Reduction of traits for genetic evaluation of linear described traits in the Old Kladruber horse

L. Vostrý[1,2], J. Přibyl[1,2], P. Šimeček[2]

[1]Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Prague, Czech Republic
[2]Institute of Animal Science, Prague-Uhříněves, Czech Republic

ABSTRACT: The estimated genetic parameters were used for selection of important conformation traits and reduction in the number of the described characteristics. For reduction in the number of described traits from the original 36 to 24, the traits were selected according to: measure of genetic similarity (cluster analysis), measure of uncertainty multidimensional quantity, value of the variance of aggregate genotype, value of the variance of selection index and correlation of trait to the first principal component of the genetic matrix. Reduction in the number of the described traits was based on the value of heritability coefficient and genetic correlations matrix. The reliabilities of selection indexes were estimated between 0.41 and 0.53. Among the three multivariate analysis methods evaluated in this study, the variance of selection index had the highest reliabilities of selection indexes. The estimation of selection index variance which omitted traits with low heritability coefficient and high genetic correlation was the most suitable for the traits selection. This procedure would enable the breeders to reduce field costs (e.g. time, labour) required for obtaining the genetic parameter estimates necessary for a specific breeding programme.

Keywords: multivariate analysis methods; reduction in the number of traits; selection index; Old Kladruber horse

The Old Kladruber horse is the most important Czech genetic resource. It was established by the end of the 18th and the onset of the 19th century. The breed existing today is a warmblood created on the basis of Old Spanish and Old Italian horses and it has been bred continuously in the Czech Republic for more than two hundred years. The breed is a robust carriage (coach) horse, which was originally used for ceremonial purposes by the Habsburg emperors and the clerical hierarchy. The population was closed against gene immigration from related breeds of Old Spanish origin in 1992. Due to imperial tradition, the breed consists of a grey and a black variety.

A linear-type description is routinely used mainly in cattle, where a number of analyses were performed (Brothostone, 1994; Veselá et al., 2005; Bouška et al., 2008). Body conformation in the horse was evaluated first by a scoring system, while a linear-type description was introduced later. The linear-type description and description of the morphological structure in horse breeds were studied e.g. by Jakubec et al. (1999, 2007) in Old Kladruber horses, Vostrý et al. (2011b) in Czech cold-blooded horses, Pretorius et al. (2004) in Friesian horses, Zechner et al. (2001) and Baban et al. (1998) in Lipizzan horses, Molina et al. (1999) in Andalusian horses, Samoré et al. (1997) in the Hafling breed, Hartman et al. (1994) in riding horses, and by Van Bergen and Van Averdonk (1993) in Shetland ponies. Koenen et al. (1994) investigated the relationship between body conformation and performance

---

Supported by the Ministry of Education, Youth and Sports of the Czech Republic (Project No. MSM 6046070901) and by the Ministry of Agriculture of the Czech Republic (Project No. MZE 0002701404).

in Dutch Warmblood riding horses. Estimations of genetic parameters in horses were also studied by Bokor et al. (2007) who concentrated on possibilities of selection in racehorses.

Understanding relations among morphologic traits, i.e. their interdependence (correlation coefficients) is very important in animal breeding so as to determine the breeding criteria and possible breeding response of selection programmes. Knowledge of genetic structure and variability of traits are priceless for further decisions which traits should be included in the selection process. Selection of important traits and constructing of the selective index generated e.g. Park et al. (2011) or Sokolović et al. (2011).

The objective of our study was the selection of important conformation traits and reduction in the number of described characteristics from the present 36 to 24 traits. The reduction in the number of evaluated traits was intentionally illustrated on the Old Kladruber breed with regard to a high number of described traits and a lower number of individuals, taking into account lower calculation requirements.

MATERIAL AND METHODS

Data

The data analysis was based on the genetic correlation matrix and coefficients of heritability of the 36 traits (Vostrý et al., 2011a). The set for genetic parameters and genetic correlations comprised 977 individuals recorded over a period of 16 years (1990–2006). Repeated measurements of 40% of the animals were included in the database (1390 observations in total). Repeated measurements were also included in the analysis of the random effect of an animal's permanent environment.

For the estimation of genetic variance and genetic covariance between all the traits BLUP animal model was used:

$$y = \mathbf{X}b + \mathbf{Z}a + \mathbf{W}pe + e$$

where:

y = vector of observations of the linear description

b = vector of fixed effects (colour variant (white, black), breed (National stud, private herds), sex (stallions and mares), age at description (1, ..., ≥ 8), year of description (m = 1995, ..., 2006), breed × colour variant interaction (1, ..., 4)

a = vector of random additive genetic effects

pe = vector of random permanent environmental effects

e = vector of random errors

$\mathbf{X}, \mathbf{Z}, \mathbf{W}$ = incidence matrices associated with b, u and pe, respectively

For more information on estimation of genetic parameters and genetic correlation see Vostrý et al. (2011a).

The traits were included in the following four categories:

(1) **Front**: height at withers (HW) (transformed from measured values), heftiness (HF), head profile (HP), neck length (NL), neck tethering (NTe), neck topside (NTo), withers length (WL), withers height (WH)

(2) **Body**: topline length (TL), topline vault (TV), loin length (LL), loin vault (LV), shoulder length (SL), shoulder slope (SS), chest girl (ChG), chest length (ChL), chest width (ChW), chest depth (ChD), breast width (BW)

(3) **Rear**: croup length (CrL), croup width (CrW), croup shape (CrS), croup profile (CrP), tail tethering (TT)

(4) **Limbs**: forelimbs - front view (FlFV), forelimbs side view (FiSV), foretoe side view (FtSV), foretoe length (FtL), forehoof side view (FhSV), forehoof width (FhW), forehoof size (FhS), hind limb side view (HLSV), hind limb rear view (HLBV), fetlock from the side (FlFS), length of the fetlock (LF), angle of hoof wall (AHW)

Reduction in the number of the described traits

Six methods were compared to reduce the number of the described traits. Except for one method, n-tuples (1, 2, ..., 12) of traits were omitted from the set of 36 traits, and significance was tested in the remaining traits. The n-tuples of traits were selected by combinatory analysis.

(1) Measure of genetic similarity between the traits

The level of genetic similarity was assessed by cluster analysis using the VARCLUS procedure (SAS, 2005). The genetic correlation matrix between all the 36 described traits provided the input data. Reduction in the number of the described traits consisted of the omission of the traits which reached a low h^2 value and at the same time were in a cluster near the other traits with a high h^2 value, which indicates the high genetic similarity of the traits concerned.

(2) Measure of uncertainty

This method is based on the assumption that after reduction in the number of traits, the variability of a multivariate variable will be described in a worse manner, and the level of uncertainty will reach lower values.

The level of uncertainty is a function which assigns a non-negative number to the genetic variance-covariance matrix and which fulfills the following requirements:

(*i*) addition of a new variable to the set does not decrease the uncertainty (while equality is attained if and only if the added variable equals a variable already included in the set multiplied by a constant)

(*ii*) uncertainty reaches its maximum for the given diagonal of the variance-covariance matrix if and only if there are just zeros outside the diagonal (e.g., given variables are independent)

The validity of condition (*i*) can be deduced from Cauchy-Schwartz inequality (see e.g. Hardy et al., 1952). The validity of condition (*ii*) is obvious.

The higher the uncertainty, the more independent are the variables in the set and the greater is their number. The uncertainty measure is defined

$$f(v)=\sqrt{tr\,(\mathbf{G})^2-\sum_{ij}g_{ij}^2}$$

where:

$tr(\mathbf{G})$ = trace of genetic variance-covariance matrix

g_{ij} = element *(ij)* of genetic variance-covariance matrix **G**

The individual traits were stepwise omitted and the lower reduction value of uncertainty ($f(v)$) was determined by their omission.

(3) Value of the variance of aggregate genotype

This method employs the first principal component to define aggregate genotype in a multi-trait selection index. The first principal component is the normalized linear combination of all the variables with maximum variance.

Therefore, for the additive genetic covariance matrix **G** the first principal component is the vector of normalized relative weights a which would maximize variance in the aggregate genotype (Arnason, 1984):

$$H = a'g$$

where:

H = value of aggregate genotype

a = the first principal component

g = vector of genetic values

Setting a equal to the first principal component of the **G** matrix results in maximization of the expression of the aggregate genotype variance (Arnason, 1984):

$$\sigma_H^2 = a'\mathbf{G}a$$

where:

σ_H^2 = variance of aggregate genotype

a = the first principal component

G = genetic variance-covariance matrix

The number of traits was reduced by the traits in which the highest value of aggregate genotype variance was determined by their omission. The selection goal was changed by this method at the same time on the basis of the omitted traits.

(4) Variance of selection index at a reduction in traits in the selection goal

Similarly to method 3, maximization of selection index variance by means of the first principal component was used. However, besides the variance-covariance genetic matrix, the variance-covariance phenotype matrix was also included in this method (Arnason, 1984):

$$\sigma_{I1}^2 = a'\mathbf{G}\mathbf{P}^{-1}\,\mathbf{G}a$$

where:

σ_{I1}^2 = variance of selection index

a = the first principal component of the matrix $\mathbf{G}\mathbf{P}^{-1}\mathbf{G}$

G = genetic covariance matrix

P = phenotype covariance matrix

As in method 3, the number of traits was reduced by traits in which the highest value of selection index variance was determined by their omission. By applying this method, the traits both in the selection index and in the selection goal were reduced. The number of traits was reduced in **G** and **P** matrices in this method.

(5) Variance of selection index while maintaining the selection goal

In this method the value of selection index variance was estimated when the number of described traits in the selection index was reduced and the number of the described traits in the selection goal was maintained. The traits to be omitted were selected on the basis of the value of selection index variance, while the values of the selection goal were maintained:

$$\sigma^2_{I2} = a'\mathbf{C}\mathbf{P}^{-1}\,\mathbf{C}a$$

where:

σ^2_{I1} = variance of selection index when the traits in selection goal are maintained

a = the first principal component of matrix **G**

C = genetic covariance matrix between the traits of matrix **P** and **G**

P = phenotype covariance matrix

As in methods 3 and 4, the number or traits was reduced by traits in which the highest value of selection index variance was determined by their omission.

(6) The value of correlation of a trait with the first principal component

This method selected those traits which had the highest value of correlation with the first principal component of the genetic variance-covariance matrix **G** (Flak, 1994):

$$r_{G,xi} = \frac{a'G_i}{\sqrt{\sigma^2_H\ \sigma^2_{xi}}}$$

where:

$r_{G,xi}$ = correlation of a trait with the first principal component

a = the first principal component of the genetic variance-covariance matrix **G**

G_i = vector of the variance-covariance genetic matrix attributed to trait i

σ^2_H = variance of aggregate genotype

σ^2_{xi} = genetic variance of trait i

Traits with a high value of $r_{G,xi}$ have a greater significance in the description of the aggregate multivariate variable than those with a low value of $r_{G,xi}$. This is the reason why 12 traits were omitted in which the lowest absolute value of correlation with the first principal component was estimated ($r_{G,xi}$).

RESULTS AND DISCUSSION

Table 1 shows that significant genetic correlations were estimated for 32% of the traits (18% – $P < 0.01$ and 14% – $P < 0.05$) and that coefficients of heritability for all the traits were estimated to be statistically significant ($P < 0.01$ or $P < 0.05$).

Reduction in the number of the described traits

Reduction in the number of the described traits from 36 to 24 was done on breeders' requests. It is advisable to omit those traits of linear type description which do not contribute to the total description of a multivariate variable in a marked way. Using multivariate methods for investigation of such a variable material as are the conformation traits was the best choice, also as concerns reduction of their number.

(1) Measure of genetic similarity between the traits

The values of correlation coefficients were used as input data for cluster analysis (Figure 1). By applying cluster analysis, 19 clusters were formed which explained 83% of the variability. The particular clusters mostly corresponded to the above-mentioned 4 groups of traits: front, body, rear and limbs. Only the traits HF, CrW, LV, and FlFV were included in the cluster of the other groups. It is similar with the trait SS, which did not, however, show such a close relationship to the other traits. On the contrary, no close correlation of NL, HP, TV, FiSV, ChD, HLBV, HLSV with the other traits was determined. Negative correlations were estimated between the traits SS, SL, LV, ChW, ChD, ChG, BW, CrW, CrL, CrP, FlFV, FiSV, NTe, NTo, HF, HLBV, FtSV, AHW, FlFS, FhSV, FhW, FhS, FtL, LF on the one hand, and the traits LL, CrS, ChL, TT, NL, TL, HLSV, HW, WL, WH, HP, TV on the other.

In order to reduce the number of the described traits, it was advisable to omit the traits with low coefficient h^2 and with high genetic similarity to the other traits. The value of 0.10 was set as the limit value of coefficient h^2. Table 1 shows that 8 traits have lower or identical values: HLSV, HLBV, SS, FlFS, SL, CrL, FlFV, and LL. Of these 8 traits only 6 traits can be omitted (SS, FlFS, SL, CrL, FlFV, and LL). The reason is that the traits HLBV and HLSV have low values of h^2 (0.04 and 0.05), but unlike the remaining 6 traits they are not highly genetically similar to the other traits. As illustrated in Figure 1, which shows the percentage of variability explained by the particular clusters on the x-axis, if these

Table 1. Coefficient of heritability (diagonal), genetic correlations (above diagonal), and phenotypic correlations (below diagonal) among 36 traits of linear type description (Vostrý et al., 2011a)

	HW	HF	HP	NL	NTe	Nto	WL	WH	TL	TV	LL	LV	SL	SS	ChG	ChL	ChW	ChD	BW	CrL	CrW	CrS	CrP	TT	FlFV	FiSV	FtSV	FtL	FhSV	FhW	FhS	HLSV	HLBV	FlFS	LF	AHW
HW	0.57	**0.71**	**0.24**	**0.28**	*0.24*	*0.14*	**0.43**	**0.43**	*−0.19*	*0.16*	*0.29*	−0.06	0.11	0.02	**0.52**	0.02	−0.11	*−0.15*	*−0.15*	**−0.44**	*−0.18*	0.04	**−0.39**	−0.07	−0.02	0.14	−0.11	0.08	−0.11	**0.22**	**0.37**	−0.07	*−0.28*	0.15	0.06	−0.05
HF	0.56	**0.49**	**0.19**	0.11	*0.30*	0.13	0.14	0.01	*−0.20*	*−0.15*	0.09	0.08	0.20	0.02	**0.62**	−0.07	**0.37**	0.13	**0.34**	*−0.28*	*0.19*	−0.04	−0.09	−0.19	0.22	0.12	0.03	−0.07	0.09	**0.35**	*0.29*	−0.07	*−0.39*	0.05	−0.09	−0.08
HP	0.11	0.09	**0.65**	*0.22*	−0.07	0.04	**0.21**	*0.18*	−0.10	0.03	0.15	−0.03	*0.25*	*0.26*	*0.13*	*0.13*	*0.12*	0.05	*0.16*	−0.13	*−0.16*	0.00	**−0.32**	*0.15*	0.13	**−0.27**	0.07	*0.14*	**0.23**	**0.24**	*0.27*	−0.07	*−0.32*	*0.23*	0.09	−0.10
NL	0.03	0.03	0.12	**0.11**	**0.44**	**0.42**	0.00	−0.08	*0.28*	−0.21	0.34	−0.06	0.06	0.02	*0.29*	0.19	−0.04	−0.42	0.00	−0.26	−0.08	0.18	−0.15	0.06	0.09	*0.33*	−0.24	0.00	−0.05	−0.04	−0.06	−0.19	−0.14	−0.10	−0.06	−0.19
NTe	0.00	0.00	0.08	0.13	**0.12**	**0.72**	**−0.43**	**−0.47**	0.04	0.02	0.02	*0.39*	0.35	0.16	*0.37*	0.03	0.21	0.09	0.24	0.10	0.31	0.07	0.03	−0.14	−0.12	0.31	**−0.46**	−0.23	*−0.34*	0.03	−0.19	−0.28	−0.05	0.19	−0.18	−0.06
NTo	0.01	0.07	0.09	0.32	0.30	**0.33**	**−0.45**	**−0.34**	0.06	0.00	0.06	0.19	**0.44**	0.06	**0.50**	0.02	**0.32**	*0.23*	**0.37**	0.23	**0.35**	0.01	0.16	−0.13	0.01	**0.52**	−0.19	0.03	−0.07	0.13	0.08	0.12	−0.01	0.21	−0.15	−0.03
WL	0.18	0.10	0.09	−0.02	−0.32	−0.15	**0.19**	**0.68**	−0.19	0.15	*0.38*	*−0.25*	*−0.32*	−0.03	*0.20*	0.00	**−0.35**	−0.09	**−0.35**	**−0.53**	**−0.32**	0.02	**−0.40**	0.16	0.07	*−0.30*	−0.06	−0.03	−0.15	−0.15	0.04	0.21	0.01	−0.16	−0.06	0.11
WH	0.16	0.05	0.06	−0.01	−0.23	−0.19	0.49	**0.23**	0.02	**0.35**	**0.48**	**−0.44**	−0.21	−0.10	0.02	0.19	**−0.41**	*−0.22*	**−0.40**	**−0.44**	**−0.53**	0.01	**−0.51**	**0.34**	0.08	−0.21	0.16	*0.22*	0.05	0.04	0.12	0.12	0.12	−0.06	0.20	−0.04
TL	0.01	−0.02	0.01	0.06	0.02	−0.04	0.00	0.01	**0.28**	**−0.43**	**0.60**	**−0.68**	−0.09	**0.36**	−0.14	**0.84**	0.02	**−0.34**	0.02	**−0.51**	0.00	*0.20*	−0.14	*0.26*	−0.07	−0.22	0.12	0.02	0.11	0.04	−0.10	−0.09	−0.06	0.00	*−0.21*	*−0.28*
TV	0.04	−0.03	−0.01	−0.01	−0.07	0.02	0.07	0.08	−0.22	**0.28**	−0.11	*0.29*	−0.13	−0.30	−0.11	**−0.38**	*−0.27*	−0.01	*−0.22*	0.15	−0.14	−0.07	−0.10	0.03	0.04	−0.07	*−0.26*	0.19	−0.18	0.04	0.05	0.10	0.39	−0.01	**0.39**	0.05
LL	0.17	0.09	0.06	0.01	−0.05	−0.09	0.15	0.16	0.31	−0.08	**0.10**	**−0.69**	−0.21	0.06	0.19	**0.68**	*−0.28*	−0.18	−0.25	**−0.68**	**−0.39**	**0.43**	**−0.52**	**0.57**	−0.20	−0.12	−0.18	0.04	−0.07	−0.03	−0.06	0.17	−0.21	−0.19	−0.12	−0.29
LV	−0.08	−0.01	−0.03	0.00	0.09	0.09	−0.11	−0.16	−0.25	0.11	−0.35	**0.15**	0.15	−0.16	0.03	**−0.60**	0.17	0.18	0.18	**0.60**	*0.27*	*−0.24*	**0.38**	**−0.51**	−0.05	0.19	*−0.31*	−0.19	−0.13	0.04	−0.14	−0.04	0.20	0.10	0.14	*0.34*
SL	−0.02	0.08	0.02	0.12	0.00	0.17	−0.01	−0.04	−0.04	−0.01	−0.02	0.06	**0.07**	0.44	0.22	−0.04	**0.39**	0.18	**0.35**	0.33	*0.29*	−0.10	0.17	−0.28	0.37	0.12	0.00	−0.09	−0.26	0.06	0.07	−0.03	0.05	0.15	−0.14	−0.17
SS	−0.08	−0.05	−0.03	0.10	0.01	0.11	−0.10	−0.05	0.02	−0.05	0.00	0.03	0.30	**0.05**	0.01	**0.52**	0.27	−0.04	0.23	−0.20	0.16	0.07	−0.11	−0.06	0.12	*−0.46*	*0.28*	0.02	−0.01	−0.11	−0.03	−0.28	−0.06	0.35	−0.24	0.22
ChG	0.46	0.46	0.04	0.01	0.10	0.17	0.00	−0.08	−0.07	−0.04	0.03	0.03	0.06	0.01	**0.40**	0.01	*0.51*	**0.43**	**0.50**	−0.19	0.40	0.17	−0.03	−0.05	*0.29*	**0.35**	−0.13	*−0.24*	0.07	0.34	0.23	0.10	−0.16	0.12	**−0.36**	−0.07
ChL	0.11	0.09	0.07	0.11	0.02	−0.01	0.10	0.08	0.52	−0.19	0.30	−0.14	−0.05	0.05	0.04	**0.23**	0.10	−0.20	0.07	**−0.54**	−0.18	**0.32**	**−0.38**	**0.35**	−0.15	*−0.28*	0.04	0.02	0.12	*0.18*	0.03	−0.02	−0.27	0.07	−0.16	−0.07
ChW	−0.08	0.15	0.06	0.04	0.16	0.20	−0.17	−0.23	−0.03	−0.11	−0.12	0.15	0.12	0.10	0.26	0.05	**0.31**	**0.34**	**0.96**	0.15	**0.55**	0.01	*0.24*	*−0.22*	**0.57**	0.19	0.14	−0.04	**0.33**	**0.40**	0.08	−0.01	−0.20	0.11	−0.13	0.04
ChD	−0.05	0.07	0.01	0.02	0.07	0.14	−0.05	−0.15	−0.17	0.00	−0.12	0.12	0.12	−0.02	0.28	−0.03	0.33	**0.27**	**0.42**	*0.30*	**0.36**	0.03	0.14	−0.01	−0.03	−0.06	0.06	**−0.33**	0.15	**0.36**	*0.26*	0.13	0.16	0.28	−0.36	0.12
BW	−0.08	0.15	0.07	0.02	0.13	0.20	−0.15	−0.23	0.00	−0.09	−0.11	0.15	0.10	0.11	0.26	0.05	0.91	0.32	**0.33**	0.12	**0.52**	−0.03	*0.21*	*−0.20*	**0.61**	0.11	0.19	0.01	**0.34**	**0.43**	0.06	−0.08	−0.09	0.15	−0.13	0.05
CrL	−0.13	−0.06	−0.04	0.05	0.08	0.14	−0.09	−0.15	−0.16	0.08	−0.22	0.34	0.16	0.07	0.03	−0.08	0.14	0.12	0.15	**0.08**	*0.28*	−0.16	**0.55**	−0.14	−0.01	0.31	−0.04	−0.04	0.04	0.00	0.03	0.14	0.16	0.09	0.08	*0.30*
CrW	−0.10	0.10	−0.03	0.08	0.14	0.21	−0.17	−0.25	−0.07	−0.07	−0.15	0.18	0.17	0.05	0.25	−0.01	0.47	0.33	0.46	0.20	**0.35**	**−0.29**	**0.71**	**−0.44**	*0.29*	0.00	0.07	*−0.34*	0.05	**0.25**	0.11	0.00	0.17	*0.33*	**−0.52**	0.00
CrS	−0.04	−0.02	0.05	0.05	0.06	0.06	0.01	−0.04	0.11	0.01	0.10	0.00	0.00	0.05	0.02	0.16	0.03	0.01	0.03	0.09	−0.04	**0.35**	**−0.38**	**0.67**	−0.10	0.20	−0.15	0.17	0.03	*−0.16*	*−0.20*	−0.03	−0.12	−0.18	*0.27*	0.02
CrP	−0.18	−0.09	−0.09	0.03	0.15	0.16	−0.15	−0.16	−0.11	−0.05	−0.21	0.16	0.07	0.07	0.00	−0.11	0.24	0.10	0.25	0.23	0.39	−0.05	**0.25**	**−0.53**	0.14	0.09	0.16	−0.08	0.18	0.09	0.09	0.18	0.24	0.20	*−0.23*	0.15
TT	−0.04	−0.05	0.11	0.00	−0.02	−0.10	0.07	0.12	0.12	0.01	0.09	−0.08	−0.09	−0.03	−0.07	0.17	−0.12	−0.10	−0.11	−0.03	−0.20	0.29	−0.07	**0.17**	−0.10	−0.02	−0.03	*0.21*	0.12	−0.14	−0.14	−0.09	−0.22	−0.10	0.16	0.02
FlFV	−0.09	0.04	0.01	0.04	0.07	0.07	−0.04	−0.05	−0.03	−0.03	−0.06	0.08	0.07	0.08	0.09	0.02	0.46	0.18	0.49	0.14	0.23	0.03	0.17	−0.09	*0.08*	0.05	*0.34*	*0.27*	0.09	0.16	−0.02	−0.14	0.14	−0.09	0.12	−0.06
FiSV	0.00	0.00	−0.02	0.09	0.07	0.11	−0.03	−0.02	−0.01	0.01	0.00	0.04	0.03	−0.05	0.04	−0.01	0.07	0.03	0.07	0.04	0.02	0.01	0.01	−0.01	0.00	*0.10*	−0.26	0.11	−0.05	−0.08	−0.03	0.23	−0.21	−0.34	0.26	−0.14
FtSV	−0.06	−0.02	0.01	0.01	−0.03	0.02	0.02	0.02	0.01	−0.01	0.02	−0.05	0.06	0.10	0.01	0.04	0.09	0.09	0.10	−0.02	0.03	−0.02	0.04	−0.02	0.10	−0.06	**0.17**	**0.49**	**0.56**	0.15	0.26	−0.27	0.09	0.34	0.12	*0.26*
FtL	0.02	−0.06	0.03	0.09	−0.03	0.03	0.04	0.10	−0.03	0.08	0.02	0.01	0.02	0.01	−0.04	0.02	−0.04	−0.11	−0.04	0.03	−0.09	0.01	−0.02	0.05	0.00	0.01	0.26	**0.21**	**0.31**	−0.02	0.11	−0.08	−0.07	0.13	**0.72**	**0.32**
FhSV	−0.02	0.03	0.03	−0.04	−0.04	0.02	0.04	0.02	0.06	−0.05	0.02	−0.06	−0.01	0.01	0.05	0.05	0.07	0.02	0.06	−0.08	0.00	0.02	0.00	0.01	0.01	−0.04	0.34	0.11	**0.16**	**0.46**	**0.43**	−0.04	−0.12	*0.37*	0.08	*0.26*
FhW	0.13	0.25	0.12	0.01	0.00	0.03	0.09	0.06	0.01	0.01	0.00	0.04	0.03	0.01	0.19	0.08	0.16	0.12	0.16	0.02	0.11	−0.02	0.00	0.00	0.08	0.00	0.17	0.04	0.17	**0.41**	**0.73**	0.07	0.01	*0.36*	0.00	0.09
FhS	0.12	0.16	0.09	0.08	−0.05	0.02	0.08	0.07	−0.04	0.00	−0.01	0.03	0.09	0.10	0.10	0.06	0.11	0.12	0.11	0.05	0.13	−0.03	−0.01	−0.04	0.07	0.05	0.10	0.08	0.08	0.40	**0.14**	0.26	−0.09	0.35	−0.01	0.20
HLSV	−0.05	−0.10	−0.01	0.02	0.01	0.06	−0.02	−0.02	−0.04	0.05	−0.03	0.02	−0.02	0.02	−0.04	−0.02	−0.02	0.01	−0.02	−0.02	0.00	−0.03	0.04	−0.02	0.00	0.06	0.05	0.03	−0.04	−0.01	−0.02	*0.05*	0.06	−0.44	0.01	0.03
HLBV	−0.09	−0.05	−0.01	0.01	0.07	0.07	−0.04	−0.02	−0.04	0.01	−0.04	0.10	0.02	0.04	0.00	−0.01	0.11	0.10	0.11	0.06	0.22	−0.02	0.15	−0.08	0.14	0.01	0.07	0.02	0.03	0.04	0.05	0.04	**0.04**	0.06	0.05	0.13
FlFS	−0.04	−0.01	0.05	0.04	0.00	0.04	0.02	−0.01	−0.03	0.05	0.00	0.05	0.02	0.05	−0.01	0.02	0.06	0.08	0.07	−0.01	0.09	−0.04	−0.06	−0.11	0.03	−0.01	0.29	0.12	0.09	0.16	0.14	−0.05	0.01	*0.06*	−0.26	*0.37*
LF	0.01	−0.08	−0.01	0.04	−0.06	0.01	0.04	0.07	0.00	0.05	0.06	0.00	−0.03	0.01	−0.08	0.01	0.00	−0.07	−0.01	0.05	−0.11	0.03	−0.02	0.04	−0.04	−0.02	0.14	0.52	0.09	0.01	0.02	0.09	−0.02	0.11	**0.20**	0.17
AHW	−0.06	−0.05	−0.02	−0.04	0.04	−0.01	−0.01	0.03	−0.02	0.08	−0.03	0.00	−0.03	−0.01	0.02	−0.04	0.02	0.00	0.02	0.01	−0.03	0.03	−0.04	0.05	0.03	0.00	0.09	0.07	0.23	0.05	−0.01	−0.06	−0.01	0.20	0.12	**0.20**

Significance: normal = not significant, italics = $P < 0.05$, bold = $P < 0.01$

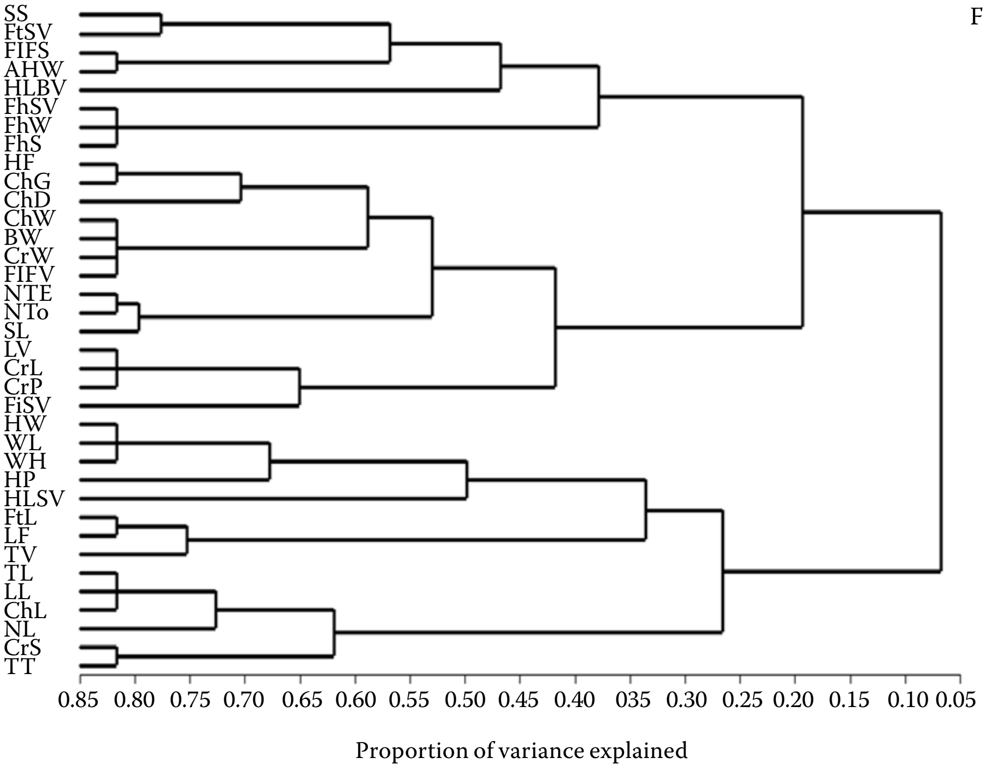

Figure 1. Cluster analysis

traits were omitted, the proportion of variability explained would decrease to 0.50 in HLBV and to 0.38 in HLSV. This could result in great undesirable changes in the selection of this population. To ensure the total of 24 described traits, another 6 traits with the lowest h^2 value and with high genetic similarity to the other described traits were omitted – NTe, LV, FhSV, TT, FhS, and WL. In these traits, h^2 ranged from 0.12 to 0.19. If the abovementioned 12 traits are omitted, the percentage of explained variability decreases at most to 78%. The cluster analysis to reduce the number of the evaluated traits and factor in plant breeding was previously used by Sokolović et al. (2011).

(2) Measure of uncertainty

Table 2 shows the values of uncertainty calculated on the basis of the omission of linear type description traits. When all the 36 traits of linear type description were included, the level of uncertainty was 16.23. Along with the traits omission, the value of uncertainty decreased evenly (ca. by 0.5%). HLBV was the first trait omitted. After the omission of 8 traits, the rate of a decrease in the value of uncertainty increased and after omitting 12 traits , the value of uncertainty decreased by ca. 1.5%. After the omission of the 12 required traits, the value of uncertainty was 14.84. These traits were selected for omission on the basis of the level of uncertainty: HLBV, SL, FhS, FlFS, SS, FtSVS, HLSV, CrL, NTe, FhSV, FtSV, and NL. As documented by the list of the traits, the traits HLBV and HLSV were selected. In these two traits no higher absolute values of genetic correlation to the other traits described were found. As shown in Table 1, the highest values of genetic correlation between these two traits did not exceed the absolute value of 0.39 for HLBV and 0.28 for HLSV. For these reasons, similar to those in cluster analysis, it is not advisable to omit these two traits from the linear type description of conformation. Comparison of the selected traits shows that this method lays greater emphasis on the value of genetic variance of traits, while the genetic relationship between the traits are less accentuated.

(3) Value of the variance of aggregate genotype

In the values of aggregate genotype (Table 2), a reduction in the number of evaluated traits leads to an increase in the values of aggregate genotype variance compared to the values of uncertainty. This is so because only the more important traits remain in the genotype. When the number of traits was reduced from 36 to 35, the value of aggregate genotype variance increased by 10%. With further reduction in the number of the traits, the increase

Table 2. The values of uncertainty, aggregate genotype variance and selection index variance with a reduction in the number of traits

Number of traits	$f(v)$	σ^2_H	σ^2_{I1}	$r^2_{I1,H}$	σ^2_{I2}	$r^2_{I2,H}$
36	16.23	2.48	1.37	0.538	1.3361	0.538
35	16.16	2.71	1.51	0.557	1.3360	0.538
34	16.09	2.89	1.63	0.563	1.3357	0.538
33	16.01	2.99	1.80	0.601	1.3351	0.538
32	15.94	3.08	1.83	0.593	1.3346	0.537
31	15.85	3.17	1.93	0.610	1.3339	0.537
30	15.76	3.28	1.98	0.605	1.3331	0.537
29	15.67	3.38	2.05	0.606	1.3317	0.536
28	15.55	3.45	2.12	0.614	1.3304	0.536
27	15.39	3.53	2.16	0.613	1.3286	0.535
26	15.22	3.60	2.19	0.609	1.3267	0.534
25	15.04	3.63	2.23	0.616	1.3240	0.533
24	14.84	3.63	2.23	0.615	1.3211	0.532

$f(v)$ = level of uncertainty, σ^2_H = aggregate genotype variance, σ^2_{I1} = selection index variance with a reduction in traits of selection goal, $r^2_{I1,H}$ = reliability between selection index and aggregate genotype with reduction in traits of selection goal, σ^2_{I2} = selection index variance, $r^2_{I2,H}$ = reliability between selection index and aggregate genotype with unreduced selection goal

in the value of aggregate genotype variance diminished until the value of aggregate genotype variance remained identical for the reduction in the number of traits from 25 to 24. Therefore, a further reduction in the number of traits was meaningless. Based on the value of aggregate genotype variance, the following traits should be omitted: LL, CrL, CrP, WL, WH, FlFS, SL, CrW, ChW, TT, FtL, and BW. The majority of the traits with low h^2 and high correlation with other traits of linear type description were omitted by this method, which satisfied the defined conditions for trait omission. The only exceptions were SS and FlFV, which reached the value $h^2 = 0.05$ or 0.08 and higher values of genetic correlation and were left in the evaluation. To reduce the number of traits to the required 24, the following traits were omitted: WL, WH, ChW, BW, CrW, CrP, TT, and FtL. These traits had a higher h^2 value than the set limit 0.10 (in CrW $h^2 = 0.35$), but they showed high values of genetic correlation with the other traits, and therefore they did not substantially contribute to the total genetic variance. This method reduced the number of the evaluated traits in the selection goal, but those traits, which did not contribute to the description of the multivariate variable in a significant way, were omitted.

(4) Variance of selection index with a reduction in traits in the selection goal

Like in the variance of aggregate genotype, by this method the reduction in the number of the described traits of conformation also led to an increase in the values of selection index variance (σ^2_{I1}) to the same extent (Table 2). On the basis of this method, these traits were omitted from the evaluation: WH, WL, CrS, CrW, SL, FlFV, TV, AHW, ChL, FlFS, LF, and FtL. Obviously, six traits were also selected on the basis of the value of aggregate genotype variance. As shown in equation (6), phenotype values were included through the matrix **P** in the estimation of selection index variance. However, these values were not adjusted for random errors of measurement. As already stated above, this result documents that in the remaining six traits environmental effects and errors of measurement act in opposite directions. The reliability between the aggregate genotype and selection index is expressed by the relation $r^2_{I1,H} = \sigma^2_I/\sigma^2_H$, where σ^2_H is the variance of the selection goal reduced by the respective n traits. Vector a maximizes the ratio of the standard deviation of the selection index from the aggregate genotype, hence maximizing the correlation between the selection index and the aggregate genotype. Therefore, a higher correlation between both variances was computed ($r^2_{I1,H}$ = 0.53 to 0.62).

(5) Variance of selection index while maintaining the selection goal

The omission of traits in the selection index while keeping to the selection goal led to a decrease in the values of selection index variance (σ^2_{I2}) compared to the two preceding methods (methods 3 and 4) – see Table 2. The following traits were omitted by this method: ChD, ChW, CrL, FhSV, FhS, FlFS, FtL, FtSV, HLSV, LL, Nte, and TL. Five traits were selected by methods 3 and 5, and two traits by methods 4 and 5. Similarly to methods 3 and 4, the first principal component *a*, which maximized the variance of the selection index, was used in this method. Because the traits from the selection index were omitted, the values of the selection index variance (σ^2_{I2}) and the values of reliability of the selection index ($r^2_{I2,H}$) were decreased. Differences between the estimated variances of the selection index (σ^2_{I2}) showed negligible changes. This means that the selected traits in the selection index do not provide any significant information for the description of the selection goal. However, the low values of reliability ($r^2_{I2,H}$), which ranged from 0.54 to 0.53 in relation to the number of omitted traits, suggest that the given selection index describes the overall selection goal with lower reliability.

Huang and Harding (1998) published that even though the first principal component (*a*) has maximum variance, selection may not achieve the maximum response. There are two additional concerns: (a) principal components may differ in different generations because phenotypic and genetic covariance matrix may be subject to change from generation to generation. This could lead to conflicting explanations of underlying genetic processes; (b) genetic eigenvalues could be negative if the genetic variance-covariance matrix is not positive definite. This happens mainly when lots of traits are being estimated at the same time. Means to overcome these problems include sampling more generations. To obtain positive definite variance-covariance matrices, the genotypic variance-covariance matrix was adjusted

Table 3. The values of the coefficient of correlation of traits with the first principal component of matrix **G**

No.	Trait	*r*	No.	Trait	*r*
1	HW	**–0.17**	19	BW	0.71
2	HF	0.26	20	CrL	0.56
3	HP	**–0.10**	21	CrW	0.80
4	NL	**–0.07**	22	CrS	–0.30
5	NTe	0.44	23	CrP	0.66
6	NTo	0.54	24	TT	–0.55
7	WL	–0.54	25	FlFV	0.29
8	WH	–0.67	26	FiSV	0.29
9	TL	–0.31	27	FtSV	**0.02**
10	TV	**–0.11**	28	FtL	**–0.24**
11	LL	–0.62	29	FhSV	**0.13**
12	LV	0.58	30	FhW	0.33
13	SL	0.45	31	FhS	**0.15**
14	SS	**0.03**	32	HLSV	**0.02**
15	ChG	0.43	33	HLBV	**0.09**
16	ChL	–0.38	34	FlFS	0.28
17	ChW	0.71	35	LF	–0.28
18	ChD	0.52	36	AHW	**0.12**

r = coefficient of correlation of a trait with the first principal component of matrix **G** by the weighted bending method (see Vostrý et al., 2011a).

(6) The value of correlation of a trait with the first principal component

The absolute values of correlation (r) between the particular traits and the first principal component ranged from 0.02 for HLSV to 0.79 for CrW (not published). Table 3 shows 12 traits with the lowest absolute value r (in bold). On the basis of this value, the following traits were omitted to obtain the required 24 described traits: HLSV, FTSV, SS, NL, HLBV, HP, TV, AHW, ThSV, FhS, HW, and FtL. The omission of these traits documents that both the traits with the highest h^2 – HP (h^2 = 0.65) and HW (h^2 = 0.57) and those with a low value of genetic correlation (r_G), with other traits (HLVS and HLBV) were omitted by this method. This method violated both conditions for trait omission, i.e. it omitted traits with a low value of r_G in relation to other traits as well as those with a high h^2 value.

Summarization of the reduction in the number of the studied traits

Table 4 shows the traits omitted by the particular methods. The FlFS trait was omitted by all used methods except for method 5. The traits FtL, FhSV, FhS, SL, and CrL were also omitted from the majority of the methods used with the exception of two methods (CrL – 4 and 5, FhS – 3 and 4, FtL – 1 and 2, and SL – 5 and 6). On the contrary, nine traits in Table 4 (BW, ChD, CrP, CrS, FtL, HP, HW, LV, TL) were omitted by only one method. This table also demonstrates that the highest number of identical traits (8) was omitted by methods 1 and 2.

Methods 3 and 4 satisfied the defined conditions for a reduction in linear type description traits in the best way. The reason is that method 1 was based on a subjective selection of traits, and method 2 placed a greater emphasis on the value of genetic variance of traits and less emphasis on correlations between the traits. The trait HLBV was omitted by method 5, which would have to be maintained in the description, in accordance with the low value of genetic correlation with the other traits. Method 6 violated both conditions for the omission of traits, i.e. it omitted both the traits with a low value of r_G (HLBV and HLSV) with the other traits and the traits with a high h^2 value (HP and HW). Methods 3 and 4 satisfied the criteria for a reduction in the number of described traits. As shown by the values of $r_{I1,H}$, these two methods are very similar. During reduction of the described traits we worked out a large matrix with very complicated mutual relationships. Altogether

Table 4. Comparison of the methods of selection

Trait	Methods					
	1	2	3	4	5	6
AHW				+		+
BW			+			
CrL	+	+	+		+	
CrP			+			
CrS				+		
CrW			+	+		
FhS	+	+			+	+
FhSV	+	+			+	+
FlFS	+	+	+	+	+	
FlFV	+	+		+		
FtL			+	+	+	+
FtSV		+			+	+
HLBV		+				+
HLSV		+			+	+
HP						+
HW						+
ChD					+	
ChL				+		+
ChW			+		+	
LF				+		
LL	+		+		+	
LV	+					
NL		+				+
NTe	+	+			+	
SL	+	+	+	+		
SS	+	+				+
TL					+	
TT	+		+			
TV				+		+
WH			+	+		
WL	+		+	+		
$r^2_{I2,H}$	0.507	0.506	0.413	0.448	0.532	0.484

31 out of 36 traits were omitted by the particular methods. Only five traits (HF, NTo, ChG, FhW, FlFS) were not omitted by any method. With the exception of traits of the rear, it was always one or two traits from each group of traits.

Because horse conformation is described on the basis of phenotype expression, the particular methods were compared as to how they complied to the unreduced selection goal. This evaluation was performed by means of correlation between the selection index and the aggregate genotype. The selection index was represented by *n*-tuples of traits remaining in the selection goal after the omission of 12 traits. The traits in the selection index were constantly changed in accordance with the given method. The selection goal was represented by the genetic covariance matrix comprising all 36 traits. The estimated reliabilities of selection indexes ($r^2_{I2,H}$) (Table 4) were calculated by method 5. As documented in Table 4, method 5 reached the highest value ($r^2_{I2,H}$) because the traits that maximized $r^2_{I2,H}$ were omitted from this method. The next highest value was equally reached by methods 1 and 2 ($r^2_{I2,H}$ = 0.507 and 0.506), even though method 2 did not lay great emphasis on genetic relations between the particular traits (r_G). In method 1 the traits were selected on the basis of subjective evaluation. Method 6, in which both conditions of omission (low h^2 value and high value of r_G) were violated, reached medium values of $r^2_{I2,H}$. On the contrary, the lowest values of $r^2_{I2,H}$ were estimated by methods which satisfied both conditions of omission – methods 3 and 4. Based on this knowledge and on the fact that the particular traits are described on the basis of phenotype expression, method 5 (Variance of selection index in the maintenance of selection goal) was chosen for reduction in the described traits, although the trait HLBV is not described by the selection index comprising the traits selected by method 5. Nevertheless, this selection index complies with the selection goal in the most reliable way. These results are consistent with the conclusions of Godshalk and Timothy (1988) that compared three multivariate analysis methods (principal factor analysis, maximum-likelihood factor analysis, and principal component analysis). Godshalk and Timothy (1988) reported that the principal component analysis had the highest correlation with index selection. The high correlation for principal component with index selection indicates the potential for using this statistical method for selection purposes.

CONCLUSION

The results documented that the reliabilities of selection indexes were estimated by tested methods between 0.41 and 0.53. Variance of selection index in the maintenance of selection goal was chosen for a reduction in the described traits. By this method, the highest value of reliability with the selection goal was estimated.

REFERENCES

Arnason T. (1984): Genetic studies on conformation and performance of Icelandic toelter horses. 3. Study on covariance matrices and breeding objectives by principal component analysis. Acta Agriculturae Scandinavica, 34, 440–449.

Baban M., Rastija T., Caput P., Knezvc I., Stipic N. (1998): Estimation of heritability of Lipizzaner horses for morphological traits by means of various methods. Czech Journal of Animal Science, 43, 299–303.

Bokor A., Blouin C., Langlois B. (2007): Possibility of selecting racehorses on jumping ability based on their steeplechase race results in France, the United Kingdom and Ireland. Journal of Animal Breeding and Genetics, 124, 124–132.

Bouška J., Štípková M., Pytloun P., Pytloun J., Kubešová M. (2008): Relationships among body condition score, milk yield and sires' breeding value for beef production efficiency in Czech Fleckvieh cattle. Czech Journal of Animal Science, 53, 453–461.

Brotherstone S. (1994): Genetic and phenotypic correlations between linear type traits and production traits in Holstein-Friesean dairy cattle. Animal Production, 59, 183–187.

Flak P. (1997): The use of bending methods for construction selection indices. In: Proc. 43rd Biometric Colloquium, 18th–21st March, Munich, Germany, 35.

Godshalk E.B., Timithy D.H. (1988): Factor and principal component analyses as alternative to index section. Theoretical and Applied Genetics, 76, 352–360.

Hardy G.H., Littlewood J.E., Pólya, G. (1952): Further Remarks on Method: The Inequality of Schwarz. §6.5 in Inequalities. 2nd Ed. Cambridge University Press, Cambridge, UK, 132–134.

Huang H., Harding J. (1998): Quantitative analysis of correlations among flower traits in gerbera. Theoretical and Applied Genetics, 97, 312–322.

Jakubec V., Schlote W., Jelinek J., Scholz A., Zalis N. (1999): Linear type trait analysis in the genetic resource of the Old Kladrub Horse. Archiv für Tierzucht, 42, 215–224.

Jakubec V., Rejfková M., Volenec J., Majzlík I., Vostrý L. (2007): Linear type trait analysis in the varieties and studs of the Old Kladrub horse. Czech Journal of Animal Science, 52, 299–307.

Koenen E.P.C., Van Veldhuizen A.E., Brascamp E.W. (1994): Genetic parameters of linear scored conformation traits and their relation with dressage and showjumping in the Dutch Warmblood Riding Horse population. Livestock Production Science, 43, 85–94.

Molina A., Valera M., Dos Santos R., Rodero A. (1999): Genetic parameters of morphofunctional traits in Andalusian horse. Livestock Production Science, 60, 295–303.

Park J.E., Lee J.R., Oh S., Lee J.W., Oh H.S., Kim H. (2011): Principal components analysis applied to genetic evaluation of racing performance of Thoroughbred race horses in Korea. Livestock Science, 135, 293–299.

Pretourius S.M., van Marle-Köster E., Mostrt B.E. (2004): Description of the Friesian Horse population of South Africa and Namibia. South Africa Journal of Animal Science, 34, 149–157.

Samore A.B., Pagnacco G., Miglior F. (1997): Genetic parameters and breeding values for linear type traits in the Hafliger horse. Livestock Production Science, 52, 105–111.

SAS (2004): SAS/STAT® 9.1 User's Guide. SAS Institute Inc., Cary, USA.

Sokolović D., Lugić Z., Radović J., Živanović T., Babić S., Simić A., Maletić R. (2011): Evaluation of morphological traits, dry matter yield and quality of *Lolium perenne* L. autochthonous populations from Serbia through multivariate analysis. Genetika, 43,129–140.

Veselá Z., Přibyl J., Šafus P., Vostrý L., Šeba K., Štolc L. (2005): Breeding value for type traits in beef cattle in the Czech Republic. Czech Journal of Animal Science, 50, 385–393.

Vostrý L., Přibyl J., Mach K., Majzlík I. (2011a): Genetic parameters estimation and breeding values prediction for linear described traits in the Old Kladruber horse. Journal of Animal and Feed Science, 20, 338–349.

Vostrý L., Čapková Z., Přibyl J., Mach K. (2011b): Analysis of Czech cold-blooded horses: genetic parameters, breeding value and the influence of inbreeding depression on linear description of conformation and type characters. Czech Journal of Animal Science, 56, 217–230.

Zechner P., Zohman F., Sölkner J., Bodi I., Habe F., Marti E., Brem G. (2001): Morphologic description of the Lipizzan horse population. Livestock Production Science, 69, 163–177.

Received: 2011–09–26
Accepted after corrections: 2011–12–02

Corresponding Author

Ing. Luboš Vostrý, Ph.D., Czech University of Life Sciences Prague, Faculty of Agrobiology,
Department of Food and Nature Resources, Kamýcká 129, 165 21 Prague 6-Suchdol, Czech Republic
Tel. +420 224 382 547, e-mail: vostry@af.czu.cz