

GENETIC DIVERSITY AND MULTIVARIATE ANALYSIS OF YIELD PERFORMANCE TRAITS IN POTATO

Vikas Mangal¹, Salej Sood^{1*}, Vinay Bhardwaj², Bhawna Dipta¹, Mehi Lal³, Kritika Singh¹, RK Singh³, Vinod Kumar¹ and Brajesh Singh¹

ABSTRACT: The present study investigated the genetic diversity and yield performance traits in 332 potato accessions evaluated in 2020 and 366 accessions in 2021 at ICAR-Central Potato Research Institute, Regional Station, Modipuram. An augmented design, was used to assess the marketable tuber yield (MTY), non-marketable tuber yield (NMTY), total tuber yield (TTY), specific gravity (SG), and plant maturity (PM). Significant genetic variation was observed, with heritability values exceeding 90% for key traits like TTY and MTY. Principal component analysis (PCA) identified marketable and total tuber yield as primary contributors to variability, which explained 63.2% of the total variance. The formation of four distinct clusters (red, blue, green, and black) in the present investigation highlights the presence of subgroups with shared traits. Hierarchical clustering highlighted the potential of combining the yield related traits of Tuberosum accessions, the genetic diversity of Andigena accessions (*Solanum tuberosum* Andigena group), and the regional adaptability of Indian varieties to develop superior potato cultivars with improved yield. Key accessions, including JEX/A-459 and Kufri Pushkar, exhibited superior yield performance, which can be a promising candidate for breeding programs.

KEYWORDS: Genetic diversity, marketable tuber yield, specific gravity, principal component analysis, hierarchical clustering.

Abbreviations: PM – Plant maturity, MT – Marketable tuber number, NMT – Non-Marketable tuber number, MTY – Marketable tuber yield, NMTY – Non-Marketable tuber yield, TTY – Total tuber yield, SG – Specific gravity.

INTRODUCTION

Potato (*Solanum tuberosum* L., $2n = 4x = 48$) originated from the Titicaca Lake basin in the Peruvian and Bolivian Andes, is the world's most important non-grain food crop after rice and wheat (Sood *et al.*, 2024). This crop contributed vital role in food

security and serves as a significant source of carbohydrates, vitamins, and minerals for millions of people worldwide (Mangal *et al.*, 2022).

Along with its important role in food security, potato have unique genetic traits and a rich cultural history, which makes this crop valuable for both agricultural and genetic research (Datta *et al.*, 2015). In India, potato cultivation holds a crucial role in agriculture, which contributes significantly to both domestic consumption and the agricultural economy. The growing demand for high

*Corresponding author; email: salej1plp@gmail.com

¹ICAR-Central Potato Research Institute, Shimla - 171 001, Himachal Pradesh, India

²ICAR-National Research Centre on Seed Spices, Ajmer, India

³ICAR-Central Potato Research Institute, Regional Station, Modipuram, Meerut - 250110, Uttar Pradesh, India

yielding varieties with improved tuber quality traits, and challenges such as climate change, pests and diseases, highlights the need for superior potato varieties through continuous breeding efforts.

Genetic variability is backbone for breeding potato varieties with improved disease resistance, better environmental adaptability and high yield (Mangal *et al.*, 2024). It provides new alleles for complex traits and strengthen the genetic base for sustainable crop development (Deperi *et al.*, 2018). For achieving this all our breeding activities mainly depends on the genetic diversity within available germplasm, which needs comprehensive evaluation of their phenotypic and genotypic diversity (Esnault *et al.*, 2014). Phenotypic diversity present among the cultivated potato varieties or germplasm is essential for breeding and adaptation to diverse environments (Abebe *et al.*, 2013). Evaluating phenotypic variation among potato accessions based on agro-morphological traits and yield attributes is essential for identifying valuable genetic traits and enhancing strategies for adaptation (Ahmadzadeh and Felenji 2011). Despite a wide genetic diversity available in the potato germplasm, most cultivated varieties have a narrow genetic base, due to prolonged selective breeding focus on few traits only (Bhardwaj *et al.*, 2023). Breeding programs aim to improve potato yield and quality by selecting accessions with desirable agronomic traits, such as high marketable or total tuber yield, diseases resistance, and adaptability to diverse environmental conditions. The genetic improvement of potatoes is complex due to its heterozygous and tetraploid nature, which complicates the inheritance patterns of key yield contributing traits. However, the extensive genetic diversity within potato populations offers significant opportunities and challenges for breeders seeking to integrate multiple desirable traits into new varieties.

To tackle all these challenges, a comprehensive evaluation of a large number of genotypes is essential. In the augmented design, check varieties are replicated across blocks, which enable the evaluation of a large number of accessions with reduced resources and useful in preliminary breeding trials to identify promising accessions for further testing. Principal component analysis (PCA) is utilized to explore the relationships among the studied traits and to identify variability in the dataset. Additionally, hierarchical clustering is employed to assess the genetic and phenotypic diversity among the accessions, which enable the identification of distinct clusters that could serve for future breeding efforts.

The present study provides valuable insights into the genetic diversity and yield potential of a diverse collection of potato accessions, which provides critical information for breeding programs focusing on developing high yielding new varieties. The use of statistical methods, such as PCA and hierarchical clustering, facilitates a deeper understanding of the complex relationships among traits and help in the identification high-yielding potato accession to enhance food security.

MATERIAL AND METHODS

Experimental site

The experiment was carried out at the ICAR-Central Potato Research Institute (CPRI), Regional Station, Modipuram, Meerut, during the rabi seasons (October to February) of 2020 and 2021. The geographical location of the station is 29.033°N latitude, 77.683°E longitude, with an elevation of approximately 237 meters above sea level. The region experiences a subtropical climate, characterized by hot summers and cool winters, which is suitable for potato cultivation.

An augmented design was used to evaluate a large number of accessions efficiently: 332 accessions in year 2020, 366 in year 2021, with 319 accessions common during both the years (Table 1). In the year 2020 studies, blocks 1 to 7 each comprised of 42 test accessions along with 4 replicated check varieties, whereas block 8 included 34 accessions and 4 checks. In the year 2021 experiment, blocks 1, 2, 4, 5 and 6 comprised of 46 test accessions and 4 checks, block 3 contained 45 test accessions and 4 checks, block 7 included 43 test accessions and 4 checks, and block 8 comprised 44 test accessions and 4 checks. All accessions used

were obtained from the ICAR-Central Potato Research Institute (ICAR-CPRI), where they are conserved in the *in-vitro* form. A total of four check varieties were used, Kufri Jyoti, Kufri Chipsona-1, Kufri Himalini, and Kufri Girdhari. The accessions used in this study belonged to three groups: CP (*Solanum tuberosum*, Tuberosum group) for high-yielding cultivated lines, JEX-A (*Solanum tuberosum*, Andigena group) for genetic diversity and stress tolerance, and Kufri (Indian varieties) bred for regional adaptability. Standard agronomic practices recommended for potato cultivation in the region were followed

Table 1. List of potato accessions used in the study.

S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.
1	CP1012	26	CP1347	51	CP1462	76	CP1642	101	CP1753	126	CP1982	151	CP2224	176	CP3079
2	CP1038	27	CP1348	52	CP1468	77	CP1646	102	CP1764	127	CP1988	152	CP2235	177	CP3081
3	CP1137	28	CP1367	53	CP1470	78	CP1653	103	CP1767	128	CP1989	153	CP2284	178	CP3096
4	CP1140	29	CP1368	54	CP1479	79	CP1659	104	CP1784	129	CP1990	154	CP2285	179	CP3098
5	CP1143	30	CP1379	55	CP1480	80	CP1662	105	CP1800	130	CP2010	155	CP2321	180	CP3102
6	CP1151	31	CP1390	56	CP1486	81	CP1664	106	CP1802	131	CP2011	156	CP2324	181	CP3103
7	CP1157	32	CP1395	57	CP1491	82	CP1667	107	CP1806	132	CP2029	157	CP2335	182	CP3106
8	CP1159	33	CP1399	58	CP1529	83	CP1668	108	CP1824	133	CP2030	158	CP2338	183	CP3116
9	CP1175	34	CP1402	59	CP1533	84	CP1669	109	CP1827	134	CP2049	159	CP2339	184	CP3124
10	CP1177	35	CP1405	60	CP1538	85	CP1672	110	CP1835	135	CP2059	160	CP2346	185	CP3145
11	CP1187	36	CP1411	61	CP1544	86	CP1673	111	CP1846	136	CP2061	161	CP2347	186	CP3153
12	CP1215	37	CP1414	62	CP 1545	87	CP1674	112	CP1854	137	CP2065	162	CP2348	187	CP3171
13	CP1218	38	CP1418	63	CP1553	88	CP1685	113	CP1864	138	CP2071	163	CP2350	188	CP3173
14	CP1225	39	CP1420	64	CP1555	89	CP1687	114	CP1868	139	CP2086	164	CP2364	189	CP3180
15	CP1235	40	CP1426	65	CP1559	90	CP1688	115	CP1869	140	CP2089	165	CP2368	190	CP3182
16	CP1246	41	CP1427	66	CP1564	91	CP1693	116	CP1871	141	CP2090	166	CP2370	191	CP3183
17	CP1263	42	CP1428	67	CP1571	92	CP1700	117	CP1873	142	CP2110	167	CP2385	192	CP3189
18	CP1291	43	CP1431	68	CP1581	93	CP1706	118	CP1881	143	CP2118	168	CP2390	193	CP3192
19	CP1302	44	CP1433	69	CP1584	94	CP1710	119	CP1884	144	CP2134	169	CP2412	194	CP3201
20	CP1304	45	CP1435	70	CP1588	95	CP1711	120	CP1889	145	CP2142	170	CP2418	195	CP3203
21	CP1319	46	CP1440	71	CP1597	96	CP1730	121	CP1918	146	CP2149	171	CP2927	196	CP3211
22	CP1325	47	CP1447	72	CP1602	97	CP1735	122	CP1922	147	CP2165	172	CP3036	197	CP3222
23	CP1326	48	CP1450	73	CP1616	98	CP1736	123	CP1926	148	CP2171	173	CP3044	198	CP3246
24	CP1330	49	CP1453	74	CP1619	99	CP1747	124	CP1971	149	CP2183	174	CP3068	199	CP3247

Genetic diversity and multivariate analysis of potato yield

S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.	S. No.	Accession No.
25	CP1335	50	CP1454	75	CP1633	100	CP1749	125	CP1974	150	CP2189	175	CP3072	200	CP3256
201	CP3261	226	CP3549	251	CP3792	276	CP4214	301	JEX/A-202	326	JEX/A-468	351	JEX/A-93	376	K. Sadabahar
202	CP3274	227	CP3575	252	CP3795	277	CP4224	302	JEX/A-21	327	JEX/A-498	352	JEX/A-947	377	Kufri Sindhuri
203	CP3295	228	CP3577	253	CP3796	278	CP4242	303	JEX/A-215	328	JEX/A-506	353	JEX/A-99	378	K. Surya
204	CP3296	229	CP3585	254	CP3797	279	CP4254	304	JEX/A-22	329	JEX/A-513	354	KCM	379	K. Swarna
205	CP3318	230	CP3587	255	CP 3799	280	CP4256	305	JEX/A-232	330	JEX/A-539	355	Kufri Anand		
206	CP3328	231	CP3588	256	CP3809	281	CP4311	306	JEX/A-26	331	JEX/A-58	356	Kufri Arun		
207	CP3329	232	CP3600	257	CP3816	282	CP4316	307	JEX/A-267	332	JEX/A-597	357	Kufri Ashoka		
208	CP3334	233	CP3625	258	CP3853	283	CP4398	308	JEX/A-275	333	JEX/A-612	358	Kufri Badshah		
209	CP3352	234	CP3632	259	CP3867	284	CP4593	309	JEX/A-288	334	JEX/A-638	359	Kufri Bahar		
210	CP3362	235	CP3634	260	CP3871	285	CP4594	310	JEX/A-296	335	JEX/A-668	360	K. Chipsona-1		
211	CP3363	236	CP3636	261	CP3880	286	CP4596	311	JEX/A-298	336	JEX/A-683	361	K. Chipsona-3		
212	CP3412	237	CP3639	262	CP3881	287	CP655	312	JEX/A-299	337	JEX/A-705	362	Kufri Dewa		
213	CP3414	238	CP3641	263	CP3885	288	CP658	313	JEX/A-30	338	JEX/A-707	363	Kufri Frysona		
214	CP3442	239	CP3646	264	CP3891	289	CP659	314	JEX/A-316	339	JEX/A-708	364	Kufri Garima		
215	CP3443	240	CP3651	265	CP3893	290	JEX/A-10	315	JEX/A-317	340	JEX/A-763	365	Kufri Gaurav		
216	CP3475	241	CP3652	266	CP3894	291	JEX/A-1016	316	JEX/A-32	341	JEX/A-79	366	Kufri Girdhari		
217	CP3486	242	CP3679	267	CP3898	292	JEX/A-1038	317	JEX/A-329	342	JEX/A-801	367	Kufri Himalini		
218	CP3491	243	CP3681	268	CP3901	293	JEX/A-1046	318	JEX/A-361	343	JEX/A-804	368	Kufri Jyoti		
219	CP3502	244	CP3690	269	CP3903	294	JEX/A-1081	319	JEX/A-379	344	JEX/A-827	369	Kufri Kanchan		
220	CP3505	245	CP3696	270	CP3917	295	JEX/A-1092	320	JEX/A-380	345	JEX/A-865	370	Kufri Karan		
221	CP3506	246	CP3718	271	CP3939	296	JEX/A-1152	321	JEX/A-390	346	JEX/A-877	371	Kufri Khyati		
222	CP3511	247	CP3738	272	CP4052	297	JEX/A-15	322	JEX/A-42	347	JEX/A-907	372	Kufri Lauvkar		
223	CP3525	248	CP3761	273	CP4096	298	JEX/A-164	323	JEX/A-45	348	JEX/A-912	373	Kufri Megha		
224	CP3527	249	CP3763	274	CP4149	299	JEX/A-19	324	JEX/A-457	349	JEX/A-918	374	Kufri Pukhraj		
225	CP3529	250	CP3768	275	CP4179	300	JEX/A-197	325	JEX/A-459	350	JEX/A-920	375	Kufri Pushkar		

throughout the growing seasons. Each accession was grown in a single-row plot of 3m length, with a row-to-row spacing of 60 cm and plant-to-plant spacing of 20 cm, which accommodate 15 plants per plot. The replicated check varieties were distributed across the blocks for statistical adjustments.

Crop management: The crop was raised in the winter season (October - February). Fields were ploughed and levelled before planting. The farmyard manure @ 10-15 t/ha and N:P:K fertilizer was applied @180:80:100 kg/ha during both the years. Full FYM, full P and K while half N fertilizer were applied at the time of field preparation. Rest half of the nitrogen was applied at the time of earthing up, 30-40 days after planting. Well sprouted tubers of size 40-60 g were planted manually across the locations. Pre-emergence herbicide, metribuzin @ 0.7-1.0 g/l was applied for weed control followed by one manual weeding within a month after planting. Pre-sowing irrigation was given for uniform germination followed by four irrigations at 10-15 days interval as per the need and soil type. The crop was sprayed with insecticides and fungicides to manage the insect pests and late blight disease at 15-20 days interval thrice after earthing up. The dehauling and harvesting were done at 90 and 105 days after planting, respectively.

Data collection and observations

1. **Marketable tuber (MT) number:** Tubers were harvested at maturity, sorted based on market standards (typically size), and counted to determine the number of marketable tubers (MT).
2. **Marketable Tuber Yield (MTY):** MT were then weighed to measure the yield, which was recorded in kg per plot basis and converted to tons per hectare.
3. **Non-Marketable tuber (NMT, no./plant):** Tubers at the time of harvesting which

did not meet proper size and weight were classified as non-marketable.

4. **Non-Marketable Tuber Yield (NMTY):** NMT tubers were counted, weighed, and the yield was recorded in kg per plot basis and extrapolated to tons per hectare.
5. **Total Tuber Yield (TTY):** The total tuber yield was determined by summing the marketable and non-marketable tuber yields, recorded in kg per plot basis and converted to tons per hectare.
6. **Specific Gravity (SG):** This trait was measured using the standard water displacement method, wherein tuber samples were immersed in water, and the volume of displaced water was measured. The formula used for specific gravity was:
$$\text{Specific gravity} = \frac{\text{Weight of tubers in air}}{\text{Weight of tubers in air} - \text{Weight of tubers in water}}$$
7. **Plant Maturity (PM):** Physiological maturity was determined by observing the natural senescence of the foliage at 90 days after planting. Maturity was recorded as a numerical percentage, which represent the proportion of foliage that had senesced.

Data Analysis

The analysis was performed using various R packages essential for multivariate data analysis and visualization. The *FactoMineR* package (Husson *et al.*, 2016) was used for PCA and *factoextra* (Kassambara & Mundt, 2017) helped to visualize the results. Package *ggplot2* (Wickham & Wickham, 2016) was used to generate quality graphs, and *corrplot* (Wei *et al.*, 2013) was used for visualizing correlation matrices. Data manipulation and cleaning were done with *tidyverse* (Wickham *et al.*, 2016) and *dplyr* (Wickham & Wickham, 2020) package in R.

RESULTS

The use of an augmented design in this experiment facilitated the evaluation of a large number of potato accessions with minimum resources and also maintained statistical robustness. The incorporation of replicated check varieties among the blocks enabled accurate variance estimation and adjustment for block effects. The results also revealed significant differences among test genotypes and checks, which demonstrated the effectiveness of this design in distinguishing superior accessions. The ANOVA demonstrated significant variation among treatments, which indicates substantial phenotypic differences in the evaluated accessions (Table 2). Significant block effects indicated environmental heterogeneity across blocks, which was effectively adjusted for result accuracy. The checks vs. accessions comparison showed the presence of superior accessions compared to standard checks (Kufri Jyoti, Kufri Chipsona-1, Kufri Himalini and Kufri Girdhari).

The genetic parameters for yield and its contributing traits in 2020 and 2021 revealed significant variability for key traits. In 2020, total tuber yield (TTY) exhibited a high genotypic coefficient of variation (GCV) of 36.03%, phenotypic coefficient of variation (PCV) of 36.58%, and a low environmental coefficient of variation (ECV) of 6.36%, with heritability (H^2) of 96.98% and genetic advance as a percentage of mean (GAM) of 73.19%, which indicate strong genetic influence and improvement potential (Table 4). The highest TTY was recorded in the accession JEX/A-459 (24.23 t/ha) and the lowest in JEX/A-21 (1.73 t/ha) (Table 3). Other accessions which exhibited high yield were CP2086 (24.15 t/ha) followed by CP3898 (23.95 t/ha), CP2142 (23.81 t/ha) and CP1330 (23.4 t/ha). Similarly, marketable tuber yield (MTY) showed a GCV of 40.6%, PCV of 41.12%, ECV of 6.52%, with $H^2 = 97.49\%$ and

Table 2. Analysis of variance (ANOVA) for agronomic traits across two years (2020 and 2021).

Source	Df	MT		MTY		NMT		NMTY		SG		TTY		PM	
		2020	2021	2020	2021	2020	2021	2020	2021	2020	2021	2020	2021	2020	2021
Treatment (ignoring Blocks)	331	365	2561.79**	3238.97ns	22.81**	15.9**	4236.58**	11141.83**	0.64**	0.9**	8.2e-05ns	8.1e-05**	22.76**	15.9**	17.54**
Treatment: Check	3	3	640.2*	4462.95ns	8.25**	5.57*	1849.08**	3220.83ns	0.38ns	0.18ns	0.0014**	0.00074**	9.21**	5.25*	27.08**
Treatment: Test vs. Check	1	1	56.91ns	40164.02**	211.08**	324.26**	1742.9**	42931.69**	2.53**	1.11ns	3.9e-05ns	0.00019**	259.84**	287.41**	137.82**
Treatment: Test	327	361	2587.08**	3126.51ns	22.37**	15.13**	4266.11**	11119.59**	0.64**	0.91**	6.9e-05ns	7.6e-05**	22.16**	15.23**	17.12**
Block (eliminating Treatments)	7	7	340.71ns	2547.17ns	0.38ns	6.19**	1284.21**	1526.79ns	0.64**	0.3ns	5e-05ns	9.9e-06ns	1.58ns	5.26**	5.36ns
Residuals	21	21	174.75	2533.42	0.56	1.4	209.65	1126.19	0.16	0.26	7.20e-05	2.10e-05	0.67	1.27	3.27

GAM = 82.69%. Accession CP 2142 exhibited highest MTY (22.94 t/ha), whereas andigena accession JEX/A-865 recorded the lowest (0.32 t/ha) MTY. The non-marketable tuber yield (NMTY) exhibited GCV = 50.72%, PCV = 58.3%, and ECV = 28.76%, with moderate heritability (75.67%) and GAM = 91.01%.

In 2021, a slight reduction in genetic control was observed. TTY had a GCV of 33.0%, PCV of 34.46%, and ECV of 9.94%, with H^2 of 91.69% and GAM = 65.19%. Indian variety Kufri Pushkar recorded the highest TTY (23.98 t/ha) and the andigena accession JEX/A-708 the lowest (2.08 t/ha). Likewise, other accessions which showed high yield were CP1989 (22.91 t/ha), CP3641 (22.28 t/ha) and Kufri Khyati (19.68 t/ha). MTY demonstrated a GCV of

38.92%, PCV of 40.86%, and ECV of 12.43%, with H^2 = 90.75% and GAM = 76.49%. Kufri Pushkar showed the highest MTY (22.0 t/ha), while CP 2011 exhibited the lowest (1.3 t/ha) MTY. The NMTY exhibited a GCV of 44.63%, PCV of 52.75%, and ECV of 28.12%, with H^2 = 71.59% and GAM = 77.91%.

Marketable tuber (MT) number in 2020 recorded GCV of 32.43% and PCV of 33.59%, with high heritability (H^2 = 93.25%) and GAM = 64.61%, while in 2021, environmental influence increased, with H^2 = 18.97%. For non-marketable tuber number (NMT), the GCV and PCV remained high across both years (68.43% and 70.18% in 2020, 61.37% and 64.73% in 2021), reflecting strong genetic control (H^2 = 95.09% in 2020, 89.87% in 2021). The specific

Table 3. Summary of yield and quality traits in potato accessions during 2020 and 2021.

Traits	Mean		Min		Max		Std. Error		Std. Deviation		CV	
	2020	2021	2020	2021	2020	2021	2020	2021	2020	2021	2020	2021
PM	NA	72.24	NA	63.12	NA	89.38	NA	0.21	NA	4.09	NA	2.5
MT	151.44	174.33	11.59	0	282.84	424.47	2.81	3.11	51.21	59.58	8.72	28.44
NMT	93.07	162.9	0	23.75	609.87	781.25	3.61	5.41	65.72	103.5	15.31	20.96
MTY	11.5	9.52	0.32	1.3	22.94	22	0.26	0.2	4.73	3.91	6.4	12.11
NMTY	1.37	1.8	0	0.21	4.56	6.71	0.05	0.05	0.84	0.96	28.14	28.37
TTY	12.87	11.33	1.73	2.08	24.23	23.98	0.26	0.2	4.71	3.89	6.24	9.74
SG	1.06	1.06	1.04	1.04	1.09	1.1	0.00049	0.00046	0.01	0.01	0.79	0.43

NA - This parameter was not measured in the year 2020

Table 4. Genetic variability, heritability, and genetic advance estimates for studied traits in potato accessions during 2020 and 2021.

Trait	GCV		PCV		ECV		H^2 (BS)		GA		GAM	
	2020	2021	2020	2021	2020	2021	2020	2021	2020	2021	2020	2021
PM		5.15		5.73		2.5		80.88		6.91		9.56
MT	32.43	13.97	33.59	32.07	8.73	28.87	93.25	18.97	97.84	21.88	64.61	12.55
NMT	68.43	61.37	70.18	64.73	15.56	20.6	95.09	89.87	128.12	195.51	137.66	120.02
MTY	40.6	38.92	41.12	40.86	6.52	12.43	97.49	90.75	9.51	7.28	82.69	76.49
NMTY	50.72	44.63	58.3	52.75	28.76	28.12	75.67	71.59	1.25	1.41	91.01	77.91
TTY	36.03	33	36.58	34.46	6.36	9.94	96.98	91.69	9.42	7.38	73.19	65.19
SG	NA	0.69	0.78	0.82	0.79	0.43	NA	72.23	NA	0.01	NA	1.22

GCV: Genotypic Coefficient of Variation, PCV: Phenotypic Coefficient of Variation, ECV: Environmental Coefficient of Variation, H^2 (BS): Broad Sense Heritability, GA: Genetic Advance, GAM: Genetic Advance as Percentage of Mean.

gravity (SG) remained stable with minimal variability, showing $H^2 = 72.23\%$ in 2021.

The year-wise adjusted mean values for the studied traits across the years 2020 and 2021 are presented in Fig. 1 The median values for marketable tuber (MT) number were slightly higher in 2021 compared to 2020, which indicate a higher marketable tuber number in the second year. For non-marketable tuber number (NMT), the median is higher in 2021, and the variability was higher in comparison to 2020. There were several outliers in both years for MT (except 2020) and NMT, which indicate few accessions had very high or few tuber numbers (Fig. 1a). The adjusted mean boxplot for plant maturity (PM) indicated that the majority of the observations fall within a narrow range, with a median value of approximately 73 (Fig. 1b). The median value for specific gravity (SG) during both

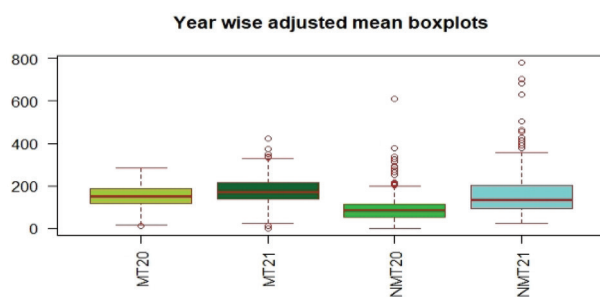


Fig. 1a. Year-wise adjusted mean boxplots for marketable tuber count (MT) and non-marketable tuber count (NMT) in 2020 (MT20, NMT20) and 2021 (MT21, NMT21). The Y-axis represents the number of tubers per plot.

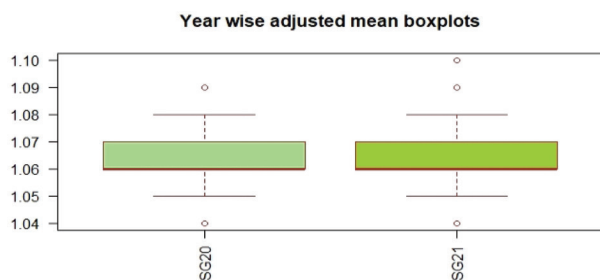


Fig. 1c. Year-wise adjusted mean boxplots for specific gravity (SG) in 2020 (SG20) and 2021 (SG21). The Y-axis represents specific gravity.

the years was nearly identical, around 1.07, which indicate consistency in specific gravity values across the years (Fig. 1c). The overall variation for SG was low, with values mostly clustering around 1.06 to 1.07. The year-wise adjusted mean boxplots for tuber yield indicated that marketable tuber yield (MTY) and total tuber yield (TTY) were slightly higher in 2020 compared to 2021, with more variability observed in 2020. Non-marketable tuber yield (NMTY) was consistently low across the years, with minimal impact on overall yield (Fig. 1d).

The bar chart (Fig. 2) illustrates the contribution of different variables to the first two principal components (Dim-1-2). It shows how much each variable contributes to the variation captured by these principal components (PCs), with MTY and TTY contributed the most (each 20-22%), followed by NMT, NMTY and MT. The other traits

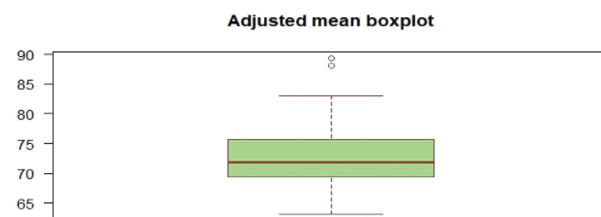


Fig. 1b. Boxplot showing the adjusted mean values for plant maturity (PM) across evaluated potato accessions. The Y-axis represents the number of days to maturity.

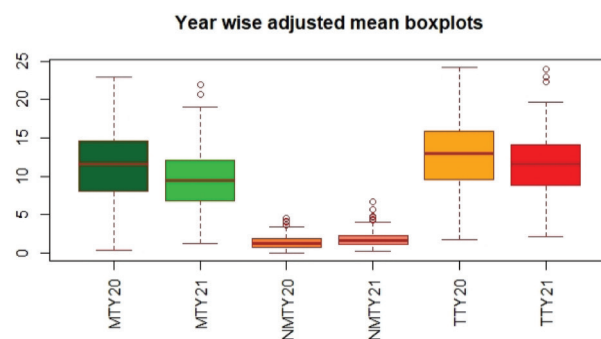


Fig. 1d. Year-wise adjusted mean boxplots for marketable tuber yield (MTY), non-marketable tuber yield (NMTY), and total tuber yield (TTY) in 2020 (suffix '20') and 2021 (suffix '21'). The Y-axis represents yield in tons per hectare (t/ha).

viz., SG and PM showed relatively lower contributions, that suggest these variables have less influence in the first two dimensions of the PCA. Complementing this, the second plot (Fig. 3) presents a correlation matrix between the PCs (Dim 1 to Dim 5) and the original variables (PM, MT, MTY, NMT, NMTY, SG, TTY). The size and colour intensity of the circles indicates the strength and direction of the correlation between each variable and principal component. Dim 1 appears to be heavily influenced by yield-related variables like MTY and TTY. Dimension 4 (Dim 4) showed strong associations with variables like SG and PM, that indicate these components capture variance related to these specific traits.

The correlation matrices for the year 2020 and 2021 in the figure 4 and figure 5 revealed the relationships between yield-related traits in potato. In both the years, a strong positive correlation was observed between MT, MTY and TTY, that showed increase in the number of marketable tubers is strongly associated with higher yields (both MTY and TTY), with MTY playing a crucial role in contributing to the total tuber yield. Similarly, NMT and NMTY were strongly positively correlated with each other in both years, but they exhibited weak negative correlations with MT, MTY, and TTY. In 2021, PM showed a strong negative correlation with MTY and TTY, indicating that

increased maturity may lead to reduced yields, a relationship not prominently seen in 2020.

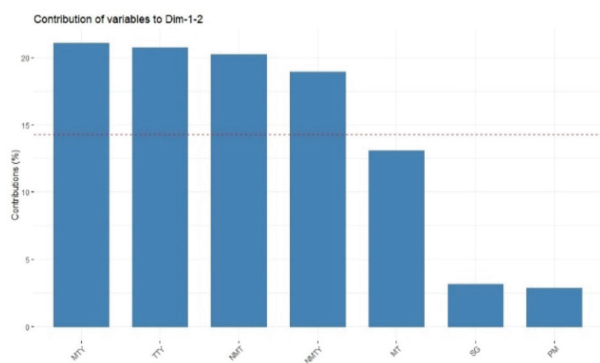


Fig. 2. Contribution of yield and its contributing traits to principal components 1 and 2

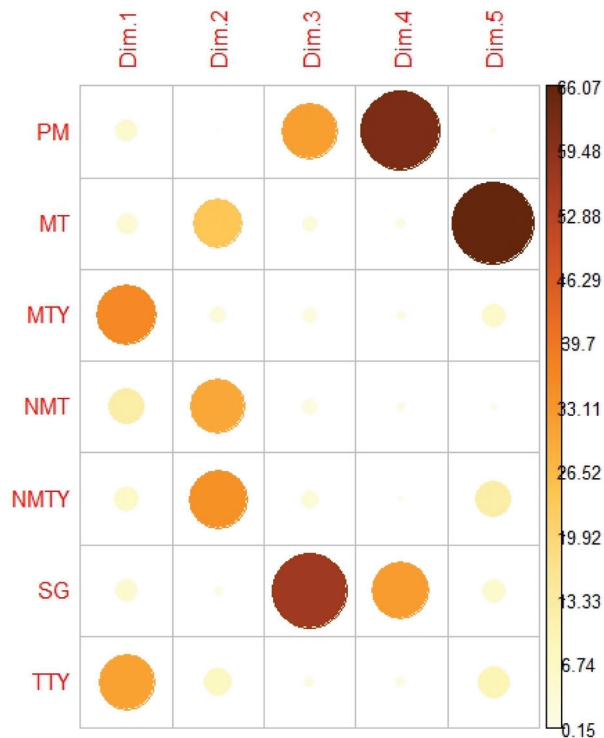


Fig. 3. Correlation between principal components and yield related traits

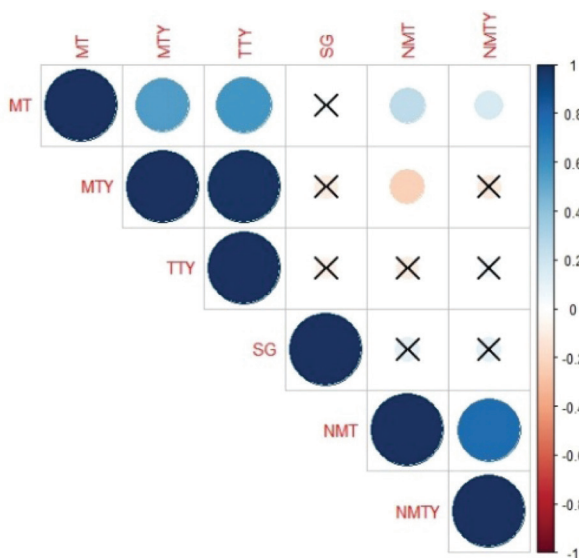


Fig. 4. Pairwise correlation matrix among studied traits (Year 2020)

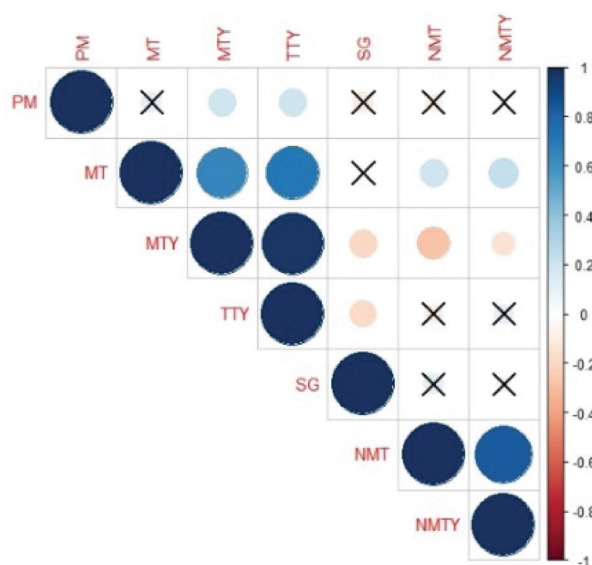


Fig. 5. Correlation matrix of yield and quality traits in potato accessions evaluated in 2021. Traits include plant maturity (PM), marketable tuber number (MT), marketable tuber yield (MTY), total tuber yield (TTY), specific gravity (SG), non-marketable tuber number (NMT), and non-marketable tuber yield (NMTY).

The first five PCs accounted for almost the entire variability in the data, with the first component explained approximately 34.94% of the total variance, the second component 28.32%, and the third 13.49%. Collectively, the first three PCs captured around 76.75% of the total variance, and by the fifth component, 98.12% of the variance was explained, which showed these components provide a comprehensive representation of the data (Table 5). First PC was heavily influenced by MTY and TTY, contributing 35.98% and 31.31% respectively. These variables were the primary drivers of the variance in this dimension, or in the overall yield performance. Second PC was significantly influenced by NMTY and NMT, with contributions of 34.58% and 29.90%, respectively. This indicates that Dim-2 captured the variability related to non-marketable tuber yield. Dim-3 was dominated by specific gravity, which contributes 56.36%, suggesting that this component was primarily

Table 5. Correlation coefficients of variables across principal components (Dim 1 to Dim 5).

Variable	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
PM	0.35	0.05	-0.55	0.76	0.05
MT	0.32	0.69	0.15	0.11	-0.62
MTY	0.94	0.23	0.15	-0.10	0.18
NMT	-0.55	0.77	-0.15	-0.10	0.06
NMTY	-0.39	0.83	-0.18	-0.07	0.27
SG	-0.35	0.15	0.73	0.55	0.17
TTY	0.88	0.39	0.12	-0.11	0.23
% of variance	34.94	28.32	13.49	13.1	8.27
Cumulative % of variance	34.94	63.26	76.75	89.85	98.12

related to the quality aspects of the tubers. The similar things were also represented by PCA biplot in our study (Fig. 6).

Hierarchical clustering of genetic accessions in a circular dendrogram:

The circular dendrogram represented the hierarchical clustering of different accessions based on their genetic or phenotypic similarity, derived from pooled data of two year (Fig. 7). The color-coded clusters in the diagram indicated the distinct groups of similar samples, with red (131 accessions), blue (96), green (44) and black (48) labels which represent the main clusters. The clustering analysis in the current investigation exhibited distinct genetic groupings of the potato accessions, with each cluster provides valuable insights for breeding and genetic improvement programs. The red cluster, comprised of 113 accessions of Tuberosum lines, 15 Andigena, and 3 Indian varieties (Kufri Megha, Kufri Lauvkar, and Kufri Bahar), which exhibited superior performance in TTY and MTY, which provides yield related traits to potato breeding programs. The blue cluster, includes 73 Tuberosum, 18 Indian varieties, and 5 Andigena, showed moderate and stable performance among studied traits, indicating suitability of these

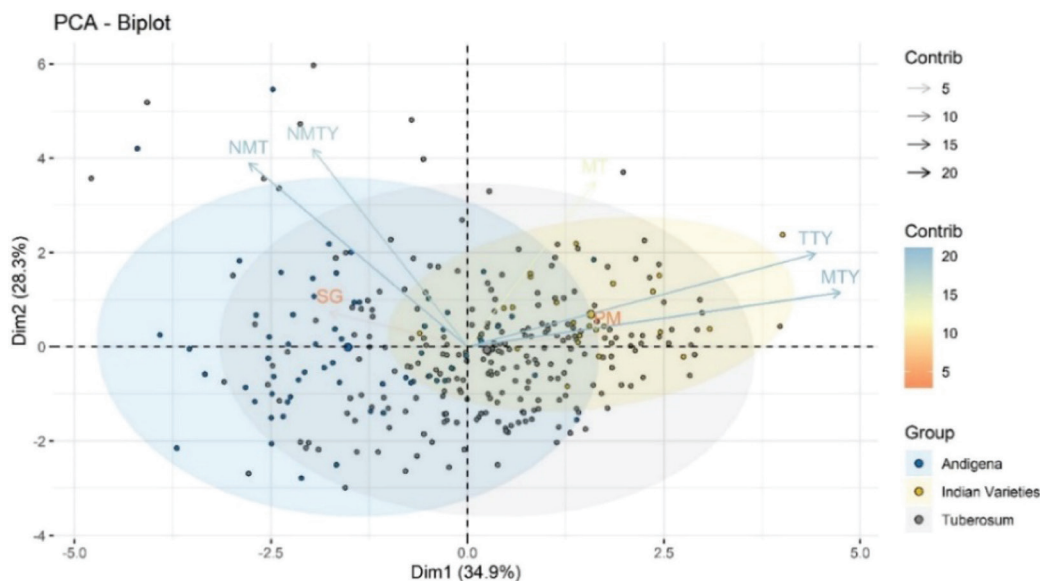


Fig. 6. PCA biplot illustrating the contribution of studied traits across potato accessions

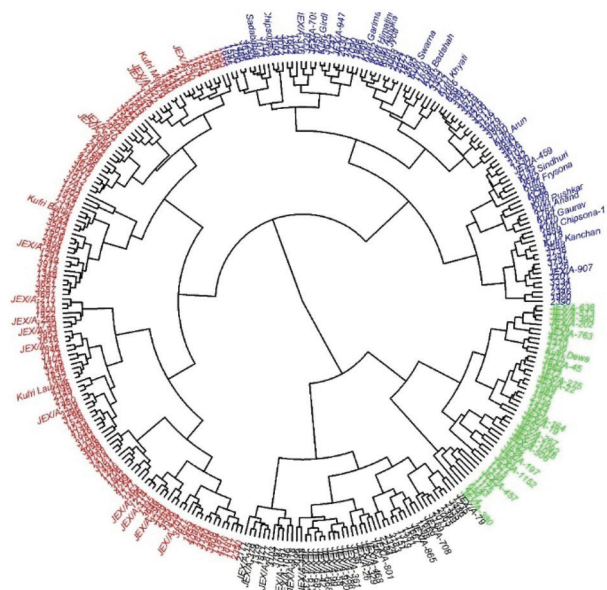


Fig. 7. Phylogenetic clustering of potato accessions based on genetic similarity

accessions for improving adaptability in potato cultivars. The green cluster, comprises of 25 Tuberosum, 18 Andigena, and one Indian variety, demonstrated significant diversity, particularly in traits like specific gravity and plant maturity. These traits are valuable for breeding programs which targets

tuber quality and maturity traits. The fourth cluster (black), comprised of 24 Andigena and 24 Tuberosum lines, demonstrated higher variability in non-marketable tuber number and yield. This clustering analysis provides an understanding the genetic or phenotypic diversity within the dataset and identifying closely related groups of samples.

DISCUSSION

The augmented design allowed for the evaluation of a large breeding material through repeated checks among blocks. This is beneficial for genetic diversity studies where a large number of unreplicated test entries need to be assessed under similar field conditions. The findings from this study provide important insights into the genetic diversity, trait variability, and stability of key potato yield and quality traits across different accessions evaluated over two years (2020 and 2021). The integration of PCA and hierarchical clustering enabled a comprehensive understanding of the underlying patterns and relationships among the studied traits.

Year-wise stability and variability in yield traits

The year-wise comparison of yield traits revealed slight variations between the year 2020 and 2021, particularly in yield related traits (MTY and TTY), with higher variability and slightly better performance observed in the year 2020. This variation could be attributed to environmental factors that may have influenced the growth and yield of the potato plants. The stability observed in specific gravity across both the years suggests that this trait is less sensitive to environmental variations and strongly influenced by genetic factors. The consistently low values of NMTY across both years indicated that most of the accessions evaluated had low non-marketable tuber yield, which is a desirable trait for potato improvement. From the current investigation it is clear that MTY and TTY are the key traits with high genetic control (H^2) and genetic advance (GAM), which makes them superior traits for selection in potato breeding programs. Superior accessions such as JEX/A-459, CP 2142, and Kufri Pushkar demonstrated excellent performance, which underscores their potential as parental lines for yield improvement.

Trait variability and principal component analysis

The PCA revealed that majority of the variability in the dataset is captured by the first two PCs, which together accounted for 63.2% of the total variance. The high contribution of yield related traits (MTY and TTY) to Dim1 (PC1) suggested that these traits are the primary contributors of yield performance across the evaluated accessions. These findings align with the previous studies by other researchers which highlighted the importance of marketable yield in their breeding programs. Seid *et al.* (2021) reported that the first six PCs accounted for 88.20% of

the total variation among 24 potato accessions, with the first PC alone contributed 34.30%, which underscores key traits for processing quality in potatoes. Tessema *et al.* (2022) used PCA to identify key traits contributing to yield variability among Ethiopian potato varieties, with the first few PCs captured the majority of phenotypic variation. Khan *et al.* (2020) in their study demonstrated that the first two PCs captured around 48.29% of the total variability, which showed complexity and breadth of phenotypic trait variation. Another PCA study showed that the first PC accounted for 33.72% of the variation, followed by the second and third which contributes 16.16% and 10.70%, respectively (Verma and Singh 2016). Ahmadizadeh and Felenji (2011) reported that the first PC accounted for 38.3% of the total variation among potato cultivars, governed by key yield-related traits.

The separation of NMTY and NMT along Dim2 indicates that these traits capture those aspect of variability which were related to factors that reduces marketable tuber yield. The minimal contribution of specific gravity and plant maturity to the first two dimensions suggested that these traits, may have less influence on the overall variability. The correlation matrix allows for an understanding of how each variable correlates with the different principal components. The correlation strength in our study varied across dimensions, which reflects multidimensional nature of the data and the distinct role of each variable in contributing to different aspects of the variance.

The PCA biplot effectively demonstrates how different potato genotypes (Andigena, Indian varieties, Tuberosum) relate to key agricultural traits (MT, MTY, TTY, NMT, NMTY, SG, PM) (Fig. 6). The Indian varieties and the tuberosum genotypes were found to be good for TTY, MTY and MT, while most of the Andigena genotypes had more of NMT and NMTY. This trait distinction clearly revealed

that the tuberosum types had good tuber size and yield and must be utilized in breeding programmes.

Genotypic diversity and clustering analysis

The hierarchical clustering analysis revealed valuable insights into the genetic and phenotypic diversity among the evaluated accessions. The formation of four distinct clusters (red, blue, green, and black) in the present investigation highlights the presence of subgroups with shared traits. The distribution of potato accessions across the four identified clusters underscores the genetic diversity within the population, which is crucial for the success of breeding programs for improve specific traits. Abebe *et al.* (2013) revealed three distinct groups among Ethiopian potato varieties in their study, which underscores significant phenotypic diversity that can be leveraged for targeted breeding programs. Through SSR markers and phenotypic traits, Dalamu *et al.* (2024) identified four major genetic clusters in native potato accessions, which highlights genetic diversity valuable for breeding programs to improve resilience and adaptability. Similarly, Datta *et al.* (2015) observed substantial genetic diversity across yield and quality traits, with distinct clusters formed based on tuber weight, tuber count, and dry matter content. Esnault *et al.* (2014) identified distinct genetic clusters within a diverse potato collection, including Chilosé Island landraces and global cultivars, which demonstrated substantial genetic diversity. Ahmadizadeh and Felenji (2011), Khan *et al.* (2013), Iqbal *et al.* (2018), Panigrahsi *et al.* (2014) categorized potato cultivars into two, five, two and seven distinct clusters respectively, based on morphological traits, which showed the effectiveness of phenotypic diversity in targeted breeding. Ghebresslassie *et al.* (2015) used single linkage clustering to classify potato accessions into three main

clusters: Group I (average yield traits), Group II (high yield potential), and Group III (lower yield but distinct morphology). This finding highlights the substantial phenotypic diversity within cultivated potatoes. Samiha *et al.* (2024) reported that hierarchical clustering of 62 accessions grouped high-yielding and robust accessions into distinct clusters, which revealed differences in tuber weight, plant height, and dry matter content across clusters. Cluster analysis by Seid *et al.* (2021) categorized 24 potato accessions into six clusters, with accessions in cluster II and VI showed superior traits, such as higher tuber yield, specific gravity, and dry matter content, which were promising candidates for processing breeding programs.

The clustering of accessions with similar yield performance or quality traits suggested that these subgroups may be valuable for targeted breeding efforts. For example, accessions in the red cluster in our study, demonstrated high marketable yield, which could be prioritized for breeding programs on this trait. Conversely, accessions in the green cluster, which showed distinct characteristics from red cluster, might be explored for their unique attributes that could contribute to broader genetic diversity in breeding populations. The dominance of Tuberosum accessions across clusters, particularly in the red and blue clusters, underscores their significance as the primary cultivated group, valued for their high yield potential, tuber uniformity, and other quality. These accessions serve as the backbone for commercial potato production and provide stable performance in diverse agroecological conditions. In contrast, the presence of Andigena accessions in all clusters, particularly in the green and black clusters, emphasizes their importance as a reservoir of genetic diversity. As a landrace species from South America, Andigena contributes traits such as stress tolerance, disease resistance, and adaptability to marginal environments.

These accessions are vital for the introduction of novel alleles into breeding populations and broadening the genetic base of potato cultivars. The inclusion of Indian varieties is significant due to their regional adaptability, market acceptability, and resistance to biotic and abiotic stresses. These varieties demonstrated the importance of integrating locally adapted germplasm into breeding programs to address the specific challenges of Indian agroecological zones. The clustering patterns revealed the potential of combining the productivity traits of *Tuberosum*, the genetic diversity of *Andigena*, and the regional adaptability of Indian varieties to develop superior potato cultivars with enhanced yield, quality and resilience. The grouping of most Indian varieties into a single cluster highlights their shared genetic background and adaptation to similar agroecological conditions, which showed their development for consistent performance in specific environments and suitability for targeted breeding programs. Certain *Andigena* accessions, like JEX/A-707, JEX/A-298, JEX/A-317, JEX/A-539 and JEX/A-32, may have been improved through geneflow or adapted to similar environmental and cultivation conditions as *Tuberosum*, which also results into comparable agronomic performance and their categorization in the red cluster.

CONCLUSION

The current investigation provides a detailed analysis of potato accessions using PCA and hierarchical clustering. The findings of this study emphasize the importance of marketable yield as an important trait of variability, the stability of specific gravity, and the potential for focused breeding efforts based on identified genotype clusters. These findings are valuable for future potato breeding programs focused on improving yield and quality traits. This study also demonstrated the utility of the augmented

design for evaluation of genetic diversity and yield performance traits in a large panel of potato accessions. The identification of traits which govern yield variability, combined with the clustering of accessions based on similar traits, provides valuable information for selection of parental lines in the breeding programs. Future studies could explore the environmental factors that contributed to the year-to-year variability observed in this study. Additionally, further genetic analysis of the identified clusters could provide insights into the underlying genetic mechanisms controlling the key traits identified in this study.

ACKNOWLEDGEMENT

We thank ICAR for financial and infrastructural support.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest

ETHICAL STATEMENT

This article does not contain any studies with human participants or animals performed by any of the authors

LITERATURE CITED

- Abebe T, Wongchaochant S, Taychasinpitak T (2013) Analysis of the phenotypic diversity within cultivated potato varieties in Ethiopia at three locations. *Kasetsart J - Nat Sci* 47:803–817
- Ahmadzadeh M, Felenji H (2011) Evaluating Diversity among Potato Cultivars Using Agro-Morphological and Yield Components in Fall Cultivation of Evaluating Diversity among Potato Cultivars Using Agro-Morphological and Yield Components in Fall Cultivation of Jiroft Area. *Am J Agric Environ Sci* 11:655–662
- Bhardwaj V, Kumar A, Sharma S, Singh, B., Poonam, Sood, S, Dipta, B, Singh, R, Siddappa, S, Thakur, AK and Dalamu, D (2023) Analysis of Genetic Diversity, Population Structure and Association Mapping for Late Blight Resistance in Potato (*Solanum tuberosum* L.) Accessions Using SSR Markers. *Agronomy* 13:294.

Vikas Mangal, Salej Sood, Vinay Bhardwaj, Bhawna Dipta, Mehi Lal, Kritika Singh, R K Singh, Vinod Kumar and Brajesh Singh

- Dalamu D, Tiwari JK, Zinta R, Bhatia, N., Kaushal N and Kumar, V (2024) Genetic Diversity and Population Structure Analyses Using Simple Sequence Repeat Markers and Phenotypic Traits in Native Potato Collection in India. *Potato Res* 67:357–381.
- Datta S, Das R, Singh D (2015) Evaluation of genetic diversity for yield and quality parameters of different potato (*Solanum tuberosum* L.) germplasm. *J Appl Nat Sci* 7:235–241.
- Deperi SL, Tagliotti ME, Bedogni MC, Manrique-Carpintero, NC, Coombs J, Zhang R, Douches D, and Huarte MA (2018) Discriminant analysis of principal components and pedigree assessment of genetic diversity and population structure in a tetraploid potato panel using SNPs. *PLoS One* 13:e0194398.
- Esnault F, Solano J, Perretant MR, Herve M, Label A, Pellé R, Dantec JP, Boutet G, Brabant P and Chauvin JE (2014) Genetic diversity analysis of a potato (*Solanum tuberosum* L.) collection including Chiloe Island landraces and a large panel of worldwide cultivars. *Plant Genet Resour Characterisation Util* 12:74–82.
- Khan MF, Tabassum N, Latif A, Khaliq A and Malik M (2013) Morphological characterization of potato (*Solanum tuberosum* L.) germplasm under rainfed environment. *African J Biotechnol* 12:3214–3223.
- Ghebresslassie BM, Githiri SM, Mehari T, Kasili RW (2015) Analysis of Diversity among Potato Accessions Grown in Eritrea Using Single Linkage Clustering. *Am J Plant Sci* 06:2122–2127.
- Husson, F., Josse, J., Le, S., Mazet, J., & Husson, M. F. (2016). Package ‘factominer’. An R package, 96(96), 698.
- Iqbal N, Erum S, Azim R, Shakil M and Khan MF (2018) Estimation of phenotypic diversity among locally available potato germplasm. *Int J Biosci* 13:10–17.
- Kassambara, A., & Mundt, F. (2017). Package ‘factoextra’. Extract and visualize the results of multivariate data analyses, 76(2).
- Khan A, Erum S, Riaz N, Shinwari MI and Ibrar Shinwari M (2020) Multivariate Analysis of Potato Genotypes for Genetic Diversity. *bioRxiv* 2020.04.14.041236.
- Mangal V, Sood S, Dalamu, Kumar V and Bhardwaj V (2022) Role of genetic resources in management of potato pests and diseases. In *Sustainable Management of Potato Pests and Diseases* (pp. 185–211). Singapore: Springer Singapore.
- Mangal V, Luthra SK, Sood S, Gupta VK, Bhardwaj V, Kumar V, Thakur AK, Dalamu, Singh R and Bradshaw JE (2024) Breeding New Cultivars for Biotic, Abiotic and Quality Traits. In *Approaches for Potato Crop Improvement and Stress Management* (pp. 119–157). Singapore: Springer Nature Singapore.
- Panigrahi KK, Sarkar KK, Baisakh B, Mohanty A (2014) Assessment of Genetic Divergence in Potato (*Solanum Tuberosum* L.) Genotypes for Yield and Yield Attributing Traits. *Int J Agric Environ Biotechnol* 7:247.
- Samaha TK, Rahman MA, Islam S, Jalal N, Islam A and Nahiyani ASM (2024) Exploring potato diversity: A comprehensive genetic and phenotypic analysis of quantitative and qualitative traits. *Czech J Genet Plant Breed* 60:126–135.
- Seid E, Mohammed W, Abebe T (2021) Genetic Diversity Assessment through Cluster and Principal Component Analysis in Potato (*Solanum tuberosum* L.) Genotypes for Processing Traits. *Int J Food Sci Agric* 5:440–447.
- Sood S, Bhardwaj V, Mangal V, Kumar A, Singh B, Dipta B, Kaundal B, Kumar V and Singh B (2024) Genome-wide association mapping to identify genetic loci governing agronomic traits and genomic prediction prospects in tetraploid potatoes. *Sci Hortic* 328:112900.
- Tessema GL, Mohammed AW, Abebe DT (2022) Genetic variability studies for tuber yield and yield attributes in Ethiopian released potato (*Solanum tuberosum* L.) varieties. *PeerJ* 10:e12860.
- Verma A, Singh D (2016) Assessment of genetic diversity and association among agro-morphological characters in potato (*Solanum tuberosum* L.). *Indian J Agric Res* 50:345–349.
- Wei T, Simko V, Levy M, Xie Y, Jin Y and Zemla J (2013) corrplot: Visualization of a correlation matrix. R package version 0.73, 230(11), 1–26.
- Wickham H and Wickham H (2016) Getting Started with ggplot2. *ggplot2: Elegant graphics for data analysis*, 11–31.
- Wickham H and Wickham MH (2020) Package ‘plyr’. Obtenido <https://cran.r-project.org/web/packages/dplyr/dplyr.Pdf>.
- Wickham H, Averick M, Bryan J, Chang W, McGowan LDA François R, ... and Yutani H (2019) Welcome to the Tidyverse. *Journal of open source software*, 4(43), 1686.

(MS Received : January 22, 2025; Accepted : July 7, 2025)