

SUPPLEMENTARY MATERIALS

to the article by M.A. Duk, V.A. Stanin, A.A. Kanapin, A.A. Samsonova, T.A. Rozhmina, M.G. Samsonova
 "Flax transposons: unraveling their impact on domestication and agronomic trait variation"

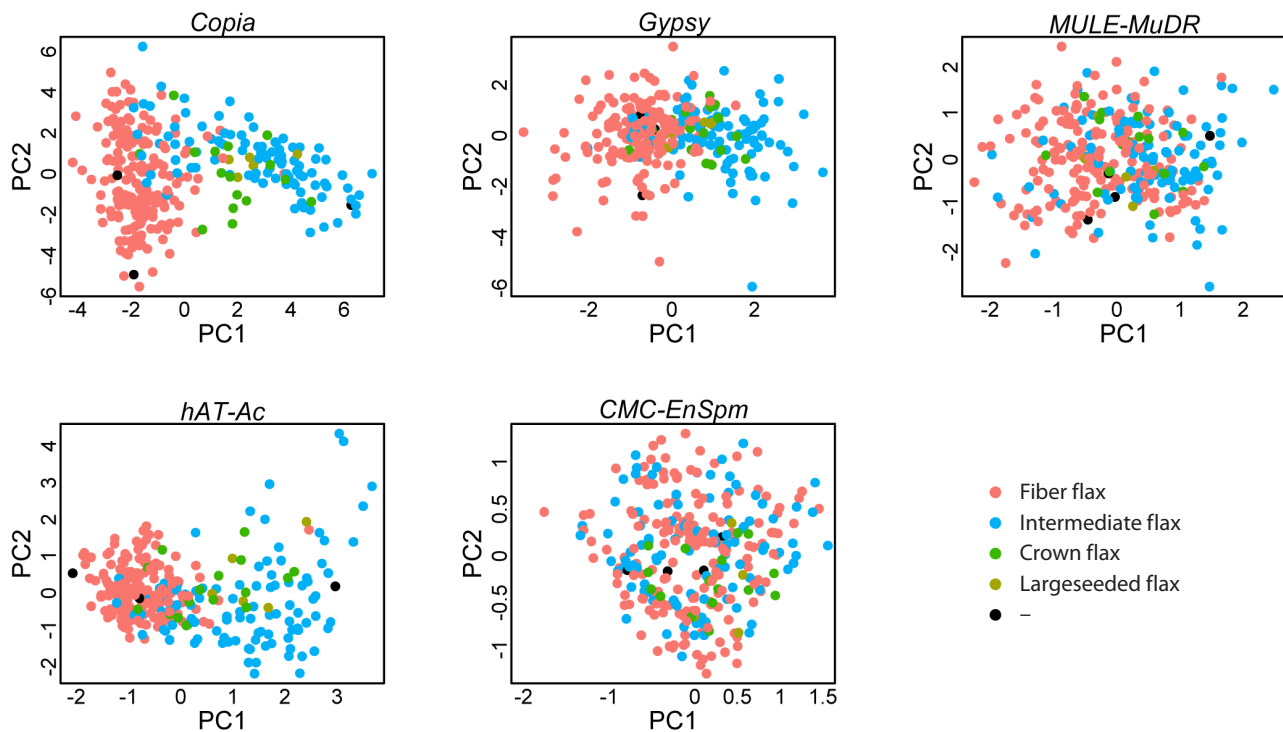


Fig. S1. Plots for first and second principal components for TE insertions of superfamilies: *Copia*, *Gypsy*, *MULE-MuDR*, *CMC-EnSpm* and *hAT-Ac*.

Accessions are colored by morphotype.

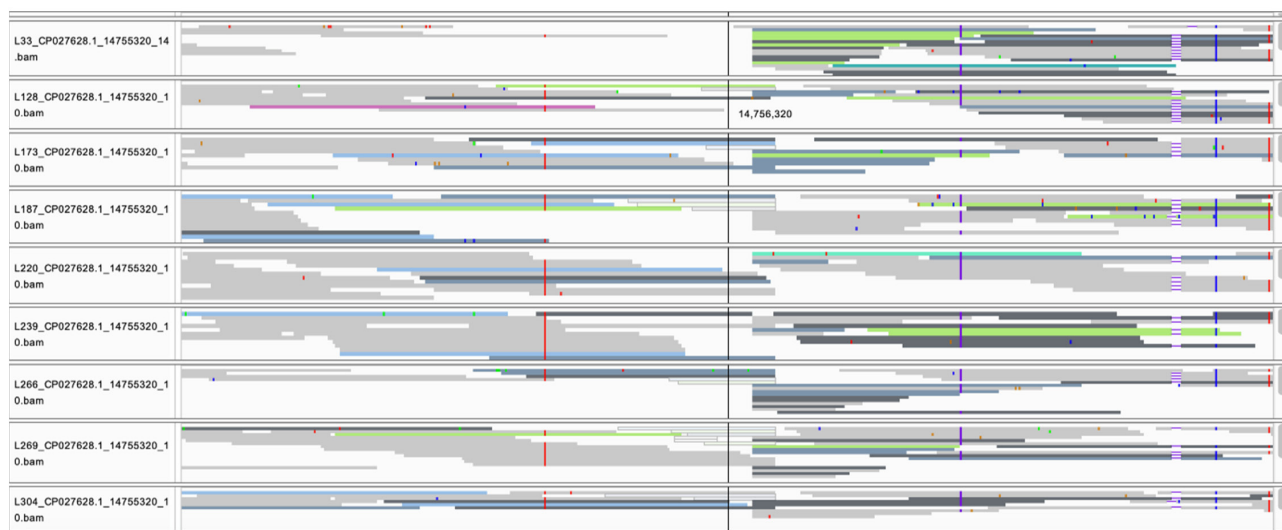


Fig. S2. Confirmation of *hAT-Ac* transposon insertion at position 14756320 on Chr4. The result is shown for nine samples. The dashed vertical line indicates the insertion site.



Fig. S3. Confirmation of *Copia* transposon insertion at position 3391752 on Chr13. The result is shown for six samples. The dashed vertical line indicates the insertion sit.



Fig. S4. Confirmation of *hAT-Tag1* transposon insertion at position 11870656 on Chr1. The result is shown for five samples. The dashed vertical line indicates the insertion sit.

Table S1. Proportion of the transposones, located in different distance from known genes

Superfamily	2,000 bp upstream	Exon	Intron	2,000 bp downstream	Far from genes	Totally in collection
All 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
<i>DNA</i> (Unspecified classification)	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
Unclassified	0.15	0.027	0.058	0.17	0.6	172,984

Table S2. List of TEs, abundant in clades I–III

We considered TE as abundant in clade, if the frequency of occurrence in one clade is no more than twice the frequency of occurrence in each of the other two clades

TE identifier	Superfamily	Number of insertion sites	Number of insertions in one accession			Abundant in clade
			Clade I	Clade II	Clade III	
LU_v2-B-R24968-Map6	<i>Helitron</i>	26	0.22	0.068	0.033	Clade I
LU_v2-B-R5502-Map4	<i>Gypsy</i>	16	0.12	0.023	0.04	Clade I
LU_v2-B-R19931-Map3	<i>Copia</i>	17	0.13	0.034	0.033	Clade I
LU_v2-B-G1344-Map3	<i>Copia</i>	33	0.22	0.057	0.087	Clade I
LU_v2-B-R3227-Map20	<i>Copia</i>	21	0.13	0.034	0.06	Clade I
LU_v2-B-R5856-Map20	<i>Gypsy</i>	39	0.25	0.068	0.11	Clade I
LU_v2-B-R2691-Map3_reversed	<i>Copia</i>	58	0.37	0.18	0.11	Clade I
LU_v2-B-R17883-Map3_reversed	<i>Copia</i>	19	0.12	0.057	0.04	Clade I
LU_v2-B-R3765-Map4	<i>hAT-Ac</i>	29	0.18	0.08	0.067	Clade I
LU_v2-B-G1822-Map20	<i>DNA</i> (Unspecified classification)	87	0.49	0.24	0.22	Clade I
LU_v2-B-R3265-Map5_reversed	<i>Copia</i>	72	0.16	0.4	0.17	Clade II
LU_v2-B-R3814-Map20	<i>Copia</i>	17	0.044	0.091	0.04	Clade II
LU_v2-B-G3342-Map7_reversed	<i>Copia</i>	21	0.015	0.011	0.13	Clade III
LU_v2-B-R3762-Map5	<i>Copia</i>	54	0.13	0.034	0.28	Clade III
LU_v2-B-R3967-Map17	<i>Copia</i>	73	0.16	0.12	0.34	Clade III

Table S3. Flax genome regions with selection signs for the accession groups

Region location	Reduction in diversity*	QTL**
Chr13:16200001-16400000	F/O, O/L	
Chr4:16600001-16800000	O/F, O/L	
Chr1:28600001-28800000	F/K, F/L	
Chr12:1000001-1200000	F/K, F/L	uq.C12-1 (PLH), uq.C12-1 (TL), QIOD-Lu12.3 (IOD), QLIN-Lu12.3 (LIN), QLIO-Lu12.3 (LIO)
Chr6:17400001-17600000	K/F, L/F	
Chr7:3800001-4000000	K/F, K/L	QPM-crc-LG7 (PM), QPAL-Lu7.3 (PAL)
Chr4:9800001-10000000	F/L, O/L	
Chr9:2100001-21200000	F/L, K/L	
Chr15:13400001-13600000	F/L, K/L	
Chr2:24800001-25000000	F/O	
Chr5:3200001-3400000		
Chr6:13800001-14000000		QOIL-Lu6.4 (OIL)
Chr8:2800001-3000000		
Chr10:200001-400000		
Chr11:6400001-6600000		
Chr8:16800001-17000000	O/F	
Chr14:18800001-19000000		
Chr3:17600001-17800000	F/K	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	K/F	
Chr6:3400001-3600000		
Chr6:15800001-16000000		
Chr9:19600001-19800000		
Chr15:10000001-10200000		
Chr2:1200001-1400000	F/L	
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	L/F	

Table S3 (end)

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

* Abbreviations: F/O – reduction in diversity in fiber cultivars as compared with oil (linseed) cultivars, O/F – reduction in diversity in oil (linseed) cultivars as compared with fiber cultivars, F/L, O/L – reduction in diversity in cultivars as compared with landraces for fiber and oil (linseed) flax respectively, L/F, L/O – reduction in diversity in landraces as compared with cultivars or fiber and oil (linseed) flax respectively, F/K – reduction in diversity in fiber cultivars as compared with kryazhs, K/F – reduction in diversity in kryazhs as compared with fiber cultivars, K/L – reduction in diversity in kryazhs as compared with fiber landraces, L/K – reduction in diversity in fiber landraces as compared with kryazhs. ** QTL, published in (You, Cloutier, 2020), abbreviations: PLH – plant height, TL – technical length of the stem, IOD – iodine content, LIN – linolenic acid content, LIO – linoleic acid content, PM – incidence of powdery mildew, PAL – palmitic acid content, OIL – oil content, MC – mucus content, OLE – oleic acid content, STE – stearic acid content, LDG – lodging, TSW – weight of a thousand seeds, FN – number of seed boxes, NB number of branches.

Table S4. TE insertions, associated with phenotypic traits, measured in 2019, 2020 (two data sets) and 2021 years

TE insertion site (superfamily)	Model	Trait, year (Effect)	Insertion location*	Closest gene (Ortholog)	Annotation (QTL)
Chr1_3002494 (<i>Copia</i>)	FarmCPU (3.6e-07), Blink (1.2e-10), SUPER (5.2e-07), GLM (7.3e-07)	Oil, 2019 (+)	intron	<i>Lus10042426</i> (<i>AT5G10770.1</i>)	Eukaryotic aspartyl protease family protein
Chr1_10250548 (-)	SUPER (3.2e-07), GLM (4.4e-07)	EFL, 2020(1) (-)	intron	<i>Lus10022657</i> (<i>AT3G04380.1</i>)	SET-domain containing protein lysine methyltransferase family protein
	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) (-)	intron	<i>Lus10009386</i> (<i>AT1G08230.2</i>)	Transmembrane amino acid transporter family protein
	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) (-)	intron	<i>Lus10008435</i> (<i>AT5G37290.1</i>)	ARM repeat superfamily protein
Chr1_20792291 (-)	FarmCPU (1.0e-07)	TW, 2019 (-)	intron	<i>Lus10015841</i> (<i>AT3G62870.1</i>)	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein
Chr1_28316405 (<i>hAT-Ac</i>)	FarmCPU (1.5e-07)	Tswgt, 2020(1) (+)	1309 us	<i>Lus10018915</i> (<i>AT1G71890.1</i>)	Major facilitator superfamily protein
Chr2_2089015 (<i>Copia</i>)	FarmCPU (1.8e-08)	Tswgt, 2021 (+)	exon	<i>Lus10016394</i> (<i>AT3G08030.1</i>)	Protein of unknown function, DUF642
Chr2_5913484 (<i>MULE-MuDR</i>)	FarmCPU (7.6e-09)	FW, 2019 (+)	6051 us	<i>Lus10038050</i> (<i>AT2G47000.1</i>)	ATP binding cassette subfamily B4
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>)	Peptidoglycan-binding LysM domain-containing protein
Chr3_15688558 (<i>Gypsy</i>)	Blink (3.4e-09)	STI, 2020(1) (+)	17020 ds	<i>Lus10018978</i> (<i>AT4G13650.1</i>)	Pentatricopeptide repeat (PPR) superfamily protein
Chr3_19617078 (<i>Gypsy</i>)	Blink (9.9e-09)	DSI, 2021 (-)	intron	<i>Lus10008232</i> (<i>AT5G46050.1</i>)	Peptide transporter 3
Chr3_20336100 (<i>Copia</i>)	Blink (1.0e-07)	STI, 2020(1) (-)	1882 ds	<i>Lus10007194</i> (<i>AT4G01690.1</i>)	Flavin containing amine oxidoreductase family
Chr3_20819951 (-)	FarmCPU (4.2e-08)	TW, 2020(1) (+)	13400 ds	<i>Lus10012620</i> (<i>AT4G10810.1</i>)	-

Table S4 (continued)

TE insertion site (superfamily)	Model	Trait, year (Effect)	Insertion location*	Closest gene (Ortholog)	Annotation (QTL)
Chr3_24121740 (<i>Copia</i>)	Blink (2.8e-07)	FC, 2021 (+)	exon	<i>Lus10017112</i> (<i>AT1G48130.1</i>)	1-cysteine peroxiredoxin 1
Chr3_24426254 (<i>MULE-MuDR</i>)	Blink (1.7e-10)	FC, 2021 (-)	190 us	<i>Lus10017063</i> (<i>AT5G49480.1</i>)	Ca ²⁺ -binding protein 1
Chr4_3247800 (<i>Gypsy</i>)	Blink (2.1e-08)	PH, 2021 (-)	11124 us	<i>Lus10029546</i> (<i>AT2G02120.1</i>)	Scorpion toxin-like knottin superfamily protein
Chr4_11952218 (<i>Copia</i>)	FarmCPU (2.0e-08)	DSI, 2020 (-)	853 ds	<i>Lus10036789</i> (<i>AT1G13980.1</i>)	sec7 domain-containing protein
Chr4_14756320 (<i>hAT-Ac</i>)	FarmCPU (6.4e-08), Blink (1.9e-11), SUPER (5.6e-08), GLM (7.4e-08)	EFL, 2020(1) (-)	2059 us	<i>Lus10041548</i> (<i>AT3G61120.1</i>)	AGAMOUS-like 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) (-)			

Table S4 (continued)

TE insertion site (superfamily)	Model	Trait, year (Effect)	Insertion location*	Closest gene (Ortholog)	Annotation (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-like 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) (-)			

Table S4 (continued)

TE insertion site (superfamily)	Model	Trait, year (Effect)	Insertion location*	Closest gene (Ortholog)	Annotation (QTL)
Chr4_14994752 (<i>Copia</i>)	SUPER (2.5e-07), GLM (6.1e-07)	TL, 2021 (-)	intron	<i>Lus10041602</i> (<i>AT3G52150.1</i>)	RNA-binding (RRM/RBD/RNP motifs) family protein; 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) (-)	intron	<i>Lus10033725</i> (<i>AT1G12470.1</i>)	Zinc ion binding
	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>)	PEBP (phosphatidyle-thanolamine-binding protein) family protein
Chr6_11780716 (<i>hAT-Ac</i>)	FarmCPU (6.1e-07)	Tswgt, 2021 (+)	1006 us	<i>Lus10017263</i> (<i>AT4G36040.1</i>)	Chaperone DnaJ-domain superfamily protein
Chr6_15308804 (-)	Blink (3.4e-07)	FW, 2020(1) (+)	6646 ds	<i>Lus10014365</i> (<i>AT5G12100.1</i>)	pentatricopeptide (PPR) repeat-containing protein
Chr6_17837550 (<i>Copia</i>)	Blink (1.9e-08)	TW, 2020(2) (+)	400 us	<i>Lus10024282</i> (<i>AT1G58250.1</i>)	Golgi-body localisation protein domain; RNA pol II promoter Fmp27 protein domain
Chr7_15829004 (<i>Copia</i>)	FarmCPU (2.5e-12)	FW, 2020(1) (-)	709 us	<i>Lus10038370</i> (<i>AT4G20050.1</i>)	Pectin lyase-like superfamily protein; 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>)	Glutaredoxin family protein
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>)	MATE efflux family protein
	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>)	Centrosomal protein
	FarmCPU (1.6e-09)	PH, 2020(2) (+)			

Table S4 (continued)

TE insertion site (superfamily)	Model	Trait, year (Effect)	Insertion location*	Closest gene (Ortholog)	Annotation (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>)	Double-stranded DNA-binding family protein
	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 (+)	intron	<i>Lus10000673</i> (<i>AT1G32450.1</i>)	Nitrate transporter 1.5
Chr8_18366764 (<i>Copia</i>)	FarmCPU (1.9e-07)	PH, 2020(2) (+)	exon	<i>Lus10007810</i> (<i>AT5G36930.2</i>)	Disease resistance protein (TIR-NBS-LRR class) family
	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>)	AUX/IAA transcriptional regulator family protein 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>)	Trehalose-6-phosphate phosphatase
Chr10_3726193 (<i>hAT-Ac</i>)	GLM (6.7e-07)	FW, 2019 (+)	9126 ds	<i>Lus10039416</i> (<i>AT1G23820.1</i>)	Spermidine synthase 1
Chr10_6544722 (<i>hAT-Ac</i>)	GLM (7.6e-07)	FW, 2019 (+)	intron	<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>)	Protein kinase protein with adenine nucleotide alpha hydrolases-like domain
	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>)	Alpha/beta-Hydrolases superfamily protein
Chr12_11155523 (<i>MULE-MuDR</i>)	FarmCPU (2.4e-07)	TW, 2019 (-)	373 us	<i>Lus10030109</i> (<i>AT4G02530.1</i>)	Chloroplast thylakoid lumen protein
Chr13_1437052 (<i>Copia</i>)	SUPER (5.4e-08), GLM (6.8e-08)	EFL, 2020(1) (-)	exon	<i>Lus10010648</i> (<i>AT3G03770.1</i>)	Leucine-rich repeat protein kinase family protein

Table S4 (end)

TE insertion site (superfamily)	Model	Trait, year (Effect)	Insertion location*	Closest gene (Ortholog)	Annotation (QTL)
Chr13_2207778 (<i>MULE-MuDR</i> , <i>hAT-Tag1</i>)	FarmCPU (6.7e-08)	TL, 2020(1) (+)	394 ds	<i>Lus10026981</i> (<i>AT5G35320.1</i>)	–
Chr13_2213745 (<i>Copia</i>)	Blink (6.2e-08)	Oil, 2020(2) (+)	271 us	<i>Lus10026978</i> (<i>AT2G26110.1</i>)	Protein of unknown function (DUF761)
Chr13_3391752 (<i>Copia</i>)	FarmCPU (1.1e-07)	PH, 2020(2) (–)	1071 ds	<i>Lus10026032</i> (<i>AT4G36670.1</i>)	Major facilitator superfamily protein
	FarmCPU (2.3e-08), Blink (2.6e-07)	NI, 2020(2) (–)			
	FarmCPU (1.5e-07)	TW, 2020(1) (–)			
Chr13_6346942 (<i>hAT-Ac</i>)	Blink (1.2e-09)	FW, 2020(2) (–)	253 us	<i>Lus10002084</i> (<i>AT5G64040.1</i>)	Photosystem I reaction center subunit PSI-N, chloroplast, putative/PSI-N, putative (PSAN)
Chr13_8552558 (<i>Copia</i>)	FarmCPU (8.9e-08)	TL, 2021 (+)	7148 ds	<i>Lus10026441</i> (<i>AT4G33060.1</i>)	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein
Chr13_11641862 (<i>Copia</i>)	Blink (2.0e-10), SUPER (7.4e-07), GLM (9.2e-07)	EFL, 2020(1) (–)	972 ds	<i>Lus10009758</i> (<i>AT4G17650.1</i>)	Polyketide cyclase/dehydrase and lipid transport protein
Chr13_14372739 (<i>Copia</i>)	GLM (4.4e-07)	PH, 2021 (–)	780 us	<i>Lus10034660</i> (<i>AT4G29680.1</i>)	Alkaline-phosphatase-like family protein
Chr14_9124491 (<i>Copia</i>)	FarmCPU (6.6e-08)	DSI, 2020 (–)	153 us	<i>Lus10005526</i> (<i>AT2G37690.1</i>)	Phosphoribosyl-aminoimidazole carboxylase, putative/AIR carboxylase, putative
Chr14_9853868 (<i>Copia</i>)	FarmCPU (5.4e-10), Blink (3.9e-07)	Nsed, 2020(2) (+)	984 ds	<i>Lus10000815</i> (<i>AT2G27410.1</i>)	Domain of unknown function (DUF313)
Chr14_16743646 (<i>Copia</i>)	Blink (2.1e-10)	FW, 2020(2) (–)	466 ds	<i>Lus10017981</i> (<i>AT1G14590.1</i>)	Nucleotide-diphospho-sugar transferase family protein

* Coordinate of the TE insertion is the center of the 50 bp window.

** Insertion effect on the trait: (+) positive, (–) negative. * ds и us – downstream or upstream from the gene, EFL – elemental fiber length, FC – fiber content, FW – fiber weight, IL – inflorescence length, NI – number of internodes, Oil – oil content, PH – plant height, Nsed – number of seeds from one plant, STI – taper of the stem, TL – technical length of the stem, TW – technical part weight, Tswgt – thousand seed weight.

Table S5. Insertions of the regions with selection signs, detected while using TE insertion sites or SNPs as markers of genetic diversity, and genes in which located TEs, used as markers

Region location (superfamily of TE used as markers)	Reduction in diversity*	Gene (Ortholog)	Annotation
Chr13:5000001-5200000 (Copia; Gypsy)	F/K, F/L	<i>Lus10009365</i> (<i>AT5G67090.1</i>)	Subtilisin-like serine endopeptidase family protein
Chr14:18600001-18800000 (Copia; Gypsy)	F/L	<i>Lus10015020</i> (<i>AT5G62190.1</i>)	DEAD box RNA helicase (PRH75)
Chr15:13400001-13600000 (hAT-Ac; MULE-MuDR; Copia)	F/L, K/L	<i>Lus10009990</i> (<i>AT2G41480.1</i>)	Peroxidase superfamily protein
		<i>Lus10009986</i> (<i>AT4G01860.1</i>)	Transducin family protein/WD-40 repeat family protein
Chr4:12400001-12600000 (hAT-Ac; Copia; MULE-MuDR)	O/L	<i>Lus10036915</i> (<i>AT1G69220.1</i>)	Protein kinase superfamily protein
		<i>Lus10036917</i> (<i>AT1G26370.1</i>)	RNA helicase family protein
Chr8:22400001-22600000 (unclassified TEs)	O/L	<i>Lus10015365</i> (<i>AT5G15140.1</i>)	Galactose mutarotase-like superfamily protein QTL: scaffold635-43971 (NB), QOLE-Lu8.1 (OLE)

* Abbreviations: F/K – reduction in diversity in fiber flax cultivars as compared with kryazhs, F/L, O/L – reduction in diversity in cultivars as compared with landraces for fiber and oil (linseed) flax respectively, K/L – reduction in diversity in kryazhs as compared with fiber landraces.