



**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA
CELULAR E MOLECULAR**

**Fitocistatinas em arroz (*Oryza Sativa*) e caracterização funcional da
fitocistatina XII**

Ana Paula Christoff

Porto Alegre, dezembro de 2012

Universidade Federal do Rio Grande do Sul

Centro de Biotecnologia

Programa de Pós-Graduação em Biologia Celular e Molecular

**Fitocistatinas em arroz (*Oryza Sativa*) e caracterização funcional da
fitocistatina XII**

Ana Paula Christoff

Dissertação submetida ao Programa de Pós-Graduação em Biologia Celular e Molecular do Centro de Biotecnologia da UFRGS como requisito parcial para obtenção do título de Mestre.

Orientador: Dr. Rogério Margis

Porto Alegre

Dezembro 2012

INSTITUIÇÃO E FONTES FINANCIADORAS

Este trabalho foi desenvolvido no Laboratório de Genomas e Populações de Plantas do Centro de Biotecnologia da UFRGS.

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico,
Ministério da Ciência e Tecnologia.

AGRADECIMENTOS

Inicialmente gostaria de agradecer ao meu orientador, Prof. **Rogério Margis**, por me aceitar em seu laboratório e confiar em mim para realização deste trabalho, além de diversos outros trabalhos paralelos, os quais contribuíram imensamente para minha formação. Pela liberdade e oportunidade de crescimento profissional e pessoal, pela confiança, paciência e pela sua excelente orientação, como exemplo de professor e pesquisador. À professora **Márcia Margis**, pela disponibilização de toda infraestrutura necessária para a realização de meus experimentos, além do acompanhamento e auxílio em vários momentos durante meu mestrado.

Aos meus colegas e amigos do **LGPP**, **NGFP**, e demais laboratórios vizinhos, sempre me ajudando no que fosse preciso, pela amizade, discussões científicas e filosóficas. Obrigada! Em especial à **Andreia Turchetto-Zolet** por estar sempre por perto, em todos os momentos de meu mestrado, desde o auxílio em experimentos, análises e revisões, pelo grande companheirismo, carinho, amizade, proporcionando, inclusive, muitas caronas para casa. Também, a todas as meninas do lab, **Ana Körbes**, **Fernanda Cruz e Franceli**, pelos laços de amizade criados, cafés e comilanças, além de toda ajuda laboratorial e psicológica disponibilizada. Agradeço principalmente à **Gisele Passaia**, pessoa sem a qual este trabalho não teria sido realizado! Muito, muito obrigada por tudo que me ensinastes, além da grande amizade, incentivo, sugestões, entre tantas outras coisas cuja descrição não cabe aqui. Também a **Andréia Caverzan**, pela grande amizade e auxílio em inúmeros momentos, do mestrado e da vida.

À eterna galera do **LabDros** (UFSM), principalmente: **Bruna**, **ten Caten**, **Francine**, **Gabriel (+Cris e Malu)**, **Gabinha**, **Lizandra**, **Mauro**, **Paloma**, **Ronaldo**, **Sinara e Valéria**, pois fazem uma falta gigantesca na vida de uma pessoa, apesar de existir Gmail e Facebook. À **Francis**, que saiu do LabDros para POA, e, um ano depois, vim atrás morar com ela. Pela grande amizade, pelo nosso dialeto peculiar e, principalmente, pela diversão infundável proporcionada

todos os dias destes últimos dois anos, juntamente com o **Caio**. Muito Obrigada por tudo!

Aos meus pais, **Ivo** e **Ivani**, por todo amor, carinho, compreensão e segurança. Por terem me ensinado os grandes valores da vida, e investido tanto em minha educação, mesmo sofrendo com a distância, desde que saí de Três Passos para cursar biologia na UFSM, em Santa Maria. Apesar da distância vocês sempre estiveram presentes, mesmo porque, conversávamos TODO dia por telefone, MSN ou Gmail, impossível conseguir dormir sem dizer: boa noite pai, boa noite mãe. Por serem meus exemplos de caráter e dedicação e pela confiança que sempre depositaram em mim. Não tenho palavras para agradecer, obrigada por tudo! Amo vocês!!! (e o meu cão também).

Enfim, agradeço ao grande amor da minha vida: **Felipe**. Que há cinco anos e meio tem me ensinado, ajudado, incentivado, obrigado, pressionado, compreendido, apoiado e confiado, em todos os aspectos pessoais e científicos de minha vida. Obrigada por tudo, toda ajuda na TCC, na dissertação de mestrado, e em breve, na tese de doutorado. Agradeço muito por ter encontrado uma pessoa como você. Te amo!

Também gostaria de agradecer a todas as pessoas e instituições que tornaram possível o desenvolvimento deste trabalho. Ao **PPGBCM** da UFRGS e seus professores. Aos professores Giancarlo Pasquali e Diego Bonatto que aceitaram fazer parte de minha comissão de acompanhamento. Aos funcionários do Programa, principalmente Sílvia e Luciano pelo auxílio durante estes dois anos.

À UFRGS, a CAPES e ao CNPq, pela estrutura e apoio financeiro para realização deste trabalho.

ÍNDICE

RESUMO	8
ABSTRACT	9
1. INTRODUÇÃO	10
1.1 O Arroz como planta modelo de estudos	10
1.2 Cistatinas e fitocistatinas.....	11
1.3 Fitocistatinas e suas proteases-alvo	14
1.4 Funções biológicas atribuídas as fitocistatinas	19
1.5 Atuação das fitocistatinas em resposta a estresses	23
1.6 Fitocistatinas em arroz.....	26
1.7 A fitocistatina XII de arroz e suas homólogas em plantas.....	30
1.8 Inibição de legumainas por fitocistatinas tipo OcXII.....	32
2. OBJETIVOS.....	34
2.1 Objetivo Geral	34
2.2 Objetivos Específicos.....	34
3. MATERIAIS E MÉTODOS	35
3.1 Caracterização <i>in silico</i> dos genes de fitocistatinas em arroz.....	35
3.2 Análises evolutivas das fitocistatinas	39
3.3 Material vegetal.....	40
3.4 Análise da expressão gênica das fitocistatinas em arroz, por RTqPCR durante os processos de germinação e crescimento de plantas.....	41
3.5 Construção do vetor para silenciamento gênico de OcXII	43
3.6 Transformação de <i>Agrobacterium tumefaciens</i>	46
3.7 Transformação genética de calos embriogênicos de arroz.....	47
3.8 Confirmação das plantas transgênicas	49
3.9 Confirmação do silenciamento de OcXII e nível de expressão gênica das demais fitocistatinas e legumainas de arroz.....	49
3.10 Alterações fenotípicas das plantas OcXII silenciadas	50
3.11 Obtenção e análise da progênie das plantas silenciadas OcXII	50
3.12 Caracterização da sequência de nucleotídeos de uma região do promotor do gene OcXII.....	51
3.13 Construção de um vetor para análise do promotor do gene OcXII em plantas de arroz	52
3.14 Transformação de calos embriogênicos de arroz para a expressão dos genes repórteres fusionados ao promotor OcXII.....	53

3.15	Análise da modulação da expressão gênica da fitocistatina XII em resposta a tratamento com ácido abscísico (ABA).....	54
4.	RESULTADOS	55
4.1	Caracterização geral das fitocistatinas de arroz	55
4.2	Expressão gênica de fitocistatinas durante a germinação e desenvolvimento de plantas de arroz.....	63
4.3	Considerações evolutivas sobre as fitocistatinas	68
4.4	Obtenção do vetor RNAi e transformação genética de plantas de arroz para silenciamento da fitocistatina XII	81
4.5	Plantas transgênicas resultantes da transformação para silenciamento da fitocistatina XII	84
4.6	Análises de plantas T1 das linhagens OcXII silenciadas	92
4.7	Análises da região promotora do gene da fitocistatina XII	95
4.8	Obtenção do vetor pHGWFS7 expressando genes repórteres de forma controlada pelo promotor da fitocistatina XII	98
4.9	Modulação da expressão da fitocistatina XII por ABA	100
5.	DISCUSSÃO	103
5.1	Caracterização das fitocistatinas em arroz	103
5.2	Considerações evolutivas sobre as fitocistatinas relacionadas com seu perfil funcional.....	109
5.3	Silenciamento gênico da fitocistatina XII de arroz.....	113
5.4	Fitocistatina XII de arroz e o estudo da sequência promotora de expressão gênica	116
6.	CONCLUSÕES E PERSPECTIVAS	118
7.	REFERÊNCIAS	121
	CURRICULUM VITAE	132

RESUMO

Cistatinas são inibidores competitivos de proteases cisteínicas. Em plantas, as fitocistatinas, constituem uma família independente, caracterizada pela presença do motivo conservado LARFAV. Elas podem estar envolvidas em uma grande variedade de processos fisiológicos, desde o controle da proteólise endógena em órgãos reprodutivos e vegetativos, até a inibição de proteases extracelulares de herbívoros, parasitas e patógenos. Algumas fitocistatinas possuem uma extensão carboxi-terminal em sua sequência, com um motivo consenso SNSL importante para a inibição de legumainas. Em arroz, existem doze genes distintos de fitocistatinas, sendo a fitocistatina XII (OcXII) a única com um domínio extra carboxi-terminal. Considera-se esta extensão extremamente importante, pois permitiria a atuação bifuncional deste inibidor de proteinase, simultaneamente inibindo dois tipos de proteases, papainas e legumainas. O objetivo principal que norteou o desenvolvimento deste trabalho foi identificar as doze fitocistatinas de arroz (*Oryza sativa*), caracterizando, principalmente, o gene da OcXII, de forma a analisar as prováveis implicações funcionais deste gene na planta. A caracterização da expressão dos doze genes codificadores das fitocistatinas revelou que os genes comumente mais expressos são OcI, OcIII e OcXII, sendo que OcXII destaca-se nas plantas em estágio final de maturação das sementes, em tecidos específicos de endosperma e células em cultura. Para germinação e crescimento de plântulas, é necessária uma redução da expressão de OcXII. O padrão evolutivo das fitocistatinas, em diversas plantas, sugere uma origem comum para estes genes, seguida de divergências e duplicações específicas em cada espécie. A OcXII é mantida em um grupo monofilético com as demais fitocistatinas com extensão C-terminal. Nos estudos funcionais sobre a fitocistatina XII de arroz, nove linhagens de plantas RNAi para o gene OcXII foram obtidas via transformação de calos de arroz (ssp. *japonica*, cv. Nipponbare). Quatro destas linhagens apresentaram nível de pelo menos 90% de silenciamento de OcXII, sem alteração na expressão gênica das demais fitocistatinas e legumainas. Duas destas linhagens tiveram alterações fenotípicas com redução da viabilidade das sementes geradas. O silenciamento de OcXII e expressão de demais fitocistatinas manteve-se inalterado na geração T1, enquanto as legumainas 1 e 3 sofreram alterações em pelo menos uma das linhagens. A transformação de plantas para estudos de localização celular com o promotor do gene OcXII, resultou em quatro linhagens transformadas, sendo que em uma delas a expressão do gene repórter *gus*, controlado pelo promotor OcXII, foi significativamente detectada nos calos em vias de regeneração de plantas. Também, pode ser verificado que a OcXII responde ao tratamento com ABA (ácido abscísico), tendo um aumento de expressão significativo após 4 horas, tanto em tecidos de folha quanto de raiz. Desta forma, a fitocistatina XII é um dos genes de fitocistatinas mais expressos durante todo o desenvolvimento das plantas de arroz, sendo o único gene com provável capacidade de inibir proteases cisteínicas do tipo legumainas.

ABSTRACT

Cystatins are cysteine proteases competitive inhibitors. Phytocystatins (PhyCys), in plants, constitute an independent family due to the presence of the LARFAV conserved domain. PhyCys are involved in a wide variety of physiological processes, since the control of the endogenous proteolysis in vegetative and reproductive organs, by inhibition of extracellular proteases from herbivores, pathogens and parasites. Some phytocystatins have an extension in their carboxy terminal sequence with a SNSL consensus domain, important for inhibiting legumains. In rice, there are twelve distinct phytocystatins genes and the phytocystatin XII (OcXII) are the only one with a carboxy-terminal additional domain. This extension would be extremely important, allowing the bifunctional performance of this proteinase inhibitor, simultaneously inhibiting two types of proteases: papains and legumains. A main goal of this work is to identify the twelve phytocystatins in rice (*Oryza sativa*), focusing at the OcXII gene in order to analyze the functional implications of this gene in the whole plant. Characterization of the gene expression pattern of the twelve phytocystatins revealed that the genes more ubiquitously expressed are OcI, OcIII and OcXII. OcXII stands out in plants in final stage of seed maturation, in endosperm specific tissues and in culture cells. For germination and seedling growth OcXII undergoes a reduction in its genic expression. The phytocystatins evolution patterns in various plants, suggests a common origin for these genes, followed by divergences and duplications, specific to each species. The OcXII, is kept in a monophyletic group with all the other phytocystatins with C-terminal extension. In the functional studies on phytocystatin XII from rice, nine lines of RNAi plants were obtained by transformation of rice's callus (*japonica* ssp. cv. Nipponbare). Four of these strains had low levels of OcXII mRNA, revealing that this gene were at least 90% silenced, any changes in gene expression for other phytocystatins or legumains were detected. Two of these transgenic lines showed phenotypic changes with reduced seed viability. The OcXII gene silencing and other phytocystatins expression remained unchanged in the T1 generation, while the first and third legumains varied in at least one of the strains. The plants transformation, for cellular localization studies with OcXII gene promoter resulted in four transgenic strains, and in one of them expression of the *gus* reporter gene, controlled by the promoter OcXII, was significantly detected in the calli undergoing the plant regeneration process. Also, it can be noticed that the OcXII gene expression responds to the treatment with ABA (abscisic acid), with a significant increase, in both tissues as leaf root, after 4 hours of treatment. Thus, phytocystatin XII is one of the most expressed genes throughout the development of rice plants, being the only gene with the probable ability to inhibit cysteine proteases of legumain type, in addition to papain proteases.

1. INTRODUÇÃO

1.1 O Arroz como planta modelo de estudos

O arroz (*Oryza sativa* L.) pertence à divisão angiosperma, à classe das monocotiledôneas e à família Poaceae. Esta planta vem sendo cultivada por mais de 7000 anos e alimenta diariamente pelo menos metade do nosso Planeta (IZAWA & SHIMAMOTO 1996), além de proporcionar uma grande fonte de renda em vistas de sua produção (590 milhões de toneladas produzidas por ano e 150 milhões de hectares utilizados para cultivo, anualmente). O Brasil é um dos maiores produtores mundiais de arroz, sendo esta produção destacada na região Sul, responsável por 60% da produção nacional, que totaliza cerca de 11 milhões de toneladas anuais (<http://sistemasdeproducao.cnptia.embrapa.br>).

Dentre as monocotiledôneas, o arroz é amplamente considerado como uma planta modelo para estudos, uma vez que diversos protocolos de manipulação genética estão bem estabelecidos, além de muitas informações moleculares, tais como sequência completa do genoma, bancos de ESTs (*Expressed sequence tags*), microarranjos, transcriptomas, bancos de mutantes, entre inúmeras outras informações estarem disponíveis. O arroz é um organismo diploide com 24 cromossomos e um genoma de 420 Mb (Megabases), sendo o menor genoma entre os cereais e com uma grande densidade gênica (GALE 1998; SHIMAMOTO & KYOZUKA 2002). A sub-espécie *japonica*, variedade Nipponbare, foi a primeira a ter o seu genoma completamente sequenciado e anotado, além de ser a planta de mais fácil transformação genética dentre os cereais utilizando *Agrobacterium tumefaciens* (SASAKI & BURR 2000). Com o avanço das tecnologias de sequenciamento, novas metodologias foram empregadas, baseadas em montagem *de novo* de todo o genoma partindo do sequenciamento por *shotgun*, a fim de aumentar a cobertura do genoma desta cultivar. O resultado foi uma cobertura de 93% do genoma, levando à predição de aproximadamente 32 mil genes, além de possuir homologia de 98% com proteínas conhecidas em milho, trigo e cevada, e também, uma grande sintonia cromossômica com estes cereais, mas com uma sintonia limitada quando comparado a *Arabidopsis thaliana* (GOFF *et al.* 2002; OUYANG *et al.* 2007).

Assim, tendo em vista a grande importância socioeconômica, o tamanho relativamente pequeno do genoma, a grande densidade gênica e a maior facilidade de manipulação genética, o arroz tornou-se um excelente modelo para estudos, possibilitando o entendimento de mecanismos biológicos básicos das plantas, necessários à sua sobrevivência e evolução.

1.2 Cistatinas e fitocistatinas

Cistatinas são inibidores competitivos de proteases cisteínicas (Cys proteases), interagindo com o sítio ativo das enzimas alvo, e bloqueando o seu acesso ao substrato. Em plantas, elas podem estar envolvidas em uma grande variedade de processos fisiológicos, desde o controle da proteólise endógena em órgãos reprodutivos e vegetativos, resposta a estresses bióticos e abióticos, até a inibição de proteases extracelulares de herbívoros, parasitas e patógenos (ZHANG, LIU, & TAKANO 2008; HWANG *et al.* 2009, 2010; CARRILLO *et al.* 2011b).

Primeiramente identificadas e caracterizadas em animais, as cistatinas se subdividem em três famílias de acordo com a homologia entre suas sequências, presença de pontes dissulfeto e a massa molecular da proteína. As cistatinas de tipo 1 ou stefinas, com uma massa molecular de aproximadamente 11kDa, possuem apenas um domínio inibitório e ausência de pontes dissulfeto; as de tipo 2 ou cistatinas propriamente ditas, também apresentam um domínio inibitório único, mas com massa molecular de 15kDa e com quatro resíduos cisteínicos formando duas pontes dissulfeto; e o último grupo, as cistatinas de tipo 3 ou kininogênios, com domínios inibitórios múltiplos semelhantes a stefinas, que possivelmente resultaram de eventos de duplicação (TURK & BODE 1991; BENCHABANE *et al.* 2010). A identificação das cistatinas baseia-se prioritariamente, na presença do domínio caracterizado como cistatina, o qual é formado por três motivos conservados: (1) um resíduo de glicina na região amino-terminal; (2) o motivo QxVxG, na região amino-terminal, e (3) um triptofano, carboxi-terminal, sendo que, estes dois últimos motivos estruturais localizam-se no primeiro e segundo *loops* entre as folhas- β da estrutura. Estes motivos são os principais responsáveis pela atividade inibitória das cistatinas, pois eles interagem

com o sítio ativo das enzimas-alvo, as proteases cisteínicas (BODE *et al.* 1988) (figura1).

Em plantas, estes inibidores de proteases cisteínicas são denominados fitocistatinas (PhyCys), pois diferem das cistatinas animais pela presença do motivo LARFAVDEHN em sua p-hélice amino-terminal e formam um *cluster* distinto das demais cistatinas identificadas em animais (MARGIS, REIS, & VILLERET 1998). As fitocistatinas são também desprovidas de sítios putativos de glicosilação e pontes dissulfeto. As pontes dissulfeto podem ser satisfatoriamente substituídas por outros tipos de interações moleculares que mantenham a estrutura funcional da proteína, como as interações hidrofóbicas dos resíduos aromáticos conservados, que ocorrem entre as folhas antiparalelas β 4- β 5 e a conformação do *loop*2 (MARGIS *et al.* 1998). Outras características comuns atribuídas às fitocistatinas por MARGIS *et al.* 1998 são: a ausência de resíduos cisteínicos em suas sequências e a conservação da distância de 17 aminoácidos entre os domínios LARFAV e QxVxG em todas as fitocistatinas. Apesar destas restrições particulares às fitocistatinas, grande parte das características moleculares das cistatinas e fitocistatinas manteve-se através do processo evolutivo dos eucariotos, uma vez que estes inibidores de proteases cisteínicas são amplamente distribuídos entre animais e plantas. Apesar das grandes divergências entre as sequências de aminoácidos, as stefinas, cistatinas e fitocistatinas possuem uma estrutura terciária conservada, formada por cinco folhas- β antiparalelas que interagem com uma das faces de uma p-hélice central (TURK & BODE 1991) (figura 1). Esta estrutura é composta pelos motivos proteicos conservados, característicos da superfamília das cistatinas, já descritos acima.

A região amino-terminal e, conseqüentemente, o resíduo de glicina é extremamente importante na determinação dos valores das constantes de associação e dissociação das proteases alvo, possibilitando uma forte atividade inibitória e aumentando a especificidade por cys proteases (MACHLEIDT *et al.* 1989; BJÖRK *et al.* 1994). Porém, foi demonstrado que para fitocistatinas este resíduo de glicina, apesar de conservado, não é estritamente necessário para a inibição das enzimas alvo (ABE *et al.* 1988).

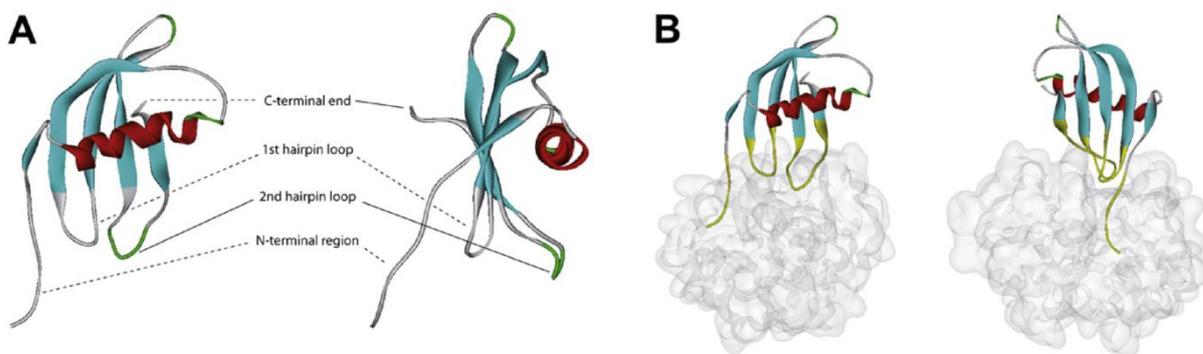


Figura 1. Modelo estrutural da fitocistatina-I de arroz (oryzacistatina). **A)** visão frontal e lateral demonstrando as cinco folhas- β (azul), a p-hélice (vermelho) onde se localiza o motivo LARRFAVDEHN, e os dois *loops* inibitórios (amarelo). **B)** visão da fitocistatina interagindo com a protease cisteínica modelo, a papaína (em cinza claro). Reproduzido de BENCHABANE *et al.* 2010.

A conservação da estrutura terciária, assim como dos motivos QxVxG e PW, nas estruturas de *loop* do inibidor, tanto em cistatinas quanto em fitocistatinas, sugere um mecanismo inibitório conservado para animais e plantas (BENCHABANE *et al.* 2010). Baseando-se no peso molecular, as fitocistatinas podem ser subdivididas em três grupos distintos, sendo a maioria incluída no grupo I e possuindo 12-16kDa. No grupo II encontram-se as fitocistatinas com aproximadamente 23kDa e que possuem uma extensão carboxi-terminal em sua sequência. O terceiro grupo inclui fitocistatinas de 80kDa, conhecidas como multicistatinas por possuírem diversos domínios de cistatinas (WANG *et al.* 2008). As fitocistatinas também podem ser agrupadas evolutivamente nos mesmos três grupos citados acima (Figura 2), de forma que as fitocistatinas do grupo I possuem apenas um domínio cistatina em sua estrutura proteica, com aproximadamente 100 resíduos, enquanto que as fitocistatinas de grupo II, com cerca de 200 resíduos, compartilham um domínio N-terminal altamente conservado com as de grupo I e um segundo domínio C-terminal também similar ao domínio cistatina. Esta estrutura em forma de dois domínios, das fitocistatinas de grupo II, é distintiva deste grupo em particular. E, ainda, no grupo III encontram-se as fitocistatinas com diversos domínios cistatinas repetidos (MARGIS-PINHEIRO *et al.* 2008; CHU *et al.* 2011).

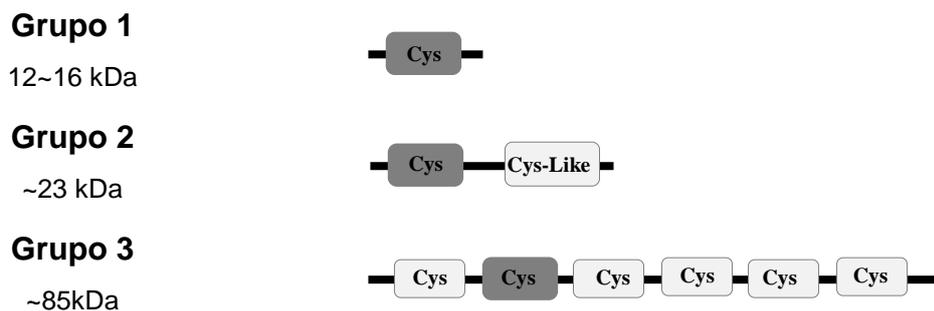


Figura 2. Três grupos de fitocistatinas baseados na massa molecular e no número de domínios cistatina presentes. Os domínios cistatina com estrutura resolvida estão demonstrados em cinza. Modificado de CHU et al. 2011.

1.3 Fitocistatinas e suas proteases-alvo

As proteases cisteínicas desempenham diversos papéis importantes através do ciclo de vida das plantas, mediante proteólise extensiva ou limitada de proteínas-alvo, a fim de desencadear a degradação ou a maturação das mesmas. Estas atividades são de extrema importância para o crescimento das plantas, armazenando e mobilizando proteínas, além de auxiliar na resposta a estresses ambientais (GRUDKOWSKA & ZAGDAŃSKA 2004). Os alvos primários das fitocistatinas, assim como das demais cistatinas, são proteases cisteínicas do tipo papaínas (família C1A) (BODE *et al.* 1988; ARAI, WATANABE, & KONDO 1991). As papaínas tem uma estrutura tridimensional de dois lobos conectados, compostos por α -hélices e folhas β , em forma de “V” e o sítio ativo localiza-se neste sulco que contém os resíduos catalíticos Cys-His (TURK & TURK 2001). Em angiospermas, o número de proteases cisteínicas da família C1A é bastante elevado, quando comparado com algas e outras plantas mais basais, o que pode significar uma grande especialização destas proteases para o desempenho de novas funções (MARTINEZ & DIAZ 2008) (tabela 1). As atividades desempenhadas pelas papaínas estão envolvidas em muitos processos fisiológicos importantes nas plantas vasculares (BEERS, JONES, & DICKERMAN 2004), podendo destacar-se a atuação sobre proteínas intracelulares, no

processamento, modificação e remobilização de proteínas de armazenamento durante a germinação das sementes (WATANABE *et al.* 1991; LEPELLEY *et al.* 2012), senescência, morte celular programada (BEERS, WOFFENDEN, & ZHAO 2000; HAYASHI *et al.* 2001; WAN *et al.* 2002), e ainda, em resposta à estresses (KHANNA-CHOPRA, SRIVALLI, & AHLAWAT 1999). Proteases C1A também são importantes na defesa contra patógenos e insetos, atuando contra as proteínas por eles produzidas e secretadas (CHEN *et al.* 1992; VAN DER HOORN & JONES 2004; DUTT *et al.* 2010). As papaínas possuem um papel de defesa contra herbívoros uma vez que elas podem ser encontradas no látex secretado pelas plantas, sendo capaz de matar ou intoxicar os organismos que dela se alimentam (KONNO *et al.* 2003). As papaínas de plantas foram classificadas em nove subfamílias (RD21A, CEP1, XCP2, XBCP3, THI1, SAG12, RD19A, AALP, CTB3) e análises bioquímicas demonstraram que estas proteínas existem em diferentes isoformas ativas, tendo atividade em variados níveis de pH e com sensibilidade alterada para diferentes inibidores (RICHAU *et al.* 2012).

Tabela 1. Número de proteases C1A (papaínas), C13 (legumaínas) e seus prováveis inibidores (cistatinas) em plantas. Reproduzido de MARTINEZ & DIAZ 2008.

Organismo	Cistatinas (com extensão C-terminal)	Papaínas	Legumaínas
<i>Osteococcus tauri</i>	0 (0)	9	0
<i>Osteococcus lucimarinus</i>	0 (0)	9	0
<i>Chlamydomonas reinhardtii</i>	1 (0)	11	1
<i>Volvox carteri</i>	1 (0)	14	1
<i>Physcomitrella patens</i>	5 (3)	11	4
<i>Selaginella moelendorffii</i>	2 (1)	19	2
<i>Oryza sativa</i>	12 (1)	45	5
<i>Arabidopsis thaliana</i>	7 (2)	32	4
<i>Populus trichocarpa</i>	8 (3)	36	5
<i>Hordeum vulgare</i>	13 (1)	32	5

Em 2007, MARTINEZ *et al.* reportaram a existência de fitocistatinas bifuncionais, pela presença de uma extensão carboxi-terminal (C-terminal) a qual apresenta o motivo proteico SNSL conservado, entre a p-hélice e a segunda folha- β da estrutura, sendo capaz de inibir proteases do tipo legumainas *in vitro*, semelhante ao mecanismo desempenhado pelos inibidores de legumainas em humanos (figura 3 e 4)

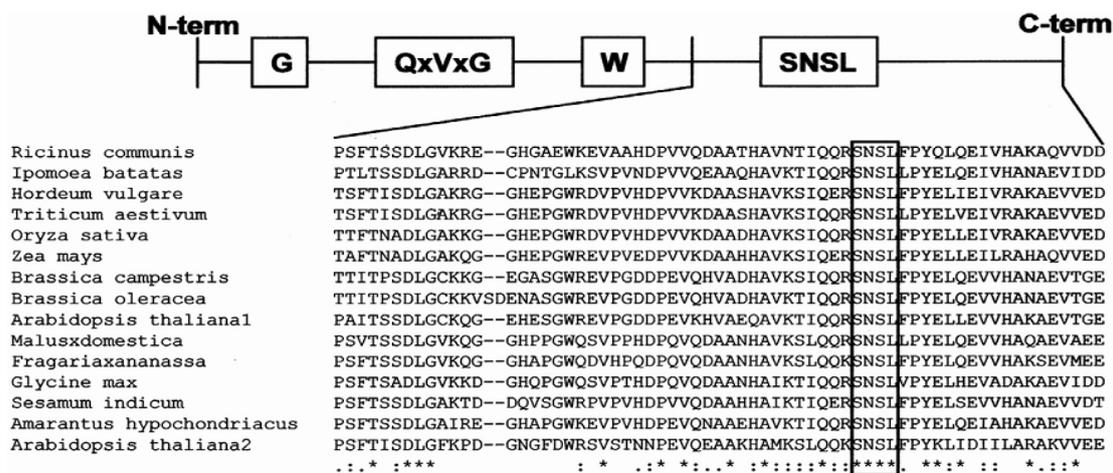


Figura 3. Fitocistatinas com a extensão carboxi-terminal, e o motivo conservado SNSL, encontradas em várias plantas. Reproduzido de MARTINEZ *et al.* 2007.

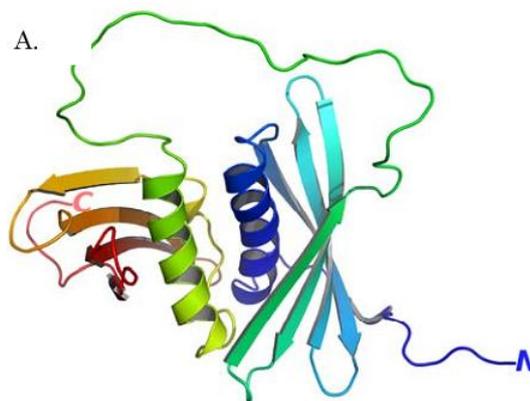


Figura 4. Modelo estrutural da tarocistatina (cistatina de taro, *Colocasia esculenta*), uma fitocistatina de grupo II, com a extensão carboxi-terminal (Ct) formando um novo domínio, capaz de inibir legumainas, pelo sítio SNSL. A região amino-terminal (Nt) mantém-se similar as demais fitocistatinas, representadas na figura 1. Reproduzido de WANG *et al.* 2008.

Estas proteínas, as legumainas, são amplamente distribuídas entre plantas e animais, sendo que em plantas seu número é bastante reduzido, quando comparado com as proteases cisteínicas da família C1A (MARTINEZ & DIAZ 2008) (Tabela 1). Legumainas são proteases cisteínicas que pertencem à família C13 (EC 3.4.22.34, MEROPS). Previamente descritas em plantas, estas proteases, também denominadas de *Vacuolar Processing Enzymes* (VPEs), são sintetizadas como propeptídeos, que por sua vez são co-transcricionalmente direcionados para o lúmen do retículo endoplasmático rugoso e, na forma de uma prolegumaina inativa, é transportada para o vacúolo ou para a parede celular (HARA-NISHIMURA, INOUE, & NISHIMURA 1991; MÜNTZ, BLATTNER, & SHUTOV 2002). Ao alcançar estes compartimentos, sob condições de pH ácido, a legumaina é auto-cataliticamente ativada (HIRAIWA, NISHIMURA, & HARA-NISHIMURA 1999; FISCHER *et al.* 2000; KUROYANAGI, NISHIMURA, & HARA-NISHIMURA 2002). Legumainas são também conhecidas como *asparaginil endopeptidases*, devido à sua habilidade de reconhecer e clivar resíduos de asparagina (Asn) ou ácido aspártico (Asp), em polipeptídeos (ABE *et al.* 1993; ISHII 1994; BECKER *et al.* 1995; HIRAIWA *et al.* 1999). Tais especificidades catalíticas colocam as legumainas como importantes componentes do processamento de diversas proteínas em suas respectivas formas maduras, para ativá-las, mobilizá-las, ou ainda estocá-las, de acordo com as necessidades metabólicas da planta (HARA-NISHIMURA *et al.* 1991).

Nas sementes de plantas, as legumainas são responsáveis pelo processamento pós-traducional de diversas proteínas para a biogênese e/ou armazenamento das mesmas, contribuindo assim para a biossíntese dos componentes vacuolares na semente e em alguns casos também em outros tecidos da planta, como endosperma, cotilédone, raiz, hipocótilo e folhas (HARA-NISHIMURA *et al.* 1995). Durante a germinação e o crescimento da plântula, legumainas ativam outras proteases cisteínicas (tipo papaínas da família C1A), as quais degradam proteínas de armazenamento para o desenvolvimento da planta (OKAMOTO & MINAMIKAWA 1999; KATO, SUTOH, & MINAMIKAWA 2003). Outro papel postulado para as legumainas é o de defesa contra patógenos por

meio da morte celular programada, e resposta hipersensitiva, devido à atividade similar a caspase observada em legumainas (HATSUGAI *et al.* 2004; ROJO *et al.* 2004). Em plantas, não é conhecida a presença de genes de caspases, portanto, outros genes, tais como as legumainas, são capazes de desempenhar função de caspase na morte celular programada (BONNEAU *et al.* 2008). As caspases são endopeptidases aspartato específicas, extremamente importantes para o processo de apoptose em células animais, mas que são funcionalmente relacionadas com as legumainas de plantas, além de possuírem similaridades na estrutura tridimensional e na especificidade pelo substrato (CHEN, RAWLINGS, & STEVENS 1998; SHUTOV *et al.* 2012).

Frequentemente, as legumainas de plantas são separadas em dois grandes grupos principais, (1) legumainas de semente e (2) vegetativas, devido ao seu padrão de expressão gênica em alguns tecidos e a homologia entre as sequências peptídicas (MÜNTZ *et al.* 2002; MÜNTZ & SHUTOV 2002). Estas características estão provavelmente relacionadas com a especificidade dos papéis desempenhados pelas legumainas, durante o ciclo de vida da planta. Entretanto, nem todas as legumainas se encaixam nesta classificação. Em *Arabidopsis thaliana*, quatro genes homólogos de legumainas estão relativamente bem descritos, pVPE e γ VPE são específicas de órgãos vegetativos, β VPE é específica de semente e a δ VPE não pertence a nenhuma destas duas subfamílias propostas, mas é expressa em um estágio central da embriogênese, antes da síntese das proteínas de armazenamento (YAMADA *et al.* 2005). A distinção das legumainas como vegetativa e de semente, pode ser associada à classificação vacuolar das plantas em vacúolos de armazenamento e vacúolos líticos. Desta forma, pVPE e γ VPE possuem expressão específica em vacúolos líticos de órgãos vegetativos e os alvos destas duas legumainas provavelmente sejam enzimas hidrolíticas vacuolares, induzidas pela senescência, já que a expressão destes genes de VPE não é detectada em tecidos de folhas saudáveis, mas aparecem naqueles em senescência, seguidos pela ativação de diversas proteases cisteínicas em resposta à senescência e degradação do tecido (HARA-NISHIMURA *et al.* 1998; KINOSHITA *et al.* 1999). Em algumas plantas, os homólogos de β VPE foram detectados em vacúolos de armazenamento proteicos,

desempenhando funções chave na maturação de proteínas (HARA-NISHIMURA & NISHIMURA 1987; HARA-NISHIMURA, TAKEUCHI, & NISHIMURA 1993; HARA-NISHIMURA *et al.* 1995; KINOSHITA *et al.* 1999). Também, em *Arabidopsis*, a expressão do gene de β VPE diminui após a germinação da semente e continua diminuindo gradativamente com o crescimento da plântula (KINOSHITA *et al.* 1999). As diferenças entre o padrão de expressão e funções destes grupos de legumainas podem estar relacionadas a diferentes elementos regulatórios na região promotora destes genes, os quais são responsáveis pela expressão espacial e temporal diferenciada das legumainas (KINOSHITA *et al.* 1999; MÜNTZ *et al.* 2002).

As atividades desempenhadas por proteases cisteínicas endógenas, sejam papaínas ou legumainas, são de extrema importância para o desenvolvimento vegetal e manutenção de sua homeostase celular, como demonstrado em vários dos trabalhos citados acima. A distribuição universal das fitocistatinas entre as plantas e a habilidade de inibição de proteases cisteínicas, juntamente com a conservação da estrutura terciária destes inibidores de proteases, sugere um papel fundamental no mecanismo de regulação de Cys proteases.

1.4 Funções biológicas atribuídas às fitocistatinas

Estudos iniciais em arroz haviam identificado apenas dois genes de fitocistatinas que foram denominados Oryzacistatina-I (Ocl) e Oryzacistatina-II (OclI) (ABE *et al.* 1987). Desde então, estas se tornaram as primeiras fitocistatinas bem caracterizadas. Tendo especificidade contra proteases cisteínicas, a expressão gênica da Ocl foi observada até no máximo duas semanas após o florescimento das plantas de arroz, e não detectada em sementes maduras, enquanto OclI é expressa constitutivamente através dos estádios de maturação e mesmo nas sementes já maduras, o que sugeriu o envolvimento destas fitocistatinas na regulação da proteólise desencadeada por proteases, além do mecanismo inibitório semelhante à cistatinas animais. Análises posteriores se seguiram com estas Oryzacistatinas, baseadas nestes primeiros estudos de caracterização e em outros que comprovassem a atividade inibitória contra papaínas (ABE *et al.* 1988; ARAI *et al.* 1991), e caracterizassem a

sequência genômica destas duas oryzacistatinas (HIROTO *et al.* 1989; KONDO *et al.* 1990, 1991). Após a identificação da Ocl e Ocll, os estudos estenderam-se para outras plantas tais como milho, trigo e soja, onde fitocistatinas também foram encontradas e sugeriu-se que estes inibidores ocorressem universalmente em plantas, regulando o catabolismo intracelular e atuando na defesa contra parasitas, uma vez que Ocl e Ocll são capazes de inibir proteases cisteínicas virais e do trato digestivo de insetos (ARAI, MATSUMOTO, & ABE 1998). Com a caracterização molecular das cistatinas em diversas plantas, notou-se que elas pertenceriam a uma nova classe de cistatinas, diferente das cistatinas de animais. Como foi então proposto, elas passaram a pertencer à classe das fitocistatinas (MARGIS *et al.* 1998).

As fitocistatinas possuem uma característica bastante importante, já abordada em diversos estudos, que é a de defesa contra patógenos, uma vez que elas são capazes de inibir as proteases secretadas por estes patógenos e auxiliar nos mecanismos de defesa da planta. Inicialmente, estudos *in vitro* e *in vivo* com a Oryzacistatina-I sendo expressa por *Escherichia coli*, apresentaram a capacidade inibitória desta proteína contra outras proteínas presentes no intestino médio de insetos, além da supressão do crescimento de *Tribolium castaneum* quando este era alimentado com o extrato da Oryzacistatina recombinante (CHEN *et al.* 1992). Aumentando-se a expressão de algumas fitocistatinas de soja (*Glycine max*), por lesões nos tecidos ou utilizando tratamento com metil jasmonato, foi possível verificar que as fitocistatinas que são induzidas nestas condições citadas, funcionam na planta hospedeira como um mecanismo de defesa contra os insetos *Diabrotica virgifera* e *Leptinotarsa decemlineata*, cuja composição do sistema digestório é principalmente de papaínas, e estas são inibidas pelas fitocistatinas da planta (ZHAO *et al.* 1996). LALITHA *et al.* 2005 confirmaram este estudo mediante experimentos *in vitro* e *in vivo*, e adicionaram o inseto *Callosobruchus maculatus* além das duas espécies citadas acima. Plantas transgênicas de *Arabidopsis thaliana*, expressando uma forma mutante de oryzacistatina (OCIΔ86), capaz de inibir a principal protease presente na glândula digestória da lesma *Deroceras reticulatum*, foram obtidas demonstrando, de fato, os efeitos antimetabólicos da inibição direta das enzimas proteolíticas do parasita

(WALKER *et al.* 1999). Atividades inibitórias semelhantes às fitocistatinas foram encontradas com as proteínas isoladas da anêmona *Actinia equina*, a planta *Coix lachryma-Jobi*, e de batata (*Solanum tuberosum*), contra o inseto *Tribolium castaneum*, quando combinada com inibidores de serino proteases (OPPERT *et al.* 2003). Desde então uma ampla gama de estudos vem sendo realizados a fim de aprimorar a utilização das fitocistatinas como inibidores de herbívoros que atacam plantações, otimizando a especificidade inibitória de cistatinas recombinantes contra pestes agrícolas (GOULET *et al.* 2008).

Além de inibir proteases cisteínicas presentes no trato digestório de herbívoros, uma série de estudos também revelou a importância na inibição de nematódeos parasitas. URWIN, ATKINSON, & WALLER 1995b demonstraram a importância da Oryzacistatina-I na resistência contra nematódeos, especificamente contra fêmeas de *Globodera pallida*, reduzindo sua fecundidade, quando expressa nos pelos das raízes de tomateiros (*Solanum lycopersicum*) e também em experimentos *in vitro*. A obtenção de plantas transgênicas de arroz, expressando a Ocl modificada molecularmente (OCIΔ86) também demonstrou-se eficiente na redução de 55% da produção de ovos pelo nematódeo *Meloidogyne incognita* (VAIN *et al.* 1998). Já em alfafa (*Medicago sativa*), plantas expressando Ocl e OclII, analisadas visualmente pela expressão concomitante do gene repórter *gus*, revelaram que a expressão destes genes é constitutiva em folhas, raízes e tecidos vasculares, sendo que sua expressão torna-se aumentada e localizada na raiz quando há a inoculação dos nematódeos (*Pratylenchus penetrans*) que lesionam este tecido e, as populações de nematódeos tornam-se cerca de 30% menores nas plantas expressando os transgenes Ocl e OclII (SAMAC & SMIGOCKI 2003). CHAN *et al.* 2010 demonstraram efetivamente que plantas de taro (*Colocasia esculenta*) superexpressando um gene de cistatina (*CeCPI*) suprime a formação de galhas, resultantes da infecção pelo nematódeo *Meloidogyne incognita*, além de diminuir consideravelmente o número de fêmeas do parasita e de ovos para a progênie.

Efeitos antivirais também já foram reportados para as fitocistatinas, testadas contra o Herpes Simplex Virus Type 1. Em células de macacos, a Ocl demonstrou-se efetiva e com efeitos antipiréticos, levantando a possibilidade de

utilização farmacológica (AOKI *et al.* 1995). Utilizando-se da informação de que o mecanismo de infecção por diversos vírus baseia-se em proteases cisteínicas, um importante estudo com a expressão constitutiva de um inibidor de proteases cisteínicas de arroz (Ocl), foi capaz de induzir a resistência contra dois tipos de vírus do tipo potyvirus (classe de vírus agronomicamente importantes, sendo transmitidos por ácaros para as plantas), e que utilizam proteases cisteínicas em suas vias de replicação (GUTIERREZ-CAMPOS *et al.* 1999). Uma vez que as fitocistatinas também são capazes de inibir proteases do tipo legumaínas, elas estão diretamente relacionadas com a hipersensibilidade em resposta a infecção por vírus, como demonstrado em tabaco (HATSUGAI *et al.* 2004), levando as células infectadas à morte celular programada, e proporcionando estratégias de defesa para a planta.

Outra ampla rede de pesquisas concentra-se em torno da aplicação das fitocistatinas como importantes antifúngicos. MARTINEZ *et al.* 2005 reportaram a importância do gene *Cyf1* de morango, que codifica uma fitocistatina, como um bom inibidor do crescimento *in vitro* dos fungos fitopatogênicos *Botrytis cinerea* e *Fusarium oxysporum*. Assim como a região amino-terminal da Tarocistatina de *Colocasia esculenta*, também é capaz de inibir o crescimento do fungo *Sclerotium rolfsii* (WANG *et al.* 2008). Estudos mais recentes continuam afirmando esta importância contra fungos fitopatogênicos, como da cistatina AhCPI, de *Amaranthus hypochondriacus*, que é capaz de deter o crescimento *in vitro* dos fungos *Fusarium oxysporum*, *Sclerotium cepivorum* e *Rhizoctonia solani*, propiciando uma atividade antifúngica (VALDES-RODRIGUEZ *et al.* 2010). Experimentos realizados com as treze cistatinas de cevada revelaram que todas possuem a capacidade de inibir, *in vitro*, diferentes tipos de fungos patogênicos, mas não são capazes de inibir bactérias patogênicas. Contudo, este mesmo estudo revelou que estas características podem não ser funcionais *in vivo*, uma vez que plantas transgênicas de *Arabidopsis thaliana*, expressando a cistatina HvCPI-6 de cevada, não eram mais resistentes a fungos e bactérias patogênicas do que as plantas controle, não transformadas (CARRILLO *et al.* 2011a). Tais estudos levantam diversas hipóteses sobre este mecanismo de inibição das

fitocistatinas e suas possíveis utilizações biotecnológicas para a transformação genética de culturas, a fim de torná-las mais resistentes a pestes e patógenos.

Os estudos que obtiveram sucesso em produzir plantas transgênicas utilizando inibidores de proteases, entre eles as fitocistatinas, para resistência à herbívoros e patógenos foram sumarizados por (SCHLÜTER *et al.* 2010) (Tabela 2), reforçando a ideia de que ainda mais estudos são necessários para demonstrar os efeitos que essas modificações trariam para o ecossistema geral e para as espécies que não são alvo destas proteases, caso as plantas expressando estas proteases recombinantes venham a se tornar comercialmente viáveis.

Tabela 2. Exemplos de estudos demonstrando o sucesso em obter plantas transgênicas resistentes a pestes e patógenos, utilizando inibidores de proteases cisteínicas. Modificado de SCHLÜTER *et al.* 2010.

Planta	Inibidor recombinante	Protease-alvo	Herbívoro-alvo
Insetos Herbívoros			
Alfafa	Oryzacistatina II	Cisteínica	<i>Phytodecta fomicata</i>
Batata	Fusões de multidomínios cistatina	Cisteínica +Aspartato	<i>Frankliniella occidentalis</i>
	Várias cistatinas de plantas e animais	Cisteínica +Aspartato	<i>Frankliniella occidentalis</i>
	Cistatina de cevada HvCPI-1 C ⁶⁸	Cisteínica	<i>Leptinotarsa decemlineata</i>
Tabaco	Esporamina + Taro cistatina	Serino+ cisteínica	<i>Helicoverpa armigera</i>
Nematóides parasitas de raízes			
Alfafa	Oryzacistatina I	Cisteínica	<i>Pratylenchus penetrans</i>
Batata	Oryzacistatina I	Cisteínica	<i>Meloidogyne incognita</i>
Tomate	Taro cistatina	Cisteínica	<i>Meloidogyne incognita</i>
Patógenos			
Tabaco	Esporamina + Taro cistatina	Serino + Cisteínica	<i>Pythium aphanidermatum</i>

1.5 Atuação das fitocistatinas em resposta a estresses

As plantas estão sujeitas a diversos estresses abióticos que modificam o ambiente natural em que vivem. Seca, excesso de salinidade, excesso de luz e temperaturas extremas são alguns exemplos e podem causar sérias perdas na agricultura (WANG, VINOGRAD, & ALTMAN 2003). Estresses ambientais podem ser definidos, em plantas, como mudanças nas condições de crescimento, dentro de seu habitat normal, que alteram ou perturbam a homeostasia metabólica do

organismo (SHULAEV *et al.* 2008) resultando em respostas moleculares, bioquímicas, fisiológicas e morfológicas que geralmente auxiliam a planta a tolerar o estresse, atingindo um novo estado de homeostase (MITTLER 2006). Assim, as relações entre as repostas gênicas e fisiológicas das plantas frente a diferentes tipos de estresses são importantes para compreender as bases deste mecanismo para sobrevivência e aclimação (SWINDELL 2006). As fitocistatinas podem ser de extrema importância, e ter sua expressão remodelada quando as plantas são submetidas a estresses bióticos e abióticos, a fim de aumentar a tolerância da planta a tais condições desfavoráveis de crescimento. Sendo assim, alguns estudos têm mostrado associações entre a expressão das fitocistatinas e os mais diferentes tipos de estresses aos quais as plantas possam ser submetidas (Tabela 3).

Tabela 3. Exemplos de fitocistatinas induzidas por estresses bióticos ou abióticos. Modificado de BENCHABANE *et al.* 2010.

Planta	Cistatina	Fonte	Tratamentos de estresse
Vias de estresse abiótico			
Amaranto	AhCPI	Pecíolo, raiz, folha	Seca, sal, frio, altas temperaturas
Arabidopsis	AtCYSa, AtCYSb	Folha, células	Seca, salinidade, frio, oxidação
Cevada	Icy	Folha, raiz	Anaerobiose, frio, altas temperaturas
Castanha	CsC	Folha, Raiz	Salinidade, frio, altas temperaturas
Feijão de corda	VuC1	Folha	Seca, ácido abscísico
Milho	CC8, CC9	Miolo	Frio
Cakile	CmC	Folha	Seca, sal
Trigo	TaMDC1	Plântula, folha e raiz	Frio, seca, sal, ácido abscísico
Vias de estresse biótico			
Castanha	Csc	Folha, raiz	Infecção por fungo, ferimento, ácido jasmônico
Batata	Multicistatina	Folha	Herbivoria
Soja	N2, R1	Folha	Ferimento, Metil-jasmonato,
Tomate	Multicistatina	Folha	Metil-jasmonato
	SICYS8 (TMC)	Folha	Metil-jasmonato
	SICIS9, SICIS10	Folha	Ácido araquidônico

Inicialmente, três genes de inibidores de proteases cisteínicas foram isolados em soja (*Glycine max*). Sabendo-se que a expressão de cada gene variava de acordo com os tecidos e os estádios de desenvolvimento da planta, foram realizados experimentos avaliando a expressão destes genes frente a ferimentos nas folhas da planta e a tratamentos com metil jasmonato, revelando que dois destes genes tinham sua expressão aumentada após estes tratamentos e outro apenas uma leve indução com o metil jasmonato (BOTELLA *et al.* 1996), assim como outros estudos utilizando metil jasmonato em outras plantas (FARMER, JOHNSON, & RYAN 1992; BOLTER 1993). Em castanha (*Castanea sativa*), PERNAS, SÁNCHEZ-MONGE, & SALCEDO, 2000 reportaram o envolvimento das cistatinas não apenas em respostas de defesa a patógenos (estresses bióticos), mas também em resposta a estresses abióticos. Em resposta a infecção por fungo, ferimentos e ácido jasmônico o mRNA da fitocistatina avaliada teve um aumento considerável nas folhas da planta, e em raízes frente aos dois últimos tratamentos. Também foi observado um grande aumento do número de transcritos, tanto em folha, quanto em raiz, quando as plântulas de castanha foram submetidas a choques salinos e de frio, além de aumentar em raízes estressadas por calor. Em cevada, o gene *Icy* que codifica uma fitocistatina, teve sua expressão detectada em embriões, endosperma em desenvolvimento, folhas e raízes, sendo que o nível de mRNA deste gene teve um aumento considerável nos tecidos vegetativos em resposta a anaerobiose, escuro e tratamento de frio em 6°C (GADDOUR *et al.* 2001). A caracterização de um gene de fitocistatina em grãos de *Amaranthus hypochondriacus* (*AhCPI*) também revelou um padrão de expressão tecido-específica, tendo uma maior expressão em sementes maduras e gradativamente decaindo durante a germinação, sendo que sua expressão nas hastes e nas raízes aumenta substancialmente em resposta a déficit hídrico, estresses por salinidade, frio e calor, sendo que o calor induz um acúmulo dos transcritos de *AhCPI* em folhas (VALDÉS-RODRÍGUEZ *et al.* 2007). Estudos mais detalhados em *Arabidopsis thaliana* (ZHANG *et al.* 2008), demonstraram que duas fitocistatinas, *AtCYSa* e *AtCYSb* (correspondendo às atuais *AtCYS3* e *AtCYS6*, respectivamente) têm perfis de expressão aumentados quando a planta é tratada com sal, seca, frio e

sofre estresse oxidativo, levando a uma tolerância da planta a estas condições adversas. Assim como os genes *AtCYS1* e *AtCYS2* possuem perfis únicos de expressão gênica durante a indução de estresses abióticos, altas temperaturas, seca e ferimentos aumentam a expressão destas duas fitocistatinas, mas os padrões temporais e espaciais desta indução diferem e apenas *AtCYS1* é induzida por frio (HWANG *et al.* 2010).

Todos estes trabalhos já publicados sugerem fortemente uma relação direta entre o aumento da expressão gênica de fitocistatinas e as respostas a estresses ambientais, além da identificação de *cis*-elementos nos promotores destes genes, os quais seriam importantes para a alteração da expressão gênica em determinadas condições ambientais. Uma observação importante retoma ao fato de que proteínas mal dobradas, desnaturadas e danificadas são eliminadas mediante degradação e substituídas por novas proteínas (GRUDKOWSKA & ZAGDAŃSKA 2004). A degradação e a síntese de novas proteínas, que dependem de enzimas proteolíticas, são componentes essenciais na resposta da planta a estresses abióticos. Desta forma, proteases cisteínicas acumulam-se quando tecidos são expostos a diferentes estresses abióticos, podendo degradar outras proteínas. Entretanto, o dano celular também leva à síntese de inibidores de proteases, capazes de regular a atividade destas proteases cisteínicas (YANG & YEH 2005). Tais mecanismos sugerem um controle fino sobre a degradação de proteínas em resposta ao estresse sofrido pela planta, a fim de contribuir para o aumento da homeostase interna do organismo, levando à sobrevivência do mesmo.

1.6 Fitocistatinas em arroz

Inicialmente, na década de 80, genes que codificam inibidores de proteinases cisteínicas de plantas, hoje denominados fitocistatinas, começaram a ser identificados em arroz. A primeira fitocistatina isolada e bem caracterizada foi a oryzacistatina-I (Ocl), encontrada em sementes de arroz (ABE, KONDO, & ARAI 1987b), sendo capaz de inibir efetivamente proteases cisteínicas do tipo papaínas. Desde então, diversos estudos seguiram-se a fim de identificar os domínios importantes para a atividade inibitória desta oryzacistatina (ABE *et al.*

1988), além de caracterizar sua sequência genômica e codificadora (HIROTO *et al.* 1989). KONDO *et al.* 1990 isolaram clones de cDNA de uma nova oryzacistatina, OclI, também encontrada em sementes de arroz, e com aproximadamente 55% de identidade com a Ocl, em nível proteico. As comparações feitas entre Ocl e OclI, neste estudo, revelaram algumas peculiaridades de cada gene, como a diferença entre a especificidade inibitória de cada um, onde Ocl inibe mais eficientemente papaínas do que OclI. Devido à detecção das proteínas referentes a Ocl e OclI em sementes maduras de arroz, foi postulado o envolvimento destes inibidores de proteases na regulação da proteólise causada por diferentes proteinases (KONDO *et al.* 1990). Seguindo da identificação funcional da OclI, a organização genômica da mesma foi caracterizada e comparada com a Ocl, as quais se mantêm similares (KONDO *et al.* 1991). Esta caracterização inicial desencadeou a busca de proteínas similares às oryzacistatinas em diversas outras plantas, além de uma contínua caracterização e validação da atividade inibitória das oryzacistatinas (ARAI *et al.* 1991; TURK & BODE 1991; URWIN, ATKINSON, & MCPHERSON 1995a), e o início da utilização destes inibidores de proteases cisteínicas, de forma recombinante, para avaliar o efeito sobre patógenos e parasitas (CHEN *et al.* 1992; URWIN *et al.* 1995b).

Durante a década de 80 e 90, não se sabia claramente quais seriam as relações evolutivas e as reais similaridades entre as cistatinas de animais e de vegetais. A oryzacistatina-I foi a primeira cistatina de plantas a ter sua estrutura cristalografada, com uma resolução de 2.8 Å (KUDO *et al.* 1998), e a estrutura tridimensional foi posteriormente melhor resolvida e mais detalhada por RMN (Ressonância Magnética Nuclear), revelando uma maior similaridade desta proteína com as cistatinas tipo 2 de animais, ao invés das de tipo 1 (stefinas) (NAGATA *et al.* 2000). Apesar de todo o esforço para uma classificação das cistatinas vegetais em tipo 2 ou tipo 1, de acordo com as de animais, um novo grupo foi proposto para a classificação destes inibidores de proteases cisteínicas de plantas, as fitocistatinas (MARGIS *et al.* 1998), baseado em relações filogenéticas, e características particulares de plantas, como já descrito anteriormente. Grande parte das descobertas realizadas em relação às

características e funções das fitocistatinas deve-se aos estudos aprofundados com as oryzacistatina I e II de arroz. Mais recentemente, em calos de arroz da variedade *nipponbare*, foi realizada a caracterização da terceira fitocistatina de *Oryza sativa*, a oryzacistatina III, altamente homóloga a oryzacistatina II, mas com algumas diferenças na região amino-terminal e em termos de organização no locus gênico. A OclIII, assim como as outras duas Ocs, possui a capacidade de inibir, *in vitro*, proteases do tipo papaínas, sendo que a constante de inibição da OclIII difere de Ocl e II (OHTSUBO *et al.* 2005). Este estudo também demonstrou que a maior eficiência de inibição de papaínas obedece a ordem Ocl > OclIII > OclII, além da OclIII ser a única das três com capacidade de inibir eficientemente a catepsina B, uma protease cisteínica amplamente encontrada em animais.

A disponibilização dos dados de sequenciamento do genoma do arroz proporcionou uma visão mais completa dos genes e suas famílias gênicas, além da possibilidade de estudos realizando inferências filogenéticas e evolutivas em geral, comparando todos os possíveis genes homólogos que possam existir no genoma deste organismo. MARTÍNEZ *et al.* 2005 utilizando os dados disponíveis do genoma do arroz encontraram e anotaram doze diferentes genes de fitocistatinas, os quais foram denominadas Ocl, OclII, OclIII, OclIV, OcV, OcVI, OcVII, OcVIII, OcIX, OcX, OcXI e OcXII de acordo com os respectivos loci gênicos: Os01g58890, Os05g41460, Os05g33880, Os01g68660, Os01g68670, Os03g11180, Os03g11170, Os03g31510, Os03g11160, Os04g28250, Os09g08100 e Os01g16430. Estes genes foram encontrados tanto no genoma da sub-espécie *japonica*, quanto em *indica*, com algumas variações de aminoácidos, apenas na região do peptídeo sinal. Segundo MARTÍNEZ *et al.* 2005, a organização genômica das fitocistatinas em arroz demonstra que três genes (Ocl, OclII e OclIII) possuem um intron entre as sequências que codificam o motivo conservado LARFAV e o sítio ativo QxVxG (figura 5). Oito sequências de arroz, (OclIV – OcXI) não possuem íntrons em suas ORFs (Open Reading Frame), enquanto que apenas um dos genes, OcXII, possui três íntrons. Estas sequências estão distribuídas em cinco cromossomos diferentes, sendo que algumas delas estão arranjadas de forma em *tandem* o que pode indicar a ocorrência de eventos

de duplicação entre estes genes no cromossomo I, entre as Ocs IV e V, e no cromossomo 3, entre as Ocs VI, VII e IX (figura 6).

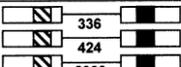
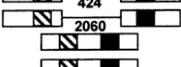
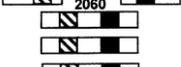
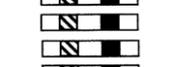
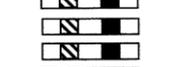
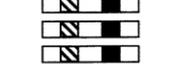
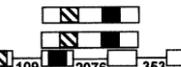
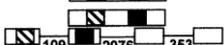
Gene name	Locus id	Predicted gene structure	Amino acids
OC-I	Os01g58890		139
OC-II	Os05g41460		156
OC-III	Os05g33880		150
OC-IV	Os01g68660		158
OC-V	Os01g68670		147
OC-VI	Os03g11180		113
OC-VII	Os03g11170		117
OC-VIII	Os03g31510		123
OC-IX	Os03g11160		114
OC-X	Os04g28250		156
OC-XI	Os09g08100		120
OC-XII	Os01g16430		250

Figura 5. Organização genômica das fitocistatinas em arroz. Os exons estão indicados pelas caixas, e os introns pelas linhas. O tamanho dos introns é demonstrado em pares de base (pb) abaixo da linha dos mesmos. O motivo LARFAV está representado nas caixas com linhas diagonais, enquanto que o motivo QxVxG está representado nas caixas pretas. Figura reproduzida de MARTÍNEZ *et al.* 2005.

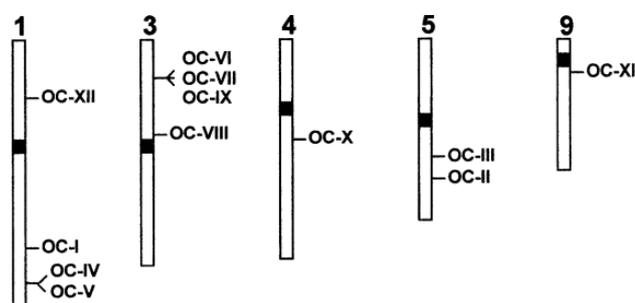


Figura 6. Distribuição cromossômica dos genes que codificam fitocistatinas em arroz. Figura reproduzida de MARTÍNEZ *et al.* 2005

Os motivos proteicos encontrados por MARTINEZ *et. al* 2005 nas sequências de proteínas das doze fitocistatinas de arroz, sugerem fortemente que todas elas possuem a estrutura tridimensional similar a Ocl, já descrita (NAGATA *et al.* 2000). Esta semelhança deve-se ao fato de que grande parte dos resíduos que constituem as cinco folhas- β e a p-hélice são conservados entre todas as proteínas, além da grande conservação dos motivos envolvidos na interação com as proteases cisteínicas. Entretanto, em três fitocistatinas de arroz OcVII, OcIX e OcXI o aminoácido triptofano (W), importante no segundo *loop* da estrutura da

proteína está ausente, sendo que este resíduo é de extrema importância para a efetiva inibição de proteases cisteínicas (KOIWA *et al.* 2001). A variação do sítio ativo destas fitocistatinas pode não estar diretamente relacionada com a perda de função das mesmas, mas sim, pode estar associada com diferentes propriedades inibitórias contra diferentes tipos de proteases cisteínicas, que foram adquiridas com o processo evolutivo das proteases e seus inibidores (MARTÍNEZ *et al.* 2005). A fim de validar a predição dos doze genes de fitocistatinas em arroz, os autores MARTINEZ *et al.* 2005, também buscaram pelo perfil de expressão destes genes em bancos de ESTs disponíveis no GenBank, contendo 299,998 ESTs de arroz em quatro diferentes tecidos: folha, raiz, flor e semente. A expressão da Oryzacistatina I foi encontrada em todos estes quatro tecidos analisados, enquanto que seis outros genes (OcIV, OcV, OcVI, OcVII, OcIX e OcXI) não foram detectados nas bibliotecas de ESTs analisadas. As demais cinco fitocistatinas apresentam um padrão mais restrito de expressão gênica, a OcII sendo expressa em folha, raiz e flor, a OcIII apenas em folha e flor, OcVIII em folha e semente, a OcX expressa em folha e raiz, e finalmente, a OcXII com a expressão detectada em folha e semente.

1.7 A fitocistatina XII de arroz e suas homólogas em plantas

A OcXII possui a maior sequência codificadora de proteínas (250 aminoácidos) entre as fitocistatinas de arroz. O transcrito do gene é composto por 4 exons e 3 introns (figura 5), e o *locus* gênico no cromossomo 1 (O1g16430) (MARTÍNEZ *et al.* 2005), sendo o único entre os doze genes descritos em arroz a possuir uma extensão carboxi-terminal. Interessantemente, esta extensão C-terminal possui motivos que são caracteristicamente reconhecidos em outras fitocistatinas (domínio cistatina, Cys, figura 2), sugerindo que a mesma tenha surgido através de uma duplicação interna de uma cistatina, seguida de uma série de divergências evolutivas levando a uma maior diversificação dos motivos típicos de fitocistatinas e originando um domínio cistatina degenerado (MARTÍNEZ *et al.* 2005). A diversificação do domínio C-terminal da OcXII, assim como de algumas outras fitocistatinas nas plantas em geral, levou à formação de um inibidor com capacidade bifuncional em plantas, podendo inibir tipos diferentes de proteases

cisteínicas, além das tipo-papaínas, comumente inibidas por fitocistatinas. Em 2007, MARTINEZ *et. al*, demonstraram a importância do motivo SNSL, na extensão C-terminal de uma fitocistatina de cevada (HvCPI-4), sendo este motivo capaz de inibir proteases cisteínicas do tipo legumaínas. Com base na similaridade entre as sequências de aminoácidos, das diferentes fitocistatinas de plantas com a extensão C-terminal (figura 3), é possível inferir que estas proteínas sejam capazes de desempenhar funções de maneira conservada em plantas. Desta forma, em arroz, este motivo encontra-se presente unicamente na fitocistatina XII, o que remete ao fato de que esta seja a única fitocistatina de arroz com capacidade para inibir proteases cisteínicas do tipo legumaínas, em arroz. Em outros organismos do reino vegetal, também foram encontrados ortólogos destas fitocistatinas com extensão C-terminal e motivo SNSL conservado: três genes em *Physcomitrella patens*, um em *Selaginella moellendorffii*, dois em *Arabidopsis thaliana*, três em *Populus trichocarpa* e um em *Hordeum vulgare* (MARTINEZ & DIAZ 2008).

Quando analisadas de forma comparativa e evolutiva, estas fitocistatinas com o domínio C-terminal extra em angiospermas estão evolutivamente relacionadas com as fitocistatinas ancestrais de algas. Contudo, em algas, a extensão C-terminal, responsável pela inibição de legumaínas, é separada do domínio cistatina característico por um *stop* codon, o qual não está mais presente nas demais fitocistatinas carboxi-estendidas de plantas, mas mesmo em algas a proteína é traduzida corretamente, indicando que a maquinaria de tradução não reconhece este códon como um sinal para o término da síntese proteica (MARTINEZ & DIAZ 2008). As fitocistatinas com extensão carboxi-terminal, quando analisadas evolutivamente com as demais fitocistatinas de estrutura canônica, podem ser identificadas em um grupo filogenético distinto (grupo II), estando mais relacionado com as fitocistatinas de grupo I e sugerindo que estas proteínas estendidas tenham se originado da duplicação de um gene ancestral deste grupo I de fitocistatinas (MARGIS-PINHEIRO *et al.* 2008). Assim, a OcXII e seu domínio C-terminal extra teriam se originado da duplicação de um gene ancestral ao grupo ao qual pertencem a OcI, OcII e OcIII.

1.8 Inibição de legumaínas por fitocistatinas tipo OcXII

As fitocistatinas, em arroz são consideradas uma família gênica composta por doze membros. Dentre estes genes, apenas a OcXII possui algumas características peculiares que ainda não foram satisfatoriamente elucidadas, principalmente relacionadas com a sua função, uma vez que este é o único dos doze genes que possui a extensão carboxi-terminal, com um motivo SNSL conservado e que teria a capacidade de inibir proteases cisteínicas do tipo legumaínas. Apenas em cevada, o ortólogo deste gene foi caracterizado e a função de inibidor de legumaínas relacionada ao motivo SNSL (MARTINEZ *et al.* 2007), sugerindo-se assim uma atividade bifuncional para este gene, inibindo papaínas e legumaínas, sendo que até então não se sabia da existência de inibidores de legumaínas em plantas.

No genoma do arroz, existem apenas cinco genes de legumaínas (Os01g37910, Os02g43010, Os04g45470, Os05g51570, Os06g01610) (MARTINEZ & DIAZ 2008) cujas funções ainda não estão devidamente estabelecidas, mas que provavelmente possuem grande importância no processo de desenvolvimento, crescimento e ciclo de vida das plantas. Assim como já descrito anteriormente em outras plantas, estas proteases cisteínicas podem atuar em diversos processos fisiológicos, desde o fornecimento primário de fontes de energia para a planta em germinação, mediante ativação de proteases do tipo papaínas, até a defesa contra parasitas e patógenos, atuando na morte celular programada das células comprometidas e auxiliando nos mecanismos de defesa da planta (MUNTZ *et al.* 2002). Desta forma, as fitocistatinas do tipo OcXII, que possuem o motivo SNSL C-terminal, são consideradas potenciais inibidores de legumaínas, sendo os únicos inibidores deste tipo de protease cisteínica já descrito em plantas (MARTINEZ *et al.* 2007).

No arroz como uma planta modelo que não apresenta redundância gênica para este provável inibidor de legumaínas (OcXII), o estudo das funções desse gene, torna-se de extrema valia para a compreensão dos mecanismos moleculares das fitocistatinas envolvidas na inibição de proteases cisteínicas do tipo legumaínas. Estas proteases, por sua vez, também se tornam importantes alvos para pesquisas, já que, em arroz, existem apenas cinco genes hipotéticos

de legumainas exercendo funções ainda não devidamente caracterizadas nesta planta. Desta forma, uma série de abordagens genéticas e moleculares em plantas de arroz, silenciadas para a fitocistatina XII, podem revelar aspectos funcionais deste gene, refletindo diretamente nas suas proteases-alvo, as legumainas, e conseqüentemente no desenvolvimento e crescimento das plantas. Há também a necessidade de estudos de genômica funcional, mediante manipulação genética de plantas, para a compreensão do padrão de expressão do gene da fitocistatina XII, revelando em que estádios de desenvolvimento, em quais tecidos e em quais condições ambientais o promotor deste gene é ativado, levando à expressão do mesmo e conseqüente inibição das legumainas-alvo. Além, de ampliar o conhecimento geral sobre todos os genes de fitocistatinas em arroz, avaliando seu padrão de expressão gênica e possíveis funções desempenhadas em estádios de desenvolvimento e condições fisiológicas específicas das plantas, auxiliando na compreensão dos mecanismos necessários para a sobrevivência do organismo.

2. OBJETIVOS

2.1 Objetivo Geral

Caracterizar os doze genes codificadores de fitocistatinas de arroz (*Oryza sativa*), em especial o gene da fitocistatina XII (OcXII), de forma a analisar as prováveis implicações funcionais deste gene na planta.

2.2 Objetivos específicos

2.2.1. Caracterizar funcional e evolutivamente os doze genes codificadores de fitocistatinas em arroz;

2.2.2. Analisar funcionalmente o gene codificador da OcXII mediante obtenção de plantas com este gene silenciado por RNAi;

2.2.2.1. Determinar o nível de silenciamento de OcXII nas plantas RNAi;

2.2.2.2. Verificar alterações no padrão de expressão gênica das demais Oryzacistatinas e legumaínas;

2.2.2.3. Observar alterações fenotípicas nas plantas;

2.2.2.4. Analisar a geração T1 das plantas silenciadas quanto ao nível de expressão gênica das Oryzacistatinas e legumaínas;

2.2.3. Analisar a região promotora de transcrição do gene da fitocistatina XII de arroz;

2.2.3.1. Clonar a sequência referente à parte do promotor do gene da fitocistatina XII em vetor para transformação de plantas, de forma que os genes repórteres GUS e GFP sejam expressos de maneira controlada pelo promotor de OcXII.

3. MATERIAIS E MÉTODOS

3.1 Caracterização *in silico* dos genes de fitocistatinas em arroz

Partindo do primeiro estudo de caracterização das fitocistatinas de arroz, realizado por MARTINEZ *et. al* 2005, os doze *loci* gênicos correspondentes às doze fitocistatinas de arroz foram buscados no banco de dados Phytozome (<http://www.phytozome.net>) (Ocl - Os01g58890, OclI - Os05g41460, OclII - Os05g33880, OclIV - Os01g68660, OclV - Os01g68670, OclVI - Os03g11180, OclVII - Os03g11170, OclVIII - Os03g31510, OclIX - Os03g11160, OclX - Os04g28250, OclXI - Os09g08100, OclXII - Os01g16430). A estrutura de exons e introns foi representada com o auxílio do programa Fancy Gene (RAMBALDI & CICCARELLI 2009), de acordo com as anotações depositadas no banco de dados do Phytozome. Para cada um dos doze genes foi realizada a busca do domínio proteico característico de cistatina no banco de dados SMART (LETUNIC, DOERKS, & BORK 2012), e também no banco InterProScan (QUEVILLON *et al.* 2005).

A caracterização da expressão gênica foi realizada primeiramente com o uso do banco de dados *Rice Oligonucleotide Array* (JUNG *et al.* 2008), onde estão depositados dados experimentais de microarranjos para genes de arroz. Os dados de expressão das fitocistatinas foram avaliados em diferentes tecidos e estádios de desenvolvimento de acordo com as *probes* disponíveis para dez dos genes de fitocistatinas de arroz (tabela 4). Complementando os experimentos avaliados no *Rice Array*, também foi realizada a busca pela expressão das fitocistatinas nos microarranjos disponibilizados pelo software *Genevestigator* (HRUZ *et al.* 2008), sendo que *probes* para os mesmos dez genes foram encontradas (tabela 4), e também puderam ser avaliados durante diferentes estádios de desenvolvimento e regiões anatômicas da planta, focando especificamente na fitocistatina XII.

Adicionalmente, a expressão gênica das fitocistatinas em arroz foi analisada utilizando-se dados disponíveis no laboratório, correspondendo ao sequenciamento do transcriptoma (RNA-seq) de plântulas de arroz (tecidos de folha e raiz separadamente). Estes dados foram gerados a partir do

sequenciamento do RNA total de plântulas de arroz (3 semanas de idade), pela plataforma Illumina (Hi-seq). Os *reads* obtidos no sequenciamento foram mapeados contra as sequências conhecidas dos transcritos das doze fitocistatinas, de forma que nenhum *mismatch* entre as sequências foi permitido (apenas pareamento perfeito). Para o mapeamento, foram utilizados *scripts* em Python e o programa Bowtie (LANGMEAD *et al.* 2009), além da utilização do programa Tablet (MILNE *et al.* 2010) para a visualização dos mapeamentos e contagem de *reads* em cada sequência. O número total de *reads*, que mapearam nas fitocistatinas, foram normalizados levando-se em consideração o número total de *reads* sequenciados em cada biblioteca (Folha e Raiz) e, posteriormente, normalizados pelo tamanho da sequência correspondente ao transcrito de cada gene.

Tabela 4. Lista de sequências utilizadas como *probes*, para cada gene das fitocistatinas de arroz, nos experimentos de microarranjos, no banco de dados do Rice Oligonucleotide Array e no Genevestigator.

Gene alvo: Ocl				
Nome	No.	Início	Fim	Sequência
Os.1316.1.S1_a_at	1	347	371	GGCCAATTCTCTGCTGGAGTTCGAG
	2	380	404	GAGTGTGAAGCAGCAAGTTGTCGCT
	3	391	415	AGCAAGTTGTCGCTGGCACTTTGTA
	4	400	424	TCGCTGGCACTTTGTA TACTATTTAC
	5	442	466	GGGATGCCAAGAAGCTCTATGAAGC
	6	459	483	TATGAAGCTAAGGTCTGGGAGAAAC
	7	475	499	GGGAGAAACCATGGATGGACTTCAA
	8	508	532	AGGAGTTCAAGCCTGTCGATGCCAG
	9	515	539	CAAGCCTGTCGATGCCAGTGCAAAT
	10	545	569	AGGCCATCTCGTATCCTATGTGTA
	11	559	583	TCCTATGTGTATCAAGTTATCAAGA
Gene alvo: OclI				
Os.7892.1.S1_a_at	1	311	335	GTTTCATGCACTACCTCACCGTCGAG
	2	325	349	TCACCGTCGAGGTGAAGGAACCCGG
	3	406	430	AGCAGCTCCAGGATTTCAAGCCCCT
	4	435	459	GACGCCACCGCCTAAACGTACATAG
	5	450	474	ACGTACATAGATCATCGTCCGGCTG

	6	544	568	CTTGCTGGAAGGTGTAATATGATGA
	7	566	590	TGAATAAAACCTGCTGCTTTGCCGG
	8	573	597	AACCTGCTGCTTTGCCGGGTCATAA
	9	581	605	GCTTTGCCGGGTCATAATAGACATA
	10	600	624	GACATATCACTTCTGTATTTCCCTAG
	11	615	639	TATTTCTAGTGCAATACACAACAT
Gene alvo: OcIII				
Os.11188.1.S1_at	1	437	461	TTAAGGCAGTCGTCGAGTTCCGTCA
	2	515	539	AAGATACTGCTGATCCCCTGTGGC
	3	545	569	GGAACGACCTGCACAACACTGAGAA
	4	573	597	TAAAGTCTCTGTTGTCCTGTCAACC
	5	596	620	CCTTTTCTCAGACATACTCGGTGTG
	6	611	635	ACTCGGTGTGAATCTAGCTCCATGG
	7	693	717	AGTAGCTGTTATCCTTTTGTGAAGT
	8	708	732	TTTGTGAAGTTTTCCCACTGTTAT
	9	817	841	GTGAACACTCCCATGTATATTCCAT
	10	840	864	ATACCTTTCGTTCCCATGCATGTAT
	11	882	906	ACCGTTGTATGGTATACCCGCTGCA
Gene alvo: OcIV				
Os.42764.1.S1_at	1	1041	1065	GAACGAGTCGTAGCACGCAGGCAGA
	2	1063	1087	AGAGAGAGGCAAGGTCACACAGGAC
	3	1075	1099	GGTCACACAGGACACAGCTGAGCTG
	4	1089	1113	CAGCTGAGCTGGTGTGAAGAATAAAA
	5	1118	1142	TACCAAGTTTCAGGTGCAGTTGTGC
	6	1151	1175	GGGTACTCGTATTCAGTAGAGAGTG
	7	1175	1199	GAACAAAGTAAGTGTACGTGTGCGT
	8	1187	1211	TGTACGTGTGCGTGTTCGCGCAATA
	9	1202	1226	TGCGGCAATAGATGTTTGAAGGCTT
	10	1217	1241	TCGAAGGCTTTTTGTGTAGAGAGAC
	11	1246	1270	GAGAGTGGCAATGAAGTACTATACT
Gene alvo: OcVI				
Os.57206.1.S1_at	1	141	165	AGGCTGGAGCCCGATAAAGAACATC
	2	202	226	ATCACGGAGAACAACCGGGTGTCCG
	3	232	256	GACGAGCTCACGTTCCACAGGGTCA
	4	365	389	AGAAGTGGAGCAACACGCGCAAGCT
	5	401	425	ACAAGAACCACAACCTTCTGAGCTCC
	6	423	447	TCCCCGGAGGGCGATCGATAGCTCT
	7	439	463	GATAGCTCTCGATCGGCAAGATTTA

	8	470	494	GAATAGCTCGAGTTGTTTTGATGTA
	9	483	507	TGTTTTGATGTAACCATTGCGTGGA
	10	499	523	TTGCGTGGATATTTTGTCCATCCAA
	11	514	538	GTCCATCCAATTGTTTTTCACATCG
Gene alvo: OcVIII				
Os.16142.1.S1_at	1	464	488	GTACGAGCAGTCGTGGACCAACACC
	2	535	559	CTCCCTCCAGATCTATCTTATCTAT
	3	548	572	TATCTTATCTATGTCTCTTCTGCTC
	4	568	592	TGCTCCATCTCGATCAGCTGTAAA
	5	583	607	AGCTGTAAATTTTCCCTGCCTAAT
	6	650	674	ATGAGCACAAGCTAAGTACCTGTAC
	7	663	687	AAGTACCTGTACTTGTTCGGTCTGA
	8	682	706	GTCTGAGTTGTGTCCCGTGTGTGC
	9	699	723	GTGTGTGCGGTTCAATCTTTTGCAT
	10	724	748	GTAAGACAGCGAGTCCACACTTGTA
	11	771	795	GGGATTGGAATAATGGTGCCCATCT
Gene alvo: OcIX				
OsAffx.3198.1.S1_at	1	64	88	GCGGCCGGCGACGAATCCTGGAAGA
	2	77	101	AATCCTGGAAGACGATCGACGCGAA
	3	83	107	GGAAGACGATCGACGCGAACGACCG
	4	89	113	CGATCGACGCGAACGACCGGCACGT
	5	99	123	GAACGACCGGCACGTCCAGGACGTC
	6	140	164	CGGAGACAGACTGGGCGTTCGGCGAC
	7	167	191	GCGGCCTCACGCTCAACACGGTGGGA
	8	173	197	TCACGCTCAACACGGTGGACGGCGC
	9	206	230	GGTTTGAGGCCGGCGTGAACACTACTA
	10	257	281	GCGTCGTCGCCAAGTACCTCAGGTT
	11	260	284	TCGTCGCCAAGTACCTCAGGTTCCA
Gene alvo: OcX				
Os.9817.1.S1_at	1	342	366	GCGTGGCCTACTACCTCAAGGTCGC
	2	438	462	TGAAGGCCTGGCTCAAATCCAAGGA
	3	473	497	TTCACGCCTGCTTCTTCTACCAAAT
	4	547	571	TACTAGTTGGACTTTACTCCCTCCG
	5	612	636	TAGTACTACGAATCAGACATCCGGT
	6	627	651	GACATCCGGTACCATAGTTTTTTAT
	7	688	712	ATAATTACGAAGCATGTGCGTGAC
	8	701	725	ATGTGCGTGACAGGCTTATACTCC
	9	709	733	GCACAGGCTTATACTCCTCCTTTTG

	10	725	749	CTCCTTTTGGAGTATATGGCATATG
	11	767	791	GGTTCCTTTGGAGTTTTACATATCG
Gene alvo: OcXI				
OsAffx.6201.1.S1_at	1	75	99	GTGGGTCGGCGGGTGAACGTGATC
	2	102	126	GGACGTTGCCGGCAACAATCAGATT
	3	115	139	AACAATCAGATTCAGCGGGTTCGGCG
	4	169	193	GGGACCAATGACAGGCTGCAGTTTCG
	5	175	199	AATGACAGGCTGCAGTTCGTCCGCG
	6	188	212	AGTTCGTCCGCGTGGTCGCAGCCGA
	7	229	253	GGGTCCAACCTATCTGGTTGTCATCG
	8	238	262	TATCTGGTTGTCATCGATGCTGCTA
	9	245	269	TTGTCATCGATGCTGCTAGCTCTCG
	10	258	282	TGCTAGCTCTCGCAAAAAGACGAGG
	11	287	311	TGTACGTGGCCGTTGTGGCTGACTT
Gene alvo: OcXII				
Os.6850.1.S1_at	1	567	591	TGGGTGGCGTGATGTTCCAGTACAT
	2	649	673	AACTCCCTGTTTCCATATGAACCTC
	3	667	691	GAACTTCTCGAGATCGTTCGTGCAA
	4	833	857	ATGAATCGAGCAGCCAGTGAACCTT
	5	849	873	GTGAACCTTGCCACATTAAGCTGTA
	6	898	922	AAAACAGTGAACATGCACCCTTCCG
	7	976	1000	ACTACACTATCCTTTTGTTCCTGTT
	8	991	1015	TGTTCCCTGTTGCTCTTAGTGTGATA
	9	1023	1047	ACTGGTCGTGTAACCTTCTCAAGGGA
	10	1045	1069	GGAACCGTGGGTCTATGCCATACAG
	11	1091	1115	GTTGATACTTCTCAAGTCACCACTC

3.2 Análises evolutivas das fitocistatinas

Os genes de fitocistatinas já identificados em arroz, por MARTINEZ *et al.* 2005, foram utilizados na busca por sequências de proteínas homólogas nos demais genomas de plantas, disponíveis no banco de dados Phytozome (<http://www.phytozome.net/>). Apenas as sequências contendo um ou dois domínios cistatinas conservados foram consideradas e alinhadas utilizando o programa MUSCLE (EDGAR 2004), implementado no software MEGA5, *Molecular Evolutionary Genetics Analysis* (TAMURA *et al.* 2011).

A construção dos alinhamentos e filogenias ocorreu de forma a analisar-se primeiramente as relações entre todas as fitocistatinas encontradas nos genomas de plantas disponíveis no Phytozome, analisando-se a região amino-terminal da proteína, com os motivos LARFAV, QxVxG, PW, e levando em consideração 109 sítios de aminoácidos para a análise evolutiva. Estas análises foram inferidas com o uso de 100.000.000 de gerações. Após a análise evolutiva global de todas as fitocistatinas encontradas, uma segunda análise foi realizada utilizando apenas as fitocistatinas com extensão carboxi-terminal. Para isto, foi realizada a análise da proteína inteira com os motivos LARFAV, QxVxG, PW e SNSL, considerando-se 230 sítios de aminoácidos. Estas análises foram realizadas verificando-se 50.000.000 de gerações.

Os alinhamentos gerados com as sequências proteicas foram manualmente inspecionados e somente posições não ambíguas foram utilizadas para análises filogenéticas no programa BEAST v.1.7.2, *Bayesian Evolutionary Analysis by Sampling Trees* (DRUMMOND *et al.* 2012). A análise bayesiana foi inferida utilizando os algoritmos de Cadeias de Markov e Monte Carlo (MCMC), o modelo de especiação de Yule e o modelo de substituição JTT (DRUMMOND *et al.* 2012). As relações evolutivas foram representadas pelos cladogramas com a maior soma de credibilidades, além dos valores dos nós serem estabelecidos pelos valores das médias destes cladogramas e demonstrados pelo valor de probabilidade posterior, calculados pelo software TreeAnnotator v.1.7.2 (DRUMMOND *et al.* 2012). As árvores filogenéticas foram visualizadas utilizando o software FigTree v.1.3.1 (RAMBAUT, A. 2008).

3.3 Material Vegetal

Plantas de arroz, *Oryza sativa* L. ssp. *japonica* cv. nipponbare, foram utilizadas para a realização do estudo. As plantas foram crescidas e mantidas em casa de vegetação com temperatura constante de 25°C e fotoperíodo de 12 horas.

3.4 Análise da expressão gênica das fitocistatinas em arroz por RTqPCR durante os processos de germinação e crescimento de plantas

Sementes de arroz foram descascadas e esterilizadas pela lavagem com hipoclorito de sódio 4% por 30 minutos sob agitação contínua. Após este processo, as sementes foram lavadas cinco vezes com água destilada autoclavada e em ambiente estéril. Por fim, 112 sementes foram colocadas em 8 frascos contendo meio MS (Sigma-Aldrich), totalizando 14 sementes por frasco. Estes frascos foram mantidos por três dias no escuro, com temperatura de 25°C e posteriormente colocados em sala de cultura *in vitro*, sob fotoperíodo de 12h luz/ 12h escuro e temperatura de 25°C.

Para a análise de expressão gênica nos diferentes estádios de desenvolvimento, o RNA total foi extraído com o reagente *Trizol* (Invitrogen) de um pool de três plantas e em quadruplicatas biológicas, nos seguintes momentos: semente madura, 12h de germinação, 24h de germinação, plântulas com 2 dias, plântulas com 4 dias, plântulas com 8 dias e folhas de plantas com 60 dias, além de folhas de plantas com 240 dias, em estágio inicial de senescência (figura 7). Este RNA foi então reversamente transcrito em cDNA utilizando um oligonucleotídeo (OligodT) complementar à cauda poli-A dos mRNAs expressos, com uma transcriptase reversa M-MLV (*Invitrogen Life Technologies*), de acordo com o protocolo do fabricante. Em seguida, experimentos de PCR em tempo Real (RTqPCR) foram realizados no aparelho *CFX384 BioRad*, utilizando o protocolo com o fluoróforo intercalante *SYBR Green*. Os experimentos foram executados com quadruplicatas biológicas e em quadruplicatas técnicas de pipetagem. Os genes constitutivos foram selecionados de acordo com os previamente testados por (JAIN *et al.* 2006), sendo que os genes *OsUbi*, *OseF1p*, foram os mais estáveis para a normalização da quantidade de mRNA presente em cada amostra, neste experimento. Para cada gene de interesse, dentre as fitocistatinas, foram projetados oligonucleotídeos específicos, utilizando o software *Primer3* (<http://frodo.wi.mit.edu/>), com temperatura média de anelamento de 60°C e visando a amplificação de produtos com tamanho de 100-200 nucleotídeos (tabela 5). A quantificação da expressão gênica relativa foi calculada com o

método matemático $2^{-\Delta\Delta Ct}$ (LIVAK, SCHMITTGEN, 2001). A significância estatística foi obtida com a análise de ANOVA, um fator, com o teste de Tuckey considerando $p < 0.05$, utilizando o software SPSS 20 (IBM corp).

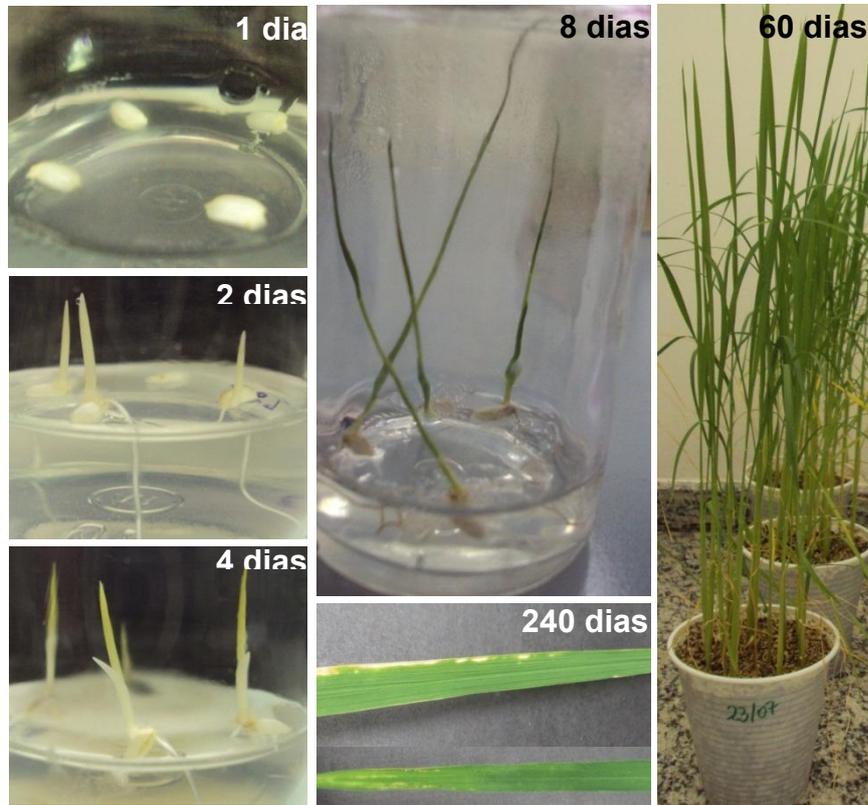


Figura 7. Estádios de desenvolvimento das plantas de arroz onde a expressão das fitocistatinas foi avaliada. Os estádios de semente seca e 12 horas de germinação não estão representados visto que estes não possuem diferenças morfológicas significativas a serem representadas. Os estádios de 24h (1 dia), 2 48h (2 dias), 96h (4 dias), 8 dias, 60 dias e 240 dias encontram-se representados pelas figuras e indicados no canto superior direito de cada uma.

Tabela 5. Sequências dos oligonucleotídeos utilizados para os experimentos de RTqPCR.

Oligonucleotídeos para RTqPCR			
Primer	Sequência	Primer	Sequência
Oc1_F	AAGGGGAGGAGAAGATGTCG	irOc12F	AGAAGGTGTACGAGGCCAAG
Oc1_R	GCTGCTTCACACTCACAAGC	irOc12R	TGGTCTGCAGCATCTTTGAC
Oc2_F	CGTCCCGTTACCCTCTTCT	Oc12_F_e1	GTCGACGAGCACAACAAGC
Oc2_R	TTCACCACCTCTCCAACCTC	Oc12_R_e2	GTTGGTGAAGGTGGTTGCAT
Oc2N_F	CGGAATCGGTGTTCACTACA	Oc12iF	CACCGGTCAAACCTCCCTGTTTCCA
Oc2N_R	ACGGTGGTGAGGTCGTTT	Oc12iR	TCCTCCTTGTTCCTCCTCTT
Oc3_F	TGAAGCCGTGGCTAGATTTT	Leg1F	GAAGGCCCAAGGGTTCTAAA
Oc3_R	CTCAGTGTTGTGCAGGTCGT	Leg1R	TGTTTGCAAAGGAACGCATA
Oc4_F	GGAAGGACGGAGGTGAGG	Leg2F	TGCAAATGCCGAAGTTAAGA
Oc4_R	CGACGCGGAGGTAGTACTTG	Leg2R	CTCCCTCAGAGCTTTCAGCTT
Oc5_F	GGCCTCCAAGCTGTATTACG	Leg3F	TAATCCTGCAAATGGCAACA
Oc5_R	GTGTTCCACGACGCAGAAG	Leg3R	TCCCTAAGAGCCCTCAGCTT
Oc6_F	CTGGAGCCCGATAAAGAACA	Leg4F	TGGACAGAAGCGTGGAGAT
Oc6_R	GATTTTCGAGGCGGTAGTTCA	Leg4R	ACCCACACTGATCCTCGAAC
Oc7_F	ATGACCATGAGGACCAGCA	Leg5F	GCCCCAACGACCATATCTT
Oc7_R	GTCACCCTCTCGAACGTCAG	Leg5R	CGAACATGCTACCCGACTC
Oc8_F	GACCCCCACATCCAGGAG	Os_eF1af	GACTTCCTTCACGATTTTCATCGTAA
Oc8_R	GTGTTGGTCCACGACTGCT	Os_eF1ar	TTTCACTCTTGGTGTGAAGCAGAT
Oc9_F	CAGCAGCCTCGTCCTGTT	Os_Ubi5f	ACCACTTCGACCGCCACTACT
Oc9_R	CGGCCTCAAACCTCTTCTC	Os_Ubi5r	ACGCCTAAGCCTGCTGGTT
Oc10_F	CGCAGAAGCAGGTCGTCT	Os_FDHF	CAAAATCAGCTGGTGCTTCTC
Oc10_R	GGTAGAAGAAGCAGGCGTGA	Os_FDHR	TTCCAATGCATTCAAAGCTG
Oc11_F	CCGGCAACAATCAGATTCA	Os_GAPDHf	GGGCTGCTAGCTTCAACATC
Oc11_R	GCTAGCAGCATCGATGACAA	Os_GAPDHR	TTGATTGCAGCCTTGATCTG
		Os_snU6f	GGGGACATCCGATAAAATTG
		Os_snu6r	AGGGGCCATGCTAATCTTCT

3.5 Construção do vetor para silenciamento gênico de OcXII

A fim de produzir plantas de arroz silenciadas para o gene da fitocistatina XII, o cDNA correspondente a este gene foi isolado a partir da extração de RNA total de folhas de arroz com o reagente Trizol (*Invitrogen Life Technologies*). Este RNA foi então reversamente transcrito em cDNA utilizando um oligonucleotídeo (OligodT) e uma transcriptase reversa M-MLV (*Invitrogen Life Technologies*), de acordo com o protocolo do fabricante. O cDNA obtido foi utilizado em uma reação

de PCR (*Polymerase chain reaction*), utilizando uma polimerase de alta fidelidade, Phusion[®] (Finnzymes), e oligonucleotídeos específicos para a amplificação de um fragmento de 138 pares de bases (pb) da região carboxi-terminal da OcXII, (Tabela 6, primer Oc12i). Este fragmento específico de 138pb foi inserido em um vetor de tecnologia Gateway, pENTR (Invitrogen). Mediante transformação por choque térmico, o vetor pENTR foi transferido para células competentes de *Escherichia coli* Top10. Para esta transformação, 1 a 2 µL do produto da recombinação foram adicionados a 80 µl de células competentes. A mistura foi então mantida por 30 minutos a 0°C, 2 minutos a 42°C e por mais 5 minutos a 0°C. A seguir, 400µL de meio LB líquido (Invitrogen) foram adicionados às células, e as mesmas incubadas por 1 hora a 37°C em agitação constante. Volumes de 200µL da cultura contendo as bactérias foram plaqueadas em meio LB sólido (Invitrogen) contendo canamicina (25 µg/mL), para a seleção das bactérias transformadas. As placas foram mantidas na estufa a 37°C por 16 horas. As colônias de bactérias que cresceram no meio contendo o antibiótico foram selecionadas e crescidas em meio LB líquido contendo canamicina (25 µg/mL), e o DNA plasmídial foi extraído utilizando a metodologia de lise alcalina em miniprep, descrita abaixo (SAMBROOK & RUSSEL 2001). Células de 1,5mL da suspensão de bactérias foram concentradas por centrifugação em 13,000x g por 2 minutos. Após o descarte do sobrenadante, 200 µL da solução A (Tris-HCl 25 mM pH 8,0; EDTA 10 mM) foram adicionadas às células e a mistura foi homogeneizada em vórtex. Em seguida, a solução B (200 µL, contendo NaOH 0,2 N e SDS 1%) foi misturada lentamente por inversão à mistura anterior. Posteriormente, 200 µL de Acetato de Potássio 5M foram adicionados às outras soluções e homogeneizados por inversão. Após centrifugação por 5 minutos a 13,000x g, 600 µL do sobrenadante foram transferidos para novo tubo. Os plasmídeos foram então precipitados utilizando 600 µL de isopropanol e centrifugação por 20 minutos em 13,000x g. O sobrenadante foi descartado e o precipitado lavado com 350 µL de etanol 70%. A mistura foi centrifugada por 5 minutos a 13,000x g. O sobrenadante foi descartado e o precipitado foi seco em estufa a 37°C, por 20 minutos. O DNA plasmídial foi resuspenso em 50 µL de água deionizada autoclavada.

As transformações das colônias de bactérias selecionadas foram confirmadas por PCR utilizando oligonucleotídeos iniciadores específicos para a construção contendo a sequência OcXII C-terminal específica. Tendo sido confirmada a presença do fragmento OcXII no vetor pENTR, o mesmo foi sequenciado pelo método de Sanger (EBI, Applied Biosystems) e posteriormente, após a confirmação da identidade da sequência nucleotídica, utilizado para a recombinação com o vetor de RNAi, pANDA (MIKI & SHIMAMOTO 2004) utilizando a LR clonase (Invitrogen) de acordo com as especificações do fabricante. No vetor pANDA (figura 8), a construção específica para o gene de interesse (OcXII) permite a formação de hairpins, os quais são clivados pela maquinaria de RNA de interferência celular, direcionando pequenos RNAs para a região C-terminal do transcrito da OcXII, assim, silenciando somente o gene de interesse, dentre as demais fitocistatinas. Para que este silenciamento ocorra, é imprescindível que a sequência do gene de interesse insira-se em dois locais do vetor binário, de maneira que uma fique invertida em relação à outra. Portanto, após a transformação do vetor pANDA o mesmo é extraído das bactérias (conforme descrito acima) e a orientação das inserções confirmada por meio de PCR utilizando uma combinação de oligonucleotídeos onde um deles anele de forma reversa no gene *gus-linker* e o outro sendo o primer reverso do fragmento OcXII clonado, resultando em um fragmento de 870pb. Outra combinação de oligonucleotídeos compreende um primer *forward* no gene *gus-linker* e também o *primer* reverso do fragmento OcXII, resultando em 968pb (tabela 6).

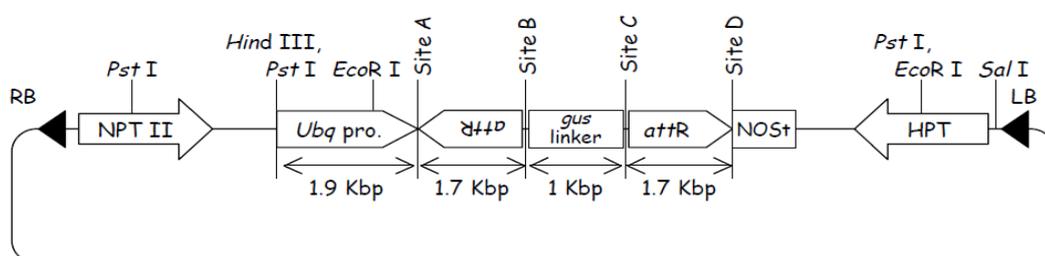


Figura 8. Estrutura do vetor pANDA, utilizado para silenciamento de genes via RNAi. Reproduzido de MIKI & SHIMAMOTO 2004.

Tabela 6. Sequências dos oligonucleotídeos utilizados para a clonagem dos fragmentos da fitocistatina XII, construção dos vetores de silenciamento e estudo do promotor.

Oligonucleotídeos utilizados para clonagens e confirmação dos vetores	
Primer	Sequência
PromoOc12F	CACCGCAGAGAGCCAGAAGAAGGA
PromoOc12R	ATGACGGTTTGGCTGTTGAG
Oc12iF	CACCGGTCAAACCTCCCTGTTTCCA
Oc12iR	TCCTCCTTGTTCCTCTT
GusLinker_F	CATGAAGATGCGGACTTACG
GusLinker_R	ATCCACGCCGTATTCGG
HPT_F	GAGCCTGACCTATTGCATCTCC
HPT_R	GGCCTCCAGAAGAAGATGTTGG
pHGWFS7 Hyg_F	AACGTTGCGGTTCTGTCACT
pHGWFS7 Egfp_R	TGGTGCAGATGAACTTCAGG

3.6 Transformação de *Agrobacterium tumefaciens*

A transformação dos calos embriogênicos de arroz (*Oryza sativa* L. ssp. *japonica*, cv. Nipponbare), dá-se por meio de *Agrobacterium tumefaciens*. Para isto, células de *A. tumefaciens*, cepa AGL1, foram utilizadas para a inserção do vetor pANDA, contendo o fragmento OcXII, por meio de eletroporação, onde 1 µL deste vetor foi adicionado a 80 µL de células eletrocompetentes da bactéria. A mistura foi eletroporada em cubeta de 0,2 cm utilizando-se um pulso elétrico de 2.500V (Eletroporador Bio Rad, Gene Pulser® II). Em seguida, 400 µL de LB líquido (Invitrogen) foram adicionados às células e o material foi transferido para um tubo de microcentrífuga de 1,5 mL e incubado por 2 horas a 28°C em agitação constante. 200 µL de células foram espalhadas em duas placas com LB sólido (Invitrogen), contendo 100 µg/mL de rifampicina e 25 µg/mL de canamicina. As placas foram mantidas a 28°C por dois dias. Após a obtenção de colônias isoladas, foi extraído o DNA plasmidial por lise alcalina (Sambrook et al 1989), descrita no item anterior (3.5), e a transformação de *A. tumefaciens* foi confirmada por PCR. Nestas reações de PCR confirmatórias utilizaram-se os mesmos

oligonucleotídeos iniciadores, específicos da sequência OcXII introduzida no vetor pANDA mais os primers do gene *gus-linker* (870 pb e 968 pb), além de oligonucleotídeos que amplificam uma região do gene de resistência a higromicina (*HPT*), resultando em um fragmento de 500 pb, e os *primers* específicos da região c-terminal clonada de OcXII (138 pb) (tabela 6).

3.7 Transformação genética de calos embriogênicos de arroz

A etapa de transformação das plantas foi realizada com calos embriogênicos de arroz, obtidos a partir de sementes maduras da cultivar Nipponbare, subespécie *japonica*. Os grãos foram descascados manualmente e, em seguida, desinfetados. A desinfecção foi feita mediante a lavagem das sementes com alvejante a 50% (hipoclorito de sódio 8%), por 30 minutos com agitação contínua. Sob condições assépticas, estas sementes foram lavadas seis vezes com água destilada estéril e colocadas em placas com meio sólido de indução de calos N6D (UPADHYAYA *et al.*, 2002). As placas, contendo em média 10 sementes, foram mantidas na luz a 25 °C por um período de 6 a 8 semanas, para que a formação de calos fosse induzida. Após a indução, os calos foram transferidos para placas contendo meio de proliferação NB (UPADHYAYA *et al.*, 2002), e colocados a 28°C no escuro em estufa. Esses calos foram cultivados neste meio de cultura e repicados (transferidos para meio novo) a cada quatro semanas.

A transformação de arroz via *A. tumefaciens* foi realizada segundo o protocolo de (UPADHYAYA *et al.*, 2002), e descrita abaixo. Para a realização do experimento, 100 calos jovens (idade máxima de 3 meses) foram selecionados e repicados em placas de petri contendo meio NB, de forma que cada placa continha aproximadamente 25 calos, os quais foram mantidos a 28°C no escuro. Três dias após a repicagem dos calos, uma colônia de *A. tumefaciens* contendo a construção RNAi (fragmento C-terminal da OcXII inserido no vetor pANDA) foi selecionada e estriada em meio LB sólido o qual permaneceu em crescimento por dois dias a 28°C. No quinto dia após a repicagem dos 100 calos foi realizada a transformação dos mesmos da seguinte maneira: as *A. tumefaciens* foram transferidas para um frasco contendo 50mL de meio NB líquido e 100 µM de acetoseringona. Posteriormente, os calos repicados há cinco dias foram

incubados nesta solução por 15 minutos e então transferidos para placas contendo meio NB sólido com 100 µM de acetoseringona. Após um período de 48 horas de co-cultivo com a *A. tumefaciens*, os calos foram lavados em solução composta por 50 mL de água estéril e 150mg/L de timentina, sendo mantidos em agitação a 28°C por 10 minutos, e este processo foi repetido por cinco vezes, a fim de eliminar completamente as bactérias dos calos de arroz. Em seguida, os calos foram transferidos para meio NB-HTC (UPADHYAYA *et al.*, 2002), composto por higromicina (50 mg/L), timentina (150 mg/L) e cefotaxima sódica (200 mg/L) (visando a seleção dos calos transformados resistentes a higromicina e a eliminação de contaminações por *A. tumefaciens*). Os calos foram mantidos na estufa incubadora, no escuro, a 28 °C por cerca de 30 a 40 dias. Após este período, os calos resistentes a higromicina mantêm sua coloração mais clara, enquanto que os não-transformados ficam mais escuros, o que permite a seleção manual dos calos resistentes para novas placas de meio NB-HTC, onde os calos permanecem novamente por cerca de um mês, no escuro a 28 °C, para que proliferem.

Após este período de seleção e proliferação dos calos transformados, os mesmos são transferidos para meio de pré-regeneração (UPADHYAYA *et al.*, 2002), onde permanecem por um período de 8 a 12 dias no escuro a 28 °C e são posteriormente transferidos para o meio de regeneração (UPADHYAYA *et al.*, 2002), o qual permanece sob fotoperíodo de 12 horas luz/12 horas escuro, com temperatura constante de 28°C. Novamente, após um período de 20 a 30 dias, os calos transformados regeneram em plantas, as quais são transferidas para meio MS (Sigma-Aldrich) contendo 50 mg/L de higromicina. O mesmo procedimento de regeneração de plantas controle (não-transformadas) é realizado paralelamente, porém as plantas não são submetidas a higromicina.

As plantas regeneradas são mantidas a 28 °C, sob fotoperíodo de 12 horas luz/12 horas escuro por pelo menos 15 dias em meio MS. Posteriormente as plantas resistentes a higromicina são transferidas para vermiculita e permanecem fechadas com sacos plásticos para aclimação por 10 dias. As plantas são mantidas com solução nutritiva (HOAGALND & ARNON, 1950), regadas três vezes por semana e, após 30 a 40 dias as mesmas são transferidas da

vermiculita para vasos com solo, os quais ficam imersos pela metade em água. Já na terra, estas plantas permanecem em temperatura constante de 28 °C com fotoperíodo de 12 horas luz/12 horas escuro.

3.8 Confirmação das plantas transgênicas

As plantas que passaram pelo processo de seleção em higromicina foram utilizadas para a confirmação da transformação genética. Para isto, foram pulverizadas em nitrogênio líquido dentro de um tubo de microcentrífua de 1,5 mL. Após a pulverização, foi adicionado em cada tubo 600µL de tampão de lise (Tris/HCl 10 mM, pH 8.0, EDTA 1 mM, SDS 1%, Protocolo PureGene CSIRO *Plant Industry*), os quais foram agitados e incubados a 65°C por 1 hora. Posteriormente, as amostras foram centrifugadas a 13.000x g por 3 minutos, onde o sobrenadante foi retirado para um tubo novo, e 200 µL de Acetato de Potássio 6M foram adicionados. Após nova centrifugação por 10 minutos, o sobrenadante foi transferido para um tubo novo, onde foram adicionados 600 µL de isopropanol, misturado por inversão 50 vezes, e centrifugado novamente por 10 minutos. O sobrenadante foi então descartado e o precipitado lavado com etanol 70%, sendo posteriormente colocado em estufa a 37°C por 15 minutos, solubilizado em água deionizada autoclavada e estocado em freezer -20 °C.

A partir do DNA extraído de cada planta, foram realizadas reações de PCR a fim de confirmar se as mesmas eram transgênicas. Para isto, foram utilizados oligonucleotídeos iniciadores específicos para o gene HPT (higromicina), e também iniciadores para a região do *gus-linker*, presente no vetor pANDA (tabela 6). Tais regiões amplificam fragmentos de 500 pb e 636 pb, respectivamente indicando que a construção foi inserida no genoma da planta.

3.9 Confirmação do silenciamento de OcXII e nível de expressão gênica das demais fitocistatinas e legumainas de arroz

A confirmação do nível de expressão gênica foi realizada pela quantificação relativa do mRNA dos genes de interesse, em amostras de tecido vegetal das plantas RNAi obtidas. Folhas foram coletadas e imediatamente congeladas em

nitrogênio líquido. O RNA foi extraído das amostras utilizando o reagente Trizol (*Invitrogen Life Technologies*). O cDNA foi obtido utilizando-se um oligonucleotídeo (OligodT) e uma transcriptase reversa M-MLV (*Invitrogen Life Technologies*), de acordo com o protocolo do fabricante, e armazenados a – 20°C. Em seguida, experimentos de PCR em tempo Real (RTqPCR) foram realizados no aparelho *CFX384 BioRad*, utilizando o protocolo com o fluoróforo intercalante *SYBR Green*. Os experimentos foram executados em quadruplicatas técnicas de pipetagem. Os genes constitutivos (OseF1p e OsFDH) foram selecionados de acordo com os previamente testados por (JAIN *et al.* 2006), para a normalização da quantidade de mRNA presente em cada amostra. Para cada gene de interesse, dentre as doze fitocistatinas e as cinco Legumainas (Leg1 - Os01g37910, Leg2 - Os02g43010, Leg3 - Os04g45470, Leg4 - Os05g51570, Leg5 - Os06g01610), foram projetados oligonucleotídeos específicos com o software Primer3 (<http://frodo.wi.mit.edu/>), com temperatura média de anelamento de 60°C e visando a amplificação de produtos com tamanho de 100-200 nucleotídeos (tabela 5). A quantificação da expressão gênica relativa foi calculada com o método matemático $2^{-\Delta\Delta Ct}$ (LIVAK, SCHMITTGEN, 2001).

3.10 Alterações fenotípicas das plantas OcXII silenciadas

As análises fenológicas das plantas T0 foram realizadas pela observação da morfologia da parte aérea e das raízes, além de medições do comprimento da parte aérea. Também foram realizadas comparações levando em consideração o número de panículas, número de sementes por panícula, número de sementes por planta, número de sementes viáveis, massa das sementes e biomassa das folhas e das sementes. A fim de avaliar as diferenças entre as plantas controle e as plantas silenciadas, análises estatísticas foram realizadas, utilizando ANOVA, um fator, e teste de Tuckey com $p < 0.05$.

3.11 Obtenção e análise da progênie das plantas silenciadas OcXII

As sementes oriundas das linhagens de plantas T0, silenciadas para a fitocistatina XII de arroz, constituem a T1 das linhagens transformadas. Assim,

estas sementes permaneceram a uma temperatura de 37°C por três dias para a quebra de dormência. Após, as sementes foram descascadas e esterilizadas da mesma forma descrita no item 3.7. As sementes estéreis foram colocadas em meio MS, no escuro, a 25 °C por três dias e, posteriormente, transferidas para sala de cultura *in vitro*, com temperatura constante de 25 °C e fotoperíodo de 12 horas luz/ 12 horas escuro, onde permaneceram por cerca de 30 dias. Após o período de crescimento *in vitro*, as plântulas foram transferidas para vermiculita e regadas a cada dois dias com solução nutritiva. A partir de tecido de folhas das plantas em vermiculita, foi realizada a extração de DNA e a reação de PCR com o gene *gus*, como já descrito no item 3.8, a fim de confirmar a transgenia das mesmas, pela presença do vetor de silenciamento. As plantas transgênicas, confirmadas pelo PCR anterior, foram submetidas à extração de RNA, a partir de uma folha de cada planta, seguida da síntese de cDNA e experimentos de RTqPCR, da forma descrita no item 3.9. As plantas cujo nível de expressão de OcXII estava significativamente diminuído foram transferidas para potes contendo solo, e imersos em água até a metade, sendo mantidas em casa de vegetação com fotoperíodo de 12 horas luz/ 12 horas escuro e temperatura de 25 °C.

3.12 Caracterização da sequência de nucleotídeos de uma região do promotor do gene OcXII

A região de 2.300 pares de bases (pb), da sequência genômica que antecede o sítio de início da tradução do gene da fitocistatina XII, foi denominada de região promotora desse gene. Esta região de 2.300pb foi utilizada para buscas de possíveis sequências de elementos regulatórios (*cis*-elementos), a fim de revelar se fatores de transcrição seriam capazes de se ligar nesta região promotora e ativar o gene OcXII, em diferentes tecidos e/ou estádios de desenvolvimento das plantas. Para esta caracterização foram utilizados os bancos de dados PlanCARE (LESCOT *et al.* 2002) e PlantPAN (CHANG *et al.* 2008), os quais possuem sequências de diversos elementos regulatórios, já descritos em plantas.

3.13 Construção de um vetor para análise do promotor do gene OcXII em plantas de arroz

A sequência de nucleotídeos de 2.138pb da região genômica, correspondendo ao promotor do gene OcXII foi amplificada, utilizando oligonucleotídeos iniciadores específicos (tabela 6). Esta sequência foi introduzida no vetor de tecnologia *Gateway* pENTR e transformada em *E. coli*. O procedimento ocorreu da mesma forma já descrita no item 3.5, porém, o vetor de destino utilizado na reação de recombinação foi o pHGWFS7 (KARIMI, INZÉ, & DEPICKER 2002) (figura 9). Este vetor, por sua vez, possibilita o estudo de promotores de genes fornecendo uma construção específica, onde o promotor do gene de interesse (OcXII) insere-se anteriormente às sequências dos genes repórteres *Egfp* (*Enhanced Green Fluorescent Protein*) e *gus* (β -glucoronidase). Desta forma, no momento em que o promotor do gene-alvo é ativado pela maquinaria celular, os genes repórteres *Egfp* e *gus* podem ser detectados, respectivamente, por microscopia ou por coloração com o substrato X-gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid, cyclohexylammonium salt). A confirmação da construção correta do vetor de expressão gênica deu-se por meio da extração de DNA plasmidial, como já descrito no item 3.5. Após, reações de PCR foram feitas com três combinações de *primers* a fim de confirmar a orientação correta do promotor em relação aos genes repórteres no vetor binário. As combinações de oligonucleotídeos foram as seguintes: 1) PromoOc12F+ PromoOc12R, resultando em um amplicon de 2.138 pb, 2) PromoOc12R+ pHGWFS7 HygF, amplificando 2.400 pb e 3) PromoOc12F+pHGWFS7 EgfpR, também com um fragmento de 2.400 pb (tabela 6).

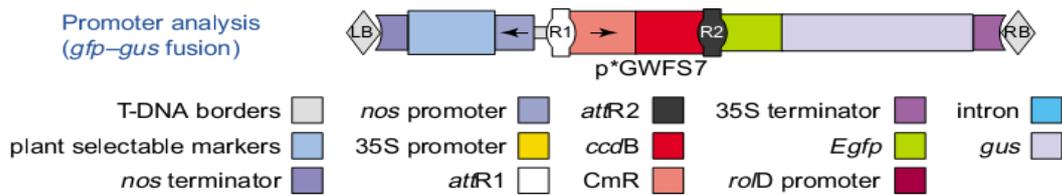


Figura 9. Vetor *Gateway* pHGWS7 utilizado para o estudo de promotores de genes mediante expressão controlada dos genes repórteres *gus* e *Egfp*. Reproduzido de KARIMI *et al.* 2002.

3.14 Transformação de calos embriogênicos de arroz para a expressão dos genes repórteres fusionados ao promotor OcXII

Assim como descrito nos itens 3.6 e 3.7, a construção expressando os genes repórteres *Egfp* e *gus*, de maneira controlada pelo promotor da fitocistatina XII, foi transformada em *A. tumefaciens*, e posteriormente co-cultivadas com calos embriogênicos de arroz, para a transformação genética dos mesmos. Os calos transformados foram igualmente selecionados em meio contendo higromicina (50 mg/L) e mantidos em proliferação por cerca de 40 dias. As linhagens transgênicas, oriundas da seleção por higromicina foram testadas para funcionalidade da construção PromotorOc12+*gus*, de forma que alguns calos de cada linhagem foram selecionados e colocados em tubos de micricentrífuga contendo o reagente X-gluc. Estes tubos, com os calos, permaneceram por 24 horas no escuro a 37 °C. As linhagens com sinal positivo para esta reação, onde os calos expressando a construção ficam azuis, foram colocados em meio de pré-regeneração e posteriormente em regeneração da mesma forma que no item 3.7. Uma vez em meio de regeneração, alguns calos de cada linhagem foram colocados novamente em X-gluc, para confirmação da presença do vetor pHGWS7 contendo o promotor da fitocistatina XII.

3.15 Análise da modulação da expressão gênica da fitocistatina XII em resposta a tratamento com ácido abscísico (ABA)

Tendo identificado alguns elementos regulatórios no promotor da fitocistatina XII, que levem a transcrição da mesma em condições determinadas, o ácido abscísico (ABA) foi escolhido para testes da modulação da expressão do gene OcXII. Primeiramente, 80 sementes de arroz foram germinadas em rolos de papel filtro, umedecidos com água, e após 10 dias transferidas para hidroponia com solução nutritiva (Hoagland & Arnon, 1950), sendo a mesma trocada a cada três dias. Quando as plantas atingiram três semanas de idade foi realizado o tratamento com 100µM de ácido abscísico (ABA, Sigma-Aldrich), onde o mesmo foi colocado diretamente na solução hidropônica. Para a extração de RNA e validação da expressão gênica das fitocistatinas de acordo com o tratamento, foram coletadas plantas nos tempos: 0h (controle), 2h, 4h e 24h após o início do tratamento. Em cada um dos tempos descritos acima, as plantas foram coletadas em quadruplicatas de amostras biológicas, sendo cada uma composta por um *pool* de três plantas. O RNA foi extraído separadamente para tecidos de folha e raiz, bem como a síntese de cDNA e experimentos de RTqPCR foram realizados de forma independente para estes dois tecidos. Os protocolos e procedimentos seguem conforme descrito no item 3.4.

4. RESULTADOS

4.1 Caracterização das fitocistatinas de arroz

A organização estrutural dos doze genes de fitocistatinas de arroz, em exons e introns encontra-se esquematizada na figura 10, sendo possível verificar que todas as doze fitocistatinas compartilham as sequências representadas pelos exons em preto, resultando em sequências proteicas altamente similares. Apenas a OcXII possui um segmento adicional na região carboxi-terminal, representada pelos dois últimos exons (em cinza), com sequência diferente dos demais genes.

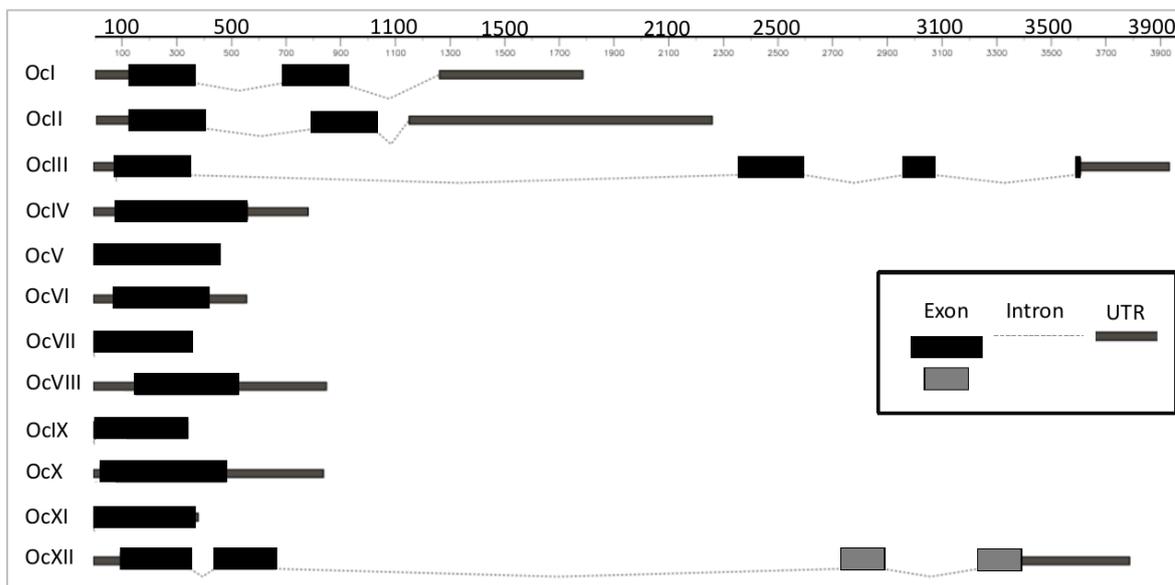
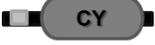


Figura 10. Estrutura gênica das fitocistatinas de arroz. Exons em preto correspondem à codificação de sequências proteicas similares, entre as doze fitocistatinas. Exons em cinza representam a extensão C-terminal, exclusiva da fitocistatina XII.

Os doze genes de fitocistatinas possuem uma sequência proteica bastante conservada, levando-se em consideração os motivos proteicos característicos de fitocistatinas, LARFAV, QxVxG e PW. A busca destes genes em bancos de dados específicos de domínios proteicos já caracterizadas, revelou que todos os doze genes, exceto OcIX, possuem a sequência que caracteriza o domínio cistatina,

conservada (tabela 7). A OcXII, além do domínio cistatina característico, apresenta também um domínio similar a cistatina (*cistatina-like*), sendo assim a única fitocistatina com dois domínios proteicos caracterizados. Todas as fitocistatinas, exceto OcII e OcVI, também possuem sequências de peptídeos sinais característicos (tabela 7).

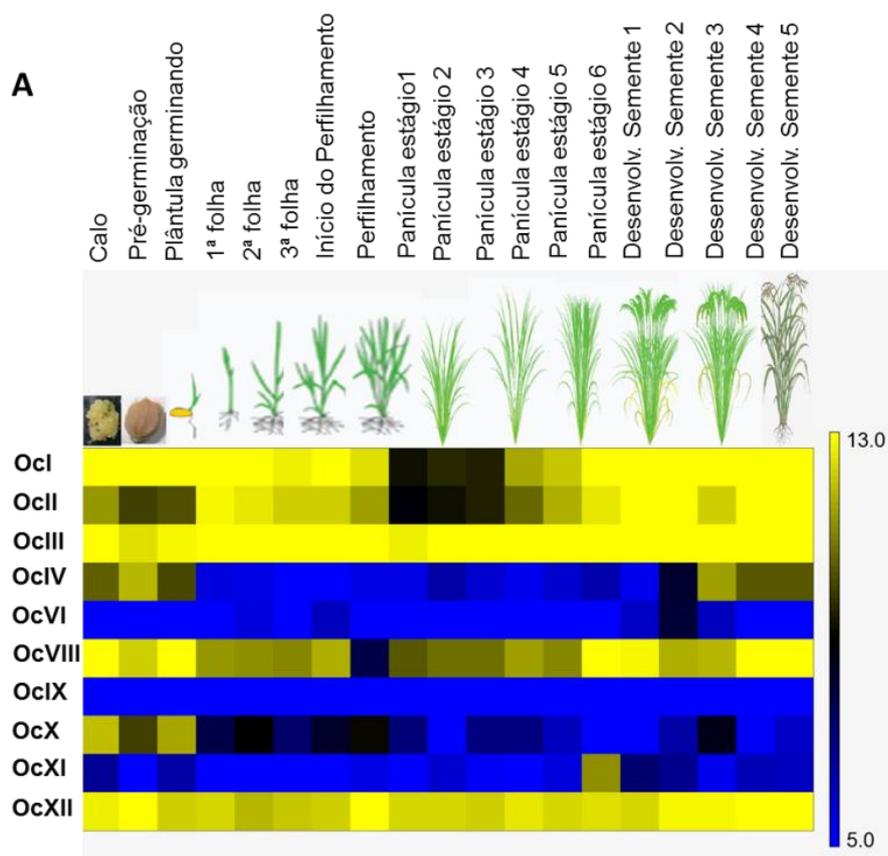
Tabela 7. Domínios proteicos conservados nas doze fitocistatinas de arroz.

Gene	Representação	Domínios	Começo	Fim
Ocl		Peptídeo sinal	1	26
		Cistatina	45	134
OcII		Cistatina	54	146
OcIII		Peptídeo sinal	1	35
		Cistatina	45	133
OcIV		Peptídeo sinal	1	24
		Cistatina	48	151
OcV		Peptídeo sinal	1	25
		Cistatina	47	142
OcVI		Cistatina	18	110
OcVII		Peptídeo sinal	1	24
		Cistatina	25	116
OcVIII		Peptídeo sinal	6	20
		Cistatina	30	120
OcIX		Peptídeo sinal	1	23
OcX		Peptídeo sinal	1	37
		Cistatina	147	232
OcXI		Peptídeo sinal	1	24
		Cistatina	25	118
OcXII		Peptídeo sinal	1	32
		Cistatina	46	136
		Cistatina-like	156	249

Os dados obtidos na análise dos experimentos de microarranjos permitiram realizar uma predição do padrão de expressão dos genes de fitocistatinas em arroz (figura 11 A). Verificou-se que os genes mais expressos, de forma geral, são: OcI, OcII, OcIII, OcVIII e OcXII. Os genes OcIV e OcX apresentaram um nível mais baixo de expressão, sendo OcIV expressa apenas em estádios iniciais de desenvolvimento da planta (calo, pré-germinação e plântula germinando) e, também, nos últimos quatro estádios de desenvolvimento de sementes nas plantas. A OcX tem seu maior nível de expressão nos estádios iniciais de

germinação da plântula. OcXI foi detectada apenas entre o último estágio de panícula e início de desenvolvimento de sementes. E as demais, OcVI e OcIX não tiveram expressão gênica detectada (figura 11 A). Entre todas as fitocistatinas, a OcIII é a que mantém sua expressão elevada e constante em todos os estádios de desenvolvimento avaliados. A Ocl é o segundo gene mais expresso, seguido dos genes OcXII e OcII (figura 11 A).

A expressão gênica da fitocistatina XII pode ser observada durante todo o desenvolvimento da planta, sendo mais intenso nos estádios de pré-germinação do calo, de perfilhamento da planta, e nos últimos estádios de desenvolvimento e maturação das sementes (figura 11 A). De forma similar, o resultado do Genevestigator mostra que a expressão da OcXII é elevada, e praticamente constante durante todo o desenvolvimento da planta. Os menores níveis de expressão da OcXII podem ser observados nos estádios iniciais de desenvolvimento da planta, sofrendo algumas oscilações entre aumento e diminuição da expressão gênica. A expressão de OcXII atinge o maior nível observado no último estágio de desenvolvimento e maturação das sementes (figura 11 B).



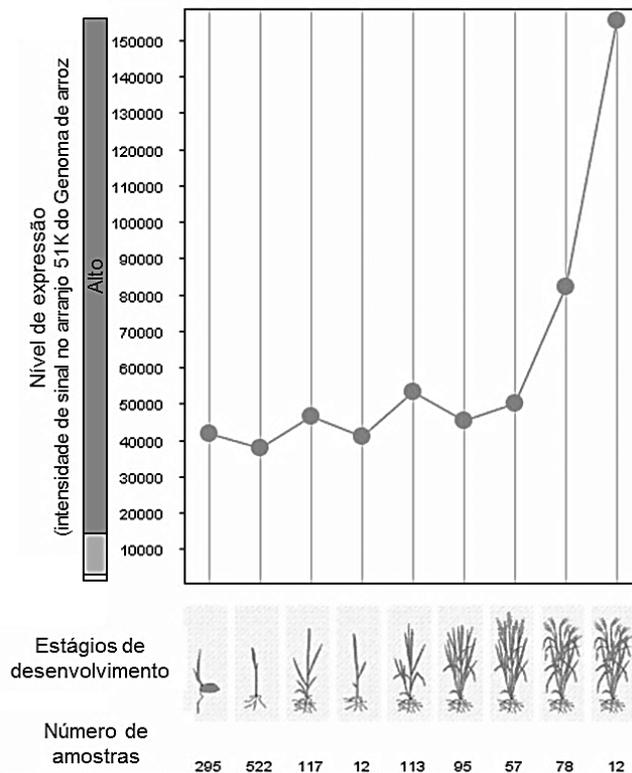


Figura 11. Expressão gênica das fitocistatinas em diferentes estádios de desenvolvimento de plantas de arroz. **A)** *Heatmap* obtido no *Rice Array*. O nível de expressão gênica é demonstrado pela intensidade de cores, indo do azul (menos expresso) para amarelo (mais expresso). Os genes não representados OcV e OcVII, não possuem *probes* para sua detecção nos experimentos avaliados. **B)** Expressão gênica da fitocistatina XII, obtida no software Genevestigator. O gene OcXII foi plotado de acordo com seu nível de expressão, caracterizado como alto (painel esquerdo do gráfico). Abaixo do gráfico estão representados os estádios de desenvolvimento das plantas e o número de amostras analisadas para experimentos em cada estádio.

Uma verificação mais detalhada da expressão destes genes de fitocistatinas foi realizada analisando-se individualmente diferentes tecidos das plantas, onde a expressão de um gene acaba não sendo mascarada pelo *pool* de RNA total e amostragem dos mRNAs mais abundantes. Desta forma, diversos tecidos foram avaliados separadamente em experimentos de microarranjos. Na observação da expressão dos genes que codificam fitocistatinas em uma planta inteira, é

possível ter-se a noção de que os genes mais expressos são, em ordem decrescente, Ocl, OclIII, OcXII, OcII, OcVIII, OcX, OcVI, sendo que OcVI, OcIX e OcXI não apresentam níveis de expressão gênica observáveis (figura 12).

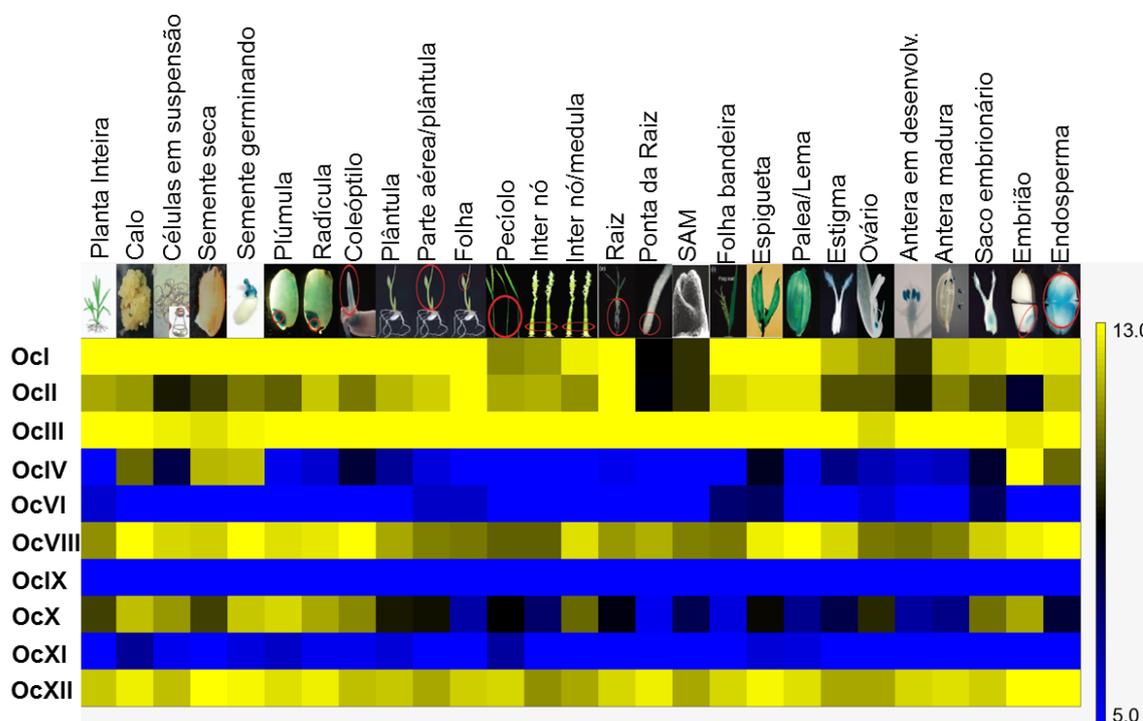


Figura 12. Expressão dos genes codificadores de fitocistatinas em tecidos específicos de plantas de arroz. A escala de cores demonstra nível crescente de expressão gênica de acordo com a intensidade da cor azul para amarelo.

Seguindo a ordem decrescente dos genes mais expressos nos diferentes tecidos, a fitocistatina III é claramente um gene com expressão ubíqua e elevada, tendo uma leve diminuição em células em suspensão, semente seca, semente germinando, ovário, e no embrião. A Ocl, também com níveis de expressão elevados, é diminuída apenas nos tecidos de pecíolo, inter nó, ponta da raiz, estigma, ovário, antera e saco embrionário. Já a OcXII por sua vez, têm sua expressão gênica detectada em todos os tecidos avaliados, variando entre níveis altos e médios. A maior expressão da fitocistatina XII pode ser destacada nos tecidos de calo, semente seca, semente germinando, radícula, ponta da raiz, espigueta, embrião e endosperma (figura12). A OcII possui seus máximos de expressão nos tecidos de folha e de raiz e aOcVIII em células de calos, semente

germinando, coleóptilo, palea/lema e no endosperma. O mRNA da fitocistatina IV encontra-se de forma mais abundante no embrião da semente de arroz, e a OcX em sementes que estão germinando. De maneira geral, as fitocistatinas possuem níveis de expressão peculiares para cada gene que as codifica, em cada tecido diferente, mas sempre mantendo certo nível basal em todos os tecidos analisados, principalmente referindo-se as Ocl, OcII, OcIII, OcVIII e OcXII. Enquanto que as demais apresentam um padrão mais tecido-específico de expressão gênica.

A OcXII, sendo o alvo principal deste estudo, teve sua expressão gênica analisada separadamente em 39 tecidos de plantas de arroz (figura13).

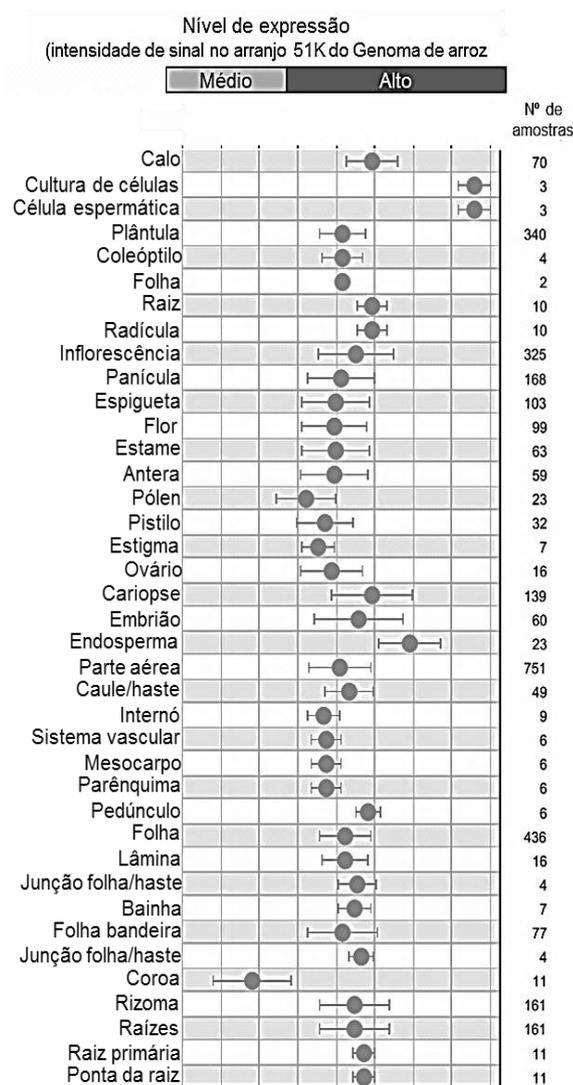


Figura 13. Expressão gênica da fitocistatina XII avaliada em experimentos de microarranjos, disponibilizados pelo software Genevestigator. Cada ponto no gráfico representa a expressão do gene codificador da OcXII no tecido avaliado, com seu respectivo desvio padrão. Os números à direita indicam a quantidade de amostras avaliadas para cada tecido específico. Acima do gráfico o nível de expressão é indicado (médio e alto), sendo que na maioria dos estádios o nível de expressão de OcXII encontra-se na faixa de alta expressão e poucos com nível médio de expressão.

A análise individual da fitocistatina XII em tecidos específicos de plantas de arroz demonstra que, de fato a expressão deste gene é elevada e constante na planta durante seu desenvolvimento. Os maiores níveis de expressão gênica de OcXII podem ser verificados nos tecidos de células em cultura e células espermáticas. O terceiro lugar onde o mRNA deste gene é mais abundante, em relação aos outros tecidos, é no endosperma. Em alguns tecidos a expressão de OcXII foi um pouco menor do que o geral observado. A menor expressão detectada, considerada em nível médio, foi na região da coroa da raiz das plantas. O segundo lugar com menor expressão de OcXII, mas ainda assim em nível alto, foi encontrado em grãos de pólen. Apesar destas pequenas variações encontradas, o perfil de expressão gênica da fitocistatina XII não parece sofrer grandes alterações, sendo mantida quase sempre em um nível alto e constante em todos os tecidos da planta analisados, segundo a escala da figura 13.

Outra abordagem utilizada para a avaliação da expressão gênica das fitocistatinas em arroz foi a análise transcriptômica de dados de sequenciamento do mRNA total, de folhas e raízes de plântulas. Esta análise possibilitou a validação de algumas observações feitas com os experimentos de microarranjos, onde foi possível detectar a presença de transcritos dos seguintes genes: Ocl, OclI, OclII, OclIII, OcVIII, OcX e OcXII, enquanto que para as OcV, OcVI e OcIX um número extremamente baixo de transcritos foi reportado. Para as OcIV, OcVII e OcXI nenhum transcrito foi encontrado (figura 14). A fitocistatina I foi a que apresentou o maior número de *reads* mapeados, sugerindo que sua expressão é relativamente maior do que os demais genes, nos dois tecidos analisados, sendo

um pouco maior em folha do que em raiz. O segundo gene a apresentar maior número de *reads* foi a OcIII, onde há ~36% a mais em raiz do que em folha, sendo que o nível de expressão em folha se equipara ao nível de expressão em raiz da OcII. A expressão em folha de OcII é de aproximadamente 20% da sua expressão em raiz. A fitocistatina XII também possui um nível de expressão gênica detectável nos dois tecidos diferentes, sendo que a expressão em folha é menor, correspondendo a cerca de aproximadamente 70% da expressão em raiz. Por fim, com uma menor abundância de transcritos analisados, as Ocs VIII e X também foram detectadas, com expressão maior em raiz do que em folha. De forma geral, entre as doze fitocistatinas, nove tiveram sua expressão gênica detectada, sendo que seis delas apresentaram um número mais elevado de *reads*, indicando uma maior expressão destes genes. Também, torna-se evidente que dentre as fitocistatinas, a OcI foi a única a ter um número maior de *reads* detectados em folha, enquanto as demais Ocs mostram-se mais abundantes nos tecidos que compõe a raiz das plantas analisadas.

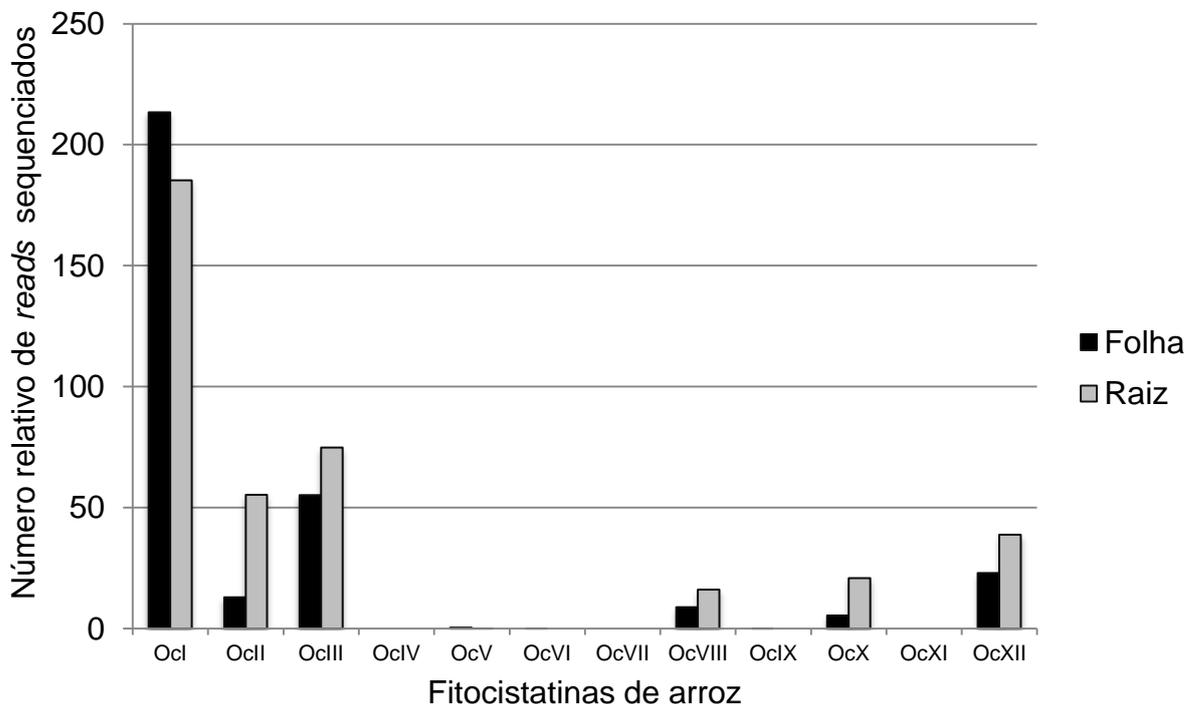
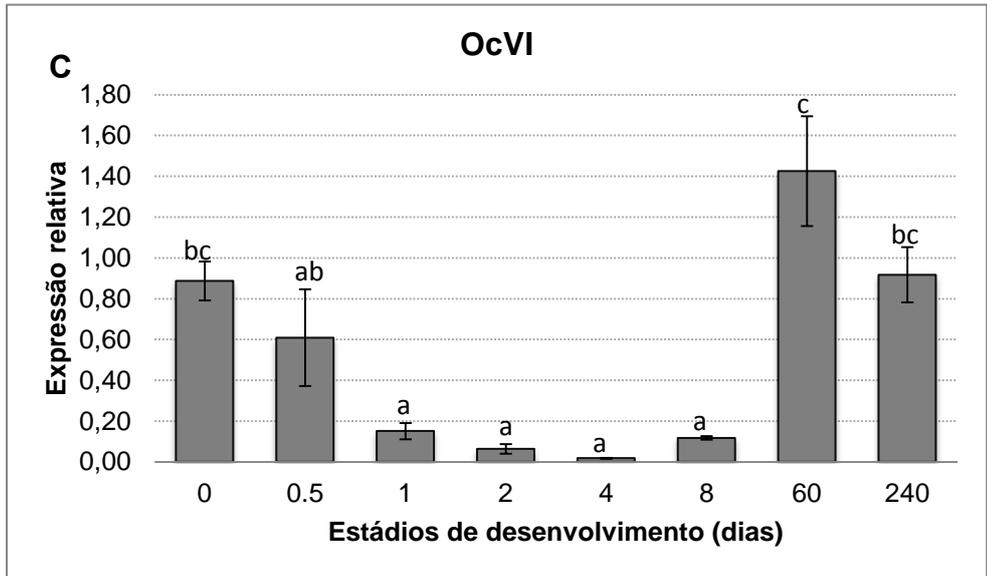
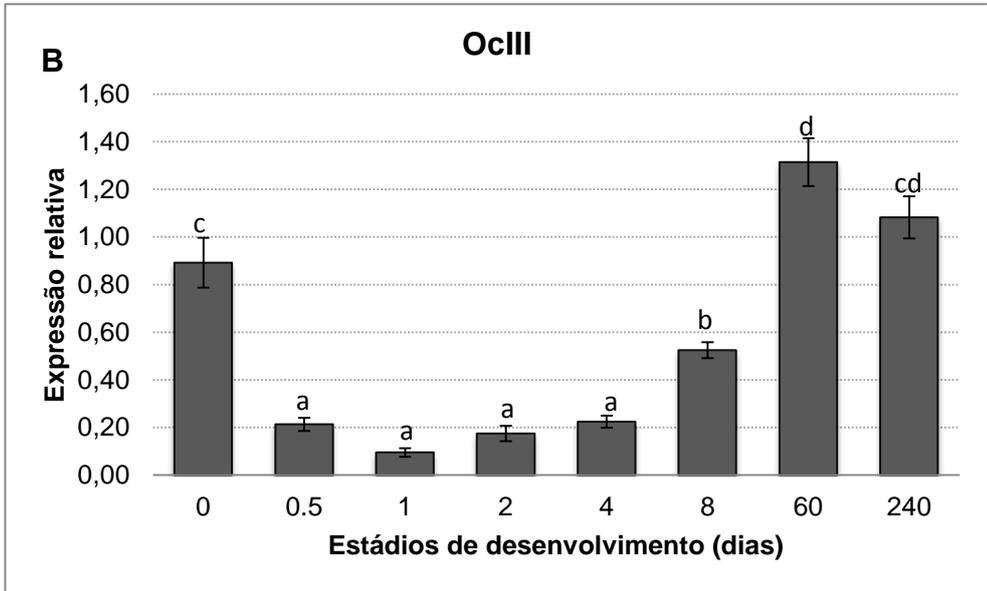
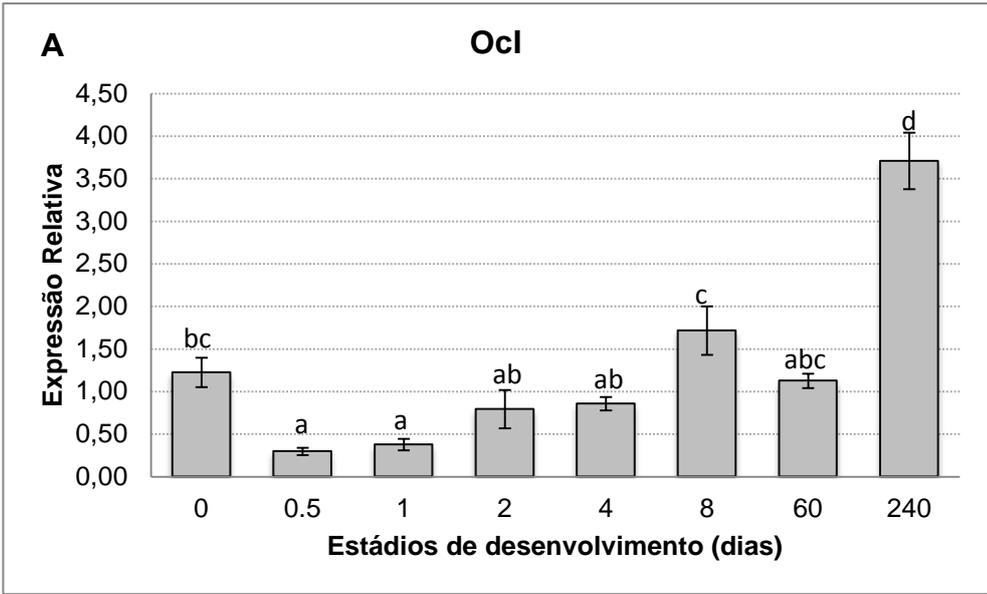


Figura 14. Análise das bibliotecas de RNAseq, demonstrado o número relativo de *reads* encontrado para cada gene de fitocistatina de arroz, em dois tecidos diferentes, folha e raiz, indicando a quantidade de transcritos destes genes.

4.2 Expressão gênica de fitocistatinas durante a germinação e desenvolvimento de plantas de arroz

Os resultados obtidos nos experimentos de microarranjos e RNAseq serviram de base para a elucidação inicial do padrão de expressão dos genes de fitocistatinas em arroz. A fim de confirmar estes resultados obtidos *in silico*, e avaliar o comportamento da expressão destes genes em períodos mais próximos, durante a germinação e crescimento da plântula, experimentos de RTqPCR foram realizados com plantas de arroz em diferentes etapas de desenvolvimento. Partindo da avaliação da quantidade de mRNA em sementes maduras (0 dias), seguindo para estádios de 12 horas de germinação (0.5 dia), 24 horas de germinação (1 dia), 48 horas de germinação (2 dias), 96 horas após a germinação (4 dias), 8 dias, 60 dias e até plantas com 240 dias de idade (figura 7). Neste experimento os únicos genes detectados foram a Ocl, OclIII, OcVI, OcVIII e OcXII, e a expressão de cada um, em cada estágio de desenvolvimento está representada separadamente na figura 15.

A fitocistatina I possui um padrão de expressão gênica onde inicialmente, em sementes maduras (0 dias), o nível de sua expressão é elevado e decai nos próximos dois estádios (0.5 e 1 dia), após este período a expressão volta a ser semelhante ao que era no primeiro estágio, sendo que no último estágio (240 dias) há um pico de expressão gênica de Ocl, de aproximadamente quatro vezes do que nos demais estádios (figura 15A). A fitocistatina III, possui uma expressão elevada em sementes maduras, e, assim como a Ocl, sofre uma queda quando a germinação inicia-se, sendo que para OclIII essa baixa expressão permanece até o período de 4 dias, quando sofre um aumento, dobrando o nível de sua expressão em oito dias. O aumento da expressão de OclIII é significativo no estágio de 60 dias, sendo até 30% maior que no estágio de semente madura, e este nível permanece alto nas plantas entrando em fase de senescência, no último estágio avaliado (figura 15B).



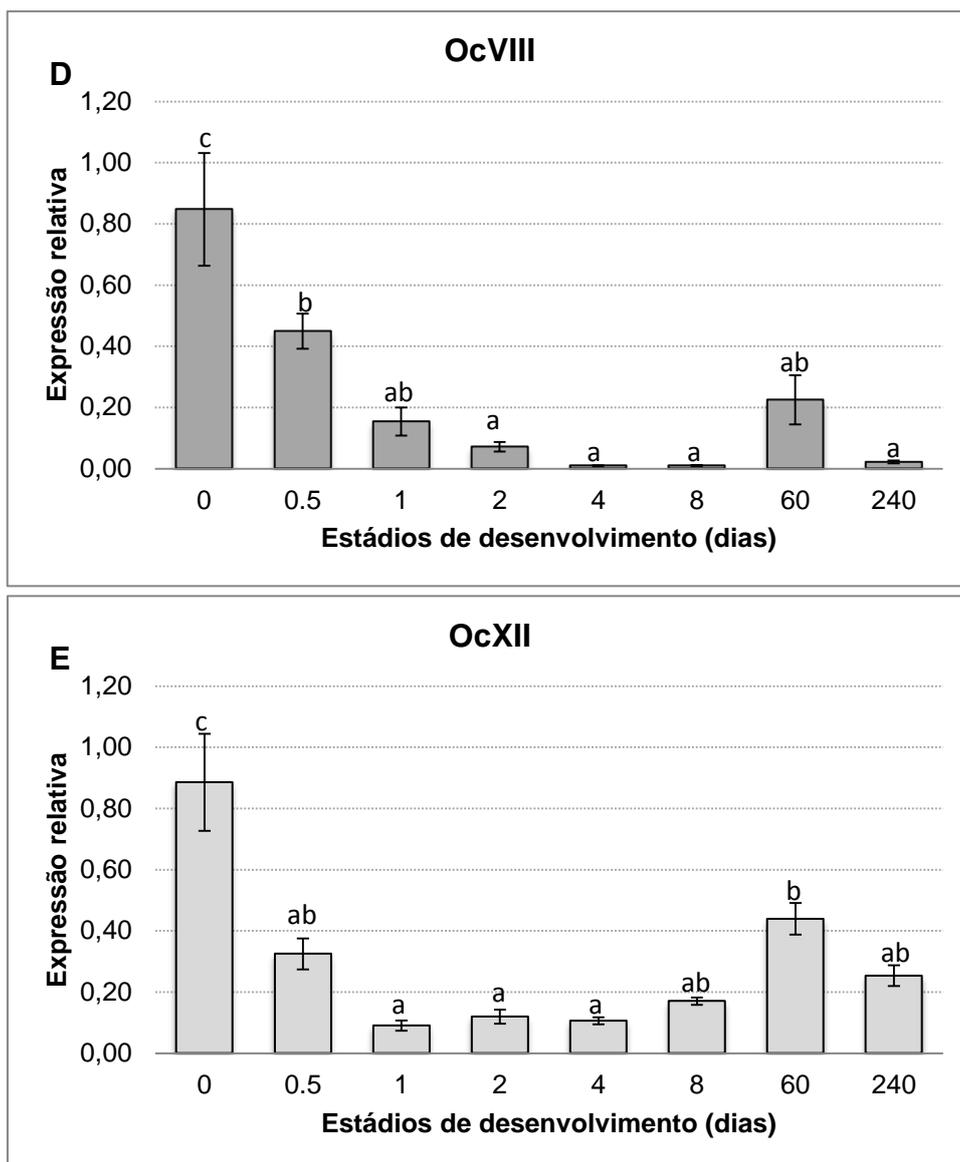


Figura 15. Expressão gênica relativa das fitocistatinas, em diferentes estádios de desenvolvimento de plantas de arroz, avaliados em dias após o início da germinação (0, 0.5, 1, 2, 4, 8, 60 e 240). Apenas os cinco genes representados tiveram sua expressão detectada por RTqPCR. **(A)** Padrão de expressão gênica da fitocistatina I. **(B)** Expressão do gene codificador da fitocistatina III. **(C)** Expressão da fitocistatina VI. **(D)** Expressão da fitocistatina VIII. **(E)** Padrão de expressão gênica da fitocistatina XII. Grupos da análise estatística referem-se ao teste de Tuckey, com significância <0.05 .

Os resultados obtidos para a expressão gênica da fitocistatina VI (figura 15 C) sugerem que a expressão desse gene permanece elevada em sementes secas, decaindo com a germinação da planta até o estágio de 8 dias. O nível volta a ser mais elevado, similar ao inicial em sementes, nas plantas mais velhas, com 60 e 240 dias. Para a fitocistatina VIII, o padrão de expressão gênica observado revela que, da mesma forma que para as outras Ocs, a OcVIII tem expressão gênica elevada em estágio de semente, neste caso, sendo a mais elevada entre todos os demais estádios. Este nível de expressão decai gradativamente com a germinação da planta, e permanece em um nível basal ao longo do desenvolvimento, considerando-se os estádios avaliados neste experimento (figura 15 D).

A fitocistatina XII, por sua vez, também mantém os padrões de expressão gênica similares aos já observados para os demais genes (figura 15 E). O maior nível de expressão deste gene encontra-se no estágio de sementes madura (0 dias), decaindo para mais da metade deste valor após 12 horas de semente em germinação (0.5 dias). Após essa queda brusca na expressão do gene, posterior ao período de apenas 12 horas, a quantidade de mRNA da OcXII permanece neste nível mais baixo durante todos os demais estádios de desenvolvimento da planta analisados, até a idade de 240 dias da planta (8 meses).

A fim de tornar mais clara a comparação da expressão destes genes de fitocistatinas durante o desenvolvimento das plantas, os resultados obtidos, apresentados na figura 15, foram plotados em um novo gráfico, onde a expressão dos genes pode ser avaliada de forma comparativa e relativa ao gene OcVIII no dia 0 (figura 16). O que pode ser percebido, de maneira geral, são os genes mais expressos Ocl, OcIII e OcXII, enquanto OcVI e OcVIII possuem níveis bem menores de expressão gênica. Os demais genes OcII, OcIV, OcV, OcVII, OcIX, OcX e OcXI, não tiveram expressão gênica detectável por esta metodologia. Também, permanece claramente visível que estes genes de fitocistatinas possuem um nível elevado de expressão gênica nas sementes de arroz, o qual diminui quando as mesmas começam a germinar e, principalmente para as Ocl, OcIII e OcVI a expressão gênica volta a aumentar nas plantas mais velhas (8 dias, 60 dias, 240 dias). A fitocistatina I foi a única a apresentar, de forma

significativa, uma elevação da expressão gênica em folhas que começam a desencadear o processo de senescência, podendo ter alguma função relacionada com esta via. Em suma, a fitocistatina XII apresentou-se como um dos três genes de fitocistatinas mais expressos durante todo o desenvolvimento das plantas analisadas. Entretanto, a mesma destaca-se nas sementes maduras de arroz, e evidentemente é necessária uma diminuição da quantidade de seu mRNA para que a planta consiga germinar, o qual não mais retorna para níveis tão elevados quanto o detectado no estágio de semente madura (0 dias).

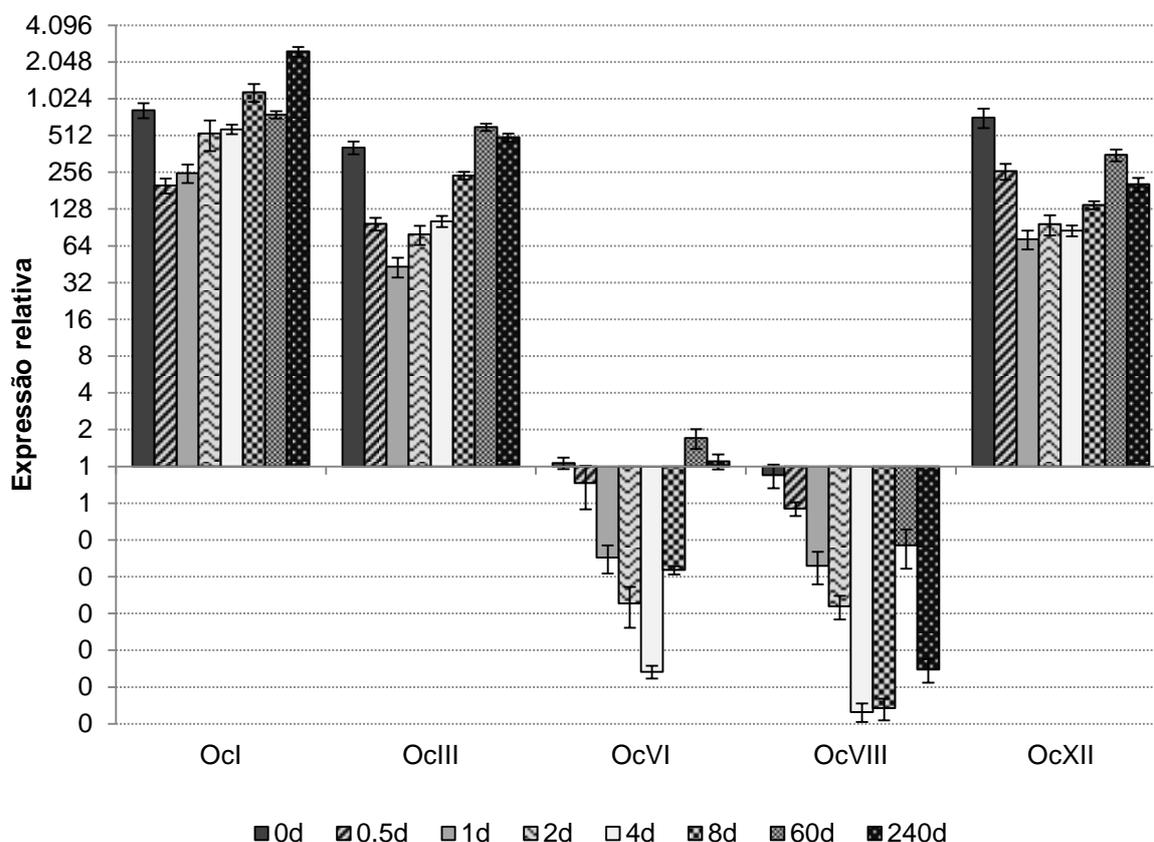


Figura 16. Expressão gênica relativa das fitocistatinas de arroz, durante diferentes estádios de desenvolvimento. Para visualização comparativa da expressão em escala, de cada gene, o cálculo de $\Delta\Delta Ct$ foi realizado para todos os genes em relação ao estágio de semente da fitocistatina VIII. A escala apresentada encontra-se em \log_2 , e a análise estatística refere-se à mesma demonstrada nos gráficos anteriores (figura 15).

4.3 Considerações evolutivas sobre as fitocistatinas

As fitocistatinas vêm sendo estudadas ao longo dos últimos 30 anos e diversas relações evolutivas têm sido propostas para estas sequências. Com o intuito de auxiliar na elucidação das origens e diversificações destes genes, principalmente na espécie de *Oryza sativa*, uma busca pelas sequências homólogas às fitocistatinas de arroz, descritas por Martinez *et al.* 2007, foi realizada no banco de dados Phytozome, em todos os genomas atuais que possuem sequências disponíveis. Considerando-se as sequências genômicas de 25 espécies diferentes, foram encontradas 204 genes de fitocistatinas, sendo que 36 destes genes são ortólogos à OcXII, apresentando uma extensão C-terminal, com o motivo proteico SNSL (tabela 8).

Tabela 8. Sequências de genes homólogos a fitocistatinas encontrados para 25 espécies de plantas diferentes, no banco de dados Phytozome. As sequências com seus respectivos códigos em negrito e marcadas com asterisco representam as fitocistatinas com extensão C-terminal, para cada espécie encontrada. Apenas os genes cujos domínios cistatinas eram conservados foram considerados.

Espécie	Código Phytozome	Acrônimo	Nº total de Fitocistatinas
<i>Oryza sativa</i>	LOC_Os01g58890	Osa_OcI	11
	LOC_Os05g41460	Osa_OcII	
	LOC_Os05g33880	Osa_OcIII	
	LOC_Os01g68660	Osa_OcIV	
	LOC_Os01g68670	Osa_OcV	
	LOC_Os03g11180	Osa_OcVI	
	LOC_Os03g11170	Osa_OcVII	
	LOC_Os03g31510	Osa_OcVIII	
	LOC_Os09g08100	Osa_OcX	
	LOC_Os04g28250	Osa_OcXI	
	LOC_Os01g16430	Osa_OcXII*	
	<i>Arabidopsis thaliana</i>	AT5G12140	
AT2G31980		AtCYS2	
AT2G40880		AtCYS3	
AT4G16500		AtCYS4	
AT5G47550		AtCYS5	
AT3G12490		AtCYS6*	
AT5G05110		AtCYS7*	
<i>Manihot esculenta</i>	cassava4.1_014862m	Mes014862*	8
	cassava4_1_018473m	Mes018473m	

	cassava4_1_019663m	Mes019663m	
	cassava4_1_019672m	Mes019672m	
	cassava4_1_019939m	Mes019939m	
	cassava4.1_023762m	Mes023762m*	
	cassava4_1_027574m	Mes027574m	
	cassava4_1_027619m	Mes027619m	
<i>Ricinus communis</i>	27574.m000225	Rco27574.m000225	5
	29686.m000868	Rco29686.m000868*	
	29785.m000949	Rco29785.m000949	
	29940.m000399	Rco29940.m000399*	
	30183.m001284	Rco30183.m001284	
<i>Linum ussitatissimum</i>	Lus10008702	Lus10008702*	4
	Lus10026117	Lus10026117*	
	Lus10026779	Lus10026779	
	Lus10041257	Lus10041257	
<i>Populus trichocarpa</i>	POPTR_0001s07020.1	Ptr0001s07020*	9
	POPTR_0001s23300.1	Ptr0001s23300	
	POPTR_0002s23690.1	Ptr0002s23690	
	POPTR_0003s19140.1	Ptr0003s19140*	
	POPTR_0006s01560.1	Ptr0006s01560	
	POPTR_0006s03170.1	Ptr0006s03170*	
	POPTR_0009s02740.1	Ptr0009s02740	
	POPTR_0014s14680.1	Ptr0014s14680	
	POPTR_0016s03070.1	Ptr0016s03070*	
<i>Medicago truncatula</i>	Medtr2g030120.1	Mtr2g030120	6
	Medtr2g092440.1	Mtr2g092440	
	Medtr3g058780.1	Mtr3g058780	
	Medtr3g104370.1	Mtr3g104370	
	Medtr5g095890.1	Mtr5g095890	
	Medtr8g044470.1	Mtr8g044470*	
<i>Phaseolus vulgaris</i>	Phvulv091000112m	Pvu091000112	10
	Phvulv091000114m	Pvu091000114	
	Phvulv091000119m	Pvu091000119	
	Phvulv091003997m	Pvu091003997	
	Phvulv091012647m	Pvu091012647	
	Phvulv091014625m	Pvu091014625	
	Phvulv091015691m	Pvu091015691	
	Phvulv091019782m	Pvu091019782*	
	Phvulv091019996m	Pvu091019996	
	Phvulv091019997m	Pvu091019997	
<i>Glycine max</i>	Glyma04g10360.1	Gma04g10360	16
	Glyma05g28250.1	Gma05g28250	
	Glyma07g39590.1	Gma07g39590	
	Glyma09g01360.1	Gma09g01360	
	Glyma09g16960.1	Gma09g16960	
	Glyma11g06850.1	Gma11g06850	
	Glyma11g36780.1	Gma11g36780	
	Glyma13g04250.1	Gma13g04250	
	Glyma13g25870.1	Gma13g25870*	
	Glyma14g04250.1	Gma14g04250	

	Glyma14g04260.1	Gma14g04260	
	Glyma15g36180.1	Gma15g36180*	
	Glyma18g00690.1	Gma18g00690	
	Glyma18g12240.1	Gma18g12240	
	Glyma19g39400.1	Gma19g39400	
	Glyma20g08800.1	Gma20g08800	
<i>Cucumis sativus</i>	Cucsa.009310.1	Csa009310	13
	Cucsa.048060.1	Csa048060	
	Cucsa.076380.1	Csa076380*	
	Cucsa.083990.1	Csa083990	
	Cucsa.100800.1	Csa100800	
	Cucsa.100820.1	Csa100820	
	Cucsa.100830.1	Csa100830	
	Cucsa.180410.1	Csa180410	
	Cucsa.180420.1	Csa180420	
	Cucsa.142550.1	Csa142550	
	Cucsa.181470.1	Csa181470	
	Cucsa.268290.1	Csa268290	
	Cucsa.365610.1	Csa365610	
<i>Prunus persica</i>	ppa010412m	Ppe010412*	7
	ppa012846m	Ppe012846	
	ppa013391m	Ppe013391	
	ppa013935m	Ppe013935	
	ppa021911m	Ppe021911	
	ppa022464m	Ppe022464	
	ppa023297m	Ppe023297	
<i>Capsella rubella</i>	Carubv10001694m	Cru10001694*	8
	Carubv10002326m	Cru10002326	
	Carubv10006049m	Cru10006049	
	Carubv10014520m	Cru10014520*	
	Carubv10024323m	Cru10024323	
	Carubv10024940m	Cru10024940	
	Carubv10025408m	Cru10025408	
	Carubv10027345m	Cru10027345	
<i>Brassica rapa</i>	Bra000210	Bra000210	22
	Bra001878	Bra001878	
	Bra001880	Bra001880	
	Bra005556	Bra005556	
	Bra005839	Bra005839*	
	Bra008912	Bra008912	
	Bra012712	Bra012712	
	Bra016965	Bra016965	
	Bra017357	Bra017357	
	Bra017358	Bra017358	
	Bra017478	Bra017478	
	Bra022141	Bra022141	
	Bra023365	Bra023365	
	Bra024935	Bra024935	
	Bra034755	Bra034755*	
	Bra038474	Bra038474	

	Bra038475	Bra038475	
	Bra038477	Bra038477	
	Bra038478	Bra038478	
	Bra038725	Bra038725*	
	Bra040651	Bra040651	
<i>Carica papaya</i>	evm.model.supercontig_34.196	Cpa34.196	8
	evm.model.supercontig_34.197	Cpa34.197	
	evm.model.supercontig_34.199	Cpa34.199	
	evm.model.supercontig_34.200	Cpa34.200	
	evm.model.supercontig_34.201	Cpa34.201	
	evm.model.supercontig_75.103	Cpa75.103	
	evm.model.supercontig_157.8	Cpa157.8	
	evm.model.supercontig_228.9	Cpa228.9*	
<i>Citrus sinensis</i>	orange1.1g027502m	Csi1g027502*	6
	orange1.1g033492m	Csi1g033492	
	orange1.1g034261m	Csi1g034261	
	orange1.1g037584m	Csi1g037584	
	orange1.1g042638m	Csi1g042638	
	orange1.1g048720m	Csi1g048720	
<i>Eucalyptus grandis</i>	Eucgr.A02184.1	EgrA02184	11
	Eucgr.A02187.1	EgrA02187	
	Eucgr.A02189.1	EgrA02189	
	Eucgr.B03762.1	Egr.B03762	
	Eucgr.E03919.1	EgrE03919	
	Eucgr.E03920.1	EgrE03920	
	Eucgr.E03921.1	EgrE03921	
	Eucgr.E03923.1	EgrE03923	
	Eucgr.J00085.1	EgrJ00085*	
	Eucgr.J00666.1	EgrJ00666	
	Eucgr.L00642.1	EgrL00642	
<i>Vitis vinifera</i>	GSVIVT01003320001	Vvi01003320001	3
	GSVIVT01025801001	Vvi01025801001*	
	GSVIVT01037389001	Vvi01037389001	
<i>Aquilegia coerulea</i>	Aquca_009_00061.1	Aco009_00061*	8
	Aquca_028_00213.1	Aco028_00213	
	Aquca_075_00041.1	Aco075_00041	
	Aquca_036_00059.1	Aco036_00059	
	Aquca_028_00208.1	Aco028_00208	
	Aquca_028_00209.1	Aco028_00209	
	Aquca_003_00453.1	Aco003_00453	
	Aquca_007_00495.1	Aco007_00495	
<i>Shorgum bicolor</i>	Sb03g010730.1	Sbi03g010730*	11
	Sb03g037370.1	Sbi03g037370	
	Sb09g024230.1	Sbi09g024230	
	Sb09g020170.1	Sbi09g020170	
	Sb09g024240.1	Sbi09g024240	
	Sb01g043320.1	Sbi01g043320	
	Sb03g043680.1	Sbi03g043680	
	Sb01g031450.1	Sbi01g031450	
	Sb01g031440.1	Sbi01g031440	

	Sb01g031460.1	Sbi01g031460	
	Sb01g031470.1	Sbi01g031470	
<i>Zea mays</i>	GRMZM2G013461_T01	Zma2G013461*	12
	GRMZM2G012160_T01	Zma2G012160	
	GRMZM2G438551_T01	Zma2G438551	
	GRMZM2G440968_T01	Zma2G440968	
	GRMZM2G024264_T01	Zma2G024264	
	GRMZM2G133620_T01	Zma2G133620	
	GRMZM2G401328_T01	Zma2G401328	
	GRMZM2G401374_T01	Zma2G401374	
	GRMZM2G359581_T01	Zma2G359581	
	GRMZM2G030717_T01	Zma2G030717	
	GRMZM2G312061_T01	Zma2G312061	
	GRMZM2G148925_T01	Zma2G148925	
<i>Setaria italica</i>	Si002179m	Sit002179*	10
	Si003312m	Sit003312	
	Si024721m	Sit024721	
	Si024132m	Sit024132	
	Si039899m	Sit039899	
	Si038060m	Sit038060	
	Si037934m	Sit037934	
	Si027571m	Sit027571	
	Si028385m	Sit028385	
	Si028295m	Sit028295	
<i>Selaginella moellendorffii</i>	85303	Smo85303*	3
	418098	Smo418098*	
	19455	Smo19455	
<i>Physcomitrella patens</i>	Pp1s238_59V6.1	Ppa1s238_59V6*	4
	Pp1s212_8V6.1	Ppa1s212_8V6*	
	Pp1s142_12V6.1	Ppa1s142_12V6	
	Pp1s238_60V6.1	Ppa1s238_60V6	
<i>Chlamidomonas reinhardtii</i>	Cre04.g225650.t1.1	Cre04.g225650	1
<i>Volvox carteri</i>	Vocar20011820m	Vca20011820	1
Número total de fitocistatinas: 204		Total de fitocistatinas com extensão C-terminal: 36	

A análise inicial foi realizada de forma a verificar a posição filogenética das fitocistatinas de arroz em grupos de sequências homólogas de plantas. O alinhamento e a análise bayesiana englobaram todas as fitocistatinas encontradas nos genomas das plantas consideradas (tabela 8), e resultaram em uma filogenia onde podemos verificar a existência de quatro grandes grupos bem suportados (figura 17). O primeiro grupo (1, preto) é composto de um único subgrupo formado por duas espécies, *Chlamidomonas reinhardtii* e *Volvox carteri*, as quais permanecem de forma basal às sequências de fitocistatinas das demais espécies analisadas. No segundo grupo formado (azul), podem ser encontradas

sequências tanto de monocotiledôneas quanto de dicotiledôneas, as quais não possuem uma clara separação, em função do baixo suporte, mas que podem ser visualizadas no subgrupo 2.1, em azul escuro (dicotiledôneas), e no subgrupo 2.2, em azul claro (monocotiledôneas, exceto *C. sativus*). Neste grupo 2, subgrupo 2.2 podem ser encontradas as sequências das fitocistatinas VI, VII, VIII e XI. O grupo 3, em verde claro, é um grupo menor, composto de apenas algumas sequências de monocotiledôneas e dicotiledôneas, incluindo as fitocistatinas IV, V e X. O grupo 4, comporta um grande número de sequências de fitocistatinas, as quais podem ser separadas em cinco subgrupos principais. No subgrupo 4.1, em lilás, podem ser identificadas as sequências que se dispõem de forma basal ao grupo 4, sendo elas encontradas nas duas espécies, *S. moellendorffii* e *P. patens*. O subgrupo 4.2, verde escuro, é formado pelas sequências das monocotiledôneas de *O. sativa*, *Z. mays*, *S. bicolor*, e *S. italica*, além de algumas sequências de dicotiledôneas como *A. thaliana*, *B. rapa*, *C. rubella* e *M. truncatula*, sendo o grupo onde encontram-se as fitocistatinas I e II de arroz. Identificado como subgrupo 4.3, em laranja, pode ser destacado o agrupamento contendo somente sequências de plantas dicotiledôneas, entre as quais *G. max*, *E. grandis*, *C. sinensis*, *M. truncatula*, *P. vulgaris*, *L. ussitatissimum*, *P. trichocarpa*, *R. comunis*, *M. esculenta*, *V. vinífera* e *C. papaya*. No subgrupo 4.4, em cor roxa, são encontradas as sequências de monocotiledôneas homólogas à OclIII de *Z. mays*, *S. bicolor* e *S. italica*, além de uma sequência de *P. persica*, uma dicotiledônea. O último subgrupo identificado é o 4.5, destacado em dois tons de vermelho, o qual reúne todas as sequências de fitocistatinas com extensão C-terminal.

Este subgrupo 4.5 pode ainda ser dividido em dois grupos menores, para melhor visualização das relações entre as sequências. O grupo 4.5.1, vermelho claro, é composto somente pelas sequências carboxi-extendidas de mono e dicotiledôneas, incluindo a fitocistatina XII. Entretanto, o grupo 4.5.2, vermelho escuro, é formado apenas por sequências de dicotiledôneas, mas possui mais algumas sequências além das carboxi-extendidas, cuja sequência N-terminal é mais relacionada com este agrupamento.

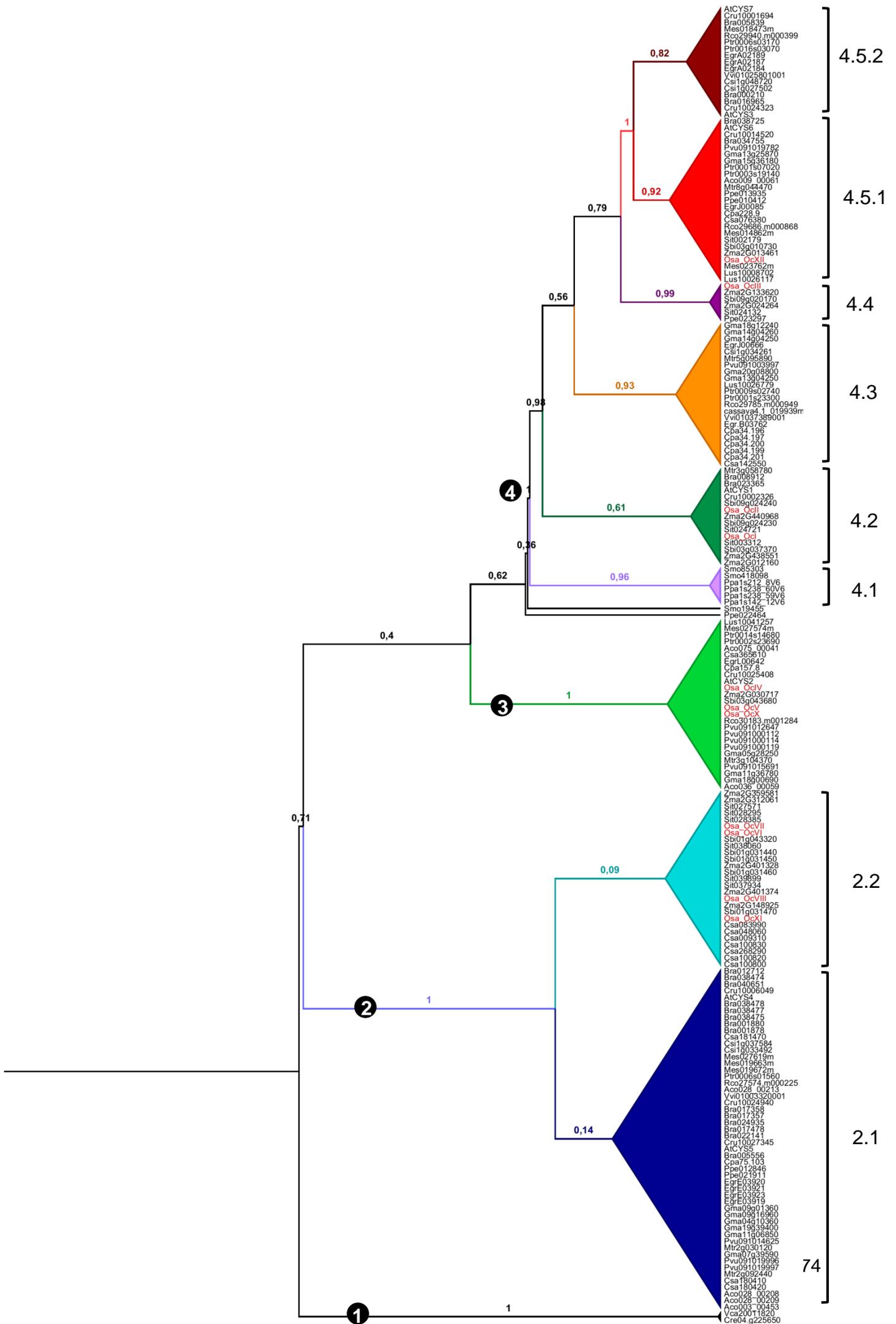


Figura 17. Análise filogenética das sequências proteicas de fitocistatinas. Os principais grupos da análise estão destacados pelas diferentes cores, e especificados pelos números nos círculos pretos em cada ramo. Os subgrupos dentro dos grupos 2 e 4 estão especificados pelas chaves numeradas, a direita de cada agrupamento. As sequências que representam as fitocistatinas de arroz estão destacadas em vermelho.

A análise filogenética das fitocistatinas pode ser visualizada mais detalhadamente na figura 18, demonstrando as relações mais estreitas entre os genes analisados, dentro de cada um dos quatro grupos maiores. Assim, é possível observar a permanência das sequências de *Chlamidomonas reinhardtii* e *Volvox carteri*, no grupo 1, de forma externa a todas as demais sequências analisadas. No grupo 2, bem suportado, não existe uma separação clara entre subgrupos de sequências, pelo baixo suporte dos ramos, apesar de haver uma aparente tendência de separação entre monocotiledôneas (grupo 2.1, exceto pela presença da espécie *C. sativus*) e dicotiledôneas (grupo 2.2). No grupo 2.2 é onde estão localizadas as fitocistatinas de arroz, sendo que OcVI e OcVII constituem um grupo de provável duplicação recente e as Ocs VIII e XI denotam outro evento de duplicação mais recente. A formação do grupo 3 é bem suportada (probabilidade posterior igual a 1), e neste grupo encontram-se diversas sequências de monocotiledôneas e dicotiledôneas, podendo ser destacada a presença das sequências de fitocistatinas de arroz, OclV, OcV e OcX, sendo as duas últimas, também resultantes de um evento de duplicação recente. O grupo 4 apresenta como principal ancestral o agrupamento identificado como 4.1, contendo as sequências de *S. moellendorffii* e *Physcomitrella patens*, agrupadas em clados de acordo com sua própria espécie. Após a separação do grupo 4.1, os demais grupos não apresentam uma clara divergência evolutiva entre si, tendo em vista o baixo suporte obtido na análise e, portanto, formando um agrupamento conciso, retratado pela identificação do grupo 4, como uma única entidade. Entretanto, fica evidenciada a formação de subgrupos de sequências dentro deste grupo 4. As sequências mais relacionadas com as fitocistatinas I e II de arroz formam o subgrupo 4.2, composto de um subclado externo de sequências de

dicotiledôneas (*A. thaliana*, *C. rubella*, *B. rapa* e *M. truncatula*) e outro clado de monocotiledôneas (*O. sativa*, *Z. mays*, *S. bicolor* e *S. italica*). Neste subclado de monocotiledôneas pode ser observada uma divisão de acordo com as sequências homólogas a Ocl, e com as sequências homólogas a OcII, sendo que, pelo menos uma sequência representante de cada espécie de monocotiledôneas encontra-se em cada clado homólogo a Ocl ou OcII. No agrupamento 4.3 podem ser visualizadas sequências exclusivas de dicotiledôneas enquanto que, no grupo 4.4 são encontradas agrupadas as sequências de monocotiledôneas (exceto uma sequência de *P. persica*), sendo estas mais relacionadas com a fitocistatina III de arroz.

O agrupamento 4.5, bem suportado (probabilidade posterior =1), é onde podem ser encontradas todas as sequências das fitocistatinas carboxi-extendidas de plantas, incluindo a fitocistatina XII, foco deste trabalho. Este resultado, remonta à formação de um grupo maior com sequências de monocotiledôneas e dicotiledôneas, incluindo a OcXII (4.5.1) e outro grupo onde estão presentes apenas sequências de dicotiledôneas (4.5.2). Nesta análise, considerando-se apenas a região amino-terminal das fitocistatinas, existem algumas peculiaridades em relação à formação do grupo 4.5.2, onde algumas outras sequências de dicotiledôneas, que não possuem extensão C-terminal, acabaram sendo incluídas neste último grupo. Assim, algumas fitocistatinas não carboxi-extendidas das espécies de *E. grandis* e *M. esculenta*, além de outras cópias, sem extensão C-terminal, das espécies de *A. thaliana*, *B. rapa*, *C. rubella*, e *C. sinensis*, foram clusterizadas no subgrupo 4.5.2, juntamente com as sequências com extensão C-terminal das espécies de *A. thaliana*, *B. rapa*, *C. rubella*, *C. sinensis*, *V. vinifera*, *R. communis* e *P. trichocarpa*.

Outro padrão interessante que pode ser observado nesta análise filogenética (figura 18) é a presença de vários genes oriundos de eventos de duplicação recente dentro de cada espécie analisada. Isto pode ser observado entre genes da mesma espécie que apresentam os ramos muito próximos e clusterizados em um mesmo agrupamento filogenético.

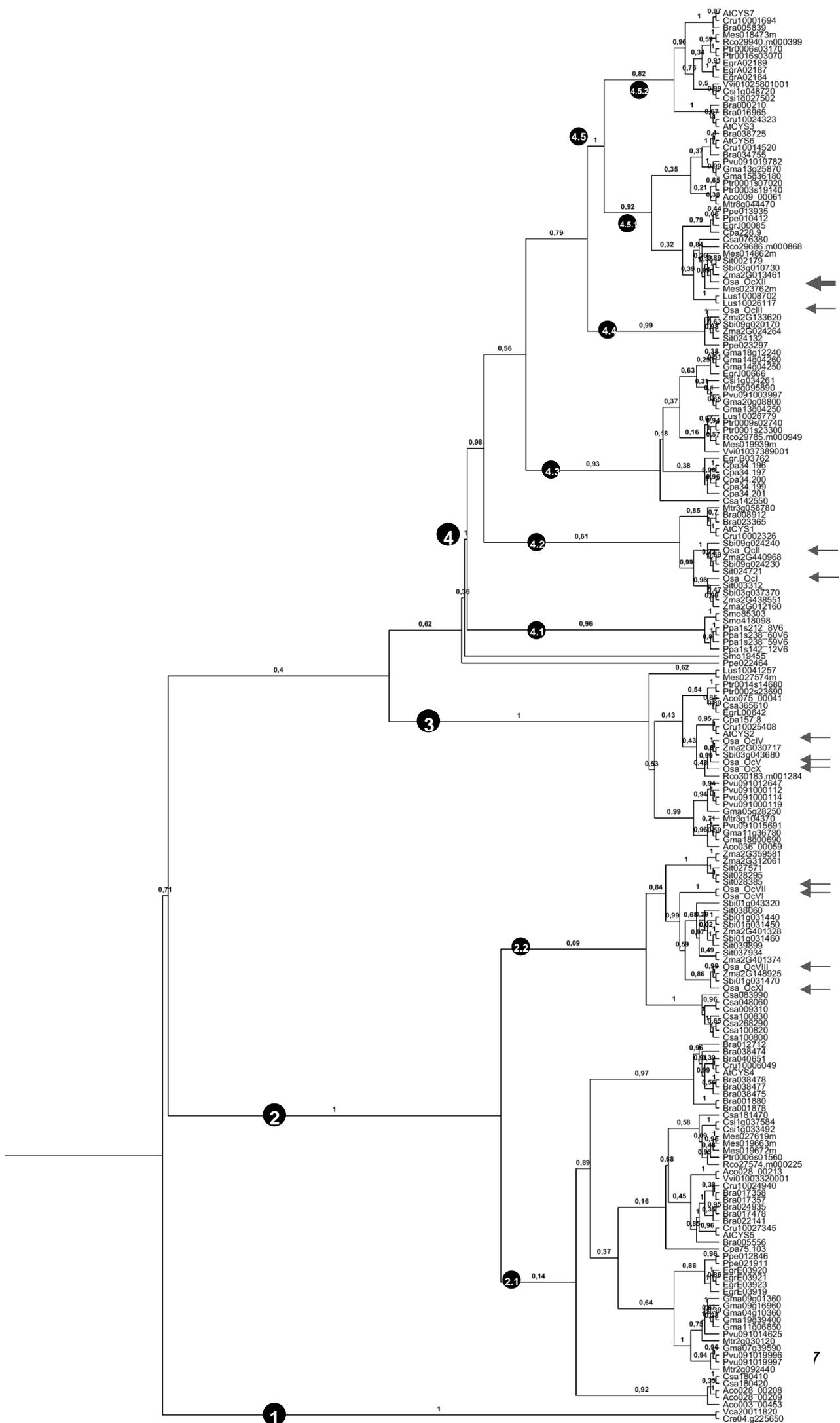


Figura 18. Análise filogenética envolvendo a região N-terminal de todas as fitocistatinas encontradas nas 25 espécies de plantas consideradas. Os valores acima dos ramos referem-se ao valor de probabilidade posterior obtido. Os quatro grupos principais, e respectivos subgrupos encontram-se destacados pelos números nos círculos pretos. As sequências das fitocistatinas de arroz encontram-se destacadas pelas setas a direita dos táxons, sendo a seta mais grossa correspondente à OcXII. Para inclusão das fitocistatinas com extensão C-terminal nesta análise global, somente a região N-terminal destas sequências de proteínas foram consideradas.

A fim de clarificar a resolução das relações filogenéticas das fitocistatinas carboxi-extendidas, uma segunda análise foi realizada apenas com os 36 genes de fitocistatinas que possuíam a extensão C-terminal em sua sequência (tabela 8). O resultado obtido sugere claramente a existência de três principais grupos bem suportados (probabilidade posterior >0.95) (figura 19). Com este resultado é possível verificar que sequências de *S. moellendorffii* e *P. patens* formam o primeiro dos grupos obtidos (1), de forma que estas sequências possuem uma relação de ancestralidade com as demais sequências de plantas vasculares, sendo estas monocotiledôneas e dicotiledôneas. No segundo grupo da análise (2), encontram-se todas as sequências carboxi-extendidas de monocotiledôneas, e grande parte das sequências de dicotiledôneas. E, o terceiro grupo (3) caracteriza-se pela presença de sequências encontradas apenas em espécies do grupo das dicotiledôneas.

Observando mais detalhadamente a análise bayesiana obtida (figura 19), é possível salientar que nas espécies *S. moellendorffii* e *P. patens*, as sequências duplicadas de cada espécie mantêm-se agrupadas, primeiramente com sua própria espécie e em seguida formam o grupo 1, mais basal em relação as demais plantas. O grupo 2 é composto de sequências tanto de monocotiledôneas quanto de dicotiledôneas, sendo que alguns subgrupos podem ser destacados. O subgrupo 2.1 é composto somente por sequências de monocotiledôneas, entre elas o arroz, milho, sorgo e *S. itálica*. No subgrupo 2.2 as sequências das espécies de *Arabidopsis*, *B. rapa* e *C. rubella* formam um grupo filogeneticamente

O último subgrupo, formado dentro do maior grupo de fitocistatinas carboxi-extendidas (grupo 2), é o 2.5 que agrupa as sequências das espécies de *E. grandis* e *C. papaya*. Externamente ao grupo 2, encontra-se um clado onde são encontradas apenas espécies de dicotiledôneas, o grupo 3. Neste último grupo, estão as sequências de fitocistatinas com extensão C-terminal das seguintes espécies, *P. trichocarpa*, *R. communis*, *C. sinensis* e *V. vinífera*, formando o subgrupo 3.1, além de sequências de *Arabidopsis*, *B. rapa* e *C. rubella*, formando o subgrupo 3.2. As espécies que formam o grupo 3, exceto *C. sinensis* e *V. vinífera*, também estão representadas no grupo 2 com outra(s) sequência(s) adicional(is) de fitocistatina(s) carboxi-extendida(s). Para melhor verificar a possível origem desta separação filogenética, o alinhamento de sequências das espécies presentes tanto no grupo 2 quanto no grupo 3 foi realizado. Algumas peculiaridades das sequências do grupo 3 podem ser destacadas, sendo elas principalmente relacionadas com substituições de aminoácidos em sítios específicos, em relação às sequências homólogas presentes no grupo 2. O alinhamento pode ser observado na figura 20, onde se encontram destacados treze posições em que aminoácidos foram trocados. Destes treze sítios, em seis deles as substituições mostram-se de forma altamente conservada, uma vez que todas as sequências do grupo 2 possuem um único aminoácido em comum, enquanto as do grupo 3 possuem outro. Os outros sete sítios de substituição são um pouco mais variáveis, já que as sequências do grupo 2 mantêm até dois tipos de aminoácidos em comum e as do grupo 3 também possuem até dois outros tipos de aminoácidos em comum. Sendo que estes tipos de aminoácidos nunca são os mesmos entre os dois grupos filogenéticos, quando considerados os mesmos sítios da sequência proteica.

4.4 Obtenção do vetor RNAi e transformação genética de plantas de arroz para silenciamento da fitocistatina XII

Mediante a síntese de *primers* específicos para a amplificação da região C-terminal, exclusiva da fitocistatina XII (figura 21 A), foi possível a obtenção de um fragmento correspondente à 138 pb, denominado fragmento Oc12i (figura 21 B). A inserção deste *amplicon* no vetor pENTR (figura 21 C) foi confirmada por reação de PCR, pela amplificação com os *primers* específicos do fragmento Oc12i (figura 21 D). Dos três clones obtidos, apenas um foi escolhido para a reação de recombinação com o vetor pANDA (figura 21 E). A confirmação da orientação correta dos fragmentos Oc12i no vetor pANDA foi realizada com combinações de *primers* resultando em amplicons de 968 pb e 870 pb respectivamente, de acordo com as figuras 21 E e F. Dos quatro vetores testados, todos possuíam as duas inserções nas corretas orientações, e um destes foi selecionado para transformação em *A. tumefaciens* e posterior transformação dos calos embriogênicos de arroz. Durante todas as etapas os amplicons obtidos foram purificados e sequenciados, e suas seqüências confirmadas contra os bancos de dados Phytozome e NCBI (*National Center of Biotechnology Information*).

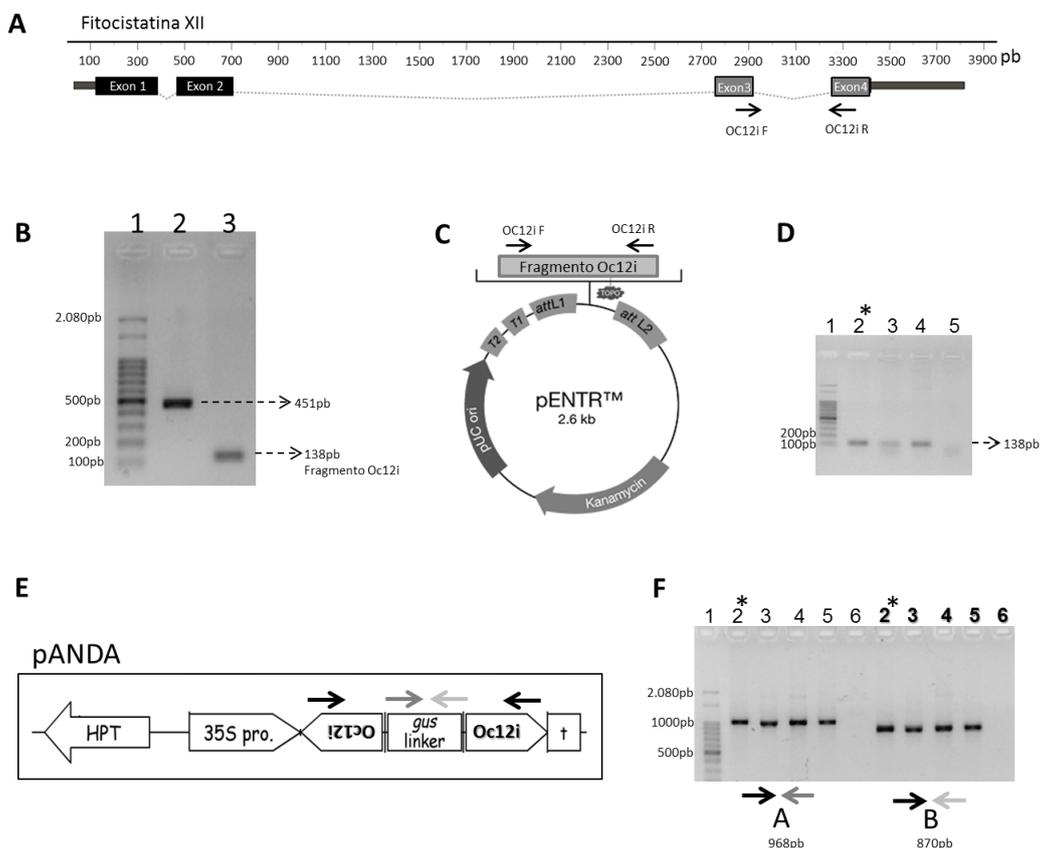


Figura 21. Construção do vetor de silenciamento gênico para a fitocistatina XII de arroz. **A.** Representação da estrutura gênica da fitocistatina XII, demonstrando o local de anelamento dos *primers* projetados para a amplificação de um fragmento específico. **B.** Resultado da amplificação com os primers Oc12i, sendo o primeiro fragmento, indicado pelo número 2, oriundo da amplificação do DNA genômico da planta e o fragmento indicado pelo número 3, resultante da amplificação no cDNA da planta. Na canaleta número 1 pode-se verificar o marcador de peso molecular, para confirmação do tamanho esperado das bandas obtidas. **C.** Representação do local de inserção do fragmento Oc12i no vetor pENTR, e os *primers* utilizados para a confirmação da inserção no vetor, o resultado da amplificação encontra-se na foto da figura **D.** O asterisco (*) indica o clone escolhido para prosseguir com o experimento. **E.** Demonstração do vetor pANDA recombinado com o fragmento Oc12i, e as setas indicam o local e sentido de anelamento dos primers utilizados para a confirmação do vetor, como demonstrado na figura **F**, com a combinação A e B de primers, de acordo com as cores das setas na figura **E**.

A obtenção dos calos de arroz transgênicos, transformados com o vetor de silenciamento RNAi está ilustrada na figura 22.

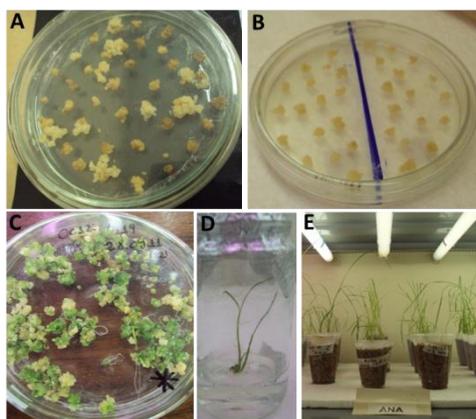


Figura 22. Transformação de calos embriogênicos de arroz para obtenção de plantas silenciadas para OcXII. **A.** Seleção dos calos transformados em higromicina. **B.** Proliferação dos calos transgênicos, cada metade da placa corresponde a um mesmo evento de transformação. **C.** Regeneração dos calos.

D. Seleção das plântulas regeneradas em higromicina. E. Crescimento das plantas em vermiculita.

4.5 Plantas transgênicas resultantes da transformação para silenciamento gênico da fitocistatina XII

Após a regeneração e seleção das plantas transformadas em higromicina, o resultado foi a obtenção de 92 plantas, correspondentes a nove linhagens transgênicas com o vetor RNAi OcXII e plantas controle, não transformadas. Conforme descrito na tabela 9 abaixo.

Tabela 9. Plantas obtidas após o processo de transformação dos calos embriogênicos com o vetor RNAi para a fitocistatina XII.

Linhagem	Plantas	Nº de Indivíduos
8	A-B-C-D-E-F-G-H	8
9	A-B-C-D-E-F-G-H-I-J-K-L-M	13
10	A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-X-Y-W-Z	26
11	A-B-C-D-E	4
14	A	1
16	A	1
19	A-B-C-D-E	5
24	A-B-C-D-E-F-G-H-I-J	10
25	A-B-C-D	4
Nipp	A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T	20
Total	9 Linhagens	92 plantas

Destas nove linhagens obtidas, algumas plantas foram selecionadas a fim de comprovar-se a transgenia das mesmas. As linhagens testadas encontram-se destacadas na figura 23. Os dois pares de *primers* utilizados para a confirmação da transgenia correspondem, primeiramente, ao gene HPT (figura 23 A), cujo *amplicon* corresponde ao tamanho de 430pb, e em seguida as mesmas amostras foram testadas com o *primer gus-linker* (figura 23 B), resultando em um fragmento amplificado de 636pb. Todas as amostras testadas apresentaram amplificação positiva, e com os tamanhos esperados, para os dois conjuntos de oligonucleotídeos, exceto a planta 10D, a qual apresentou amplificação positiva para a região do HPT, mas não com a região do *gus-linker*.

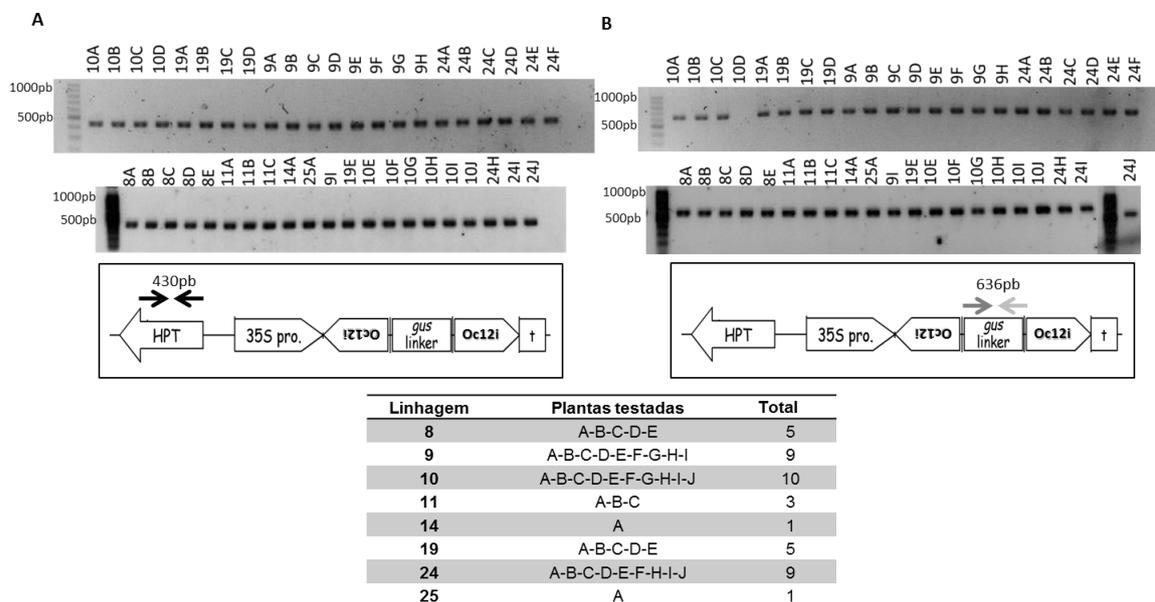
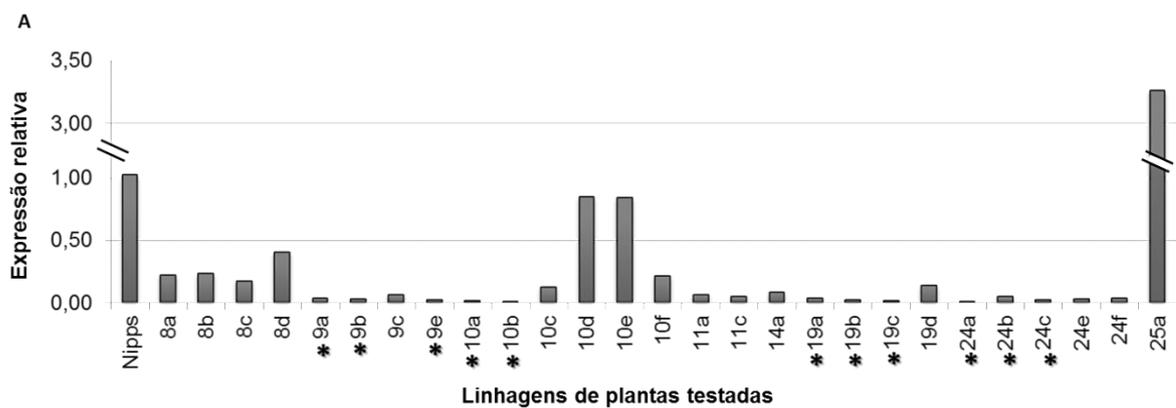


Figura 23. Linhagens de plantas T0, testadas para a confirmação da transgenia, e presença do vetor RNAi para a OcXII. **A.** Amplificação da região correspondente ao gene de resistência à higromicina (HPT), parte do vetor pANDA, integrado no genoma das plantas. O fragmento resultante corresponde a uma banda de 430 pb, conforme indicado no esquema abaixo da foto do gel. **B.** DNA das plantas testado para amplificação de 636 pb do gene *gus-linker*, conforme esquema abaixo da foto do gel. As setas indicam a posição e orientação dos *primers* utilizados. A tabela abaixo indica as linhagens e número total de plantas com o DNA testado.

Após a confirmação da transgenia, as plantas foram testadas para verificação do nível de silenciamento da fitocistatina XII. A análise por RTqPCR revelou que em grande parte das plantas transgênicas, o nível de mRNA da OcXII estava diminuído (figura 24), em comparação com o nível de expressão do gene nas plantas controle (Nipp). Plantas de quatro linhagens (L9, L10, L19 e L24) que apresentaram um maior nível de silenciamento, com cerca de 90% do mRNA do gene alvo reduzido, foram selecionadas para análises posteriores e estão indicadas na figura 24 B. Nas demais linhagens, L8, L11 e L14, também houve redução da expressão gênica da fitocistatina XII, porém de forma menos significativa. Algumas plantas não tiveram o nível do mRNA afetado de forma significativa (10d e 10e). Outra linhagem (25 a) parece ter uma superexpressão do gene e não sobreviveu por mais de três meses.



Linhagem	Plantas selecionadas com o menor nível de silenciamento de OcXII
9	A-B-E
10	A-B
19	A-B-C
24	A-B-C

Figura 24. Análise por RTqPCR do nível de silenciamento do gene OcXII nas plantas transgênicas (T0). **A.** Níveis de expressão gênica de OcXII nas diferentes plantas de cada linhagem. As plantas indicadas com asterisco (*) foram escolhidas por apresentarem o maior nível de silenciamento de OcXII, e estão indicadas na tabela abaixo do gráfico.

As plantas das quatro linhagens selecionadas (L9, L10, L19 e L24) foram submetidas a experimentos de RTqPCR a fim de verificar a alteração da expressão gênica nos demais onze genes que codificam fitocistatinas. No conjunto das quatro linhagens não foi observada variação significativa entre o nível de expressão das fitocistatinas em plantas controle (Nipp) e plantas OcXII RNAi, exceto para a fitocistatina XII, que apresenta-se efetivamente silenciada. As linhagens L10 e L24 apresentaram uma pequena redução no nível de expressão da OcVI (figura 25). Para melhor visualização do nível de expressão gênica das fitocistatinas entre plantas normais e RNAi, os dados da figura 25 foram plotados em escala, de forma que todos foram normalizados em relação as plantas Nipp da OcX, como controle endógeno (figura 26). Assim, torna-se possível a visualização dos genes que são de fato mais expressos nas plantas (Ocl, OcIII, OcVI, OcXI e OcXII), validando o resultado de que não existem variações significativas entre a expressão das fitocistatinas nas plantas transformadas em relação às controle, exceto OcXII e OcVI, como já relatado. Neste experimento, não foi possível detectar transcritos referentes aos genes das fitocistatinas II, V, VII e IX, assim como em experimentos anteriores.

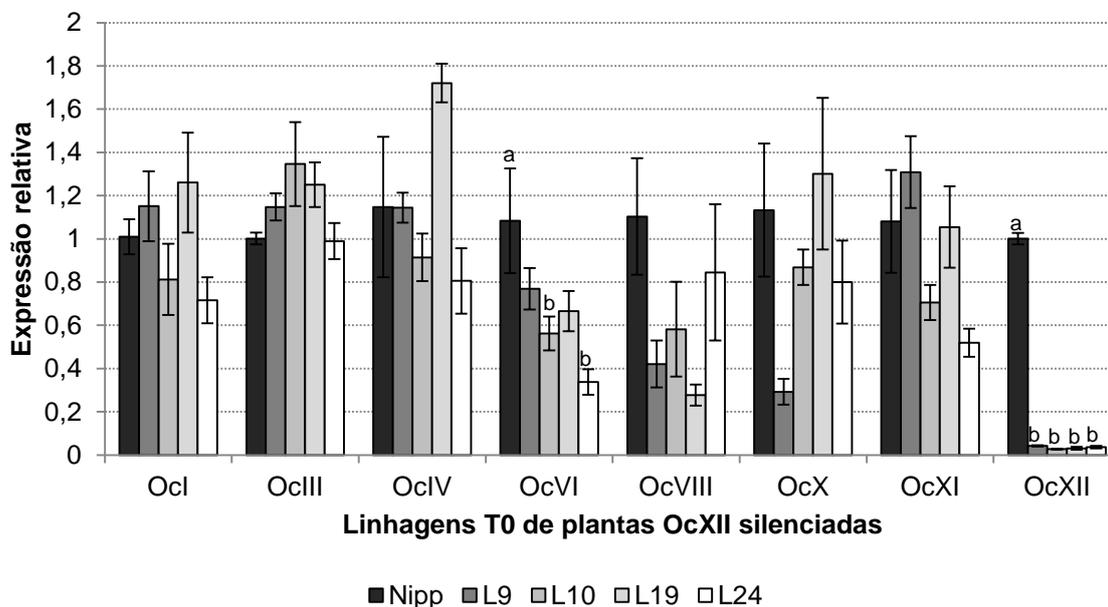


Figura 25. Expressão relativa das fitocistatinas nas plantas controle (Nipp), em relação às quatro linhagens de plantas OcXII RNAi (L9, L10, L19 e L24). Os grupos estatísticos referem-se ao teste de Tuckey com $p < 0.05$. As fitocistatinas II, V, VII e IX não foram detectadas em nível de mRNA.

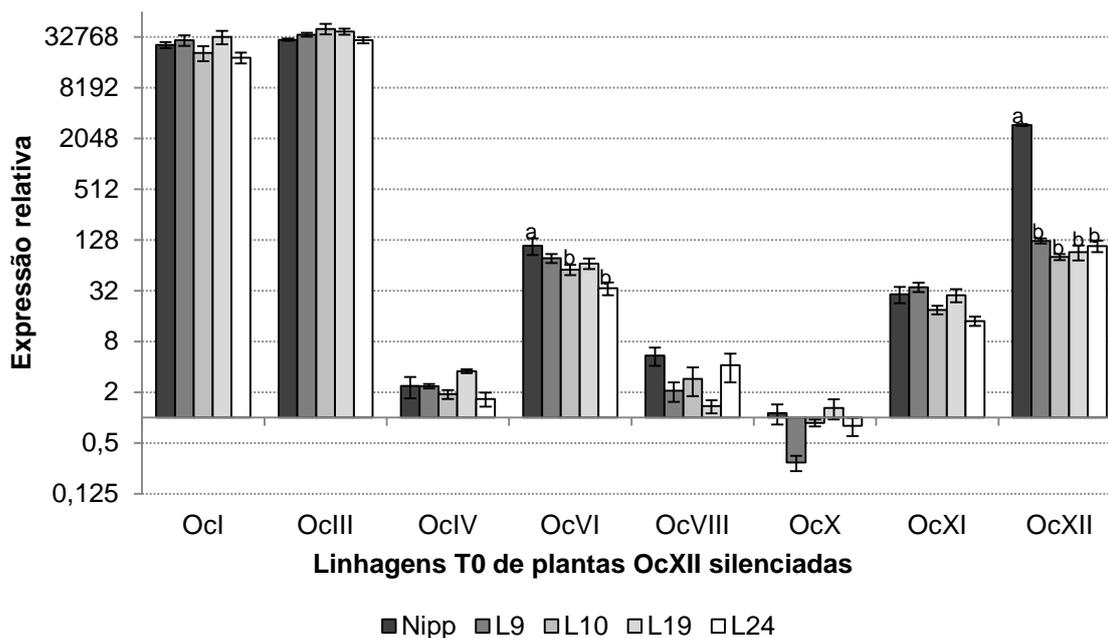


Figura 26. Expressão relativa das fitocistatinas nas plantas controle (Nipp), em comparação com as plantas silenciadas das linhagens L9, L10, L19 e L24. Os dados estão plotados no gráfico de forma a demonstrar o nível proporcional de expressão de cada gene das fitocistatinas, nos tecidos de folha analisados.

A verificação do nível de silenciamento de OcXII, juntamente com o nível de expressão do vetor RNAi, foi realizada para as quatro linhagens de plantas (L9, L10, L19 e L24). Três conjuntos de oligonucleotídeos foram utilizados (figuras 27 A e B). O resultado obtido demonstra que a expressão do transgene (*hairpin*) é relativamente variável entre as linhagens, sendo maior na L9, intermediário em L19 e L24, e menor na L10 (figura 27 C). Entretanto, esta variação não gera diferenças significativas no silenciamento de OcXII entre as quatro linhagens avaliadas, uma vez que o nível de mRNA de OcXII detectado não varia entre elas.

Adicionalmente, a expressão gênica relativa dos genes de legumainas de arroz foi validada, uma vez que estas proteases seriam os alvos primários da fitocistatina XII, e, devido ao silenciamento desta, poderiam sofrer algum tipo de alteração em nível de expressão gênica. O resultado obtido revela que não há variação significativa em nível de expressão gênica das legumainas, nas plantas OcXII silenciadas, em relação às plantas controle (figura 28).

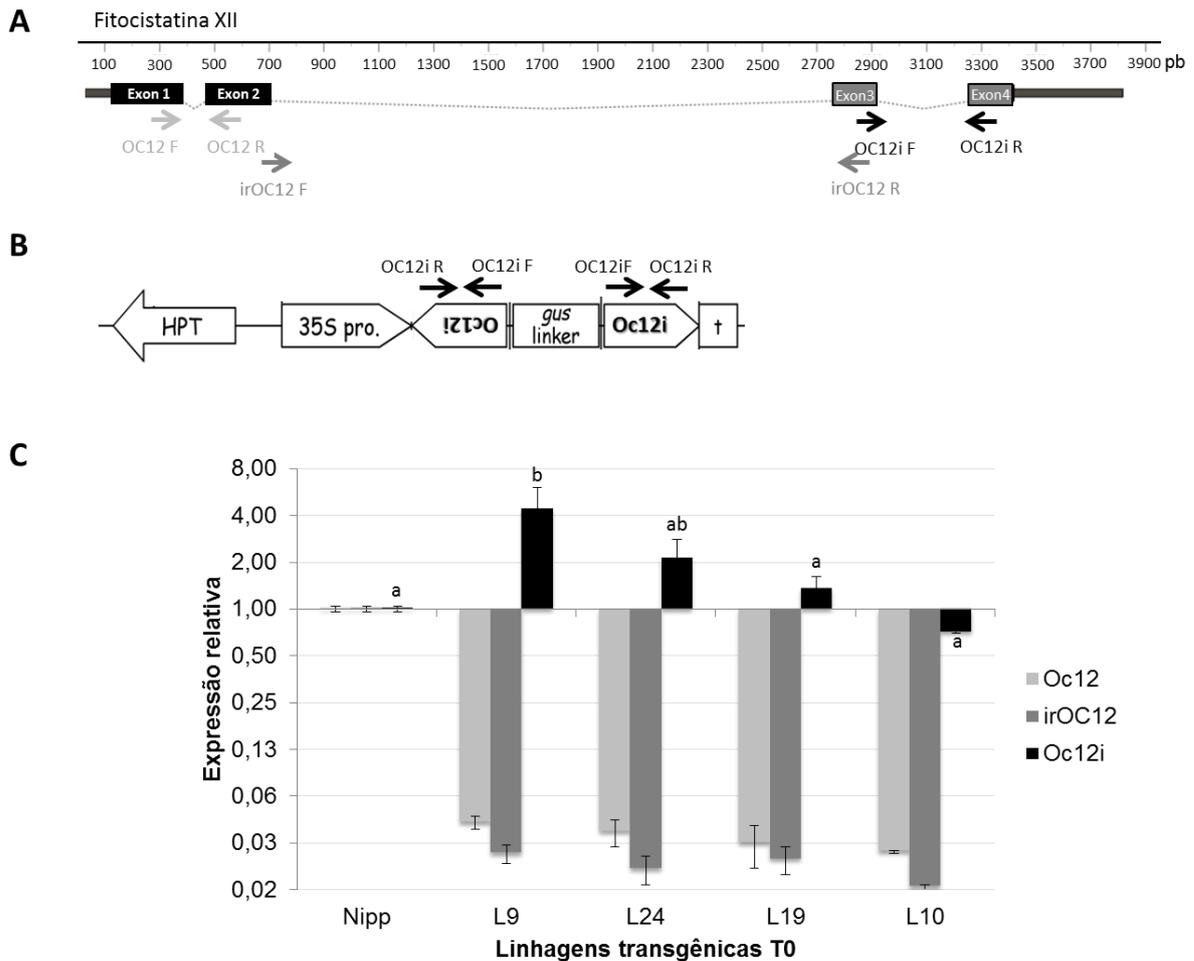


Figura 27. Silenciamento da fitocistatina XII em comparação ao nível de expressão do vetor RNAi. **(A)** Conjunto de *primers* utilizados para amplificação da fitocistatina XII, sendo que o último (Oc12iF+Oc12iR), também amplifica o fragmento Oc12i inserido no vetor de silenciamento **(B)**. RTqPCR das linhagens L9, L10, L19 e L24. Cada barra corresponde a um dos três conjuntos de primers utilizados, sendo as duas primeiras (Oc12 e irOC12) correspondentes ao nível de silenciamento da OcXII e último (Oc12i), representante da expressão do vetor RNAi. Análise estatística referente ao teste de Tuckey $p < 0.05$.

As mesmas quatro linhagens de plantas silenciadas para a fitocistatina XII (L9, L10, L19 e L24), foram analisadas quanto à presença de características morfológicas alteradas. O fenótipo das plantas não apresentou variações altamente significativas visualmente, o que pode ser verificado na figura 29.

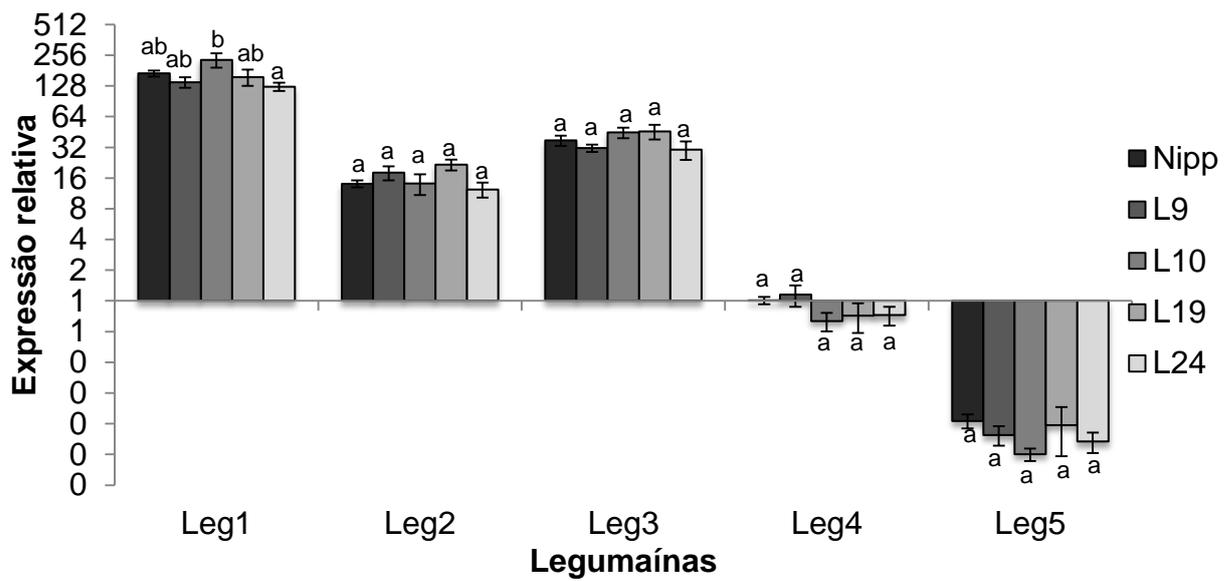


Figura 28. Expressão gênica relativa dos cinco genes de legumáinas de arroz, em plantas controle (Nipp), e linhagens silenciadas para OcXII (L9, L10, L19 e L24). Análise estatística referente ao teste de Tuckey com $p < 0.05$.



Figura 29. Fenótipo de plantas OcXII silenciadas e plantas controle, considerando-se a parte aérea, as raízes e as panículas das respectivas linhagens.

Para uma melhor quantificação da existência ou não de diferenças fenotípicas entre as plantas, foram realizadas medições e comparações entre as plantas, onde foram consideradas oito características fenotípicas principais (figura 30). Em relação à altura das plantas, L9 e L10 não são diferentes da planta controle (Nipp), enquanto que as linhagens L19 e L24 são aproximadamente 10 a 15% menores que as Nipps (figura 30 A). O índice de viabilidade das plantas também é diferente apenas para as linhagens L19 e L24, sendo ~60% menor que nas plantas controle (figura 30 B).

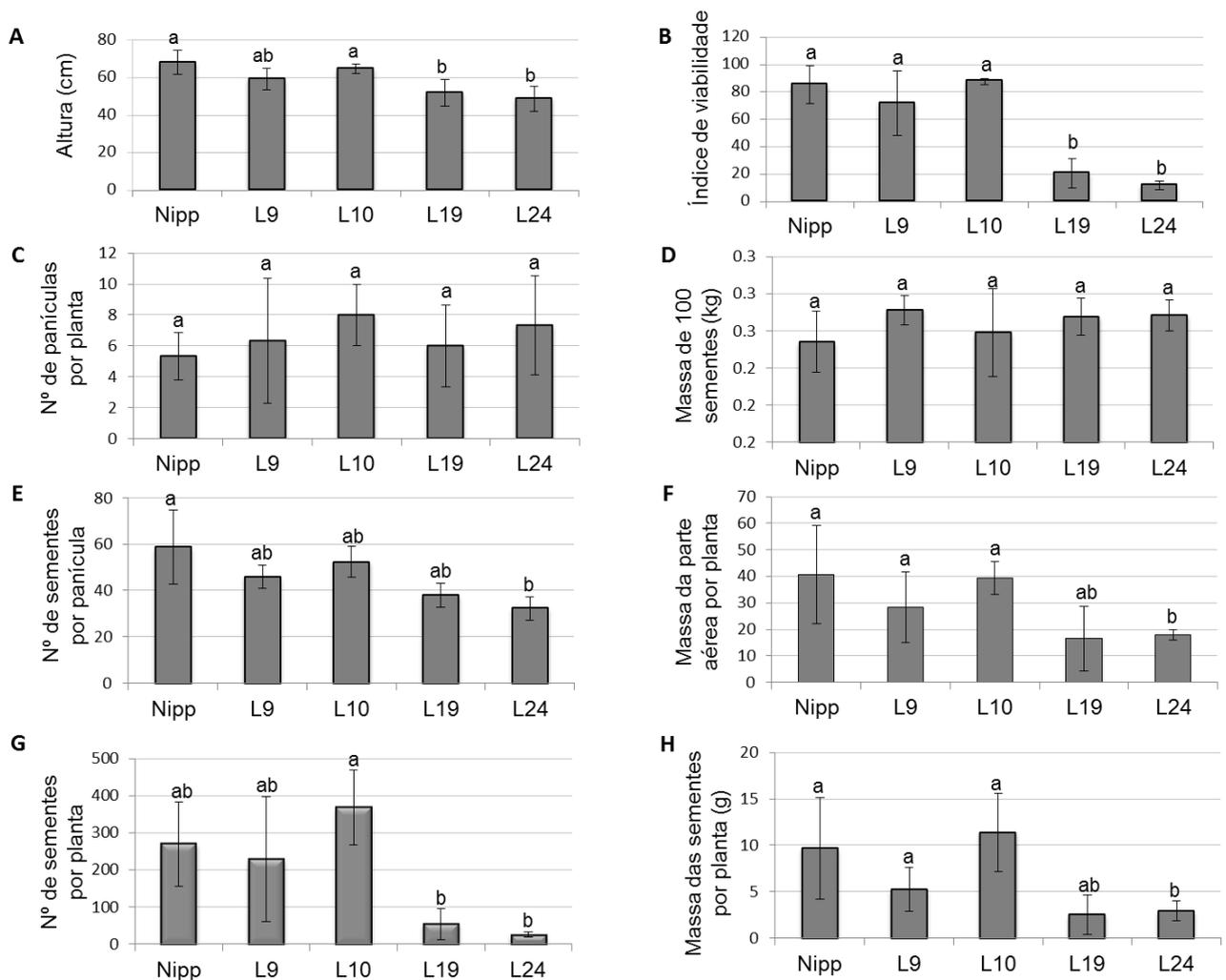


Figura 30. Análises fenológicas das plantas controle (Nipp), e linhagens de plantas OcXII silenciadas (L9, L10, L19 e L24). Análise estatística realizada com o uso de ANOVA, um fator, teste de Tuckey $p < 0.05$.

O número de panículas por planta, assim como a massa das sementes não apresentaram variações entre as plantas controle e as silenciadas (figura 30 C e D). No entanto, quando analisado o número de sementes por panícula (figura 30 E), a massa da parte aérea por planta (figura 30 F) e a massa de sementes por planta (figura 30 H), a linhagem 24 sofre reduções significativas, quando comparada com as plantas controle. Por fim, o número de sementes analisado por planta (figura 30 G), não revela diferenças significativas entre as plantas, apesar das linhagens 19 e 24 produzirem um número extremamente reduzido de sementes.

4.6 Análises de plantas T1 das linhagens OcXII silenciadas

As sementes obtidas das linhagens analisadas anteriormente, Nipp, L9, L10, L19 e L24, foram coletadas e colocadas para geminar, a fim de verificar se as plantas T1 permaneciam transgênicas e silenciadas para a fitocistatina XII. Conforme indicado na tabela 10, foram obtidas 12 plantas T1 para a linhagem controle Nipp, 12 plantas para a L9, 13 plantas para a L10, apenas uma planta da L19 e três da L24. Amostras de folha destas plantas foram coletadas e o DNA das mesmas extraído e analisado para a presença do transgene, mediante reação de PCR para o gene *gus-linker*, do vetor de RNAi (figura 31). Esta análise possibilitou a confirmação da transformação genética passada para a primeira geração das plantas transgênicas OcXII silenciadas. Apenas uma planta da linhagem 10 não apresentou amplificação para o gene *gus*.

Tabela 10. Obtenção da primeira geração de plantas transformadas (T1), silenciadas para a OcXII. Número de sementes e plantas obtidas para cada linhagem. Na última coluna está sumarizado o resultado da figura 31, para a presença do transgene no genoma das plantas.

Linhagem T0	Sementes T1	Plantas T1	Amplificação positiva para o transgene
Nipp	100	12	0
9	100	12	12
10	100	13	12
19	20	1	1
24	13	3	3

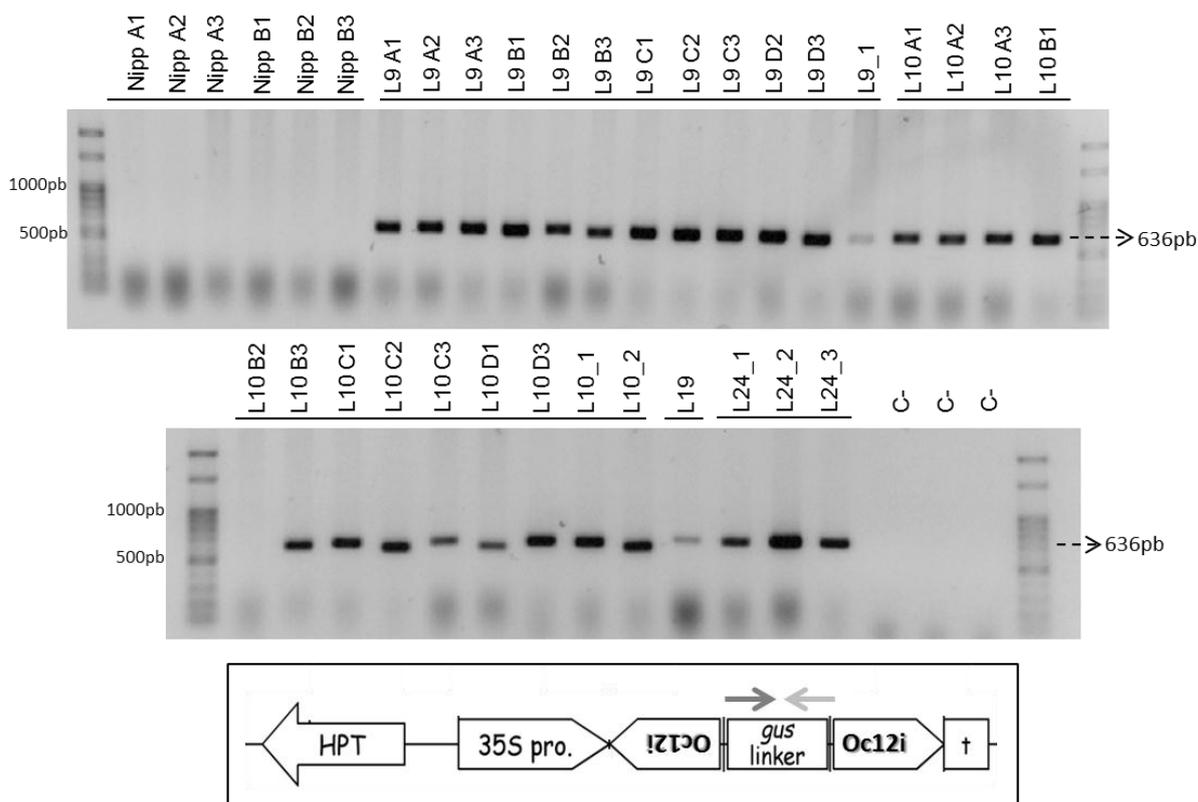
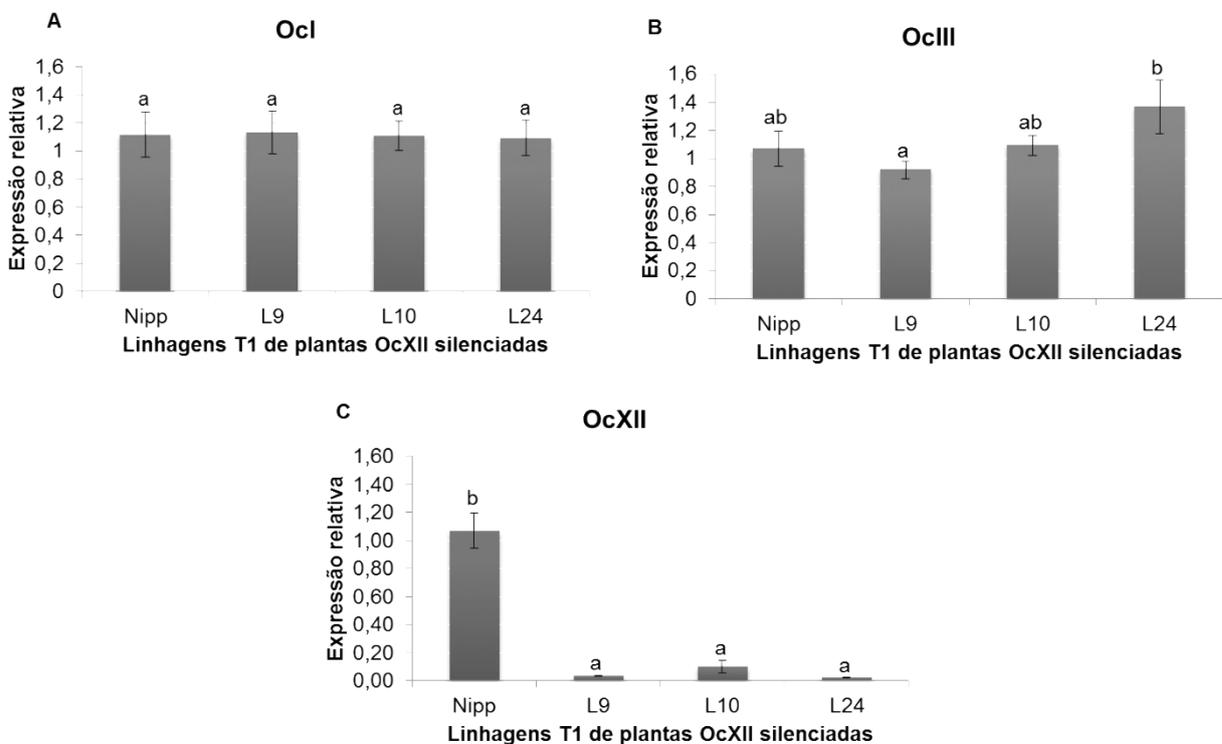


Figura 31. Amplificação do gene *gus-linker*, presente no vetor RNAi OcXII, como esquematizado abaixo da foto do gel, utilizando o DNA das plantas obtidas na geração T1. A confirmação da banda de 636pb ocorreu em todas as plantas transgênicas, exceto na planta B2 da L10. As plantas controle, Nipp, permanecem sem amplificação do fragmento, como esperado.

Após a confirmação da transgenia das plantas, foram realizados experimentos de RTqPCR para verificação de alterações no nível de expressão gênica das demais fitocistatinas e legumainas. Nesta parte do estudo, foram testados apenas os genes com maior relevância de expressão gênica, as fitocistatinas Ocl, OcIII e OcXII, além das legumainas Leg1, Leg2, Leg3 e Leg4. O resultado observado demonstra que não existem variações significativas na expressão gênica relativa das fitocistatinas, exceto OcXII, que se mantém silenciada (figura 32 A, B e C). As legumainas Leg2 e Leg4 não apresentaram alterações significativas entre as plantas transformadas e as plantas controle (figura 32 E e G). Entretanto, a Leg1 mostrou-se com uma expressão significativamente reduzida (aproximadamente 60%) na linhagem 24, não tendo variações entre as demais linhagens e as plantas controle (figura 32 D). A Leg3 sofreu alterações de expressão gênica na L9 e na L24, sendo na L9 o dobro da Nipp, e na L24 três vezes maior que na Nipp (figura 32 F).



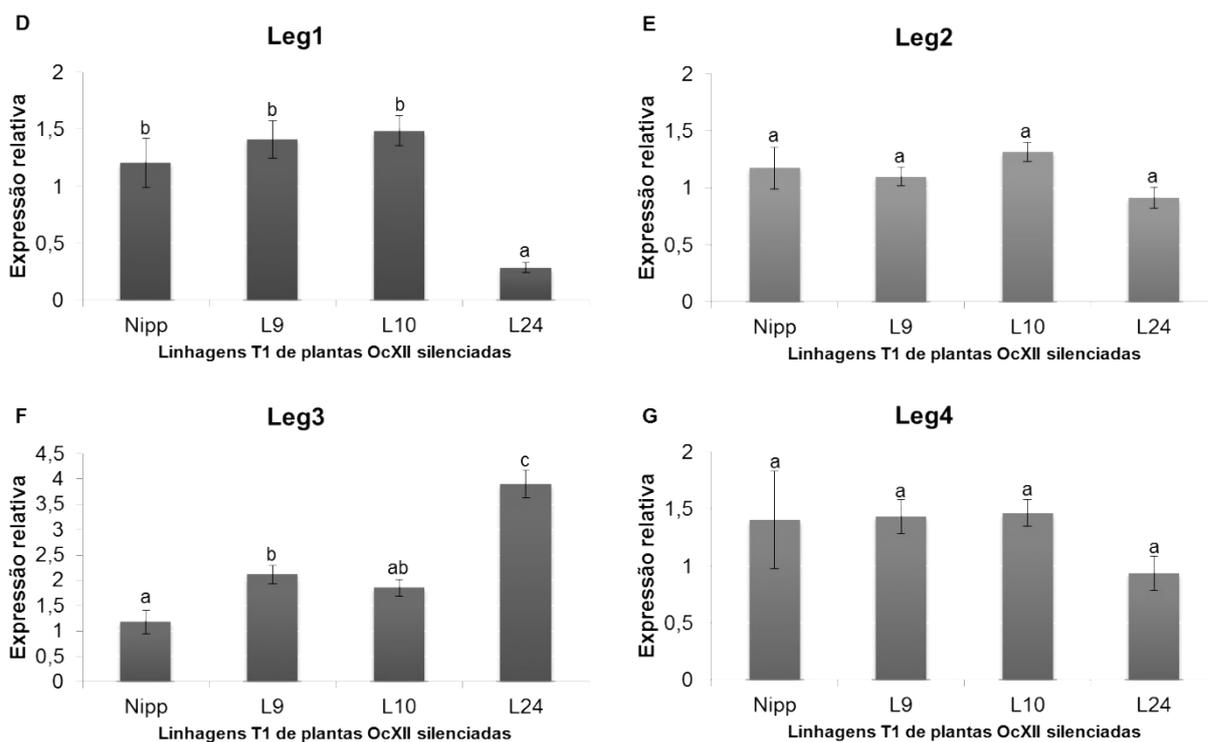


Figura 32. Expressão gênica relativa das fitocistatinas e legumainas na geração T1, oriunda de plantas OcXII silenciadas. Apenas as linhagens com pelo menos três plantas foram consideradas para a análise (tabela 10), e a estatística obtida pelo teste de Tuckey com $p < 0.05$.

4.7 Análises da região promotora de transcrição do gene da fitocistatina XII

A análise *in silico* de parte da região promotora de transcrição da OcXII, resultou na identificação de sítios de ligação (*cis*-elementos) para 20 fatores de transcrição distintos (tabela 11). Tais *cis*-elementos identificados estão relacionados à ativação da transcrição de genes via ligação de fatores de transcrição, relacionados à resposta por ácido abscísico, expressão específica de endosperma, resistência a doenças, resposta à seca, giberilina, luz, entre outros descritos na tabela 12.

Tabela 11. *Cis*-elementos encontrados na região promotora do gene da fitocistatina XII. Análise referente a 2.300 pb da região genômica que antecede o sítio de início de transcrição do gene.

Sítios de ligação de fatores de transcrição na região promotora da fitocistatina XII					
Fator	Sítio	Sentido	Fator	Sítio	Sentido
ACGTABREMOTIFA2OSEM	1089	-	GCN4OSGLUB1	1134	-
	2189	-	GT1CONSENSUS	1068	-
ACGTOSGLUB1	136	+		1606	-
AMYBOX1	523	-		1626	-
ANAEROCONSENSUS	1396	+		1707	-
BIHD1OS	100	-		1728	-
	1379	+		1753	-
	194	-		1869	-
	288	-		1885	+
	447	-		418	+
	689	-		504	-
BP5OSWX	1550	-	IRO2OS	1090	-
CAREOSREP1	1011	-		2190	-
	1302	+	POLASIG2	1385	+
	2213	+	PROLAMINBOXOSGLUB1	274	-
CGACGOSAMY3	1128	+	SITEIIATCYTC	1029	-
	1646	-		1415	+
	2091	-		1431	-
	2156	+		1487	-
	2268	-		1501	-
	299	-		1543	+
	728	+		1573	-
	841	-		1617	-
DRECRTCOREAT	1054	-		2039	+
	1328	-		2068	-
	2092	-		2082	+
E2FCONSENSUS	1664	-		2113	-
	415	-		2129	+
	547	+		234	+
GARE1OSREP1	523	-		TATABOXPAL	1667
GATABOX	1070	-	WRKY710S	100	+
	1267	-		1044	-
	1628	-		1125	+
	1688	-		1134	+
	1715	-		1380	-
	1719	-		194	+
	1880	+		2255	-
	1938	-		288	+
	1942	-		355	-
	1978	+		447	+
	2058	+		637	+
	428	+		698	+
	768	-			

Tabela 12. Descrição da função relacionada com o fator de transcrição encontrado para os *cis*-elementos presentes na região promotora de OcXII.

Fator de transcrição	Função relacionada
ACGTABREMOTIFA2OSEM	Elemento responsivo a Ácido abscísico
ACGTOSGLUB1	Elemento necessário para a expressão específica em endosperma
AMYBOX1	Sequência conservada de regiões 5' da p-amilase
ANAEROCONSENSUS	Motivo de genes induzidos por anaerobiose
BIHD1OS	Ligação do fator de transcrição com homeodomínios BELL para resposta de resistência a doenças
BP5OSWX	Ligação da proteína MYC
CAREOSREP1	Elemento encontrado em regiões promotoras de genes de proteases cisteínicas
CGACGOSAMY3	Acoplamento do elemento G-box
DRECRTCOREAT	Elemento responsivo à desidratação
E2FCONSENSUS	Motivo de ligação do fator de transcrição E2F
GARE1OSREP1	Elemento responsivo a giberilina
GATABOX	Necessário para alta expressão, resposta a luz e expressão tecido-específica
GCN4OSGLUB1	Expressão específica de endosperma
GT1CONSENSUS	Elemento encontrado em genes regulados por luz
IRO2OS	Indução pela deficiência de ferro
POLASIG2	Sinal Poli-A
PROLAMINBOXOSGLUB1	Elemento necessário para a regulação quantitativa do gene GluB-1, em endosperma
SITEIIATCYTC	Encontrado em regiões promotoras de genes de citocromo e genes nucleares envolvidos na maquinaria de fosforilação oxidativa
TATABOXPAL	Sítio de ligação do fator de transcrição OsTBP2
WRKY710S	Ligação do fator repressor transcricional da via de sinalização de giberilina

4.8 Obtenção do vetor pHGWS7 expressando genes repórteres de forma controlada pelo promotor da fitocistatina XII

A fim de investigar o padrão de expressão do gene da fitocistatina XII *in vivo*, localizando células e tecidos específicos onde o promotor deste gene é ativado, levando à transcrição do mesmo, foi realizada a amplificação da região genômica correspondente a uma região do promotor OcXII (figura 33A). O fragmento amplificado corresponde a um *amplicon* de 2.138 pb (figura 33B), o qual foi inserido em vetor *Gateway* pENTR (figura 33C). Após a obtenção das colônias de *E.coli* contendo o vetor pENTR transformado, os plasmídeos das mesmas foram extraídos e confirmados para a inserção do promotor pela reação de PCR com os primers PromoOc12 indicados na figura 33C, e resultando nas amplificações apresentadas na figura 33D. Destas 20 colônias testadas, apenas uma foi selecionada e sequenciada para confirmação do inserto. Posteriormente à confirmação da sequência nucleotídica inserida, este vetor pENTR foi utilizado para recombinação com o vetor de expressão pHGWS7. A recombinação do vetor foi confirmada por reações de PCR com três conjuntos de *primers* indicados na figura 33 E, os quais resultaram nos respectivos fragmentos: A) 2.138 pb, B) e C) 2.300 pb (figura 33F), confirmando a presença e orientação correta da inserção do fragmento promotor OcXII no vetor de expressão gênica.

Apenas um dos plasmídeos confirmados na figura 33 F foi selecionado para dar continuidade ao experimento. Assim, o plasmídeo de número 5 foi utilizado para transformação de *A. tumefaciens* e de calos embriogênicos de arroz. A transformação dos calos resultou em quatro linhagens resistentes à higromicina, as quais foram testadas para a funcionalidade do vetor de expressão pela incubação de alguns calos com o substrato X-gluc por 24 horas. Das quatro linhagens testadas, duas apresentaram uma fraca marcação em azul (L1 e L3) (figura 34 A). A linhagem 5 foi onde houve a maior marcação dos calos em azul e a linhagem 4 não apresentou sinais de marcação

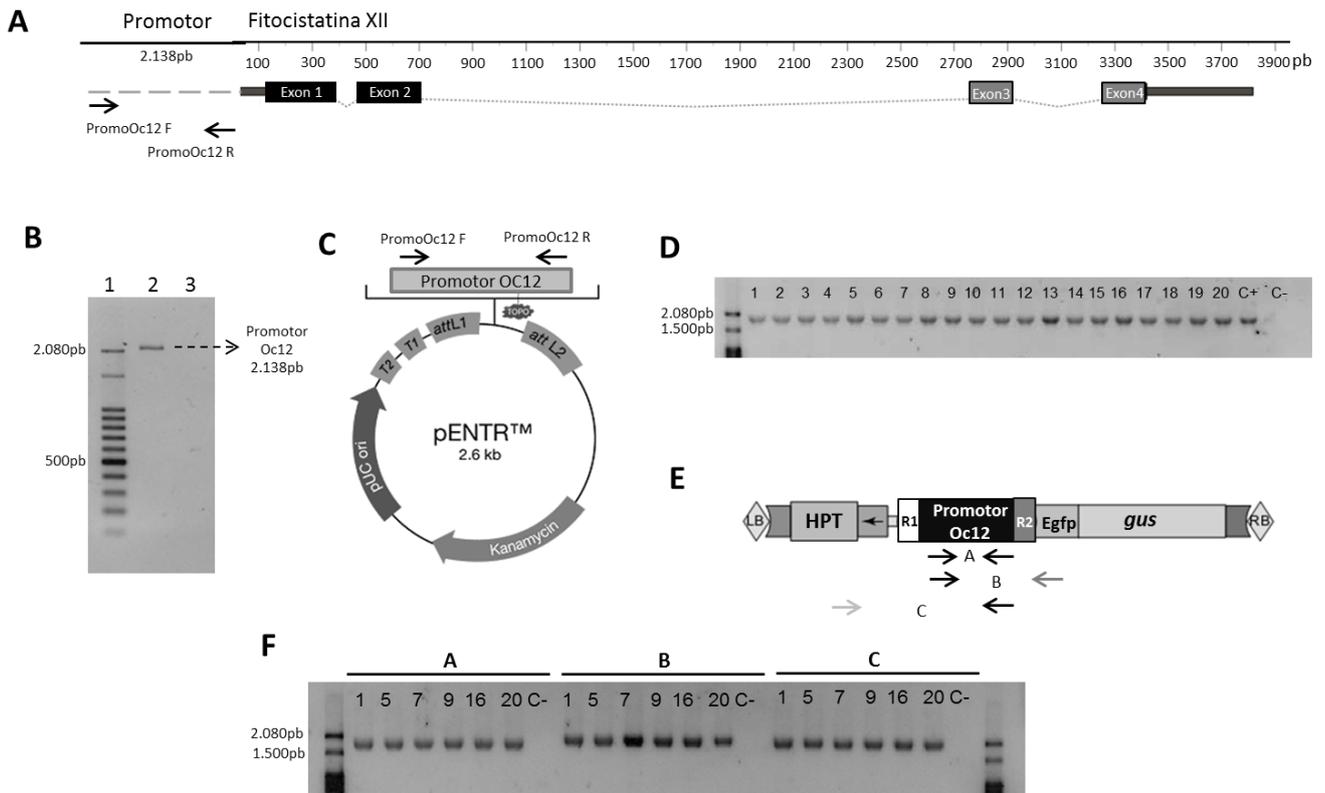


Figura 33. Obtenção do vetor pHGWFS7 com o promotor OcXII controlando a expressão dos genes repórteres *gus* e *Egfp*. (A) Estrutura genômica da fitocistatina XII, demonstrando a região amplificada com os *primers* PromoOc12. (B) Fragmento resultante da amplificação da região demonstrada na figura A, representado na coluna 2. Marcador molecular e controle negativo representados nas colunas 1 e 3, respectivamente. (C) Vetor pENTR representando o local de inserção do fragmento Promotor Oc12, e as setas acima indicam os *primers* utilizados para a amplificação do fragmento inserido, como pode ser verificado na figura (D). (E) Vetor pHGWFS7 com o promotor OcXII inserido. As combinações de setas A, B e C representam os conjuntos de *primers* utilizados para a amplificação e confirmação da orientação correta do fragmento no vetor, o que pode ser visualizado no esquema da figura (F).

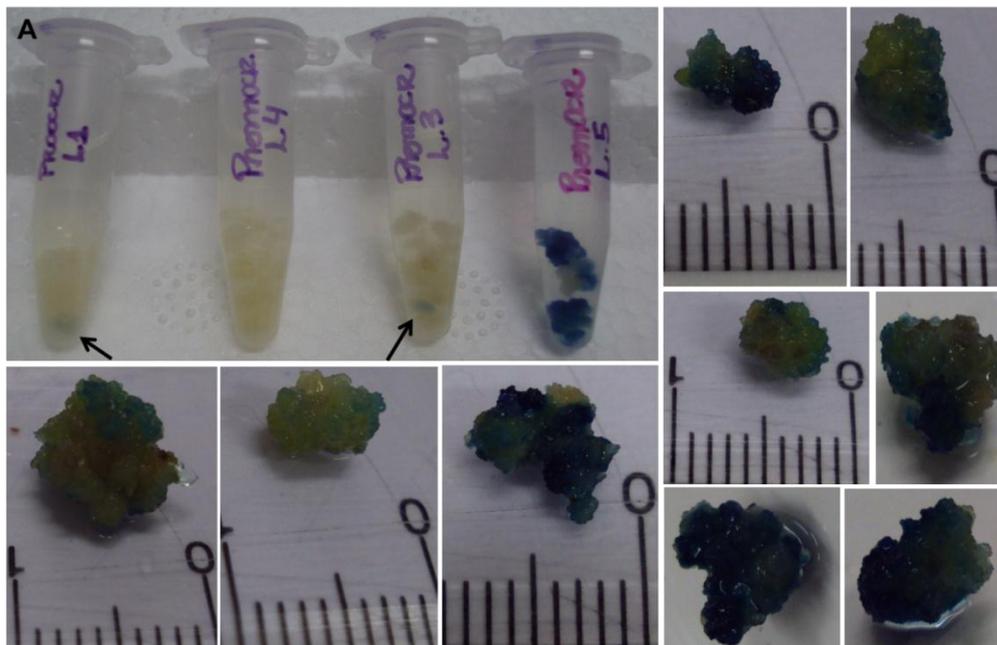


Figura 34. Calos transformados com a construção do vetor pHGWS7+Promotor Oc12. **(A)** Quatro linhagens testadas para a expressão do promotor Oc12, utilizando o substrato X-gluc. As demais figuras menores referem-se a calos da linhagem 5, que encontravam-se em meio de regeneração de plantas.

4.9 Modulação da expressão da fitocistatina XII por ABA

O ácido abscísico (ABA) é um fitormônio vegetal envolvido em diversos processos fisiológicos, tais como a síntese de proteínas de armazenamento em sementes, dormência da semente, inibição da fase de transição do crescimento embrionário para o crescimento germinativo além de estar envolvido em diversas respostas a estresses ambientais, como seca, alta salinidade, frio, entre outros (Ye, Jia, & Zhang 2012). Estando o ácido abscísico envolvido com estes processos, aos quais as fitocistatinas já foram também relacionadas, foi realizado o experimento com este fitormônio a fim de revelar se o mesmo era capaz de modular a expressão gênica da fitocistatina XII de arroz, em dois tecidos diferentes, folha e raiz.

A fitocistatina XII foi avaliada em três diferentes períodos de tratamento, a fim de analisar uma resposta do gene em curto prazo (2 e 4 horas), e a longo

prazo (24 horas). O resultado obtido evidencia que o gene da OcXII responde ao tratamento com o fitormônio ABA, e que esta é uma resposta rápida, uma vez que os níveis de expressão do gene aumentam gradativamente até o período de quatro horas de tratamento e após 24 horas os mesmos já estão retornando aos seus níveis basais (figura 35). Na raiz a resposta é mais rápida do que em folhas, como pode ser observado na análise feita após duas horas de tratamento. Após quatro horas de exposição a 100 μ M de ABA, os níveis de expressão de OcXII estão aumentados, de forma semelhante, tanto nas folhas quanto nas raízes das plantas analisadas. Este aumento, após quatro horas de tratamento, corresponde aproximadamente a oito vezes a expressão gênica verificada nas plantas controle, antes das mesmas receberem o tratamento (tempo 0). A análise da expressão de OcXII, 24 horas após a exposição ao fitormônio, demonstra que os níveis de mRNA voltam a cair, se assemelhando novamente aos níveis basais observadas no tempo 0. Porém, esta queda da expressão ocorre de forma em que o nível de mRNA detectado em raiz já está, novamente, bastante similar ao controle, enquanto que, em folhas o nível de expressão ainda está um pouco mais elevado, não diferindo estatisticamente do nível basal (tempo 0), nem do nível após quatro horas de tratamento.

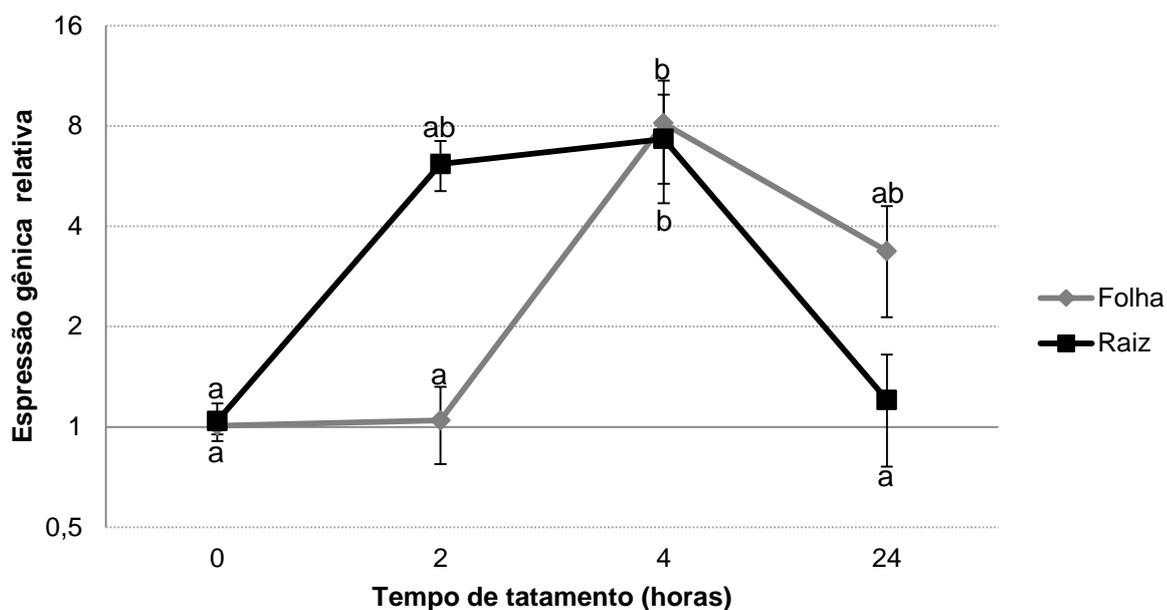


Figura 35. Expressão relativa da fitocistatina XII em raízes e folhas de plantas submetidas a tratamento com 100µM de ABA. A expressão relativa da OcXII foi avaliada nos quatro pontos marcados no gráfico, correspondendo aos tempos de 0, 2, 4 e 24 horas, nos quais amostras foram coletadas para validação da modulação da expressão gênica por ABA. A análise estatística refere-se ao Teste de Tuckey considerando $p < 0.05$.

5. DISCUSSÃO

5.1 Caracterização das fitocistatinas em arroz

As fitocistatinas, como inibidoras de proteases cisteínicas, são há algum tempo conhecidas em plantas. Em arroz, elas foram descritas por MARTÍNEZ *et al.* 2005, e desde então ainda não haviam sido realizados estudos para a caracterização funcional da totalidade destes genes na espécie *Oryza sativa*. Apenas os genes Ocl, OclI e OclII possuem estudos de caracterização mais aprofundados avaliando suas possíveis funções e expressões gênicas (ABE *et al.* 1987b; HIROTO *et al.* 1989; ARAI *et al.* 1991; OHTSUBO *et al.* 2005). O foco do trabalho aqui realizado envolve o gene da fitocistatina XII, porém, a caracterização deste no contexto de sua família gênica torna-se importante para uma visualização mais ampla das funções desempenhadas pela OcXII nas plantas de arroz.

Neste presente trabalho, com os resultados da caracterização gênica foi possível confirmar que o gene da OcXII era diferente dos demais onze genes da família, tendo uma sequência extra na sua região C-terminal. Como pode ser observado (figura 10), todos os genes de fitocistatinas compartilham sequências altamente conservadas. Apesar da estrutura de alguns genes ser mais variável, em se tratando do número de exons e introns, a sequência proteica codificada por eles não varia significativamente entre os *loci* gênicos. A fitocistatina XII é uma exceção, pois além dos exons em preto (região N-terminal), compartilhados com as demais 11 fitocistatinas, possui a sequência C-terminal extra representada pelos exons em cinza na figura 10. Esta sequência, apesar de remontar alguma similaridade com o domínio cistatina (tabela 7), possui algumas diferenças, entre elas, principalmente, a presença do motivo SNSL, que anteriormente foi caracterizado por MARTINEZ *et al.* 2007 em cevada, como um motivo importante para inibição de uma outra classe de proteases cisteínicas, as legumainas. Todas as doze fitocistatinas de arroz apresentam o motivo LARFAV característico, além dos motivos cistatinas conservados, QxVxG e PW. A exceção torna-se a fitocistatina IX, que apesar de ainda ser reconhecida pelos bancos de dados como

uma proteína similar a fitocistatinas, ela já não possui os motivos citados acima, que a caracterizam como uma inibidora de proteases cisteínicas.

Analisando-se os doze genes de fitocistatinas quanto aos seus padrões de expressão gênica, o único que não foi detectado por nenhum dos métodos utilizados foi a OcVII (figuras 11a, 12, 14 e 16). Enquanto que as demais fitocistatinas apresentaram um padrão variável entre os diversos métodos empregados, tendo em vista os tecidos específicos analisados. A comparação dos níveis de mRNA entre os tecidos de folha, que foram analisados tanto por microarranjos quanto por RNAseq e RTqPCR, demonstra que os genes OcI, OcIII e OcXII são os genes mais expressos neste tecido, além de OcVIII também ser detectado, em níveis mais baixos, pelas três diferentes metodologias. Quando a expressão destes quatro genes é verificada nos diferentes tecidos dos experimentos de microarranjos, também se percebe que os mesmos estão presentes de forma ubíqua nos demais tecidos da planta, mas apresentando algumas variações tecido-específica de expressão gênica (figura 12). O gene da OcII é consideravelmente elevado nos diferentes tecidos avaliados nos experimentos de microarranjos. Entretanto, a expressão deste gene possui um padrão divergente onde, nos microarranjos o nível de sua expressão é equivalente em folhas e raízes de plântulas (figura 12), mas na análise de RNAseq sua detecção foi de aproximadamente quatro vezes maior em raízes do que em folhas (figura 14). OcII não foi detectado em nenhum dos estádios de desenvolvimento analisados por RTqPCR (figura 16). Sendo assim, os resultados para este gene apresentam discrepâncias entre as técnicas utilizadas para detecção de sua expressão, mesmo quando são considerados os mesmos tecidos e estádios de desenvolvimento. Os demais genes OcIV, OcV, OcVI, OcIX, OcX e Oc XI foram detectados apenas com uma ou outra das metodologias empregadas, em apenas alguns tecidos e com um nível de expressão mais reduzido. Tais resultados indicam que os prováveis genes com maior relevância para as plantas de arroz, de um modo geral, são as fitocistatinas I, III, VIII e XII, pois estes genes possuem uma expressão mais constante e elevada, enquanto que as demais fitocistatinas não são tão expressas no decorrer do desenvolvimento da planta, sendo detectadas em apenas alguns tecidos

específicos (exceto OcV, OcVII, OcIX cujos níveis de mRNA não foram satisfatoriamente detectáveis). Um único estudo avaliando a expressão das fitocistatinas por ESTs, em tecidos de folha, raiz, flor e sementes de arroz, demonstrou, de forma similar, que a expressão da Ocl era detectada em todos os tecidos avaliados, OcII apenas em folha, flor e raiz, OcIII em folha e flor, OcVIII em folha e semente, OcX em folha e raiz e OcXII em folha e semente. Os demais genes não foram detectados (MARTINEZ *et. al*, 2005).

Os principais efeitos da inibição endógena pelas fitocistatinas estão relacionados com os processos de regulação de proteases cisteínicas em órgãos de armazenamento da planta (BENCHABANE *et al.* 2010). Um dos primeiros trabalhos realizados com a oryzacystatina já demonstrava que a expressão deste gene era bastante elevada durante a formação das sementes de arroz, onde também era possível correlacionar seu nível de mRNA com o aumento de mRNA de gluteína, uma das principais proteínas de armazenamento (ABE *et al.* 1987a). Assim, sabe-se que as fitocistatinas são ativamente sintetizadas nas sementes em desenvolvimento, ou eventualmente, em órgãos vegetativos para promover a deposição das proteínas de armazenamento. Segundo o modelo proposto por BENCHABANE *et al.* 2010, há um aumento na quantidade de fitocistatinas, em nível proteico, durante formação das sementes, quando ocorre a deposição das proteínas de armazenamento sendo que o nível de proteína das fitocistatinas permanece alto durante o período de dormência da semente, para proteção destas proteínas de armazenamento. No momento em que o início de germinação é sinalizado, o nível proteico de fitocistatinas diminui, de forma a aumentar a atividade das proteases cisteínicas e desencadear a hidrólise e mobilização das proteínas armazenadas. Quando a expressão gênica das doze fitocistatinas de arroz foi testada durante estádios de desenvolvimento e germinação de plantas, apenas os genes Ocl, OcIII, OcVI, OcVIII e OcXII tiveram seu mRNA detectado (figura 15 e 16). O padrão de expressão destes cinco genes revela de forma similar, que para as fitocistatinas de arroz há uma maior expressão gênica nas sementes maduras, decaindo com o início da germinação. Tal modo de expressão é condizente com as principais funções postuladas para as fitocistatinas e descritas no modelo acima. Também pode ser verificado que

alguns genes (Ocl, OclIII e OcVI), após passarem por um período de redução de expressão gênica, durante o desenvolvimento inicial das plantas, tem seu nível de expressão elevado em folhas de plantas adultas e entrando em senescência (figura 15 a, b e c). A expressão gênica destas últimas três fitocistatinas pode ser relacionada com a necessidade de inibição ou diminuição de outros processos fisiológicos, em tecidos vegetativos, aos quais as proteases cisteínicas estão envolvidas. Além da clara importância das proteases cisteínicas para o desenvolvimento e crescimento de plantas, elas também desempenham papéis cruciais durante a senescência e morte celular programada, sendo também importantes em vias de resposta a estresses bióticos e abióticos (GRUDKOWSKA & ZAGDAŃSKA 2004). Os processos de envelhecimento e senescência de folhas das plantas são caracterizados pelo aumento da expressão de alguns genes específicos. Como já demonstrado, proteases cisteínicas, assim como seus inibidores, possuem expressão elevada e coordenada em tecidos senescentes, de forma que o aumento do nível de mRNA de fitocistatinas nestes locais, sugerem prováveis funções na prevenção da morte celular não desejada em certos tecidos da folha (HUANG *et al.* 2001; TAJIMA *et al.* 2011). Em uma contextualização do nível de expressão gênica das fitocistatinas de arroz, nos diferentes estádios de desenvolvimento das plantas, pode-se verificar que as fitocistatinas mais expressas na planta são os genes Ocl, OclIII e OcXII (figura 16). Ocl e OclIII apresentam grandes elevações da expressão gênica em plantas mais velhas, demonstrando a possibilidade de serem os genes mais estritamente relacionados em desacelerar e controlar os processos de senescência e morte celular das plantas.

A fitocistatina XII teve seu padrão de expressão gênica mais detalhadamente analisado para melhor elucidação das suas possíveis implicações funcionais. Com capacidade de ser um inibidor bifuncional de proteases cisteínicas, como descrito para o gene ortólogo em cevada (MARTINEZ *et al.* 2007), além da OcXII regular processos fisiológicos desencadeados pelas papaínas, ela também pode inibir processos relacionados com as legumaínas. As análises de microarranjos na figura 11A, demonstram que este gene da OcXII tem seu nível de mRNA detectado em todos os estádios de desenvolvimento das plantas, sendo que

existem algumas variações na intensidade de expressão, principalmente nos estádios de pós-germinação e crescimento da planta. Tais variações são corroboradas pelo resultado apresentado na figura 11b, onde fica claramente visível que a expressão de OcXII é extremamente elevada nos estádios finais de maturação das sementes. Considerando-se os tecidos específicos avaliados (figura 12), a presença da OcXII é ubíqua e pouco variável entre níveis médios e altos de expressão gênica. É notável que a expressão deste gene se sobressai em sementes, principalmente no tecido de endosperma (figura 13). A análise comparativa da expressão de OcXII, por RTqPCR, durante a germinação corrobora com o maior nível de mRNA, sendo detectado na semente madura e este nível de expressão decai consideravelmente nas primeiras doze horas, após o início da germinação e permanece neste nível mais basal durante todo o desenvolvimento das plantas de arroz analisadas (figura 15E). O perfil de expressão gênica encontrado para a fitocistatina XII possibilita a afirmação de que uma maior quantidade da proteína codificada por este gene é necessária para a manutenção do processo de dormência das sementes, mas no momento em que o início da germinação é desencadeado, o nível de seus transcritos diminui de forma a proporcionar a redução da quantidade deste inibidor de proteases cisteínicas para ativação enzimática das proteases-alvo, levando à degradação das proteínas de armazenamento e mobilização de nutrientes. Porém, para uma melhor caracterização funcional dos genes de fitocistatinas, principalmente OcXII, é crucial a identificação de suas proteases-alvo, para a realização de uma correlação mais acurada. Dados semelhantes com o gene ortólogo de *A. thaliana*, AtCYS6, indicam que tanto transcritos, quanto a proteína codificada por este gene, já não é mais detectada em plantas com mais de 14 dias de idade, sugerindo que AtCYS6 tem um maior acúmulo em sementes e plântulas do que em plantas adultas (HWANG *et al.* 2009). Contudo, no presente estudo realizado com arroz, foi possível detectar a transcrição do gene OcXII em estádios mais avançados de desenvolvimentos das plantas, porém com um nível de mRNA bem menor do que nas sementes das mesmas. De forma similar, estudos com o gene ortólogo de OcXII em cacau (*Theobroma cacao*, TcCYS4) mostraram que a proteína codificada por este gene estava presente, em baixos níveis, em tecidos

de folhas maduras, mas quando estas folhas eram infectadas com o fungo *Moniliophthora perniciosa*, a proteína não era mais detectada. Tais observações levaram à hipótese de que o gene TcCYS4 estaria envolvido na prevenção da morte celular em folhas saudáveis, inibindo a atividade de proteases cisteínicas envolvidas neste processo, enquanto que em folhas infectadas, na ausência de TcCYS4 as proteases cisteínicas atuariam na morte celular (PIROVANI *et al.* 2010).

Em outros estudos realizados com cereais, é demonstrado que a ausência de inibidores das proteases cisteínicas em células de endosperma, permite que as proteases degradem as proteínas de armazenamento da semente, prejudicando a germinação e o desenvolvimento das plantas (SHEWRY & HALFORD 2002; MARTINEZ *et al.* 2009) Desta forma, justifica-se a necessidade de uma expressão gênica elevada das fitocistatinas nos estádios de maturação e manutenção da dormência das sementes. Entretanto, as proteases cisteínicas também estão diretamente relacionadas com os processos de morte celular na planta (SOLOMON *et al.* 1999). As proteases cisteínicas conhecidas como legumaínas são também expressas em tecidos vegetativos das plantas, além de sementes, e apresentam funções de processamento e degradação de outras proteínas em tecidos vegetativos, principalmente naqueles entrando em senescência, ativando proteases cisteínicas do tipo papaínas (MÜNTZ *et al.* 2002). Em arroz, o único gene com a capacidade potencial de inibir as proteases do tipo legumaínas seria a fitocistatina XII, baseado no fato de que suas proteínas ortólogas, em cevada e cana-de-açúcar, são as únicas descritas até o momento como inibidores endógenos de legumaínas, devido a sua extensão carboxi-terminal com o motivo proteico SNSL conservado (MARTINEZ *et al.* 2007; SANTOS-SILVA *et al.* 2012). Esse motivo proteico também é importante em algumas cistatinas de animais com capacidade de inibir legumaínas (ALVAREZ-FERNANDEZ *et al.* 1999). Assim, os resultados de expressão gênica obtidos revelam que para a germinação das plantas é importante que haja uma redução da expressão gênica de OcXII, aumentando a atividade das legumaínas que por sua vez auxiliam na ativação de papaínas para degradação das proteínas de armazenamento e remobilização de nutrientes para a planta. De forma similar, a

redução dos níveis dos inibidores Ocl, OclIII e OcVIII, na etapa de germinação e desenvolvimento da plântula, propicia o ambiente adequado para as papaínas exercerem sua função. Quanto à expressão destes inibidores em outros tecidos da planta adulta, é possível acreditar que haja a necessidade deles para a manutenção do nível de proteólise adequado para a situação fisiológica em que as plantas se encontram. Entretanto, existe a necessidade da caracterização dos alvos específicos de cada um destes inibidores de proteases para melhor entendimento da função dos mesmos.

5.2 Considerações evolutivas sobre as fitocistatinas relacionadas com seu perfil funcional em arroz

Os 204 genes identificados com os domínios cistatina conservados, nos 23 genomas de Viridiplantae e dois genomas de Chloropytha, possibilitaram a verificação do padrão evolutivo destas fitocistatinas (figura 17 e 18). A análise filogenética resultante pode ser dividida em quatro clados principais, onde no primeiro deles encontram-se os genes de espécies menos derivadas evolutivamente, *V. carteri* e *C. reinhardtii*, os quais mantêm relações de ancestralidade com os demais genes de Viridiplantae. Desta forma, como já observado em trabalhos anteriores utilizando diferentes metodologias de análise evolutiva, e um menor número de genes de diferentes espécies, as cistatinas destas duas espécies permanecem de forma basal aos demais genes analisados (MARTINEZ & DIAZ 2008). É válido ressaltar que as sequências proteicas dos genes destas duas Chloropyceae, não possuem a sequência consenso LARFAV, e estes genomas também não apresentam versões carboxi-extendidas de fitocistatinas. Este fato vem de encontro à afirmação feita em trabalhos anteriores de que as fitocistatinas possuem uma origem evolutiva comum, mantendo-se como um grupo filogeneticamente relacionado em Viridiplantae e diferente das cistatinas de animais (MARGIS *et al.* 1998; MARTINEZ & DIAZ 2008; KORDIS & TURK 2009).

As verdadeiras fitocistatinas (Viridiplantae) podem ser divididas em mais três grupos principais (grupos 2, 3 e 4; figuras 17 e 18). No grupo 2 encontram-se

sequências de dicotiledôneas (subgrupo 2.1), e de monocotiledôneas (subgrupo 2.2). Neste subgrupo 2.2 o posicionamento das fitocistatinas VI e VII indicam que uma delas é originária de um evento de duplicação recente deste *locus* no genoma de arroz. Tal afirmação é suportada também pela localização em tandem dos mesmos no cromossomo 3 (figura 6) (MARTÍNEZ *et al.* 2005). Neste mesmo *cluster* 2.2, também estão as sequências das fitocistatinas VIII e XI, mais estritamente relacionadas entre si, podendo ter uma origem comum também de uma duplicação mais ancestral. A proximidade das OcVI, OcVII, OcVIII e OcXI neste grupo 2 da filogenia corrobora com resultados prévios obtidos (MARGIS-PINHEIRO *et al.* 2008) e que também incluem neste grupo a OcIX, a qual não foi utilizada neste presente estudo por não apresentar o domínio cistatina conservado. Comparando-se os padrões evolutivos com a expressão gênica avaliada para estas quatro fitocistatinas, é possível verificar que apenas um dos genes de cada série de duplicação (OcVI e OcVIII) possui níveis de mRNA significativamente detectáveis, levando à inferência de que os outros dois genes (OcVII e OcXI) são prováveis pseudogenes oriundos de duplicações. As sequências das fitocistatinas IV, V e X estão agrupadas em ramos muito próximos do grupo 3, na filogenia (figura 18), como já observado em outros estudos (MARTÍNEZ *et al.* 2005; MARGIS-PINHEIRO *et al.* 2008), além de encontrarem-se em um grupo monofilético apenas de monocotiledôneas. A expressão gênica das Ocs IV e X, nos resultados de microarranjos, é baixa, porém o padrão de expressão é bastante similar (figura 11a). A OcV não é expressa. Assim, o posicionamento filogenético destes genes também infere que os mesmos originaram-se durante a diversificação do genoma de arroz, tendo surgido nesta planta provavelmente por eventos de duplicação gênica, sendo que as OcIV e OcX ainda têm funções mais relacionadas, enquanto que OcV é provavelmente um pseudogene originado de OcIV, devido ao posicionamento sequencial no cromossomo 1 de arroz (figura 6) (MARTÍNEZ *et al.* 2005), e ausência de expressão gênica.

O grupo 4 da análise filogenética (figuras 17 e 18) é onde se encontram as fitocistatinas de arroz que apresentaram o maior nível de expressão gênica neste estudo (OcI, OcIII e OcXII). Este grupo pode ser subdividido em cinco subgrupos,

conforme indicado nas figuras, e é possível verificar que as sequências de espécies mais basais (*S. moellendorffii* e *P. patens*), posicionam-se de forma externa às demais sequências de Viridiplantae, deste grupo em particular, formando o subgrupo 4.1. Desta forma, as sequências do grupo 4 apresentam uma relação de ancestralidade mais estreita, tendo provavelmente se originado na diversificação inicial das plantas, enquanto que as demais sequências do grupo 2 e 3 devem ter se originado nas diversificações internas de cada genoma, através de duplicações dos genes pré-existentes. No subgrupo 4.2 as sequências das fitocistatinas de arroz Ocl e Ocll encontram-se agrupadas em outros dois subclados compostos apenas por sequências de monocotiledôneas, indicando que as mesmas podem ter-se originado de uma duplicação ancestral ao surgimento das monocotiledôneas, e vêm divergindo separadamente desde então. A Oclll, localizada no subgrupo 4.4 também é clusterizada monofileticamente com outras sequências de monocotiledôneas, inferindo que este gene vêm sendo mantido desde a origem deste grupo. E, por fim, o *cluster* da OcXII (4.5) é onde agrupam-se todas as sequências com extensão carboxi-terminal de plantas, indicando que de fato, as sequências com este segmento adicional surgiram antes da grande diversificação de Viridiplantae e mantiveram-se nos genomas das plantas, sendo que em alguns genomas houveram duplicações e diversificações posteriores e independentes. Curiosamente, algumas fitocistatinas carboxi-extendidas de dicotiledôneas formaram um agrupamento separado, juntamente com outras sequências, das mesmas espécies, que não possuem a extensão C-terminal (subgrupo 4.5.2). Estas relações podem indicar que no(s) ancestral(is) das espécies de Viridiplantae originou-se a versão bi-funcional das fitocistatinas, através da duplicação e diversificação do domínio N-terminal de uma fitocistatina (MARTINEZ *et al.* 2007; MARTINEZ & DIAZ 2008; MARGIS-PINHEIRO *et al.* 2008). Durante as diversificações internas dos genomas de cada espécie, duplicações deste gene bi-funcional ocorreram nos genomas de algumas dicotiledôneas (subgrupo 4.5.2), sendo que também podem ter ocorrido duplicações apenas do domínio N-terminal, as quais mantêm-se agrupadas com as fitocistatinas carboxi-extendidas apenas pela similaridade de seus domínios N-terminais.

Sabendo-se que as fitocistatinas carboxi-extendidas possuem relações monofiléticas de ancestralidade, uma análise avaliando a sequência completa destas proteínas (N-terminal e C-terminal) foi realizada para melhor elucidação das relações evolutivas. Este resultado (figura 19) corrobora com a topologia filogenética obtida na análise da região N-terminal (figuras 17 e 18). A fitocistatina XII, no grupo 2, subgrupo 2.1 é monofileticamente relacionada com as demais sequências carboxi-extendidas de monocotiledôneas e também mantém relações de ancestralidade comum com as demais sequências de dicotiledôneas deste mesmo grupo 2. Interessantemente, algumas sequências com extensão C-terminal de dicotiledôneas continuaram formando um grupo a parte (grupo 3). No alinhamento das sequências deste grupo (figura 20), é possível verificar que a separação deste grupo deve-se a pelo menos treze sítios identificados com substituição de aminoácidos entre sequências do grupo 2 e 3. A análise filogenética obtida permite a postulação de duas principais hipóteses para o padrão evolutivo observado. Na primeira delas, as sequências que formam o grupo 3 podem ter-se originado posteriormente à diversificação individual dos genomas das dicotiledôneas por duplicação dos genes pré-existentes. A segunda hipótese, leva em consideração a presença de dois genes carboxi-extendidos nas espécies basais (grupo 1). Assim, o provável ancestral das espécies de Viridiplantae também possuiria duas sequências de fitocistatinas carboxi-extendidas. Durante o processo evolutivo das espécies, o ancestral que deu origem às monocotiledôneas deve ter perdido uma destas sequências, tendo em vista que os genomas das espécies atuais não mais apresentam uma segunda cópia deste gene. Por outro lado, o ancestral de dicotiledôneas manteve as duas cópias carboxi-extendidas, sendo uma delas mais divergente e apresentando substituições nucleotídicas, tais como na figura 20. Entretanto, esta segunda cópia foi mantida apenas nos genomas de algumas espécies, enquanto a mesma foi perdida em outras dicotiledôneas.

Cistatinas bi-funcionais também já foram reportadas em metazoários, além de plantas, porém, em metazoários elas não apresentam um domínio separado responsável pela função de inibição de proteases cisteínicas da família C13, como em plantas (ALVAREZ-FERNANDEZ *et al.* 1999; KORDIS & TURK 2009).

De forma geral, estudos envolvendo diversos genes de cistatinas revelam que o surgimento da superfamília das cistatinas ocorreu já no ancestral das linhagens eucariotas, sendo que elas passaram por processos evolutivos dinâmicos e complexos, através de diversas duplicações de genes e domínios (KORDIS & TURK 2009). As proteínas originárias deste evento inicial seguiram em evolução paralela entre os genomas de plantas e animais (BENCHABANE *et al.* 2010).

5.3 Silenciamento gênico da fitocistatina XII de arroz

A fitocistatina XII de arroz, por ser o único gene nesta planta com a provável capacidade inibidora bifuncional, de papaínas e legumaínas, foi escolhida como alvo para silenciamento gênico. Assim, neste trabalho, foram produzidas e parcialmente caracterizadas plantas transgênicas com o gene OcXII silenciado em nível de RNA mensageiro.

As plantas silenciadas obtidas apresentaram grande redução na expressão gênica de OcXII, sendo que quatro linhagens (L9, L10, L19 e L24) apresentaram uma redução nos níveis de mRNA de OcXII de pelo menos 90% (figura 24). A análise fenotípica destas plantas não resultou em grandes variações entre as plantas controle e as silenciadas, exceto quando se compara a altura das plantas e a viabilidade das sementes, as quais são significativamente reduzidas nas linhagens 19 e 24 (figura 30). Também existem algumas peculiaridades na L24, a qual apresenta menor massa da parte aérea, menor número de sementes por panícula e menor massa das sementes por planta. As sementes destas quatro linhagens de plantas compõem a T1 analisada no estudo, revelando que o silenciamento da OcXII continua específico, não afetando outros genes codificadores de fitocistatinas (figura 32A, B e C). Porém, dois genes de legumaínas mostraram-se alterados na T1 da L24 (figura 32 D e F). A legumaína 1 sofreu uma diminuição significativa, enquanto que a legumaína 3 teve seu nível de mRNA bastante elevado na L24, entretanto a leg 3 também aumentou sua expressão na L9. Tais alterações da expressão gênica das legumaínas podem estar indicando que exista um mecanismo de *feedback* na regulação de sua expressão, onde a ausência do inibidor destas proteases causou um desbalanço no processo regulatório, levando ao aumento ou diminuição transcricional de

algumas legumaínas. Entretanto, análises mais detalhadas em nível proteico, e também a verificação do efeito de posição e inserção do transgene no genoma destas plantas, precisam ser realizadas, uma vez que estas alterações não podem ser atribuídas ao nível de silenciamento, ou expressão diferencial do hairpin de silenciamento (figura 27).

Alguns estudos já foram realizados em outras plantas com os genes ortólogos à OcXII. Em *A. thaliana*, um primeiro estudo realizado com o gene AtCYSb (conhecido também como AtCYS6), demonstra que a super-expressão deste gene estimula o crescimento das plantas em condições adversas, tais como alta concentração de sal (ZHANG *et al.* 2008). Outros mutantes específicos para o gene AtCYS6 de *Arabidopsis* foram obtidos, de forma que foram analisadas plantas *knock-out* para este gene, e outras superexpressando o mesmo. Neste estudo realizado por HWANG *et al.* 2009, as principais conclusões indicam que a superexpressão de AtCYS6 diminui a velocidade de germinação e crescimento das plantas, além de atrasar o alongamento da raiz primária, enquanto que o silenciamento do mesmo aumenta a taxa de germinação, e a raiz primária cresce mais rápido. Ainda, experimentos medindo a atividade enzimática de proteases cisteínicas foram realizados com estas plantas mutantes. As plantas superexpressando AtCYS6 apresentaram uma leve diminuição na atividade das proteases cisteínicas, sendo que esta atividade tinha um aumento após a germinação das plantas. Contudo, este estudo realizado acima pode ter perdido algumas informações importantes sobre a importância funcional desta fitocistatina bi-funcional (AtCYS6) uma vez que, este gene não é único em *A. thaliana*. Uma outra cópia, AtCYS7, também possui os dois domínios cistatinas (MARTÍNEZ *et al.* 2005), e portanto, poderia estar compensando o fenótipo de silenciamento do gene AtCYS6 nas plantas. Ainda, para a medição da atividade proteica foi utilizado um substrato proteico específico para proteases cisteínicas do tipo papaínas (N α -Benzoyl-DL-arginine β -naphthylamide hydrochloride – BANA). Entretanto, o gene AtCYS6, assim como a OcXII, de arroz, são os únicos prováveis inibidores de proteases do tipo das legumaínas, então, seria adequado usar um substrato para medição da atividade destas proteases, especificamente. Mas, assim como em arroz a fitocistatina XII é um dos genes mais expressos

(figura 16) e o inibidor de proteases cisteínicas por ele codificado deve ser capaz de inibir, também, proteases do tipo papaínas, AtCYS6 é importante na inibição de papaínas, e como demonstrado, apesar da existência de outros genes com função específica de inibição de papaínas, a alteração de AtCYS6 leva ao desbalanço da atividade normal destas proteases, implicando em alterações na germinação e crescimento das plantas (HWANG *et al.* 2009). As linhagens transgênicas L9 e L10, por serem silenciadas para OcXII e não apresentarem alterações no nível de expressão das demais fitocistatinas, nem das legumaínas, além de também não possuírem diferenças morfológicas significativas em relação às plantas controle, tiveram sua germinação e desenvolvimento acompanhados paralelamente com plantas controle. Apesar destas alterações observadas em *A. thaliana*, o silenciamento da OcXII em arroz não acarretou em mudanças na taxa de germinação e desenvolvimento das plantas transformadas (dados não mostrados).

A caracterização das plantas silenciadas ainda deve ser concluída, levando-se em consideração principalmente os estáios de germinação e senescência, uma vez que as legumaínas podem estar mais envolvidas nestes processos, e estes genes serão confirmados como alvos da OcXII. O estudo do inibidor bi-funcional de Cana-de-açúcar CaneCPI-3 (GIANOTTI *et al.* 2006), possibilitou a posterior caracterização bioquímica do mesmo. Nesta caracterização foi possível obter-se a confirmação de que as legumaínas de cana são de fato inibidas por este gene com extensão carboxi-terminal, CaneCPI-3, e que a capacidade inibitória contra legumaínas é perdida quando a asparagina do sítio SNSL é mutada (SANTOS-SILVA *et al.* 2012). Já existem também indícios de que, apesar das fitocistatinas serem relacionadas com processos de defesa contra patógenos em plantas, a extensão carboxi-terminal dos genes de fitocistatinas, tais como a OcXII, não estaria envolvida em inibição de infecção por fungos, e não teria atividade inibitória contra papaínas, ao contrário da região amino-terminal destes inibidores, os quais possuem estas funções (WANG *et al.* 2008). Estudos de cristalização proteica com a cistatina carboxi-extendida de taro (*Colocasia esculenta*), revelaram que a região amino-terminal destes inibidores de proteases cisteínicas é capaz de interagir com papaínas, enquanto que apenas dois resíduos da região

carboxi-terminal foram detectados na estrutura fitocistatina-papaína (CHU *et al.* 2011). Entretanto, neste mesmo estudo foi revelada a existência de uma grande flexibilidade entre os domínios N-terminal e C-terminal, o que sugere que esta flexibilidade exista justamente para acomodação de proteases do tipo leguminaínas interagindo com o domínio C-terminal e papaínas interagindo com o N-terminal. Assim, grande parte dos estudos, até o momento, foram realizados com a região N-terminal destas fitocistatinas extendidas, por ser a região mais similar aos demais genes, deixando em aberto uma ampla área de investigação sobre a importância deste segundo domínio nas fitocistatinas, o qual foi mantido evolutivamente em todo o grupo Viridiplantae, identificado até o momento.

5.4 Fitocistatina XII de arroz e o estudo da sequência promotora de expressão gênica

A análise da região que antecede o início do sítio de transcrição do gene da OcXII possibilitou a identificação de *cis-elementos* que podem atuar na regulação da expressão gênica da fitocistatina XII, em diferentes situações fisiológicas. Os elementos encontrados são relacionados principalmente com respostas à ABA, anaerobiose, resistência a doenças, seca, giberilina, luz, além de outros elementos indutores de expressão gênica específica de endosperma e outros encontrados em regiões promotoras de proteases cisteínicas (tabelas 11 e 12). Alguns destes elementos também foram encontrados no promotor do gene similar a OcXII em trigo, como por exemplo elementos responsivos à luz, ABA, infecção por fungos, específicos de endosperma, além de outros *cis-elementos* como aqueles envolvidos em transcrição meristemática, responsividade a metil-jasmonato e ácido salicílico (DUTT *et al.* 2010). A presença de alguns destes *cis-elementos*, na região promotora de fitocistatinas carboxi-extendidas parece estar evolutivamente relacionada entre as plantas. Em cana-de-açúcar, o gene CaneCPI-3 é altamente responsivo ao ABA, tendo seus níveis de mRNA significativamente aumentados, enquanto que tratamentos com giberilina (GA), ácido salicílico (SA), e ácido jasmônico (JA), não desencadearam alterações na expressão gênica desta fitocistatina (SANTOS-SILVA *et al.* 2012)

Através dos estudos de expressão gênica com o promotor da fitocistatina XII fusionado ao vetor de expressão, controlando os genes repórteres *gus* e *egfp*, foi possível a observação da alta expressão deste gene em calos de arroz de uma das linhagens obtidas (figura 34). Em uma última análise realizada neste trabalho, pode ser confirmada a responsividade do promotor do gene OcXII ao aumento de ABA, levando à uma maior transcrição deste gene, como parte da resposta da planta (figura 35). Neste experimento é possível verificar-se que há um aumento significativo da expressão gênica após um período de quatro horas de tratamento com 100 μ M de ABA, tanto em raízes quanto nas folhas da planta. Tendo em vista que o tratamento foi aplicado na solução em contato com as raízes das plantas, é interessante verificar a cinética de aumento da expressão gênica, sendo que este aumento é mais rápido em raízes, e mais lento em folhas. Da mesma forma, 24 horas após o início do tratamento, os níveis de mRNA da fitocistatina XII estão voltando ao seu estado basal, mais rapidamente em raízes do que em folhas.

Dois outros estudos referentes à localização da expressão de genes ortólogos à OcXII, e resposta a fitormônios, foram realizados em *A. thaliana*. No primeiro estudo, foi verificada a presença de *cis-elementos* responsivos a fitormônios relacionados à germinação e desenvolvimento das plantas. Assim plantas transgênicas utilizando o promotor do gene AtCYS6 fusionado com o gene repórter *gus* foram obtidas (HWANG *et al.* 2009). Primeiramente, neste estudo, foi verificada a expressão predominante do gene AtCYS6 em cotilédones de plantas e tecidos meristemáticos de raiz. Quando estas plantas eram tratadas com GA observou-se que o nível de expressão do gene *gus* diminuía nos três primeiros dias de germinação, assim como nas plantas controle, enquanto que nas plantas tratadas com ABA, a expressão de *gus* era detectada até 12 dias após o tratamento, inclusive retardando a germinação e o desenvolvimento da plântula (HWANG *et al.* 2009). Outro estudo realizado com a fitocistatina carboxi-extendida AtCYS6 frente a diferentes estresses abióticos demonstrou que a expressão deste gene é elevada em tratamentos com 100 mM de NaCl, é pouco induzida em tratamentos de seca com 500 mM de manitol e não responde ao estresse oxidativo induzido por paraquat, mas aumenta quando o estresse oxidativo é desencadeado por H₂O₂. Em baixas temperaturas a expressão de

AtCYS6 aumenta, sendo que o nível de seu mRNA atinge um pico no período de 24 horas após o início do tratamento. Não foram visualizadas alterações da expressão gênica em folhas tratadas com sal, seca ou frio, mas houve aumento da expressão gênica em raízes submetidas às condições de frio e aumento da concentração de ABA (ZHANG *et. al*, 2008).

De forma geral, a expressão da fitocistatina XII, assim como dos demais genes com extensão C-terminal em outras plantas, tem grande relevância nos processos de germinação, desenvolvimento e resposta a estresses, desencadeadas das plantas, como demonstrado pelos experimentos realizados, principalmente com ácido abscísico, e pela presença de *cis-elementos* que podem induzir a transcrição do gene especificamente no endosperma das sementes. Entretanto, as diferentes plantas estudadas apresentam variações quanto ao padrão de resposta aos tratamentos analisados, por isto, é necessária uma avaliação mais específica em arroz, para que possam ser formuladas hipóteses concretas para o padrão de expressão gênica da fitocistatina XII.

6. CONCLUSÕES E PERSPECTIVAS

Neste trabalho foi realizada a caracterização dos doze genes que codificam fitocistatinas em arroz (*Oryza sativa*), de acordo com as estruturas gênicas e proteicas, além da elucidação do padrão de expressão gênica, em diferentes tecidos e estádios de desenvolvimento das plantas. Também foram realizadas análises filogenéticas com as sequências proteicas de fitocistatinas encontradas em plantas que possuem o seu genoma sequenciado, a fim de contribuir para o conhecimento evolutivo desta família gênica, com enfoque principal nas fitocistatinas de arroz. Estes estudos serviram de base para a contextualização evolutiva e funcional da fitocistatina XII de arroz, a qual foi utilizada em análises funcionais mediante a obtenção de plantas silenciadas para este gene, e expressando o promotor do mesmo, fusionado com genes repórteres.

Os genes das fitocistatinas de arroz, em termos estruturais, são bastante similares, exceção feita à OcXII que apresenta uma extensão C-terminal na sua

sequência, que codifica um domínio proteico similar a cistatina, contendo o sítio conservado SNSL, com provável capacidade de inibir legumainas.

As análises de expressão gênica revelaram que os genes mais ubiquamente expressos em arroz são Ocl, OclIII e OcXII, enquanto que os demais apresentam algumas discrepâncias entre os métodos utilizados para detecção da expressão, provavelmente devido ao seu nível mais reduzido de mRNA ou padrão mais tecido-específico de transcrição. Apenas a OcVII não teve sua expressão gênica detectada por nenhuma metodologia.

Com os métodos filogenéticos empregados foi possível verificar o padrão evolutivo das fitocistatinas de arroz, de acordo com a sequência evolutiva das demais Viridiplantae. Assim, todos os genes de fitocistatinas de plantas apresentam uma origem monofilética, sendo que os genes Ocl, OclII, OclIII e OcXII permanecem em um mesmo clado evolutivo, onde OcXII encontra-se agrupada com os demais genes carboxi-extendidos de outras plantas. Esta observação sugere que as fitocistatinas com extensão C-terminal vem sendo evolutivamente mantidas desde os primórdios da origem evolutiva das Viridiplantae. Os demais genes de fitocistatinas de arroz apresentam um padrão mais divergente, o que demonstra a possibilidade destes terem se originado independentemente dentro dos genomas individuais de cada espécie de planta, através de processos de duplicações gênicas. Esta hipótese também pode ser corroborada ao ser confrontada com os dados de expressão gênica, uma vez que os genes mais expressos, e provavelmente sob uma maior pressão de seleção evolutiva são os genes Ocl, OclIII e OcXII.

O silenciamento gênico da OcXII, não resultou em alterações fenotípicas significativas para as plantas em um primeiro momento, mas as perspectivas de continuidade deste trabalho são baseadas em estudo e caracterização das linhagens transgênicas obtidas. Principalmente em relação à caracterização da alteração da atividade das proteases-alvo de OcXII, as legumainas. Além de identificar outras possíveis proteases que interagem com o domínio C-terminal da OcXII, e que estejam sofrendo alterações devido ao silenciamento deste gene nas plantas. Existe também a necessidade da caracterização individual das linhagens de plantas silenciadas quanto ao número de cópias inseridas no genoma das

plantas, e também a determinação do local de inserção do transgene no genoma das plantas. Estão previstos estudos de superexpressão do gene da fitocistatina XII, para avaliação fenotípica, avaliação de alterações na expressão gênica global, em paralelo com as plantas silenciadas, e também análise de proteases-alvo alteradas nas plantas. Ainda, correlacionando o domínio C-terminal da fitocistatina XII com a inibição de proteases cisteínicas do tipo legumainas, deverão ser realizados estudos de caracterização destes genes nas plantas, tanto em nível de expressão gênica, quanto também em nível evolutivo, procurando por possíveis relações funcionais. Estes genes de legumainas poderão ter sua função caracterizada nas plantas através de experimentos de silenciamento ou superexpressão de forma transiente em plantas de arroz.

As análises envolvendo os locais de expressão do gene, baseado em estudos do promotor de OcXII deverão ter continuidade, a fim de avaliar o padrão de expressão do gene em tecidos de plantas adultas. Além da possibilidade de confirmação dos dados de ativação e transcrição gênica em tecidos específicos, comparando-se com os estudos de microarranjos e RTqPCR. Também será possível verificar como o gene da OcXII se comporta frente ao estímulo por estresses ambientais, simulados por tratamentos específicos, como o ácido abscísico, giberilina, ácido jasmônico, ácido salicílico, excesso de sal, seca, calor, entre outros.

7. REFERÊNCIAS

- ABE, K., KONDO, H. AND ARAI, S. (1987). Purification and characterization of a rice cysteine proteinase inhibitor. *Agric. Biol. Chem.* **51**,2763-2768
- ABE, K., EMORI, Y., KONDO, H., ARAI, S. & SUZUKI, K. (1988) The NH₂-terminal 21 Amino Acid Residues Are Not Essential for the Papain-inhibitory Activity of Oryzacystatin , a Member of the Cystatin Superfamily. *The Journal of biological chemistry*, **263**, 7655–7659.
- ABE, K., EMORI, Y., KONDO, H., SUZUKI, K. & ARAI, S. (1987a) Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). Homology with animal cystatins and transient expression in the ripening process of rice. *Journal of Biological Chemistry*, **262**, 16793–16797.
- ABE, K., KONDO, H. & ARAI, S. (1987b) Purification and characterization of a rice cysteine proteinase inhibitor. *Agricultural and biological chemistry*, **51**, 2763–2768.
- ABE, Y., SHIRANE, K., YOKOSAWA, H., MATSUSHITA, H., MITTA, M., KATO, I. & ISHII, S. (1993) Asparaginyl Endopeptidase of Jack Bean Seeds. *Biological Chemistry*, **268**, 3525–3529.
- ALVAREZ-FERNANDEZ, M., BARRETT, A J., GERHARTZ, B., DANDO, P.M., NI, J. & ABRAHAMSON, M. (1999) Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. *The Journal of biological chemistry*, **274**, 19195–203.
- AOKI, H., AKAIKE, T., ABE, K., KURODA, M., ARAI, S., OKAMURA, R., NEGI, A. & MAEDA, H. (1995) Antiviral effect of oryzacystatin, a proteinase inhibitor in rice, against herpes simplex virus type 1 in vitro and in vivo. *Antimicrobial agents and chemotherapy*, **39**, 846–9.
- ARAI, S., MATSUMOTO, I. & ABE, K. (1998) Phytocystatins and their target enzymes: from molecular biology to practical application: a review. *Journal of food biochemistry*, **22**, 287–299.
- ARAI, S., WATANABE, H. & KONDO, H. (1991) Papain activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *Journal of biochemistry*, **109**, 294–298.
- BECKER, C., SHUTOV, A D., NONG, V.H., SENYUK, V.I., JUNG, R., HORSTMANN, C., FISCHER, J., NIELSEN, N.C. & MÜNTZ, K. (1995) Purification, cDNA cloning and characterization of proteinase B, an asparagine-specific endopeptidase from germinating vetch (*Vicia sativa* L.) seeds. *European journal of biochemistry / FEBS*, **228**, 456–62.
- BEERS, E.P., JONES, A.M. & DICKERMAN, A.W. (2004) The S8 serine, C1A cysteine and A1 aspartic protease families in Arabidopsis. *Phytochemistry*, **65**, 43–58.

- BEERS, E.P., WOFFENDEN, B.J. & ZHAO, C. (2000) Plant proteolytic enzymes: possible roles during programmed cell death. *Plant molecular biology*, **44**, 399–415.
- BENCHABANE, M., SCHLÜTER, U., VORSTER, J., GOULET, M.-C. & MICHAUD, D. (2010) Plant cystatins. *Biochimie*, **92**, 1657–66.
- BJÖRK, I., POL, E., RAUB-SEGALL, E., ABRAHAMSON, M., ROWAN, A. D. & MORT, J.S. (1994) Differential changes in the association and dissociation rate constants for binding of cystatins to target proteinases occurring on N-terminal truncation of the inhibitors indicate that the interaction mechanism varies with different enzymes. *The Biochemical journal*, **299** (Pt 1, 219–25.
- BODE, W., ENGH, R., MUSIL, D., THIELE, U., HUBER, R., KARSHIKOV, A., BRZIN, J., KOS, J. & TURK, V. (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *The EMBO journal*, **7**, 2593.
- BOLTER, C.J. (1993) Methyl Jasmonate Induces Papain Inhibitor(s) in Tomato Leaves. *Plant physiology*, **103**, 1347–1353.
- BONNEAU, L., GE, Y., DRURY, G.E. & GALLOIS, P. (2008) What happened to plant caspases? *Journal of experimental botany*, **59**, 491–9.
- BOTELLA, M. A, XU, Y., PRABHA, T.N., ZHAO, Y., NARASIMHAN, M.L., WILSON, K. A, NIELSEN, S.S., BRESSAN, R. A & HASEGAWA, P.M. (1996) Differential expression of soybean cysteine proteinase inhibitor genes during development and in response to wounding and methyl jasmonate. *Plant physiology*, **112**, 1201–10.
- CARRILLO, L., HERRERO, I., CAMBRA, I., SÁNCHEZ-MONGE, R., DIAZ, I. & MARTINEZ, M. (2011a) Differential in vitro and in vivo effect of barley cysteine and serine protease inhibitors on phytopathogenic microorganisms. *Plant physiology and biochemistry: PPB / Societe francaise de physiologie vegetale*, 1–10.
- CARRILLO, L., MARTINEZ, M., RAMESSAR, K., CAMBRA, I., CASTAÑERA, P., ORTEGO, F. & DÍAZ, I. (2011b) Expression of a barley cystatin gene in maize enhances resistance against phytophagous mites by altering their cysteine-proteases. *Plant cell reports*, **30**, 101–12.
- CHAN, Y.-L., YANG, A.-H., CHEN, J.-T., YEH, K.-W. & CHAN, M.-T. (2010) Heterologous expression of taro cystatin protects transgenic tomato against *Meloidogyne incognita* infection by means of interfering sex determination and suppressing gall formation. *Plant cell reports*, **29**, 231–8.
- CHANG, W.-C., LEE, T.-Y., HUANG, H.-D., HUANG, H.-Y. & PAN, R.-L. (2008) PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC genomics*, **9**, 561.
- CHEN, M.S., JOHNSON, B., WEN, L., MUTHUKRISHNAN, S., KRAMER, K.J., MORGAN, T.D. & REECK, G.R. (1992) Rice cystatin: bacterial expression, purification, cysteine proteinase inhibitory activity, and insect growth

- suppressing activity of a truncated form of the protein. *Protein expression and purification*, **3**, 41–9.
- CHEN, J., RAWLINGS, N. & STEVENS, R. (1998) Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases. *FEBS letters*, **441**, 361–365.
- CHU, M.-H., LIU, K.-L., WU, H.-Y., YEH, K.-W. & CHENG, Y.-S. (2011) Crystal structure of tarocystatin-papain complex: implications for the inhibition property of group-2 phytocystatins. *Planta*, **234**, 243–54.
- DRUMMOND, A.J., SUCHARD, M. A, XIE, D. & RAMBAUT, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution*, **29**, 1969–1973.
- DUTT, S., SINGH, V.K., MARLA, S.S. & KUMAR, A. (2010) In silico analysis of sequential, structural and functional diversity of wheat cystatins and its implication in plant defense. *Genomics, proteomics & bioinformatics / Beijing Genomics Institute*, **8**, 42–56.
- EDGAR, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, **32**, 1792–7.
- FARMER, E., JOHNSON, R. & RYAN, C. (1992) Regulation of expression of proteinase inhibitor genes by methyl jasmonate and jasmonic acid. *Plant Physiology*, 995–1002.
- FISCHER, J., BECKER, C., HILLMER, S., HORSTMANN, C., NEUBOHN, B., SCHLERETH, A., SENYUK, V., SHUTOV, A. & MÜNTZ, K. (2000) The families of papain- and legumain-like cysteine proteinases from embryonic axes and cotyledons of *Vicia* seeds: developmental patterns, intracellular localization and functions in globulin proteolysis. *Plant molecular biology*, **43**, 83–101.
- GADDOUR, K., VICENTE-CARBAJOSA, J., LARA, P., ISABEL-LAMONEDA, I., DÍAZ, I. & CARBONERO, P. (2001) A constitutive cystatin-encoding gene from barley (*Icy*) responds differentially to abiotic stimuli. *Plant molecular biology*, **45**, 599–608.
- GALE, M.D. (1998) Plant Comparative Genetics after 10 Years. *Science*, **282**, 656–659.
- GIANOTTI, A., RIOS, W.M., SOARES-COSTA, A., NOGAROTO, V., CARMONA, A.K., OLIVA, M.L.V., ANDRADE, S.S. & HENRIQUE-SILVA, F. (2006) Recombinant expression, purification, and functional analysis of two novel cystatins from sugarcane (*Saccharum officinarum*). *Protein expression and purification*, **47**, 483–9.
- GOFF, S. A, RICKE, D., LAN, T.-H., PRESTING, G., WANG, R., DUNN, M., *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science (New York, N.Y.)*, **296**, 92–100.
- GOULET, M.-C., DALLAIRE, C., VAILLANCOURT, L.-P., KHALF, M., BADRI, A.M., PRERADOV, A., DUCEPPE, M.-O., GOULET, C., CLOUTIER, C. & MICHAUD, D. (2008) Tailoring the specificity of a plant cystatin toward

- herbivorous insect digestive cysteine proteases by single mutations at positively selected amino acid sites. *Plant physiology*, **146**, 1010–9.
- GRUDKOWSKA, M. & ZAGDAŃSKA, B. (2004) Multifunctional role of plant cysteine proteinases. *Acta biochimica Polonica*, **51**, 609–24.
- GUTIERREZ-CAMPOS, R., TORRES-ACOSTA, J. A., SAUCEDO-ARIAS, L.J. & GOMEZ-LIM, M. A. (1999) The use of cysteine proteinase inhibitors to engineer resistance against potyviruses in transgenic tobacco plants. *Nature biotechnology*, **17**, 1223–6.
- HARA-NISHIMURA, I., INOUE, K. & NISHIMURA, M. (1991) A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms. *FEBS letters*, **294**, 89–93.
- HARA-NISHIMURA, I., KINOSHITA, T., HIRAIWA, N. & NISHIMURA, M. (1998) Vacuolar processing enzymes in protein-storage vacuoles and lytic vacuoles. *Journal of Plant Physiology*, **152**, 668–674.
- HARA-NISHIMURA, I. & NISHIMURA, M. (1987) Proglobulin processing enzyme in vacuoles isolated from developing pumpkin cotyledons. *Plant physiology*, **85**, 440–5.
- HARA-NISHIMURA, I., SHIMADA, T., HIRAIWA, N. & NISHIMURA, M. (1995) Vacuolar Processing Enzyme Responsible for Maturation of Seed Proteins. *Journal of Plant Physiology*, **145**, 632–640.
- HARA-NISHIMURA, I., TAKEUCHI, Y. & NISHIMURA, M. (1993) Molecular characterization of a vacuolar processing enzyme related to a putative cysteine proteinase of *Schistosoma mansoni*. *The Plant cell*, **5**, 1651–9.
- HATSUGAI, N., KUROYANAGI, M., YAMADA, K., MESHII, T., TSUDA, S., KONDO, M., NISHIMURA, M. & HARA-NISHIMURA, I. (2004) A plant vacuolar protease, VPE, mediates virus-induced hypersensitive cell death. *Science (New York, N.Y.)*, **305**, 855–8.
- HAYASHI, Y., YAMADA, K., SHIMADA, T., MATSUSHIMA, R., NISHIZAWA, N.K., NISHIMURA, M. & HARA-NISHIMURA, I. (2001) A proteinase-storing body that prepares for cell death or stresses in the epidermal cells of *Arabidopsis*. *Plant & cell physiology*, **42**, 894–9.
- HIRAIWA, N., NISHIMURA, M. & HARA-NISHIMURA, I. (1999) Vacuolar processing enzyme is self-catalytically activated by sequential removal of the C-terminal and N-terminal propeptides. *FEBS letters*, **447**, 213–6.
- HIROTO, K., YASUFUMI, E., KEIKO, A. & KOICHI, S. (1989) Cloning and sequence analysis of the genomic DNA fragment encoding oryzacystatin. *Gene*, **81**, 259–265.
- VAN DER HOORN, R. A. L. & JONES, J.D.G. (2004) The plant proteolytic machinery and its role in defence. *Current opinion in plant biology*, **7**, 400–7.
- HRUZ, T., LAULE, O., SZABO, G., WESSENDORP, F., BLEULER, S., OERTLE, L., WIDMAYER, P., GRUISSEM, W. & ZIMMERMANN, P. (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Advances in bioinformatics*, **2008**, 420747.

- HUANG, Y.-J., TO, K.-Y., YAP, M.-N., CHIANG, W.-J., SUEN, D.-F. & CHEN, S.-C.G. (2001) Cloning and characterization of leaf senescence up-regulated genes in sweet potato. *Physiologia plantarum*, **113**, 384–391.
- HWANG, J.E., HONG, J.K., JE, J.H., LEE, K.O., KIM, D.Y., LEE, S.Y. & LIM, C.O. (2009) Regulation of seed germination and seedling growth by an Arabidopsis phytocystatin isoform, AtCYS6. *Plant cell reports*, **28**, 1623–32.
- HWANG, J.E., HONG, J.K., LIM, C.J., CHEN, H., JE, J., YANG, K.A., KIM, D.Y., CHOI, Y.J., LEE, S.Y. & LIM, C.O. (2010) Distinct expression patterns of two Arabidopsis phytocystatin genes, AtCYS1 and AtCYS2, during development and abiotic stresses. *Plant cell reports*, **29**, 905–15.
- IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.
- ISHII, S.I. (1994) Legumain: Asparaginyl endopeptidase. *Methods in enzymology*, **244**, 604–615.
- IZAWA, T. & SHIMAMOTO, K. (1996) Becoming a model plant: The importance of rice to plant science. *Trends in Plant Science*, **1**, 95–99.
- JAIN, M., NIJHAWAN, A., TYAGI, A.K. & KHURANA, J.P. (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochemical and biophysical research communications*, **345**, 646–51.
- JUNG, K.-H., DARDICK, C., BARTLEY, L.E., CAO, P., PHETSOM, J., CANLAS, P., SEO, Y.-S., SHULTZ, M., OUYANG, S., YUAN, Q., FRANK, B.C., LY, E., ZHENG, L., JIA, Y., HSIA, A.-P., AN, K., CHOU, H.-H., ROCKE, D., LEE, G.C., SCHNABLE, P.S., AN, G., BUELL, C.R. & RONALD, P.C. (2008) Refinement of light-responsive transcript lists using rice oligonucleotide arrays: evaluation of gene-redundancy. *PLoS one*, **3**, e3337.
- KARIMI, M., INZÉ, D. & DEPICKER, A. (2002) GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in plant science*, **7**, 193–195.
- KATO, H., SUTOH, K. & MINAMIKAWA, T. (2003) Identification, cDNA cloning and possible roles of seed-specific rice asparaginyl endopeptidase, REP-2. *Planta*, **217**, 676–85.
- KHANNA-CHOPRA, R., SRIVALLI, B. & AHLAWAT, Y.S. (1999) Drought induces many forms of cysteine proteases not observed during natural senescence. *Biochemical and biophysical research communications*, **255**, 324–7.
- KINOSHITA, T., YAMADA, K., HIRAIWA, N., KONDO, M., NISHIMURA, M. & HARA-NISHIMURA, I. (1999) Vacuolar processing enzyme is up-regulated in the lytic vacuoles of vegetative tissues during senescence and under various stressed conditions. *The Plant journal : for cell and molecular biology*, **19**, 43–53.
- KOIWA, H., D'URZO, M.P., ASSFALG-MACHLEIDT, I., ZHU-SALZMAN, K., SHADE, R.E., AN, H., MURDOCK, L.L., MACHLEIDT, W., BRESSAN, R. A & HASEGAWA, P.M. (2001) Phage display selection of hairpin loop

- soyacystatin variants that mediate high affinity inhibition of a cysteine proteinase. *The Plant journal : for cell and molecular biology*, **27**, 383–91.
- KONDO, H., ABE, K., EMORI, Y. & ARAI, S. (1991) organization of oryzacystatin-II, a new cystatin superfamily member of plant origin, is closely related to that of oryzacystatin-I but different from those of animal cystatins. *FEBS letters*, **278**, 87–90.
- KONDO, H., ABE, K., NISHIMURA, I., WATANABE, H., EMORI, Y. & ARAI, S. (1990) Two Distinct Cystatin Species in Rice Seeds with Different Specificities against Cysteine Proteinases. *Journal of Biological Chemistry*, **265**, 15832–15837.
- KONNO, K., HIRAYAMA, C., NAKAMURA, M., TATEISHI, K., TAMURA, Y., HATTORI, M. & KOHNO, K. (2003) Papain protects papaya trees from herbivorous insects: role of cysteine proteases in latex. *The Plant Journal*, **37**, 370–378.
- KORDIS, D. & TURK, V. (2009) Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes. *BMC evolutionary biology*, **9**, 266.
- KUDO, N., NISHIYAMA, M., SASAKI, H., ABE, K., ARAI, S. & TANOKURA, M. (1998) Crystallization and preliminary X-ray diffraction studies of a rice cysteine proteinase inhibitor, oryzacystatin-I. *Journal of biochemistry*, **123**, 568.
- KUROYANAGI, M., NISHIMURA, M. & HARA-NISHIMURA, I. (2002) Activation of Arabidopsis vacuolar processing enzyme by self-catalytic removal of an auto-inhibitory domain of the C-terminal propeptide. *Plant & cell physiology*, **43**, 143–51.
- LALITHA, S., SHADE, R.E., MURDOCK, L.L., BRESSAN, R. A, HASEGAWA, P.M. & NIELSEN, S.S. (2005) Effectiveness of recombinant soybean cysteine proteinase inhibitors against selected crop pests. *Comparative biochemistry and physiology. Toxicology & pharmacology : CBP*, **140**, 227–35.
- LANGMEAD, B., TRAPNELL, C., POP, M. & SALZBERG, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*, **10**, R25.
- LEPELLEY, M., AMOR, M.B., MARTINEAU, N., CHEMINADE, G., CAILLET, V. & MCCARTHY, J. (2012) Coffee cysteine proteinases and related inhibitors with high expression during grain maturation and germination. *BMC Plant Biology*, **12**, 31.
- LESCOT, M., DÉHAIS, P., THIJS, G., MARCHAL, K., MOREAU, Y., VAN DE PEER, Y., ROUZÉ, P. & ROMBAUTS, S. (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research*, **30**, 325–7.
- LETUNIC, I., DOERKS, T. & BORK, P. (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic acids research*, **40**, D302–5.
- MACHLEIDT, W., THIELE, U., LABER, B., ASSFALG-MACHLEIDT, I., ESTERL, A., WIEGAND, G., KOS, J., TURK, V. & BODE, W. (1989) Mechanism of

- inhibition of papain by chicken egg white cystatin. Inhibition constants of N-terminally truncated forms and cyanogen bromide fragments of the inhibitor. *FEBS letters*, **243**, 234–8.
- MARGIS, R., REIS, E.M. & VILLERET, V. (1998) Structural and phylogenetic relationships among plant and animal cystatins. *Archives of biochemistry and biophysics*, **359**, 24–30.
- MARGIS-PINHEIRO, M., ZOLET, A.C.T., LOSS, G., PASQUALI, G. & MARGIS, R. (2008) Molecular evolution and diversification of plant cysteine proteinase inhibitors: new insights after the poplar genome. *Molecular phylogenetics and evolution*, **49**, 349–55.
- MARTINEZ, M., ABRAHAM, Z., GAMBARDELLA, M., ECHAIDE, M., CARBONERO, P. & DIAZ, I. (2005) The strawberry gene *Cyf1* encodes a phytocystatin with antifungal properties. *Journal of experimental botany*, **56**, 1821–9.
- MARTINEZ, M., CAMBRA, I., CARRILLO, L., DIAZ-MENDOZA, M. & DIAZ, I. (2009) Characterization of the entire cystatin gene family in barley and their target cathepsin L-like cysteine-proteases, partners in the hordein mobilization during seed germination. *Plant physiology*, **151**, 1531–45.
- MARTINEZ, M. & DIAZ, I. (2008) The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. *BMC evolutionary biology*, **8**, 198.
- MARTINEZ, M., DIAZ-MENDOZA, M., CARRILLO, L. & DIAZ, I. (2007) Carboxy terminal extended phytocystatins are bifunctional inhibitors of papain and legumain cysteine proteinases. *FEBS letters*, **581**, 2914–8.
- MARTÍNEZ, M., ABRAHAM, Z., CARBONERO, P. & DÍAZ, I. (2005) Comparative phylogenetic analysis of cystatin gene families from arabidopsis, rice and barley. *Molecular genetics and genomics: MGG*, **273**, 423–32.
- MIKI, D. & SHIMAMOTO, K. (2004) Simple RNAi vectors for stable and transient suppression of gene function in rice. *Plant & cell physiology*, **45**, 490–5.
- MILNE, I., BAYER, M., CARDLE, L., SHAW, P., STEPHEN, G., WRIGHT, F. & MARSHALL, D. (2010) Tablet--next generation sequence assembly visualization. *Bioinformatics (Oxford, England)*, **26**, 401–2.
- MITTLER, R. (2006) Abiotic stress, the field environment and stress combination. *Trends in plant science*, **11**, 15–9.
- MÜNTZ, K., BLATTNER, F.R. & SHUTOV, A.D. (2002) Legumains-a family of asparagine-specific cysteine endopeptidases involved in propolypeptide processing and protein breakdown in plants. *Journal of plant physiology*, **159**, 1281–1293.
- MÜNTZ, K. & SHUTOV, A.D. (2002) Legumains and their functions in plants. *Trends in Plant Science*, **7**, 340–344.
- NAGATA, K., KUDO, N., ABE, K. & ARAI, S. (2000) Three-dimensional solution structure of oryzacystatin-I, a cysteine proteinase inhibitor of the rice, *Oryza sativa* L. *japonica*. *Biochemistry*, **39**, 14753–14760.

- OHTSUBO, S., KOBAYASHI, H., NORO, W., TANIGUCHI, M. & SAITOH, E. (2005) Molecular cloning and characterization of oryzacystatin-III, a novel member of phytocystatin in rice (*Oryza sativa* L. *japonica*). *Journal of agricultural and food chemistry*, **53**, 5218–24.
- OKAMOTO, T. & MINAMIKAWA, T. (1999) Molecular cloning and characterization of Vigna mungo processing enzyme 1 (VmPE-1), an asparaginyl endopeptidase possibly involved in post-translational processing of a vacuolar cysteine endopeptidase (SH-EP). *Plant molecular biology*, **39**, 63–73.
- OPPERT, B., MORGAN, T.D., HARTZER, K., LENARCIC, B., GALESA, K., BRZIN, J., TURK, V., YOZA, K., OHTSUBO, K. & KRAMER, K.J. (2003) Effects of proteinase inhibitors on digestive proteinases and growth of the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Comparative biochemistry and physiology. Toxicology & pharmacology: CBP*, **134**, 481–90.
- OUYANG, S., ZHU, W., HAMILTON, J., LIN, H., CAMPBELL, M., CHILDS, K., THIBAUD-NISSEN, F., MALEK, R.L., LEE, Y., ZHENG, L., ORVIS, J., HAAS, B., WORTMAN, J. & BUELL, C.R. (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic acids research*, **35**, D883–7.
- PERNAS, M., SÁNCHEZ-MONGE, R. & SALCEDO, G. (2000) Biotic and abiotic stress can induce cystatin expression in chestnut. *FEBS letters*, **467**, 206–10.
- PIROVANI, C.P., DA SILVA SANTIAGO, A., DOS SANTOS, L.S., MICHELI, F., MARGIS, R., DA SILVA GESTEIRA, A., ALVIM, F.C., PEREIRA, G.A.G. & DE MATTOS CASCARDO, J.C. (2010) Theobroma cacao cystatins impair *Moniliophthora perniciosa* mycelial growth and are involved in postponing cell death symptoms. *Planta*, **232**, 1485–97.
- QUEVILLON, E., SILVENTOINEN, V., PILLAI, S., HARTE, N., MULDER, N., APWEILER, R. & LOPEZ, R. (2005) InterProScan: protein domains identifier. *Nucleic acids research*, **33**, W116–20.
- RAMBALDI, D. & CICCARELLI, F.D. (2009) FancyGene: dynamic visualization of gene structures and protein domain architectures on genomic loci. *Bioinformatics (Oxford, England)*, **25**, 2281–2.
- RAMBAUT, A. (2008) FigTree v1.1.1: Tree figure drawing tool. Available: <http://tree.bio.ed.ac.uk/software/figtree/>. Accessed 2 June 2012.
- RICHAU, K.H., KASCHANI, F., VERDOES, M., PANSURIYA, T.C., NIESSEN, S., STÜBER, K., COLBY, T., OVERKLEEF, H.S., BOGYO, M. & VAN DER HOORN, R.A.L. (2012) Subclassification and Biochemical Analysis of Plant Papain-like Cysteine Proteases displays Subfamily-specific Characteristics. *Plant physiology*, **158**, 1583–1599.
- ROJO, E., MARTÍN, R., CARTER, C., ZOUHAR, J., PAN, S., PLOTNIKOVA, J., JIN, H., PANEQUE, M., SÁNCHEZ-SERRANO, J.J., BAKER, B., AUSUBEL, F.M. & RAIKHEL, N.V. (2004) VPEgamma exhibits a caspase-like activity that contributes to defense against pathogens. *Current biology*, **14**, 1897–906.

- SAMAC, D. A & SMIGOCKI, A.C. (2003) Expression of Oryzacystatin I and II in Alfalfa Increases Resistance to the Root-Lesion Nematode. *Phytopathology*, **93**, 799–804.
- SAMBROOK, J. & RUSSEL, D.W. (2001) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, ISBN 0-87969-577-3, New York.
- SANTOS-SILVA, L.K., SOARES-COSTA, A., GERALD, L.T.S., MENEGHIN, S.P. & HENRIQUE-SILVA, F. (2012) Recombinant expression and biochemical characterization of sugarcane legumain. *Plant physiology and biochemistry: PPB / Societe francaise de physiologie vegetale*, **57C**, 181–192.
- SASAKI, T. & BURR, B. (2000) International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. *Current opinion in plant biology*, **3**, 138–41.
- SCHLÜTER, U., BENCHABANE, M., MUNGER, A., KIGGUNDU, A., VORSTER, J., GOULET, M.-C., CLOUTIER, C. & MICHAUD, D. (2010) Recombinant protease inhibitors for herbivore pest control: a multitrophic perspective. *Journal of experimental botany*, **61**, 4169–83.
- SHEWRY, P.R. & HALFORD, N.G. (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *Journal of experimental botany*, **53**, 947–58.
- SHIMAMOTO, K. & KYOZUKA, J. (2002) Rice as a model for comparative genomics of plants. *Annual review of plant biology*, **53**, 399–419.
- SHULAEV, V., CORTES, D., MILLER, G. & MITTLER, R. (2008) Metabolomics for plant stress response. *Physiologia plantarum*, **132**, 199–208.
- SHUTOV, A.D., BLATTNER, F.R., KAKHOVSKAYA, I. A & MÜNTZ, K. (2012) New aspects of the molecular evolution of legumains, Asn-specific cysteine proteinases. *Journal of plant physiology*, **169**, 319–21.
- SOLOMON, M., BELENGHI, B., DELLEDONNE, M., MENACHEM, E. & LEVINE, A. (1999) The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *The Plant cell*, **11**, 431–44.
- SWINDELL, W.R. (2006) The association among gene expression responses to nine abiotic stress treatments in *Arabidopsis thaliana*. *Genetics*, **174**, 1811–24.
- TAJIMA, T., YAMAGUCHI, A., MATSUSHIMA, S., SATOH, M., HAYASAKA, S., YOSHIMATSU, K. & SHIOI, Y. (2011) Biochemical and molecular characterization of senescence-related cysteine protease-cystatin complex from spinach leaf. *Physiologia plantarum*, **141**, 97–116.
- TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. & KUMAR, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, **28**, 2731–9.
- TURK, V. & BODE, W. (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS letters*, **285**, 213–219.

- TURK, V. & TURK, B. (2001) Lysosomal cysteine proteases: facts and opportunities. *The EMBO journal*, **20**, 4629–4633.
- UPADHYAYA, N.M., ZHOU, X-R., ZHU, Q-H., EAMENS, A., WANG, M-B., WATERHOUSE, P.M., DENNIS, E.S. (2002) Transgenic Rice. In *Transgenic Cereals*, ed O'Brien L, Henry RJ, AACC, Minnesota: 28-87
- URWIN, P.E., ATKINSON, H.J. & MCPHERSON, M.J. (1995a) Involvement of the NH₂-terminal region of oryzacystatin-I in cysteine proteinase inhibition. *Protein engineering*, **8**, 1303–7.
- URWIN, P., ATKINSON, H. & WALLER, D. (1995b) Engineered oryzacystatin I expressed in transgenic hairy roots confers resistance to *Globodera pallida*. *The Plant Journal*, **8**, 121–131.
- VAIN, P., WORLAND, B., CLARKE, M, C., RICHARD, G., BEAVIS, M., LIU, H., KOHLI, A., LEECH, M., SNAPE, J., CHRISTOU, P. & ATKINSON, H. (1998) Expression of an engineered cysteine proteinase inhibitor (Oryzacystatin-ID86) for nematode resistance in transgenic rice plants. *Theor Appl Genet*, **96**, 266–271.
- VALDES-RODRIGUEZ, S., CEDRO-TANDA, A., AGUILAR-HERNANDEZ, V., CORTES-ONOFRE, E., BLANCO-LABRA, A. & GUERRERO-RANGEL, A. (2010) Recombinant amaranth cystatin (AhCPI) inhibits the growth of phytopathogenic fungi. *Plant physiology and biochemistry: PPB / Société française de physiologie végétale*, **48**, 469–75.
- VALDÉS-RODRÍGUEZ, S., GUERRERO-RANGEL, A., MELGOZA-VILLAGÓMEZ, C., CHAGOLLA-LÓPEZ, A., DELGADO-VARGAS, F., MARTÍNEZ-GALLARDO, N., SÁNCHEZ-HERNÁNDEZ, C. & DÉLANO-FRIER, J. (2007) Cloning of a cDNA encoding a cystatin from grain amaranth (*Amaranthus hypochondriacus*) showing a tissue-specific expression that is modified by germination and abiotic stress. *Plant physiology and biochemistry: PPB / Société française de physiologie végétale*, **45**, 790–8.
- WALKER, A J., URWIN, P.E., ATKINSON, H.J., BRAIN, P., GLEN, D.M. & SHEWRY, P.R. (1999) Transgenic Arabidopsis leaf tissue expressing a modified oryzacystatin shows resistance to the field slug *Deroceras reticulatum* (Müller). *Transgenic research*, **8**, 95–103.
- WAN, L., XIA, Q., QIU, X. & SELVARAJ, G. (2002) Early stages of seed development in *Brassica napus*: a seed coat-specific cysteine proteinase associated with programmed cell death of the inner integument. *The Plant journal: for cell and molecular biology*, **30**, 1–10.
- WANG, K.-M., KUMAR, S., CHENG, Y.-S., VENKATAGIRI, S., YANG, A.-H. & YE, K.-W. (2008) Characterization of inhibitory mechanism and antifungal activity between group-1 and group-2 phytocystatins from taro (*Colocasia esculenta*). *The FEBS journal*, **275**, 4980–9.
- WANG, W., VINOCUR, B. & ALTMAN, A. (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta*, **218**, 1–14.

- WATANABE, H., ABE, K., EMORI, Y. & HOSOYAMA, H. (1991) Molecular cloning and gibberellin-induced expression of multiple cysteine proteinases of rice seeds (oryzains). *Journal of Biological*, **266**, 16897–16902.
- YAMADA, K., SHIMADA, T., NISHIMURA, M. & HARA-NISHIMURA, I. (2005) A VPE family supporting various vacuolar functions in plants. *Physiologia Plantarum*, **123**, 369–375.
- YANG, A H. & YEH, K.W. (2005) Molecular cloning, recombinant gene expression, and antifungal activity of cystatin from taro (*Colocasia esculenta* cv. Kaosiung no. 1). *Planta*, **221**, 493–501.
- YE, N., JIA, L. & ZHANG, J. (2012) ABA signal in rice under stress conditions. *Rice*, **5**, 1.
- ZHANG, X., LIU, S. & TAKANO, T. (2008) Two cysteine proteinase inhibitors from *Arabidopsis thaliana*, AtCYSa and AtCYSb, increasing the salt, drought, oxidation and cold tolerance. *Plant Molecular Biology*, **68**, 131–143.
- ZHAO, Y., BOTELLA, M. A, SUBRAMANIAN, L., NIU, X., NIELSEN, S.S., BRESSAN, R. A & HASEGAWA, P.M. (1996) Two wound-inducible soybean cysteine proteinase inhibitors have greater insect digestive proteinase inhibitory activities than a constitutive homolog. *Plant physiology*, **111**, 1299–306.

CURRICULUM VITAE

CHRISTOFF, A. P.

1. DADOS PESSOAIS

Nome: Ana Paula Christoff

Local e data de nascimento: Três Passos, RS, Brasil. 21/08/1988

Endereço profissional: Universidade Federal do Rio Grande do Sul
Av. Bento Gonçalves, 9500 Prédio 43431, Sala 213 - Setor IV -
Campus do Vale da UFRGS. CEP 91501970. Cx.P. 15005
Porto Alegre/RS

Telefone profissional: 3308-7766

Email: anachff@gmail.com

2. FORMAÇÃO

2011 Mestrado em Programa de Pós-Graduação em Biologia Celular e Molecular.
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil
Título: Proteinases cisteínicas de origem vegetal, Legumainas, e seus inibidores: as fitocistatinas
Orientador: Rogerio Margis
Bolsista do (a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

2007 - 2010 Graduação em Ciências Biológicas.
Universidade Federal de Santa Maria, UFSM, Santa Maria, Brasil
Orientador: Elgion Lucio da Silva Loreto
Bolsista do (a): Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul

3. ESTÁGIOS

05/2007 - 08/2007 Estágio, Hospital Universitário, HUSM, Santa Maria/RS
Diagnóstico molecular - setor Hemato-Oncologia

4. PROJETOS DE PESQUISA

2007 - 2007 **Prevenção do câncer de pele: vendo a ação da radiação ultravioleta na sala de aula**
Situação: Concluído
Natureza: Projeto de extensão
Alunos envolvidos: Graduação (2); Mestrado acadêmico (1);

Integrantes: Ana Paula Christoff; Lenira Maria Nunes Sepel; Élgion Lúcio da Silva Loreto (Responsável); Luciele Ceolin; Francele de Abreu Carlan
Financiador(es): Fundo de Incentivo à Extensão-FIEX

2008 - 2009

Caracterização molecular e funcional do elemento gypsy de ZAPRIONUS INDIANUS

Situação: Concluído Natureza: Projetos de pesquisa
Alunos envolvidos: Graduação (1);
Integrantes: Ana Paula Christoff; Élgion Lúcio da Silva Loreto (Responsável)
Financiador(es): Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul-FAPERGS

2007 - 2010

O Transposon Tip100 em espécies de IPOMOEA nativas do Rio Grande do Sul

Situação: Concluído Natureza: Projetos de pesquisa
Alunos envolvidos: Graduação (1);
Integrantes: Ana Paula Christoff; Lenira Maria Nunes Sepel; Élgion Lúcio da Silva Loreto (Responsável)
Financiador(es): Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul-FAPERGS

2011 - Atual

Proteinases cisteínicas de origem vegetal e seus inibidores: as Fitocistatinas

Situação: Em andamento Natureza: Projetos de pesquisa
Integrantes: Ana Paula Christoff; Rogerio Margis (Responsável)

2011 - Atual

MicroRNAs e silenciamento gênico pós-transcricional em plantas

Descrição: Identificação do papel de diferentes miRNAs e de suas famílias na regulação pós-transcricional da expressão gênica em arroz, soja, outras oleaginosas e espécies nativas. Emprego da metodologia de RNAi na produção de plantas transgênicas em estudos de genômica funcional.
Situação: Em andamento Natureza: Projetos de pesquisa
Alunos envolvidos: Mestrado acadêmico (1); Doutorado (7);
Integrantes: Ana Paula Christoff; Guilherme Loss; Ana Paula Korbes; Marcia Margis-Pinheiro; Rogerio Margis (Responsável); Andréia Caverzan; Franceli Rodrigues Kulcheski; Frank Lino Guzman Escudero; Guilherme Cordenonsi da Fonseca; João Braga de Abreu Neto; Vanessa Galli

5. PRÊMIOS E DISTINÇÕES

- 2010** Prêmio Iniciação Científica de melhor trabalho na área de Genética, Evolução e Melhoramento de Plantas, Sociedade Brasileira de Genética
- 2009** Menção Honrosa Prof. Antônio Rodrigues Cordeiro, II Simpósio Brasileiro de Genética Molecular de Plantas

6. ARTIGOS COMPLETOS PUBLICADOS

1. Christoff, Ana-Paula, Loreto, Elgion L.S., Sepel, Lenira M.N.

Evolutionary history of the Tip100 transposon in the genus Ipomoea. *Genetics and Molecular Biology (Impresso)*. , v.35, p.460 - 465, 2012.

2. Viero, Cláudia, Camassola, Melissa, Bellagamba, Bruno, Ikuta, Nilo, CHRISTOFF, A. P., Meirelles, Lindolfo da Silva, Ayres, Raquel, Margis, Rogério, Nardi, Nance Beyer

Molecular Analysis of the Differentiation Potential of Murine Mesenchymal Stem Cells from Tissues of Endodermal or Mesodermal Origin. *Stem Cells and Development*. , v.v. 00, p.111005012553004 - , 2011.

7. RESUMOS E TRABALHOS APRESENTADOS EM CONGRESSOS

1. TURCHETTO-ZOLET, A. C., CHRISTOFF, A. P., LOSS, G., KORBES, A. P., CAGLIARI, A., MARASCHIN, F. S., MARGIS-PINHEIRO, M., MARGIS, R. Diacylglycerol acetyltransferase (DGAT) enzymes: a phylogenetic analysis and comparative gene expression in soybean In: 6th International Crop Science Congress, 2012, Bento Gonçalves.

6th International Crop Science Congress. , 2012.

2. MARGIS, R., CHRISTOFF, A. P., OLIVEIRA, LFV

Diversity of nucleotide edition by addition on soybean miRNAs In: Plant RNA workshop, 2012, Vienna.

Plant RNA workshop. , 2012.

3. CHRISTOFF, A. P., OLIVEIRA, LFV, MARGIS, R.

Is there a rule for miRNA nucleotide addition? In: 6th International Crop Science Congress, 2012, Bento Gonçalves.

6th International Crop Science Congress. , 2012.

4. CHRISTOFF, A. P., LORETO, E. L. S., SEPEL, L. M. N.

Análise Evolutiva do Elemento de Transposição Tip100 em espécies de IPOMOEA In: XVII Encontro de Geneticistas do Rio Grande do Sul, 2010, Vale Vêneto.

Anais do XVII Encontro de Geneticistas do Rio Grande do Sul. , 2010.

5. CHRISTOFF, A. P., LORETO, E. L. S., SEPEL, L. M. N.
Distribuição do Transposon Tip100 em espécies de Ipomoea In: 25ª Jornada Acadêmica Integrada, 2010, Santa Maria.
Anais da 25ª JAI. , 2010.

6. CHRISTOFF, A. P., LORETO, E. L. S., SEPEL, L. M. N.
Relações filogenéticas do transposon Tip100 entre espécies do gênero Ipomoea In: 56º Congresso Brasileiro de Genética, 2010, Guarujá.
Anais do 56º Congresso Brasileiro de Genética. , 2010.

7. CHRISTOFF, A. P., LORETO, E. L. S.
Caracterização Molecular da ORF1 do Elemento Gypsy de Zaprionus Indianus In: VI Simpósio de Ecologia, Genética e Evolução de Drosophila, 2009, Florianópolis.
Anais VI Simpósio de Ecologia, Genética e Evolução de Drosophila

8. CHRISTOFF, A. P., LORETO, E. L. S., SEPEL, L. M. N.
Evolução do Elemento Transponível Tip100 em espécies do gênero IPOMOEA In: II Simpósio de Biodiversidade, 2009, Santa Maria.
Anais do II Simpósio de Biodiversidade. , 2009.

9. CHRISTOFF, A. P., LORETO, E. L. S., SEPEL, L. M. N.
Isolamento e Caracterização de Marcadores ITS em Espécies do Gênero IPOMOEA In: 24 Jornada Acadêmica Integrada, 2009, Santa Maria.
Anais da 23 Jornada Acadêmica Integrada. , 2009.

10. CHRISTOFF, A. P., OLIVEIRA, LFV, LORETO, E. L. S., SEPEL, L. M. N.
The transposon TIP100 in the genus IPOMOEA In: II Simpósio Brasileiro de Genética Molecular de Plantas, 2009, Búzios.
anais do II Simpósio Brasileiro de Genética Molecular de Plantas. , 2009.

11. CHRISTOFF, A. P., OLIVEIRA, LFV, LORETO, E. L. S., SEPEL, L. M. N.
Análise do Elemento Transponível Tip100 em Espécies do Gênero Ipomoea Nativas do Rio Grande do Sul In: 23º Jornada Acadêmica Integrada, 2008, Santa Maria.
Anais da 23º Jornada Acadêmica Integrada. , 2008.

12. CHRISTOFF, A. P., OLIVEIRA, LFV, LORETO, E. L. S., SEPEL, L. M. N.
Investigação da presença de TIP100 em espécies do gênero Ipomoea nativas do Rio Grande do Sul In: 54º Congresso Brasileiro de Genética, 2008, Salvador.
Anais do 54º Congresso Brasileiro de Genética. , 2008.

Apresentação de trabalho e palestra

1. CHRISTOFF, A. P., LORETO, E. L. S., SEPEL, L. M. N.
Evolução do Elemento Transponível Tip100 em espécies do gênero IPOMOEA, 2009. (Simpósio,Apresentação de Trabalho)