Selecting the ideotype of improved rice cultivars using multiple regression and multivariate models

Hassan Haghshenas, Afshin Soltani, Abbas Ghanbari Malidarreh, Hossein Ajam Norouzi and Salman Dastan[d]

ELECTRONIC SUPPLEMENTAL MATERIAL

Materials and methods

Statistical analysis

In order to determine the most important traits and demonstrate the contribution share of different traits for the formation and the determination of PY, the method of variable selection and multiple regressions was used. The stepwise regression analysis with forward selection method was performed. Through this method, the relationship between the PY and all the traits were quantitatively determined; from the 38 investigated traits. After that, the best models of one to ten variables were specified. Moreover, the measures of minimum, maximum, average, and the level of significance for the six traits were selected and the effect on the PY was presented. All the traits influencing the PY were considered as independent variables, and the PY as a dependent variable and, eventually, the best regression models or production models, including one-to-*p* variables (six variables in here), were specified by forward selection (abbreviated as maxr). The best regression model was selected based on the highest R^2. Furthermore, for multiple alignments, the models were controlled by investigating the variance inflation factor (Soltani et al. 2000). Then, the best regression models of the six variables were selected in the step seven. The reason for selection this step was that by increasing the number of variables from one to six, the changes in R^2 remained significantly constant ($R^2 = 0.64^{**}$). Later, the mentioned equation was investigated and analyzed, and, by deriving the component correlation between the equation components, the positive and the negative relationship, and correlation of the components with each other were evaluated. Finally, the traits' specifications, in the form of average and best models, which can be placed in the PY regression model, entered the production model of six variables. In order to determine the PY model (production model), the relationships between all the variables were measured and the PY was evaluated using the regression method (Soltani et al. 2016). The final model was obtained through the controlled trial-and-error method, which can quantify the effect of the PY limitations. The average PY was calculated by the model by putting the observed average variables (Xs) in the fields under study in the PY model. By placing the best observed value of the variables in the PY model, the maximum obtainable PY was calculated. The difference between these two variables has been considered PY changes. The difference between the multiplication of the average observed value for each variable by its coefficient and the multiplication of the best observed value for the same variable by the coefficient of the same variable presents the value of the PY variation for that variable. The ratio of PY variation for each variable to the total PY variation show its share in creating PY changes and is presented in percentage. Different procedures of the software SAS (version 9.1) were used for analysis.

Analysis of canonical correlation

The relationship between canonical variables and main variables is evaluated with correlation coefficients between them, which is generally called structural factors (Khattree and Naik 2000; Johnson and Wichern 2002). To investigate this correlation, the correlation between VAR/WITH variables (investigated traits) and its main canonical variables was evaluated for five pairs of primary variables. These pairs of variables provided the greatest correlation with VAR/WITH variables (investigated traits).

Principle component analysis (PCA) and biplot analysis

The principal component analysis method explained by Harman (1976) was followed in the extraction of the components. The percentage variances explained by each component were determined (Tadesse and Bekele 2001). Principal component analysis were performed using the software SPSS (version 16) for all the traits of rice cultivars. Biplots were generated using the software Minitab (version 16) that runs in a windows environment, an earlier version of which was described in Yan (2001). Up-to-date information on GGE biplot is available at http://www.ggebiplot.com.

Table S1. Description of improved rice cultivars characteristics using in the experiment

Cultivar	Growth condition	Maturity condition	Paddy yield (kg ha^{-1})	Quality condition	Tolerance to stress	Origin
Dasht	Semi-dwarf	Late maturity	7500	Normal quality	Tolerance	Iran
Amol 3	Semi-dwarf	Late maturity	8000	Normal quality	Tolerance	Iran
Ghaem	Semi-dwarf	Late maturity	6000	Normal quality	Tolerance	Iran
Pardis	Semi-dwarf	Late maturity	6000	Normal quality	Tolerance	Iran
Pazhouhesh	Semi-dwarf	Late maturity	6650	Normal quality	Tolerance	Iran
Keshvari	Semi-dwarf	Late maturity	7500	Normal quality	Tolerance	Iran
Kados	Semi-dwarf	Late maturity	6500	Normal quality	A little sensitive	Iran
Nemat	Dwarf	Late maturity	7250	Normal quality	Tolerance	Iran
Fajr	Dwarf	Late maturity	6250	Normal quality	Tolerance	Iran
Khazar	Dwarf	Late maturity	6300	Normal quality	Tolerance	Iran
Sahel	Dwarf	Late maturity	6750	Normal quality	Tolerance	Iran
Shiroodi	Dwarf	Late maturity	7750	Normal quality	Tolerance	Iran
Neda	Dwarf	Late maturity	7750	Normal quality	Tolerance	Iran

Table S2. List of investigated traits affecting the yield in regression models

Traits	Unit	Min.	Mean	Max.	pr>F
Days to germination (DG)	day	2.5	4.1	6	**
Days to transplanting (DTP)	day	25	29.2	34	**
Days to tillering (DT)	days since seeding in the nursery	35	39.7	45	**
Days to stem elongation (DS)	days since seeding in the nursery	51	63.8	76	**
Days to heading (DH)	days since seeding in the nursery	84	93.8	108	**
Days to pollination (DP)	days since seeding in the nursery	85	95	108	**
Days to physiological maturity (DPM)	days since seeding in the nursery	99	112	122	**
Days to harvesting date (DHD)	days since seeding in the nursery	112	120.8	131	**
Leaf area per plant (LA)	m^2	0.11	0.2	0.4	**
Leaf area index (LAI)	-	2.95	6	9.3	**
Leaf number (LN)	no.	44	79.8	130	**
Flag leaf length (FLL)	cm	20	30.2	42	**
Panicle length (PL)	cm	22.6	27.1	32.2	**
Stem length (SL)	cm	54	76.5	130	**
Plant height (PH)	cm	8.25	103.6	157.5	**
Number of tiller per hill (TTH)	no.	7.4	13.9	21.4	**
Number of fertile tiller per hill (FTH)	no.	7	13.3	19.6	**
Fertile tiller percentage per hill (FTP)	%	81.53	96.1	100	**
Number of infertile tiller per hill (ITH)	no.	0	0.6	2.4	**
Tillering coefficients (TC)	%	3.7	6.9	10.7	**
Panicle number per plant (PNP)	no.	7	13.4	19.6	**
Panicle per square meter (PM)	no.	175	333.5	490	**
Total spikelet per panicle (TS)	no.	65.8	126	207.2	**
Number of filled spikelet per panicle (FS)	no.	61.1	97.9	165	**
Filled spikelet percentage (FSP)	%	56.45	79.9	94.9	**
Blank spikelet per panicle (BS)	no.	3.8	28.1	90.1	**
Thousand grain weight (TGW)	gr.	22	25.3	30	**
Leaf and shoot dry weight in pollination stage (LSWP)	gr.	6	42.4	83.5	**
Panicle dry weigh in pollination stage (PWP)	gr.	2	9	20.5	**
Total plant weight in pollination stage (TW)	gr.	0	50.6	98.5	**
Leaf and stem weight in physiological maturity stage (LSWM)	gr.	18.5	42.6	70	**
Panicle dry weight in physiological maturity stage (PWM)	gr.	17.5	33.6	56	**
Allocation efficiency of dry matter to panicle (ALP)	gr.	0.1	0.2	0.3	**
Maximum dry matter accumulation (DMA)	gr.	45	76.3	123	**
Paddy yield (PY)	kg ha^{-1}	5600	6624.1	8170	**
Straw yield (SY)	kg ha^{-1}	4750	7296.79	9800	*
Biological yield (BY)	kg ha^{-1}	10500	13921.15	17160	**
Harvest index (HI)	%	39.8	47.77	56.10	**

The last column is the probability of significant F-test by ANOVA to compare cultivars.

ns, * and **: non-significant and significant in 5% and 1% probability level, respectively.

Table S3. Mean comparison of all investigated traits of 13 improved rice cultivars.

Traits	Unit	Dasht	Amol 3	Ghaem	Pardis	Pazhouhesh	Keshvari	Kados	Nemat	Fajr	Sahel	Shiroodi	Neda	Khazar	LSD 0.05
DG	day	5.0a	5.0a	4.0c	4.0c	4.0c	4.5b	3.8d	4.0c	4.0c	4.0c	3.8d	3.8d	4.0c	0.00
DTP	day	26.5g	27.0f	29.0d	29.5c	28.0e	29.0d	26.0h	32.0a	30.0b	30.0b	32.0a	32.0a	29.0d	0.00
DT	day	41.0a	40.8ab	39.3bcd	41.0a	38.8cde	37.5e	37.8de	40.2abc	39.7abc	40.5ab	41.2a	39.8abc	37.8de	1.57
DS	day	59.8fe	60.5e	64.0cd	68.3a	64.5c	64.3c	59.0g	68.3a	64.5c	63.5d	66.8b	66.3b	59.7fg	0.72
DH	day	97.0cd	98.2b	87.8i	90.2g	87.2i	92.7f	90.8g	97.2cd	99.2a	88.7h	97.5bc	96.7de	96.2e	0.68
DP	day	99.7b	101.0a	88.7j	90.3h	87.7k	93.8f	91.3g	97.5cd	99.8b	89.3i	98.0c	97.0de	96.8e	0.50
DPM	day	113.5f	121.2a	105.5j	105.5j	101.7k	116.0d	111.8g	117.2c	108.8h	106.8i	117.0c	117.7b	115.5e	0.43
DHD	day	125.0b	122.5c	117.8e	117.7e	114.5f	118.3e	120.0d	125.3b	121.7c	114.0f	125.2b	127.0a	121.8c	1.11
LA	m^2	0.24cd	0.28b	0.24c	0.21e	0.24cd	0.22de	0.20ef	0.26bc	0.34a	0.24cd	0.24c	0.20ef	0.19f	0.02
LAI	-	5.9cd	6.9b	6.0c	5.3e	5.9cd	5.5de	5.0ef	6.5bc	8.6a	6.0cd	6.0c	5.1ef	4.7f	0.54
LN	no.	53.3f	70.2e	83.2cd	71.3e	85.7c	74.3e	69.0e	76.7de	99.8ab	95.2b	94.5b	107.5a	56.7f	8.24
FLL	cm	35.2a	32.2bc	27.8efg	26.7fg	25.8g	32.5bc	32.8ab	30.3cd	25.8g	32.5bc	29.5de	28.7efg	33.33ab	2.39
PL	cm	26.5e	29.1b	24.3f	27.4de	24.1f	27.9cd	29.2b	26.8de	26.4e	26.5e	28.8bc	23.6f	31.2a	1.14
SL	cm	84.6b	75.8cde	79.9bc	73.2ef	67.0gh	78.7cd	99.0a	66.7gh	73.7def	68.9fg	67.2g	62.2h	98.2a	5.01
PH	cm	111.0b	104.8cd	104.2cd	100.6de	91.1g	106.6bc	128.3a	93.4g	100.2efg	95.4fg	96.0efg	85.8h	129.4a	5.06
TTH	no.	9.1f	12.5de	14.2bc	13.4cd	16.9a	11.6e	14.5bc	15.2b	14.4bc	18.1a	14.9b	17.3a	8.6f	1.25
FTP	%	95.3b-e	94.6cde	98.1ab	99.3a	95.6bcd	97.7ab	99.3a	95.7bcd	92.3e	97.3abc	96.4abc	93.0de	94.5cde	2.98
TC	%	4.6f	6.3de	7.1bc	6.7cd	8.4a	5.8e	7.2bc	7.6b	7.2bc	9.0a	7.4b	8.6a	4.3f	0.63
PM	no.	217.5f	294.2e	345.8cd	332.5d	402.5b	282.5e	358.8cd	367.5c	333.3d	438.3a	357.9cd	401.7b	203.3f	27.9
TS	no.	186.0 a	167.0b	84.9g	137.9c	79.5g	141.2c	119.9d	133.5c	103.3ef	111.7de	133.8c	97.6f	142.0c	12.22
FSP	%	77.1de	70.5f	90.9ab	80.7cd	93.2a	81.5c	78.4cde	71.7f	81.8c	87.1b	70.8f	79.3cde	75.8e	3.87
BS	no.	42.0b	52.1a	7.8f	26.7cd	5.5f	25.8cd	28.6c	39.2b	20.3de	15.1e	38.3b	20.1de	43.9b	7.25
TGW	gr.	25.0f	23.5g	25.1f	22.8h	23.8g	22.2i	23.5g	27.3b	26.3d	29.3a	27.0c	26.0e	27.3bc	0.26
LSWP	gr.	62.5a	33.7fgh	24.8i	30.3ghi	26.5hi	53.1bc	39.7def	46.0cd	56.7ab	39.2def	41.2de	61.1a	36.5efg	7.44
PWP	gr.	12.2a	11.3ab	5.0g	7.3ef	5.7f	9.5bcd	11.3ab	10.9abc	8.3de	8.9cde	9.5bcd	9.4cde	7.2ef	2.02
TW	gr.	74.7a	44.9def	29.8h	37.6fgh	32.2gh	62.6abc	51.0cde	56.9bcd	65.0ab	48.1def	50.7cde	61.0bc	43.7fgh	12.17
LSWM	gr.	61.8a	34.9def	39.2cde	31.1f	32.9fe	44.3bc	47.0b	36.2def	35.0def	41.0bcd	64.4a	47.2b	39.4cde	7.23
PWM	gr.	30.7def	32.6cde	34.9bcd	35.0bcd	29.2ef	40.5a	30.5def	27.4f	35.5bc	37.5ab	39.0ab	31.6c-f	32.7cde	4.58
ALP	gr.	0.16def	0.24a	0.17cde	0.20bc	0.18b-e	0.15ef	0.21b	0.19bcd	0.13f	0.18b-e	0.20bcd	0.15ef	0.17cde	0.04
DMA	gr.	92.5b	67.5efg	74.1def	66.0fg	62.1g	84.8bc	77.5cde	63.6g	70.5d-g	78.5cd	103.4a	78.8cd	72.1d-g	10.41
PY	kg ha^{-1}	6850abc	6272de	6063e	6317cde	6345cde	6958ab	6645bcd	6773bcd	6347cde	7127ab	7152ab	7365a	5903e	576.62
BY	kg ha^{-1}	7147bc	7660ab	6550c	7692ab	7193bc	7033bc	7150bc	7218bc	6975bc	7125bc	7250bc	8625a	7240bc	1017.30
HI	%	49.2abc	45.1d	48.7a-d	45.0d	47.1a-d	49.8ab	48.5a-d	48.6a-d	47.7a-d	50.2a	49.6ab	46.2bcd	45.5cd	3.86

*: Values within a column followed by same letter are not significantly different at LSD (P≤0.05). *: Refer to Table S2 for abbreviation description.

Table S4. Simple correlation analysis between VAR variables (phonological traits) and WITH variables (other investigated traits: agronomic traits, PY and HI) in 13 improved rice cultivars.

Correlation	LAI	LN	FLL	PL	PH	TTH	FTP	PM	TS	FSP	TGW	LSWP	PWP	LSWM	PWM	ALP	PY	HI
DG	0.21^{*}	-0.54^{**}	0.45^{*}	0.14	0.15	-0.56^{**}	-0.12	-0.58^{**}	0.71^{**}	-0.22^{*}	-0.35^{*}	0.19	0.39^{*}	0.08	-0.03	0.19	-0.10	-0.03
DTP	0.10	0.64^{**}	-0.37^{*}	-0.23^{*}	-0.62^{**}	0.44^{*}	-0.22^{*}	0.41^{*}	-0.32^{*}	-0.15	0.55^{**}	0.19	-0.15	0.03	0.19	-0.28^{*}	0.36^{*}	0.07
DT	0.34^{*}	0.16	0.02	-0.06	-0.34^{*}	0.14	-0.13	0.13	0.27^{*}	-0.18	0.19	0.08	0.19	0.10	0.06	0.13	0.25^{*}	-0.07
DS	0.14	0.54^{**}	-0.60^{**}	-0.43^{*}	-0.75^{**}	0.50^{**}	0.05	0.50^{**}	-0.34^{*}	0.05	0.11	-0.01	-0.25^{*}	-0.18	0.13	-0.15	0.24^{*}	0.03
DH	0.36^{*}	-0.02	0.21^{*}	0.33^{*}	0.03	-0.34^{*}	-0.57^{**}	-0.41^{*}	0.51^{**}	-0.77^{**}	0.23^{*}	0.61^{**}	0.52^{**}	0.30^{*}	-0.04	-0.08	0.16	-0.09
DP	0.37^{*}	-0.11	0.28^{*}	0.33^{*}	0.06	-0.41^{*}	-0.55^{**}	-0.48^{**}	0.60^{**}	-0.77^{**}	0.17	0.60^{**}	0.56^{**}	0.31^{*}	-0.05	-0.04	0.12	-0.08
DPM	-0.05	-0.15	0.49^{**}	0.44^{*}	0.13	-0.33^{*}	-0.29^{*}	-0.36^{*}	0.58^{**}	-0.83^{**}	0.07	0.45^{*}	0.63^{**}	0.36^{*}	-0.01	0.19	0.27^{*}	-0.05
DHD	0.07	-0.02	0.21	0.12	-0.02	-0.26^{*}	-0.42^{*}	-0.31^{*}	0.40^{*}	-0.72^{**}	0.18	0.59^{**}	0.50^{**}	0.45^{*}	-0.23	-0.09	0.24^{*}	-0.09

* and ** show the probability at 5 and 1 percent level, respectively.

*: Refer to Table S2 for abbreviation description.

Table S5. Canonical correlations and their significant probability level.

Canonical variables	Canonical correlation	Square of canonical correlation	Eigenvalue	Proportion	Cumulative	Pr > F
1	0.98	0.98	43.93	0.43	0.43	<0.0001
2	0.98	0.97	27.69	0.27	0.70	<0.0001
3	0.97	0.94	16.47	0.16	0.86	<0.0001
4	0.94	0.89	7.92	0.08	0.93	<0.0001
5	0.87	0.76	3.24	0.03	0.96	0.0071

Table S6. Correlation between the VAR variables (phonological traits) and the canonical variables of the WITH variables (other investigated: agronomic traits, PY and HI).

Phenological traits	Canonical variables (CV)				
(Variables VAR)	CV_1	CV_2	CV_3	CV_4	CV_5
DG	-0.45*	0.63**	0.04	-0.44*	0.22*
DTP	0.85**	0.09	0.19	0.25*	-0.28*
DT	0.22*	0.50**	-0.08	-0.08	0.15
DS	0.53**	0.37*	-0.05	-0.46*	-0.39*
DH	0.22*	0.25*	0.82**	-0.37*	0.06
DP	0.12	0.32*	0.78**	-0.43*	0.09
DPM	-0.03	0.17	0.79**	-0.13	-0.07
DHD	0.14	0.11	0.72**	-0.40*	-0.34*
Standardized variance	0.17	0.13	0.31	0.12	0.05

* and ** show the probability at 5 and 1 percent level, respectively.

*: Refer to Table S2 for abbreviation description.

Table S7. Correlation between the WITH variables (other investigated: agronomic traits, PY and HI) and the canonical variables of the VAR variables (phonological traits).

Other traits	Canonical variables (CV)				
(WITH variables)	CV_1	CV_2	CV_3	CV_4	CV_5
LAI	0.27*	0.43*	0.15	-0.12	0.37*
LN	0.64**	-0.05	-0.03	0.31*	-0.09
FLL	-0.33*	-0.07	0.17	-0.35*	0.21*
PL	-0.24*	-0.18	0.49*	0.07	0.48*
PH	-0.54**	-0.49*	0.18	-0.23*	0.31*
TTH	0.39*	-0.10	-0.25*	0.50**	-0.12
FTP	-0.36*	-0.07	-0.34*	0.43*	-0.15
PM	0.34*	-0.11	-0.29*	0.55**	-0.14
TS	-0.36*	0.44*	0.38*	-0.34*	0.20*
FSP	0.04	-0.20*	-0.82**	0.07	-0.02
TGW	0.78**	-0.24*	0.02	-0.20*	0.14
LSWP	0.17	0.13	0.43*	-0.44*	-0.10
PWP	-0.28*	0.21*	0.53**	-0.17	0.05
LSWM	0.02	-0.09	0.22*	-0.34*	-0.21*
PWM	0.26*	0.12	-0.04	0.26*	0.30*
ALP	-0.43*	0.15	0.11	0.36*	0.12
PY	0.24*	0.09	0.12	0.09	-0.12
HI	0.12	-0.004	-0.10	0.04	-0.03
Standardized variance	0.14	0.05	0.11	0.10	0.04

* and ** show the probability at 5 and 1 percent level, respectively.

*: Refer to Table S2 for abbreviation description.

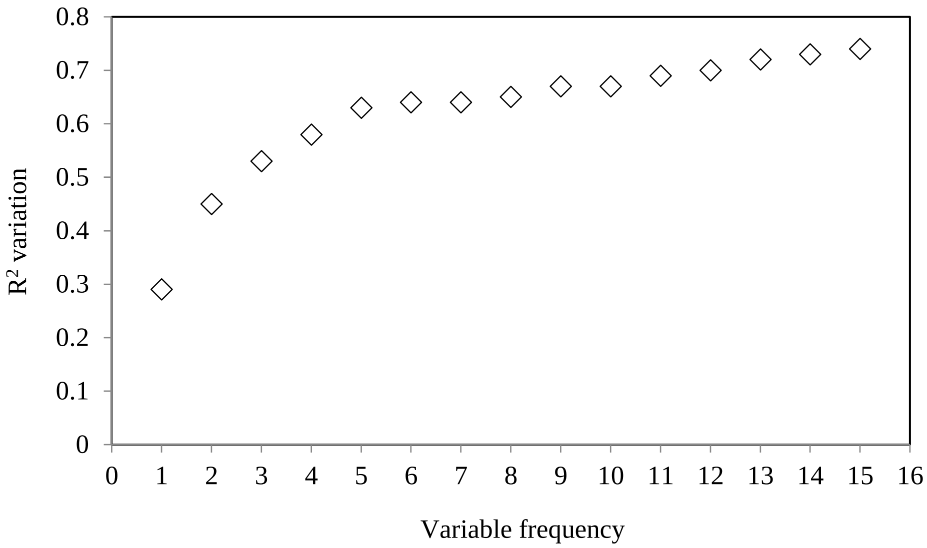

Figure S1. R^2 changes due to the increased number of variables that affect yield.

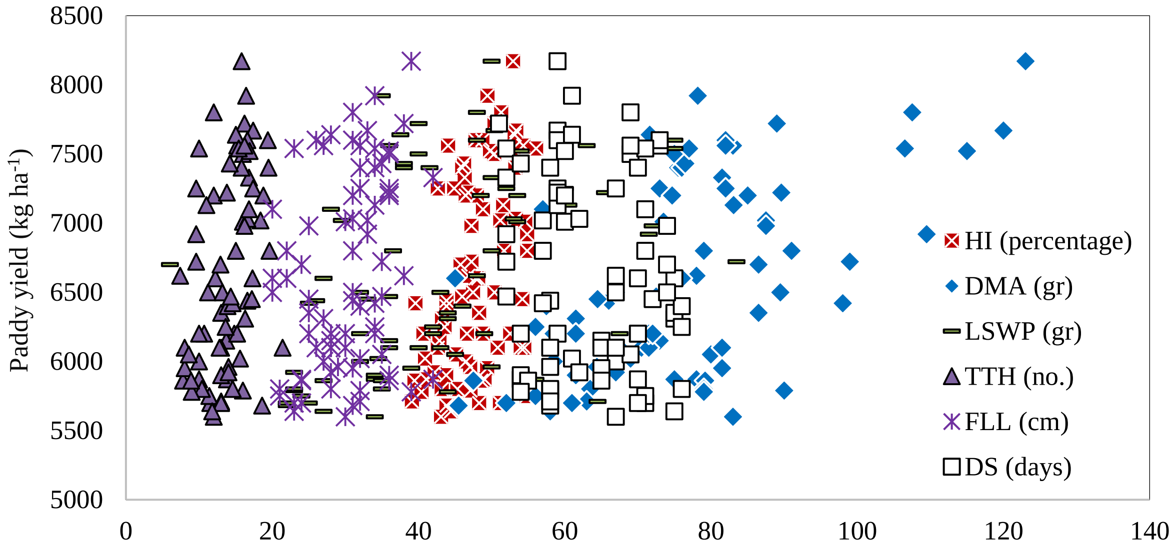

Figure S2. Correlation between paddy yield (PY) with DS, FLL, THH, LSWP, DMA and HI.

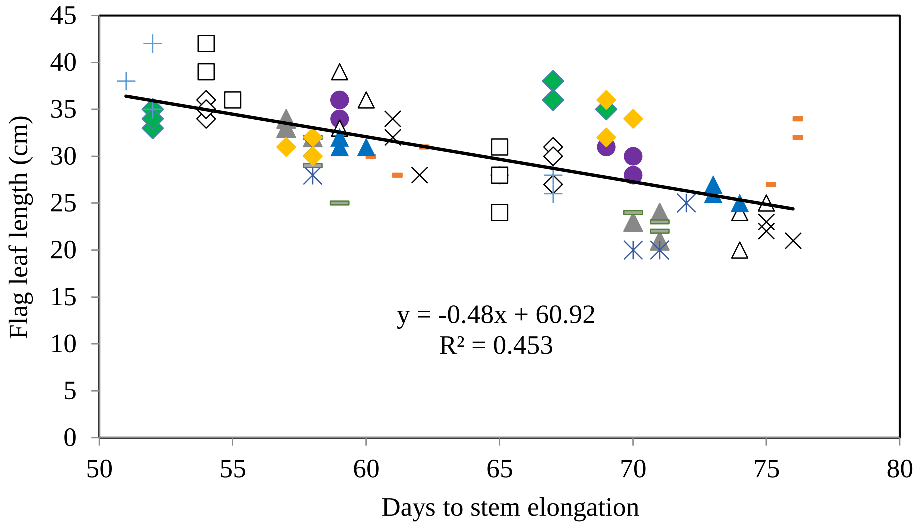

Figure S3. The relationship between flag leaf length (FLL) and the number of days to stem elongation (DS).

Results and discussion

Mean comparison analysis (GLM) of investigated traits in 13 improved cultivars

According to the findings, the vegetative and reproductive periods of *cv.* 'Neda' were higher than other cultivars. In this term, *cv.* 'Shiroodi', *cv.* 'Dasht' and *cv.* 'Nemat' stood ranks next. The lowest growth periods were observed in *cv.* 'Pazhouhesh' (Table S3). Mean comparison of agronomic traits showed that among genotypes, *cv.* 'Fajr' had the highest LA (0.34 m^2) and LAI (8.6), but *cv.* 'Neda' demonstrated the most LN (107.5 leaves). The highest FLL and TS belonged to *cv.* 'Dasht', the main cause of which was the higher FLA of this cultivar. In addition, the highest BS belonged to *cv.* 'Amol 3'. Furthermore, *cv.* 'Pardis' and *cv.* 'Kados' had the highest FTP (99.3%). The maximum TTH belonged to 'Pazhouhesh', Sahel and Neda cultivars. Between all cultivars, *cv.* 'Khazar' and *cv.* 'Kados' achieved the highest PL, SL and PH (Table S3). Mean comparison shows differences of all investigated traits between the three cultivars. Genotypes mean comparison showed that the highest TC belonged to *cv.* 'Phazhouhesh', *cv.* 'Sahel', and *cv.* 'Neda'. The high PM and TGW in *cv.* 'Sahel' has its genetic potential. The most TS (186 spikelets) was obtained in *cv.* 'Dasht', but FSP was highest in *cv.* 'Phazouhesht' (Table S3).

The highest LSWP, PWP, TW, and LSWM were observed for *cv.* 'Dasht'. The maximum PWM (40.5 gr) was related to *cv.* Keshvari', but the highest ALP (0.24 gr) was belong to *cv.* 'Amol 3'. In addition, *cv.* 'Shiroodi' had the highest DMA (103.4 gr). All the investigated cultivars were varied in terms of phenological traits, morphological indices, agronomic traits and PY components which cause variable function and PY (Table S3).

Mean comparison of cultivars showed that the highest PY (7365 kg ha^{-1}) and BY (8625 kg ha^{-1}) belonged to *cv.* 'Neda. In addition, *cv.* 'Shiroodi' and *cv.* 'Keshvari' got ranks next. The least PY (6063 kg ha^{-1}) was attributed to *cv.* 'Ghaem' (Table S3). In terms of PY and BY, cultivars had a high variation. Mean comparison showed that *cv.* 'Sahel' had higher HI (50.2%) (Table S3).

Correlation between the VAR variables (phenological traits) and the canonical variables of WITH variables (agronomic traits, PY, and HI)

From first to fifth variables was observed 17%, 13%, 31%, 12%, and 5% of the variation in this group (Table S6). The highest correlation in the first canonical variable was related to DTP ($r = 0.85^{**}$), and DS ($r = 0.53^{**}$) got rank next. Moreover, the first canonical variable explained significant positive correlation with DT and DH, but DG ($r = 0.45^{*}$) revealed significant negative correlation. The maximum correlation in the second canonical variable was related to DG ($r = 0.63^{**}$) and DT ($r = 0.50^{**}$). The second canonical variables demonstrated significant positive correlation with DS, DH and DP (Table S6). The most correlation in third canonical variables was related to DH ($r = 0.82^{**}$), DP ($r = 0.78^{**}$), DPM ($r = 0.79^{**}$), and DHD ($r = 0.72^{**}$), respectively. The highest correlation of fourth and fifth canonical variables was related to DTP ($r = 0.25^{*}$) and DG ($r = 0.22^{*}$), respectively. The fourth canonical variables demonstrated significant negative correlation DG, DS, DH, DP, and DHD. The fifth canonical variables demonstrated significant negative correlation with DTP, DS, and DHD (Table S6).

Correlation between the WITH variables (agronomic traits, PY, and HI) and the canonical variables of VAR variables (phenological traits)

This group revealed 14%, 5%, 11%, 10%, and 4% of the variation from first to fifth variables (Table S7). The maximum correlation in the first canonical variable was related to TGW ($r = 0.80^{**}$), and LN ($r = 0.64^{**}$) stood rank next. In addition, the first canonical variable explained significant positive correlation with LAI, TTH, PM, PWM, and PY, other traits (FLL, PL, PH, FTP, TS, PWP, and ALP) shows significant negative correlation. The highest correlation in the second canonical variable was related to TS ($r = 0.44^{*}$) and LAI ($r = 0.43^{*}$). These canonical variables showed significant positive correlation with TS, LSWP, and LSWM, but other traits (TTH, FTP, PM, and FSP) demonstrated significant negative correlation (Table S7). The highest correlation of fourth and fifth canonical variables was related to PM ($r = 0.55^{**}$) and PL ($r = 0.45^{*}$), respectively. The fourth canonical variables showed significant positive correlation with LN, PM, PWM and ALP, but FLL, PH, TS, TGW, LSWP, and LSWM shows significant negative correlation. The fifth canonical variables demonstrated significant positive correlation with FLL, PH, TS, and PWM, but LSWM ($R = -0.21^{*}$) shows significant negative correlation (Table S7).

References

Harman HH. 1976. Modern factor analysis. 3rd Ed. University of Chicago Press, Chicago, pp. 376.

Johnson RA, Wichern DW. 2002. Applied Multivariate Statistical Analysis. Prentice-Hall, Upper Saddle River, NJ.

Khattree R., Naik DN. 2000. Multivariate Data Reduction and Discrimination with SAS Software. SAS Institute, Cary, NC.

Soltani A, Hajjarpour A, Vadez V. 2016. Analysis of chickpea yield gap and water-limited potential yield in Iran. Field Crops Res. 185:21-30. doi:10.1016/j.fcr.2015.10.015

Soltani A, Rezaei A, Khajehpour M, Mirlohi A. 2000. Communication and share of various morphological and physiological traits in determining of grain yield sorghum. J Agr Sci Nat Resour. 4:85-94.

Tadesse W, Bekele E. 2001. Factor analysis of yield in grasspea (*Lathyrus sativus* L.). Lathyrus Lathyrism Newsletter 2:416-421.

Yan W. 2001. GGE biplot-a windows application for graphical analysis of multi environment trial data and other types of two-way data. Agron J. 93: 1111–1118.doi:10.2134/agronj2001.9351111x