUS3A</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NsGUS3B</td>
<td>1</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SlGUS1A</td>
<td>1</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SlGUS1B</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SlGUS2</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SlGUS3A</td>
<td>1</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SlGUS3B</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Phylogenetic analysis

Phylogenetic analysis was conducted on the determined *N. tabacum* GUS
sequences and on that reported in the literature or derived by genome bank
analysis.

As it can be seen in fig. 4.6, plant GUSs can be clustered in three groups: α, β
and γ. The group α comprehends GUS1 and GUS2 genes, the group β
comprehends GUS1 genes and the group γ GUS3 genes. In the group α is
possible to identify four main subgroups (A, B, C, D). The subgroup A
comprehends Solanaceae GUS2 and GUS1, subdivided in two clusters. The
subgroup B comprehends GUS1 and GUS2 genes of different families. The
subgroup C comprehends GUS1 and GUS2 genes of Arabidopsis
(Brassicaceae). The subgroup D comprehends GUS1 and GUS2 genes of the
monocots Poaceae and Musaceae.
In the group γ there are three subgroups (A, B, C). The subgroup A comprehends genes of Poaceae and Musaceae. The subgroup B comprehends GUS3B of Solanaceae. The subgroup C comprehends different GUS3 of different families and a cluster formed by GUS3A of Solanaceae.

This analysis of the phylogenetic tree indicates that GUS1A and GUS2 are paralog genes, possibly derived from a process of duplication of a GUS gene, which evolved separately form GUS1B and GUS3 genes. The observation that GUS1A and GUS2 in Solanaceae cluster separately and in the other subgroups (Arabidopsis, Poaceae, other families) are both present indicates that the duplication event repeated more times. Also GUS3A and GUS3B are paralog genes, which derived from a process of duplication and further data are required to verify if, in this case, there was one or more duplication events. It might be interesting to understand the selective pressure that acted to duplicate GUS genes in Angiosperms.

Separate phylogenetic analysis of genomic GUS sequences of Solanaceae (fig. 4.10) and Arabidopsis (fig. 4.11 ) confirmed the clustering of GUS genes and indicated that the cluster of GUS1A is less distant from the cluster of GUS1B-GUS2 genes than from those of GUS3 gene.

Overall analyses of the phylogenetic tree indicate confusion in the denomination of GUS genes. In fact, no Arabidopsis GUS is included in the GUS1A cluster, while the two genes named GUS1 and GUS2 in Arabidopsis are paralog genes, derived probably by GUS duplication in the Brassicaceae family. GUS1A is clustering in a quite different group with respect to GUS1B. GUS1A genes ware deposited in GenBank classified as GUS1 on the bases of homology with Arabidopsis GUS genes.

To avoid confusion, we propose that GUS1A could be named GUS4 and the same name (GUS1 or GUS2) could be given to all the GUS genes present in the group α.

In alternative, the names of GUS1 and GUS2 genes should be changed when their phylogenetic evolution is clearer, leaving the same name to the GUS gene ancestor to the entire group α and giving a different name to all the other GUSs derived from gene duplication in the different subgroups.

Gymnosperm, bryophyte and algae were not included in the analysis because of the scarcity of available sequences.
However, it was possible to identify two sequences of gymnosperm homologue to GUS genes in *Picea glauca* (BT109195) and *Picea sitchensis* (EF676824) and one sequence of briophytes in the genome of *Physcomitrella patens* (XM_002512068).

The determination of GUS sequences in gymnosperms, vascular seedless plants, bryophytes and *Charophyta* will allow rooting the phylogenetic tree, allowing inferences on GUS evolution in land plants.
Fig. 4.6 Phylogenetic trees obtained by distance matrix (kimura-2-parameter model) /UPGMA comparison of the β-glucuronidase/heparanase cDNA sequences of plant species evaluated in this study. Original dataset was bootstrapped 1000 times and the obtained values are showed at corresponding node. Sequence accession numbers and Nicotiana sequences are reported in appendix. In some cases gene name is followed by a letter only in order to distinguish, when necessary, one gene from the other. This classification does not have relationship with that adopted for Solanaceae. Detailed explanation is reported in the text.
Fig. 4.7 Close-up of tree reported in fig. 4.6. In evidence GUS genes of Solanaceae (A) and Arabidopsis (B).
Fig. 4.8 Close-up of tree reported in fig. 4.6. In evidence GUS genes of Solanaceae (A) and Arabidopsis (B).
Fig. 4.9 Close-up of tree reported in fig. 4.6. In evidence GUS genes of Solanaceae.
Fig. 4.10 Phylogenetic trees obtained by distance matrix (kimura-2-parameter model) /UPGMA comparison of the β-glucuronidase/heparanase genomic sequences of Solanaceae species evaluated in this study. Original dataset (genomic sequences) was bootstrapped 1000 times and the obtained values are showed at corresponding node. GUS genomic sequences of *N. tomentosiformis* and *N. sylvestris* were obtained as explained in materials and method. Other Solanaceae GUS genomic sequences were obtained from GenBank.
Fig. 4.11 Phylogenetic trees obtained by distance matrix (Kimura-2-parameter model) / UPGMA comparison of the β-glucuronidase/heparanase genomic sequences of Arabidopsis species evaluated in this study. Original dataset (genomic sequences) was bootstrapped 1000 times and the obtained values are showed along branches. Arabidopsis GUS genomic sequences were obtained from GenBank.


MATERIALS AND METHODS

Plant material

*Nicotiana tabacum* (cv. Samson) and *Nicotiana alata* Link et Otto plants were grown in controlled environmental conditions under a 12-hour photoperiod at 26/22°C day/night temperature. Light was provided by 400 W Philips HDK/400 lamps. *N. tabacum* plants were also grown from May to September in the garden.

*N. alata* plants transformed with the construct LAT52 promoter – *E. Coli GUS* gene were kindly donated by E. Newbigin and M. Lush of the Plant Cell Biology Research Centre (University of Melbourne).

*N. tabacum* (cv. Samson) seeds were surface sterilized with 96% (v/v) ethanol for 2 min and with mixture 1 part bleach to 3 parts sterile water for 10 min. After five washes, 5 min each, with sterile distilled water, seeds were sown on plates. *Nicotiana* seeds were sown on plates containing MS medium (4.7 g/L), 3% sucrose, 0.8 % agar, pH 5.8. Seedlings were grown in growth chambers with the following conditions: 14 h light (27°C) and 10 h darkness (25°C). In some cases *Nicotiana* seeds were grown under continuous light at 25°C (Sudan et al 2006).

*Solanum lycopersicum* cv. Heinz 1706, kindly donated by Rich Ozminkowski and Claudio Leggeri of Heinz Company, were grown at the same sterile conditions of *N. tabacum*.

*Arabidopsis thaliana* (cv. Columbia) wild type was used for the different experiments. Seeds were surface sterilized as indicate before. Seeds were germinated and grown on plates containing MS medium (4.7 g/L), 1 % sucrose and 1 % agar for plant culture. The plated seeds were left in dark conditions at 4°C for 48 h to ensure uniform germination, and then moved to a growth chamber at 22° (±2) C, under 16 h light and 8 h dark. In some cases was maintained condition of continuous light.

Plants were also grown in controlled environmental conditions under a 12-hour photoperiod at 22 (±2)°C temperature. Light was provided by 400 W Philips HDK/400 lamps.
Pollination techniques

Pollen of transformed and wild type *N. alata* was used to pollinate pistils of *N. tabacum* in a compatible cross. Flowers attached to plants were emasculated 2 days before anthesis, and covered by a gauze layer to avoid cross-contamination. At anthesis the flower were detached from plants and placed with their cut ends in water and maintained at environment temperature. The stigma was covered by a drop of maize oil and pollinated by a stick with pollen collected from plants of *N. alata*. Then pistils were removed from flower at different time from pollination according to the needs of the assay.

The presence of pollen tubes were verified by staining with aniline blue (0.1% in K$_3$PO$_4$ 0.1 M), that selectively stain callose, typical component of pollen tubes wall.

Pollen germination and growth

The medium used for *N. tabacum* pollen was 1.62 mM H$_3$BO$_3$, 1.25 mM Ca(NO$_3$)$_2$ & H$_2$O, 2.97 mM KNO$_3$, 1.65 mM MgSO$_4$·7H$_2$O, 12 % sucrose (Brewbaker & Kwack 1963). A new medium was developed for *N. alata* pollen: 12.5% PEG 6000, 5% Suc, 1.0 mm CaCl$_2$, 1.0 mm KCl, 0.8 mm MgSO$_4$, 1.6 mm H$_3$BO$_3$, 0.03% casein acid hydrolysate, 30 µM CuSO$_4$, 70 µM 2-thiouracil and 25 mM MES pH 5.9 (Lush et al 1997, Read et al 1993).

In vitro. Pollen grains were pre-hydrated in a Petri dish on a glass slides put on imbibed Whatman filter paper. They were suspended, at a concentration of 5 mg/mL, for 1 h in the appropriate pollen medium. 20 µL of suspension were put on a poly-D-lysine (2mg/mL poli-D-lysine hydrobromide) coated glass slides or well. Then pollen were maintained in the dark at room temperature.

In some cases 1% agarose low melting was uniformly add, spread and, when solidified, covered with appropriate pollen medium.

Pollens tubes length was measured using ImageJ.

Semi-in vivo. Pollinated pistils of *N. tabacum* were collected at different times after pollination.
Tab. 1 Position reached by pollen tubes during their growth in style at different times after pollination according previous experiments conducted in our laboratory (data not published)

<table>
<thead>
<tr>
<th>Time after pollination (h)</th>
<th>Range of distance from base of the stigma (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.1</td>
</tr>
<tr>
<td>8</td>
<td>0.1-0.3</td>
</tr>
<tr>
<td>10</td>
<td>0.3-0.5</td>
</tr>
<tr>
<td>12</td>
<td>0.3-0.7</td>
</tr>
<tr>
<td>14</td>
<td>0.5-0.7</td>
</tr>
<tr>
<td>16</td>
<td>0.7-1.1</td>
</tr>
<tr>
<td>18</td>
<td>1.1-1.3</td>
</tr>
<tr>
<td>20</td>
<td>1.3-1.5</td>
</tr>
<tr>
<td>22</td>
<td>1.3-1.7</td>
</tr>
<tr>
<td>24</td>
<td>1.5-1.9</td>
</tr>
<tr>
<td>26</td>
<td>1.7-2.3</td>
</tr>
<tr>
<td>28</td>
<td>2.0-2.5</td>
</tr>
<tr>
<td>30</td>
<td>2.3-2.7</td>
</tr>
<tr>
<td>32</td>
<td>2.5-3.0</td>
</tr>
<tr>
<td>34</td>
<td>2.7-3.3</td>
</tr>
<tr>
<td>36</td>
<td>2.9-3.5</td>
</tr>
</tbody>
</table>

The pistils were cut at a distance from stigma selected in function of the position reached by pollen tubes during their growth in style after pollination, (Tab. 1). Then pistils were put vertically in vials containing the appropriate medium.

Pistils were maintained at room temperature in condition of humidity.

The emergence of pollen tubes from cut styles was followed during time.

**Hystochemical GUS assay**

Hystochemical GUS assay was performed by incubating plant materials with X-glu staining solutions. 0,1 M X-glu (5-Bromo-4-chloro-3-indolyl-beta-D-glucoside, Sigma-Aldrich) solution was prepared in DMSO. X-glu was then diluted at a final concentration of 1 mM in 10 mM EDTA, 0,5 mM potassium ferrocyanide, 0,5mM potassium ferricyanide, 0.1 M phosphate, pH 7 (Naleway 1992).
Endogenous GUS activity was detected by using 0.1 M acetate buffer, pH 4.5, instead of phosphate buffer. The samples were vacuum infiltrated for 30 min (ca. 200 mbar) to facilitate penetration of the assay buffer. Samples were incubated 6 h at 37°C and then overnight at room temperature. The stained samples were immersed in absolute ethanol to remove chlorophyll. The ethanol solution was changed at least three times at 1 h intervals. Samples were observed under a stereomicroscope (Leica) and epifluorescence microscope (Olympus IX 51 Microscope) in transmitted light or fluorescent light.

ELF®97 β-D-glucuronide (ELF®97) (Life Technologies) was used as alternative substrate for histochemical assay (Zhou et al 1996). ELF®97 was dissolved at a concentration of 0.1 mM in 0.1 M phosphate buffer (pH 7). Samples were incubated with the substrate at room temperature in dark condition. Samples were observed with epifluorescence microscope using FITC filter (Olympus IX 51 Microscope).

**Immunohistochemical GUS assay**

Pollen tubes growth in vitro were fixed in 3% formaldehyde for 30 minutes at environment temperature, washed three times with phosphate-buffered saline solution (PBS) and incubated with cellulase (2% w/v in PBS) (Sigma-Aldrich) for 3 minutes in dark condition. GUS immunohistochemical assay was performed by using Anti-β-Glucuronidase (N-Terminal) antibody produced in rabbit (Sigma-Aldrich) and Tyramide Signal Amplification kit with HRP—goat anti-rabbit IgG and Alexa Fluor® 488 (Life Technologies). Samples were observed with epifluorescence microscope (Olympus IX 51 Microscope). Specificity of Anti-β-Glucuronidase (N-Terminal) antibody in detecting E. coli GUS present in pollen extract was confirmed using western blot.

**DNA and RNA extraction**

Genomic DNA was extracted from fresh leaves (60 mg of tissue) using Invisorb® Spin Plant Mini Kit (Stratec molecular, Berlin) according to the manufacturer’s protocol.
For RNA extraction, *N. tabacum* and *A. thaliana* samples were grinded in liquid nitrogen and in some cases (pollen and roots) in the presence of quartz sands. Extraction was then performed by Trizol® Reagent method (Invitrogen-Life Technologies), following manufacturer’s protocol. After extraction, RNA was treated with DNase I, following manufacturer’s protocol, to eliminate any genomic contamination (Ambion-Life Technologies) and resuspended in 10 mM Tris-HCl buffer pH 7.5.

RNA integrity was controlled by gel electrophoresis. RNA was quantified by spectroscopic technique measuring absorbance at 260 nm and considering that an absorbance of 1 unit at 260 nm corresponds to 40 μg/mL. RNA purity was determined by measuring the $A_{260}/A_{280}$ ratio.

**Reverse transcription and PCR amplification**

RNA obtained were reverse transcribed into first-strand cDNA with SuperScript® III First-Strand Synthesis System for RT-PCR and oligo dT primer (Invitrogen-Life Technologies).

Gene specific primers for cDNA amplification were synthesized by Life Technonologies (Carlsbad, USA) and IDT (Coralville, USA). The sequence was generally determined by the software Primer 3 (developed by Steve Rozen Helen J. Skaletsky, 1996, 1997) free available on-line. When necessary primers were designed by hand and checked by OligoCalc (Kibbe 2007) and PerlPrimer v. 1.1.21.

When possible primers were selected in such a way that the 5' and the 3' primers span different exons, so that the amplification product obtained from the cDNA would be of different length from that obtained from any contaminant genomic DNA comprising intron sequences.

Specificity of primer was verified by sequencing PCR products.

All sequencing reactions were performed by BMR Genomics (Italy, Padova).
According to the GUSs and Actin gene sequences reported in literature (Bréhélin et al 2003, Woo et al 2007), the following gene specific primers for A. thaliana were obtained:

AtGUS1 FW 5'-TTTGGTCCGAGGTTTCTACG-3'
AtGUS1 BW 5'-TTCAAAATATCCGTCCGAAAGC-3'
AtGUS2 FW 5'-GCTACGGGTTTACGCACATT-3'
AtGUS2 BW 5'-CAACACACCGTTTTCTGGTG-3'
AtGUS3 FW 5'-CTGGACCAAAGAGGCAAAAG-3'
AtGUS3 BW 5'-TTGTCTTGCACAATGGGTGT-3'
AtUGT85A1 FW 5'-GGGTGGAACTCGATATTGGA-3'
AtUGT85A1 BW 5'-CTAAGCGCTGCCACTCTACC-3'
AtUGT85A2 FW 5'-GAGGATGTTGGCAAGTTGGT-3'
AtUGT85A2 BW 5'-CAACACACCGTTTTCTGGTG-3'
AtUGT85A3 FW 5'-TTCCCAAGAGAACCTTGTTGA-3'
AtUGT85A3 BW 5'-TTGTCTTGCACAATGGGTGT-3'
AtUGT85A4 FW 5'-TCGGGAGTCTAACGGTTTTG-3'
AtUGT85A4 BW 5'-ACACCGGCGTACAAACTCTC-3'
AtUGT85A7 FW 5'-AGAGTCTCGCTGGTGGTG-3'
AtUGT85A7 BW 5'-AGATTGATGACCGACGACC-3'

Semiquantitative PCR was conducted using Taq DNA Polymerase recombinant (Invitrogen-Life Technologies) and the following parameters: 95°C for 4 min, 28 cycles of 95°C for 50 sec 55°C for 40 sec and 72°C for 1 min and the last step of 72°C for 5 min. Actin gene was use as control and internal standard for semiquantitative evaluation of genes expression. The following primers for N. tabacum were obtained by using predicted sequences derived from EST (NCBI) analysis (see following paragraph sequence precicion):

NtGUS1 FW 5'-GACACAAGCACATTTATCCC-3'
NtGUS1 BW 5'-GGTGGTTGCACACTAGATCG-3'
NtGUS2 FW 5'-CTTCTATTAGCACCAGGAGG-3'
NtGUS2 BW 5'-GCCAAGCTGATCTAAGTACC-3'
NtGUS3 FW 5'-TCGGGAGTCTACGTTTTG-3'
NtGUS3 BW 5'-ACACCGGCGTACAAACTCTC-3'
NtACT FW 5'-CAGGAGAGAGGTTACATG-3'
NtACT BW 5'-TGGAGTTGAGGTAGTCTC-3'.

As internal control was used actin gene (Accession number GQ339768).
PCR reactions were performed using Taq DNA Polymerase Recombinant (Life Technologies -Invitrogen).

The following reaction conditions were used: initial denaturation 95°C for 1 min., denaturation 95°C for 45 sec., annealing 55°C for 45 sec., extension 72°C for 1 min., final extension 72°C for 5 min, 4°C ad libitum. Steps from denaturation to extension were repeated for 35 cycles. Gel electrophoresis of PCR products were conducted in some cases with MetaPhor agarose (Cambrex Bio Science) instead traditional agarose in order to obtain an optimal separation of different bands.

**Multi-probe in situ hybridization in Arabidopsis**

Multi-probe *in situ* hybridization were performed according to the method previously published (Bruno et al. 2011). Addition of tyramide amplification stage to multi-probe method is only at preliminary level so protocol isn’t reported here. Synthesis of labelled RNA probes were obtained by using in vitro transcription by RNA polymerase T7 or SP6 (DIG, Biotin, FITC RNA labeling Mix, Roche) in the presence of Digoxigenin-11-UTP, Biotin-16-UTP or Fluoroscein-12-UTP (Hejatko et al. 2006). cDNA was used as template and the primers were the same used in semiquantitative PCR, with addition of an extension (CCAAGCTTCTAATACGACTCACTATAGGGAGA) at 3’ end of forward primers.

**Predicted sequences verification**

PCR reaction were performed using Platinum Taq Polymerase High Fidelity (Life Technologies-Invitrogen) according to the manufacturer’s protocol. The amplification products were excised from the gel and sent to be sequenced by BMR Genomics (Italy, Padova).

The sequence of *NtGUS2* was obtained also by cDNA amplification with the following primers:

FW 5’TGACTCGGTTCAAACAGTGTT-3’

BW 5’-ACTAGTGATTAGCGTGGTGC-3’. 
The 3’ end part of \textit{NiGUS2} sequence was completed using 5’/3’ RACE Kit 2nd Generation (Roche) and the following primer: 5’-GGTACTTAGATCAGCTTGGC-3’.

\textbf{Sequence retrieval from databases}

A research in different databases was performed to retrieve β-glucuronidase (heparanase) sequences of different plant species: GenBank, Solanum Genomics Network (http://solgenomics.net), Genome Database for Rosaceae (www.rosaceae.org), and The Banana Genome Hub (http://banana-genome.cirad.fr). Orthologous GUS genes were obtained by using as query in BLAST research the known sequences of \textit{Arabidopsis thaliana} genes (Woo et al 2007).

\textbf{Sequence prediction}

\textit{Nicotiana tabacum} \textit{GUS} gene sequences, before the availability of progenitors genome, were predicted by CLUSTAL Omega (EBI) aligning the homologue genes of \textit{Arabidopsis thaliana} and \textit{Solanum lycopersicum} (previous retrieved from database) with the Expressed Sequence Tag (NCBI and Solanum Genomics Network) of \textit{Nicotiana tabacum}.

<table>
<thead>
<tr>
<th>Gene</th>
<th>\textit{Arabidopsis thaliana}</th>
<th>\textit{Solanum lycopersicum}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{GUS1}</td>
<td>NM_125518</td>
<td>XM_004241543</td>
</tr>
<tr>
<td>\textit{GUS2}</td>
<td>NM_120865</td>
<td>XM_004235932</td>
</tr>
<tr>
<td>\textit{GUS3}</td>
<td>NM_180762</td>
<td>XM_004249554</td>
</tr>
</tbody>
</table>

\textit{GUS} gene sequences of \textit{Nicotiana tomentosiformis} and \textit{Nicotiana sylvestris} were identified by aligning their genomics sequences with cDNA sequences of \textit{S. lycopersicum} obtained from GenBank (see tab). Sequences were aligned using CLUSTAL Omega (EBI) and BLAST (NCBI).
Phylogenetic analysis

The phylogenetic reconstruction was performed by the software package MEGA 5.2 (Tamura et al 2011). The original cDNA sequences were first aligned by ClustalW. The alignment was resampled by bootstrap 1000 times. Distance matrices were obtained according to the Kimura-2 parameters method (Kimura 1980) and clustering was obtained by the UPGMA method. Alternatively, the tree was built by Maximum-Likelihood (ML) with the general time reversible model and the results displayed by Neighbour-Joining (NJ).

Proteins alignment, prediction of protein sequence, protein sorting and post-translational modifications

The protein sequences were predicted by ORF Finder (NCBI) and ExPASy (EMBL-EBI). Protein sequences were then aligned by Clustal Omega (EMBL-EBI).

Protein sorting was predicted using different tools: TargetP (Emanuelsson et al 2000), ESLpred2 (Garg & Raghava 2008), Psort (Nakai & Horton 1999), WolfPsort (Horton et al 2007), CELLO (Yu et al 2006), N-terminal hydrophobic signal peptide SignalP 4.1 (Petersen et al 2011) and PrediSi (by Karsten Hiller, Institute for Microbiology, Technical University of Braunschweig).

Post-translational modifications were predicted by means of ScanProsite (Sigrist et al 2013) and Plant-Specific Myristoylation Predictor (Podell & Gribskov 2004).

Motif analysis

The presence of conserved motifs present in the protein sequences was investigated by MEME 4.6.1 and MAST motif search software (http://meme.sdsc.edu/meme/cgi-bin/meme.cgi) with the following parameters: any number of repetition as distribution of motif occurrences, maximum numbers of different motifs to find= 20, minimum motif width as 6 and a maximum motif width set to 50.
Functional annotation of these motifs was analyzed by InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/).

**Identification of regulatory elements involved in transcription**

Regulatory elements involved in transcription were identified according an *in silico* hypothesis-driven approach presented by Bernard and co-workers (Bernard et al 2010). In particular the authors explore, in *A. thaliana* and *Oryza sativa*, the bioinformatics-based evidence that sequences other than the TATA-box and TATA-variants but located in the same region relative to the TSS may be functional core-promoter elements.


**General Conclusions and Perspectives**

*Gus as a reporter gene in pollen*

We demonstrated that, on the contrary of what previously hypothesized (Gerola et al 2000), LAT52 expression is not differentially regulated during pollen tube growth. The previous observations, which led to hypothesize regulation of LAT52 promoter activity, were based on the use of Gus as a reporter gene and artefacts affected histochemical detection of Gus activity. In fact we observed that the localization of Gus activity by X-glu histochemical assay can be affected by the dispersion of the soluble intermediate reaction product and by the presence of Gus inhibitors. This is particularly true in the case of Gus histochemical detection in pollen tubes. ELF® 97 results to be a better substrate for histochemical detection of Gus activity with respect to X-glu. However its industrial production has been discontinued.

*Endogenous Gus activity in pollen*

*N. tabacum* pollen germination and elongation are inhibited by treatment with saccharolactone, an inhibitor of Gus activity, thus supporting the hypothesis previously reported in literature (Sudan et al 2006) that Gus activity is involved in cell growth processes and, in particular, in apical growth. Further studies for the determination of Gus roles in pollen will involve the selective inhibition of Gus genes expression and also the immunomodulation of Gus enzymes. The finding of gametophytic expression of Gus genes paves the way for potential application to plant breeding based on gametophytic selection also.
**GUS genes expression**

GUS activity was histochemically detectable in all plant organs of *N. tabacum* and *A. thaliana* seedlings. Specific primers were determined for *GUS1, GUS2* and *GUS3* genes and, by PCR amplification, their expression has been demonstrated in all the organs of both plant species. The expression patterns of *GUS* genes in *Arabidopsis thaliana* was investigated in root tip by whole mount multi probe *in situ* hybridization technique.

A specific expression of *GUS3* was observed in root border like cells. It was hypothesized that this gene might play a role in the cell wall remodelling required for the detachment of border like cells from the root tip. Particularly interesting was also *GUS2* and *GUS1* expression in the root cup meristem. This might indicate their involvement, in tandem action with *UGT* genes, in regulation of mitosis, according to the previously proposed model (Woo et al 2007). To this regard experiments are planned to investigate the expression of the different *UGT* genes in the root tip of *Arabidopsis*.

**GUS genes in plants**

By analysis of genome data bank of *Solanum lycopersicum*, and of the two progenitors of *N. tabacum* (*N. sylvestris* and *N. tomentosiformis*) we identified five different *GUS* genes, two more than those characterized in *A. thaliana* (Woo et al 2007). In *N. tabacum* we got the complete sequence of *GUS2* gene and partial sequences of two *GUS1* and two *GUS3* genes.

Comparison of *GUS* genes actually present in the different species of *Nicotiana* will shed light on the genomic processes that have led to the evolution of these species.

Phylogenetic analysis of *GUS* genes in angiosperms show that *GUS* genes cluster in three groups: one formed by the two *GUS3* genes, one that includes *GUS1* and *GUS2* genes of *Arabidopsis* and one formed by a *GUS* gene that is not present in *Arabidopsis* and which has been erroneously named *GUS1* in *Solanum*. 84
Analysis of the phylogenetic group which includes GUS1 and GUS2 of Arabidopsis reveal that the two genes derive from a “recent” gene duplication which took place independently several times inside the group. This might explain the similar expression pattern of GUS1 and GUS2 observed by in situ hybridization in A. thaliana.

Further phylogenetic studies are needed to support a revision of the classification GUS genes, now based on those identify in A. thaliana. The analysis of GUS genes of ancestral plants (from Gymnosperms to Charophyta) will allow a better comprehension of the evolution of GUS genes in land plants.


ACKNOWLEDGEMENTS

Thanks to
- Prof. Paolo Gerola, my supervisor during the Ph.D. course, for his guidance, teachings and above all for his patience;
- Prof. Giorgio Binelli for the supervision of the study of GUS genes in Solanaceae and of the phylogenetic analysis;
- all the people that worked and currently works in laboratory of botany: Drs. Tamara Corinti, Massimo Zilio, Filippo Spriano, Francesca Roggiani, Andrea Nicoli, Mariana Blagojevic and Alice di Pierro;
- Prof. Beatrice Bitonti's group at the University of Calabria, Department of Biology, Ecology and Earth Science, with whom I collaborated to perform multiprobe in situ hybridization study; in particular I thank Drs. Leonardo Bruno, Adriana Chiappetta, Olimpia Gagliardi, Natasha Spadafora, Domenico Iaria, Antonella Muto, Tiziana Ventura, Mr. Federico Ferraro and Mr. Giovanni Bruni;
- Prof. Alessandra Moscatelli and Dr. Elisabetta Onelli of University of Milan, Department of Bioscience for their counsels regarding pollen culture;
- Prof. Francesco Acquati and Dr. Laura Monti for their support in determining gene sequences;
- Drs. Maurizio Brivio, Maristella Mastore, Alberto Vianelli, Simona Binda Rossetti and Mr. Marco Lanfranchi for their counsels and support;
- Dr. Luisa Guidali for the assistance in confocal microscopy observation;
- all the administrative and technical personnel, do not mentioned before, without their work it is not possible to conduct research activity;
- Drs. Ed Newbigin and Mary Lush of the Plant Cell Biology Research Centre (University of Melbourne) for the gift of N. alata seed transformed with the construct LAT52-GUS and N. alata wild type seeds.
APPENDIX
<table>
<thead>
<tr>
<th>Species</th>
<th>Accession nucleotide</th>
<th>Accession protein</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis lyrata</td>
<td>XM_002866376</td>
<td>XP_002866422</td>
<td>Arabidopsis lyrata GUSA</td>
</tr>
<tr>
<td>Arabidopsis lyrata</td>
<td>XM_002870363.1</td>
<td>XP_002870409.1</td>
<td>Arabidopsis lyrata GUSB</td>
</tr>
<tr>
<td>Arabidopsis lyrata</td>
<td>XM_002871245.1</td>
<td>XP_002871291.1</td>
<td>Arabidopsis lyrata GUSC</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>NM_125518</td>
<td>NP_200933</td>
<td>Arabidopsis thaliana GUS1</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>NM_120865</td>
<td>NP_196400</td>
<td>Arabidopsis thaliana GUS2</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>NM_180762.2</td>
<td>NP_851093.1</td>
<td>Arabidopsis thaliana GUS3</td>
</tr>
<tr>
<td>Brachypodium distachyon</td>
<td>XM_003578543.1</td>
<td>XP_003578591.1</td>
<td>Brachypodium distachyon GUS2</td>
</tr>
<tr>
<td>Brachypodium distachyon</td>
<td>XM_003570281.1</td>
<td>XP_003570329.1</td>
<td>Brachypodium distachyon GUS3</td>
</tr>
<tr>
<td>Cicer arietinum</td>
<td>XM_004508644.1</td>
<td>XP_004508419.1</td>
<td>Cicer arietinum GUS3</td>
</tr>
<tr>
<td>Cucumis sativus</td>
<td>XM_004135393.1</td>
<td>XP_004135441.1</td>
<td>Cucumis sativus GUS2A</td>
</tr>
<tr>
<td>Cucumis sativus</td>
<td>XM_004157412.1</td>
<td>XP_004157460.1</td>
<td>Cucumis sativus GUS2B</td>
</tr>
<tr>
<td>Cucumis sativus</td>
<td>XM_004159341.1</td>
<td>XP_004159389.1</td>
<td>Cucumis sativus GUS3</td>
</tr>
<tr>
<td>Fragaria vesca</td>
<td>XM_004299278.1</td>
<td>XP_004299326.1</td>
<td>Fragaria vesca GUS2</td>
</tr>
<tr>
<td>Fragaria vesca</td>
<td>XM_004298345.1</td>
<td>XP_004298393.1</td>
<td>Fragaria vesca GUS1</td>
</tr>
<tr>
<td>Fragaria vesca</td>
<td>XR_184038.1</td>
<td>XP_004292974.1</td>
<td>Fragaria vesca GUS3</td>
</tr>
<tr>
<td>Glycine max</td>
<td>XM_003527051.1</td>
<td>XP_003540675</td>
<td>Glycine max GUS1B</td>
</tr>
<tr>
<td>Glycine max</td>
<td>XM_003520056.1</td>
<td>XP_003540764.1</td>
<td>Glycine max GUS1A</td>
</tr>
<tr>
<td>Glycine max</td>
<td>XM_003546530.1</td>
<td>XP_003546578.1</td>
<td>Glycine max GUS3</td>
</tr>
<tr>
<td>Hordeum vulgare</td>
<td>AJ495772</td>
<td>CAD42650</td>
<td>Hordeum vulgare GUS</td>
</tr>
<tr>
<td>Lotus japonicus</td>
<td>BT140493.1</td>
<td>AFK40288.1</td>
<td>Lotus japonicus GUS</td>
</tr>
<tr>
<td>Medicago truncatula</td>
<td>XM_003609426</td>
<td>XP_003609474</td>
<td>Medicago truncatula GUSB</td>
</tr>
<tr>
<td>Medicago truncatula</td>
<td>XM_003613964.1</td>
<td>XP_003614012.1</td>
<td>Medicago truncatula GUSA</td>
</tr>
<tr>
<td>Megachile rotundata</td>
<td>XM_003708560</td>
<td>XP_00370860</td>
<td>Megachile rotundata GUS</td>
</tr>
<tr>
<td>Musa acuminata</td>
<td>GSMUA_Achr4T32620_001*</td>
<td></td>
<td>Musa acuminata GUS2</td>
</tr>
<tr>
<td>Musa acuminata</td>
<td>GSMUA_Achr7T17860_001*</td>
<td></td>
<td>Musa acuminata GUS3</td>
</tr>
<tr>
<td>Organism</td>
<td>Accession</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------</td>
<td>------------------------------</td>
<td></td>
</tr>
<tr>
<td>Nicotiana benthamiana</td>
<td>NbS00028845g0015.1</td>
<td>Nicotiana benthamiana GUS2</td>
<td></td>
</tr>
<tr>
<td>Nicotiana benthamiana</td>
<td>NbS00010101g0016.1</td>
<td>Nicotiana benthamiana GUS1</td>
<td></td>
</tr>
<tr>
<td>Nicotiana benthamiana</td>
<td>NbS00010101g0016.1</td>
<td>Nicotiana benthamiana GUS3</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>NM_001054955.1</td>
<td>Oryza sativa GUSB</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>NP_001048420.1</td>
<td>Oryza sativa GUS1</td>
<td></td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>ABF94603.1</td>
<td>Phyllostachys edulis GUS</td>
<td></td>
</tr>
<tr>
<td>Phyllostachys edulis</td>
<td>FP091857.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Populus trichocarpa</td>
<td>XM_002330977</td>
<td>Populus trichocarpa GUSA</td>
<td></td>
</tr>
<tr>
<td>Populus trichocarpa</td>
<td>XP_002321464</td>
<td>Populus trichocarpa GUSB</td>
<td></td>
</tr>
<tr>
<td>Populus trichocarpa</td>
<td>XP_002315010.1</td>
<td>Populus trichocarpa GUSC</td>
<td></td>
</tr>
<tr>
<td>Prunus persica</td>
<td>ppa003804m**</td>
<td>Prunus persica GUS</td>
<td></td>
</tr>
<tr>
<td>Scutellaria baicalensis</td>
<td>AB040072</td>
<td>Scutellaria baicalensis GUS</td>
<td></td>
</tr>
<tr>
<td>Setaria italica</td>
<td>XM_004954196.1</td>
<td>Setaria italica GUS3</td>
<td></td>
</tr>
<tr>
<td>Setaria italica</td>
<td>XP_004962990.1</td>
<td>Setaria italica GUS1</td>
<td></td>
</tr>
<tr>
<td>Solanum lycopersicum</td>
<td>XM_004235932.1</td>
<td>Solanum lycopersicum GUS2</td>
<td></td>
</tr>
<tr>
<td>Solanum lycopersicum</td>
<td>XP_004235980.1</td>
<td>Solanum lycopersicum GUS1B</td>
<td></td>
</tr>
<tr>
<td>Solanum lycopersicum</td>
<td>XP_004241591.1</td>
<td>Solanum lycopersicum GUS1</td>
<td></td>
</tr>
<tr>
<td>Solanum lycopersicum</td>
<td>XP_004239465.1</td>
<td>Solanum lycopersicum GUS1A</td>
<td></td>
</tr>
<tr>
<td>Solanum lycopersicum</td>
<td>XP_004249602.1</td>
<td>Solanum lycopersicum GUS3B</td>
<td></td>
</tr>
<tr>
<td>Solanum lycopersicum</td>
<td>XP_004242732.1</td>
<td>Solanum lycopersicum GUS3A</td>
<td></td>
</tr>
<tr>
<td>Solanum tuberosum</td>
<td>PGSC0003DMC400042675</td>
<td>Solanum tuberosum GUSA</td>
<td></td>
</tr>
<tr>
<td>Solanum tuberosum</td>
<td>PGSC0003DMC400019475</td>
<td>Solanum tuberosum GUSB</td>
<td></td>
</tr>
<tr>
<td>Sorghum bicolor</td>
<td>XM_002452946</td>
<td>Sorghum bicolor GUS</td>
<td></td>
</tr>
<tr>
<td>Triticum aestivum</td>
<td>AK334371.1</td>
<td>Triticum aestivum GUS</td>
<td></td>
</tr>
<tr>
<td>Vitis vinifera</td>
<td>XM_002283224</td>
<td>Vitis vinifera GUS3A</td>
<td></td>
</tr>
<tr>
<td>Vitis vinifera</td>
<td>XP_002283260</td>
<td>Vitis vinifera GUS3B</td>
<td></td>
</tr>
<tr>
<td>Vitis vinifera</td>
<td>XP_002283254.2</td>
<td>Vitis vinifera GUS2</td>
<td></td>
</tr>
<tr>
<td>Vitis vinifera</td>
<td>XP_002284470</td>
<td>Vitis vinifera GUS1</td>
<td></td>
</tr>
<tr>
<td>Vitis vinifera</td>
<td>XP_002274743.1</td>
<td>Xenopus tropicalis putative GUS</td>
<td></td>
</tr>
<tr>
<td>Xenopus tropicalis</td>
<td>BCI60516</td>
<td>Xenopus tropicalis putative GUS</td>
<td></td>
</tr>
<tr>
<td>Zea mays</td>
<td>EU972798</td>
<td>Zea mays GUS1</td>
<td></td>
</tr>
<tr>
<td>Zea mays</td>
<td>ACG44916</td>
<td>Zea mays GUS3</td>
<td></td>
</tr>
<tr>
<td>Zea mays</td>
<td>NM_001158611</td>
<td>Zea mays GUS3</td>
<td></td>
</tr>
</tbody>
</table>

* The Banana Genome Hub, ** Solanum Genomics Network, *** Genome Database for Rosaceae.
Annex 2

Nicotiana tomentosiformis and Nicotiana sylvestris GUS coding DNA sequences

>NsGUS1A
ATGGATTCAGAAGACTAGGTGTTTGTGTTTGTGTTTGTGTTTTGTGTTTGC
TAATTCAGATGAATTAAGGTGATGTAAGAAAAAGAGGAGCTGACATATATAGCTAAACAGATGATA
ATTTCCATATGTCGACCTTTTGAATTTGGTGCCGAGAAGAAAGATTGTGATTTATACAAATGTCCCA
TGGGAAAAATCGGTGATCCCTTAAATCTGGACTGAAAAAAGATTTCTTCAAACACGGCGTTAAGC
AGCATTTCATCCAATAGGATCAAGATGAGGATGTTTTCGCTACACAGTACACGTAAAAATACAAAG
TTGGAAAAATCCTCAAAAGCAATGCTCAAAATTTTGAAGAAAAATCATAGTGCCCTTTTGATTT
GCTGAAGGACTCTTCCACATGAGATAGATGGATAGAATGCTAAAATCTTACAAAAAAAAAGG
GTGGGCTGCAATACCATCTTATATGATGCTGCAATGTTGAGAGATGACAGATGAGATGATGATG
ATACTCTTTGGGAGAGAGTGGAGAGTGGAGGAGGGTGC

>NsGUS1B
ATGGGATTCACCACCTTTTGTGTTAATCTTTTCTGGCAGATAGTCTGCTGATTTTTGCGCAAAAT
AGTATGAGCATACAGCTTTTGTGTTGCTGACATATATAGCTAAACAGATGATA
ATTTCCATATGTCGACCTTTTGAATTTGGTGCCGAGAAGAAAGATTGTGATTTATACAAATGTCCCA
TGGGAAAAATCGGTGATCCCTTAAATCTGGACTGAAAAAAGATTTCTTCAAACACGGCGTTAAGC
AGCATTTCATCCAATAGGATCAAGATGAGGATGTTTTCGCTACACAGTACACGTAAAAATACAAAG
TTGGAAAAATCCTCAAAAGCAATGCTCAAAATTTTGAAGAAAAATCATAGTGCCCTTTTGATTT
GCTGAAGGACTCTTCCACATGAGATAGATGGATAGAATGCTAAAATCTTACAAAAAAAAAGG
GTGGGCTGCAATACCATCTTATATGATGCTGCAATGTTGAGAGATGACAGATGAGATGATGATG
ATACTCTTTGGGAGAGAGTGGAGAGTGGAGGAGGGTGC
CAGAATGTACTCAAGACGTCTGTACTTTCAGTGTGTTGTGGGAGCTGGAGGGGC
TTAACACAGCGGTCCGACCATGTACTAATAGCTTTTTGCTTCTGATATTGGACC
AGCTTGGGATGTCGCTGCTTTATGTACTAAAGCACATCTGTCGACAGACACTGATTGGTGA
AACTATGGTTTTACACACTACTACACAACTTCGAGATTATGACCTATTTGGAC
GGCAACACTACCATAGCTGAGGAGTTCGTCCTTAACTTAGTGAATCAGAAAGAGCA
CAGACTCATAGCCACAGAAGAGCTCTCCAAATTTGCAAGTTAAAGTAGAGATAGACAT
CAAATACAGCAAGAGAATACCATCTTTTACAGCAGAAGATAGGAGATTACAGATCGAG
CTGCTGAAATGCAACACTACATAGCTTATTGCTGGAATTATGCATTTAATTGAGCCCTAT
ATTTGTAATTCAACAGACACATTAACAGTAGCTCATTCTCTCTATTGGGATTAACATACAT
CATGATATTATTTTGGCTCTTCTGAGCTGA

>NsGUS3ch10B
ATGGCTTTTGTGTTCTACTGACTCTTTTGAGGTGATTTCTCTGCTATTGCT
AAGCTGTTGCTTCTTCTTCTAAATATGGTATTAGAACACTGGGAAATCTCTCCTCTCTT
TAATCTGTTATAAGCAGAAAGTTTTTGTGGAATAGCAGAATTAGGTTGCGCTTCT
TTGGACCAGGTTTCTGGAGTTAATGGTACTTTTGCTGTTAAGGAAAGAGATTC
GGATGATGTTTTTTGCTGCTTGTAGCTGAA

>NtomGUS1A
ATGGATTCAAGAACTAGGTTGTTTTGTTTGGAGTTATTTCTCTGCTATTGCTG
AACTATGGTTTTGTTTGGTTGTTCTTTAGTGTTTTTCTTCTGATTATGCGCTCTCTCTT
TTGCGACGTTTGAGGAAGCTGCAATGCAACAACTTCAAGAGAAAGAGATTC
GGATGATGTTTTTTGCTGCTTGTAGCTGAA

>GtATGTTTTACACACATTTTCTTCTTCAATTAGATTTTGGGATTAACATACAT
CATGATATTATTTTGGCTCTTCTGAGCTGA

>AtGUS
ATGGATTCAAGAACTAGGTTGTTTTGTTTGGAGTTATTTCTCTGCTATTGCTG
AACTATGGTTTTGTTTGGTTGTTCTTTAGTGTTTTTCTTCTGATTATGCGCTCTCTCTT
TTGCGACGTTTGAGGAAGCTGCAATGCAACAACTTCAAGAGAAAGAGATTC
GGATGATGTTTTTTGCTGCTTGTAGCTGAA
>NtomGUS3A
ATGGGAGCTTTCTACATATGACATGGATGAGTTAACAACACCTTTTCAGAAGACAGGAGC
ACTCTGGATCTGCTTTGCTGATGCTTATGATTGGAGACGAGGATCAATAGGCTGACATGCGTG
GAGGAAATTGGGATTCCAGCAATGCCCTCGATTTCATAAAATACACTGTTGCCAAGGGCTAC
CAGATACACTCATGGGAATTGG
AAATGAATTGAGTGGTCAGGGAATTGGTGCAAGTGTTAA
TGCTGTACAGTATGGAAAAGATGTTATCCATCTGACAATCTCATAGACCAAGTATACAAGA
ATTTCGACCAACGTCCTCTTTTCATTGCACCAGGAGGATTCTATAGTCCTGAGTGGTTCAGC
AAGCTCCTTGAGGTTTCAGGGGCTGGCATAGTCAATGTCTTGACGCATCATATTTATAATCT
TGGCCCAGGGTCTGAAGGAAGCAAGCTTGTA
GATAAAGTTTTAAATCCTCAATACCTGAATA
AAATATCAGACACATTTGGTAATCTTACTCAAACCATAAAAATGAAGGGTCCTTGGGCTTCA
GCTTGGGTTGAGAGGACTCCTAACAACAGGCTGGTCTATAGTGTCTATGCTCCTAACCCCT
CTGATATTTCATTTCTGAGTTCAAGCTTGGCAGCAGACGACCATCACCTCAATACCTCTCT
CCAGTTGTAGATAA
NtomGUS3A
>NtomGUS3B
ATGGCTTTTTTGTTCACGCTCTCTTTGGGGGTGTATCTTCTCGGCTATTTATGTTGTACAGCAG
CCAAATGGAACGCAATGAGCAGAGGCTTCTTTATTTACATGTGAGGATTACCCACTCTTTTCG
TCTCTTTTTCTTTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Annex 3

Alignment of amino acid sequences and identification of some conserved motif

SbGUS
-MGFQVWQGLCVCFSFLF-I----------------CGVIGEETTVIKIEENPVAQTIDEN 44
S1GUS3B
-MQMASSFKCLLVLFL--LSLYFTQISRAGD-----------IVEEGILYIDGVSSIAKIIDKD 51
NsGUS3B
---MASLFSCEFLEYLL-LLLNSQCSRAESSSKSSAFKSIYIDGSAAIKAKEID 56
NtGUS3B
---MASLFSCEFGLCFL--LLLVSQCSRAE-----------EGTLYIDGSAAIAKIDED 47
AtGUS3
---MA-YRQILAVLFCVQFLDCTV------------SSAVEENGTVFYGRAVGTIDED 47
S1GUS3A
---MGSLFLQGAVLWLFLFSLRF-IC------------GTK--AQGTVFIDKIAIGRIIDR 45
NsGUS3A
---MGSLFLQKMLWICFLSLRL-IC-----------GYSEVAQGVSFDGKNAIGRTDNN 47
NtGUS3A
---MGSLFLQKAVLWICFLSRF----GYSEVAQGVSFDGKNAIGRTDNN 47
S1GUS1A
MVSSRFLLFCLFVSS---CLSNSDELKVKVGSIAQTDDN 44
NsGUS1A
---MDSRTRFLCFLVLS---CLSNSDELKVKVSAIATDDN 42
NtGUS1A
---MDSRTRFLCFLVLS---CLSADKLKVGVTSIAKTDDN 42
S1GUS1B
-MGFP---PLLIFLA-----LCPAF------------LAQIVEETEFLIDGVKIAETDNN 41
NsGUS1B
-MGFP---PPLLIFLA-----VCPAF------------LAQIVEDTALLIDTVKIAETDAN 41
NtGUS1B
-MGFP---PPLLIFLA-----LCPAF------------LAQIVEDTALIDTVKIAETDAN 41
S1GUS2
-MNFR-----TLLIFVA-----LCPAL------------SAIISQVIQLTIDASSTKVAWTDDN 41
NsGUS2
-MDFR-----TLLIFLA-----LCPAF------------SAQTVEDTELMDTVSRIAWTDDN 41
NtGUS2
-MDFR-----TLLIFLA-----LCPAF------------LAQTGEDTELMDTIVSVKIAWTDDN 41
NtGUS2
-MDFR-----TLLIFLA-----LCPAF------------LAQTGEDTELMDTIVSVKIAWTDDN 41
AtGUS2
-MGFR-----VCVIVVFLGCLLLVPEKT-----------MAQEMKRAISIVIQGARRVCETDNN 46
AtGUS1
-MGFN-----VVVFSLCLLLLLPPVT-----------FGSNMERTTLVIGSRTRIAETDNN 43

Signal peptide
SbGUS
YVCATLDWPTKCNYGNCWGKSSFLNL------DLNNIIRNAVK------EFAP 89
S1GUS3B
FICATLDWPPTKCDYGTCWGNASLLNL------DLSNKVLVNAIR------AFSP 96
NsGUS3B
FICATLDWPSNKCDYGTCWGNSSLLNL------DLSNKVLVNAIR------AFSP 101
NtomGUS3B
FICATLDWPSNKCDYGTCWGNSSLLNLVTPLSLFDLSNKVLVNAIR------AFSP 99
AtGUS3
FICATLDWPPKEKCDYSGCSWDHASILNL------DLNNVILQNAIK------AFAP 92
S1GUS3A
FICATLDWPPKEKCDYTCAWDYSSFLNL------DLNNIIFLANAIK------AFSP 90
NsGUS3A
FICATLDWPPKEKCDYTCAWDNSLNASLNL------DLNNIIFLANIKDLHNYYAALSP 100
NtomGUS3A
FICATLDWPPKEKCDYTCAWDNSLNASLNL------DLNNIIFLANIKDLHNYYAALSP 100
S1GUS1A
FICATLDWPPENKCDYNQCPWKGAGILNL------DLKNRILTNAVK------AFNP 89
NsGUS1A
FICATLDWPPENKCDYNQCPWKGAGILNL------DLKNKILTNAVK------AFNP 87
NtomGUS1A
FICATLDWPPENKCDYNQCPWKGAGILNL------DLKNKILTNAVK------AFNP 87
S1GUS1B
YVCATLDWPPKEKCNHNDPWCWGSTSLNL------DLSHPFLANSIQ------AFNH 86
NsGUS1B
YVCATLDWPPKEKCDYNPCFWYTSLMNLL------DLSHPFLANSIQ------AFNP 86
NtomGUS1B
YVCATLDWPPKEKCNYCEPWGTSLMNLL------DLSHPFLANSIK------AFNP 86
S1GUS2
YICATLDWPPKEKCNYQCWPWGASILNL------DLTHPYLVNAIR------AFKG 86
NsGUS2
YICATLDWPPKEKCNYQCWPWDLASILNL------DLSHPFLENAIR------AFKG 86
NtomGUS2
YICATLDWPPKEKCNYQCWPWDLASILNL------DLSHPFLENAIR------AFKG 86
NtGUS2
YICATLDWPPKEKCNYQCWPWDLASILNL------DLSHPFLENAIQ------AFKG 86
AtGUS2
FVCATLDWPHDKCNYDQCWPYGSSVINM------DLTRPLTKAIK------AFKP 91
AtGUS1
FICATLDWPPKEKCNYDQCWPYGASILNL------NLASPLLLAKAIQ------AFRT 88

::***:* ** *:** * * :..:*: * * : * *:  *
SbGUS          LKLRFGGTLDRLVYQTSRDEPCDSTFYNNNTNLDFSHACLSDLRDEINQFILETG-S 148
S1GUS3B        LTIIRLGGLQDKLIVYQTMHDQKQCPCLPFFHDDELFKFTQGCLPLSRWDELNEFFKKTG-A 155
NsGUS3B        LKLIRLGGLQDKLIVYQTAHEYQPCSPFLLNLKELFGFTQGCLPLSRWDELQFFKKTG-A 160
NtomGUS3B      LKLIRLGGLQDKLIVYQTAHYEPCSPFLLNLKELFGFTQGCLPLSRWDELNEIFKKTG-A 158
AtGUS3         LKLIRIGGLQDIVIYETPDSKQPCLPTKNSILFGYQGCLPMRWWDELNAFFKKTG-T 151
S1GUS3A        LKLIRLGGLQDKLIVYQTAHEYPCSPFLLNLKELFGFTQGCLPLSRWDELNAFFKKTG-A 149
NsGUS3A        LKLIRLGGLQDKLIVYQTAHEYPCSPFLLNLKELFGFTQGCLPLSRWDELNAFFKKTG-A 159
NtomGUS3A      LKLIRLGGLQDKLIVYQTAHEYPCSPFLLNLKELFGFTQGCLPLSRWDELNAFFKKTG-A 159
S1GUS1A        LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKHDGLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145
NsGUS1A        LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKHDGLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145
NtomGUS1A      LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKHDGLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145
S1GUS2         LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKFDGLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145
NsGUS2         LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKFQDGFLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145
NtomGUS2       LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKFQDGFLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145
NtGUS2         LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKFQDGFLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145
AtGUS2         LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKFQDGFLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145
AtGUS1         LRLRLGLQNRRVYQKGGYKGNLIPSSCPQTPKFQDGFLFGFSGCGLPMRSDLWDELNNLFKKTG-A 145

* :*:*:*:*:*          :*          :*          :*          :*          :*          :*          :*
SbGUS          EAVFGLNALRGKTVEIKGIIKDGQYLGETTTAVGEWDYSNSKFLIEYSLKKGYKHIRGWT  208
SlGUS3B        KVTFGLNTLNGKKIAS---------DGRTALGDWDSSNAESLIRYTVSRYGNY-IHGWE  203
NsGUS3B        KVTFGLNALNGKTIAP--------NG-SALGDWDSSNAESFIRYTVSRYGTY-ILGWE  207
NtomGUS3B      KVTFGLNALNGKTIAP--------NG-SALGDWDSSNAESFIRYTVSRYGTY-ILGWE  205
AtGUS3         KVIFGLNALSRSIKS--------NG-EAIGAWNYTNAESFIRTAENNYT-IDGWE  198
SlGUS3A        SIIFGLNALYGRSVHP--------DS-LSVGAWDSNAESLIRYTVKGYD-IHGWE  196
NsGUS3A        SIIFGLNALYGRSVHP--------DG-LSVGAWDSNAESLIRYTVKGYA-IHGWE  206
NtomGUS3A      SIIFGLNALYGRSVHP--------DG-LSMGAWDSNAESLIRYTVKGYI-IHGWE  206
SlGUS1A        AITFSFNALIGRIPSDE-------NDTLWVGDWNHYNAKSLMKYTLNKGYK-IDSYE  198
NsGUS1A        AITFSFNALIGRIPSDE-------D-DTLWEGDNQFNANAKSLMKYTIRKGYK-IDSYE  196
NtomGUS1A      AITFSFNALLGRIQSDE-------D-DTLWEGDNQYNAKSLMKYTVDGGYK-IDSYE  195
SlGUS1B        LVTFGLNALRGRQRTS-------KRVWEGNWDSSNAHDFIDYTYSKGYQ-IHSWE  192
NsGUS1B        LVTFGLNALHRQRAG-------KRVWGNWDSSNARDFINYTISKGYQ-IHSWE  192
NtomGUS1B      LVTFGLNALHRQRTG-------KRLWGNWDSSNAHDFINYTISRGYQ-IHSWE  192
SlGUS2         LVTFGLNALGRQRQAN-------RHAWVGNWDSSNALNFIYTVAKGYH-IHSWE  192
NsGUS2         LVTFGLNALGRRRIN-------RHAWGNWDSSNARDFMKYTVAKGYQ-IHSWE  192
NtomGUS2       LVTFGLNALGRRRIN-------RHAWGNWDSSNALDFIKYTVAKGYQ-IHSWE  192
AtGUS2         VVTFGLNALRGHKLR-------GKAWGDNHINTQDLNVTYSKGYV-IDSWEL  197
AtGUS1         IVTFGLNALHRKMN-------GTAWGGDWTHTQDFMNYTYSKGYA-IDSWE  194

*:::*:*:*:  *:::*: ::::*: *: *: *:*:*:*:*

heparanase motif
Acid-base residue (Glu212 in S. baicalensis)
Tyr residue important for glycosyl hydrolase activity
Nucleophilic residue
<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>SbGUS</td>
<td>IPNPDYYSALLWHRLMGSKVLKTEIVGTKNVYIYAHCAKKSNGITMLVLNHGDGESSVKIS</td>
<td>444</td>
</tr>
<tr>
<td>S1GUS3B</td>
<td>HPNPDYYGALLWHRLMGRLSVLSTQFQGKLRYSAYAHCSKSSSEGIALMLINMHSSITVINIS</td>
<td>439</td>
</tr>
<tr>
<td>NsGUS3B</td>
<td>HPNPDYYGALLWHRLMGRSVLFSTQFQGTKLRTYHCSKSSDGiTLILNMGDGATVNS</td>
<td>442</td>
</tr>
<tr>
<td>NtomGUS3B</td>
<td>VPNPDYYGALLWHRLMGSVLSTQFQGTKLRTYHCSKSSDGITLLINMDGATVNS</td>
<td>440</td>
</tr>
<tr>
<td>AtGUS3</td>
<td>TNPDYYSALLWRQLMGARKLFTTSFGKTIRSYHCARQSKGIGTVLMLNLDNNTTVAK</td>
<td>434</td>
</tr>
<tr>
<td>S1GUS3A</td>
<td>EPNPDYYSALLWHLMGNSVLFSTSGTTLKRAYHCAKQSQQGIGTVLMLNLDGNTTIPR</td>
<td>432</td>
</tr>
<tr>
<td>NsGUS3A</td>
<td>VPNPDYYSALLWHRLGNSVLFSTSGTTLKRAYHCAKQSQQGIGTVLMLNLDGNTTMSAN</td>
<td>442</td>
</tr>
<tr>
<td>NtomGUS3A</td>
<td>VPNPDYYSALLWHRLGNSVLFSTSGTTLKRAYHCAKQSQQGIGTVLMLNLDGNTTIRAN</td>
<td>442</td>
</tr>
<tr>
<td>S1GUS1A</td>
<td>IPNPDYYSALLWHGMKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFDS</td>
<td>434</td>
</tr>
<tr>
<td>NsGUS1A</td>
<td>VPNPDYYSALLWHGMKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>432</td>
</tr>
<tr>
<td>NtomGUS1A</td>
<td>VPNPDYYSALLWHLMGKSVMSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>431</td>
</tr>
<tr>
<td>S1GUS1B</td>
<td>VPNPDYYSALLWHGMKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFDVS</td>
<td>434</td>
</tr>
<tr>
<td>NsGUS1B</td>
<td>VPNPDYYSALLWHLMGKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>429</td>
</tr>
<tr>
<td>NtomGUS1B</td>
<td>VPNPDYYSALLWHLMGKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTEVS</td>
<td>429</td>
</tr>
<tr>
<td>S1GUS2</td>
<td>VPNPDYYSALLWHLMGKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>430</td>
</tr>
<tr>
<td>NsGUS2</td>
<td>VPNPDYYSALLWHLMGKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>430</td>
</tr>
<tr>
<td>NtomGUS2</td>
<td>VPNPDYYSALLWHLMGKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>430</td>
</tr>
<tr>
<td>NtGUS2</td>
<td>VPNPDYYSALLWHLMGKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>430</td>
</tr>
<tr>
<td>AtGUS2</td>
<td>VPNPDYYSALLWHLMGKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>435</td>
</tr>
<tr>
<td>AtGUS1</td>
<td>VPNPDYYSALLWHLMGKNSVLFSTSGTTLKRAYHCAKQSGVTLLINMDKSTTFEVS</td>
<td>431</td>
</tr>
</tbody>
</table>

*Note: The sequences are aligned for comparison.*
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SbGUS</td>
<td>-</td>
<td>527</td>
</tr>
<tr>
<td>S1GUS3B</td>
<td>M</td>
<td>542</td>
</tr>
<tr>
<td>NsGUS3B</td>
<td>M</td>
<td>545</td>
</tr>
<tr>
<td>NtomGUS3B</td>
<td>M</td>
<td>543</td>
</tr>
<tr>
<td>AtGUS3</td>
<td>-</td>
<td>536</td>
</tr>
<tr>
<td>S1GUS3A</td>
<td>-</td>
<td>536</td>
</tr>
<tr>
<td>NsGUS3A</td>
<td>-</td>
<td>546</td>
</tr>
<tr>
<td>NtomGUS3A</td>
<td>-</td>
<td>546</td>
</tr>
<tr>
<td>S1GUS1A</td>
<td>-</td>
<td>532</td>
</tr>
<tr>
<td>NsGUS1A</td>
<td>-</td>
<td>517</td>
</tr>
<tr>
<td>NtomGUS1A</td>
<td>-</td>
<td>516</td>
</tr>
<tr>
<td>S1GUS1B</td>
<td>-</td>
<td>539</td>
</tr>
<tr>
<td>NsGUS1B</td>
<td>-</td>
<td>539</td>
</tr>
<tr>
<td>NtomGUS1B</td>
<td>-</td>
<td>539</td>
</tr>
<tr>
<td>S1GUS2</td>
<td>-</td>
<td>540</td>
</tr>
<tr>
<td>NsGUS2</td>
<td>-</td>
<td>541</td>
</tr>
<tr>
<td>NtomGUS2</td>
<td>-</td>
<td>541</td>
</tr>
<tr>
<td>NtGUS2</td>
<td>-</td>
<td>541</td>
</tr>
<tr>
<td>AtGUS2</td>
<td>-</td>
<td>543</td>
</tr>
<tr>
<td>AtGUS1</td>
<td>-</td>
<td>539</td>
</tr>
</tbody>
</table>