Indian J. Genet. Plant Breed., (2024); 84(3): 461-470
https://doi.org/10.31742/ISGPB.84.3.17
ISSN: 0975-6906
www.isgpb.org

RESEARCH ARTICLE

Enhancing tobacco (*Nicotiana tabaccum* L.) breeding efficiency utilizing GBLUP through SSR markers for superior parental selection based on leaf quality traits

Parviz Goodarzi Mokri, Reza Darvishzadeh*, Bahram Maleki Zanjani[1], Hamid Hatami Maleki[2*] and Hossein Zeinalzadeh-Tabrizi[3]

Abstract

Tobacco (*Nicotiana tabaccum* L.) is considered to be an industrial and medicinal plant that plays an important role in the economies of most countries. The present study demonstrated how the genomic best linear unbiased predictor (GBLUP) method could determine the future breeding potential of a tobacco panel by means of 26 SSR fingerprinting data. A set of 71 genotypes of tobacco considering 11 agro-morphological and leaf chloride content of a qualitative character were assessed during two consecutive years under field conditions. Results revealed that GBLUP could efficiently predict the breeding value (BV) of studied characters. Considering the total ranks of each genotype across studied characters, genotypes, C.H.T.269-12e", C.H.T.266-6, SS298-2, C.H.T.209.12e, Triumph, and Ohdaruma had the highest predicted BVs and, therefore, these genotypes are good candidates for parental selection. Based on BVs data, the studied characters were classified into groups whose chemical characteristics were distinguished from others. Cluster analysis of this tobacco panel based on BVs leads to four heterotic groups, and the combination of their information with the total ranks of each genotype across studied characters can guide tobacco breeders in selecting desirable and effective parents.

Keywords: Best linear unbiased prediction, heterotic groups, oriental tobacco, SSR markers

Introduction

Tobacco (*Nicotiana tabacum* L.) is an allopolyploid species from the *Solanaceae* family with 2n=4x=48 chromosomes. It is regarded as both an industrial and medicinal crop due to its leaves, which are consumed in the form of smoke (Chaplin 1975), as well as its nicotine and alkaloid content (Tso 2006). According to Berlowitz et al. (2020), teas made from tobacco leaves were used against intestinal worms, as a laxative to induce vomiting (emetic), as an expectorant, for fainting and dizziness, as well as for headaches. Tobacco leaves are applied to cuts as an antiseptic and to stop bleeding. The global tobacco production was around 6,502 million kg grown in an area of 3.43 mha in 2017. It is grown on less than 1% of the world's agricultural land and on a wide variety of soils and climates (Lencucha et al. 2022). The area under cultivation and the production of tobacco in Iran are 9500 ha and the production is 19,200 t/ha, respectively (Mirkarimi et al. 2021). Tobacco is classified according to several characteristics, including growth type. The oriental-type tobacco is a type of tobacco that is sun-cured with a small leaf, a delicate texture, mild smoke, and a pervasive odor. Oriental-type tobacco has the ability to grow in low-fertility soils and is grown in Iran, Turkey, Greece, Bulgaria, Lebanon, and the Republic of Macedonia (Davis and Nielson 1999). So, frequently, oriental-type tobacco was grown and implemented as the major constituent of

Department of Plant Production and Genetics, Faculty of Agriculture, Urmia University, Urmia, Iran.

[1]Department of Agronomy and Plant Breeding Faculty of Agriculture, Zanjan University, Zanjan, Iran.

[2]Department of Plant Production and Genetics, Faculty of Agriculture, University of Maragheh, Maragheh, Iran.

[3]Department of Horticulture and Agronomy, Faculty of Agriculture, Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan.

***Corresponding Author:** Hamid Hatami Maleki, Department of Plant Production and Genetics, Faculty of Agriculture, University of Maragheh, Maragheh, Iran, E-Mail: hatamimaleki@maragheh.ac.ir, r.darvishzadeh@urmia.acir

How to cite this article: Mokri P.G., Darvishzadeh R., Zanjani B.M., Maleki H.H. and Zeinalzadeh-Tabrizi H. 2024. Enhancing tobacco (*Nicotiana tabaccum* L.) breeding efficiency utilizing GBLUP and SSR markers for superior parental selection based on leaf quality traits. Indian J. Genet. Plant Breed., **84**(3): 461-470.

Source of support: Nil

Conflict of interest: None.

Received: March 2024 **Revised:** June 2024 **Accepted:** July 2024

blend cigarette stocks. Therefore, characters related to the smoking quality of tobacco are of special importance. Many factors, including physical (combustibility, elasticity, moisture, and color), chemical (micro and macro elements level in leaves), and organoleptic (odor and taste), affect the tobacco leaf quality (Yang *et al.* 2007). Along with agro-morphological characteristics that affect tobacco lead yield, chloride, as a chemical micronutrient, has positive effects on the quality of oriental-type tobacco (Darvishzadeh *et al.* 2011). Albeit, small amounts of chloride (below 1.5%) are needed for plant growth and improving yield and quality characteristics such as color, moisture content, elasticity, and leaf burning capacity (Mcevoy 1957; Chari 1995), but larger amounts of chloride (more than 2%) have many adverse effects and decrease the quality of tobacco leaves. Thus, chloride accumulation levels in leaves have a determinative role in tobacco quality (Akehurst 1981; Guardiola *et al.* 1987). According to literature (Darvishzadeh and Alavi 2011), chloride accumulation in oriental-type tobacco is varied and depends on the type of genotype, as well as several genes are engaged to control its accumulation.

The existence of genetic variability is a prerequisite for establishing any plant breeding program through selection, but information about the genetic variance, especially the additive variance of an interested trait, is vital. Additive genetic effect is referred to as breeding value, or the value of genes to progeny. So, the breeding value (BV) is the sum of the average effects of alleles passed from parents to progeny (Falconer and Mackay 1996). Therefore, by selecting based on breeding value, the efficiency of selection can be improved (Quintal *et al.*2017). For the main autogamous species, such as tobacco, the key objective of breeding programs has been to commercialize pure-line cultivars. Hybrids are, however, used with a view to combining simply inherited traits into single genotypes rather than as a means of exploiting true heterosis. Anyhow, Since an artificial mating design is used to produce commercial tobacco seed, the possibility of improving tobacco performance by exploiting heterosis should be considered. In this regard, the general and specific combining abilities of the lines is need to be estimated. So, a tobacco breeder has to identify lines possessing suitable general combining ability as representative of additive gene effect to complement each other when crossed. The best linear unbiased prediction (BLUP) approach (Henderson 1985) has been regularly used for the prediction of the breeding value BV instead of traditional mating systems, which require several crosses. In the mixed model equations (MME), the BLUP technique uses the relationship information in matrix A. The matrix A can be calculated by means of coancestory information, but these coancestory coefficients are not well determined in the self-pollinated plants (Bauer *et al.* 2006). Hence, genetic similarities through molecular markers have been used to compute matrix A (Bernardo 1993, 1994). Nowadays, genomic best linear unbiased prediction (GBLUP) is applied for the prediction of breeding values of several characteristics in crop plants such as maize (Cantelmo *et al.* 2017), wheat (Bonnet *et al.* 2020), rice (Chung and Liao 2020), Asiatic cotton (Vineeth et al. 2022) and potato (Sood et al. 2020; Sood et al. 2022). To our knowledge, there are few reports about GBLUP in tobacco. Hence, in the current study, the GBLUP method was imposed on the two-year data of agro-morphological characteristics as well as

Table 1. List of studied tobacco genotypes

S.No.	Genotype	Origin	S.No.	Genotype	Origin	S.No.	Genotype	Origin
1	Ts 8	-	26	Jahrom14	Iran	51	SPT 430	Iran
2	F.K.40-1	-	27	C.H.T.269-12e	Iran	52	SPT 432	Iran
3	Samsun 959	Turkey	28	Matianus	Iran	53	SPT 433	Iran
4	Samsun dere	Turkey	29	Nevrokop	Bulgaria	54	SPT 434	Iran
5	Tyk-Kula	Iran	30	Mutant 3	Iran	55	SPT 436	Iran
6	Alborz23	Iran	31	C.H.T.209.12e	Iran	56	SPT 439	Iran
7	ss-289-2	Iran	32	Xanthi	Iran	57	SPT 441	Iran
8	Basma 12-2	Zimbabwe	33	C.H.T.283-8	Iran	58	P.D.324	Iran
9	Basma 16-10	Zimbabwe	34	C.H.T.266-6	Iran	59	P.D.325	Iran
10	Basma 104-1	Zimbabwe	35	C.H.T.273-38	Iran	60	P.D.328	Iran
11	Basma 181-8	Zimbabwe	36	Pobeda 1	Russian	61	P.D.329	Iran
12	K.B		37	L 17	Bulgaria	62	P.D.336	Iran
13	G.D.165	Bulgaria	38	Melkin 261	Turkey	63	P.D.345	Iran
14	Pobeda 2	Russian	39	H.T.I	-	64	P.D.364	Iran

15	Kramograd N.H.H. 659	Bulgaria	40	Triumph	Iran	65	P.D.371	Iran
16	Immni 3000	Australia	41	Basma.S.31	Belgium	66	P.D.381	Iran
17	kharmanli 163	Iran	42	SPT 403	Iran	67	Mutant 4	Iran
18	Izmir	Turkey	43	SPT 405	Iran	68	C.H.T.209.12e × F.K.40-1	Iran
19	Ploudive 58	Bulgaria	44	SPT 406	Iran	69	T-B-22	
20	T.K.23	-	45	SPT 408	Iran	70	Krumovgraid	Bulgaria
21	Pz17	-	46	SPT 409	Iran	71	Ohdaruma	Yugoslavia
22	OR-205	Iran	47	SPT 410	Iran			
23	OR-379	Iran	48	SPT 412	Iran			
24	Trabozan	Turkey	49	SPT 413	Iran			
25	Line 20	Iran	50	SPT 420	Iran			

= Not known

leaf chloride along with SSR genotyping information with the goal of selecting parents to be used in future tobacco breeding programs.

Materials and methods

Experimental methods and variable measurements

The field experiment was conducted at the research farm of the Urmia Tobacco Research Institute of Iran. The randomized complete block design with three replications for each genotype was utilized to study the agro-morphological traits of tobacco germplasm (Table 1).

In this research, 11 agro-morphologic traits, including stem diameter (SD), days to 50% flowering (D50F), leaf number (LN), plant height (PH), green leaf yield (GLY), dry leaf yield (DLY), chlorine (Cl), leaf width (LW), one green leaf weight (GLW), leaf length (LL) along with the leaf chlorine content (Cl) of 71 oriental-type and semi oriental-type tobacco genotypes were taken and the data were recorded. In the field condition, The seeds of each genotype were planted in the field at two locations, viz., the Tobacco Research Center of Urmia (West Azarbaijan province of Iran) and Anghaneh village, which is located near the Urmia lake for two years (2019-2020). Planting was done indirectly in the form of seedling transplanting. Approximately 5 grams of seeds per m^2 were sown for each genotype separately. The seed beds were prepared according to the oriental-type tobacco custom. The seedlings were transplanted to plots when their average height was about 12 cm. Each plot comprised three 5 m rows, with a spacing of 65×6 ×20 cm. The ripe leaves were harvested three times and were sun-cured according to oriental tobacco. To measure the leaf chloride content a random sample of 20 leaves was taken from each plot, and the percentage of chloride was determined as defined by CORESTA (Cooperation Center for Scientific Research Relative to Tobacco). After 50% of the plants in each plot bloomed, three plots were selected randomly, and agro-morphological characters were recorded.

Genomic DNA extraction and SSR fingerprinting

Genomic DNA was extracted from leaf samples of plants according to the method of Dellaporta *et al.* (1983). DNA fingerprinting of the studied tobacco germplasm was done using 26 simple sequence repeats (SSR) primer pairs (Supplementary Table S1) from the tobacco SSR linkage map (Bindler *et al.* 2007). DNA sample concentration was assessed using spectrophotometry at 260 nm (BioPhotometer 6131; Eppendorf, Hamburg, Germany). DNA integrity was evaluated by electrophoresis on 0.8% (w/v) agarose gels in 0.5X TBE buffer (45 mM Tris base, 45 mM boric acid, 1-mM EDTA pH 8.0) using 1-µL of DNA. Samples displaying a smear on the gel were excluded from further analysis. PCR was conducted in a 20 µL volume using a 96-well Eppendorf Mastercycler Gradient (Type 5331, Eppendorf AG, Hamburg, Germany). The reaction mixture comprised 2.5 mM of each primer, 0.4 Unit of Taq DNA polymerase (Cinna Gen Inc., Tehran, Iran), 100 µM of each dNTP (BioFluxbiotech), 2 µL 10X PCR buffer, 2 mM $MgCl_2$ (CinnaGen, Tehran, Iran), ddH_2O, and 25 ng template DNA. Amplification consisted of 35 cycles, including denaturation at 94°C for 1-minute, annealing at 55 °C for 1 min, and extension at 72°C for 1.5 minutes. Additionally, initial denaturation at 94°C for 4 min and final extension at 72°C for 10 minutes were performed. The reaction products were mixed with an equal volume of formamide dyes (98% formamide, 10 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol) and resolved in a 3% (w/v) agarose gel in 0.5X TBE buffer. The gel was stained with 1.0 Pg mL^{-1} ethidium bromide and visualized under UV light using a GelDoc image analysis system (Gel Logic 212 PRO, USA).

Data analysis

To facilitate the prediction of breeding values (BVs) corresponding to the examined traits, the genetic relationships among the studied tobacco genotypes (individuals) were quantified through the computation of the kinship matrix (also known as the A matrix), utilizing SSR fingerprinting data within the TASSEL software framework. Subsequently, the BVs for eleven agro-morphological characteristics, in addition to leaf chloride content, were estimated employing a mixed linear model approach as delineated. Briefly, the mixed linear model is:

$$Y = Xb + Zu + e$$

Let Y be the observation vector, b and u be vectors of fixed and random effects, X and Z be incidence matrices for fixed and random effects, and e be a vector of residuals. In the mixed linear model, fixed and random effects are estimated using BLUE (best linear unbiased estimation) and BLUP (best linear unbiased prediction), respectively.

Vectors e and u are random effects with a normal distribution with a mean of zero and deviation of $VAR\begin{bmatrix}u\\e\end{bmatrix} = \begin{bmatrix}G & 0\\0 & R\end{bmatrix}$. We typically assume $R = VAR(e) = \sigma_e^2 I_n$ and $G = VAR(u) = \sigma_u^2 I_t$. Index t and n indicate the number of levels of random effects (genotype or treatment) and the number of observations in the identity matrix (I), respectively (Yang 2010). σ_u^2 is random effects variance and σ_e^2 is residual variance. In practice, BLUP and BLUE must be replaced with empirical BLUP and BLUE. In other words, variance components in G and R must be replaced with their estimation and calculated using restricted maximum likelihood (RELM) (Patterson and Thompson 1971).

BLUP and BLUE will be estimated based on Henderson's Mixed Model Equations (1990):

$$\begin{bmatrix}X'R^{-1}X & X'R^{-1}Z\\Z'R^{-1}X & Z'R^{-1}Z + G^{-1}\end{bmatrix}\begin{bmatrix}\hat{\beta}\\\hat{u}\end{bmatrix} = \begin{bmatrix}X'R^{-1}Y\\Z'R^{-1}Y\end{bmatrix}$$

Where R and G are $R = \sigma_e^2 I_n$ and $G = \sigma_u^2 I_t$.

If Henderson's Mixed Model Equations multiplied by σ_e^2 and the number of repeats for genotypes be considered unequal, then equations will change to the below form (Bernardo and Yu 2007):

$$\begin{bmatrix}X'r^{-1}X & X'r^{-1}Z\\Z'r^{-1}X & Z'r^{-1}Z + \theta^{-1}\end{bmatrix}\begin{bmatrix}\hat{\beta}\\\hat{u}\end{bmatrix} = \begin{bmatrix}X'r^{-1}Y\\Z'r^{-1}Y\end{bmatrix}$$

Where $\theta^{-1} = A^{-1}(\frac{V_e}{V_g}), VAR(u) = \sigma_u^2 I_t \approx AV_A$ and $VAR(e) = \sigma_e^2 I_n \approx r_n \sigma_e^2$.

A is a t × t matrix (t= number of genotypes) of kinship coefficient that indicates genetic covariance structure among the individuals. R is an identity matrix provided that the number of genotypes is the same; if not, R will be an n × n matrix (n= number of observations) in which extra-diameter elements be equal to zero and diagonal elements be the inverse of the number of genotype's repeats (for example, genotype one's the inverse of the number of repeats in the first set (year × place), two sets (year × place), up to the end of sets). Where V_A and V_e are genetic and residual variance, respectively.

Cluster analysis was implemented to determine the similarity or dissimilarity of genotypes and their characters based on BV and to clarify relationships among them using the K-means algorithm. In this way, "factoextra" (Kassambara and Mundt 2020) as an R package was implemented. Then the optimal number of clusters was determined using the Elbow index.

Results

Predicted BVs pertaining to 11 agro-morphologic characteristics, as well as leaf chloride content, are shown in Table 2. Genotypes were ranked based on the predicted BVs of each character (except chloride) in such a way that rank 1 was assigned to the genotype with the highest predicted breeding value. This procedure for Cl content was done in reverse because Cl content is an undesirable character, so rank 1 was assigned to the genotype with the least predicted BV. In terms of leaf Cl content, genotypes Pobeda 2 and P.D.329, with predicted BVs of 0.24 and -0.25, had the highest positive and lowest negative values, respectively. Regarding character LN, genotypes Line20 and SPT409 with BV values of 5.91 and -23.82, had the highest and lowest values, respectively. In this research, genotypes Ohdaruma and P.D.336 possessed the highest and lowest values of BV for characters LL and LW, and their values ranged between 8.44 and -4.93 for character LL and 5.05 and -3.29 for character LW. For the characters PH and SD, genotypes Samsundere and SPT439 with BV values of 28.89 and -59.85 for PH and genotypes C.H.T.209.12e and SPT433 with BV values of 1.99 and -1.62 for SD were detected. Results revealed that genotypes C.H.T.209.12e and SPT409 with BV values of 36.42 and -33.58 had the greatest and the poorest potential to be selected as a parent for breeding character D50F. In both yield-related characters, including GLY and DLY, genotypes Triumph and SPT413 had the highest (1.52) and the lowest (-0.74) BV values, respectively. The sum of the ranks for each genotype across all studied characters was considered the total rank of the genotype. In terms of total ranks, genotypes C.H.T.269-12e, C.H.T.266-6, SS298-2, C.H.T.209.12e, Triumph, and Ohdaruma with ranks of 74, 81, 90, 94, 115, and 134 had the highest predicted BVs, respectively. On the other hand, genotypes SPT403, SPT413, SPT436, Tyk-Kula, SPT408, and Xanthi "with total ranks of 623, 563, 551, 545, 544, and 543 had the lowest predicted BVs.

Table 2. Predicted breeding value and corresponding ranks of the genotypes for studied characteristics in the tobacco germplasm

Genotype	CI	Rank	LN	Rank	LL	Rank	LW	Rank	GLW	Rank	GLY	Rank	DLY	Rank	PH	Rank	SD	Rank	D50F	Rank	Sum
Ts 8	0.11	54	-3.97	20	3.85	10	3.27	5	3.45	10	4.73	3	0.77	5	26.15	3	1.04	5	5.07	8	123
F.K.40-1	0.07	49	-4.58	27	0.92	23	0.05	39	-0.50	47	-0.67	42	-0.09	44	-3.51	39	0.08	20	-6.13	30	360
Samsun 959	-0.10	14	-2.25	14	-0.17	38	-1.94	65	-2.42	67	-1.06	48	-0.32	62	-12.36	46	-0.05	23	-3.19	20	397
Samsun dere	-0.03	28	3.29	4	-0.10	35	0.22	33	-0.88	54	-0.27	37	0.00	36	28.89	1	0.33	10	-3.17	19	257
Tyk-Kula	0.02	36	-10.36	50	-1.25	53	-1.49	60	-0.56	48	-3.33	69	-0.60	70	-23.38	54	-0.57	51	-14.32	54	545
Alborz23	0.03	38	0.52	10	0.99	22	0.02	43	-0.78	52	0.97	23	0.15	24	9.85	24	0.40	9	5.35	7	252
ss-289-2	0.05	42	2.70	6	6.25	2	2.02	12	3.35	11	4.54	4	0.87	4	28.47	2	1.15	4	15.26	3	90
Basma 12-2	-0.19	6	-7.09	37	2.43	14	1.08	18	1.95	13	1.09	21	0.32	15	10.64	22	0.28	11	-3.52	21	178
Basma 16-10	-0.21	4	-9.25	48	0.20	32	0.15	35	0.34	30	0.30	29	0.22	22	-3.95	40	-0.79	60	-9.23	35	335
Bbasma 104-1	-0.21	5	-7.43	39	0.80	26	0.73	22	0.46	27	0.16	32	0.25	19	13.85	15	-0.35	34	-10.93	43	262
Basma 181-8	-0.14	9	1.60	8	-0.30	40	-2.75	69	-2.43	68	-1.12	50	-0.09	45	14.12	14	0.14	17	2.37	10	330
K.B	0.02	34	5.39	2	1.09	20	0.20	34	1.60	16	2.45	12	0.52	10	8.98	26	-0.34	33	0.41	14	201
G.D.165	0.16	66	-7.86	41	-0.61	45	-0.48	51	0.13	33	-1.47	54	-0.28	60	-36.19	61	-0.81	61	-10.59	42	514
Pobeda 2	0.24	71	2.84	5	-1.67	56	-0.83	55	-0.36	42	1.39	18	0.12	26	8.36	27	-0.63	54	-0.43	16	370
Kramograd N.H.H. 659	0.06	44	-8.72	47	-0.32	41	-0.27	47	-1.30	58	0.39	28	-0.03	39	-20.07	50	-0.51	47	-12.04	48	449
Immni 3000	0.06	45	-5.32	30	0.26	31	-1.14	57	-1.00	55	2.02	15	0.24	20	13.02	16	0.12	18	-3.05	18	305
kharmanli 163	0.11	52	-5.36	31	1.90	17	2.31	9	1.52	17	0.89	24	0.03	35	25.79	4	0.17	15	-9.28	37	241
Izmir	-0.03	29	-5.43	33	-2.94	66	-1.84	62	-1.59	61	-1.31	52	-0.13	48	6.39	31	-0.55	50	-12.45	50	482
Ploudive 58	0.21	69	-4.16	21	1.70	18	0.06	38	0.92	20	3.09	9	0.42	12	6.77	30	0.17	14	3.27	9	240
T.K.23	0.16	63	-3.55	18	0.00	34	-0.23	45	0.10	34	-0.80	44	-0.17	50	-13.09	48	-0.44	38	-9.89	41	415
Pz17	0.12	57	-11.06	53	-0.16	37	-0.17	44	-0.09	37	0.68	25	0.05	32	-21.51	51	-0.36	35	-9.73	39	410
OR-205	-0.10	15	-5.38	32	-0.82	48	0.08	36	-0.22	40	1.31	19	0.30	16	12.81	18	-0.47	43	-6.61	31	298
OR-379	0.16	65	-6.47	36	-0.70	46	-0.42	49	-0.59	50	0.14	33	0.07	30	17.45	11	-0.25	28	-8.03	33	381
Trabozan	0.13	59	-0.41	11	2.49	13	0.28	30	1.14	18	2.71	11	0.53	8	7.85	28	0.16	16	-5.36	27	221
Line 20	0.15	62	5.91	1	-1.62	55	0.03	40	-0.14	38	1.02	22	0.23	21	19.11	8	0.10	19	-4.99	24	290
Jahrom14	-0.12	11	-10.77	52	0.28	30	0.58	25	-0.58	49	-2.22	65	-0.47	67	11.90	19	-0.18	26	-9.26	36	380

------- continued

Genotype	CI	Rank	LN	Rank	LL	Rank	LW	Rank	GLW	Rank	GLY	Rank	DLY	Rank	PH	Rank	SD	Rank	D50F	Rank	Sum
C.H.T.269-12e	-0.05	26	2.24	7	5.50	5	1.65	15	4.39	3	4.46	5	0.88	3	19.91	6	1.67	2	35.13	2	74
Matianus	0.16	64	-4.28	24	-1.49	54	0.27	31	-0.28	41	0.47	27	0.13	25	-23.03	53	-0.33	32	-11.12	45	396
nevrokop	-0.14	10	-7.24	38	0.09	33	0.29	29	0.45	28	-0.64	40	0.06	31	-1.46	37	-0.54	49	-9.69	38	333
Mutant 3	0.02	35	-11.39	54	-0.47	43	0.98	19	-0.04	35	0.26	31	-0.18	51	2.91	33	-0.29	31	-11.59	46	378
C.H.T.209.12e	0.02	37	-0.46	12	5.36	7	2.94	6	4.03	7	3.33	6	0.58	7	17.80	10	1.99	1	36.42	1	94
Xanthi	0.11	53	-6.12	34	-2.51	62	-2.00	67	-1.67	62	-2.46	66	-0.32	61	-4.03	41	-0.58	53	-10.99	44	543
C.H.T.283-8	-0.06	21	-2.87	15	0.76	27	0.64	23	-0.17	39	2.30	13	0.52	9	-2.37	38	-0.06	25	-0.29	15	225
C.H.T.266-6	-0.16	8	-4.18	22	5.42	6	2.69	7	3.69	8	3.28	7	0.72	6	21.33	5	0.62	7	7.69	5	81
C.H.T.273-38	0.13	58	-1.73	13	0.86	25	0.44	28	0.31	31	2.79	10	0.45	11	18.58	9	0.44	8	2.15	11	204
Pobeda 1	0.10	51	4.33	3	-2.60	63	-1.21	58	-0.71	51	-0.22	36	-0.07	42	11.38	21	-0.75	56	-4.17	23	404
L 17	0.01	33	-8.60	45	1.96	16	0.52	26	0.57	24	2.15	14	0.25	18	5.50	32	0.20	13	-1.98	17	238
Melkin 261	0.11	55	-8.51	43	-0.56	44	0.83	20	-0.39	44	-0.30	38	0.03	34	-12.43	47	-0.90	66	-13.84	53	444
H.T.I	0.03	39	-10.44	51	-0.83	49	-1.24	59	-1.15	56	-0.55	39	-0.09	43	-32.97	57	-0.41	37	-18.19	60	490
triumph	0.07	46	-3.81	19	6.08	3	2.22	11	4.83	2	8.63	1	1.52	1	9.40	25	1.39	3	9.48	4	115
Basma.S.31	0.05	43	-3.07	16	-4.45	70	-2.92	70	-3.52	70	-0.99	46	-0.01	37	2.25	34	-0.78	59	-9.79	40	485
SPT 403	0.11	56	-20.29	65	-1.19	52	-1.64	61	-2.05	64	-1.99	62	-0.34	64	-44.19	67	-0.89	65	-29.92	67	623
SPT 405	0.18	68	-15.60	57	-0.28	39	0.02	42	0.48	25	0.07	35	-0.04	40	-21.79	52	-0.57	52	-23.47	63	473
SPT 406	-0.22	2	-18.52	61	4.40	9	3.71	3	3.48	9	-0.64	41	-0.10	46	-36.64	62	-0.48	44	-22.04	61	338
SPT 408	0.14	61	-19.65	63	-0.85	50	0.63	24	-0.08	36	-1.53	55	-0.20	52	-45.10	68	-1.16	69	-26.75	66	544
SPT 409	-0.05	23	-23.82	71	3.41	11	1.45	16	2.43	12	-3.76	70	-0.59	69	-49.92	69	-0.68	55	-33.58	71	467
SPT 410	0.18	67	-17.70	59	-0.37	42	0.76	21	0.82	21	1.59	17	0.19	23	-27.91	56	-0.38	36	-15.36	56	398
SPT 412	0.21	70	-18.28	60	-2.82	65	-0.23	46	-0.38	43	1.27	20	0.12	27	-35.45	59	-0.52	48	-17.51	58	496
SPT 413	-0.05	24	-22.29	68	-0.73	47	-0.56	52	0.81	22	-4.34	71	-0.74	71	-57.22	70	-1.36	70	-31.58	68	563
SPT 420	0.07	47	-19.42	62	5.55	4	2.42	8	4.06	6	-1.41	53	-0.16	49	-36.10	60	-0.50	45	-22.89	62	396
SPT 430	-0.08	19	-17.36	58	2.74	12	3.32	4	4.22	5	-2.16	64	-0.28	59	-34.03	58	-1.01	67	-23.82	64	410

------- continued

Genotype	CI	Rank	LN	Rank	LL	Rank	LW	Rank	GLW	Rank	GLY	Rank	DLY	Rank	PH	Rank	SD	Rank	D50F	Rank	Sum
SPT 432	0.08	50	-20.12	64	4.93	8	3.92	2	4.22	4	0.57	26	0.09	29	-37.79	63	-0.25	27	-14.33	55	328
SPT 433	-0.07	20	-22.37	69	0.90	24	1.71	14	1.89	15	-3.24	68	-0.53	68	-41.77	65	-1.62	71	-32.76	70	484
SPT 434	0.07	48	-21.28	67	1.25	19	2.26	10	1.90	14	-1.97	61	-0.33	63	-39.57	64	-0.88	64	-32.51	69	479
SPT 436	-0.02	30	-20.76	66	-3.10	67	-0.85	56	-0.85	53	-1.66	57	-0.25	58	-42.65	66	-0.44	39	-17.51	59	551
SPT 439	-0.06	22	-23.40	70	-1.76	58	1.12	17	0.46	26	-1.53	56	-0.21	53	-59.85	71	-1.02	68	-26.51	65	506
SPT 441	-0.05	25	-14.69	56	0.33	28	1.83	13	0.16	32	-1.93	60	-0.23	57	-11.95	45	-0.28	30	-13.28	51	397
P.D.324	-0.10	16	-9.71	49	-0.14	36	0.22	32	0.39	29	-2.14	63	-0.23	56	-6.81	42	-0.83	62	-5.48	28	413
P.D.325	-0.12	12	-6.45	35	-3.91	68	-1.94	66	-2.36	66	-0.99	47	-0.05	41	-14.59	49	-0.77	58	-15.84	57	499
P.D.328	-0.21	3	-7.47	40	-1.72	57	-0.57	53	-1.36	59	-3.07	67	-0.40	65	1.59	35	-0.76	57	-12.18	49	485
P.D.329	-0.25	1	-8.67	46	-2.49	61	-0.73	54	-1.21	57	0.11	34	0.10	28	-8.94	44	-0.27	29	-6.01	29	383
P.D.336	-0.08	18	-4.39	26	-4.93	71	-3.29	71	-3.58	71	-0.94	45	-0.11	47	-7.00	43	-0.51	46	-12.02	47	485
P.D.345	-0.04	27	-4.68	28	-3.99	69	-2.22	68	-3.27	69	0.28	30	0.27	17	15.15	13	-0.85	63	-5.24	26	410
P.D.364	-0.10	17	-4.38	25	-2.26	60	-1.91	64	-2.05	63	-1.08	49	0.04	33	0.47	36	-0.46	40	-8.87	34	421
P.D.371	-0.18	7	-3.11	17	-2.64	64	-1.87	63	-2.17	65	-0.69	43	-0.02	38	19.82	7	-0.46	41	-5.19	25	370
P.D.381	-0.11	13	-7.95	42	-2.02	59	-0.45	50	-1.56	60	-1.74	59	-0.41	66	16.78	12	-0.47	42	-7.95	32	435
Mutant 4	0.00	31	-8.59	44	-1.05	51	0.49	27	0.93	19	3.17	8	0.34	14	10.49	23	-0.05	24	-3.61	22	263
C.H.T.209.12e×F.K.40-1	0.04	40	0.97	9	2.12	15	0.07	37	0.58	23	1.88	16	0.39	13	11.58	20	0.24	12	6.26	6	191
T-B-22	0.14	60	-4.22	23	1.09	21	-0.40	48	-0.42	45	-1.18	51	-0.22	55	-25.46	55	0.01	21	-13.46	52	431
krumovgraid	0.00	32	-4.83	29	0.32	29	0.02	41	-0.49	46	-1.68	58	-0.22	54	6.98	29	-0.04	22	0.93	13	353
Ohdaruma	0.04	41	-11.70	55	8.44	1	5.05	1	6.76	1	4.87	2	1.07	2	12.92	17	0.81	6	1.50	12	138

Cluster analysis on BV data using the K-means algorithm produced 4 groups, which is proved by the Elbow index (Fig. 1). Results showed that the 71 studied tobacco genotypes were classified into four separate groups (Fig. 2). According to the generated dendrogram, the largest number of genotypes (36 genotypes) with the highest and negative values of BV for Cl, GLW, LL, and LW were placed in cluster 4. Cluster 3 (9 genotypes) also had negative as well as the highest values of BVs for GLY, PH, C50F, and SD. As shown in Fig. 2, in most cases, cluster 2 (21 genotypes) had negative and low values of BVs, while cluster 1 (6 genotypes) had positive and highest values of BV for the majority of studied characteristics except for Cl, LN, and PH. Classification of measured characteristics showed that characters LL, LW, and GLW were placed in cluster 1, Cl separately in cluster 2, GLY and DLY in cluster 3, and PH, SD, LN, and C50F in cluster 4 (Fig. 2).

Discussion

Crossing between plant genotypes is an inseparable subject that each plant breeder may face. Hence, the breeder must determine the combining ability of under studies plant genotypes. Albeit there are several methods for identifying the genotypes combining abilities such as diallel top-cross, they are time-consuming techniques (due to crossing), especially when the experimental plant material isin a large scale. Currently, BLUP (best linear unbiased predictor) which firstly was introduced by animal breeders, has been implemented by plant breeders for the estimation of additive genetic effect (breeding value) of interested traits in understudies germplasm as well as cross performance. The efficiency of BLUP (ABLUP)for the prediction of BV related to several characters in crops was proven (Roudbari et al. 2017; Bemejo et al. 2020). About BLUP, it is mandatory to know about the pedigree of the studied genotype to construct a kinship matrix and likewise, in self-pollinated plants (Bauer et al. 2006) the well-done coancestry analysis is confusing. Today, with the appearance and development of molecular markers, this kinship matrix could be estimated by DNA markers and by this way, breeders could overcome the coancestry coefficient analysis problems, especially in self-pollinated plants.The recent approach, named as GBLUP, which implemented in the present work for the prediction of BVs for understudied self-pollinated tobacco individuals. Herein, by using GBLUP genotypes, P. D. 328 as well as P. D. 329 has been detected with superior BVs for Cl character and this finding is paralleled to the results of Darvishzadehand Alavi (2011), which used diallel mating design for estimation of additive genetic effects and identification of best parental

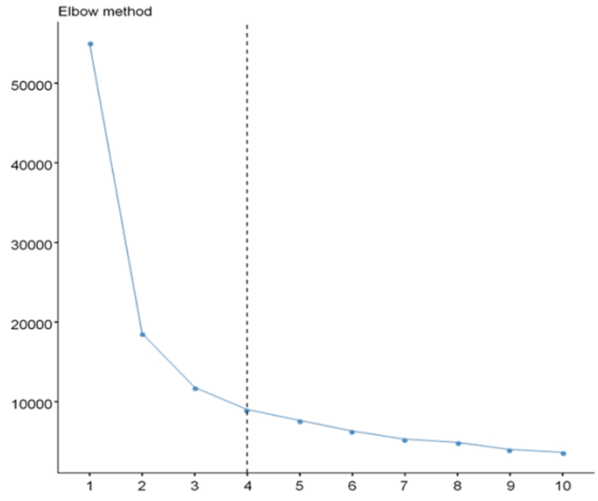

Fig. 1. Determining the optimal number of clusters for genotypes using Elbow method

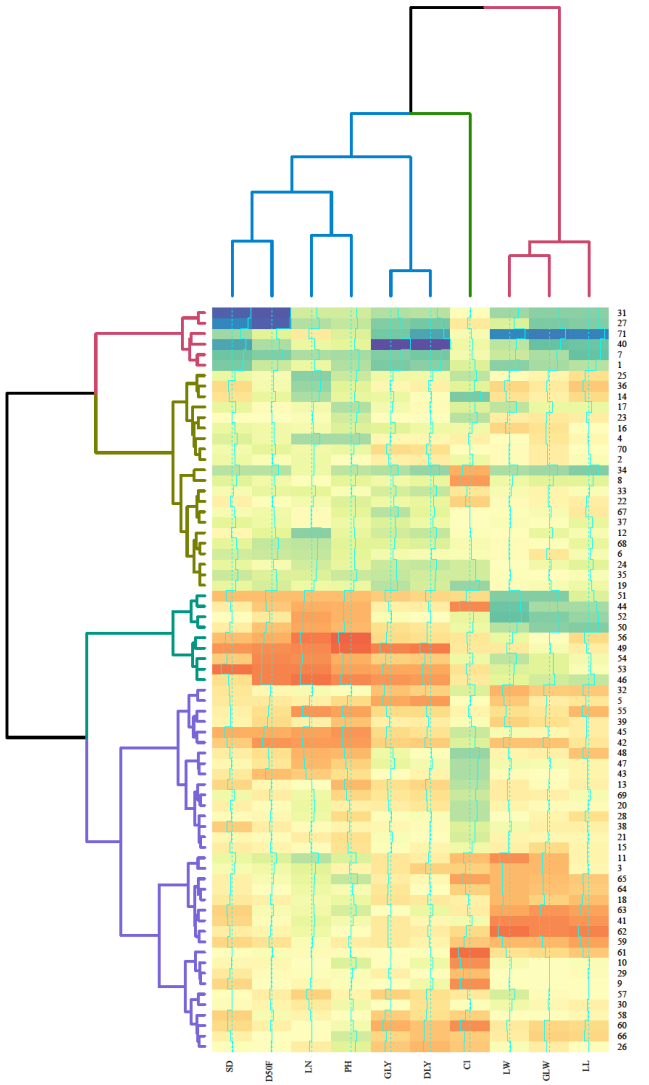

Fig. 2. Dendrogram of 71 tobacco genotypes as well as 11 studied characteristics based on predicted breeding values SD= Stem diameter, D50F, Days to 50% flowering, LN= leaf number, PH = Plant height, GLY = Green leaf yield, DLY = Dry leaf yield, Cl = Chlorine, LW = Leaf width, GLW = One green leaf weight, and LL = Leaf length

lines in oriental-type tobacco. Similarly, there are also reports (Patel et al. 2012; Seyyed Nazari et al. 2016) that used diallel crossing system to predict the GCA (general combining ability) of genotypes for agro-morphological characters in oriental and semi oriental-type tobacco. For instance, Seyyed Nazari et al. (2016) revealed that genotypes, Kromovgraid for dry weight of leaf, number of leaves, and length of stem, B.S.31 for fresh weight of leaf and length of stem, SPT 406 and SPT 410 for width of leaf, G.D. 165 for number of leaves and Xanthi for diameter of stem are the best parents according to their GCA values. Interestingly, the above-mentioned tobacco genotypes were also inspected in the present study and similarly recognized with moderate and acceptable BV values through the GBLUP approach. In plant breeding, parental line selection has been done with two aims, including identifying suitable parents for commercial hybrid varieties as well as identifying suitable parents to develop inbred lines for subsequent breeding cycles (Chung and Liao 2022). For achieving the first goal in tobacco, It is concluded that the use of genomic selection in tobacco can decrease cycle time and costs in hybrid breeding, particularly by rapidly establishing heterotic pools (involving distant genotypes), reducing testcrossing, and limiting the loss of genetic variance as also observed earlier by Labroo et al. (2021).

Considering the 11 agro-morphologic characters accompanied by leaf Cl content as well as the sum of ranks for each genotype, the genotypes C.H.T.269-12e, C.H.T.266-6, SS 289-2, C.H.T.209.12e, Triumph, and Ohdaruma by having remarkable BVs can efficiently transfer their genotypic values to the next generation (Piepho *et al.* 2008) and would be selected as potentially parental genotypes to develop new populations in tobacco breeding programs. In this project, character classification represents the ability of BVs as suitable discriminators, which accurately separates highly heritable characters like morphological traits from other ones (chemical characteristics). So, it seems that BV prediction will be well done for each quantitative and quality characteristic of tobacco, and BV directly reflects its genetic effects (Villumsen and Janss 2009). As a result, identified heterotic groups of studied tobacco germplasm will be efficiently used as progenitors in the construction of the mapping population because of the transmission of parental distance to the progeny and the establishment of good segregation. Similarly, Hatami Maleki *et al.* (2013) screened tobacco germplasm through simple phenotypic values. They crossed some parental genotypes for establishing mapping populations and reported narrow genetic variability among mapping individuals, which led to some problems in QTL analysis. Therefore, future research must incorporate genomic tools into tobacco breeding programs to accelerate genetic improvement and ensure the continued economic and medicinal significance of this vital crop.

Supplimentary material

Supplimentary Table S1 is presented which can be accessed at www.isgpb.org.

Authors' contribution

Conceptualization of research (RD, HHM); Designing of the experiments (RD, HHM); Contribution of experimental materials (PGM, RD, BMZ); Execution of field/lab experiments and data collection (PGM, RD); Analysis of data and interpretation (RD, HHM, HZT); Preparation of the manuscript (RD, HHM).

Acknowledgments

The authors wish to express their sincere gratitude to the Urmia Tobacco Research Institute of Iran for their invaluable support and resources provided throughout the course of this study. Additionally, our thanks extend to Urmia University for their academic support and collaboration.

References

Akehurst B. C. 1981. Tropical Agricultural Series (Vol. 164). New York: Longman Inc. p.

Bauer A. M., Reetz T. C. and Léon J. 2006. Estimation of breeding values of inbred lines using best linear unbiased prediction (BLUP) and genetic similarities. Crop Sci., **46**(6): 2685–2691. doi:10.2135/cropsci2006.01.0019

Berlowitz I., Torres E. G., Walt H., Wolf U., Maake C. and Martin-Soelch C. 2020. Tobacco is the chief medicinal plant in my work: Therapeutic uses of tobacco in Peruvian Amazonian medicine exemplified by the work of a maestro tabaquero. Front. Pharmacol., **11**: 594591.

Bermejo C., Cazzola F., Maglia F. and Cointry E. 2020. Selection of parents and estimation of genetic parameters using BLUP and molecular methods for lentil (*Lens culinaris* Medik.) breeding program in Argentina. Exp. Agricul., **56**(1): 12–25. doi:10.1017/s0014479719000061

Bernardo R. 1993. Estimation of coefficient of coancestry using molecular markers in maize. Theor. Appl. Genet., **85**(8): 1055–1062. doi:10.1007/BF00215047

Bernardo R. 1994. Prediction of maize single-cross performance using RFLPs and information from related hybrids. Crop Sci., **34**(1): 20–25. doi:10.2135/cropsci1994.0011183x003400010003x

Bernardo R. and Yu J. 2007. Prospects for genome wide selection for quantitative traits in maize. Crop Sci., **47**: 1082–1090.

Bindler G., Plieske J., Bakaher N., Gunduz I., Ivanov N., Van der Hoeven R. and Donini P. 2011. A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. Theor. Appl. Genet., **123**(2): 219–230. doi:10.1007/s00122-011-1578-8

Cantelmo N. F., Von Pinho R. G. and Balestre M. 2017. Genome-wide prediction for maize single-cross hybrids using the GBLUP model and validation in different crop seasons. Mol. Breed., **37**(4): doi: 10.1007/s11032-017-0651-7

Chung P. Y. and Liao C. T. 2022. Selection of parental lines for plant breeding via genomic prediction. Front. Plant Sci., **13**: 934767.

Chaplin J. F. 1975. Genetic influence on chemical constituents of tobacco leaf and smoke. BeiträgeZurTabakforschung International/Contributions to Tobacco Research, **8**(4): doi:

10.2478/cttr-2013-0385

Chari M. S. 1995. Role of research in the improvement of productivity and quality of Indian flue-cured Virginia tobacco (Vol. 26). Rajahmundry Indian; Central Tobacco Research Institute.

Chung P.-Y. and Liao C.-T. 2020. Identification of superior parental lines for biparental crossing via genomic prediction. PLOS ONE, **15**(12): e0243159. doi:10.1371/journal.pone.0243159

Darvishzadeh, Reza, Alavi S. R. and Sarafi A. 2011. Genetic variability for chlorine concentration in oriental tobacco genotypes. Archiv Fur Acker- Und Pflanzenbau Und Bodenkunde, **57**(2): 167–177. doi:10.1080/03650340903286414

Darvishzadeh R. and Alavi R. 2011. Genetic analysis of chloride concentration in oriental tobacco genotypes. J. Pl. Nutri., **34**(7): 1070–1078. doi:10.1080/01904167.2011.555588

Davis D. L. and Nielson M. T. 1999. Tobacco: Production, Chemistry and Technology. Blackwell Sci., 467.

Dellaporta S. L. and Hicks W. J. 1983. A plant DNA mini preparation: version II. Plant Mol. Biol. Reporter, **1**: 19–21.

Falconer D. S. 2022. Introduction to Quantitative Genetics (classic reprint). London, England: Forgotten Books.

Guardiola J. M., Perez O. and Diaz L. 1987. Effect of chlorine and potassium on combustibility from fine plantations. *T*obacco, **10**: 29–43.

Hatami M., Hamid, Karimzadeh G., Darvishzadeh R., Naghavi M., Reza and Sarrafi A. 2013. Identification of QTLs associated with low chloride accumulation in oriental tobacco. Genetika, **45**(3): 855–864. doi:10.2298/gensr1303855m

Henderson C. R. 1985. Best linear unbiased prediction of nonadditive genetic merits in non-inbred populations. J. Animal Sci., **60**: 111–117.

Kassambara A. and Mundt F. 2020. Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R Package Version 1.0.7. https://CRAN.R-project.org/package=factoextra

Labroo M. R., Studer A. J. and Rutkoski J. E. 2021. Heterosis and hybrid crop breeding: a multidisciplinary review. Front. Genet., **12**: 643761. DOI: 10.3389/fgene.2021.643761

Lencucha R., Drope J., Magati P. and Sahadewo G. A. 2022. Tobacco farming: overcoming an understated impediment to comprehensive tobacco control. Tobacco Control, **31**(2): 308-312.

McEvoy E. T. 1957. The growth and mineral content of flue-cured tobacco as influenced by reaction of nutrient solutions with ionic forms of nitrogen. Canad. J. Soil Sci., **37**(2): 79–83. doi:10.4141/cjss57-012

Mirkarimi S. R., Ardakani Z. and Rostamian R. 2021. Economic and environmental assessment of tobacco production in Northern Iran. Ind. Crops Products, **161**: 113171.

Patel J. N., Patel A. D., Chakraborty S. and Sasidharan N. 2012. Heterosis and relative heterosis in tobacco (*Nicotiana tobacco* L.). Animal Env. Sci., **2**: 270–275.

Patterson H. D. and Thompson R. 1971. Recovery of Inter-Block Information when Block Sizes are Unequal. Biometrika, **58**(3): 545. doi: 10.2307/2334389

Piepho H. P., Möhring J., Melchinger A. E. and Büchse A. 2008. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica, **161**(1–2): 209–228. doi: 10.1007/s10681-007-9449-8

Quintal S. S. R., Viana A. P., Campos B. M., Vivas M. and Amaral Júnior A. T. D. O. 2017. Selection via mixed models in segregating guava families based on yield and quality traits. RevistaBrasileira de Fruticultura, **39**(2): doi: 10.1590/0100-29452017866

Seyyed-Nazari R., Ghadimzadeh M., Darvishzadeh R. and Alavi S. R. 2016. Diallel analysis for estimation of genetic parameters in oriental tabacco genotypes. Genetika, **48**: 125–137.

Tso T. C. 2006. Tobacco research and its relevance to science, medicine and industry. BeiträgeZurTabakforschung International/Contributions to Tobacco Research, **22**(3): 133–145. doi: 10.2478/cttr-2013-0824

Villumsen T. M. and Janss L. 2009. Bayesian genomic selection: the effect of haplotype length and priors. BMC Proceedings, 3 Suppl 1, S11. doi: 10.1186/1753-6561-3-s1-s11

Yang B. C., Xiao B. G., Xj C. and Shi C. H. 2007. Assessing the genetic diversity of tobacco germplasm using inter simple sequence repeat and inter-retro-transposon amplification polymorphism markers. Ann. Appl. Biol., **50**: 393–433.

Yang R.-C. 2010. Towards understanding and use of mixed-model analysis of agricultural experiments. Canad. J. Plant Sci., **90**(5): 605–627. doi:10.4141/cjps10049

Vineeth T.V., Prasad I. , Chinchmalatpure A.R., Lokeshkumar B. M., Kumar S., Ravikiran K. T. and Sharma P.C. 2022. Weighted average absolute scores of BLUPs (WAASB) based selection of stable Asiatic cotton genotypes for the salt affected Vertisols of India. Indian J. Genet. Plant Breed., **82**(1): 104-108. doi: 10.31742/IJGPB.82.1.15

Slater A.T., Cogan N.O.I., Forster J.W., Hayes B.J. and Daetwyler H.D. 2016. Improving genetic gain with genomic selection in autotetraploid potato. The Plant Genome, **9**(3): https://doi.org/10.3835/plantgenome2016.02.002

Sood S., Bhardwaj V., Kaushik S. K., Sharma S., Dalamu, Lal M. and Kumar M. 2022. Identification of superior parental lines based on BLUP breeding values in potato (*Solanum tuberosum* L.). Indian J. Genet. Plant Breed., **82**(2): 200-207. doi: 10.31742/IJGPB.82.2.9

Sood S., Lin Z., Caruana B., Slater A.T. and Daetwyler H.D. 2020. Making the most of all data: combining non-genotyped and genotyped potato individuals with HBLUP. The Plant Genome, **13**(3): e20056.doi: 10.1002/tpg2.20056

Supplementary Table S1. Sequences, size and linkage group of SSR primers

Primer	5¢®3¢	(bp)Duplication Fragments Size	Linkage Group
PT30014	F: TGCCGTGTAAATTTCATTTGG R: AGGATTCCTAACGTGTATTATGTTCT	205	11
PT30172	AAACAACGTCGAAGCATTTG ACGCATGAAATTGTAAGGGC	216	4
PT30202	TCGAAACCTCGAGGACAGTT TATCCAAATCTCCAAAGCCC	225	7
PT30250	GAACACACGTTCGTCATTGG ATAAGTCCCTTTAATTTAATTGCG	177	10
PT30165	ACCTCTGTGGCCGTAAGCTA CCTCTACTTCAACAGGGTAAGAAA	224	19
PT30241	AAGTCTCGTGTGGTTGCTTT AAAGGGCAATGTGTCTAGCTC	199	15
PT30027	CCGAGAGTTGCATTTGAATTT AGGGTTCTACGCAAGAGATTG	225	13
PT30021	CATTTGAACATGGTTGGCTG CTCAACTCTCGTCGCTCTTG	224	4
PT30034	GACGAAACTGAGGATATTCCAAA TGGAAACAAAGCCATTACCC	216	22
PT20343	GGAACACCACCACCATAA GGAGCTCAGGTTCCAATG	322	4
PT30285	CATCATGGCAAGTCACCATC TGCTGGAAATTAGCGAGGTT	177	18
PT30126	GTGATTCCAGCGGAAGACAT TTCGAAATAAGTACCTAGAGTCGG	208	10
PT30008	CGTTGCTTAGTCTCGCACTG GGTTGATCCGACACTATTACGA	192	11
PT30159	GCATGCATATGAACATGGGA TTTGACATCTCTACTCTTCCGTTT	197	14b
PT30205	GGTCGATCCACAATTTAAACG GCACTTGCTCCTTTGTACCC	193	3b
PT30260	GGTAGGGTGGAACAAATTTATCA AATATGGTCTATGCCCGCAA	225	8a
PT30292	AAGACAGATTGGTGCGGAAC AGCACTTGGACAGGCGAATA	156	7
PT30319	ACAACAACTACGTTAGTGTGAGAAA TCATGTGTGCCAAGCTCTTC	181	11
PT30324	TGCTCTGCGTTAGAACAGGA CGACGAGAGAAGATTAGTGAAAGA	151	12
PT30046	GATAGGTAGATTATCCTCTGCAACA GGTGCTAGCAACATCATCAAA	182	13
PT30061	TCGTCCATTTCTTTCTCTCTCA CATAAATAGTTGCTCATTCAATCG	182	11
PT30067	AAGCCTGGTCAGTTATCCCA ATTCGCACCACTTAATCCCA	204	2
PT30075	CGATCGGGTCGTTACACAAT CCCATCAGGTTGTTGGGTTA	195	11
PT30094	AACAAGAACGACGGTTACGC GGGTCATGCGTTCGAATTAT	201	18
PT30110	TTGTACGTTCCTCGCTGATG GGCCGACAATAAAGTGGCT	213	21
PT30132	CCTAACAGCATTTGCTACCCA GATGGACAAGAGTGGCCTTT	216	10