

УДК 631.527:631.524.7

ПРИМЕНЕНИЕ КЛАСТЕРНОГО АНАЛИЗА ДЛЯ РАЗДЕЛЕНИЯ ОБРАЗЦОВ КОЛЛЕКЦИИ ЛЬНА-ДОЛГУНЦА ПО ПРИЗНАКАМ КАЧЕСТВА ВОЛОКНА

М. А. ЛИТАРНАЯ

РУП «Институт льна»,
аг. Устье, Республика Беларусь, 221003, e-mail: malarittaML@mail.ru

(Поступила в редакцию 18.01.2019)

В статье обсуждены результаты применения кластерного анализа методом *k*-средних и целесообразность его использования в селекционном процессе. За 2011-2013 гг. было изучено 57 коллекционных образцов льна-долгунца различного эколого-географического происхождения и проведена оценка качества длинного трепаного волокна. Использование кластерного анализа методом *k*-средних позволило определить различия между коллекционными образцами льна-долгунца и разделить их на три обособленных кластера, сгруппировав по характеру проявления признаков. В первом кластере образцы представляют ценность по комплексу изучаемых признаков. Образцы второго кластера характеризуются высокими показателями горстевой длины, разрывной нагрузки и номера длинного трепаного волокна. По горстевой длине выделены образцы Гамма, Глиnum, Rod 829, L-1120 (65,7 см). В среднем за годы исследований разрывная нагрузка волокна образцов варьировала от 195,0 Н до 292,3 Н. Наибольшей крепостью обладали образцы Hercules (290,0 Н), Гамма (291,3 Н), Глиnum (292,3 Н). В зависимости от генотипа показатели гибкости варьировали в пределах 33,0-51,7 мм, наибольшее значение признака отмечено у форм льна-долгунца Українській ранній (50,3 мм) и Мрія (51,7 мм). По метрическому номеру (тонине) образец Мрія (215,1 мм/мг) достоверно превосходил сорт Оршанский 2 (191,9 мм/мг). Таким образом использование кластерного анализа позволило группировать образцы с учетом их пригодности для дальнейшей селекционной работы методом гибридизации. При подборе родительских компонентов с целью получения форм льна, характеризующиеся высокими показателями качества длинного трепаного волокна, целесообразно использовать образцы Гамма, Глазур, Глиnum, Глухівській юбілейний, К-6, L-1120, Мрія, Bertelin, Hercules, nameless (вн-70).

Ключевые слова: лен-долгунец, образец, показатели качества, волокно, кластерный анализ.

The article discusses the results of cluster analysis using the *k*-means method and the feasibility of its use in the breeding process. In 2011-2013, 57 collection samples of flax of different ecological and geographical origin were studied and the quality of the long frayed fiber was assessed. The use of cluster analysis by *k*-means allowed to determine the differences between the collection samples of flax and divide them into three separate clusters, grouped by the nature of the signs. In the first cluster, the samples are valuable for the complex of studied features. The samples of the second cluster are characterized by high indices of the mountain length, breaking load and number of the long frayed fiber. For gostevoi the length of selected samples of Gamma, Glinum, Rod 829, L-1120 (65,7 cm). On average, during the years of research, the breaking load of fiber samples ranged from 195.0 N to 292.3 N. The greatest strength possessed samples Hercules (290,0 N), Gamma (291,3 N), Glinum (292,3 N). Depending on the genotype, the flexibility indices varied within 33.0-51.7 mm, the highest value of the trait was observed in the forms of flax of Ukrain's'kij rannij (50.3 mm) and Mriya (51.7 mm). At the metric number (fineness) sample of Mriya (of 215.1 mm/mg) were significantly superior to grade Orshanskij 2 (191.9 mm/mg). Thus, the use of cluster analysis allows you to group samples, taking into account their suitability for further breeding work by hybridization. In the selection of parental components for the purpose of receiving forms of flax characterized by high indicators of quality long trebanog fiber, it is advisable to use samples of Gamma, icing on it., Glinum, Gluhivs'kij yubilejnyj, K-6, L-1120, Mriya, Bertelin, Hercules, nameless (HV-70).

Key words: flax, sample, quality indicators, fiber, cluster analysis.

Введение

Селекция всегда связана с отбором. Одним из условий организации успешного селекционного процесса является умение выделять лучшие генотипы растений среди большего разнообразия их форм: будь то местный материал, подвергшиеся длительному естественному или сознательному отбору, или мировые ассортименты, включающие наилучшие мировые стандартные сорта. Исходный материал, собранный из разных стран мира, в условиях, существенно отличающихся от условий его дальнейшего использования в той или иной зоне, несет в себе новую наследственную информацию, на этой основе можно получить конкурентоспособную популяцию [1, 2].

Проблеме качества получаемой продукции посвящена вся земледельческая наука. Решается она двумя главными путями: созданием (по возможности) оптимальных условий роста и развития растений и выведением новых сортов, хорошо приспособленных к местным условиям и отвечающим требованиям потребителя. Если первому пути отвечает специфика таких агрономических дисциплин, как земледелие, растениеводство, агрохимия, в задачу которых

входит изучение и воздействие на условия возделывания сельскохозяйственных культур. Второй путь решается селекцией, которая разрабатывает способы воздействия на растения для изменения их наследственности в нужном направлении и создает новый исходный материал, отвечающий условиям произрастания и требованиям производства [10].

Исходя из потребности промышленного сектора производства в урожайных сортах с высоким качеством льноволокна, важнейшей задачей селекции является создание сортов с высокими прядильными свойствами льноволокна, определяющими конкурентоспособность продукции.

В селекции накоплен огромный материал по изучению растений, но работа с ним имеет свои особенности. Нередко исследователи проверяют сотни и тысячи образцов, чтобы найти те, которые имеют ценные генетические признаки и будут полезны для селекции в конкретном регионе, поэтому применение и внедрение в практику селекционных исследований новых методов анализа информации, является на сегодняшний день важной задачей. В настоящее время особый интерес представляет кластерный анализ в селекции растений. Кластерный анализ используется при изучении генетического родства; установлении изменчивости хозяйственно-полезных признаков сортов растений под воздействием разнообразных условий среды; дифференциации, идентификации, создании баз данных сортов сельскохозяйственных культур на молекулярно-генетическом уровне [3, 4, 5, 11].

Достоинство кластерного анализа в том, что он позволяет группировать объекты не по одному параметру, а по целому набору признаков. Кроме того, он в отличие от большинства математико-статистических методов, не накладывает особых ограничений на вид рассматриваемых объектов, и позволяет рассматривать множество исходных данных. Это имеет большое значение, когда признаки имеют разнообразные показатели.

ТАКИМ ОБРАЗОМ, В ЗАДАЧУ НАШИХ ИССЛЕДОВАНИЙ ВХОДИЛО: ИЗУЧИТЬ КОЛЛЕКЦИОННЫЕ ОБРАЗЦЫ И С ПОМОЩЬЮ МЕТОДА КЛАСТЕРИЗАЦИИ ВЫДЕЛИТЬ ИСТОЧНИКИ с высоким качеством длинного трепано волокна для целей селекции.

Основная часть

Исследования по изучению коллекции проводили на опытных участках РУП «Институт льна» путем закладки полевых опытов, проведения сопутствующих наблюдений и учетов согласно методическим указаниям по изучению коллекции льна (*Linum usitatissimum* L.) [6]. Почва опытных участков дерново-подзолистая, среднесуглинистая, развивающаяся на лессовидном суглинке, подстилаемой с глубины 1 м моренным суглинком, со следующими агрохимическими показателями (табл. 1), по благоприятным предшественникам согласно отраслевого регламента [7].

Таблица 1. Агрохимические показатели почвы

Показатель	Годы		
	2011	2012	2013
Кислотность почвы, pH (KCl)	5,8	5,2	5,9
Содержание гумуса, %	1,8	1,9	1,8
Содержание подвижного фосфора, мг/кг почвы	248,0	228,8	168,3
Содержание обменного калия, мг/кг почвы	219,0	157,5	193,8

Метеорологические условия в годы проведения исследований характеризовались различным температурным режимом, значительной периодичностью и характером выпадения осадков, гидротермический коэффициент в 2011 году составил 1,4 (год характеризовался как влажный), в 2012 году – 1,24 (год слабозасушливый), в 2013 году – 0,92 (год засушливый), что способствовало более объективной оценке коллекционного материала [8].

В качестве исходного материала для исследований было отобрано 57 коллекционных образцов различного эколого-географического происхождения, которые изучали по следующим признакам: горстевая длина (см), цвет (группа), гибкость (мм), разрывная нагрузка (Н), метрический номер (мм/мг), средний номер волокна, относительная разрывная нагрузка расчетная (сН/текс).

Инструментальную оценку качества волокна проводили в лаборатории качества льнопродукции института. Полученные данные были подвержены статистической обработке в пакете программ Statistica 10.0 в модуле Cluster Analysis методом k-средних [9], для проведения которого значения признаков образцов коллекции были стандартизированы (переменные имеют нулевое среднее и единичную дисперсию).

Применение программы Statistica позволило распределить данные образцы коллекции на три обособленных кластера. Результаты классификации можно видеть на рисунке 1, где горизонтальная ось представляет изучаемые признаки, вертикальная – соответствующие этим признакам баллы.

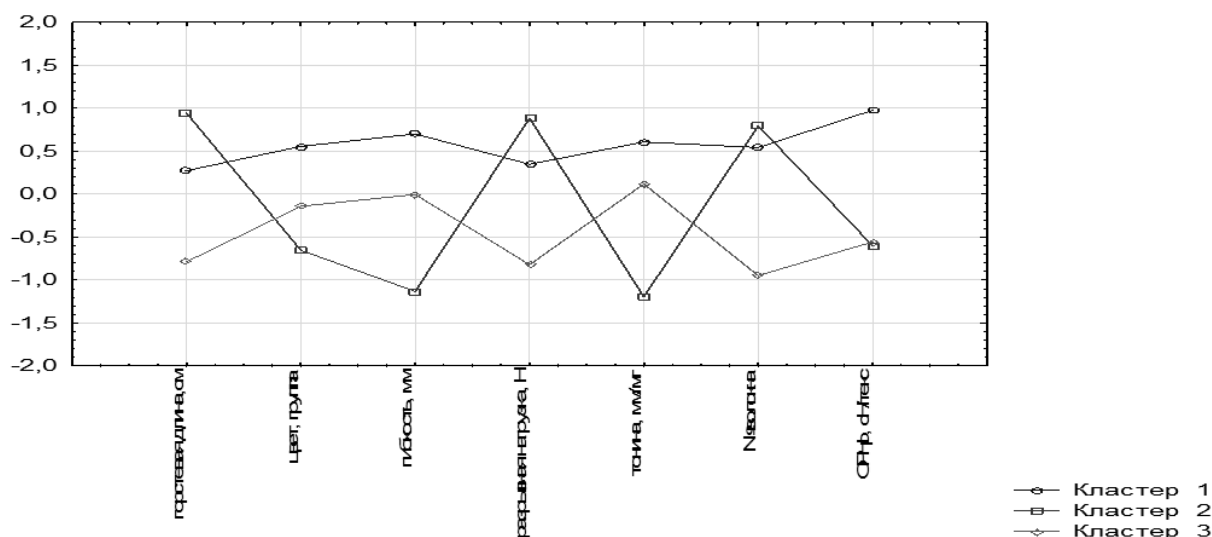


Рис. 1. График средних значений кластеров

В табл. 2 представлен дисперсионный анализ признаков качества длинного трепаного волокна, где установлены достоверные отличия между образцами коллекции по всем изучаемым признакам по годам исследований.

Таблица 2. Дисперсионный анализ признаков качества длинного трепаного волокна (среднее 2011–2013 гг.)

Признак	Межгрупповая		Внутригрупповая		F-статистика	значим. P
	SS	df	SS	df		
горстевая длина, см	27,47542	2	28,52458	54	26,007	0,000000
цвет, группа	12,50086	2	43,49914	54	7,759	0,001091
гибкость, мм	26,97280	2	29,02720	54	25,089	0,000000
разрывная нагрузка, Н	28,20032	2	27,79968	54	27,389	0,000000
метрический номер, мм/мг	26,63697	2	29,36303	54	24,493	0,000000
№ волокна	35,06339	2	20,93661	54	45,218	0,000000
ОРНр, сН/текс	31,82693	2	24,17307	54	35,549	0,000000

Кластерный анализ позволяет рассматривать достаточно большой объем информации и сжимать большие массивы информации, делать их компактными и наглядными. Средние значения признаков выделенных кластеров достоверно различаются между собой (табл. 3), что подтверждает правильность их формирования. В первый кластер вошли образцы, которые представляют ценность по комплексу изучаемых признаков. Образцы второго кластера представляют ценность по горстевой длине, разрывной нагрузке и номеру длинного трепаного волокна.

Таблица 3. Средние значения признаков выделенных кластеров (среднее 2011–2013 гг.)

Признак	Среднее значение кластеров			Общее среднее	НСР _{0,05}
	Кластер 1	Кластер 2	Кластер 3		
горстевая длина, см	59,4	63,0	53,7	58,7	2,2
цвет, группа	3,7	3,3	3,4	3,5	0,2
гибкость, мм	45,3	37,4	42,3	41,6	2,5
разрывная нагрузка, Н	251,0	263,9	223,0	246,0	13,0
метрический номер, мм/мг	177,9	137,9	167,0	160,9	11,8
№ волокна	11,7	11,9	10,3	11,3	0,5
ОРНр, сН/текс	14,1	13,0	13,0	13,4	0,4

Горстевая длина варьировала у образцов коллекции в пределах 44,3–65,7 см. Достоверно выше в сравнении со стандартом Оршанский 2 она была у образцов Silva (63,0 см), ВНИИЛ-9, Алей (63,3 см), Мрія, Глазур, К-6 (64,0 см), Ярок (64,3 см), С-108, Bertelin, (64,7 см), Могилевский (65 см). Наибольшее значение этого признака отмечено у образцов Гамма, Гліум, Rod 829, Л-1120 (65,7 см) (табл. 4).

Цвет волокна характеризует степень зрелости и чистоты волокна, зависит от технологии возделывания и уборки посевов, погодных условий, технологии приготовления льнотресты и других факторов. Среди изучаемых генотипов коллекции лишь 12,2 % имели по цвету четвертую группу – это Велижский кряж (вн-25), Sebeco 7411 NL, nameless (вн-70), L.Sussex, ВНИИЛ-8, Hercules и Глухівській юбілейний.

Крепость пряжи, а также ее обрывность в технологическом процессе зависит от разрывной нагрузки. В среднем за годы исследований разрывная нагрузка волокна образцов варьировала от 195,0 до 292,3 Н. Достоверно большую, чем у стандарта Оршанский 2 (272,7 Н), разрывную нагрузку имели образцы Hercules (290,0 Н), Гамма (291,3 Н), Гліум (292,3 Н). В пределах НСР находились образцы № 881 Concurrent (284,3 Н), ВНИИЛ-9 (281,7 Н), Гладіатор (273,0 Н).

Таблица 4. Образцы коллекции льна-долгунца, выделившиеся по качеству длинного трепаного волокна (среднее 2011–2013 гг.)

Название образца	Кластер	Горстевая длина, см	Цвет, группа	Гибкость, мм	Разрывная нагрузка, Н	Тонина, мм/мг	№ волокна	ОРНр, сН/текс
Bertelin	1	64,7	3,0	46,3	258,0	166,3	12,0	14,2
Hercules	1	61,7	4,0	48,0	290,0	181,6	12,7	15,2
L. Sussex	1	60,3	4,0	45,0	239,0	177,4	11,7	13,8
nameless	1	60,3	4,0	46,7	251,7	197,2	12,3	14,5
Rod 829	2	65,7	3,0	33,0	249,3	158,4	11,7	12,5
Silva	2	63,0	3,3	35,3	263,3	141,7	11,7	12,8
Алей	2	63,3	3,6	41,2	261,6	148,3	12,4	13,5
Альтгаузен	1	56,0	3,7	46,7	269,0	179,2	11,7	14,6
Веста	2	61,0	3,3	36,7	267,7	144,6	11,7	13,1
ВНИИЛ-9	1	63,3	3,7	37,3	281,7	177,1	12,3	13,9
Гамма	2	65,7	3,7	39,0	291,3	115,7	13,0	13,4
Глазур	1	64,0	3,3	44,7	236,0	159,9	12,0	13,5
Гліну́м	2	65,7	2,7	35,0	292,3	113,1	12,0	13,0
Глухівський юбілейний	1	61,3	4,0	49,3	221,3	163,0	12,0	13,7
К-6	2	64,0	3,7	43,7	250,7	139,2	12,3	13,4
Л-1120	1	65,7	3,3	45,0	246,0	190,9	12,5	14,1
Могилевский	2	65,0	3,6	38,2	237,9	156,8	12,2	12,8
Мрія	1	64,0	3,7	51,7	253,0	215,1	12,7	15,2
Оршанский 2	1	60,7	3,7	40,7	272,7	191,9	12,0	14,2
С-108	2	64,7	3,0	36,7	253,3	131,7	12,0	12,6
Ярок	2	64,3	3,3	36,8	250,0	127,1	11,8	12,5

Очень важно, чтобы волокно обладало одновременно хорошими показателями как прочности, так и гибкости, так как заметное снижение одного из них отрицательно сказывается на прядильных свойствах волокна. Волокно, не обладающее достаточной гибкостью, непригодно для получения тонкой пряжи. Достоверно большую гибкость, чем у стандарта (40,7 мм) имели почти половина или 42,1 % изученных нами образцов. В зависимости от генотипа показатели гибкости варьировали в пределах 33,0-51,7 мм. Наибольшее значение признака отмечено у следующих образцов: Українській ранній (50,3 мм) и Мрія (51,7 мм).

По метрическому номеру (тонине) был выделен лишь один образец Мрія (215,1 мм/мг), достоверно превосходивший сорт Оршанский 2 (191,9 мм/мг).

Заключение

Применение кластерного анализа методом *k*-средних позволяет группировать образцы с учетом их пригодности для дальнейшей селекционной работы методом гибридизации. Для подбора родительских компонентов целесообразно использовать образцы Гамма, Глазур, Гліну́м, Глухівський юбілейний, К-6, Л-1120, Мрія, Bertelin, Hercules, nameless (вн-70), обладающие высокими показателями длинного трепаного волокна, имеющие номер волокна 12-13.

ЛИТЕРАТУРА

1. Лемеш, В. А. Микросателлитный анализ стародавних белорусских сортов льна / В. А. Лемеш, М. В. Богданова, В. И. Сакович // Льноводство: реалии и перспективы: материалы международной научной конференции, аг. Устье, 27–28 июня 2013г. / РУП «Институт льна»; под ред. И.А. Голуб [и др.]. – Могилев, 2013. – С. 93–96.
2. Труды по прикладной ботанике, генетике и селекции / ред.кол. Н. И. Дзюбенко, О. П. Митрофанова [и др.]. – Т. 164 – СПб., 2007. – 400 с.
3. Касьяненко, А. Н. Использование многомерного статистического анализа в селекции растений / А. Н. Касьяненко // Тез.докл. Всесоюз. совещ. – Симферополь-Ялта, 1989. – С. 38–39.
4. Руанет, В. В. Использование искусственных нейронных сетей для решения частных задач генетики и селекции: дис. ... канд. биол. наук: 03.00.15 / В. В. Руанет. – М., 2003. – 108 с.
5. Kuruvadi, S. Multivariate analysis of genetic divergence in wheat. / Turrialba. – 38, № 4. – P. 267–271.
6. Методические указания по изучению коллекции льна (*Linum usitatissimum* L.) / В. З. Богдан [и др.]. – РУП «Институт льна», Устье, 2011. – 13 с.
7. Отраслевой регламент. Возделывание льна-долгунца. Типовые технологические процессы / Гусаков В.Г. [и др.]. – Минск. Ин-т системных исследований, 2012. – 48 с.
8. Агрометеорологический бюллетень / ГУ «Республиканский гидрометеорологический центр»; редактор Н. В. Мельчакова, начальник И. А. Полишук. – 2011–2015 гг.
9. Буреева, Н. Н. Многомерный статистический анализ с использованием ППП «STATISTICA» / Н. Н. Буреева // Учебно-методический материал по программе повышения квалификации «Применение программных средств в научных исследованиях и преподавании математики и механики». – Нижний Новгород, 2007. – 112 с.
10. Богдан, В. З. Создание и изучение исходного материала льна-долгунца для селекции сортов раннеспелой биологической группы: дис. ... канд. с/х наук: 06.01.05 / В. З. Богдан. – Жоди́но, 2002. – 116 с.
11. Девис, Дж. С. Статистический анализ данных в геологии / Дж. Девис. – М.: Недра, 1990. – Т. 1. – 319 с., Т. 2. – 427 с.