

ПРИМЕНЕНИЕ МЕТОДА 3VmrMLM ДЛЯ ПОИСКА НОВЫХ ГЕНОМНЫХ ВАРИАНТОВ, АССОЦИИРОВАННЫХ С ХАРАКТЕРИСТИКАМИ ВОЛОКНА У ЛЬНА

© 2025 г. М.А. Дук*, **, #, А.А. Канапин*, М.П. Банкин*, М.Г. Самсонова*

*Санкт-Петербургский политехнический университет Петра Великого,
Политехническая ул., 29, Санкт-Петербург, 195251, Россия

**Физико-технический институт им. А.Ф. Иоффе, Политехническая ул., 26, Санкт-Петербург, 194021, Россия

#E-mail: duk@mail.ioffe.ru

Поступила в редакцию 27.09.2024 г.

После доработки 27.09.2024 г.

Принята к публикации 04.10.2024 г.

Лен — важная сельскохозяйственная культура, выращиваемая для получения масла или волокна. Льняное волокно используется в различных областях промышленности, и выведение новых сортов льна-долгунца с лучшими характеристиками волокна представляет интерес. Полногеномный поиск ассоциаций (GWAS) позволяет найти варианты, ассоциированные с характеристиками, важными для качества волокна, однако различия в данных, связанные с разными погодными условиями при выращивании в разные годы, понижают мощность методов GWAS. Метод 3VmrMLM допускает поиск вариантов для данных, измеренных в нескольких средах, позволяя найти новые варианты, не найденные другими методами. В качестве разных сред были взяты измерения в разные годы, и метод позволил найти всего 205 вариантов, характерных для всех или нескольких сред, 37 из которых попадали в тело известных генов с важными функциями. Наличие эффекта некоторых вариантов на характеристики волокна было также подтверждено на независимой выборке растений.

Ключевые слова: геномные ассоциации, лен, содержание волокна, длина элементарного волокна, GWAS, 3VmrMLM.

DOI: 10.31857/S0006302925010181, EDN: LUNBPB

Лен (*Linum usitatissimum*) — одна из важнейших сельскохозяйственных культур, используемая как для получения волокна, так и для получения масла. Льняное волокно, в частности, используется в различных отраслях промышленности для создания одежды, тканей и композитов [1]. В настоящее время композиты из льняных волокон широко используются в автомобильной промышленности [2]. Элементарное волокно представляет собой отдельную клетку большой длины с толстой клеточной стенкой [3]. Клетки происходят из апикальной меристемы и развиваются в две стадии: растяжение и утолщение стенки. Утолщение клеточной стенки включает отложение нескольких слоев изнутри первичной клеточной стенки и начинается после растяжения. Отложенные слои содержат большое количество цел-

люлозы, волокна льна собираются в пучки, расположенные на периферии стебля [1], каждый пучок состоит из 15–50 элементарных волокон.

По типу конечного продукта, для которого выращивается растение, сорта культурного льна можно условно разделить на два типа: масличный лен и лен-долгунец. Для получения волокна были выведены различные сорта льна-долгунца, отличающиеся повышенной высотой растения и меньшей ветвистостью, так как элементарное волокно обычно разрывается в местах прикрепления листа [4], тогда как для получения масла, наоборот, требуются растения с большим количеством семян, соответственно, большим количеством соцветий и большей ветвистостью. Иными словами, желаемые характеристики растения могут быть противоположными в зависимости от конечной цели выращивания.

Для выведения новых сортов льна-долгунца важно исследовать генетические варианты, связанные с лучшими характеристиками волокна, такими как длина элементарного волокна, вес

Сокращения: GWAS — метод полногеномного поиска ассоциаций, QTN — однонуклеотидные маркеры количественных признаков (quantitative trait nucleotides), QEI — однонуклеотидные маркеры количественных признаков (QTN-by-environment interactions).

и содержание волокна, высота растения и диаметр стебля. Полногеномный поиск ассоциаций (GWAS – genome-wide association search) позволяет найти ассоциации между однонуклеотидными полиморфизмами (SNP – single nucleotide polymorphisms) и признаками, предоставляя значимые варианты – QTN (quantitative trait nucleotides), ассоциированные с измеренными фенотипическими признаками с учетом структуры популяции. В ранее опубликованном исследовании [5] были получены варианты, ассоциированные с несколькими признаками, указывающими на качество волокна, с помощью методов, представленных в пакете GATK.

Тем не менее, величина некоторых фенотипических характеристик может сильно зависеть от условий выращивания и отличаться при выращивании в различных условиях окружающей среды. Различная погода, температурный режим и влажность в разные годы могут создавать различные условия, которые могут повлиять на полученные фенотипические данные, поэтому применение нового метода 3VmrMLM [6] в режиме, рассчитанном на использование данных, измеренных в разных условиях окружающей среды, может предоставить новые генетические варианты, не найденные другими методами.

В данной работе мы применили метод 3VmrMLM в режиме Multi_env к данным коллекции 297 образцов культурного льна, состоящей как из льна-долгунца, так и из масличного льна, связанным с различными характеристиками волокна, измеренными в 2019–2021 гг., а также проверили влияние найденных вариантов с большим эффектом на величину соответствующего признака на независимой выборке.

МАТЕРИАЛЫ И МЕТОДЫ

Генотипические и фенотипические данные. В этом исследовании использовалась коллекция льна, состоящая из 297 образцов культурного льна из коллекции Федерального научного центра лубяных культур (Торжок, Россия). 180 образцов принадлежали к сортам льна-долгунца и 117 – к сортам масличного льна.

Были рассмотрены 5 фенотипических признаков, связанных с качеством и количеством волокна у культурного льна, а именно: средний диаметр стебля, средняя длина элементарного волокна, содержание волокна, вес волокна и высота растения. Признаки были измерены в 2019 г., 2020 г. (2 измерения) и 2021 г.

Подготовка данных. ДНК была выделена из листьев с помощью набора DNeasy Plant Mini Kit (Qiagen, Нидерланды). Секвенирование ДНК проводили в BGI с использованием протокола Illumina, генерирующего парные прочтения дли-

ной 150 пар оснований. Для сопоставления со сборкой генома NCBI ASM22429v2 использовали bwa-mem [26], для вызова вариантов использовали NGSEP версии 4.0 [27].

Анализ геномных ассоциаций. Для поиска геномных ассоциаций (GWAS) использовали пакет 3VmrMLM [6] в режиме Multi_env, для необходимого преобразования данных для использования пакета применялись TASSEL [28] и PLINK [29] со стандартными настройками согласно рекомендациям авторов метода.

Подтверждение значимости и эффекта найденных ассоциаций. Для подтверждения значимости сравнивались значения фенотипических признаков в группах растений, содержащих (генотип 0/1 или 1/1) и не содержащих альтернативный аллель (генотип 0/0), с использованием критерия Манна–Уитни, *p*-value получены с помощью соответствующих функций в R.

В случае, если носители альтернативного аллеля имели большее значение соответствующего признака, чем носители референсного, эффект QTN считался положительным, в противном случае – отрицательным. Для QEI эффект считался отдельно для каждого года, в случае, если различие в значении признака у носителей референсного и альтернативного аллеля не имело статистической значимости в данных соответствующего года, считалось, что QEI не имеет эффекта на признак.

Пересечения с генами. Была использована аннотация генов льна, предоставленная группой С. Клотье (S. Cloutier) [7].

РЕЗУЛЬТАТЫ

Были рассмотрены 5 фенотипических признаков, связанных с качеством и количеством волокна у культурного льна, измеренные для коллекции льна в 2019–2021 гг. На рис. 1 показан график корреляций между признаками, измеренными в разные годы. Можно видеть, что такие признаки, как высота растения и содержание волокна, коррелируют друг с другом при измерении в разные годы, тогда как диаметр стебля, длина волокна и вес волокна, измеренные в разные годы, не всегда одинаково коррелируют друг с другом и с другими признаками, связанными с качеством волокна. Из-за возможных разных результатов в разные годы имеет смысл использовать метод 3VmrMLM в режиме Multi_env, разработанном для получения генетических ассоциаций в разных средах.

В результате применения метода 3VmrMLM в режиме Multi_env было получено 124 QTN (quantitative trait nucleotides, однонуклеотидных маркеров количественных признаков) и 83 QEI ((QTN)-by-environment interactions, однонуклеотидных

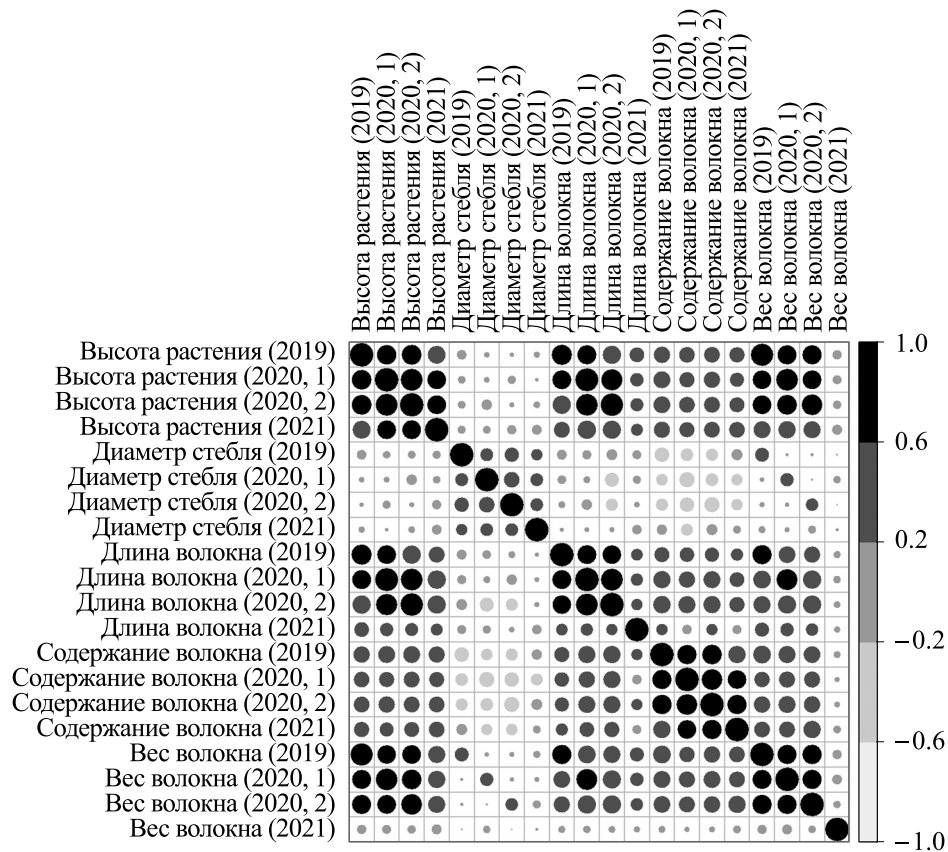


Рис. 1. Корреляции между признаками разных лет.

маркеров количественных признаков, с действием, опосредованным средой). Для подтверждения наличия эффекта влияния альтернативного аллеля на

величину признака был применен статистический тест. Найденные варианты обнаруживались на всех хромосомах (рис. 2). Больше всего QTN было расположено на 2-й, 4-й и 9-й хромосомах, больше всего QEI было расположено на 3-й и 4-й хромосомах.

Найденные QTN суммарно объясняли от 7.96% вариации для диаметра стебля до 22.15% вариации для веса волокна, тогда как QEI объясняли от 5.35% вариации для высоты растения до 32.84% вариации для веса волокна. Отдельные маркеры объясняли до 3.86% вариации признака, как можно видеть в табл. 1.

Из 124 QTN 28 попали в тело генов льна с известными функциями, из них 5 также попали в локусы количественных признаков (QTL), опубликованные в работе [7], список генов с их функциями приведен в табл. 2. 9 из 83 QEI попали в тело гена с известной функцией, кроме того, один из них попал в известный QTL, список генов показан в табл. 3.

Функции генов, в которые попали найденные варианты, связаны с различными этапами обмена веществ и энергетического обмена, ростом клеток, устойчивости к болезням и циркадными ритмами. Среди генов можно отметить различные факторы транскрипции и ферменты, такие как

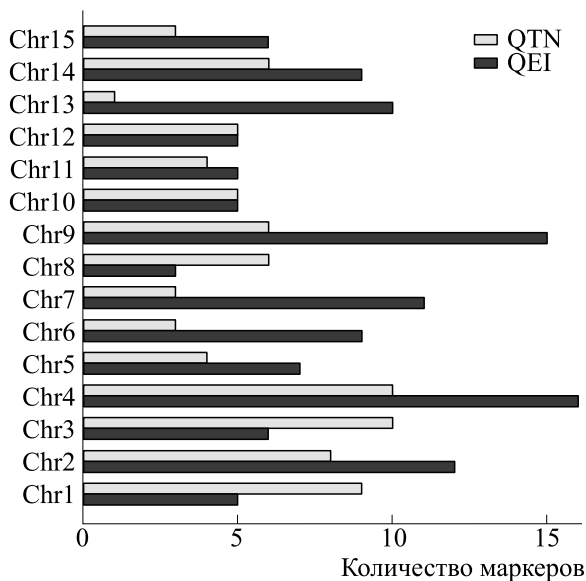


Рис. 2. Количество QTN и QEI на каждой хромосоме.

Таблица 1. Доля объясненной вариации у найденных маркеров

Признак	Тип маркера	Маркер с макс. R^2	Макс. R^2 , %	Суммарный R^2 , % для признака
Средний диаметр стебля	QTN	Chr8_9833708	1.5	7.96
	QEI	Chr3_26025687	3.86	11.05
Длина элементарного волокна	QTN	Chr2_2085748	2.47	20.7
	QEI	Chr1_10296600	2.01	24.56
Содержание волокна	QTN	Chr14_4576818	1.62	21.89
	QEI	Chr4_8501283	3.33	17.25
Вес волокна	QTN	Chr4_14603747	3.85	22.15
	QEI	Chr13_6748481	2.15	32.84
Высота растения	QTN	Chr6_60141	1.79	14.06
	QEI	Chr5_6119675	2.48	5.35

трансферазы, гидролазы и протеинкиназы, влияющие на скорость транскрипции генов и обмена веществ. Скорость обмена веществ оказывает важное влияние на рост и удлинение клеток, что является основным процессом при образовании элементарного волокна у льна. Большинство QTL, в которые попадали найденные варианты, были связаны с количеством семян и скороспелостью.

Можно обратить внимание, что QTN, объясняющие большой процент вариации, например, Chr4_14603747, объясняющий 3.85% вариации, и QEI Chr4_896405, объясняющий 2.12% вариации в величине веса волокна, попадают в тело известных генов, кроме того, QTN Chr4_14603747 также попадает в регион QTL, связанный с высотой растения, что указывает на большую важность этого варианта.

В табл. 4 показаны также варианты, не попавшие в известные гены, однако расположенные вблизи маркеров, ассоциированных с качеством волокна у льна, ранее опубликованных в работе [5]. Наличие близких друг к другу маркеров, найденных разными методами, может указывать на важные регуляторные области на 2-й и 3-й хромосоме, связанные с искомыми характеристиками волокна.

Для проверки аллельного эффекта была рассмотрена также и независимо секвенированная выборка из 100 образцов льна. Было проведено сравнение между носителями референсного и альтернативного аллеля по соответствующим признакам. На рис. 3 показаны графики для некоторых вариантов, имевших подтвержденный аллельный эффект на обеих выборках. В верхней части графиков показано значение p -value статистического теста. Можно видеть, что данные ва-

рианты, объяснявшие большой процент вариации в фенотипических данных, имели аналогичный эффект и на независимо секвенированной выборке.

ОБСУЖДЕНИЕ

В этой статье мы использовали программу 3VmrMLM для полногеномного поиска ассоциаций между однонуклеотидными полиморфизмами и признаками, связанными с качеством волокна, измеренными в 2019–2021 гг. Всего было обнаружено 124 QTN и 83 QEI, из которых 28 QTN и 9 QEI соответственно попали в тело генов с известными функциями.

Среди них можно отметить QTN Chr11_4337123 и Chr13_4673119 (см. табл. 2), оказывающие положительное влияние на длину элементарного волокна, которые попали соответственно в гены *Lus10026344* и *Lus10019679*. Ортолог первого у *Arabidopsis*, *AT2G21660.1*, влияет на рост и устойчивость в условиях сильного солевого и обезвоживающего стресса, а также обеспечивает устойчивость к замораживанию, в частности посредством регуляции открытия устьиц [8]. Известно, что растения с повышенной экспрессией этого гена демонстрируют замедленное прорастание и уменьшение роста растения. Ортолог второго, *AT2G22125.1*, является регулятором микротрубочкового цитоскелета [9], который регулирует первичный биосинтез клеточной стенки и организацию микрофибрилл целлюлозы [10, 11], что напрямую важно для характеристик волокна. Кроме того, на независимой выборке из 100 образцов также наблюдалась увеличенная длина элементарного волокна у носителей альтернативного аллеля в этих позициях (рис. 3).

Таблица 2. QTN, попавшие в тело известных генов

Признак	QTN	R^2 , %	Эффект	Название гена	Ортолог <i>Arabidopsis</i>	Описание гена
Средний диаметр стебля	Chr1_20418276	0.55	отр.	Lus10015886	AT4G00060.1	Белок семейства нуклеотидил-трансфераз
	Chr13_18748528	0.40	полож.	Lus10030551	AT3G48730.1	Глутамат-1-полуальдегид 2,1-аминомутаза 2
Длина элемент. волокна	Chr3_25464392	1.06	отр.	Lus10037556	AT1G79830.1	Кандидат голгина 5
	Chr4_15011100 (<i>QPLH-Lu4.3 (PLH)</i> ; <i>QDTM-Lu4.1 (DTM)</i>)	0.71	отр.	Lus10041603	AT5G06120.2	Суперсемейство белков с повторами ARM
	Chr6_3439867	0.89	полож.	Lus10019433	AT2G26695.1	Ran BP2/NZF суперсемейство белков, подобных цинковым пальцам
	Chr9_21214242	0.62	полож.	Lus10002632	AT5G18230.1	Регулятор транскрипции семейства белков NOT2/NOT3/NOT5
	Chr11_4337123	0.54	полож.	Lus 0026344	AT2G21660.1	Реакция на холод, циркадный ритм и связывание РНК 2
	Chr13_673119	0.30	полож.	Lus10019679	AT2G22125.1	Связывание
	Chr15_8457029	0.49	отр.	Lus10040913	AT4G04180.1	Р-петля, содержащая суперсемейство белков нуклеозид-трифосфат-гидролаз
Содержание волокна	Chr1_4014281	0.20	полож.	Lus10008627	AT2G02800.1	Протеинкиназа 2В
	Chr5_11983994	0.70	полож.	Lus10035293	AT5G39110.1	RmlC-подобный белок суперсемейства купинов
	Chr7_4641364 (<i>QPAL-Lu7.3 (PAL)</i>)	0.55	полож.	Lus10040144	AT4G17050.1	Уреидоглицин-аминогидролаза
	Chr8_20266637	0.29	полож.	Lus10024639	AT4G23440.1	Белок устойчивости к болезням (класс TIR-NBS)
	Chr9_2232548	0.44	отр.	Lus10017513	AT5G41410.1	Белок семейства POX (растительный гомеобокс)
	Chr12_1903929 (<i>QIOD-Lu12.3 (IOD)</i> / <i>QLIN-Lu12.3 (LIN)</i> / <i>QLIO-Lu12.3 (LIO)</i>)	1.45	полож.	Lus10023319	AT4G18240.1	Крахмалсинтаза 4
	Chr14_4576818	1.62	полож.	Lus 0000492	AT3G22830.1	Фактор транскрипции теплового шока A6В

Таблица 2. Окончание

Признак	QTN	R^2 , %	Эффект	Название гена	Ортолог <i>Arabidopsis</i>	Описание гена
Вес волокна	Chr3_24977526	1.05	отр.	Lus10037673	AT1G72050.2	Фактор транскрипции ША
	Chr4_14603747 (<i>QYLD-Lu4.1 (YLD)/QPLH-Lu4.3 (PLH)/QDTM-Lu4.1 (DTM)</i>)	3.85	отр.	Lus10041516	AT5G13980.1	Белок семейства 38 гликозилгидролаз
	Chr5_16671940	0.59	полож.	Lus10039645	AT5G65760.1	Белок семейства сериновых карбокси-пептидаз S28
Высота растения	Chr4_14808531 (<i>QYLD-Lu4.1 (YLD)/QPLH-Lu4.3 (PLH) QDTM-Lu4.1 (DTM)</i>)	1.02	отр.	Lus10041560	AT3G01910.1	Сульфитоксидаза
	Chr5_3947583	0.28	отр.	Lus10008279	AT1G32900.1	Белок суперсемейства UDP-гликозил-трансфераз
	Chr6_306908	0.96	отр.	Lus10010943	AT1G22400.1	Белок суперсемейства UDP-гликозил-трансфераз
	Chr6_3439859	0.62	полож.	Lus10019433	AT2G26695.1	Ran BP2/NZF суперсемейство белков, подобных цинковым пальцам
	Chr9_18403628	0.32	отр.	Lus10029677	AT1G21780.1	Белок, содержащий домен ВТВ/POZ
	Chr11_7738933	0.22	отр.	Lus10036466	AT1G72200.1	Суперсемейство белков RING/U-box
	Chr13_2137850	0.14	полож.	Lus10026989	AT2G24370.1	Протеинкиназный белок с доменом, подобным адениннуклеотид-альфа-гидролазе
	Chr14_3027258	0.26	полож.	Lus10004884	AT2G27550.1	Центрорадиальный белок
	Chr15_9777313	0.91	отр.	Lus10041208	AT1G31690.1	Белок семейства аминоксидаз меди

Примечание. В скобках после названия QTN также при наличии указан регион QTL с признаком, опубликованный в работе [7], в который попадает найденный QTN. Обозначения признаков QTL соответствуют обозначениям у авторов: PLH – высота растения; DTM – количество дней до созревания; LIN, LIO, PAL – содержание жирных кислот: линоленовой, линолевой, пальмитиновой; IOD – содержание йода; YLD – количество семян.

Таблица 3. QEI, попавшие в тело известных генов

Признак	QEI	R^2 , %	Эффект (2019 / 2020,1 / 2020,2 / 2021)	Название гена	Ортолог <i>Arabidopsis</i>	Описание гена
Длина элемент. волокна	Chr2_5973219	0,74	нет / отр. / отр. / полож.	Lus10038040	AT3G62190.1	Шаперон DnaJ-домен суперсемейства белков
	Chr8_21053023	1,41	нет / нет / нет / отр.	Lus10033967	AT5G18620.2	Фактор ремоделирования хроматина 17
	Chr15_9737862	0,26	нет / отр. / отр. / нет	Lus10041200	AT1G05440.1	С-8 стеролизомеразы
Вес волокна	Chr1_8041967	1,81	отр. / отр. / отр. / нет	Lus10035589	AT3G53090.1	Убиквитин-протеинлигаза 7
	Chr2_1279412	1,15	отр. / отр. / отр. / нет	Lus10019354	AT3G14640.1	Цитохром P450, семейство 72, подсемейство A, полипептид 10
	Chr4_896405	2,12	нет / нет / нет / полож.	Lus10030321	AT5G13550.1	Сульфатный транспортер 4.1
	Chr9_3660315	1,13	отр. / отр. / отр. / нет	Lus10040460	AT4G34490.1	Циклаза-ассоциированный белок 1
Высота растения	Chr8_22526065 (<i>QOLE-Lu8.1 (OLE)</i>)	0,73	нет / полож. / полож. / полож.	Lus10015350	AT4G12010.1	Семейство белков устойчивости к болезням (класс TIR-NBS-LRR)
	Chr15_9029961	0,40	нет / полож. / полож. / нет	Lus10041041	AT4G10120.1	Белок семейства сахарозо-фосфатсинтаз

Примечание. В скобках после названия QTN также при наличии указан регион QTL с признаком, опубликованный в работе [7], в который попадает найденный QTN. Обозначения признаков QTL соответствуют обозначениям у авторов: OLE – содержание олеиновой кислоты.

Таблица 4. QTN/QEI, расположенные на расстоянии менее 20 kb от ранее опубликованных QTN, ассоциированных с качеством волокна

Тип	Признак	Маркер	Эффект	QTN [5]	Расстояние
QTN	Длина элементарного волокна	Chr2_3842366	отр.	Chr2:17384359	3331
				Chr2:17384362	3328
QEI	Содержание волокна	Chr4_182259	нет / нет / нет / полож.	Chr4:184242	1983
QEI	Вес волокна	Chr3_5902620	нет / полож. / полож. / отр.	Chr3:5883916	18704
QEI	Содержание волокна	Chr3_5077437	полож. / нет / нет / нет	Chr3:5077352	85

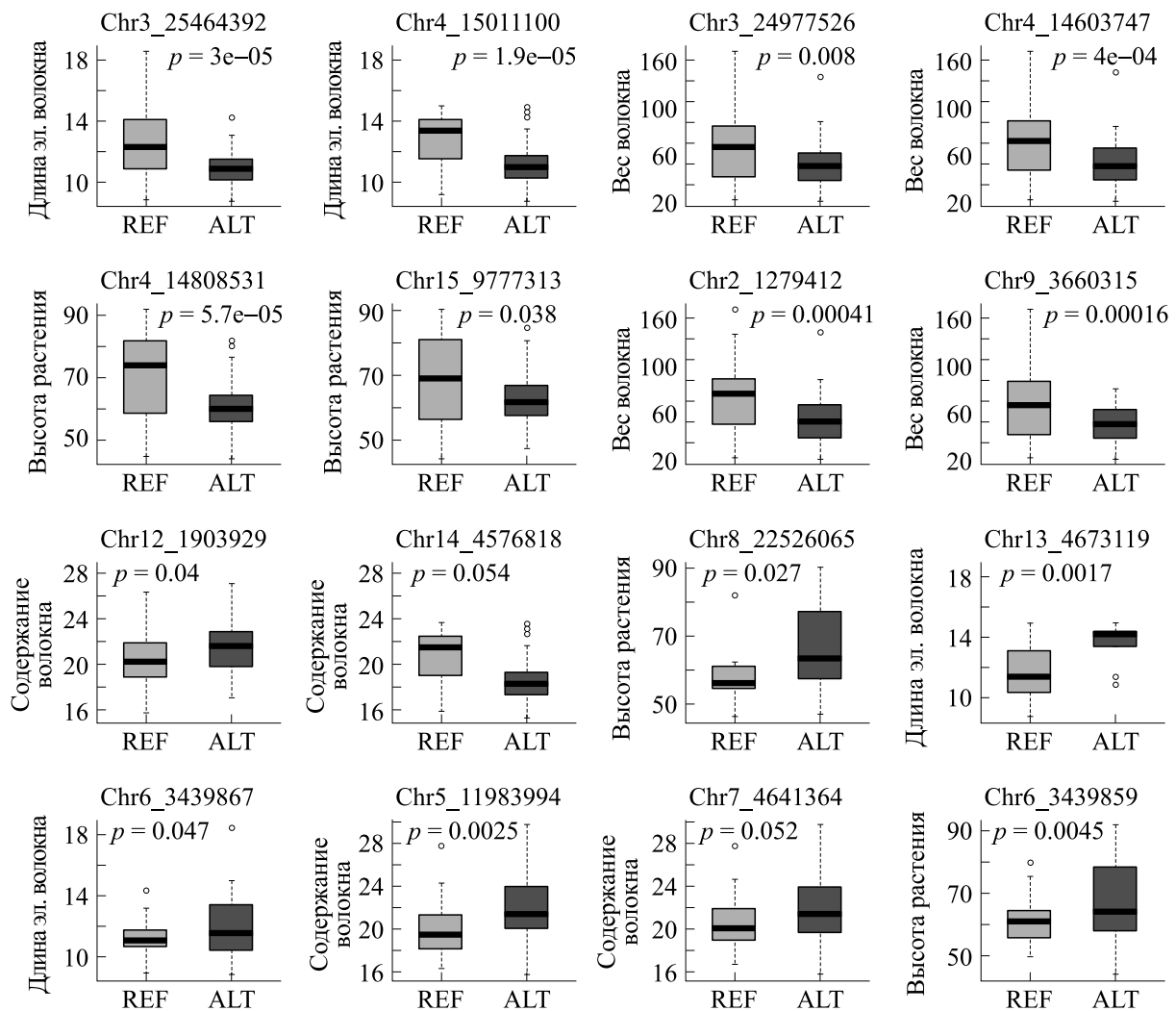


Рис. 3. Сравнение величины соответствующих признаков между носителями референсного (REF) и альтернативного (ALT) аллеля в независимой выборке из 100 образцов.

На содержание волокна оказывал положительное влияние QTN Chr7_4641364 (табл. 2), попадающий в ген *Lus10040144* (ортолог *AT4G17050.1*), который участвует в катаболизме пуриновых нуклеотидов и обеспечивает полное расщепление пуринов [12, 13]. Расщепление окисленных пуринов имеет важное значение для азотного обмена у растений [14]. Повышенное содержание волокна у носителей этого варианта наблюдалось и в независимой выборке (рис. 3).

Многие QTN и QEI оказывали отрицательный эффект на характеристики волокна, некоторые из которых связаны с ранним зацветанием. Например, можно отметить QTN Chr9_2232548 (табл. 2), попадающий в ген *Lus10017513*, ортолог которого, *AT5G41410.1*, участвует в формировании завязей семян и сигнальных путях ауксина и цитокинина [15]. Также как и QEI

Chr8_21053023 (табл. 3), который попадал в ген *Lus10033967* (ортолог *AT5G18620.2*), необходимый для поддержания вегетативной фазы растения, способный предотвращать раннюю активацию вегетативно-репродуктивного перехода [16]. Иными словами, раннее зацветание и формирование семян может отрицательно сказаться на длине стебля, которая необходима для хороших характеристик волокна. На это же указывает и попадание найденных QTN, в частности расположенных на 4-й хромосоме Chr4_15011100, Chr4_14603747 и Chr4_14808531 (табл. 2), оказывающих отрицательное влияние на характеристики волокна, в локус, связанный с ранним зацветанием и количеством семян [7].

На характеристики растений, необходимые для качества волокна, также оказывают влияние однонуклеотидные полиморфизмы, попадающие

в гены, связанные с обменом крахмала. Например, QTN Chr12_1903929 (табл. 2), оказывающий положительное влияние на содержание волокна, попадает в ген *Lus10023319*, ортолог которого *AT4G18240.1* участвует в формировании гранул крахмала, также известно, что растения с повышенной экспрессией этого гена демонстрируют более высокую скорость роста [17]. Также QTN Chr5_3947583 (табл. 2), оказывающий отрицательное влияние на высоту растения, попадает в ген *Lus10008279*, ортолог которого *AT1G32900.1*, кодирует белок, который требуется для синтеза амилозы [18] и регулирует распад крахмальных гранул в течение ночи [19].

Так же, как и гены, участвующие в обмене крахмала, есть гены, связанные с энергетическим обменом и общим качеством работы клеток, в которые попадают найденные однонуклеотидные полиморфизмы, оказывающие отрицательный эффект на вес волокна. Например, QTN Chr3_24977526 (табл. 2) попадает в ген *Lus10037673*, ортолог которого, *AT1G72050.2*, кодирует незаменимый рибосомный белок, необходимый для формирования важного компонента большой рибосомальной субъединицы [20]. QE1 Chr9_3660315 (табл. 3) попадает в ген *Lus10040460* (ортолог *AT4G34490.1*), который участвует в энергетическом обмене и регуляции удлинения и деления клеток [21].

Влияние на высоту растения – необходимую характеристику льна-долгунца, выращиваемого для получения волокна, – оказывают также и мутации в генах, связанных с работой растительных гормонов. Например, QTN Chr6_306908 (табл. 2), оказывающий отрицательное влияние на высоту растения, попадает в ген *Lus10010943*, ортолог которого *AT1G22400.1* участвует в метаболизме растительных гормонов цитокининов [22], а QTN Chr9_18403628 попадает в ген *Lus10029677* (ортолог *AT1G21780.1*), участвующий в работе комплекса лигазы белка E3 убиквитин (CUL3-RBX1-BTB), который регулирует рост корня растения и модулирует выделение гормона роста этилена [23]. QTN Chr15_9777313 (табл. 2) попадает в ген *Lus10041208* (ортолог *AT1G31690.1*), который регулирует выработку оксида азота – ключевой сигнальной молекулы, регулирующей широкий спектр физиологических процессов, включая реакцию на солевой стресс, а также модулирует рост первичного корня [24]. Отрицательное влияние вариантов Chr6_306908 и Chr15_9777313 также подтверждается и на независимой выборке (рис. 3).

Также высота растения в целом может быть связана с общим здоровьем и способностью справляться с патогенами, поэтому, например, QE1 Chr8_22526065 (табл. 3), который оказывал положительное влияние на высоту растения, по-

падает в ген *Lus10015350*, ортолог которого *AT4G12010.1* участвует в работе иммунитета растения [25], что также наблюдается и на независимой выборке (рис. 3).

Найденные варианты могут быть полезны в дальнейшей работе по улучшению сортов льна-долгунца на основе характеристик качества их волокна.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при финансовой поддержке Российского научного фонда (грант № 23-16-00037).

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов, связанных с изложенными в статье данными.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит описания собственных исследований с участием людей или животных в качестве объектов.

СПИСОК ЛИТЕРАТУРЫ

1. Goudenhooft C., Bourmaud A., and Baley C. Flax (*Linum usitatissimum* L.) fibers for composite reinforcement: exploring the link between plant growth, cell walls development, and fiber properties. *Front. Plant Sci.*, **10**, 411 (2019). DOI: 10.3389/fpls.2019.00411
2. Fogorasi M. and Barbu I. The potential of natural fibres for automotive sector – review. *IOP Conf. Ser. Mater. Sci. Eng.*, **252**, 012044 (2017). DOI: 10.1088/1757-899X/252/1/012044
3. Goudenhooft C., Bourmaud A., and Baley C. Varietal selection of flax over time: Evolution of plant architecture related to influence on the mechanical properties of fibers. *Ind. Crop. Prod.*, **97**, 56–64 (2017).
4. Rozhmina T., Bankin M., Samsonova A., Kanapin A., and Samsonova M. A comprehensive dataset of flax (*Linum usitatissimum* L.) phenotypes. *Data Brief*, **37**, 107224 (2021). DOI: 10.1016/j.dib.2021.107224
5. Kanapin A., Rozhmina T., Bankin M., Surkova S., Duk M., Osyagina E., and Samsonova M. Genetic determinants of fiber-associated traits in flax identified by omics data integration. *Int. J. Mol. Sci.*, **23**, 14536 (2022). DOI: 10.3390/ijms232314536
6. Li M., Zhang Y. W., Xiang Y., Liu M. H., and Zhang Y. M. PIVmrMLM: The R and C++ tools associated with 3VmrMLM, a comprehensive GWAS method for dissecting quantitative traits. *Mol. Plant.*, **15** (8), 1251–1253 (2022). DOI: 10.1016/j.molp.2022.06.002, PMID: 35684963
7. You F. and Cloutier S. Mapping quantitative trait loci onto chromosome-scale pseudomolecules in flax. *Methods Protoc.*, **3** (2), 28 (2020). DOI: 10.3390/mps3020028

8. Yan Y., Ham B. K., Chong Y. H., Yeh S. D., and Lucas W. J. A plant SMALL RNA-BINDING PROTEIN 1 family mediates cell-to-cell trafficking of RNAi Signals. *Mol. Plant.*, **13** (2), 321–335 (2020). DOI: 10.1016/j.molp.2019.12.001
9. Mei Y., Gao H. B., Yuan M., and Xue H. W. The *Arabidopsis* ARCP protein, CSI1, which is required for microtubule stability, is necessary for root and anther development. *Plant Cell*, **24** (3), 1066–1080 (2012). DOI: 10.1105/tpc.111.095059
10. Gu Y., Kaplinsky N., Bringmann M., Cobb A., Carroll A., Sampathkumar A., Baskin T. I., Persson S., and Somerville C. R. Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. *Proc. Natl. Acad. Sci. USA*, **107** (29), 12866–12871 (2010). DOI: 10.1073/pnas.1007092107
11. Landrein B., Lathe R., Bringmann M., Vouillot C., Ivakov A., Boudaoud A., Persson S., and Hamant O. Impaired cellulose synthase guidance leads to stem torsion and twists phyllotactic patterns in *Arabidopsis*. *Curr. Biol.*, **23** (10), 895–900 (2013). DOI: 10.1016/j.cub.2013.04.013
12. Werner A. K., Romeis T., and Witte C. P. Ureide catabolism in *Arabidopsis thaliana* and *Escherichia coli*. *Nature Chem. Biol.*, **6** (1), 19–21 (2010). DOI: 10.1038/nchembio.265
13. Shin I., Percudani R., and Rhee S. Structural and functional insights into (S)-ureidoglycine aminohydrolase, key enzyme of purine catabolism in *Arabidopsis thaliana*. *J. Biol. Chem.*, **287** (22), 18796–18805 (2012). DOI: 10.1074/jbc.M111.331819
14. Serventi F., Ramazzina I., Lamberto I., Puggioni V., Gatti R., and Percudani R. Chemical basis of nitrogen recovery through the ureide pathway: formation and hydrolysis of S-ureidoglycine in plants and bacteria. *ACS Chem. Biol.*, **5** (2), 203–214 (2010). DOI: 10.1021/cb900248n, PMID: 20038185
15. Bencivenga S., Simonini S., Benková E., and Colombo L. The transcription factors BEL1 and SPL are required for cytokinin and auxin signaling during ovule development in *Arabidopsis*. *Plant Cell*, **24** (7), 2886–2897 (2012). DOI: 10.1105/tpc.112.100164
16. Li G., Zhang J., Li J., Yang Z., Huang H., and Xu L. Imitation Switch chromatin remodeling factors and their interacting RINGLET proteins act together in controlling the plant vegetative phase in *Arabidopsis*. *Plant J.*, **72** (2), 261–270 (2012). DOI: 10.1111/j.1365-313X.2012.05074.x
17. Gámez-Arjona F. M., Li J., Raynaud S., Baroja-Fernández E., Muñoz F. J., Ovecka M., Ragel P., Bahaji A., Pozueta-Romero J., and Mérida A. Enhancing the expression of starch synthase class IV results in increased levels of both transitory and long-term storage starch. *Plant Biotechnol. J.*, **9** (9), 1049–1060 (2011). DOI: 10.1111/j.1467-7652.2011.00626.x
18. Seung D., Soyk S., Coiro M., Maier B. A., Eicke S., and Zeeman S. C. PROTEIN TARGETING TO STARCH is required for localising GRANULE-BOUND STARCH SYNTHASE to starch granules and for normal amylose synthesis in *Arabidopsis*. *PLoS Biol.*, **13** (2), e1002080 (2015). DOI: 10.1371/journal.pbio.1002080
19. Smith S. M., Fulton D. C., Chia T., Thorneycroft D., Chapple A., Dunstan H., Hylton C., Zeeman S. C., and Smith A. M. Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in *Arabidopsis*. *Plant Physiol.*, **136** (1), 2687–2699 (2004). DOI: 10.1104/pp.104.044347
20. Layat E., Cotterell S., Vaillant I., Yukawa Y., Tutois S., and Tourmente S. Transcript levels, alternative splicing and proteolytic cleavage of TFIIF control 5S rRNA accumulation during *Arabidopsis thaliana* development. *Plant J.*, **71** (1), 35–44 (2012). DOI: 10.1111/j.1365-313X.2012.04948.x
21. Barrero R. A., Umeda M., Yamamura S., and Uchimiyi H. Arabidopsis CAP regulates the actin cytoskeleton necessary for plant cell elongation and division. *Plant Cell*, **14** (1), 149–163 (2002). DOI: 10.1105/tpc.010301
22. Hou B., Lim E. K., Higgins G. S., and Bowles D. J. N-Glucosylation of cytokinins by glycosyltransferases of *Arabidopsis thaliana*. *J. Biol. Chem.*, **279** (46), 47822–47832 (2004). DOI: 10.1074/jbc.M409569200
23. Thomann A., Lechner E., Hansen M., Dumbiauskas E., Parmentier Y., Kieber J., Scheres B., and Genschik P. Arabidopsis CULLIN3 genes regulate primary root growth and patterning by ethylene-dependent and -independent mechanisms. *PLoS Genet.*, **5** (1), e1000328 (2009). DOI: 10.1371/journal.pgen.1000328
24. Groß F., Rudolf E. E., Thiele B., Durner J., and Astier J. Copper amine oxidase 8 regulates arginine-dependent nitric oxide production in *Arabidopsis thaliana*. *J. Exp. Bot.*, **68** (9), 2149–2162 (2017). DOI: 10.1093/jxb/erx105
25. Lolle S., Greeff C., Petersen K., Roux M., Jensen M. K., Bressendorff S., Rodriguez E., Sømark K., Mundy J., and Petersen M. Matching NLR immune receptors to autoimmunity in *camta3* mutants using antimorphic NLR alleles. *Cell Host Microbe*, **21** (4), 518–529 (2017). DOI: 10.1016/j.chom.2017.03.005
26. Li H. and Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25** (14), 1754–1760 (2009). DOI: 10.1093/bioinformatics/btp324
27. Tello D., Gil J., Loaiza C. D., Riascos J. J., Cardozo N., and Duitama J. NGSEP3: accurate variant calling across species and sequencing protocols. *Bioinformatics*, **35** (22), 4716–4723 (2019). DOI: 10.1093/bioinformatics/btz275
28. Bradbury P. J., Zhang Z., Kroon D. E., Casstevens T. M., Ramdoss Y., and Buckler E. S. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, **23** (19), 2633–2635 (2007). DOI: 10.1093/bioinformatics/btm308
29. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M. A., Bender D., Maller J., Sklar P., de Bakker P. I., Daly M. J., and Sham P. C. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81** (3), 559–575 (2007). DOI: 10.1086/519795

The 3VmrMLM Method Provides New Genomic Variants Associated with Fiber Characteristics in Flax

M.A. Duk^{*, **}, A.A. Kanapin^{*}, M.P. Bankin^{*}, and M.G. Samsonova^{*}

^{}Peter the Great St. Petersburg Polytechnic University, Polytekhnicheskaya ul. 29, St. Petersburg, 195251 Russia*

*^{**}Ioffe Institute, Polytekhnicheskaya ul. 26, St. Petersburg, 194021 Russia*

Flax is an important agricultural crop grown for oil and fiber. Flax fiber is used in various industries, and breeding new flax varieties with better fiber characteristics is subject of interest. Genome-wide association studies (GWAS) can find variants associated with traits important for fiber quality, but differences in data due to different growing conditions in different years reduce the power of GWAS methods. The 3VmrMLM method allows searching for variants in data measured in several environments, allowing finding new variants not found by other methods. Measurements in different years were taken as different environments, and the method found a total of 205 variants characteristic of all or several environments, 37 of which fell into the body of known genes with important functions, the effect of some variants on fiber characteristics was also confirmed in an independent set of plants.

Keywords: genomic variants, flax, fiber content, fiber length, GWAS, 3VmrMLM