JMBFS

Journal of Microbiology, Biotechnology and Food Sciences

International peer-reviewed scientific online journal

Published by Faculty of Biotechnology and Food Sciences

# UTILIZATION OF CDDP MARKERS IN ANALYSIS OF GENETIC VARIABILITY OF *ARACHIS HYPOGAEA* L.

*Lucia Klongová[1], Matúš Kysel'[1], Veronika Fialková[1], Eloy Fernández Cusimamani[2], Adam Kováčik[3], Jana Žiarovská*[3]*

***Address(es):***
[1] Slovak University of Agriculture in Nitra, Research Centre of AgroBioTech; Tr. A. Hlinku 2, 94976, Nitra, Slovak Republic.
[2] Czech University of Life Sciences Prague, Faculty of Tropical AgriSciences, Department of Crop Sciences and Agroforestry, Kamýcká129, 165 21, Prague 6, Czech Republic.
[3] Slovak University of Agriculture in Nitra, Faculty of Agrobiology and Food Resources, Institute of Plant and Environmental Sciences; Tr. A. Hlinku 2, 94976, Nitra, Slovak Republic.

*Corresponding author: jana.ziarovska@uniag.sk

https://doi.org/10.55251/jmbfs.9884

ARTICLE INFO

Received 10. 2. 2023
Revised 16. 5. 2023
Accepted 16. 5. 2023
Published 1. 8. 2023

Regular article

OPEN ACCESS

**ABSTRACT**

*Arachis hypogaea* L. – peanut (syn. groundnut) belong botanically to annual herbs, more specifically to legumes (family *Fabaceae*). Cultivated from is characterized as allotetraploid specie and that has a large and genome with high level of complexity as a result of natural hybridization of wild diploid species *Arachis cardenasii* (nn) and *Arachis batizocoi*. Individual analysis that are aimed to description of the genetic similarity and variability of wild and cultivated *Arachis* species are substantial for the knowledge of intaspecific relationships. In this study, conserved DNA-derived polymorphism technique was used to analyse genetic variability of natural genotypes of peanut. CDDP is a simple PCR based method that was proved to be efficient in the studies of plant populations polymorphism. Primers used in CDDP target conserved sequences of plant functional genes that are involved in response to abiotic and biotic stresses. Here, bulked DNA samples of twenty-one natural populations were analysed by with five CDDP primer combinations. Selected primers produced a total of 260 amplicons, among which 107 (41.2%) amplicons were polymorphic. The average number of obtained amplicons per primer was about 51.8. Amplified polymorphic fragments of analysed genotypes ranged from 13 to 33 and percentage of polymorphism ranged from 35 % to 47 %. UPGMA cluster analysis was performed on the base of prepared binary matrices and obtained clustering has grouped the analyzed peanut populations into five major groups with further subgroups. The results of CCDP fingerprinting shown, that this marker technique provide sufficient distinguishing of a wild population of *Arachis hypogaea* L.. Analysed genotypes of peanut were separated into different groups based on genetic diversity, however, the level of polymorphisms that detected used markers among populations of *Arachis hypogaea* L. was low.

**Keywords:** *Arachis hypogaea* L., CDDP fingerprints, genetic polymorphism, structure analysis

## INTRODUCTION

*Arachis hypogea* L. commonly named as the peanut is one of the most popular oilseeds and important legume crops for human consumption., it which grown in the subtropical, tropical and warm temperate areas of Asia, Africa, and the Americas **(Govindarajan *et al.*, 2014)**. According to all available information, peanuts come from the region of Bolivia and Argentina **(Bertioli *et al.*, 2011)**. *A. hypogaea* is a cultivated peanut, that was probably a result of hybridizing between *A. cardenasii* (nn) and *A. batizocoi* **(Krapovickas & Gregory, 2007)**, as both parents grew nearby in Bolivia. This cultivated peanut is an allotetraploid (2n=4X=40) plant species with a large and complex genome belonging to the family *Fabaceae* **(Sanders, 2003)**. Peanuts are a rich source of nutrition with a high content of oils (50%), proteins (25%), carbohydrates (20%), fibre (5%), vitamin E, calcium, phosphorus, magnesium, zinc and several other micronutrients and minerals **(Settaluri *et al.*, 2012; Toomer, 2018)**, and have the potential to be an economically beneficial food supplement **(Variath & Janila, 2017)**. The largest producer of peanuts is China, with an annual production of 17.77 million tons, which is almost half of the world's production. India is the second largest producer of peanuts and produces 6.3 million tons annually. Other important peanut producers are Nigeria, Senegal, Chad, the USA, Argentina, etc. **(*Peanut Production by Country 2023*, n.d.)**. There are studies focused mainly on the morphological and agronomic characteristics of *A. hypogaea*, and their resistance to abiotic and biotic stresses or oleic acid content **(REN *et al.*, 2011; Upadhyaya *et al.*, n.d., 2006)**.

Peanuts are characterized by significant morphological variations, but the level of genetic variability is low. The understanding of genetic diversity is an important step for their efficient use in the peanuts breeding program **(Gantait *et al.*, 2019)**. Different DNA marker systems have been used to study genetic resources, phylogenetic relationships, and analysis of peanut genetic diversity. The lack of polymorphism at the genetic level was noticed using randomly amplified polymorphic DNA (RAPD) **(Dwivedi *et al.*, 2001; Raina *et al.*, 2001; Samaha *et al.*, 2019)**, amplified fragment length polymorphisms (AFLP) **(He & Prakash, 2001; Herselman, 2003)**, restriction fragment length polymorphisms (RFLP) **(Kochert *et al.*, 1991)**, specific SSR markers (**Cuc *et al.*, 2008; Ren *et al.*, 2014)**. The PBA and iPBS marker technique for *A. hypogaea* L. pro-vided a higher level of polymorphism, ranging from 60 % to 78 % **(Montero-Torres *et al.*, 2020; Klongová *et al.*, 2021)**.

CDDP molecular markers are an efficient technique based on genome-wide amplification of functional genes of conserved DNA sequences **(Collard & Mackill, 2009)**. CDDP marker system has been used in several plant species such as chickpea **(Hajibarat *et al.*, 2015)**, rice **(Collard & Mackill, 2009)** , date palm **(Atia *et al.*, 2017)**, *Chrysanthemum* **(Li *et al.*, 2014)**, *Paeonia suffruticosa* **(Wang *et al.*, 2014)**, *Rosa rugose* **(Jiang & Zang, 2018)**, wheat **(Kysel' *et al.*, 2021)** and apple **(Bilčíková *et al.*, 2021)** and these studies were demonstrated their utility, efficiency and reproducibility of genetic diversity estimation. Unlike the RAPD marker method, the CDDP method has better reproducibility because longer primers and higher annealing temperatures are used for amplification **(Collard & Mackill, 2009)**.

Principal coordinate analysis (PCoA) is standardly applied in ecological studies of peanuts or studies extending into geographic relationships between peanut accessions **(Abdella *et al.*, 2022; Stalker *et al.*, 1995; Zou *et al.*, 2020)**. PCoA plots show the distribution of objects in 2D space because of the correlation matrix X with geographic distances and Y with genetic distances. If the dendrograms present visibly distinguishable cluster of genotypes, PCoA should be able to capture them as populations. The most common mistake that researchers make is the use of a dissimilarity matrix, which is obtained by transforming the original similarity matrix. In such a case, autocorrelation occurs, but it also requires an association between identifiable objects **(Lagendre and Foltin, 2010; Lagendre *et al.*, 2015)**. PCoA are often performed together with the analysis of molecular variance (AMOVA) and the correlation through Mantel test. They search the genetic structure of a set of objects based on their geographic structure. AMOVA indicates whether the degree of genetic diversity in populations and subpopulations of objects or between them corresponds to geographic diversity. Geographic distance assumes genetic distance in object populations **(Bontrager & Angert, 2018)**. The Mantel test verifies correlational relationships between objects, which, however, should not associate with each other. The essence and correctness of

using the Mantel test are explained by **Legendre and Foltin (2010)** and **Legendre *et al.* (2015)**.

In this study, the genetic diversity of natural genetic resources of *A. hypogaea* L. from 21 regions in Bolivia was analyzed using five gene targeted CDDP markers. To our knowledge, this marker technique has not yet been applied to the *A. hypogaea* L. population.

## MATERIAL AND METHODS

### Plant material

The biological material was collected as an original plant source in Bolivia at a total of 21 locations marked with numbers 1 – 20 (Figure 1, Table 1). The seeds were transported to the Faculty of Tropical AgriSciences, Czech University of Life Sciences in Prague; Czech Republic where they were planted in pots. Young plants were transported to the Research Centre Agrobiotech, Slovak University of Agriculture in Nitra, Slovakia.

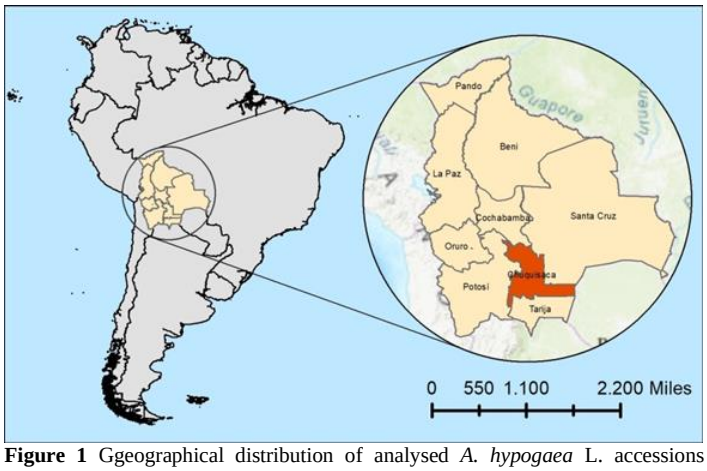


**Figure 1** Ggeographical distribution of analysed *A. hypogaea* L. accessions collected from Chuquisaca, Bolivia

**Table 1** Accessions of *A. hypogaea* L. analysed in this study

| No. | Accession | No. | Accession |
|---|---|---|---|
| 1 | CMC-003 | 12 | CMC-006 |
| 2 | CMC-008 | 13 | CMC-019 |
| 3 | CMC-004 | 14 | CMC-015 |
| 4 | CMC-005 | 15 | CMC-002 |
| 5 | CMC-016 | 16 | CMC-009 |
| 6 | CMC-018 | 17 | CMC-017 |
| 7 | CMC-010 | 18 | T1 |
| 8 | JJV-002 | 19 | T5AN1890 |
| 9 | CMC-012 | 20 | CMC-013 |
| 10 | CMC-014 | 21 | Pico loro |
| 11 | CMC-001 | - | - |

### DNA extraction

Bulked gDNA was extracted from five leaves of individual plants using the GeneJET Plant Genomic DNA Purification Mini Kit (ThermoScientific) according to the manufacturer's instructions. DNA was quantified with a NanoPhotometerTM (IMPLEN) spec-trophotometer.

### CDDP markers

The genetic diversity among studied *A. hypogaea* L. samples was evaluated using the CDDP technique. Markers selected for analysis are gene specific and matched the conserved regions of DNA. The primers are designed to target well-characterized genes, in this case, were used for PCR amplification of functional regions of the WRKY gene. In our study, five WRKY CDDP primers as reported by Collard and Mackill **(Collard & Mackill, 2009)** were used. Table 2 summarises the CDDP marker technique. PCR mixtures contained 2X EliZyme HS Robust MIX (MasterMix), 400 nM of each primer, and 10 ng of template DNA. The final volume was 10 µl. Amplification was performed using thermal cycler TProfessional Basic gradient XL (Biometra). Following time and temperature profile was used for PCR: denaturation at 95 °C for 5 min; 35 cycles consisting of 95 °C for 45 s, 54 °C for 45 s, 72 °C for 90 s; and the final extension was held at 72 °C for 10 min. Amplicons were separated on 3% agarose gels prepared with 1× TBE buffer and amplicons were stained with GelRed® Nucleic Acid Gel Stain (Biotium). The PCR fragments were visualized under UV light using a BDAdigital system 30 (Analytik Jena).

### Data analysis

The software GelAnalyzer 1.0 was used for image analysis of electrophoretically separated fragments and then transformed into a binary matrix. The scorable visualized fragments on gels were classified according to size as present (1) or absent (0). The binary vector (0/1) was generated from this data. Each genotype represented by a binary vector was deprived of all biological values by transformation and is now represented by a binary profile. Biological values were proxy with statistical analyses for dendrogram construction, ANOVA, correlation by multiple Mantel Test, descriptive statistics, and polymorphism calculation. The similarity and dissimilarity was calculated based on the Jaccard coefficient and the dendrograms were constructed using the unweighted pair group method with arithmetic mean (UPGMA) method using the online accessible software ***(DendroUPGMA: Dendrogram construction using the UPGMA algorithm*, n.d.)**. Another statistic software GenAlex (a statistical tool for MS Excel interface) was used to transform an original as a binary set providing converted data especially default for statistical tools of PCoA (Principal coordinates analysis).

The basic Mantel statistic is characterized as the sum of the products of the corresponding elements of the matrices $Z = \sum_{ij} X_{ij} Y_{ij}$, (Σ_ij the double sum over all i; and all j where i ≠ j). Z can take on any value depending on the exact nature of X and Y, that is why one usually uses a normalized Mantel coefficient, calculated as the correlation between the pairwise elements of X and Y.

The total number of amplified bands (TNB), the number of polymorphic bands (NPB), the percentage of polymorphic bands (PPB), the polymorphism information content (PIC), the marker index (MI) and the resolving power (Rp) were calculated to evaluate the efficiency of each used CDDP primer. The polymorphism information content (PIC) was calculated by using the $\mathrm{PIC} = 1 - [\mathrm{f}^2 + (1 - \mathrm{f})^2]$ formula, where f is the marker frequency in the data set **(De Riek *et al.*, 2001; Roldán-Ruiz *et al.*, 2000)**. The distinguishing ability of the primers among analysed 21 samples was assessed according to their value of resolving power (Rp), as $Rp = \sum Ib$, where Ib is band informativeness, $Ib = 1 - [2x|0{,}5 - f|]$ **(Prevost & Wilkinson, 1999)**.

**Table 2** Information about CDDP primers used to detect genetic variability in natural genetic resources of *Arachis hypogaea* L.

| Gene | Gene function | Primer name | Primer sequence (5′ to 3′) | Length | GC (%) | Reference |
|---|---|---|---|---|---|---|
| WRKY | Transcription factors for developmental and physiological roles | WRKY- F1 | TGGCGSAAGTA CGGCCAG | 18 | 67 | **(Xie *et al.*, 2005)** |
| | | WRKY- R1 | GTGGTTGTGCT TGCC | 15 | 60 | |
| | | WRKY- R2 | GCCCTCGTASG TSGT | 15 | 67 | |
| | | WRKY- R3 | GCASGTGTGCT CGCC | 15 | 73 | |
| | | WRKY- R2B | TGSTGSATGCTC CCG | 15 | 67 | |
| | | WRKY- R3B | CCGCTCGTGTG SACG | 15 | 73 | |

## RESULTS

### CDDP markers – efficiency and polymorphism

CDDP fingerprints have been extensively used to analyse genetic diversity within plant species. It is a simple, efficient, and reproducible method to study polymorphism.

Here, five CDDP primer combinations were used to analyze genetic diversity in 21 accessions of *A. hypogaea* L. These primers amplified pure and reproduced fragments. Total number of amplified bands (TNB), number of polymorphic bands (NPB), percentage of polymorphic bands (PPB) and polymorphism information content (PIC) were calculated (Table 3) for individual CDDP primers with average values as well. 260 bands with a range of their length from 21 bp up to 2300 bp were amplified by five CDDP primers for the 21 accessions, and 107 of them were polymorphic. The total number of amplified bands was from 36 (primer WRKY-R1) to 78 (primer WRKY-R3b), with a mean of 51.8 amplicoms per primer. The number of polymorphic bands varies between 13 (primer WRKY-R3a) and 33 (primer WRKY-R3b) with an average of 21.4 bands per primer. Polymorphic information content value was from 0.23 to 0.38, with an average of 0.29, demonstrating that the CDDP primers employed in the study were efficient and polymorphic. The primer WRKY-R2a had the lowest PIC value of 0.23. Moreover, the resolving power (RP) values ranged between 7.81 (primer WRKY-R1) and 21.14 (primer WRKY-R3b). All CDDP primers shown a low level of polymorphism, with an average of 41.2 %. WRKY-R1 primer shown the highest level of polymorphism at 47 % and primer WRKY-R3a was the lowest at 35%.

**Table 3** Parameters of genetic diversity utilized in *A. hypogaea* L. accessions using 5 CDDP primers

| Marker | TNB | NPB | PPB (%) | PIC | Rp |
|---|---|---|---|---|---|
| WRKY- R1 | 36 | 17 | 47 | 0,32 | 7,81 |
| WRKY- R2a | 48 | 21 | 43,8 | 0,23 | 11,43 |
| WRKY- R3a | 37 | 13 | 35 | 0,38 | 10,00 |
| WRKY- R2b | 61 | 23 | 37,7 | 0,26 | 16,67 |
| WRKY- R3b | 78 | 33 | 42,3 | 0,26 | 21,14 |
| Mean | 51,8 | 21,4 | 41,2 | 0,29 | - |

Note: TNB - total number of amplified bands; NPB - number of polymorphic bands; PPB (%) – the percentage of polymorphic bands; PIC - polymorphism information content; Rp – resolving power

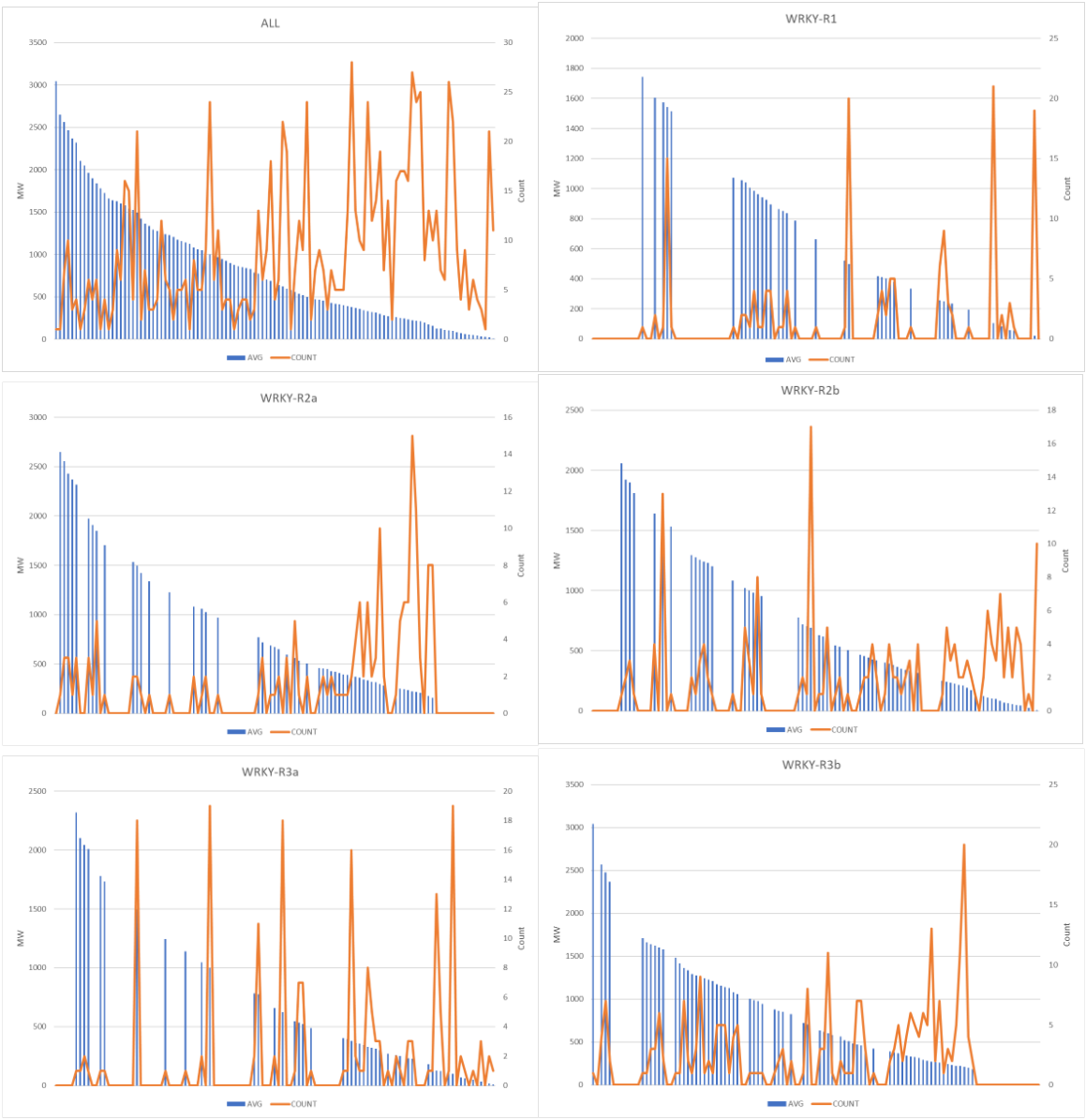


**Figure 2 (a-f)** Structure of distribution of separated amplicons

### CDDP markers – ANOVA and descriptive statistics

The analysis of the polymorphism of the binary vector profiles shown significant variability across the set of monitored genotypes of *A. hypogea* L., which, however, was not reflected in the analysis of the binary vector profiles using the ANOVA method. The differences between the profiles reached significance only in the partial set of the WRKY-R2a primer pair of the CDDP marker technique. With the help of descriptive statistics, it is possible to observe differences in the distribution of separated amplicons, which, by mutual accumulation, compensated for the potential variability at the level of polymorphism. As shown in Figure 2a-f, primer pairs WRKY-R1, WRKY-R2b, and WRKY-R3a successfully captured DNA amplicons not only along the entire length of the profile but also evenly distributed. On the other hand, the profiles of primer pairs WRKY-R2a and WRKY-R3b were significantly dominated by short amplicons, which was also reflected in the ANOVA analysis. Based on amplicon separation, the use of primer pairs WRKY-R2b and WRKY-R3a appear to be the most reproducible (Figures 2a-e). From the

point of view of other parameters of descriptive statistics, when comparing individual primer pairs, primer pairs WRKY-R2b and WRKY-R3a were strikingly similar in 4 out of 5 parameters. It was with them that the highest variabilities according to means (arithmetic, geometric, harmonic) were achieved, while according to the maximum, the first place in variability went to the primer pair WRKY-R3b. The least variable means in all 5 parameters was presented in the binary profile of primer pair WRKY-R1. To monitor the distribution of separated amplicons, the arithmetic mean of the vector profiles of individual primer pairs was chosen (WRKY-R1 – 500-600 bp; WRKY-R2a – 200-400 and 600-1000 bp; WRKY-R2b and WRKY-R3a – 400-1000 bp; and WRKY-R3b – 500-900 bp). The maximum values for the part of the vector profiles were around 500 bp. It can be stated about the primer pair WRKY-R3b was successful in detecting amplicons with a minimum of around 200 bp, like WRKY-R2a. In contrast to it, the amplicons were distributed in the range of 600-900 (geometric mean) and 500-700 bp (harmonic mean). Regarding the position of individual vector profiles of *A. hypogea* L. genotypes, it can be stated that the profiles of genotypes 4 and 19 (primer pair WRKY-R1) had the most significant influence in descriptive statistics; 1, 3, 7, 8, 10, 12, 13, 16, 17, and 19 (WRKY-R2a); 7, 8, 15, 17, and 19 (F-R2b); 5, 7, 8, 15, and 19 (WRKY-R3a); and 3, 5, 7, 8, 10, 11, 12, 14, 17, 18, 19, and 21 (WRKY-R3b). To amplicon separations, only genotype profiles 12, 16, and 19 were responsible for shorter amplicons as can be seen in Figure 3.
Each figure presents the composition of vector profiles with different primer pairs. Descriptive statistics were based on original data with biological values. Abbreviations AVG reflect arithmetic mean and COUNT reflects the number of fragments.

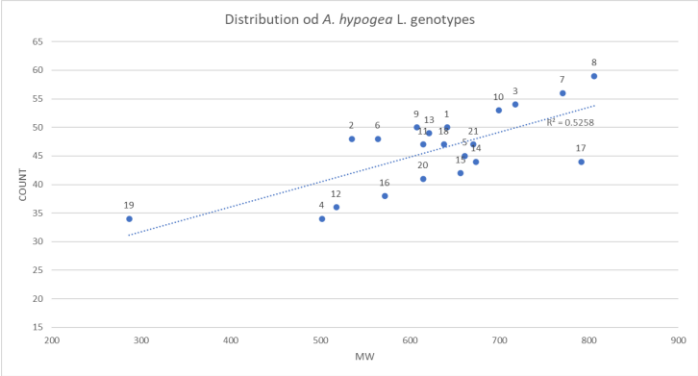


**Figure 3** Composition of monitored genotypes (*A. hypogea* L.)

**Correlation analysis of *A. hypogaea* accessions by Mantel test**

From the perspective of descriptive statistics, the distance between the vector profiles of primer pairs WRKY-R2b and WRKY-R3a is the shortest, and the vector profiles of primer pair WRKY-R1 are the most divergent. To identify these relationships, a correlation analysis was run between the detection capabilities of individual primer pairs according to Mantel. In addition to the correlation of the matrices of binary vector profiles, the Mantel test provided us with information on which of the two matrices being compared dominates. Dominance is in favour of the X-axis if the values of the X-axis are concentrated at one pole of the axis significantly more than the values of the Y-axis and vice versa. Based on the data, it is possible to present 4 groups of correlation relations (Figure 4). The X-axis dominance correlation relationship is hidden in the vector profile matrices obtained by primer pairs WRKY-R2a and WRKY-R2b against WRKY-R3a.

**Table 4** Parameters of genetic diversity utilized in *A. hypogaea*, L accessions using 5 CDDP primers

| Correlation pair | Rxy | $R^2$ | y |
|---|---|---|---|
| WRKY-R1-WRKY-R2a | -0.088 | 0.0077 | -0.0794x + 0.9036 |
| WRKY-R1-WRKY-R2b | 0.255 | 0.0653 | 0.1921x + 0.7367 |
| WRKY-R1-WRKY-R3a | 0.283 | 0.0802 | 0.3777x + 0.3924 |
| WRKY-R1-WRKY-R3b | 0.124 | 0.0153 | 0.099x + 0.7667 |
| WRKY-R2a-WRKY-R2b | 0.400 | 0.1598 | 0.3317x + 0.5775 |
| WRKY-R2a-WRKY-R3a | 0.034 | 0.0012 | 0.0502x + 0.5927 |
| WRKY-R2a-WRKY-R3b | 0.400 | 0.1603 | 0.3534x + 0.5291 |
| WRKY-R2b-WRKY-R3a | 0.380 | 0.1448 | 0.6748x + 0.0549 |
| WRKY-R2b-WRKY-R3b | 0.421 | 0.1769 | 0.4475x + 0.4454 |
| WRKY-R3a-WRKY-R3b | 0.270 | 0.073 | 0.1621x + 0.7274 |

Note: Rxy – Pearson product-moment correlation coefficient r; R2 - redundancy between two sets of ordination scores; y - Formula of linear distribution

The dominance of the Y axis occurred in the matrix profiles of primer pairs WRKY-R2a, WRKY-R2b, and WRKY-R3b against WRKY-R1, and WRKY-R3b against WRKY-R3a. The primer pairs WRKY-R2a, WRKY-R2b, and WRKY-R3b are characterized by similar dominance. The lowest mutual dominance is recorded between the primer pairs WRKY-R1 and WRKY-R3a. Table 4 shows the values of Rxy, which, if they have positive values, indicate a positive correlation of the independent matrices Y and X. The values of Rxy around point 0 aim at the unexpected uncorrelated state of the vector profiles of the corresponding correlation pair. Correlation itself (positive, negative) does not affect the presence of dominance.

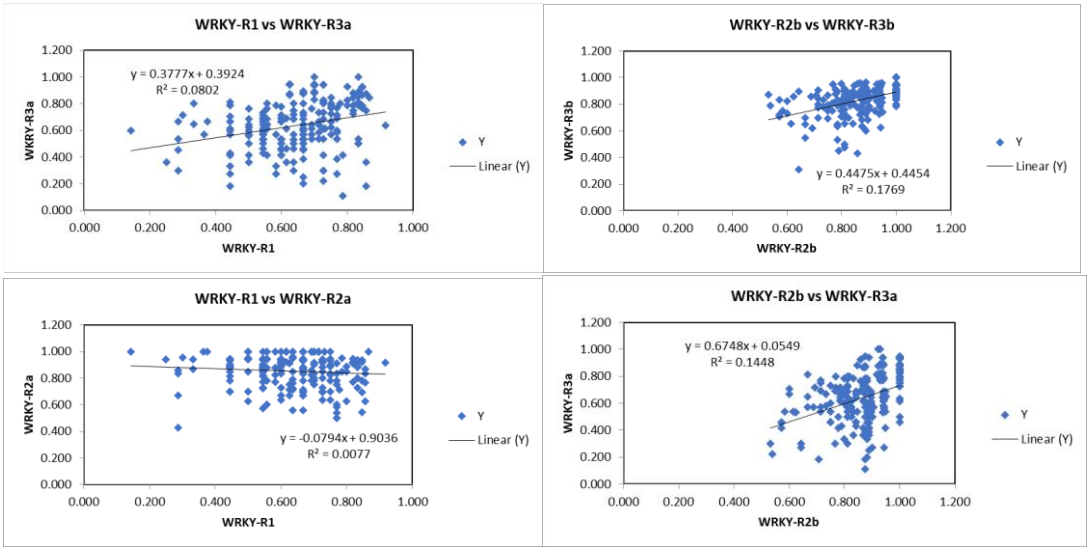


**Figure 4** Diagrams of correlation pairs by Mantel test (*A. hypogea* L.) presented 4 different relations

**PCoA analysis of *A. hypogaea* accessions**

Using the analysis of main coordinates, the division of the observed set of wild-growing genotypes was observed. Analysis was performed on genetic distance and covariance for each primer pair separately. Genetic distance is based on distance matrices based on Jaccard distance matrices with standardized data. Coordinate 1 explains from 0.071% to 0.088% of genetic variability and coordinate 2 explains from 0.059% to 0.077% of genetic variability. For genetic variance, PCoA relied on covariance matrices with standardized data based on Jaccard dissimilarity association matrices. Coordinate 1 ex-plains from 0.312% to 0.389% of genetic variability and coordinate 2 explains from 0.265% to 0.344% of genetic variability. Genetic variability is distributed among several dimensions. If the individual genotypes accumulated coordinately, it would indicate a correlation between the genotypes. PCoA analysis failed to generate clusters of genotypes, but only zones of accumulated genotypes (Figure 5). Here we present 5 potential coordination systems for tested genotypes. Coordination of genotypes with primer pairs WRKY-R1, WRKY-R2a, and WRKY-R3b generated 2 zones. The first zone included genotypes 3, 7, 19, 6, and 16 (WRKY-R1); 2, 6, 5, 17, 14, 7, 13, and 9 (WRKY-R2a) and 12, 20, and 15 (WRKY-R3b), which performed slightly because of geographically different conditions. The second zone of genotypes has already moved significantly away and included genotypes 4, 5, 12, 10, and 8 (WRKY-R1); 19, 3, and 18 (WRKY-R2a) and 17, 2, 9, 3, 13, 21, 16, and 19 (WRKY-R3b). In

this case, it is a geographical difference without correlation with genetic variability. For the remaining 2 primer pairs (WRKY-R2b and WRKY-R3a), PCoA analysis recognized only 1 zone with genotypes 10, 2, 13, 21, 3, 7, 11, 18, 6, and 1 (WRKY-R2b) and 13, 1, 16, 4, 18, 20, 6, 14, 11, 8, 21, and 3 (WRKY-R3a). Thus, a weak correlation of individual range between genetic variability and geographic distance is presented. The genetic variability of the genotypes does not show a significant geographical distance, even though they are free-growing *A. hypogea* L. plants. The originally cultured genotypes regained their genetic variability (after a few generations) typical of wild individuals.

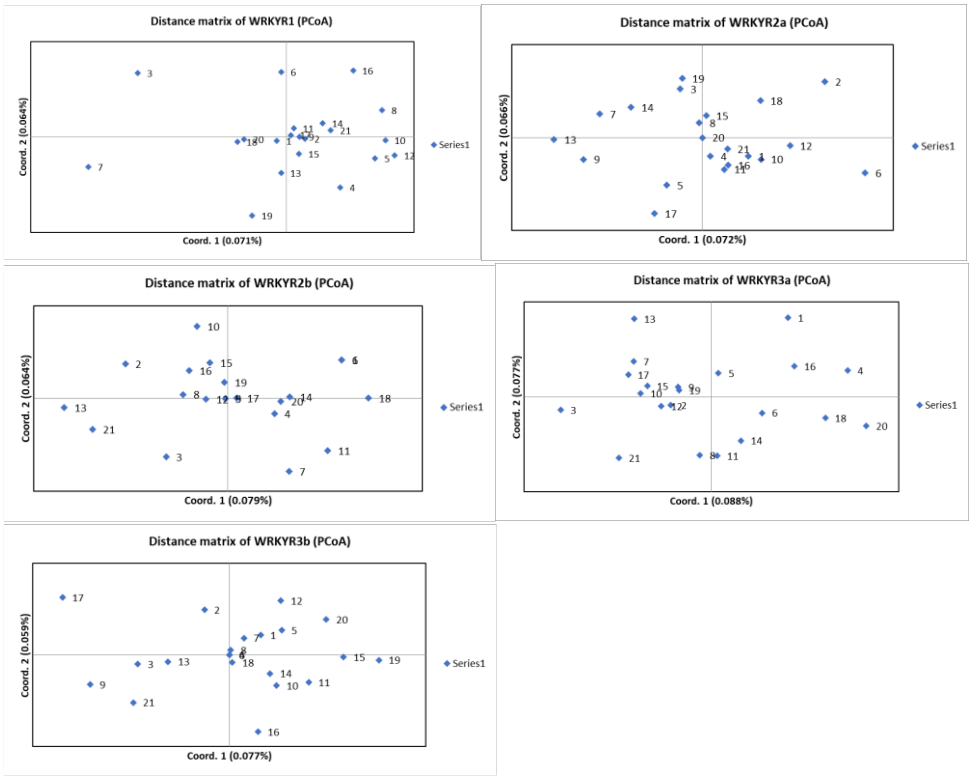


**Figure 5** Diagrams of coordination systems of *A. hypogea* L. genotypes by PCoA

**Cluster analysis of *A. hypogaea* accessions**

The Jaccard´s genetic similarity coefficient for all 21 *A. hypogaea* L. accessions based on five CDDP primers ranged from 0.09 (genotypes 4 and 15) to 0.48 (genotypes 15 and 20). Based on the UPGMA dendrogram, the analyzed *A. hypogaea* L. accessions could be divided into five main groups and multiple sub-groups when the cophenetic correlation coefficient (r) is 0.90 indicating high reliability of the clustering result (Figure 6). This UPGMA dendrogram was constructed utilizing the observed data from five CDDP molecular markers utilized in this study. This UPGMA dendrogram was constructed using the obtained data from five CDDP molecular markers utilized in this study.

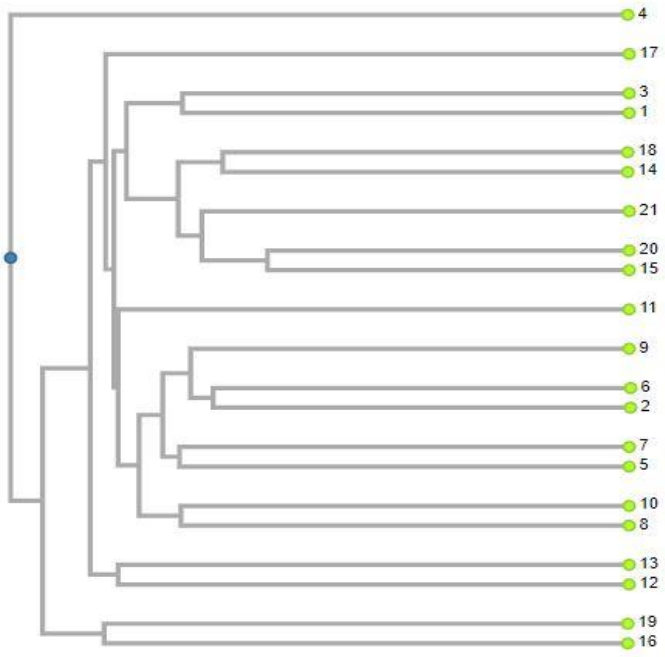


**Figure 6** Dendrogram of Jaccard genetic similarity values among 21 analysed *Arachis hypogaea* L. genotypes using all 5 CDDP marker

**DISCUSSION**

Is known that the cultivated peanuts originate in two species (*A. duranensis* and *A. ipaensis*) that are typical in southern Bolivia to northwestern Argentina. Based on natural localization, and archaeological analysis, it was suggested that the eastern slopes of Cordillera is the most probably area for the origin of *A. hypogaea* L. due to the favourable environment for peanut growth **(Stalker *et al.*, 1995)**. In our study were used 21 genotypes of *A. hypogaea* L. were gathered from the Bolivian region. Population studies aim to investigate interspecies and intraspecies structure on a regional level based on genetic data gained with 2 major groups of systems microsatel-lites and gene-targeted markers **(Abdella *et al.*, 2022; Moretzsohn *et al.*, 2004; Stalker *et al.*, 1995; Zou *et al.*, 2020)**. This challenge was expecting uniformity of geographical coordination for all sampled genotypes. Calculation of polymorphism, efficiency, correlation, structure, and clustering provide both confirming as well as refuting results from CDDP gene-targeted markers (WRKY). Assessment of genetic variability provides useful information for long-term sustainability and is an essential resource for plant conservation and breeding programs. Selection of suitable markers for genetic diversity analysis is extremely important and should be reliable in analyzing diversity. Pozcai et al 2013 reported, that the CDDP markers have high reproducibility compared to another amplified DNA markers. Conserved regions of DNA contain the same primer insertion sites but differ in gene distribution, generating polymorphic bands. In this study, the CDDP marker technique was successfully used to analyze the genetic diversity of 21 accessions of *Arachis hypogaea*, indicating that CDDP could be an effective tool for genetic analyses.

Based on data of each WRKY primer pair separately, were performed UPGMA clustering analysis with 5 different constructions of a dendrogram for CDDP markers so the recognition of a population structure failed even though their cophenetic correlation coefficient was highly similar (r = 88.77-93.23%). x All CDDP markers shown a low level of polymorphism, with an average of 41.2 %. WRKY-R1 primer shown the highest level of polymorphism at 47 % and primer WRKY-R3a was the lowest at 35%. This level of polymorphism suggests low genetic diversity in accessions of *A. hypogaea* L., the same as in the study, where were used newly developed microsatellite markers **(Moretzsohn *et al.*, 2004)** and the notion was corroborated that cultivated peanut presents a relatively reduced variation at the DNA level. Compared with previous studies, it has been shown that 90% polymorphism was found using CDDP markers in 143 *Solanum dulcamara* samples **(Poczai *et al.*, n.d.)** and 92.53% polymorphism was found using 19 CDDP primers in 53 *chrysanthemum* cultivars **(Li et al., 2014)**. In 120 different *Rosa rugosa* plants from northeastern China, the percentage of

polymorphic loci found using 13 selected CDDP primers was 94.53% **(Jiang & Zang, 2018)**.

Many researchers performing PCoA provide the first two components explaining genetic distances, variability, or similarity with high eigenvalues. The crucial part for higher eigenvalues is gathering an appropriate set of objects; over 100 **(Abdella *et al.*, 2022; Stalker *et al.*, 1995; Zou *et al.*, 2020)**. Based on that fact, 21 genotypes are not enough vast set for such analysis. That's why the significance of Mantel test ($p > 0.05$) does not indicate to reflect the original genetic distance matrix. All genotypes of *A. hypogaea* L. form a solid group and could be tracked into the common origin which collaborates with interspecies studies **(Moretzsohn *et al.*, 2004)** but is in contrast with intraspecies **(Abdella *et al.*, 2022; Moretzsohn *et al.*, 2004)**. The genetic structure of 384 peanut accessions that were divided into 2 populations showed 10.2% of south american accessions grouped in the first and 42.5% in the second group meanwhile 7 populations showed 20.9%, 75.5%, 5.7%, 12%, and 7.7% from south american accessions of peanut **(Zou *et al.*, 2020)**. More than 20% of accessions (4,578) were clustered into more than 1 cluster. Our results generated 5 clusters for 21 genotypes which consider 4.2 genotypes per group for the population. The average size of the population in our study is even lower than SE. **Legendre and Fortin, 2010** have shown the loss of power in studies of relationships between variables and rectangular data tables, which are turned into distance matrices. In situations where the question or hypothesis is formulated in terms of the raw data, Mantel tests should not be used.

## CONCLUSION

This study was aimed at the evaluation of genetic diversity and genetic similarity among *Arachis hypogaea* L. natural accessions, which could be provided important data to plant conservation and breeding programs. The CDDP marker technique used in this study provide sufficient distinguishment of *Arachis hypogaea* L. accessions. Based on CDDP fingerprints, analyzed genotypes were separated into different groups based on genetic diversity, but the detected level of polymorphisms for used markers among populations of *Arachis hypogaea* L. was low. These separated groups collide with 2 groupings of genotypes based on descriptive statistics. Geographic data could provide better insight into when such differences in clustering occurred. A slight explanation should associate with primer pairs' disproportional detection. A Set of primer pairs showed different detection dominance (high/low) shading with high TNB, NPB, PPB (%), Rp, and low PIC. If the aim of the study is population structure, dominant primer pairs are the better option of choice. If a study is more focused on monitoring or screening the population, low dominant CDDP primer pairs can provide a meaningful overview.

**Acknowledgments:** This publication was supported by the Operational Program Integrated Infrastructure within the project: Demand-driven Research for the Sustainable and Innovative Food, Drive4SIFood 313011V336, cofinanced by the European Regional Development Fund

## REFERENCES

Abdella, M., Kebede, M. & Feyissa, T. (2022). Genetic Diversity of Ethiopian Groundnut (*Arachis hypogaea* L.) Accessions. *East African Journal of Sciences*, *16*(1), 15–30.

Atia, M. A. M., Sakr, M. M. & Adawy, S. S. (2017). Assessing Date Palm Genetic Diversity Using Different Molecular Markers. *Methods in Molecular Biology (Clifton, N.J.)*, *1638*, 125–142. https://doi.org/10.1007/978-1-4939-7159-6_12

Bertioli, D. J., Seijo, G., Freitas, F. O., Valls, J. F. M., Leal-Bertioli, S. C. M. & Moretzsohn, M. C. (2011). An overview of peanut and its wild relatives. *Plant Genetic Resources*, *9*(1), 134–149. https://doi.org/10.1017/S1479262110000444

Bilčíková, J., Farkasová, S. & Žiarovská, J. (2021). Genetic variability of commercially important apple varieties (*Malus* x *domestica* Borkh.) assessed by CDDP markers. *Acta Fytotechnica et Zootechnica*, *24*(6), Article 6. https://afz.fapz.uniag.sk/legacy/journal/index.php/on_line/article/view/793/../793.html

Bontrager, M. & Angert, A. L. (2018). *Genetic differentiation is determined by geographic distance in Clarkia pulchella* (s. 374454). bioRxiv. https://doi.org/10.1101/374454

Collard, B. C. Y. & Mackill, D. J. (2009). Conserved DNA-Derived Polymorphism (CDDP): A Simple and Novel Method for Generating DNA Markers in Plants. *Plant Molecular Biology Reporter*, *27*(4), 558–562. https://doi.org/10.1007/s11105-009-0118-z

Cuc, L. M., Mace, E. S., Crouch, J. H., Quang, V. D., Long, T. D. & Varshney, R. K. (2008). Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). *BMC Plant Biology*, *8*(1), 55. https://doi.org/10.1186/1471-2229-8-55

De Riek, J., Calsyn, E., Everaert, I., Van Bockstaele, E. & De Loose, M. (2001). AFLP based alternatives for the assessment of Distinctness, Uniformity and Stability of sugar beet varieties: *Theoretical and Applied Genetics*, *103*(8), 1254–1265. https://doi.org/10.1007/s001220100710

*DendroUPGMA: Dendrogram construction using the UPGMA algorithm*. (n.d.). Cit 13. november 2022, z http://genomes.urv.cat/UPGMA/

Dwivedi, S. L., Gurtu, S., Chandra, S., Yuejin, W. & Nigam, S. N. (2001). Assessment of genetic diversity among selected groundnut germplasm. I: RAPD analysis. *Plant Breeding*, *120*(4), 345–349. https://doi.org/10.1046/j.1439-0523.2001.00613.x

Gantait, S., Panigrahi, J., Patel, I. C., Labrooy, C., Rathnakumar, A. L. & Yasin, J. K. (2019). Peanut (*Arachis hypogaea* L.) Breeding. V J. M. Al-Khayri, S. M. Jain & D. V. Johnson (Ed.), *Advances in Plant Breeding Strategies: Nut and Beverage Crops: Volume 4* (s. 253–299). Springer International Publishing. https://doi.org/10.1007/978-3-030-23112-5_8

Govindarajan, D. S., Kamalakkannan, K., Subiramani, S., G, P. K., Packiaraj, G., Sadhasivam, V., Swamynathan, R., Senthil Kumar, T. & Narayanasamy, J. (2014). Analysis of genetic variability in *Arachis hypogaea* L. using RAPD markers. *The Journal of the Swamy Botanical Club*, *31*, 29–36.

Hajibarat, Z., Saidi, A., Hajibarat, Z. & Talebi, R. (2015). Characterization of genetic diversity in chickpea using SSR markers, Start Codon Targeted Polymorphism (SCoT) and Conserved DNA-Derived Polymorphism (CDDP). *Physiology and Molecular Biology of Plants: An International Journal of Functional Plant Biology*, *21*(3), 365–373. https://doi.org/10.1007/s12298-015-0306-2

He, G. & Prakash, C. (2001). Evaluation of genetic relationships among botanical varieties of cultivated peanut (*Arachis hypogaea* L.) using AFLP markers. *Genetic Resources and Crop Evolution*, *48*(4), 347–352. https://doi.org/10.1023/A:1012019600318

Herselman, L. (2003). Genetic variation among Southern African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. *Euphytica*, *133*(3), 319–327. https://doi.org/10.1023/A:1025769212187

Jiang, L. & Zang, D. (2018). Analysis of genetic relationships in *Rosa rugosa* using conserved DNA-derived polymorphism markers. *Biotechnology & Biotechnological Equipment*, *32*(1), 88–94. https://doi.org/10.1080/13102818.2017.1407255

Klongová, L., Speváková, I., Žiarovská, J., Montero-Torres, J., Romero-Ortega, S., Pozzo, T., Fernandéz, E. (2021) Analýza variability genetických zdrojov podzemnice olejnej využitím dĺžkového polymorfizu medzi PBS miestami retrotranspozómov. *Recenzovaný zborník príspevkov z vedeckej konferencie Mladí vedci – bezpečnosť potravinového reťazca*, 43–47, ISBN 978-80-89738-30-4

Kochert, G., Halward, T., Branch, W. D. & Simpson, C. E. (1991). RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *81*(5), 565–570. https://doi.org/10.1007/BF00226719

Krapovickas, A. & Gregory, W. C. (2007). Taxonomía del género *Arachis* (Leguminosae). *Bonplandia*, *16*, 1. https://doi.org/10.30972/bon.160158

Kyseľ, M., Žiarovská, J., Štefúnová, V., Ražná, K. & Švec, M. (2021). Marker profiling of wheat with different drought torerance by CDDP. *Journal of Microbiology, Biotechnology and Food Sciences, 2021*(vol. 10), 1220–1222.

Li, T., Guo, J.-E., Zheng, C.-S., Sun, X. & Sun, X.-Z. (2014). Genetic diversity and construction of fingerprinting of chrysanthemum cultivars by CDDP markers. *Beijing Linye Daxue Xuebao/Journal of Beijing Forestry University*, *36*, 94–101. https://doi.org/10.13332/j.cnki.jbfu.2014.04.018

Montero-Torres, J., Droppa, M., Pozzo, T., Acebey-Aldunate, R., Fernández, E.C., Žiarovská, J. (2020) Characteristics of PBA profiling parkers in the analysis of *Arachis hypogaea* L. genome. *In Food and nutrition security and its resilience to global crises. Witzen-hausen German Institute for Tropical and Subtropical Agriculture.* 40, ISBN 978-3-945266-09-0.

Moretzsohn, M. de C., Hopkins, M. S., Mitchell, S. E., Kresovich, S., Valls, J. F. M. & Ferreira, M. E. (2004b). Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. *BMC Plant Biology*, *4*(1), 11. https://doi.org/10.1186/1471-2229-4-11

*Peanut Production by Country 2023*. (n.d.). Cit 05. január 2023, z https://worldpopulationreview.com/country-rankings/peanut-production-by-country

Poczai, P., Varga, I., Bell, N. E. & Hyvonen, J. (n.d.). *Genetic diversity assessment of bittersweet (Solanum dulcamara, Solanaceae) germplasm using conserved DNA-derived polymorphism and intron-targeting markers*. Cit 06. január 2023, z https://core.ac.uk/display/19207179

Prevost, A. & Wilkinson, M. J. (1999). A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics*, *98*(1), 107–112. https://doi.org/10.1007/s001220051046

Raina, S. N., Rani, V., Kojima, T., Ogihara, Y., Singh, K. P. & Devarumath, R. M. (2001). RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. *Genome*, *44*(5), 763–772.

Ren, X., Jiang, H., Yan, Z., Chen, Y., Zhou, X., Huang, L., Lei, Y., Huang, J., Yan, L., Qi, Y., Wei, W. & Liao, B. (2014). Genetic diversity and population structure of the major peanut (*Arachis hypogaea* L.) cultivars grown in China by SSR markers. *PloS One*, *9*(2), e88091. https://doi.org/10.1371/journal.pone.0088091

REN, X.-P., JIANG, H.-F., LIAO, B.-S., ZHANG, X.-J., LEI, Y., HUANG, G.-Q., YAN, L.-Y. & CHEN, Y.-N. (2011). Distributing and genetic diversity of high oleic acid germplasm in peanut (*Arachis Hypogaea* L.) core collection of China. *Journal of Plant Genetic Resources*, *12*.

Roldán-Ruiz, I., Dendauw, J., Van Bockstaele, E., Depicker, A. & De Loose, M. (2000). AFLP markers reveal high polymorphic rates in ryegrasses (Lolium spp.). *Molecular Breeding*, *6*(2), 125–134. https://doi.org/10.1023/A:1009680614564

Samaha, G. M., Ahmed, M. A. & Abd El-Hameid, A. R. (2019). Assessment of growth and productivity of five peanut cultivars and genetic diversity using RAPD markers. *Bulletin of the National Research Centre*, *43*(1), 168. https://doi.org/10.1186/s42269-019-0201-z

Sanders, T. H. (2003). PEANUTS. V B. Caballero (Ed.), *Encyclopedia of Food Sciences and Nutrition (Second Edition)* (s. 4420–4427). Academic Press. https://doi.org/10.1016/B0-12-227055-X/00897-X

Settaluri, V. S., Kandala, C. V. K., Puppala, N. & Sundaram, J. (2012). Peanuts and Their Nutritional Aspects—A Review. *Food and Nutrition Sciences*, *3*(12), 1644–1650. https://doi.org/10.4236/fns.2012.312215

Stalker, H. T., Dhesi, J. S. & Kochert, G. (1995). Genetic diversity within the species Arachis duranensis Krapov. &W.C. Gregory, a possible progenitor of cultivated peanut. *Genome*, *38*(6), 1201–1212. https://doi.org/10.1139/g95-158

Toomer, O. T. (2018). Nutritional chemistry of the peanut (*Arachis hypogaea*). *Critical Reviews in Food Science and Nutrition*, *58*(17), 3042–3053. https://doi.org/10.1080/10408398.2017.1339015

Upadhyaya, H. D., Ortiz, R., Bramel, P. J. & Singh, S. (n.d.). *Development of a groundnut core collection using taxonomical, geographical and morphological descriptors*.

Upadhyaya, H. D., Reddy, L. J., Gowda, C. L. L. & Singh, S. (2006). Identification of diverse groundnut germplasm: Sources of early maturity in a core collection. *Field Crops Research*, *97*(2–3), 261–271. https://doi.org/10.1016/j.fcr.2005.10.010

Variath, M. T. & Janila, P. (2017). Economic and Academic Importance of Peanut. V R. K. Varshney, M. K. Pandey & N. Puppala (Ed.), *The Peanut Genome* (s. 7–26). Springer International Publishing. https://doi.org/10.1007/978-3-319-63935-2_2

Wang, X., Fan, H., Li, Y., Sun, X., Sun, X., Wang, W. & Zheng, C. (2014). Analysis of genetic relationships in tree peony of different colors using conserved DNA-derived polymorphism markers. *Scientia Horticulturae*, *175*, 68–73. https://doi.org/10.1016/j.scienta.2014.05.026

Xie, Z., Zhang, Z.-L., Zou, X., Huang, J., Ruas, P., Thompson, D. & Shen, Q. J. (2005). Annotations and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signaling in aleurone cells. *Plant Physiology*, *137*(1), 176–189. https://doi.org/10.1104/pp.104.054312

Zou, K., Kim, K.-S., Kim, K., Kang, D., Park, Y.-H., Sun, H., Ha, B.-K., Ha, J. & Jun, T.-H. (2020). Genetic Diversity and Genome-Wide Association Study of Seed Aspect Ratio Using a High-Density SNP Array in Peanut (*Arachis hypogaea* L.). *Genes*, *12*(1), 2. https://doi.org/10.3390/genes12010002