

# Correlation and sequential path analysis of oil yield and related characteristics in camelina under seasonal variations

Merve Göre<sup>1</sup> , Hossein Zeinalzadeh-Tabrizi<sup>2,\*</sup>  and Orhan Kurt<sup>3</sup> 

<sup>1</sup> Department of Medicinal and Aromatic Plants, Ödemis Vocational Training School, Ege University, Izmir, Turkey

<sup>2</sup> Department of Horticulture and Agronomy, Faculty of Agriculture, Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan

<sup>3</sup> Department of Field Crops, Faculty of Agriculture, Ondokuz Mayıs University, Samsun, Turkey

Received 24 May 2022 – Accepted 12 December 2022

**Abstract** – The objectives of the current study were to determine the usefulness of sequential path analysis in camelina to obtain information about the relationship between yield and yield components and to evaluate their relative importance in camelina oil yield under summer and winter cultivation. A split-plot design, with two varieties as the main plot and four sowing times as the subplot, was carried out over two growing seasons (2017–2019) in Samsun, Turkey. Sequential path analysis revealed that, as first-order predictors, grain yield and oil content displayed the most significant and positive direct effects on oil yield in both summer and winter cultivation. The sequential path analysis of second-order variables over the first-order variable revealed that seed number per pod and pod number explained approximately 90% of the variation of the grain yield in summer cultivation and branch number explained approximately 67% of the variation grain yield in winter cultivation. These results indicated that grain yield, as a main predictor of oil yield, had different pathways to affect oil yield in the summer and winter seasons. A higher magnitude of seed number per pod compared to pod number in this study indicated that selecting for higher grain yield can be done indirectly using plants with lower pod number and higher seed number per pod in the summer season. Moreover, branch number was the only trait that had a direct negative effect on grain yield in the winter season, indicating that plants with lower branch number should be selected for higher grain yield. Different environmental factors, including the seasonal cultivation of camelina in this study, were found to be a key factor in improving oil yield and, hence, should be considered as criteria indices in camelina breeding programs in the future.

**Keywords:** camelina / multicollinearity / path analysis / spring cultivation / winter cultivation

**Résumé – Corrélation et analyse séquentielle du rendement en huile et des caractéristiques connexes chez la caméline en fonction de la saison.** Les objectifs de la présente étude étaient de déterminer l'utilité de l'analyse séquentielle de chemin pour la caméline afin d'obtenir des informations sur la relation entre le rendement et les composantes du rendement et d'évaluer leur importance relative dans le rendement en huile de la caméline en culture d'été et d'hiver. Un dispositif expérimental en *split-plot*, avec deux variétés comme bloc principal et quatre périodes de semis comme sous-blocs, a été réalisé sur deux saisons de croissance (2017–2019) à Samsun, en Turquie. L'analyse séquentielle de chemin a révélé que, en tant que prédicteurs de premier ordre, le rendement en graines et la teneur en huile ont affiché les effets directs les plus significatifs et les plus positifs sur le rendement en huile dans les cultures d'été et d'hiver. L'analyse du chemin séquentiel des variables de second ordre sur la variable de premier ordre a montré que le nombre de graines par gousse et le nombre de gousses expliquent environ 90 % de la variation du rendement en graines de la culture d'été et que le nombre de branches explique environ 67 % de la variation du rendement en graines de la culture d'hiver. Ces résultats indiquent que le rendement en graines, en tant que prédicteur principal du rendement en huile, suivait des voies différentes pour affecter le rendement en huile en été et en hiver. Une variation plus élevée du nombre de graines par gousse par rapport au nombre de gousses dans cette étude a indiqué que la sélection pour un meilleur rendement en graines peut se faire indirectement en utilisant des plantes avec un nombre de gousses plus faible et un nombre de graines par

\*Corresponding author: [h.zeinelzade@manas.edu.kg](mailto:h.zeinelzade@manas.edu.kg)

gousse plus élevé pendant la saison d'été. De plus, le nombre de branches est le seul trait qui a eu un effet négatif direct sur le rendement en graines pendant la saison d'hiver, ce qui indique que les plantes avec un nombre de branches plus faible devraient être sélectionnées en vue d'un meilleur rendement. Différents facteurs environnementaux, y compris la culture saisonnière de la caméline dans cette étude, se sont avérés être clés dans l'amélioration du rendement en huile et, par conséquent, devraient être considérés comme des indices de critères dans les programmes de sélection de la caméline à l'avenir.

**Mots clés** : caméline / multicolinéarité / analyse de chemin / culture de printemps / culture d'hiver

## 1 Introduction

Camelina is a member of the *Brassicaceae* family, and *Camelina sativa* is the only species of economic importance out of 7 species (*C. sativa*, *C. laxa*, *C. rumelica*, *C. microcarpa*, *C. hispida*, *C. anomala* and *C. alpkoensis*) belonging to the genus *Camelina* (Davis, 1970; Göre, 2015). Camelina is a native plant of Northern Europe and Central Asia, and many remains from archaeological excavations have revealed that this plant has been cultivated in Europe for at least 3000 years (Zubr, 1997). Canola was preferred instead of camelina because it is a higher-yielding oil plant after camelina was cultivated economically as an important oil crop in Europe until the 1930s and in North America until the 1950s (Berti *et al.*, 2016). However, interest in camelina has increased again due to its superior fatty acid ratio compared to other industrial plants, especially Omega-3, its suitability for low-input agricultural systems and its potential to be used as a biofuel (Bujnovský *et al.*, 2020; Lohaus *et al.*, 2020). More than 90% of the fatty acids in camelina oil are unsaturated fatty acids and the concentration of saturated fatty acids is around 8–10%. Polyunsaturated fatty acids constitute an important part (about 60%) of unsaturated fatty acids in camelina (Murphy, 2016). They consist of 35–45% linolenic acid (C18:3; n-3; Omega-3) and 15–20% linoleic acid (C18:2; n-6; Omega-6). The concentration of monounsaturated fatty acids is approximately 36%, and these fatty acids are primarily composed of oleic acid (C18:1n-9) and eicosenoic acid (C20:n-9).

Camelina oil contains more monounsaturated fatty acids than soybean, sunflower and cotton. On the other hand, camelina oil contains polyunsaturated fatty acids close to canola, soybean and sunflower (Günç Ergönül and Aksoylu Özbek, 2020). The combination of these superior oil profile properties with desirable agricultural properties such as short life cycle (Kagale *et al.*, 2014), high adaptation to different environmental conditions (Singh *et al.*, 2015) and low input requirement (Manca *et al.*, 2013) has made camelina a preferred resource. Since the main goal of a breeding program is to obtain high yields of good quality, it is important to know the relationship between the various characteristics that have a direct and indirect effect on the yield.

Physiological events affect the development of plant characters. Yield in plant production results from physiological events occurring in plants. However, it is affected by many factors such as genetic potential, environmental conditions and cultivation technique applications such as fertilization, irrigation, and sowing time (Gill and Narang, 1993). Understanding the relationship between yield and other factors affecting yield is highly likely to lead to higher yields.

The relationship between yield and its components determines the appropriate properties used in yield improvement studies. Simple correlation and path analysis can be used to reveal this relationship. Correlation and path analysis also provide information to researchers about the genetic relationship between yield and the characters that contribute to the yield, and this is useful in the development of reproductive strategies of plants (Mohammadi *et al.*, 2003; Asghari-Zakaria *et al.*, 2007; Feyzian *et al.*, 2009; Maleki *et al.*, 2011; Abdolinasab *et al.*, 2020). Correlation simply measures the relationships between yield and other traits but does not provide a factual concept between dependent and independent variables. Therefore, path analysis is commonly employed in plant breeding to evaluate the correlations between yield and yield components (Tuncturk and Ciftci, 2005).

Multiple regression for path analysis assumes that each character used as a predictor variable is unrelated to the others in the dataset. In reality, yield-related characteristics are closely interconnected, frequently resulting in substantial multicollinearity. Because of this, it is difficult to determine and evaluate the actual contribution of a given character. A new way of grouping variables into different order paths was first used in crop plants by Samonte *et al.* (1998).

Sequential path analysis aims to help researchers understand the relationships between traits that develop next in the life cycle of plants. Del Moral *et al.* (2005) reported that yield and its components are not accurately reflected by simple correlation analyses and path analyses are very helpful in clarifying relationships and effects. Vollmann *et al.* (2007) showed that camelina yield is a complex polygenic character and there are strong relationships between yield and yield components. Simple correlation analysis between the yield and yield components of camelina has been performed by various researchers (Katar *et al.*, 2012; Guy *et al.*, 2014; Jiang *et al.*, 2016; Zanetti *et al.*, 2017; Leclère *et al.*, 2021). Nevertheless, no literature was found using “path analysis” or “sequential path analysis” to reveal the relationships between yield and yield components on camelina, and also the contribution of yield components in seed and oil yields under different seasons remained unknown. The objectives of the current study were: (i) to determine the usefulness of sequential path analysis in camelina, (ii) to obtain information about the relationship between yield and yield components and (iii) to evaluate their relative importance on camelina yield under summer and winter cultivation. In addition, it is to determine the important agricultural characteristics that have direct and indirect effects on the yield that can be used as selection criteria in camelina breeding programs.

## 2 Material and method

### 2.1 Plant material

In this research, camelina genotypes of Swedish origin PI-304269 and Danish origin PI-650142, obtained from USDA and found to have the best performance in adaptation studies in the research area, were used as plant material.

### 2.2 Soil and climate information

The soil structure of the experimental area is moderate in terms of organic matter and phosphorus content and rich in potassium content. In addition, the soil pH level of the experimental area is 7.71, and the soil situation is slightly alkaline (Tab. 1).

It is seen that the temperature, relative humidity, and day length averages of the planting seasons are higher than the averages of the long years. However, the total amount of precipitation is lower than the average for the long years (Tab. 2). It is also seen that the precipitation in the first summer planting flowering season period (July 2017) is negligible (0.4 mm), while the precipitation amount in the second summer planting season is less than half of the average precipitation of long years. In addition, the climate data other than the total precipitation amount in the summer planting season of 2018 are above the average of 2017 and many years, while the relative humidity and day length is higher in the winter planting season. Accordingly, it can be said that the climate data discussed in both growing seasons have a limiting structure in terms of plant cultivation.

### 2.3 Experimentation

The plots for the experiments were arranged in a split-plot design with three replicates, in which varieties were set up as the main plot, whereas sowing times as the sub-plot. The sowing dates for the summer and winter seasons were made with 10-day intervals. When the dates for sowing was arranged at May 1st, 11, 21 and 31 during the 2017 and 2018 summer seasons, the dates for sowing was arranged at 24 October, 3 November, 13 November and 23 November during the 2017 and 2018 winter seasons. The experimental area was empty in the previous year, the field was plowed twice according to the cultivation of camelina, and suitable sowing intervals were opened. The plot size was 3 × 1 m with a spacing of 20 cm row distance and 5 cm plant distance in each row. Camelina genotypes were sown by hand. In both years of the experiment period, 40 kg ha<sup>-1</sup> of ammonium nitrate fertilizer was applied before flowering. Weed control was carried out twice, before the pre-flowering and during the full-flowering stage. Field capacity irrigation was carried out on 1st July in the first year and on 12 June and 11 July in the second year for summer vegetation. At harvest, 0.5 m from the beginning and end of each plot and the remaining area after throwing a row from both sides, was harvested by hand when the plants were 90% mature. Yield and yield components were calculated on ten plants randomly taken from the harvest area of each parcel during the harvest.

**Table 1.** Information on soil structure on experimental area.

Analyse type	Result	Situation
Lime (%)	1.72	Limy
Organic matter (%)	2.83	Medium
Potassium (K <sub>2</sub> O) (kg ha <sup>-1</sup> )	7.59	High
Phosphorus (P <sub>2</sub> O <sub>5</sub> ) (kg ha <sup>-1</sup> )	0.66	Medium
Total salt (%)	0.02	Salt-free
pH	7.71	Slightly alkaline
Saturation (%)	79	Clayey

### 2.4 Data collection

In this experiment, various agronomic traits were recorded, including root length (RL), root weight (RW), plant height (PH), height of first pod (HFP), branch number (BN), pod number (PN), pod weight (PW), seed number per pod (SNP), seed weight per plant (SW), biological yield (BIO), stem ratio (SR), dry matter (DM), grain yield (GY), oil content (OC), oil yield (OY), total seed weight (TSW) and harvest index (HI).

#### 2.4.1 Biological yield (BIO) (g/plant)

Ten plants taken randomly from each plot were placed in a separate paper bag and labeled, and then dried in an oven at 82 °C for 48 hours. At the end of the period, dried plant samples were weighed on a precision scale (g) and recorded as biological weight.

#### 2.4.2 Stem ratio (SR) (g g<sup>-1</sup>)

The data obtained by the ratio of stem dry weight to biological weight in ten plant samples taken randomly from each plot was recorded as stem ratio.

#### 2.4.3 Dry matter (DM) (g/day)

Ten plant samples taken randomly from each plot were dried in an oven at 83 °C for 48 hours and then weighed and their biological weights were calculated. The following equation is used to calculate the amount of dry matter per unit time. The time expression in the relation refers to the number of days from sowing to sampling.

$$\text{Dry matter} [\text{Biological weight}/\text{Time}].$$

#### 2.4.4 Grain yield (GY) (g/plant)

In 10 plant samples taken randomly from each plot during the harvest period, seeds obtained from all capsules on the plant were weighed on a precision scale (g) and recorded as grain yield per plant.

#### 2.4.5 Oil content (OC) (%)

Oil analysis was performed using Ankom XT15 semi-automatic Soxhlet device in accordance with AOAC (Cunniff and Washington, 1997). The XT4 filter bags were tared, 1–1.5 g of finely ground seed sample was placed in the bag and

**Table 2.** Climate data of summer sowing season (May–August 2017 and 2018), winter sowing season (October–June 2018 and 2019) and long years (LY) (1968–2019).

Months	Temperature (°C)				Rainfall (mm)				Relative humidity (%)				Day length (hour)			
	2017	2018	2019	LY	2017	2018	2019	LY	2017	2018	2019	LY	2017	2018	2019	LY
January	–	7.2	7.5	6.4	–	102.4	94.8	89.7	–	91.5	85.7	82.3	–	3.0	3.4	2.7
February	–	8.6	6.8	7.2	–	45.0	28.4	41.3	–	94.0	96.8	85.1	–	4.2	4.9	3.2
March	–	10.1	7.6	8.5	–	115.4	34.2	57.5	–	96.5	92.1	87.2	–	5.0	7.4	3.6
April	–	12.2	10.7	11.4	–	7.6	35.0	50.3	–	92.5	98.7	90.0	–	8.9	8.1	4.7
May	15.0	17.6	17.0	16.2	61.8	24.8	40.0	53.6	92.2	97.9	87.3	90.7	8.3	8.5	9.8	6.2
June	20.8	21.9	23.0	21.1	36.4	8.0	12.0	56.2	87.0	90.7	76.0	87.2	9.8	10.4	10.1	8.3
July	23.7	24.5	22.7	24.0	0.4	16.4	56.4	50.1	78.4	86.4	68.9	83.7	9.6	9.9	10.7	8.8
August	24.2	24.6	23.1	24.7	14.2	18.6	29.8	34.0	80.6	83.4	73.0	81.9	9.1	9.4	9.61	8.2
September	21.1	20.4	–	20.6	11.0	75.0	–	54.0	78.3	92.3	–	83.2	7.2	7.3	–	6.3
October	14.7	16.8	–	15.8	30.0	45.4	–	74.7	81.4	97.0	–	87.3	6.1	6.3	–	4.6
November	11.1	11.8	–	12.0	64.2	59.0	–	70.9	85.9	98.4	–	85.4	3.5	3.7	–	3.8
December	10.0	8.0	–	8.5	145.0	35.8	–	116.6	79.3	97.0	–	83.5	3.2	3.3	–	2.7

	Monthly average temperature (°C)			Total rainfall (mm)			Monthly average relative humidity (%)			Average day length (hour)		
	First year	Second year	LY	First year	Second year	LY	First year	Second year	LY	First year	Second year	LY
<b>Summer</b>	20.9	22.1	21.5	112.8	67.8	193.9	84.5	89.6	85.9	9.2	9.6	7.9
<b>Winter</b>	12.6	12.1	11.9	542.4	459.6	610.8	89.9	92.1	86.5	5.9	6.3	4.4

weighed. The sample number is written on the bags with a special pen that is not affected by chemicals and does not leave any residue. XT4 filter bags were sealed using a heated press device. The samples were dried in an oven at 105 °C for 150 minutes in order to lose their moisture content, and then cooled to room temperature in a desiccator and weighed. After this process, XT4 bags filled with samples were extracted in Ankom XT15 device at 90 °C for 70 minutes. At the end of the period, the XT4 bags were kept in an oven at 105 °C for 30 minutes, cooled in the desiccator and then weighed again on a precision balance. At the end of all these processes, the oil rate was calculated using the following equation:

$$\text{Oil ratio (\%)} = \frac{(1. \text{oven value} - 2. \text{oven value}) \times 100}{\text{First sample weight}}$$

#### 2.4.6 Oil yield (OY) (kg/da)

Considering the oil rate and grain yield per decare, the oil yield per decare was calculated using the equation below:

$$\text{Oil content (\%)} = \frac{\text{Oil ratio(\%)} \times \text{grain yield (kg/da)}}{100}$$

#### 2.4.7 Harvest index (HI) (%)

After removing the marginal effect from each plot, the plants in the middle lines were harvested and weighed. It was calculated using the following equation:

$$\text{Harvest index} = \frac{\text{Grain yield}}{\text{Biological yield}} \times 100.$$

## 2.5 Statistical analysis

Before the analysis of variance, outliers and the normality of the data were tested with the Grubbs test and the Anderson-Darling test. SPSS 24.0 statistical software was used to calculate Pearson correlation coefficients between various pairs of traits. Average data for planting dates were used to perform correlation and path analysis. Therefore, two data sets of summer and winter cultivation were subjected to the respective analyses.

In the current study, two distinguished types of path analysis (simple and complex models) were compared. Complex models (sequential path analysis) set up traits at different ontological levels, with relations (that is, co-relations or cause-and-effect relationships) between them meant to reflect possible biological relations, while simple models (usual path analysis) set up all traits except the dependent one at the same ontological level, making for their being treated as co-related (Kozak and Azevedo, 2014).

Initially, the conventional path analysis was employed, and all traits were treated as the initial predictor traits for the oil yield (OY). Then, for the first and second-order paths of the predictor variables to be organized based on their respective contributions to the total variance of the respective variable (OY) and least collinearity, sequential stepwise multiple

regressions were carried out. The “variance inflation factor” (Hair *et al.*, 1984), which is the inverse of the “tolerance” value, was used to quantify the extent of multicollinearity in each component path using SPSS 24.0 statistical software. After constructing a sequential path diagram, path analysis was done using AMOS 24.0. First-order predictors described the associations with OY, followed by second-order predictors explaining the relationships with the first-order predictors.

### 3 Results and discussion

#### 3.1 Phenotypic correlations

All traits in winter camelina cultivation in the present study, except SR and HI, were measured higher than in summer cultivation (unpublished data) due to the longer vegetative growth period of camelina. Climate data in the current study showed that in winter cultivation of camelina, monthly average temperature, total rainfall, and monthly average relative humidity were higher and the average day length was lower compared to summer cultivation (Tab. 2). So, the earlier autumn sowing and winter cultivation of camelina not only allowed the plants to intercept and absorb more solar radiation for a longer growing period, more water and relative humidity for the crop, but also prevented end-season drought and heat stress during the seed filling period. These results confirmed previous findings on camelina that weather conditions have a significant impact on seed production and that milder temperatures during the growing season result in higher seed yields (Zubr, 1997; Zanetti *et al.*, 2017; Zanetti *et al.*, 2021).

Phenotypic correlation coefficients of different traits in summer and winter cultivation of camelina are presented in Table 3. The OY strongly followed the GY in both summer and winter cultivation. The strong and positive correlation between grain and oil yield (0.977\*\*) was due to the calculation of oil yield by multiplying grain yield by oil content. A strong positive correlation between OY and GY was also reported in Katar *et al.* (2012) and Angelini *et al.* (2020) studies. SNP had positive and significant correlations with OY (0.734\*) and GY (0.780\*) in summer cultivation, whereas this trait had negatively non-significant correlations with OY (−0.678) and GY (−0.573) in winter cultivation (Tab. 3). Angelini *et al.* (2020) also reported a positive and significant correlation of SNP with OY and GY.

Morphological differences between summer and winter-grown camelina have been previously noted by several researchers (Berti *et al.*, 2011; Katar *et al.*, 2012; Guy *et al.*, 2014; Anderson *et al.*, 2018; Kinay *et al.*, 2019; Angelini *et al.*, 2020; Kurt and Göre, 2020; Zanetti *et al.*, 2020).

According to Guy *et al.* (2014), spring planting in the Pacific Northwest (PNW) of the United States yielded better seed performance than fall planting. In contrast, in Turkey’s ecological conditions, research by Katar *et al.* (2012) and Kinay *et al.* (2019) found that winter sowings outperformed summer sowings in yield and yield components, which were similar to this study.

Camelina seed oil content rose when the flowering and seed-filling stages occurred at lower temperatures (*i.e.*, autumn sowing), according to Righini *et al.* (2019). Zanetti *et al.* (2020) stated that camelina’s longer period of vegetative development during the winter cultivation is likely the

explanation for taller and usually bigger plants. Carbohydrates and nitrogen are often transferred from vegetative tissues to reproductive organs to boost seed yield when crops have a prolonged vegetative phase (Bouchet *et al.*, 2016). It has been reported that seed yield and oil content are higher when camelina is grown at colder growing season temperatures (especially during seed development) (Obour *et al.*, 2017). Obeng *et al.* (2019) revealed that GY was negatively associated with heat stress index, suggesting that heat stress during flowering and seed development reduced camelina’s seed production.

The correlation between camelina seed yield and yield components to better understand the relationship between yield-related traits has been reported by several researchers (Katar *et al.*, 2012; Guy *et al.*, 2014; Hossain *et al.*, 2019; Angelini *et al.*, 2020). PN had negatively non-significant correlations with OY (−0.229) and GY (−0.142) in summer cultivation, while had negatively significant correlations with OY (−0.841\*\*) and GY (−0.769\*) in winter cultivation (Tab. 3). A similar result was also reported in Jewett (2013) study, which found that the camelina yields are unaffected by PN. Hossain *et al.* (2019) stated that the contribution of one pod per plant to seed yield in camelina was low. It can be said that this is due to the fact that the camelina plant produces about three to four times more pods than the canola plant. Therefore photosynthetic accumulation is less in camelina pods than in canola pods.

BN showed a positively non-significant correlation with PN (0.589) in summer cultivation and a significant positive correlation (0.865\*\*) in winter cultivation (Tab. 3). This result was in agreement with the findings of Angelini *et al.* (2020). More branching is produced to compensate for lower plant density, as shown by Hossain *et al.* (2019). Thus, more BN has led to lower plant density and increased PN. This was further corroborated by the absence of a clear relationship between camelina yield and the number of plants per unit area (Angelini *et al.*, 2020). In contrast, Jewett (2013) reported that camelina yield was most strongly influenced by plant density per hectare. He indicated that increasing the planting density of the field is the quickest and most straightforward approach to boost camelina productivity. However, in a dryland agricultural system, a higher density may have a negative tradeoff in increased crop water demand. He concluded that if breeders are interested in improving seed output, it would be prudent to adopt a variety with a thousand seed weight, as this is highly heritable and genotype-dependent.

OC showed a non-significant and positive correlation with GY (0.284) in summer cultivation, while had a non-significant and negative correlation with GY (−0.377) in winter cultivation (Tab. 3). This might be attributed to camelina’s substantially higher vegetative development period in winter cultivation than previously reported by Zanetti *et al.* (2020). HI in winter grown camelina is relatively low compared to other *Brassicaceae*, such as oilseed rape (Fan *et al.*, 2017). Guy *et al.* (2014) concluded that OC was shown to be inversely associated with seed yield. They stated that OC, the larger of the two components of oil yield in their study, did not appear to be necessarily negatively connected with high seed yield, suggesting that the environment might impact seed oil content more than genotype selection. As a matter of fact, it has been reported that the yield performance of camelina is mainly

**Table 3.** Phenotypic correlation coefficients between oil yield and related traits in summer (upper diagonal) and winter (lower diagonal) cultivation of camelina.

RL	RW	PH	HFP	BN	PN	PW	SNP	SW	BIO	SR	DM	GY	OC	OY	TSW	HI
RL	1	0.629	-0.523	-0.512	0.488	0.503	0.472	0.174	0.108	-0.127	0.542	-0.124	-0.600	-0.248	-0.261	0.327
RW	0.450	1	-0.313	-0.756*	0.946**	0.633	-0.059	-0.244	0.431	-0.075	0.492	-0.531	-0.617	-0.624	-0.236	0.386
PH	-0.259	-0.681	1	0.760*	-0.232	0.073	0.136	0.563	0.386	0.845**	0.157	0.515	0.691	0.613	-0.075	-0.318
HFP	-0.300	-0.638	0.810*	1	-0.729*	-0.363	0.262	0.443	0.563	0.637	-0.044	0.633	0.772*	0.762*	0.158	-0.517
BN	0.051	0.514	-0.121	0.237	1	0.589	-0.258	-0.353	0.262	-0.066	0.299	-0.612	-0.501	-0.679	-0.044	0.194
PN	0.426	0.760*	-0.196	-0.025	0.865**	1	0.489	0.394	0.705	0.403	0.739*	-0.142	-0.123	-0.229	-0.287	0.362
PW	0.142	0.388	-0.061	0.253	0.648	0.693	1	0.824*	0.697	0.496	0.755*	0.509	0.022	0.433	-0.408	0.201
SNP	0.325	0.750*	-0.347	-0.152	0.584	0.784*	0.853**	1	0.909**	0.653	0.602	0.780*	0.261	0.734*	-0.625	0.304
SW	0.036	-0.012	0.074	0.321	0.159	0.201	0.830*	0.599	0.709*	0.776*	0.650	0.688	0.278	0.653	-0.385	-0.027
BIO	0.150	0.684	-0.202	-0.096	0.630	0.790*	0.860**	0.906**	0.583	0.726*	0.965**	0.190	-0.062	0.137	-0.443	0.203
SR	-0.235	-0.881**	0.878**	0.726*	-0.461	-0.563	-0.285	-0.560	0.070	0.517	0.594	0.359	0.521	0.418	-0.027	-0.353
DM	0.059	0.424	-0.021	0.191	0.587	0.655	0.943**	0.849**	0.799*	-0.257	1	0.063	-0.192	-0.010	-0.373	0.221
GY	0.151	-0.470	0.075	-0.123	-0.847**	-0.769*	-0.686	-0.573	-0.250	0.495	-0.640	1	0.284	0.977**	-0.580	0.149
OC	-0.554	-0.449	0.438	0.540	0.246	-0.077	0.152	-0.284	0.171	0.303	0.219	-0.377	1	0.457	0.397	-0.398
OY	-0.047	-0.698	0.262	0.145	-0.783*	-0.841**	-0.610	-0.678	-0.122	0.676	-0.552	0.930**	-0.031	1	-0.438	0.029
TSW	-0.353	-0.426	0.419	0.857**	0.418	0.073	0.496	0.107	0.545	0.403	0.415	-0.248	0.516	0.031	1	-0.779*
HI	-0.697	-0.512	0.557	0.696	0.223	-0.194	-0.131	-0.364	-0.136	0.503	-0.043	-0.075	0.617	0.184	0.642	1

**Table 4.** Direct effects of first-order predictor variables on oil yield and measures of collinearity in model 1 (all predictor variables as first-order variables over oil yield as response variable).

Predictor variables	Summer cultivation			Winter cultivation		
	Direct effect	Tolerance	VIF	Direct effect	Tolerance	VIF
RL	0.038	0.445	2.247	0.000	0.461	2.168
RW	−0.028	0.481	2.079	0.004	0.000	2515.267
PH	0.006	0.412	2.428	0.000	0.005	192.656
HFP	0.0130	0.217	4.603	−0.001	0.013	77.466
BN	−0.098	0.271	3.691	0.000	0.034	29.523
PN	−0.076	0.972	1.028	0.000	0.064	15.614
PW	−0.056	0.725	1.379	0.001	0.029	34.776
SNP	−0.094	0.389	2.570	0.001	0.022	46.056
SW	−0.314**	0.085	11.798	0.037	0.029	34.776
BIO	0.201**	0.139	7.170	0.000	0.037	27.358
SR	−0.022	0.680	1.471	0.027	0.003	354.668
DM	−0.032	0.948	1.055	0.000	0.088	25.494
GY	1.060**	0.235	4.254	1.096**	0.011	94.497
OC	0.192**	0.475	2.103	0.320**	0.137	7.326
TSW	0.059	0.322	3.106	0.122	0.150	6.653
HI	−0.095	0.555	1.802	0.000	0.088	11.306

RL: root length; RW: root weight; PH: plant height; HFP: height of first pod; BN: branch number; PN: pod number; PW: pod weight; SNP: seed number per pod; SW: seed weight; BIO: biological yield; SR: stem ratio; DM: dry matter; GY: grain yield; OC: oil content; OY: oil yield; TSW: total seed weight; HI: harvest index.

affected by environmental factors and the effect of genetic factors on yield is less (Obour *et al.*, 2017; Zanetti *et al.*, 2017; Righini *et al.*, 2019).

Negative correlations have been observed between GY and TSW in both growing seasons (Tab. 3). The latter outcome was similar to Gehringer *et al.* (2006), Obour *et al.* (2017), Neupane *et al.* (2018) and Obeng *et al.* (2019) studies, which suggest that TSW may be of little utility in predicting seed yields in camelina. This is also consistent with the findings of Vollmann *et al.* (2007), who discovered a negative association between camelina seed yield and 1000-seed weight in research conducted over three growing seasons. Therefore, TSW may not be a good predictor of grain yield in camelina.

The negative correlation of BN with GY observed in both growing seasons in this study was also previously reported in Hossain *et al.* (2019). Katar *et al.* (2012) reported a significant and positive correlation of BN with GY and OY. They concluded that this difference was related to yearly variations that significantly affected the branch number per plant.

PH had non-significant and positive correlations with GY, OC and OY in summer (0.515, 0.691, and 0.613) and winter (0.075, 0.438, and 0.262) cultivations, respectively (Tab. 3). A positive correlation between PH and GY was also reported in Gehringer *et al.* (2006) and Neupane *et al.* (2020) studies. Results by Guy *et al.* (2014) revealed that a strong positive relationship was observed between HI and SY in their study.

There was a high correlation coefficient between yield components in the current study, which meant that it was impossible to determine the exact contribution of each component to total oil production because of mixed or muddled effects. Path analysis is used to understand the magnitude and direction of component characteristics' direct

and indirect contributions to yield since correlation coefficients do not convey the entire picture when the causative factors are interconnected and interdependent.

### 3.2 Conventional and sequential path analysis

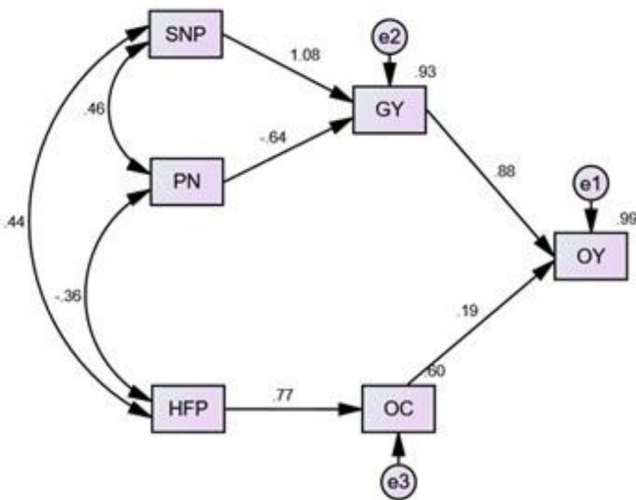
The conventional path analysis demonstrated collinearity between independent traits, demonstrating the first model's insufficiency to reflect each trait's actual contribution in both summer and winter cultivation of camelina (Tab. 4). Table 4 summarizes the direct impacts of predictor traits on camelina oil yield using conventional path analysis and indicators of collinearity across two growing seasons. However, there was considerable collinearity between several traits, notably those with high direct impacts on oil yield, despite the fact that the multicollinearity of some of the traits was good (VIF < 10 and tolerance > 0.1) in both summer and winter cultivation. For example, the VIF for SW was 11.798 in summer cultivation and 2515.267 for RW during winter cultivation, respectively (Tab. 4). That is because agronomic traits are highly correlated (Tab. 4). Multicollinearity was found to be severe in matrices with characteristics with VIFs greater than 10 and TOL lower than 0.1 (Mansfield and Helms, 1982). In this regard, the VIF of SW in summer cultivation and all traits except RL, OC and TSW in winter cultivation showed high collinearity (Tab. 4).

In this study, the use of sequential path analysis was found to be effective in reducing the collinearity of attributes (Tab. 5). Compared to conventional path analysis, sequential path analysis simplifies the relationship between traits and their contribution to oil yield. The findings revealed a considerable reduction in the variance-influence factor values of the first

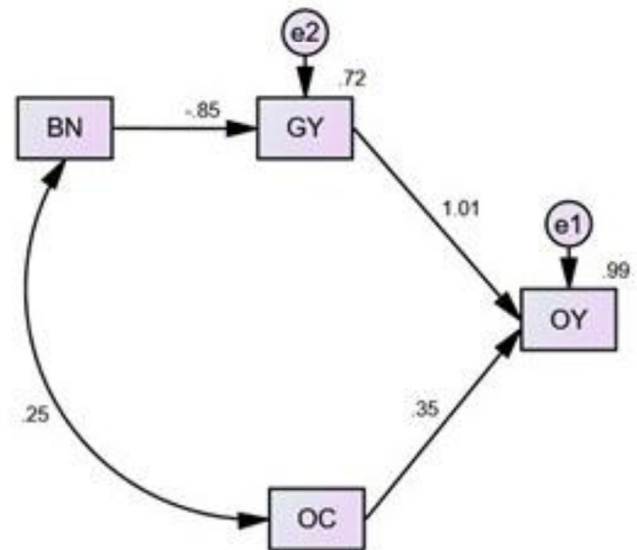
**Table 5.** Direct effects, tolerance, and variance inflation factor (VIF) values for the predictor variables in model 2 (predictors grouped into first- and second-order variables).

Response variable	Predictor variables	Adjusted $R^2$	Tolerance	VIF	Direct effect	Indirect effect via	
<b>Summer cultivation</b>							
OY	GY	0.986	0.920	1.087	0.922**	GY	OC
	OC		0.920		1.087	0.195**	–
GY	SNP	0.902	0.787	1.271	1.076**	SNP	PN
	PN		0.787		1.271	–0.639**	–
OC	HFP	0.529	1.000	1.000	0.772*	–	–
<b>Winter cultivation</b>							
OY	GY	0.978	0.858	1.166	1.071**	GY	OC
	OC		0.858		1.166	0.373**	–
GY	BN	0.671	1.000	1.000	–0.847**	–	–

RW: root weight; PH: plant height; HFP: height of first pod; BN: branch number; GY: grain yield; OC: oil content; OY: oil yield; TSW: total seed weight.



**Fig. 1.** Sequential path model indicates the interrelationships between oil yield and related traits in the summer cultivation of camelina.



**Fig. 2.** Sequential path model indicates the interrelationships between oil yield and related traits in the winter cultivation of camelina.

model compared to the second model in both summer and winter cultivation (Tab. 5). Moreover, the VIF lower than 10 and TOL greater than 0.1 displayed by the first- and second-order predictors in the sequential path analysis elucidated the interrelationships between the OY-related characteristics. Using stepwise regression analysis, variable collinearity was reduced, and by lowering effect mixing, each variable’s real participation rate was accurately determined in different paths. Studies have shown that sequencing path analysis is more effective than conventional path analysis (Mohammadi *et al.*, 2003; Asghari-Zakaria *et al.*, 2007; Dalkani *et al.*, 2011). GY and OC were chosen as first-order predictors to explain OY variance in both summer and winter cultivation (Tab. 5, Figs. 1 and 2). Given that GY mainly determines OY, these variables account for about 98% of the variance in OY.

GY and OC displayed the most significant and positive direct effects on OY (0.922 and 0.195, respectively) (Tab. 5). Both seasons had different magnitudes of impacts, but their direction remained the same. Their indirect effects on OY were found to be positive, non-significant and negligible in summer cultivation. In contrast, the indirect effect of OC through GY was found to be negative (–0.404), non-significant and moderate in winter cultivation. This was because of a low, negative, and non-significant correlation (–0.377) between the two characters OC and OY in winter cultivation (Tab. 4), indicating that indirect selection for higher GY generally has lower OC in winter cultivation (Tab. 5).



According to the stepwise regression of second-order variables over the first-order variables, SNP and PN account for about 90% of the GY's variation in summer cultivation, and BN explained about 67% of the variation of the GY in winter cultivation (Tab. 5). As a main predictor of OY, these results indicated that GY had different pathways in the summer and winter seasons (Figs. 1 and 2). Mohammadi *et al.* (2003) concluded that character relationships revealed by path analysis might be impacted by a variety of factors, including the germplasm utilized, the attributes evaluated for analysis, the environment(s) used for assessment, and the statistical approaches used to resolve the correlations.

In the summer season, SNP had the most considerable positive direct effect on GY (1.076\*\*) with a non-significant, low and negative indirect effect *via* PN (−0.295). Selecting for GY can be done indirectly using plants with lower PN and higher SNP in the summer season (Tab. 5), despite the fact that PN had a negative direct effect on GY (−0.639\*\*) and a non-significant, moderate, and positive indirect effect *via* SNP (0.497). In Hossain *et al.* (2019) study using multiple regression analysis revealed similar results, and they concluded that the SNP significantly affected seed yield rather than PN. In the winter season, BN was the only trait that had a direct negative effect on GY (−0.847\*\*), indicating that plants with lower BN should be selected for higher GY. Darapuneni *et al.* (2014) reported that seed yield in flax was negatively correlated with pods per tiller. They found that the positive direct effect of tiller number on yield was significant enough to outweigh the negative indirect effect of pods per tiller, resulting in a positive effect of tiller number on seed yield.

The Chi-square approach was used to validate the sequential path model. The expected and observed correlation coefficients were compared. Results of Chi-square analysis showed low and non-significant values (0.087<sup>ns</sup>, 0.247<sup>ns</sup>) for the two datasets (summer and winter cultivation) used in the construction of the path.

The diagrams of sequential path analysis for both summer and winter cultivation are shown in Figures 1 and 2. Improved insight into the interdependencies between oil yield-related characteristics in camelina was achieved through the categorization of predictors into first- and second-order predictors.

## 4 Conclusion

Climatic and environmental factors such as seasonal cultivation can influence camelina seed and oil yield. In the present study, stepwise regression analysis was used to reduce the collinearity measures of all variables, allowing for the determination of the actual contributions of each predictor variable in distinct path components with minimum confounding effects and interference. The sequential path analysis of second-order variables over the first-order variable by stepwise regression revealed that SNP and PN explained approximately 90% of the variation of the GY in summer cultivation and BN explained approximately 67% of the variation GY in winter cultivation. These results indicated that GY, as a main predictor of OY, had different pathways to affect oil yield in the summer and winter seasons. A higher magnitude of SNP compared to PN in this study indicated that selecting for higher GY can be done indirectly using

plants with lower PN and higher SNP in the summer season. Moreover, BN was the only trait that had a direct negative effect on GY in the winter season, indicating that plants with lower BN should be selected for higher GY. Different environmental factors, including the seasonal cultivation of camelina in this study, were found to be a key factor in improving oil yield and, hence, should be considered as criteria indices in camelina breeding programs in the future.

## Conflicts of interest

The authors declare that they have no conflicts of interest in relation to this article.

## References

- Abdolinassab M, Rahimi M, Karatas A, Ercisli S. 2020. Sequential path analysis and relationships between fruit yield in watermelon. *Pakistan Journal of Agricultural Sciences* 57: 1425–1430.
- Anderson JV, Horvath DP, Doğramaci M, *et al.* 2018. Expression of FLOWERING LOCUS C and a frameshift mutation of this gene on chromosome 20 differentiate a summer and winter annual biotype of *Camelina sativa*. *Plant Direct* 2: 1–14.
- Angelini LG, Abou Chehade L, Foschi L, Tavarini S. 2020. Performance and potentiality of camelina (*Camelina sativa* L. Crantz) genotypes in response to sowing date under Mediterranean environment. *Agronomy* 10: 1929.
- Asghari-Zakaria R, Fathi M, Hasan-Panah D. 2007. Sequential path analysis of yield components in potato. *Potato Research* 49: 273–279.
- Berti M, Gesch R, Eynck C, Anderson J, Cermak S. 2016. Camelina uses, genetics, genomics, production, and management. *Industrial Crops and Products* 94: 690–710.
- Berti M, Wilckens R, Fischer S, Solis A, Johnson B. 2011. Seeding date influence on camelina seed yield, yield components, and oil content in Chile. *Industrial Crops and Products* 34: 1358–1365.
- Bouchet A-S., Laperche A, Bissuel-Belaygue C, Snowdon R, Nesi N, Stahl A. 2016. Nitrogen use efficiency in rapeseed. A review. *Agronomy for Sustainable Development* 36: 1–20.
- Bujnovský R, Holíčková M, Ondrejčíková P. 2020. Spring *Camelina sativa* – Perspective cultivation as biofuel feedstock in Slovakia. *Industrial Crops and Products* 154: 112634.
- Cunniff P, Washington D. 1997. Official methods of analysis of aoac international. *J AOAC Int* 80: 127A.
- Dalkani M, Darvishzadeh R, Hassani A. 2011. Correlation and sequential path analysis in Ajowan (*Carum copticum* L.). *Journal of Medicinal Plants Research* 5: 211–216.
- Darapuneni MK, Morgan GD, Ibrahim AMH, Duncan RW. 2014. Association of flax seed yield and its components in Southeast Texas using path coefficient and biplot analyses. *Journal of Crop Improvement* 28: 1–16.
- Davis PH. 1970. Flora of Turkey and the East Aegean Islands. Vol. 3. Edinburgh University Press.
- del Moral LFG, Rharrabti Y, Elhani S, Martos V, Royo C. 2005. Yield formation in mediterranean durum wheats under two contrasting water regimes based on path-coefficient analysis. *Euphytica* 146: 203–212.
- Fan J, McConkey B, Janzen H, Townley-Smith L, Wang H. 2017. Harvest index-yield relationship for estimating crop residue in cold continental climates. *Field Crops Research* 204: 153–157.

- Feyzian E, Dehghani H, Rezaei A, Jalali M. 2009. Correlation and sequential path model for some yield-related traits in melon (*Cucumis melo* L.). *Journal of Agricultural Science and Technology* 11: 341–353.
- Gehring A, Friedt W, Lühs W, Snowdon R. 2006. Genetic mapping of agronomic traits in false flax (*Camelina sativa* subsp. *sativa*). *Genome* 49: 1555–1563.
- Gill M, Narang R. 1993. Yield analysis in gobhi sarson (*Brassica napus* subsp. *Oleifera* var. *Annua*) to irrigation and nitrogen. *Indian Journal of Agronomy* 38: 257–265.
- Göre M. 2015. A research to establish effects of explant sources and plant growth regulators on camelina [*Camelina sativa* (L.) crantz] tiller and plant induction. MS.c thesis. Ondokuz Mayıs University, Institute of Science.
- Günç Ergönül P, Aksoylu Özbek Z. 2020. Cold pressed camelina (*Camelina sativa* L.) seed oil. In: Green technology, bioactive compounds, functionality, and applications. Elsevier.
- Guy SO, Wysocki DJ, Schillinger WF, *et al.* 2014. Camelina: adaptation and performance of genotypes. *Field Crops Research* 155: 224–232.
- Hair JF, Anderson RE, Tatham RL, Black WC. 1984. Multivariate data analysis with readings, 1995. Tulsa, OK: Petroleum Publishing.
- Hossain Z, Johnson EN, Wang L, Blackshaw RE, Cutforth H, Gan Y. 2019. Plant establishment, yield and yield components of Brassicaceae oilseeds as potential biofuel feedstock. *Industrial Crops and Products* 141: 111800.
- Jewett FG. 2013. Camelina variety performance for yield, yield components and oil characteristics. MS.c. Colorado State University.
- Jiang Y, Li J, Caldwell CD. 2016. Glucosinolate content of camelina genotypes as affected by applied nitrogen and sulphur. *Crop Science* 56: 3250–3262.
- Kagale S, Koh C, Nixon J, *et al.* 2014. The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. *Nature Communications* 5: 1–11.
- Katar D, Arslan Y, Subasi İ. 2012. Genotypic variations on yield, yield components and oil quality in some Camelina (*Camelina sativa* (L.) Crantz) genotypes. *Turkish Journal of Field Crops* 17: 105–110.
- Kinay A, Yilmaz G, Ayisigi S, Dokulen S. 2019. Yield and quality parameters of winter and summer-sown different camelina (*Camelina sativa* L.) genotypes. *Turkish Journal of Field Crops* 24: 164–169.
- Kozak M, Azevedo RA. 2014. Sequential path analysis: what does “sequential” mean? *Scientia Agricola* 71: 525–527.
- Kurt O, Göre M. 2020. Effects of sowing date and genotype on oil content and main fatty acid composition in camelina [*Camelina sativa* L. (Crantz)]. *Turkish Journal of Field Crops* 25: 227–235.
- Leclère M, Lorent A-R., Jeuffroy M-H., Butier A, Chatain C, Loyce C. 2021. Diagnosis of camelina seed yield and quality across an on-farm experimental network. *European Journal of Agronomy* 122: 126190.
- Lohaas RH, Neupane D, Mengistu MA, Solomon JKQ, Cushman JC. 2020. Five-year field trial of eight *Camelina sativa* cultivars for biomass to be used in biofuel under irrigated conditions in a semi-arid climate. *Agronomy* 10: 562.
- Maleki HH, Karimzadeh G, Darvishzadeh R, Sarrafi A. 2011. Correlation and sequential path analysis of some agronomic traits in tobacco (*Nicotiana tabacum* L.) to improve dry leaf yield. *Australian Journal of Crop Science* 5: 1644–1648.
- Manca A, Pecchia P, Mapelli S, Masella P, Galasso I. 2013. Evaluation of genetic diversity in a *Camelina sativa* (L.) Crantz collection using microsatellite markers and biochemical traits. *Genetic Resources and Crop Evolution* 60: 1223–1236.
- Mansfield ER, Helms BP. 1982. Detecting multicollinearity. *The American Statistician* 36: 158–160.
- Mohammadi S, Prasanna B, Singh N. 2003. Sequential path model for determining interrelationships among grain yield and related characters in maize. *Crop Science* 43: 1690–1697.
- Murphy EJ. 2016. Camelina (*Camelina sativa*). In: McKeon TA, Hayes DG, Hildebrand DF, Weselake RJ, eds. Industrial oil crops. Elsevier.
- Neupane D, Solomon JK, Mclennon E, Davison J, Lawry T. 2020. Camelina production parameters response to different irrigation regimes. *Industrial Crops and Products* 148: 112286.
- Neupane D, Solomon JKQ, Davison J, Lawry T. 2018. Nitrogen source and rate effects on grain and potential biodiesel production of camelina in the semiarid environment of northern Nevada. *GCB Bioenergy* 10: 861–876.
- Obeng E, Obour AK, Nelson NO, *et al.* 2019. Seed yield and oil quality as affected by camelina cultivar and planting date. *Journal of Crop Improvement* 33: 202–222.
- Obour AK, Obeng E, Mohammed YA, *et al.* 2017. Camelina seed yield and fatty acids as influenced by genotype and environment. *Agronomy Journal* 109: 947–956.
- Samonte SPB, Wilson L, McClung A. 1998. Path analyses of yield and yield-related traits of fifteen diverse rice genotypes. *Crop Science* 38: 1130–1136.
- Righini D, Zanetti F, Martínez-Force E, Mandrioli M, Toschi TG, Monti A. 2019. Shifting sowing of camelina from spring to autumn enhances the oil quality for bio-based applications in response to temperature and seed carbon stock. *Industrial Crops and Products* 137: 66–73.
- Singh R, Bollina V, Higgins EE, *et al.* 2015. Single-nucleotide polymorphism identification and genotyping in *Camelina sativa*. *Molecular Breeding* 35: 1–13.
- Tunçtürk M, Ciftçi V. 2005. Selection criteria for potato (*Solanum tuberosum* L.) breeding. *Asian Journal of Plant Sciences* 4: 27–30.
- Vollmann J, Moritz T, Kargl C, Baumgartner S, Wagentristl H. 2007. Agronomic evaluation of camelina genotypes selected for seed quality characteristics. *Industrial Crops and Products* 26: 270–277.
- Zanetti F, Alberghini B, Marjanović Jeromela A, *et al.* 2021. Camelina, an ancient oilseed crop actively contributing to the rural renaissance in Europe. A review. *Agronomy for Sustainable Development* 41: 1–18.
- Zanetti F, Eynck C, Christou M, *et al.* 2017. Agronomic performance and seed quality attributes of Camelina (*Camelina sativa* L. crantz) in multi-environment trials across Europe and Canada. *Industrial Crops and Products* 107: 602–608.
- Zanetti F, Gesch RW, Walia MK, Johnson JMF, Monti A. 2020. Winter camelina root characteristics and yield performance under contrasting environmental conditions. *Field Crops Research* 252.
- Zubr J. 1997. Oil-seed crop: *Camelina sativa*. *Industrial Crops and Products* 6: 113–119.