Article

Statistical Multivariate Methods for the Selection of High-Yielding Rapeseed Lines with Varied Seed Coat Color

Agnieszka Łopatyńska [1] , Joanna Wolko [1,*] , Jan Bocianowski [2] , Adrian Cyplik [2] and Katarzyna Gacek [1,†]

1 Plant Breeding and Acclimatization Institute—National Research Institute, Poznań Division, Department of Oilseed Crops, Strzeszyńska 36, 60-479 Poznań, Poland
2 Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland
* Correspondence: j.wolko@ihar.edu.pl; Tel.: +48-61-846-42-02
† Current address: IHAR Group, Plant Breeding Smolice Ltd., Smolice 146, 63-740 Kobylin, Poland.

Abstract: A crucial aim for rapeseed breeders is increasing the seed yield through improving yield- and silique-related traits. In this research, traits including the number of seeds per silique, silique length, silique width, and thousand-seed weight combined with seed coat color were investigated. Yellow-seeded lines of oilseed rape have a thinner seed coat, lower fiber content, and higher protein and oil content than traditional varieties. Hybrids of yellow- and black-seeded *Brassica napus*, which combine the advantages of both types of seeds, are tested for their suitability in breeding. Therefore, the aim of this study was to assess a population of 78 doubled haploid lines with different seed coat colors to select genotypes connecting a light color of seeds with good yielding characteristics. Multivariate methods such as analysis of variance, correlation analysis, regression analysis, canonical variate analysis, Mahalanobis distances, and estimation of gene effects were used. The results showed that all tested genotypes were significantly different regarding all five studied traits. The multidimensional analyses gave an accurate overview of the relationship between the examined features. The color and weight of seeds were considered the most important for breeding rapeseed with better seed properties. Although a negative correlation between these two characteristics occurs quite commonly, it was not present in our plant material, which allowed us to select lines with a-light seed color and high thousand-seed weight, which is a significant achievement. As a result, two doubled haploid lines were chosen as valuable breeding material for creating oilseed rape varieties with improved biochemical characteristics of seeds.

Keywords: *Brassica napus*; multivariate methods; thousand-seed weight; yellow-seeded rapeseed

1. Introduction

Improving the seed yield in oilseed rape (*Brassica napus* L.) is a key task for breeders. The yield is one of the most important and complex traits in crop species determined by direct and indirect yield-related traits. The main features that define the seed yield in *Brassica napus* include the number of siliques per plant (NSPP), number of seeds per silique (NSPS), and thousand-seed weight (TSW) [1]. The number of siliques per plant and seeds per silique have the greatest impact on the amount of seeds produced by the plant; therefore, improving the NSPS without decreasing the seed weight is an important breeding goal in many crop species [2]. It may be complicated due to the negative correlation that has been observed between NSPS and TSW [3].

Various studies show that the seed yield is also determined indirectly by silique-related traits such as the silique length (SL) and silique width (SW). The silique length is an important yield trait and positively correlates with the number of seeds per silique and seed weight [4]. Previous studies have proved that the enhancement of silique length could effectively improve the seed yield [5–7]. Since the number of siliques per plant is very susceptible to the genotype and environment [8], improving the seed yield through

coordination between NSPS, SL, SW, and TSW is a crucial aim in rapeseed breeding. For this reason, these are the features that we focused on in this study.

Yellow- and light-color-seeded lines of oilseed rape are characterized by a thinner seed coat, lower fiber content, and higher protein and oil content than traditional rapeseed [9]. Due to the properties of yellow seeds, they could be better exploited, as the by-product of obtaining oil from them is a protein-rich rapeseed meal with good digestibility. This makes it different from the extracted meal obtained from black-seeded lines, which due to the high fiber content cannot be used, for example, in poultry farming [10]. Despite many positive features, yellow-seeded lines are not commonly cultivated, because they have many disadvantages, such as seeds sprouting in the siliques and self-pressing of oil from the seeds during storage [11]. In addition, the yield of yellow-seeded lines is usually lower than that of black-seeded [12]. Furthermore, Piotrowska et al. [13] observed in their research a reduced seed yield of yellow lines, but the oil content in their seeds was higher. However, the overall yield of oil from the yellow lines studied by them was lower than that of the reference cultivar—precisely because of the lower seed yield.

Summarizing all aspects, valuable hybrids of yellow- and black-seeded *Brassica napus*, which combine the advantages of both types of seeds, such as high yield and high oil and protein content, are searched and tested for their suitability in breeding. However, this requires a non-standard approach, such as using more complicated mathematical methods. Therefore, the aim of this study was to assess the population of DH lines with different seed coat color in order to select lines combining these components. For this purpose, multidimensional analyses were used, which allowed for the assessment of the tested lines in terms of many features simultaneously. Multivariate statistical methods, such as canonical variate analysis or Mahalanobis D^2 statistics, simplify the dimensionality of multi-feature object comparisons [14,15]. Using these methods allowed us to choose from the studied lines those combining the most desirable traits, which is a valuable achievement with promising results.

2. Materials and Methods

2.1. Plant Material

The plant material used in this study consisted of 78 doubled haploid (DH) lines of winter oilseed rape and their parental forms. The DH population (M305 × Z114) was developed from F_1 plants of a cross between M305 (black-seeded) and Z114 (yellow-seeded) *Brassica napus* DH lines selected for contrasting seed coat color and seed fiber content.

The DH lines were produced in the Laboratory of Plant Tissue Culture, Plant Breeding and Acclimatization Institute—National Research Institute in Poznań, Poland using the isolated microspore culture method according to the procedure described by Cegielska-Taras et al. [16]. Microspores were isolated from unopened flower buds of donor plants (M305 × Z114 hybrids), sterilized, and macerated in NLN medium. The microspore pellet was plated on Petri dishes for 10 days in the dark at 30 °C. The dishes with microspore-derived embryos (MDEs) visible to the naked eye were then placed on an illuminated rotary shaker at 25 °C. Twenty-one-day-old MDEs were plated on Petri dishes with B5 medium, and then conversion of the embryos to plantlets took place. The dishes were then placed for 14 days at 1 °C under 8 h light. After this time, the dishes were transferred to the conditions of the culture room: 24 °C temperature and 16 h lighting. Over the next three weeks, the development of normal shoots from the apical meristem of the embryo was observed, as well as the development of roots. The young plants were transferred to the soil.

2.2. Field Trial

The M305 × Z114 mapping population and parental lines were cultivated in the Plant Breeding and Acclimatization Institute—National Research Institute experimental field in Poznań, Poland (52°45′ N, 16°90′ E) during the growing season of 2017/2018. The field trial was carried out in a randomized block design with three replicates of double 2 m

long rows with 30 cm spacing. In each row 100 seeds were sown. The field was managed with standard methods. Accordingly, fertilizer was applied in the fall: K 90 kg/ha, P 60 kg/ha, N 18 kg/ha, and S 18 kg/ha. In the spring, N 26 kg/ha and S 13 kg/ha were used. Appropriate plant protection products were also applied in the fall and spring.

The silique length, silique width, and number of seeds per silique were estimated on 20 siliques from each replicate. The siliques were collected at the stage of mature seeds from the main branch of five random plants and then dried. The thousand-seed weight was estimated as the average of three measurements from the mixed seeds in each replication. The seed coat color was determined on the same seed samples with the Hunter Labs spectrocolorimeter (Colorflex) and classified using a 0 (black) to 5 (yellow) scale [17]. The results of detailed analyses of the content of oil, protein, fiber (ADF and NDF), and glucosinolates in the seeds of the studied population in two growing seasons (2015/2016 and 2017/2018), investigated with a near-infrared reflectance spectroscopy (NIRS), have been previously published in Wolko et al.'s [18] study. The studied DH population was also sequenced and the obtained SNPs markers were used for QTL genetic mapping to search for loci associated with the biochemical components of seeds and seed coat color. Several candidate genes connected to seed fiber and color were identified as a result, which was described in a publication by Gacek et al. [19].

2.3. Statistical Analysis

The normality of the distributions for the studied traits (silique length, silique width, number of seeds per silique, seed coat color, and thousand-seed weight) was tested using Shapiro–Wilk's normality test [20]. Multivariate analysis of variance (MANOVA) was performed based on the following model using a MANOVA procedure: $Y = XT + E$, where Y is an ($n \times p$)-dimensional matrix of observations, where n is the total number of observations and p is the number of traits; X is an ($n \times k$)-dimensional matrix of design, where k is the number of genotypes; T is a ($k \times p$)-dimensional matrix of unknown effects; and E is an ($n \times p$)-dimensional matrix of residuals. Next, one-way analyses of variance (ANOVA) were carried out to determine the effects of the DH lines on the variability of the examined traits, for each trait independently. The arithmetical means and standard deviations of the traits were calculated. Moreover, Fisher's least significant differences (LSDs) and coefficients of variation (CVs) were also estimated at the significance level $\alpha = 0.05$. Homogeneous groups for the analyzed traits were determined based on the least significant differences. The relationships among the observed traits were assessed based on Pearson's correlation. The results were also analyzed using multivariate methods. The effect of the silique length, silique width, number of seeds per silique, and seed coat color on thousand-seed weight was checked using multivariate regression analysis. The canonical variate analysis was applied to present a multitrait assessment of similarity for the tested DH lines and their parental forms in a lower number of dimensions with the least possible loss of information [21]. It allows illustrating the variation in genotypes in terms of all the observed traits in a graphic form. The Mahalanobis distance was suggested as a measure of "polytrait" genotype similarity [22], whose significance was verified by means of the critical value D_α, called "the least significant distance" [23]. The differences between the analyzed genotypes were verified with a cluster analysis using the nearest-neighbor method and Euclidean distances [24].

Estimation and testing of the additive gene effects and additive-by-additive interaction of homozygous loci (epistasis) effects as well as additive gene effects were performed based on the methods presented by Bocianowski and Krajewski [25], Bocianowski [26], and Bocianowski et al. [27].

All the analyses were conducted using the GenStat (18th edition) statistical software package.

3. Results

The results of the MANOVA indicated that all studied genotypes were significantly different regarding all five quantitative traits (Wilk's λ = 0.01067; $F_{79;160}$ = 2.95; $p < 0.0001$). The results of analysis of variance for these traits [silique length ($F_{79;160}$ = 2.03), silique width ($F_{79;160}$ = 1.44), number of seeds per silique ($F_{79;160}$ = 1.96), seed coat color ($F_{79;160}$ = 12.86), and thousand-seed weight ($F_{79;160}$ = 1.68)] showed variability of the tested genotypes at significance level α = 0.05.

The mean values and standard deviations for the observed traits indicated high variability among the 80 tested genotypes, for which significant differences were found in terms of all the analyzed quantitative traits (Table 1), and on this basis homogeneous groups were distinguished (Table S1). The silique length varied from 41.98 mm (for DH 07) to 69.63 mm (for DH 69) with an average of 53.66 mm and allowed for the distinction of 17 homogeneous groups. The values of silique width ranged from 3.12 mm (for DH 05) to 4.26 mm (for DH 30), with an average of 3.65 mm, and divided the studied genotypes into 11 groups. The number of seeds per silique varied from 8.5 (for DH 07) to 18.88 (for DH 01), with an average of 15.44, and allowed the distinction of 20 homogeneous groups. The values of seed coat color ranged from 0.1 (for DH 26) to 4.75 (for parental form Z114), with an average of 1.52. This trait distinguished the greatest number of groups, which was 32. The thousand-seed weight varied from 3.72 g (for DH 77) to 6.57 g (for DH 08), with an average of 5.01 g, and divided the studied DH lines and parental forms into 18 homogeneous groups.

Table 1. Basic values and parameters for five quantitative traits in all tested genotypes of *Brassica napus*.

	Silique Length [mm]	Silique Width [mm]	Number of Seeds per Silique	Seed Coat Color [0–5]	Thousand-Seed Weight [g]
Mean	53.66	3.65	15.44	1.52	5.01
Min.	41.98	3.12	8.50	0.10	3.72
Max.	69.63	4.26	18.88	4.75	6.57
s.d.	3.88	0.29	2.15	0.42	0.64
LSD$_{0.05}$	7.72	0.53	3.92	0.84	1.19
CV [%]	7.24	7.88	13.90	27.52	12.82

Table 1 also contains the coefficients of variation (CVs). The lowest and similar variability was observed for silique length and silique width (7.24% and 7.88%, respectively). Moreover, a similar but slightly higher variability occurred in the number of seeds per silique and thousand-seed weight (13.90% and 12.82%, respectively). The highest coefficient of variation occurred for the seed coat color (27.52%), in which the largest number of homogeneous groups was distinguished.

The correlation coefficients between all pairs of observed traits are presented in the form of a heatmap in Figure 1. This analysis indicated a statistically significant positive correlation between the silique length and number of seeds per silique ($r = 0.582$), silique length and thousand-seed weight ($r = 0.296$), as well as silique width and thousand-seed weight ($r = 0.412$). At the same time, a statistically significant negative correlation was observed between the silique length and seed coat color ($r = -0.289$).

The effect of the silique length, silique width, number of seeds per silique and seed coat color on the thousand-seed weight was evaluated using a multivariate regression analysis (Table 2). The thousand-seed weight was significantly positively determined by the silique length and silique width, but negatively by the number of seeds per silique. These three quantitative traits accounted for 28.90% of the thousand-seed weight variability. The effect of the seed coat color on the thousand-seed weight was not significant.
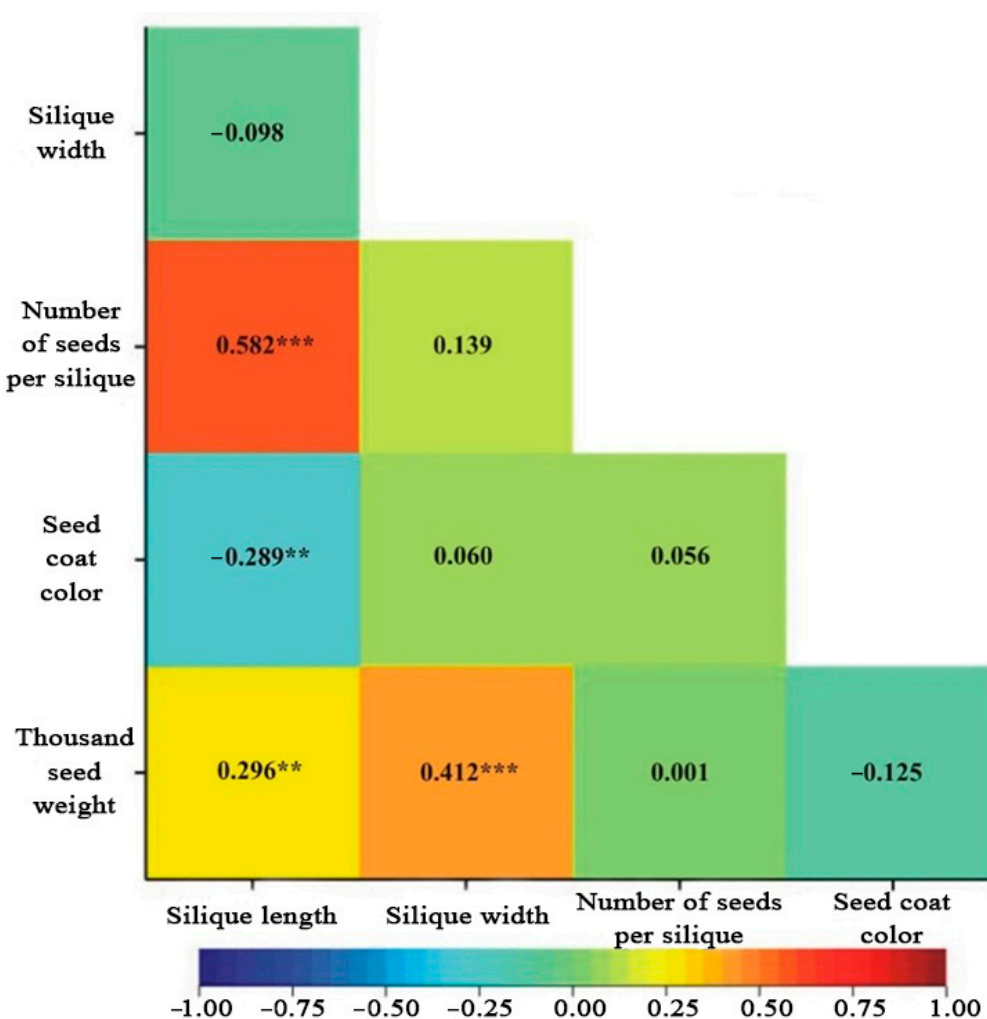
Figure 1. Heatmap for linear Pearson's correlation coefficients between observed traits of *Brassica napus*. ** $p < 0.01$; *** $p < 0.001$.

Table 2. Traits significantly affecting the thousand-seed weight.

Parameter	Estimate	Standard Error	*t*-Statistic
Constant	−0.229	1.043	−0.22 ns
Silique length [mm]	0.058	0.017	3.381 **
Silique width [mm]	0.875	0.208	4.205 ***
Number of seeds per silique	−0.066	0.026	−2.572 *
Seed coat color [0–5]	−0.010	0.050	−0.202 ns
Percentage variance accounted		28.90%	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns—not significant.

Individual traits were of varying importance and had different shares in the joint multivariate variation of the studied genotypes. The analysis of the first two canonical variates for the DH population and parental lines regarding the five quantitative traits is shown in Figure 2. In the graph, the coordinates of the point for a particular genotype were the values for the first and second canonical variate, respectively. The studied population of the DH lines (apart from a few) was quite evenly distributed around the center of the coordinate system. The parental lines were located at opposite ends of the V_1 axis. The first two canonical variates accounted for 78.16% of the total variability among the individual genotypes. The most significant positive, linear relationship with the first canonical variate was found for the silique length, while the negative, for the seed coat color (Table 3). The

second canonical variate was significantly positively correlated with the silique length and thousand-seed weight, but negatively with the silique width. The silique length was the only trait positively correlated with both variates. The very strong (close to 1) correlation between the seed coat color and the first canonical variate means that the studied genotypes are distributed along this axis regarding this trait.
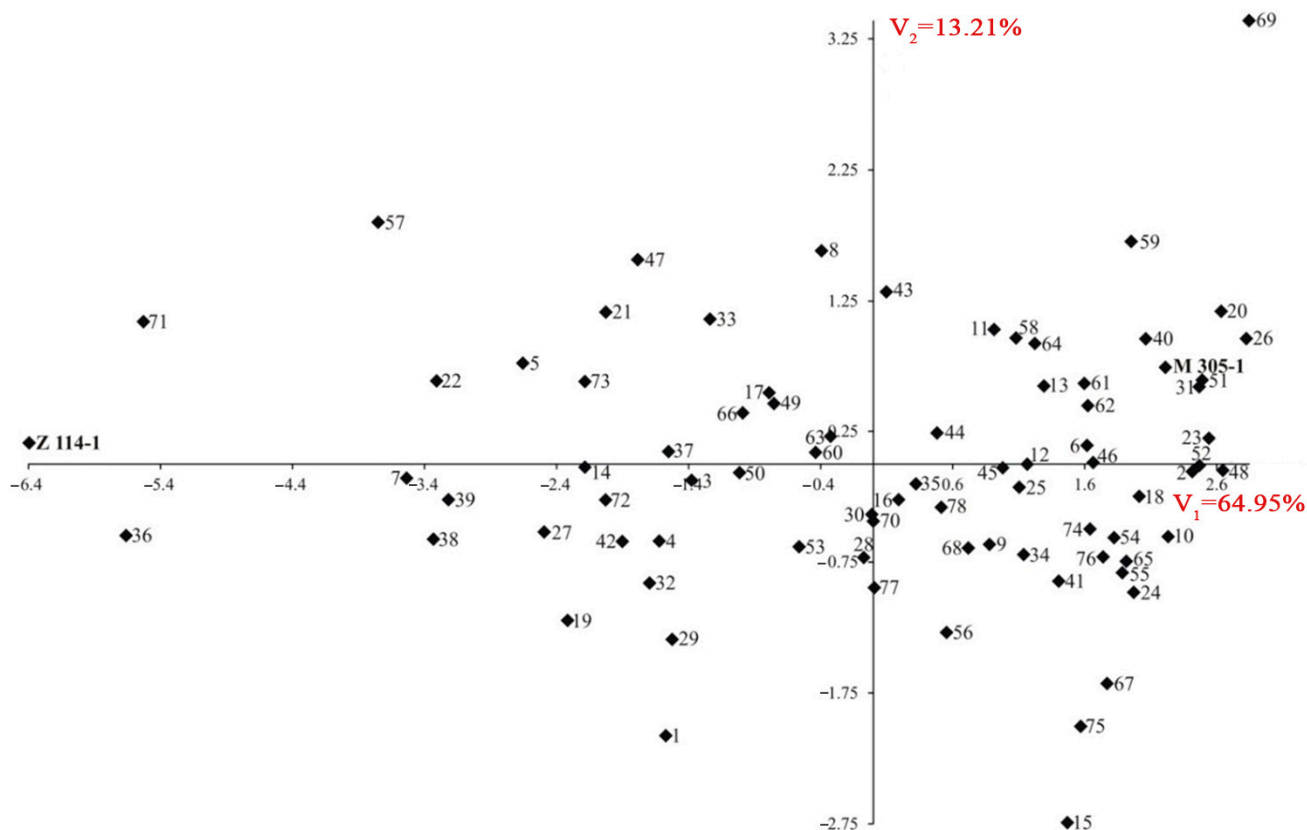

Figure 2. Distribution of the studied *Brassica napus* DH lines and their parental forms in the space of the first two canonical variates.

Table 3. Correlation coefficients between the first two canonical variates and studied traits.

Trait	First Canonical Variate	Second Canonical Variate
Silique length [mm]	0.289 **	0.64 ***
Silique width [mm]	−0.074	−0.374 ***
Number of seeds per silique	−0.129	−0.101
Seed coat color [0–5]	−0.995 ***	0.058
Thousand-seed weight [g]	0.12	0.378 ***
Percentage of explained multivariate variability	64.95%	13.21%

** $p < 0.01$; *** $p < 0.001$.

The greatest variation in terms of all five traits jointly measured with Mahalanobis distances was found for DH 69 and Z114 (the distance between them amounted to 9.926). The greatest similarity was found between DH 11 and DH 64 (0.502) (data not shown).

In the dendrogram presented in Figure 3, all the examined *B. napus* genotypes were divided into four groups as a result of agglomeration grouping using the Euclidean distance method. The first group (I) comprised two lines, DH 07 and DH 42, while the second one (II) consisted of five genotypes (DH 05, DH 36, DH 47, DH 71, and the yellow-seeded

parental line Z114). The third group (III) contained three lines, DH 08, DH 30, and DH 72, while the fourth group (IV) contained the rest of the genotypes and was divided into four subgroups: IVA consisted of 11 DH lines (DH 12, DH 24, DH 26, DH 31, DH 43, DH 46, DH 51, DH 61, DH 74, DH 75, DH 78) and the black–seeded parental line M305; IVB, four DH lines (DH 13, DH 20, DH 59, DH 69); IVC, 32 DH lines; and IVD comprised 22 DH lines.
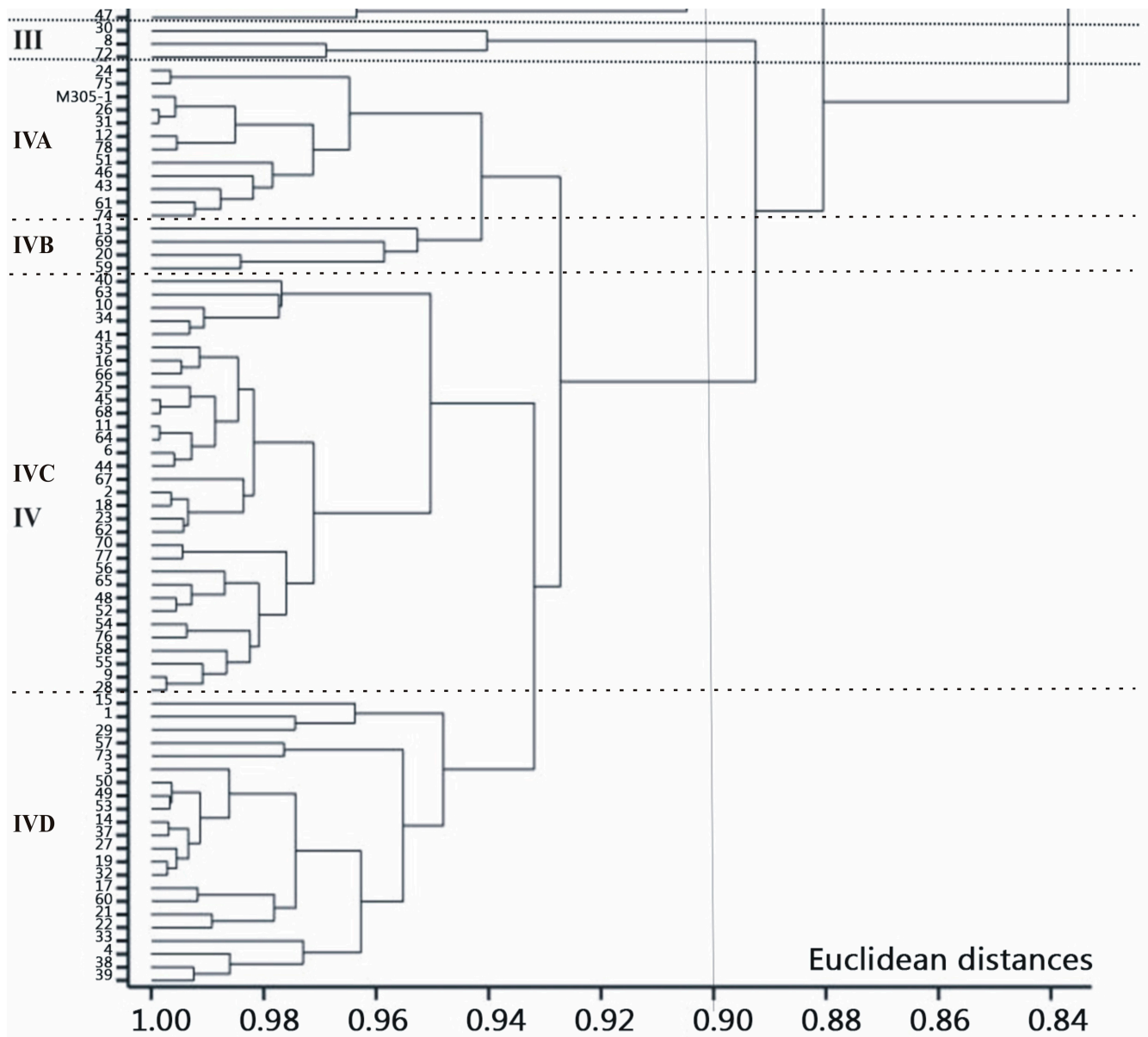
Figure 3. Dendrogram of cluster groupings of *Brassica napus* DH lines and their parental forms based on all five quantitative traits.

The additive gene action effect estimated based on the parental forms was significant only for seed coat color and based on DH lines it was significant for all five traits (Table 4). This effect for four observed features based on DH lines was larger than the parameter estimated based on the parental lines, and it was only on a similar level for seed coat color. The estimates of epistasis effects for all traits were statistically significant, but only positive for seed coat color.

Table 4. Estimates of additive and epistasis effects for observed traits of *Brassica napus*.

Effect	Silique Length [mm]	Silique Width [mm]	Number of Seeds per Silique	Seed Coat Color [0–5]	Thousand-Seed Weight [g]
			parental forms		
$a_{Parents}$ [1]	4.24	0.005	0.235	2.174 ***	0.337
			doubled haploid lines		
a_{DH} [2]	13.83 ***	0.570 *	5.190 ***	2.13 ***	1.423 **
aa_{DH} [3]	−39.83 ***	−3.087 ***	−10.233 ***	0.634 **	−3.586 ***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. [1] the total additive effect estimated based on parental observations. [2] the total additive effect estimated based on doubled haploid lines observations. [3] the total additive-by-additive interaction effect estimated based on doubled haploid line observations.

4. Discussion

Breeding programs for winter oilseed rape cultivars are focused on improving the seed and oil yield; therefore, information about the genetic variability of traits is required [28]. The aim of this study was to estimate the variability of the silique- and seed-related traits in relation to the seed coat color using multivariate statistical methods. To achieve this goal, a population of DH lines with different seed color (from yellow to black) was used as a plant material. We decided to evaluate these lines because yellow seeds have a better biochemical composition than black seeds (mainly a lower fiber content). The mean content of neutral detergent fiber (NDF) in the yellow-seeded Z114 line was 17%, while in black-seeded M305 line it was 23%. The mean content of acid detergent fiber (ADF) was about 10% in the Z114 line and 17% in line M305 [18]. At the same time, yellow-seeded lines may be characterized by a lower yield [13] and a lower thousand-seed weight, which can be observed in the yellow-seeded parent of our DH population (3.97–4.81 g according to the year of research), while the thousand-seed weight of the M305 line ranged from 4.97 to 5.52 g [18]. Therefore, multivariate analysis was used to find lines with a light seed color, high thousand-seed weight, and good values of silique-related traits at the same time. In order to check whether the studied genotypes have sufficient variability for such statistical methods, an analysis of variance was initially performed. The results of the ANOVA showed variability of all the tested genotypes at a significance level $\alpha = 0.05$, and MANOVA indicated that all studied genotypes were significantly different regarding all five quantitative traits jointly. Thus, it confirmed that the tested population is suitable for carrying out the assumed multidimensional analyses.

The studied population of DH lines had very diverse seed coat colors, having a range from 0.1 to 4.75 on a scale from 0 to 5 (from black to light-yellow). Bartkowiak-Broda et al. [29], examining hybrids between yellow- and black-seeded lines, observed a seed coat color range from 1.7 to 5.0, which means they were slightly lighter than our material. The thousand-seed weight of the yellow lines they evaluated (between 3.5 and 5.2 g) was lower than in this study (3.72–6.57 g). The mean color value for our 78 DH lines was 1.52, while other authors noted it on the level of 3.5 for 58 yellow-seeded and seven black-seeded lines jointly [30]. The average thousand-seed weight observed by these authors (4.1 g) was much lower than in our genotypes (5.01 g). Therefore, compared to other studies, the population we observed had a wider range of seed coat color, but also heavier seeds.

The seed yield is a very complex feature that consists of direct and indirect yield-related traits. The number of siliques per plant, number of seeds per silique, and thousand-seed weight have the greatest impact on the size of the yield [7,31]. The last two features are related to indirect yield traits such as the length and width of the siliques. The silique length is positively correlated with the number of seeds per silique (the longer the siliques, the more seeds) and the silique width, with the thousand-seed weight (the wider the siliques, the bigger the seeds) [6,32,33]. We focused on these four features in our research. The values of silique length in our DH population (41.98–69.63 mm) had a narrower range than that observed by Wang et al. [7] in their DH population in four environments (19–118 mm). The

same situation occurred in reference to the silique width (3.12–4.26 mm in our study and 1.3–6.7 mm in Wang's study) and for the number of seeds per silique (8.50–18.88 and 0.6–39, respectively). These results show that our plant material, despite its origin from parents with different seed color, was more homogeneous in terms of these characteristics. A wider range of trait values tested in Wang's research is probably due to the fact that he conducted field trials in four environments, and we only in one. The features related to the structure of the crop are sensitive to environmental conditions. A relationship similar to that in our results for the three examined traits was also observed by Cai et al. [34] (38.4–93.9 mm for silique length, 10.35–27.66 for number of seeds per silique, and 2.20–5.49 g for thousand-seed weight). By analyzing these traits, it can be seen that our DH lines have shorter siliques with fewer seeds but a greater thousand-seed weight (average values: SL 53.66 mm, NSPS 15.44, TSW 5.01 g) than that observed by Cai et al. [34] (average values: SL 38.4 mm, NSPS 10.35, TSW 2.20 g) and Fu et al. [35] (average values: SL 56.7 mm, TSW 3.47 g).

The above-mentioned relationships are reflected in the calculated values of the standard deviation—they are lower in our study than in Cai et al.'s [34] for features such as silique length (3.88 and 7.8, respectively) or the number of seeds per silique (2.15 and 3.26, respectively) and higher for TSW (0.64 and 0.59, respectively). Analyzing the coefficients of variation, we can distinguish two features with significantly lower variability (silique length and silique width), two with a medium level (number of seeds per silique and thousand-seed weight), and one feature with a significantly higher variability than the others (seed coat color). Bartkowiak-Broda et al. [29], examining the variability of yellow- and black-seeded hybrids, noted a slightly lower coefficient of variation for the seed coat color (20.9%) and thousand-seed weight (8.2%) than in our study. All studied genotypes formed various numbers of homogeneous groups in terms of the tested traits. The largest number of groups was distinguished for the seed coat color, which indicates that for this trait, the objects were the most diverse. This is proved by the value of the coefficient of variation, which is the highest for this feature (27.52%). The high variability of the seed color, confirmed in this population, enables the selection of lines with the desired biochemical characteristics of the seeds and good yield-related traits.

In our research there was a statistically significant positive correlation between the silique length and number of seeds per silique, which is a commonly observed relationship [6,32,36]. We also noted a significant and positive correlation between the silique length and thousand-seed weight, similar to Fu et al. [35]. It indicates that longer siliques have heavier seeds. The opposite results, with a negative correlation for SL and TSW, were obtained by Zhang et al. [6]. Additionally, a statistically significant negative correlation was present in our study between the silique length and seed coat color, which suggests that the genotypes with darker seeds had longer siliques. The same result was achieved by Myszka et al. [30]. Another positive correlation was observed between the silique width and thousand-seed weight, indicating that wider siliques contain larger seeds. There was no statistically significant correlation between the seed coat color and thousand-seed weight. This means that genotypes in our population have varied seed sizes regardless of color, which allows the selection of genotypes connecting the desirable features of yellow seed color and high thousand-seed weight.

A regression analysis explains how the value of the dependent variable changes when any independent variable is modified. In most oilseed rape studies, regression analysis is evaluated for the seed yield per plot or plant. For example, Elliott et al. [37] and Sharafi et al. [38] observed a significant positive effect of the thousand-seed weight on the seed yield, which means that an increase in the TSW causes an increase in the yield. In our research, a linear regression was performed for the thousand-seed weight, because we wanted to check whether the color of the seeds affects their weight. As in our earlier studies [39], carried out on other rapeseed populations, the regression analysis showed that the thousand-seed weight was significantly positively affected by the silique length and negatively by the number of seeds per silique. An additional feature positively influencing the TSW in the current research was the silique width. There was no significant effect of

the seed coat color on the thousand-seed weight. Due to this, it is possible to select lines with large seeds and a light seed coat color at the same time.

Using canonical variate analysis, the tested DH lines and their parental lines were characterized in terms of five quantitative traits. The first two canonical variates together explained 78.16% of the total variability between the individual genotypes, which means that there was a small loss of information. A similar value (72.55%) was observed by Bocianowski et al. [40], who analyzed rapeseed hybrids and their parental lines in terms of 19 traits. Both canonical variates were significantly positively correlated with the silique length and only the second with thousand-seed weight. The seed color had a significant negative impact on the first variate, and the silique width on the second one. The conducted canonical variate analysis allowed for the assessment of the distribution of objects in terms of their features in the space of the first two canonical variates. As the first variate explained a significant percentage of the observed variability (64.95%), as well as the correlation value between it and the seed color being very high (close to -1), the studied genotypes are arranged along the V_1 axis in terms of this trait with high accuracy. Most of the studied DH lines were quite evenly distributed around the center of the coordinate system, while the parental lines were placed at the opposite ends of the V_1 axis, which reflects the difference in the color of their seeds. Genotypes characterized by a light seed color and low thousand-seed weight are placed in the bottom left quarter of the coordinate system, and genotypes with light seeds but a higher TSW in the upper left part. In the upper right part of the graph there are genotypes characterized by darker and heavier seeds, and in the last quarter (bottom right), lines with dark seeds and a low thousand-seed weight are located. This arrangement of lines in the coordinate system facilitates the selection of genotypes with the desired characteristics. Two lines in the upper left quarter attract attention—DH 71 (color 4.4, TSW 5.5 g) and DH 57 (color 3.5, TSW 5.6 g). They have a similar color level, but a higher thousand-seed weight than the yellow-seeded parental line Z114 (color 4.75, TSW 4.85 g). The position of these two lines in the upper left quarter also indicates their potential for long siliques, as this feature is highly significantly correlated with the V_2 variate (0.64).

The Mahalanobis distances between the examined objects allow us to assess the phenotypic similarity in the context of all examined traits. The Mahalanobis distances calculated in our study for all five quantitative traits ranged from 0.502 to 9.926. Very similar values were noticed by Parvin et al. [41] and Mili et al. [42] in *Brassica napus* genotypes— from 0.304 to 8.145 and from 0.378 to 12.433, respectively. An additional method in diversity studies is analysis of the specific clusters created on the dendrogram of similarity. The non-hierarchical Euclidean method of clustering is precise and critically identifies subclusters of the main groups at different levels, thus supporting the selection of diverse breeding materials [43]. The genotypes we studied were divided into four main clusters. From the fourth group, four subgroups were distinguished. As expected, the parental lines were placed in two separate groups. Most of the studied lines (69) were in group IV together with the black-seeded parent. In the second group of five genotypes, along with the yellow-seeded parent, there was line DH 71, which stands out from our plant material with a light color of seeds and a high thousand-seed weight.

The estimation of genetic parameters has a significant role in oilseed rape breeding. Our results indicate the importance of both additive and epistasis gene effects for all studied traits on the basis of doubled haploid lines observations. Additionally, the additive effect estimated based on parental observations was statistically significant for seed coat color. The importance of both additive and epistasis gene effects for the silique length and number of seeds per silique was also noted by Bocianowski et al. [44]. Luo et al. [45] observed that non-additive effects had a great influence on heritability and epistasis and noted the importance of environmental interactions. The significant additive and epistasis gene effects found in this study for all traits mean that they are regulated in a complex manner involving many genes with small effects and also by gene-by-gene interactions.

5. Conclusions

This study proves that multivariate statistical methods are useful for selecting breeding material in oilseed rape populations. In our opinion, out of the five traits we studied, the seed color and weight of seeds are the most important from the point of view of rapeseed breeding with better seed properties, combining the advantages of black- and yellow-seeded genotypes. Contrary to some published studies, in our population there was no negative correlation between the color of seeds and their weight, which allows us to select lines with desirable traits. The multidimensional analysis revealed an exact overview of the relationship between the examined features. As a result, two doubled haploid lines (DH 57 and DH 71) were selected as a valuable breeding material that could be used to create oilseed rape varieties with improved biochemical properties of seeds. Taking into account the parental genotypes used in this study, obtaining lines with such bright, large, and well-formed seeds is a significant achievement. The low fiber content in these yellow-seeded lines would provide an extracted meal with a better quality and better digestibility for farm animals. Rapeseed meal is already used as a supplement to protein feed, but the high fiber content limits its use in poultry farming. The use of yellow-seeded lines with a reduced fiber content would ensure a more comprehensive use of rapeseed not only as an oil plant, but also as a valuable source of feed protein.

References

1. Chen, W.; Zhang, Y.; Liu, X.; Chen, B.; Tu, J.; Fu, T. Detection of QTL for six yield-related traits in oilseed rape (*Brassica napus*) using DH and immortalized F_2 populations. *Theor. Appl. Genet.* **2007**, *115*, 849–858. [CrossRef]
2. Jiang, H.L.; Hong, J.; Jiang, Y.T.; Yu, S.X.; Zhang, Y.J.; Shi, J.X.; Lin, W.H. Genome-Wide Association Analysis Identifies Candidate Genes Regulating Seed Number Per Silique in *Arabidopsis thaliana*. *Plants* **2020**, *9*, 585. [CrossRef]
3. Wolko, J.; Dobrzycka, A.; Bocianowski, J.; Bartkowiak-Broda, I. Estimation of heterosis for yield-related traits for single cross and three-way cross hybrids of oilseed rape (*Brassica napus* L.). *Euphytica* **2019**, *215*, 156. [CrossRef]
4. Wang, H.; Zaman, Q.U.; Huang, W.; Mei, D.; Liu, J.; Wang, W.; Ding, B.; Hao, M.; Fu, L.; Cheng, H.; et al. QTL and Candidate Gene Identification for Silique Length Based on High-Dense Genetic Map in *Brassica napus* L. *Front. Plant Sci.* **2019**, *10*, 1579. [CrossRef] [PubMed]
5. Udall, J.A.; Quijada, P.A.; Lambert, B.; Osborn, T.C. Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 2. Identification of alleles from unadapted germplasm. *Theor. Appl. Genet.* **2006**, *113*, 597–609. [CrossRef] [PubMed]
6. Zhang, L.W.; Yang, G.S.; Liu, P.W.; Hong, D.F.; Li, S.P.; He, Q.B. Genetic and correlation analysis of silique-traits in *Brassica napus* L. by quantitative trait locus mapping. *Theor. Appl. Genet.* **2011**, *122*, 21–31. [CrossRef]
7. Wang, X.; Chen, L.; Wang, A.; Wang, H.; Tian, J.; Zhao, X.; Chao, H.; Zhao, Y.; Zhao, W.; Xiang, J.; et al. Quantitative trait loci analysis and genome-wide comparison for silique related traits in *Brassica napus*. *BMC Plant Biol.* **2016**, *16*, 71. [CrossRef]
8. Zając, T.; Klimek-Kopyra, A.; Oleksy, A.; Lorenc-Kozik, A.; Ratajczak, K. Analysis of yield and plant traits of oilseed rape (*Brassica napus* L.) cultivated in temperate region in light of the possibilities of sowing in arid areas. *Acta Agrobot.* **2016**, *69*, 1696. [CrossRef]

9. Rahman, M.; McVetty, P.B.E. A review of *Brassica* seed color. *Can. J. Plant Sci.* **2011**, *91*, 437–446. [CrossRef]

10. Nesi, N.; Delourme, R.; Bregeon, M.; Falentin, C.; Renard, M. Genetic and molecular approaches to improve nutritional value of *Brassica napus* L. seed. *Comptes Rendus Biol.* **2008**, *331*, 763–771. [CrossRef]

11. Gawrysiak-Witulska, M.; Rudzińska, M.; Siger, A.; Bartkowiak-Broda, I. A high drying temperature causes degradation of sterols and tocopherols in yellow-seeded *Brassica napus* oils. *Eur. J. Lipid Sci. Technol.* **2015**, *117*, 483–490. [CrossRef]

12. Rakow, G.; Relf-Eckstein, J.A.; Raney, J.P. Rapeseed genetic research to improve its agronomic performance and seed quality. *Helia* **2007**, *30*, 199–206. [CrossRef]

13. Piotrowska, A.; Krótka, K.; Krzymański, J. Wartość gospodarcza żółtonasiennych linii rzepaku ozimego (*Brassica napus* L.). Economical value of yellow-seeded lines of winter oilseed rape (*Brassica napus* L.). *Rośliny Oleiste Oilseed Crops* **2000**, *21*, 359–368.

14. Daoyu, Z.; Lawes, G.S. Manova and discriminant analysis of phenotypic data as a guide for parent selection in kiwifruit (*Actinidia deliciosa*) breeding. *Euphytica* **2000**, *114*, 151–157. [CrossRef]

15. Vaylay, R.; van Santen, E. Application of canonical discriminant analysis for the assessment of genetic variation in tall fescue. *Crop Sci.* **2002**, *42*, 534–539. [CrossRef]

16. Cegielska-Taras, T.; Tykarska, T.; Szala, L.; Kuras, M.; Krzymański, J. Direct plant development from microspore-derived embryos of winter oilseed rape *Brassica napus* L. ssp *oleifera* (DC.) Metzger. *Euphytica* **2002**, *124*, 341–347. [CrossRef]

17. Michalski, K. Seed color assessment in rapeseed seeds using Color and Near Infrared Reflectance Spectrometers. *Rośliny Oleiste Oilseed Crops* **2009**, *30*, 119–132.

18. Wolko, J.; Dobrzycka, A.; Bocianowski, J.; Szala, L.; Cegielska-Taras, T.; Bartkowiak-Broda, I.; Gacek, K. Genetic variation of traits affecting meal quality in black × yellow seeded doubled haploid population of winter oilseed rape. *Agron. Res.* **2020**, *18*, 2259–2270. [CrossRef]

19. Gacek, K.; Bayer, P.E.; Anderson, R.; Severn-Ellis, A.A.; Wolko, J.; Łopatyńska, A.; Matuszczak, M.; Bocianowski, J.; Edwards, D.; Batley, J. QTL Genetic Mapping Study for Traits Affecting Meal Quality in Winter Oilseed Rape (*Brassica napus* L.). *Genes* **2021**, *12*, 1235. [CrossRef] [PubMed]

20. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]

21. Rencher, A.C. Interpretation of canonical discriminant functions, canonical variates, and principal components. *Am. Stat.* **1992**, *46*, 217–225. [CrossRef]

22. Seidler-Łożykowska, K.; Bocianowski, J. Evaluation of variability of morphological traits of selected caraway (*Carum carvi* L.) genotypes. *Ind. Crops Prod.* **2012**, *35*, 140–145. [CrossRef]

23. Mahalanobis, P.C. On the generalized distance in statistics. *Proc. Natl. Acad. Sci. India A* **1936**, *12*, 49–55. [CrossRef]

24. Skomra, U.; Bocianowski, J.; Agacka, M. Agro-morphological differentiation between European hop (*Humulus lupulus* L.) cultivars in relation to their origin. *J. Food Agric. Environ.* **2013**, *11*, 1123–1128.

25. Bocianowski, J.; Krajewski, P. Comparison of the genetic additive effect estimators based on phenotypic observations and on molecular marker data. *Euphytica* **2009**, *165*, 113–122. [CrossRef]

26. Bocianowski, J. Analytical and numerical comparisons of two methods of estimation of additive × additive interaction of QTL effects. *Sci. Agric.* **2012**, *69*, 240–246. [CrossRef]

27. Bocianowski, J.; Warzecha, T.; Nowosad, K.; Bathelt, R. Genotype by environment interaction using AMMI model and estimation of additive and epistasis gene effects for 1000-kernel weight in spring barley (*Hordeum vulgare* L.). *J. Appl. Genet.* **2019**, *60*, 127–135. [CrossRef]

28. Liersch, A.; Bocianowski, J.; Nowosad, K.; Mikołajczyk, K.; Spasibionek, S.; Wielebski, F.; Matuszczak, M.; Szała, L.; Cegielska-Taras, T.; Sosnowska, K.; et al. Effect of Genotype × Environment Interaction for Seed Traits in Winter Oilseed Rape (*Brassica napus* L.). *Agriculture* **2020**, *10*, 607. [CrossRef]

29. Bartkowiak-Broda, I.; Piotrowska, A.; Hernacki, B.; Cegielska-Taras, T.; Michalski, K. Development of germplasm of yellow seeded winter oilseed rape (*Brassica napus* L. var. *oleifera*). In Proceedings of the GCIRC Technical Meeting, Manesar, India, 28–29 September 2009.

30. Myszka, K.; Boros, D.; Piotrowska, A.; Bartkowiak-Broda, I. Porównanie składu chemicznego śrut rzepakowych uzyskanych z rzepaku ozimego (*Brassica napus* L.) o zróżnicowanej barwie nasion. Comparison of chemical composition of rapeseed meals derived from winter rapeseed (*Brassica napus* L.) differing in colour of seeds. *Rośliny Oleiste Oilseed Crops* **2011**, *32*, 257–268.

31. Zhu, Y.; Ye, J.; Zhan, J.; Zheng, X.; Zhang, J.; Shi, J.; Wang, X.; Liu, G.; Wang, H. Validation and Characterization of a Seed Number Per Silique Quantitative Trait Locus qSN.A7 in Rapeseed (*Brassica napus* L.). *Front. Plant Sci.* **2020**, *11*, 68. [CrossRef]

32. Zhao, W.; Zhang, L.; Chao, H.; Wang, H.; Ta, N.; Li, H.; Wang, X.; Li, S.; Xiang, J.; Zhang, L.; et al. Genome-wide identification of silique-related traits based on high-density genetic linkage map in *Brassica napus*. *Mol. Breeding* **2019**, *39*, 86. [CrossRef]

33. Shi, J.Q.; Li, R.Y.; Qiu, D.; Jiang, C.C.; Long, Y.; Morgan, C.; Bancroft, I.; Zhao, J.Y.; Meng, J.L. Unraveling the complex trait of crop yield with quantitative trait loci mapping in *Brassica napus*. *Genetics* **2009**, *182*, 851–861. [CrossRef] [PubMed]

34. Cai, D.; Xiao, Y.; Yang, W.; Ye, W.; Wang, B.; Younas, M.; Wu, J.; Liu, K. Association mapping of six yield-related traits in rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* **2014**, *127*, 85–96. [CrossRef]

35. Fu, Y.; Wei, D.; Dong, H.; He, Y.; Cui, Y.; Mei, J.; Wan, H.; Li, J.; Snowdon, R.; Friedt, W.; et al. Comparative quantitative trait loci for silique length and seed weight in *Brassica napus*. *Sci. Rep.* **2015**, *5*, 14407. [CrossRef] [PubMed]

36. Diepenbrock, W. Yield analysis of winter oilseed rape (*Brassica napus* L.): A review. *Field Crop Res.* **2000**, *67*, 35–49. [CrossRef]

37. Elliott, R.H.; Mann, L.W.; Olfert, O.O. Effects of seed size and seed weight on seedling establishment, seedling vigour and tolerance of summer turnip rape (*Brassica rapa*) to flea beetles, *Phyllotreta* spp. *Can. J. Plant Sci.* **2007**, *87*, 385–393. [CrossRef]

38. Sharafi, Y.; Majidi, M.M.; Jafarzadeh, M.; Mirlohi, A. Multivariate Analysis of Genetic Variation in Winter Rapeseed (*Brassica napus* L.) Cultivars. *J. Agric. Sci. Technol.* **2015**, *17*, 1319–1331.

39. Łopatyńska, A.; Bocianowski, J.; Cyplik, A.; Wolko, J. Multidimensional Analysis of Diversity in DH Lines and Hybrids of Winter Oilseed Rape (*Brassica napus* L.). *Agronomy* **2021**, *11*, 645. [CrossRef]

40. Bocianowski, J.; Liersch, A.; Bartkowiak-Broda, I. Badanie zmienności fenotypowej mieszańców F1 CMS *ogura* rzepaku ozimego i ich form rodzicielskich za pomocą statystycznych metod wielowymiarowych. Investigation of phenotypic distance of F1 CMS *ogura* winter oilseed rape hybrids and parental lines using multivariate statistical methods. *Rośliny Oleiste Oilseed Crops* **2009**, *30*, 161–184.

41. Parvin, E.; Mahmud, F.; Bhuiyan, S.R.; Haque, M. Multivariate Analysis of Genetic Variation in Rapeseed (*Brassica napus* L.). *Agric. Food Sci. Res.* **2019**, *6*, 1–8. [CrossRef]

42. Mili, S.M.S.K.; Monalesa, N.; Rahman, M.J.; Afsana, N.; Zeba, N.; Mahmud, F. Genetic diversity analysis in F5 population of *Brassica napus* L. *Int. J. Biosci.* **2016**, *9*, 177–184. [CrossRef]

43. Singh, S.; Dwivedi, A.K.; Ashutosh, K.O.; Kumar, K. Genetic Divergence Analysis in Indian mustard (*Brassica juncea* L.). *Int. J. Curr. Microbiol. App. Sci.* **2018**, *7*, 2496–2503. [CrossRef]

44. Bocianowski, J.; Nowosad, K.; Dobrzycka, A.; Wolko, J. Estimation of additive and epistatic gene effects of doubled haploid lines of winter oilseed rape (*Brassica napus* L.). *Euphytica* **2017**, *213*, 122. [CrossRef]

45. Luo, X.; Ding, Y.; Zhang, L.; Yue, Y.; Snyder, J.H.; Ma, C.; Zhu, J. Genomic Prediction of Genotypic Effects with Epistasis and Environment Interactions for Yield-Related Traits of Rapeseed (*Brassica napus* L.). *Front. Genet.* **2017**, *8*, 15. [CrossRef] [PubMed]