Universidade Federal do Rio Grande do Sul

Centro de Biotecnologia

Programa de Pós-Graduação em Biologia Celular e Molecular (PPGBCM)

Laboratório de Fisiologia Vegetal, Departamento de Botânica

Evolução e expressão de dois genes desconhecidos de arroz (*Oryza sativa* L.) possivelmente envolvidos com a homeostase de ferro

Aluno: Victor Hugo Rolla Fiorentini

Orientador: Felipe Klein Ricachenevsky

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

Porto Alegre

Março de 2024

INSTITUIÇÕES E FONTES FINANCIADORAS

Este trabalho foi desenvolvido no Laboratório de Fisiologia Vegetal do Departamento de Botânica e no Núcleo de Genômica Funcional de Plantas, do Departamento de Genética do Instituto de Biociências da Universidade Federal do Rio Grande do Sul (UFRGS). O aluno recebeu bolsa de estudos concedida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). O Laboratório conta com recursos do CNPq e FAPERGS para execução de seus projetos.

Agradecimentos

Gostaria de começar agradecendo meus pais, João Tadeu e Bianca, já que a ajuda deles começou há mais de 27 anos. Como tenho limite de espaço, literalmente não cabe em palavras o quanto das minhas conquistas eu devo ao esforço de vocês, mas saibam que não passou despercebido! Espero deixá-los orgulhosos. Abraço apertado!

Agradeço aos meus irmãos, João Pedro e Arthur, por sempre apoiarem e demonstrarem interesse na minha carreira.

Agradeço ao meu orientador, Felipe Ricachenevsky, cuja carreira é tão repleta de conquistas que me inspira a ser um profissional melhor e cuja vida social é tão repleta de amor, respeito e dignidade que me inspira a ser uma pessoa melhor. Prof, a proposta de aperfeiçoamento pessoal do mestrado tomou um significado muito mais profundo sob sua orientação. Obrigado!

Quero agradecer também a todos os colegas da fisiologia vegetal, CBiot e genética que sempre se disponibilizaram para ajudar nos momentos difíceis, seja para tirar uma dúvida, para carregar geladeira de um lado pro outro ou para ouvir um desabafo. Agradeço pelos momentos de descontração, o sentimento de leveza proporcionado por vocês foi essencial para que eu me sentisse acolhido numa cidade onde eu não conhecia ninguém. Não posso deixar de mencionar a Raq, que tem a paciência de um monge tibetano, sempre aturando minhas baboseiras, o Lucas, nosso querido emburrado que tem o maior coração de todos, o Jover, por sua habilidade de tornar qualquer situação leve e alegre, a Angie, por todas as risadas incontroláveis e por sempre estar suave na nave, leve na neve e de boa na lagoa. Parte dessa conquista é de vocês! Eu não teria chegado aqui sem o apoio de vocês, muito obrigado.

Agradeço aos discentes do programa, por se esforçarem para melhorar a qualidade das pesquisas, assim como a experiência do aluno no programa. Só tenho elogios aos profissionais da secretaria, nunca fui tão bem atendido. Silvia, você é um amor de pessoa e não tenho dúvida que a frase que você mais escuta é “Silvinha, salvou minha vida”. Agradeço às instituições financiadoras por tornarem meu trabalho uma possibilidade.

Índice

Lista de Abreviaturas

ROS - do inglês, *Reactive Oxygen Species*

Fe^{2+} - íon ferroso

Fe^{3+} - íon férrico

Fe - ferro

FRO - do inglês, *Ferric Reductase Oxidase*

IRT - do inglês, *Iron-Regulated Transporter*

YS - do inglês, *Yellow Stripe*

NAAT - do inglês, *Nicotianamine Aminotransferase*

T-DNA - do inglês, *Transfer-Deoxyribonucleic Acid*

YSL - do inglês, *Yellow Stripe-like*

bHLH - do inglês, *Basic helix-loop-helix*

At - *Arabidopsis thaliana*

FIT - do inglês, *FER-like iron-deficiency-induced transcription factor*

Sl - *Solanum lycopersicum*

FER - do inglês, *Ferritin*

Os - *Oryza sativa*

IRO - do inglês, *Iron-related transcription factor*

NAS - do inglês, *Nicotianamine Synthase*

DMAS - do inglês, *Deoxymugineic Acid Synthase*

VIT - do inglês, *Vacuolar Iron Transporter*

TRG - do inglês, *Taxonomically Restricted Genes*

BLAST - do inglês, *Basic Local Alignment Search Tool*

PSI-BLAST - do inglês, *Position-Specific Iterated-Basic Local Alignment Search Tool*

QQS - do inglês, *Qua-Quine Starch*

IMA - do inglês, *IRON MAN*

IDE1 - do inglês, *IRON DEFICIENCY ELEMENT-BINDING FACTOR*

HRZ - do inglês, *HAEMERYTHRIN MOTIF-CONTAINING REALLY INTERESTING NEW GENE (RING) AND ZINC-FINGER PROTEIN*

MIR - do inglês, *Mitochondrial Iron-Regulated*

IDEF - do inglês, *Iron Deficiency-Responsive Cis-Acting Element Binding Factor*

RNA-seq - do inglês, *Ribonucleic Acid Sequencing*

RT-qPCR - do inglês, *Quantitative Reverse Transcription Polymerase Chain Reaction*

ID - do inglês, *Identification*

CDS - do inglês, *coding sequence*

Resumo

O principal objetivo deste trabalho é contribuir para a compreensão da função dos transcritos e confirmar os padrões de expressão de dois genes não documentados do arroz (*Oryza sativa*), LOC_Os11g15624 e LOC_Os04g45510. Dados públicos de RNA-seq e outras publicações envolvendo análise do nível de expressão gênica sugerem que ambos os genes expressam proteínas hipotéticas funcionais, enquanto LOC_Os11g15624 pode participar da resposta de deficiência de ferro da planta de arroz e LOC_Os04g45510, da resposta de excesso de ferro, embora não haja mais informações sobre esses genes na literatura. Como objetivo secundário deste trabalho, pretendemos esclarecer o processo evolutivo e analisar a presença de genes parálogos e ortólogos relacionados a ambos os genes em arroz. Para isso, realizamos análises de sintenia e colinearidade entre os genes de interesse em arroz e seus respectivos genes homólogos, além de propor árvores filogenéticas contendo a evolução dos genes de interesse. Para confirmar os padrões de expressão e tentar esclarecer a função de ambos os genes de interesse, linhagens transgênicas de superexpressão de arroz estão sendo cultivadas e serão submetidas a diferentes tratamentos, incluindo excesso e deficiência de ferro. Pretendemos medir os parâmetros moleculares e fisiológicos das plantas transgênicas.

Abstract

The main objective of this work is to contribute to the understanding of the function of transcripts and confirm the expression patterns of two undocumented genes from rice (*Oryza sativa*), LOC_Os11g15624 and LOC_Os04g45510. Public RNA-seq data and other publications involving gene expression level analysis suggest that both genes express hypothetical functional proteins, while LOC_Os11g15624 may participate in the iron deficiency response of the rice plants and LOC_Os04g45510 in the iron excess response. Although there is no further information about these genes in the literature. As a secondary objective of this work, we intend to clarify the evolutionary process and analyze the presence of paralogous and orthologous genes related to both rice genes. To this end, we performed synteny and collinearity analyzes between the rice genes of interest and their respective homologous genes, in addition to proposing phylogenetic trees containing the evolution of the genes of interest. To confirm the expression patterns and try to clarify the function of both genes of interest, transgenic rice overexpression lines are being cultivated and will be subjected to different treatments, including iron excess and deficiency. We intend to measure the molecular and physiological parameters of transgenic plants.

1. Introdução

a. Arroz: impacto econômico e modelo de pesquisa

O arroz (*Oryza sativa* L.) está entre as três maiores culturas mundiais, representando 30% dos cereais cultivados, sendo acompanhado por trigo e milho (Gnanamanickam, 2009). Comparativamente, o arroz é a mais importante para a alimentação da população, tendo aproximadamente 80% da sua produção total destinada ao consumo humano e sendo parte da base alimentar de mais da metade da população mundial (Gnanamanickam, 2009; Rice almanac, 2013; Wassmann *et al.*, 2010). Seu impacto como fonte de alimento no cenário global é tão significativo que 2004 foi declarado como o ano do arroz pela ONU (Gnanamanickam, 2009).

O Brasil está entre os 10 maiores produtores mundiais de arroz e é o maior produtor da América Latina (Rice almanac, 2013). A safra de 2023/24 rendeu 10.755,5 mil toneladas, ocupando um território de 1.565,4 mil hectares, possuindo um rendimento de 6.871 kg/hectare (Conab, 2024). A maior parte da produção anual é dedicada ao consumo interno e o restante é destinado à exportação, sendo o Rio Grande do Sul o estado com a maior produção de arroz do Brasil, responsável por cerca de 75% da produção nacional (Coêlho, 2021; SANTOS *et al.*, 2006).

O arroz também apresenta valor como planta modelo para o estudo de outras monocotiledôneas (de Oliveira *et al.*, 2020). Seu alto grau de sintenia com genes de outras espécies da família *Poaceae* e a abundância de dados públicos disponíveis para a comunidade científica, como sequência referência para o genoma das duas subespécies mais utilizadas, *japonica* e *indica*, sequenciamento de diversas cultivares e banco de dados de mutantes o colocam como um excelente modelo para a pesquisa, próximo à *Arabidopsis thaliana* (de Oliveira *et al.*, 2020; Li *et al.*, 2017; Stein *et al.*, 2018). Portanto, o arroz é uma espécie na qual os trabalhos podem ter impacto tanto na pesquisa básica quanto aplicada.

b. Evolução do gênero *Oryza*

O gênero *Oryza*, pertencente a tribo *Oryzeae*, subfamília *Ehrhartoideae*, família *Poaceae*, possui aproximadamente 24 espécies com 11 tipos de genomas diferentes (AA, BB, CC, BBCC, CCDD, EE, FF, GG, KKLL, HHJJ e HHKK) cuja especiação iniciou há aproximadamente 15 milhões de anos (de Oliveira *et al.*, 2020; Kellogg, 2009; Menguer;

Sperotto; Ricachenevsky, 2017). A domesticação do arroz iniciou há aproximadamente nove mil anos a partir de seus presumidos ancestrais *O. nivara* e *O. rufipogon* (de Oliveira *et al.*, 2020; Menguer; Sperotto; Ricachenevsky, 2017, Stein *et al.*, 2018). O gênero *Oryza* possui duas espécies domesticadas: *O. sativa*, incluindo as subespécies *japonica* e *indica*; e *O. glaberrima* (Ishimaru *et al.*, 2009; Menguer; Sperotto; Ricachenevsky, 2017). Diversas espécies do gênero *Oryza* já possuem seu genoma referência sequenciado e disponibilizado em bancos de dados públicos: *O. sativa* ssp. *japonica e O. sativa* ssp. *indica, O. glaberrima, O. brachyantha, O. nivara, O. rufipogon, O. barthii, O. glumaepatula, O. meridionalis, O. punctata* e *O. australiensis* (Chen *et al.*, 2013; Phillips *et al.*, 2022; Sasaki, 2005; Stein *et al.*, 2018; Yu *et al.*, 2002; Zhang *et al.*, 2015). A abundância de recursos e dados disponíveis sobre o gênero *Oryza*, como genomas sequenciados, acessos para distribuição, linhagens de introgressão e linhagens mutantes e transgênicas de espécies do gênero fornecem oportunidade para análises transcricionais de diversas espécies, assim como do cruzamento entre elas, possibilitando ao gênero *Oryza* ser considerado um "gênero-modelo" (Wairich *et al.*, 2021).

c. Ferro: importância e efeitos do excesso e deficiência

A homeostase do ferro é essencial para a maioria de organismos e tipo celulares conecidos. Em humanos, apesar do ferro possuir funções importantes, o uso mais importante está relacionado à síntese de hemoglobina (Camaschella, 2019). A OMS reconhece que, em relação à saúde mundial, o iodo, a vitamina A e o ferro são os nutrientes mais importantes, pois sua deficiência acarreta grande risco para a saúde (WHO, 2023). A deficiência de ferro na alimentação leva à anemia, que afeta mais de um bilhão de indivíduos, principalmente crianças, mulheres grávidas e populações de baixa renda (Camaschella, 2019; WHO, 2023). No Brasil, 95% dos casos de anemia estão relacionados à deficiência de ferro na dieta (Senge-PR, 2021).

Para as plantas não é diferente: o ferro também é essencial. Ele participa de diversos processos biológicos, sendo cofator de enzimas, assimilação de nitrogênio, respiração, fotossíntese, entre outros processos celulares (Rasheed *et al.*, 2020). No entanto, o excesso de ferro também pode ser danoso, já que ele participa das reações de Fenton e seu excesso facilita o acúmulo de ROS, induzindo danos a estruturas celulares (Wairich *et al.*, 2021). Por isso, as plantas desenvolveram mecanismos para a regulação da homeostase do ferro, já que tanto deficiência como excesso são deletérios (Becker; Asch, 2005; Kar; Panda, 2020).

Tanto a deficiência quanto o excesso do ferro acarretam mudanças semelhantes no desenvolvimento da planta, reduzindo desenvolvimento das raízes, parte aérea e produção de grãos (Rasheed *et al.*, 2020; Zhang, X. *et al.*, 2019). Como sintomas que diferenciam os dois estresses, a deficiência de ferro acarreta sintomas como clorose e taxa fotossintética reduzida (Zhang, X. *et al.*, 2019). Já a toxicidade resulta no bronzeamento foliar e redução do perfilhamento, podendo levar à morte da planta (Rasheed *et al.*, 2020). Ainda assim, em plantas, a deficiência de ferro é um problema muito mais comum, dada a baixa disponibilidade de ferro tipicamente encontrada em solos naturais e utilizados na agricultura (Gnanamanickam, 2009; Zhang *et al.*, 2019).

O arroz é suscetível à toxicidade de ferro, especialmente quando cultivado sob condições de alagamento (Stein *et al.*, 2019). As condições de hipóxia (baixa concentração de oxigênio) e baixo pH facilitam a redução do Fe^{+3} para Fe^{+2}, forma mais solúvel que se acumula na solução do solo, levando à absorção excessiva (Guerinot; Yi, 1994; Stein *et al.*, 2019). A toxicidade por ferro é um dos distúrbios nutricionais mais comuns para arroz, podendo reduzir a produção em 12 - 100%, dependendo do genótipo, das propriedades do solo em questão e da severidade da toxicidade (Stein *et al.*, 2019).

No entanto, apesar do ferro ser um elemento relativamente abundante no solo, dependendo das propriedades do solo, ele pode se encontrar indisponível para a planta (Wairich *et al.*, 2019). Como o ferro é necessário em diversas etapas da fotossíntese, a sua deficiência se torna fator limitante para o desenvolvimento e produção de diversas espécies de plantas. A deficiência de ferro está presente em aproximadamente 30% dos solos globais (Pradhan *et al.*, 2020). No cultivo do arroz, a deficiência ocorre em solos calcários e/ou aerados com pH neutro ou básico, já que o ferro forma óxi-hidróxidos que são insolúveis, atingindo concentrações entre 10^{-14} e 10^{-17} M, sendo que as concentrações ideais são entre 10^{-4} e 10^{-9} M (Guerinot; Yi, 1994). Nessas concentrações a planta de arroz se encontra em deficiência de ferro, apresentando os sintomas mencionados. Portanto, a correta regulação da concentração de ferro é necessária para o funcionamento de processos vitais em plantas. Por isso, as plantas têm mecanismos de homeostase que surgiram ao longo da evolução, com a finalidade de garantir a manutenção da concentração de ferro nos tecidos da planta.

d. Deficiência de Ferro em arroz

As plantas apresentam duas estratégias para absorção de ferro. A estratégia I consiste na secreção de prótons (H^+) por células de raiz na rizosfera por meio da atividade de H^+-ATPase, diminuido o pH e aumentando a solubilidade do Fe^{3+}, forma mais comumente encontrada em solos aerados (Wairich *et al.*, 2019; Wang *et al.*, 2020). O Fe^{3+} é reduzido para Fe^{2+} por uma proteína de membrana da família FRO (*Ferric Reductase Oxidase*) e transportado através da membrana plasmática das células de raiz por um transportador de alta afinidade da família IRT (*Iron-Regulated Transporter*) (Kar; Panda, 2020; Wairich *et al.*, 2019; Wang *et al.*, 2020). A estratégia II consiste na secreção de fitosideróforos (moléculas sintetizadas a partir de S-adenosil-metionina) para a rizosfera, que por sua vez formam complexos Fe(III)-fitosideróforo solúveis que são transportados para o citosol de células da raiz através de transportadores específicos da família YS (*Yellow Stripe*) (Kar; Panda, 2020; Wairich *et al.*, 2019; Wang *et al.*, 2020).

A estratégia I é comum a todas as plantas, exceto as plantas da família *Poaceae*; já a estratégia II é comum das plantas da família *Poaceae* (Kar; Panda, 2020; Wairich *et al.*, 2019). O arroz, no entanto, não se encaixa no padrão, utilizando uma estratégia combinada, na qual a estratégia II completa e parte da estratégia I ocorrem simultaneamente (Kar; Panda, 2020; Wairich *et al.*, 2019). Como o arroz pertence à família *Poaceae*, era esperada a ocorrência da estratégia II somente, porém o arroz expressa também um transportador de Fe^{2+}, *OsIRT1*, componente da estratégia I (Cheng *et al.*, 2007). Em um estudo com uma linhagem de arroz mutante para o gene NAAT (*Nicotianamine Aminotransferase*), mutação que leva à redução da produção de fitosideróforos e absorção ineficiente de Fe^{3+}, que a planta mutante ainda se desenvolve quando suprida com Fe^{2+} (Cheng *et al.*, 2007). Também foi mostrado que linhagens com inserção de T-DNA na sequência do gene *OsYSL15* conseguiam se desenvolver (Hindt; Guerinot, 2012). A expressão de componentes da estratégia I pelo arroz foi inicialmente proposta como uma adaptação resultante da sua ocorrência em solos alagados, nos quais a forma Fe^{2+} é predominante (Ishimaru *et al.*, 2006; Wairich *et al.*, 2019). No entanto, um trabalho do nosso grupo mostrou que plantas do grupo de genomas AA do gênero *Oryza*, quando sob deficiência de ferro, induzem tanto genes que fazem parte da estratégia II quanto *OsIRT1*. Estes dados sugerem que a estratégia combinada surgiu antes da domesticação de *Oryza sativa* (Wairich et al 2019)

Para lidar com flutuações na disponibilidade de nutrientes, as plantas estabeleceram mecanismos de regulação da expressão dos genes que codificam as proteínas transportadoras, assim como enzimas que participam das estratégias de absorção de ferro (Lichtblau *et al.*, 2023; Wairich *et al.*, 2019). O controle da expressão dos genes mencionados depende da atuação de fatores de transcrição bHLH (*Basic helix-loop-helix*) (Lichtblau *et al.*, 2023). Em *A. thaliana,* é bem descrito o funcionamento da rede que inclui o fator de transcrição *AtFIT* (*FER-like iron-deficiency-induced transcription factor*) (Hindt; Guerinot, 2012). *AtFIT* é um ortólogo funcional de um gene de tomate que codifica um fator de transcrição da família bHLH, *SlFER*, que também participa da resposta de deficiência de ferro (Yuan *et al.*, 2005). *AtFIT* é induzido em condições de deficiência de ferro e regula a expressão de dois genes também relacionados à resposta de deficiência de ferro em *A. thaliana*: *AtFRO2* e *AtIRT1* (Colangelo; Guerinot, 2004; Jakoby *et al.*, 2004). Linhagens de superexpressão de *AtFIT* não apresentam níveis de expressão elevados de *AtFRO2* nem *AtIRT1*, porém linhagens de superexpressão de *AtFIT* e *AtbHLH38* ou *AtbHLH39* apresentam níveis de expressão elevados de *AtFRO2* e *AtIRT1*, assim como maior acúmulo de ferro, quando comparadas com plantas selvagens, sugerindo que *AtFIT* interage com *AtbHLH38* e *AtbHLH39* para controlar a expressão de proteínas relacionadas com a estratégia I de absorção de ferro (Colangelo; Guerinot, 2004; Jakoby *et al.*, 2004; Yuan *et al.*, 2008). O gene ortólogo de *AtFIT* em arroz, *OsFIT*, foi caracterizado, desempenhando função similar (Liang *et al.*, 2020). Diferente de *AtFIT*, o *OsFIT* é expresso tanto em raiz quanto em parte aérea e linhagens que superexpressam somente *OsFIT* já apresentam indução dos genes envolvidos na estratégia I e II de captação de ferro (Hindt; Guerinot, 2012; Liang *et al.*, 2020). A indução dos genes envolvidos na estratégia I e II está atrelada à interação entre *OsFIT* e *OsIRO2* (*Iron-related transcription factor 2*), fator de transcrição bHLH que possui papel central na indução dos genes envolvidos na expressão de proteínas relacionadas à estratégia II de captação de ferro (Kar; Panda, 2020; Liang *et al.*, 2020; Wang *et al.*, 2020). *OsIRO2* é induzido em condições de deficiência de ferro, sendo expresso tanto em raiz quanto em parte aérea (Hindt; Guerinot, 2012; Ogo *et al.*, 2007). Foi documentado que *OsIRO2* regula positivamente os genes *OsNAS1*, *OsNAS2* (*Nicotianamine Synthase 2*), *OsNAAT1* e *OsDMAS1* (*Deoxymugineic Acid Synthase 1*), genes envolvidos na síntese de fitosideróforos, assim como o gene *OsYSL15*, que codifica a proteína transportadora do complexo Fe-fitosideróforo (Hindt; Guerinot, 2012; Ogo *et al.*, 2007). Além disso, foi documentado que *OsIRO2* não regula a expressão de *OsIRT1*, sugerindo que o papel regulatório de *OsIRO2* é específico para genes envolvidos na estratégia II de captação de ferro (Hindt; Guerinot, 2012; Ogo *et al.*, 2007). Linhagens que superexpressam *OsIRO2* demonstraram tolerância a condições de deficiência de ferro, se desenvolvendo em solos calcários sem prejudicar o transporte e acúmulo de ferro nos grãos (Hindt; Guerinot, 2012; Ogo *et al.*, 2011).

e. Excesso de Ferro em Arroz

Plantas de arroz também devem ser capazes de resistir a cenários de concentrações excessivas de ferro no solo. A captação excessiva de ferro pela planta é prejudicial, já que ele participa das reações de Fenton, facilitando a formação de ROS que danificam estruturas celulares, como o DNA, membranas e proteínas (Stein *et al.*, 2019; Wairich *et al.*, 2021). Alguns mecanismos de tolerância já foram propostos para arroz. A primeira estratégia a ser apresentada, estratégia I, consiste na formação de uma barreira oxidativa na rizosfera a qual limita o transporte de íons na raiz (Becker; Asch, 2005; Rasheed *et al.*, 2020; Wairich *et al.*, 2021). Essa estratégia limita a quantidade de Fe^{2+} que é transportado para a parte aérea (tecido sensível à sua presença) (Becker; Asch, 2005; Rasheed *et al.*, 2020). Utilizando o oxigênio transportado pelo aerênquima, o Fe^{2+} é oxidado, formando depósitos de óxidos de ferro na região da rizosfera, que ao serem acumulados, formam uma placa de ferro que atua como uma barreira física e limita a passagem de Fe^{2+} para o interior da raiz (Becker; Asch, 2005; Rasheed *et al.*, 2020; Wairich *et al.*, 2021). Quando o Fe^{2+} ultrapassa a barreira oxidativa, a planta de arroz ainda é capaz de excluí-lo da membrana celular de células da endoderme da raiz, já que, para ser carregado no xilema, ele seria obrigado a passar pela membrana celular destas células devido à presença das estrias de Caspary (supondo que a planta possui um sistema radicular não danificado) (Becker; Asch, 2005).

A estratégia II consiste no depósito de ferro em órgãos, tecidos e organelas da planta onde o ferro não é prejudicial, principalmente na raiz e colmo (Becker; Asch, 2005). O ferro pode ser armazenado no apoplasto e vacúolos ou no interior de cavidades de proteínas ferritina (Becker; Asch, 2005; Briat *et al.*, 2010; Silveira *et al.*, 2009). Em arroz, o transporte e acúmulo de ferro em vacúolos é facilitado pela ação de transportadores vacuolares da família VIT (*Vacuolar Iron Transporter*), como *OsVIT1* e *OsVIT2*, em situações de excesso de ferro (Zhang *et al.*, 2012). As proteínas ferritina, expressas a partir dos genes *OsFER1* e *OsFER2* em arroz, são capazes de acumular até 4500 átomos de ferro em seu interior, sendo que existe evidência sugerindo que este seja o principal papel da ferritina (Dey *et al.*, 2023; Stein; Ricachenevsky; Fett, 2009). Algumas plantas são capazes de regular o pH do apoplasto, sendo que a acidificação do apoplasto favorece a captação de Fe^{2+}, limitando seu transporte para outros tecidos (Becker; Asch, 2005).

A estratégia III envolve a detoxificação de ROS e outras espécies de radicais livres geradas, garantindo tolerância ao excesso de ferro nos tecidos aéreos da planta (Majerus *et al.*, 2007; Majerus; Bertin; Lutts, 2009). A ação antioxidante dessa estratégia está relacionada à atividade de diversas moléculas. Dentre as principais não enzimáticas: glutationa, ácido ascórbico, compostos fenólicos, alcalóides e α-tocoferóis; e enzimáticas: superóxido dismutase, catalase, ascorbato peroxidase, glutationa redutase, dehidroascorbato redutase, glutationa peroxidase, guaiacol peroxidase e glutationa-S-transferase (Dufey *et al.*, 2015; Majerus; Bertin; Lutts, 2009).

f. Genes órfãos, TRG e novidades evolutivas

Com o avanço da bioinformática e a disponibilidade de genomas sequenciados em bancos de dados públicos, o estudo sobre o surgimento de novos genes tem se tornado cada vez mais comum. Ao contrário do que se pensava, genes novos não são originados somente por eventos de mutação, duplicação e transferência horizontal de genes previamente existentes (Singh; Syrkin Wurtele, 2020). Alguns trabalhos sugerem a possibilidade do surgimento de genes novos a partir de regiões previamente não codificantes do DNA, nomeados genes *de novo* (Singh; Syrkin Wurtele, 2020; Zhang, L. *et al.*, 2019). Genes *de novo* são originados a partir de regiões não codificantes, costumam possuir sequência curta, baixa expressão e, comparando os números de genes conhecidos e sem função conhecida em genomas próximos, parecem ser gerados e eliminados a uma taxa constante durante a evolução dos genomas (Schlötterer, 2015; Zhang, L. *et al.*, 2019).

O projeto de sequenciamento do genoma de leveduras foi uma das inspirações para o levantamento da hipótese sobre a existência de uma nova categoria de genes, os chamados genes órfãos, com um terço dos genes identificados classificados nessa categoria

(Tautz; Domazet-Lošo, 2011). Genes órfãos são definidos como genes que não participam de nenhuma família gênica e não possuem homologia com genes de espécies taxonomicamente próximas ou distantes (Andersson; Jerlström-Hultqvist; Näsvall, 2015; Singh; Syrkin Wurtele, 2020; Vakirlis; Carvunis; McLysaght, 2020). Por exemplo: um gene sem similaridade detectável com outros, e presente em apenas um único genoma. Por ser uma área de estudo relativamente recente, os termos e suas definições não estão completamente solidificados (Schlötterer, 2015). Por fim, existem poucos genes considerados órfãos que já foram caracterizados, o que pode ser resultado da falta de similaridade com sequências de outros genes, tornando sua detecção mais difícil, hipótese sustentada por análise *in silico* sugerindo que até 80% dos genes órfãos podem passar despercebidos durante a anotação de um novo genoma (Li *et al.*, 2021; Singh; Syrkin Wurtele, 2020).

O conceito de genes órfãos pode ser confundido com o conceito de TRG (*Taxonomically Restricted Genes*). Diferente de genes órfãos, TRG são genes que fazem parte de famílias gênicas pequenas e possuem homologia com genes presentes em grupos restritos de espécies (Vakirlis; Carvunis; McLysaght, 2020). O conceito de genes órfãos parece ser menos abrangente e o de TRG mais abrangente, porém não existem métodos quantitativos claros e estabelecidos para definir a abrangência de cada conceito (Schlötterer, 2015). As ferramentas para detecção de similaridade entre sequências mais aceitas são o BLAST (*Basic Local Alignment Search Tool*) e o PSI-BLAST (*Position-Specific Iterated BLAST*), porém a confiabilidade do resultado não depende somente da funcionalidade das ferramentas mencionadas, que possuem limitações; depende também da qualidade da anotação do genoma e dos genes (Tautz; Domazet-Lošo, 2011).

Estudos sugerem diferentes mecanismos para o surgimento de genes órfãos e TRG; porém existem duas possibilidades mais aceitas: eles podem surgir pela divergência de um gene preexistente que sofreu muitos eventos de mutação, e não apresenta mais similaridade com a família gênica de origem; ou podem surgir *de novo*, sendo originados por mutações acumuladas em regiões anteriormente não codificantes (Andersson; Jerlström-Hultqvist; Näsvall, 2015; Singh; Syrkin Wurtele, 2020; Tautz; Domazet-Lošo, 2011). O estudo de genes *de novo* é recente, então o conhecimento adquirido até agora não representa fielmente as características de todos os genes *de novo*, porém a maioria dos que já foram identificados codificam proteínas que se ligam a proteínas conservadas, como fatores de transcrição ou receptores (Andersson; Jerlström-Hultqvist; Näsvall, 2015; Schlötterer, 2015; Vakirlis; Carvunis; McLysaght, 2020; Zhang, L. *et al.*, 2019). Por surgirem com relativa alta frequência e majoritariamente interagirem com redes de expressão gênica, síntese proteica e cascatas de sinalização previamente existentes, o estudo de genes *de novo* pode ser importante para o entendimento de como certas espécies sobrevivem a mudanças no ambiente (Andersson; Jerlström-Hultqvist; Näsvall, 2015; Tautz; Domazet-Lošo, 2011; Zhang, L. *et al.*, 2019).

Como exemplo, a análise do surgimento do gene *QQS* (*Qua-Quine Starch*), em *Arabidopsis thaliana*, um gene *de novo* relacionado ao metabolismo de amido, levantou questionamentos intrigantes (Li *et al.*, 2009). A falta de similaridade entre a sequência proteica do gene *QQS* e genes de outras espécies sugere que ele seja uma novidade genética e um gene órfão exclusivo de *A. thaliana*, sendo assim classificado em diversos estudos posteriores à sua caracterização (Li; Wurtele, 2015; O'Conner *et al.*, 2018; Silveira *et al.*, 2013). Foi observado que os padrões de metilação do DNA, assim como a expressão de *QQS* variam significativamente entre acessos de *A. thaliana* e essas mudanças eram herdadas pela prole (Silveira *et al.*, 2013). Para aprofundar a análise, foram estudadas mutações epigenéticas presentes em diferentes acessos da espécie. Entende-se por mutação epigenética, qualquer alteração na expressão gênica ou fenotípica, podendo ser herdada ou não pela prole, que não envolve mudança na sequência de pares de base do DNA (Goldberg; Allis; Bernstein, 2007). Foi estabelecida uma relação inversamente proporcional entre o nível de metilação do promotor e a expressão de *QQS* e foi mostrado que variações epigenéticas relacionadas à sua expressão podem ser herdadas (Silveira *et al.*, 2013). Existem poucos exemplos de variações epigenéticas que podem ser herdadas, tornando o gene *QQS* uma exceção em comparação com outros genes cuja regulação pode ser controlada por mutações epigenéticas induzidas em laboratório mas para os quais não se encontra variação epigenética em diferentes acessos da espécie na natureza (Silveira *et al.*, 2013). Algumas hipóteses foram formuladas para tentar explicar esse comportamento ímpar de *QQS*: uma delas é que a pressão evolutiva ambiental não muda significativamente entre indivíduos com mutações epigenéticas diferentes relacionadas à expressão de *QQS* (por ele não possuir papel crítico para a sobrevivência e reprodução do indivíduo), permitindo sobrevivência dos indivíduos e propagação das variações; a outra é que genes *de novo* possuem pré-disposição para variações epigenéticas herdáveis (Silveira *et al.*, 2013). A variação epigenética herdável pode ser vista como algo benéfico para genes *de novo*, já que mutações epigenéticas podem ser revertidas com maior facilidade do que mutações na sequência de pares de base do DNA e o mecanismo regulatório do gene *de novo* precisa ser desenvolvido junto à sequência que será expressa, ao contrário de genes originados por duplicação (que já possuem maquinaria regulatória estabelecida) (Silveira *et al.*, 2013). A sequência do gene *QQS* foi introduzida no genoma de outras espécies, como soja e milho para verificar a interação de um gene órfão com redes de sinalização, expressão gênica e interação de proteínas conservadas e rendeu resultados positivos, uma vez que as linhagens transgênicas demonstraram desbalanço na proporção de proteína e amido em tecidos e padrões similares aos de *A. thaliana* (Li *et al.*, 2009; O'Conner *et al.*, 2018). Os resultados mencionados sugerem que o estudo de genes *de novo* e genes órfãos pode levar ao desenvolvimento de novas técnicas e estratégias de manipulação de características interessantes em plantas.

g. Genes com função desconhecida relacionados à homeostase de ferro

A caracterização de proteínas transportadoras, fatores de transcrição e outras proteínas com função na manutenção da homeostase de nutrientes em plantas é importante, pois além de ampliar o conhecimento da área, permite o surgimento de novas estratégias de biofortificação, seja envolvendo técnicas de engenharia genética, cruzamento selecionado de plantas ou controle nutricional do solo ou da própria planta. Existem diversos genes ou famílias gênicas com função desconhecida cuja caracterização trouxe avanço no conhecimento sobre a homeostase de metais. Um exemplo é a família peptídica IRON MAN (IMA), identificada inicialmente em *Arabidopsis thaliana*, e presente em outras espécies de Magnoliophyta (Grillet *et al.*, 2018). A família IMA é composta por peptídeos de sequência curta que regulam positivamente a resposta de deficiência de ferro da planta e possuem um motivo na região C-terminal altamente conservado em espécies de angiospermas (Grillet *et al.*, 2018; Peng *et al.*, 2022). Em *Arabidopsis*, a superexpressão dos genes da família IMA ocasionou indução de genes relacionados à aquisição de ferro, gerando maior acúmulo de ferro em tecidos da planta; já o *knock-out* simultâneo dos genes da família IMA inibiu a resposta de aquisição de ferro, ocasionando sintomas de deficiência de ferro na planta, como clorose severa (Grillet *et al.*, 2018). Foram documentados 8 genes da família IMA em *Arabidopsis*, 14 em trigo e 2 em arroz (*OsIMA1* e *OsIMA2*) (Wang *et al.*, 2023). A superexpressão dos genes *OsIMA1* e *OsIMA2* em arroz resultou em plantas tolerantes a condições de deficiência de ferro com maior acúmulo do metal na parte aérea e sementes (Kobayashi; Nagano; Nishizawa, 2021). A superexpressão dos genes mencionados em arroz também ocasionou a indução exacerbada dos genes relacionados à resposta de deficiência de ferro, exceto por *OsYSL2*, mostrando o papel central dos genes da família IMA na aquisição e mobilização de ferro (Kobayashi; Nagano; Nishizawa, 2021). Em arroz, o mecanismo pelo qual os genes da família IMA atuam na aquisição e mobilização de ferro não está totalmente descrito, porém sua regulação relacionada com a expressão de alguns fatores de transcrição e hormônios conhecidos já foi elucidada. A expressão dos genes IMA em arroz parece ser induzida pelos fatores de transcrição *OsbHLH058/059* e *OsIDEF1* (*IRON DEFICIENCY ELEMENT-BINDING FACTOR 1*), assim como auxinas, citocininas e ácido jasmônico; e inibida pelo fator de transcrição *OsIRO3* e ácido abscísico, enquanto sua degradação é auxiliada pelas ubiquitina ligases *OsHRZ1* e *2* (*HAEMERYTHRIN MOTIF-CONTAINING REALLY INTERESTING NEW GENE (RING) AND ZINC-FINGER PROTEIN 1 AND 2*) (Grillet *et al.*, 2018; Kobayashi; Nagano; Nishizawa, 2021; Peng *et al.*, 2022; Wang *et al.*, 2023).

Outro exemplo é o gene *MIR* (*Mitochondrial Iron-Regulated*), um gene de arroz relacionado à homeostase de ferro que até recentemente era considerado um gene órfão, estando presente apenas no genoma de arroz (Ishimaru et al 2009). Em arroz, o gene MIR é induzido em condições de deficiência de ferro e reprimido em excesso e sua expressão é regulada pelo fator de transcrição IDEF1, conhecido por regular a transcrição de outros genes relacionados à resposta de deficiência de ferro da planta (de Oliveira *et al.*, 2020; Ishimaru *et al.*, 2009). Plantas mutantes para o gene *OsMIR* apresentaram controle da concentração de ferro prejudicado, induzindo genes de absorção mesmo não estando em condições de deficiência, e consequentemente acumulando maiores concentrações de ferro na parte aérea e raiz (de Oliveira *et al.*, 2020; Ishimaru *et al.*, 2009). Análises filogenéticas mais recentes sugerem que *OsMIR* é novidade genética que está presente no genoma de outras espécies do gênero *Oryza* com genoma AA (de Oliveira *et al.*, 2020). Além disso, foi demonstrado que MIR tem origem recente, a partir de uma sequência não-codificante, que por sua vez é originária de um éxon de um gene de rafinose sintase (de Oliveira *et al.*, 2020).

h. Identificação dos genes de interesse

O presente estudo tem como foco dois genes desconhecidos de *O. sativa* ssp. *japonica*, LOC_Os11g15624 e LOC_Os04g45510. De acordo com a base de dados do *Rice Genome Annotation Project*, ambos genes expressam proteína hipotética. LOC_Os11g15624 está localizado no cromossomo 11, possui sequência com 390 nucleotídeos e três variantes de *splicing* alternativo. Já LOC_Os04g45510 está localizado no cromossomo 4, com sequência de 342 nucleotídeos e uma única variante de *splicing* que possui um exon. Utilizando trabalhos de transcriptomas da literatura, foi observado que LOC_Os11g15624 é induzido em condições de deficiência de ferro e reprimido em condições de excesso; e que LOC_Os04g45510 é induzido em excesso e reprimido em deficiência (Drogue *et al.*, 2014; Gonçalves *et al.*, 2020; Wairich *et al.*, 2019, 2021). Os padrões de expressão dos genes de interesse em relação a concentrações de ferro nos levou a hipotetizar que eles estão envolvidos com a manutenção da homeostase de ferro e despertou o interesse da investigação sobre a função e evolução dos genes mencionados.

2. Objetivos

a. Objetivo geral

Caracterizar a história evolutiva de dois genes desconhecidos de *O. sativa* ssp *japonica*, LOC_Os11g15624 e LOC_Os04g45510, possivelmente envolvidos com a homeostase de ferro, e caracterizar a expressão de LOC_Os11g15624.

b. Objetivos específicos:

- ↳ Identificar sequências homólogas e construir uma árvore filogenética de genes relacionados a LOC_Os11g15624.1;
- ↳ Identificar sequências homólogas e construir uma árvore filogenética de genes relacionados a LOC_Os04g45510.1;
- ↳ Avaliar relações de sintenia e colinearidade entre sequências homólogas a LOC_Os11g15624.1;
- ↳ Avaliar relações de sintenia e colinearidade entre sequências homólogas a LOC_Os04g45510.1;
- ↳ Avaliar a expressão gênica de LOC_Os11g15624.1 e sequências homólogas em resposta a deficiência de ferro em plantas do gênero *Oryza*;
- ↳ Clonar a sequência codificante de LOC_Os11g15624.1 em vetor para superexpressão em plantas de arroz;
- ↳ Clonar a sequência codificante de LOC_Os04g45510.1 em vetor para superexpressão em plantas de arroz.
- ↳ Transformar calos de arroz para superexpressar LOC_Os11g15624.1 e regenerar linhagens contendo o transgene.

3. Capítulo I

Manuscrito em preparação, a ser submetido.

Into the unknown: identification and characterization of newly evolved genes in rice (*Oryza sativa* L.)

Victor Hugo Rolla Fiorentini[1], Lucas Roani Ponte[1], João Pedro do Carmo Filgueiras[2], Edgar Luis Waschburger[2], Jover da Silva Alves[1], Angie Geraldine Sierra Rativa[1], Raquel Vargas Olsson[1], Charley Christian Staats[3], Márcia Margis-Pinheiro[2], Andreia Turchetto-Zolet[2], Felipe Klein Ricachenevsky[1,4]

[1] Universidade Federal do Rio Grande do Sul, Laboratório de Fisiologia Vegetal
[2] Departamento de Genética
[3] Centro de Biotecnologia
[4] Departamento de Botânica

Abstract

Rice (*Oryza sativa* L.) is a model species and an economically important crop as a staple to feed an increasing human population. Iron (Fe) is essential for plants, and rice is affected by both Fe deficiency and Fe excess stresses. Several genes in the rice genome have no functional annotation associated to them, and a subset of them are taxonomically restricted genes. In this work, we aim at understanding the molecular history of two uncharacterized genes that are regulated by Fe concentrations, LOC_Os11g15624 and LOC_Os04g45510, and to functionally characterize them. Public RNA-seq data show that LOC_Os11g15624 is up-regulated by Fe deficiency, while LOC_Os04g45510 is up-regulated by Fe excess. We used predicted protein sequences to find similar sequences present in genomes of plant species as well as in rice itself. We identified 28 and 30 homologous proteins to LOC_Os11g15624 and LOC_Os04g45510, respectively. Phylogenetic trees and sequences identified for each gene family suggest that LOC_Os11g15624 evolved recently in the *Oryza* genus, while LOC_Os04g45510 is found in monocots and evolved before *Poaceae* split from other lineages. We found that there is conserved synteny and collinearity between sequences in *Oryza* genomes. Gene expression analyses showed that

LOC_Os11g15624 is up-regulated by Fe deficiency in roots and shoots of cultivated rice. Moreover, we cloned both genes in plasmids for overexpression in rice plants. We regenerated one lineage potentially overexpressing LOC_Os11g15624, which will be functionally characterized. Our work will allow us to shed light on the function of these two uncharacterized genes and their role in iron homeostasis.

Key-words: rice, iron deficiency, iron excess, taxonomically restricted genes, genes of unknown function.

1. Introduction

Rice (*Oryza sativa L.*) is among the three largest crops in the world, representing 30% of cultivated cereals, followed by *Triticum aestivum* and *Zea mays* (Gnanamanickam, 2009). Comparing the three, rice is the most important for feeding the population, since approximately 80% of its total grain production is dedicated to human consumption and it is part of the food base for more than half of the world's population (Gnanamanickam, 2009; Rice almanac, 2013; Wassmann *et al.*, 2010). Brazil is among the 10 largest rice producers in the world and is the main producer outside of Asia (Rice almanac, 2013). The 2023/24 harvest yielded 10,755.5 tons, occupying an area of 1,565.4 hectares, with a yield of 6,871 kg/hectare, according to the Conab database (Conab, 2024). Most of the annual production is dedicated to domestic consumption, while Rio Grande do Sul is responsible for more than 70% of the annual production (Coêlho, 2021; SANTOS *et al.*, 2006).

Rice is also a model plant for monocots (de Oliveira *et al.*, 2020). Its high degree of synteny with genes from other species from the *Poaceae* family and abundance of public data available, such as reference sequence for the genome of the two most employed in research subspecies, *japonica* and *indica*, sequencing of several cultivars and mutant databases make it an excellent model plant for research, second only to *Arabidopsis thaliana* (de Oliveira *et al.*, 2020; Li *et al.*, 2017; Stein *et al.*, 2018).

Iron is essential for the survival of almost all organisms. Although iron has other important functions, in humans the greatest use of iron is related to hemoglobin synthesis (Camaschella, 2019). The World Health Organization (WHO) recognizes that iodine, vitamin A and iron are the most important nutrients to global health, and their deficiency pose a risk to human health (WHO, 2023). Diets that are iron deficient can cause Iron

Deficiency Anemia (IDA) (Camaschella, 2019; WHO, 2023), and iron deficiency-related anemia affects more than one billion individuals, mainly children, pregnant women and low-income populations (WHO, 2023). In Brazil, 95% of anemia cases are related to iron deficiency in the diet (Senge-PR, 2021). It is also estimated that a third of young children in Brazil are anemic (Nogueira-de-Almeida *et al.*, 2021), demonstrating that providing adequate iron amounts in human's diet should be a target for research.

For plants, iron is also essential, as it is involved in several biological processes, being a cofactor for enzymes, participating in nitrogen assimilation, respiration and photosynthesis (Rasheed *et al.*, 2020). Approximately 30% of global soils are deficient in iron and even in fertile soils it might not be available for plants depending on soil characteristics (Pradhan *et al.*, 2020; Wairich *et al.*, 2019). Plants have developed mechanisms for iron uptake, which include two strategies (for reviews please see Hindt; Guerinot, 2012; Pradhan *et al.*, 2020; Zhang *et al.*, 2019). Strategy I is common to all plants, except plants from the *Poaceae* family, while strategy II is common to plants from the *Poaceae* family (Kar; Panda, 2020; Lichtblau *et al.*, 2023; Wairich *et al.*, 2019). Rice is an exception to the rule, using a combined strategy, with strategy II and added elements of strategy I (Wairich *et al.*, 2019).

Rice production can also be affected by iron excess. High iron concentrations inside the cell lead to Reactive Oxygen Species (ROS) formation that damage cellular structures (Stein *et al.*, 2019; Wairich *et al.*, 2021). Three iron excess tolerance mechanisms have already been proposed for rice, involving the limitation of iron uptake, detoxification of ROS and other free radical species and storage of iron in tissues, organelles or molecules of the plant where iron is not harmful (Becker; Asch, 2005; Kar; Panda, 2020; Rasheed *et al.*, 2020). Iron excess can lead to yield reduction (Stein *et al.*, 2019). Therefore, iron concentration in rice plants must be kept within a narrow range.

The genus *Oryza* is part of the tribe Oryzeae, subfamily Ehrhartoideae, family *Poaceae* and has approximately 27 species with 11 different genome types (AA, BB, CC, BBCC, CCDD, EE, FF, GG, KKLL, HHJJ, and HHKK) whose speciation began approximately 15 million years ago (de Oliveira *et al.*, 2020; Kellogg, 2009; Menguer; Sperotto; Ricachenevsky, 2017). The domestication of rice occurred approximately nine thousand years ago from its ancestors *O. nivara* and *O. rufipogon* (de Oliveira *et al.*, 2020; Menguer; Sperotto; Ricachenevsky, 2017). Several species of the *Oryza* genus already have their reference genome sequenced and made available in public databases: *O. sativa ssp. japonica* and *O. sativa ssp. japonica indica, O. glaberrima, O. brachyantha, O. nivara, O. rufipogon, O. barthii, O. glumaepatula, O. meridionalis, O. punctata* and *O. australiensis* (Chen *et al.*, 2013; Phillips *et al.*, 2022; Sasaki, 2005; Stein *et al.*, 2018; Yu *et al.*, 2002; Zhang *et al.*, 2015). Due to that, *Oryza* is becoming an interesting model genus for *in silico* and functional analyses.

With the advancement of bioinformatics and availability of sequenced genomes in public databases, the study of the evolution of newly evolved genes became more feasible. Recent studies suggest that new genes can emerge from previously non-coding regions of DNA, which are therefore named *de novo* genes (Andersson; Jerlström-Hultqvist; Näsvall, 2015; Schlötterer, 2015; Singh; Syrkin Wurtele, 2020). *De novo* genes tend to have a short sequence and low expression level. By comparing the number of genes with unknown functions in closely related genomes, newly evolved genes appear to arise and disappear at a constant rate (Vakirlis; Carvunis; McLysaght, 2020; Zhang, L. *et al.*, 2019). It is hypothesized that the arise of *de novo* genes (among other hypotheses) might be linked to the emergence of orphan genes, defined as genes that do not participate in any gene family nor have homology with genes from taxonomically close or distant species (Andersson; Jerlström-Hultqvist; Näsvall, 2015; Singh; Syrkin Wurtele, 2020; Tautz; Domazet-Lošo, 2011); as well as TRG (*Taxonomically Restricted Genes*), which are genes that can participate in small gene families and have homology with restricted groups of species (Schlötterer, 2015; Vakirlis; Carvunis; McLysaght, 2020). The study of *de novo* genes is recent, and the majority of those that have already been identified seem to encode proteins that bind to conserved proteins, such as transcription factors or receptors (Andersson; Jerlström-Hultqvist; Näsvall, 2015; Schlötterer, 2015; Vakirlis; Carvunis; McLysaght, 2020; Zhang, L. *et al.*, 2019). Because they arise with relatively high frequency and mostly interact with previously existing gene expression networks, protein production and reaction cascades, the study of *de novo* genes can be important for understanding how certain species survive changes in the environment (Andersson; Jerlström-Hultqvist; Näsvall, 2015; Tautz; Domazet-Lošo, 2011; Zhang, L. *et al.*, 2019).

Our work aims at understanding the evolutionary history of two *O. sativa* unknown genes, LOC_Os11g15624 and LOC_Os04g45510, as well as to functionally characterize them. We found that these genes evolved at different moments in the monocots lineages, with LOC_Os11g15624.1 being restricted to closely related species to cultivated rice, whereas LOC_Os04g45510.1 being more ancient, sharing similarity to genes at the base of monocots split. We also found that LOC_Os11g15624 is a genetic novelty with multiple copies in the genome of analyzed species; while the LOC_Os04g45510 has fewer copies in the genomes of each species. Future aims include the functional characterization of both genes in rice.

2. Materials and Methods

2.1. Plant material

Seeds of *Oryza sativa* L. (Nipponbare cultivar), *O. meridionalis* (MER 1 - 93266) and *O. rufipogon* (BRA00004909-8 accession from EMBRAPA Rice & Beans) were dehusked, sterilized using a 2% hypochlorite solution, rinsed and left soaked in distilled water overnight at 28 ^{o}C. After that, seeds were placed in Petri dishes with filter paper and distilled water at 25 ^{o}C with a photoperiod of 16/8 hours for five days or until most of the shoots reached the top of the Petri dish. Seedlings were then transferred to trays with vermiculite watered with complete nutritive solution and maintained at 25 ^{o}C with a photoperiod of 16/8 hours of light and darkness and 150 μmol m^{-2} s^{-1} of light intensity. When most plants reached the two-leaf stage, they were transferred to the hydroponic system with complete nutrient solution for ten days or until all plants had reached the three-leaf stage. Treatments (iron deficiency, iron excess and control) were then introduced and plants were collected for RT-qPCR on the seventh day of treatment. The complete nutrient solution was composed of 0.1 mM KCl; 0.1mM KH_2PO_4; 2 mM $Ca(NO_3)$; 0.5 mM $MgSO_4$; 0.7 mM K_2SO_4; 0.1 mM Fe-EDTA; 0.5 μM $ZnSO_4$; 10 μM H_3BO_3; 0.2 μM $CuSO_4$; 0.01 μM $(NH_4)_6Mo_7O_{24}$; 0.5 μM $MnSO_4$ (Ricachenevsky *et al.*, 2011). Iron deficiency solution omitted Fe-EDTA. Iron excess solution contained 2 mM of $FeSO_4$ $7H_2O$. All nutrient solutions had their pH adjusted to 5.4.

2.2. BLAST analysis

BLAST searches were conducted using the genes of interest's protein sequences as annotated in *Rice Genome Annotation Project* (http://rice.uga.edu) with the objective to investigate the existence of homology between the genes of interest and genes from other species, as well as genes from rice. Our search included the Ensembl (https://plants.ensembl.org/index.html) and Phytozome (https://phytozome-next.jgi.doe.gov) database. The genes of interest's protein sequences were used as query and the BLAST searches included all species from the Phytozome database and all monocot species from the Ensembl database. As a cutoff criteria we used E-values lower or equal to 10E-14. Redundancy between BLAST results from different databases was excluded.

2.3. Phylogenetic tree generation and alignments

For the alignment of nucleotide sequences, we used MEGA software version 11.0.11 (Tamura; Stecher; Kumar, 2021) and MAFFT software version 7 (Katoh; Standley, 2013). MAFFT alignments were conducted employing the FFT-NS-1, FFT-NS-2, FFT-NS-i or L-INS-i strategy. For the generation of phylogenetic trees we used IQTREE software (model of substitution: K2P+G4 and bootstrap 1000 times) (Nguyen *et al.*, 2015) and Fig tree v1.4.4 (Rambaut, 2010) to visualize results. During previous analyses, using sequences recovered by the original annotation from the databases, we recovered homology only in the last exon of the LOC_Os11g15624.1 gene (which contains three exons). However, during BLASTn analyses, we also found a high conservation of the second exon for all sequences, as well as for the first exon. Therefore, we used the SPLIGN tool (Kapustin *et al.*, 2008), with the LOC_Os11g15624.1 nucleotide sequence as a reference, to recover the regions of homology between these genes, as in most cases, the homologous portions referring to the first and second exons were originally "lost", as they were annotated as introns or were divided between two neighboring genes. By adopting this methodology, we lost the coding information for some of the genes analyzed, nevertheless we decided to implement it in order to have a larger homologous region to be compared. Even after doing so, the evolutionary relationship between the recovered clades had low support, which might be explained by the high level of similarity among analyzed sequences.

2.4. Synteny and collinearity analyses

The synteny and collinearity analysis were performed using the Simple Synteny software version 1.6 (Veltri; Wight; Crouch, 2016). The input consisted of chromosomal nucleotide sequences and gene sequences. The gene sequence file contained the genomic nucleotide sequence of either LOC_Os11g15624 or LOC_Os04g45510 and the first five genes located up and downstream of the locus from the *Oryza sativa* ssp. *japonica* genome. Chromosomal sequence files contained chromosomal nucleotide sequence for chromosomes where genes were found. Nucleotide gene and chromosomal sequences were obtained from the Ensembl database. ID's used are relative to the *Oryza sativa* ssp. *japonica* genome and were repeated on hits from other species to indicate probable orthologous relationship with the *Oryza sativa* ssp. *japonica* ortholog. All settings were unchanged except for the "Minimum Query Coverage Cutoff" setting, which was adjusted to 10%. Selection of contigs was done manually. Selection of the most probable ortholog gene in species from the *Oryza* genus other than *Oryza sativa* was oriented by the score result given by the Simple Synteny software and by the Ensembl BLASTn tool. The cutoff for considering a gene as a possible ortholog was 10E-5 or lower E-value in BLAST results. Genes from the BLAST analysis with E-values higher than the cutoff were ignored. Each gene was allowed only one copy (with the highest score) per species. Chromosomal sequence IDs:

Oryza_sativa.IRGSP-1.0.dna.chromosome.11;
Oryza_sativa.IRGSP-1.0.dna.chromosome.4;
Oryza_rufipogon.OR_W1943.dna.chromosome.11;
Oryza_barthii.O.barthii_v1.dna.chromosome.11.fa;
Oryza_glaberrima.Oryza_glaberrima_V1.dna.chromosome.11.fa;
Oryza_glumipatula.Oryza_glumaepatula_v1.5.dna.chromosome.11.fa;
Oryza_meridionalis.Oryza_meridionalis_v1.3.dna.chromosome.8.fa;
Oryza_nivara.Oryza_nivara_v1.0.dna.chromosome.10.fa;
Oryza_barthii.O.barthii_v1.dna.chromosome.4;
Oryza_brachyantha.Oryza_brachyantha.v1.4b.dna.chromosome.4;
Oryza_glaberrima.Oryza_glaberrima_V1.dna.chromosome.4;
Oryza_glumipatula.Oryza_glumaepatula_v1.5.dna.chromosome.4;
Oryza_meridionalis.Oryza_meridionalis_v1.3.dna.chromosome.3;
Oryza_nivara.Oryza_nivara_v1.0.dna.chromosome.4;
Oryza_rufipogon.OR_W1943.dna.chromosome.4;
Oryza_punctata.Oryza_punctata_v1.2.dna.chromosome.4.fa.

2.5. RT-qPCR

Plant's roots and shoots were collected on the seventh day of treatment and immediately frozen in liquid nitrogen and stored at -80 ºC. Each sample was composed of roots or shoots derived from three plants (3 biological replicates and 4 technical replicates). RNA was extracted using TRIzol™ (Invitrogen™) according to the manufacturer's instructions. RNA extractions were quantified by Nanodrop® and treated with DNase I enzyme (Invitrogen™) for genomic DNA removal. cDNA synthesis was performed with 500 ng of RNA from root tissue or 1500 ng of RNA from shoot tissue using M-MLV Reverse Transcriptase (200 U/µL) (Invitrogen™) and stored at -20 ºC. RT-qPCR reaction was carried out with a final volume of 20 µL per well, with 10 µL of cDNA diluted 50 times and 10 µL of PCR solution. The PCR solution contained 49.7 µl of H_2O Milli-Q, 26.5 µl of 10X PCR buffer, 15.9 µl of $MgCl_2$ 50mM, 5.3 µl of primer 10 mM (forward and reverse), 2.7µl dNTPs 10 mM, 26.5 µl of SYBR Green® and 0.7 µl of Platinum Taq DNA polymerase (5 U/µl, Invitrogen®). The RT-qPCR reaction was executed with the StepOne Real-Time PCR Systems (AppliedBiosystems, Foster City, CA, USA) equipment. Expression data was obtained according to the Livak; Schmittgen, (2001) method and PCR efficiency was assessed as described in Ricachenevsky *et al.*, (2011). One specific primer pair was designed to amplify LOC_Os11g15624.1 (F: CTCATCCCTGCCCTCAAACC; R: GTCCACTTCAGTTCCTGCGA) and LOC_Os04g45510.1 (F: CCAAGGTGAGGAAGGAGTGG; R: CATGTACTTGTCCGTCCGGT) from all species evaluated as well as IRO2 (F: CTCCCATCGTTTCGGCTACCT; R: GCTGGGCACTCCTCGTTGATC) and YSL15, as marker genes for iron deficiency response (F: AACATAAGGGGGACTGGTAC; R: TGATTACCGCAATGATGCTTAG) and UBQ5, as normalizer gene (F: AAGCTCGCCGTGCTC; R: AGTAGTGGCGGTCGAAGT).

2.6. Vector cloning

We commercially synthesized LOC_Os11g15624.1 and LOC_Os04g45510.1 CDS sequences (from GenScript) containing Gateway-compatible flanking sequences cloned in pUC57 vector. The target sequence was transferred to the expression vector in plants, pANIC6A, using LR Clonase II (Thermo Fisher Scientific Inc.) (Mann *et al.*, 2012). Both constructs were transformed into *E. coli* strain DH5α. The cloning steps were confirmed by

PCR. For calli transformation, the vector was transformed into *Agrobacterium tumefaciens* strain EHA105 by electroporation as described (Kámán-Tóth *et al.*, 2018).

2.7. Generation of LOC_Os11g15624.1 overexpression lines

We used *A. tumefaciens*-mediated rice transformation as described (Ozawa, 2009) with minor modifications. Nipponbare seeds were sterilized in 70% ethanol solution for 30 min, then introduced in a 4% sodium hypochlorite solution for 30 min and rinsed in sterile distilled water five times. Sterile seeds were placed on solid N6D media for four weeks. After that, healthy calli were selected and transferred to solid NB media for five days (before transformation). *A. tumefaciens* transformed with the expression vector were cultivated for two days in solid LB selective media, with the addition of kanamycin and rifampicin. The co-cultured calli and Agrobacterium were cultivated in solid DKND medium supplemented with glucose (5 g/L), L-cysteine (100 mg/L) and acetosyringone (15 mg/L) for two days before being transferred to the solid DKND selective medium, supplemented with hygromycin (50mg/ml), timentin (300mg/ml) and cefotaxin (200mg/ml). Regenerated shoots were transferred to solid MS medium containing sucrose (30 g/L) and hygromycin (50mg/ml). All other steps were performed as described (Ozawa, 2009).

2.8. Statistical analysis

The experiments were conducted according to a random distribution arrangement. The data obtained were subjected to statistical analysis using the t test and ANOVA ($p \leq 0.05$). Statistical analyses were performed using GraphPad Prism software, version 8.00 (http://www.graphpad.com).

3. Results and Discussion

3.1. Public transcriptome data

We found LOC_Os11g15624 and LOC_Os04g45510 as possible iron regulated uncharacterized genes based on their different expression values fond on transcriptome data of *O. sativa* ssp. *japonica* plants under control, iron excess and deficiency conditions (Wairich *et al.*, 2019, 2021) (Table 1). Other transcriptome-related studies also documented the opposite iron-dependent expression pattern for the genes of interest

(Drogue *et al.*, 2014; Gonçalves *et al.*, 2020). We hypothesized that these genes might be involved in *O. sativa* response to iron starvation and toxicity and decided to investigate if the expressed proteins established homology with proteins from the *O. sativa* genome itself and other species as well as confirm their expression in *O. sativa* and other *Oryza* species.

Table 1 - Transcriptome analysis data of *O. sativa* ssp. *japonica* from Wairich *et al.*, 2019, 2021.

Gene ID	**Treatment**	**q value**	**Log2FC**	**CC (mean)**	**Treatment (mean)**	**Sample size (n)**
LOC_Os11g15624	+Fe	0,0142008	-7,4914298	32,666667	0	3
LOC_Os11g15624	-Fe	0,0146708	3,8152655	472,54873	6652,0569	2
LOC_Os04g45510	+Fe	0	7,3848288	2	341	3
LOC_Os04g45510	-Fe	0,18	-Inf	10,747509	0	2

Gene IDs were converted to the MSU format. "CC" and "Treatment" values are presented as the mean value between samples and are relative to the number of transcripts per million. The Log2FC (fold change) was calculated as follows: log2 of average transcript amount under Fe treatment / average transcript amount under control treatment. Positive Log2FC values reflect up-regulation of a gene under a specific treatment, while negative values reflect down-regulation.

3.2. Alternative splicing analysis

As mentioned before, the LOC_Os11g15624 gene has three alternative splicing variants, so we analyzed public RNA-seq data to focus our work on the most represented variant in nature. According to RNA-seq data referring to *O. sativa* ssp. *japonica* cultivated in control conditions, the variant with the highest number of reads matches the exon positions of LOC_Os11g15624.1 (data not shown), so we chose to focus our attention on that variant for the remainder of our work.

3.3. BLAST analysis

Our BLASTp search with LOC_Os11g15624.1 protein sequence as query found 28 genes that met the cutoff criteria (Supplementary Table 1). All species belonged to the *Oryza* genus. *O. sativa* and *O. glaberrima* genome had 5 hits (LOC_Os11g15624.1 included), *O. rufipogon* had 6 hits, *O. glumaepatula* and *O. nivara* had 4 hits, *O. barthii* had 2 hits and *O. punctata* and *O. meridionalis* had 1 hit.

Our BLASTp search with LOC_Os04g45510.1 protein sequence as query found 30 genes that met the cutoff criteria (Supplementary Table 2). All species were monocots, with 25 hits from the *Poaceae* family, 3 hits from the *Joinvilleaceae* family, 1 hit from the *Bromeliaceae* family and 1 hit from the *Acoraceae* family. With the exception of *Joinvillea ascendens* genome, which had 3 hits, all other species only had 1 hit.

3.4. Phylogenetic tree

For the LOC_Os11g15624.1 phylogenetic tree, we used sequences retrieved with the SPLIGN tool and focused on sequences that displayed homology between at least two exons. Sequences that only showed homology with one exon were not included.

As shown below (Figure 1), the LOC_Os11g15624.1 gene appears to be recent in evolutionary terms. Its ancestral homology seems to be restricted to the *Oryza* genus. Furthermore, we observed that LOC_Os11g15624.1 homologs are present in AA genome *Oryza* species and possibly a few more distantly related species (*O. punctata* – BB genome). Analyzing Figure 1, we noticed that *O. sativa* and *O. rufipogon* gene sequences show high similarity. This is supported by the fact that *O. rufipogon* is one of the *O. sativa*'s presumed ancestors (Menguer; Sperotto; Ricachenevsky, 2017). *O. glaberrima* and *O. barthii*"s sequences also have high similarity among themselves and to *O. sativa* and *O. rufipogon*'s sequences. It is interesting to notice that sequences with high similarity are on the same chromosome. *O. nivara* sequence is in between, as most of its sequences appear closely related to the species, but its sequences are in different chromosomes most of the time. *O. glumaepatula, O. longistaminata, O. punctata* and *O. meridionalis* sequences seem to be more phylogenetically distant. We also found multiple copies (with at least two homologous exons) present in each species. Our results do not allow us to suggest a hypothesis on how several copies were established in the genomes of such closely related species. However, we demonstrated that LOC_Os11g15624.1 is a recently evolved gene, which seems to be generating multiple copies in the genomes of *Oryza*.

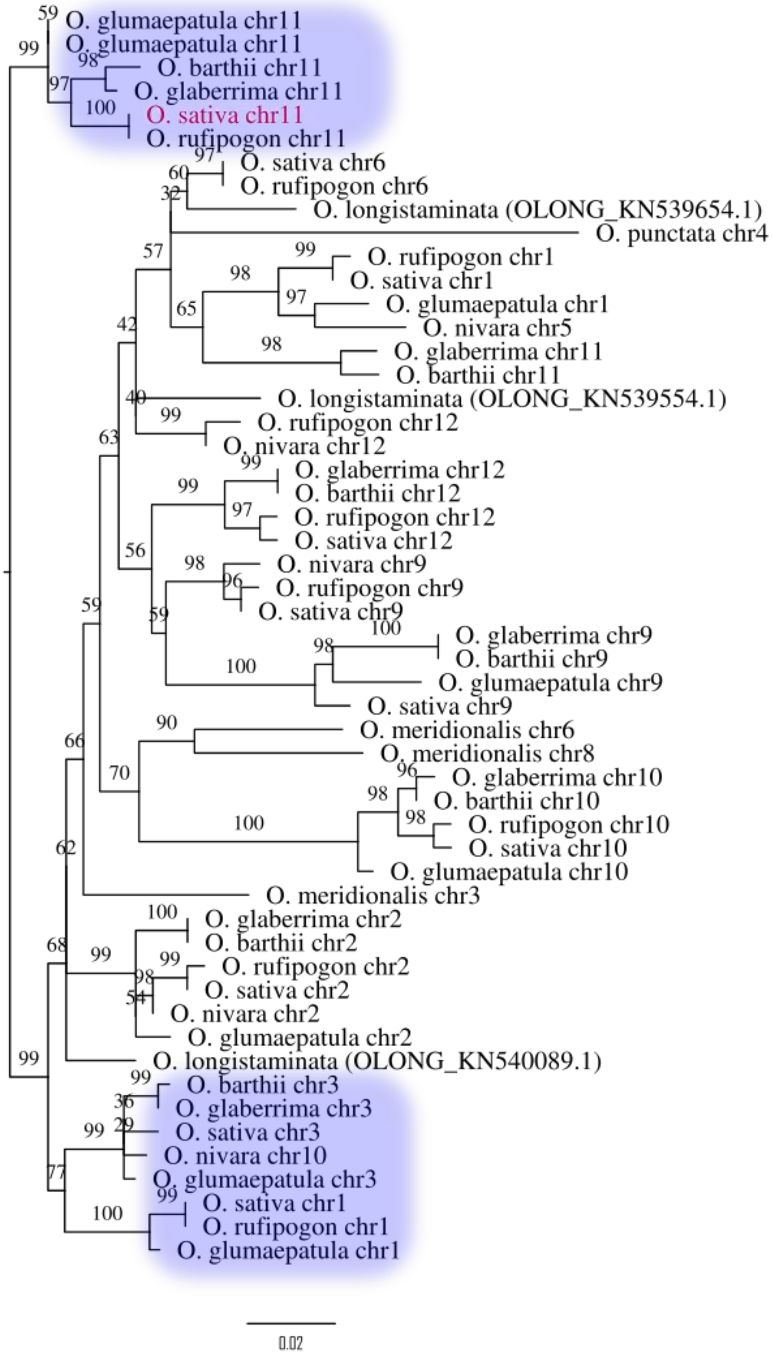
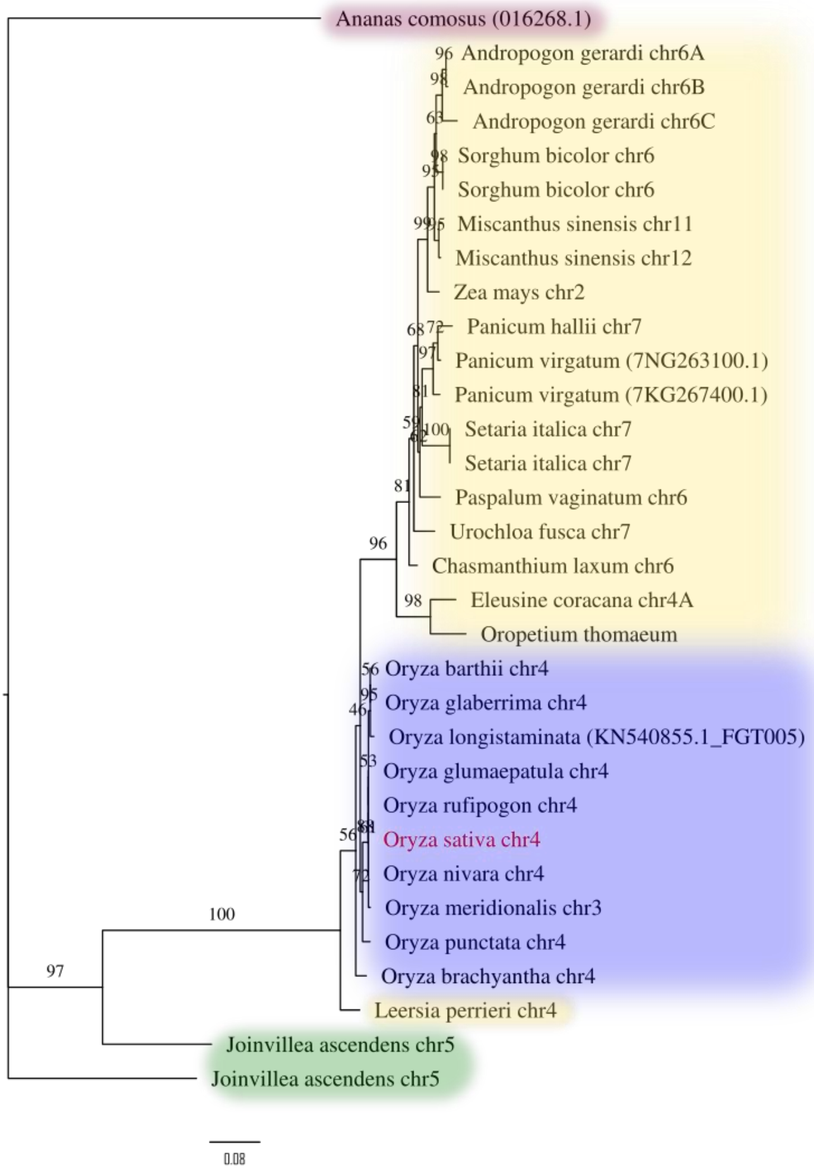


Figure 1 – Proposed phylogenetic tree for the evolution of the LOC_Os11g15624.1 gene. The branch containing the gene of interest was chosen as a rooting point. The scale and distance between species is calculated based on the number of substitutions per site. Numbers on branches represent the bootstrap value. The LOC_Os11g15624.1 gene is written in red font. Sequences highlighted in purple presented homology

with 3 exons. The tree was generated through the IQTREE software and visualized through the Fig tree software.

For LOC_Os04g45510.1 phylogeny we chose to work with two possibilities of trees (Figure 2). In Figure 2a we included outgroups that showed lower similarity with LOC_Os04g45510.1 so we could be more certain of tree rooting. In Figure 2b we excluded outgroup sequences to generate a tree with greater branch support. LOC_Os04g45510.1 gene seems to have evolved earlier in *Poaceae*, displaying fewer copies in the genomes of each species, since it is found in monocots outside *Poaceae* (Figure 2b). Interestingly, LOC_Os04g45510.1 has only one copy in *O. sativa*, a feature that is shared with most species in which the gene has homologous sequences. Except for *O. meridionalis*, all *Oryza* species display their copy of LOC_Os04g45510.1 orthologous gene in chromosome 4. This shows once again the level of conservation between species that share homology with LOC_Os04g45510.1, suggesting that, contrary to LOC_Os11g15624.1, it seems to be conserved within the same genomic environment and as a single copy gene.

a)



Ananas comosus (016268.1)
Andropogon gerardi chr6A
Andropogon gerardi chr6B
Andropogon gerardi chr6C
Sorghum bicolor chr6
Sorghum bicolor chr6
Miscanthus sinensis chr11
Miscanthus sinensis chr12
Zea mays chr2
Panicum hallii chr7
Panicum virgatum (7NG263100.1)
Panicum virgatum (7KG267400.1)
Setaria italica chr7
Setaria italica chr7
Paspalum vaginatum chr6
Urochloa fusca chr7
Chasmanthium laxum chr6
Eleusine coracana chr4A
Oropetium thomaeum
Oryza barthii chr4
Oryza glaberrima chr4
Oryza longistaminata (KN540855.1_FGT005)
Oryza glumaepatula chr4
Oryza rufipogon chr4
Oryza sativa chr4
Oryza nivara chr4
Oryza meridionalis chr3
Oryza punctata chr4
Oryza brachyantha chr4
Leersia perrieri chr4
Joinvillea ascendens chr5
Joinvillea ascendens chr5
96
98
63
98
95
99
95
68
72
97
81
59
100
81
96
98
56
95
46
53
56
72
100
97
0.08

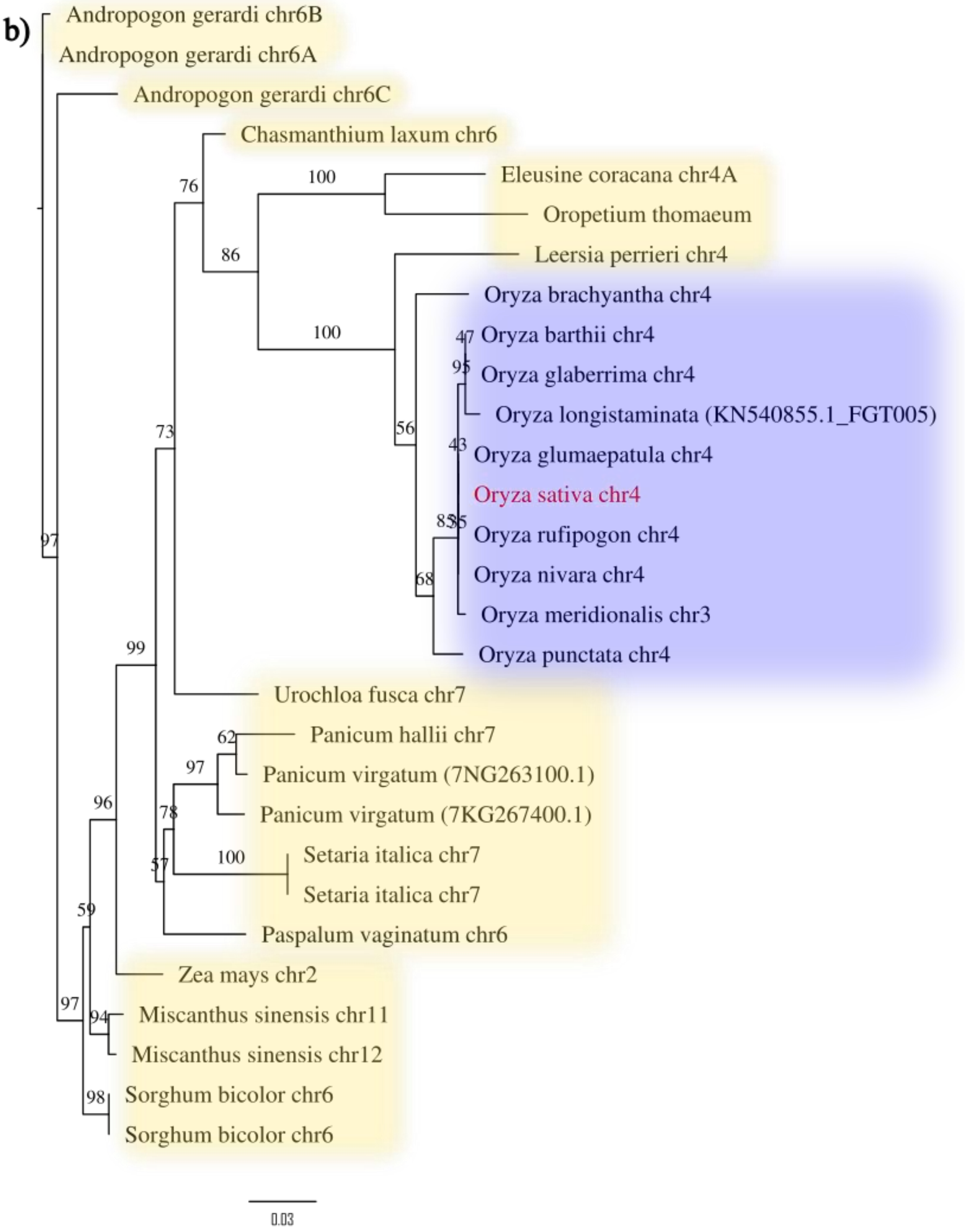
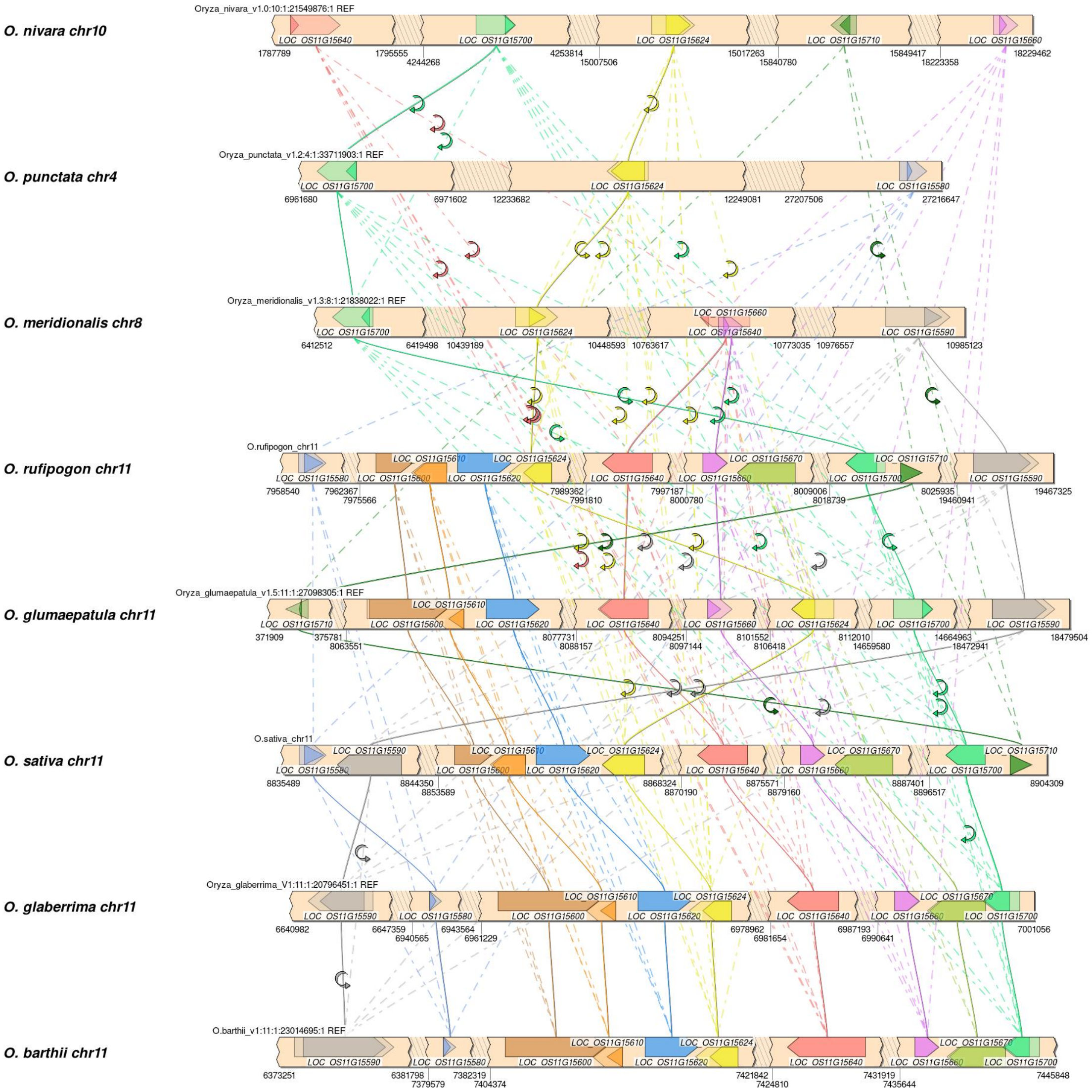


Figure 2 – Proposed phylogenetic tree for the evolution of the LOC_Os04g45510.1 gene. (a) and (b) have different rooting points (*Ananas comosus* and *Andropogon gerardi*, respectively). The scale and distance between species is calculated based on the number of substitutions per site. Numbers on branches represent the bootstrap value. The LOC_Os11g15624.1 gene is written in red font. Sequences highlighted in purple belong to the *Oryza* genus (and *Poaceae* family), yellow to *Poaceae*, green to *Joinvilleaceae* and red to *Bromeliaceae*. The tree was generated through the IQTREE software and visualized through the Fig tree software.

3.5. Synteny and Collinearity

The synteny and collinearity analysis were carried out to understand how the gene environment was altered during the evolution of the two uncharacterized genes. We chose to include the chromosomes that contained the best BLAST hit for LOC_Os11g15624.1. As shown below (Figure 3), *O. rufipogon, O. glumaepatula, O. glaberrima* and *O. barthii* all had their best BLAST hit (for LOC_Os11g15624.1) in chromosome 11, displaying a syntenic relationship with *O. sativa*. Considering the LOC_Os11g15624.1 gene and the ten *in tandem* genes, *O. rufipogon, O. glaberrima* and *O. barthii* were syntenic and collinear with *O. sativa*. On the other hand, *O. meridionalis, O. nivara* and *O. punctata* had their best BLAST hit (for LOC_Os11g15624.1) in different chromosomes (8, 10 and 4, respectively) and most of the hits for the *in tandem* genes were not located in that chromosome. Except for *O. punctata*, which is part of the BB or BBCC genome group (Menguer; Sperotto; Ricachenevsky, 2017), all other analyzed species are of the AA genome group (Menguer; Sperotto; Ricachenevsky, 2017), so we expected a clearer syntenic and collinear relationship between all species, especially between *O. sativa, O. rufipogon* and *O. nivara* which are phylogenetic close species (Gao *et al.*, 2019; Nishikawa; Vaughan; Kadowaki, 2005) and *O. rufipogon* and *O. nivara* are O. sativa's presumed ancestors (Menguer; Sperotto; Ricachenevsky, 2017). Therefore, our data suggest that, besides being a newly evolved gene that has emerged recently within the *Oryza* genus, LOC_Os11g15624 might be changing genomic location, suggesting that its genomic context might not yet be fixed.


O. nivara chr10
Oryza_nivara_v1.0:10:1:21549876:1 REF
LOC_OS11G15640
LOC_OS11G15700
LOC_OS11G15624
LOC_OS11G15710
LOC_OS11G15660
1787789
1795555
4244268
4253814
15007506
15017263
15840780
15849417
18223358
18229462
O. punctata chr4
Oryza_punctata_v1.2:4:1:33711903:1 REF
LOC_OS11G15700
LOC_OS11G15624
LOC_OS11G15580
6961680
6971602
12233682
12249081
27207506
27216647
O. meridionalis chr8
Oryza_meridionalis_v1.3:8:1:21838022:1 REF
LOC_OS11G15700
LOC_OS11G15624
LOC_OS11G15660
LOC_OS11G15640
LOC_OS11G15590
6412512
6419498
10439189
10448593
10763617
10773035
10976557
10985123
O. rufipogon chr11
O.rufipogon_chr11
LOC_OS11G15580
LOC_OS11G15610
LOC_OS11G15600
LOC_OS11G15624
LOC_OS11G15620
LOC_OS11G15640
LOC_OS11G15660
LOC_OS11G15670
LOC_OS11G15710
LOC_OS11G15700
LOC_OS11G15590
7958540
7962367
7975566
7989362
7991810
7997187
8000780
8009006
8018739
8025935
19460941
19467325
O. glumaepatula chr11
Oryza_glumaepatula_v1.5:11:1:27098305:1 REF
LOC_OS11G15710
LOC_OS11G15610
LOC_OS11G15600
LOC_OS11G15620
LOC_OS11G15640
LOC_OS11G15660
LOC_OS11G15624
LOC_OS11G15700
LOC_OS11G15590
371909
375781
8063551
8077731
8088157
8094251
8097144
8101552
8106418
8112010
14659580
14664963
18472941
18479504
O. sativa chr11
O.sativa_chr11
LOC_OS11G15590
LOC_OS11G15580
LOC_OS11G15610
LOC_OS11G15600
LOC_OS11G15624
LOC_OS11G15620
LOC_OS11G15640
LOC_OS11G15670
LOC_OS11G15660
LOC_OS11G15710
LOC_OS11G15700
8835489
8844350
8853589
8868324
8870190
8875571
8879160
8887401
8896517
8904309
O. glaberrima chr11
Oryza_glaberrima_V1:11:1:20796451:1 REF
LOC_OS11G15590
LOC_OS11G15580
LOC_OS11G15610
LOC_OS11G15600
LOC_OS11G15624
LOC_OS11G15620
LOC_OS11G15640
LOC_OS11G15670
LOC_OS11G15660
LOC_OS11G15700
6640982
6647359
6940565
6943564
6961229
6978962
6981654
6987193
6990641
7001056
O. barthii chr11
O.barthii_v1:11:1:23014695:1 REF
LOC_OS11G15590
LOC_OS11G15580
LOC_OS11G15610
LOC_OS11G15600
LOC_OS11G15624
LOC_OS11G15620
LOC_OS11G15640
LOC_OS11G15670
LOC_OS11G15660
LOC_OS11G15700
6373251
6381798
7379579
7382319
7404374
7421842
7424810
7431919
7435644
7445848

● LOC_Os11g15624 - gene of interest

● LOC_Os11g15620 - OsFBX420

● LOC_Os11g15610 - hypothetical protein

● LOC_Os11g15600 - hypothetical protein

● LOC_Os11g15590 - expressed protein

● LOC_Os11g15580 - hypothetical protein

● LOC_Os11g15640 - expressed protein

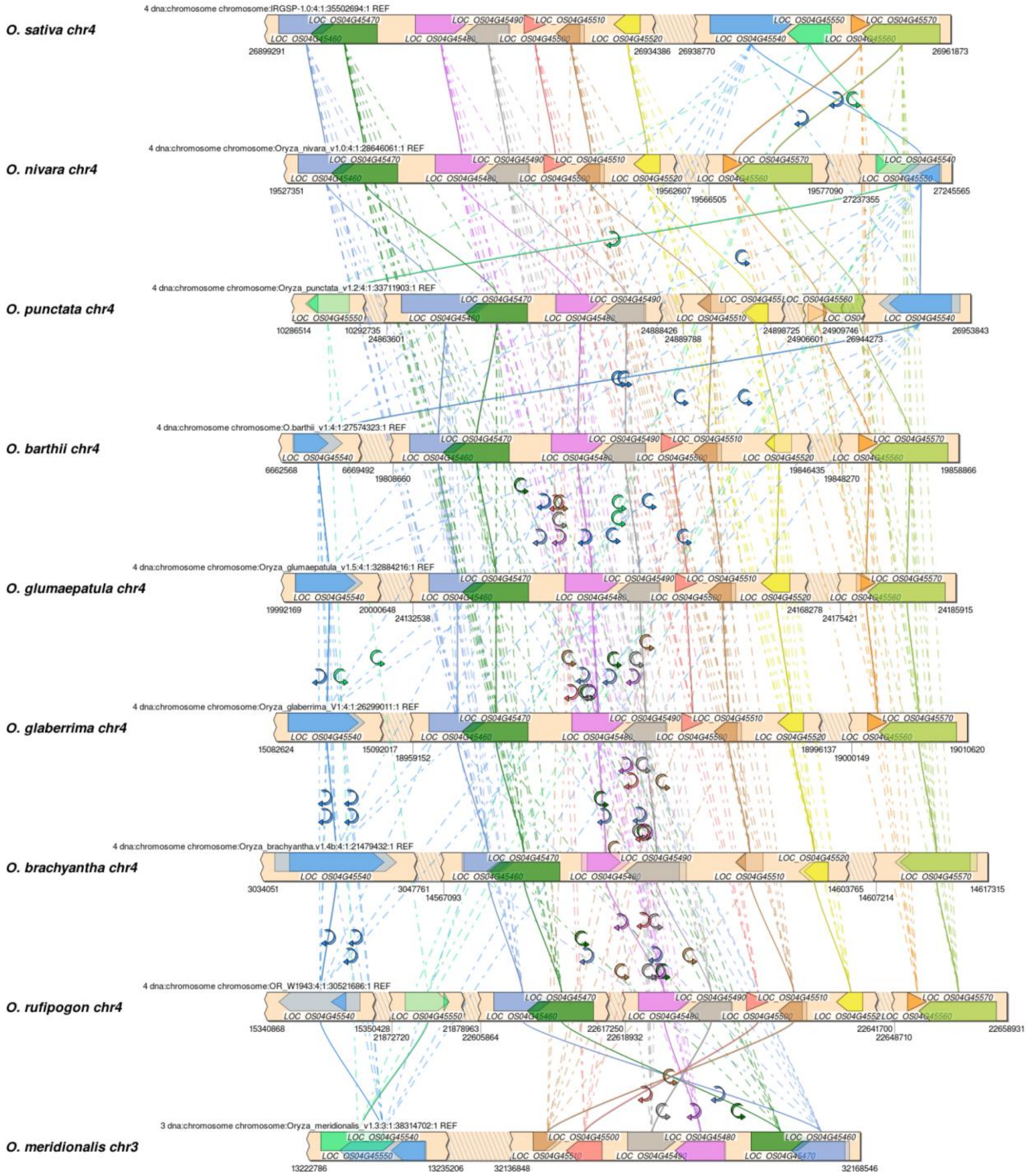
● LOC_Os11g15660 - hypothetical protein

● LOC_Os11g15670 - NBS-LRR disease resistance protein

● LOC_Os11g15700 - NB-ARC domain containing protein

● LOC_Os11g15710 - hypothetical protein

Figure 3 – Synteny and Collinearity analysis result generated by the Symple Synteny software. Species names and chromosome numbers are shown at the right side of the figure. Gene IDs are referent to the *Oryza sativa* ssp. *japonica* reference genome. The lower portion of the figure shows the relation between gene ID and the color they are represented by, as well as the gene product description acquired from the *Rice Genome Annotation Project* database. The shaded portion of the colored boxes represents the length of matching nucleotide sequences of the BLAST hit. The translucent portion is referent to the length of the annotated gene. Arrows indicate inversion of transcription direction. The LOC_Os11g15624.1 gene is shown in brown.


O. sativa chr4
4 dna:chromosome chromosome:IRGSP-1.0:4:1:35502694:1 REF
LOC_OS04G45470
LOC_OS04G45490
LOC_OS04G45510
LOC_OS04G45550
LOC_OS04G45570
LOC_OS04G45460
LOC_OS04G45480
LOC_OS04G45500
LOC_OS04G45520
LOC_OS04G45540
LOC_OS04G45560
26899291
26934386
26938770
26961873
O. nivara chr4
4 dna:chromosome chromosome:Oryza_nivara_v1.0:4:1:28646061:1 REF
LOC_OS04G45470
LOC_OS04G45490
LOC_OS04G45510
LOC_OS04G45570
LOC_OS04G45540
LOC_OS04G45460
LOC_OS04G45480
LOC_OS04G45500
LOC_OS04G45520
LOC_OS04G45560
LOC_OS04G45550
19527351
19562607
19566505
19577090
27237355
27245565
O. punctata chr4
4 dna:chromosome chromosome:Oryza_punctata_v1.2:4:1:33711903:1 REF
LOC_OS04G45470
LOC_OS04G45490
LOC_OS04G4552
LOC_OS04G45560
LOC_OS04G45550
LOC_OS04G45460
LOC_OS04G45480
LOC_OS04G45510
LOC_OS04
LOC_OS04G45540
10286514
10292735
24863601
24888426
24889788
24898725
24906601
24909746
26944273
26953843
O. barthii chr4
4 dna:chromosome chromosome:O.barthii_v1:4:1:27574323:1 REF
LOC_OS04G45470
LOC_OS04G45490
LOC_OS04G45510
LOC_OS04G45570
LOC_OS04G45540
LOC_OS04G45460
LOC_OS04G45480
LOC_OS04G45500
LOC_OS04G45520
LOC_OS04G45560
6662568
6669492
19808660
19846435
19848270
19858866
O. glumaepatula chr4
4 dna:chromosome chromosome:Oryza_glumaepatula_v1.5:4:1:32884216:1 REF
LOC_OS04G45470
LOC_OS04G45490
LOC_OS04G45510
LOC_OS04G45570
LOC_OS04G45540
LOC_OS04G45460
LOC_OS04G45480
LOC_OS04G45500
LOC_OS04G45520
LOC_OS04G45560
19992169
20000648
24132538
24168278
24175421
24185915
O. glaberrima chr4
4 dna:chromosome chromosome:Oryza_glaberrima_V1:4:1:26299011:1 REF
LOC_OS04G45470
LOC_OS04G45490
LOC_OS04G45510
LOC_OS04G45570
LOC_OS04G45540
LOC_OS04G45460
LOC_OS04G45480
LOC_OS04G45500
LOC_OS04G45520
LOC_OS04G45560
15082624
15092017
18959152
18996137
19000149
19010620
O. brachyantha chr4
4 dna:chromosome chromosome:Oryza_brachyantha.v1.4b:4:1:21479432:1 REF
LOC_OS04G45470
LOC_OS04G45490
LOC_OS04G45520
LOC_OS04G45540
LOC_OS04G45460
LOC_OS04G45480
LOC_OS04G45510
LOC_OS04G45570
3034051
3047761
14567093
14603765
14607214
14617315
O. rufipogon chr4
4 dna:chromosome chromosome:OR_W1943:4:1:30521686:1 REF
LOC_OS04G45470
LOC_OS04G45490
LOC_OS04G45510
LOC_OS04G45570
LOC_OS04G45540
LOC_OS04G45550
LOC_OS04G45460
LOC_OS04G45480
LOC_OS04G45500
LOC_OS04G4552
LOC_OS04G45560
15340868
15350428
21872720
21878963
22605864
22617250
22618932
22641700
22648710
22658931
O. meridionalis chr3
3 dna:chromosome chromosome:Oryza_meridionalis_v1.3:3:1:38314702:1 REF
LOC_OS04G45540
LOC_OS04G45500
LOC_OS04G45480
LOC_OS04G45460
LOC_OS04G45550
LOC_OS04G45510
LOC_OS04G45490
LOC_OS04G45470
13222786
13235206
32136848
32168546

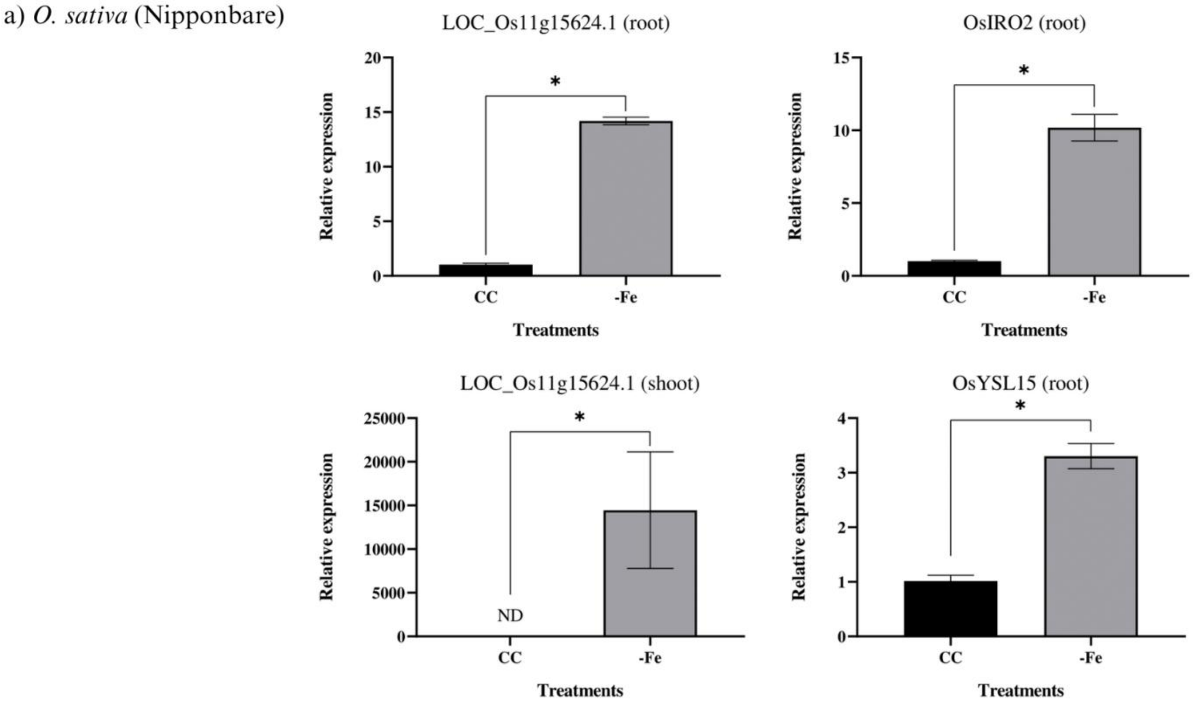
● LOC_Os04g45510 - gene of interest
● LOC_Os04g45500 - expressed protein
● LOC_Os04g45490 - elongation factor
● LOC_Os04g45480 - heat shock protein STI
● LOC_Os04g45470 - vacuolar-processing enzyme precursor
● LOC_Os04g45460 - cysteine-rich repeat secretory protein precursor
● LOC_Os04g45520 - integral membrane protein
● LOC_Os04g45540 - retrotransposon protein
● LOC_Os04g45550 - retrotransposon protein
● LOC_Os04g45560 - hypothetical protein
● LOC_Os04g45570 - expressed protein

Figure 4 – Synteny and Collinearity analysis result generated by the Symple Synteny software. Species names and chromosome numbers are shown at the right side of the figure. Gene IDs are referent to the *Oryza sativa* ssp. *japonica* reference genome. The lower portion of the figure shows the relation between gene ID and the color they are represented by, as well as the gene product description acquired from the *Rice Genome Annotation Project* database. The shaded portion of the coloured boxes represents the length of matching nucleotide sequences of the BLAST hit. The translucent portion is referent to the length of the annotated gene. Arrows indicate inversion of transcription direction. The LOC_Os04g45510.1 gene is shown in yellow.

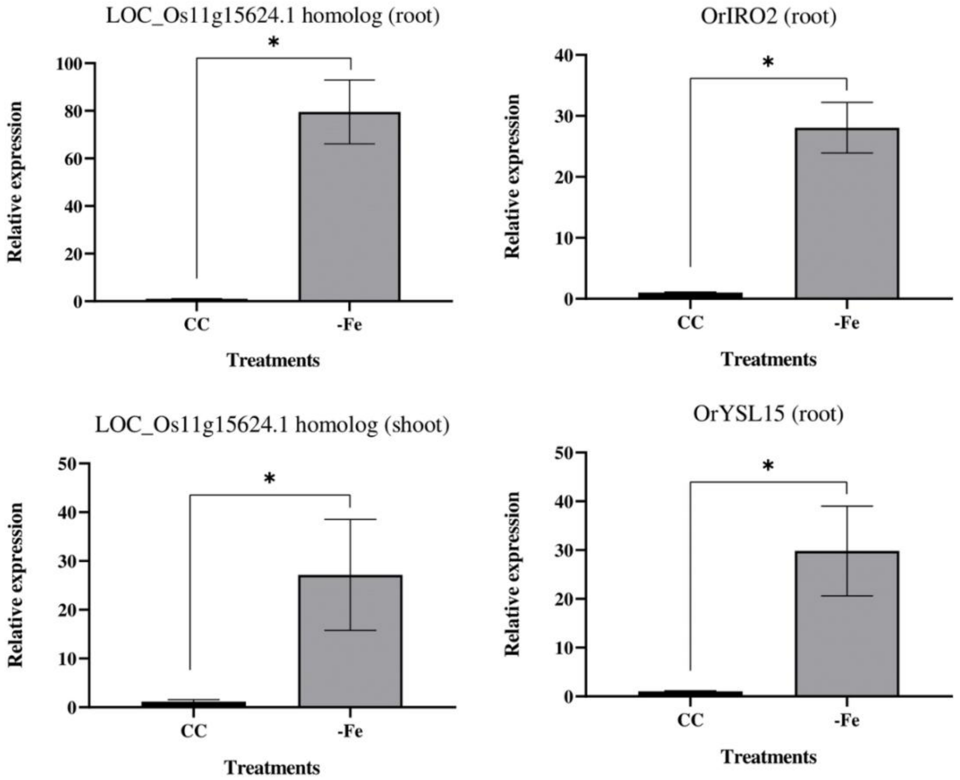
When compared to LOC_Os11g15624.1, the LOC_Os04g45510.1 analysis displayed a clearer synteny and collinearity relationship between *O. sativa* and other *Oryza* species (Figures 4). *O. sativa, O. glumaepatula, O. rufipogon, O. nivara, O. glaberrima* and *O. barthii* seem to share syntenic and collinear relationships. *O. brachyantha* and *O. punctata* also share syntenic and collinear relationship (with the exception of two genes) with *O. sativa*. *O. meridionalis*, on the other hand, had the weakest syntenic and collinear relationship with *O. sativa* and was the only species that did not have the best BLAST hit for LOC_Os04g45510.1 in chromosome 4 (the best BLAST hit was located in chromosome 3). It is interesting to notice that among the AA genome species included in the analysis *O. sativa, O. glumaepatula, O. rufipogon, O. nivara, O. glaberrima* and *O. barthii* displayed cleaer syntenic and collinear relationships, whereas *O. meridionalis* had the same gene in a different genomic context. This result shows that the region surrounding LOC_Os04g45510.1 is highly conserved in *Oryza* genus AA type genome species and that the LOC_Os04g45510.1 genomic sequence might have been translocated to the third chromosome of *O. meridionalis* during its speciation, breaking synteny and collinearity with other species. Considering that *O. meridionalis* was the first AA genome species to separate from the others, it is likely that it occurred after the split.

3.6. Gene expression

In order to confirm that LOC_Os11g15624.1 is expressed in Nipponbare and its orthologs in *Oryza* genus species, as well as confirm that iron deficiency positively regulates their expression, we cultivated plants from *Oryza sativa*, *O. rufipogon* and *O. meridionalis* in control and iron deficient conditions for seven days. We used known iron deficiency up-regulated genes as markers, *OsIRO2* and *OsYSL15*, as already described (Wairich *et al.*, 2019). As observed in Figure 5, *OsIRO2* and *OsYSL15* from all three species were up-regulated in roots of plants under iron deficiency compared to control plants (Figure 5a-c). Therefore, we confirmed that our treatments resulted in iron deficiency.



b) *O. rufipogon*



c) *O. meridionalis*

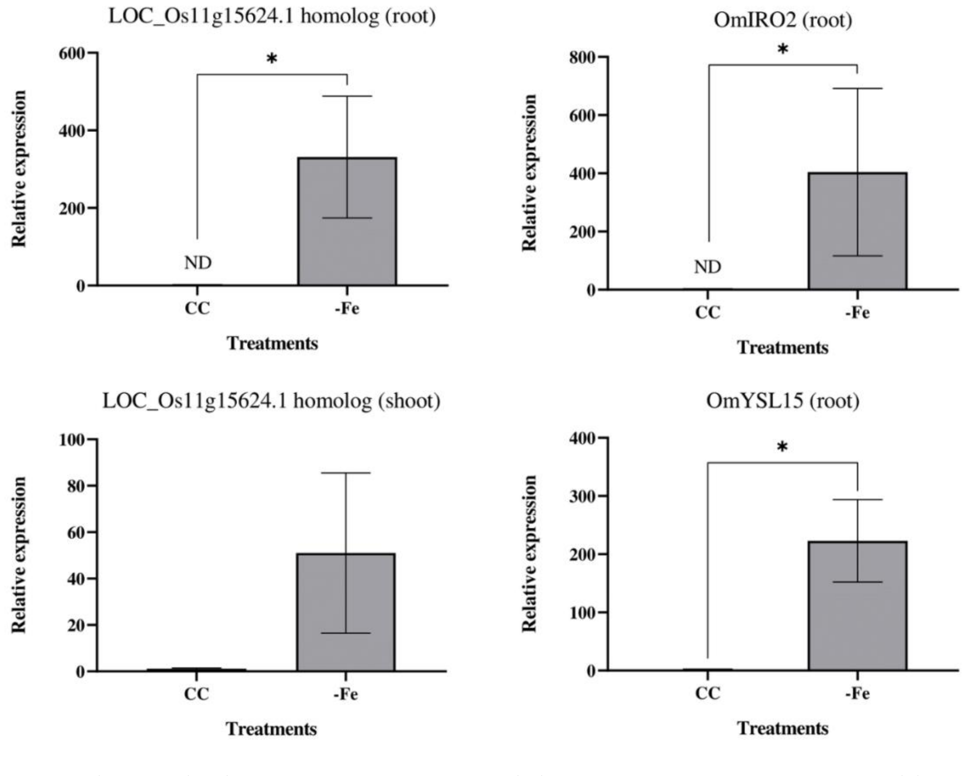


Figure 5 – RT-qPCR expression results for OsIRO2, OsYSL15 and the LOC_Os11g15624.1 gene and its homologs in (a) *O. sativa* (Nipponbare), (b) *O. rufipogon* and (c) *O. meridionalis*. Gene expression values are relative to OsUBQ5. Control (CC) and iron deficiency (-Fe) treatments are shown in the x axis.

Statistically significant different values ($p \leq 0,05$) were indicated by an asterisk (*). Relative expression values below 1 were annotated as ND (not detected).

As shown above (Figure 5), LOC_Os11g15624.1 and its homologs in *O. rufipogon* and *O. meridionalis* were up regulated under iron deficiency conditions, both in roots and shoots. LOC_Os11g15624.1 relative expression is much higher in iron deficiency treatment than in control treatment. In all species, except for *Oryza sativa*, relative expression values for LOC_Os11g15624.1 were higher in roots compared to shoots. The expression pattern of LOC_Os11g15624.1 suggests that the expressed protein has a role in rice iron starvation response.

4. Transformation of rice calli

LOC_Os11g15624.1 and LOC_Os04g45510.1 CDS sequences containing Gateway-compatible flanking sequences were cloned in pUC57 vector. The target sequence was transferred to pANIC6A. Both constructs were transformed into *E. coli* strain DH5α. The cloning steps were confirmed by PCR (Figure 6). For calli transformation, the vector was transformed into *Agrobacterium tumefaciens* strain EHA105 by electroporation for future *A. tumefaciens*-mediated rice transformation.

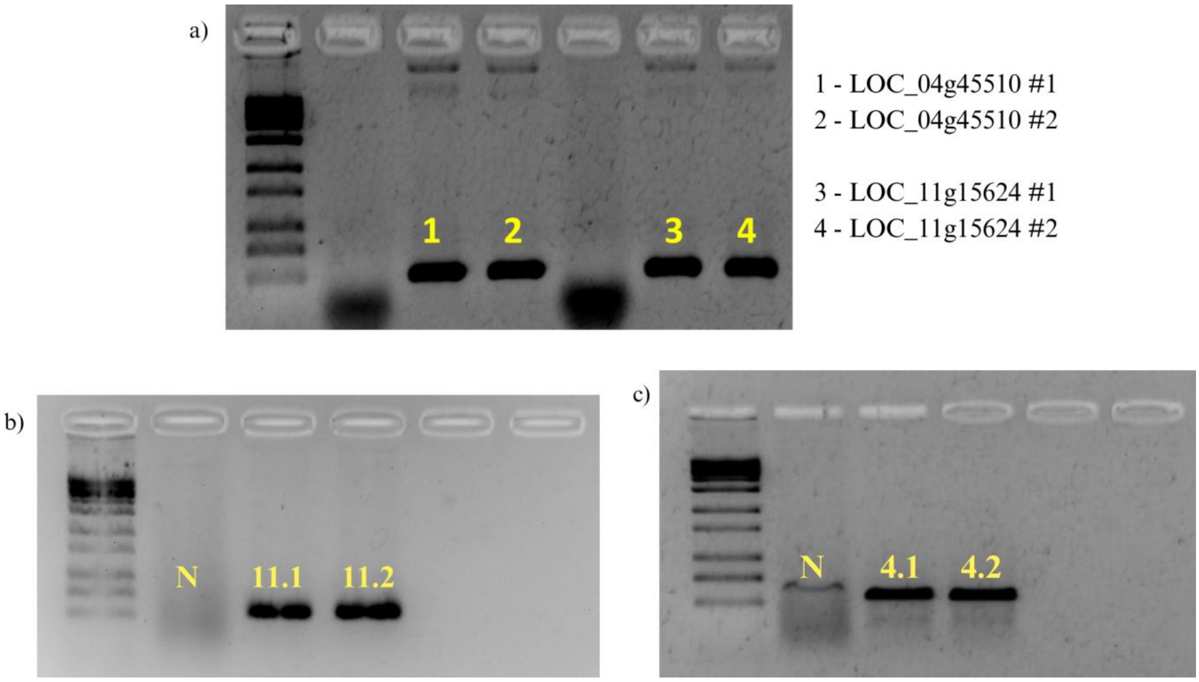


Figure 6 - (a) - Agar gel for PCR confirmation of transformation of DH5α (4 independent transformation events) with pANIC6A expression vector containing either LOC_Os11g15624.1 or LOC_Os04g45510.1 CDS sequence. Primers are shown in Supplementary Table 3. Amplicon size is approximately 500 bp. (b)

and (c) – Agar gels for PCR confirmation of transformation (4 independent transformation events - 11.1, 11.2, 4.1 and 4.2) of *Agrobacterium tumefaciens* strain EHA105 with pANIC6A expression vector containing either LOC_Os11g15624.1 (11.1 and 11.2) or LOC_Os04g45510.1 (4.1 and 4.2) CDS sequences by electroporation. The N labeled bands represent the negative controls. Primers used here are the same as described in the RT-qPCR methodology. Amplicon sizes are approximately 500 bp.

Transformation of rice plants with LOC_Os11g15624.1 and LOC_Os04g45510.1 genes are in progress. We recently regenerated the first transgenic line overexpressing LOC_Os11g15624.1 gene. We generated seeds and will soon analyze gene expression and hygromycin resistance in this particular line.

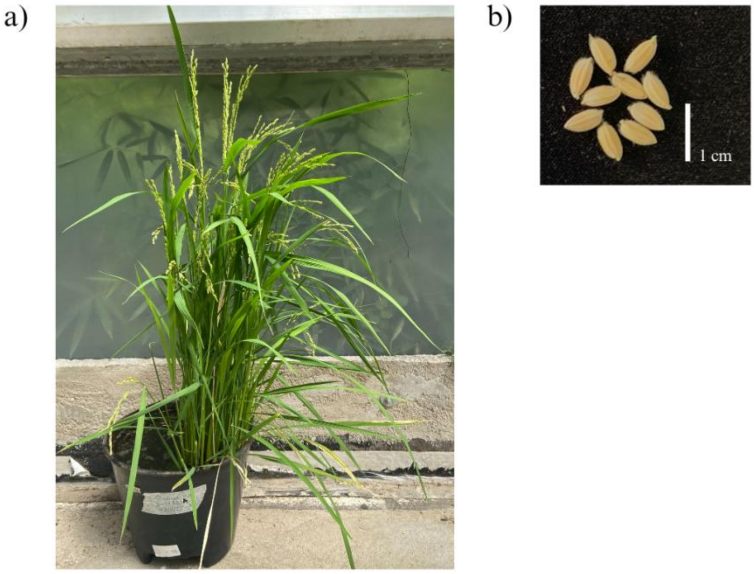


Figure 7 - (a) Nipponbare plant LOC_Os11g15624.1 overexpression transgenic lineage 11.23 first clone being cultivated in soil. (b) first seeds to be collected from the transgenic plant. Scale bar indicating 1 cm.

5. Future perspectives

Rice plant calluses for overexpression were transformed with *Agrobacterium* and are currently in the selection step. After regeneration, plants will be cultivated and submitted to different treatments (involving different concentrations of iron in the soil) with the objective to induce iron excess and starvation responses, allowing the analysis of the effects of the overexpression of both genes (separately) on the rice plant's response to iron stress and the confirmation of their expression patterns. We are also performing experiments using *Oryza* species in order to verify whether the expression of LOC_Os04g45510.1 is iron regulated.

6. Conclusion

Our ongoing work successfully confirmed the regulation of expression of the LOC_Os11g15624.1 gene under iron deficiency and control conditions in rice and its homologs in *O. rufipogon* and *O. meridionalis*. We also provide a better understanding of the evolution of LOC_Os11g15624.1, LOC_Os04g45510.1 and its homologs, showing that they are evolutionary novelties that arise at different times during monocot, *Poaceae* and *Oryza* evolution. The LOC_Os11g15624.1 gene seems to be a genetic novelty and we propose that it is restricted to a few species from the *Oryza* genus, and is still changing location and copy numbers in the genomes. The LOC_Os04g45510.1 gene, in contrast, seems to have evolved earlier in the evolution of monocots, but has restricted expansion in copy number.

7. References

ANDERSSON, D. I.; JERLSTRÖM-HULTQVIST, J.; NÄSVALL, J. Evolution of New Functions De Novo and from Preexisting Genes. **Cold Spring Harbor Perspectives in Biology**, [*s. l.*], v. 7, n. 6, p. a017996, 2015.

BECKER, M.; ASCH, F. Iron toxicity in rice—conditions and management concepts. **Journal of Plant Nutrition and Soil Science**, [*s. l.*], v. 168, n. 4, p. 558–573, 2005.

CAMASCHELLA, C. Iron deficiency. **Blood**, [*s. l.*], v. 133, n. 1, p. 30–39, 2019.

CHEN, J. *et al.* Whole-genome sequencing of Oryza brachyantha reveals mechanisms underlying Oryza genome evolution. **Nature Communications**, [*s. l.*], v. 4, n. 1, p. 1595, 2013.

COÊLHO, J. D. Arroz: produção e mercado. [*s. l.*], 2021. Disponível em: https://bnb.gov.br/s482-dspace/handle/123456789/s1dspp01.dmz.bnb:8443/s482-dspace/handle/123456789/698. Acesso em: 29 fev. 2024.

CONAB. **Boletim da Safra de Grãos**. [*s. l.*], 2024. Disponível em: http://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos. Acesso em: 7 fev. 2024.

DE OLIVEIRA, B. H. N. *et al.* The Mitochondrial Iron-Regulated (MIR) gene is Oryza genus specific and evolved before speciation within the Oryza sativa complex. **Planta**, [*s. l.*], v. 251, n. 5, p. 94, 2020.

DROGUE, B. *et al.* Plant root transcriptome profiling reveals a strain-dependent response during Azospirillum-rice cooperation. **Frontiers in Plant Science**, [*s. l.*], v. 5, 2014. Disponível em: https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2014.00607. Acesso em: 24 fev. 2024.

GAO, L.-Z. *et al.* Evolution of Oryza chloroplast genomes promoted adaptation to diverse ecological habitats. **Communications Biology**, [*s. l.*], v. 2, n. 1, p. 1–13, 2019.

GNANAMANICKAM, S. S. Rice and Its Importance to Human Life. *In*: GNANAMANICKAM, S. S. (org.). **Biological Control of Rice Diseases**. Dordrecht:

Springer Netherlands, 2009. (Progress in Biological Control). p. 1–11. Disponível em: https://doi.org/10.1007/978-90-481-2465-7_1. Acesso em: 7 fev. 2024.

GONÇALVES, B. X. *et al.* Phosphate starvation responses in crop roots: from well-known players to novel candidates. **Environmental and Experimental Botany**, [*s. l.*], v. 178, p. 104162, 2020.

HINDT, M. N.; GUERINOT, M. L. Getting a sense for signals: Regulation of the plant iron deficiency response. **Biochimica et Biophysica Acta (BBA) - Molecular Cell Research**, [*s. l.*], v. 1823, n. 9, Cell Biology of Metals, p. 1521–1530, 2012.

KÁMÁN-TÓTH, E. *et al.* A simplified and efficient Agrobacterium tumefaciens electroporation method. **3 Biotech**, [*s. l.*], v. 8, n. 3, p. 148, 2018.

KAPUSTIN, Y. *et al.* Splign: algorithms for computing spliced alignments with identification of paralogs. **Biology Direct**, [*s. l.*], v. 3, n. 1, p. 20, 2008.

KAR, S.; PANDA, S. K. Iron Homeostasis in Rice: Deficit and Excess. **Proceedings of the National Academy of Sciences, India Section B: Biological Sciences**, [*s. l.*], v. 90, n. 2, p. 227–235, 2020.

KATOH, K.; STANDLEY, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. **Molecular Biology and Evolution**, [*s. l.*], v. 30, n. 4, p. 772–780, 2013.

KELLOGG, E. A. The Evolutionary History of Ehrhartoideae, Oryzeae, and Oryza. **Rice**, [*s. l.*], v. 2, n. 1, p. 1–14, 2009.

LI, G. *et al.* The Sequences of 1504 Mutants in the Model Rice Variety Kitaake Facilitate Rapid Functional Genomic Studies. **The Plant Cell**, [*s. l.*], v. 29, n. 6, p. 1218–1231, 2017.

LICHTBLAU, D. M. *et al.* **The small iron-deficiency-induced protein OLIVIA and its relation to the bHLH transcription factor POPEYE**. [*S. l.*]: bioRxiv, 2023. Disponível em: https://www.biorxiv.org/content/10.1101/2023.12.05.570225v1. Acesso em: 7 fev. 2024.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. **Methods**, [*s. l.*], v. 25, n. 4, p. 402–408, 2001.

MANN, D. G. J. *et al.* Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (Panicum virgatum L.) and other monocot species. **Plant Biotechnology Journal**, [*s. l.*], v. 10, n. 2, p. 226–236, 2012.

MENGUER, P. K.; SPEROTTO, R. A.; RICACHENEVSKY, F. K. A walk on the wild side: Oryza species as source for rice abiotic stress tolerance. **Genetics and Molecular Biology**, [*s. l.*], v. 40, n. 1 suppl 1, p. 238–252, 2017.

NGUYEN, L.-T. *et al.* IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. **Molecular Biology and Evolution**, [*s. l.*], v. 32, n. 1, p. 268–274, 2015.

NISHIKAWA, T.; VAUGHAN, D. A.; KADOWAKI, K. Phylogenetic analysis of Oryza species, based on simple sequence repeats and their flanking nucleotide sequences from the mitochondrial and chloroplast genomes. **Theoretical and Applied Genetics**, [*s. l.*], v. 110, n. 4, p. 696–705, 2005.

NOGUEIRA-DE-ALMEIDA, C. A. *et al.* Prevalence of childhood anaemia in Brazil: still a serious health problem: a systematic review and meta-analysis. **Public Health Nutrition**, [*s. l.*], v. 24, n. 18, p. 6450–6465, 2021.

OZAWA, K. Establishment of a high efficiency Agrobacterium-mediated transformation system of rice (Oryza sativa L.). **Plant Science**, [*s. l.*], v. 176, n. 4, p. 522–527, 2009.

PHILLIPS, A. L. *et al.* The first long-read nuclear genome assembly of Oryza australiensis,

a wild rice from northern Australia. **Scientific Reports**, [*s. l.*], v. 12, n. 1, p. 10823, 2022.
PRADHAN, S. K. *et al.* Genetic regulation of homeostasis, uptake, bio-fortification and efficiency enhancement of iron in rice. **Environmental and Experimental Botany**, [*s. l.*], v. 177, p. 104066, 2020.
RAMBAUT, A. **FigTree**. Versão v1.3.1. Edinburgh: Institute of Evolutionary Biology, University of Edinburgh, 2010.
RASHEED, A. *et al.* IRON TOXICITY, TOLERANCE AND QUANTITATIVE TRAIT LOCI MAPPING IN RICE; A REVIEW. **Applied Ecology and Environmental Research**, [*s. l.*], v. 18, n. 6, p. 7483–7498, 2020.
RICACHENEVSKY, F. K. *et al.* ZINC-INDUCED FACILITATOR-LIKE family in plants: lineage-specific expansion in monocotyledons and conserved genomic and expression features among rice (Oryza sativa) paralogs. **BMC Plant Biology**, [*s. l.*], v. 11, n. 1, p. 20, 2011.
RICE ALMANAC: SOURCE BOOK FOR THE MOST IMPORTANT ECONOMIC ACTIVITIES ON EARTH. Fourth Editioned. Los Baños, Philippines: IRRI, 2013.
SANTOS, A. B. dos *et al.* **A cultura do arroz no Brasil.** [*S. l.*]: Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2006., 2006. Disponível em: http://www.infoteca.cnptia.embrapa.br/handle/doc/193176. Acesso em: 29 fev. 2024.
SASAKI, T. The map-based sequence of the rice genome. **Nature**, [*s. l.*], v. 436, n. 7052, p. 793–800, 2005.
SCHLÖTTERER, C. Genes from scratch – the evolutionary fate of de novo genes. **Trends in Genetics**, [*s. l.*], v. 31, n. 4, p. 215–219, 2015.
SENGE-PR. Carência de ferro no organismo já faz 40% da população mundial ser anêmica e índice pode aumentar com pandemia. *In*: SENGE-PR. 5 ago. 2021. Disponível em: https://www.senge-pr.org.br/noticia/carencia-de-ferro-no-organismo-ja-faz-40-da-populacao-mundial-ser-anemica-e-indice-pode-aumentar-com-pandemia/. Acesso em: 26 jan. 2024.
SINGH, U.; SYRKIN WURTELE, E. How new genes are born. **eLife**, [*s. l.*], v. 9, p. e55136, 2020.
STEIN, J. C. *et al.* Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. **Nature Genetics**, [*s. l.*], v. 50, n. 2, p. 285–296, 2018.
STEIN, R. J. *et al.* Genotype Variation in Rice (Oryza sativa L.) Tolerance to Fe Toxicity Might Be Linked to Root Cell Wall Lignification. **Frontiers in plant science**, [*s. l.*], v. 10, p. 746, 2019.
TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. **Molecular Biology and Evolution**, [*s. l.*], v. 38, n. 7, p. 3022–3027, 2021.
TAUTZ, D.; DOMAZET-LOŠO, T. The evolutionary origin of orphan genes. **Nature Reviews Genetics**, [*s. l.*], v. 12, n. 10, p. 692–702, 2011.
VAKIRLIS, N.; CARVUNIS, A.-R.; MCLYSAGHT, A. Synteny-based analyses indicate that sequence divergence is not the main source of orphan genes. **eLife**, [*s. l.*], v. 9, p. e53500, 2020.
VELTRI, D.; WIGHT, M. M.; CROUCH, J. A. SimpleSynteny: a web-based tool for visualization of microsynteny across multiple species. **Nucleic Acids Research**, [*s. l.*], v. 44, n. W1, p. W41-45, 2016.
WAIRICH, A. *et al.* Chromosomal introgressions from Oryza meridionalis into domesticated rice Oryza sativa result in iron tolerance. **Journal of Experimental Botany**,

[*s. l.*], v. 72, n. 6, p. 2242–2259, 2021.
WAIRICH, A. *et al.* The Combined Strategy for iron uptake is not exclusive to domesticated rice (Oryza sativa). **Scientific Reports**, [*s. l.*], v. 9, n. 1, p. 16144, 2019.
WASSMANN, R. *et al.* Rice and global climate change. *In*: [*S. l.: s. n.*], 2010. p. 22.
WHO. **Fact sheets - Malnutrition**. [*S. l.*], 2023. Disponível em: https://www.who.int/news-room/fact-sheets/detail/malnutrition. Acesso em: 26 jan. 2024.
YU, J. *et al.* A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. indica). **Science**, [*s. l.*], v. 296, n. 5565, p. 79–92, 2002.
ZHANG, Y. *et al.* Genome and Comparative Transcriptomics of African Wild Rice Oryza longistaminata Provide Insights into Molecular Mechanism of Rhizomatousness and Self-Incompatibility. **Molecular Plant**, [*s. l.*], v. 8, n. 11, p. 1683–1686, 2015.
ZHANG, L. *et al.* Rapid evolution of protein diversity by de novo origination in Oryza. **Nature Ecology & Evolution**, [*s. l.*], v. 3, n. 4, p. 679–690, 2019.
ZHANG, X. *et al.* The Adaptive Mechanism of Plants to Iron Deficiency via Iron Uptake, Transport, and Homeostasis. **International Journal of Molecular Sciences**, [*s. l.*], v. 20, n. 10, p. 2424, 2019.

8. Supplementary data

Supplementary Table 1 - BLASTp analysis results with LOC_Os11g15624.1 protein sequence as query. Data in red is referent to the gene of interest and was retrieved from the Phytozome database. Data in black was retrieved from the Phytozome database. Data in green was retrieved from the Ensembl database.

Gene	Transcript	Location	Genome	e-value	Identity (%)	CDS length (nuc)	Protein length (aa)	Transcripts	Family	Class
LOC_Os11g15624	LOC_Os11g15624.1	Chr11:8864460..8867281 reverse	Oryza sativa v7.0	2.93e-92	100	390	130	3	Poaceae	Monocot
LOC_Os10g11889	LOC_Os10g11889.1	Chr10:6597052..6599560 reverse	Oryza sativa v7.0	1.10e-27	83	180	60	2	Poaceae	Monocot
OGLUM11G08570	OGLUM11G08570.1	Chromosome 11: 8,108,277-8,109,301 reverse strand	Oryza_glumaepatula_v1.5	3.1e-50	98.7	246	81	1	Poaceae	Monocot
OGLUM01G19470	OGLUM01G19470.1	Chromosome 1: 19,024,583-19,026,571 reverse strand	Oryza_glumaepatula_v1.5	8.4e-18	84.1	597	198	2	Poaceae	Monocot

OGLUM03G23070	OGLUM03G23070.1	Chromosome 3: 19,192,248-19,193,678 forward strand	Oryza_glumaepatula_v1.5	9.7e-12	87.1	348	115	1	Poaceae	Monocot
OGLUM01G19470	OGLUM01G19470.2	Chromosome 1: 19,024,583-19,026,011 reverse strand	Oryza_glumaepatula_v1.5	3.6e-11	83.9	348	115	2	Poaceae	Monocot
ONIVA09G02080	ONIVA09G02080.1	Chromosome 9: 2,917,118-2,923,713 reverse strand	Oryza nivara v1.0	2.3e-27	92	672	137	2	Poaceae	Monocot
ONIVA01G14700	ONIVA01G14700.1	Chromosome 1: 11,806,959-11,807,674 forward strand	Oryza nivara v1.0	5.7e-25	81.5	324	107	1	Poaceae	Monocot
ONIVA09G02080	ONIVA09G02080.2	Chromosome 9: 2,917,120-2,923,8	Oryza nivara v1.0	1e-21	95	465	40	2	Poaceae	Monocot

		30 reverse strand								
ONIVA10G14310	ONIVA10G14310.1	Chromosome 10: 15,011,349-15,012,771 forward strand	Oryza nivara v1.0	1.4e-11	90	366	121	1	Poaceae	Monocot
OPUNC04G05940	OPUNC04G05940.1	Chromosome 4: 12,240,439-12,241,004 reverse strand	Oryza_punctata_v1.2	1.3e-12	83.9	321	106	1	Poaceae	Monocot
OBART03G22350	OBART03G22350.1	Chromosome 3: 16,872,762-16,874,750 forward strand	Oryza barthii_v1	7.8e-20	90.9	594	197	1	Poaceae	Monocot
OBART11G08880	OBART11G08880.1	Chromosome 11: 7,419,411-7,421,436	Oryza barthii_v1	1.3e-14	100	624	207	1	Poaceae	Monocot

		reverse strand								
ORUFI12G07700	ORUFI12G07700.1	Chromosome 12: 6,216,051-6,218,442 reverse strand	Oryza rufipogon W1943	1.2e-28	94.1	309	102	1	Poaceae	Monocot
ORUFI01G40710	ORUFI01G40710.1	Chromosome 1: 33,646,860-33,647,575 reverse strand	Oryza rufipogon W1943	4.1e-26	86.3	261	86	1	Poaceae	Monocot
ORUFI01G25050	ORUFI01G25050.1	Chromosome 1: 21,800,608-21,801,125 forward strand	Oryza rufipogon W1943	1.5e-24	75.4	225	74	1	Poaceae	Monocot
ORUFI11G09410	ORUFI11G09410.1	Chromosome 11: 7,982,774-7,987,367 reverse strand	Oryza rufipogon W1943	8.3e-17	89.7	294	97	1	Poaceae	Monocot

ORUFI11G09420	ORUFI11G09420.1	Chromosome 11: 7,987,461-7,988,879 reverse strand	Oryza rufipogon W1943	1.5e-14	100	483	160	1	Poaceae	Monocot
ORUFI01G18530	ORUFI01G18530.1	Chromosome 1: 15,384,399-15,386,440 reverse strand	Oryza rufipogon W1943	2.5e-10	80.6	372	123	1	Poaceae	Monocot
OMERI08G08860	OMERI08G08860.1	Chromosome 8: 10,442,888-10,444,873 forward strand	Oryza_meridionalis_v1.3	2.4e-10	80	351	116	1	Poaceae	Monocot
ORGLA11G0078900	ORGLA11G0078900.1	Chromosome 11: 6,976,393-6,978,759 reverse strand	Oryza_glaberrima_V1	8e-87	96.1	390	129	1	Poaceae	Monocot
ORGLA03G0208200	ORGLA03G0208200.1	Chromosome 3: 16,784,	Oryza_glaberrima_V1	2.4e-46	88.1	255	84	1	Poaceae	Monocot

		989-16,786,444 forward strand								
ORGLA12G0064600	ORGLA12G0064600.1	Chromosome 12: 5,644,797-5,646,150 reverse strand	Oryza_glaberrima_V1	1.8e-36	86.8	207	68	1	Poaceae	Monocot
ORGLA02G0192900	ORGLA02G0192900.1	Chromosome 2: 18,694,410-18,695,821 reverse strand	Oryza_glaberrima_V1	1.2e-35	86.8	207	68	1	Poaceae	Monocot
ORGLA10G0039700	ORGLA10G0039700.1	Chromosome 10: 5,488,347-5,489,713 reverse strand	Oryza_glaberrima_V1	2.6e-27	81.0	180	59	1	Poaceae	Monocot
LOC_Os06g19095	LOC_Os06g19095.1	Chromosome 6: 10,865,395-10,867,828	Oryza sativa Japonica	3.9e-36	86.8	1236	68	1	Poaceae	Monocot

		reverse strand								
LOC_Os03g31730	LOC_Os03g31730.1	Chromosome 3: 18,119,653-18,121,181 forward strand	Oryza sativa Japonica	4.3e-29	87.5	1529	59	1	Poaceae	Monocot
Os01g0495701	Os01t0495701-00	Chromosome 1: 16,980,463-16,981,993 reverse strand	Oryza sativa Japonica	6e-29	85.7	1531	59	1	Poaceae	Monocot

Supplementary Table 2 - BLASTp analysis results with LOC_Os04g45510.1 protein sequence as query. Data in red is referent to the gene of interest and was retrieved from the Phytozome database. Data in black was retrieved from the Phytozome database. Data in green was retrieved from the Ensembl database.

Gene	**Transcript**	**Location**	**Genome**	**e-value**	**Identity (%)**	**CDS length (nuc)**	**Protein length (aa)**	**Transcripts**	**Family**	**Class**
LOC_Os04g45510	LOC_Os04g45510.1	Chr4:26925797..26927110 reverse	Oryza sativa v7.0	1.19e-63	100	342	114	1	Poaceae	Monocot
Oropetium_20150105_12226	Oropetium_20150105_12226A	Oropetium_genomic_2014111	Oropetium thomaeum v1.0	1.51e-51	77	348	116	1	Poaceae	Monocot

		2_029:1732993..1733341 reverse								
Misin12G158000	Misin12G158000.2.p	Chr12:62806423..62807519 forward	Miscanthus sinensis v7.1	4.56e-58	88	345	115	2	Poaceae	Monocot
Urofu.7G214600	Urofu.7G214600.1.p	Chr07:23296729..23298361 reverse	Urochloa fusca v1.1	6.83e-58	85	351	117	1	Poaceae	Monocot
Chala.06G181600	Chala.06G181600.1.p	Chr06:56333508..56333850 reverse	Chasmanthium laxum v1.1	1.24e-57	86	342	114	1	Poaceae	Monocot
Misin11G157500	Misin11G157500.1.p	Chr11:60835993..60837176 reverse	Miscanthus sinensis v7.1	3.67e-57	88	345	115	1	Poaceae	Monocot
Pavir.7KG267400	Pavir.7KG267400.1.p	Chr07K:40964373..40966326 reverse	Panicum virgatum v5.1	4.79e-55	85	339	113	1	Poaceae	Monocot
Sevir.7G192500	Sevir.7G192500.1.p	Chr_07:25085744..25087219 reverse	Setaria viridis v2.1	3.00e-54	82	354	118	1	Poaceae	Monocot

Seita.7G182700	Seita.7G182700.1.p	scaffold_7:26024566..26025939 reverse	Setaria italica v2.2	3.00e-54	82	354	118	1	Poaceae	Monocot
ELECO.r07.4AG0309100	ELECO.r07.4AG0309100.1	4A:7445041..7445380 forward	Eleusine coracana v1.1	7.51e-54	79	339	113	1	Poaceae	Monocot
Pavag06G172700	Pavag06G172700.1.p	Chr06:33464488..33466495 reverse	Paspalum vaginatum v3.1	2.78e-47	85	537	179	1	Poaceae	Monocot
Pahal.7G235400	Pahal.7G235400.1.p	Chr07:40588160..40589466 reverse	Panicum hallii v3.2	2.04e-43	83	354	118	1	Poaceae	Monocot
Joasc.05G086200	Joasc.05G086200.1.p	Chr05:63208757..63209908 reverse	Joinvillea ascendens v1.1	6.60e-47	71	357	119	1	Joinvilleaceae	Monocot
Joasc.05G086300	Joasc.05G086300.1.p	Chr05:63225480..63225813 reverse	Joinvillea ascendens v1.2	1.47e-36	60	333	111	1	Joinvilleaceae	Monocot
Aco016268	Aco016268.1	LG03:11319372..11319741 reverse	Ananas comosus v3	6.74e-31	50	369	123	1	Bromeliaceae	Monocot

Joasc.05G086100	Joasc.05G086100.1.p	Chr05:63206609..63207741 reverse	Joinvillea ascendens v1.1	5.26e-21	46	423	141	2	Joinvilleaceae	Monocot
Acora.08G183000	Acora.08G183000.1.p	Chr08:24181596..24183134 forward	Acorus americanus v1.1	3.71e-8	39	357	119	1	Acoraceae	Monocot
OB04G27120	OB04G27120.1	Chromosome 4: 14,592,627-14,593,232 reverse strand	Oryza_brachyantha.v1.4b	1.7e-41	100	606	201	1	Poaceae	Monocot
OBART04G20110	OBART04G20110.1	Chromosome 4: 19,836,415-19,836,756 reverse strand	Oryza barthii v1	2.5e-41	100	342	113	1	Poaceae	Monocot
ORGLA04G0172000	ORGLA04G0172000.1	Chromosome 4: 18,986,920-18,987,261 reverse strand	Oryza_glaberrima_V1	1.8e-41	100	342	113	1	Poaceae	Monocot
AMDW01015182.1_FG001	AMDW01015182.1_FGT001	Scaffold AMDW010151	Oryza longistaminata	1.7e-41	100	216	71	1	Poaceae	Monocot

		82.1: 25-240 reverse strand	v1.0							
OLONGI [KN540855 .1_FG005]	KN540855 .1_FGT00 5	Scaffold KN540 855.1: 8,718-9,059 reverse strand	Oryza longista minata v1.1	1.7e-41	100	342	113	1	Poaceae	Monocot
OGLUM04 G20000	OGLUM0 4G20000.1	Chromosome 4: 24,159, 781-24,160, 122 reverse strand	Oryza_ glumae patula_ v1.5	2.9e-41	100	342	113	1	Poaceae	Monocot
ONIVA04 G18310	ONIVA04 G18310.1	Chromosome 4: 19,554, 111-19,554, 452 reverse strand	Oryza_ nivara_ v1.0	3e-41	100	342	113	1	Poaceae	Monocot
OPUNC04 G17740	OPUNC04 G17740.1	Chromosome 4: 24,890, 582-24,890, 923 reverse strand	Oryza_ punctata _v1.2	2.7e-41	100	342	113	1	Poaceae	Monocot

ORUFI04G21670	ORUFI04G21670.1	Chromosome 4: 22,633,302-22,633,643 reverse strand	Oryza rufipogon W1943	2.9e-41	100	342	113	1	Poaceae	Monocot
OMERI03G30720	OMERI03G30720.1	Chromosome 3: 32,140,077-32,140,418 forward strand	Oryza_meridionalis_v1.3	2.8e-41	100	342	113	1	Poaceae	Monocot
Zm00001eb075410	Zm00001eb075410_T001	Chromosome 2: 23,811,785-23,813,151 forward strand	Zea mays	4.4e-37	95.5	733	113	1	Poaceae	Monocot
SORBI_3006G161600	EES12567	Chromosome 6: 51,938,989-51,940,700 reverse strand	Sorghum bicolor	2.2e-37	92.8	1712	114	1	Poaceae	Monocot
LPERR04G16260	LPERR04G16260.1	Chromosome 4: 16,060,939-16,061,	Leersia perrieri v1.4	6.8e-41	98.6	333	110	1	Poaceae	Monocot

		271 reverse strand								

Supplementary Table 3 - Primer sequences used to amplify LOC_Os04g45510 and LOC_Os11g15624 sequences from *E. coli* strain DH5α transformed with pANIC6A vector.

Primer	Sequence
pUBI1 FWD	GCGGTCGTTCATTCGTTCTAG
LOC_Os04g45510 pUBI1 RVS	GCGTCGTAGAATCCGAAC
LOC_Os11g15624 pUBI1 RVS	TGGATTTGGCCAGTGACC

Supplementary Table 4 - Nucleotide sequences used as query for the sinteny and collinearity analyses.

Gene ID	Sequence
LOC_Os11g15624	TTTTCTCTGTTCCTTTTTCTCCTCCTTCTTCTTCCTATGGAAACACCGAGCCC TCCAGCTCATCCCTGCCCTCAAACCTCTCCAAACTGCCGCCAAGGTCACTG GCCAAATCCACCGTCGCCGGCCCTTTTTCTGCCACAACAGCAGGCCATCA ACTTCGCCGCTCATCAGTGAGGCTTCGTTCTCTCTTTTCCCTCTTTTTCAGC TTTCGTATCGAGCTTTTCCGAAGCTCCCATGGGCTGTCATGGCCATCACCA TGGCCAACACATCCAGAATTTCCTCTTTATCAACTAACCACAAAAATGGAA CCTTCTCATCCGGGAACGCGTTTTGACGCGCTCACCCGCCCGCGTTCGCCA CCGGAGCTGCCACTTCCATGCCCGGTGTCACCATAGCTCCAGATCGCGTTC CGGCTAGCTTTGTTGTCTTCCAGCCCAGCGTACACGTCCTTTGGGGGTGTC AGCCGGTGGAACACGTCATGCTATCGTCTCCGTCGCCGTTCCCTCACTTAA GTTCTGGAGCCTTTCTCTCCACTTCTGTTAATCACGTGAAAAGTCACGGGA AGAAGATAAACAGTGCCGGCCAGCATTGTTTCCTATTGGTTGAACCCAAC CAACCCCATCAAAACCACCTAAATAGTGCCCAACCACAATAAAACCAGTC CAAACAATTCTATTAGGTTGTTTTCCTCCTGTTCAAGCTCTAAAACACCAA ATCATTTTCAGTTTTCACCTTTGTTATCTATAGCGTTACTCCCAAACTGCAA CCACTATTACACATGCCATAAATCATAACCCTTGGATAAATTGTTTGCCAT TCCAAATCAAGTTAACTCTCAATTTATCATTTTCTTTTGTTCGGGCTCTGAT TCACAATATTTATTTTATCGTTTATTTGTAGTATTTTTATTATAGGCTTATTT GGTCATTTCTTTGTGTACGTACGTTAGATCACGTTGAAGATCCCACCGGAC TCGCAGGAACTGAAGTGGACCAAGAAGGCTATGTTGAAGCTTCCGAAGGC CACGTTCCACGAGGCAAGCCACAGTTCCAATCTTTGATCATATTGAAACCT AGTAGAGAACCACTTAGTTATTTTATGCATTGAGTATTTATAAAATTATTT TGCGTATGAATTATTTTTGAATAAACTTGTTAGTTTAGTTATTACCCATTAT TACATTGTTATTATTGGCCAAATCCATGATTGCTATATCACGCCCTTTGATA ACCTAGGTGTAGATTTATATACTTCACAATAATATTCTAGTGGATAAGACC ACCAAATATCTATACTCATATACATGGTTTACCCGTCGTCACGTGGTGGCG TGTCAGGTGATATAAAATATGTTATTTGATATGAGTCACTCGCCTGGAGTG ATATGGAAGATTAATTCCGGGTGAAGATGGATGTTGTGGTGTGTGAGGCC GTGCATATTGTTATAATTATTTTGTGGAAAGTGTATATTTGCACTTGAAGT ATGCACGGGCCTGTGCTAGTTTTGGTGGCTTTCCCCAGATATAAGGACTAC CCCCTTATGCCATGAAAAATGATGTAATTTGCGCCCATCCACAGGACTGAG TAATCAGACTTAGCCTAGTACATGTCCCACTATATGAAAGACTGTAGGGTG GCTCTATGAGCAGGTTTGGGGTGGACATGTGGAAGGGGCTTTGGCCAGGT TGACATGATCCACAACGGTGGAACCCGTGAAAGCAGCGACCGTCTAGCCT

	TGCTTAGGCCCAGTAAGTGCCCGACTGCTAGGACGGATCAATATCCCACT GGGCAGTAACTAGTAGTAAGTGGGACTCTCGAACGGCTATGTTGTCCTCA GCCTAGCTAGCTGTTGACATCAGGTGGGTAAAGCTGGGCAGTCTCTGTAG AGATGAGTCTTAAATCATTCAAATGTGCCGTGCTTCCGGTTAAGAAGCTAT GTCATTGGGTGATTCCTATTTATCATGGTATGAGTAGATTTTAAATTAAAA TATGGAATATAATGAAGATGAGGATGAGTGGTGATTAATATTGCTACCGT TTTAAAATGATGATTTACAATTGTTTGCGAATAAAATGTTTGATAAAGTAG GGAAAAACTATTTATAATATGCTTGTTATGAAACCTCCTTGATGAACTATA TTCTTTGTACGCCTTACATATATTCTATTTGTATTATTATGGATGTGGCTTC TGGGTACCTCGGTACTCACCTCTTGCATAAAATGTGGTTCAGGCGATTAGT TTTTCCCTGAGGAGGCACAGGAAGGGGAAGAATTCTATGGTGCTGACGAG TTCTTGTTATAAGTCGATCTGCTCTCAGGCGTTGCCTGTGGTGTTTTTATCC TATTTTGTGTTCGGGTCTTTTGTTTGTAGGGAGTGTGTTTGGCCTGTCGGCC CCTAAAACCTGTCTTAGTTATCTTTTATGTAATAATTTCTGTAAAAAGTTGG GTATATGTATTGAAAGCTGTTCTGTATATCAATCTGTGCGCGATCGAAATC CTGGACGGGTGGTTTCGGGATACCATACTCAAACCGGGTATCCCCACAAT CTTTCACTCAAAATTTGGATGGAGGGGCGGCAGACAAAGAATAGGATGGG AAACAGAGGCAGGCTAAACTGATGGCGAAGAGTTTGAAGACGGCGATGG AAACCTGTGTTTTGATGCTTGAAATGCCAATATTTTAGTCCTGTAATAACT GTTTGTTTTCCTTTTTTTTCTCGTCTTATCTGTAAGAACTAGATGTTCTGGA ACTCTTGTAATCATCTGCGTTTTATTTAATGAAAACGGGGGAGATACCCTC CCTTTCTAAAAAAAAATA
LOC_Os11g15580	CACTGTCGTTTTGATCAGATATGATATGGTCATATATGCCCATATCTCCAG GCTTTTCTTAGGGCGTTAGCAGTGCGCCCTTCCTTCCAAGGAGCCCAAGGG CATACACGTAGGAAAAGG CACGGAGCTTAACAAATGGGGCTTCTAGCTTTACAATGCGCCGTGCTTAAG GTGGGGCCC ACCAGCAAACTTTTTCCCCGGTCGGCCCCACCACGACGCCGCTCATCCCTT TTCGGCTTT TCCGCTCCATCGCTTTCTCCGCCGCTCTGTCGTGACTTTCGCTCCTAACCCC TTCTCCGC CGCTTCTTCCTCCCGCTCACTCGGCGTATCGGCGACCCCCCTCCCCCTCCCC CGCCCCCT CCTCCTCTCTGTCCTCTCTACACCGGCGCACGGCGGCGACCGCCCTACCAC TCCACTTTC TCTGTCCGACGGCATGGCGGCACCGATCTGGCGGCGCTTGGGAGAAGAAG CGACAGCCGC CTCTCCCCTCCGGCGCCGGGCAAACGAGCTCGGCGCCGTGGGTTTGAGCA AGGGCGGCGA CGAGGCTCCCTCTCCGGCGCGGATCTGCCCTAGGGCGAAGCGGCGGCGGC GGCCTCGGTG ATGGAGTTGAGCAGAACGGTGGTTGTTGGCGGGGAGGTCGCCGTCAAGCT GCTCCTGCTG CTTGGAACAGAGGACGACGGCGGGGTTGACGGTTCGGAACGGGGACGAC AACAAGGTTAG GAGCCAGATCTGCCACCGCCTCTTCCACCGCTGCCACCATCTTCCCTTCTC CCCAGGCCG CCGACGTCACTAGATCCGGCGCTGCTTTGGGGGCGACGAAGCACGGGGCC ATGCCCTCTC CTAATGCGCCATGCGCCCCGTTCCTCTCCTCGTTCTGACGCGTGCATGCCA AAACTGACG AAAAACGCGTGCCCAAGTTTGGCTCCGCGGGCTTCAGGTTGGTTGTGCCA GACGTATGCG TGGCGTGGAATCTCCAACGTGGCGCGGCGGGGAAGAACGCCACGGCGACT CTCGGCGACG CACTGTGAATGCTCTTAGTGTTATCACCATTGCTGGAACAATCAACATAAA AAGCTGCAC ATGTCGGACACAGGAAGTCTGCAACAATCAACTGACAAACACACGGTCAC

LOC_Os11g15590	TGCAAACTGC CCAACGAATGGCATCTTGGTATTAGCCTGTGATAATCAACGAGAGGCATC TGTCTAGAGC CTCAAAGAGCATTCCTACATCACCACAATCTGGGCCTTGTTTGATTTATAT TATTAAGCG AATAATTGCTGCTTATTATCTAGCTGATTATAGGGATAAATTATATAGTTT GAGAAATAA AGGTAACAAATAGCTTAAGAAAACATATTTTTCGAAACGAGCAAACAATC CATTTCAATA ATCCAGTGAATAATCTAAAAGCATAAGAAGGAAAATGGAAGCCTCGAGCT CTTGCTAACA ATAAAGGCAACCTAAAATCCATGGATGGCTAGTTTCAAAATTGTAAAATA AATCAATTAA TGGGGAGTGAATTGAAGATCGTCCGGATGTAGGTTTACGAAAGAACCAAA CTAGTGCAAC AGCATATGCTCGACCGGGGCTTGTAGTTTTGACTAACTATAGGATCAAAAC AAGCTATTG AAACCTACTTTTTTTTTTTTGGGGGGGGGGGGGGCTGGAAAAATGGATCAC ACACTTTGT ATCTCAACCCTTCATCTATTTATTATGCCCATGTATAGTGCACACAAGCCA CAATTTCTA CAATCCTACATCCCTATAAATACAATCCTTCTAGGAATACATTTTATTCCTT ACAAA TCTCGATGAGTTCTACAATTTTTCTGTTTATGATATTTGCATTTGAAACATA AAAATCAG CCTAAAATTAAATTTAAAGTTCTCATAAGTTTAAGGTGAAATGTCCATTTT GTCCATAAA ATTTCGTAATTACCCTCAACACACCAGGGACAGTTTGGTAAATTCGCATAG GATTGCTTC TTATCCATATACTCTATTCTCAAAACCCTAGCCAAACTCCATCCATCCACT CAGCCGCCG CCACCATCTACTCACTCTCCTCCATCCACTGCGCCGCCTGCATCTCCTCCCC CAGCCACC TCTTCGCCCTTCCCCGCCTCACACCCCCACCCAGGGAAGGATGGTGGCGCG CGGATCCGC CGCCTCCCGTCCTCGGGAGCGACGGCAGCGCATGAATCCAGCGGCCTCCG TCCTCGGGAG CGGCCTGGCGTGAGTGGGAGCACCGGCAGCACGCGGATCCGCCTGCCTTC GCCCTCCGGA GCAACGGCGGAGGGCTTAGGAAGGCGGGAGCGAGCGGCTACGGTGGCGG CGGCTGCGGCA ACGACGGCGATGGTGGCTGCAACGACGACGACGACGGCGACAGCTGCGA CGGTAGCTGCG ATGACGACGATGATGATGACCAGCGGCGGCCGGCGACTTGCAATCGCGAG GCGGGCAGGT GGTGTGGGGCCGCGGGGGCGCGCCGGCGCGGGGCAGGGGTGCATGGCAA CGGGCCAGCGA CGCGTCGGCGCCTCAGCACAGGATCTGTACGCCGTCGCGGCGGCCAGCGA GTCGGCAACC AGGAGCACTGGACCAGCAATGGTTGCGCCGGAATGCCAACCAAAGGTCTA GCTCCAAGCG TCGTTCACTCATCTGTCCATGTGCACGGCTGACGGCTGGTATATTCAGTGA GTATTTCTA GGATCTCTCTCTGGTTTGGTTTTTCTCTTAGCAGAGTGAGATTAATTGAAA AGTTGAAAC CTAGAATGAGTAAACTCAGTTCAGTTCAAGGCATATTTTTGTGATTTGTAT GCAGAGAAT TGAGATCATGTGTCTTTACTTCCTGGTTCAATTAGCAATGAGTAAAAATGA

AGCGCTGCA
ATTAGGAATGATTGATTATATTATTGTGAACAAGTTTTTATTTGACTCAA
CTATACTAT
GACATGTGGGCACACCCATGACAAAAATCTGACCCTTGTTTCTGCTAAAT
TCGAATGCG
CTGTTAAGTTTAAGTACACTGGGTTTGTTTGCTTAGCTTGTTCTTCTGATCG
TATAGTCA
CTAGAACAGGTGCTTTACTGGTACAGGTGGACCTAATTACTTTTATCATAA
TTCATGAGT
GAATTGGAGGCATCGCCACAACTCAAATTGACTTGAGGATCATAACTCAA
ATTGACTTGA
GGTTCATAGCTTGGACAAGTGAGCAATCCATCCAAGGAAATTAACTAACT
GAACCAAAGG
GTGGAATTTGATGGCACAGCTTTTGGGCTTTCTGTTTTGCAATGGGCTTCCT
GATAAGTT
CTTTCAAGAAAAATGCATGGTTAACTAATTATAGTTATGAAGATGTTTATG
CCTGCCCAA
CCGATCAGAGCTGTTGTTTACTGCTTTGATCTCTACTATGTACCAAACTTAC
CACATGCA
TTCTTCTTGAAAACAAACACCTTTTTTTTTGTTAATGTTAGGTGTGGTAGCA
TTAATCTG
AACTCAGTTTAATCCAAGTATATAGATCATGATATCTCCAAGGAAGAAATT
TTTTCTAGC
TTTTTTTTTGTTAATATTAGGTGTGGTAGCATTAATCTGAACTCAGTTTAAT
TAACCCAA
GTATATAGATCATGATATCTCCAAGGAAGAATTTTTTTCTAGCTTGGATCC
TTTAGTGCA
GTCAGTAGAAGAAGTAGGAAGGGAAGCTACTACTGCATCTCCACCTCAGT
CTGTAATTCA
TAAAGCTGAAAATGTCACTGGTAAAAGGGAAGCATCACCTCCTGAGAAGA
AAGGTCATGG
CTAGCATCTCTCTGACTCCTTCTGTAATCCAATATAATGTTTGTCACAGAA
ATATCCCTT
TAAGAATGTAAATAAGATGTGAGATTTTTACTAAATTGGCTATAAACTTTC
TCAAAGAGT
ACATGTTCCTTAAAATCATGTTCCTTAAAACCCTACTCAACTAACAGATTT
GCCTAGCAA
AACCGGTTACTAGTCAGAAGCTTCTGCAACTCTAATGCAATGTTGTGATCC
TTTCTTCTG
ATCCTTTCTTCATCTGTGTTGCAGAATGTGACTACAGGAATGAGGGCCTCG
GGACTGCAA
GATCAGGTGGCACTGGTGAGCATCGGCGGGTGGAGCCCTAACCAGGGCGT
GCCGCAGCCG
GCCATCATCCTCAACCCTTCCTTCTCGCGCCACAGACAGGGCCTTTGCCCT
GGGTGGCCA
TATTGGGCGGCGGCAATGGGATTGAGCAACCGCACTTTAAGTACAGAACT
GGAGCCAATG
GAATTGAGCACTACTGAACAAGGCAAGTAACTTGGATTCTTGTTCTTTTGT
CTTTATATG
TGCTATGCTTCATTCATATCAAGTGCAGGATCTTTTTATTGACACCTGTAA
ATGTTTTAT
GTTAGGCACCCTCACTGGCAGTAGACAAGTGGAGCCAAGGGGGGAAGAC
CATGGGCACTG
CAATTGACAGTTAGCCCACTGCAATTGATCTCTTTAAGGAGTTGCATTGCA
GCAAAACTA
AATGTTTTAGTGAACCTGAGAAGAAAGCTATAGTAAGTACTAAACTATTA
CTTCTTCTTT
CCTTCTTTTATTAAGTACCTTATGCAATGAACCTTTATTTATATAATAATGC

AAAACTGA
AACCTTGCTTTGATTGCTAGGAGGATATGCATGCCCGGGAAGTCCTCACTT
CACCATCAG
TAGAAGATGGGCAGCCAGCTAAGACTTCTATTGAAGCTGTTTCCAAAGTG
CTGCGGAAGT
CGAACACATTTTTGCGCAACGTCGGAATTTAGCAGCTCGCTGCCAAGACA
ACCAATGTGA
TGAAAGAGATACAAGTTGAACTAGATGCAAAGAAGCTAGAGTCTGCAGTA
CTTCAGGAGG
AGCTTGAAAGATTGAAGGCACGGGTACTTGTGTAGTTGGATGGTGGATGG
ATCATTTTGT
ATTCATGTTTTGTGCTTTTGTATGGAACATGTTGGAAACTGTATTCTGGAA
CTTGCTTTT
GTATGGGCTGGAAAATTAATTATGGCTGGAATGTGATTATAATATAATTAT
GCAACTGGC
TATGATTTATATTCTGTGATTGCTTTTTAATTGGGCTGGCCCATAATATGGG
CTAGAATT
GTACTGTTGGGCTATAATGAAAAAAAAACGTAAAAGCAAATTTGTAGCTG
GGTCAAAAAA
AAATATTTGGGCCAATTGTAATTGAAAATAAACAAGCCCATGCTATATGG
GCTGAATGTG
AATTGCCACGTCATTTGCCATATCACCCGCCTTGTCATCTGCAACATCACA
TGTCATCGT
ACACGTGGCGTTATTCAGTTCAATTTGTTATGACGAATTTGTCTCGTCATTC
ATCAATGA
CAGGAGAACTATCGTCACTGTATCAATGACGATGTCCTCTGGTGGTCACAG
AAGCTGGTT
AGTGACCCACCTTCTGTGATGACAACCATTTCGTCACTGCGGGCAGTTAGT
GACAGGAAA
CAAGCTGTTATCGTCATAATGTAATCGTCATAGAAGCCCAAAATTCCTGTA
GTGAGGGTT
AGGAGATTGAGTCCATATGTAGTGAGGATTTTAGTTTGATCCAATGGTTAG
AAATTTATT
TGCTTTCTTTTTAAAAAAAAATAAACCGTGCACCTTCTATATCTGTAGTAA
GCTAGTCTA
ATTTTCTTTAACACTAGGCATATATGACACACCATGAAGAGCACCTCCAGT
CTTGCTACT
GTGCTTGGAGACCCAACATTGCTATTGTCTTCACCATTGCTAGAGCATCAC
CACTACCTA
CCTCTGTCATGGCAGCAATCATACCTCCGTCGAGCTCTGCCTCATGCCCTC
ATCATCCCG
GTGGCAACATCCATATAGTCAATCGAGACCTGGCCTTTCGAGGCTTCTCAG
CAGAGCAGA
GACATAGATTCAGTTGGATCTGAACTTGGGGAAACAAATCAGTCTCCCTCC
CTCTCTCTC
TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTCTCTCTCTCTCTCTCTCTC
TGCCCC
CCTCCCCTTTGCACTTAGTGCTTGCATATCGAGCTATTTGGTGATCATATCG
TTGTATGC
TTATCTGTTGTTGATCACCTCCTT
ATTCACATTGTTCAACACTCAACTTCAAATATCGAATCATTTCACCATGTC
AGCGTCACA
TGTATGGTGCTCCATGATGAAAGATCATTTGGATGGTCTAATTCCATCCAA
AAGGGCCTT
LOC_Os11g15600
CAGGTACTGGTGCGAACAGTCCGTGTGTGTTCAACCAAGCACTTGCACCGT
CACAACATC
ATCAGTCTTTCAGTGCCAAGCCCCCGCTTCCCTTCTCCTATTCATTGTCCTT
TGCATAGA

CAATCTTGCCAATCCTTCAACACATATTGCCATCTTCAGGCACTTATAGCC
TTGTATCAC
CATGAAACCTCATTGAACAACATCTTCATCATCATCATCTTCATCATTGGA
CTATTTTCC
TACCTATACACTCAATAAGCACATTGATTCCTTTATCATTTGTCATCAATTA
GCCATAAC
CCACTAGGGTCCTATATGCTCTTTCACATGATAGCATCACATCCAGGCTAG
TAGGCTAGG
CGTCGGGCCATCCCACTTGACCCACTAAGACACATGGAACAAGGCCCACA
TGGGAAGCTA
GTTGGGAAGCAACGAACGCCCTAGACTGCCTAGAACTACACACATGCCTA
GGACCAGACC
ATGGGGAAGCTTCTCATGATCCTGCATGTGGCGATGCATGAGGCAATCGA
CAGCTTTTCA
CAAAAAGACTACCTAGGAGGCGGCTCGTGGCTATGCTTGGGGCTGAGGGC
TGGTCAAAAG
GCAATATCTGGCAGTGTGCAGGGTGACCTAGGTCCGCCACCCATGTGGAG
GTTGTCTGGA
AGGTGGCACATGGGCACGGGGTAACTCGAAGTGGCCCATGGGAAGGTTGC
CCAGGAGGTG
ATGCGTGGCGGCACTCAGGACGAGAGTCGGTCTACTTCGCTGATTTTCTTG
ATGCCCGCA
TATGTGGAAATTCCTTATGCTAGGGAAGATGTGAGGACCACACCCTCCCCC
TTTCCCGGG
AGCCCCATGCCCTGACACTAGGTGTACATGCAAGGAGTGTGCTCACTCAC
CTCGTCATGT
CGCATTAATTAGTGATGCTCACGGTGGATCAAGACATCAAATGAAGGTGA
TTAGGTGAAT
TATGCATCGGTCAAACATTTGTCTTTGACCGCCCCATTCTATTGAAAGGGA
GGCACAGTC
TCTTTGGTGGTAGGCTCGAGGCAGAGCTGCACAACATGGCTAGAAGGAAT
GATCCAGTGT
CGGACGGCATGACGACAGCTACAGTGCAGATAGTGATGTCTGTCTGTGGT
GCCTTGGCAG
CAGCCCCCACAATGGTTAGGTGGACTTAACACCTACGGGGGTATATAAGC
TAGCAAGGCT
TACTCAACAAGAGGTTAGAGATAGTTGAGTACATAAGAACACTCCAAGGC
CTGAGGAGCT
CTAGACAAACCAGCGTGAGCACTACACACAATGTGAAACAACTTCCTCAT
ATCAATCAAG
AAAGTACGAATAGGGTTATTATCCTCTGGGAAGGCCCAAACCTCGGTAAA
TTGTGTGCGC
TGTGGGATTGCTCAATGCCAAAGCTGTGCAGTCATAGCCTTAAATCCCCAT
CACCAAACA
CCAGGTTTTCCAGATCAAGTTCAGCCTACAATCATGCATCGAGACACCCCC
AGTCTAACT
GATTGTGATTATTATTTCACCACCACAGCAGGTGTAGTACTTTTGAAATAC
GATTGGTGA
ATTTTGTCCAGCACAAATTTATTACTGACTTGTTGGGGATTGATGGCTTAG
TTGCATTAT
GACGTGGACACAAATGGCAAGAGGCATAGTGAAAGCATGTCAACGATGT
GTGACTTTCCT
CAGAGATGTTGCCACCGAGATGATGAGGACATGAGGAATACCTCAAGGGA
GGTGTGACTG
GTACCATGAGAGAGGTAGGTAGTGGTGATGTTCTAGAAATGGCACTGTTG
GGCCTCCCAG
AACTATAACAAGGGTGGAGGTGCTCTTTTGTGATGTTCCATGTATGACTTA
GTGTAAAAA

LOC_Os11g15610	AAAAAGTTTGACCAGCTTACAACAAATATAGAAGGTACATGATTTATTTTATAAAAAAAG CAAACAAATTTCTACCCATTACTAGCACAGGGCTTTTGCGGATCAGTTGGATCAGATTTT TTTTCACTAAAAACCCCCAAAATCAGTTTTGGGAAGGAAATATGATAGCTATTGGTGATG CTCTTAGGTGCTCCAGACTTATGTAACACCCTGAAAAGTTCGAATCTAAAACTTAGAAAC AAAATTTTGGAATAAACAAAAATGTTTTTCGAATTTGCATCATCATGGTATTTATTGGAT TTTCCTTTTCTTGTTATCCAATGCCAAAAATTAAAATAAAAAAAGTAAAAGGTTAATTCA AAAGAAATCAAATCCAAATTTAAATAAAGTAAAAGGTTAATTCAAAAGAAATCAAATCCA AATTTTAAGTTATAATATGGTACTCAAACAATATTTATTAAATAACCATGATTATGA AGCCGCCCAAGAAATGGAAAAAGTGTGAGGAGGTGCTGCAGATAAAGCAGAAAAATGAAG ACGGTAAGGAAGGAATCAGTGGAGATATGAGTAGTGGTGGCTGGGGTGACCGGAGTCCTT GCTGGTGGGGTGGGGAGGGCAAAGGCGGCGGCGGGGAAGGAGGGAGGAGGAGGGCTT GCTGAGGTATGGGGTGTGGGCTAATTAGGGTAAGGAGATTGAGTCCATATCTAGTACACA ATGTAGTTTGATCCAATGGTTAGAAATGAATTAGTTTTTTTAAAAAAAACATGTGCCTT CTATATCTGCGGTAAGCTGGTCAAACTTTTTTTAATACTAGTCATATATGACACCGCAA AGAGCACCTCGACCCTTTGTACTGTTCTTGGAGGCCTAACATTGCTATTGTCTTCTCTAT TGCTAGAGCTTCACTATTACACACCTCTCTTGTGGCACTAGTCACACTTTCCCTGACAGC GGTCGGTCAGTTAGTCAAGGTTGGGTGTGGCCTTTACAGGGAATGGTTTGTGCGCACGGT GGTATGCCAAGTGCGGTGATGTTGGTGGTCTAGCGAGTTGCAACATCACAAGAAAGCAAC ATGCACAATGATAAATTGAGAAAAAGGACGTGCGAGTTGATTCTCACCGTGCAACAGTTA GGGAGCAGTACACCAGGCGGCAACACGAACAAAACTATGGGTGGAGAGGTTCCAGTCGGT GAAGGTCGAGCACGCGTTTCATCCTCGCGCTCTCGGCGGCCAGCAAGGAAGCGGGAGCTT GTAAACGATTTGTCTGGCTCGTGGCTCGTCGAAACGGCGTTGGCTCCCACTTGTCCTTGT CTGGGCACTGCCTTGAAGTGGCATCTAGGAGAAGTGGTGGCGTGCGGCTAGCTGGCAATG GGCAAGGAGAGTGGTGATGAGACCACAACACGGGCAGTGGTGGGGGAGGGGGGAGGGGTG GGGTACCTTTTCCGGTCATGAAGGCGACATGTGCGTGCCAAGTGGGAAAGGTGGAGTTGA GCCGGGCACATTGTTTTATAGACGCGGACGGCTTGACTCACTAGTTTGAAGGGCTAGGTG ATGTAGTGGCGGTCATCGTTCTCGCGCTCATGTGGTTGATGCAAGCGGGTATGGAAGGTG GGGCTGGGCAATATGACGGTCGCGTCCACTTTCTGTGCGAAACCTAGATGAGGAAAGTGC GGCGATGGTTTGTGTGTGCAGTGCATCGGGGCAGCGTAGACAAGGGACTCCTGATGGTGT

CAGCGCCCACCATCGATGCACTTGACCCAGAACAGTGGCTGCAAGGTGTG
GTCATCATGG
TAAAGGCGAGGCAACATAAGTCTTCATTGTGGGGCCCACATGAATGGTAG
TGGATCTGAG
GTGGATGGTATGGACAATGGTGCTTGGGAGCCGTGCACATTTGGAGAGGC
TCACCGCAGC
TTGGCCTGTTGTGACATGGCCTTGTGTGAGTGGCACGAGGCGTGGTGCGGC
ACACCTCGG
TGCCAAGGCCACGGCGTGGATGGATGCGGCTGTGTGATGGTGATACATTG
GTGGGGTGGC
AGGAAGAGGCTAACGCGGGGAGGAATGGTTGGGATGCCAAAGTTATTGCC
ATGGGGCCCA
CTGGTGACACGCGCTACACGGTGCAAGAGCGGTCCGGCAAATGTGATGCA
CGCGGTATGG
CTGAGCACGGGCGGTGCATGGACACAGCTGTGGTGGTGCGAAGCTTTTCC
CACCCTGGTC
TGGTGCGTCACAACGGCGTGTAGGCGATGTGGCTTCAAGGCTTGTGGCAT
AGCATACGAG
AGGAGTAGGGCGGACAAATGACGCTTGGGGAGCAGGACTCAACAATGGC
TTGGCGGCGTG
GTGTGCCCAGCTAGGATTCGGTGCCAAGGTCACGGCGTGGATGGATGCGG
CTGTGTGATG
GTGATACATTGGTGGGGTGGCAGGAAGAGGCTAACGCGGGGAGGAATGG
TTGAAATGCCA
AAGTTATTGCCATGGGGCCCACTGGTGACACGCGCTACACGGTGCAAGAG
CGGTCCGGCA
AATGTGGTGCATGCGGTATGGCTGAGCACGGGCGGTGCATGGACACAGCT
GTGGTGGTGC
GAAGCTTTTCCCACCCTGGTCTGGTGCGTCACAACGGTGTGTAGGCGATGT
GGCTTCAAG
GCTTGTGGCATAGCATACGAGAGGAGTAGGGCGGACAAATGACGCTTGGG
GAGCAGGACT
CAACAATGGCTTGGCGGCGTGGTGTGCCCAGCTAGGATTCGAGCAACTGC
ACGGTGGTCT
CTGATGGTGTCGGGGCATGCGAGCCATGTGAATTTTTCTTGCGAGCTCGTC
GGCGGTAGG
GGCACTATAGACTTTTCTGCACG
GGACATGCATTCTTGGTACTGAAGATTCATTTTCCATTTAGAAAGGAAAAC
TGAATTGCT
GTAGCAGGGTGCATTTCTCTTTCCTACTTAGGCATATTGCCACCACCGTAG
TGAAAGTGC
TTCCTCAGCATACATAGATCTAGTTGTGAATGGTCATGCACAATTCCTGCT
TATTTGGTA
GCAGGCATAGTTAAAGAGAGGAGCACTTTAGATTGGTCATTAGCTCTCCT
AACATAGTTA
GAGAACCATCGACCTTGATGTGGATCAAGCAAGAAGGTGTTAGGTGTTAA
AATAGCTCTA
LOC_Os11g15620 GCCTGGTAGAATGAATTTTACTTATTTTTATTTTTTATTAGCATCTTCTCAA
AAGATATA
AGGCCTCATTTTATGTGTAGAACCGAGGTCTCCAATTATTGGGGAAGGCCT
TTCCATGCA
CTGTGCAAGAAAAAGGGAGATATCTCATCTGCCCTTACACATTGCAGCCG
GCAACGTCGG
GTAGATAAGTTGAGTTCTTGCAATGAAAATATTAGGTTATCATGCTTATTT
TCCCTCTTA
TTTCTTTTGTACATTGTGTACAACTGTCCATGGCTCTGATGCTGGGAATATG
TTTACAGA
ATGCCTGAGAAGTTTACAAATGAAGGAAACAAGAAAGTGGTCATGGCTGA

TGGGGCAGAC
CACATAAGCGCTCTCCCGGATGAGCTCCTCCAATACCTGCTCTCCTTCATT
CCATCACGT
GAGGCCGTGCAGACGTGTGTGCTCTCCCAGCGTTGGCGCCACCAATGGAA
GTATGTGCCT
GCCCTTCGAATAAACGACGTCGACAGTTTCTACAGTGTTCAGCAGCTGAAC
AAGTTTGTC
CACCACCTTCTGCTGCACCGCAAACGGACTCCTCTGTATGTATGTGAGCTT
GATTCTTTT
CGCAATGGGGAAGTTGCAGAATGGTACCAATATGCTGTATCATGTGAAGT
TGAGATGCTT
CAGGTAGACACTGCCCATTCAGCTGATTATTGTCGATTACCCGAGATGGCT
ATCACCTCC
AACCACTTGACAACACTAGAATTTAGTGGTGTGCAGCTGGGGCATATCTCT
CTAGATTTT
TCAGGTTGCCTGAAATTGGAGGTATTAGTGATGCGCGGGTGTAAAATCTTG
GTGCAGAAG
ATATTGTCCCAATCAGTTAGAAGTTTGAGCATAACTCAGTGCAATTTTGAG
CTCAATACT
CGCACTCATATTTCAGCTCCTAGTCTTATTTCTTTAGAACTTGCTGATATTC
TGGGCTGG
ACTCCTGCACTTGAGAGATTGCCATCATTATCAACAGCATTTGTCAGGCTG
GATGATCGT
TGTGAAGACTATTGCTTACACAGTTACTATGGGGATTGTGGCGATCAAGTT
TCATGTGGA
AAATACTGCACTAGATTCTATGATGTCCATGATGATGATTGTGTGCTTCTT
GGTGGTCTG
TCGAACGTCACCAATTTAGAGTTGATAACTTCACCTAAAGTGGTATGTTTG
CATGTATAC
TAACTCCTAGTATCATTTATGCCATATGCTACATTTTTTTATTCTAATTAGA
GTTATATA
TACATTTTCTATCTTGCAAAGGTCTATGCATGCAACCTGTCTATGATCAGT
AAATAATAC
ATGCAAAGTATTGACTAGCCTATTCCCTGTGCTTTTCTACAGATTAAAAAG
ATGAACTCA
ATTTTGAATATTAAAAAAAGTTACCGTTAGACGTAAAATGTTTTGTATATT
ATTTGATAA
AATGCATATATACATGAGTGGAAGACGGAATTCTTAAACAGATCACAATA
ACTTTTAAAT
TTGATTATCTCTAAATCACAAAGAACTCTTTCTTGTAAGGAGTGGAGTTGG
ATTATGTTT
CTACGCAGATCTTCCTAATGTTCTTTTTTTTTTGTGAATGCCAATACGGTAC
GTTGGTAT
CAAGTACTTCTAACAGCACAATATTCAAATGCATGGATTTGAGATTAATTA
GTGTTGTAT
CCATTGAAGGATGTGTATTTAGCTGGTATGACTCATTTAAGTATTTTATAA
AAACATAAA
CATGCAAAATCTTGTACTACAAAACATAAAATAAATCCTGCAGAATGCTC
ATATTTAATC
CAGTATGGGTCAAAACTTTTGGAGTTTTGATGTTAAATTGTTCATATTCTGT
TGTGCTTC
ATGCAGTTCATTGTCAGAAAGGATTTGCTTATGCGCCCAACGTTTAGCAAG
CTAAAAACA
TTATTGCTTAATGTTTCGGATGCGGATGCTGGCTTCGGCCCATTGGTTTAC
ATTTTGAGG
AGCTCTCCAGTACTAGAGAAGCTTACTCTTCAACTGTATGAGGTAAAAGA
GATTGTATTC
TTTGATATAAAAACTTCTTGTTTATTAAAGCTATTTTTGTAGTTTTCCTATT

	TATATTAAAGAGTACTTTGCTTGCTACAGGAACCAAAAGCTAAGATAGAAACAGATGGAAGCTGCAACCTAGAGGAGCAATTGGTGGCATCCAAGAAACTGAAAGTAGTTGAAATCAAATATTCGAAAATTGTAGTGCTTTGCAGAGTTTTACAGATTCTCAACACTTGTGGTGTACCTCGCAAGAAAATTAACATTGAACGTACGGAGTTGTGGTCTTTTGGAAGTAAGTTAACTGCAGACATTCATCATTTCTCTCATATCTCCCATGTCAGCTTGTGTGTATCTACCATCTAGCAAACTGAAACTTCACAACATTATTAGACAGTGTCCAGACTAGAAATATATGTTAGCTTACAACATAATCTGGATGGATGAGTAAAGTTATGATCATCTCTATGCTTATTCATGTCCCTTAGGGCTATGCAATACATTCATGTTAGCTTGGTGCTTTTCAATTTTATCTGAGCCTTCCTGTTTGTCTTAATGTGACTTTTGATGTGCTTACTTCTTTTTTATATATGCTTACACAGGTCATTTCAGTTTCGAGCAGACTTAACAGTTCAGCTGTCTCGTTTGTCAAAATCTGAAGTGATTTACCTGGTCAAGACATGCATATGTTTTGTGTAAGAAGTTCAGAAACACCTATCAAGTACAATAAGAATGGATGTAATGGTATAAGTTTCATGTAATTGAACAATCCAGTATTGTGTGATGATTACCTGTAGAACTTGTGCCTTGTGGCCTATATATATATGCTCTGTGACATGCTTCTGTCTTGACATTGTGTGGACATAACTTTTACTTGCTAGTCCTATGGTGTTTTAGTAGGATTTAGCATTTCTAGCAAGAGATCCATCTGTAGGAATCATTCAGCATCTTTAGCATTTCTCTTTGCACAGAAACTTCAACCAGAAGCATGAATGCAGGATCTATATGTGAATTAGCTTATTCAAACAGCTTGCCAGGCCATATGTTTACCCATCACTACTACCATCCATTACTGTACAATTCCTTTGTTGGCTTATTTCTGTGTTTCTGGAAACAGTAGAATTTAGAAATGAACAAAGGAAAATGTTAGAGGAACAAAATAATCGATACAGAGGTACAGTCCGAGGGCAGGTACAATAGCAGGCTATAAGCCGGCT
LOC_Os11g15710	ATAAACATATTCTGAGCGGAATCACCGGCGACGGCTTCCTCCACCGCGCTCCGCCGGGCTCAACGCGGGGATGTCGATGGCAGTAGGGCAGCGTCTCGCCAGGCAACAGTCGTGATGATCCTCGCACCGCCGCCGACAACACAGCAGAAGCAGCAGGGGATTGGGTTCGGCGGGAGCTGGAACAGCAGGAGACAGGAACAGGCGCGACAGCGGCGGCTCGTTAGGACCAAGCGCCGTCGCTGGGAGCACGGGAGGCTAGCATCCATGGAGGCGAGCTCGGTAGCATGCCGACATCGATGTGCAGGACACTCATCATCTGGCCACCATGACCTTTGTACTGAGATCTGCGCCGCCCACAACGGGGACCTCCAGGTCGCCGTCTCCTCCATCACCGCCGGGCAGCGAGCCAAGGGAGTGGTGCTGCTGCCGGTAATTGCGGTATGCCAATACCAACATGAGCAGTGAGCCAACGAAATCAAGCAACATGTTCAGATTCCGCCACGAGCAATATGATCAGATTCCTCGTTTTGCATGCTGGCCACGTACACGTC

LOC_Os11g15700	GCGGATCCGTGCAAGCAGTCGCGTCGGGGAAAGCAGCGCTTACGTCACCTGCACGCAGAC ATGGAGCAGCTGCCACAGCCACCATCCATAGCGCACACCGAGCTGCTATGTCGCCCCTAC TCCTCCCTTCTCAGTTTTCAACCTCAGCGGGTTGACGCGGGAGCGGTCGAGACGGATCTC GACCACACAAGAGATGGGCATGAATGGGTAGAGGAAGGCGAGGCCAACGAGGAGCGCAAG GCCGCGAGGAAGAGGAGGAGAGGGAGCGTGCGGCCAGCGGACGAGGACGGGGATGAACAG AGCGTCGGCGGACACCCTGACCTCAAGAAGACGCTGACGTGGACACTCTAACGTGTAGGG CCCACATTGACTCAGCCGCCTCGTCGGACAAAACCGGAGTCAAAATAAACCATCGAAGGA CCAAAAGTTAACGGTTTTATAAGTTGAAAGATGTAATATATCCGGTTTTGTGGTTGGGGG ATGATTTTGTAACTCGATGACAAGTTGAGGGACCTTCGGTGTACTTTTTCCTCTACTTTT CAACCTTTTCAAACAAACTTACGGCCCATCGTTTGGCAGTATATTTGCAAACGAAAAATA ATTTGTAAATAAAATTTTTATATATGTGTTCTTAGTGATCTAAAAGCAAAGTCTGAAAAA TAAACTTCGATAAAAAAAAGAAAAAGTACGAATTACCCCCTCAACTATTGCTCTCGTCCG ATCTACCCCCCTAAATCACAATACCAGACATTTTTAACCCTGAACTTTTCAAACCGGATG AATTACCCCCTAACTCAGCTTGAAGCGGTTTTGATCCCACGTGGCACATACGTGGCAGTC CAGTCAACCAAACATAATAATAATAAAATAGTGGGCCCCTCTTGTCAGCATCTTCTCTCT TCTGCTTTCTTCCACCTCCTTCCCCTCCCATCTATCTCTCTCTCTCCAACG TGGCTTACCGAAGAAAGTCATCTCAGAAATCAAACATACGTCGGTGCTAGTCTCCGCTGC CCCTTTCTCTTGGTCCAACTTACATGTGATGTTTCAAGGAAGGTAGGCTAGAAGCCTAGA AGGTGGCCTTTGTGGTTAATGTGTTGTTTCATATCTCTTTCTCTTATCCTCAGTGAGCTT CAGAGTTTGTTTTTTTTCTGTCCTTGACATGGCCATTTTGTTAGAACTGCAGAGTTTCAT ACCTACTTCCTGATCGAAGCATCTTATCTCTAGCTGTCCCATAGAACTGCATAGCTTCTT AATTATATCTCCACCGCTCAAGCGCAAACATAGTTTCTCTGCAAGTCTCCGGTGTCTGTT GGTCCTGCTGCTAGATTTGGTGTTGCCCATATATGTCGGGAACTGCAGCTTGACTTCATC CTAAGCAGCAGGCGGGAGGATCTCTTAGCTGTTCTTGTGTAAGTACAATCATTGTTCTCC TCAGACTCGCCGCCAACCAATTTGATCCCTGTTTATGATAAGTTCTGCAAAAACTCTTAC TTGCTTCAAGCCTCCTCTGAGGACCGACAGGCTACAGGGTGCACACCAAACCGCAAAAAG ATGGAGGTTGCCATTGACTCAGCACGCTGGGTGGTTGACAAAGCACTAAGCCCGTTATCA GGGGGCTTGTTAGAGACATGGGCAGCAAGTTCTGAGCTTGGCGTGAACATCGATGCCATC AAGATGGAGCTGCTGTATGCCAAGGGGATGCTTGAGAATGCACAAGGCAGGGAGATCCGC AGCCAGGCGCTAAAGGAGTTGCTGTAGAAGCTACAACAGCTAGCGTTTGA

TGCCGATGAT
GTTCTGGATGAACTGGACTACTTCCGCATCCATGATGAGCTCAAAGGCACC
TATGAGGCA
GCTGATGTGGATGATGCAAGTTGCTTCCGTGGAATCCTTGTCAATTCTCGC
CACACTGCT
AAAGCTGTGGGTAAATCGCTCTCGTGCTTTCATACTGATGCTTACCATGGT
GACCCAGAT
AACCCAAGAACTGAAGCAAGGCAGCATGTCCACATGCTTAGTAAGTGCCT
TCCATGCTAC
TCTTGTCCACTTGTCCATGAAAATGATGCTGATGATGCAGTAAAATCACAA
AAATGGTGG
CAGAAGCACTGCGGTGCGTGTTGTTGCAGTCCAAATCCAGGGCGCGATCG
TGCAAATGAA
GCACCAATGCATAAGTTTGACAGGGTGGGAATATCAACAAGGATGAAGCA
TATCACGGAA
CAACTACAACCCATATGTGCTAAGGTCTCCGTCATTCTTAACATGGAGATG
CTGGGTAGT
AAAAGCAATACTCAGGACAGTACTACGAGCCAGCGAATAACCATCTCAGA
AAGTGTTGAT
CCTAAGTTATATGGACGGAACAAGATGAAGGATAAAATTATACGTGATAT
AACAAACGGT
ATATACGCTGAACAGGACCTCTCTGTCTTGTCACTATTTGGTCCAGGTGGC
ATTGGGAAA
ACAACACTCGTGCAATACATATACAACAACCAAGAAGTGCATAGCCATTT
CCAAGCAACG
ATTTGGGTATGTGTGTCCTTCAATTTCAATGTTTCAATGCTGATACAACAA
ATTAAGGAT
CAAATCCCTGAAGTTGATGGTGAAAATGGAACTGCTGAAGACCATATTGA
ACAAAGATTG
AAATCTAAAAGATTTTTGCTGATCCTGGATGACATGTGGAAATGTGATGGT
GAGGATATG
TGGAAAAGGCTGTTAGTTCCTTTTAGAAAAAGCCAAGCTAAAGGTAATGT
GGTTATAGTT
ACGACTCGATTTCCTGCTTTAGCTGAAATGGTTAATACAATGGACCACCCA
ATAGAGTTG
GAACGTTTAGAACAAGAGGAGTTTATGCAATTGTTTGAAGCATGTGTTTTT
GGTGAGGCA
AAAGCACTGTCACGCCCGGAATTTCTATCCAAAATTCCAAACGCTTACATA
TGTGTGAAC
CCTCGTCCAGAAATCAGCCGAGGCTCACAATAACAAATTGATAATAGAGT
ACAATTATTA
CTCTAATTAATAAGCGAATAAAATGTCATTACAGAGGTAGATAGTTCCTCT
CAATCAATA
AAGATCTAAGCAGCGGAAAAAATAAGATAAACGGCGCAGACGACTCCAC
TCTACAGGCAG
CTTGACCAGGGCTACACCTAATCCTCCACACCATCAGCATCATTGTAGAAC
TCTTCCTCT
GATGAATGATTGCAAGGTGAGTATATGATATACTCAGCAAGCCACGCAGC
AAATATGCAA
GTGCACAAGGATACCAAAGAATGGCATAATATAGGGCTCATTTGCAAAAG
CAGCATTTAG
CAAACATTTGAGAATTATAAAACAGTAGAGTAATTAATCATCAATTTTAAT
CAACACTAA
ACAGCACACCCATGCTGCTCAGGCCCAACCATCCTGAACAACCATACCCG
GCTGCACAGA
TCTATCTCCAAACCAGGAATATACCATTCCAAACCAGGAGCTAAACAAAT
TATTACCAGT
TATAGCATCTTTTATTATGGTGAGAAGTGTGAAACTAATCACGAAAGACAT

	TGTTAGACC CGCCCATAACCGCGGGCACGACTATTCGAATAG
LOC_Os11g15670	TGAAATTTTATCTGGCCCCATTTTTATATTGGAATTTATTAATTAAAATTAGATCTTCTC TCCTAGGCTTTTAAAATAAATTCTAATAATTCTAATTAAACAACAATTTATATATTTAGG ATTTTTAGGGTGTGACAAGCACCCTGGCAAGATCATAGTGAACTACTTGACATTGGGAAA AAAAATAATTGGCAAACTGAAGGGTTTCCCACTTGCAGCAAAAACTGTTGGAAGGTTACT AAGAAATAACCTTACTTTAGATCATTGGAGAAGAGTTCTAGCAAGTAAAGAATGGGAATT ACAAACAGGTCATAACGATATTATGCCTGCACTGAAACTGAGCTATGTTTATCTTCCTTT TCATCTGCAGAAATGTTTTTCCTATTGTGGCTTATTTCCTCATGATTACGAGTTCGATAG TGGAGAGCTGATTCACCTGTGGATGGGATTAAACATTTTATGCTGTAATGGACAGAAAAC ATTTGAAGAAATAGGATTAAGCTATCTTGATGACTTGGTTGATCATGGGTTTATGAAGAA AAATGAAAAGGAGGACACCCTAACTATGTGATGCATGAGCTACTACAAGACTTAGCACGA ATGGTGTCATTGTATGAATGTTATACAATAGATTGCTCTCATGAAAGATTCAGAGAAATC CCTCCTGATGTCCGTCACTTGTCTATCGTTATGGGTGGTGCAGAGGAAGATGCGTCAATA AATGAAACTTTTCATGGAAATGTGATTCTTATAATGAAGAGATTTAAAGTTGAAAATCTA CACACTTTGATGATATTTGGATGTTATTACCGAAGCATGGCGAATACCTTTGGAGATGTG TTCAGAAAAGCAAAGGCTCTCCGTGTTCTTCGTTTGTCTACTATGTACTATCCTGTCGAT CATATACTGCATAACTTTTCCGTGCTTATGCATCTCCGCTACCTAAAGCTTGGGTCAGAG TATGATAAGATATCCCCACCGAGATGTATATCTAGATTCTATAAGTTGATTGTCCTTGTT CTGAAAGATTGGAAAGGAAGTATCAATTTGCCTGTAGATATGAGCAACCTTGCAAGATTG CGGCATTTTATTGTCTCACACGATGAAACTCACTCAAAGATTTGTGAGGTAGGAAAGCTG CAAACACTCCAAGAGTTAAGAAGATTTGAAGTCAACAGAGAGAAAAGTGGTTTTGAAATA AAACAGCTGGGACATTTGATACAGCTCAGTGGATCACTTAGTATCTGTAATCTGGAGAAA ATGCAAGCGAAAGAAGCAGATGAAGTGAATCTGTTGTCAAAAAATAGCTTAAAGAAGCTA ACATTAGAGTGGGATGTTCAGCGATCCCAGAAGGAACCTGACAAAGAACAACATATTCTT AACGTTCTTCGACCTCATGACAATCTCCAAGAGCTATGCATCAGAGGACATGGAGGCCAT AGTTGCCCACCATGGCTTGGTAGCAAACTTTCAGTAAAAAACCTACAATCCCTTCATCTT GACACCGTCAACTGGACTGTCTTCCCACCACTCGGGGAATTTTGGCTCCCTAAAGAACCT GGTCAGGAATATCTTCGTTCTGTCCAAGGAAAAAGCTTTCAAAACTTGAAAACTCTAGAA CTTGTTGGACTCACAAAATTGGAAAAATGGGTTCACAATGACAAATTTCTCCTGTTTTCC

CTTTTAGAAACTTTCATAATTAGAGATTGCCCAGAACTGGTAGAGTTACCA
GTTTCTCAA
TATGCTAGTCAGAAATTTAAGCAAGATGTGATGATTGATCTGTTTCCTAAG
ATGCAAGAG
GTCAGAATTGCTGACTGCCCAAAATTGGAATCCCTACCACTTATTCCATGG
ACTGATACT
TTGCTCACTGTTGACATGAAGAATGTGGGTTCAAGTTTGGAAAAGTTAGTT
TACTCCACT
AAAAGTTCATCATCCAAACTATTGTTGGAAATCAAAGAAGATCATCATCT
GGAATGCCTT
GATGAGATGGTTGCCTTTCACAATTTAAGTAAAATACATGAGTTGGAGGTC
TCAAAAAGC
CCACCTTTGATGAATAAACACCTTCACTGGTTAACGTCACTGAAAACACTC
AAGATAAGT
GAATCAAGTATTACATTGCCGCTACTGGGAGGGCCAGATGATGAGAAAGA
CACACTTGCT
TTGGAACGCCTTGAGATTAAGAACTGTAGTGCCAATGGGAAGGAATTGAC
CCAATTTCTA
TTGCAACTGCCTAAGCTCTCCTTCTTCAGGATGTCTTCTTGTCAGAACGTA
ACTAGCATG
GGTGTAATGGCGCTGCTGGCCACAGCTGAACCAACATCAATGCCATCGTC
TTCGACTTCA
AGTAATGAAACAGGATCACAACTACAAATAGAAGAAGTGGGTGATGAAG
GAGGGCTGCTG
TTATTCCCAAAACATCTCACCATTTCCCTGCGTGAATTACGGATTACCATG
AACCCAGGA
TTGAGCCTGTTGGCTTCCCTTCCTCCTGAAAATAACTCCAGGCCAGGAGGT
CTTCACAAT
CTGCACTCCCTGCAAACACTTTTCATAAGAGGGTGCCCCAAACTTCTCTCA
GCCTACTCC
TCATCCTCCTCTTATTGTTTTCCGTTTCCATCTACACTGGACAGCCTCAGAA
TTGAGGAT
GTTGAGGATATGCATACGTTCGCGCCCCTCTCAAACCTGACTAGTCTCACC
TATTTATTC
GTAGAGAATTGTAGGAAGGATTTAAGAGGAGAAGGTTTATGGACTTTCTT
CACTCAAGGC
TGCCTCACCCGTCTGTGCGTTTATCGCAGTCCAAATTTCTTCGACAATTTGG
TGCCTCAT
CAACAAGAAGAGCTTCCTGCATACTGCAAAATTGAAATGCTCCGGACAGA
TGATATTGCA
GGGGTCCTTGTGACACCTATCTGCAGACTCTTTTCTTCTTCGCTTAATGTTT
TAGGCCTA
TGTTCAAACAAAGAGATTGTGAGCTTCACAAAAGAGCAAGAGAAAGCACT
TGAACTCATC
ACCTCCCTCCAGGACCTCTGCTTTTTTCACAATGAGAAGCTGCAGTCTCTC
CCTGCAGAC
CTTCGTGGACTGAACAATCTAAGGATATTGGAAATTCTGAGGTGCTCAGCC
ATCAGGTCT
CTGCCCAAGAATGCCTTTCCAAATTCTCTGCAGAAAATAAATGTTGATCGT
CGCTGCAGT
GAGGAGCTACAACATCACTGCATTATGTTGGAAGGGGTCACAGTAAATAT
AGACCGACCA
GTTAACACAAATTTATAATGCAGAGTTCTTCTCGGGAATTTTATATTTTAG
TCTTGTAGA
CTTACCATAACTTGCGCATCCATCTACAGATTGTATCTACTCTTCGAATAT
AAGACATGT
ATTTGAATACTATGTAAGTTTTTCTTTGCTCTGAATGACAGAAAGAACATT
TCAATGCTT

LOC_Os11g15660	CTCAAACATGGACAATTCTGCAATTATCATTTAGGATTATAATTTGTTTGACCCCTCTGGAGTCTGGACCATGCTAATAGATATGTGCGGTGCTGTGCTCAAATCTTGCTATATTTACACTTCTGCATATCTTGTGTATTTTTATTTGGATCCTGGTACCATTCATGCATCATAGCATCATATGCTTTAATTCACATGGTAATATATTTTCCTTGCGTTGGCTCTGTCCCCTAATGTGCCATACAAATGTAAAGATTGTACTGCTAGTGAGCGTAGATTGTACAAAGATTGGGGGTTACTTCCATTATCTTTTTTTCTTTTAAAAGGAATTTCATTTGCTGCTAGTCCGGTGTTTTCATATTTATAGTTCCTCTGTTCTTCAAAAGAACACATTTGTTTTCATCTCATTTCATCAGTTCCACCATCTCCACCAGGCAGCACGTGAGATTGAATTCAGATGAATTCCATTTAATCTTACTGTTTCCATTTGTTAAATATGAAACCAAGCTATCAACATGGTTTGAATATTTAATATTTGGTTTGTAATTCATACCATACAAATTAATAATTAACTTTTCCTTCATTTAGCTTCCTGTCTTGTGCATAATATTTATGGGAATGATATGGATAATGTATTCTTGTAAATTGTACATATTTAATTTAATTCTCCATTGTGACGATTCACGGTTTATCAAGCAATAGCAATCATAACATCCAGCATGATGCAAAAGTTTTAAAAAGTTCAAAATTTTAGTGATTTTTATTATCAGGAATTTTCAGGATGTTACAGACACGCCCGGTGGGACCGATTAGAAGTCAGAATCGCAATACATTCTCTCATGGCTGAGAAACTTGCCACCATCACCGTCTTCGGTTGGCAAAGTGGTCCCTCCCAAGACTAATAATCTAAAGCTTGGTGGGTTTGAAGTTAATGAGGGAAACATTGTTCCCATCACCGTAGACAAGCTAACACCTGAGCAAAAGAAGGAGCTTGATCTGATGATGCAGCAAACACAACACTAGTTCTTGAACTCATTCATGAAGACCCGCAAGGGAACAGTTTTTCAAAAATATAAGGTAAAATTAGTTGCTGATGTTCCTGGAACAGGTTCCTCCAAGGATGGAGAAGTACAACAAGTTCCAGCTGGTACAACTGAATTAGGTGGTAAAGGTGCTGCAGAGGGCTCAGGAGATAAAGGTGATGGTTCCCAGGGGGTGCAAGGTGAGGATCTTGGTCCATATGGTAATACATCACAATTACAGTTTAATAATTTCCAAGATCGGATTGACTACGCCATGCAACATGCTTTGATCAATCAATACCACCTTCGCGAAACTCCGGCCATCCCCTCGAACGCTCGCCGTCCACTCTCGCCGCCGTCCACCAACGCCTACCCTTGGCCACTCCTCCACTTCTCTCACCTCCCCTTGCGTCTCCCCCGTCTCCAATCTCCTCCACCCGCCCTGGAACCCTAGCTCCTCGAGACTAGCTCCCCACAGGTGGTGTAGAAACGGAGTCAGGAGAGTCGACTCAAAGGTGTTCGACTGATCTCCAAAATCGCCATTGTTGGATCTCGACACGTGTTGCCCTTCCACCTCTCTCTCCCCCACTATTGGATCGGGCTGAAAACCATGGGCTTCCTCCCTCCCCTGATCAAAAACTCTCTCGCGCATTGAGGAGCACTCTCCCAACTCCCGAAGGCCATCCCCTCCGCCGCGAATCACCGCCGGTGATCGTCGGAGCGCCG

LOC_Os11g15640	CGGGTGACTCCCTCGTCGCCTCTCCTATGGAAACCCTAGCCATCGGGTGGATTGATATATATGTATCACACATCACGTCGATCGATCGAGAGTGATGTGCATATTGAAAATTGGTTGGTCCCTACCAAATGGCTATATGCAGAAGATTCCATAAGGTTGTTACGAAGTTGTTCAAAGTAACCTTGCATGGGTGGAGATTGGGTAATAGGGTACATTCCCATAGTTGCTACTTGCTACAGCACGCTGTACCATTTCATCAAGACAAATGACCAACAAAATTGTAGAAGGCAAACACTTTACTGTTCTTCTTGAAAAAAAAATATGTTTCTAACAAGTTTTACCTCTAGTTGAAGATCGAAACTGATGCAGCTCAACCTGAAAGCAAGAACTCGTCGAAACAAAACGAAAGTAAAAGGGTGGCGATGCGCCGAGATTGTATTGAACGTGTGTTAGATTGATTACATGGGGCTCGGGGTCTATTTATACCCGAGATTACAAAATATGTCCATGTCGGATACGACTCTTATCTCTAACAAACTCTAATATACCATATGTCTTTGCGGCATACTTTTGCCCAAACATATCTCTAAGGAAAATACATAAAATATCCTAATTAATAGATACAACTGCCTTCTCAAGACTCTATCCATGCGTGGCAATCCTCATGAAGCACATTAATCACAACCCGACAACATATACCTCATCATATTGTCAGATCTGACTCGATCGGCTAACCCCATCTGACTCGAACTCTGTTGATCCCAGTCGTAGCCGATTCGGACTCCACCCGATCATTTCTCTGTCTCCGGAGCGATCTCCATCTCGAATTCCGCTTCGATTCCATCTTCATCTCCGATACTTGGATTACCAAATTCGGCCGTTAACAGTTGGTACGAAGCGGCTCTCTAATGGCAAATCAAGGCAGAGGATGGAGGCCGCATCCACCTTTCAGGCCACAGGGGGAAGGGGACGTTTCTGCCCTTTGCAGGTTGGGCGACAGGAGCCGCAAGGTCCATATCCGCCACCTCCAGATCTGAATCGAGGAGTAGAGGATGGGAAGGCGAAGCATGAGGAAATTCTGCAAGCTGCTGAGGAGAAGGGGAAACAGGGGAAGGACACGCTGGAAACTGAAGGGAGGAAGGAAGCAGGTCCAAATCTGCAGGAACCAAGTGGTAAGATGGATTCTGTCAGTTCAGGATTGGAAGCAAAGCCAAATCAGCAACAAAATCGAGATGAGTCCAAGCATCAGGTTGGTGAATCCCATTCCCAGCAGAGGATTTTCTGTTCTAAATGTAGGAGTTACAGTCATTTGTCTAGAGATTGCAAATTGTCAGTGTTCTGTATAATATTTGCAAAAGAGACTCATAGGACAGAGGATTGCATTAGGAAGAATCAGAAGAAACCTGTAGCTAAAGTAGTAGGATGTGCTGCTCCAGGTCTGGGCTATGTGCTGATTCAGAGTGCTAAGGGTATAGTGCAAAAAGAGCATGTTAACTCTTTGGCTATTGTGACAATTATGTGGGGAGGTGATCTTACTGAGCAGGCTTTGGAGGAAGTATTCACTCCACATTTCAAGTGGAACTGGACTTGGAAGGCCAGGATGCAGGAGAAAGGAGTGTCCCATATGAGATTCCCAAATAAGATGAAATTGGATGAGTTAGCTAAATTTGACTCAATTAGAGTAAAAGGAACCTCTGTACATGTGAAA

GTGAAGAAATGGACTCAAGAAGCTGAGGCAGTGGGGATGTTGCATGAAGT
GTGGGTGACT
GTGGAGGGGGTACCTGATGAAATGAAGGATTATGACCCTCTACATGAAGT
GGGTTCCAAT
CTAGGACCTGTGATAGAAGTGGATATGGTAACTCTAAAGACAAAGGATGT
GGTCAGAATT
AGAGTAGTTAGAGTAGGAATGATGACACTAAAAAGCTTGCCTCTGACCAT
GACTCTGGTG
ACACCTAAGTTGCTGGTTTACAAAGCTCATTTCAAGTTGGAGCAGATTGTG
GAGTTGGGA
TGGTTTAGAGATTGTGCACAGGAAAAGAGAGCTGTAGAGGTGGCACAGCA
AAATGAACCT
AGCAACATAGATCACATGCAGAGGAATAAGAAACCAAGGAAAGAAGAAG
CTGCGTCAAGC
ATGGGGGTTTTGGAAAAGGCAGGGGACAAGAAGAAAACCATAGTTGTGG
AAGAGGATTCA
GACAATGAAAGTGCTCAAGCAGAATTGGTAAAGCTCAAACAGATGGAGAT
AGATAGGGGAG
CTGGCTCTGAGAGTACAACTAGAAGAACAGTTTAAAGTCAAGCAAATTCA
ACTTGATAAA
ATGTAGCAGGAGGGATGATAGAGGGTGGAGTCAGAGAAGAAGCATAGGC
TGGGGAGGAGG
AAAATGAGAAAACAAAGAAGACACAAGAGGGGATGGCAGATAAAAGTGG
TACTAGTCGGA
GAGTGGAAAATATAAAAATGACAGTGAGGAGGTGATTTATGATGAAGATG
AACCAAGAGT
GCTGCTGGTGGGAAGTGAGGAGATCATGGAATCACAAGAATCAAGTGACT
TTGCTGTAGC
TGTGGGAGTTGTTCTGTCTCAGACTGATGAAGACATGACAGAGGAGAAGA
GGAAGAAAAG
TCTAAGACTTATGGAGAAGGAAGACAAGAAGGTGGCTGATGCTGCAGCTG
AAAGGAAAGA
GGCTCTTAATGCTTTCATCAACAAAGGTAACAAACAAATCCTATCTATAGT
TGATGAAAG
TAATTTTTCTCTAGCTCAAATGGCTAGTGTGTTAGGTGTTAATTTAGGTTGT
GCTCTAGA
AGCTGTGGATACAAATCTGAACTTAATCAAAAATCTGGAATTGGCTAGGA
TAAATCTGTT
CTTGAAAGAGAAACTAAAAACTTCTCAGAAAGGGGAGTGTGAAAATGTTG
GGGAGAACTT
GAGTGACATAGACATAAACATAGATGAAGTCTTACATTCAGATGATTGTA
TGTCTGATTT
TGATTATGCTGAAAACATGTTGAGGCTTTCTCAAAGCAATTCATCAGGAAA
AAAGAGGAG
GAAAAATAATATATGAAACCTTGAGATTTTCAAGGTCACTCCTAAGGCTG
CGTTCGAGCG
CAGGAGAAAGAGAGAAAGATAGTTTTGTTTTCCGCGCGCACGCTTCCCGA
ACTACTAAAC
GGTGTTTTTTTTGCAAAAAATTTCTATAGGAAAGTTGCTTTAAAAAATCAT
ATTAATCCA
TTTTTGAAGTTTAAAATAATTAATACTCAATTAATCATGTGCTAATGGCTC
ACCTCGTTT
TGCGTATCTTCCCAATTTCCTCAATCCCCTCCTTCTCAAACACACCCTAAGT
GTAGTGGC
AAAAATGAGAAGAAAAAACAAAAAAGAAATCATGAAAGGTTTGTTTTGG
AATATAAGAGG
TATGGGTGATCATGATAAAATAAAGCATCTCAATGATTTAATAAAAGAGC
ATAAAGTTGA

LOC_Os04g45510	TTTCATTGGCATCCAAGAAACTATGAAGCAGGATTTCACTGCTAAAGACTTGGATTTAAT AGGAGGAGAGGGTAAGTTTAGTTGGAATTGGACCATTCCCAGGGGTAGATCTGGGGGTAT GTTGGTGGGAATAAATGAAGACAAAATGGAAATTATGGAAGTTTCCAAGAAGAATTTTTT GTTACATGTTAAAGTA GCTTCTTGCGGGGGAATTAATGGGCCTGTAATATCATAAACCATATTTTATTTCTGAGAG TACATAATATTATTATCCATGACAATGGCGCTTGTAAGTTCATTAATATCCATTACATAA AAACAAAAGGGTAAGTTCAAGTGGAATAATCTCTCCAACTGTAACACAAAGAAAAAGAGA GATACTCCTAGATGAGTTTTCAGTTCAGGCTATTTGATTCCCTCTGTCTTCCTTGTGCTC ACACACCTAGTACTACACGATGCAATGCTCTTTTACAACAAATATTCTTGTGGTTGCCAC GTGCCTCGCCAGTAATTTAGAGGCTAAGATTCTAGCTATGCATCGAACCATTCAAGCCTT CCCGTGTCAGCCACGCGCTCGGGATTAATATAATGCTAATATGTCCATCACATGAGCTGG ATACATCTAACTTGGCCATCCTCCTTTTAGCACCCATTTGCACTAATTAGTCCATCCTTG ACTTGGAAAATTTGTGGTCTTTCCATAGCCAAATGACAAAGAGGTGACAAAGGATTAGGT GAGAGTTTACCTAGCTTGTGCCATGATTATAAATAGCTATGAGCCCCTAGATATGTTCTG CCCATACCATTACCATAGACGCAGGTCTTTCTCAGGAAGTGAAGAGAGCAGCAGAGTTGC ATATCTTGCAGGGTTTATCAGGTAAGCTCTATATAGGTGGCTTCTTCTCTTCTGAGCTCT TCATTTTTATTTGATTTTCCTTTTGCTAGTGTCATTTTTCTGCTTCTAGAAAATGGAGAA TACTCATGAATGTCATTCTGCTCCTGTTTCAATACAAGTCCTTGACCTGTTCTCTAACTA ATAAGGAGCTAGGAAGCTTTGATCTGCAGAAGATTTTCCGGCACTATACGGCATGTTGTT GATCATCTGACAATGTTGCCACAATTTCATGAAGTTCTCTCTGTAACCTTTGTTTCTTTC CCAATCTAACAAAATCTCGGCCGGCAATCTTCAGTCAGCCCGGCGAAAACACGATTCCAT AGAGTTTATCTTGCGGTTGAATTTCCAAAACTGAAGAATTTAATGCCGATCTTTCAGATA TGATCAGTATTTCAACCATCGGTATATGAAATCAGAATGCGCTAATCATCAGTATAAGCT GAAATGGAACCAATAATGCATCTTACTCATGTGCATTTCCTTGTCTGTTATAAAAGTTAT GCCCTTTACAGCTTTACTAATGTCTAATGCATGAAATAGCAGAAATTTTCTAACACCTGG TGCTACTGTAATGCAGTTTGCGGCGTAAAAGGAGCAGTGTGTAGCAATTAACTGAAACTC AAATCGTCGACCGCTAATTAACTGAGGTATTCAGACAAAGAATAACAGGCATCATGCAAC GGTGGACGCGCAACGGGATGCTGAACTTCTCGCAGCCGGCGTCGTCGTCGCCGTTCGGAT TCTACGACGCCGGCCACGGGGCGTCGAGCAGCGGCGGCAGCCGGTCGCACGGCTGCAGGA GCCAGAACAAGGACCAGCGCAACCTGGAGAAGAACCTGACCAAGGTGAG

GAAGGAGTGGA
TGAAGGTGAAGGAGGAGATGGGCTACGCGAGGCTCCTCAGCGAGCACCTC
AGCGAGACGG
TGACGGAGACCGACCGGAAGGTGGCTGCCATGCTCGAGGAGCTCGACCGG
ACGGACAAGT
ACATGCAGGATATACTGTCGTCCCAGCAGAAATAGTAGTGACCCATGACA
TCAGATATTC
AGATGCGTATTTTGATGGATTTCTATGTATTTACCTAGCAGCTAGCTAGAT
CGTCAGCTT
CTTTTCTGGGCATTTTTGCCATTGTGGTATTCGGTTTTTTAGCAGATCGATG
TCCCGGTA
CAGGGTATATATAAATCATATGTTACCGATATGTTCTGAAATGTCGTTAAA
ATTCTTCAG
CTTCGCCTCCCTCTATACATGCCACTTCCTTGGTTAGTTGGTTTGCTTGATG
AGTCATCT
CATCCGATTACCCATTACGTAAGATAAAAGCTAGCAGCATATATATGCCAT
TAATCTGCA
CAAGGAGCAAGTATTTGAAACGCAGGAATTTTTCCGGCCAGTTCTTCTGAA
AAAACAAAA
CAAATATATATAGCTAGGTTCTTGTAAAAAATATTTATAAAAAAAAACTG
AAGTTATTGC
GTTTCAAAGAGAACGAGGTAACCATCGCGTTGGATCAAGTTCCTCATGTAT
TTTCGGAGC
GATTAGCTTTCATTAGATTTAATGGCACGCATTGTGCGTGCTGATCGAACA
GTACAACCT
GTACTACACACACAGTACTACGTGGTAGCAGTACAAGAGGACTCTATAAG
TGGAAGCATG
GCAAAGGCGACGGACTAACTTTGACCCGACCAACAAATCCCCAATTAGCA
AACAAGGGAA
ACGTTAGCCCTTTCTCTTCTTCTCTTTACCGTGCAAGCTAAGTCCAGGAAA
GTGAAAGCT
CAGGGTGTCGATTACATCATAGCTGACGATTTCTGGGGGGTGGTACGTAGT
AG
ACCCAATTTCGCGTGTTTTCGCTTGGCAAAATAGTACTACTATAGTTGCCA
CTTTTTTTA
CGGCGTACGCGTGCCACCGGATGCATTCGTCCGACATTTTGATGGCCTGCT
ATCAGGCCA
CGTAGTTGGCAGCTCTACCTGAAGCGAGAGAAATTAAGCCTGGTAGATAC
CGGAGAAGCC
ACGTATCTCAGTTCGATGGGAACTTTTAACGGGTGACGCTCACGGCAAGA
ATATACAGAT
TATACCGGTTGAGATACCGTCCGGTTGACAACGTACATCTGCATGCTGAGG
CTGCAGCAC
CACAAGATCCATGTTTCAGAATCTCTCGTGATGGTTACAATTTTCCCTTTTG
ATTTTGGC

LOC_Os04g45500 AACTGAATTCGCAGCAGCAAACCGGTTTATATATCGCTCTCTTATGATTAT
TTTTAGACG
TGAAATAGTCTACCATGAACGATGCTTAAAAGTCGGAATAAGTTTACTTTA
GATCCATTC
ATTTGTCTCGGTGGTTTGGATGGCAGTTTCGACGAACGTGGCACCTGCGTG
ACTAGTTTG
ACTTGATCCTCGTCCTACGTAACGGTAACGTGGCATTTAAAACAGAATAGC
GGAACCCAC
ATGTCAGTGGGTCTCACCCTCCTCTTCTCCATGGACGGCGGCGCCTTCTTCT
CCCCTCTC
CTCTTATCACCGTGGGTTGAAAGGGAGGAGATGGAGACCTTGCTCCTTTGG
TCATAATCA
ATGGAACTCGCCATGGGCGTCGTCGTCGCCGCTTCCCGGATCGACGAGAC

	GACGCCACGTGTTCTGCCGCTAGCCTCCACCCTTCCTTGACCCTTCCCTCCGACCAATCTCTCTCCCTTGTCCCCAACCTCTTTGCCTGTCTCCCCCCTCCACCACTCCCTCCTCGTCAATGTCGATGTCACCATCGCTGCCCCTAACTCGCAGCCACCGCCGTCTTGATCTGGTCCGCCCACCGAGACTGCTTGAGAACGAGCAGGTGAGGGAAGTGGAGGTCGAGGATGACGCCAGGGGGGAGGGAAAGAGTGCGGGGAGGCTAGCTAAGGATAGAGCGGTCAGCGAGGCTGAAGCGGACGATGATTGTCTCCCCCTCAACCACCACTCCCTCCTCGCTGATGTTACCGCTACGGACCCCAACTCACCACCACTGTTGCCTTTGATCTGGTCAGCCCACCGAGATTGCTCGAGAGCGGGCGGGGGAAGGAAGTGGAGGTCGAGGACGACACGAGGGGTGCAGGGAGGTTGAATGGGGATGGCACGTTCGGCGAGGTCAAGGCGGACGACGACCAGCAGATTTTGTGCCCTCCTCATCGATGCCACCACCGCCTCCACGTCGCCAGGTGATCCCTCGTCCTGCCGCTCCTCCCCTCGTGGTGGGGGTGGGGCAGGATGCCACCGCCGCTATCCTCTACTTGTCGGGGATGAGCGGGCGTTGCCCATGTTCGGCGCCGCCGTCCATGGAGATTGAGAGGAGGATGGGACCCAATGACATGTGGGCCTCACATTTTTATAAAAATGAAATTTATATTTTGCTTCCAATGCCACATCACCGCCATGTAGGACAAAGACCAAGTCAAACCAGCTATGTATGCGTCACGTCAACTGAAATCGTCGTCCAAACCACCTTGGGATCTTATTTGCACTGATTTTGCTAGTTGAGGGACGCGTCATATTTGGTTTTACGATCCAGGGATGAATTTCAGACTAGACGACAAATTAAGGGATCTAAAGTGAACTTATTCCTCAGAAGTCAGATGTGGTGTCACTCGTCGGATTGGAAAAATGGGCTTCTCAGCCCATTTCGATGTTTTGGGGGCTTTTGTCCCGGGAGTAGAGTTACAGCATTTTGGGCCGTCCGTGCCAACAGCAAGCGTGTGCTGGGGTAGAGATGGACCTGACCTGCCAGAGACTAAGCACACCTGGCCCAGCTTACTTCTGAAAGA
LOC_Os04g45490	TAAAGAGTAATTGGTTTAGAAGCTCTTCTGGAGCTCATAACTACTCATGGCATTGATTTCTGTAAATGGTTGTTGTAGTACTAGAATATCTTCCCAATAGAACAAACTGAGAAGCAAGGAATAGAGAGATGCGAGGATATGGCTCCATAAACATCTAAAATAGAATATACACATAGCTATGTAATTAAATGAATGTCCTCTCCATGTTTGTGTAGTTAGCAAAAAGAATATGGTCGATGGCATTTTGCTCTTACACAATAGACTTGTATGCATGTAACTTATCTTGTCGATGCAACATATTGAATAAATGCTTTACTTTCTTTTAATTGCTGTCTTTTACGCATCTCCTAATGAATATTGTGCAGAAGTGAGTTTCTATTTTGCAGTAAAGTTTGTCATTAAGTTAATTACATAAGATCTGGAAGCCAAGTCTCTTGGATTCCTTTTCACTTTTCATTTTGTAGATTCACAGTTTTCAACCAACAATCAAATTCACATGGCTTTCTTCAACCTCATCGCATTCACATTCACATCATTGCA

CCATCATTTTACTTTTGTATACAGCATCGGCAAGTGCCTTTGCAATGGCAT
TCCCGCCAC
ATATCGGTCATCATGGCACATATAGATGCAGGAAAGACAACAACTACAGA
ACGTGTTTTG
TACTACACAGGAAGGAACTACAAAATTGGTGAGGTCCACGAGGGAACAG
CTACGATGGAC
TGGATGGAGCAAGAACAAGAGAGAGGAATAACCATTACGTCTGCAGCAA
CTACTGCTTTC
TGGAACAAACATAGAATCAACATCATCGATACTCCTGGGCATGTCGACTT
CACTCTTGAG
GTTGAGCGTGCTCTCAGGGTGTTGGATGGAGCTATATGTCTATTTGACAGT
GTTGCTGGG
GTAGAACCACAATCTGAAACCGTGTGGCGCCAAGCTGATAAATATGGGGT
TCCGAGAATA
TGTTTTGTAAACAAAATGGACCGCCTTGGTGCTAACTTCTTTAGAACTAGA
GACATGATA
ATCGCAAATTTAGGTGCGAAGCCTTTGGTAATACAGTTGCCCATTGGTTCA
GAGGACAAT
TTCCAAGGAGTTATCGATCTAGTCAAGATGAAGGCTATTGTATGGACAGG
AGAGGAACTT
GGTGCAAAATTTTCTTATCAGGATATACCTGCTGATCTCCAAGAGATGGCT
GATGACTAC
AAAGCTCAGATGATGGAAACTATTGTAGAGTTGGATGATGATGTTATGGA
GACTTATCTA
GAAGGTGGCGAAATAGACGAGGAAACTGTCAAGAGATTGATCCGAAAGG
GAACAATATCT
GCCAGCTTTGTTCCTGTTCTATGTGGTTCGGCCTTTAAAAACAAGGGTGTC
CAACCACTG
CTTGATGCTGTTGTTGATTACTTGCCATCACCACTTGACCTACCACCAATG
AAGGGCACT
GATCCAGAGGACCCTGAACTGATCCTTGAAAGAAAGCCTAGTGATGATGA
GCCATTTTCT
GGGTTAGCTTTCAAGATCATGACTGACCCATTTGTGGGATCATTGACATTT
GTTCGCATA
TACTCTGGGAAGCTGGTTGCTGGTTCATATGTTCTCAATGCAAACAAAGAT
AAGAAGGAA
AGAATCGGAAGGCTTCTAGAGATGCATGCGAACAGTAAGGAAGATATAAC
AGTTGCTGTG
GCAGGTGACATAGTAGCTCTTGCTGGTCTCAAAGATACAATTACTGGTGA
AACTCTTAGT
GACCCAGATAAGCCTGTTGTGCTTGAACGCATGGAATTTCCTGATCCTGTC
ATTAAAGTT
GCTATTGAACCAAAGACCAAAGCTGATGCTGACAAAATGGCAACCGGACT
AATCAAGCTT
GCTCAAGAAGACCCATCATTCCACTTCTCTAGAGATGAGGAGACCAACCA
GACTGTTATT
GAAGGAATGGGAGAATTGCACCTTGACATCATTGTTGACAGATTAAAGAG
AGAATTCAGG
GTAAGTATTCTGATTAGCTTTATTGTTTCTGTACATGCCCCACTCCCCCACG
TAATTACT
CTACATTCATGCGCAAATCATATTGAACGGTTGTTATTGTGCTTATAAACC
ACACCATTT
CCTATGTGCTTAGATTTTCCCTTTTCCTCAGCTAAGCTAAGAGAAACTACC
TGATAAGAT
AGTAAAATGTTCTTTTGGTTAGCTTATTTATTTATGTTCTATGACTTAATAG
CTCGGTTC
TTCAATTATGCCACTTATCTCACCAAATAATCTTATTAAAGTCCTAGGAAT
AGAAGCATT

CCACTAGTTTTTTGTATTCATAATTGCACCCAAGAGTTTTAACTTGAGCTCT
TAAATATG
CAACTGCACAGGTTGAAGCAAATGTTGGAGCTCCGCAAGTCAACTACCGT
GAAAGTATTT
CCAAGATTTCAGAAGTACAATACGTTCACAAGAAACAGTCTGGTGGATCT
GGGCAATTTG
CGGACATTATTGTGCGATTCGAGCCTTTGGAGGCTGGAAGTGGGTATGAA
TTCAAGAGTG
AAATAAAGGGAGGGGCAGTTCCAAAAGAATATGTACCAGGCGTGATGAA
GGGAATTGAGG
AAAGCTTGCCCAATGGTGTCCTTGCTGGTTACCCTGTTGTGGACCTCCGAG
CGGTGCTGG
TTGATGGCTCATACCATGATGTTGATTCAAGTGTCCTGGCATTCCAAATTG
CAGCCAGAG
GTGCATTCCGTGAAGGATTGAGGAAGGCTGGTCCAAGACTCCTGGAACCT
ATAATGAAAG
TTGAGGTGATAACCCCTGAGGAGCATTTGGGTGATGTTATTGGTGATTTGA
ACTCCAGAA
GAGGACAGGTCAATAGCTTCGGGGATAAACCTGGTGGCCTGAAGGTAGCA
CAAAATATAT
CCTTTCCCTTTATTAGCCATACAGTTGAAATGTGGCCTGAGACTTACAGTG
GTATATTAT
TATTTGATGCAGGTTGTCGATGCTTTCGTGCCACTTGCGGAGATGTTCCAG
TACGTCAGC
ACGCTGAGAGGAATGACCAAGGGGCGCGCGTCTTACACGATGCAGCTCGC
CAAGTTCGAC
GTCGTCCCACAGCACATCCAGAACGAGCTCTCTGCGGCGAAACAGGAGGA
AGCTAGTACT
GCTTAATTTTTTTTCCAGATTTTGGTAGATATACTAAACGGCGACACCGGA
GCAGGTGGA
TAATAGGACATCCAAGTGTATGTAGCTTTGTAGCGTCAAGCCCCTCGGGG
AGATATTGTT
TGCCATTGTTTTGGACCTGATAATGGCCTTGTCAAGTTCTTTTCTTCTTTTT
GAAGACTG
CCTTGTTAAGGTTTTCAGAAGAGAAAGCGGTACCTATAAAATTTTTGTTTC
CTTGCTCTA
GCTGAGACGTCCTCTTATTTCTGCCCTTCTGCGTTTCGTTCTGCTTGTTACA
ATGGTGTG
GATAGATTGTGTGCAGTTATTTATATTTTCTTTGTTTTATATTATAAGTCGT
TTGACTTT
TTTCTTAATCAAAATTCTTTAGATTTGACCAAGTTTATAGAAAAAAATTTA
GCAAATATA
TACAACACTAAATTAGTTTCATTAAATTTAGCATAAAATGTATTTTGATAT
TATATTTGT
TTTGTATTGAAAATATCAGTATATTTTTCTATAAACTTAATCAAACTTAAA
AAATTTAAC
TAAGAAAAAAAATTCAAACAACTTACAAAATGAACGGAGGGAGCATATA
TTATAACATAT
TATAAGTTATCTAAGCAATGGCATATAAAGCTGCTTTTAGAGAAATGTAAC
TAGTTAGAG
AGCAGCGATACGTAAAATTCTTAAACCTATATAGGCTGAGAAGAACACCT
CCTAAACTAA
GCCAAGAGAGCGCGTGTAGAACTCCATTGGTTGCTCTAAAAAATCAAACC
AAAGCCAAAT
TAAAATTTTACAATTTTAGTTTTAAAGTTAATTTTAAGATATTTTAACGTAG
TTTTTTTT
CCAAC

LOC_Os04g45480	ATTTAATTTGCATTACTAAATATTATCGTAGCGTTAGCACGGGTCAATTACTAGTGCAAA TAAGATGAATTCCCATCCAATTCACTGTTTCAGTTTCACCATGCAAATTCATGTTGTTGC ATACGCCGATCCAAACATGAGGCACATTTACATTTGGTTCACCGCCGGAATAAACTGTAC CATTTTTTTTGCTAAATCTGTAGTATTATCAGATTTGGTAAAGACAAAAACATTTCGTTT GGGTTGATCTGAGTTCTTAGCAAATTTCGGTGTAGTTTTATATGTGGAAGGTTCCGGATT TCAGTATGGTAGCATGACAAAATTGTGATGGCAAGAGCTTATCTGCCTTCTCTTTATGTC AAGATTAAATTCAGCAACACTAAAATTTGATGATGTATTTTACTATTTTTACAACAAACC AAACATAGCTATGATTTTAATGTTCTCCCAATTGCATCTTCAGATCTTCTACTGGTAGTC GCCGTGCAATCTCGCGAGTTCCCGGCTTTCTCCCACCCTCCGGAATCTTCCAGAACCGTT GAGATCGCGCGACTCTTGAGAACCGTGACCTTTATATACTTTCTCCCCCTCGACGCCCCC GTCCAAACACGCACCAAGCCACCAACCGATTCCGATCCAACGATCTCGATCGCCCGCTCC GCTTCGCTTCCCGATCGATCAGCCCCGCCGCCGCGCGCGATGGCCGACGAGGCGAAGGCC AAGGGCAACGCCGCGTTCTCCGCCGGCCGCTTCGAGGAGGCGGCGGCGCACTTCACCGAC GCCATCGCGCTCGCGCCCGACAACCACGTCCTCTACTCCAACCGATCGGCGGCCTACGCC TCGCTCCACCGCTACCCGGAGGCGCTCGCCGACGCCGAGAGGACCGTCGCGCTCAGGCCC GACTGGGCCAAGGGTTGCTCCCGCCTCGGCGCCGCGCGCCTCGGCCTCGGCGACGCCGCG GGCGCCGTGGCGGCCTACGAGAAGGGGCTCGCCCTCGAGCCGTCGAACGGGGCACTCAAG GACGGCCTCGCCCATGCCCGCCAGGCGCGCCGCCCGGCGCCGGCGTCCGGCGCCGGCGCC ATCGGCAAGGTGTTCCAGGGCCCCGAGCTCTGGAGCAGGATGGCCGCCGACCCCACCACG CGCCCCTACCTCGACCAGCCGGATTTCATGCGGATGCTGCGCGACGTGCAGCGGAATCCC AGCAGCCTCAACAACTACCTCTCCGATCCCCGGATGGTCCAGGTGCTCAGCCTCATGCTT AACCTCAGGCTACCGAACAATGACGCGCCTCCGCGGCCGCCGGCGCAGTCGACTCCGCCG CCTCCGCCGCAGCAGCAGCAGCACCAGCCGGAGACGAAGGCGAGGGAACCTGAACCCGAG CCCGAGCCGATGGAGGTGACAGAGGAGGAGAAGGAGCGGAAGGAAAGGAAGGCCGCAGCG CAGGAGGAGAAAGAAGCGGGCAACGCCGCTTACAAGAAGGATTTTGAGACGGCGATCCAA CACTACACAAAAGCCATGGAGCTCGACGATGAGGACGTCTCCTACCTCACCAACCGAGCC GCCGTTTACCTTGAGATGGGAAAGGTGAGTGTTTCTTTTTGTACTGTGCTAATTGTTGTC CTCCATTGATTTCTTTCGAGAAAGTTCAGATTATTGTGTGCTTTGGTGAAGAATTAGTAC TTCCCCATGAAATCGATTAATGTGATTACTTACCTAGAGGTTTAGTGTGATTATACTTAC

CTGGTTTAGCATATACCAGCAGTTCTAGATTGCAGTGCTCTAATTTTAGAC
GACAGCCAA
TTGAATGTTGTGAGGAAAGGAATATGCGTTTGCTATGCTTGCACTCAGCAG
CACAGGTCG
AGATGCCTTGCCTCGTTTAAGAAAGTGATGCATTGTCTGTTTCAAATAAGA
AAGTGATGA
ATTATAGCCATGTAAATGATTGTGATGCAAGGGGCGGTTCAACTATTGGAT
CTTTTAGCG
TTATATATGTTGTAATGTTATGTCCTAAAAGCTAATATGCCAATATTTTCCC
CAATTTGC
AGTATGATGAGTGCATCAACGACTGTGATAAGGCTGTTGAGAGGGGAAGG
GAACTTCATG
CTGATTTCAAGATCATCTCTAGGGCACTGACCAGAAAGGGAACTGCTCTTG
CTAAGATTG
CTAAGTGCTTCAAAGATTATGATGTTGCCATTGAGACTTACCAAAAGGCTC
TGACTGAGC
ACAGAAACCCAGATACCCTGAAGAAACTGAATGATGCTGAAATAGCAAA
GAAGGAGCTAG
AGCAGCAAGAGTACTATGATCCAAAAATAGCTGATGAGGAGCGAGAGAA
AGGTATGTTGT
GGCAAATCTGTTATTATCACTGTGTTTCTTGTTGCAGGTGGTGATGAGGAT
GAAACTTTG
ATTTTGGTCCACTCATGTGCAGGTAATGAATTCTTCAAGCAGCAGAAATAC
CCAGAAGCA
GTGAAGCATTACAGTGAAGCTCTCAGGAGAAACCCCAAGGATCCAAGGGT
AAATTCTTTG
TTCTTACTCAATTGCTTAAGCACAGCAGCTGTTTAAAGCCTGTACTTCACT
GCAATGTTG
GTTTGCATGAGTACATCTGACGGAGAACAGATTGGATAATGTGAATTGGT
CAAACATCTG
GTTTTTTTTCTCCATAGTATTTGAAAATGCTAAGCATGCTTCCTCATTGTTG
TTCTATGA
ATCTATGGTGGGTTTGGTCTCAGTGGTTCTAAGCTTAGGTTACATGGACTG
TATTTCAGG
TATACAGCAACAGGGCCGCCTGTTATACCAAGCTTGGAGCAATGCCTGAA
GGTCTCAAAG
ATGCTGAAAAATGCATTGAACTAGATCCTACATTCTCCAAGGGATACACA
AGGAAAGGTG
CAATCCAATTTTTCATGAAAGAATATGACAAAGCACTGGAAACCTATCAG
GCCGGCCTAA
AGCATGACCCGAACAACCAGGAGTTGCTTGATGGTGTTAGGAGGTAAGCA
ATGCTTAGCG
AAGGAGATCGAGCACTGTCGCCTTTCCACTGTGCAGCGTCCTTACTTAACT
GTGCCCGTC
TTTTTTCAGATGTGTCCAGCGGATCAACAAGGCTAGCAGAGGGGAACTGA
GCCAGGAGGA
ATTGCAAGAAAGACAGGTAAGGCTCTCCTTCCAGTAATCCAGTTCCAACC
ACAGTGGTGC
CCTAGGTGCATCAACTGGGGCCCTGTTTAGATCTCTGACAAAATTTAACAC
CCTGTCACA
TTGAATGTTTGAATACATGTATAGAGTATTAAATATAGACGAAAAAATTA
ACTAATTACA
CAGATTGCGTGTGAATTGCGAAATGAATCTTTTAAACCTAATTGCGTTATG
ATTTGACAA
TGTGGTGCTACCGTAAACATTTGCTAATAACGGATTAATTAGGCTTAATAA
ATTCGTCCC
ACGGTTTATAGGCGGATTCTGTAATTTGTTTTTAGACTATGTTTAATACTTC
AAATGTGT

LOC_Os04g45470	GTCCGTATATCTGATGTGACATACCAAAACTTTACCCCTCTGGACCTAAACACAACTTGG TTTGGTAAGATAAGAGCAGTGATGTTAACTTAGTGTCCGCTTCTATTTTGCAGAACAAGG CTATGCAGGACCCAGAAATCCAGAACATTCTGAAGGATCCTATCATGCAGCAGGTAAACA TAATTAGATGACCTGTTGAATTGAAGTGTTGGTAGGTGATGGATTATTCTGGTAGCCTAG TGACCATCGTAATGCTGTTCTGCAGGTGTTGACCGACTTCCAGGAGAATCCGAAGGCTGC TCAGGCGCATCTCAAGAACCCCGGAGTGATGCAGAAGATCCAGAAGCTAGTAAGTGCAGG GATAGTCCAAATGAAGTAGACATGCTGAAAGGGCCAGGGATAGAGTAATCTGATCGGTGT GCCTGGAAACTTCTGAATGCAAAATTCGTCGTCGAATTTTTGTACAGTACAATCGTGTTT ATGCAAATTCCGATTGGTCATGGCTGAGCCGAATTTCCATACTGCTGATCTTTTACTTTT GACATGACCCTATAACCGATTTTAACCATTCGTCTTATTGATATATACATATTTATTTTA AATTTTTTAATAAGATGAATAGGGAAATGTTATAAAAAAATTAATGGTGTCCTATGCAGA AAAATGCCTAATCAGTCACCACGACCTAAAAGTCGCGATCAGAAAATAATTTGTGTAAAA CTTTCATATACGTCTTCAGTGTTGACCATATGTTTATTTTAAAAGTCGCGATCAGAAAAT AATTTGTGTAAAACTTTCATATACGTCTTCAGTGTTGACCATATGTTTATTTGTAAGTGT GTTTTTAGAGAAGGGTATTTTTACCCGGTCTCTACATCCAACCGGATATATGCATGCTTT TAAATAACCTAGTCTGAAATTCGCTCCTACGGAGATTTGAACTCAGGACCTTAGGGTGCT ACTCAAGTCACTGCAACCACTAACCTACATGCCCTTTCACATGTTTATTGTAACCGTAAT GGTTTATTAGCGATCAGAAAATAATTTGTCTAAAACTTTTATATATGTGTTTTTAAGTCA ATGTTGGAAAAAAAACTACGTTAAAATATCTTAAAATTAACTTTAAAACTAAAATTGTAA AATTTTAATTTGGCTTTGGTTTGATTT ATTTCATTTGGGCCTCTAAAGTAGAAAGGTCTCAACTTTTGACGTTCCACCGGTAAACAG CAGCAGCAGCAGCAAGGCACCTACTCCTAATCGTCTCATGTAACACGCGAGTTGATCGTG ATCATTCCTCCATCACGAAATATCATCTCACGCAAAAGCACGAAGTCCACCTGCCACGGT TCTGGAGCTCCGCTCCTTTTTCCCAAACACAGCAGAGTTTCAGACTTTTCAGCTTCGGTT TCCCTTGTTTCTGATCTTTCTTTCTCCTTCATCAAGGCACGGAACAATGCGATCTCCGCG CGGTGTTTTGCATTTTGCAACCGAAATCTCGCGAAAACCCAAACCCCTCCCGTTCACCGT GCCGGGACTGTACACGTTCTGCGGCCACAAGCCATGGGATTTCCAAATCACGAAGTACAA TAAATGGAAGAAACCACAAGTTTGTTCCACGCAATGCCCCCGCGTGATCCACGCGACCAC ACTTGCGAGGAGCGCGCGAGCGAAAGCTGCCGCGATGGCGAGACGACGTGCCGCGGGTCA CGTGTCCTCGTGGCACGCTAACTCGGGAGTGTACTTAAGGCCGCGTCTCAC

CAACTACTC
GGAGTAGTTTGGCTGCTTGACCTTTGTACTGTGCTTGTGCTGTGGTCGCGG
ACGACGATG
GCTGCGCGGTGCTGGGTATGGGGCTTCGTCGTCGCGCTCCTGGCTGTGGCG
GCGGCGGCG
GATGGGGAGGAGGAGGAGGGGAAGTGGGAGCCGCTGATTCGGATGCCGA
CGGAGGAAGGG
GACGACGCTGAGGCTGCTGCTCCCGCTCCTGCTCCTGCGGCGGCGGATTAC
GGGGGGACG
AGGTGGGCGGTGCTCGTCGCCGGCTCCTCCGGCTACGGGAACTACCGGCA
CCAGGTGAGC
TCGCGGTGATGGAGTTGGGTGTGTTAGGGTCTAGGGGCGGGGGCGGCGGT
AGAGATTAGA
GACTAGTAGAAAGGAAAGGAACTGCCTTTGGGGGAAAATTGAAGCTTTTG
GCAAATTCCC
CTTTCTTGTGGGCGTTTTTGCAGCTTGTTTGGATTCGCTTCCCGATTCGGAG
GGTGTTTA
CTGTTAACATTCTTGCTGCAATGCTGCTTCCAATTTGTGGTAGATTTTACTA
ATTCCTAA
GGCAACCCTTGCTCCAAGGCTGTATGTTCTTCTACCAAAGCTCCTGCTATC
TGACTACTG
AAATCGGCCTATGGACATGTTCTATTTGTTCAGAGGCATGGCTATCCAATC
CGTATTGCT
ACTTGCTTCGCTTAAATTTTGAAAGCTCCATACAGTTATGTTTTCGATAAG
AGAGGAACA
ATCTCTAAGTCAATAATTAGGTGGGAGGCATATTGGTTGGCAACATCAGTT
TTGATCGGG
CGAAGATTTTTTAAATTGCTATAAATATGACGACTACGACACTACAAATAA
TTTATTGTC
CTTCCTGCATGATTTATGTTTGTTTATCTTCTAATATATTGACGTGCAATCC
TTCCGCGC
GTTCGAGAAAAAAAGATTTATGTTTGTTTAGTGTGTGATTGATCAAATACG
AACAGTTCT
TAAGATAACTGAACAATTATATGAGTGCTTGTTTAATGGACGTTGGTGTTT
CCTGGTGCA
TTGCAGGCCGATGTGTGCCATGCGTACCAGATTCTGCAGAAGGGAGGAGT
GAAGGAGGAG
AACATTGTGGTGTTTATGTATGATGACATTGCCCATAACATTCTCAATCCA
AGGCCTGGG
ACCATCATCAACCATCCTAAAGGTGGAGATGTTTATGCTGGTGTTCCAAAG
GTAGCATTT
TCCTTTAATTGTAGTTTCAAGAAAAAGAGGGTGCAAGTAGTTCTAGTTTCA
GTCTTCAAA
TTCGATTTCTTTTTCTAGTCATGTTGTATGGATCTCTAATGCTAGTTTTGTA
TATAATTT
GAATGTGTCTGTATTCATGCAAGGCTGATCAACTTAACATGCATTACAGGA
CTACACTGG
TCACCAGGTCACCACTGAGAACTTCTTTGCTGTTCTCTTGGGCAATAAAAC
CGCAGTTAC
TGGAGGGAGTGGGAAGGTTATAGACAGCAAACCAGAGGATCACATCTTCA
TCTATTACTC
AGATCATGGGGGTCCTGGAGTTCTTGGTGCATACTCCAGCATATGATATTT
TTCTTAGTT
AATTTTGTATGGAACTTAGCTTACAATTAACATATTTTAGGTATGCCTAAC
CTGCCGTAT
CTTTATGCTGGTGATTTTATTAAAGTGTTACAAAAGAAACATGCTTCCAAC
AGCTACTCG
AAAATGGTATAAGATCATTTTCCCCTTTTCCCTTCCATTTTTTATAATCCAA

AATGATGT
ACTTAGACTGAAATTAAACCCTAATTTCTGCAGGTTATATATGTCGAAGCA
TGTGAAAGT
GGCAGTATCTTCGAGGGCTTAATGCCAGAAAATCTTAATATTTATGTCACA
ACAGCATCC
AATGCAGTTGAGAATAGTTGGGGAACATACTGCCCTGGGGAGGAACCATC
ACCTCCTCCT
GAATATATTACATGTCTAGGTGACATGTACAGTGTTGCTTGGATGGAGGAC
AGGTAAGCT
TCATTTAGTATCTCTTAGGTGCGACATCTTGTGGATTTCTGTTCTTCTATAT
AATCTTTT
TCGGTGTGTCACAGGTATGCCTTAAGGGCTTATGTAATTTACTTTTCAGTTT
GCAATACT
TACAATAGTTTATACCAGCTACAGTCATGACTGAACATTGTTGGTCCACTC
TATGATGTA
CATGCTCAGTTTATCTTTTTACGCCCTGGAAATTTGTTTCTGCGTTCTGATT
TGGGACCA
CTCTCAAAAGTTGTCTATTTGTCTTTCAACATCTGTACAAATACACAGTAC
ATGTTTTAA
TATCTGTGTGGTAAGAAGTACTGCCACAAGGCATAATGTATCAACATGTCT
TCCTGTAAT
ATTAGTCACCTGACTAAACATAAGTAATCCACAAGTCACTGTTCTTCATAA
TGGCAACAA
AGAAAAGCCTGTAAAATATAGGTGAAGAAAAGTATCAATTGATAGAACA
ACCAGGTGACA
AATAACAAGTGCATGTATTGCTTTGCTGCTTTACACATACTTTTACACTGT
AGCAATTAT
GCTGTGATGCTTGATTCTGAAAGTTTTTGTCCTTTTGACAGTGAGACTCAT
AATCTAAAG
AAGGAAACTATCGAGGATCAGTATGAGCTGGTAAGTCTTGAGTGCACTTT
CTCGATATTG
ATTCATTGTTTTCTACTCACTGCTTTCTTGGGCAGGTTAAAAAAAGAACAT
CAAATGCAA
ATAAGTTAAATGAGGGCTCTCATGTCATGGAATATGGTGACAAGACATTC
AAGGATGAGA
AGCTCTTCCTCTATCAAGGTTTTAATCCTGCAAATGGCAACATCACAAATG
AATTGATTT
GGCCAGTACCAAAGGCTACAGTCAATCAAAGAGATGCCGATCTTCTTTTC
ATGTGGAAGA
GGGTAAATTCTAGCTGTCTTCAGTTTATTCATTTCCGTCCTTCCTTTCTTTA
GAAGTCCA
AATTGTATGTATATAATACTGAGTTGTCGTAAACTAGGGTTCTTCCACTAC
TATAAGTTG
GTCGGCGAATATTAATATAAAACTAGCATGTTAACTGAAACATGATCACA
CACAAGTTAC
ATACTTTGGATGAAACATATAGATTATGATTTTTATTTTCCTTTTTCTTGCT
GCCGAATA
GAATGGACAGAGGCATGGCCCCTTGATGCATTGACATATATTGCATTCTGA
CATTATGTT
GTTAGGATAATGGTGTTGAATTTGCTAGAGTAGTTTCTTTCATGTTGATTTG
GTTGACAT
CGCCTGTAAGAATTGGCATCATTATCTCTGCATCATTATCTCTTTTATTGTT
AACTGGGG
CTCTCTTGTTTAGTTCTGATCTTTATTTTCTACTAGAACAAATGCCCGTGTG
TTGCAACG
GGTGAAGGCTATTTTAATCTTATTATTGTATACGGTTTAACTAAGGTGAAA
TTGACTGTG
GGAATTTGATTGGATATTTTCTTTTTTAAGAAAATCATGAGCTGCAATTAG

GAGTCTGCA
TGCAAGATTTTTTAAAGAGATTTCTTATACGACTCCTTATGTATTTCTAAAA
GCAAACGA
ACTTAAAATCCAACTCAAACACGGATTTGTATTTCCAAAAGCGAACGGAC
TTTGAAAACC
GACTCAAACACGGATGACGTACTAAAATACCGGCAAAAACATCTTCAATT
TTTATAATAG
TAGAGATTGCTCATGACCTTGGATATTTATTCTATGGACACTTCACCCTTTC
TTGCCTCT
ACTTGTGTATGGAAATAATGGAATACATTTCATATTTGAAAGGGTATAGTG
TGATATGCT
TCGTATTTAAAAAGGCCATCACATGGTGTACGTGTGACAGCTGATAGCAT
GATACTATGT
GAGCTCCATTTCCGAGAACAGGACTAAGTTAAGATGTCTGAATTTTTAAGC
CAGAAATAT
TTTCAGAGCATATGACAGTTTCATGCTAATGTCTGTATTTACTTTTTCAGTA
TGAGCAGT
TGAATGGGGTGTCTGAAGACAAGCTGAGGGCTCTTAGGGAGATAGAAGAC
ACCATAGCAC
ACAGGAAGCATCTTGACAGCAGTATCGATTTCATCGGGAAGCTTGTGTTTG
GTTTTGAAA
ATGGGCCTTTAGCCCTTGAGGCTGCAAGAAGCTCTGGTCAACCATTAGTCG
ACAACTGGG
ATTGTTTGAAGAAGATGGTGAGTACAAGGCTCTTTTTTTTAACTCTGATAT
AAGCCATCT
ATAAATATTTCTTGAGTTATTGTCTTCTTGTTGCCTGTTGGCCTTAATTGCA
CAATGCTA
TCCTTGGTTCCTTGCCAATTGTCAGGTGCGAATTTTTGAATCTCAATGTGG
ATCACTCAC
TCAGTATGGCATGAAATACATGAGAGCATTTGCAAACATATGCAACAACG
GTGTCTCTGA
GGCCAAAATGATGGAAGCAAGTATCAACGCTTGCGGCCGTTACAACTCGG
CGAGATGGAG
CCCAATGACTGAAGGAGGACACAGTGCTTGATCCCCGGCTATATACAGGT
AAAATTGAAT
ACCAAAGGCCCCGGAGGTTTGCAGACCTCATTTGTGTGAACTTCTGTATAT
ACCATGGAG
ACAACAATATAGCTAGCGACCCAACAATATGGCCCCTTGTGTGATGGCTA
ATCAGTAGGT
GTCATATGCTTGTTGTACCCAACAATATGGCTTGTGTGTGATGTGGCCCTC
TGTAAATTA
CTCGATTAGCAATCGATCCAAGAATGATTATGACCAAATATGTGCATGAT
GTTGTCTTCA
GACTTGATAATAAACAAGTAGGCCAATTATTAGTGGACGAAAAAACGATG
ACCCAAAGTA
AGGTAGTGCTGCTCCGATAGCCCTACCAAGATCTCGCTTACATATCTATTG
CTGGAAAAA
TGATATTACAAGCGATATGAAGCTCTTGCCAACGTCATGCAAGCGAAGCT
CTCTTTACAA
GTCATTCTCAGTTTCTCACACAGAATGCGAGGACTGAAACTGAAACTCCTT
CCAGACATC
AAACGAAGCTAGCCTGCAGTTTAGAGTGATACATCATGTTTATATCAGGTT
AGTGAAGCT
TGCATATGGAATGAAAGATAACGTGCCAGAGGAGCAGATACAGAGATTCT
CCTACTAATT
TATGATCTAGTAGTGGTCTCTCCAAAGACAACAAAACAGAGAGGATTTCC
AAGTCAAGGG
GCTTACATGATTTTCTAAGGAAAGAGATGAACACCACGAGGATCGCCAGC

	CCTGCCAAGATGCCTACGATTATTGCCACAGTCCGCCCGACGTCATCTTGCGAATAATCTGCAGCACAAAGCCTACTATTATCACCAGAATCCAAAGGAAATCAGACAGAGATTAGAGTAGCACGGCACATATGTACAGGACTACAAAAAGCAGAGTCACTTGCAGAGTGCCCCTTTCCGTTTTGTTTATCGCGAATATATTCATACACGTTCGCTACCACGTTAGTTTGTGTGCAGCAAAGGACCCCTTCATGTCAACTTTAGGGTTGCGAGCTTGTTGGCTGGAATACAGAATTACAAAGGTATCATCCGTACCAGCCACGGACGCAACCGCACCCAAGTGAAAAGGGCCGCAGCCTGCAGGAGCCGCGATGAAATCAGCTGTATGCCTGTGTGGAAATGGCAGTTGCTTATTGGCGAAGCAAAAATACTCAAAATCCTGTTCCCATCCCCGTATCCCGGATTCCCAGTCTAGACTGAAGAGGCAGAGCATTCAGTCAAACAGCCTATTGCTCGAGTGCACTCGTCGAAACGGCATTGGTTATCTTCAGACCATATGCCACAGATTATGATAGGTTGCATCCTGTCACTGCCCATACCGTCTTACCCCAATCTGGTACAGCACAAGGATAACTGAATGTACCGAAGCAATCATTCAGAGCATAGTTATCGGGAGGGACCAAAGCCAGTAGACCACTGACCAGAGTCACCCGACCAACCGCTACGAGGTCCAACCGGTGCAAGGAGGATGCATCGCCGTGTCGAGTCGACGTATAAATTCCAGAGACGCCGCGCGCGCGCACCACCACAGACCATCGAGGAAGCGTTAGCTTAACTAAAAAGCTCGCATTTATTCTGGGTGCTCCCCACCACCTTCTCTTCGCTCTCGCGACAGGGACTCGGAAAAAAGCATTCGCCGCTTTG
LOC_Os04g45460	CTTCGACCCAGCCAAGCCTGGGCGTTGCGTAGCGCGCTTCAAGCCAGCAGCTAGCTAGCGCGAGAGCCGTCAGTTCAGTCATGCATCAGCTCGACCGGCTTTGCTACGCCGTCGTCGTGCTCGTCGTGATCGGCGCGACGGTCGCCGAGGCGGCCACCGGCACGTTCATCTACGCCGGGTGCTCGCCGTCCAAGTACCAGCCGGGCACGCCGTTCGAGGGCAACCTCAACTCGCTCCTCGCCTCCATCGCCAACGCGGCGCCCAACGGCGGGTACAACAGCTTCACGGCCGGGAGCAACGGCACCGGGGACGGCGCGGCGGCGTACGGGCTGTACCAGTGCCGCGGCGACCTGGGCAACGCCGACTGCGCGGCGTGCGTGCGGGACGCCGTGGGGCAGCTCAACGAGGTGTGCGCCGCCGCGTACGCGGCGTCGCTGCAGCTGGAGGGGTGCTACGTGCGCTACGACAGCAGCAACTTCGTCGGCCAGCCCGACAACGCCATGGTGTACCGCAAGTGCAGCACCAGCACCAGCGGCGACGGCGACTTCCTCAAGAACCGGGACGCCGTCCTCGCCGCGCTCCAGGGCGGCCTCGCCAACGGGTACAAGGTGAGCAGCTCCGGCAACGTGCAGGGTGTGTCGCAGTGCCTCGGCGACCTCGCCGCCGGCGACTGCACGACGTGCCTTGCGCAGGCGGTCGGCCAGCTCAAGGGCACCTGCGGCACGTCGCTGGCCGCCGACGTGTACCTCGCGCAGTGCTACGTCAGGTACTGGGCCAACG

GCTTCTACTTCCGCCCGTCACAAGGTACGCCCAATAGTGCAACACGAATAAGAATTTTTC
AGCCTTGCGTTTCAGAGAGCAAGAACCCGAACCAATTCTCCAGCAACCTATCCTGATCAA
GTGATCAAGCTTACCATTTGCAAGTTGCCAATTGTCCTAAGCTATCTCGGTTAGATGGTG
TAAGTTTAACCGTGTAAAGTACAGTATGGAGAGTCCTTAGCTGAGCAAGTTGTGACGGGA
ATTGATTCCGATGGGCAATGGGTAAACTTACCAAAGTTGATAGAGAGATAAGGTAGAATG
GGTAAACTACATATAGTTAAATTTATTTAGCCACCTATGGCTATCGATCAAATTGCTCAG
CTTGCTCAGCTAAGGACGAGAGCAATCACTTGATTGAAGGCAGCCACGCTACTGACAAAC
CAAAATTTAGCTTGCTGGCTATAGACTACTGCCTGTAAGCCAACAACCCCCCATAACTTA
CAGACTTATATCTAAGTTACAGTGTCAGTGTCATATATGCAGGAAGCTCCAACGACACAG
TTCAGTCATTTTCCTGCAATAAAGAAAAGAGTAAATTTCACAAAACTACACGTTTTATAG
TCTAAGTTGCAGAAAACCACACACATTTTGACACTTGGCACTTAAGTACATATATTTTGG
TAGTGTAGTTTTACAAAACCACACTATCAATGAATGGATTCACCCGCGAGATGACGTGAC
TGTTTCACCTAAGACGAGGACATGTCATCTTCGAGTGTGGAACAACCACGTCCTCATCCT
AGATGGAACAACCACATCATCTCACAGGTGAATTCATTCATCGATAGTGTGATTTTGTGA
AATTAAACTACTAAAATATATGTACTTAAGTGCCAAGTGTCAAAATATGTGTGGTTTTCT
GTAACTTGAACCATAAAACACATGTTGTTTTGTAAAATTTACTCTAATAAAATTTGCACG
GCTAATCACCTCGGCTTTTCTGGTGGACTGCATGCAAATGACAGTATCATTAGACTCACT
TTGGTGGGACATCAGCATTAATCAGAATACCAGATGATCAAGTGGCAGTAACAAAAGGGG
CTATTCTTTCCGATCCTAGAAAAAGGAAGAGTAATAAAGATGATTCTAGTACTGAAAAGT
TGATGTCAAGGGACAGTAACAAACTAACAATGCAAATTATTCTCTCTGATCCAACAGAGC
TGACTTGTTAAATATCCTGAGATTCGTATTTCCTTCAACGTCTCTTGTGCAGGGATGCAA
AGAGCTTCCTACAGGACAGTGCTTAAACCGTAAGCTAGATAATTGTCCACTGCTTTTCTT
GTACTGGATATACCGATTAGAAAGGGGTTCTCTGACCCAGTGATGGCCACTAAACAAAAT
TCCAATCAATCTAAGCATATCAAATTTATTTAGGACGGCGAGCGGCGGCTGATAAGCTTA
GCTTAGCGCCTTAGCCACTCAACCACTTCCCATGCTGAATATATTCGCGATAAACAAAAC
GGAAAGGGGCACTCTGCAAGTGACTCTGCTTTTTGTAGTCCTGTACATATGTGCCGTGCT
ACTCTAATCTCTGTCTGATTTCCTTTGGATTCTGGTGATAATAGTAGGCTTTGTGCTGCA
GATTATTCGCAAGATGACGTCGGGCGGACTGTGGCAATAATCGTAGGCATCTTGGCAGGG
CTGGCGATCCTCGTGGTGTTCATCTCTTTCCTTAGAAAATCATGTAAGCCCCTTGACTTG

```
GAAATCCTCTCTGTTTTGTTGTCTTTGGAGAGACCACTACTAGATCATAAATTAGTAGGA
GAATCTCTGTATCTGCTCCTCTGGCACGTTATCTTTCATTCCATATGCAAGCTTCACTAA
CCTGATATAAACATGATGTATCACTCTAAACTGCAGGCTAGCTTCGTTTGATGTCTGGAA
GGAGTTTCAGTTTCAGTCCTCGCATTCTGTGTGAGAAACTGAGAATGACTTGTAAAGAGA
GCTTCGCTTGCATGACGTTGGCAAGAGCTTCATATCGCTTGTAATATCATTTTTCCAGCA
ATAGATATGTAAGCGAGATCTTGGTAGGGCTATCGGAGCAGCACTACCTTACTTTGGGTC
ATCGTTTTTTCGTCCACTAATAATTGGCCTACTTGTTTATTATCAAGTCTGAAGACAACA
TCATGCACATATTTGGTCATAATCATTCTTGGATCGATTGCTAATCGAGTAATTTACAGA
GGGCCACATCACACACAAGCCATATTGTTGGGTACAACAAGCATATGACACCTACTGATT
AGCCATCACACAAGGGGCCATATTGTTGGGTCGCTAGCTATATTGTTGTCTCCATGGTAT
ATACAGAAGTTCACACAAATGAGGTCTGCAAACCTCCGGGGCCTTTGGTATTCAATTTTA
CCTGTATATAGCCGGGGATCAAGCACTGTGTCCTCCTTCAGTCATTGGGCTCCATCTCGC
CGAGTTGTAACGGCCGCAAGCGTTGATACTTGCTTCCATCATTTTGGCCTCAGAGACACC
GTTGTTGCATATGTTTGCAAATGCTCTCATGTATTTCATGCCATACTGAGTGAGTGATCC
ACATTGAGATTCAAAAATTCGCACCTGACAATTGGCAAGGAACCAAGGATAGCATTGTGC
AATTAAGGCCAACAGGCAACAAGAAGACAATAACTCAAGAAATATTTATAGATGGCTTAT
ATCAGAGTTAAAAAAAAGAGCCTTGTACTCACCATCTTCTTCAAACAATCCCAGTTGTCG
ACTAATGGTTGACCAGAGCTTCTTGCAGCCTCAAGGGCTAAAGGCCCATTTTCAAAACCA
AACACAAGCTTCCCGATGAAATCGATACTGCTGTCAAGATGCTTCCTGTGTGCTATGGTG
TCTTCTATCTCCCTAAGAGCCCTCAGCTTGTCTTCAGACACCCCATTCAACTGCTCATAC
TGAAAAAGTAAATACAGACATTAGCATGAAACTGTCATATGCTCTGAAAATATTTCTGGC
TTAAAAATTCAGACATCTTAACTTAGTCCTGTTCTCGGAAATGGAGCTCACATAGTATCA
TGCTATCAGCTGTCACACGTACACCATGTGATGGCCTTTTTAAATACGAAGCATATCACA
CTATACCCTTTCAAATATGAAATGTATTCCATTATTTCCATACACAAGTAGAGGCAAGAA
AGGGTGAAGTGTCCATAGAATAAATATCCAAGGTCATGAGCAATCTCTACTATTATAAAA
ATTGAAGATGTTTTTGCCGGTATTTTAGTACGTCATCCGTGTTTGAGTCGGTTTTCAAAG
TCCGTTCGCTTTTGGAAATACAAATCCGTGTTTGAGTTGGATTTTAAGTTCGTTTGCTTT
TAGAAATACATAAGGAGTCGTATAAGAAATCTCTTTAAAAAATCTTGCATGCAGACTCCT
AATTGCAGCTCATGATTTTCTTAAAAAAGAAAATATCCAATCAAATTCCCACAGTCAATT
```

	TCACCTTAGTTAA
LOC_Os04g45520	TAGGGAATATTTTCTTTTGGGGACCATCTCATCCCACCCATCCTCAAACAAAATACCCTTGCAATCCAACACGACCCATCCCATCCATCGAACCAAGTATGTATCGCGTGATAACATAATGTAGAGATGCAGCCTAACCCACATGTTAGTGAGAGAGAGAGGGGGACGGATTGGATTAGTTCACCACTACTTAAGGTGTGTTTGTAAGGAATATTTTCTTTTGGGGACCATCTCATCTCACCCACCCTCAAACACCCTGACGGTCCAGCACGATCCACCCCATCTATCGAACCATGTGAGTATTGTGTGATGACATAATACGGAGTTGCAGCCAGCCGTGAGGCACATATGCTTTGTACACGACGAGACGATGGTTAGGGGATCAAACGAGACGGACGGCGAGAAACGCATTGTCACTGTCACATCGTCGTCCGGGTCTATCTCAAAGTGGTCCCGGTCCACGTCGGCATCCGATCCACAAGTCAGCGAGGGGCGCGTGGCCGGCACACGCCGAGCGATCTGGGCCCTCCATTTCGCCAACCCTAGCGGCCTCACCCGAGTCACCCCTATAAATGCCGGTGGCCGAGGCGTGTGGCCGCCACACTGCGATCCCCATCCATCTTCTCCGCACACGCTACACTCCTCTCGATCGATCTTCTCCTCCCGTCTCGATCAGCTGATTCGCCCAGCCTCCGGTAGCTTGCTTGTGCTTCGTATCCATGGATCCTCACCACAGCTCCCACGTCCATGCCGGAGCCGCCACTCCTAACTGCACGTGCAGGTTGCCCAAGGAGGAGGATCTAGCCGCCGCCGTCGTCGTCGTCGACGTCGAGGCGGCTCCTCACTACACGTCGTTGAAGGCGGCCGAGGCCGACGTCGTCCCGAGAGGAGTCGGCGGCGGCGGCGGCGGTGGCGAAGGGGACGGTGGCGAGGAGTGAACTACATGGCCCGCGCGCAGTGGCTCCGCGCCGCGGTCCTCGGCGCCAACGACGGGCTCGTCTCCGTGGCGTCCCTCATGATCGGCATCGGCGCCGTGAACGAAAACAACAAGGCCATGCTCGTGTCCGGGCTGGCCGGGCTCGTGGCCGGCGCGTGCAGCATGGCCATCGGCGAGTTCGTGTCCGTGTACGCGCAGTACGACATCGAGGTGACGCAGATCGAGCGCGACGGCGACATCGACGGCGCCGATGCCGCGGCCGCGAGGGAGAAACTGCCGAGCCCGACGCAGGCCGCGTTCGCGTCGGCGCTGGCGTTCGCGATCGGCGGCCTCCTGCCGTTGCTGACGAGCGGGTTCATCAAGCCGTGGGGCCCCAGGGTCGGCGTCGTGTGCGCCGCGAGCAGCGTCGGGCTGGCCGGCTTCGGCGCCGCGGGCGGGTACCTCGGCGGCGCCAACATGGTGAGGTCCGGCACTAGGGTCCTCCTCGGCGGCTGGCTCGCCATGCTCATCACCTACGCCGTGCTCCGGCTGTTCGCGACGATCTTTCACGGCATGAACATCTCGTCGTCGGCGTGATGATGCGTGTGACGCAGAAATCGATCTATCTGGAGATGAAAAAGAAAATAATCAAGGGGTTGTGCTAGCTAGCTGTTTAATTAACTACTCTTGGTTTGCCAGTAGTACTGTACCTAGTGGCTGATTATTAACTAGCGAATCCCTGGATACCGAACTGGGAATAATCA

LOC_Os04g45540	GGGGGTGTAAGTGTAATGGTTGTTGTGAATTTGTGATGAGGCAAGACCAG GCAAATGGTC TCAAATGAGAATATTGGTGTTAATTATTGTATACCGATTCCCTTCTCACTTT GCCTCCCG TTTGTTTTTGATTTGTCTGTGTCTCTGTGACAGTTCAGAAACTAGCTGATGC TCTTGGTT GCGAATATCCAGGTGTGCCACTGGCACGTTGAATTGTTTTTGGTTGATTGG CCAGCAAAT GCCAAAGAATGCGTTGAAGGTGCAGCCGGGATTGCCAATGGTCAAAGGAA CTTGTACCAA TGGAAAAGAACAGTTTAGAACACAGAAAACAAGAACAGTGCACAGATTG TCATATGCCCC CTCAGAGAGTCCAGTAAAAGGCCCAGTTCATCTTGGGCCTATTTATGCCAC GAAAGCCCA TCTTCATAAGGAATATAGGCCTTCTTCTTCCACACTAAAGTTCAGATTTCTT TATATAGA ATTAAAGTTACAGACTTGCGAAACTTACACTACAGAAAATTACTTTTGAGG TGTAAAAAA ACAAGCACTCGACCGATCCAACGAAAATATCTGGTGATCACAGATCTGAT GATCTGTTAG CCTCGAGATATTCTGCACCTGATTTCTCGATCTTATAGCATAGAAAATTAT CAATCCAGT GACCCGGTTGCACAAATGGAAGGTAGAATAATATAAGAAAAAGGACG CCATTGCAAATCTCCGCCCATTAATGAGGGAGGACAGGGTTATCCCGGTG TCAGGCGAAA GTACGAGCCAAGTTTTGGTCAAAGTAGGATGGGTCCCCACCCAAAAGAAG GGTAGGTAGA AATGGTGTATGTGGTATCCCCTTGAGATATAAAAGGAGGACCATGCCTAC GAAGAGGGGG GACCAAAAAAAAGGAGGGACGGAAAAAAAGACTAGAGGAGACCTGAGC AAGTAGACTGGA GAAAAGAGGAGTGAGAGGGGCTCTAGAGTGGGAACTCTCTTGCTCTCCAA CTCTTTGTAA CTCCTTCATACACAGATCCACCAGAACACAGGAGTAGGGTATTACGCTTCT CAGCGGCCC GAACCTGTCTACATCGCCCGTGTCTTGCGCATTTACTCTCTCGCGAAACAT TCCCCAGAC CGATCGAGGGCGAAGAACCTCACTTAGCGCCCCCGGCCGAACCGGCAAAG GGGGGCCTGC GCGGTCTCCCGGTGAGGAGCCCCACGCTCCGTCATCTGGCGCGCCAGGTA GGGGGCGCGC GTGGGTTTTTCTAAGCCCTCGTCCTCTTTCGTGTGCGCCAAGCTTCTCCCGC TCAGCCCA ATGGCCGAACGGGCAGTGGCGCACAACCTCTCTCCGAGCGCTAGCGGTGA CGACAGCGAG CAGAACCCGCGTCGTCGAGCTCGTACTCCACCGCCCCCCTCTCGCCGAAGT CTTAGGCGG GAGGAGGCGCTCGAGGGGGTCGAAAGATCCGCGACTTCGCCGCACGCGG GCGATGGAGAA GGACGGCGAGACGGGGAGCGTCGTCTCCTCGTCTACGGTGATGGCAGCAC GCCCCAGGGC GCGCTCCAGGCTGCGGGCGCGCTCTTGCGTCACCCGCCCGTCGTCCCTGAT CCAGAGACT CCAACCCGCCGTTGGCTAGATGACGTGGCCAACTTGGTCATGACGGCGCA GCAGCGCCTA GGCGCCGGCGGGCGGTCCTCCGCCACCAAGGCCACTGGCGTTGCTACTGC CGGCTCCGTG TCGTCGAGGCGAAGGGCGCGGCGAGCGGCAGCCGTTACCCGCCGTTCATC

TGCCATTCCC
TCGTCAACGCCTCCCACCCAAGAAGATGTGCGTGAGGGGCCGGATGCCCG
CCTCAATATC
GAGCGCCGGCGCAATGATCGCCGTGCTGCCCACGCATCGGAGGGCGCTTC
CTCGTCCGGA
GCGCCACTTCGGTCCGGGCACGGGGGCCAGCCCCCCGTGTCCCCGGTTGG
CGGTGCCGGC
TGTCGAGCCTTTGTGGCGAGCCTTCGGAACGTCCGTTGGCCCCCAAGGTTC
CGGCCCACC
ATCACCGAAAAATACGACGAAAGCGTCAACCCCGCCGAGTTTCTCCAGAT
CTACACGACT
GGGATCGAGGCCGCCGGGGGTGACGACAGGGTCATGGCGAATTTCTTCCC
CATGGCCCTG
AAAGGACAGGCGCGGGATTGGCTAATGAACTTGCCACCCGCGTCGGTCCA
CTCTTGGGAG
GATTTGTGCCAGCAGTTCACCATGAACTTCCAGGGCACATATCCGCGCCCA
GGTGAAGAA
GCGGACTTACACGCGGTACAGCGAGGGGACGATGAGTCTCTTTGCTCGTA
CATTCAGCGG
TTTTGTCAAGTCCGCAACACCATACCGTGCATCCTTGCACACGCGGTGATC
TACGCGTTC
AGGGGGGGCGTGCGGCACAACCGCATGCTTGAGAAGATCGCCTCCAAGGA
GCCCCAGACC
ACTGCGGAGCTTTTTCAGCTCGCGGACCGTGTGGCTCGCAAGGAGGAGGC
ATGGACCTGG
AACTCCACCGGTTCCGGTGTGGCAGCTTCGGCCGCCCCCGGATCCGCCGCT
CGGTCAGGG
CGGCGGGACAGGAGGAGGAAGAAGAGGTCGGCCCATTCCGACGACGAGG
GTCACGTCCTC
GCCGTCGAGGGCGCTTCGCGGGCCACCCGAAAGGGGAGGCCTGCGAGCG
ACAAAAAGAAG
GAGGCCGGCGCTCCCAGCAGGGAGCGCCCGACCAGCAAGTGGTGCACCGT
GCACAACACC
TCCCTCCACGACCTCGCGGACTGCCGCGCGGTCAAAAGTTTGGCTGAGCG
GACGAGGAAG
TGGGAGGAGGAGAGGAGGCAGGAGCGCCGTGAGGGCAAGGCCCCGGCGG
CTCCCTCCAGC
AACCGGCGAAGTGAGGCCAAGCAGAAGGCCCCCGCCGAGAACATTGACG
ATGGCGACGAT
GACCTAGGGTTCCAGGAGCCCGGGGCCACCATCGCCACCGTCGATGGGGG
AGCACGCGCT
CACGCTTCTCGTCGGAGCCTCAAGGCCATGAAGCGGAGCTTCTGGCCGCG
GCCCCTACCC
ATGAGGCGACGCGGCGGGCGCGGTGGTCGGAGGTTGCCCTCACCTTCGAC
CAGACCGACC
ACCCGCCATGCATTGCCCGGGGAGGACAGATCGCATGGTGGTTTCCCCCA
CCGTTTGCAA
CGTGAAGTTGGGGCGTGTCCTCATAGATGGGGGTGCAGCCCTCAACATCC
TTTCCCCCGC
AGCCTTTGATGCCATCAAGGCCCCGGGGATGGTGCTCCGGCCGTCCCAGC
CGATCATCGG
TGTGACGCCGGGGCACACTTGGCCGCTAGGTCATATCGACCTCCCGGTCAC
CTTCGGTGG
TTCCGCCAACTTCCGCACGGAGCGGGTGAACTTCGATGTGGCGGACCTCA
GTCTGCCCTA
CAACGCGGTCCTAGGGAGGCCCGCGTTGGTGAAGTTCATGGTGGCGGTCC
ACTACGCCTA
TCTCCAGATGAAGATGCCGGGCCCCGGTGGCCCCATCACCGTCCATGGCG

ACCTCAAAGT
CGCGCTTGCTTGCATGGAGCAGCGCGGACCACCTCGCCGCGGCATCCA
AGCCCGCGGG
TGGCGACGAGAGGCTCAGCACCTCCGTCCCAGCCGCCCCGAGGCATCGGA
TGATCACGTG
CGACGAGGTCCCGATCAAGCAGGTTGCCCTGGGCGACGGCCCGTCCAAGA
CCACACGAAT
CGGTGGTCTTCTGGACGATAAATAGGAAGACGCGCTCGTCTCCTTCCTGCG
GGCAAACGC
CGACGTCTTCGCCTGGAGACCGGCGGACATGACCGGGGTCCCCAGGGAGG
TGATTGAGCA
CCGTCTCGCCGTACGGCCGGGCGCAAGGCCGGTCCGGCAGAAAGTGCGGT
GGCAGGCCCC
GGAACGTCAAGCCTTCATCCGCGAGGAGGTGGCGCGACTTCTGGAGGCCA
GATTCATCCA
AGAGGTCATCCATCCGGAGTGGCTGGCGAACCCGGTCGTCGTTCCCAAGG
CGAACGGCAA
GCTTCGGATGTGCATCGACTACACCGACCTTAACAAGGCATGTCCTAAGG
ACCCTTACCC
CCTGCCACGCATAGATCAGATTGTCGACTCCACCGCGGGGTGCGACCTTTT
GTGTTTTCT
AGATGCATACTCTGGTTATCATCAGATTCGCATGGCTAGGGAGGATGAGG
AAAAAACTGC
ATTCATTACCCCCATAGGAACTTATTGTTATACGACAATGCCTTTCGGGTT
AAAGAATGC
AGGTCCTACTTTTCAACGTACTACTCGAATTTCTTTGGGTAGCCAAATAGG
ACGTAATGT
TGAGGCCTATGTTGATGACTTGGTTGTGAAAACACGCAACCAGGAGACCT
TACTTTCGGA
TCTAGCGGAAACTTTCGAAAGTCTCCGCACCGCCCGCATAAAGCTGAACC
CCGACAAGTG
CGTGTTCGGCGTGCCTGCGGGCATGCTTCTCGGGTTCTTGGTCTCTGCCCG
AGGCATTGA
GGCCAACCCCGAGAAGATACGAGCTATAGAGCGGATGCGCCCCCCAGCAA
GCTTAGGGAT
GTGCAATGCGTCACTGGCTGTATGGCCGCCCTAAGTCGGTTCATATCGAGG
CTGGGAGAG
AAGGCGCTACCCCTATTTAAACTCGTCAAACGCTCCGGGCCGTTTATCTGG
ACGGAGGAG
GCTGAACGGGCCCTCACTCAGTTGAAGGCGTACCTCAGTTCTCCCCCAGTC
CTGGTCGCC
CCGGAGCCAGATGAGCCGTTGCTACTCTACCTGGCGGGAACCCCGCAGGT
GGTCAGCGCG
GCGTTGGTCATGGAGCGCGACGAGGACGACCCTCGCTTCGCGCATCCCCA
CCCTGTGTCG
ACTCGACCCGGGAGAGAGCAGGGAGGAGAGGCCCCCGAGCCGAACGGTG
GCCCAAGGCCC
CCGACGGCCGGAGCTGGCCCTCCGCCCGCTTGTCCGAAGGTGCCAGGCGC
CCCCGACCCC
CAGGATGGCCACGGGGCCACTGCGGGAAGGCCGCGCCTGTCGCCCTCTGA
CCCCGAGGTC
ATCGGCACAGAAGCCGAGTGCGCCCCGCGCGGCCTCTCGGAGCGTGACGA
GGACGATCCT
CGCTCTGCGCGTCCCCACCCTGTGTCGATTCGACCCGGGAGAGAGCAGGG
AGGAGAGGCC
CCCGAGCCGAACGGTGGCCTAAGGCCCCCGACGGCCGGAGCTGGCCCTCC
GCCCGCTTGT
CCGACGGTGCTAGGCGCCCCCGACCCCCAGGATGGCCCCGGGGCCACTGC

```
AGGAAGGCCG
CGCCTGTCGCCCTCCGACCCCGAGGTCGTCGGCACAGAAGCTGAGTGCGC
CCCGCGCGGC
CTCTCGGACGAAGAGCGCCCTGGAAATGCGGCCCCTAGTGAAGAAGATCG
GCCCCGCCGA
AAGGTGTAGCGGCCCGTCTACTTCGTTAGCGAGGCCCTTCGGGACGCCAA
GACCCGATAC
CCTCAGGCCTAGAAGATGCTTTACGCTATTCTGATGGCTTCGAGGAAGTTG
CGCCATTAT
TTCCAGGCGCATCAGGTCACGGTGGTTACGTCCTACCCCCTCGGCCAAATC
TTGCATAAC
CGAGAGGGTACTGGACGGGTGGTGAAATGGGCAATCGAACTCTCTGAGTT
CGATTTGCAC
TTTGAACCACGCCACGCAGTCAAGAGCCAGGCCCTCGCCGACTTTGTGGC
AGAGTGGACA
CCGGCTCCCGAGCCCGTCTCCATCCCCGAGGCCAGCTCGGGCCCTTCGCAG
CTGCCTCAC
ACCGCCTATTGGGTGATGCAGTTCGACGGCTCCCTGTCTCTTCAGGGTGCC
GGTGCGGGG
GTCACGTTGACCTCGCCGAGCGGAGATGTCCTCAAATACCTGGTCCGCCTC
GACTTTCGG
GCAACCAACAATATGGCAGAGTACGAGGGACTCCTTGCGGGACTCAGAGT
GGCGGCTGGA
CTGGGGATCCGCCGCCTCCTGGTGTTAGGCGACTCTCAGTTGGTCGTTAAC
CAGGTCTCC
AAGGAGTACCAGTGCTCTGACCCGCAGATGGATGCTTACGTGCGCCAGGT
GCGGCGTATG
GAGCGCCATTTTGACGGGATAGAGCTTCGGCACGTGCCCAGACGGGATAA
CGCGGTCGCC
GACGAACTCTCAAGGCTCGCTTCCTCGCGAGCCCAGACCCCACCGGGCGC
CTTTGAAGAA
AGGCTTGCCCAGCCGTCGGCGCGACCCGACCCCCTAGGGGAGACGGACGC
GCCTGAACGG
CCCCCGAGGCCCGTCAGAGTCCAGGCCTCGGGACCCGAGGGGAGCGCTCC
CAGCTCCCTT
AGATTGATTGCTTGGATCGCTGAGATCCAGGCGTACCTCGCAGATAAGAC
TCTACCCGAG
GACCGCGAAGGGAGTGAACGCGTCCAGCGCATCTCCAAGCGCTACGTGCT
GGTGGAAGGG
ACCCTCTATCGGCACGCGGCTAACGGAGTCCTCCTGAAGTGCATTCCTCGG
GAATAAGGC
GTCGAGCTTCTTGCCGATATCCATGAAGGCGAGTGCGGAGCCCACTCTGCC
TCACGCACC
TTGGTTGGCAAGGCCTTTCGCCAGGGTTTCTATTGGCCGACGACTGTCAAC
GATGCGGTC
GACCTGGTCCGGCGGTGCAGGGCGTGCCAGTTCCACGCCAAGCAAACCCA
TCAGCCGGCC
CAAGCCCTGCAGACCATACCTCTTTCATGGCCGTTTGCTGTCTGGGGGCTC
GATATCCTG
GGTCCGTTTAGACGGGCCCCGGGCGGGTTCGAGTATCTGTATGTCGCTATC
GACAAGTTC
ACTAAGTGGCCCGAGGCTTATCCGGTCACCAAGATCGATAAGCACTCCGC
ACTTAAATTC
ATCAGGGGCATCACTGCTCGGTTTGGGGTGCCTAACCGTATCATTACGGAT
AACGGCACC
CAGTTCACTAGTGAACTCTTCGGTGACTACTGCGAAGACATGGGCATCAA
GCTCTGCTTC
GCCTCCCCCGCCCACCCCAGAAGCAATGGCCAAGTGGAGCGCGCCAACGC
```

GGAAATCCTC
AAGGGCCTCAAAACCAAGACCTTCAACATCCTCAAGAAGCATGGCGACTC
GTGGATCGAG
GAGTTGCCAGCGGTGCTTTGGGCGAACCAAACCACGCCAAGCCGAGCAAC
CGGGGAAACG
CCTTTCTTCCTCGTCTACGGCGCGGAGGCGGTTCTCCCATCCGAGCTCACC
CTGAGGTCC
CCTCGGGCCACCATGTACTGCGAGGCTGATCAAGATCAGCTTCGCAGAGA
TGACCTCGAC
TACTTGGAAGAGCGAAGGCGGCGCGCGGCCCTCCGAGCCGCGCGCTACCA
GCAGAGCCTG
CGGCGCTACCATCAGCGCCACGTCCGGGCCCGATCACTCTGCATCGACGA
CCTCGTCCTA
CGCCGCGTCCAAACGCGTGCTGGACTGAGCAAGCTCTCACCAATGTGGGA
GGGTCCGTAT
CGAGTGATCGGCGTCCCCCGGCCGGGCTCCGTTCGGCTGGCCACGGGCGA
CGGCACAGAG
CTGCCTAACCCGTGGAACATCAAACACCTTCGTCGCTTCTACCCCTAGTGG
GGGTCGGGT
GTCAGGTCTCGGGCTCGGGCCAACCCCGCACCCCGCCCGGGCGTGCAGGG
TTGGCCGGGG
GCTACCACATGCATAAGCTTATCTTTTTCTCTTCCGTATTTCAATAAAAGC
ATTTTCGA
TTTCCTAAGGACTGTGTCTGTGCCATTGTTTCCTTTCGAAGAATCTTGGCTT
GAATTAGC
ACGCTCGTGCTCGATCCTGCCCTCAGTGGCCCGGGCCGGCTAAAATCGCCA
AACGGGGCC
CAGAACCGAGCCGTGCCCCGGGGCGTGGTGAACTCCGGGGGGGAATCGGC
AAAGGCCCAC
AATCGGGTCCCCATAACCCTCGGTCACCGCGCTCCCGGCCTGGCCAGTCTC
GACACGTGG
ATCTCCCCAGGCACTAGCCCCGACCCGCGCCGTTTGAGGCTGGGACCTGG
GCACGAGTCC
TCGTTCAAGCCAAGACACAGAAGGACCGCGGCACAGACCATCCTCGTCTC
ATATACACAA
CGAACAGAGCTGACACAAAAATTTCCTCTTTATTTGATTGCAGATTGAATT
AGTTTATAC
AAGAAGGGGGCCCGAAGGCCTCAGAAGAAGAAAGAAAAAGCTATTTAA
GGCTTTCTTCAAGACCGCTACATTACTTGAGTAGCGCGCGAATTGGGACGA
GCCGGTTCG
TTTCGCCTGGAGATGTGATGATGGTGCATCGGTCGGCGAGCGTAGTTAAC
GCGCGCATGA
TATGTCCCATCTCGACTGCCACGGCCGCACATCCGCCTCGTGCGCCGCAAA
CGGGCGAGC
AGGATTAATGGAGGAGTGGGCGCGAGAAACGAGGGGGCGAGATAGTGGG
CGCGGGAAGCG
AGGAGCCGGGCATAGCGTTTGGCAAGGGTATAAAGACGCCGAGGAAGAG
ATTCGTTTCCT
LOC_Os04g45550
TCCTCTCGCCATTATTCCCCTTGCCTTCGCCGCTTGCGCTCTAACTTCTCGT
TTCCTGTG
CCCTACCCTCGCCGCACTCACTCGCTCTCAAACCCCTCCTCCGCCGGCGTT
ATGGCACGG
GGCTCCGCACCGCTCGACGGTAGCGTGCTGCCGCCTTCTCGCATCGTGAGC
GAGAGGCAG
GCCGGGCTGCCGCGCCGATTCATGCCGGAATCTGCCACCGGCCGGGAGGT
GGTTGCGCTG
GGTGAGGGACGCCCGGCGCCACACTACCCGGGGCGGTCCGTATTCTTCCTT
CCCTTCGCA

ATGGCAGGGCTGGTCCCGTCATTCTCTTCTTTCTTCATGGATGTTCTGGAGCTTTATGAT
CTCCAGATGGCGCACCTCACCCCCAACGCAGTAATGACGTTGGCCATCTTCGCGCACCTG
TGCGAGATGTTCATCGGGGTGCGCCCATCTCTCCGGTTGTTCCGGTGGTTCTTCACCGTA
CAGCCGGTGTCGCCACCGTCGGTAGTCGGTGGCTGCTACTTCCAGCCGCGGGGGCCGGTG
CTGAACCGCTACATTCCCTGCGTCCTTCGCAAGAAGTGGGATGACTGGAAGAGTGACTGG
TTCTACACCCCTCTCGCCGACGAAGCGCGCCTCCGACTTCCGAGCCAGCCCCCGGCGCAG
GCCTCCAGCTGGCGGGCGCCGGTAGATCTGGGAGACGGCTACGACGCCGTTCTTGACCGC
CTGGCTGGTTTGCGGTCCCAGGGGCTCACGGGGGCCATGGTGTTCGGCGACTACCTTCGT
CGCCGGATCGCGCCGCTCCAGCGGCACGCCCGGGGCGCCTGGGAGTACACCGGGCCCGAG
GACATCATGAGGACCCACCAGTGGCGCAAATGGGACTGGGACCCTGAGGACTTCAGAATA
GTGGTCCAACGGGTGCTGAATCTCAGCTCCGTGGAGGCATCCTTCATCTCTCAAGGGGTC
CTTCCCCTCTCCTGCGATCCAGAGCGCGCCAACATCCTGACCATCATGCAGGAGGTCGGG
GCGTCGGGGGAGCGGGCCCCCCGAGGCCACGATGGCGCGGGCGGGAGCCGCAGGGGGGAA
CAATCTACCCCGGGAGGGGGTCGTGCTTCTGGGCCCCGTGACGGGGGCCCGGGGGGCAGC
CGCCCTACCGACGCCCGGGGGAAGAGGAAACAGGAGGATACCCCCTCCCCATCTCCTCCC
CGAGGGGGCGGGGCGGTGCGTGCCAGCAGCAGGCGCCCGGAGGGTGCTGTGCCGACGTCG
CGGCCCGAGGGGAAGCGAAAAAAGAAGCGGCTCCACAAGATGGGGGGGACCGAACCATAC
CGGGGAAACCTCATCTCCCCCCCAAGGTGGTCGTTTAGCCGACCTCCTCGCAGGTTCGTC
CCTGCGCCCACTGCGGCCGTATTCATTCTCCCATCTCCCAAGGCAAGTTTTCACTCACCC
GTTTTTCGTCTTCTAGTTTTTCTTCTACCTCAGCGAGATCCCGTCGCGCCCTTCCCGCCA
TTCCAAGTCCGGCCGGTCTGAGGCCGAGGAGACGGCGACCGCGGAGGCCCGGAGGCGGGG
AGCTGACCGGCGAGAAACCGCAAACCGCCTCCGGGAAGCCGAGGAAGCTGCCCAGGAGGC
TGTCCGGGCTTGCCAGGCTGAGGAAGCCGCTTGGGAGGAGGCCCGACTCCACCGGGCCGA
GGAGTCCGTCCGGGAGACGGAGGCGGCGCGCGCACGCCGGGCGGCCTGCCAAGCCTCCGA
CCCGACTCCGCCCGAGGCGACGACGACTTCCGAGGCCACCCACGACGAGGCCGCGGGCGC
ACCGCTCGGGCCCGACCTCTCGGGCGACGCTCAGGACGGGCCAGCTTCGGGGGACGTCCC
TGAGTCCGGCACGTCGATCGGTGGGACGAGCCGTGCGGCATCCTTACCGAGGCGGCTCTC
CCCCATGCCTTCCGCTGCCCCACTGAGCGCGGAGCCCCTTCTGCAGGCCTTGGCCACCGC
GAACACCACGGTGTTGGACGGGTTAAGTGCCCAGATGGAGGTCCTGCAGGCAGAGCGGGC

	GGAGCTCGACGCGGCGTGGGCGCGTGTCGAAGAGGGGCGACGCTCGGTGGAGGCCATGGT GGAGGCGGGCCGCAAGGCACACCGCCGGCACGCCTCAGAGCTCGAGGCCCGTCGTAGGGA CCTGGCGGAGATCGCCAGGGAGGTGGAGGAGGAGCGGGAGACTGCCCTCATTGCCACCGC CGTTTTTAACGCGGCGCGGGACGACCTCTGCCTCCAATACGGGAGCCGGGAGGCGGAGCT AGAGAAGAAGCTCAACGCGTCCCGGCAGCGTGAAGAGGCCCTTGAGGCTCGCGCCACGGT GCTGGAGGAGCGCGCCCGCGCCCTGGACACGAAGGAAAGGGAACTGGCGGGTCGCGAGGC CGCCGTCGCCATCCGGGAGGCAACGCTGGCGGCGCACGAGGCCGCCTGCGCCGAAGAGGA GTCCGCACTCCGCCTCCGCGAGGACGCGCTCACCGAGCGGGAGCGAGCTCTCGAGGAGGC CGAGGCCGCGGCGCAGCGGCTGGCGGACAGCCTGTCCCTCCGCGGGGCAGCGCGGGAGGA GCAGGCGCGCCGTAATCTGGAAGGCGCCCGTGCCGAGAGGGCCGCACTGGATCAGCGGGC CGCTGAGCTCGAGGCGCGGGCGAAGGAGCTGGACGCGAGGGCGCGCAGCGGCGGGGCAGC CGCGGGCGACGGCGATCTAGCCGCCCGCCTCGCAGCTGCCGAACACACCATCGCCGAGCT GCAGGGCGCGCTGGACTCGTCCGCCGGGGAGGTCGAGGCCCTCCGCTTGGCAGGCGAGGT AGGGCCCGGCATGTTTCGGGACGCCGTCTCCCGTCTAGACCGCGCCGGGCGGCAGGCGGG CCTCTGGGGCGGGCGGATCGCGAAATATGCCGCCAACCAGGGGGGCCTCGCCCAGCGCCT CTCAGAGATGGCCGGGACCCTCCAGCGGCTCCCCGAGGAGCTCGAGGAGACGATCAAGTC ATCCTCGAGGGACCTCGCCCGGGGGGCGGTGGAGCTCGTACTGGCAAGCTACTAGGCCAG GGACCCGGACTTCTCCCCGTGGGCAGCGCTGGACGAGTTCTCTCCTGGAACCGAGGATGG CGCGCGCGCACAGGTCCGGGACGCCGCCGTCCACATCGTCCACAGCTTCGAGGGTACGGC CCCTCGGCTCGCGTTCGCCCTCGACTCCGACGAGGAGGGCGGCGACGGTGGTGCGGACGA CAGCGACGACGAGGCCGGCGACCCGGGCGCGTCGGCGTGAGCCCCCAGGCCCCCGCCAAT CTTAAATAGCTTTTTCTTTCTTCTTCTGAGGCCTTCGGGCCCCCTTCTTGTATAAACTAA TTCAATCTGCAATCAAATAAAGAGGAAATTTTTGTGTCAGCTCTGTTCGTTGTGTATATG AGACGAGGATGGTCTGTGCCGCGGTCCTTCTGTGTCTTGGCTTGAACGAGGACTCGTGCC CAGGTCCCAGCCTCAAACGGCGCGGGTCGGGGCTAGTGCCTGGGGAGATCCACGTGTCGA GACTGGCCAGGCCGGGAGCGCGGTGACCGAGGGTTATGGGGACCCGATTGTGGGCCTTTG CCGATTCCCCCCCGGAGTTCACCACGCCCCGGGGCACGGCTCGGTTCTGGGCCCC
LOC_Os04g45560	TGAGGTGGCTAACGTGGCATCCACAAGGCAACCCAGTAAGTGACTACCTCCCATCTTCTT CCTACACCACCCCCCTCTCTCCCCTCACTCCCTCCTATGGGCACTCCACTTGTGGGTAAG

GTTGCAAGGGGGTAGGCCCGTGAACCCACTTACAATACAAAAAAGCGGGT
AATCCCGCGT
AATTGGGTACGCGGGCCTGCCCCTTGCACCCCTACTTGTGGGTCACCGACG
GCACAACCG
GTGGCTAGCAGGCGACAGAGGAGGAGGAGCTCGAGTACAACCTAAGCCA
CCACACTCCTC
CTATTGAACTTGGGAAAGTGGTGACGAAGATCTCACAACACGAGTGAAGG
TACTCAAACA
TCGTGCGGGATTCATATGGGTCATCATGACAACAACCGTCTGAAGGCAGT
GGGAGGTGGA
GAGGAAGACCTCGATTAGTGTAGAGGGAAAGTTCTCAAGGTAGTACTTTA
GGTTGAGGCG
GATGGAGTGTTGCCTTCAAATGCATTGCATAAGAGGGAAAAGAGGGATGC
CGCTATAGCA
AGGAAGGAGTTGGTCGTTGTGACATCCCCTGAGGAGAGGGCGGTACCAAG
AAAGGAGACG
ATGGAGGGGAGGACGCTACCTAGCAGGGAAGTCGGAAGCTGTGGCGGCA
AAGAGGGCCTT
GGAGAGTTGTGCATGGATGAGAGACGACGTGGTGAGGAGCTGGAGGAGG
TTCGCTTTGAT
GGTGATGTAGTTGAAGGGCGAGAGGTTGATGGCGTTGTCACCAGGAGGTC
TTAGGCGGAG
AAGGAGATTCTTAAATTGGATGGAAGCAAGGAGCCACATCCGCGGGTCAT
GGCGCAGGGA
AGAGATGAGGCCCAACAACGTGAGCGTGAAAAGCCCAAGGAGCACAAGG
AGGACGCCAAG
CTTTGCTTGGTTGCATGACGTGTGACGGTGTCAAGGGAGAATGATTACGTG
AACCATCTC
GTCAGAGTGTCGAGCCTCTCCAGAGCCATCTTCATTCCCGTTGTACGGTGA
TATGGCAGA
AGGCAACTAGGGTTTGGTGGAACAAAGGAGATCAAGAGAGGTGCTGGAG
CTCCCGCAAGA
GAGGGAGGTGAGACTAAGAGAAGGGAGTAGTTGAGGAAAAAGGTGAGAG
ATAGCCATGTA
CATATCGATGGAGTTGTTTTAATATTCATTCGCGATATACCCGGTTTAAAG
GGATTAAAA
TGGACTCCAACCCACCCAAAGAGCGCGCGTGCAATTTGAAGGTAGAAATG
GGCTTTCATG
GCAGTACTTTCCTGCGGAAACCGGCCCAAGGCTACAGTTTTCGGGCCCCA
GACCCAGAAT
GCGAGGTACCAAGGCCTGCCTAGTCACTGAAAACGCCGCTTCTCTTCTGCG
AGCCTGCAC
TAAAAAAAGCCAGAAGAATTTTGCCCCAGATCCATACCAATATCAACACA
AGTTCTTCTC
CAAATTGCTGGCATTAGAACTTTGCAAATTTTAAATTATGCACTGTATGGC
AAACGGGGA
ACGGGGTTCCTCTTGCCTTAGTTCCAACAATGCAAAATTTATTTGAAAGTC
ATCGGTTCA
CTAGTTCATCACTGGCAGAATTAATTATACCATCCTCATACCGCCTATTCA
CATCTCTAT
CTCAAATAGGAGTACAGAACATTAAGCCTTTGATTTCTGGATTACGAGCA
ATATTTTGAA
GTTTGAAGATAGATACTTTAATTTACCGAA
GGATGTGTTTGAAAGAAGGGGATAAAAAATTTAAGAAAATGGACAAAATT
AGATCCACCG
LOC_Os04g45570 TTAGCATATAATTAATTAAGTAGTTATTATTTTAAATTTTTAAAAAATAAA
TTAACAATT
TTTTAAAATAACTTTTCTATCCACTTTTTTTAAAAAACTCCCGTTCAGCAAT

CCGGAAGT
GTAAGCGGAAAATGAGGGATTTAATTTTCTTTATTCTTCTAACGAACACAA
ATATACAAC
AGTAACTCCTCTGAAGCTGTATACTAAGATCACATGATGAATGATACTAA
AATTTGACAA
TATTCCGGCGTCCATTGAAAGATTCCTTTGCCATATGTTCATATACGTCGA
ATGGCCAAT
ACATTAAACAAAAATGAAGAAGGGCTGACACGTACCTATATTCCGGGAAC
GATTTTATGA
GTTGAGAAAGAAAAATTACGTTGAAACTCAACCATAAAAAAATTTCAGTT
AACTTTAATA
TCTTAAGGTATCTTAAACGAGCACGAAAATAATCTCGTTTCGTCGTCGCCA
CCATTATAA
TCTCATCATGACCACACCAACAGCGACCCTTTAACCCCAGAAATCCAAAA
CCAAACCAGC
GAGCGATCCTTAATCCTTGCAGATAAACAACTCTACTCCTCTCACCCCCCG
CATCCATCG
AACTCCAACCTCTTCCTCTCCTCTTCTTCCCTCCCATCGACGCGCAGATCTA
TCCCCCGT
CGCCGTGCGCTTCCCTCCGATCGCGAGCAGATCGGATTTATGCGGCCGTAC
TGCTTCTGC
TGCTGCTGCTGCGAGGAGCGCCAGCCATGTCGTGGAGCTTGAGCCGTGGT
GAGGGGGAGG
AGGCGGGCGGCGGCGGCGGCGGCGCTGGGGTCGGGGTTCGGCGGTCGCA
ATGTCCAGCAT
GTCGGACTCGGGGACGGGCGGCCGCGGGGGCGCGGAGCTGATGGTGGAG
CAGTTCCACCT
CAAGGTGCTCCATGCCGTGCTCGCCGTGCGCGGGCCGCGCCCGCTGCAGC
CGGCCGCGTC
GGCGTCGTTCAGGCGGCGGGACAGGTGGTTCCACCTCCCGCTCCACGATC
CGCAGCCTCC
GCCCGCGGCGGAGGGAGTGGAGGCGCCGGAGGCGGGGGAGCCGCTCGTG
GTGGACATCCT
CCTGGCTCACGCCGCAGCCGGCGGCGGCGGCGGCGGCGGAGCGGGCGGG
GAGGTGGTGGA
GAGATGGACGGTGGTGTGCGAGCCCTGGCCGGACGCCGCGGCGGGCGAG
GGGATACCCGT
GAACCGGGCGTACAAGCGCTGCATGACGATGCTTCGGTCGGTGTACGCCA
CGCTCCGCTT
CCTGCCGGCGTACCGCGTCTTCCGCCTCCTCTGCGCCAACCAATCCTACAA
CTACGAGAT
GGTCCACCGCGTCGGCTCCTTCGCCGTGCCGCTCTCCCGCGACGAGGAGGC
CGCCATGCG
CTCCTACCAGTTCGTCCCCGTCGAGACGCAGCACGGCCGCCTCGTCGTCTC
TGTCCAGTA
CCTCCCCAGCCTCGCTGCCTTCAATCTCGAGATCTCCTCCTTGTCGCCCTCC
ATGCTGAT
CGCCGACTACGTGGGCAGCCCTGCTGCCGAGCCAATGCGTGCCTTCCCAG
CTTCGCTCAC
AGGAGCCACAGGTTCCGCTTTCCCGCAGGCGCTGTCCAATCAGCCGCAGC
GCCCGCACAG
CTGGGCGACGCCAGCGCTCTGGCCGCAGGCGCCTAGGCAGCAAGCTAGAT
TTTCGCCACC
TCATTTGCTCAACGCATCACCCACTCCATCGCCACCCAATTTTCCCAGCGG
GTATCTTCA
ATCTCGTCCTAAGGGAGGGTCTGCACCAATGAGCATACCACAGGTGGGTG
ACAGGAGGAG
CCCAATTCACCGGCCGATCACATTGCCACCAACGTCTCCGAGGAGAGTAG

GGGAGACAGG
AACATCTAGTGCACAACAGTCTCCATCAGAGAGATGCCCCTCGTTTGGGA
GGGCAGATGG
ATTTCGAATAATGGATCCCTATGCAAGCTTGTCGCCAGGACGCAAGGTGA
TCCATGGTTT
TAAAGCCTCCAATATGCTTTCAACTTATATTTGATACTTGTGTATGCCTTAT
CTTCATCA
CTGCTCAATACTGGATCCTTCATGAATCATATAACTTTTGAGAGTGAGTGT
AAGGCAATT
CATCTGGTATATCTGAAGTAGACTTTTCTGTTTTCATGTGGCTGAGTAGAG
CCTACTGGC
ATTTACCTTGGAAACTAGGAAACAAACTCTACCAAGAGCTATTTCAAGTC
AATGTTTAGC
AGTACATATTCACAAGTTTAAGTTGTTATGTTAAGTTTAGCTTCATGATATT
GTGCAACA
AACCATGTTTGATAGTGGAATGTAAGCCATTATTAGTACTGGAGTTGAAGC
AGCCTAAAT
ACTGTTTGTTAGCATATACACCATTTAGCTCTCATCTCTCCTTTTTTTTTTGG
CAAAAAA
TAGACTACATATCAGGGTCTCTTATTAGTTGTGAACTGTTTCGATTGTCTCT
TGGTTCTT
GCTGAGACCATATTATTTGTTGAAAAGAAATGCGTTCATAGTACTTGTTGA
AACAGAAGA
ATGGTTGTTGAGTTAGTTATTTTCTGAGGCTTACCCAAAAGCAGCTTTTGA
TCTTCAAAT
GGTTGGATATTTGGATTACCCTTGTCTGCTGCAATATATGTCTGCTCTTGTA
CTTTTAAT
TTTTATTTAGTGGATTGCCTTTATGGGACTTAATCTTATCCTCTGCATTAGC
TATGGATA
ACTTATTTATCTTCCTCTAGCTATTGAAGCCAGGAAAAAAATATGTATTTG
AGTTGGGTC
TGCGTTGACACCGTTTGACCCATGTGCAGGGCAAGGACACTAAGGACGAA
TCTGGCAGGT
TCTCTGCACTCTCCTCCTGTGATTCCCCTCGACAAGATGATATAGATGATG
CAGATTATC
CCTTTGCTGTGGATGATGTTGATACACCGAGCTCCCAGCCTGGGTATGTAT
GGTAGTAAC
TTAAATATTCTATCTATAGATTTGATGAATTTCGCCATTTGCTAGGGCAGT
AGTGCATCA
GATAAAGCCAGAAACGAATTACATGAAGTCTTATCTATTCATATGAAATA
CGAAGTACTA
GCTCCTATCCGTTATATTTTGGGAGGGAGGCAGTACTAAGTTGATTTAGTA
TCCATAGAT
CTAGTTGTTAAAGAATGCTAGAGGCTATCCGGTTACACAAATATTTGATCG
TTAGAGCTA
TCAAATCCTATTATGTTGCACATTATTTGATTTTTTTTTCAAAAGTTTAAA
TATGACAT
TAAGGGGACTAAGGATTTGAAAACCTTTTGGACATGTTATTTTGTCATACT
ACAAAACTA
GTGTTAGAATAACTGGTAGAGATAGCAAAAAGGAATAATAAAAATGAGG
GTTATTGGCTT
AGTTAACATGGACTCGGTGTGCCTTGATCGACAGCTCTACCAATATGTAGT
TTATTGCAA
AGATCTAGATTTTGCATAGTTACCAGTTCTTAAGTTTCTTTGCTTTATAGTT
AGAGTGAA
CACAATCATATGTCAGAGCTAACAGTTGCTACAAATGCAAAGCACAAATT
TGGACTTGAC
TCGATGACTTTTTCTTTTTGCAACTTCTAAGTTCTAACCTCTATGTAACCTG

```
CCTCTTAG
TCACAATCAACAAGTTTTAACATATTAAAATGTTTTTTAGACAAAAGGCCA
AAGCCAAGC
TATTTATAGAAAGCATAGGTAGATAGAAGAGTGTAGTCGGGATTACCGGC
CATCCTGGTA
AAGTTTCTGGAACAAAATAGCCACTAGGCTACAACTCGACAATGATTAAC
ACACTGGCCA
ACAACCAAAAGCGATATCAAGATACAACCAGAACGGCAGAAAGAAACTA
ACAAAGCACTT
ATAAAAGATTAAATTGTACTAGTGGTCATCAAACACATATCAAACATGTC
CTACATGTGC
AAGTATGTTTGTTATGGCTTGGCCCTGGACTTAAATGGCAAAGGGTGTAGT
TGTGCTCGA
CAGGAGGGTGCAGCCTGAACCCTCACCTCCCTGGAACCCTGGAAGAGAAC
AGAGTGACAG
GACAGAAGAGAGTAGGCTAAGTTCTTGCTTGTGTTTAATAATGAATGCCA
CCTCTGTTGT
ATAGAGAAGGTCGTTTGCAATATCCAAATAATGTTCCTAACGATCCTCAAC
AATAGCTAA
ACAGACCTAACCAAGAAAATGATACGCAAACTAACATTTAACTGATCAAC
GGCACTGATT
AATCCTGTAGTAACACATAATCTGGCTTGCTGGGTCGACCGAACTTATGTC
CTCACATAA
CAAGGTAATGGTCACTCAAACCTATGCCATCACAGTAGCTGGCATGGCAT
TGATGCCTAT
GGGTTAGGCAAATTTGCTTGCTAATGACTTCTGTGAACACCTAATGAAGTT
TATTTACCT
ATTACACCAAAATCTAAGCTAGTTGAACTACTTGCACGTGTGAGATAAGTC
TGCAGAAGA
GCTCCAGGAAAGTACAATTCAGCACTTTATTTTCTTAATGACAATCCATAC
AATATGGGC
CTTCCCCTACAGTTCACCATAAAAATCAATTACAGAACCATTGAAAACAG
AAATCTACAA
CCCCAACCGGCAGTCACTGTACCACCATATTGAGTGAACAACTTGAATGG
CGAGAAATGC
TACATCTGACAAGCTAGATGGATAGCACATGGGAGAGGGATGATGATTAT
TCAGGGATGC
TGATTGCTGAGTATCTGGTTTGTGGTCGAGGGTCAAATTTCAGGCCTTAGG
AAGAGTGAA
TTGCTCAAATGGAGTTTTCCTTTTACTTTTATTACACTCGTACATAAATTAA
ATAGTAAA
TTTCATAGAACCCCCGGTATTATGGTCCAAATTGCAGAAAACCACGGGGG
TTTTGGCACG
TGACACATAACCTTGGGTATTATATTATAGTACTATAGTTTTATAAAACCA
CACTGTCAA
GGCGTCAACAAATTTAACTTGACATTGTAGCAAAATAAAAAATGCTTAAC
TTAGATGGCT
AAAGTTTTTTATGATAGATAAAAAAATTCATTAATTCGAGTTTTGAGATTT
GATATTAGA
AACTCATACTTATGATGAAACAAAGAACAATTGGTTAACAATGAGGTTTT
GTGAAACTCT
AGAACCATGATATCCGGGGTTGTGCCACATGCCAAAACCTTTGTGGTTTTG
TTCAACAAG
GACCATAATACTAGTGGCACGGTTCTAATTTCCCCAAATTCTAGTGGCACG
TAGCCGATA
TCGTGAATAGTAGTGGCATTTATCTAAAATGGCATATTTGTAGTGGCATGT
CCTAATACA
TAAGGTTATGTGAAATTTACTCTAAATTAAACATCTAGTTTTGCTTGTTTGG
```

GCGGTCAT
TGTTATGTGGTGCCAAAATAGGATAATTATATTCAATGCATTGAGCTACAT
TTGTTGCGT
GATAGTCTATATGGTGTATTGAATAAGATGATTAGTCTATTTTACTATTCC
CAGTGCCAT
CAGATATTCTTTTGACGCAGCACCCTATATTAACAGGAGATTCAGATCTAA
TATAACCTT
ATCTGGATATGTGCCTCTATTATTATTCACCCCATTCATGTTTATCTTTGTT
GAGCTGTA
CCCATGGCTTTTGATAAAATATAACAGTTATGTATGTCCTCTTGTATTTCTA
CTATTCGG
TTCTTTGTTTCCACTACTGAGTTTCTATGGTTCTAACACCGTTTCTATATGC
ACTATATA
ATATACAGGAGCAGTGATGGGAAGGAAGCTAGAGATCAGGCAAGTTCGTC
GTCCCATAAA
TCACAAGATGCCGCAGTTGGCTCTCTGGTGCACCTGCTCAAAACTGCACGC
CCTCTGCGA
AACTCTAATTGTCCATCTCAAGCATCAGCGGTTGAATCTAGTGAAGCAGCC
TCAACCAGT
TCTGTCGTGTCCCGTAGGAAATCTGATGCACTCGAAGAGCTCCAATCATTT
AAAGAAATC
AAGGAGAGGTTGATGTCCCGGAGTAGAGCAAAACAGCAAGAGCCACCAG
AAAAACCATAG
CACCGTTCTTTCCTTCGTAAACTTCACAGGTGAAGATCCCCTACCACAAAG
CGGGCTATG
CGGTGAGCTCAGTGACGTTATGCTAAGGAATTCGTGTGTATATAACTGATG
GACTGATGC
TTTTAGTTAAAGAGAAGATACGAGTTACCACTGATTTCTGCGAGAATCCCT
CGCCATTCT
TGAACTCTAGAAGAATTACGCCACTGTTTATACGCTTTTATACGTAGGTGG
GAGCCTCTT
GACGAATATAGCAGCTCCAGCGTGTGCGTGTGTAATTATGTGACTGTACTG
ACTACTGTG
CCGAGTGTTCGGTAAATTAAAGTATCTATCTTCAAACTTCAAAATATTGCT
CGTAATCCA
GAAATCAAAGGCTTAATGTTCTGTACTCCTATTTGAGATAGAGATGTGAAT
AGGCGGTAT
GAGGATGGTATAATTAATTCTGCCAGTGATGAACTAGTGAACCGATGACT
TTCAAATAAA
TTTTGCATTGTTGGAACTAAGGCAAGAGGAACCCCGTTCCCCGTTTGCCAT
ACAGTGCAT
AATTTAAAATTTGCAAAGTTCTAATGCCAGCAATTTGGAGAAGAACTTGTG
TTGATATTG
GTATGGATCTGGGGCAAAATTCTTCTGGCTTTTTTTAGTGCAGGCTCGCAG
AAGAGAAGC
GGCGTTTTCAGTGACTAGGCAGGCCTTGGTACCTCGCATTCTGGGTCTGGG
GCCCGAAAA
CTGTAGCCTTGGGCCGGTTTCCGCAGGAAAGTACTGCCATGAAAGCCCATT
TCTACCTTC
AAATTGCACGCGCGCTCTTTGGGTGGGTTGGAGTCCATTTTAATCCCTTTA
AACCGGGTA
TATCGCGAATGAATATTAAAACAACTCCATCGATATGTACATGGCTATCTC
TCACCTTTT
TCCTCAACTACTCCCTTCTCTTAGTCTCACCTCCCTCTCTTGCGGGAGCTCC
AGCACCTC
TCTTGATCTCCTTTGTTCCAC

4. Discussão Geral

Nossa análise de BLAST levantou 28 genes candidatos homólogos de LOC_Os11g15624.1, incluindo espécies de monocotiledôneas da família *Poaceae* e gênero *Oryza* (*O. sativa*, *O. glaberrima*, *O. rufipogon*, *O. glumaepatula, O. nivara*, *O. barthii*, *O. punctata* e *O. meridionalis*); e 30 genes candidatos homólogos de LOC_Os04g45510.1, incluindo espécies de monocotiledôneas, sendo 25 da família *Poaceae,* 3 da família *Joinvilleaceae*, 1 da família *Bromeliaceae* e da família *Acoraceae*.

Após análise de nossas árvores filogenéticas, propomos que LOC_Os11g15624.1 é uma novidade evolutiva restrita ao gênero *Oryza*, presente em espécies de genoma AA e possivelmente algumas espécies mais distantes (*O. punctata*) com múltiplas cópias pelo genoma. Nossos resultados sugerem que LOC_Os11g15624.1 possui características de um TRG. Além disso, propomos que LOC_Os04g45510.1 é um gene com ancestralidade mais distante na família *Poaceae*, apresentando uma única cópia no genoma de cada espécie e alto grau de conservação (principalmente entre espécies do gênero *Oryza*).

Nós observamos alto grau de sintenia e colinearidade na região genômica envolta de LOC_Os11g15624.1 entre *O. sativa* e *O. rufipogon, O. glaberrima* e *O. barthii.* Analisando a região genômica envolta de LOC_Os04g45510.1, observamos maior grau de sintenia e colinearidade entre *O. sativa* e *O. glumaepatula, O. rufipogon, O. nivara, O. glaberrima* e *O. barthii.*

Nós confirmamos a expressão de LOC_Os11g15624.1 por RT-qPCR tanto em parte aérea quanto em raiz das espécies *O. sativa*, *O. rufipogon* e *O. meridionalis*. A transformação de plantas de arroz com os genes de interesse está em progresso, sendo que geramos nossos primeiros clones da linhagem transgênica 11.23 (superexpressando LOC_Os11g15624.1) a ser confirmado por sequenciamento e resistência à higromicina. Estamos coletando as primeiras sementes e planejamos investigar os efeitos fenotípicos e genéticos da superexpressão do gene de interesse.

5. Referências

ANDERSSON, D. I.; JERLSTRÖM-HULTQVIST, J.; NÄSVALL, J. Evolution of New Functions De Novo and from Preexisting Genes. Cold Spring Harbor Perspectives in Biology, [s. l.], v. 7, n. 6, p. a017996, 2015.

BECKER, M.; ASCH, F. Iron toxicity in rice—conditions and management concepts. Journal of Plant Nutrition and Soil Science, [s. l.], v. 168, n. 4, p. 558–573, 2005.

BLAST SEARCH - ORYZA_SATIVA - ENSEMBL GENOMES 54. [s. l.], [s. d.]. Disponível em: https://plants.ensembl.org/Multi/Tools/Blast. Acesso em: 6 out. 2022.

BRIAT, J.-F. et al. Ferritins and iron storage in plants. Biochimica et Biophysica Acta (BBA) - General Subjects, [s. l.], v. 1800, n. 8, Ferritin: Structures, Properties and Applications, p. 806–814, 2010.

CAMASCHELLA, C. Iron deficiency. Blood, [s. l.], v. 133, n. 1, p. 30–39, 2019.

CHEN, J. et al. Whole-genome sequencing of Oryza brachyantha reveals mechanisms underlying Oryza genome evolution. Nature Communications, [s. l.], v. 4, n. 1, p. 1595, 2013.

CHENG, L. et al. Mutation in Nicotianamine Aminotransferase Stimulated the Fe(II) Acquisition System and Led to Iron Accumulation in Rice. Plant Physiology, [s. l.], v. 145, n. 4, p. 1647–1657, 2007.

COÊLHO, J. D. Arroz: produção e mercado. [s. l.], 2021. Disponível em: https://bnb.gov.br/s482-dspace/handle/123456789/s1dspp01.dmz.bnb:8443/s482-dspace/handle/123456789/698. Acesso em: 29 fev. 2024.

COLANGELO, E. P.; GUERINOT, M. L. The Essential Basic Helix-Loop-Helix Protein FIT1 Is Required for the Iron Deficiency Response. The Plant Cell, [s. l.], v. 16, n. 12, p. 3400–3412, 2004.

CONAB. Boletim da Safra de Grãos. [S. l.], 2024. Disponível em: http://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos. Acesso em: 7 fev. 2024.

DE OLIVEIRA, B. H. N. et al. The Mitochondrial Iron-Regulated (MIR) gene is Oryza genus specific and evolved before speciation within the Oryza sativa complex. Planta, [s. l.], v. 251, n. 5, p. 94, 2020.

DEY, S. et al. Iron pulsing, a cost effective and affordable seed invigoration technique for iron bio-fortification and nutritional enrichment of rice grains. Plant Growth Regulation, [s. l.], v. 100, n. 2, p. 545–559, 2023.

DROGUE, B. et al. Plant root transcriptome profiling reveals a strain-dependent response during Azospirillum-rice cooperation. Frontiers in Plant Science, [s. l.], v. 5, 2014. Disponível em: https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2014.00607. Acesso em: 24 fev. 2024.

DUFEY, I. et al. Construction of an integrated map through comparative studies allows the identification of candidate regions for resistance to ferrous iron toxicity in rice. Euphytica, [s. l.], v. 203, n. 1, p. 59–69, 2015.

ENSEMBL PLANTS. [s. l.], [s. d.]. Disponível em: https://plants.ensembl.org/index.html. Acesso em: 6 out. 2022.

FREI, M. et al. Responses of rice to chronic and acute iron toxicity: genotypic differences and biofortification aspects. Plant and Soil, [s. l.], v. 408, n. 1, p. 149–161, 2016.

GAO, L.-Z. et al. Evolution of Oryza chloroplast genomes promoted adaptation to diverse ecological habitats. Communications Biology, [s. l.], v. 2, n. 1, p. 1–13, 2019.

GNANAMANICKAM, S. S. Rice and Its Importance to Human Life. In: GNANAMANICKAM, S. S. (org.). Biological Control of Rice Diseases. Dordrecht: Springer Netherlands, 2009. (Progress in Biological Control). p. 1–11. Disponível em: https://doi.org/10.1007/978-90-481-2465-7_1. Acesso em: 7 fev. 2024.

GOLDBERG, A. D.; ALLIS, C. D.; BERNSTEIN, E. Epigenetics: A Landscape Takes Shape. Cell, [s. l.], v. 128, n. 4, p. 635–638, 2007.

GONÇALVES, B. X. et al. Phosphate starvation responses in crop roots: from well-known players to novel candidates. Environmental and Experimental Botany, [s. l.], v. 178, p. 104162, 2020.

GRILLET, L. et al. IRON MAN is a ubiquitous family of peptides that control iron transport in plants. Nature Plants, [s. l.], v. 4, n. 11, p. 953–963, 2018.

GUERINOT, M. L.; YI, Y. Iron: Nutritious, Noxious, and Not Readily Available. Plant Physiology, [s. l.], v. 104, n. 3, p. 815–820, 1994.

HINDT, M. N.; GUERINOT, M. L. Getting a sense for signals: Regulation of the plant iron deficiency response. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, [s. l.], v. 1823, n. 9, Cell Biology of Metals, p. 1521–1530, 2012.

ISHIMARU, Y. et al. Rice plants take up iron as an Fe3+-phytosiderophore and as Fe2+. The Plant Journal, [s. l.], v. 45, n. 3, p. 335–346, 2006.

ISHIMARU, Y. et al. Rice-Specific Mitochondrial Iron-Regulated Gene (MIR) Plays an Important Role in Iron Homeostasis. Molecular Plant, [s. l.], v. 2, n. 5, p. 1059–1066, 2009.

JAKOBY, M. et al. FRU (BHLH029) is required for induction of iron mobilization genes in Arabidopsis thaliana. FEBS Letters, [s. l.], v. 577, n. 3, p. 528–534, 2004.

KÁMÁN-TÓTH, E. et al. A simplified and efficient Agrobacterium tumefaciens electroporation method. 3 Biotech, [s. l.], v. 8, n. 3, p. 148, 2018.

KAPUSTIN, Y. et al. Splign: algorithms for computing spliced alignments with identification of paralogs. Biology Direct, [s. l.], v. 3, n. 1, p. 20, 2008.

KAR, S.; PANDA, S. K. Iron Homeostasis in Rice: Deficit and Excess. Proceedings of the National Academy of Sciences, India Section B: Biological Sciences, [s. l.], v. 90, n. 2, p. 227–235, 2020.

KATOH, K.; STANDLEY, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution, [s. l.], v. 30, n. 4, p. 772–780, 2013.

KELLOGG, E. A. The Evolutionary History of Ehrhartoideae, Oryzeae, and Oryza. Rice, [s. l.], v. 2, n. 1, p. 1–14, 2009.

KOBAYASHI, T.; NAGANO, A. J.; NISHIZAWA, N. K. Iron deficiency-inducible peptide-coding genes OsIMA1 and OsIMA2 positively regulate a major pathway of iron uptake and translocation in rice. Journal of Experimental Botany, [s. l.], v. 72, n. 6, p. 2196–2211, 2021.

LARCHER, W. Physiological Plant Ecology. [S. l.: s. n.], 2003. Disponível em: https://link.springer.com/book/9783540435167. Acesso em: 7 fev. 2024.

LI, J. et al. Foster thy young: Enhanced prediction of orphan genes in assembled genomes. [S. l.]: bioRxiv, 2021. Disponível em: https://www.biorxiv.org/content/10.1101/2019.12.17.880294v3. Acesso em: 22 fev. 2024.

LI, L. et al. Identification of the novel protein QQS as a component of the starch metabolic network in Arabidopsis leaves. The Plant Journal, [s. l.], v. 58, n. 3, p. 485–498, 2009.

LI, G. et al. The Sequences of 1504 Mutants in the Model Rice Variety Kitaake Facilitate Rapid Functional Genomic Studies. The Plant Cell, [s. l.], v. 29, n. 6, p. 1218–1231, 2017.

LI, L.; WURTELE, E. S. The QQS orphan gene of Arabidopsis modulates carbon and nitrogen allocation in soybean. Plant Biotechnology Journal, [s. l.], v. 13, n. 2, p. 177–187, 2015.

LIANG, G. et al. Oryza sativa FER-LIKE FE DEFICIENCY-INDUCED TRANSCRIPTION FACTOR (OsFIT/OsbHLH156) interacts with OsIRO2 to regulate iron homeostasis. Journal of Integrative Plant Biology, [s. l.], v. 62, n. 5, p. 668–689, 2020.

LICHTBLAU, D. M. et al. The small iron-deficiency-induced protein OLIVIA and its relation to the bHLH transcription factor POPEYE. [S. l.]: bioRxiv, 2023. Disponível em: https://www.biorxiv.org/content/10.1101/2023.12.05.570225v1. Acesso em: 7 fev. 2024.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods, [s. l.], v. 25, n. 4, p. 402–408, 2001.

MAJERUS, V. et al. Organ-dependent responses of the african rice to short-term iron toxicity: Ferritin regulation and antioxidative responses. Biologia Plantarum, [s. l.], v. 51, n. 2, p. 303–312, 2007.

MAJERUS, V.; BERTIN, P.; LUTTS, S. Abscisic acid and oxidative stress implications in overall ferritin synthesis by African rice (Oryza glaberrima Steud.) seedlings exposed to short term iron toxicity. Plant and Soil, [s. l.], v. 324, n. 1, p. 253–265, 2009.

MANN, D. G. J. et al. Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (Panicum virgatum L.) and other monocot species. Plant Biotechnology Journal, [s. l.], v. 10, n. 2, p. 226–236, 2012.

MENGUER, P. K.; SPEROTTO, R. A.; RICACHENEVSKY, F. K. A walk on the wild side: Oryza species as source for rice abiotic stress tolerance. Genetics and Molecular Biology, [s. l.], v. 40, n. 1 suppl 1, p. 238–252, 2017.

NGUYEN, L.-T. et al. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution, [s. l.], v. 32, n. 1, p. 268–274, 2015.

NISHIKAWA, T.; VAUGHAN, D. A.; KADOWAKI, K. Phylogenetic analysis of Oryza species, based on simple sequence repeats and their flanking nucleotide sequences from the mitochondrial and chloroplast genomes. Theoretical and Applied Genetics, [s. l.], v. 110, n. 4, p. 696–705, 2005.

NOGUEIRA-DE-ALMEIDA, C. A. et al. Prevalence of childhood anaemia in Brazil: still a serious health problem: a systematic review and meta-analysis. Public Health Nutrition, [s. l.], v. 24, n. 18, p. 6450–6465, 2021.

O'CONNER, S. et al. From Arabidopsis to Crops: The Arabidopsis QQS Orphan Gene Modulates Nitrogen Allocation Across Species. In: SHRAWAT, A.; ZAYED, A.; LIGHTFOOT, D. A. (org.). Engineering Nitrogen Utilization in Crop Plants. Cham: Springer International Publishing, 2018. p. 95–117. Disponível em: https://doi.org/10.1007/978-3-319-92958-3_6. Acesso em: 20 fev. 2024.

OGO, Y. et al. OsIRO2 is responsible for iron utilization in rice and improves growth and yield in calcareous soil. Plant Molecular Biology, [s. l.], v. 75, n. 6, p. 593–605, 2011.

OGO, Y. et al. The rice bHLH protein OsIRO2 is an essential regulator of the genes involved in Fe uptake under Fe-deficient conditions. The Plant Journal, [s. l.], v. 51, n. 3, p. 366–377, 2007.

OZAWA, K. Establishment of a high efficiency Agrobacterium-mediated transformation system of rice (Oryza sativa L.). Plant Science, [s. l.], v. 176, n. 4, p. 522–527, 2009.

PENG, F. et al. IRONMAN peptide interacts with OsHRZ1 and OsHRZ2 to maintain Fe homeostasis in rice. Journal of Experimental Botany, [s. l.], v. 73, n. 18, p. 6463–6474, 2022.

PHILLIPS, A. L. et al. The first long-read nuclear genome assembly of Oryza australiensis, a wild rice from northern Australia. Scientific Reports, [s. l.], v. 12, n. 1, p. 10823, 2022.

PRADHAN, S. K. et al. Genetic regulation of homeostasis, uptake, bio-fortification and efficiency enhancement of iron in rice. Environmental and Experimental Botany, [s. l.], v. 177, p. 104066, 2020.

RAMBAUT, A. FigTree. Versão v1.3.1. Edinburgh: Institute of Evolutionary Biology, University of Edinburgh, 2010.

RASHEED, A. et al. IRON TOXICITY, TOLERANCE AND QUANTITATIVE TRAIT LOCI MAPPING IN RICE; A REVIEW. Applied Ecology and Environmental Research, [s. l.], v. 18, n. 6, p. 7483–7498, 2020.

RICACHENEVSKY, F. K. et al. ZINC-INDUCED FACILITATOR-LIKE family in plants: lineage-specific expansion in monocotyledons and conserved genomic and expression features among rice (Oryza sativa) paralogs. BMC Plant Biology, [s. l.], v. 11, n. 1, p. 20, 2011.

RICE ALMANAC: SOURCE BOOK FOR THE MOST IMPORTANT ECONOMIC ACTIVITIES ON EARTH. Fourth Editioned. Los Baños, Philippines: IRRI, 2013.

RICE GENOME ANNOTATION PROJECT. [S. l.], [s. d.]. Disponível em: http://rice.uga.edu/cgi-bin/ORF_infopage.cgi. Acesso em: 3 out. 2022.

ROGERS, E. E.; GUERINOT, M. L. FRD3, a Member of the Multidrug and Toxin Efflux Family, Controls Iron Deficiency Responses in Arabidopsis. The Plant Cell, [s. l.], v. 14, n. 8, p. 1787–1799, 2002.

SANTOS, A. B. dos et al. A cultura do arroz no Brasil. [S. l.]: Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2006., 2006. Disponível em: http://www.infoteca.cnptia.embrapa.br/handle/doc/193176. Acesso em: 29 fev. 2024.

SASAKI, T. The map-based sequence of the rice genome. Nature, [s. l.], v. 436, n. 7052, p. 793–800, 2005.

SCHLÖTTERER, C. Genes from scratch – the evolutionary fate of de novo genes. Trends in Genetics, [s. l.], v. 31, n. 4, p. 215–219, 2015.

SENGE-PR. Carência de ferro no organismo já faz 40% da população mundial ser anêmica e índice pode aumentar com pandemia. In: SENGE-PR. 5 ago. 2021. Disponível em: https://www.senge-pr.org.br/noticia/carencia-de-ferro-no-organismo-ja-faz-40-da-populacao-mundial-ser-anemica-e-indice-pode-aumentar-com-pandemia/. Acesso em: 26 jan. 2024.

SILVEIRA, A. B. et al. Extensive Natural Epigenetic Variation at a De Novo Originated Gene. PLOS Genetics, [s. l.], v. 9, n. 4, p. e1003437, 2013.

SILVEIRA, V. C. da et al. Role of ferritin in the rice tolerance to iron overload. Scientia Agricola, [s. l.], v. 66, p. 549–555, 2009.

SINGH, U.; SYRKIN WURTELE, E. How new genes are born. eLife, [s. l.], v. 9, p. e55136, 2020.

STEIN, J. C. et al. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. Nature Genetics, [s. l.], v. 50, n. 2, p. 285–296, 2018.

STEIN, R. J. et al. Genotype Variation in Rice (Oryza sativa L.) Tolerance to Fe Toxicity Might Be Linked to Root Cell Wall Lignification. Frontiers in plant science, [s. l.], v. 10, p. 746, 2019.

STEIN, R. J.; RICACHENEVSKY, F. K.; FETT, J. P. Differential regulation of the two rice ferritin genes (OsFER1 and OsFER2). Plant Science, [s. l.], v. 177, n. 6, p. 563–569, 2009.

TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution, [s. l.], v. 38, n. 7, p. 3022–3027, 2021.

TAUTZ, D.; DOMAZET-LOŠO, T. The evolutionary origin of orphan genes. Nature Reviews Genetics, [s. l.], v. 12, n. 10, p. 692–702, 2011.

VAKIRLIS, N.; CARVUNIS, A.-R.; MCLYSAGHT, A. Synteny-based analyses indicate that sequence divergence is not the main source of orphan genes. eLife, [s. l.], v. 9, p. e53500, 2020.

VAUGHAN, D. A. et al. On the Phylogeny and Biogeography of the Genus Oryza. Breeding Science, [s. l.], v. 55, n. 2, p. 113–122, 2005.

VELTRI, D.; WIGHT, M. M.; CROUCH, J. A. SimpleSynteny: a web-based tool for visualization of microsynteny across multiple species. Nucleic Acids Research, [s. l.], v. 44, n. W1, p. W41-45, 2016.

WAIRICH, A. et al. Chromosomal introgressions from Oryza meridionalis into domesticated rice Oryza sativa result in iron tolerance. Journal of Experimental Botany, [s. l.], v. 72, n. 6, p. 2242–2259, 2021.

WAIRICH, A. et al. The Combined Strategy for iron uptake is not exclusive to domesticated rice (Oryza sativa). Scientific Reports, [s. l.], v. 9, n. 1, p. 16144, 2019.

WANG, R. et al. IMA peptides function in iron homeostasis and cadmium resistance. Plant Science, [s. l.], v. 336, p. 111868, 2023.

WANG, W. et al. OsIRO3 Plays an Essential Role in Iron Deficiency Responses and Regulates Iron Homeostasis in Rice. Plants, [s. l.], v. 9, n. 9, p. 1095, 2020.

WASSMANN, R. et al. Rice and global climate change. In: [S. l.: s. n.], 2010. p. 22.

WHO. Fact sheets - Malnutrition. [S. l.], 2023. Disponível em: https://www.who.int/news-room/fact-sheets/detail/malnutrition. Acesso em: 26 jan. 2024.

YU, J. et al. A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. indica). Science, [s. l.], v. 296, n. 5565, p. 79–92, 2002.

YUAN, Y. X. et al. AtbHLH29 of Arabidopsis thaliana is a functional ortholog of tomato FER involved in controlling iron acquisition in strategy I plants. Cell Research, [s. l.], v. 15, n. 8, p. 613–621, 2005.

YUAN, Y. et al. FIT interacts with AtbHLH38 and AtbHLH39 in regulating iron uptake gene expression for iron homeostasis in Arabidopsis. Cell Research, [s. l.], v. 18, n. 3, p. 385–397, 2008.

ZHANG, Y. et al. Genome and Comparative Transcriptomics of African Wild Rice Oryza longistaminata Provide Insights into Molecular Mechanism of Rhizomatousness and Self-Incompatibility. Molecular Plant, [s. l.], v. 8, n. 11, p. 1683–1686, 2015.

ZHANG, L. et al. Rapid evolution of protein diversity by de novo origination in Oryza. Nature Ecology & Evolution, [s. l.], v. 3, n. 4, p. 679–690, 2019.

ZHANG, X. et al. The Adaptive Mechanism of Plants to Iron Deficiency via Iron Uptake, Transport, and Homeostasis. International Journal of Molecular Sciences, [s. l.], v. 20, n. 10, p. 2424, 2019.

ZHANG, Y. et al. Vacuolar membrane transporters OsVIT1 and OsVIT2 modulate iron translocation between flag leaves and seeds in rice. The Plant Journal, [s. l.], v. 72, n. 3, p. 400–410, 2012.