

С.И. Вакула¹, Н.В. Анисимова¹, В.В. Титок², В.Н. Леонтьев³, Л.В. Хотылева¹, А.В. Кильчевский¹

МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ ГЕНЕТИЧЕСКОЙ ИЗМЕНЧИВОСТИ ЛЬНА МАСЛИЧНОГО

¹Институт генетики и цитологии НАН Беларуси

Республика Беларусь, 220072, г. Минск, ул. Академическая, 27; e-mail: svettera@yandex.ru

²ГНУ «Центральный ботанический сад НАН Беларуси»

Республика Беларусь, 220012, г. Минск, ул. Сурганова, 2в

³Белорусский государственный технологический университет

Республика Беларусь, 227006, г. Минск, ул. Свердлова, 13а

На основании оценки корреляционных связей в коллекции сортов льна масличного различного эколого-географического происхождения выделены четыре группы признаков, характеризующих: 1) продуктивность; 2) состав семян; 3) жирнокислотную композицию масла и 4) уровень накопления запасных соединений фосфора. Выделенные комплексы признаков использованы для оценки *k*-кластерной структуры исследуемого сортифта. Применение AMOVA для разделения меж- и внутригрупповой изменчивости RAPD- и ISSR-локусов позволило оценить уровень молекулярно-генетической дивергенции между кластерами сортов.

Ключевые слова: лен масличный, корреляционная плеяда, кластерный анализ, AMOVA.

Введение

Масличный морфотип льна культурного (*Linum usitatissimum* L.) выращивают, главным образом, для получения семян, используемых в производстве круп, быстросохнущих олиф и жидких сиккативов, косметических и лекарственных препаратов. Для культуры льна масличного показана широкая фенотипическая изменчивость признаков продуктивности и качества семян. В зависимости от сорта и условий произрастания высота растений льна может составлять 30–80 см, вес тысячи семян – 5–9 г [1], концентрация α -линоленовой кислоты (АЛК) в масле семян – от 3 до 70% [2]. Химический состав семян льна определяет основное направление их хозяйственного использования. Так, содержание АЛК в кулинарном льняном масле составляет около 3%, а в техническом – более 50% [3]. Создание сортов с комплексом хозяйственно-ценных признаков, отвечающих требованиям различных отраслей промышленности, – основная задача современной селекции льна масличного.

В основу любого селекционного процесса заложены оценка изменчивости исходного материала и поиск путей ее использования. В этом отношении возможности простого фенотипического скрининга ограничены как полигенной приро-

дой количественных признаков, так и эффектами генетической комплементации и эпистаза. Целью данного исследования являлась оценка изменчивости коллекции сортов льна масличного различного эколого-географического происхождения и идентификация сортов-доноров, обладающих благоприятным сочетанием хозяйственно-ценных признаков. Подходы к комплексной оценке изменчивости включали поиск ассоциаций признаков и анализ закономерностей их проявления в структуре коллекционного генофонда.

Материалы и методы

Материал исследования – коллекция 25 образцов льна масличного различного эколого-географического происхождения: Antares, Mivast, Atalante (Франция); Blue Chip (Болгария); Glenelg (Австралия); Deep Pink (Нидерланды); Linota, SU-1-10, Omega, (США); K-5827 (Уругвай); Gold Flax, McGregor, Somme, Л-6582, K-6570, Flanders (Канада); Raluca, Sandra (Чехия); Cyan (Польша); K-2398 (Китай); Воронежский, K-5627, Небесный (Россия); ЛМ-1, ЛМ-2, (Беларусь). Полевой опыт (2005–2009 гг.) и учет морфологических признаков проводили согласно «Методическим рекомендациям по селекции и семеноводству льна-долгунца» [4].

Содержание масла в семенах льна устанавливали по сухому остатку после экстракции в аппарате Сокслета [5]. Определение жирных кислот проводили по методу Welch с модификациями [6]. Для оценки содержания запасного белка в образцах семян использовали метод Кьельдаля [7]. Выход растворимых полисахаридов из водного экстракта семян рассчитывали гравиметрически, после осаждения 98%-ным этиловым спиртом [8]. Для оценки содержания в образцах семян фитина и неорганических фосфатов (P_i) использовали методы Латта и Эскин [9] и Лоури-Лопеса [10], соответственно. Термогравиметрический анализ (термоанализатор TA-4000 Mettler Toledo STARe System, Швейцария) образцов семян проводили в интервале 25–500 °С при скорости нагревания 5 °С/мин и расходе воздуха 200 мл/мин. Кривые потери массы рассчитывали при помощи программного обеспечения STARe [11]. Для элементного анализа зольного остатка семян использовали электронный микроскоп JSM-5610 LV, оснащенный системой химического анализа EDX JED-2201 JEOL (Япония) [12]. Исследования молекулярной гетерогенности коллекции проводили на основе RAPD- (8 праймеров) и ISSR-ПЦП (12 праймеров). Для статистической обработки экспериментальных данных использовали MS Excel 2007 (описательная статистика), Statistica 10.0 (корреляционный анализ и итеративная кластеризация по методу *k*-средних), SPSS Statistics (таблицы сопряженности), Arlequin 3.5 [13] (анализ молекулярной дисперсии AMOVA).

Результаты и обсуждение

Проведена оценка характера проявления 25 хозяйственно-ценных признаков у 25 образцов льна масличного из селекционных регионов, охватывающих страны Северной Америки, Западной и Восточной Европы, Азию. Оценка мер среднего уровня (среднее арифметическое, мода, медиана, асимметрия и т.д.) показала соответствие эмпирического распределения признаков закону нормального распределения. Наибольшую вариабельность в коллекции сортов льна проявляют признаки семенной продуктивности ($V^0\%$ для массы и количества семян с растения составил соответственно 27,6 и 27,5), содержание полиненасыщенных жирных кислот ($V^0\%=60,0$) и растворимых полисахаридов ($V^0\%=26,2$).

Для выявления значимых ассоциаций признаков в коллекции сортов льна масличного использован непараметрический корреляционный анализ Спирмена. Между 25 признаками показано наличие 59 статистически достоверных связей, 36 из которых представляют обратные зависимости. Сила связи между исследуемыми признаками варьирует от слабой ($r=0,2$ между содержанием пальмитиновой и стеариновой кислот) до тесной ($r>0,9$ между признаками формы и размера семян, содержанием α -линоленовой и линолевой кислот). Группы признаков, описывающие продуктивность и качество семян льна, относительно независимы. Достоверная ассоциация показана только между массой 1000 семян и содержанием масла, что отражает сопряженность отбора этих признаков при окультуривании льна масличного [1, 3]. На основании полученных данных исследуемые признаки разделены на 4 корреляционные плеяды (см. рис.), объединяющие признаки: 1) морфологии семян (площадь, периметр, длина, ширина, фактор формы и удлинение) и продуктивности (техническая длина стебля, количество коробочек, семян с растения и семян в коробочке, масса семян с растения, масса 1000 семян); 2) содержания в семенах масла, белка, зольных веществ и растворимых полисахаридов (далее – состав семян); 3) жирнокислотного состава масла (йодное число, содержание α -линоленовой, линолевой, олеиновой, стеариновой и пальмитиновой кислот); 4) содержания запасных соединений фосфора (процентное содержание фосфора в золе семян, содержание фитина, неорганического фосфата (P_i) и коэффициент фитин/P_i).

Проведен анализ сходства сортов коллекции по характеру проявления выделенных комплексов признаков. Для решения задачи оптимального разбиения коллекции на группы фенотипически близких сортов использовали итеративный алгоритм кластеризации (метод *k*-средних) [14], реализованный в пакете Statistica 10.0. Количество *k*-кластеров для разбиения определяли по результатам усечения иерархии в графике объединения сортов агломеративного кластерного анализа. Из тринадцати признаков, описывающих морфологию семян и продуктивность,

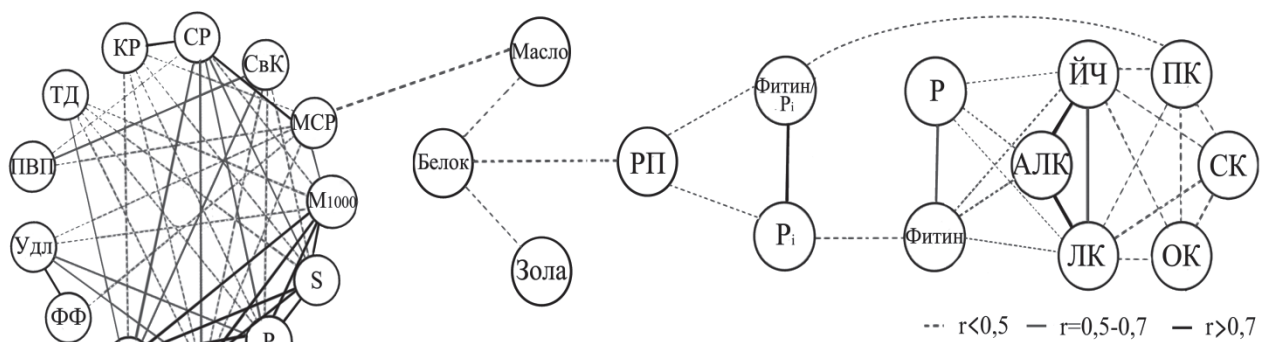


Рисунок. Значимые коэффициенты корреляции, показанные для 25 хозяйственно-ценных признаков льна масличного

Признаки продуктивности: ПВП – продолжительность вегетационного периода, с; ТД – техническая длина стебля, см; КР – количество коробочек с растения, шт.; СВК – количество семян с растения, шт.; СВК – количество семян в коробочке, шт.; МСР – масса семян с растения, г; М1000 – масса 1000 семян, г; S – площадь семян, мм²; Р – периметр семян, мм; Дл – длина семян, мм; Шр – ширина семян, мм; ФФ – фактор формы; Уд – удлинение семян. **Состав семян, % от массы:** Зола (зольные вещества); Белок; Масло; РП (растворимые полисахариды). **Концентрации жирных кислот в масле, %:** АЛК – α -линоленовой; ЛК – линолевой; ОК – олеиновой; СК – стеариновой; ПК – пальмитиновой; ЙЧ – йодное число масла. **Запасные соединения фосфора:** Фитин мг/г; Р_i – неорганический фосфат мг/г; Р – фосфор зольного остатка, % массы.

итерировано 5 *k*-кластеров, объединяющих фенотипически близкие сорта льна (табл. 1). Наиболее дистантный *k*-кластер сформирован высокопродуктивным, мелкосемянным сортом Flanders. Анализ комплексов признаков состава семян и жирнокислотной композиции масла разбил коллекцию сортов на пять и семь групп, соответственно. Высокое содержание растворимых полисахаридов и масла определяют удаленность сорта Flanders от других *k*-кластеров, сформированных по сходству состава семян. На основании проведенной оценки выделен кластер, объединяющий сорта со средним накоплением масла 45%, а также кластер сортов, перспективных для селекции на повышение кормовой ценности семян (содержание белка – более 21,0%). Уникальное соотношение жирных кислот отмечено в семенах сортов McGregor (эквивалентный уровень линолевой и олеиновой кислот) и Gold Flax (низколиноленовый сорт). Согласно сходству профилей накопления запасных соединений фосфора в коллекции выделено 6 групп сортов. Низкие уровни фосфорсодержащих соединений накапливают семена сорта Gold Flax. В один *k*-кластер объединены и наиболее высокофитиновые сорта BlueChip и ЛМ-2.

С использованием *F*-критерия подтверждена достоверность различий среднекластерных значений (при $\alpha \leq 0,05$) для 22 признаков. Для комплекса признаков продуктивности статистически не значим вклад признака «фактор формы». Средние величины содержания зольных компонентов и пальмитиновой кислоты в группах сортов, сгруппированных по признакам состава семян и композиции масла, соответственно, также достоверно не различаются. Исходя из значений евклидовых дистанций между центрами *k*-кластеров (ЕЕ), максимальной дивергенцией характеризуются группы сортов льна, раскластеризованные по сходству состава семян (ЕЕ_{мин} = 1,62). Низкий уровень расхождения отмечен для *k*-кластеров, сгруппированных на основании признаков продуктивности и формы семян (ЕЕ_{мин} = 0,83).

При генотипировании сортов коллекции с использованием маркерных систем RAPD и ISSR получено 218 ампликонов, 120 из которых полиморфны. Информационный индекс (PIC) 20 доминантных мультилокусных маркеров варьировал от 0,06 (ISSR 8a и ISSR 9) до 0,33 (Opaf16). В среднем, для анализируемой коллекции частота фиксированных рецессивных локусов составила 30%, что указывает на ее сравнительно высокое генетическое разнообразие [15, 16].

Таблица 1

Структура рабочей коллекции льна масличного

№	Продуктивность		Состав семян		
	Сорта	\bar{X}	Сорта	\bar{X}	
Дистантные <i>k</i> -кластеры	1	ВП=95,1 с.; ТД=43,1 см, МСП=0,60 г, М1000= 7,22 г	1, 7, 13, 14, 23, 24, 25,	Зола 4,0%; Белок 20,5%; Масло 43,2%; РП 22,0%	6
	2	ВП=85,6 с., ТД=37,8 см, МСП=0,50 г, М1000= 6,8 г	3, 4, 11, 19, 21	Зола 4,1%; Белок 20,6%; Масло 41,9%; РП 14,3%	1, 2, 3, 5, 7, 9, 10, 12, 13, 20, 21, 22
	3	ВП=96 с., ТД=44,6 см, МСП=0,55 г, М1000= 5,8 г	2, 8, 10, 12, 15, 16, 17, 18, 20, 22	Зола 4,1%; Белок 21,2%; Масло 43,1%; РП 9,7%	8, 11, 14, 15, 16
	4	ВП=105 с., ТД=39,0 см, МСП=0,84 г, М1000= 5,5 г	6	Зола 4,0%; Белок 21,2%; Масло 39,8%; РП 9,5%	17, 18
	5	ВП=85,6 с., ТД=50,1 см, 6МСП=0,44 г, М1000= 5,3 г	5, 9	Зола 4,0%; Белок 21,4%; Масло 41,6%; РП 14,8%	4, 19, 23, 24, 25
	Жирнокислотная композиция масла		Содержание запасных соединений фосфора		
	Сорта	\bar{X}	Сорта	\bar{X}	
Дистантные <i>k</i> -кластеры	6	АЛК=3%; ЛК= 50,9; ЙЧ=130	8	Фитин=32,3 ^a ; P _i =1,7 ^a ; Фитин/P _i =18,6;	4, 17, 21, 25
	7	АЛК=52,7%; ЛК=14,0%; ОК=18,2%; ЙЧ=181	1, 6, 15, 16, 23	Фитин=32,7; P _i =2,1; Фитин/P _i =15,7;	5, 6, 7, 11, 16, 18, 19, 23
	8	АЛК=51,2%; ЛК= 14,2%; ОК=19,8; ЙЧ=179	2, 3, 4, 18, 19, 20, 24	Фитин=35,6; P _i =2,33; Фитин/P _i =16;	3, 24
	9	АЛК=46,6%; ЛК=14,5%; ОК=22,9%; ЙЧ=170	11, 21	Фитин=33,6; P _i =1,8; Фитин/P _i =18,5;	14, 15, 20, 22
	10	АЛК=49,3%; ЛК=13,9%; ОК=20,9%; ЙЧ=175	5, 13, 14, 25	Фитин=29,7; P _i =1,6; Фитин/P _i =18,4;	8
	11	АЛК=47,3%; ЛК=15,6%; ОК=21,4%; ЙЧ=173	7, 9, 12, 17, 22	Фитин=33,8; P _i =1,8; Фитин/P _i =14,5;	1, 2, 9, 10, 12, 13
	12	АЛК=49,2%; ЛК=ОК=17,2%; ЙЧ=177	10		

Примечание. Условные обозначения – смотри примечания к рисунку. 1 – Antares; 2 – Atalante; 3 – Blue Chip; 4 – Cyan; 5 – Deep Pink; 6 – Flanders; 7 – Glenelg; 8 – Gold Flax; 9 – Linota; 10 – McGregor; 11 – Mivast; 12 – Omega; 13 – Raluca; 14 – Sandra; 15 – Somme; 16 – SU-1-10; 17 – Воронежский; 18 – К-2398; 19 – К-5627; 20 – К-5827; 21 – К-6570; 22 – Л-6582; 23 – ЛМ-1; 24 – ЛМ-2; 25 – Небесный.

Данные молекулярно-генетического полиморфизма соотнесены со структурой фенотипической изменчивости. Сортовой состав *k*-кластеров по признакам продуктивности, состава семян, композиции масла и содержания запасных соединений фосфора рассматривали как градации четырех независимых факторов дисперсионного анализа АМОВА. Для оценки значимости каждого компонента изменчивости использовали процедуру перестановок (данные не приведены). Результаты оценки доли молекулярно-генетических различий, связанных с разбиением сортифта кол-

лекции на *k*-кластеры представлены в табл. 2.

Наиболее высокий процент молекулярно-генетических различий показан между *k*-кластерами, сгруппированными по сходству признаков жирнокислотного состава масла (19,65%), наиболее низкий – между *k*-кластерами состава семян (7,75%). На основании результатов АМОВА рассчитаны генетические расстояния между *k*-кластерами, так называемые *Fst* для пар популяций. За исключением дистанции между односортовыми *k*-кластерами по композиции масла (Gold Flax и McGregor), все *Fst* значимо отличались от нуля.

Таблица 2

Распределение изменчивости ДНК-локусов по *k*-кластерам сортов коллекции льна масличного

Фактор изменчивости	Между <i>k</i> -кластерами, %	Внутри <i>k</i> -кластеров, %	Межкластерные дистанции (среднее)		r
			EE	<i>Fst</i>	
Продуктивность	19,43	80,57	0,32	1,54	0,94**
Состав семени	7,75	92,25	3,78	0,35	0,81**
ЖК состав масла	19,65	80,35	2,76	0,24	0,17
Запасные соединения Р	17,46	82,54	0,26	2,66	0,84**

Проведена оценка зависимости генетических расстояний, выраженных в *Fst*, и соответствующих им евклидовых дистанций между центрами *k*-кластеров. Для оценки силы связи использовали коэффициенты корреляции Спирмена. Для факторов «продуктивность», «состав семян», «содержание запасных соединений фосфора» показана статистически значимая (при $\alpha \leq 0,01$) зависимость между долей генетических различий (*Fst*) и фенотипической удаленностью *k*-кластеров (EE) (табл. 2). Уровень количественных различий между *k*-кластерами сортов, дивергентных по жирнокислотной композиции масла, не отражен в их генетических дистанциях, что может быть связано с ограниченностью использованного набора ДНК-маркеров, влиянием условий среды на состав масла льна или другими факторами.

Заключение

Проведена оценка связи между фенотипической структурой коллекции сортов льна масличного по комплексам признаков продуктивности, состава семян, содержания запасных соединений фосфора, жирнокислотной композиции масла и генетической гетерогенностью коллекции сортов, оцененной с использованием RAPD- и ISSR-маркеров. Доля генетической изменчивости, выявленная между четырьмя группами *k*-кластеров с использованием AMOVA, варьировала от 7,75% для фактора «состав семян» до 19,65% для фактора «композиция масла». Значимая корреляция межкластерных расстояний, выраженных в EE (фенотипическая дистанция) и *Fst* (генетическая дистанция), может свидетельствовать о накоплении в кластерах аллельных вариантов, связанных с фенотипическим проявлением комплексов признаков продуктивности, состава семян и содержания запасных соединений фосфора. Селекционная работа с

выделенными в структуре коллекции группами фенотипически и генотипически сходных сортов может способствовать дальнейшему накоплению благоприятных сочетаний аллелей, а включение их в скрещивания – созданию сортов льна масличного, отвечающих требованиям различных отраслей промышленности.

Список использованных источников

1. Черноморская, Н.М. К вопросу о внутривидовой классификации льна обыкновенного (*Linum usitatissimum* L.) / Н.М. Черноморская, А.К. Станкевич // Селекция и генетика технических культур. – 1987. – Т. 113. – С. 53–64.
2. Diederichsen, A. Cultivated flax and the genus *Linum* L. Taxonomy and germplasm conservation / A. Diederichsen, K. Richards // Flax The genus *Linum*. – London, 2003. – P. 23–33.
3. Duguid, S.D. Flax / S.D. Duguid // Oil Crops. Handbook of Plant Breeding. – 4-th ed. – NY., 2009. – P. 233–256.
4. Методические указания по проведению полевых опытов со льном-долгунцом – Л: ВНИИЛ, 1978. – 72 с.
5. Комплексный анализ состава семян льна масличного – оптимизация подходов для селекционной практики / В.В. Титок [и др.] // Труды БГТУ: Химия, технология органических веществ и биотехнология. – 2014. – № 4, (168). – С. 187–193.
6. Welch, R.W. A Micro-method for the Estimation of Oil Content and Composition in Seed Crops / R.W. Welch // J. Sci. Food Agr. – 1977. – Vol. 28, № 4. – P. 635–638.
7. Kjeldahl, J. Neue Methode zur Bestimmung des Stickstoffs in organischen Körpern (New method for the determination of nitrogen in organic substances) / J. Kjeldahl // Zeitschrift für analytische Chemie – 1883. – Vol. 22, № 1. – P. 366–383.

8. Оленников, Д.Н. Исследование процесса экстракции полисахаридов семян льна (*Linum usitatissimum* L.) / Д.Н. Оленников, Л.М. Танхаева // Химия растительного сырья. – 2007. – № 4. – С. 85–90.
9. Latta, M. A simple and rapid colorimetric method for phytate determination / M. Latta, M. Eskin // J. Agr. Food Chem. – 1980. – Vol. 28, № 6. – P. 1313–1315.
10. Скулачев, В.П. Соотношение окисления и фосфорилирования в дыхательной цепи / В.П. Скулачев. – М.: Наука, 1962. – 153 с.
11. Термический анализ и сканирующая электронная микроскопия с электронно-зондовым микроанализом в комплексных исследованиях структуры биологических объектов: [для сельскохозяйственных культур] / В.Н. Леонтьев [и др.] // Материалы, технологии, инструменты. – 2005. – Т. 10, № 4. – С. 109–115.
12. Analysis of Structural and Qualitative Features of Phytin Deposition in Ripe Flax Seeds / V. Titok [et al.] // Cytology and Genetics. – 2015. – Vol. 49, № 1. – P. 40–44.
13. Excoffier, L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows / L. Excoffier, H.E.L. Lischer // Molecular Ecology Resources. – 2010. – № 10. – P. 564–567.
14. Боровиков, В.П. Популярное введение в современный анализ данных в системе STATISTICA / В.П. Боровиков. – М.: Горячая линия – Телеком, 2015. – 288 с.
15. RAPD Analysis of 54 North American Flax Cultivars / Y.-B. Fu [et al.] // Crop Sci. – 2003. – Vol. 43, № 4, – P. 1510–1115.
16. Лемеш, В.А. Молекулярные маркеры в изучении генетических ресурсов / В.А. Лемеш // Молекулярная и прикладная генетика: сб. науч. тр. – 2008. – Т. 8. – С. 94–104.

S.I. Vakula¹, N.V. Anisimova¹, V.V. Titok², V.N. Leontiev³, L.V. Khotyleva¹, A.V. Kilcheuski¹

MULTIVARIATE STATISTICAL ANALYSIS OF THE FLAXSEED GENETIC VARIABILITY

¹Institute of Genetics and Cytology of the National Academy of Sciences of Belarus
Minsk BY-220072, Republic of Belarus

²The Central Botanical Garden of the National Academy of Sciences of Belarus
Minsk BY-220012, Republic of Belarus

³Belarusian State Technological University
Minsk BY-227006, Republic of Belarus

Four groups describing 1) the yield, 2) seed composition, 3) fatty acid ratio and 4) reserved phosphorus were identified based on their correlative relations in the collection of flaxseed varieties. These complexes were used for characterizing structure of k-means clusters for studied flaxseed varieties. For the estimation of molecular-genetic divergence level between k-clusters we estimated RAPD and ISSR variability between and among extracted groups with the use of AMOVA.

Key words: flaxseed, correlation pleiade, cluster analysis, AMOVA.

Дата поступления статьи 23 февраля 2016 г.