

ПРИЛОЖЕНИЕ

к статье М.А. Дук, В.А. Станина, А.А. Канапина, А.А. Самсоновой, Т.А. Рожминой, М.Г. Самсоновой
«Транспозоны льна: роль в генетическом разнообразии, окультуривании и детерминации хозяйственно ценных признаков»

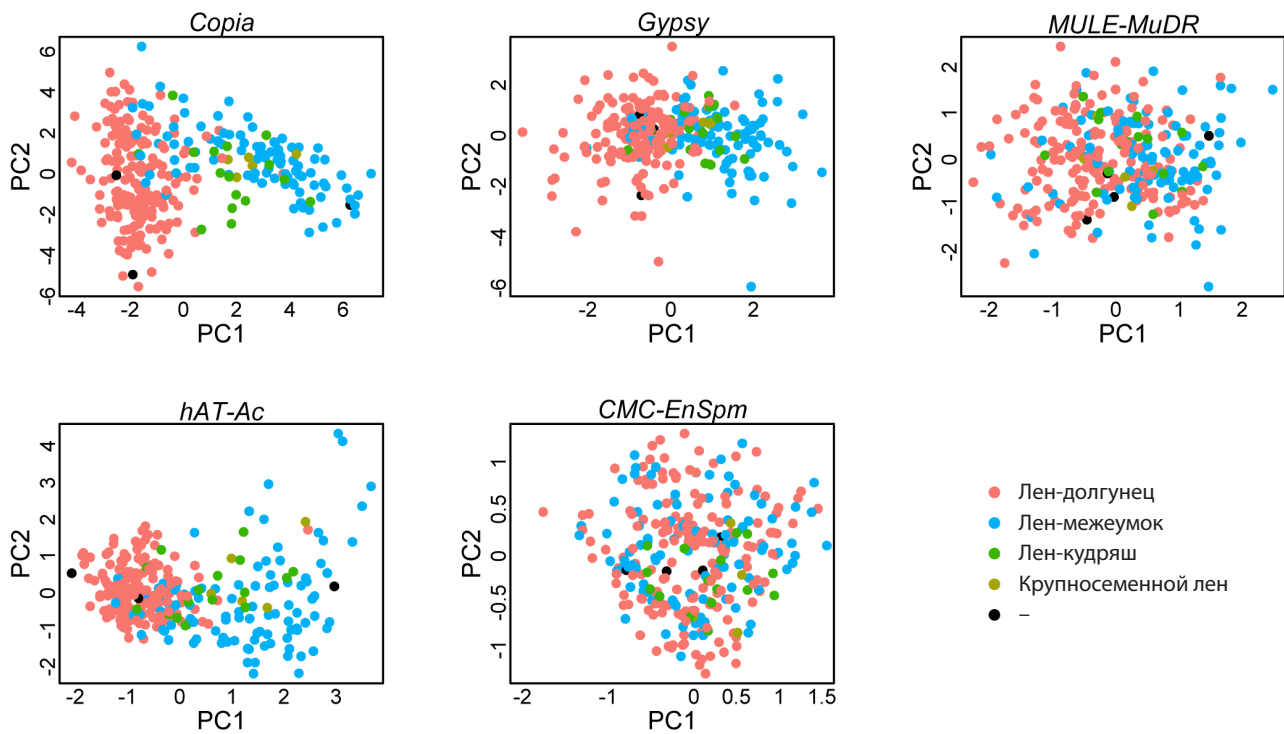


Рис. S1. Графики первых двух главных компонент для инсерций TE суперсемейств *Copia*, *Gypsy*, *MULE-MuDR*, *CMC-EnSpm* и семейства *hAT-Ac*.
Образцы раскрашены по морфотипу.

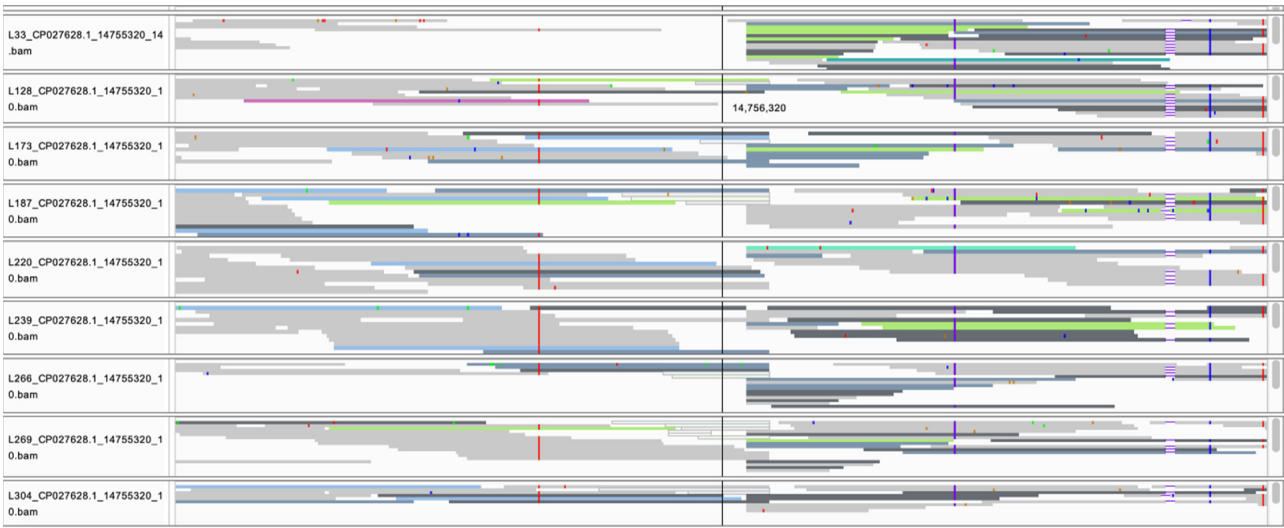


Рис. S2. Подтверждение вставки транспозона *hAT-Ac* в позиции 14756320 Chr4_.
Результат показан для девяти образцов. Штриховая вертикальная линия обозначает место вставки.



Рис. S3. Подтверждение вставки транспозона *Sorja* в позиции 3391752 Chr13.
Результат показан для шести образцов. Вертикальная линия обозначает место вставки.

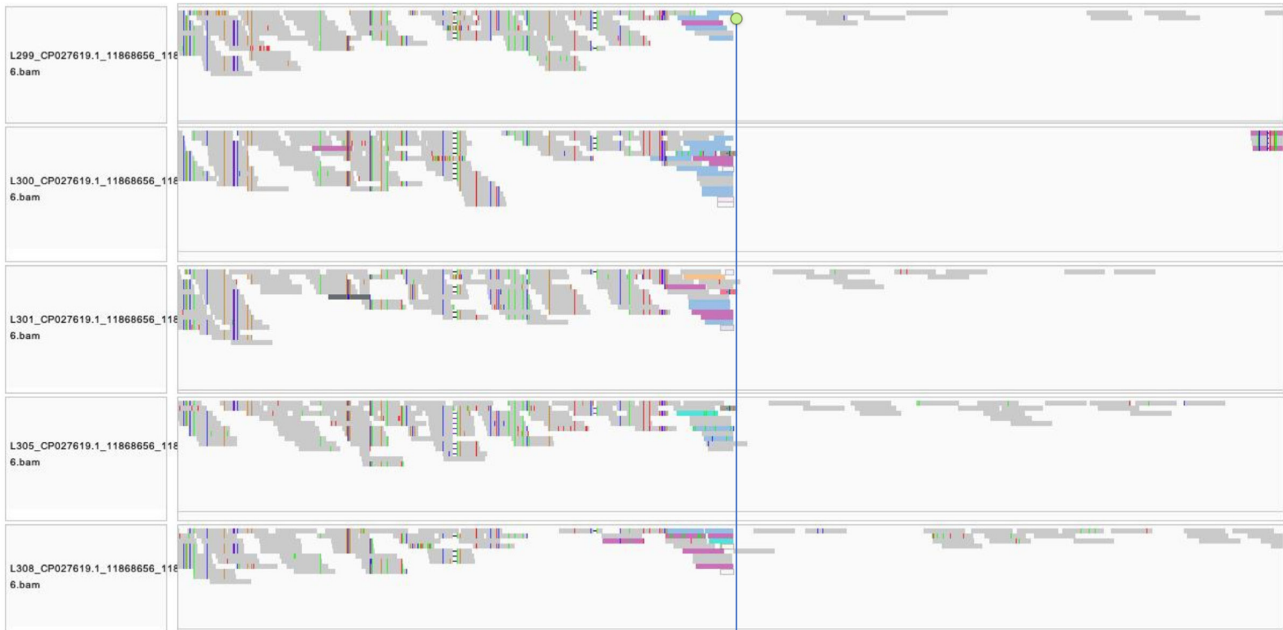


Рис. S4. Подтверждение вставки транспозона *hAT-Tag1* в позиции 11870656 Chr1.
Результат показан для пяти образцов. Вертикальная линия обозначает место вставки.

Таблица S1. Доля транспозонов, расположенных на разном расстоянии относительно известных генов

| Суперсемейство | 2000 п. н. перед геном | Экзон | Интрон | 2000 п. н. после гена | Далеко от генов | Всего в коллекции |
|----------------------------------|---------------------------|---------|--------|--------------------------|--------------------|----------------------|
| Все TE | 0.12 | 0.026 | 0.063 | 0.12 | 0.67 | 588 480 |
| <i>Copia</i> | 0.1 | 0.038 | 0.088 | 0.09 | 0.68 | 241 412 |
| <i>Gypsy</i> | 0.054 | 0.0035 | 0.029 | 0.059 | 0.85 | 85 774 |
| <i>MULE-MuDR</i> | 0.17 | 0.018 | 0.043 | 0.17 | 0.6 | 36 253 |
| <i>hAT-Ac</i> | 0.18 | 0.035 | 0.055 | 0.25 | 0.48 | 20 475 |
| DNA (неуточненной классификации) | 0.25 | 0.0028 | 0.056 | 0.29 | 0.4 | 14 945 |
| <i>L1</i> | 0.22 | 0.015 | 0.046 | 0.21 | 0.51 | 5 724 |
| <i>CMC-EnSpm</i> | 0.1 | 0.00038 | 0.003 | 0.16 | 0.74 | 5 216 |
| <i>hAT-Tag1</i> | 0.25 | 0.013 | 0.1 | 0.2 | 0.43 | 4 758 |
| <i>L2</i> | 0.096 | 0.0053 | 0.027 | 0.34 | 0.53 | 401 |
| <i>Satellite</i> | 0.016 | 0 | 0.0054 | 0.07 | 0.91 | 264 |
| <i>rRNA</i> | 0.15 | 0.011 | 0.045 | 0.51 | 0.28 | 217 |
| <i>Helitron</i> | 0.57 | 0.14 | 0 | 0 | 0.29 | 34 |
| <i>hAT-Tip100</i> | 0 | 0 | 0 | 0.21 | 0.79 | 22 |
| Неклассифицированные | 0.15 | 0.027 | 0.058 | 0.17 | 0.6 | 172 984 |

Таблица S2. Список транспозонов, которыми обогащены клады I–III (обогащенными в клade считали TE, частота появления которых в одной из клад более чем в два раза превышала частоты появления в каждой из двух других клад)

| Идентификатор последовательности TE | Суперсемейство | Число сайтов инсерции | Доля вставок на один образец | | | Характерен для клад |
|--|----------------------------------|--------------------------|------------------------------|----------|-----------|------------------------|
| | | | Клада I | Клада II | Клада III | |
| LU_v2-B-R24968-Map6 | <i>Helitron</i> | 26 | 0.22 | 0.068 | 0.033 | Клада I |
| LU_v2-B-R5502-Map4 | <i>Gypsy</i> | 16 | 0.12 | 0.023 | 0.04 | Клада I |
| LU_v2-B-R19931-Map3 | <i>Copia</i> | 17 | 0.13 | 0.034 | 0.033 | Клада I |
| LU_v2-B-G1344-Map3 | <i>Copia</i> | 33 | 0.22 | 0.057 | 0.087 | Клада I |
| LU_v2-B-R3227-Map20 | <i>Copia</i> | 21 | 0.13 | 0.034 | 0.06 | Клада I |
| LU_v2-B-R5856-Map20 | <i>Gypsy</i> | 39 | 0.25 | 0.068 | 0.11 | Клада I |
| LU_v2-B-R2691-Map3_reversed | <i>Copia</i> | 58 | 0.37 | 0.18 | 0.11 | Клада I |
| LU_v2-B-R17883-Map3_reversed | <i>Copia</i> | 19 | 0.12 | 0.057 | 0.04 | Клада I |
| LU_v2-B-R3765-Map4 | <i>hAT-Ac</i> | 29 | 0.18 | 0.08 | 0.067 | Клада I |
| LU_v2-B-G1822-Map20 | DNA (неуточненной классификации) | 87 | 0.49 | 0.24 | 0.22 | Клада I |
| LU_v2-B-R3265-Map5_reversed | <i>Copia</i> | 72 | 0.16 | 0.4 | 0.17 | Клада II |
| LU_v2-B-R3814-Map20 | <i>Copia</i> | 17 | 0.044 | 0.091 | 0.04 | Клада II |
| LU_v2-B-G3342-Map7_reversed | <i>Copia</i> | 21 | 0.015 | 0.011 | 0.13 | Клада III |
| LU_v2-B-R3762-Map5 | <i>Copia</i> | 54 | 0.13 | 0.034 | 0.28 | Клада III |
| LU_v2-B-R3967-Map17 | <i>Copia</i> | 73 | 0.16 | 0.12 | 0.34 | Клада III |

Таблица S3. Районы генома льна со следами селекции у групп образцов

| Положение участка | Уменьшение разнообразия* | QTL ** |
|-------------------------|-----------------------------|--|
| Chr13:16200001-16400000 | Д/М, М/С | |
| Chr4:16600001-16800000 | М/Д, М/С | |
| Chr1:28600001-28800000 | Д/К, Д/С | |
| Chr12:1000001-1200000 | Д/К, Д/С | uq.C12-1 (PLH), uq.C12-1 (TL), QIOD-Lu12.3 (IOD), QLIN-Lu12.3 (LIN), QLIO-Lu12.3 (LIO) |
| Chr6:17400001-17600000 | К/Д, С/Д | |
| Chr7:3800001-4000000 | К/Д, К/С | QPM-crc-LG7 (PM), QPAL-Lu7.3 (PAL) |
| Chr4:9800001-10000000 | Д/С, М/С | |
| Chr9:21000001-21200000 | Д/С, К/С | |
| Chr15:13400001-13600000 | Д/С, К/С | |
| Chr2:24800001-25000000 | Д/М | |
| Chr5:3200001-3400000 | | |
| Chr6:13800001-14000000 | | QOIL-Lu6.4 (OIL) |
| Chr8:2800001-3000000 | | |
| Chr10:200001-400000 | | |
| Chr11:6400001-6600000 | | |
| Chr8:16800001-17000000 | М/Д | |
| Chr14:18800001-19000000 | | |
| Chr3:17600001-17800000 | Д/К | Lu3-25559600 (MC) |
| Chr8:21600001-21800000 | | QOLE-Lu8.1 (OLE) |
| Chr9:4600001-4800000 | | |
| Chr9:20400001-20600000 | | QSTE-Lu9.2 (STE) |
| Chr12:2400001-2600000 | | QIOD-Lu12.3 (IOD), QLIN-Lu12.3 (LIN), QLIO-Lu12.3 (LIO) |
| Chr13:5000001-5200000 | | |
| Chr13:5400001-5600000 | | |
| Chr14:1200001-1400000 | | |
| Chr1:400001-600000 | К/Д | |
| Chr6:3400001-3600000 | | |
| Chr6:15800001-16000000 | | |
| Chr9:19600001-19800000 | | |
| Chr15:10000001-10200000 | | |
| Chr2:1200001-1400000 | Д/С | |
| Chr6:13600001-13800000 | | Lu2564 (LDG), QOIL-Lu6.4 (OIL) |
| Chr8:22000001-22200000 | | QOLE-Lu8.1 (OLE) |
| Chr9:21200001-21400000 | | |
| Chr11:1400001-1600000 | | |
| Chr12:400001-600000 | | QIOD-Lu12.3 (IOD), QLIN-Lu12.3 (LIN), QLIO-Lu12.3 (LIO) |
| Chr13:4400001-4600000 | | |
| Chr14:4200001-4400000 | | |
| Chr14:17800001-18000000 | | |
| Chr14:18600001-18800000 | | |
| Chr1:1-200000 | С/Д | |

Окончание табл. S3

| Положение участка | Уменьшение разнообразия* | QTL** |
|-------------------------|--------------------------|---|
| Chr2:21800001-22000000 | K/C | scaffold43-1111162 (TSW), QOIL-Lu2.1 (OIL) |
| Chr4:11600001-11800000 | | |
| Chr6:17000001-17200000 | | |
| Chr10:1-200000 | | |
| Chr11:15800001-16000000 | | |
| Chr12:20600001-20800000 | C/K | QLio-LG3.1 (LIO), QLin-LG3.1 (LIN), Marker4371 (PLH), scaffold156-761294 (FN) |
| Chr3:6000001-6200000 | | |
| Chr5:4200001-4400000 | | |
| Chr11:5200001-5400000 | M/C | scaffold635-43971 (NB), QOLE-Lu8.1 (OLE) |
| Chr1:8800001-9000000 | | |
| Chr3:24400001-24600000 | | |
| Chr4:12400001-12600000 | | |
| Chr8:5600001-5800000 | | |
| Chr8:22400001-22600000 | | |
| Chr12:2600001-2800000 | | |
| Chr13:5800001-6000000 | | |
| Chr14:13600001-13800000 | C/M | QIOD-Lu12.3 (IOD), QLin-Lu12.3 (LIN), QLIO-Lu12.3 (LIO) |
| Chr1:6600001-6800000 | | |
| Chr7:3400001-3600000 | | |
| Chr7:16200001-16400000 | | QPAL-Lu7.3 (PAL) |
| | | QIOD-Lu7.2 (IOD), QLin-Lu7.2 (LIN), QLIO-Lu7.2 (LIO) |

* Д/М – уменьшение разнообразия у льна-долгунца в сравнении с масличным льном; М/Д – уменьшение разнообразия у масличного льна в сравнении со льном-долгунцом; Д/С, М/С – уменьшение разнообразия у культурных сортов в сравнении со староместными сортами у льна-долгунца и масличного льна соответственно; С/Д, С/М – уменьшение разнообразия у староместных сортов в сравнении с культурными сортами у льна-долгунца и у масличного льна соответственно; Д/К – уменьшение разнообразия у культурных сортов льна-долгунца в сравнении с кряжами; К/Д – уменьшение разнообразия у кряжей в сравнении с культурными сортами льна-долгунца; К/С – уменьшение разнообразия у кряжей в сравнении со староместными сортами льна-долгунца; С/К – уменьшение разнообразия у староместных сортов льна-долгунца в сравнении с кряжами.

** QTL, опубликованные в (You, Cloutier, 2020). PLH – высота растения; TL – техническая длина стебля; IOD – содержание йода; LIN – содержание линоленовой кислоты; LIO – содержание линолевой кислоты; PM – заболеваемость мучнистой росой; PAL – содержание пальмитиновой кислоты; OIL – содержание масла; MC – содержание семенной слизи; OLE – содержание олеиновой кислоты; STE – содержание стеариновой кислоты; LDG – формирование корня; TSW – вес 1000 семян; FN – число семенных коробочек; NB – количество ветвей.

Таблица S4. Инсерции TE, ассоциированные с фенотипическими признаками, измеренными в 2019, 2020 (два измерения) и 2021 гг.

| Сайт инсерции TE (суперсемейство) | Модель (<i>p</i> -value) | Признак, год (эффект) | Место вставки* | Ближайший ген (ортолог) | Аннотация (QTL) |
|------------------------------------|--|-----------------------|----------------|---|---|
| Chr1_3002494 (<i>Copia</i>) | FarmCPU (3.6e-07), Blink (1.2e-10), SUPER (5.2e-07), GLM (7.3e-07) | Oil, 2019 (+) | Интрон | <i>Lus10042426</i> (<i>AT5G10770.1</i>) | Белок семейства эукариотических аспартилпротеаз |
| Chr1_10250548 (–) | SUPER (3.2e-07), GLM (4.4e-07) | EFL, 2020(1) (–) | Интрон | <i>Lus10022657</i> (<i>AT3G04380.1</i>) | Белок семейства лизинметил-трансфераз, содержащий SET-домен |
| | SUPER (4.2e-07), GLM (5.6e-07) | TL, 2020(1) (–) | | | |
| Chr1_11870656 (<i>hAT-Tag1</i>) | SUPER (5.2e-07), GLM (6.5e-07) | TL, 2020(1) (–) | Интрон | <i>Lus10009386</i> (<i>AT1G08230.2</i>) | Белок семейства трансмембранных переносчиков аминокислот |
| | SUPER (4.4e-07), GLM (5.0e-07) | TL, 2020(2) (–) | | | |
| | SUPER (7.4e-07), GLM (8.2e-07) | NI, 2020(1) (–) | | | |
| | GLM (6.4e-07), Blink (1.7e-08) | PH, 2020(2) (–) | | | |
| Chr1_13735350 (<i>MULE-MuDR</i>) | SUPER (3.3e-08), GLM (4.3e-08) | TL, 2020(2) (–) | Интрон | <i>Lus10008435</i> (<i>AT5G37290.1</i>) | Белок суперсемейства повторов ARM |
| Chr1_20792291 (–) | FarmCPU (1.0e-07) | TW, 2019 (–) | Интрон | <i>Lus10015841</i> (<i>AT3G62870.1</i>) | Рибосомный белок семейства L7Ae/L30e/S12e/Gadd45 |
| Chr1_28316405 (<i>hAT-Ac</i>) | FarmCPU (1.5e-07) | Tswgt, 2020(1) (+) | 1309 us | <i>Lus10018915</i> (<i>AT1G71890.1</i>) | Главный белок-посредник суперсемейства |
| Chr2_2089015 (<i>Copia</i>) | FarmCPU (1.8e-08) | Tswgt, 2021 (+) | Экзон | <i>Lus10016394</i> (<i>AT3G08030.1</i>) | Белок с неизвестной функцией, DUF642 |
| Chr2_5913484 (<i>MULE-MuDR</i>) | FarmCPU (7.6e-09) | FW, 2019 (+) | 6051 us | <i>Lus10038050</i> (<i>AT2G47000.1</i>) | Подсемейство В4 АТФ-связывающей кассеты |
| Chr3_522289 (<i>Copia</i>) | Blink (3.3e-09), SUPER (6.7e-07), GLM (8.5e-07) | TW, 2020(2) (+) | 1318 us | <i>Lus10013432</i> (<i>AT5G23130.1</i>) | Белок, связывающий пептидогликан и содержащий домен LysM |
| Chr3_15688558 (<i>Gypsy</i>) | Blink (3.4e-09) | STI, 2020(1) (+) | 17020 ds | <i>Lus10018978</i> (<i>AT4G13650.1</i>) | Суперсемейство белков пентатрико-пептидных повторов (PPR) |
| Chr3_19617078 (<i>Gypsy</i>) | Blink (9.9e-09) | DSI, 2021 (–) | Интрон | <i>Lus10008232</i> (<i>AT5G46050.1</i>) | Пептидный транспортер 3 |
| Chr3_20336100 (<i>Copia</i>) | Blink (1.0e-07) | STI, 2020(1) (–) | 1882 ds | <i>Lus10007194</i> (<i>AT4G01690.1</i>) | Семейство флавин-содержащих аминоксидо-редуктаз |
| Chr3_20819951 (–) | FarmCPU (4.2e-08) | TW, 2020(1) (+) | 13400 ds | <i>Lus10012620</i> (<i>AT4G10810.1</i>) | – |

Продолжение табл. S4

| Сайт инсерции TE (суперсемейство) | Модель (p-value) | Признак, год (эффект) | Место вставки* | Ближайший ген (ортолог) | Аннотация (QTL) |
|--------------------------------------|--|-----------------------------|-------------------|------------------------------|--|
| Chr3_24121740 (Copia) | Blink (2.8e-07) | FC, 2021 (+) | Экзон | Lus10017112 (AT1G48130.1) | 1-Цистеин пероксиредоксин 1 |
| Chr3_24426254 (MULE-MuDR) | Blink (1.7e-10) | FC, 2021 (-) | 190 us | Lus10017063 (AT5G49480.1) | Ca2+-связывающий белок 1 |
| Chr4_3247800 (Gypsy) | Blink (2.1e-08) | PH, 2021 (-) | 11124 us | Lus10029546 (AT2G02120.1) | Белок суперсемейства ноттинов, подобный токсину скорпиона |
| Chr4_11952218 (Copia) | FarmCPU (2.0e-08) | DSI, 2020 (-) | 853 ds | Lus10036789 (AT1G13980.1) | Белок, содержащий домен sec7 |
| Chr4_14756320 (hAT-Ac) | FarmCPU (6.4e-08), Blink (1.9e-11), SUPER (5.6e-08), GLM (7.4e-08) | EFL, 2020(1) (-) | 2059 us | Lus10041548 (AT3G61120.1) | AGAMOUS-подобный 13, QTL: QYLD-Lu4.1 (YLD), QPLH-Lu4.3 (PLH), QDTM-Lu4.1 (DTM) |
| | FarmCPU (2.3e-10), Blink (4.9e-10), SUPER (1.1e-07), GLM (1.4e-07), MLMM (6.0e-07) | EFL, 2020(2) (-) | | | |
| | FarmCPU (4.9e-09), Blink (3.6e-12), SUPER (9.6e-10), GLM (2.6e-10), MLM (8.9e-08), MLMM (1.8e-08) | FW, 2019 (-) | | | |
| | FarmCPU (3.0e-08), Blink (1.8e-10), SUPER (3.5e-07), GLM (4.3e-07) | FW, 2020(1) (-) | | | |
| | Blink (1.6e-11), SUPER (7.8e-07), GLM (9.3e-07) | FW, 2020(2) (-) | | | |
| | FarmCPU (3.3e-07), Blink (1.8e-08), SUPER (7.3e-07), GLM (8.3e-07) | IL, 2021 (+) | | | |
| | Blink (9.6e-10) | IL, 2019 (+) | | | |
| | FarmCPU (7.4e-12), Blink (4.8e-17), SUPER (1.6e-11), GLM (6.5e-12), MLM (7.7e-09), MLMM (1.4e-10) | NI, 2019 (-) | | | |
| | FarmCPU (6.7e-11), Blink (3.3e-08), SUPER (4.2e-11), GLM (7.2e-11), MLM (5.9e-08), MLMM (1.1e-08) | NI, 2020(1) (-) | | | |
| | | | | | |

Продолжение табл. S4

| Сайт инсерции TE (суперсемейство) | Модель (<i>p</i> -value) | Признак, год (эффект) | Место вставки* | Ближайший ген (ортолог) | Аннотация (QTL) |
|--------------------------------------|--|-----------------------------|-------------------|--|--|
| Chr4_14756320 (<i>hAT-Ac</i>) | FarmCPU (2.6e-10), Blink (8.5e-16), SUPER (1.4e-10), GLM (2.1e-10), MLM (8.2e-08), MLMM (1.3e-09) | NI, 2020(2) (-) | 2059 us | <i>Lus10041548</i> (<i>AT3G61120.1</i>) | AGAMOUS-подобный 13, QTL: QYLD-Lu4.1 (YLD), QPLH-Lu4.3 (PLH), QDTM-Lu4.1 (DTM) |
| | FarmCPU (1.4e-07), Blink (1.5e-08), GLM (1.1e-08) | PH, 2019 (-) | | | |
| | FarmCPU (1.4e-10), Blink (2.9e-12), GLM (4.0e-08), MLMM (1.0e-07) | PH, 2020(1) (-) | | | |
| | FarmCPU (8.5e-12), Blink (3.9e-11), GLM (3.2e-09), MLMM (2.8e-08) | PH, 2020(2) (-) | | | |
| | Blink (5.6e-12), SUPER (3.7e-12), GLM (5.2e-12), MLM (9.6e-09), MLMM (9.6e-10) | TL 2019 (-) | | | |
| | FarmCPU (2.5e-10), Blink (1.1e-08), SUPER (1.4e-09), GLM (1.8e-09), MLMM (4.2e-08) | TL 2020(1) (-) | | | |
| | FarmCPU (2.0e-10), Blink (8.1e-11), SUPER (3.9e-10), GLM (5.5e-09), MLM (2.0e-07), MLMM (2.5e-09) | TL 2020(2) (-) | | | |
| | Blink (5.6e-10), SUPER (3.5e-10), GLM (2.9e-10), MLM (3.4e-07), MLMM (1.3e-08) | TL 2021 (-) | | | |
| | FarmCPU (1.6e-10), Blink (5.7e-11), SUPER (7.9e-07), MLMM (6.0e-07) | FC, 2019 (-) | | | |
| | SUPER (3.0e-07), GLM (2.3e-07) | TW, 2019 (-) | | | |
| | FarmCPU (5.3e-08), Blink (5.4e-09), SUPER (5.9e-07), GLM (7.2e-07) | TW, 2020(1) (-) | | | |
| | Blink (9.1e-09), SUPER (1.3e-07), GLM (1.7e-07) | TW, 2020(2) (-) | | | |

Продолжение табл. S4

| Сайт инсерции TE (суперсемейство) | Модель (<i>p</i> -value) | Признак, год (эффект) | Место вставки* | Ближайший ген (ортолог) | Аннотация (QTL) |
|--------------------------------------|---|-----------------------------|-------------------|--|--|
| Chr4_14994752 (<i>Copia</i>) | SUPER (2.5e-07), GLM (6.1e-07) | TL, 2021 (–) | Инtron | <i>Lus10041602</i> (<i>AT3G52150.1</i>) | РНК-связывающий белок (мотивы RRM/RBD/RNP); QTL: QPLH-Lu4.3 (PLH), QDTM-Lu4.1 (DTM) |
| | SUPER (3.9e-07), GLM (5.2e-07) | NI, 2020(1) (–) | | | |
| Chr6_1356894 (<i>Copia</i>) | SUPER (6.5e-07), GLM (8.1e-07) | TL, 2020(1) (–) | Инtron | <i>Lus10033725</i> (<i>AT1G12470.1</i>) | Связывание ионов цинка |
| | SUPER (8.8e-07) | NI, 2020(1) (–) | | | |
| Chr6_11075470 (<i>Copia</i>) | FarmCPU (2.7e-07) | TL, 2020(2) (+) | 2456 us | <i>Lus10002212</i> (<i>AT5G01300.1</i>) | Белок семейства РЕВР (фосфатидил-этаноламин-связывающий белок) |
| Chr6_11780716 (<i>hAT-Ac</i>) | FarmCPU (6.1e-07) | Tswgt, 2021 (+) | 1006 us | <i>Lus10017263</i> (<i>AT4G36040.1</i>) | Белок суперсемейства шапероновых доменов DnaJ |
| Chr6_15308804 (–) | Blink (3.4e-07) | FW, 2020(1) (+) | 6646 ds | <i>Lus10014365</i> (<i>AT5G12100.1</i>) | Белок, содержащий пентатрикопеп-тидный повтор (PPR) |
| Chr6_17837550 (<i>Copia</i>) | Blink (1.9e-08) | TW, 2020(2) (+) | 400 us | <i>Lus10024282</i> (<i>AT1G58250.1</i>) | Домен белка локализации Гольджи; промотор РНК-полимеразы II; Домен белка Fmp27 |
| Chr7_15829004 (<i>Copia</i>) | FarmCPU (2.5e-12) | FW, 2020(1) (–) | 709 us | <i>Lus10038370</i> (<i>AT4G20050.1</i>) | Суперсемейство белков, подобных пектинлиазе; QTL: QIOD-Lu7.2 (IOD), QLIN-Lu7.2 (LIN), QLIO-Lu7.2 (LIO) |
| Chr8_2036044 (<i>Copia</i>) | FarmCPU (3.9e-07) | FC, 2021 (+) | 7932 ds | <i>Lus10017029</i> (<i>AT5G03870.1</i>) | Белок семейства глутаредоксинов |
| Chr8_4775776 (–) | SUPER (1.8e-07), GLM (2.7e-07), Blink (7.2e-10) | EFL, 2020(1) (–) | 34 us | <i>Lus10023944</i> (<i>AT4G38380.1</i>) | Белок семейства МАТЕ-эффлюкс |
| | SUPER (2.2e-07), GLM (2.4e-07), Blink (5.0e-10) | TW, 2020(1) (+) | | | |
| | GLM (1.5e-07), Blink | PH, 2020(1) (–) | | | |
| | SUPER (1.8e-08), GLM (2.4e-08) | TL, 2020(1) (–) | | | |
| Chr8_12565858 (<i>hAT-Ac</i>) | FarmCPU (1.5e-09) | NI, 2020(2) (+) | 1199 ds | <i>Lus10025285</i> (<i>AT2G37370.1</i>) | Центросомальный белок |
| | FarmCPU (1.6e-09) | PH, 2020(2) (+) | | | |

Продолжение табл. S4

| Сайт инсерции TE (суперсемейство) | Модель (<i>p</i> -value) | Признак, год (эффект) | Место вставки* | Ближайший ген (ортолог) | Аннотация (QTL) |
|--|---|-----------------------------|-------------------|--|---|
| Chr8_16467884 (<i>MULE-MuDR</i>) | SUPER (4.8e-07), GLM (1.7e-07) | TW, 2019 (-) | 972 ds | <i>Lus10004540</i> (<i>AT1G29850.2</i>) | Белок семейства двуцепочечных ДНК-связывающих белков |
| | SUPER (6.6e-07), GLM (1.7e-07) | FW, 2019 (-) | | | |
| Chr8_16923232 (<i>Copia</i>) | FarmCPU (5.9e-13), Blink (8.8e-10), SUPER (8.4e-07), GLM (9.1e-07) | FC, 2021 (+) | Интрон | <i>Lus10000673</i> (<i>AT1G32450.1</i>) | Нитратный транспортер 1,5 |
| Chr8_18366764 (<i>Copia</i>) | FarmCPU (1.9e-07) | PH, 2020(2) (+) | Экзон | <i>Lus10007810</i> (<i>AT5G36930.2</i>) | Семейство белков устойчивости к болезням (класс TIR-NBS-LRR) |
| | FarmCPU (5.3e-07) | NI, 2020(2) (+) | | | |
| Chr8_22233522 (<i>Copia</i>) | Blink (1.6e-07) | PH, 2020(2) (+) | 2260 us | <i>Lus10039488</i> (<i>AT5G43700.1</i>) | Белок семейства регуляторов транскрипции AUX/IAA QTL: QOLE-Lu8.1 (OLE) |
| | Blink (8.0e-07) | NI, 2020(2) (+) | | | |
| Chr9_3613834 (-) | GLM (6.7e-07) | TW, 2019 (-) | 989 ds | <i>Lus10040473</i> (<i>AT5G10780.1</i>) | - |
| Chr10_2903372 (<i>hAT-Ac</i>) | FarmCPU (4.4e-12) | Oil, 2021 (+) | 3771 ds | <i>Lus10039339</i> (<i>AT1G78090.1</i>) | Трегалозо-6-фосфатфосфатаза |
| Chr10_3726193 (<i>hAT-Ac</i>) | GLM (6.7e-07) | FW, 2019 (+) | 9126 ds | <i>Lus10039416</i> (<i>AT1G23820.1</i>) | Спермидинсинтаза 1 |
| Chr10_6544722 (<i>hAT-Ac</i>) | GLM (7.6e-07) | FW, 2019 (+) | Интрон | <i>Lus10015082</i> (<i>AT2G27830.1</i>) | - |
| Chr11_14960430 (<i>Copia</i>) | FarmCPU (5.6e-07) | TW, 2020(2) (-) | 3421 ds | <i>Lus10008566</i> (<i>AT1G55200.1</i>) | Белок протеинкиназы с доменом, подобным адениннуклеотид-альфа- гидролазе |
| | FarmCPU (1.1e-08) | TL, 2020(1) (-) | | | |
| | FarmCPU (2.6e-07) | NI, 2020(1) (-) | | | |
| Chr12_10287266 (<i>Gypsy</i>) | FarmCPU (2.0e-07) | NI, 2019 (-) | 7166 ds | <i>Lus10024219</i> (<i>AT1G74640.1</i>) | Белок суперсемейства альфа/бета-гидролаз |
| Chr12_11155523 (<i>MULE-MuDR</i>) | FarmCPU (2.4e-07) | TW, 2019 (-) | 373 us | <i>Lus10030109</i> (<i>AT4G02530.1</i>) | Белок просвета тилакоида хлоропласта |
| Chr13_1437052 (<i>Copia</i>) | SUPER (5.4e-08), GLM (6.8e-08) | EFL, 2020(1) (-) | Экзон | <i>Lus10010648</i> (<i>AT3G03770.1</i>) | Белок с богатым лейцином повтором, белок семейства киназ |

Окончание табл. S4

| Сайт инсерции TE (суперсемейство) | Модель (p-value) | Признак, год (эффект) | Место вставки* | Ближайший ген (ортолог) | Аннотация (QTL) |
|---|---|-----------------------------|-------------------|------------------------------|---|
| Chr13_2207778 (MULE-MuDR, hAT-Tag1) | FarmCPU (6.7e-08) | TL, 2020(1) (+) | 394 ds | Lus10026981 (AT5G35320.1) | – |
| Chr13_2213745 (Copia) | Blink (6.2e-08) | Oil, 2020(2) (+) | 271 us | Lus10026978 (AT2G26110.1) | Белок с неизвестной функцией (DUF761) |
| Chr13_3391752 (Copia) | FarmCPU (1.1e-07) | PH, 2020(2) (–) | 1071 ds | Lus10026032 (AT4G36670.1) | Главный белок-посредник суперсемейства |
| | FarmCPU (2.3e-08), Blink (2.6e-07) | NI, 2020(2) (–) | | | |
| | FarmCPU (1.5e-07) | TW, 2020(1) (–) | | | |
| Chr13_6346942 (hAT-Ac) | Blink (1.2e-09) | FW, 2020(2) (–) | 253 us | Lus10002084 (AT5G64040.1) | Субъединица реакционного центра фотосистемы I PSI-N, хлоропласт, предполагаемая/PSI-N, предполагаемая (PSAN) |
| Chr13_8552558 (Copia) | FarmCPU (8.9e-08) | TL, 2021 (+) | 7148 ds | Lus10026441 (AT4G33060.1) | Белок семейства циклофилин- подобных пептидил-пролил-цис- транс-изомераз |
| Chr13_11641862 (Copia) | Blink (2.0e-10), SUPER (7.4e-07), GLM (9.2e-07) | EFL, 2020(1) (–) | 972 ds | Lus10009758 (AT4G17650.1) | Поли-кетидциклаза/дегидраза и белок переноса липидов |
| Chr13_14372739 (Copia) | GLM (4.4e-07) | PH, 2021 (–) | 780 us | Lus10034660 (AT4G29680.1) | Белок семейства щелочно- фосфатазоподобных белков |
| Chr14_9124491 (Copia) | FarmCPU (6.6e-08) | DSI, 2020 (–) | 153 us | Lus10005526 (AT2G37690.1) | Фосфорибозил-аминоимидазол- карбоксилаза, предполагаемая/AIR- карбоксилаза, предполагаемая |
| Chr14_9853868 (Copia) | FarmCPU (5.4e-10), Blink (3.9e-07) | Nsed, 2020(2) (+) | 984 ds | Lus10000815 (AT2G27410.1) | Домен неизвестной функции (DUF313) |
| Chr14_16743646 (Copia) | Blink (2.1e-10) | FW, 2020(2) (–) | 466 ds | Lus10017981 (AT1G14590.1) | Белок семейства нуклеотид- дифосфосахарных трансфераз |

* Координата инсерции TE – середина окна длиной в 50 п. н.
** Эффект инсерции на признак: (+) – положительный, (–) – отрицательный. QTL, если есть, * ds и us – вниз и вверх по течению от гена; EFL – длина элементарного волокна; FC – содержание волокна; FW – вес волокна; IL – длина соцветия; NI – число междоузлий; Oil – содержание масла в семенах; PH – высота растения; Nsed – количество семян с растения; STI – сбежистость; TL – техническая длина стебля; TW – вес технической части растения; Tswgt – вес 1000 семян.

Таблица S5. Пересечение районов, испытавших влияние отбора и обнаруженных при использовании в качестве маркеров генетического разнообразия либо сайтов инсерции TE, либо однонуклеотидных полиморфизмов и гены, в которые попадали определяющие участок транспозоны

| Положение участка (суперсемейство определяющих участок TE) | Уменьшение разнообразия* | Ген (ортолог) | Аннотация |
|--|-----------------------------|-------------------------------------|--|
| Chr13:5000001-5200000 (Copia; Gypsy) | Д/К, Д/С | <i>Lus10009365</i> (AT5G67090.1) | Белок семейства субтилизин-подобных сериновых эндопептидаз |
| Chr14:18600001-18800000 (Copia; Gypsy) | Д/С | <i>Lus10015020</i> (AT5G62190.1) | РНК-хеликаза DEAD-box (PRH75) |
| Chr15:13400001-13600000 (hAT-Ac; MULE-MuDR; Copia) | Д/С, К/С | <i>Lus10009990</i> (AT2G41480.1) | Белок суперсемейства пероксидаз |
| | | <i>Lus10009986</i> (AT4G01860.1) | Белок семейства трансдуцинов/белок семейства повторов WD-40 |
| Chr4:12400001-12600000 (hAT-Ac; Copia; MULE-MuDR) | М/С | <i>Lus10036915</i> (AT1G69220.1) | Белок суперсемейства протеинкиназ |
| | | <i>Lus10036917</i> (AT1G26370.1) | Белок семейства РНК-хеликаз |
| Chr8:22400001-22600000 (–) | М/С | <i>Lus10015365</i> (AT5G15140.1) | Белок суперсемейства подобных галактозомутаротазе QTL: scaffold635-43971 (NB), QOLE-Lu8.1 (OLE) |

* Д/К – уменьшение разнообразия у культурных сортов льна-долгунца в сравнении с кряжами; Д/С, М/С – уменьшение разнообразия у культурных сортов в сравнении со староместными сортами у льна-долгунца и масличного льна соответственно; К/С – уменьшение разнообразия у кряжей в сравнении со староместными сортами льна-долгунца.