

ASSESSMENT OF *LATHYRUS* SPECIES ACCESSION VARIABILITY USING VISUAL AND STATISTICAL METHODS

BOGNA ZAWIEJA¹, WOJCIECH RYBIŃSKI², KAMILA NOWOSAD³ AND JAN BOCIANOWSKI^{1*}

¹Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland

²Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland, Strzeszyńska 34, 60-479 Poznań, Poland

³Department of Genetics, Plant Breeding and Seed Production, Wrocław University of Environmental and Life Sciences, Grunwaldzki 24A, 53-363 Wrocław, Poland

Introduction

Lathyrus species exhibit a considerable potential for crop rotation, improving soil physical conditions, reducing disease and weed infestation, thus reducing production costs (Vaz Patto *et al.*, 2006). Moreover, their seeds are characterized by high protein contents. The genus *Lathyrus* comprises approximately 200 species and subspecies found both in the Old and the New World.

Only species grass pea (*Lathyrus sativus* L.) has been grown for food. Some lesser-known species are cultivated both for animal feed and food, for example *L. cicera*, *L. clymenum*, *L. ochrus* (for grain, but mainly for forage production) and *L. tingitanus* as forage species. Other species, especially *L. hirsutus* and *L. aphaca*, have been only occasionally grown for human consumption as well as other purposes.

In this presentation variability of accessions of seven *Lathyrus* species was investigated as regards five traits (pod length, pod width, number of seeds in the pod, weight of seeds from the pod and 100 seeds weight).

Data

The research material consisted of seven *Lathyrus* collection species: *L. sativus* (S), *L. cicera* (C), *L. clymenum* (CL), *L. ochrus* (O), *L. tingitanus* (T), *L. hirsutus* (HS) and *L. aphaca* (AP). The seeds were obtained from the Gene Bank in Gatersleben (Germany), except for *L. sativus*, which was derived from the collection of grain legumes at the Institute of Plant Genetics, the Polish Academy of Sciences in Poznań. The field trial was established in the random block design with fifteen replications. In the phase of full maturity 3 pods from the main stem of randomly selected plants were harvested from each plot. Species variability was analyzed in terms of five traits: Pod Length (cm), Pod Width (cm), the Number of Seeds per Pod, Weight of Seeds per Pod (g) and 100 Seeds Weight (g).

Methods

- Coefficient of variation (CV)** – a simple and comprehensive absolute measure of variability for traits in the researched population (Zawieja & Pilarczyk, 2005).
- Visual statistical methods** – Modified Andrews curves (Khattri & Naik 2002). The Linear Discriminant Analysis (Friedman, 1989) and Nonlinear Kernel Discriminant Analysis (Mika *et al.*, 1999). Two kernel functions were used: sigmoid $k(\mathbf{x}, \mathbf{y}) = \tanh(\mathbf{a}\mathbf{x}'\mathbf{y} + c)$ and Chi-Square $k(\mathbf{x}, \mathbf{y}) = 1 - \sum_{i=1}^m \frac{(x_i - y_i)^2}{[x_i + y_i]}$, where \mathbf{x} and \mathbf{y} denote vectors of observations; a, c – constants (they are selected so that the kernel matrix was positive definite), n – the number of observations. In this paper the data were normalized before analysis, thus the parameters $a = 1$ and $c = 0$.
- Test of hypothesis on equality of coefficients of variation** – Miller method (Miller & Feltz, 1997).

Results

The **variation coefficients** min.: 0% max.: 135.5% (Fig. 1). For pod length (LP) and pod width (WP) the CVs were smaller (from 0.5% to 16%). In turn, a large range of CVs was shown for the AP, C, CL and HS accessions as regards PL. Similarly, in the case of pod width for accession CL a large range of CVs was recorded. The CVs for HS was the most varied for seed traits.

Andrews curves (Fig. 2A) for accessions. Curves for HS8 and HS4 differed markedly from the other curves. It may have meant that variability of these accessions was greater than that of others (Fig. 1). **Andrews curves as a function of traits (Fig. 2).** Among them the curve for WSP was exceptional, while the curves for NSP and 100SW were the most similar (Fig. 2). The curves representing PL and PW are also very similar. **Andrews curves for traits of each tested species separately (Fig. 3).** Curves representing pod length and pod width in C, HS, O, S and T were similar. For O, S and T the curves for WSP, NSP and 100SW were similar. The trait of WSP was highlighted in accessions HS and C. **Andrews curves for accessions in terms of traits (Fig. 4).** The curves for HS were exceptional for all traits, but mainly as regards seed traits. Besides, CL was singled out in terms of PL and O as regards PW, NSP and 100SW.

LDA method gave no satisfactory results. The points representing the accessions of all species were quite mixed. On the basis of the Wilks' lambda statistics it was concluded that all discriminant values are equal to zero ($p > 0.05$ for all discriminant variables). **The NKDA method allowed to separate groups.** Some of the species were separated. The application of the kernel sigmoid resulted in a division into five groups (Figs. 5 and 6). Two species were classified to the first group from the left: T and CL. Only AP was classified to the second group. Two species were included in the third group: C and S. Finally, O and HS produced separate groups. Four groups could be visualized when the Chi-square kernel was applied. Separate groups were formed by CL and HS, the second group comprised C, AP and S, whereas the last group contained T and two accessions of O. A different curvature of the space resulted in the separation of T and CL..

The hypothesis on the equality of the CVs. The accessions for which the CV exceeded 33%, were not taken into account. Most accessions of the tested species were uniform to a similar degree (Table 1).

Table 1. p-value of the Miller test

Species	Trait				
	Pod length	Pod width	Seeds no. from the pod	Seeds weight from the pod	100 seeds weight
<i>Lathyrus aphaca</i>	0.1194	0.2545	0.7894	0.8797	0.0280
<i>Lathyrus cicera</i>	0.0562	0.4814	0.6174	0.4987	0.9446
<i>Lathyrus clymenum</i>	0.3382	0.2064	0.4432	0.4931	0.0402
<i>Lathyrus hirsutus</i>	0.0361	0.3053	0.0961	0.4736	0.0397
<i>Lathyrus ochrus</i>	0.8411	0.8499	0.8266	0.8126	0.4241
<i>Lathyrus sativus</i>	0.9855	0.0347	0.7109	0.5358	0.6180
<i>Lathyrus tingitanus</i>	0.4150	0.6494	0.3710	0.8425	0.2646

Conclusions

- Variability of accessions is very small in terms of pod traits.
- In the case of PL the accessions of O, S and T were characterized by low variability, while the accession of HS was characterized by the highest variability. No uniformity was shown for HS by the Miller test.
- For PW the accessions of C, S and T were characterized by low variability, while the accession of CL was characterized by the highest variability (see shape of Andrews curves, Fig. 4). However, a lack of uniformity was detected only for S (Tab. 1).
- For the seed traits the accession of HS was characterized by the highest variability.
- Pod traits and seed traits were clearly different (Fig. 3) in most of the species. Moreover, it was shown that WSP was distinguished in the case of all species (Fig. 2) it was especially visible for AP and HS.
- Despite the observed large differences between CVs of seed traits, no significant differences were found for the NSP and WSP using the Miller test; only in the case of 100SW the CVs differed significantly for AP, CL and HS.
- The application of the NKDA method demonstrated differences in variability between the accessions of HS, T as well as some accessions of O.
- The NKDA method showed that variability in accessions S, C and AP was comparable.
- The significant differences among genotypes of AP, CL and HS indicate potential to obtain new cultivars differing in terms of 100SW from established genotypes.

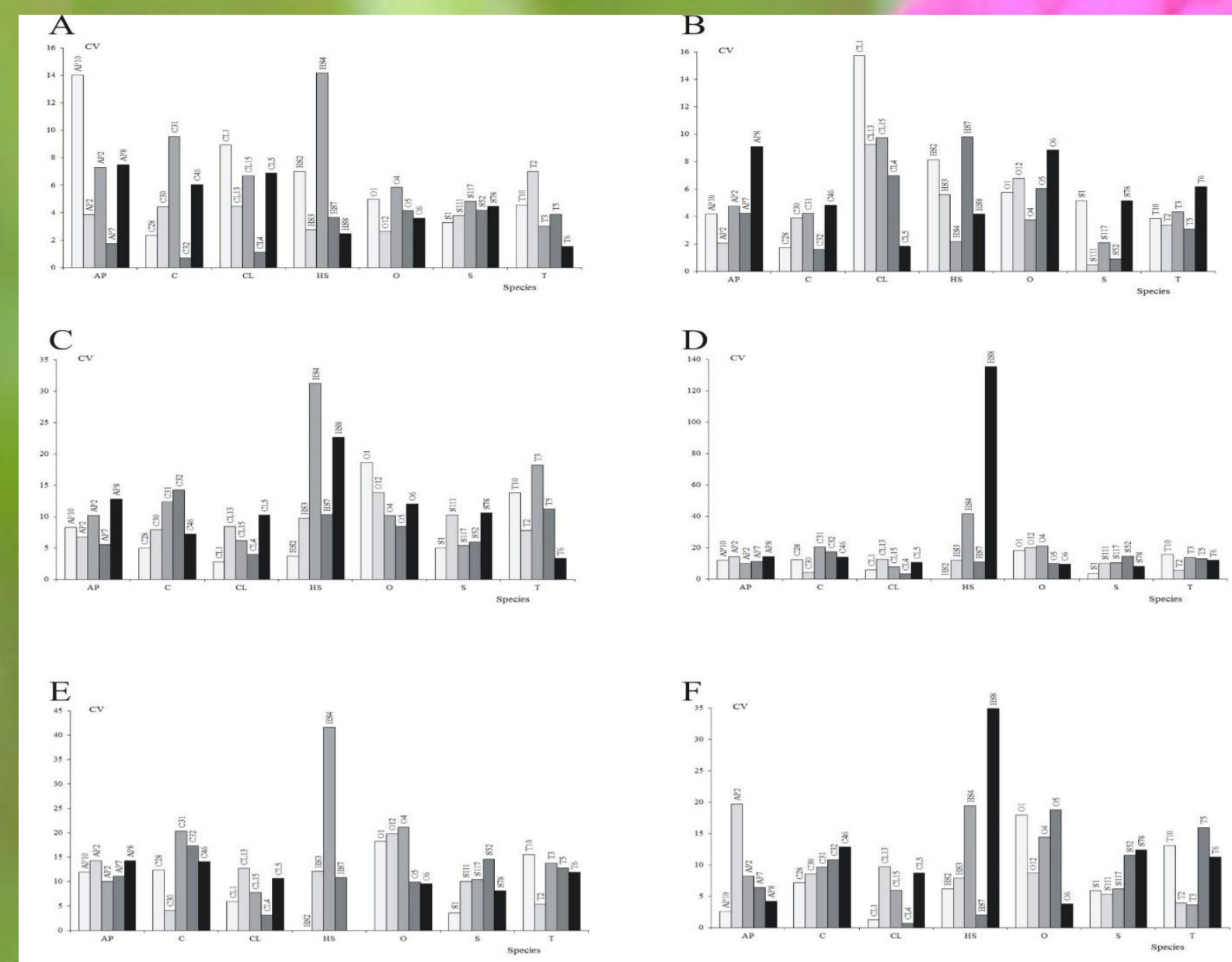


Fig. 1. Coefficients of variation for all tested objects (S – *L. sativus*, C – *L. cicera*, CL – *L. clymenum*, O – *L. ochrus*, T – *L. tingitanus*, HS – *L. hirsutus*, AP – *L. aphaca*): A – pod length,

B – pod width, C – number of seeds in the pod, D – weight of seeds from the pod, E – weight of seeds from the pod except for genotype HS8, F – 100 seeds weight

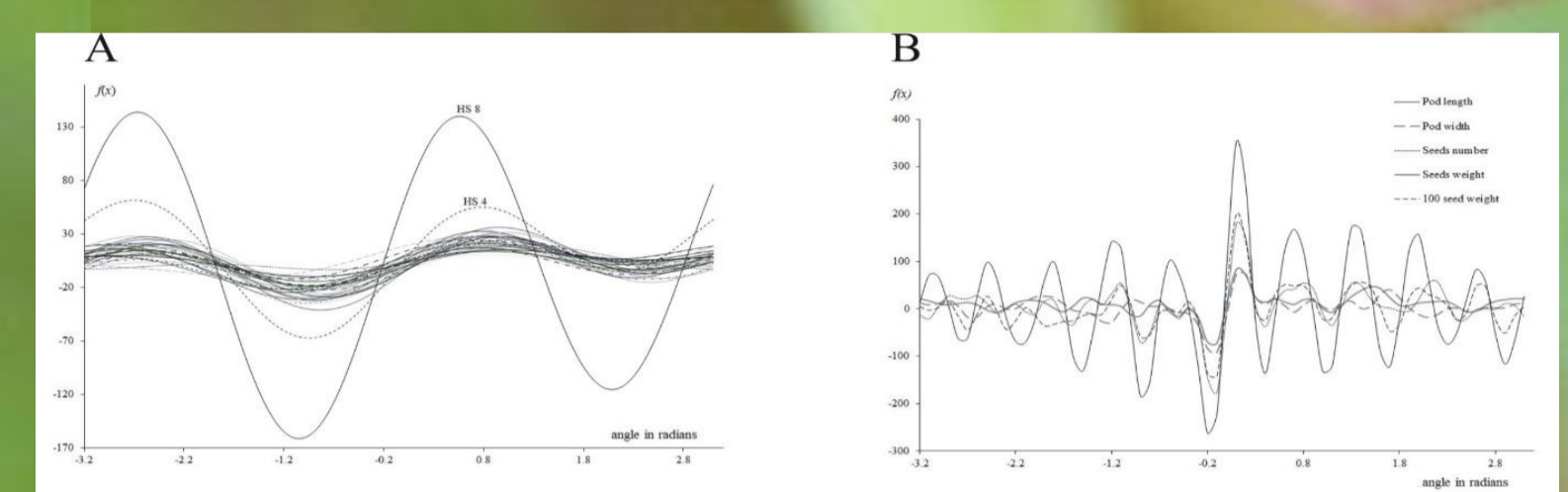


Fig. 2. Andrews curves: A – of genotypes, B – of traits for each species separately.

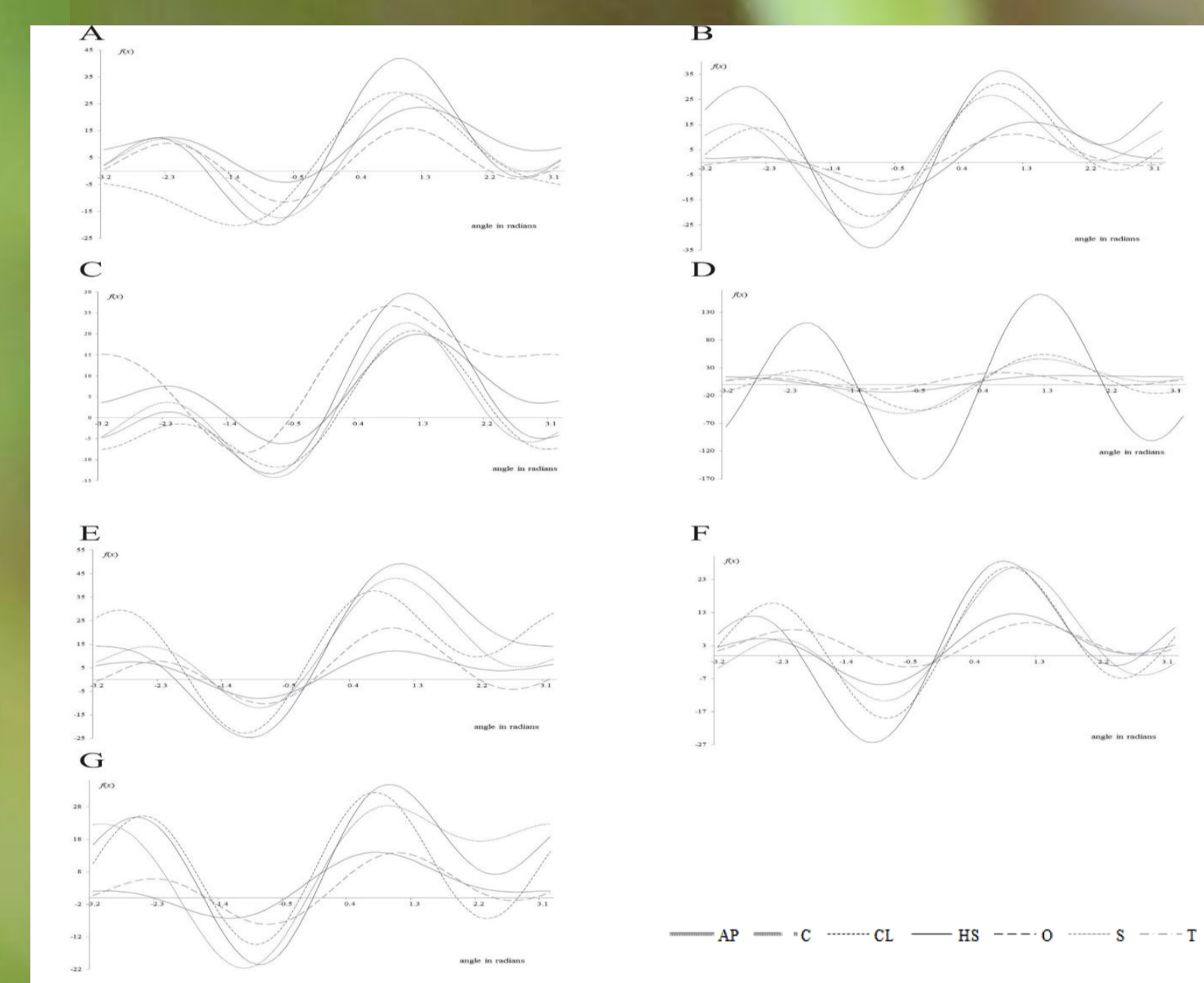


Fig. 3. Andrews curves of traits for each species separately (black solid line – seed weight; black dashed line – 100 pods weight; black dotted line – seed number in the pod; grey solid line – pod length; grey dashed line – pod width): A – *L. aphaca*, B – *L. cicera*, C – *L. clymenum*, D – *L. hirsutus*, E – *L. ochrus*, F – *L. sativus*, G – *L. tingitanus*.

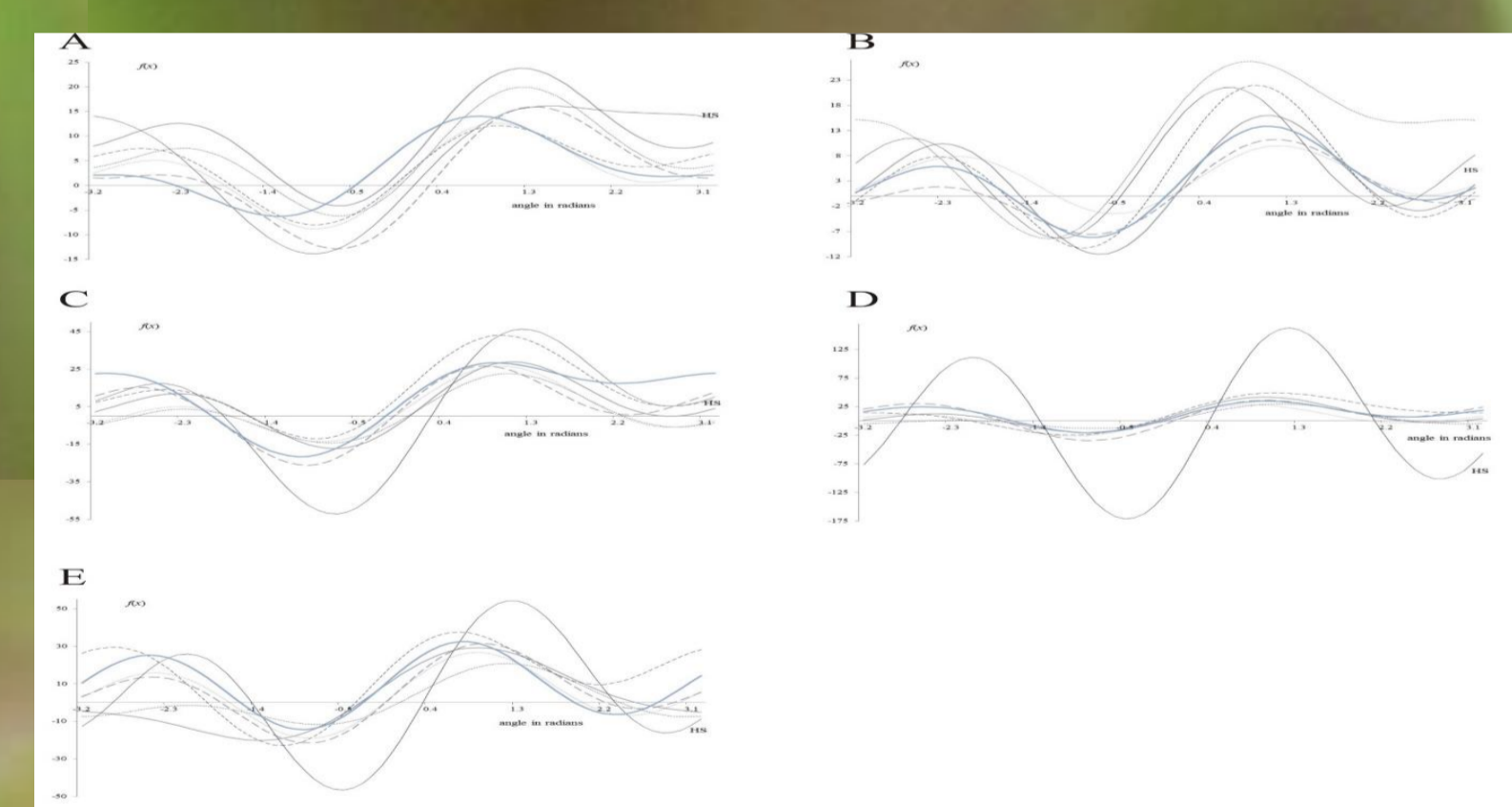


Fig. 4. Andrews curves for species (A – pod length, B – pod width, C – seed number from the pod, D – seed weight, E – 100 seeds weight)

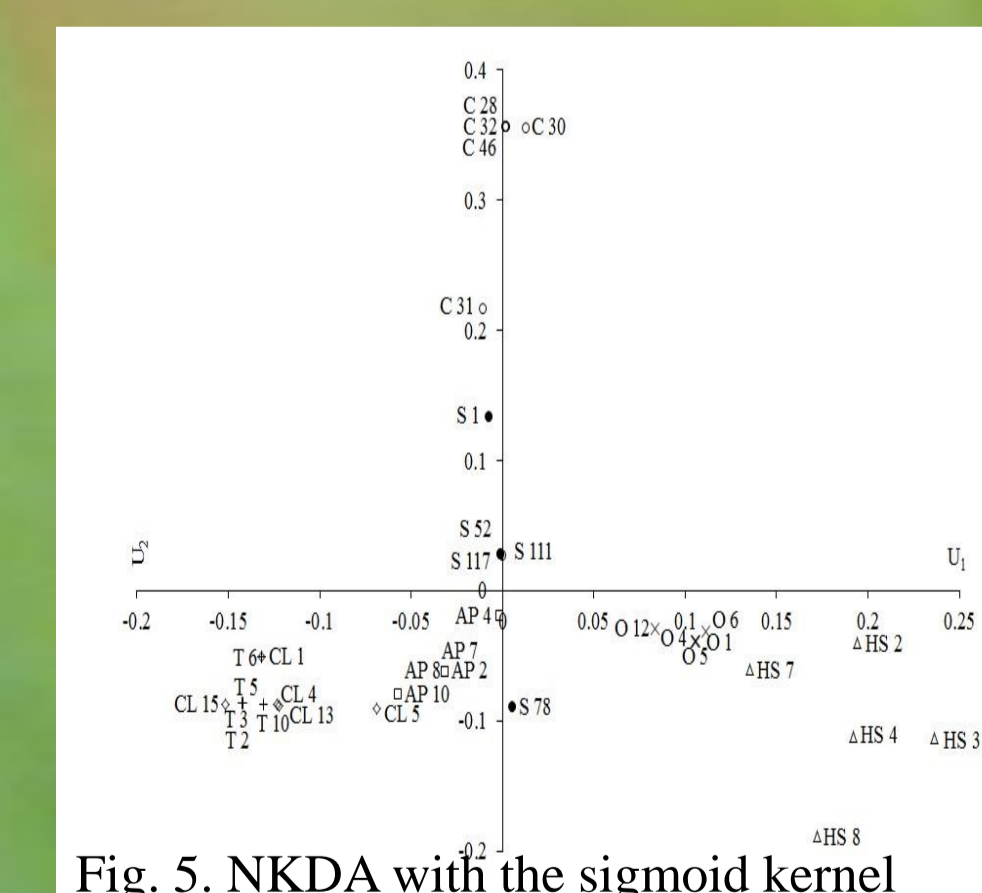


Fig. 5. NKDA with the sigmoid kernel

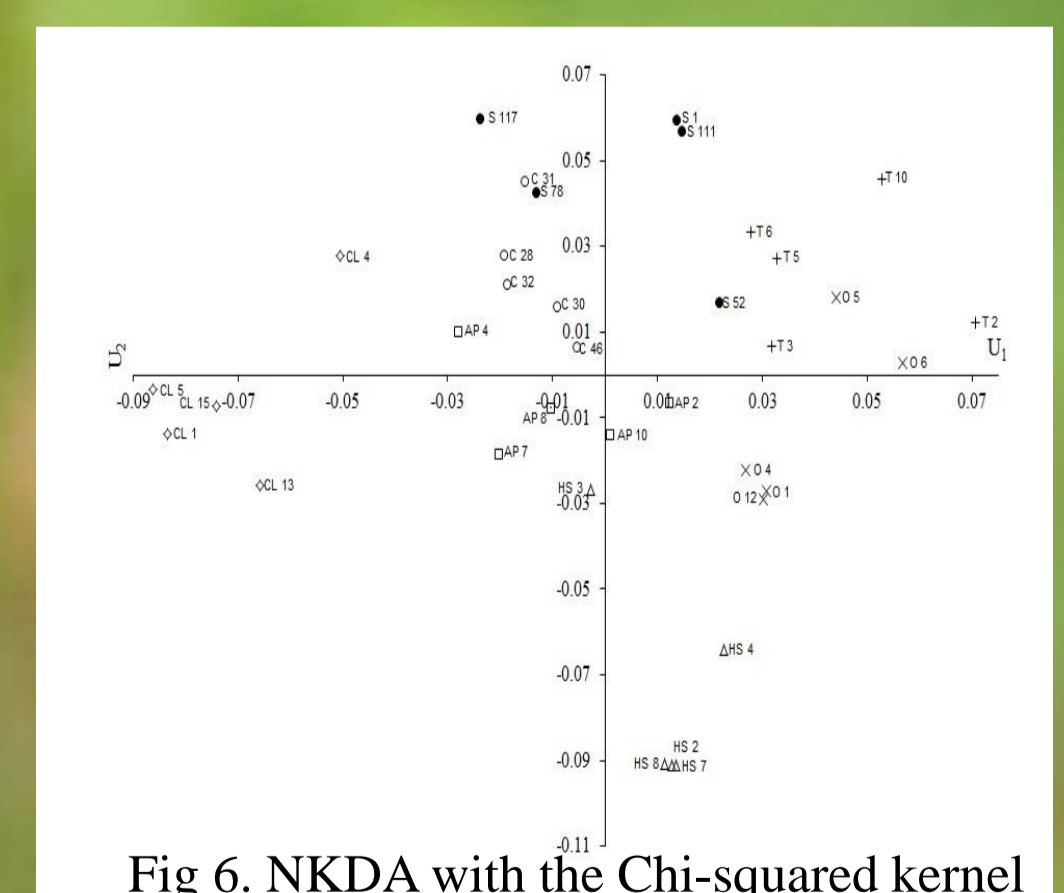


Fig. 6. NKDA with the Chi-squared kernel

Generally

Uniformity of genotypes may be pre-tested using Andrews curves. Genotypes, which curves have a clearly greater amplitude, are characterized by high CV values. The lines of a similar amplitude and waveform indicate the genotypes with similar variability for all the analyzed traits. A similar pre-analysis may be performed in terms of traits, variability of genotypes is the largest in terms of the trait of the largest amplitude. The proposed visual methods made it possible to assess variability of the objects in terms of all analyzed traits and immediately exclude non-uniform cases as well as identify those that were uniform to a similar extent.

References

- Vaz Patto, M.C., B. Skiba, E.C.K. Pang, S.J. Ochatt, F. Lambein and D. Rubiales. 2006. *Lathyrus* improvement for resistance against biotic and abiotic stresses: From classical breeding to marker assisted selection. *Euphytica*, 147: 133-147.
- Zawieja, B. and W. Pilarczyk. 2005. The comparison of traditional UPOV uniformity criterion and new approach based on Bennett's test for coefficients of variation. *Colloq. Biometr.*, 35: 155-163.
- Khattri, R. and D.N. Naik. 2002. Association in contingency tables, correspondence analysis and (modified) Andrews plots. In: Huber-Carol, C., N. Balakrishnan, M.S. Nikulin and M. Mesbah (Eds.), *Goodness-of-fit Tests and Model Validity*, Birkhäuser, Boston, pp. 311-326.
- Mika, S., G. Rätsch, J. Weston, B. Schölkopf and K.R. Müller. 1999. Fisher discriminant analysis with kernels. In: Hu, Y.H., J. Larsen, E. Wilson and S. Douglas (Eds.), *Neural Networks for Signal Processing*, IX: 41-48.
- Miller, G.E. and C.J. Feltz. 1997. Asymptotic inference for coefficients of variation. *Commun. Stat. Theory*, 26: 715-726.