Received: 28 Oct 2022 | Revised: 12 Apr 2023 | Accepted: 18 Apr 2023

Electronic Journal of Plant Breeding

Research Note

Principal Component Analysis in advanced breeding lines of oat (*A. sativa* × *A. sterilis*)

Nagesh Bichewar[1*], A. K. Mehta[1], Kadthala Bhargava[2] and S. Ramakrishana[1]

[1]Department of Genetics and Plant Breeding, College of Agriculture, Jabalpur, JNKVV, Jabalpur-482004, Madhya Pradesh, India
[2]Department of Genetics and Plant Breeding, College of Agriculture, Rajendranagar, PJTSAU, Hyderabad-500030, Telangana, India.
***E-Mail:** nageshbichewar168@gmail.com

Abstract
One hundred advanced oat lines involving two checks were assessed for genetic variability and diversity based on 15 agro-morphological traits by using principal component analysis. Only four out of 19 principal components (PCs) exhibited more than one eigen value and 71.42 % variability. PC-1 accounted for 47.28 % of overall variability followed by PC-2, which exhibited 9.05% variability while PC-3 and PC-4 accounted for 8.15% and 6.95% variability respectively. PC3 contains most of the traits that contributes significantly to the yield hence, lines from PC3 can be selected for rapid improvement of yield. Based on PC scores, genotypes JO (IS) 12-4-9, JO (IS) 12-5-5, JO (IS) 12-5-14, JO (IS) 12-6-19, JO (IS) 12-8-2, JO (IS) 12-8-4, JO (IS) 12-8-5, JO (IS) 12-8-11 and JO (IS) 12-8-18 are the major sources of variation in the oat population and this information would be useful for selecting potential inbred lines for future oat improvement programme.

Keywords: Inbreds, Principal Component Analysis (PCA) and Oat.

Oat (*Avena sativa*) is a cereal crop which was originated in Mediterranean and Middle East region. The crop is considered as one of the oldest crop cultivated by mankind and domesticated after wheat and barley and (Murphy and Hoffman, 1992). The global area and production of oat is about 27m ha and 40 metric tonnes respectively (http://eagri.org). The Russian Federation, Canada, Finland and Poland are ranked as the top five countries for world oat production (Mohanta *et al*., 2020). In India, about 5.0 lakh ha is under oat cultivation. In *rabi*, the crop is raised mainly for fodder in north western and central parts of the country and the crop is now even expanding in the eastern regions of the country as well (Hilli and Kapoor, 2021). It is the seventh most economically significant cereal crop following corn, rice, wheat, barley, sorghum and millets. The crop, though cultivated primarily for grain purpose, is also used as feed for livestock. Under suitable climate, the crop is excellent source of hay and can be used for grazing and provide quality silage too. It is also the only cereal containing globulin or legume-like protein, avenalin, as the major (80%) storage protein (Ahmad *et al*., 2016). There is significant quantity of proteins, starch, unsaturated fatty acids and dietary fibres in whole grain of oat as well as it holds micronutrients like vitamin E, foliates, zinc, iron, selenium, copper, manganese, carotenoids, betaine, choline, sulphur consisting amino acids, lignins, lignans, and alkylresorcinols (Flander *et al*., 2007) and is unique source of polyphenols (avenanthramides) as well (Singh *et al*., 2013). As the crop is winter hardy, it is best suitable for temperate regions (Kumari and Jindal, 2019) but the stages like seed formation and maturity are sensitive to water shortage and heat (Murphy and Hoffman, 1992).

https://doi.org/10.37992/2023.1402.049

The first step in any breeding programme is to determine genetic variation for agronomically important traits among germplasm in order to identify valuable genotypes for further consideration. Morphological measurements and phenotypic characterisation can be used to assess genetic variability (Greene *et al*., 2008). Relevant statistical techniques provide suitable bio-agronomic characterisation which helps in early assessment and classification of oat genotypes, licensing plant breeders to identify and select beneficial genetic resources for straight utilization in breeding programmes (Achleitner *et al*., 2008; Rezai and Frey,1990). Principal component analysis (PCA) can be used to establish relationships between various accessions based on genetic variation in germplasm for agronomic traits.

The concept of PCA was earlier laid by Pearson (1901) and after that improved by Hotelling (1933). PCA is a vital tool in contemporary data analysis since it offers a simple, non-parametric approach for extracting information from confusing data sets. PCA offers a blueprint for minimizing a compound data set to a lower aspect, revealing hidden, simplified structures that are frequently found within complex data sets. It lowers the magnitude of the data while the majority of variation of the data set is retained. Hence, PCA is effective technique in present-day data evaluation which minimizes the data containing a greater number of correlated variables into a lesser number of new variables by combining linearly the variables that consider for majority of variation included in the earlier variables and also helps ranking the germplasm on the basis of PC scores.

The experiment was conducted during *Rabi,* 2019-20 in Randomized Block Design (RBD) in three replications with each genotype sown in three rows per replication under All India Coordinated Research Project on Forage Crops, at Seed Breeding Farm, Department of Genetics & Plant Breeding, College of Agriculture, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur (M.P.) to assess the genetic variability and diversity among 98 advanced breeding lines generated from interspecific hybridization between *A. sativa* × *A. sterilis* (F_8 generation) of oat along with two checks. Observation on fifteen agro-morphological traits such as Stem width(SW), Leaf width(LW), Days to 50% flowering (DF), Days to maturity(DM), Plant height(PH), Number of tillers per plant(NTP), Number of leaves per tiller(NLT), Flag leaf length(FLL), Internode length(INL), Flag leaf width(FLW), Number of spikelets per panicle (NSP), Number of florets per panicle (NFP), Panicle weight(PW), 1000 seed weight(1000SW), Seed yield per plant(SYP) were recorded by selecting five random plants per within each replication plot except for Days to 50% Flowering and Days to Maturity as these traits were assessed on plot basis. The genetic variability for each trait was estimated using principal component analysis. Genotypic means were used in the PCA and analysis was done by using WINDOSTAT (version 9.1). Details of advance breeding lines and checks used in the study are given in **Table 1**.

PCA for yield and yield attributing traits was carried out in 98 advanced oat lines and two checks. Out of the total 15 principal components (PCs), only four recorded eigen value greater than one and contributed about 71.42 %

Table 1. List of oat genotypes included in the study

JO (IS) 12-4-1	JO (IS) 12-5-1	JO (IS) 12-6-1	JO (IS) 12-7-3	JO (IS) 12-8-3
JO (IS) 12-4-2	JO (IS) 12-5-2	JO (IS) 12-6-2	JO (IS) 12-7-4	JO (IS) 12-8-4
JO (IS) 12-4-3	JO (IS) 12-5-3	JO (IS) 12-6-3	JO (IS) 12-7-5	JO (IS) 12-8-5
JO (IS) 12-4-4	JO (IS) 12-5-4	JO (IS) 12-6-4	JO (IS) 12-7-6	JO (IS) 12-8-6
JO (IS) 12-4-5	JO (IS) 12-5-5	JO (IS) 12-6-5	JO (IS) 12-7-7	JO (IS) 12-8-7
JO (IS) 12-4-6	JO (IS) 12-5-6	JO (IS) 12-6-6	JO (IS) 12-7-8	JO (IS) 12-8-8
JO (IS) 12-4-7	JO (IS) 12-5-7	JO (IS) 12-6-7	JO (IS) 12-7-9	JO (IS) 12-8-9
JO (IS) 12-4-8	JO (IS) 12-5-8	JO (IS) 12-6-8	JO (IS) 12-7-10	JO (IS) 12-8-10
JO (IS) 12-4-9	JO (IS) 12-5-9	JO (IS) 12-6-9	JO (IS) 12-7-11	JO (IS) 12-8-11
JO (IS) 12-4-10	JO (IS) 12-5-12	JO (IS) 12-6-10	JO (IS) 12-7-12	JO (IS) 12-8-12
JO (IS) 12-4-11	JO (IS) 12-5-13	JO (IS) 12-6-11	JO (IS) 12-7-13	JO (IS) 12-8-13
JO (IS) 12-4-12	JO (IS) 12-5-14	JO (IS) 12-6-12	JO (IS) 12-7-14	JO (IS) 12-8-14
JO (IS) 12-4-13	JO (IS) 12-5-15	JO (IS) 12-6-13	JO (IS) 12-7-15	JO (IS) 12-8-15
JO (IS) 12-4-14	JO (IS) 12-5-16	JO (IS) 12-6-14	JO (IS) 12-7-16	JO (IS) 12-8-16
JO (IS) 12-4-15	JO (IS) 12-5-17	JO (IS) 12-6-15	JO (IS) 12-7-17	JO (IS) 12-8-17
JO (IS) 12-4-16	JO (IS) 12-5-18	JO (IS) 12-6-16	JO (IS) 12-7-18	JO (IS) 12-8-18
JO (IS) 12-4-17	JO (IS) 12-5-19	JO (IS) 12-6-17	JO (IS) 12-7-19	JO (IS) 12-8-19
JO (IS) 12-4-18	JO (IS) 12-5-20	JO (IS) 12-6-18	JO (IS) 12-7-20	JO (IS) 12-8-20
JO (IS) 12-4-19	JO (IS) 12-5-1	JO (IS) 12-6-19	JO (IS) 12-8-1	Kent
JO (IS) 12-4-20	JO (IS) 12-5-2	JO (IS) 12-6-20	JO (IS) 12-8-2	JO 5

https://doi.org/10.37992/2023.1402.049

variability among the traits studied (**Table 2**). The PC-1 accounted for maximum share of total variability i.e. 47.28 % in the set of all variables whereas the remaining components showed lesser (PC-2, PC-3 and PC-4 exhibited 9.05%, 8.15% and 6.95% of total variation respectively) for the traits under study (**Table 2**). Similar findings were reported by Kumari and Jindal (2019). Scree plot (**Fig. 1**) showed the percentage of variance related with every trait by plotting a graph between eigen values and different PCs. PC1 contributed 47.28 % variability with eigen value 7.0913. In scree plot, a semi-curve line was obtained because the first component usually explains much of the variability, the next few components explain a moderate amount, and the latter components only explain a small fraction of the overall variability. When compared to other three PCs, PC1 exhibited the greatest variation. So, selection of lines from this PC is useful.

Table 2. Eigen values, variability and cumulative percentage of advanced lines of oat *(sativa × A. sterilis*)

Principal component (PC)	Eigen value	Variability (%)	Cumulative %
PC1	7.0913	47.28	47.28
PC2	1.3576	9.05	56.23
PC3	1.2220	8.15	64.47
PC4	1.0424	6.95	71.42
PC5	0.8510	5.67	77.10
PC6	0.7553	5.03	82.13
PC7	0.6095	4.06	86.19
PC8	0.5794	3.86	90.06
PC9	0.4355	2.90	92.96
PC10	0.3324	2.22	95.18
PC11	0.2525	1.68	96.86
PC12	0.2131	1.42	98.28
PC13	0.1147	0.76	99.05
PC14	0.0829	0.55	99.60
PC15	0.0603	0.40	1.00

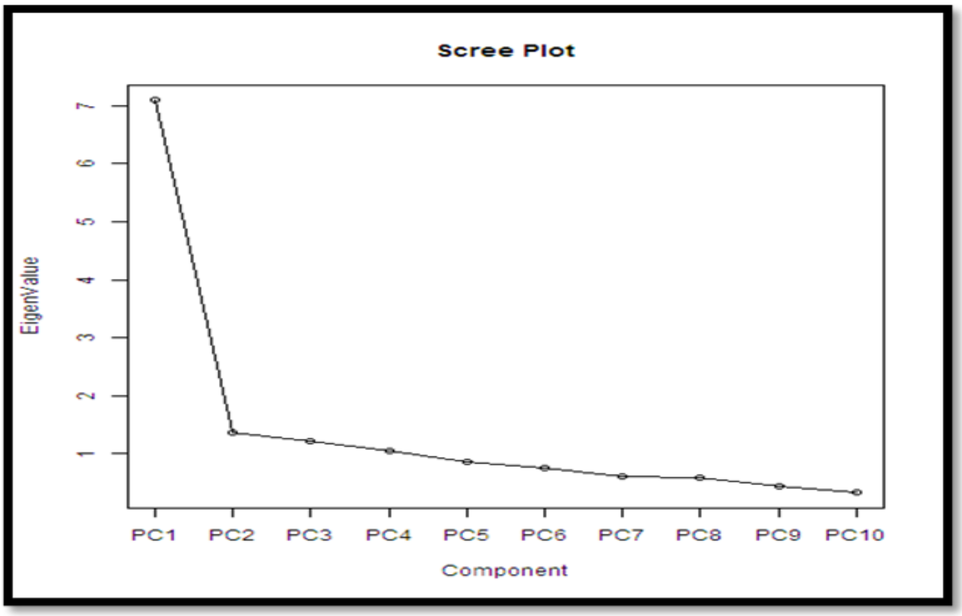

Fig. 1. Scree plot diagram of Eigen values constructed on fifteen traits recorded in 98 oat advanced lines

Rotated component matrix depicted in **Table 3**. showed that PC1 accounted for the highest variability (47.28%) and includes traits such as days to 50% flowering, days to maturity, number of tillers per plant and number of leaves per tiller. The PC2 was observed to have the highest value for plant height, while most of the characters were observed to be in negative direction.

PC3 includes yield contributing traits like stem width, flag leaf width, number of spikelets per panicle, number of florets per panicle, 1000 seed weight and seed yield per plant. As shown in **Table 4**. PC3 contained the traits which significantly contributes to the yield hence, lines from PC3 can be selected for bringing out quick uplifting of dependent traits i.e., yield. Similar findings were reported by Krishna *et al*. (2014). Similarly, high contribution in positive direction was observed for leaf width, internode length, flag leaf length and panicle weight in PC4.

High score for a specific genotype in a specific component indicated high values for the variables in that specific genotype (Singh and Chaudhary, 1977). The lines JO (IS) 12-8-1(5.3725), JO (IS) 12-8-2 (4.1748), JO (IS) 12-8-3 (5.0273) along with other genotypes showed PC score greater than one in PC 1 (**Table 5**) and as per the interpretation of rotated component matrix (**Table 4**) traits falling under PC 1were days to 50 percent flowering, days to maturity, number of tillers per plant and number of leaves per tillers. Hence, the lines JO (IS) 12-8-1(5.3725), JO (IS) 12-8-2 (4.1748), JO (IS) 12-8-3 (5.0273) along with other genotypes can be used for improvement of traits such as days to 50 percent flowering, days to maturity, number of tillers per plant and number of leaves per tillers. Similarly, lines JO (IS) 12-4-1 (1.7531), JO (IS) 12-4-2 (1.775), JO (IS) 12-4-3 (3.3433) along with other genotypes performed better for the trait plant height.

The lines JO (IS) 12-4-7(1.3228), JO (IS) 12-4-9 (2.2294), JO (IS) 12-4-11 (1.5902) along with other genotypes performed better for traits such as stem width, flag leaf width, number of spikelets per panicle, number of florets per panicle, 1000 seed weight and seed yield per plant. Similarly, JO (IS) 12-4- (1.61), JO (IS) 12-4-8 (1.0369), JO (IS) 12-4-13 (1.4943), JO (IS) 12-4-14 (1.343) along with other genotypes performed better for traits like leaf width, internode length, flag leaf length and panicle weight.

The genotypes like JO (IS) 12-4-9, JO (IS) 12-5-5, JO (IS) 12-5-14, JO (IS) 12-6-19, JO (IS) 12-8-2, JO (IS) 12-8-4, JO (IS) 12-8-5, JO (IS) 12-8-11 and JO (IS) 12-8-18 fell

Table 3. Rotated component matrix for different traits of advanced lines of oat (*A. sativa ×A. sterilis*)

Character	Principal Components			
	PC1	PC 2	PC3	PC4
Stem width(cm)	-0.1654	0.3180	0.3389	0.0837
Leaf width(cm)	-0.0407	-0.1656	0.1656	0.6293
Days to 50% flowering	0.2623	-0.2974	-0.1417	-0.3116
Days to maturity	0.2109	-0.6374	0.0206	0.0862
Plant height(cm)	-0.3343	0.3288	-0.0745	-0.0039
Number of tillers per plant	0.3580	-0.0060	0.0809	-0.0460
Number of leaves per tiller	0.3469	-0.0037	0.1212	-0.0599
Internode length(cm)	-0.2786	-0.0577	-0.1743	0.3548
Flag leaf length(cm)	0.0855	-0.3473	-0.4279	0.5088
Flag leaf width(cm)	0.1340	-0.3076	0.5110	-0.3090
Panicle weight(g)	-0.3264	0.0496	-0.0086	0.5507
Number of spikelets per panicle	-0.3159	-0.1710	0.3028	-0.0768
Number of florets per panicle	-0.3119	-0.1588	0.3388	-0.0681
1000 seed weight(g)	-0.3073	-0.1920	0.3425	-0.0871
Seed yield per plant(g)	-0.2668	-0.2518	0.3420	-0.3220

Table 4. Interpretation of rotated component matrix for the traits having highest value in each PC

PCs	Characters
PC 1	Days to 50% flowering, Days to maturity, Number of tillers per plant and Number of leaves per tiller
PC 2	Plant height
PC 3	Stem width, Flag leaf width, Number of spikelets per panicle, Number of florets per panicle, 1000 seed weight and Seed yield per plant
PC4	Leaf width, Internode length, Flag leaf length and Panicle weight

Table 5. List of advanced lines of oat *(A.sativa×A.sterilis)* with >1 PC scores

Genotypes	Principal components
JO (IS) 12-8-1(5.3725), JO (IS) 12-8-2 (4.1748), JO (IS) 12-8-3 (5.0273), JO (IS) 12-8-4 (5.3725), JO (IS) 12-8-5 (4.5414), JO (IS) 12-8-6 (5.8929), JO (IS) 12-8-7 (5.2634), JO (IS) 12-8-8 (5.1812), JO (IS) 12-8-9 (5.3909), JO (IS) 12-8-10 (5.3445), JO (IS) 12-8-11 (5.2023), JO (IS) 12-8-12(5.177), JO (IS) 12-8-13 (5.0979), JO (IS) 12-8-14 (5.1364), JO (IS) 12-8-15 (5.0635), JO (IS) 12-8-16 (4.8122), JO (IS) 12-8-17 (5.0416), JO (IS) 12-8-18 (4.7032), JO (IS) 12-8-19 (4.5501), JO (IS) 12-8-20 (5.1221).	PC 1
JO (IS) 12-4-1 (1.7531), J JO (IS) 12-4-2 (1.775), JO (IS) 12-4-3 (3.3433), JO (IS) 12-4-4 (1.4376), JO (IS) 12-4-9 (1.5681), JO (IS) 12-4-10 (1.0576), JO (IS) 12-4-13 (1.3224), JO (IS) 12-4-15 (2.2782), JO (IS) 12-4-16 (1.6802), JO (IS) 12-4-18 (1.9581), JO (IS) 12-4-19 (1.5349), JO (IS) 12-5-14 (2.5916), JO (IS) 12-5-17 (1.2553), JO (IS) 12-6-6 (1.2651), JO (IS) 12-6-18 (1.6448), JO (IS) 12-7-16 (1.5995), JO (IS) 12-8-2(1.7502) and JO (IS) 12-8-14 (1.0771)	PC 2
JO (IS) 12-4-7(1.3228), JO (IS) 12-4-9 (2.2294), JO (IS) 12-4-11 (1.5902), JO (IS) 12-4-20 (1.5347), JO (IS) 12-5-3 (2.753), JO (IS) 12-5-5 (1.0384), JO (IS) 12-5-6 (1.8921), JO (IS) 12-5-14 (1.4622), JO (IS) 12-5-15 (1.3508), JO (IS) 12-6-7 (1.2238), JO (IS) 12-6-11 (1.4184), JO (IS) 12-6-19 (1.6613), JO (IS) 12-7-12 (1.0594), JO (IS) 12-7-18 (1.0759), JO (IS) 12-8-2 (1.4715), JO (IS) 12-8-4 (1.1414), JO (IS) 12-8-5 (1.4203), JO (IS) 12-8-11 (1.6865), JO (IS) 12-8-18 (1.182) and Kent (1.8357)	PC 3
JO (IS) 12-4- (1.61), JO (IS) 12-4-8 (1.0369), JO (IS) 12-4-13 (1.4943), JO (IS) 12-4-14 (1.343), JO (IS) 12-4-15 (1.5867), JO (IS) 12-4-16 (1.2445), JO (IS) 12-4-17 (1.1568), JO (IS) 12-5-5 (1.6621), JO (IS) 12-5-7 (1.6474), JO (IS) 12-5-8 (1.4391), JO (IS) 12-5-13 (1.4265), JO (IS) 12-6-15 (1.2951), JO (IS) 12-6-17 (2.2328), JO (IS) 12-6-19 (1.1312), JO (IS) 12-7-6 (1.1511), JO (IS) 12-7-15 (1.9866), JO (IS) 12-7-19 (1.1964), JO (IS) 12-8-16 (2.1886), JO (IS) 12-8-17 (1.8851) and JO (IS) 12-8-18 (1.4509)	PC 4
JO (IS) 12-8-14	PC 1 and PC 2
JO (IS) 12-8-4, JO (IS) 12-8-5 and JO (IS) 12-8-11	PC 1 and PC 3
JO (IS) 12-8-16 and JO (IS) 12-8-17	PC 1 and PC 4
JO (IS) 12-4-9 and JO (IS) 12-5-14	PC 2 and PC 3
JO (IS) 12-4-13, JO (IS) 12-4-15 and JO(IS) 12-4-16	PC 2 and PC 4
JO (IS) 12-5-5 and JO (IS) 12-6-19	PC 3 and PC 4
JO (IS) 12-8-18	PC 1, PC 3 and PC 4
JO (IS) 12-8-2	PC 1, PC 2 and PC 3

in yield and other yield related traits PCs with high PC scores (**Table 5**) which indicates these lines contained high value for yield and for yield attributing traits also. Based on PC scores, lines JO (IS) 12-8-2 and JO (IS) 12-8-18 fell under three PCs, so these lines can be used in breeding programme for the improvement of maximum traits. Thus, the results of principal component analysis have shown that population contained high amount of genetic variation and also identified the traits contributing towards genetic diversity. Hence, the results can significantly help for designing the intensive selection procedures for rapid improvement of various agro-morphological traits analysed in this study.

REFERENCES

Achleitner, A., Tinker, N.A., Zechner, E. and Buerstmayr, H. 2008. Genetic diversity among oat varieties of worldwide origin and associations of AFLP markers with quantitative traits. *Theoretical and Applied Genetics*, **117**(7): 1041-1053. [Cross ref]

Ahmad, M., Wani, B.A., Mehraj, U., Dar, Z.A., Lone, A.A. and Rather, M.A. 2016. Genotype x environmental interaction and stability analysis for grain quality and yield in oats (*Avena sativa* L.). *Electronic Journal of Plant Breeding*, **7**(4):1132-1135. [Cross ref]

Flander, L., Salmenkallio-Marttila, M., Suortti, T. and Autio, K. 2007. Optimization of ingredients and baking process for improved wholemeal oat bread quality. *LWT-Food Science and Technology*, **40**(5):860-870. [Cross ref]

Greene, N.V., Kenworthy, K.E., Quesenberry, K.H., Unruh, J.B. and Sartain, J.B. 2008. Diversity and relatedness of common carpet grass germplasm. *Crop science*, **48**(6): 2298-2304. [Cross ref]

Hilli, H.J. and Kapoor, R. 2021. Oat floral biology and factors affecting seed set. *Bhartiya Krishi Anusandhan Patrika*, **36**(1): 79-80. [Cross ref]

Hotelling, H. 1933. Analysis of complex statistical variables into principal components. *Journal of Educational Psychology*, **24**: 417. [Cross ref]

Krishna, A., Ahmed, S., Pandey, H.C. and Kumar, V. 2014. Correlation, path and diversity analysis of oat (*Avena sativa* L.) genotypes for grain and fodder yield. *Journal of Plant Science & Research*, **1**(2):1-9.

Kumari, T. and Jindal, Y. 2019. Genetic diversity and variability analysis in oats (*Avena sp*) genotypes. *Electronic Journal of Plant Breeding*, **10**(1):1-8. [Cross ref]

Mohanta, T.K., Mohanta, Y.K., Yadav, D., Hashem, A., Abd_Allah, E.F. and Al-Harrasi, A. 2020. Global Trends in Phytohormone Research: Google trends analysis revealed African countries have higher demand for phytohormone information. *Plants*, **9**(9): 1248. [Cross ref]

Murphy, J.P. and Hoffman, L.A. 1992. The origin, history and production of oat. *Oat science and technology*, **33**:1-28. [Cross ref]

Pearson, K. 1901. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, **2**:559. [Cross ref]

Rezai, A. and Frey, K.J. 1990. Multivariate analysis of variation among wild oat accessions-seed traits. *Euphytica*, **49**(2):111-119. [Cross ref]

Singh, R., De, S. and Belkheir, A. 2013. *Avena sativa* (Oat), a potential neutraceutical and therapeutic agent: an overview. *Critical reviews in food science and nutrition*, **53**(2):126-144. [Cross ref]

Singh, R.K. and Chaudhary, B.D. 1977. Biometrical methods in quantitative genetic analysis. Kalyani Publishers. New Delhi 266.

https://doi.org/10.37992/2023.1402.049