

ARTIGOS COMPLETOS	
RESUMOS DE PESQUISA	





ARTIGOS COMPLETOS

ANÁLISE IN SILICO DA FAMÍLIA UDP-GLICOSILTRANSFERASE (UGT) NO GENOMA DE PHASEOLUS VULGARI.	S
L	192

ANÁLISE IN SILICO DA FAMÍLIA UDP-GLICOSILTRANSFERASE (UGT) NO GENOMA DE PHASEOLUS VULGARIS L.

Luis Gustavo Gomes Lobo, Tiago Benedito dos Santos

Universidade do Oeste Paulista – UNOESTE, Presidente Prudente, SP. E-mail: gustavo2620@outlook.com

RESUMO

As UDP-glicosiltransferases (UGTs; EC 2.4.1.x) são representadas por uma família multigênica altamente diversa e podem ser amplamente encontrados em animais, plantas e microrganismos. Em seu papel catalítico nas plantas, os UGTs alteram a atividade biológica de compostos de moléculas pequenas, transferindo doadores de glicosil com receptores e, em seguida, participam e influenciam no crescimento e o desenvolvimento das plantas, podendo desencadear a formação de metabólitos secundários e a resposta a estresses abióticos. Na literatura existem alguns estudos sobre a caracterização dessa família multigênica em algumas espécies de plantas, contudo, não há sua descrição em *Phaseolus vulgaris* L. O objetivo desse estudo foi identificar e caracterizar através das metodologias de bioinformática os genes *UGTs* no genoma de *Phaseolus vulgaris* L. Nossas análises resultaram na identificação de 68 genes putativos UGTs, o quais foram agrupados em 17 grupos filogenéticos principais, distribuídos em todos os 10 cromossomos. O perfil transcricional permitiu a catalogar alguns genes putativos para futuras validações moleculares. As informações geradas por esse estudo facilitarão a triagem de genes candidatos e posterior caracterização de sua especificidade e função no feijoeiro.

Palavras-chave: Bioinformática, feijão, PvUGT, expressão de genes.

IN SILICO ANALYSIS OF THE UDP-GLYCOSYLTRANSFERASE (UGT) FAMILY IN THE GENOME OF *PHASEOLUS VULGARIS* L.

ABSTRACT

UDP-glycosyltransferases (UGTs; EC 2.4.1.x) are represented by a highly diverse multigene family and can be found widely in animals, plants and microorganisms. In their catalytic role in plants, UGTs alter the biological activity of small molecule compounds, transferring glycosyl donors with receptors and then participating and influencing plant growth and development, which can trigger the formation of secondary metabolites and response to abiotic stresses. In the literature there are some studies on the characterization of this multigene family in some plant species, however, there is no description in *Phaseolus vulgaris* L. The aim of this study was to identify and characterize, through bioinformatics methodologies, the *UGTs* genes in the genome of Phaseolus vulgaris L Our analyzes resulted in the identification of 68 putative *UGTs* genes, which were grouped into 17 major phylogenetic groups, distributed across all 10 chromosomes. The transcriptional profile allowed to catalog some putative genes for future molecular validations. The information generated by this study will facilitate the screening of candidate genes and subsequent characterization of their specificity and function in common bean. **Keywords:** Bioinformatics, common bean, PvUGT, gene expression.

INTRODUÇÃO

As glicosiltransferases (GTs; EC 2.4.x.y), é considerada uma superfamília de enzimas que catalisam a transferência de porções de açúcar de moléculas doadoras ativadas para moléculas aceptadoras específicas, formando ligações de glicosídicos, que estão envolvidas na biossíntese de dissacarídeos, oligossacarídeos e polissacarídeos, e estão amplamente presentes em todos os organismos vivos (MEECH et al., 2019). Até o presente momento (Agosto 2021), um total de 114 superfamílias de glicosiltransferase foram incluídos no banco de dados CAZy (http://www.cazy.org/GlycosylTransferases.html/), e são divididas com base na especificidade dos substratos, similaridade das sequências de aminoácidos e especificidade

catalítica. A uridina difosfato glicosiltransferase (UGT) é a maior família de glicosiltransferase presentes nas espécies de plantas (YANG et al., 2020), também conhecida como família da glicosiltransferase 1 (GT1). UGTs usam doadores de açúcar difosfato de uridina (UDP) (UDP-glicose, UDP-galactose, UDP-arabinose, UDP-ramnose, UDP-xilose ou UDP-ácido glucurônico) para catalisar a reação de glicosilação. Existe um motivo PSPG altamente conservado na extremidade carboxila da proteína UGT como uma característica unificada que é uma das poucas regiões com similaridade de sequência significativa. O motivo PSPG é o local de ligação do nucleotídeo-difosfato-açúcar da enzima UGT (BOWLES et al., 2005) e contém 44 resíduos de aminoácidos. A maioria dos UGTs está intimamente associada à glicosilação de metabólitos secundários importantes, como flavonóides, terpenóides e esteróis pertencentes a esta família (VOGT; JONES, 2000; LE ROY et al., 2016).

As UGTs foram identificadas em várias espécies de plantas, desde planta inferior, como *Chlamydomonas reinhardtii*, a plantas superiores, como a *Vitis vinifera* (YONEKURA-SAKAKIBARA; HANADA, 2011; WEI et al., 2021). Adicionalmente, mais de 100 UGTs foram identificados em *Arabidopsis thaliana* (Li et al., 2001), 180 em *Oryza sativa* (CAPUTI et al., 2012), 147 em *Zea mays* (LI et al., 2014) e 179 em *Triticum aestivum* L. (HE et al., 2018), respectivamente.

Embora na literatura existem alguns relatos sobre a caracterização dessa família multigênica em algumas espécies de plantas importantes, todavia, não há descrição em *Phaseolus vulgaris* L. O objetivo deste estudo foi identificar e caracterizar através das ferramentas da bioinformática os genes da família da glicosiltransferase (*UGTs*) presentes no genoma de *P. vulgaris* L. As informações geradas por esse estudo irá permitir a triagem de genes candidatos e posterior caracterização de sua especificidade e função no feijoeiro.

MATERIAL E MÉTODOS

O ponto de partida para as buscas dos genes candidatos UGTs de P. vulgaris L., foi através de palavra-chave (UGT-Glycosyltransferase) no banco de dados Phytozome (https://phytozome.jgi.doe.gov/ GOODSTEIN et al., 2012). Todas as sequências referentes aos genes UGTs de P. vulgaris L. (genômicas, aminoácidos e CDS - Coding DNA Sequence/sequência codificadora) foram baixadas e posteriormente arquivadas em documento no formato FASTA para análises in silico posteriores. Em seguida, todas as sequências de proteínas foram verificadas quanto a presença de possíveis domínios conservados. Prosseguiu-se também o confronto das sequências no banco de dados do NCBI (National Center for Biotechnology Information/ ALTSHUL et al., 1997), utilizando as ferramentas BlastX e BlastP, tendo como primícia verificar sua especificidade. Com o banco de dados NCBI Conserved Domains (https: //www.ncbi.nlm.nih. gov/ Structure / cdd / wrpsb.cgi), as sequências PvUGTs candidatas foram confirmadas como membro da família UGT. A plataforma Phytozome também foi utilizada para a obtenção das sequências homólogas de UGTs de outras espécies de plantas para análises in silico adicionais. O comprimento, peso molecular (kDa) e o ponto isoelétrico (pl) de cada proteína UGT foram calculados pelo programa online ExPASy (https://web.expasy.org/compute pi/ GASTEIGER et al., 2003). A localização de proteínas UGT foi prevista usando a ferramenta de análise Plant-mPLoc subcelular (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/; CHOU; SHEN, 2010). As estruturas do genes UGTs (éxon/íntron) foram geradas do Gene através programa Structure Display (GSDS) (http://gsds.cbi.pku.edu.cn/ HU et al., 2015). A plataforma online MapGene2chromosome v2.1 (MG2C; http://mg2c.iask.in/mg2c_v2.1/) foi usada para mapear os genes PvUGTs de acordo com suas distâncias relativas e posições cromossômicas. As sequências putativas de aminoácidos de UGTs de P. vulgaris L. e A. thaliana foram alinhadas no programa ClustalW, e uma árvore filogenética foi construída usando o método de máxima verossimilhança com 1000 réplicas de bootstrap no software MEGA 7. 0 (KUMAR et al., 2016) utilizando o método de Neighbour-Joining (NJ), com teste de confiabilidade bootstrap de 1000 repetições. O perfil de expressão tecido-específico dos genes UGTs de P. vulgaris L. foram determinados empregando perfis transcricionais obtidos pelos valores de FPKM (do inglês Fragments per kilobase of exon per million fragments mapped) depositados banco de dados no Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Pvulgaris). O perfil transcricional de cada gene alvo e de cada respectiva biblioteca foi expresso e apresentado em figura (Heatmap) com o auxílio da ferramenta *CIMMiner* (http://discover.nci.nih.gov/cimminer).

RESULTADOS E DISCUSSÃO

O processo de glicosilação é catalisado por enzimas específicas como as glicosiltransferases (GTs), que são altamente divergentes, polifiléticas e pertencem a uma família multigênica encontrada em todos os organismos vivos (LAO et al., 2014). Nas plantas, as GTs alteram a atividade biológica de compostos de moléculas pequenas, transferindo doadores de glicosil com receptores e, em seguida, participam e influenciam do crescimento e desenvolvimento das plantas, na formação de metabólitos secundários e na resposta a estresses ambientais (VOGT; JONES, 2000; LE ROY et al., 2016). Neste estudo, foram identificados 68 genes *UGTs* no genoma de *P. vulgaris* L. Em seguida, nomeamos os 68 genes *UGTs* do feijão de *PvUGT1* a *PvUGT68* de acordo com a distribuição dos genes no cromossomo (Tabela 1). O número de membros da família UGT identificados no feijoeiro foi menor quando comparado com outras espécies de plantas como *A. thaliana* (107; LI et al., 2001), *O. sativa* (180; CAPUTI et al., 2012), *Prunus persica* (168; WU et al., 2017), *Citrus grandis* (145; WU et al., 2020) e *Vitis vinifera* (228; WEI et al., 2021).

Tabela 1. Lista dos genes putativos UGTs no genoma de P. vulgaris L. e suas principais características físico-química.

Gene	ID - Genoma P. vulgaris L.	Localização subcelular	(aa) ^a	kDa ^b	pl ^c	GRAVY ^d	Localização subcelular ^e	Homólogo - <i>A. thaliana</i> ^f
PvUGT1	Phvul.001G182100.1	Chr01:4403493344036926	498	56.21	5.89	-0.158	Membrana Celular	AT2G36760.1
PvUGT2	Phvul.001G182200.1	Chr01:4404663644048436	499	56.76	5.83	-0.096	Membrana Celular	AT2G36760.1
PvUGT3	Phvul.001G182300.1	Chr01:4405155944053154	486	54.65	5.34	-0.068	Membrana Celular/Cloroplasto	AT2G36750.1
PvUGT4	Phvul.001G182400.1	Chr01:4405595644057854	489	55.11	6.38	-0.148	Membrana Celular/Cloroplasto	AT2G36750.1
PvUGT5	Phvul.001G182500.1	Chr01:4406031544062011	490	55.23	5.24	-0.104	Membrana Celular	AT2G36800.1
PvUGT6	Phvul.001G182600.1	Chr01:4406553244066998	488	54.71	5.36	-0.019	Membrana Celular	AT2G36760.1
PvUGT7	Phvul.001G182700.1	Chr01:4406974544071211	488	55.09	5.97	-0.161	Membrana Celular	AT2G36760.1
PvUGT8	Phvul.001G182800.1	Chr01:4407506244076933	501	56.34	6.13	-0.192	Membrana Celular	AT2G36780.1
PvUGT9	Phvul.001G182900.1	Chr01:4407929844082051	489	54.98	5.69	-0.103	Membrana Celular/Cloroplasto	AT2G36800.1
PvUGT10	Phvul.001G183000.1	Chr01:4408740544088865	486	54.14	6.86	-0.036	Cloroplasto	AT3G53150.1
PvUGT11	Phvul.001G238800.1	Chr01:4921232049215816	460	51.39	5.32	-0.029	Membrana Celular	AT3G55700.1
PvUGT12	Phvul.001G238900.1	Chr01:4921944749222079	466	51.38	5.73	-0.086	Membrana Celular	AT3G55700.1
PvUGT13	Phvul.002G026900.1	Chr02:28194422821243	481	52.40	5.49	-0.000	Membrana Celular/Cloroplasto	AT5G26310.1
PvUGT14	Phvul.002G096500.1	Chr02:1888149918883234	474	52.46	5.41	0.037	Cloroplasto	AT3G50740.1
PvUGT15	Phvul.002G124600.1	Chr02:2636024026361703	487	54.87	8.09	-0.194	Cloroplasto	AT4G19460.1
PvUGT16	Phvul.003G021900.1	Chr03:20811982082628	476	51.53	5.79	0.101	Membrana Celular/Cloroplasto	AT5G26310.1
PvUGT17	Phvul.003G148300.1	Chr03:3589034635892820	490	54.05	5.60	-0.127	Cloroplasto	AT5G17050.1
PvUGT18	Phvul.003G277200.1	Chr03:5141950351421308	504	57.11	7.06	-0.244	Cloroplasto	AT4G19460.1
PvUGT19	Phvul.004G035300.1	Chr04:41189284121127	451	50.03	6.34	-0.075	Cloroplasto	AT1G22380.1
PvUGT20	Phvul.004G035500.1	Chr04:41362204138119	452	51.31	5.72	-0.130	Membrana Celular/Cloroplasto	AT3G02100.1
PvUGT21	Phvul.004G041000.2	Chr04:48533504854871	384	43.03	5.98	-0.104	Cloroplasto	AT4G01070.1
PvUGT22	Phvul.004G086500.1	Chr04:2814013428142230	460	51.88	5.37	-0.172	Membrana Celular	AT2G43820.1
PvUGT23	Phvul.004G137900.2	Chr04:4372646543728519	480	52.97	5.72	0.047	Membrana Celular	AT3G16520.3
PvUGT24	Phvul.004G138100.1	Chr04:4373962443741007	250	27.67	5.20	0.131	Cloroplasto	AT1G78270.1
PvUGT25	Phvul.004G138200.1	Chr04:4374439843745981	431	47.54	5.26	-0.009	Membrana Celular	AT3G16520.3
PvUGT26	Phvul.004G138400.2	Chr04:4377081343772491	478	52.37	5.79	0.046	Membrana Celular	AT2G29730.1
PvUGT27	Phvul.004G138450.1	Chr04:4378446743786112	429	46.96	5.45	0.1	Membrana Celular	AT2G29730.1
PvUGT28	Phvul.005G005400.1	Chr05:435376441968	356	39.21	5.36	-0.173	Cloroplasto	AT3G16520.3
PvUGT29	Phvul.005G005500.1	Chr05:445357447636	476	52.16	5.72	0.025	Cloroplasto	AT3G16520.3
PvUGT30	Phvul.005G005600.1	Chr05:451528457448	477	52.34	5.72	0.034	Cloroplasto	AT3G16520.3
PvUGT31	Phvul.006G017600.1	Chr06:46906764696021	457	51.52	5.33	-0.211	Membrana Celular/Cloroplasto	AT1G22360.1
PvUGT32	Phvul.006G017500.1	Chr06:48041364806473	482	54.46	5.89	-0.220	Membrana Celular/Cloroplasto	AT1G22360.1
PvUGT33	Phvul.006G017400.1	Chr06:81970948201997	481	54.55	5.37	-0.225	Cloroplasto	AT1G22360.1

PvUGT34	Phvul.006G212950.1	Chr06:3058547430586349	86	97.12	5.64	-0.241	Núcleo	AT5G24750.1
PvUGT35	Phvul.006G212700.1	Chr06:3062508430631089	503	56.50	5.98	0.005	Cloroplasto	AT5G24750.1
PvUGT36	Phvul.007G039200.1	Chr07:32081073209549	480	53.27	5.87	0.030	Membrana Celular	AT4G15480.1
PvUGT37	Phvul.007G039300.1	Chr07:32170853218754	486	53.89	6.11	-0.119	Membrana Celular	AT4G15480.1
PvUGT38	Phvul.007G048100.1	Chr07:39470223952641	447	50.09	5.51	-0.006	Membrana Celular/Cloroplasto	AT2G16890.2
PvUGT39	Phvul.007G152800.1	Chr07:2583744525839378	491	54.89	5.96	-0.214	Membrana Celular	AT2G36770.1
PvUGT40	Phvul.008G026100.1	Chr08:21483872150817	494	55.73	6.27	-0.108	Membrana Celular	AT2G36780.1
PvUGT41	Phvul.008G140800.1	Chr08:2452386524525465	464	51.06	5.76	0.040	Cloroplasto	AT4G01070.1
PvUGT42	Phvul.008G140700.1	Chr08:2455928024560686	468	51.38	5.25	-0.052	Cloroplasto	AT4G01070.1
PvUGT43	Phvul.008G140600.1	Chr08:2458382124585266	481	52.71	6.26	-0.083	Membrana Celular/Cloroplasto/Citoplasma/Mitocôndria/Núcleo	AT4G01070.1
PvUGT44	Phvul.008G140500.1	Chr08:2459604524597703	433	47.33	5.72	0.131	Membrana Celular/Cloroplasto/Citoplasma/Núcleo	AT4G01070.1
PvUGT45	Phvul.008G151900.1	Chr08:2551070925512322	462	52.02	5.44	-0.097	Membrana Celular	AT2G16890.2
PvUGT46	Phvul.008G256600.1	Chr08:6040135260405162	457	51.79	6.68	-0.103	Membrana Celular/Cloroplasto	AT2G30150.2
PvUGT47	Phvul.008G256800.1	Chr08:6041664560425170	462	52.40	6.68	-0.069	Membrana Celular/Cloroplasto	AT2G30140.1
PvUGT48	Phvul.008G257200.1	Chr08:6044683860448796	459	51.01	5.59	-0.115	Cloroplasto	AT2G30150.2
PvUGT49	Phvul.008G262000.1	Chr08:6089173360893139	468	52.14	5.22	0.044	Membrana Celular/Cloroplasto	AT2G29730.1
PvUGT50	Phvul.008G262100.1	Chr08:6089459560896049	484	54.11	6.40	-0.069	Membrana Celular/Cloroplasto	AT2G29730.1
PvUGT51	Phvul.008G290200.1	Chr08:6281350462814895	463	51.32	5.94	-0.041	Membrana Celular	AT3G21780.1
PvUGT52	Phvul.008G290300.1	Chr08:6281764362819034	463	51.19	5.90	-0.009	Membrana Celular/Cloroplasto	AT3G21760.1
PvUGT53	Phvul.008G291000.1	Chr08:6285077762853665	472	53.11	7.56	-0.160	Membrana Celular/Cloroplasto/Núcleo	AT3G22250.1
PvUGT54	Phvul.009G120600.1	Chr09:1840841818410170	480	53.37	5.51	-0.034	Cloroplasto	AT4G01070.1
PvUGT55	Phvul.009G218800.1	Chr09:3307145833075241	467	51.77	5.63	0.002	Membrana Celular/Cloroplasto	AT3G21760.1
PvUGT56	Phvul.010G036900.1	Chr10:54285485429998	437	50.20	6.14	-0.195	Membrana Celular	AT2G22590.1
PvUGT57	Phvul.010G037600.1	Chr10:54893815490706	441	50.21	5.90	-0.175	Membrana Celular	AT2G22590.1
PvUGT58	Phvul.010G042200.1	Chr10:64144556416112	465	51.84	5.80	-0.015	Membrana Celular/Cloroplasto/Citoplasma	AT4G01070.1
PvUGT59	Phvul.010G042300.1	Chr10:64236406425061	473	52.54	6.26	-0.034	Cloroplasto	AT4G01070.1
PvUGT60	Phvul.010G042400.1	Chr10:64386006440018	472	52.49	6.55	-0.020	Cloroplasto/Núcleo	AT4G01070.1
PvUGT61	Phvul.010G042500.1	Chr10:64506726452093	473	52.74	6.40	-0.108	Cloroplasto	AT4G01070.1
PvUGT62	Phvul.010G042600.1	Chr10:64696736471315	478	53.25	7.33	-0.090	Cloroplasto	AT4G01070.1
PvUGT63	Phvul.010G053500.1	Chr10:82198238221642	452	50.38	5.40	0.017	Cloroplasto	AT4G01070.1
PvUGT64	Phvul.010G053700.1	Chr10:82633348265065	482	53.11	5.83	-0.075	Membrana Celular/Cloroplasto	AT4G01070.1
PvUGT65	Phvul.010G071900.1	Chr10:3225775132259884	485	53.78	6.18	-0.029	Cloroplasto	AT2G18570.1
PvUGT66	Phvul.011G151100.1	Chr11:4277612742778032	451	50.26	6.45	-0.104	Membrana Celular/Cloroplasto	AT1G73880.1
PvUGT67	Phvul.011G151200.1	Chr11:4282417142826011	466	51.47	5.98	-0.034	Membrana Celular	AT1G73880.1
PvUGT68	Phvul.011G158500.1	Chr11:4515133645153794	480	54.70	5.57	-0.184	Membrana Celular/Cloroplasto	AT1G22380.1

^aTamanho da proteína; ^bPeso molecular; ^cPonto isoétrico; ^dÍndice de propriedade hidrofílica/hidrofóbica da proteína; ^ePredição da localização subcelular da proteína; [†]Sequências homológas com *A. thaliana* (TAIR: <u>https://www.arabidopsis.org/</u>).

Adicionalmente, podemos ver claramente que embora haja grandes diferenças na quantidade de glicosiltransferases nas diferentes espécies de plantas, a proporção de genes inteiros na planta é diferente também (revisado por SONG et al., 2019). As propriedades físico-químicas de cada gene *PvUGTs* foi calculada com a ferramenta *ExPASy*. O comprimento de aminoácido previstos das proteínas UGTs variou de 86 aa (PvUGT34) a 504 aa (PvUGT18), e o peso molecular previsto da proteína UGT ficou entre 27.67 (PvUGT24) e 97.12 kDa (PvUGT34), e o ponto isoelétrico (pl) variou de 5.20 (PvUGT24) a 8.09 (PvUGT15) (Tabela 1). Embora o gene *PvUGT34* não esteja com sua sequência completa, ao analisa-la no banco de dados *NCBI Conserved Domains*, verificou-se a presença do domínio glicosiltransferase (Tabela 2), por esse motivo a incluímos em algumas análises. Os resultados da localização subcelular demonstraram que 45 e 44 membros PvUGTs foram previstos no cloroplasto e na membrana celular, respectivamente. O restante dos membros PvUGTs foram previstos no núcleo (5), citoplasma (3), mitocôndria (1) (Tabela 1). Essa diversificação sobre a predição subcelular dos membros da família UGTs também é descrita em outros estudos (WU et al., 2017; WU et al., 2020).

A análise da estrutura de um determinado gene é um importante parâmetro, pois fornece informações relevantes sobre a função, organização e evolução. Para caracterizar a diversidade estrutural dos genes neste estudo, as sequências *CDS* e suas sequências de DNA genômico correspondentes foram comparadas. Um diagrama de organização éxon-íntron foi gerado com a ferramenta de bioinformática *GSDS* e as informações quanto ao número de íntrons foram apresentadas na Tabela 2. Nota-se uma variação de zero, um, dois e 14 íntrons entre os genes analisados. Importante mencionar que os íntrons são uma parte importante dos genes, embora não participem da codificação de proteínas, os eventos de ganho ou perda de íntrons e a posição de inserção dos íntrons em relação às sequências de proteínas são geralmente considerados como as principais pistas para entender a evolução ou diversificação de uma família gênica (ROGOZIN et al., 2005).

Tabela 2. Descrição da estrutura dos genes *PvUGTs* e domínio conservado das proteínas.

Gene	Íntron	Descrição - NCBI Conserved Domains search
PvUGT1	0	Glycosyltransferase GTB-type
PvUGT2	0	Glycosyltransferase_GTB-type
PvUGT3	0	Glycosyltransferase_GTB-type
PvUGT4	0	Glycosyltransferase_GTB-type
PvUGT5	0	Glycosyltransferase_GTB-type
PvUGT6	0	Glycosyltransferase_GTB-type
PvUGT7	0	Glycosyltransferase_GTB-type
PvUGT8	0	Glycosyltransferase_GTB-type
PvUGT9	0	Glycosyltransferase_GTB-type
PVUGT10	0	Glycosyltransferase_GIB-type
PVUGT11 DvIIGT12	1	Glycosyltransferase GTB-type
PVUGT12 PvUGT13	0	Glycosyltransferase_GTB-type
PvUGT14	0	Glycosyltransferase GTB-type
PvUGT15	0	Glycosyltransferase GTB-type
PvUGT16	0	Glycosyltransferase_GTB-type
PvUGT17	1	Glycosyltransferase_GTB-type
PvUGT18	0	Glycosyltransferase_GTB-type
PvUGT19	1	Glycosyltransferase_GTB-type
PvUGT20	1	Glycosyltransferase_GTB-type
PvUGT21	0	Glycosyltransferase_GTB-type
PvUGT22	1	Glycosyltransterase_GTB-type
PVUG123	U 2	Glycosyltransferase_GTB type
PVUG124	2	Glycosyltransferase GTB-type
PVLIGT26	0	Glycosyltransferase GTB-type
PvUGT27	0	Glycosyltransferase GTB-type
PvUGT28	1	Glycosyltransferase GTB-type
PvUGT29	1	Glycosyltransferase_GTB-type
PvUGT30	1	Glycosyltransferase_GTB-type
PvUGT31	1	Glycosyltransferase_GTB-type
PvUGT32	1	Glycosyltransferase_GTB-type
PvUGT33	1	Glycosyltransferase_GTB-type
PvUGT34	1	Glycosyltransferase_GTB-type
PvUGT35	14	Glycosyltransferase_GTB-type
PVUG136	0	Glycosyltransferase_GIB-type
PVUGIS7	2	Glycosyltransferase_GTB-type
PvUGT30	0	Glycosyltransferase_GTB-type
PvUGT40	0	Glycosyltransferase GTB-type
PvUGT41	0	Glycosyltransferase_GTB-type
PvUGT42	0	Glycosyltransferase_GTB-type
PvUGT43	0	Glycosyltransferase_GTB-type
PvUGT44	0	Glycosyltransferase_GTB-type
PvUGT45	0	Glycosyltransferase_GTB-type
PvUGT46	1	Glycosyltransferase_GTB-type
PVUGT47	2	Glycosyltransferase_GTB-type
PVUG148	1	Glycosyltransferase GTR-type
PVLIGT50	0	Glycosyltransferase GTR-type
PvUGT51	0	Glycosyltransferase GTB-type
PvUGT52	0	Glycosyltransferase_GTB-type
PvUGT53	1	Glycosyltransferase_GTB-type
PvUGT54	0	Glycosyltransferase_GTB-type
PvUGT55	0	Glycosyltransferase_GTB-type
PvUGT56	0	Glycosyltransferase_GTB-type
PvUGT57	0	Glycosyltransferase_GTB-type
PvUGT58	0	Glycosyltransferase_GTB-type
PvUGT59	0	Glycosyltransterase_GTB-type
PVUG160	U	Glycosyltransferase_GTB-type
PVUGI61	0	Glycosyltransferase_GTB_type
PVUGT62	1	Glycosyltransferase GTR-type
PvUGT64	0	Glycosyltransferase GTB-type
PvUGT65	1	Glycosyltransferase GTB-type
PvUGT66	1	Glycosyltransferase_GTB-type
PvUGT67	0	Glycosyltransferase_GTB-type
PvUGT68	1	Glycosyltransferase GTB-type



O mapeamento dos genes *PvUGTs* de *P. vulgaris* L. foi analisada usando a ferramenta *MG2C*. Com base nas informações de anotação do genoma de *P. vulgaris* L. (ver Tabela 1), os 68 genes foram mapeados e distribuídos entre todos os 10 cromossomos. O cromossomo oito contém a maioria dos genes *PvUGTs* (14 genes) ancorados, enquanto que o cromossomo nove tem o menor número de genes *PvUGTs* (dois genes – *PvUGTs54* e *PvUGT55*) (Figura 3). Uma vez que a distribuição dos genes *PvUGTs* em cada cromossomo foi desigual, e a maioria estava agrupada, isto é um indicativo que há uma duplicação do gene no processo evolutivo para essa classe gênica em *P. vulgaris* L.



Figura 3. Distribuição dos genes *PvUGTs* entre os 10 cromossomos de *P. vulgaris* L. Barras verticais representam os cromossomos de *P. vulgaris* L. O número do cromossomo está no topo de cada cromossomo. A barra de escala à esquerda está representada a escala do comprimento dos cromossomos (Mb).

Os PvUGTs identificados foram submetidos à análise filogenética para ver seu padrão de agrupamento e relações genéticas. Um total de 67 sequências de proteínas UGT de feijão, 19 UGTs de *Arabidopsis* (AtUGT71B1, AtUGT72B1, AtUGT73B1, AtUGT74B1, AtUGT75B1, AtUGT76B1, AtUGT78D1, AtUGT79B1, AtUGT82A1, AtUGT83A1, AtUGT84A1, AtUGT85A1, AtUGT86A1, AtUGT87A1, AtUGT88A1, AtUGT89A2, AtUGT90A1, AtUGT91A1, AtUGT92A1) 3 UGTs de *Z. mays* (GRMZM2G042865, GRMZM2G120016, GRMZM5G834303) foram utilizados para a realização da análise filogenética (Figura 4). Retiramos o PvUGT34 desta análise por não apresentar a sequência completa. As proteínas foram classificadas em 17 subgrupos distintos, A – Q, sendo suas representações descriminadas pelas diferentes cores representação da figura 4. De modo particularizado, as 67 sequências PvUGTs foram divididos em 17 subgrupos distintos e apresentou um grupo externo (*outgroup*), os quais os números de membros nos diferentes subgrupos foram distribuídos desigualmente. O Grupo E é o maior grupo, com 31 membros,



seguido do Grupo D com 12 membros (Figura 4). Nem todos os grupos apresentaram sequências PvUGTs. A exemplo disto podemos mencionar os Grupos M, O, P, Q, respectivamente. Curiosamente, as sequências PvUGT15, PvUGT18 e PvUGT35, formaram um grupo externo, não participando da configuração de nenhum grupo. Isto pode ser um indicativo de que embora as sequências de aminoácidos apresentem o domínio característico dessa família, sua função no metabolismo de feijão pode ser diferente. Estudos adicionais se faz necessário para a confirmação dessa hipótese. Em estudos anteriores, as sequências UGTs de *Arabidopsis* foi usado pela primeira vez para estudar a evolução. Sua árvore filogenética UGT foi dividida em 14 grupos diferentes (A – N) (LI et al. 2001; ROSS, LIM, BOWLES, 2001). Mais tarde, três outros grupos (O, P, Q) foram descobertos durante o estudo de glicosiltransferases de *Z. mays* (LI et al. 2014). Nossa classificação quanto as distribuições dos grupos foram consistentes com os estudos acima mencionado.





Figura 4. Análise filogenética de UGTs de P. vulgaris L.

A análise do padrão de expressão é uma ferramenta extremamente útil para prever as funções biológicas dos genes e permite traçar o perfil de expressão espaço-temporal com base em dados de transcriptoma. Neste estudo, obtivemos os valores (*FPKM*) do *RNAseq* oriundo das seguintes bibliotecas de *P. vulgaris* L.: botão floral, flor, vagem verde madura, folhas, nódulos, raiz 10 (10 dias após o plantio), raiz

18 a 22 de outubro de 2021 Anais do ENEPE ISSN 1677-6321	Unoeste	

19 (19 dias após o plantio), caule 10 (10 dias após o plantio), caule 19 (19 dias após o plantio), vagem jovem e trifólio jovem, os quais estão depositados no banco de dados *Phytozome* (Figura 5).



Figura 5. Perfis de expressão *in silico* dos genes *PvUGTs*. O *heatmap* foi calculado por agrupamento hierárquico de genes *PvUGTs*. A abundância de alto e baixo transcritos dos genes *PvUGTs* é representada pela escala de cores.

Os dados de padrão de expressão in silico, expostos neste estudo, apresenta uma variação quanto a expressão de alguns genes PvUGTs, bem como o tecido o qual apresentou os transcritos (Figura 5). Com relação a ordem dos respectivos genes, os mesmos encontram-se na escala X da figura ilustrativa heatmap, respectivamente: 1:PvUGT32; 2:PvUGT33; 3:PvUGT49; 4:PvUGT5; 5:PvUGT51; 6:PvUGT3; 7:PvUGT22; 8:PvUGT9; 9:PvUGT67; 10:PvUGT4; 11:PvUGT42; 12:PvUGT46; 13:PvUGT61; 14:PvUGT62; 15:PvUGT2; 16:PvUGT39; 17:PvUGT64; 18:PvUGT23; 19:PvUGT30; 20:PvUGT14; 21:PvUGT56; 22:PvUGT11; 23:PvUGT26; 24:PvUGT28; 25:PvUGT47; 26:PvUGT44; 27:PvUGT17; 28:PvUGT45; 29:PvUGT52; 30:PvUGT66; 31:PvUGT25; 32:PvUGT31; 33:PvUGT57; 34:PvUGT68; 35:PvUGT27; 36:PvUGT53; 37:PvUGT59; 38:PvUGT50; 39:PvUGT20; 40:PvUGT21; 41:PvUGT34; 42:PvUGT18; 43:PvUGT35; 44:PvUGT16; 45:PvUGT24; 46:PvUGT29; 47:PvUGT10; 48:PvUGT38; 49:PvUGT6; 50:PvUGT63: 51:PvUGT15; 52:PvUGT54; 53:PvUGT60; 54:PvUGT8; 55:PvUGT40; 56:PvUGT7; 57:PvUGT12; 58:PvUGT43; 59:PvUGT19; 60:PvUGT36; 61:PvUGT48; 62:PvUGT13; 63:PvUGT65; 64:PvUGT58; 65:PvUGT41; 66:PvUGT55; 67:PvUGT1; 68:PvUGT37. Nota-se, por exemplo, que alguns genes foram altamente responsivos nas raizes 10 e 19. Ademais, outros genes apresentaram alta expressão na flor. Dessa forma, a abordagem da caracterização in silico nos diferentes tecidos presentes nesse estudo, estabelece um acervo científico para o entendimento das vias onde a UGTs podem ser atuantes.

CONCLUSÃO

Neste estudo, 68 genes putativos da família UGTs em *P. vulgaris* L. foram identificados e empregouse várias ferramentas da bioinformática para realizar uma caracterização mais apurada. No geral, este estudo analisou sistematicamente a família do gene UGTs em *P. vulgaris* L. e pela primeira vez este estudo fornece algumas informações valiosas para estudos funcionais adicionais dos genes UGTs no futuro.





REFERÊNCIAS

ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. Gapped BLAST and PSIBLAST: a new generation of protein database search programs. Nucleic Acids Research, v.25, p.3389-3402, 1997. <u>https://doi.org/10.1093/nar/25.17.3389</u>

BOWLES. D.; ISAYENKOVA, J.; LIM, EK.; et al. Glycosyltransferases: managers of small molecules. Curr Opin Plant Biol. 8, 254–263, 2005. <u>https://doi.org/10.1016/j.pbi.2005.03.007</u>

CAPUTI, L.; MALNOY, M.; GOREMYKIN, V.; NIKIFOROVA, S.; MARTENS, S. A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. Plant J. 69(6), 1030–42, 2012. <u>https://doi.org/10.1111/j.1365-313X.2011.04853.x</u>

GOODSTEIN, D.M.; SHU, S.; HOWSON, R.; NEUPANE, R.; HAYES, R.D.; FAZO, J.; ROKHSAR, D.S. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res., v.40, n.D1, p.D1178-D1186, 2012. https://doi.org/10.1093/nar/gkr944

HE, Y.; AHMAD, D.; ZHANG, X.; ZHANG, Y.; WU, L.; JIANG, P.; et al. Genome-wide analysis of family-1 UDP glycosyltransferases (UGT) and identification of UGT genes for FHB resistance in wheat (*Triticum aestivum* L.). BMC Plant Biol. 18(1), 67, 2018. <u>https://doi.org/10.1186/s12870-018-1286-5</u>

LE ROY, J.; HUSS, B.; CREACH, A.; HAWKINS, S.; NEUTELINGS, G. Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. Front. Plant Sci. 7, 735, 2016. https://doi.org/10.3389/fpls.2016.00735

LI, Y.; BALDAUF, S.; LIM, EK.; BOWLES, DJ. Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. J Biol Chem 276:4338–4343, 2001. <u>https://doi.org/10.1074/jbc.M007447200</u>

LI, Y.; LI, P.; WANG, Y.; DONG, R.; YU, H.; HOU, B. Genome-wide identification and phylogenetic analysis of Family-1 UDP glycosyltransferases in maize (*Zea mays*). Planta. 239, 265–1279, 2014. https://doi.org/10.1007/s00425-014-2050-1

MEECH, R.; HU, D.G.; MCKINNON, R.A. The UDP-glycosyltransferase (UGT) superfamily: New members, new functions, and novel paradigms. Physiol. Rev. 99, 1153–1222, 2019. https://doi.org/10.1152/physrev.00058.2017.

ROGOZIN, IB.; LYONSWEILER, J.; KOONIN, EV. Intron sliding in conserved gene families. Trends Genet. 16, 430–2, 2000. <u>https://doi.org/10.1016/S0168-9525(00)02096-5</u>

ROSS, J.; LI, Y.; LIM, E.; BOWLES, DJ. Higher plant glycosyltransferases. Genome Biol 2, REVIEWS3004, 2001. https://doi.org/10.1186/gb-2001-2-2-reviews3004

SONG, Z.; NIU, L.; YANG, Q.; et al. Genome-wide identification and characterization of UGT family in pigeonpea (*Cajanus cajan*) and expression analysis in abiotic stress. Trees 33, 987–1002, 2019. https://doi.org/10.1007/s00468-019-01833-6

VOGT, T.; AND JONES, P. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci. 5, 380–386, 2000. <u>https://doi.org/10.1016/S1360-1385(00)01720-9</u>

WEI, Y.; MU, H.; XU, G.; WANG, Y.; LI, Y.; LI, S.; & WANG, L. Genome-Wide Analysis and Functional Characterization of the UDP-Glycosyltransferase Family in Grapes. *Horticulturae*, 7(8), 204, 2021. https://doi.org/10.3390/horticulturae7080204

203



YANG, C.; LI, C.; WEI, W.; et al. The unprecedented diversity of UGT94-family UDP-glycosyltransferases in Panax plants and their contribution to ginsenoside biosynthesis. Sci Rep 10, 15394, 2020. https://doi.org/10.1038/s41598-020-72278-y

YONEKURA-SAKAKIBARA, K.; AND HANADA, K. An evolutionary view of functional diversity in family 1 glycosyltransferases. Plant J. 66, 182–193, 2011. <u>https://doi.org/10.1111/j.1365-313X.2011.04493.x</u>

GASTEIGER, E.; HOOGLAND, C.; GATTIKER, A.; DUVAUD, S.; WILKINS, M.R.; APPEL, R.D.; BAIROCH, A. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The Proteomics protocols handbook. Humana Press, p.571–607, 2005. <u>https://doi.org/10.1385/1-59259-890-0:571</u>

HU, B.; JIN, J.; GUO, A-Y.; ZHANG, H.; LUO, J.G. GGSDS 2.0: an upgraded gene feature visualization server. Bioinformatics, v.31, n.8, p.1296-1297, 2015. <u>https://doi.org/10.1093/bioinformatics/btu817</u>

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: molecular evolutionary genetics analysis version 7.0 forbiggerdatasets.MolecularBiologyEvolution,v.33,n.7,p.1870-1874,2016.https://doi.org/10.1093/molbev/msw054

Wu BP, Gao LX, Gao J, Xu YY, Liu HR, Cao XM, Zhang B, Chen KS. Genome-wide identification, expression patterns, and functional analysis of UDP glycosyltransferase family in peach (Prunus persica L. Batsch). Front Plant Sci. 2017;8:389. <u>https://doi.org/10.3389/fpls.2017.00389</u>

Wu, B.; Liu, X.; Xu, K. Genome-wide characterization, evolution and expression profiling of UDP-glycosyltransferase family in pomelo (Citrus grandis) fruit. BMC Plant Biol. 2020, 20, 459. https://doi.org/10.1186/s12870-020-02655-2

Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV: Analysis of evolution of exon-intron structure ofeukaryoticgenes.BriefBioinform2005,6(2):118-134.10.1093/bib/6.2.118.https://doi.org/10.1093/bib/6.2.118

Chou KC, Shen HB: Plant-mPLoc: A top-down strategy to augment the power for predicting plant protein subcellular localization. PLoS ONE 2010, 5: e11335. 10.1371/journal.pone.0011335. https://doi.org/10.1371/journal.pone.0011335



RESUMOS DE PESQUISA

INFECÇÃO POR CÂNDIDA ALBICANS NO CÉREBRO DE DROSOPHILA MELANOGASTER LEVA A FORMAÇÃO) DE
PLACAS BETA AMILOIDES NA DOENÇA DE ALZHEIMER	. 206

Pesquisa (ENAPI)

Comunicação oral on-line

UNIVERSIDADE DO OESTE PAULISTA - UNOESTE Ciências Biológicas Genética

INFECÇÃO POR CÂNDIDA ALBICANS NO CÉREBRO DE DROSOPHILA MELANOGASTER LEVA A FORMAÇÃO DE PLACAS BETA AMILOIDES NA DOENÇA DE ALZHEIMER.

MARIELLE AMARAL GOMES DA CRUZ VICTOR CAMPOS MAYARA DE OLIVEIRA VIDOTTO FIGUEIREDO JOYCE MARINHO DE SOUZA CARLOS ANTONIO COUTO LIMA

A doença de Alzheimer (DA) é caracterizada como uma patologia neurodegenerativa progressiva e irreversível, devido ao acumulo de peptídeos Beta amilóide derivado da proteína percursora amilóide (APP) através de clivagens proteolíticas por B- e Y- secretases. Necropsias de pacientes acometidos pela DA indicaram a presença de fungos levando a hipótese que a neurodegeneração possa ter sido causada pela inflamação fúngica, além disso a Cândida albicans foi encontrada em 89,6% do soro de pacientes com DA. Avaliar o efeito da infecção de C.albicans no cérebro de drosophilas e o efeito na formação de placas beta amilóide. Utilizamos alfinete entomológico mergulhado no inóculo de C. albicans e, introduzido no cérebro das moscas anestesiadas através do olho esquerdo. A reprodutibilidade foi monitorada pelo aparecimento de um ponto de melanização no local da injeção, para controle negativo os animais sofreram apenas a injuria tecidual. Com isso monitoramos a motilidade utilizando o teste de escalada; a técnica de RTqPCR para avaliar a expressão de componentes do sistema imunológico de insetos; ensaio de sobrevivência e por fim a marcação com Tioflavina-S das placas Beta-amilóides. Durante o estudo de sobrevida notamos a injeção cause letalidade, entretanto animais infectados apresentam letalidade aumentada e motilidade reduzida, indicando um comprometimento neuromotor acentuado. Moscas infectados aumentaram a expressão de mRNA de Drosomicina, CecropinA1, AttacinC, Diptericina B, Metchnikowin, principais genes relacionados a resposta imune, possivelmente em resposta a presença de C. albicans. Surpreendentemente, a marcação com Tioflavina-S demostrou um aumento na marcação de placas Beta-amilóides tempo dependente. Ou seja animais mais velhos infectados apresentavam uma maior marcação. Embora não avaliamos os níveis de colonização no cérebro, demostramos o início do processo inflamatório oriundo da presença do fungo que originou as alterações apresentadas com a presença das placas maneira generalizada nestes animais Concluímos que a infecção da Cândida está associada ao aumento de agregados de placas Beta-amilóide no cérebro de Drosophilas e a possibilidade de avaliar mecanismos subjacentes na correlação entre fungos e a DA. Órgão de fomento financiador da pesquisa: Universidade do Oeste Paulista