

MARCIN STUDNICKI¹**WIESŁAW MĄDRY**¹**TADEUSZ ŚMIAŁOWSKI**²¹ Katedra Doświadczalnictwa i Bioinformatyki, SGGW w Warszawie² Instytut Hodowli i Aklimatyzacji Roślin, Zakład Roślin Zbożowych w Krakowie

Wielocechowa analiza różnorodności fenotypowej w kolekcji roboczej pszenicy jarej

Multivariate analysis of phenotypic diversity in Polish spring wheat collection

Sukces programów hodowlanych roślin jest uzależniony między innymi od wiedzy o zmienności fenotypowej i genetycznej posiadanych kolekcji zasobów genowych oraz stopnia jej wykorzystania. Zastosowanie wielocechowych metod statystycznych jest ważnym składnikiem oceny różnorodności oraz klasyfikacji i typologii obiektów w obrębie zgromadzonych kolekcji zasobów genowych danego gatunku roślin lub materiałów hodowlanych w kolekcjach roboczych. Celem tej pracy jest ocena różnorodności fenotypowej 149 obiektów pszenicy jarej pod względem 12 cech ilościowych. Obiekty te zostały zgromadzone w kolekcji roboczej Zakładu Oceny Jakości i Metod Hodowli Zbóż Instytutu Hodowli i Aklimatyzacji Roślin w Krakowie. Niekompletną serię wieloletnią doświadczeń polowych przeprowadzono w miejscowości Kończewice w województwie Kujawsko-Pomorskim w latach 1996–1998. Niekompletność serii i danych sklasyfikowanych dwukierunkowo wynikała stąd, że niektóre obiekty badano nie we wszystkich latach. W pracy wykorzystano analizę skupień Warda i analizę składowych głównych, które pozwalają na komplementarne wnioski. Pierwsze cztery składowe główne wyjaśniły 67% ogólnej zmienności wielocechowej między badanymi obiektami. Pierwsza składowa wyjaśniająca 24% zmienności wielocechowej, była ona skorelowana głównie z plonem ziarna z poletka oraz masą ziarna z kłosa. Badane obiekty podzielono na 9 grup jednorodnych wielocechowo. Przedstawiono charakterystykę wielocechową obiektów w wydzielonych grupach oraz ocenę wielocechowego podobieństwa tych grup.

Słowa kluczowe: analiza składowych głównych, analiza skupień, BLUP, cechy rolnicze, kolekcja robocza, pszenica jara, różnorodność fenotypowa

The success of crop improvement programs depends among others on the knowledge of both phenotypic and genetic diversity within the maintained plant genetic resources and from the degree of their utilization. The use of multivariate statistical methods is an important strategy for assessment of within- collection diversity, classification and typology of accessions. The purpose of the paper was to evaluate phenotypic diversity for selected quantitative agronomic traits in a spring wheat collection (so-called “working” collection) maintained by the Department Unit of Cereal Breeding and Quality Evaluation of the Plant Breeding and Acclimatization Institute in Kraków. A field trial including 149 accessions (released cultivars and advanced breeding lines) was conducted across three seasons (1996–

1999) in Kończewice, Kujawy-Pomerania province. In each year 12 agronomic quantitative traits were evaluated. The data were put into incomplete two-way accession \times year classification for each trait. Two complementary multivariate methods including cluster analysis and principal component analysis were used. The first four principal components captured 67% of the total multivariate variability among the accessions. The first principal component accounted for 24% of the total variability and it was mostly correlated with grain yield and weight of grain per ear as the traits with the largest contribution to the multivariate variability among the accessions. The accessions were grouped into nine clusters based on Ward's clustering method. The clusters (groups) were then characterized with regard to all the studied traits, and multivariate similarities among the groups were described using the biplot in principal component analysis.

Key words: agronomic traits, BLUP, cluster analysis, genetic resources, phenotypic diversity, principal component analysis, spring wheat

WSTĘP

Ważnym czynnikiem rozwoju programów hodowlanych roślin jest stopień wykorzystania różnorodności biologicznej (genetycznej i fenotypowej) w kolekcjach dotychczas zgromadzonych zasobów genowych. Jednym z ważnych elementów strategii wykorzystania obiektów w tych kolekcjach jest ocena genetyczna i fenotypowa ich różnorodności. W wielu przypadkach zadanie to jest ograniczone tylko do pozyskania danych paszportowych dla obiektów w kolekcjach. Kompleksowa, tzn. genetyczna i fenotypowa ocena różnorodności kolekcji roślinnych zasobów genowych wymaga współpracy specjalistów z wielu dziedzin, tj. genetyków, hodowców, biochemików, fitopatologów, fizjologów roślin, biometryków i innych (Mazurkiewicz i Struś, 1997; Brown-Guedira i in., 2000; Mohammadi i Prasanna, 2003; Mądry, 2007; Sorrells, 2007; Damania, 2008). Wkład biometrii do tej współpracy polega na opracowywaniu, doskonaleniu i zastosowaniu efektywnych metod statystycznych (Mohammadi i Prasanna 2003, Crossa i Franco, 2004; Martinez-Calvo i in., 2007; Mądry, 2007).

W ramach systematycznych i zintegrowanych prac nad gromadzeniem, przechowywaniem, oceną oraz wykorzystaniem pozyskanych lub wytworzonych dotychczas obiektów pszenicy w Polsce, utworzono w 1970 roku kolekcję roboczą przy Zakładzie Roślin Zbożowych IHAR. Do tego czasu polskie ośrodki hodowli prowadziły własne kolekcje (Mazurkiewicz i Struś, 1997). Kolekcja robocza jest to zbiór wartościowych obiektów genetycznie utrwalonych, które mogą być źródłem materiałów wyjściowych do hodowli nowych odmian (Mohammadi i Prasanna, 2003; Damania, 2008). W skład takiej kolekcji wchodzi zarówno odmiany zarejestrowane w kraju, odmiany pochodzące z zagranicznych ośrodków hodowli oraz wartościowe linie zdeponowane przez hodowców. Zgromadzone obiekty są badane pod względem wielu rolniczych cech ilościowych w okresie kilku lat w jednej miejscowości (Kubicka i in., 2004; Hartung i Piepho, 2005; Mądry, 2007). W kolejnych latach badań obserwuje się obiekty na pojedynczych poletkach. Dane z takiej serii lat stanowią niekompletną klasyfikację dwukierunkową krzyżową z jedną obserwacją w podklasie w niepustych podklasach. Niekompletność danych wynika z tego, że w poszczególnych latach nie są badane wszystkie obiekty.

Wielocechowa analiza statystyczna danych z obserwacji fenotypowych cech obiektów w kolekcji, wykonywana przy użyciu metod wielowymiarowych, jest bardzo wartościowo-

wym podejściem do oceny różnorodności, pozwalającym na pogłębioną i wyczerpującą ocenę kolekcji zasobów genowych roślin uprawnych (Mohammadi i Prasanna, 2003; Crossa i Franco, 2004; Martinez-Calvo i in., 2007; Mądry, 2007). Umożliwia ona klasyfikację obiektów zgromadzonych w kolekcji na grupy jednorodne pod względem wielu cech jednocześnie oraz ich charakterystykę, zwaną też typologią, które ułatwiają jasne opisanie podobieństwa pomiędzy obiektami. Za pomocą metod wielowymiarowych można określić także związki pomiędzy cechami, względną moc dyskryminacyjną (rozdzielającą obiekty) badanych cech i charakter podobieństwa pomiędzy obiektami pod względem najważniejszych cech, czyli tych o największej mocy dyskryminacyjnej.

Wielowymiarowe metody statystyczne z sukcesem zostały zastosowane do klasyfikacji i charakterystyki zmienności fenotypowej kolekcji wielu gatunków roślin, między innymi pszenicy (DeLacy i in., 2000; Hailu i in., 2006), jęczmienia (Žáková i Benkova, 2006; Ruiz i in., 1997; Manjunatha i in., 2007), żyta (Kubicka, 2004), czy amarantusa (Mapes i in., 1996).

Celem tej pracy jest ocena różnorodności fenotypowej pod względem 12 rolniczych cech ilościowych obiektów w kolekcji roboczej pszenicy jarej, zgromadzonej w Instytucie Hodowli i Aklimatyzacji Roślin, za pomocą wielowymiarowych metod statystycznych, tj. analizy skupień i analizy składowych głównych.

MATERIAŁ I METODY

Materiał badawczy i doświadczenia polowe

Materiał badawczy obejmował 149 obiektów kolekcji roboczej pszenicy jarej (tab. 3), zgromadzonej i ocenianej w Zakładzie Oceny Jakości i Metod Hodowli Zbóż Instytutu Hodowli i Aklimatyzacji Roślin w Krakowie. Doświadczenia polowe przeprowadzono w latach 1996–1998 w Kończewicach, w województwie kujawsko-pomorskim. Ziarniaki z każdego obiektu wysiano na pojedynczych poletkach o powierzchni 10 m².

W trakcie doświadczenia na poletkach obserwowano 12 następujących rolniczych cech ilościowych: liczbę dni od 1.IV. do kłoszenia i dojrzałości oraz wysokość roślin w łanie w centymetrach. Pomiary te dokonywano na podstawie obserwacji 50 roślin losowo wybranych z poletka. W czasie wegetacji roślin określono porażenie chorobami liści oraz porażenie chorobami kłosa w 9-stopniowej skali bonitacyjnej, w której 1 — oznacza największe porażenie chorobą, a 9 — najmniejsze porażenie. Po zbiorze określono plon ziarna w dt/ha i jego składowe — liczbę ziaren z kłosa, masę ziarna z kłosa w g i masę 1000 ziaren w g. Oceniano także cechy technologiczne ziarna, takie jak procentowa zawartość białka w ziarnie — oznaczona metodą bliskiej podczerwieni NIR, liczba opadania — oznaczona metodą Hagberga-Pertena, liczba sedymentacji — oznaczona metodą Axforda z wykorzystaniem siarczanu dodecylosodowego SDS.

Analiza statystyczna danych

Jednocechowa analiza zmienności kolekcji

Obserwacje fenotypowe badanych cech u 149 obiektów były analizowane na podstawie modelu mieszanego analizy wariancji dla niekompletnej klasyfikacji podwójnej z jedną

obserwacją w podklasie o postaci obiekty x lata. Model ten ma następującą postać (Jahufer i in., 1997; Piepho i Mohring 2005; Littell i in., 2006):

$$y_{ij} = m + g_i + r_j + e_{ij}$$

gdzie:

y_{ij} jest obserwacją danej cechy u i -tego obiektu ($i = 1, \dots, a$) w j -tym roku ($j = 1, \dots, b$), m jest średnią ogólną, g_i jest losowym efektem głównym i -tego obiektu, r_j jest stałym efektem głównym j -tego roku, e_{ij} jest składnikiem losowym, obejmującym efekt interakcji i -tego obiektu ($i = 1, \dots, a$) w j -tym roku oraz błąd doświadczalny.

Analizę wariacji według przyjętego powyżej modelu mieszanego wykonano za pomocą pakietu statystycznego SAS 9.1, wykorzystując procedurę GLM (SAS Institute Inc. 2004, Littell i in. 2006).

Losowe efekty główne dla obiektów g_i dla każdej cechy, ze względu na niekompletność danych, zostały oszacowane za pomocą najlepszego nieobciążonego liniowego predyktora, oznaczanego w skrócie BLUP (ang. the Best Linear Unbiased Predictor) (Robinson 1991, Jahufer i in. 1997, Hartung i Piepho 2005, Hartung i in. 2006). Do wyznaczenia estymatorów BLUP zastosowano metodę największej wiarygodności z restrycją REML (ang. Restricted Maximum Likelihood). Estymatory BLUP dla efektów losowych, dotyczących poziomów danego czynnika, posiadają mniejszy standardowy błąd oszacowania, niż najlepsze nieobciążone liniowe estymatory typu BLUE (ang. the Best Linear Unbiased Estimators), stosowane dla efektów stałych, dotyczących poziomów tego samego czynnika (Robinson, 1991; Hartung i Piepho, 2005). Obliczenia numeryczne wykonano za pomocą pakietu statystycznego SAS 9.1, wykorzystując procedurę MIXED, pozwalającą wyznaczyć estymatory typu BLUP dla efektów losowych badanych obiektów (Littell i in., 2006; SAS Institute Inc., 2004).

Dla potrzeb jednocechowej i wielocechowej analizy zmienności fenotypowej obiektów zgromadzonych w kolekcji pszenicy jarej wyznaczono oceny średnich genotypowych, posługując się estymatorem:

$$\hat{m}_i = \hat{m} + \hat{g}_i$$

gdzie:

\hat{m}_i jest oceną nieobciążoną średniej genotypowej dla danej cechy (oceną średniej cechy w populacji lat) dotyczącej i -tego obiektu, \hat{m} jest oceną średniej ogólnej, \hat{g}_i jest oceną BLUP losowego efektu głównego dotyczącego i -tego obiektu.

Oceny średnich genotypowych \hat{m}_i dla obiektów posłużyły do określenia podstawowych charakterystyk (średnia, wariancja, współczynnik zmienności CV, minimalna i maksymalna wartość) obserwowanych cech w obrębie zbioru badanych obiektów. Średnie genotypowe \hat{m}_i będą wykorzystywane także w wielocechowej analizie zmienności rozpatrywanej kolekcji.

Wielocechowa analiza różnorodności kolekcji

Analiza skupień

Badane obiekty podzielono na grupy (skupienia) o małej wielocechowej różnorodności genotypowej, posługując się metodą Warda analizy skupień (Jahufer i in., 1997; Casler i van Santen, 2000; Mohammadi i Prasanna, 2003; Crossa i Franco, 2004; Mądry, 2007). W

tej hierarchicznej metodzie analizy skupień w czasie aglomeracji wykorzystywane jest podejście analizy wariancji, oparte na minimalizacji sumy kwadratów odchyłeń wewnątrz grup (ang. within-group error sum of squares). W każdym kroku aglomeracji następuje połączenie takich dwóch grup, dla których suma kwadratów odchyłeń jest najmniejsza spośród wszystkich możliwych do utworzenia grup (Johnson i Wichern, 2002). Ta metoda charakteryzuje się podziałem obiektów na grupy o podobnej i małej liczbie relatywnie wyrównanych (jednorodnych) obiektów. Z tego powodu jest ona uznana za bardzo efektywną przy podziale obiektów na grupy jednorodne i najczęściej stosowana w ocenie zarówno genotypowej, jak i fenotypowej zmienności kolekcji roślinnych zasobów genowych (Mohammadi i Prasanna, 2003; Crossa i Franco, 2004; Mądry, 2007). Jako miarę podobieństwa pomiędzy obiektami zastosowano kwadrat odległości euklidesowej, wyznaczonej na podstawie standaryzowanych ocen średnich genotypowych, dotyczących rozpatrywanych cech (Mohammadi i Prasanna, 2003; Crossa i Franco, 2004; Mądry, 2007). Ważnym problemem w analizie skupień jest wyznaczenie właściwej (rozsądnej merytorycznie i uzasadnionej statystycznie) liczby grup, często określanym jako miejsce przecięcia dendrogramu. W celu wyznaczenia uzasadnionej statystycznie liczby skupień, zastosowano w pracy kryterium oparte na statystyce pseudo F i pseudo t^2 . Kryteria te nie pozwoliły jednak jednoznacznie podjąć decyzję o liczbie skupień. Zastosowano drugie kryterium zaproponowane przez Caslera i van Santena (2000) oraz Upadhyaya i wsp. (2003). Jest ono oparte na relacji międzygrupowej sumy kwadratów odchyłeń do całkowitej sumy kwadratów odchyłeń dla wszystkich badanych cech. Przyjmuje się zwykle, że obiekty mogą być podzielone na taką liczbę skupień, aby wewnątrzgrupowa suma kwadratów wyjaśniła znaczącą część całkowitej sumy kwadratów odchyłeń, zwykle około 75%.

Analiza składowych głównych

Ta metoda jest stosowana do:

- syntetycznej oceny związków pomiędzy badanymi cechami ilościowymi, polegającej na wykryciu tzw. czynników — składowych głównych, rozumianych jako zespoły cech wzajemnie skorelowanych,
- określenia względnej mocy dyskryminacyjnej (rozdzielającej obiekty) każdej z badanych cech, prowadzącej do wykrycia cech o największym znaczeniu w wielocechowej zmienności między obiektami,
- charakterystyki podobieństwa pomiędzy obiektami pod względem najważniejszych cech, czyli tych o największej mocy dyskryminacyjnej za pomocą narzędzi graficznych, tj. wykresu rozmieszczenia obiektów w układzie dwóch lub trzech składowych głównych (Ruiz i in., 1997; Assafa i in., 2001, 2003; Johnson i Wichern, 2002; Mohammadi i Prasanna, 2003; Kubicka i in., 2004; Mądry, 2007).

Analiza składowych głównych została przeprowadzona za pomocą procedury PRINCOMP pakietu statystycznego SAS 9.1 (SAS Institute Inc., 2004).

WYNIKI I DYSKUSJA

Analiza wariancji i parametry zmienności cech

Wyniki analizy wariancji dla 12 rozpatrywanych cech, opartej na modelu mieszanym, przedstawiono w tabeli 1. Stwierdzono istotną zmienność genotypową obiektów dla każdej z rozpatrywanych cech.

Tabela 1

Analiza wariancji dla 12 cech rolniczych badanych obiektów w kolekcji roboczej pszenicy jarej
Analysis of variance for 12 agronomic traits of studied accessions in a spring wheat collection

Źródła zmienności Sources of variation	D.F.	Średnie kwadraty Mean squares											
		X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂
Obiekty Object	148	211,93*	5881,32*	165,11*	15,69*	22,09*	191,96*	443,69*	32,53*	2,16*	362,91*	1,80*	1,35*
Lata Years	2	17,78*	26179,82*	11224,65*	3529,45*	2383,28*	8649,51*	104,59*	386,55*	0,080*	31,96*	62,09*	4,29*
Resztowa Residual	148	0,56	1787,30	64,64	2,84	3,16	53,41	14,64	9,91	0,03	5,51	1,02	0,61

* Istotne przy $\alpha = 0,01$; Significant at $\alpha = 0,01$

D.F. — Stopnie swobody; Degrees of freedom

Tabela 2

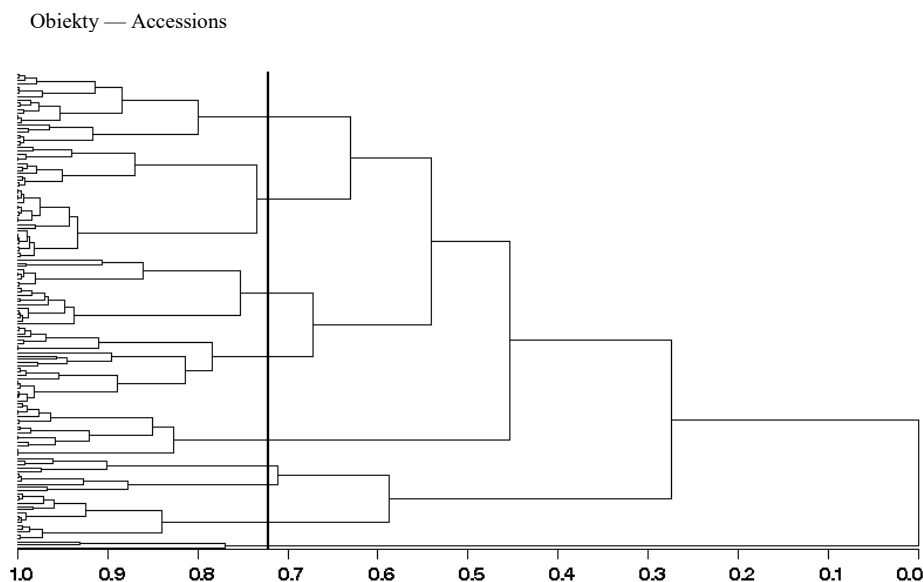
Charakterystyka zmienności kolekcji roboczej pszenicy jarej pod względem 12 cech rolniczych
Variability characterization for 12 agronomic traits in a spring wheat collection

Cecha Trait	Średnia Mean	Wariancja Variance	Współczynnik zmienności (%) Coefficient of variability (%)	Min. Min	Maks. Max
Zawartość białka w ziarnie Grain protein content	12,96	0,25	3,87	11,73	15,04
Liczba opadania Falling number	295,45	1285,49	12,14	174,53	351,00
Test sedimentacji Sedimentation test	45,37	28,10	11,68	28,66	61,75
Dni do dojrzałości Days to maturity	102,31	5,18	2,23	97,55	107,65
Dni do kłoszenia Days to heading	48,19	7,66	5,74	41,92	54,80
Plon ziarna Grain yield	72,64	47,56	9,49	46,65	87,09
Wysokość rośliny Plant height	95,40	37,09	6,38	70,42	110,97
Liczba ziaren w kłosie Grain number per ear	41,52	7,36	6,53	33,98	49,78
Masa ziaren w kłosie Grain weight per ear	1,84	0,02	7,23	1,34	2,20
Masa 1000 ziaren Weight of 1000 grains	44,47	10,14	7,16	37,27	53,74
Porażenie chorobami liści Resistance to leaf diseases	5,57	0,18	7,69	4,27	6,71
Porażenie chorobami kłosa Resistance to ear diseases	6,21	0,20	7,13	4,75	7,24

Na podstawie ocen odchyłeń standardowych i współczynników zmienności dla średnich genotypowych, otrzymanych za pomocą BLUP-ów (tab. 2) możemy wnioskować, że największą zmiennością charakteryzowała się liczba opadania, (współczynnik zmienności był równy 12%), liczba sedymentacji, (współczynnik zmienności był równy 11,7%), a także plon ziarna z poletka, (współczynnik zmienności był równy 9,5%). Najmniejszą zmiennością charakteryzowały się następujące cechy: liczba dni do kłoszenia, (współczynnik zmienności był równy 5,7%), zawartość białka w ziarnie, (współczynnik zmienności kształtował się na poziomie 3,9%) oraz liczba dni do dojrzałości pełnej, (współczynnik zmienności był równy 2,2%).

Grupowanie obiektów wielocechowo jednorodnych

Na rysunku 1 przedstawiono dendrogram, obrazujący proces grupowania ocenianych obiektów za pomocą metody Warda, wraz z miejscem przecięcia jego gałęzi na poziomie 71% całkowitej sumy kwadratów odchyłeń cech (zmienności całkowitej), wyjaśnionej przez skupienia. Podobny poziom wyjaśnionej zmienności przez skupienia stanowił kryterium do wydzielania skupień w pracach Caslera i van Santena (2000) oraz Upadhyaya i wsp. (2003). Zatem 149 obiektów podzielono na 9 grup jednorodnych pod względem wszystkich 12 ocenianych cech.



Udział międzygrupowej zmienności cechy w całkowitej zmienności obiektów
The proportion of variance accounted for by the clusters

Rys. 1. Dendrogram uzyskany za pomocą metody Warda analizy skupień dla 149 obiektów pszenicy jarej (linia pionowa obrazuje przecięcie gałęzi na poziomie 71% zmienności całkowitej wyjaśnionej przez skupienia)

Fig. 1. Dendrogram generated by Ward cluster analysis for 149 accessions of spring wheat (vertical line shows a cutting point at the level of 71% variability explained by clusters)

W tabeli 3 zaprezentowano skład obiektów w poszczególnych grupach jednorodnych, w których znalazły się obiekty pochodzące z różnych źródeł. Liczebność wydzielonych grup wahała się od 35 obiektów w grupie 1, do 3 obiektów w grupie 9. W każdej grupie znajdowały się zarówno linie hodowlane, wytworzone w różnych ośrodkach hodowlanych oraz odmiany zarejestrowane w kraju i zagranicą. Oznacza to, że hodowcy w różnych ośrodkach hodowlanych korzystali z tej samej lub spokrewnionej bazy genetycznej pszenicy. Do stwierdzenia faktycznego spokrewnienia niektórych obiektów w tej kolekcji konieczne jest badanie ich podobieństwa genetycznego za pomocą markerów molekularnych (Brown-Guedira i in., 2000; Mohammadi i Prasanna, 2003; Mądry, 2007; Damania, 2008) lub prześledzenie rodowodów zgromadzonych obiektów.

Tabela 3

Grupy jednorodne obiektów pszenicy jarej wydzielone na podstawie 12 cech rolniczych za pomocą metody Warda analizy skupień
Homogeneous groups of spring wheat accessions revealed on the basis of 12 agronomic traits using the Ward clustering method

Grupa Group	Nazwa obiektu Name of accession
1	CSW92, ETA, Frisal, HEC 1341/90, HEC 2527/91, HEC 347/90, Henika, Ibwsn 383/81, Jasna, KOC 1887/84, KOC 1997/92, KOC 279, KOC 3612/88, KOC 3708/94, KOC 4162/90, KOH 1749, KOH 1809, KOH 1813, KOH 1831, KOH 2365, KOH 393, LONA, Munk, NAD 196, NAD 297, Nemaes, Omega, Planet, SOA 2896/87, Thasos, Tinos, Turbo, VDH 3022/89, WYS. z KEN.LEOP., Zena
2	ALSN 11, ALSN 26, ANDUSIL, Cadenza, Contesa, CSW 91/B, IBWSN 144/79, KOC 1660/93, KOC 2083/90, KOH 2242, KOH 2254, KOH 263/15/89, KOH 272/11/83, KOH 391/16/82, RAH 15393, RAH 3794, RAH 63593, RAH 66493, Sigma, SOA 2531/85, SOA 280/R/91, Sunnan, Troy, Waller
3	Alkora, Erwin, HEC 425/93, HEC 493, HEC 587/87, IGNA, KOC 2391/89, KOC 2612/88, KOC 2775/87, KOC 2823/92, KOC 3024/90, KOC 3317/89, KOH 1671, KOH 2182, KOH 253/13/88, KOH 60/13/88, KOH 800
4	6/b40, AL SN-11, ALBIS, COMBI, FAVORIT, HEC 1941/92, HEC 1945/91, HEC 2530/91, HENA, Hermes, IBWSN 260/81, JOAN DOLLAR, KOC 2892, KOC 3277/92, KOC 3395, KOH 1094, KOH 696, KRJ 7/95, M 104/77, M 159/75, LANDU
5	2/B/19, Artis, Attis, Devon, Dragon, Hanno, HEC 1396/80, HEC 393, HEC 594, HEC 961/88, Hezja, Klaros, KOC 1177/92, KOC 2815/89, KOC 4798/90, KOH 1373, KOH 1895, Quatro, Ralle, Remia, Santa, Torka, VDH 83/2014/1/3
6	Banti, Ismena, Kenya Tembo, KOC 3516/86, Naxos, S-55
7	Helia, KOC 4479/90, KOH 1766, KOH 1767, KOH 1962, KOH 1995, KOH 2384, KOH 2391, KOH 495 / Nawra, KOH 596, KOH 742, KOH 797, M 151/76, M 279/74, Rubino
8	Arkada, IBWSN 390/81, M 203/78, M 70/78, Remus
9	ALSN 25, KRJ 4/95, KRJ 9/95

Analiza składowych głównych

Wyniki analizy składowych głównych na podstawie standaryzowanych wartości BLUP-ów dla średnich genotypowych, dotyczących badanych cech, przedstawiono w tabeli 4. Do dalszej analizy wybrano cztery pierwsze składowe główne, ponieważ dla nich wartości własne były większe od 1 (Assefa i in., 2003; Mohammadi i Prasanna, 2003; Kubicka i in., 2004; Okpul i in., 2004; Mądry, 2007).

Cztery pierwsze składowe główne wyjaśniły 67% ogólnej zmienności międzyobiektowej dla wszystkich rozpatrywanych cech. Pierwsza z tych składowych wyjaśniła 24% łącznej zmienności BLUP-ów dla efektów genotypowych, natomiast dwie pierwsze

wyjaśniły łącznie 44% tej zmienności. Trzy pierwsze składowe wyjaśniły łącznie 56% ogólnej zmienności obiektów.

Tabela 4

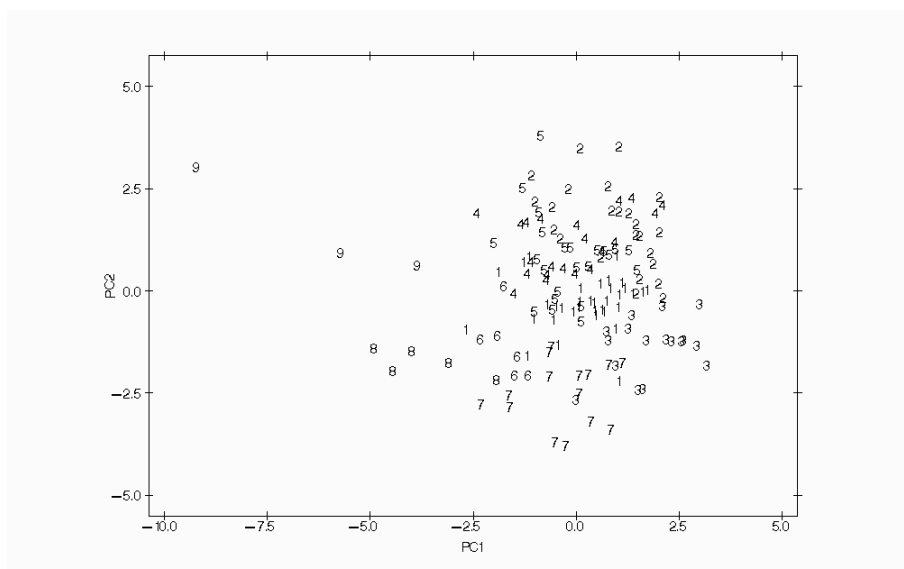
Współczynniki korelacji między cechami a każdą z czterech pierwszych składowych głównych (PC) oraz wartości własne i procent wyjaśnionej zmienności przez te składowe główne
Correlation coefficients between traits and each of the first four principal components (PC), eigenvalues and parts of multivariate variability explained by these components

Cechy — Traits	PC 1	PC 2	PC 3	PC 4
Zawartość białka w ziarnie — Grain protein content	-0,30	0,25	0,35	-0,05
Liczba opadania — Falling number	0,32	-0,01	-0,40	0,29
Test sedymentacji — Sedimentation test	-0,25	-0,02	0,07	0,21
Dni do dojrzałości — Days to maturity	0,26	0,49	0,06	0,00
Dni do kłoszenia — Days to heading	0,31	0,45	-0,06	0,18
Plon ziarna — Grain yield	0,41	-0,15	-0,06	0,20
Wysokość rośliny — Plant height	0,13	0,20	0,61	-0,03
Liczba ziaren w kłosie — Grain number per ear	0,30	-0,05	-0,10	-0,69
Masa ziaren w kłosie — Grain weight per ear	0,37	-0,32	0,28	-0,34
Masa 1000 ziaren — Weight of 1000 grains	0,15	-0,37	0,46	0,34
Porażenie chorobami liści — Resistance to leaf diseases	0,32	0,31	0,14	0,13
Porażenie chorobami kłosa — Resistance to ear diseases	0,24	-0,31	0,05	0,27
Wartość własna Eigenvalues	2,93	2,42	1,39	1,30
Wyjaśniona część zmienności wielocechowej obiektów Explained part of multivariate variability of accessions	0,24	0,20	0,12	0,11
Skumulowana część zmienności wielocechowej Cumulative part of multivariate variability	0,24	0,45	0,56	0,67

Pierwsza składowa główna była stosunkowo silnie skorelowana dodatnio z plonem ziarna ($r = 0,41$), masą ziarna z kłos ($r = 0,37$), liczbą opadania ($r = 0,32$), a także porażeniem przez choroby liści ($r = 0,32$). Wskazuje to, że wymienione cztery cechy były wzajemnie dość silnie skorelowane dodatnio. Są one oznaką istnienia czynnika genetycznego (także w części środowiskowego), który w znacznym stopniu warunkuje jednocześnie wszystkie te cztery cechy. Ten czynnik może obejmować poligeny sprzężone lub/oraz działające plejotropowo. Cechy najsilniej skorelowane z pierwszą składową główną miały największy udział w wielocechowym zróżnicowaniu obiektów (tab. 5, rys. 2 i 3). Stosunkowo duże wartości pierwszej składowej głównej dla rozpatrywanych obiektów oznaczają dość często duże wartości dodatnio skorelowanych z nią cech, zaś małe wartości ujemnie skorelowanych z nią cech. Natomiast małe wartości pierwszej składowej głównej dla obiektów, wskazują dość często na małe wartości dodatnio skorelowanych z nią cech, zaś duże wartości ujemnie skorelowanych z nią cech.

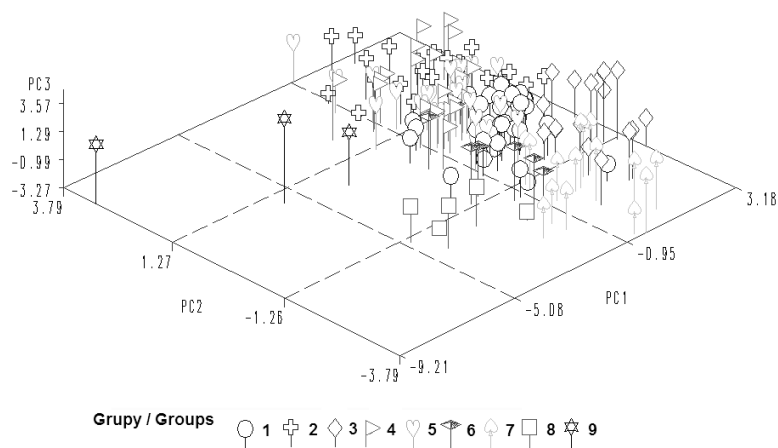
Dla drugiej składowej głównej stwierdzono dodatnią korelację z liczbą dni do dojrzałości ($r = 0,49$) oraz liczbą dni do kłoszenia ($r = 0,45$), natomiast ujemną korelację ($r = -0,37$) dla masy 1000 ziaren. Wskazuje to na ujemne korelacje między długością każdej z dwóch faz rozwojowych a masą 1000 ziaren. Stwierdzony związek odmienny w stosunku do wielu badań, potwierdzających korelacje dodatnie między długością faz rozwojowych a masą 1000 ziaren u zbóż (DeLacy i in., 2000; Hailu i in., 2006). Zatem obiekty o relatywnie dużych wartościach drugiej składowej głównej wyróżniały się często relatywnie

długim okresem wegetacji oraz charakteryzowały się niską masą 1000 ziaren (grupa 2), i odwrotnie (grupa 7), co dobrze obrazuje tabela 5.



Rys. 2. Rozmieszczenie 9 grup obiektów pszenicy jarej w układzie dwóch pierwszych składowych głównych

Fig. 2. Scatter plot for 9 groups of spring wheat in axes of the first two principal components



Rys. 3. Rozmieszczenie 9 grup obiektów pszenicy jarej w układzie trzech pierwszych składowych głównych

Fig. 3. Scatter plot for 9 groups of spring wheat in axes of the first three principal components

Tabela 5

Charakterystyka wielocechowa wydzielonych grup obiektów pszenicy jarej
Characteristics of separated accession groups for 12 agronomic traits in a spring wheat collection

Cechy Traits	Grupy Groups					
	I		II		III	
	średnia mean	wariancja variance	średnia mean	wariancja variance	średnia mean	wariancja variance
1	2	3	4	5	6	7
X ₁	12,83	0,15	12,87	0,11	12,60	0,13
X ₂	312,64	317,29	308,85	797,63	303,69	1334,21
X ₃	45,06	15,65	40,39	30,86	43,38	26,40
X ₄	101,68	1,73	104,58	1,97	101,41	1,65
X ₅	48,13	4,32	51,17	3,25	47,13	5,70
X ₆	74,81	16,20	71,73	34,90	79,34	10,72
X ₇	93,09	23,44	95,44	36,71	98,01	18,49
X ₈	42,34	2,94	42,05	3,31	44,90	3,51
X ₉	1,83	0,00	1,80	0,01	2,04	0,01
X ₁₀	42,93	3,22	42,48	9,35	46,08	6,32
X ₁₁	5,45	0,08	5,84	0,08	5,71	0,11
X ₁₂	6,31	0,21	6,01	0,12	6,40	0,11
	IV		V		VI	
X ₁	13,41	0,10	12,95	0,29	12,94	0,27
X ₂	282,37	1076,95	298,23	445,50	220,63	897,81
X ₃	48,98	17,34	48,16	13,71	41,56	7,19
X ₄	104,36	4,11	103,06	2,84	99,65	2,46
X ₅	49,02	2,57	49,36	2,90	43,80	3,85
X ₆	69,85	38,08	73,70	19,99	70,51	89,74
X ₇	98,58	25,56	97,67	16,43	96,43	19,71
X ₈	42,67	6,87	38,46	2,51	39,87	1,68
X ₉	1,87	0,01	1,73	0,01	1,79	0,00
X ₁₀	44,04	4,67	45,05	6,74	45,21	2,69
X ₁₁	5,70	0,05	5,84	0,11	5,44	0,18
X ₁₂	5,97	0,09	6,26	0,14	6,33	0,13
	13,41	0,10	12,95	0,29	12,94	0,27
	VII		VIII		IX	
X ₁	12,89	0,07	12,99	0,34	14,51	0,40
X ₂	295,79	703,82	266,64	1991,35	207,20	2726,18
X ₃	46,91	19,04	46,39	5,71	51,75	75,33
X ₄	100,03	0,86	98,75	0,34	99,13	0,61
X ₅	45,94	3,44	43,32	1,13	44,34	0,64
X ₆	72,74	26,64	61,44	36,30	50,50	41,05
X ₇	91,55	25,02	82,97	58,15	105,65	23,87
X ₈	40,35	2,26	38,52	1,02	38,35	20,42
X ₉	1,95	0,01	1,73	0,00	1,61	0,06
X ₁₀	49,54	7,39	44,87	6,79	40,62	10,41
X ₁₁	5,14	0,22	4,73	0,13	4,75	0,01
X ₁₂	6,66	0,04	5,80	0,03	5,09	0,17

Trzecia składowa główna była dodatnio skorelowana z wysokością roślin ($r = 0,61$), ujemnie zaś z liczbą opadania ($r = -0,40$). To świadczy, że obiekty o relatywnie wysokich źdźbłach, posiadały często relatywnie niską wartość liczby opadania. Natomiast czwarta składowa główna była dość silnie skorelowana ujemnie z liczbą ziaren w kłosie ($r = -0,69$).

Brak silnych korelacji pomiędzy składowymi głównymi a cechami, świadczą o tym, że nie ma wśród badanych cech takich, które mają stosunkowo duży udział w wielocechowym zróżnicowaniu obiektów. Żadna z badanych cech nie posiadała dominującego udziału w ogólnej zmienności obiektów.

Z rysunków 2 i 3 wynika, że większość obiektów zlokalizowana jest w okolicach wartości średnich dla trzech pierwszych składowych głównych, zwłaszcza pierwsza, czwarta i piąta grupa obiektów. Obiekty w tych grupach odznaczały się średnimi wartościami dla cech znacząco skorelowanych z pierwszą i drugą składową główną.

Grupa 2 obiektów charakteryzowała się dużymi dodatnimi wartościami pierwszej i drugiej składowej głównej, stwierdzono u nich relatywnie duże wartości dla liczby opadania, liczby dni do kłoszenia i dojrzałości, plonu ziarna, masy ziarna z kłosa i porażenia chorobami liści (małe porażenie), a także poniżej średniej ogólnej wartości dla masy 1000 ziaren. Potwierdzają to średnie grupowe tych cech, podane w tabeli 5. Te relacje wartości cech obiektów w grupie 2 wynikają z siły i znaku korelacji z pierwszymi dwiema składowymi głównymi (tab. 4). Obiekty w grupie 9 (najbardziej odstające od pozostałych) miały najniższe (ujemne) wartości pierwszej składowej głównej i jedne z najwyższych wartości drugiej składowej głównej. Wyróżniały się one relatywnie długim okresem do kłoszenia i dojrzałości i niską masą 1000 ziaren oraz relatywnie dużą wysokością roślin, co przedstawia tabela 5. Grupa 3 zawierała przede wszystkim obiekty o dużej liczbie opadania i plonie ziarna, które posiadają najwyższe wartości dla pierwszej składowej głównej (cechy stosunkowo silnie skorelowane z tą składową główną), a także krótkim okresem wegetacji (niskie wartości dla drugiej składowej głównej). Niskimi roślinami były przede wszystkim obiekty należące do grupy 1 oraz 8, obiekty te posiadały niskie wartości dla cech ujemnie skorelowanych z trzecią składową główną.

Wśród 149 ocenianych obiektów w kolekcji roboczej występują takie, które znacznie odbiegają od średnich wartości zarówno dla pierwszej, drugiej jak i trzeciej składowej głównej. Analizując rysunki 2 i 3 możemy takie obiekty wyłonić. Należały one przeważnie do grupy 3, 8 i 9. Obiekty z tych grup różnią się najbardziej od innych badanych obiektów w kolekcji, odznaczając się kontrastowymi poziomami cech najsilniej skorelowanych z dwiema pierwszymi składowymi głównymi, tzn. cech o relatywnie największej mocy dyskryminacyjnej (tj. plon ziarna, masa ziaren z kłosa, masa 1000 ziaren, liczba opadania, liczba dni do kłoszenia i dojrzałości, a także oporność na porażenie przez choroby liści). Mogą one zostać wykorzystane jako komponenty rodzicielskie w trakcie hodowli nowych odmian pszenicy jarej. Przede wszystkim są to następujące obiekty Alkora, Erwin, HEC 425/93, HEC 493, HEC 587/87, KRJ 4/95, KRJ 9/95, M 203/78, M 70/78, Remus.

WNIOSKI

Wykorzystanie najlepszych liniowych nieobciążonych predyktorów (estymatorów BLUP) dla średnich genotypowych badanych cech na podstawie danych niekompletnych w tablicy dwukierunkowej o postaci obiekty \times lata, pozwala efektywnie i poprawnie pod względem statystycznym, zastosować wielocechowe metody oceny fenotypowej

różnorodności obiektów, zgromadzonych w kolekcjach zasobów genowych lub kolekcjach roboczych.

Obydwie metody wielocechowe (analiza skupień i analiza składowych głównych), uzupełniają się wzajemnie (są komplementarne) w efektywnej ocenie różnorodności fenotypowej obiektów, zgromadzonych w kolekcjach zasobów genowych. Taka ocena dostarcza informacji, które mogą być wykorzystane do lepszego zarządzania kolekcją zasobów genowych lub kolekcją roboczą, a także ułatwia wykorzystanie obiektów z takich kolekcji w hodowli roślin do wyboru form rodzicielskich.

Badana kolekcja robocza pszenicy jarej odznaczała się dużą zmiennością fenotypową obiektów dla rozpatrywanych cech rolniczych - zastosowanie analizy skupień i analizy składowych głównych pozwalają na wyróżnienie grup obiektów interesujących pod względem cech plonotwórczych i odpornościowych.

Cennymi obiektami, które z mogą przyczynić się do postępu hodowlanego pszenicy jarej są Alkora, Erwin, HEC 425/93, HEC 493, HEC 587/87, KRJ 4/95, KRJ 9/95, M 203/78, M 70/78, Remus. Cechy takie jak plon ziarna, masa ziarna z kłosa, liczba opadania i stopień porażenia przez choroby liścia posiadały największy udział w wielocехowej zmienności zgromadzonych w kolekcji roboczej pszenicy jarej obiektów.

LITERATURA

- Assefa K., Tefera H., Merker A., Kefyalew T., Hundera F. 2001. Quantitative trait diversity in tef (*Eragrostis tef* (Zucc.) Trotter) germplasm from Central and Northern Ethiopia. *Genetic Resources and Crop Evolution* 48: 53 — 61.
- Bennett S. J. 2000. Genetic variation of five species of *Trifolium* L. from south-west Turkey. *Genetic Resources and Crop Evolution* 47: 81 — 91.
- Brown-Guedira G. L., Thompson J. A., Nelson R. L., Warburton M. L. 2000. Evaluation of genetic diversity of soybean introductions and North American Ancestors using RAPD and SSR markers. *Crop Sci.* 40: 815 — 823.
- Casler M. D., van Santen E. 2000. Patterns of variation in a collection of meadow fescue accessions. *Crop Sci.* 40: 248 — 255.
- Crossa J., Franco J. 2004. Statistical methods for classifying genotypes. *Euphytica* 153: 19 — 37.
- Damania A. D. 2008. History, achievements, and current status of genetic resources conservation. *Agron. J.* 100: 9 — 21.
- DeLacy J.H., Skovmand B., Huerta J. 2000. Characterization of Mexican wheat landraces using agronomically useful attributes. *Genetic Resources and Crop Evolution* 47:591 — 602.
- Hailu F., Merker A., Harjit-Singh., Belay G., Johansson E. (2006) Multivariate analysis of diversity of tetraploid wheat germplasm from Ethiopia. *Genetic Resources and Crop Evolution* 53: 1089 — 1098
- Hartung K., Piepho H. P. 2005. A threshold model for multiyear genebank data based on different rating scales. *Crop Sci.* 45: 1045 — 1051.
- Hartung K., Piepho H. P., Knupffer H. 2006. Analysis of genebank evaluation data by using geostatistical methods. *Genetic Resources and Crop Evolution* 53: 737 — 751.
- Jahufer M. Z. Z., Cooper M., Harch B. D. 1997. Pattern analysis of the diversity of morphological plant attributes and herbage yield in a world collection of white clover (*Trifolium repens* L.) germplasm characterised in a summer moisture stress environment of Australia. *Genetic resources and crop evolution* 44: 289 — 300.
- Johnson R. A., Wichern D. W. 2002. Applied multivariate statistical analysis. Prentice – Hall, Inc. Upper Saddle River, NJ.

- Kubicka H., Mądry W., Sieczko L., Komar L., Puchalski J. 2004. Wielowymiarowa analiza różnorodności genotypowej linii wsobnych żyta ozimego (*Secale cereale* L.) dla cech rolniczych i fenologicznych. Zesz. Probl. Post. Nauk Rol. Seria A 497: 375 — 389.
- Littell R. C., Milligan G.A., Stroup W. W., Wolfinger R.D., Schabenberger O. 2006. SAS for mixed models. SAS Institute Inc. Cary, NC.
- Manjunatha T., Bisht I. S., Bhat K. V., Singh B. P. 2007. Genetic diversity in barley (*Hordeum vulgare* L. ssp. *vulgare*) landraces from Uttaranchal Himalaya of India. Genetic resources and crop evolution 54: 55 — 65.
- Mapes C., Caballero J., Espitia E., Bye R.A. (1996) Morphophysiological variation in some Mexican species of vegetable *Amaranthus*: evolutionary tendencies under domestication. Genetic resources and crop evolution 43: 283 — 290.
- Martinez-Calvo J., Gisbert A. D., Carmen Alamar M., Hernandorena R., Romeo C. 2007. Study of a germplasm collection of loquat (*Eriobotrya japonica* Lindl.) by multivariate analysis. Genetic Resources and Crop Evolution 55: 695 — 703.
- Mazurkiewicz B., Struś M. (1997) Kolekcje robocze pszenicy ozimej jako czynnik postępu hodowlanego. Biuletyn IHAR 204: 81 — 87.
- Mądry W. 2007. Metody statystyczne do oceny różnorodności fenotypowej dla cech ilościowych w kolekcjach roślinnych zasobów genowych. Zesz. Probl. Post. Nauk Rol. 517:21 — 41.
- Mohammadi S. A., Prasanna B. M. 2003. Analysis of genetic diversity in crop plants — salient statistical tools and considerations. Crop Sci. 43: 1235 — 1248.
- Okpul T., Singh D., Gunua T., Wagih M. E. 2004. Assessment of diversity using agro-morphological traits for selecting a core sample of Papua New Guinea taro (*Colocasia esculenta* (L.) Schott) collection. Genetic Resources and Crop Evolution 51: 671 — 678.
- Piepho H. P., Mohring J. 2005. Best Linear Unbiased Prediction of cultivar effects for subdivided target regions. Crop Sci. 45:1151 — 1159.
- Robinson G. K. 1991. That BLUP is a good thing: the estimation of random effects Statistical Science 6:15 — 51.
- Ruiz M., Ravela F., Carrillo J. M. 1997. Analysis of the discriminating power of agro-morphological and biochemical descriptors in a sample of the Spanish collection of barley (*Hordeum vulgare* L.). Genetic Resources and Crop Evolution 44: 247 — 255.
- SAS Institute Inc. (2004) SAS OnlineDoc® 9.1.3. Cary, NC.
- Sorrells M. E. 2007. Application of new knowledge, technologies, and strategies to wheat improvement. Euphytica 157: 299 — 306.
- Upadhyaya H. D., Ortiz R., Bramel P. J., Singh S. 2003. Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. Genetic Resources and Crop Evolution 50:139 — 148.
- Žáková M., Benková M. 2006. Characterization of spring barley accessions based on multivariate analysis. Commun. Biometry Crop Sci. 1 (2): 124 — 134.