Article

Characterization of Oilseed Crop Noug (*Guizotia abyssinica*) Using Agro-Morphological Traits

Adane Gebeyehu [1,2,*], Cecilia Hammenhag [1], Rodomiro Ortiz [1], Kassahun Tesfaye [2] and Mulatu Geleta [1]

1 Department of Plant Breeding, Swedish University of Agricultural Sciences, P.O. Box 190, 23422 Lomma, Sweden; cecilia.hammenhag@slu.se (C.H.); rodomiro.ortiz@slu.se (R.O.); mulatu.geleta.dida@slu.se (M.G.)

2 Ethiopian Biotechnology Institute, Addis Ababa P.O. Box 5954, Ethiopia; kassahuntesfaye@yahoo.com

* Correspondence: adane.gebeyehu.demissie@slu.se or adyamrot@gmail.com

Abstract: Noug (*Guizotia abyssinica*) is an outcrossing oilseed crop that serves as a source of edible oil and other nutrients although its seed yield is generally low. The analysis of agro-morphological traits in relation to seed and oil yields is important for improving the productivity of this crop. The present study aimed at assessing the variation and heritability of quantitative (10) and qualitative (6) traits in noug based on 60 landrace accessions collected from wide geographic area in Ethiopia. The field trial was conducted at two sites in Ethiopia using a square lattice design. The analysis of variance revealed significant variation ($p < 0.05$) among these accessions. The highest broad-sense heritability (H^2) was recorded for days to 10% flowering (DTF10; 85.4%), whereas number of capitula per plant (NCPP) and number of seeds per capitulum (NSPC) showed medium heritability ($H^2 = 38.5\%$ and 31.6%, respectively). NCPP and NSPC showed a highly significant ($p < 0.01$) positive and negative genotypic correlation with days to flowering, respectively. These three traits showed very low genotypic coefficient of variation (<1%). In the case of qualitative traits, small capitulum, large flower, green leaf, green stem and course leaf margin showed significant association with higher number of seeds per plant (NSPP). Euclidean distance-based cluster analysis revealed that the clustering pattern of the accessions poorly correlates with the geographic distance between sample collecting sites. Similarly, no clear clustering pattern of accessions was revealed by principal component analysis (PCA) that explained 66.3% and 53.6% of the total variation of quantitative traits and qualitative traits, respectively. The oil content of these accessions was previously investigated and accessions with high oil content show large differences in terms of days to flowering, NSPP and thousand seed weight (TSW). Among the accessions included in this research, *Hr_B21*; *Gj_C17*, *Sh_I4* and *Gr_F15 Gj_G18* and *Tg-R13* are top ranking, as they have at least one the following highly desirable traits: early maturity, high oil content, NSPP and TSW. Hence, crossbreeding of their selected genotypes would lead to the development of new cultivars that combine early maturity and both high seed and oil yields.

Keywords: cluster analysis; *Guizotia abyssinica*; heritability; qualitative traits; quantitative traits; seed yield

1. Introduction

Ethiopia is well known as a center of diversity for several crops, including noug (*Guizotia abyssinica* (L.f.) Cass), teff (*Eragrostis tef* (Zucc.) Trotter), enset (*Ensete ventricosum* (Welw.) Cheesman), Ethiopian mustard (*Brassica carinata* A. Braun), sorghum (*Sorghum bicolor* (L.) Moench.), durum wheat (*Triticum durum* Desf.) and barley (*Hordeum vulgare* L.), among others. Noug is an annual diploid ($2n = 2x = 30$) [1] oilseed crop exhibiting homomorphic type of sporophytic self-incompatibility [2,3]. It is the only cultivated species in the genus *Guizotia* that belongs to the subtribe Milleriinae of the tribe Heliantheae in the family Compositae [3,4]. In Africa, it is mainly cultivated in Ethiopia but it is also grown to

some extent in countries such as Sudan, Uganda, Tanzania, Malawi and Zimbabwe [5,6]. In Asia, it is a minor oilseed crop mainly cultivated in India and to some extent in Nepal, Bangladesh and Bhutan. Other countries where it is grown as minor crop are the West Indies in the Caribbean and USA [5]. It can grow up to a height of 1.4 to 2 m although short types (about 35 cm) are also known [6]. The number of branches per plant show large variation and is usually low when the environmental conditions are unfavorable and plant density is high.

Based on the time it takes to reach maturity, noug grown in Ethiopia is classified into three groups: abat, mesno and bunegne. Abat is a late-maturing type, which is planted in June and harvested in December; mesno is a type suitable for cultivation in waterlogged soils from late September to January; and bunegne is an early-maturing type suitable for cultivation in lowland areas from July to October [7]. Generally, the crop flowers within 50 to 100 days and matures within about 120 to 180 days after sowing and produces up to 40 capitula per plant [8,9].

Noug is among major sources of edible oil in Ethiopia, mainly produced by smallholders. It accounts for about 30% of the countries oilseed production and 26% of the produce is retained for home consumption [10]. Traditionally, noug seed oil is extracted for home consumption from slightly roasted seeds by grinding with a pestle and mortar, adding hot water and stirring it until the oil floats to the surface, which is followed by scooping the oil off. The press-cake after oil extraction is an excellent poultry and livestock feed as it contains 33 to 37% protein and is rich in inorganic constituents and crude fibers [6]. In non-food sectors, noug oil can be used for illumination, anointing, painting and cleaning machineries as well as for pharmaceutical purposes and making soaps [9].

Noug is grown across different regions in Ethiopia although more than 90% of its production is concentrated in the highlands of Amhara and Oromia regions. Wollega and Shewa are the major noug-producing area in the Oromia region whereas Gojjam and Gondar are the top noug producing areas in the Amhara region [9,11]. According to a CSA [12] report, there was 3.3% increase in area of noug production and 6.9% increase in total production from 2016/2017 to 2017/2018 with an average yield of 1.1 t ha^{-1}. In the year 2018/2019, its yield increased form 1.1 t ha^{-1} to 1.49 t ha^{-1} with a 3.2% increase in yield [12]. The crop has high genetic diversity, particularly in Ethiopia [13], where it has been domesticated [14]. Geleta [15] and Geleta et al.[16] reported the genetic diversity of 70 Ethiopian noug collections using random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) markers and revealed the presence of higher genetic diversity within populations than among populations, which suggests wide genetic basis for noug improvement. However, its seed yield remains low, with national average productivity of about 1.1 t ha^{-1}, [12] with only a slight increase in 2018/2019 cropping season [12] despite breeding efforts to increase seed yield. Major factors contributing to low seed yield include low response to inputs, seed shattering due to indeterminate growth habit, lodging, bird damage, parasitic weeds such as dodder (*Cuscuta campestris*), insect pests and diseases [9,17]. In India, noug serves as source of edible oil and contributes about 3% of oilseed production [18]. In the United States and Europe, noug seed is used for feeding birds, especially finches [6,19].

Germplasm collected from diverse environments are often heterogeneous and usually exhibit considerable genetic variation for desirable traits. Heritable variation may be controlled by few major genes (e.g., flower size and stem color,) or multiple additive genes (e.g., number of flowers per plant and plant height) although the intensity of the gene expression vary with environments [20,21]. Trait characteristics in plant genotypes develop as a result of gene expression and its interaction with environments during growth and development. Hence, understanding the variation and heritability of desirable agro-morphological traits is essential for crop improvement through utilization of existing genetic resources. The objectives of the present study were therefore to characterize Ethiopian noug accessions based on their agro-morphological traits and assess their genetic

variation as well as to identify traits that associates with seed yield and accessions that have a high yield potential for further improvement.

2. Materials and Methods

2.1. Plant Material

Sixty-four noug landrace accessions collected from its major growing areas across different agro-ecological zones in Ethiopia were initially included in this study. However, four accessions failed during the field trial and hence 60 accessions were used for data analysis (Supplementary Materials Table S1).

2.2. Experimental Sites and Design

The experiment was carried out at two experimental sites: (1) Holeta Agricultural Research Center, which is located 30 km west of Addis Ababa, 2400 m above sea level (masl), 9°00′ N and 38°30′ E having a soil type of nitosols and vertisols; and (2) Debre Zeit Agricultural Research Center, which is located 47 km southeast of Addis Ababa, 1900 masl, 08°44′ N and 38°58′ E having a soil type of vertisols (Figure 1). Holeta receives an average of 1144 mm annual rainfall with 6 °C and 22 °C minimum and maximum temperature, respectively, whereas Debre Zeit receives an average of 851 mm annual rainfall with 9 °C and 28 °C minimum and maximum temperature, respectively.
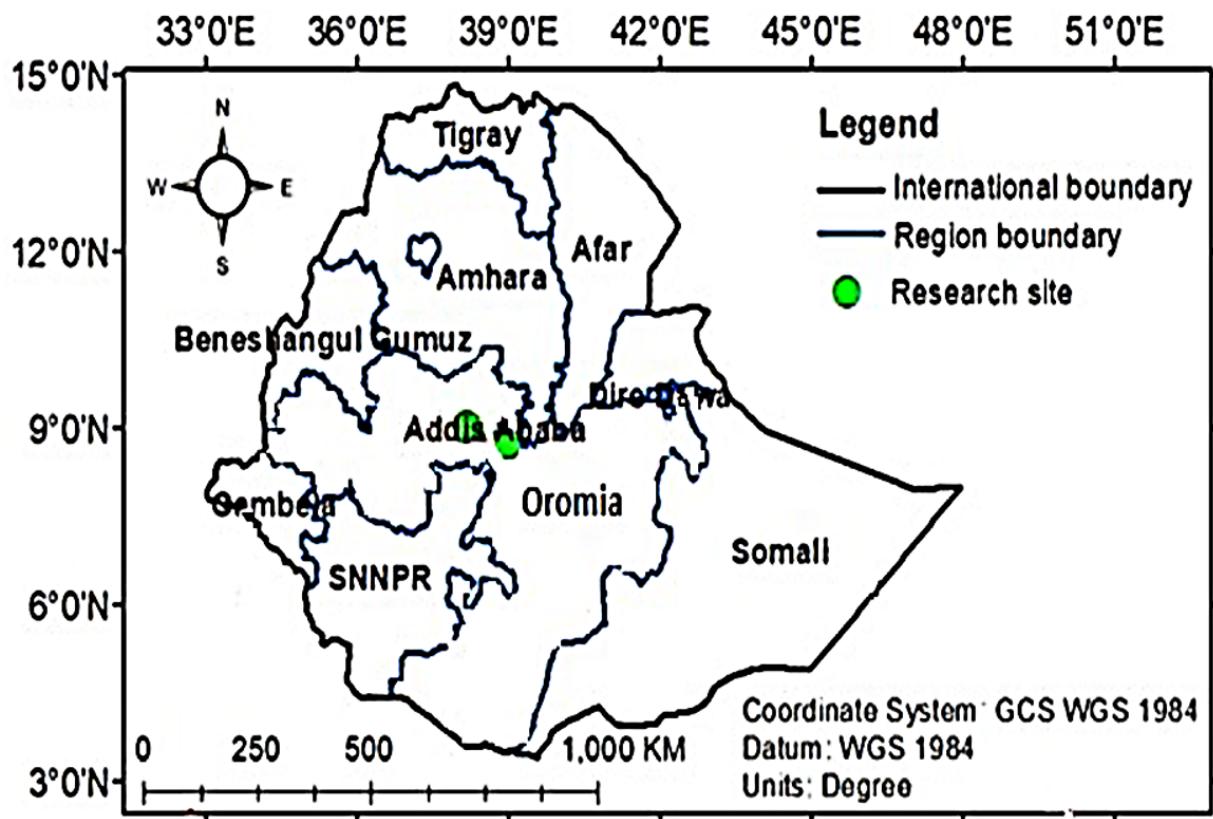

Figure 1. Map of Ethiopia showing the field trial sites.

The field trial was conducted from June to December in two replications at each site using a lattice square experimental design (8 blocks × 8 plots). The size of each plot was 9 m^2 (3 × 3 m). In each plot, seeds were planted in 10 rows and the spacing between plants within a row was 10 cm, whereas the spacing between rows was 30 cm. The distance between plots within a replicate was 1 m, whereas the distance between the two replicates in each site was 5 m. The accessions were randomly assigned to plots within a replicate and randomization was done for each replicate separately.

2.3. Recording Phenotypic Data

Data were recorded for each trait based on standard noug morphological descriptors [9,22]. The traits include 10 quantitative traits; namely, days to 10% flowering (DTF10), days to 50% flowering (DTF50), days to 90% flowering (DTF90), number of nodes per plant (NNPP), number of capitula per plant (NCPP), number of primary branches (NPB), plant height (PH), leaf length (LL), leaf width (LW) and number of seeds per capitula (NSPC) as well as six qualitative traits; flower size (FS), capitulum size (CS), leaf color (LC), leaf margin (LM), stem color (SC) and stem hairiness (SH) (Table 1). The data for DTF10, DFF50 and DTF90 were collected at a plot level whereas data for the other traits were collected from 50 plants in each plot (five randomly selected plants from each of the 10 rows).

Table 1. Description of the ten quantitative and six qualitative/categorical traits of noug included in the present study.

Trait	Trait Code	Description
Quantitative		
Days to 10% flowering	DTF10	Number of days from planting to the flowering of 10% of plants on a plot
Days to 50% flowering	DTF50	Number of days from planting to the flowering of 50% of plants on a plot
Days to 90% flowering	DTF90	Number of days from planting to the flowering of 90% of plants on a plot
Leaf length	LL	The length of leaves in the middle part of a stem measured from base to tip in cm (average of five leaves per plant)
Leaf width	LW	The width (at the center) of leaves in the middle part of a stem measured in cm (average of five leaves per plant)
Number of capitula per plant	NCPP	Total number of capitula on each plant (count)
Number of nodes per plant	NNPP	Total number of nodes on each plant (count)
Number of primary branches	NPB	Total number of primary branches on each plant (count)
Number of seeds per capitulum	NSPC	An average number of seeds per capitulum, calculated based on five capitula from each plant
Plant height	PH	Height of a plant measured in cm at full flowering stage
Qualitative/Categorical		
Flower size	FS	Three categories were given based on the diameter of fully developed flowers: 1 = small (<3 cm); 3 = medium (3 to 4 cm); 5 = large (>4 cm)
Capitulum size	CS	Three categories were given based on the cross-sectional circumference of mature capitulum: 1 = small (<3 cm); 3 = medium (3 to 4 cm); 5 = large (>4 cm)
Leaf color	LC	Three categories were given based on the intensity of green color of mature leaves: 1= light green; 3 = green; 5 = deep green
Leaf margin	LM	Three categories were given based on the indent size of the margins of mature leaves: 1 = fine; 3 = medium; 5 = course
Stem color	SC	Two categories were given based on the color of the middle portion of a stem: 1 = green; 2 = purple
Stem hairiness	SH	Three categories were given based on the level of hairiness of the middle part of stems: 1 = no or little observable hair; 3 = medium; 5 = clearly visible dense hair

2.4. Data Analysis

Quantitative trait data were subjected to analysis of variance (ANOVA) using the GenStat 16th edition statistical package (https://www.vsni.co.uk/software/genstat/ accessed on 24 July 2021). Broad-sense heritability (H^2), cluster analysis and genetic correlations were computed by META-R version 6.0. The association between two environments each having two replications and accessions were analyzed using Tukey method at 95% confidence level and the six qualitative traits using Chi-square (χ^2) test and Kruskal-Wallis

test. The pattern of variation among accessions were examined by principal component analysis (PCA) using MINITAB v18.

2.5. Estimation of Variance Components

The genotypic and phenotypic variances for quantitative traits were estimated as phenotypic and genotypic coefficients of variation (PCV and GCV, respectively hereafter) expressed as percentage of the corresponding phenotypic and genotypic standard deviations. The PCV and GCV values were ranked as low (0–10%), medium (10–20%) and high (>20%) as suggested by Sivasubramanian and Madhavamenon [23]. The variance components were estimated as follows:

$$\text{Genotypic variance } (\sigma^2_G) = (MS_G - MSe)/r$$

$$\text{Environmental variance } (\sigma^2_e) = MS_e/r$$

$$\text{Phenotypic variance } (\sigma^2 p) = \sigma^2_G + \sigma^2_e$$

where MS_G is the mean sum of square of genotype; MS_e is mean sum of square of the error and r is number of replications. The coefficients of variation were calculated according to the below equations:

$$\text{Phenotypic coefficient of variation PCV} = \frac{\sigma^2 p}{x} \times 100$$

$$\text{Genotypic coefficient of variation GCV} = \frac{\sigma^2 g}{x} \times 100$$

where: σp = phenotypic standard deviation, σg = genotypic standard deviation, x = grand mean for character under study.

2.6. Broad-Sense Heritability (H^2) and Genetic Correlations as Computed by META-R

The broad-sense heritability (H^2) of a trait is the percentage of genotypic to phenotypic variance at an individual environment and was categorized as low (<30%), moderate (30–60%) and high (>60%) as described by Robinson et al. [24]. The H^2 was calculated for quantitative traits according to Alvarado et al. [25] as:

$$H^2 = \sigma^2 g / \left(\sigma^2 g + \sigma^2 \varepsilon / nRep \right)$$

where $\sigma^2 g$ and $\sigma^2 \varepsilon$ are the genotype and error variance components, respectively, and nRep is the number of replicates.

When there is a combined genotype by environment (G × E) interaction, broad-sense heritability was estimated for quantitative traits as:

$$H^2 = \frac{\sigma^2 g}{\sigma^2 g + \sigma^2 g\varepsilon / nLoc + \sigma^2 \varepsilon / (nLoc \times nRep)}$$

where the new term $\sigma^2 g\varepsilon$ is the G × E interaction variance component and nLoc is the number of environments in the analysis.

3. Results

3.1. Analysis of Variance and Mean Performance of the Accessions

The analysis of variance (ANOVA) revealed highly significant differences ($p \leq 0.01$) among the accessions for most quantitative traits (Table 2). A wide variability was observed for DTF10, DTF50, DTF90, PH, NCPP and NSPC. DTF10 ranged from 60 to 102 days with a mean value of 77 days, DTF50 ranged from 63 to 107 days with a mean value of 81 days and DTF90 ranged from 65 to 109 days with a mean value of 84 days. The PH ranged from 29.6 to 144.7 cm, NCPP ranged from 10.8 to 87.1 and NSPC ranged from 39.4 to 96.9 with

mean values of 87.15 cm, 94.2 cm and 64.5 cm, respectively. The number of seeds per plant ranged from 101 to 3481 (Figure 2) and NPB ranged from 5 to 17 with a mean value of 10.8 (Table 3).

Table 2. Analysis of variance for quantitative traits of 60 noug accessions.

Traits	Sources of Variation				Coefficient of Variation (%)
	MSG (59)	**MSE (1)**	**MSG*E (59)**	**MSR (120)**	
DTF10	179.2 ***	18,357.5 ***	22.1 *	14.4	4.9
DTF50	188.0 ***	20,813.4 ***	26.9 ***	12.6	4.4
DTF90	185.8 ***	21,606.8 ***	33.4 ***	12.7	4.2
NNPP	4.7 ***	97.3 ***	0.7 **	0.4	6.1
PH	610.0 ***	91.2	370.1	290.4	18.1
NPB	9.0 ***	18.5 *	4.8	3.9	18.1
NCPP	243.5	7695.4 ***	235	188	35.1
LL	0.6 *	16.3 ***	0.4	0.4	24.3
LW	1.2	3.2	1.2	1.0	33.4
NSPC	103.7 *	20,419.2 ***	104.0 *	63.0	12.3

*, ** and *** = significant at 0.05, 0.01 and 0.001 significance level, respectively, MS_G = mean square due to genotypes, MS_E = mean square due to environment, MS_{G*E} = mean square due to G × E interactions, MS_R = mean square due to residual, numbers in parenthesis = degrees of freedom, DTF10 = days to 10% flowering, DTF50 = days to 50% flowering, DTF90 = days to 90% flowering, PH = plant height, NNPP = number of nodes per plant, NCPP = number of capitula per plant, NPB = Number of primary branches, LL = leaf length, LW = leaf width, NSPC = Number of seeds per capitulum.
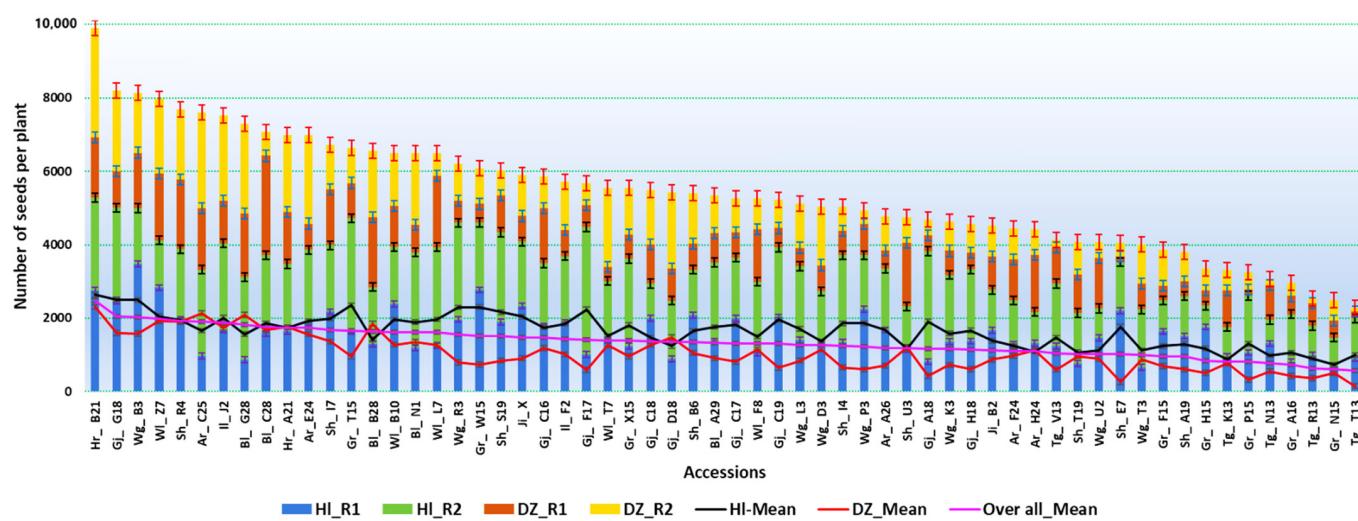

Figure 2. Over all mean number of seeds per plant across locations. Hl_R1 = Holeta Replication 1, Hl_R2 = Holeta Replication 2, DZ_R1 = Debre Zeit Replication 1, DZ_R2 = Debre Zeit Replication 2, Hl_Mean = Holeta Mean, DZ_Mean = Debre Zeit–Mean.

Chi-square test-based associations between the two environments each having two replications and the 60 accessions for desirable characters of the six qualitative traits are given in Table 3. The grouping of the 60 accessions, as revealed using Tukey method at 95% confidence level for the desirable qualitative characters is provided in Table S2. These accessions formed 5, 9, 14, 7 and 9 groups based on the desirable characters that showed significant positive correlation with number of seeds per plant (NSPP; a seed yield related trait) in five of the six qualitative traits, i.e., large flower size, green leaf, green stem, small capitula and course leaf margins. The percentage of large flowers (FS-5) varied from 7% (accession *Sh_I7*) to 73% (accession *Tg_R13*), green leaf (LC-3) from 53% (accession *Tg_T13*) to 99% (accession *Wl_F8*), green stem (SC-1) from 12% (accession *Tg_K13*) to 65% (accession *Ar_H24*), small capitula (CS-1) from 0 (accessions *Gr_H15*, *Tg_K13*, *Tg_R13*, *Tg_T13* and *Tg_V13*) to 79% (accession *Sh_E7*) and course leaf margins (LM-5) from 2% (accession *Sh_E7*) to 74% (accession *Gj_C17*) (Table 3 and Table S2).

Table 3. Chi-square (χ^2) test (degrees of freedom = 177) for association of environments and accessions (two environments, two replications each environment and sixty accessions) for desirable characters of the six qualitative traits (first four columns). Observed (Obs.) and expected (Exp.) frequencies and contribution to χ^2 for accessions with the lowest and highest sum of observed frequencies (percent) across the two sites are shown (last seven columns).

Character	Pearson χ^2	LR χ^2	*p*-Value	Acc		DZ_R1	DZ_R2	HL_R1	HL_R2	All
CS-1	1629.7	1802.6	<0.001	Gj_D18	Obs. freq.	0	2	6	0	8
					Exp. freq.	1.0	1.8	2.7	2.5	
					χ^2	1.0	0.01	4.1	2.5	
				Sh_E7	Obs. freq.	78	71	78	90	317
					Exp. freq.	41.0	70.0	106.5	99.5	
					χ^2	33.4	0.01	7.6	0.9	
FS-5	1162.3	1251.9	<0.001	Sh_I7	Obs. freq.	16	0	5	8	29
					Exp. freq.	4.7	6.5	8.9	8.9	
					χ^2	27.1	6.5	1.7	0.1	
				Tg_R13	Obs. freq.	55	48	94	95	292
					Exp. freq.	47.4	65.7	89.3	89.6	
					χ^2	1.22	4.79	0.25	0.33	
LC-3	227.9	231.1	<0.01	Tg_T13	Obs. freq.	44	62	44	62	212
					Exp, freq.	51.3	50.8	55.6	54.3	
					χ^2	1.0	2.4	2.4	1.1	
				Wg_T3	Obs. freq.	100	100	100	100	400
					Exp. freq.	96.8	95.9	104.9	102.4	
					χ^2	0.1	0.2	0.2	0.1	
LM-5	804.9	881.8	<0.001	Sh_E7	Obs. freq.	0	6	3	0	9
					Exp. freq.	2.3	2.0	2.3	2.3	
					χ^2	2.3	7.9	0.2	2.3	
				Gj_C17	Obs. freq.	84	62	88	62	296
					Exp. freq.	76.79	66.08	77.43	75.69	
					χ^2	0.68	0.25	1.44	2.48	
SC-1	817.1	926.6	<0.001	Tg_K13	Obs. freq.	8	33	10	0	51
					Exp. freq.	12.3	13.0	12.6	13.1	
					χ^2	1.5	30.7	0.5	13.1	
				Ar_H24	Obs. freq.	70	64	72	53	259
					Exp. freq.	62.5	66.1	64.1	66.7	
					χ^2	0.9	0.1	1.0	2.7	
SH-3	706.6	772.3	<0.001	Wg_B3	Obs. freq.	8	17	4	23	52
					Exp. freq.	13.1	13.1	13.1	12.7	
					χ^2	2.0	1.1	6.3	8.4	
				Wl_F8	Obs. freq.	90	86	52	56	284
					Exp. freq.	71.8	71.6	71.5	69.1	
					χ^2	4.6	2.9	5.3	2.5	

LR χ^2 = likelihood ratio chi-square; Acc = Accession; Obs freq = observed frequency; Exp. Freq. = expected frequency; χ^2 = contribution to chi-square; DZ_R1; Debre Zeit replication-1; DZ_R2; Debre Zeit replication-2; HL_R1; Holeta replication-1; HL_R2; Debre Zeit replication-2. CS-1: Capitulum size = small; FS-5: Flower size = large; LC-3: Leaf color = green; LM-5: Leaf margin = course; SC-1: Stem color = green; SH-3: Stem hairiness = medium.

3.2. Genotypic and Phenotypic Correlation

The genotypic and phenotypic correlations analyzed based on data from the 60 accessions between all pairs of quantitative traits are provided in Table 4. The traits; DTF10, DTF50, DTF90 and NNPP showed highly significant positive genotypic and phenotypic correlation between them. NCPP showed highly significant negative genotypic correlation with all quantitative traits except with number of primary branches (NPB) and leaf width (LW), which were highly significant positive phenotypic correlation. Interestingly, NSPC showed low but significant positive phenotypic correlation with NCPP.

Table 4. Genotypic (below diagonal) and phenotypic correlation (above diagonal) coefficient among the 10 quantitative traits.

Traits	DTF10	DTF50	DTF90	NNPP	PH	NPB	NCPP	LL	LW	NSPC
DTF10		0.91 ***	0.94 ***	0.59 ***	0.22	0.38	−0.04	0.23	−0.06	−0.23 ***
DTF50	0.99 ***		0.95 ***	0.57 ***	0.23 *	0.36	−0.03	0.24	−0.08	−0.30 **
DTF90	0.99 ***	0.99 ***		0.57 ***	0.25 *	0.35	−0.10	0.21	−0.06	−0.30 ***
NNPP	0.68 ***	0.72 ***	0.74 ***		0.41 ***	0.73 ***	0.06	0.32 *	−0.01	−0.09 *
PH	0.25 ***	0.38	0.39	0.44 ***		0.50	0.27 *	0.65 ***	0.63 ***	−0.11
NPB	0.51	0.55 ***	0.52 ***	0.89 ***	0.29 ***		0.43	0.53 **	0.37	0.12
NCPP	0.11 ***	0.60 ***	0.24 ***	0.36 ***	0.65 ***	0.99 ***		0.35 ***	0.25 ***	0.24 *
LL	0.38	0.57 ***	0.47	0.48 ***	0.72 ***	0.48 ***	0.33 ***		0.66 ***	−0.11
LW	−0.14 ***	0.01	−0.03 ***	−0.26 ***	0.50	−0.21 ***	0.40 ***	0.42 ***		−0.03
NSPC	−0.36 ***	−0.61 ***	−0.54 ***	−0.05 ***	−0.08 ***	0.41 ***	−0.77 ***	−0.04 ***	0.49 ***	

*, ** and *** = significant at 0.05, 0.01 and 0.001 significance level respectively, DTF10 = days to 10% flowering, DTF50 = days to 50% flowering, DTF90 = days to 90% flowering, PH = plant height, NNPP = number of nodes per plant, NCPP = number of capitulum per plant, NPB = number of primary branches, LL = leaf length, LW = leaf width, NSPC = number of seeds per capitulum.

3.3. Estimation of Variance Components and Broad-Sense Heritability

As suggested by Sivasubramanian and Madhavamenon [23], medium (10–20%) genotypic coefficient of variation (GCV) was recorded only for LW (12.35%) whereas other 9 quantitative traits had low (<10%) GCV values. The lowest GCV values were recorded for NCPP (0.45%) followed by NSPC (0.53%), DTF50 (0.55%), DTF90 (0.56%) and DTF10 (0.69%) respectively. On the other hand, high (>20%) phenotypic coefficient of variation (PCV) values were recorded for LW (36.74%) and LL (21.74%), whereas other traits had low PCV, in which the lowest PCV was recorded for DTF90 (0.63%) (Table 5).

Table 5. Estimates of phenotypic values, variance components and broad-sense heritability (H^2) for the 10 quantitative traits measured in 60 noug accessions.

Traits	Minimum	Maximum	Mean ± SD	Genotypic Variance	Phenotypic Variance	H2 (%)	GCV (%)	PCV (%)
DTF10	60.00	102.0	77.45 ± 3.80	0.2914	0.3413	85.38	0.69	0.75
DTF50	63.00	107.0	81.00 ± 3.60	0.2031	0.2823	71.94	0.55	0.65
DTF90	65.00	109.0	84.07 ± 3.60	0.2195	0.2844	77.14	0.56	0.63
NNPP	6.60	13.82	10.05 ± 0.61	0.4468	0.5676	78.71	6.65	7.49
PH	29.60	144.7	87.15 ± 13.71	0.1586	0.5575	28.45	1.02	1.91
NPB	5.10	16.82	10.85 ± 1.97	0.1611	0.4234	38.05	3.69	5.99
NCPP	10.78	87.08	94.22 ± 17.04	0.1808	0.4691	38.54	0.45	0.72
LL	1.04	4.66	2.54 ± 0.62	0.0506	0.3050	16.59	8.85	21.74
LW	1.00	6.00	3.01 ± 1.01	0.1383	1.2228	11.31	12.35	36.73
NSPC	39.40	96.92	64.52 ± 7.94	0.1154	0.3649	31.62	0.53	0.93

DTF10 = days to 10% flowering, DTF50 = days to 50% flowering, DTF90 = days to 90% flowering, PH = plant height, NNPP = number of nodes per plant, NCPP = number of capitula per plant, NPB = Number of primary branches, LL= leaf length, LW = leaf width, NSPC = number of seeds per capitulum, SD = standard deviation; PCV = phenotypic coefficient of variation, GCV = genotypic coefficient of variation, H^2 = broad-sense heritability.

Broad-sense heritability (H^2) estimates of the quantitative traits are given in Table 5. According to Robinson et al. [24] and Alvarado et al. [25] classification of broad-sense heritability, low heritability (<30%) was recorded for PH (28.45%), LL (16.59%) and LW (11.31%); medium heritability (30–60%) was recorded for NPB (38.05%), NCPP (38.54%) and NSPC (31.62%) and high heritability (>60%) was recorded for DTF10 (85.38%), DTF50 (71.94%), DTF90 (77.14%) and NNPP (78.71%) (Table 5).

3.4. Cluster Analysis and Principal Component Analysis

Euclidian distance matrices between pairs of the 60 accessions were generated for quantitative and qualitative traits separately. The two distance matrices were used for UPGMA based hierarchical cluster analysis (Figure 3). The cluster analysis grouped the

60 noug accessions into five distinct clusters; three major clusters (3,4,5) and two minor cluster (1,2) for quantitative traits (Figure 3A) and four distinct clusters; three major clusters (2,3,4) and one minor cluster (1) for qualitative traits (Figure 3B). Principal component (PC) biplots of quantitative and qualitative traits were generated for the first two PCs (Figure 4). The first two PCs, which contributed to 66.3% of the variation among the quantitative traits (Figure 4A) and 53.6% of the variation among the qualitative traits (Figure 4B) tested, were the most significant PCs that explained majority of the variations among traits.
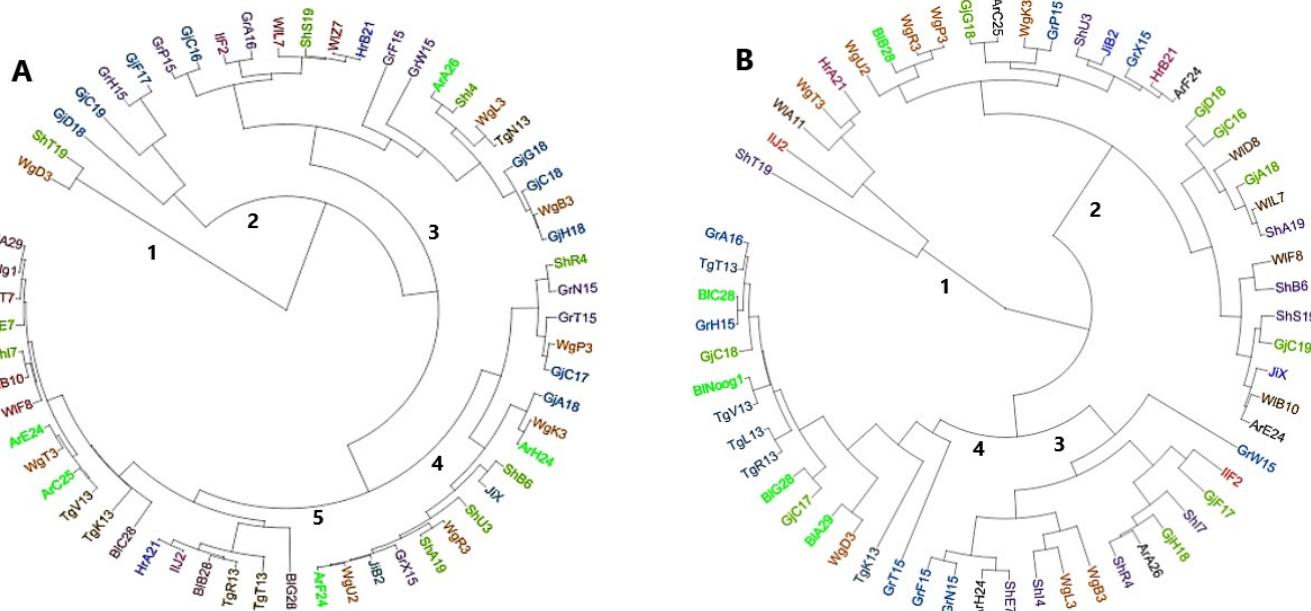
Figure 3. Dendrogram of 60 noug germplasm accessions constructed with UPGMA using Euclidian distance showing genetic relationships between qualitative (**A**) and quantitative (**B**) traits. Note: the underscore in accession codes were excluded for simplicity (e.g., the full code of "ShS19" in this figure is "Sh_S19").
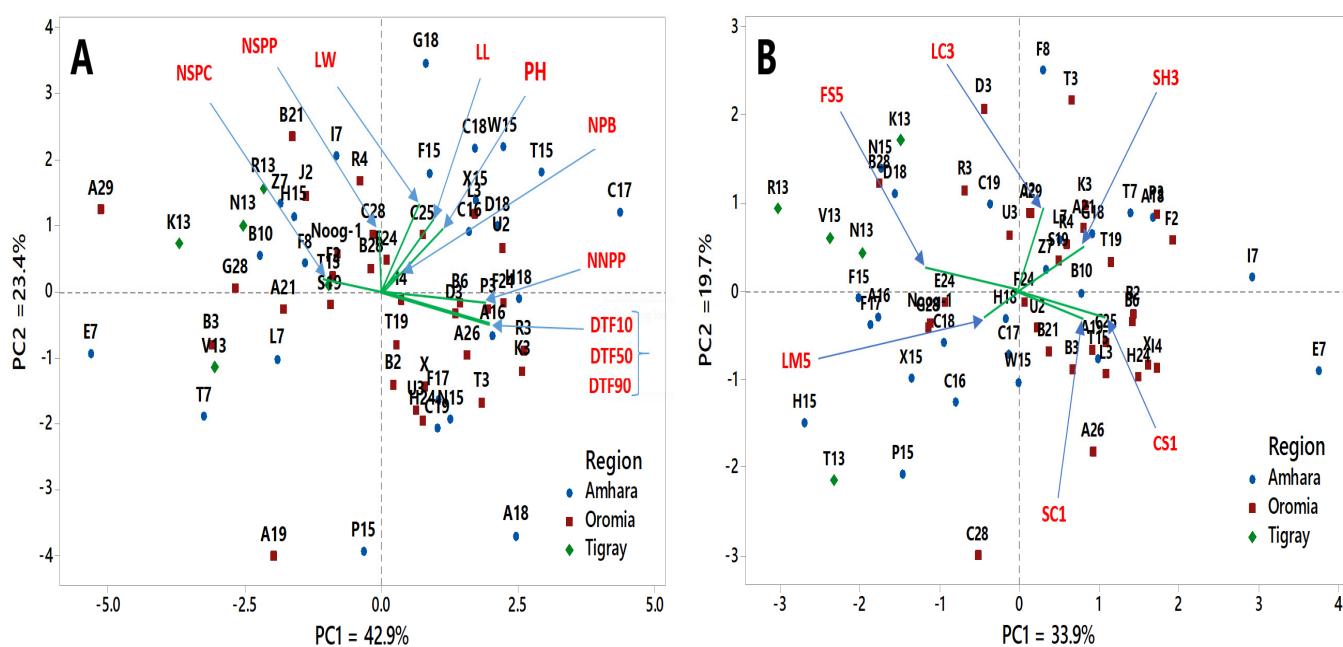
Figure 4. PCA biplots of the first two principal components (PC1 and PC2) combining the score plots of the 60 noug accessions showing their grouping pattern and the corresponding loading plots based on (**A**) the 10 quantitative traits and

(**B**) the frequency of the desirable characteristics of the six qualitative traits. (**A**) DTF10 = days to 10% flowering, DTF50 = days to 50% flowering, DTF90 = days to 90% flowering, PH = plant height, NNPP = number of nodes per plant, NCPP = number of capitulum per plant, NPB = number of primary branches, LL = leaf length, LW = leaf width, NSPC = number of seeds per capitulum. (**B**) CS-1: Capitulum size = small; FS-5: Flower size = large; LC-3: Leaf color = green; LM-5: Leaf margin = course; SC-1: Stem color = green; SH-3: Stem hairiness = medium. Note: Only characters to the right of the underscore in Accession codes were used for sake of simplicity (e.g., the full accession code of "A19" in this figure = "Sh_A19") (see Table S1).

4. Discussion

Analysis of qualitative and quantitative agro-morphological traits is very important for efficient representation of crops' genetic diversity in its ex-situ conserved gene pools as well as for identification of suitable germplasm for use in plant breeding programs. Quantitative traits are controlled by many genes –often with minor effects– and highly influenced by environmental factors. The effects of environmental factors on crop phenotypes are estimated by the magnitude of the differences between the genotypic and phenotypic coefficients of variation; small differences reflect a large genotypic effect, whereas large differences suggest a high environmental influence [26,27]. In the present study, the analysis of variance revealed significant differences among the accessions for most of the traits studied. Panda and Sial [28] also reported a wide range of variation for most of the traits they investigated in noug grown in India.

The observed phenotypic variability among the tested noug accessions suggests the potential of the crop for genetic improvement through selection and crossbreeding of superior genotypes. The genotypic coefficient of variation (GCV) and phenotypic coefficient of variation (PCV) analysis help to group the phenotypic variability into heritable and non-heritable components depending on the level of differences between the two parameters and hence is useful for the identification of superior genotypes for crossbreeding. As suggested by Sivasubramanian and Madhavamenon [23], moderate GCV was recorded only for LW (12.35%) whereas other traits had low (<10%) GCV. On the other hand, high (>20%) PCV was recorded for LW (36.74%) and LL (21.74%) whereas the lowest PCV was recorded for DTF90 (0.63%). The result of the present study is in agreement with the results reported in Amsalu [29] where low PCV and GCV values were recorded for days to flowering, days to maturity, plant height, seed yield and thousand seed weight (TSW); and partly in agreement with Bhoite et al. [30], who found low PCV and GCV values for DTF50 and PH and high PCV and GCV values for NPB, NCPP and NSPC. Low PCV and GCV values were also obtained for DTF50 and days to maturity by Ahmad et al. [31] and Tiwari et al. [32] in agreement with the present results.

Heritability values are very important in predicting the expected progress to be attained through the process of selection. The GCV accompanied by heritability estimates provides an estimate of the amount of genetic advance to be expected through phenotypic selection [33]. Among the 10 quantitative traits included in this study, the highest broad-sense heritability (>60%) was recorded for DTF10 (85.38%) followed by NNPP (78.71%) and the lowest heritability (<30%) was recorded for LW (11.31%) and LL (16.59%) as per Robinson et al. [24] and Alvarado et al. [25] classification of broad-sense heritability.

As results of the present study indicates, LL and LW showed high phenotypic variance and low heritability indicate the environmental effect on their expression is high and hence are not suitable for selecting germplasm of interest. Whereas, low PCV and GCV values and high heritability for many of the traits such as days to flowering and NNPP indicate low influence of the environment on the variation of these traits than the other traits. Early maturing accessions are preferred in areas having frequent drought due to inadequate rainfall. In the present study about 21.7% of the noug accessions matured earlier than 88 days. For example, *Ar_C24* and *Ar_E24* (40% of the accessions from Arsi); *Bl_A29*, *Bl_B28*, *Bl_C28*, *Bl_G28* and *Bl_Noog1* (all accessions from Bale); *Gj_C18*, *Gj_D18*, *Gj_F17* and *Gj_G18* (44% of the accessions from Gojjam); *Gn_F15*, *Gn_H15*, *Gn_P15* and *Gn_W15* (50% of the accessions from Gondar); *Sh_A19*, *Sh_B6*, *Sh_E7*, *Sh_I4* and *Sh_T19* (55% of accessions from Shewa) and *Wg_B3* and *Wg_L3* (25% of the accessions from Wolega) were

early maturing but showing high seed yield. Hence, these accessions are preferred in areas where the rainy season is relatively short. It is therefore very important to consider days to flowering together with NCPP and NSPC in noug breeding programs as days to flowering is related to earliness, a very important trait in the tropics where there is short seasonal rainfall and NCPP and NSPC are related to oil yield in noug. Number of seeds per plant (a seed yield trait) is a product of NCPP and NSPC and hence these traits directly contribute to seed yield, which in turn linked to oil yield.

As yield and quality are complex multigenic traits, knowledge of the extent of the genotypic and phenotypic correlation between traits is useful in the identification of most important plant characteristics contributing to desirable traits. The existence of strong positive phenotypic and genotypic correlation between different traits is useful for indirect selection of complex traits such as seed yield and oil content in noug. Highly heritable traits that have higher genetic correlation with seed yield and oil content could be used for selecting desirable genotypes for use in noug breeding program. The broad sense heritability of DTF10, DTF50, DTF90 and NNPP were 85.38%, 71.94%, 77.14% and 78.71%, respectively and hence they are highly heritable. Research conducted by Weyessa [7] and Tiwari et al. [32] also showed high heritability for days to flowering., which is highly important trait in tropical regions where drought is a problem. The ultimate goal of noug breeding as an oilseed crop is to attain increased seed and oil yields per unit area. DTF10, DTF50, DTF90 and NNPP showed highly significant negative phenotypic and genotypic correlations with number of seeds per capitulum, which is a seed yield related trait. This suggests that early maturing types produce more seeds per capitulum in average as compared to late maturing types. Hence, selecting such plants having larger number of capitula per plant is an interesting approach to increase seed and oil yield in noug.

In the principal component analysis, the first two principal components (PCs) explained 66.3% of the total variation in the case of the quantitative traits (Figure 4A) and 53.6% in the case of the qualitative traits (Figure 4B). These values are lower than those reported in Pulate et al. [34] in similar study in noug where the first two PCs explained 94% of the total variation for quantitative and 60.8% of the total variation for qualitative traits, respectively. Our analysis showed the absence of clear grouping pattern of the accessions for both quantitative and qualitative data, as the accessions were spread along both PC axes. However, it was clear that the accessions in the far right and far left were separated along PC1 axis, which was highly influenced by days to flowering, number of nods per plants and number of seeds per capitulum as depicted by the loadings of the traits. For example, the cluster of nine accessions (*A16, A26, B6, D3, F24, H18, K3, R3* and *P3*; see Table S1 for full accession codes) were formed mainly due to their late maturing character and higher number of nodes per plant although they were from distant areas in Amhara and Oromia regions.

The five accessions from Tigray can be generally regarded as early maturing and are top ranking in number of seeds per capitulum; and hence can serve as sources of alleles for these characteristics in crossbreeding. As depicted in Figure 4A, days to maturity does not have strong influence on NSPC and NCPP (yield component traits) and hence it is possible to develop high yielding early maturing noug cultivars. Accession *Hr_B21* (shown as *B21* in Figure 4) from West Harerge, Oromia region is an excellent example for this as it is an early maturing type, which came on top in terms of average number of seeds per plant (Table S1). Interestingly, this accession had above average thousand seed weight (TSW) (4.2 gm; the average being 3.6 gm) although it had slightly lower than average oil content (35%; the average being 38%) (Table S1). The results clearly show that there is no strong trade-off between NSPP and seed weight. Seed size (weight) is a highly heritable trait in noug. Hence, noug breeding based on NSPP and seed weight can lead to improved seed yield. In the case of qualitative traits, PC1 of PCA explained 33.9% of the total variation mainly due to the influence of higher proportion of CS1, SC1, FS5 and LM5, but flower size and capitulum size were the most important in separating genotypes along the PC1 axis. Similar to the case of quantitative traits, accessions from Tigray showed similarity

in having higher proportion of large flower size (FS5) whereas accessions from Amhara and Oromia showed wider variation. However, the number of seeds per plant for Tigray accessions is intermediate compared to the other accessions (Table S1).

The 60 noug accessions were grouped into five clusters based on quantitative (Figure 3A) and four clusters based on qualitative (Figure 3B) traits. However, the clustering pattern did not follow the geographic proximity of the collection sites of the accessions except in few cases where smaller groups were formed within a cluster. The clustering pattern based on quantitative traits was somewhat different from that of qualitative traits suggesting a poor correlation between the two groups of traits. Such difference in clustering pattern of different *Guizotia* species and populations was also observed in molecular markers based studies (Random amplified polymorphic DNA vs Amplified Fragment Length Polymorphism) [15,35]. A recent study using transcriptome based SNP markers revealed a weak population structure in noug grown in Ethiopia due to population admixture, which is mainly a result of strong gene flow between populations via pollen and a gradual nation-wide germplasm exchange [36]. The poor clustering pattern of the accessions in terms of geographic proximity in the present study also suggest strong gene flow between populations throughout the country.

Among the sixty accessions used in this study, *Gj_C17*, *Sh_I4* and *Gr_F15* are top ranking in oil content with 54%, 52% and 51%, respectively (Table S1; [37]). The codes used for these accessions in Geleta et al. [37] were *Gj-2*, *Sh-2* and *Gr-1*, respectively. *Gj_C17* is late maturing whereas the other two have intermediate maturity (Table S1). However, they had slightly lower than average number of seeds per plant (Table S1). Other interesting accessions are *Gj_G18* (from Amhara region) and *Tg-R13* (from Tigray region). *Gj_G18* is characterized by intermediate maturity (DTF90 = 83), above average oil content (42%), average TSW (3.9 gm) and high NSPP (2327 seeds). Whereas *Tg-R13* is characterized by early maturity (DTF90 = 71), high oil content (49%), high TSW (4.7 gm) and low NSPP (965 seeds). Overall, the average oil content and NSPP of early, intermediate and late maturing groups did not show significant differences, thereby suggesting the absence of strong association between maturity time, oil content and seed yield in noug. Hence, crossbreeding of selected genotypes from these accessions should be seriously considered for improving noug in both seed and oil yield.

5. Conclusions

The present study revealed the presence of significant phenotypic and genotypic variation in Ethiopian noug gene pool as reflected based on both qualitative and quantitative traits. The lack of clear clustering pattern of the accessions with regard to their geographic proximity suggests strong gene flow between the landrace populations across a wide geographic region. The present study showed that days to flowering is a highly heritable trait whereas NCPP and NSPC are moderately heritable in noug. Since days to flowering (maturity) does not have strong influence on number of seeds per plant, it is possible to develop high yielding early maturing noug cultivars suitable for areas where there is a shortage of rainfall during plant growing season. Although there are significant positive and negative correlations between several traits included in the present study, there is no strong trade-off among the most important traits, such as earliness, seed yield and oil content. Hence, crossbreeding of selected genotypes with desirable traits from various landrace populations should be considered in noug breeding programs. During selection of genotypes, strong emphasis should be given for traits such as NCPP, NSPC, oil content and days to maturity for improvement of seed and oil yield through crossbreeding.

References

1. Dagne, K. Meiosis in interspecific hybrids and genomic interrelationships in *Guizotia* Cass.(Compositae). *Hereditas* **1994**, *121*, 119–129. [CrossRef]
2. Geleta, M.; Asfaw, Z.; Bekele, E.; Teshome, A. Edible oil crops and their integration with the major cereals in North Shewa and South Welo, Central Highlands of Ethiopia: An ethnobotanical perspective. *Hereditas* **2002**, *137*, 29–40. [CrossRef]
3. Geleta, M.; Bryngelsson, T. Population genetics of self-incompatibility and developing self-compatible genotypes in niger (*Guizotia abyssinica*). *Euphytica* **2010**, *176*, 417–430. [CrossRef]
4. Baagoe, J. The genus *Guizotia* (Compositae). A Taxonomic Revision. *Bot. Tidsskr.* **1974**, *69*, 1–39.
5. Geleta, M.; Ortiz, R. The importance of *Guizotia abyssinica* (niger) for sustainable food security in Ethiopia. *Genet. Resour. Crop Evol.* **2013**, *60*, 1763–1770. [CrossRef]
6. Seegeler, C.J.P. *Oil Plants in Ethiopia: Their Taxonomy and Agricultural Significance*; Center of Agricultural Publishing and Documentation: Wageningen, The Netherlands, 1983; pp. 122–146.
7. Bulcha, W. Guizotia abyssinica *(Lf) Cass. Record from PROTA4U*; van der Vossen, H.A.M., Mkamilo, G.S., Eds.; PROTA (Plant Resources of Tropical Africa/Ressources végétales de l'Afrique tropicale): Wageningen, The Netherlands, 2007.
8. Riley, K.; Belayneh, H. *Niger Seed:* Guizotia abyssinica *Cass, in Oil Crops in the World: Their Breeding and Utilization*; McGraw Hill Publishing Company: New York, NY, USA, 1989; pp. 394–403.
9. Getinet, A.; Sharma, S. *Niger,* Guizotia abyssinica *(Lf) Cass*; Bioversity International: Rome, Italy, 1996; Volume 5.
10. *Oilseeds Research Strategy 2016–2023*; Ethiopian Institute of Agricultural Research: Addis Ababa, Ethiopia, 2017.
11. Petros, Y. Genetic Diversity and Oil Quality of Guizotia Cass (Asteraceae). Ph.D. Thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden, 2008.
12. CSA, R. *The Federal Democratic Republic of Ethiopia Report on Area and Production of Major Crops*, Addis Ababa, Ethiopia, 2016.
13. Geleta, M.; Bryngelsson, T.; Bekele, E.; Dagne, K. Assessment of genetic diversity of *Guizotia abyssinica* (Lf) Cass (Asteraceae) from Ethiopia using amplified fragment length polymorphism. *Plant Genet. Resour.* **2008**, *6*, 41–51. [CrossRef]
14. Dempewolf, H.; Tesfaye, M.; Teshome, A.; Bjorkman, A.; Andrew, R.; Scascitelli, M.; Black, S.; Bekele, E.; Engels, J.M.M.; Cronk, Q.; et al. Patterns of domestication in the Ethiopian oil-seed crop noug (*Guizotia abyssinica*). *Evol. Appl.* **2015**, *8*, 464–475. [CrossRef] [PubMed]
15. Geleta, M. Genetic Diversity, Phylogenetics and Molecular Systematics of Guizotia Cass (Asteraceae). Ph.D. Thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden, 2007.
16. Geleta, M.; Bryngelsson, T.; Bekele, E.; Dagne, K. AFLP and RAPD analyses of genetic diversity of wild and/or weedy Guizotia (Asteraceae) from Ethiopia. *Hereditas* **2007**, *144*, 53–62. [CrossRef]
17. Teklewold, A.; Wakjira, A. Seed filling and oil accumulation in noug [*Guizotia abyssinica* (LF) Cass.]. *Sinet Ethiop. J. Sci.* **2004**, *27*, 25–32.
18. Adarsh, M.N.; Kumari, P.; Devi, S. A review of *Guizotia abyssinica*: A multipurpose plant with an economic prospective. *J. Ind. Pollut. Control* **2014**, *30*, 277–280.
19. Kandel, H.J.; Porter, P.M.; Johnson, B.L.; Henson, R.A.; Hanson, B.K.; Weisberg, S.; LeGare, D.G. Plant population influences niger seed yield in the northern Great Plains. *Crop Sci.* **2004**, *44*, 190–197. [CrossRef]
20. Hill, W.G. Understanding and using quantitative genetic variation. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 73–85. [CrossRef]
21. Xu, Y.; Boddupalli, P.; Zou, C.; Lu, Y.; Xie, C.; Zhang, X.; Prasanna, B.M.; Olsen, M.S. Enhancing genetic gain in the era of molecular breeding. *J. Exp. Bot.* **2017**, *68*, 2641–2666. [CrossRef]
22. Genet, T.; Belete, K. Phenotypic diversity in the Ethiopian noug germplasm. *Afr. Crop Sci. J.* **2000**, *8*, 137–143. [CrossRef]

23. Sivasubramanian, S.; Menon, M. Heterosis and inbreeding depression in rice. *Madras Agric. J.* **1973**, *60*, 1139–1140.

24. Robinson, H.F.; Comstock, R.E.; Harvey, P.H. Estimates of Heritability and the Degree of Dominance in Corn 1. *Agron. J.* **1949**, *41*, 353–359. [CrossRef]

25. Alvarado, G.; Rodríguez, F.M.; Pacheco, A.; Burgueño, J.; Crossa, J.; Vargas, M.; Pérez-Rodríguez, P.; Lopez-Cruz, M.A. META-R: A software to analyze data from multi-environment plant breeding trials. *Crop. J.* **2020**, *8*, 745–756. [CrossRef]

26. Buzatti, R.S.D.O.; Pfeilsticker, T.R.; Muniz, A.C.; Ellis, V.A.; De Souza, R.P.; Lemos-Filho, J.P.; Lovato, M.B. Disentangling the environmental factors that shape genetic and phenotypic leaf trait variation in the tree Qualea grandiflora across the Brazilian savanna. *Front. Plant Sci.* **2019**, *10*, 1580. [CrossRef] [PubMed]

27. Ji, M.; Deng, J.; Yao, B.; Chen, R.; Fan, Z.; Guan, J.; Li, X.; Wu, F.; Niklas, K.J. Ecogeographical variation of 12 morphological traits within Pinus tabulaeformis: The effects of environmental factors and demographic histories. *J. Plant Ecol.* **2017**, *10*, 386–396.

28. Panda, S.; Sial, P. Assessment of existing genetic variability and yield component analysis in noug (*Guizotia abyssinica* (L.f.) Cass). *Indian J. Innov. Dev.* **2012**, *1*, 511–514.

29. Amsalu, F. Genetic Variability Correlation and Path Coefficient Analysis of Niger Seed (*Guizotia abyssinica* [Lf] Cass.) Genotypes. *Int. J. Res. Stud. Agric. Sci.* **2020**, *6*, 8–16.

30. Bhoite, K.D.; Kusalkar, D.V.; Patil, H.M. Genetic studies of yield and its component of Niger [(*Guizotia abyssinica* (Lf) Cass.] in Western Ghat Zone of Maharashtra. *J. Pharmacogn. Phytochem.* **2020**, *9*, 1182–1184.

31. Ahmad, E.; Singh, M.N.; Paul, A.; Ansari, A.M.; Sigh, D.N. Genetic studies of yield and yield component of niger (*Guizotia abyssinica* Cass.) in rainfed condition of western plateau of Jharkhand. *Int. J. Trop. Agric.* **2016**, *34*, 1127.

32. Tiwari, V.N.; Ahirwar, A.D.; Rai, G.K. Estimation of genetic parameters of variability for yield and its attributing traits in niger [*Guizotia abyssinica* (LF) Cass.]. *Plant Arch.* **2016**, *16*, 157–158.

33. Wright, S. Correlation and causation. *J. Agric. Res.* **1921**, *20*, 557–580.

34. Pulate, S.C.; Patil, H.S.; Patil, M.R. Multivariate analysis of genetic divergence among niger genotypes in relation to seed oil quality traits. *Bioscan* **2013**, *8*, 829–833.

35. Geleta, M.; Bryngelsson, T.; Bekele, E.; Dagne, K. Comparative analysis of genetic relationship and diagnostic markers of several taxa of Guizotia Cass (Asteraceae) as revealed by AFLPs and RAPDs. *Plant Syst. Evol.* **2007**, *265*, 221–239. [CrossRef]

36. Tsehay, S.; Ortiz, R.; Johansson, E.; Bekele, E.; Tesfaye, K.; Hammenhag, C.; Geleta, M. New Transcriptome-Based SNP Markers for Noug (*Guizotia abyssinica*) and Their Conversion to KASP Markers for Population Genetics Analyses. *Genes* **2020**, *11*, 1373. [CrossRef]

37. Geleta, M.; Stymne, S.; Bryngelsson, T. Variation and inheritance of oil content and fatty acid composition in niger (*Guizotia abyssinica*). *J. Food Compos. Anal.* **2011**, *24*, 995–1003. [CrossRef]