

УДК 633.854.78 : 631.51 : 519.24

ИДЕНТИФИКАЦИЯ РАСТЕНИЙ ИНБРЕДНЫХ ЛИНИЙ ПОДСОЛ- НЕЧНИКА ПО ПРИЗНАКАМ ОПРЕДЕЛИТЕЛЯ UPOV

Гронин В. В. – младший научный сотрудник

*Всероссийский научно-исследовательский институт биологической
защиты растений*

Международной организацией по защите прав селекционеров (UPOV) разработан определитель для проведения оценки на отличимость, однородность и стабильность линий и гибридов подсолнечника. Нами изучена возможность использования количественных признаков определителя UPOV в условиях сильной средовой изменчивости. Показана возможность использования канонического дискриминантного анализа для идентификации отдельных растений инбредных линий подсолнечника.

Перечисляя требования к идеальному типу сорта пшеницы, Н. И. Вавилов [1, с. 434, 435] в один ряд с признаками высокой продуктивности, качества урожая, устойчивости растений к разным стрессам поставил также и "наличие в данном сорте удобных отличительных морфологических признаков для апробации и отделения от других сортов" (подчеркнуто автором). Особенно возрастает значение надежного описания генотипов для поддержания в процессе семеноводства чистоты и идентичности при высоком уровне генетического улучшения, когда уменьшаются сортовые различия, но повышаются требования к генетической чистоте [2, с. 188]. Еще более актуальным стало выявление надежных отличительных признаков после вступления России в 1998 году в международную организацию по защите прав селекционеров UPOV, так как патентная защита сорта или гибрида невозможна без наличия у него таких признаков.

Организацией UPOV разработан специальный определитель для проведения оценки на отличимость, однородность и стабильность (ООС) родительских линий и гибридов подсолнечника [3]. Таблица признаков определителя содержит дискретные и морфометрические признаки, оцениваемые глазомерно. Для всех дискретных признаков пригоден визуальный способ оценки. Различия по таким признакам, как, например, наличие или отсутствие антоциановой окраски гипокотыля, зеленый цвет листа, бугристость листовой пластинки, окраска язычковых цветков, расположение листочков обертки по отношению к корзинке легко выявляются внешним осмотром и позволяют надежно распознавать принадлежность отдельных растений к соответствующим родительским линиям и межлинейным гибридам. Поэтому качественные признаки более предпочтительны для использования в целях отличимости. В ряде случаев дискретных признаков недостаточно для определения принадлежности растений подсолнечника к конкретным сортообразцам. Поэтому неизбежно использование количественных признаков определителя, которые подвержены модификационной изменчивости.

Чтобы изучить возможность использования количественных признаков определителя UPOV в условиях сильной средовой изменчивости, был заложен опыт, в котором неконтролируемая модификационная изменчивость усиливалась экспериментальным варьированием густоты стояния растений: 20 тыс./га, 40 тыс./га, 70 тыс./га. Выращивались растения 5 родительских линий межлинейных гибридов: ВК 499, ВК 653, ВК 678, ВК 389, ВК 392. Изучались следующие количественные признаки, предлагаемые определителем UPOV: размер листа, угол между нижними боковыми жилками листа, длина язычковых цветков, густота язычковых цветков, длина вершины листочка обертки, высота растения, размер корзинки. Все признаки были измерены на 20-ти растениях каждой инбредной линии на каждой густоте в каждый год исследования. Размер листа измерен как

длина листовой пластинки, максимальная ширина листовой пластинки, расстояние от максимальной ширины до вершины листовой пластинки и длина черешка, густота язычковых цветков – как число язычковых цветков на 10 см периметра корзинки, а размер корзинки – как диаметр и периметр корзинки.

Для каждой инбредной линии вычислены средние величины всех признаков, дисперсии, стандартные отклонения на каждой густоте и для средних между густотами [4]. Все линии сравнены между собой по средним с использованием критерия Стьюдента на каждой густоте, а также по средним между густотами. По многим признакам выявлены достоверные различия между средними на уровнях значимости 0,01 и 0,05.

Чтобы рассчитать вероятность ошибочных определений отдельных растений вычислены коэффициенты дискриминации Любищева (K) для каждого сравнения (K равен отношению квадрата разности средних к сумме их дисперсий) [4, с. 14].

Оказалось, что для большинства признаков величина K составила от 0 до 3, чему соответствует отношение числа ошибочных определений к общему числу испытаний от 1:2 до 1:9. Только для отдельных пар инбредных линий по некоторым признакам получены достаточно большие значения K (табл. 1).

Таблица 1 – Эффективность количественных признаков определителя на примере родительских линий селекции ВНИИМК

Сравниваемые линии	Признаки	2001 год			2002 год		
		t	K	О	t	K	О
ВК 499 и ВК 653	Длина вершины листочка обертки	13,0*	3,3	1:9	9,2*	4,2	1:13

ВК 653 и ВК 392	Число язычковых цветков на 10 см периметра корзинки	6,8*	0,8	1:4	8,9*	4,0	1:13
ВК 653 и ВК 389	Длина вершины листочка обертки	10,5*	2,0	1:6	10,7*	5,8	1:24
ВК 653 и ВК 678	Длина листовой пластинки	12,4*	2,7	1:9	10,6*	5,6	1:24
	Расстояние от максимальной ширины до вершины листа	12,0*	2,6	1:9	10,8*	5,8	1:24
	Угол между боковыми жилками листа	13,1*	3,2	1:9	9,5*	4,5	1:18
	Длина вершины листочка обертки	15,0*	3,9	1:13	14,7*	10,8	1:104
	Высота растения	15,7*	4,4	1:13	9,4*	4,4	1:13
<p>* – различия достоверны на уровне значимости $p=0,99$,</p> <p>О – отношение количества ошибочных определений к общему числу испытаний.</p>							

Изучая вариабельность коэффициента дискриминации по годам, мы обнаружили, что величина K нестабильна. Для линий ВК 389 и ВК 653 по признаку "длина вершины листочка обертки" в один из годов исследования $K=5,8$, по данным второго года $K=2,0$. Для линий ВК 653 и ВК 678 по признаку "длина листовой пластинки" K равен 5,6 и 2,7, по признаку "расстояние от максимальной ширины до вершины листа" получены значения K : 5,8 и 2,5. Для этих значений K отношение ошибочных определений к общему числу испытаний варьировало от 1:24 до 1:6. Для пары линий ВК 653 и ВК 678 по признаку "длина вершины листочка обертки" в один год исследования получено $K=10,8$, а во второй год K был равен 3,9. Под влия-

нием условий года величина K для этого признака изменилась почти в 3 раза, а отношение ошибочных определений к общему числу испытаний – в восемь раз (от 1:104 при величинах дисперсий для ВК 653 $\sigma^2=0,22$ и для ВК 678 $\sigma^2=0,10$ до 1:13 для ВК 653 $\sigma^2=0,26$, для ВК 678 $\sigma^2=0,07$). Только для признака "высота растения" этой же пары линий величина K была примерно одинаковой в оба года – 4,4 (отношение ошибочных определений к общему числу испытаний 1:13).

Таким образом, точность идентификации растений по простым мерным признакам может значительно меняться даже в тех случаях, когда разность между средними составляет более 100 % от меньшей величины, как для линий ВК 653 и ВК 678 по признаку "длина вершины листочка обертки".

Если разница между средними составляет от 35 до 70 %, то теоретически ошибочное определение будет встречаться от одного раза на 18 до одного раза на 24 определения. В большинстве же случаев число ошибочных определений при идентификации отдельных растений может составлять от 10 до 50 %.

Эти результаты свидетельствуют о необходимости поиска соотношений морфометрических признаков, которые были бы нечувствительны к модифицирующему воздействию среды.

Поэтому была поставлена задача найти такие соотношения для 5-ти изучаемых линий. Поиск таких соотношений осуществлялся методом дискриминантного анализа. Дискриминантный анализ представляет собой статистический аппарат для изучения различий между двумя и более группами объектов по отношению к нескольким переменным одновременно, он успешно применялся в различных областях науки: психология, социология, судебное производство, медицина, животноводство [6, с. 80].

С использованием пакета прикладных программ SPSS 10.0 для дискриминации всех изучаемых линий путем комплексирования признаков вычислялись канонические дискриминантные функции. Путем анализа мер существенности собственных значений (λ) дискриминантных функций производился отбор признаков для использования в качестве переменных в линейных комбинациях. Так как изучалось 5 инбредных линий, были вычислены 4 дискриминантные функции с использованием 12 переменных. Например, первая и вторая дискриминантные функции имеют вид:

$$Df 1 = -8,316 + 0,276 X 1 - 0,285 X 2 + 0,108 X 3 + 0,462 X 4 - 0,764 X 5 + 0,108 X 6 + 0,272 X 7 - 1,206 X 8 + 1,604 X 9 - 0,042 X 10 + 0,246 X 11 - 0,006 X 12;$$

$$Df 2 = -14,989 - 0,145 X 1 + 0,102 X 2 + 0,023 X 3 - 1,043 X 4 + 4,221 X 5 + 0,314 X 6 + 0,711 X 7 - 0,020 X 8 - 0,938 X 9 + 0,014 X 10 + 0,193 X 11 + 0,033 X 12;$$

где: X_1 – длина листовой пластинки, X_2 – ширина листовой пластинки, X_3 – расстояние от максимальной ширины до вершины листовой пластинки, X_4 – длина язычковых цветков, X_5 – максимальная ширина язычковых цветков, X_6 – число язычковых цветков на 10 см периметра корзинки, X_7 – длина листочка обертки, X_8 – ширина листочка обертки, X_9 – длина вершины листочка обертки, X_{10} – высота растения, X_{11} – число листьев, X_{12} – диаметр корзинки.

Собственные значения всех четырех функций и меры их существенности представлены в таблице 2.

Таблица 2 – Собственные значения и меры их существенности для четырех дискриминантных функций

Функция	Собственное значение λ	Относительное процентное содержание	Каноническая корреляция
<i>Df 1</i>	7,123	47,6	0,936
<i>Df 2</i>	4,159	27,8	0,898

<i>Df</i> 3	2,688	18,0	0,854
<i>Df</i> 4	1,000	6,7	0,707

Из таблицы следует, что *Df* 1 и *Df* 2 достаточно эффективно разделяют изучаемые родительские линии (рис. 1), так как собственные значения функций *Df* 1 и *Df* 2 (7,123 и 4,159) много больше, чем λ для *Df* 3 и *Df* 4. Однако для количественного сравнения дискриминантных возможностей функций необходимо рассмотреть относительное процентное содержание: функция *Df* 1 содержит 47,6 % общих дискриминантных возможностей, а функция *Df* 2 – 27,8 %. Коэффициент канонической корреляции является мерой связи между изучаемыми группами и дискриминантной функцией. Высокие значения коэффициентов канонической корреляции свидетельствуют о высокой степени зависимости между каждой из функций и изучаемыми объектами, то есть растениями инбредных линий.

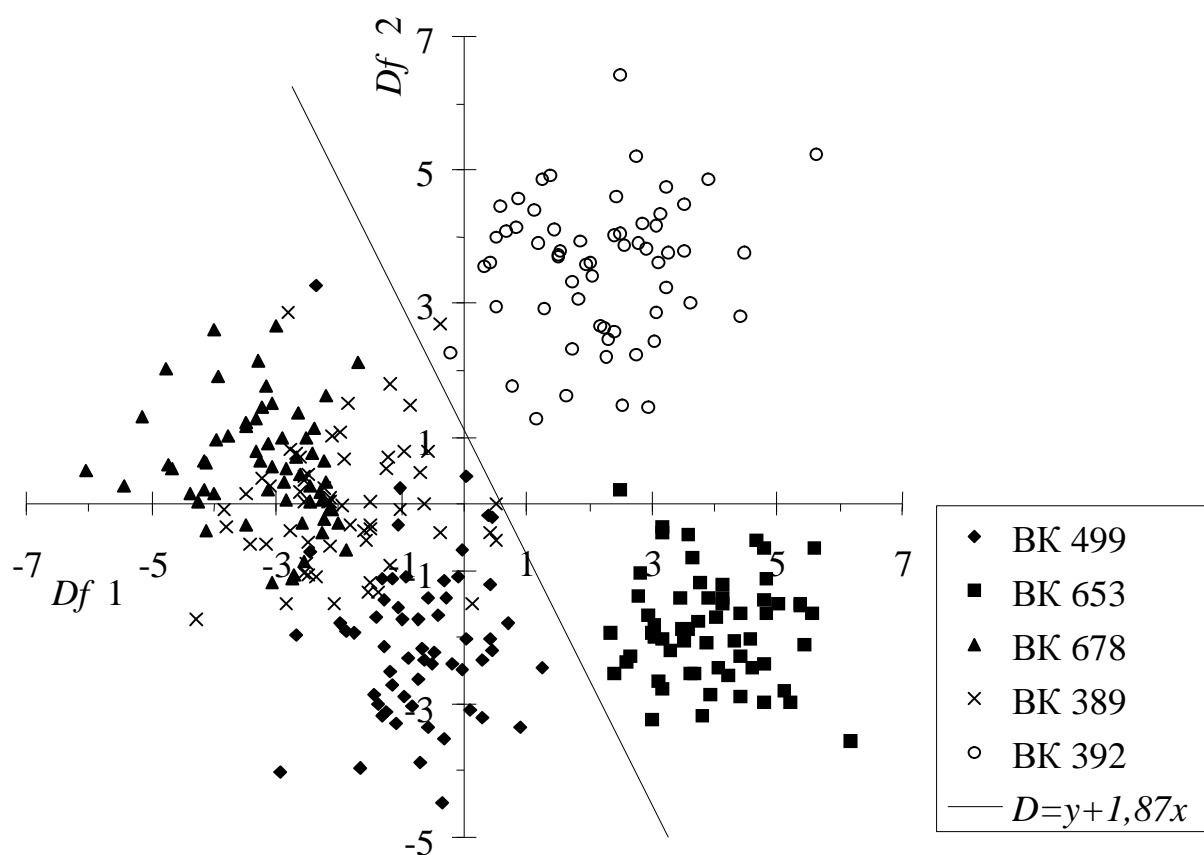


Рисунок 1 – Дискриминация растений 5-ти родительских линий подсолнечника в пространстве дискриминантных функций $Df 1$ и $Df 2$

Эффективность 1-й и 2-й дискриминантных функций подтверждается λ -статистикой Уилкса (рассмотрение λ Уилкса необходимо при работе с выборками) [6, с. 108]. λ Уилкса служит для измерения остаточной дискриминантной способности и изменяется от 0 до 1. Остаточная дискриминантная способность – это способность переменных различать классы, если исключить информацию, полученную с помощью ранее вычисленных функций. Для $Df 1$ и $Df 2$ значения λ Уилкса близки к 0 и намного меньше, чем значения λ Уилкса для 3-й и 4-й функций (табл. 3). Это означает, что центроиды классов хорошо разделены. На основе λ Уилкса получены значения теста значимости χ^2 . Значения χ^2 свидетельствуют о достоверности различий между классами по соответствующим функциям на уровне значимости $p=0,0001$.

Таблица 3 – Значения λ Уилкса и меры их существенности

Функция	λ Уилкса	χ^2	Число степеней свободы	p
$Df 1$	0,003	1665,59	48	0,0001
$Df 2$	0,026	1057,10	33	0,0001
$Df 3$	0,136	580,48	20	0,0001
$Df 4$	0,500	201,36	9	0,0001

Таким образом, для дискриминации растений 5-ти инбредных линий достаточны 1-я и 2-я дискриминантные функции.

В пространстве 1-й и 2-й дискриминантных функций инбредные линии ВК 392 и ВК 653 отделились от остальных (рис. 1). Прямая $D = y + 1,87x$ на графике позволяет разграничить области отдельных классов ($x=Df 1$, $y=Df 2$). Для инбредных линий ВК 392 и ВК 653 значение D

будет больше 1,11, а для ВК 389, ВК 678 и ВК 499 значение D будет меньше 1,11. Инбредные линии ВК 392 и ВК 653 разделились по Df 2, то есть оси ординат, разграничительное значение равно 1. Коэффициент дискриминации Любищева для этой пары инбредных линий равен 16,97, чему соответствует 1 ошибочное определение на 740 испытаний. Однако область инбредной линии ВК 389 на графике сильно перекрывается с областями ВК 678 и ВК 499. Поэтому эти три родительских линии объединены в условную группу. Рассчитаны коэффициенты дискриминации Любищева между этой группой и инбредными линиями ВК 392 и ВК 653. Для ВК 392 и группы из трех родительских линий $K=12,16$, то есть 1 ошибочное определение на 139 испытаний, а для ВК 653 и группы – $K=11,22$, соответственно, 1 ошибочное определение на 104 испытания. Стандартизированной мерой ошибок для любого количества классов является τ -статистика [6, с. 120]. Если считать группу ВК 389, ВК 678 и ВК 499 за один класс, то $\tau = 0,995$, значит, классификация с помощью дискриминантных функций дает на 99,5 % меньше ошибок, чем случайная классификация.

С помощью 1-й и 2-й дискриминантных функций не удалось различить инбредные линии ВК 389, ВК 678 и ВК 499. Поэтому был проведен дискриминантный анализ для этой группы с новым отбором переменных для линейных комбинаций. В таблицах 4 и 5 представлены основные статистические параметры этого анализа, свидетельствующие о значимости различий между классами в пространстве вычисленных функций. Различия между данными сортообразцами достоверны на уровне $p=0,0001$.

Таблица 4 – Собственные значения дискриминантных функций и меры существенности для группы линий ВК 389, ВК 678 и ВК 499

Функция	Собственное значение λ	Относительное процентное содержание	Каноническая корреляция
---------	--------------------------------	-------------------------------------	-------------------------

<i>Df I</i>	5,154	71,1	0,915
<i>Df II</i>	2,096	28,9	0,823

Таблица 5 – Значения λ Уилкса и меры их существенности для группы ВК 389, ВК 678 и ВК 499

Функция	λ Уилкса	χ^2	Число степеней свободы	p
<i>Df I</i>	0,052	509,85	18	0,0001
<i>Df II</i>	0,323	195,50	8	0,0001

Дискриминантные функции имеют вид:

$$DfI = -3,165 - 0,534 X1 + 0,561 X2 - 0,298 X3 - 0,047 X6 + 0,447 X7 - 1,381 X8 + 0,033 X10 + 0,013 X13 + 0,034 X14;$$

$$DfII = -5,710 + 0,289 X1 + 0,206 X2 - 0,383 X3 + 0,116 X6 + 0,456 X7 + 1,221 X8 + 0,034 X10 - 0,469 X13 - 0,027 X14,$$

где $X13$ – длина черешка листа, $X14$ – угол между нижними боковыми жилками листа, измеренный в градусах.

На графике (рис. 2) видно, что растения инбредной линии ВК 678 отделились от растений ВК 389 и ВК 499 без трансгрессии. Разграничительное значение между ВК 678 и ВК 389 по $Df I$ равно 1,5. Для растений ВК 678 значение $Df I$ больше 1,5, а для растений ВК 389 – меньше 1,5. Коэффициент дискриминации Любищева для пары ВК 678 и ВК 389 равен 10,18, чему соответствует 1 ошибочное определение на 76 испытаний. Линии ВК 678 и ВК 499 разделились по $Df I$ разграничительным значением 0,62. Растения линии ВК 678 имеют значение $Df I > 0,62$, а растения ВК 499 – $Df I < 0,62$. Для пары ВК 678 и ВК 499 коэффициент дискриминации Лю-

бищева $K=16,57$, значит, ошибочное определение будет встречаться 1 раз на 416 испытаний. Линии ВК 389 и ВК 499 разделились по $Df II$ с небольшой трансгрессией, разграничительное значение – 0.

Растения линии ВК 389 имеют значение $Df II$ больше 0, а растения линии ВК 499 – меньше 0. Коэффициент дискриминации Любищева для пары линий ВК 389 и ВК 499 по $Df II$ равен 5,05. Такому значению K соответствует 1 ошибочное определение на 18 испытаний. Для данного анализа $\tau = 0,95$. Это означает, что классификация с помощью дискриминантных функций дает на 95 % меньше ошибок, чем случайная классификация.

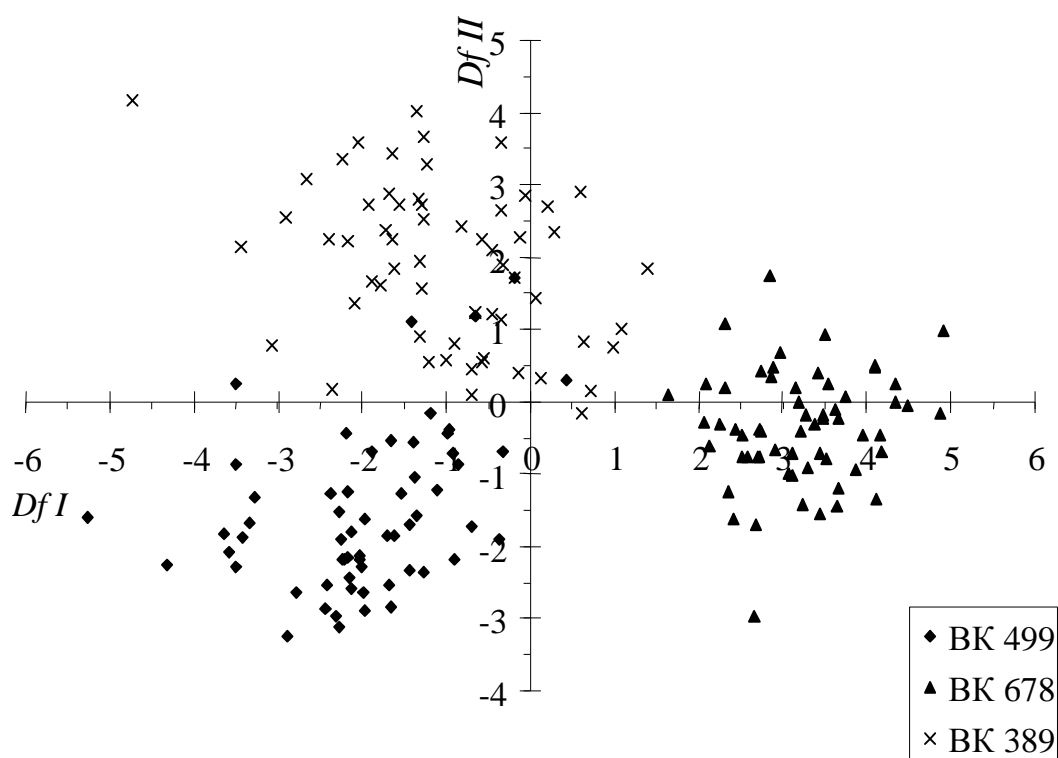


Рисунок 2 – Дискриминация растений 3-х родительских линий с помощью дискриминантных функций $Df I$ и $Df II$

Таким образом, показана ненадежность количественных признаков определителя UPOV для идентификации отдельных растений на примере 5-ти инбредных линий. Чтобы эффективно использовать различные количественные признаки в целях отличимости отдельных растений необходимо вести поиск соотношений этих признаков. Показана возможность ис-

пользования дискриминантного анализа для конструирования таких соотношений количественных признаков.

Список литературы

1. Вавилов, Н. И. Избранные сочинения. Генетика и селекция / Н. И. Вавилов. – М. : Колос, 1966. – 559 с.
2. Carcia, C. J. C. Importancia y usos de la descripcion varietal en sorgo / C. J. C. Carcia. – Revista Chapingo, 1985. – v. 10. – № 47–49. P. 187–193.
3. Guidelines for the conduct of tests for distinctness, uniformity and stability. TG/81/5(proj.) Sunflower, 2000-02-18.
4. Лакин, Г. Ф. Биометрия : учеб. пособие для биол. спец. вузов. – 4-е изд., перераб. и доп. – М. : Высш. шк., 1990. – 352 с.
5. Прикладная математика в биологии. – М. : Изд-во Моск. ун-та, 1979. – 152 с.
6. Ким, Дж.-О. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка : пер. с англ. / под ред. И. С. Енюкова. – М. : Финансы и статистика, 1989. – 215 с.