

УДК 631.527:633.34
МРНТИ 68.35.37

DOI: <https://doi.org/10.37788/2025-2/103-114>

Н.Ф. Григорчук^{1*}, В.А. Аремьева¹, С.М. Абитаева¹

¹ТОО «Опытное хозяйство масличных культур», Казахстан

^{*}(e-mail: nataly.grigorchuk@gmail.com)

Цифровизация селекционного процесса и применение кластерного анализа в развитии селекции сои

Аннотация

Основная проблема: Не смотря на то что, Казахстан ежегодно наращивает посевные площади сои ее возделывание затруднено в северных и восточных регионах из за ограниченного количества сортов адаптированных к климатическим условиям. Традиционные методы подбора родительских форм для создания новых ультраранних и ранних сортов сои требуют значительных временных и ресурсных затрат, поэтому современные селекционные технологии направлены на повышение точности и эффективности гибридизации. В условиях интенсификации сельского хозяйства и необходимости адаптации сортов сои к конкретным агроэкологическим условиям, кластерный анализ зарекомендовал себя как эффективный инструмент оценки и отбора сортового материала. Его применение позволило не только структурировать изученные образцы сои по хозяйственно ценным признакам, но и выделить наиболее перспективные формы для последующего использования в селекционной работе. В статье приведены результаты исследования 102 коллекционных образцов сои различного эколого-географического происхождения по основным хозяйственно ценным признакам с целью определения критериев отбора источников и доноров в селекции на высокую продуктивность, скороспелость, высокие биохимические показатели при помощи кластерного анализа методом Варда с использованием программы Statistica v.13.

Цель: Провести кластерный анализ 102 коллекционных образцов сои для систематизации по хозяйственно ценным признакам и определения генетически перспективных родительских форм, с целью повышения эффективности селекционного процесса и ускорения создания новых сортов, адаптированных к условиям северных и восточных областей Казахстана.

Методы: Исследование проводилось с применением метода иерархической кластеризации по методу Варда. Статистическая обработка проводилась в Statistica v.13.

Результаты и их значимость: В результате изучения 102 коллекционных образцов сои с помощью кластерного анализа было выделено пять кластеров, различающихся по комплексу хозяйственно ценных признаков. Из первого кластера выделилось четыре образца как источники и доноры высокой урожайности и высокого содержания белка в семенах. Из второго кластера шесть образцов как источники и доноры скороспелости, пять образцов как источники высокого содержания белка в семенах. Из третьего кластера четыре скороспелых образца, три образца по урожайности, пять – по содержанию белка, три – по содержанию жира. Из четвертого кластера выделилось по уровню урожайности три образца, два – по массе 1000 семян, два – по содержанию белка и четыре – по высоте прикрепления нижних бобов. В пятом кластере семь образцов сои по урожайности и три – по высоте прикрепления нижних бобов. Кластерный анализ подтвердил свою эффективность как инструмент цифровизации селекционного процесса, что способствует ускоренному созданию адаптированных сортов сои и расширению ареала её возделывания в Казахстане.

Ключевые слова: соя, селекция, хозяйственно ценные признаки, образец, кластерный анализ, цифровизация, Ward's метод, дендрограмма, цифровые технологии.

Введение

Соя является стратегически важной культурой, востребованной в пищевой промышленности, животноводстве и производстве растительных масел [1]. В условиях Казахстана актуальной проблемой является расширение ареала её возделывания, связанная с

разнообразием климатических условий, ограниченной влагообеспеченностью и резкими колебаниями температуры в различных регионах [2].

По данным Министерства сельского хозяйства Республики Казахстан в 2024 году посевные площади сои составили около 100 тыс. гектаров. В Восточно-Казахстанской области, в частности в Глубоковском, Зайсанском и Шемонаихинском районах посевы сои занимают 1,3 тыс. гектаров. Основные площади посевов сои (более 85 %) сосредоточены в Алматинской области, где благоприятные климатические условия для возделывания сои [3].

Для расширения посевных площадей в восточных и северных областях необходимо создание новых сортов сои, которые обладают широкой экологической адаптивностью, устойчивостью к абиотическим стрессам и способностью стабильно формировать высокие урожаи в неблагоприятных условиях.

Традиционные методы подбора родительских форм для создания новых ультраранних и ранних сортов сои требуют значительных временных и ресурсных затрат, поэтому современные селекционные технологии направлены на повышение точности и эффективности гибридизации [4, 5, 6]. Создание конкурентоспособных сортов сои с нужными хозяйственно-ценными признаками, такими как короткий вегетационный период, высокая урожайность, повышенное содержание белка и жира в семенах, высокое прикрепление нижних бобов является приоритетной задачей селекции.

С развитием цифровых технологий большую роль играет кластерный анализ, который позволяет анализировать и систематизировать большой объем данных, группировать сорта по хозяйственно ценным признакам и выявлять перспективные комбинации для скрещивания [7]. Такой подход повышает обоснованность выбора родительских форм, способствует более рациональному планированию гибридизации и ускоряет создание новых сортов, адаптированных к специфическим региональным условиям Казахстана.

Одним из наиболее эффективных методов кластерного анализа является иерархическая кластеризация, в частности, метод Варда (Ward's method) [8]. Этот метод характеризуется способностью формировать группы сортов с минимальной внутрикластерной дисперсией и максимальной разницей между кластерами. Главная задача метода – минимизировать увеличение внутрикластерной дисперсии на каждом этапе объединения, что обеспечивает более точное и информативное разделение сортов на однородные группы [9]. Результаты кластеризации удобно визуализировать с помощью дендрограммы – древовидной диаграммы, которая наглядно отображает процесс объединения кластеров. По дендрограмме определяется оптимальное количество кластеров, выявляются наиболее близкие по хозяйственно ценным признакам сорта, что значительно упрощает выбор перспективных родительских пар для гибридизации [10].

Применение кластерного анализа в селекции сои позволяет:

- оптимизировать выбор родительских форм. Исходя из выделенных групп, можно выбирать сорта, обладающие контрастными или дополняющими характеристиками, что повышает вероятность успешной гибридизации.

- уменьшить временные затраты, и заранее исключить малоперспективные комбинации.

- упрощает анализ данных. Появляется возможность систематического подхода к гибридизации, что позволяет анализировать большое количество объемов данных и находить закономерности в наследовании признаков.

- увеличить вероятность получения высокопродуктивных форм. При выборе родительских форм из разных кластеров существует высокая вероятность появления гетерозиса.

- выбирать образцы с высоким содержанием белка и жира в семенах и коротким вегетационным периодом.

- создать высокопродуктивные конкурентоспособные сорта, которые соответствуют требованиям современного сельского хозяйства [11].

В селекции сои кластерный анализ применяется для определения генетической близости сортов, выявления групп со схожими характеристиками и подбора родительских пар с комплементарными признаками. Для гибридизации подбираются родительские формы из разных кластеров, чтобы усилить генетическое разнообразие. Для сохранения хозяйственно ценных признаков подбираются родительские формы из одного кластера. На основе дендрограммы можно выделить сорта с высоким содержанием белка и жира в семенах для применения в пищевой и масложировой промышленности, с коротким вегетационным

периодом – для выращивания в северных и восточных областях [12]. Внедрение в производство новых сортов сои с комплексом хозяйственно ценных признаков, адаптированных к неблагоприятным факторам окружающей среды, позволит расширить ареал возделывания сои в Казахстане.

Целью исследования является системная оценка сортов, объединение их в группы на основе хозяйственно ценных признаков с использованием кластерного анализа, для обоснованного использования генетического разнообразия и повышения эффективности их использования в селекционном процессе.

Материалы и методы

Исследования проводились на опытном поле научного севооборота ТОО «Опытное хозяйство масличных культур» в лаборатории селекции сои на протяжении 2019-2023 гг. Хозяйство расположено в южной части Глубоковского района, в предгорной зоне Восточно-Казахстанской области. Климатические условия хозяйства характеризуются резко-континентальным климатом. Количество осадков и температурный режим за вегетационный период представлены в таблицах 1 и 2. Почвы преимущественно мощные слабо выщелоченные и типичные черноземы с содержанием гумуса – 5-17 %.

Таблица 1 – Количество осадков за вегетационный период, мм.
Данные ТОО «ОХМК», 2019-2023гг.

Год	Месяц					Сумма осадков за вегетационный период, мм
	Май	Июнь	Июль	Август	Сентябрь	
2019	60,0	166,0	111,0	51,0	11,0	399,0
2020	4,6	31,0	39,6	39,4	66,2	180,8
2021	4,6	86,6	55,1	67,0	10,8	224,1
2022	10,6	66,0	43,2	12,0	0,0	131,8
2023	34,8	9,0	12,2	84,4	103,0	243,4
Ср.многолетние	46,0	59,0	64,0	47,1	32,0	248,1

Таблица 2 – Температура воздуха в течении вегетационного периода, °С
Данные ТОО «ОХМК», 2019-2023гг.

Год	Месяц					Среднемесячная температура за вегетационный период, °С
	Май	Июнь	Июль	Август	Сентябрь	
2019	10,1	20,8	23,9	21,6	13,7	18,0
2020	16,8	17,3	20,7	18,3	10,6	16,8
2021	14,9	17,9	20,7	18,2	14,2	17,2
2022	17,9	19,1	19,8	21,2	15,5	18,7
2023	12,9	20,4	21,3	19,2	11,90	17,1
Ср.многолетние	13,7	18,9	21,2	19,1	12,9	17,2

Изучались 102 коллекционных образца сои из различных эколого-географических зон – Казахстана, России, Украины, Белорусии, Китая, Канады, США и Европы. Названия сортов сои представлены в зашифрованном виде. Посев осуществлялся вручную, площадь делянки – 8 м², глубина заделки семян – 5 см. Контролем служил сорт сои- Берлік KB.

Полевые испытания, учеты и наблюдения проводились по методике ВИР [13]. Для определения урожайности применялся измерительно-весовой метод [14]. Анализ сои на содержание белка и жира в семенах проводился на инфракрасном анализаторе Инфраскан-3150. Статистическая обработка данных в соответствии с методикой полевого опыта Доспехова Б.А. [15], с использованием программ Excel, Statistica 6.0. Анализ хозяйственно ценных признаков сортов сои проводился с использованием метода иерархической кластеризации по методу Варда (Ward's method) [8], реализованного в пакете Statistica v.13 (StatSoft, DellInc.). В качестве

меры сходства использовалось евклидово расстояние, что позволило объективно оценить степень различия между сортами по множеству хозяйственно ценных признаков [10].

Результаты

В рамках настоящего исследования был проведён кластерный анализ 102 коллекционных образцов сои, выращенных в условиях опытного поля ТОО «ОХМК». В качестве исходных данных были использованы хозяйственно ценные признаки, включая продолжительность вегетационного периода, урожайность семян, масса 1000 семян, высота прикрепления нижних бобов и содержание белка и жира в семенах. Все показатели были нормированы для обеспечения сопоставимости данных.

Кластеризация осуществлялась методом иерархической агломерации с использованием меры расстояния Евклида и метода связи Варда. Согласно оценке степени различия между образцами по множеству хозяйственно ценных признаков на дендрограмме выделились 5 кластеров, что свидетельствует о наличии устойчивых групп образцов с различной степенью схожести по изученным признакам.

На основе иерархической кластеризации методом Варда были выделены пять групп образцов сои, объединённых по схожим признакам (табл. 3). Эти группы помогают определить потенциальные родительские пары для гибридизации с целью улучшения ключевых характеристик.

Таблица 3 – Распределение сортов по кластерам и группам спелости.

Кластер	Ультраранние (до 90)	Раннеспелые (91–100)	Ранние (101–110)	Среднеранние (111–120)
1	-	12	2	-
2	9	10	-	-
3	2	33	-	-
4	-	1	14	-
5	-	1	16	2

Из таблицы 3 видно распределение образцов по кластерам и группам спелости. В первом кластере выделено 12 образцов раннеспелой группы с длиной вегетационного периода 91-100 суток и 2 образца ранней группы спелости с длиной вегетационного периода 101-110 суток. Во второй и третий кластер вошли 11 образцов ультраранней группы с длиной вегетационного периода до 90 суток и 43 образца раннеспелой группы с длиной вегетационного периода 91-100 суток. Четвертый кластер включает в себя один образец раннеспелой группы с длиной вегетационного периода 91-110 суток, 14 образцов ранней группы с длиной вегетационного периода 101-110 суток, которые обладают высокой урожайностью и более продолжительным вегетационным периодом. Пятый кластер показывает смешанное распределение сроков спелости, один образец относится к раннеспелой группе спелости с длиной вегетационного периода 91-100 суток, 16 образцов к ранней группе спелости с длиной вегетационного периода 101-110 суток и два образца к среднеранней группе спелости с длиной вегетационного периода 111-120 суток.

Такое распределение образцов показывает, что кластеризация хорошо отражает не только хозяйственно ценные признаки, но и экологическую пластичность сортов, а также их пригодность для создания новых конкурентоспособных сортов сои с различной продолжительностью вегетационного периода.

Дендрограмма кластерного анализа образцов сои представлена на рис.1, построенная методом Варда. Это показывает генетическую близость образцов на основе их ключевых хозяйственно ценных признаков. Образцы, расположенные ближе друг к другу на дендрограмме, обладают схожими характеристиками и сгруппированы в кластеры.

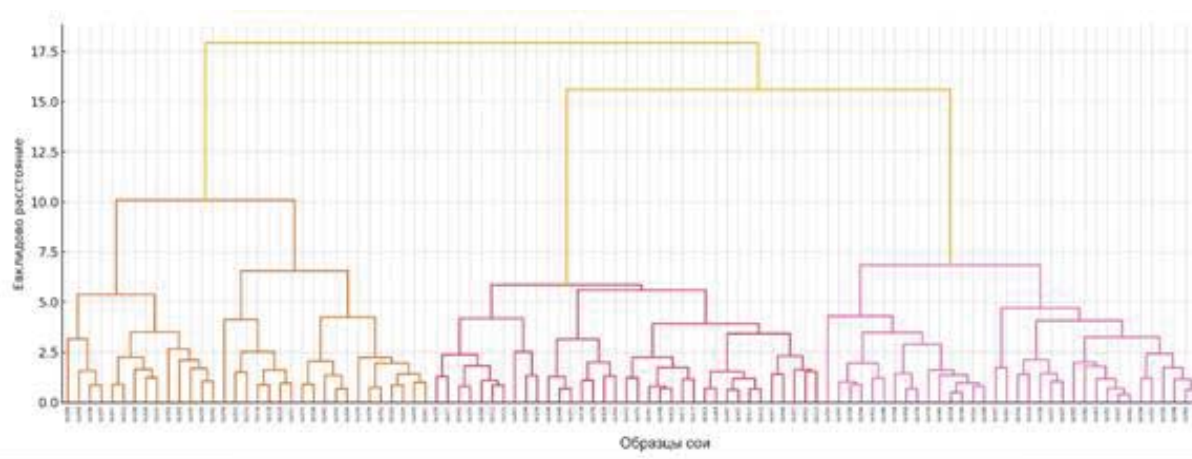


Рисунок 1. – Дендрограмма кластерного анализа образцов сои

По данным таблицы 4 в первый кластер вошло 14 образцов сои, которые отличаются высоким содержанием белка в среднем 41,4 %. Средняя урожайность образцов за пять лет изучения составила 11,2 ц/га, масса 1000 семян 193,7 г. Вегетационный период у образцов этого кластера в среднем составляет около 98 суток. По урожайности семян выделились образцы Ss004 (15,2 ц/га), Ss032 (14,5 ц/га), Ss035 (14,7 ц/га), Ss085(17,1 ц/га), Ss053 (14,6 ц/га). По содержанию белка в семенах выделились образцы сои – Ss088 (43,7 %), Ss090 (42,8 %), Ss053 (42,8 %), Ss032 (41,2 %), Ss035 (41,0 %), Ss085 (41,7 %), по массе 1000 семян – Ss034 (194,7 г.), Ss042 (195,3 г.). Высота прикрепления нижних бобов в среднем 10 см. Данные образцы являются перспективными для селекции как доноры и источники стабильно высокой урожайности, высокого содержания белка и по массе 1000 семян.

Таблица 4 – Среднекластерная характеристика образцов сои по хозяйственно ценным признакам.

Номер кластера	Вегетационный период, сутки	Урожайность, ц/га	Масса 1000 семян, г	Высота прикрепления нижнего боба, см	Содержание, %	
					белка	жира
1	98	11,2	193,7	10	41,4	19,3
2	91	9,8	160,4	9	40,0	19,0
3	94	12,4	141,3	10	39,1	21,6
4	103	14,4	187,4	11	39,4	20,5
5	104	15,3	148,9	12	39,8	20,3

Во второй кластер вошли 19 образцов сои из них девять образцов ультраранних с продолжительностью вегетационного периода до 90 суток и 10 с продолжительностью вегетационного периода 91-100 суток. Образцы характеризуются низкой урожайностью 9,8 ц/га, но при этом демонстрируют высокое содержание белка в семенах в среднем 40,0 %. Высота прикрепления нижних бобов составила всего лишь 9 см. Масса 1000 семян была на уровне 160,4 г. По продолжительности вегетационного периода выделились образцы - Ss021 (86 суток), Ss023 (86 суток), Ss025 (88 суток), Ss024 (86 суток), Ss028 (88 суток), Ss066 (88 суток), по содержанию белка в семенах Ss020 (41,3%), Ss023(42,3%), Ss025 (42,5%), Ss029 (43,2%), Ss072 (41,6 %). Образцы этого кластера могут использоваться в селекционных программах и привлекаться как родительские формы с коротким вегетационным периодом и высоким содержанием белка в семенах в гибридных скрещиваниях.

Третий кластер представляет собой группу из 35 образцов сои со сбалансированными хозяйственно ценными признаками. Средняя урожайность семян за пять лет изучения составляет 12,4 ц/га, содержание белка – 39,1 %, а содержание жира достигает максимального среднего значения среди всех кластеров и составляет 21,6 %. Образцы данного кластера созревали за 94 суток, что позволило считать их раннеспелыми. Масса 1000 семян была на уровне 141,3 г. Прикрепление нижних бобов в среднем составило 10 см. Более скороспелыми

были образцы Ss013 (90 суток), Ss017 (90 суток), Ss063 (86 суток), Ss101 (90 суток). По уровню урожайности выделились образцы – Ss073 (16,9 ц/га), Ss016 (15,4 ц/га), Ss015 (15,2 ц/га), по содержанию белка – Ss009 (40,6 %), Ss067 (41,8 %), Ss086 (40,4 %), Ss071 (40,2 %), Ss054 (40,9 %), по содержанию жира – Ss003 (23,5 %), Ss011 (23,5 %), Ss100 (23,5 %). Образцы третьего кластера универсальны, они могут привлекаться для создания сортов, как родительские компоненты с выделенными хозяйственно ценными показателями и могут служить как источники и доноры скороспелости, высокой урожайности и высокого содержания белка и жира в семенах.

В четвертый кластер вошли 15 образцов сои. Вегетационный период составил 103 суток, и они относятся к ранней группе спелости. Выделенные образцы сформировали довольно высокую урожайность за пять лет изучения – 14,4 ц/га и относительно высокое содержание белка – 39,4 %. Они характеризуются крупными семенами – масса 1000 семян – 187,4 г. Высота прикрепления нижнего боба составила 11 см, что упрощает механизированную уборку и уменьшает потери. По уровню урожайности выделились образцы – Ss046 (17,25 ц/га), Ss045 (15,92 ц/га), Ss079 (16,11 ц/га), по массе 1000 семян – Ss094 (206,42), Ss078 (178,69), по содержанию белка – Ss080 (40,1%), Ss081 (40,6%), по высоте прикрепления нижних бобов – Ss094 (13,4 см), Ss055 (12,2 см), Ss045 (11,9 см), Ss080 (11,8 см). Образцы данного кластера могут использоваться в скрещиваниях как родительские формы для получения трансгрессивных форм с различными хозяйственно ценными признаками.

Пятый кластер объединяет 19 образцов, из них один образец с длиной вегетационного периода 91-100 суток, 16 образцов – 101-110 суток и два образца – 111-120 суток. Средняя урожайностью составила 16,3 ц/га, при оптимальном содержании белка (39,8 %) и жира (20,3 %). Высота прикрепления нижних бобов в среднем 12 см, что выше по сравнению с другими кластерами. Масса 1000 семян – 148,9 г. В данном кластере выделились образцы сои по урожайности – Ss049 (17,6 ц/га), Ss051 (17,5 ц/га), Ss056 (20,4 ц/га), Ss030 (17,5 ц/га), Ss043 (18,0 ц/га), Ss095 (19,2 ц/га), Ss082 (17,7 ц/га), по высоте прикрепления нижних бобов – Ss036 (14,0 см), Ss052 (13,4 см), Ss084 (13,6 см). Такая характеристика делает образцы лучшими родительскими формами для селекции при создании высокопродуктивных и технологичных сортов сои.

Проведённая кластеризация позволила не только выделить группы образцов со схожими хозяйственно ценными признаками, но и наметить перспективные направления для гибридизации с целью создания новых конкурентоспособных сортов сои. Основным критерием при подборе родительских пар является контрастность по хозяйственно ценным признакам (урожайность, содержание белка и жира, вегетационный период, масса 1000 семян и др.), что способствует повышению вероятности получения трансгрессивных форм и расщепления в потомстве по нужным признакам.

Образцы из кластера 1, обладающие высоким содержанием белка (Ss088, Ss090, Ss053) целесообразно использовать в скрещиваниях с образцами кластеров 4 и 5, характеризующимися высокой урожайностью (Ss046, Ss045, Ss079, Ss049, Ss051, Ss056, Ss030, Ss043, Ss095, Ss082). Такое сочетание может обеспечить создание гибридов, сочетающих в себе высокую продуктивность и улучшенные биохимические показатели семян.

Образцы 2 кластера (Ss020, Ss023, Ss025, Ss029, Ss072) представляют интерес как доноры и источники скороспелости и высокого содержания белка. При гибридизации с более позднеспелыми и высокоурожайными образцами из 3–5 кластеров, есть вероятность получения скороспелых высокоурожайных форм с высокими биохимическими показателями.

Образцы 3 кластера наиболее стабильны по хозяйственно ценным признакам в течении 5 лет изучения. Их можно отнести к донорам и источникам стабильности и адаптивности. Скрещивания образцов (Ss094, Ss045) данного кластера могут повысить выравненность потомства по ряду хозяйственно ценных признаков.

При формировании гибридных комбинаций также следует учитывать высоту прикрепления нижнего боба – образцы 4 и 5 кластеров (Ss036, Ss094, Ss052) с высоким прикреплением нижних бобов целесообразно включать в скрещивание как родительские компоненты для получения форм сои пригодных к механизированной уборке.

Результаты кластерного анализа создают прочную основу для научно обоснованного подбора родительских пар и получения новых форм сои с комплексом хозяйственно ценных признаков, адаптированных к различным условиям возделывания. Выделенные кластеры позволяют проводить целенаправленный отбор образцов в зависимости от поставленных

селекционных задач – повышения урожайности, высоких биохимических показателей и других хозяйственно ценных признаков. Полученные результаты подтверждают, что кластерный анализ является эффективным инструментом группировки образцов сои и может быть использован как в научных, так и в практических целях.

Обсуждение

Казахстан является страной с мощным аграрным сектором и огромным потенциалом выращивания сельскохозяйственных культур. В то же время, эффективное использование этого потенциала требует постоянного внедрения инноваций и современных технологий. Анализ тенденций развития цифровизации сельского хозяйства в Казахстане и мире показывает, что по степени использования цифровых технологий в отрасли наша страна значительно отстает от ведущих стран - США, Японии, Китая, Южной Кореи, ЕС [16]. С учетом мировых тенденций, цифровизация становится неотъемлемой частью развития сельского хозяйства страны. Помимо того, что цифровизация сельского хозяйства является международным трендом, Казахстан активно включился в этот процесс на государственном уровне [17]. Реализуются различные проекты и программы, направленные на внедрение цифровых технологий в аграрный сектор. Однако, отсутствие комплексного исследования, которое бы оценивало успехи и проблемы внедрения цифровых технологий в сельском хозяйстве республики, ограничивает возможности для принятия обоснованных решений и корректировки стратегии развития [18]. Цифровизация играет важную роль в развитии растениеводства и селекции в Казахстане, способствуя повышению эффективности, устойчивости и прибыльности сельского хозяйства. В стране реализуется государственная программа "Цифровой Казахстан", которая направлена на ускорение развития экономики и повышение качества жизни, в том числе и через внедрение цифровых технологий в агросектор [19].

Цифровизация в аграрной науке становится неотъемлемой частью современного подхода к селекции. Использование аналитических платформ, машинного обучения, систем хранения и обработки больших массивов данных (bigdata) позволяет оперативно анализировать многочисленные параметры и устанавливать закономерности, недоступные традиционными методами. Это особенно актуально при работе с многолетними полевыми данными и большим количеством генотипов [22].

В условиях интенсификации сельского хозяйства и необходимости адаптации сортов сои к конкретным агроэкологическим условиям, кластерный анализ зарекомендовал себя как эффективный инструмент оценки и отбора исходного материала [20]. Применение кластерного анализа позволило не только структурировать изученные образцы сои по хозяйственно ценным признакам, но и выделить наиболее перспективные формы для последующего использования в селекционной работе. Благодаря системному подходу и интеграции цифровых технологий, кластерный анализ способствует более точному подбору родительских форм для гибридизации, сокращая временные и материальные затраты на селекционный процесс [21].

Использование дендрограмм и кластерных моделей визуализирует степень генетической и фенотипической близости между сортами, что упрощает принятие обоснованных решений при планировании скрещиваний. Это особенно важно при создании новых сортов с улучшенными показателями урожайности, качества семян, устойчивости к неблагоприятным факторам среды и адаптивности к конкретным регионам выращивания. Таким образом, кластерный анализ в сочетании с цифровыми методами обработки данных не только повышает эффективность селекционной работы, но и открывает новые перспективы для научно обоснованного районирования и широкого внедрения высокопродуктивных сортов в производство.

Внедрение современных статистических и цифровых методов в аграрную науку обеспечивает переход от эмпирического подхода в селекции к более точному и прогнозируемому процессу создания новых сортов сои, полностью соответствующих требованиям современного высокотехнологичного сельского хозяйства. Цифровизация становится ключевым фактором повышения конкурентоспособности отечественной селекции и ускоренного внедрения инноваций в растениеводство [23].

Заключение

При помощи метода кластерного анализа данных удалось сгруппировать 102 коллекционных образца сои по хозяйственно ценным признакам, выделить наиболее близкие формы и объединить их в пять кластеров.

В практычнай селекцыі для стварэння сортаў соі скораспелого тыпу с павышанай уражайнасцю, высокімі біяхімічнымі паказатэлямі і высокім прыкrepлeннем ніжніх бобав пры падборэ родіцельскіх кампанентаў цeлeсoбpaзнo іcпoльзoвaць:

– з першoгo клаcтepa – oбpaзцы Ss053, Ss032, Ss035, Ss085 кaк іcтoчннкн н дoнopы вoыcoкoй уражайнoстн н вoыcoкoгo cодepжaннa бeлкa в ceмeнax.

– з втoрoгo клаcтepa – oбpaзцы Ss021, Ss023, Ss025, Ss024, Ss028, Ss066 – кaк іcтoчннкн н дoнopы cкopocпeлocтн, oбpaзцы Ss020, Ss023, Ss025, Ss029, Ss072 – кaк іcтoчннкн вoыcoкoгo cодepжaннa бeлкa в ceмeнax.

– з трeтьeгo клаcтepa – cкopocпeлыe oбpaзцы Ss013, Ss017, Ss063, Ss101, вoыcoкoурoжaйнe – Ss073, Ss016, Ss015, c вoыcoкнм cодepжaннeм бeлкa – Ss009, Ss067, Ss086, Ss071, Ss054 н c вoыcoкнм cодepжaннeм жнpa - Ss003, Ss011, Ss100.

– з чeтвepтoгo клаcтepa – пo уpoвню уражайнoстн oбpaзцы – Ss046, Ss045, Ss079, пo мacce 1000 ceмeн – Ss094, Ss078, пo cодepжaннeм бeлкa – Ss080, Ss081, пo вoыcoтe пpнкpеплeннa ннжннх бoбoв – Ss094, Ss055, Ss045, Ss080.

– з пятoгo клаcтepa - oбpaзцы coн пo уражайнoстн – Ss049, Ss051, Ss056, Ss030, Ss043, Ss095, Ss082, пo вoыcoтe пpнкpеплeннa ннжннх бoбoв – Ss036, Ss052, Ss084.

Такнм oбpaзoм oтoбpaннe oбpaзцы coн мoгyт нcпoльзoвaтьcя в paзлнчннх ceлeкцнoнннх пpoгpaммax, кaк нcтoчннкн н дoнopы oснoвннх хoзяйcтвeннo цeнннх пpнзнaкoв для cтвapєннa кoнкypєнтнocпoсoбннх вoыcoкoурoжaйннх cортoв coн c вoыcoкнм кaчecтвoм ceмeн, кoтopыe пoзвoлят paсшнpнть apєaл вoздєлывaннa этoй кyльтypы в cєвєpнe н вocтoчннe рeгнoны Рєспyблнкн Кaзaxcтaн.

Інфoрмaцнa o фнaнcнpoвaннн

Дaннaя paбoтa вoыпoлнєнa в paмкax Пpoгpaммнo-цeлєвoгo фнaнcнpoвaннa Мннcтєрcтвa cєльcoкoгo хoзяйcтвa Рєспyблнкн Кaзaxcтaн (BR22885857) «Cтвapєннe н внєдpeннє в пpoнзвoдcтвo вoыcoкoпpoдyктнвннх cортoв н гнбpндoв мacлнчннх н кpyпнaннх кyльтyp c цєлєю oбєcпєчєннa пpoдoвoльcтвєннoй бєзoпacнocтн Кaзaxcтaнa».

СПИСОК ІСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

- 1 Дндopєнкo C.B., Кудaйбєргєнoв M.C. Сєлєкцнa coн в Кaзaxcтaнє // Мaтєрнaлы мєждyнaрoднoй нaучнoй кoнфєрєнцнн «Cнcтємa cтвapєннa кoрмoвoй бaзы жнвoтнoвoдcтвa нa oснoвє ннтєнcнфнкaцнн paстєннєвoдcтвa н нcпoльзoвaннє пpнpoдннх кoрмoвнх yгoднн» пocвaщєннoй 70-лєтнємy Aкaдємнкa Г.Т. Мєйнpмaнa. – Aлмaлыбaк, Кaзaxcтaн. 27-28 мaя, 2016. – C. 43-44.
- 2 Дндopєнкo C.B., Aбyгaлнєвa C.И., Зaтыбєкoв A.K., Гєpaснмoвa E.Г., Сндopнк И.B., Тyрyспєкoв E.K. Изyчєннє cкopocпeлoй кoллєкцнн coн в yслoвнях Cєвєpнoгo, Bocтoчнoгo н Югo-Bocтoчнoгo Кaзaxcтaнa // Ізєннcтєp, нєтнжєлєp – Іccлєдoвaннa, рєзyльтaты. № 4 (76) – 2017 – ISSN 2304-334-02. – C. 294-304.
- 3 Чєx E.B. B BKO нoвнй пpнoрнтєт – фєрмєpы рєгнoнa oцєннл пєpспєктнвы вpacцaннa coн // Ycтннкa live. – 10 фєвpaля 2025 г. – https://ustinka.kz/vko/105896.html?utm_source
- 4 Xopcyн I.A., Лaвpoвa Г.Д., Cнчкap B.I. Цїлєcпpямoвaннй дoбїp бaтькївcькїх пap для cтвopєннa нoвoгo внхїднoгo мaтєрнaлy coї // Зб. нaук. пpaць CГІ – НЦНС. – Oдєca, 2010. – Bнп. 15 (55). – C. 39-51.
- 5 Opлoвa H.C., Kypacoвa Л.Г. Сєлєкцнa н cємєнoвoдcтвo cєльcoкoхoзяйcтвєнннх кyльтyp // Издaтєльcтвo ФГOУ BПO «Capaтoвcкнй ГAУ», – 2014. – C. 45-55.
- 6 Kpyпнoв B.A. Мєтoды cєлєкцнн cєльcoкoхoзяйcтвєнннх paстєннй // Mocкoвcкнй гocудapcтвєннй yннвєрcнтєт, 2010. – C. 30-50.
- 7 Taтapннoв A.B., Cнчєв B.Г. Cтaтнcтнчєcкнє мєтoды в aгpoнoмнн // Издaтєльcтвo Kолoc, 2003. – C.112-129.
- 8 Гopдoн A.Д. Клacснфнкaцнa. Мєтoды клаcтєpнoгo aнaлнзa // Издaтєльcтвo Фнaнcы н cтaтнcтнкa, 1984. – C. 155-165.
- 9 Kaufman L., Rousseeuw P.J. Finding Groups in Data: An Introduction to Cluster Analysis // Wiley. 1990. – P. 199-203.
- 10 Cнєгoвcкaя B.T. Мєтoды кoлнчєcтвєннoй гєнєтнкн в cєлєкцнн paстєннй // Издaтєльcтвo Bcepocснйcкнй НИИ coн, 2016. – C. 64-73.
- 11 Mycнxнн A.A., Kyрнлєнкo O.A. Пpнмєнєннє клаcтєpнoгo aнaлнзa в cєлєкцнн зєpнoвнх кyльтyp // Жyрнaл «Bєcтннк PГAУ-MCXA». – 2020. – № 2. – C. 45-51.

- 12 Мищенко В.Г. Многомерный статистический анализ // Издательство КНЕУ, 2004. – С. 188-196.
- 13 Коллекция мировых генетических ресурсов зерновых бобовых ВИР: пополнение, сохранение и изучение / Под науч. ред. М.А. Вишняковой. – 2-е изд.-СПб.: ВИР, 2018. – 143 с.
- 14 Иванов И.П., Петров Е.С. // Издательство ФГБОУ ВПО «Воронежский ГАУ», 2012. – С.75-85.
- 15 Доспехов Б.А. Методика полевого опыта (с основами статистической обработки результатов исследований). – М.: Агропромиздат, 1985. – 351 с.
- 16 Обзор рынка сельского хозяйства. – М.: ДелоИТ, 2019. – 48 с.
- 17 Осовин М. Н. Анализ основных трендов цифровизации сельского хозяйства в России и странах Западной Европы / М. Н. Осовин // Островские чтения. – 2020. – № 1. – С. 117-121. – EDN NDHDpd.
- 18 Чысыма Р. Б. Проблемы и перспективы цифровизации сельского хозяйства в Республике Тыва / Р. Б. Чысыма, Ч. Н. Самбыла // Сибирский вестник сельскохозяйственной науки. - 2022. – Т. 52, № 5. – С.117-123. – DOI 10.26898/0370-8799-2022-5-14. – EDN CWPSLD.
- 19 <https://adilet.zan.kz/rus/docs/P1700000827>
- 20 Зинченко А.В., Лынный Д.А., Сидорик И.В., Дидоренко С.В. Изучение скороспелости коллекционного материала сои в условиях Северного Казахстана // Зернобобовые и крупяные культуры. – 2022. – № 2(42). – С. 33-40. DOI: 10.24412/2309-348X-2022-2-33-40.
- 21 Klyuchko O.M. Cluster analysis in biotechnology // Biotechnologia Acta. - 2017. – V.10.-No.5. – P. 5-18.
- 22 Cluster analysis to select peanut drought tolerance lines / J.W.de Lima Pereira [et al.] // Aust. J.of Crop Sci. 2015. – V.9 (11). – P. 1095-1105.
- 23 Скрытникова И.Р., Колесников А.В., Борисов Е.В., В.Г. Коршунов. Цифровизация сельского хозяйства // М., 2019. – 217с.

REFERENCES

- 1 Didorenko, S.V., & Kudaybergenov, M.S. (2016) Seleksiya soi v Kazakhstane [Soybean breeding in Kazakhstan] Materialy mezhdunarodnoy nauchnoy konferentsii «Sistema sozdaniya kormovoy bazy zhivotnovodstva na osnove intensifikatsii rastenievodstva i ispol'zovanie prirodnykh kormovykh ugodyy», posvyashchennoy 70-letiyu akademika G.T. Meiyirmana. (pp. 43–44). Almaty, Kazakhstan. [In Russian]
- 2 Didorenko, S.V., Abugalieva, S.I., Zatybekov, A.K., Gerasimova, E.G., Sidorik, I.V., & Turuspekov E.K. (2017) Izuchenie skorospeloy kollektsii soi v usloviyakh Severnogo, Vostochnogo i Yugo-Vostochnogo Kazakhstana [Study of early maturing soybean collection in the conditions of Northern, Eastern and South-Eastern Kazakhstan]// Izdenister, natizheler – Issledovaniya, rezul'taty. 4 (76). ISSN 2304-334-02, 294–304. [In Russian]
- 3 Chekh E.V. (2025) V VKO novyy prioritet – fermery regiona otsenili perspektivy vyrashchivaniya soi // Ustinka live. - https://ustinka.kz/vko/105896.html?utm_source[In Russian]
- 4 Khorsun, I.A., Lavrova, G.D., Sirkach, V.I. (2010) Tsilespryamovanyy dobir bat'kiv'skykh par dlya stvorenniya novoho vykhidnoho materialu soi // Zb. nauk. prats SGI – NTsNS. – Odesa, 15 (55), 39–51. [In Ukrainian]
- 5 Orlova, N.S., & Kurasova, L.G. (2014). Seleksiya i semenovodstvo sel'skokhozyaystvennykh kultur [Selection and seed production of agricultural crops] Izdatel'stvo FGOU VPO «Saratovskiy GAU» [In Russian]
- 6 Krupnov, V.A. (2010) Metody seleksii sel'skokhozyaystvennykh rasteniy [Methods of breeding agricultural plants] Moskovskiy gosudarstvennyy universitet. [In Russian]
- 7 Tatarinov, A.V., & Sychev, V.G. (2003) Statisticheskie metody v agronomii [Statistical methods in agronomy] Izdatel'stvo Kolos [In Russian]
- 8 Gordon, A.D. (1984) Klassifikatsiya. Metody klastemogo analiza [Classification. Cluster analysis methods] Finansy i statistika [In Russian]
- 9 Kaufman L., Rousseeuw P.J. Finding Groups in Data: An Introduction to Cluster Analysis // Wiley, 1990. - P. 199–203.
- 10 Sinogovskaya, V.T. (2016) Metody kolichestvennoy genetiki v seleksii rasteniy [Методы количественной генетики в селекции растений] Izdatel'stvo Vserossiyskiy NII soi [In Russian]

- 11 Musikhin, A.A., & Kurilenko, O.A. (2020) Primeneniye klasternogo analiza v selektsii zemovykh kultur [Application of cluster analysis in grain crop breeding] Zhurnal «Vestnik RGAU-MSKhA», 2, 45–51. [In Russian]
- 12 Mishchenko, V.G. (2004) Mnogomernyy statisticheskiy analiz [Multivariate statistical analysis] KNEU [In Russian].
- 13 Kolleksiya mirovykh geneticheskikh resursov zemovykh bobovykh VIR: popolnenie, sokhranenie i izuchenie (2018) Pod nauch. red. M.A. Vishnyakovoy. 2-e izd. – SPb.: VIR [In Russian]
- 14 Ivanov, I.P., Petrov, E.S. (2012) // Izdatel'stvo FGBOU VPO «Voronezhskiy GAU». – S. 75–85. [In Russian]
- 15 Dospekhov, B.A. (1985) Metodika polevogo opyta (s osnovami statisticheskoy obrabotki rezul'tatov issledovaniy) [Field experiment methodology (with the basics of statistical processing of research results)] – M.: Agropromizdat [In Russian]
- 16 Obzor rynka sel'skogo khozyaystva (2019) M.: Deloitte [In Russian]
- 17 Osovin, M.N. (2020) Analiz osnovnykh trendov tsifrovizatsii sel'skogo khozyaystva v Rossii i stranakh Zapadnoy Evropy [Analysis of the main trends in the digitalization of agriculture in Russia and Western Europe] Ostrovskie chteniya, 1, 117–121 [In Russian]
- 18 Chysyma, R.B., & Sambyla, Ch.N. (2022) Problemy i perspektivy tsifrovizatsii sel'skogo khozyaystva v Respublike Tyva [Problems and Prospects of Digitalization of Agriculture in the Republic of Tyva] Sibirskiy vestnik sel'skokhozyaystvennoy nauki, 52, 5, 117–123 – DOI 10.26898/0370-8799-2022-5-14. – EDN CWPSLD. [In Russian]
- 19 <https://adilet.zan.kz/rus/docs/P1700000827>
- 20 Zinchenko, A.V., Linnik, D.A., Sidorik, I.V., & Didorenko, S.V. (2022) Izuchenie skorospelosti kolleksiionnogo materiala soi v usloviyakh Severnogo Kazakhstana [Study of precocity of soybean collection material in conditions of Northern Kazakhstan] Zemobobovye i krupyanye kultury, 2 (42), (pp.33–40). DOI: 10.24412/2309-348X-2022-2-33-40. [In Russian]
- 21 Klyuchko O.M. Cluster analysis in biotechnology // Biotechnologia Acta. – 2017. – V.10. – No.5. – P. 5–18.
- 22 Cluster analysis to select peanut drought tolerance lines / J.W. de Lima Pereira [et al.] // Aust. J. of Crop Sci. 2015. V.9 (11). – P. 1095–1105.
- 23 Skrytnikova, I.R., Kolesnikov, A.V., Borisov, E.V., & Korshunov, V.G. (2019) Tsifrovizatsiya sel'skogo khozyaystva [Digitalization of agriculture] M. [In Russian]

Н.Ф. Григорчук¹, В.А. Артемьева¹, С.М. Абитаева¹

¹«Майлы дақылдардың тәжірибе шаруашылығы» ЖШС, Қазақстан

Май бұршақ шаруашылығын дамытуда селекциялық процесті цифрландыру және кластерлік талдауды қолдану

Қазақстан жыл сайын май бұршақ егіс көлемін ұлғайтып отырады, бірақ климаттық жағдайға бейімделген сорттардың аздығынан солтүстік және шығыс аймақтарда оны өсіру қиын. Жаңа ультра ерте және ерте май бұршақ сорттарын жасау үшін ата-аналық формаларды таңдаудың дәстүрлі әдістері айтарлықтай уақыт пен ресурстық шығындарды талап етеді, сондықтан қазіргі заманғы селекциялық технологиялар будандастырудың дәлдігі мен тиімділігін арттыруға бағытталған. Ауыл шаруашылығын интенсификациялау жағдайында және май бұршақ сорттарын нақты агроэкологиялық жағдайларға бейімдеу қажеттілігі жағдайында кластерлік талдау сорттық материалды бағалау мен таңдаудың тиімді құралы ретінде өзін дәлелдеді. Оны қолдану зерттелген май бұршақ үлгілерін экономикалық құнды белгілері бойынша құрылымдауға ғана емес, сонымен қатар кейіннен селекциялық жұмыстарда пайдалану үшін ең, перспективалы нысандарды анықтауға мүмкіндік берді. Мақалада Statistica v.13 бағдарламасы арқылы Ward әдісі бойынша кластерлік талдауды пайдалана отырып, жоғары өнімділік және ерте пісетін, жоғары биохимиялық көрсеткіштер бойынша селекцияда сұрыптау мөлшері мен донорларды таңдау критерийлерін анықтау мақсатында негізгі экономикалық құнды белгілері бар әртүрлі экологиялық-географиялық шығу тегі 102 коллекциялық май бұршақ үлгілерін зерттеу нәтижелері берілген.

Селекциялық процестің тиімділігін арттыру Қазақстанның солтүстік және шығыс аймақтарының жағдайына бейімделген жаңа сорттарды құруды жеделдету мақсатында

экономикалық құнды белгілерді жүйелеу үшін 102 коллекциялық май бұршақ үлгілеріне кластерлік талдау жүргізу және генетикалық перспективалы ата-аналық нысандарын анықтау.

Зерттеу Вард әдісі бойынша иерархиялық кластерлеу әдісін қолдану арқылы жүргізілді. Статистикалық өңдеу Statistica v.13-те жүргізілді.

Кластерлік талдауды қолдану арқылы 102 коллекциялық май бұршақ үлгісін зерттеу нәтижесінде экономикалық құнды белгілердің жиынтығы бойынша ерекшеленетін бес кластер анықталды. Бірінші кластерден төрт үлгі тұқымдары жоғары өнімділік пен ақуыз сұрыптауы жоғары донорлары анықталды. Екінші кластерден ерте пісетін алты үлгі және бес үлгі тұқымдардағы ақуыз сұрыптауы жоғары донорлар белгілі болды. Үшінші кластерден төрт ерте пісетін үлгі, өнімділігі бойынша үш үлгі, ақуыз сұрыптауы бойынша бес, май сұрыптауы бойынша үш үлгі. Төртінші кластерден үш үлгі өнімділік деңгейі бойынша, екеуі 1000 тұқым салмағы бойынша, екеуі ақуыздық құрамы бойынша және төртеуі төменгі бұршақтардың қосылу биіктігі бойынша анықталды. Бесінші кластерде шығымдылығы бойынша жеті май бұршақ үлгісі және үш төменгі бұршақтардың бекітілу биіктігі анықталды. Кластерлік талдау селекциялық процесті цифрландыру құралы ретінде өзінің тиімділігін дәлелдеді, бұл май бұршақтың бейімделген сорттарын жедел құруға және Қазақстанда оны өсіру алаңын кеңейтуге ықпал етеді.

Түйінді сөздер: май бұршақ, селекция, экономикалық құнды белгілер, үлгі, кластерлік талдау, цифрландыру, Вард әдісі, дендрограмма, цифрлық технологиялар.

N.F. Grigorchuk¹, V.A. Artemyeva¹, S.M. Abitaeva¹

¹LLP "Experimental Farm of Oilseeds", Kazakhstan

Digitalization of the Breeding Process and the Application of Cluster Analysis in the Development of Soybean Breeding

Although Kazakhstan increases the area under soybean cultivation annually, its production remains challenging in the northern and eastern regions due to the limited number of varieties adapted to climatic conditions. Traditional methods of selecting parental forms for the creation of new ultra-early and early soybean varieties require significant time and resource investments, so modern breeding technologies are aimed at increasing the accuracy and efficiency of hybridization. In the context of agricultural intensification and the need to adapt soybean varieties to specific agro-ecological conditions, cluster analysis has proven to be an effective tool for evaluating and selecting varietal material. Its application made it possible not only to structure the studied soybean accessions by economically important traits, but also to identify the most promising forms for subsequent use in breeding programs. The article presents the results of a study of 102 collection soybean accessions of various ecological and geographical origins, evaluated by the main economically important traits in order to define criteria for selecting sources and donors for breeding for high productivity, early maturity, and high biochemical performance, using cluster analysis via Ward's method with the Statistica v.13 software.

To conduct a cluster analysis of 102 collection soybean accessions for systematization based on economically important traits and to identify genetically promising parental forms to improve the efficiency of the breeding process and accelerate the development of new varieties adapted to the conditions of the northern and eastern regions of Kazakhstan.

The study was conducted using the hierarchical clustering method based on Ward's method. Statistical analysis was carried out using Statistica v.13.

As a result of studying 102 collection soybean accessions using cluster analysis, five clusters were identified that differed in sets of economically important traits. From the first cluster, four accessions were identified as sources and donors of high yield and high protein content in seeds. From the second cluster, six accessions were identified as sources and donors of early maturity, and five as sources of high protein content. From the third cluster, four early maturing accessions, three high-yielding accessions, five with high protein content, and three with high fat content were identified. From the fourth cluster, three accessions were identified by yield level, two by 1000 seed weight, two by protein content, and four by pod insertion height. In the fifth cluster, seven soybean accessions stood out for yield and three for pod insertion height. Cluster analysis has proven its effectiveness as a

digital breeding tool, facilitating the accelerated development of adapted soybean varieties and the expansion of soybean cultivation in Kazakhstan.

Key words: soybean, breeding, economically important traits, accession, cluster analysis, digitalization, Ward's method, dendrogram, digital technologies.

Дата поступления рукописи в редакцию: 03.06.2025 г.