

федеральное государственное бюджетное образовательное учреждение
высшего образования
"Санкт-Петербургский государственный университет" (СПбГУ)

БАЗА ДАННЫХ:

**«База данных для экспресс-выявления участвующих в синтезе
флавоноидов генов» (FLVs)**

Листов 6

Авторы:

**Зыкин Павел Александрович
Андреева Елена Александровна
Цветкова Наталья Владимировна**

Санкт-Петербург

2022

«База данных для экспресс-выявления участвующих в синтезе флавоноидов генов» (FLVs)

Материалы, идентифицирующие базу данных

База данных содержит 1 таблицу, состоящую из 2х частей: «Genes», «Pathways».

Таблица «Genes» содержит набор данных об экспрессии генов, полученный в результате научного исследования, посвященного изучению синтеза флавоноидных соединений в линиях ржи с окрашенным и неокрашенным зерном. Всего 12 полей, 49596 записей.

Таблица «Pathways» содержит набор данных о группах генов из базы GeneOntology, обогащенных среди дифференциально экспрессирующихся генов, выявленных в результате научного исследования. Всего 7 полей, 133 записи.

Общая структура базы данных

База данных FLVs	
Таблица Genes	Таблица Pathways
<ol style="list-style-type: none"> 1. gene — идентификатор гена 2. transcript — порядковый номер транскрипта 3. short_name — название белкового продукта гена 4. description — полное описание гена в соответствии с БД SwissProt 5. CDS — нуклеотидная последовательность белок-кодирующего участка 6. AA — аминокислотная последовательность белкового продукта 7. ya1 — уровень экспрессии соответствующего гена в зерновках неокрашенной линии 7, в транскриптах гена на миллион всех транскриптов (TPM) 8. ya2 — уровень экспрессии соответствующего гена в зерновках неокрашенной линии vi4, TPM 9. ya3 — уровень экспрессии соответствующего гена в зерновках неокрашенной линии Ф-56 ржи, TPM 10. ya4 — уровень экспрессии соответствующего гена в зерновках неокрашенной линии ржи RMu12/2020, TPM 11. ya5 — уровень экспрессии соответствующего гена в зерновках неокрашенной линии RMu12/2021, TPM 12. pathways — метаболический путь, в котором участвует данный ген 	<ol style="list-style-type: none"> 1. ya1 — кластеризация по общему количеству транскриптов, относящимся к транскрипционным факторам, регулирующих синтез флавоноидов, в зерновках неокрашенной линии 7 в транскриптах гена на миллион всех транскриптов (TPM) 2. ya2 — кластеризация по общему количеству транскриптов, относящимся к транскрипционным факторам, регулирующих синтез флавоноидов, в зерновках неокрашенной линии vi4, TPM 3. ya3 — кластеризация по общему количеству транскриптов, относящимся к транскрипционным факторам, регулирующих синтез флавоноидов, в зерновках неокрашенной линии Ф-56 ржи, TPM 4. ya4 — кластеризация по общему количеству транскриптов, относящимся к транскрипционным факторам, регулирующих синтез флавоноидов, в зерновках неокрашенной линии ржи RMu12/2020, TPM 5. ya5 — кластеризация по общему количеству транскриптов, относящимся к транскрипционным факторам, регулирующих синтез флавоноидов, в зерновках неокрашенной линии RMu12/2021, TPM 6. chr — локализация на хромосоме 7. function — функция белкового продукта

Таблицы или их фрагменты, по размеру достаточные для идентификации базы данных.

Таблица Genes

A	B	C	D	E	F	G	H	I	J	K	L
gene	transcript	Short name	Descriptor	QDS	AA	ya1	ya2	ya3	ya4	ya5	pathways
NODE_16076_length	NODE_16076	E3 ubiquitin-protein ligase RF12	ARAT1	CGTGAACCA	REPTKRTSA	3.9	1.3	4.2	4.0	1.1	E3 ubiquitin-protein ligase RF12
NODE_16076_length	NODE_16076	E3 ubiquitin-protein ligase RF12	ARAT1	CGTGAACCA	MSPRAISR	3.9	1.3	4.2	4.0	1.1	E3 ubiquitin-protein ligase RF12
NODE_250975_length	NODE_25097	UDP-glucose flavonoid 3-O-gluc	UF07_FR	TCGGCCGC	MLVSGVAR	0.0	0.0	0.0	0.0	0.1	UDP-glucose flavonoid 3-O-gluc
SECCE1Rv1G0000220.1	NODE_1017	Flavonol 3-sulfotransferase	F3ST_FLAC1	CTTTGTGA	MAAHPSVSI	0.0	0.0	0.0	0.0	0.5	Flavonol 3-sulfotransferase
SECCE1Rv1G0000220.1	NODE_1018	Flavonol 3-sulfotransferase	F3ST_FLAC1	CTTTGTGA	MAAHPSVSI	0.0	0.0	0.0	0.0	1.1	Flavonol 3-sulfotransferase
SECCE1Rv1G0000310.1	NODE_1266	Malonyl-coenzyme A:anthocya	3MAT_DAH	TTAACATT	MDPPQLCV	0.0	0.0	0.0	0.0	0.0	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000310.1	NODE_1266	Malonyl-coenzyme A:anthocya	3MAT_DAH	TTAACATT	MDPPQLCV	0.0	0.0	0.0	0.0	1.4	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000310.1	NODE_1277	Malonyl-coenzyme A:anthocya	3MAT_DAH	TTAACATT	MDPPQLCV	0.0	0.0	0.0	0.0	0.0	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000310.1	NODE_1444	Malonyl-coenzyme A:anthocya	3MAT_DAH	CTGCCGAT	MDPPLHLCV	2.7	0.0	0.0	0.0	0.0	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000310.1	NODE_1459	Malonyl-coenzyme A:anthocya	3MAT_DAH	CCATGTGAT	MDPPQLCV	0.0	0.0	0.0	0.0	0.0	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000310.1	NODE_1459	Malonyl-coenzyme A:anthocya	3MAT_DAH	CCATGTGAT	MDPPQLCV	0.0	0.0	0.0	0.0	0.0	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000310.1	NODE_1567	Malonyl-coenzyme A:anthocya	3MAT_DAH	AGTTGAGC	MDPPQLCV	0.0	0.0	0.3	0.4	0.6	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000310.1	NODE_1647	Malonyl-coenzyme A:anthocya	3MAT_DAH	GCCGAGAT	PETSLPLTF	0.0	0.0	0.0	0.2	0.4	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000310.1	NODE_3132	Malonyl-coenzyme A:anthocya	3MAT_DAH	GCAGCTTG	LCVLENVAV	0.0	4.7	0.0	0.0	0.1	Malonyl-coenzyme A:anthocya
SECCE1Rv1G0000440.1	NODE_1046	Serine carboxypeptidase-like 19	SCP19_ARA	CGGCCCAT	MHILEPRCD	2.1	0.0	0.0	0.0	0.0	Serine carboxypeptidase-like 19
SECCE1Rv1G0000440.1	NODE_2384	Putative serine carboxypeptid	SCP52_ARA	CCATTCCT	SVQSIYLIK	0.1	0.0	0.0	0.0	0.0	Putative serine carboxypeptid
SECCE1Rv1G0000440.1	NODE_2384	Putative serine carboxypeptid	SCP52_ARA	CCATTCCT	SVQSIYLIK	0.1	0.0	0.0	0.0	0.4	Putative serine carboxypeptid
SECCE1Rv1G0000440.1	NODE_3280	Putative serine carboxypeptid	SCP52_ARA	TGCCGCTG	YHFSLITKV	0.0	10.6	0.0	0.0	0.3	Putative serine carboxypeptid
SECCE1Rv1G0000450.1	NODE_2069	Anthocyanidin 3-O-glucosyltr	UF0G_HOR	TGTGAACCA	MAAAASRA	0.1	0.0	5.9	0.2	0.5	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000450.1	NODE_2069	Anthocyanidin 3-O-glucosyltr	UF0G_HOR	TGTGAACCA	MAAAASRA	0.1	0.0	5.9	0.2	0.5	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000450.1	NODE_2315	Anthocyanidin 3-O-glucosyltr	UF0G_HOR	GCTCCGGG	LPAGFDLE	0.0	0.0	1.2	0.1	0.0	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000450.1	NODE_2542	Anthocyanidin 3-O-glucosyltr	UF0G_HOR	CCTGACCG	APHVVITF	1.3	0.0	26.6	12.0	0.6	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000470.1	NODE_2170	Anthocyanidin 3-O-glucosyltr	UF0G_HOR	TACAAATG	MQCPLSPN	0.0	0.0	0.0	1.2	0.0	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000470.1	NODE_2170	Anthocyanidin 3-O-glucosyltr	UF0G_HOR	TACAAATG	MARWIMSS	0.0	0.0	0.0	1.2	0.0	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000470.1	NODE_2290	Anthocyanidin 3-O-glucosyltr	UF0G_HOR	TACAAATG	MLEEYLVV	0.0	0.0	1.2	13.6	0.6	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000470.1	NODE_2503	Anthocyanidin 3-O-glucosyltr	UF0G1_MAR	GGAGCAAG	ARAVAYAF	0.0	0.0	5.2	0.0	0.1	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000470.1	NODE_2503	Anthocyanidin 3-O-glucosyltr	UF0G1_MAR	CGCCCGCG	ARAVAYAF	0.0	0.0	0.0	3.9	0.0	Anthocyanidin 3-O-glucosyltr
SECCE1Rv1G0000470.1	NODE_3345	Anthocyanidin 3-O-glucosyltr	UF0G_HOR	GCGGCCGC	MLATTMEL	0.0	0.0	0.0	0.1	0.0	Anthocyanidin 3-O-glucosyltr

Перечень полей таблицы “Genes”

Название поля	Тип поля
gene	Текстовый
transcript	Текстовый
Short name	Текстовый
Description	Текстовый
CDS	Текстовый
AA	Текстовый
ya1	Числовой
ya2	Числовой
ya3	Числовой
ya4	Числовой
ya5	Числовой
pathways	Текстовый

Перечень полей таблицы “Pathways”

Название поля	Тип поля
ya1	Числовой
ya2	Числовой
ya3	Числовой
ya4	Числовой
ya5	Числовой
Chr	Текстовый