

MARKER-ASSISTED SCREENING FOR POTENTIAL DUPLICATES AMONG COMMON BEAN GENE BANK COLLECTION

Merim Halilović¹, Arnela Okić¹, Almira Konjić¹, Silvio Šimon², Jasmin Grahić^{*1}

Original scientific paper

Abstract

The characterization and preservation of local plant genetic resources are crucial for maintaining agrobiodiversity and enhancing crop improvement initiatives. This study aimed to evaluate the genetic diversity and identify potential duplicates among 21 local common bean (*Phaseolus vulgaris* L.) accessions stored in the Gene Bank of the Faculty of Agriculture and Food Sciences at the University of Sarajevo. A total of 26 SSR (single sequence repeat) markers were utilized to genotype the accessions, revealing an average of 4.0 alleles per locus. The observed heterozygosity (H_o) averaged 0.043, while the expected heterozygosity (H_e) was 0.536. The polymorphic information content (PIC) values ranged from 0.171 to 0.859, with 11 loci classified as highly informative ($PIC > 0.5$). Cluster analysis based on Jaccard's distance and the UPGMA method indicated that no two accessions shared identical allele profiles, confirming the absence of genetic duplicates within the analyzed group. These findings highlight the significance of conserving and further characterizing local common bean germplasm as a vital resource for breeding programs focused on improving yield, stress resilience, and adaptation. Additionally, several SSR loci, especially PVat007, BM143, and GATS91, were identified as highly informative and suitable for future diversity and selection studies.

Keywords: *common bean, SSR markers, genetic diversity, Gene bank, landraces, duplicate identification*

INTRODUCTION

The common bean (*Phaseolus vulgaris* L.) is a crucial food legume globally, serving as a key source of dietary protein, minerals, and vitamins, particularly in traditional farming systems. In Bosnia and Herzegovina (B&H), the common bean is cultivated across a range of agroecological zones, where farmers have historically preserved diverse landraces that are well-adapted to local environmental conditions and cultural practices. These traditional varieties represent a valuable reservoir of genetic diversity,

¹ University of Sarajevo - Faculty of Agriculture and Food Sciences, Sarajevo, B&H

² Ministry of Agriculture Forestry and Fisheries Republic of Croatia, Zagreb, Croatia

Correspondence: j.grahic@ppf.unsa.ba

which is vital for food security, especially in the context of climate change and resource-limited agriculture (Grahić *et al.*, 2013; Grahić *et al.*, 2018). According to the 2023 report "Harvested area and yield of major crops", issued by the Agency for Statistics of Bosnia and Herzegovina, the dry common bean was harvested from 2,782 hectares, with an average yield of 3.51 t/ha (Agency for Statistics of B&H, 2023). These figures underscore the continued importance of this crop within smallholder and traditional farming systems across the country. Despite its significance, the genetic diversity of common bean landraces in B&H has not been extensively characterized using molecular markers. Previous studies have mainly focused on morphological traits (Grahić *et al.*, 2013) and their effectiveness in determining the origins of gene pools (Grahić *et al.*, 2018). While these methods provide valuable insights, their resolution is limited. Microsatellite markers (SSRs) are recognized for their efficacy in characterizing genetic diversity and remain an important tool for identifying distinct and potentially duplicated accessions (Carović-Stanko *et al.*, 2017; Savić *et al.*, 2021).

SSR markers are highly effective for characterizing common bean germplasm due to their considerable polymorphism, reproducibility, and codominant inheritance. They have been extensively employed in the Balkans and Europe to differentiate between landraces and commercial varieties of different agricultural crops, detect admixed genotypes, and inform conservation strategies (Maras *et al.*, 2015; Carović-Stanko *et al.*, 2017; Pipan and Meglič, 2019). A recent study focusing on B&H common bean germplasm, utilizing SSR analysis, confirmed the presence of both Andean and Mesoamerican gene pools within the national collection, shedding light on historical hybridization events and significant intraspecific diversity (Bajrić *et al.*, 2021). Given the increasing threat of genetic erosion due to the replacement of traditional varieties with commercial cultivars, it is essential to characterize and conserve local landraces. The Gene Bank at the Faculty of Agriculture and Food Sciences, University of Sarajevo, has successfully gathered and preserved a substantial number of common bean accessions from across Bosnia and Herzegovina. However, there remains uncertainty regarding the extent to which these accessions represent genetically distinct materials and the potential existence of duplicates.

This study employs 26 SSR markers to assess genetic diversity and identify possible duplicates among 21 local common bean accessions held in the Sarajevo Gene Bank. The results will enhance the efficient conservation, management, and utilization of local common bean genetic resources for breeding programs, thereby promoting sustainable agricultural development.

MATERIALS AND METHODS

For molecular analysis, 3 to 5 seeds from each of the 21 common bean (*Phaseolus vulgaris* L.) accessions stored in the Gene Bank of the Faculty of Agriculture and Food Sciences University of Sarajevo were individually sown in pots filled with commercial substrate. Genomic DNA was extracted from the young green leaves of the resulting

seedlings using the peqGOLD Plant DNA Kit (Peqlab, Germany), following the manufacturer's instructions.

A total of 26 primer pairs were utilized for SSR marker amplification (as detailed in Table 1). The PCR reactions were conducted using a Veriti™ Thermal Cycler (Applied Biosystems®, Foster City, California, USA), according to the protocol described by Carović-Stanko *et al.* (2017). Each 20 µL PCR reaction mixture contained 2 pmol of a tailed forward primer, 8 pmol of a reverse primer, 8 pmol of a FAM-labeled M13 primer, 1× PCR buffer, 4 pmol of each dNTP, 0.5 U of Taq™ HS DNA Polymerase (Takara Bio Inc.), and 5 ng of genomic DNA.

Table 1. Microsatellite markers used in this study The table shows the marker name, sequence, and their repetitive motifs.

Markers	Forward and reverse PCR primer	Repetitive motifs
BM157	F: ACTTAACAAGGAATAGCCACACA	(GA) ₁₆
	R: GTTAATTGTTTCCAATATCAACCTG	
BM172	F: TGTAACGACGGCCAGTATGCTGTAGCTCAAACAGGGCACT	(GA) ₂₃
	R: GCAATACCGCCATGAGAGAT	
BM143	F: TGTAACGACGGCCAGTATGCGGGAAATGAACAGAGGAAA	(GA) ₃₅
	R: ATGTTGGGAACCTTTAGTGTG	
GATS91	F: TGTAACGACGGCCAGTATGCGAGTGCAGGAAAGCGAGTAGAG	(GA) ₁₇
	R: TCCGTGTTCTCTGTCTGTG	
Pvctt001	F: TGTAACGACGGCCAGTATGCGAGGTGTTCACTATTGTCACTGC	(CTT) ₃ (T) ₃ (CTT) ₆
	R: TTCATGGATGGTGGAGGAACAG	
BM210	F: TGTAACGACGGCCAGTATGCACCACTGCAATCCTCATCTTTG	(CT) ₁₅
	R: CCCTCATCTCCATTCTTATCG	
BMd20	F: TGTAACGACGGCCAGTATGCGTTGCCACCGGTGATAATCT	(TA) ₅
	R: GTGAGGCAAGAAGCCTTCAA	
BMd12	F: TGTAACGACGGCCAGTATGCCATCAACAAGGACAGCCTCA	(AGC) ₇
	R: GCAGCTGGCGGGTAAACAG	
BM151	F: TGTAACGACGGCCAGTATGCCACAACAAGAACCTCCT	(TC) ₁₄
	R: TTATGTATTAGACCACATTACTTCC	
PVag001	F: TGTAACGACGGCCAGTATGCCAATCCTCTCTCTCATTTCCAATC	(GA) ₁₁
	R: GACCTTGAAGTCGGTGTCTTT	
PVat007	F: TGTAACGACGGCCAGTATGCAGTTAAATTATACGAGGTAGCCTAAATC	(AT) ₁₂
	R: CATTCCCTTCACACATTCACCG	
BMd42	F: TGTAACGACGGCCAGTATGCTCATAGAAGATTGTGGAAGCA	(AT) ₅
	R: TGAGACACGTACGAGGCTGTAT	
BMd45	F: TGTAACGACGGCCAGTATGCGGTTGGGAAGCCTCATAACG	(AG) ₅
	R: ATCTTCGACCCACCTTGCT	
BMb247	F: ATCCTAGGGAGTCATGAAGG	(AT) ₁₃
	R: AGAATTGTAACACACCGAC	
BMb96	F: CATAAAGCACGTCACCTCAA	(CA) ₁₁
	R: GCCTTGGACACTACCATTT	
BMd22	F: GGTCACTTCGGAGCATTC	(TC) ₆
	R: CGGGAAATGGAAGTCACAGT	
BMb508	F: TTGAGACAAATGACTCACCA	(CA) ₁₁
	R: CGTGTCTCTTTAAACAATCC	
BMd47	F: TGTAACGACGGCCAGTATGCACCTGGTCCCTCAAACCAAT	(AT) ₅
	R: CAATGGAGCACCAAAGATCA	
BMb469	F: CATTCATGTGAACCTTTTCATT	(AGA) ₁₀
	R: ATTGTTTGGTTTGTGCTTCT	
BMd46	F: GGCTGACAACAACCTCTGCAC	(TCT) ₄
	R: CTGGCATAGGTTGTCTCTTC	
BM197	F: TGGACTGGTCGATACGAAGC	(GT) ₈
	R: CCCAGAAAGATTGAGAACACCAC	
BMb267	F: TGAGCATCTCTACTTGGTT	(GAA) ₉
	R: AATCTCGCCTCTCTCTCTTT	

BMd53	F: TGTAACGACGGCCAGTATGCTGCTGACCAAGGAAATTCAG	(GTA) ₅
	R: GGAGGAGGCTTAAGCACAAA	
PVag003	F: TCACGTACGAGTTGAATCTCAGGAT	(AG) ₈
	R: GGTGTCGGAGAGGTTAAGGTTG	
BMb174	F: TTGAAACAAATCAGACCCCTC	(TTA) ₁₀
	R: ATACATAGATGCAAGAGCGA	
BMd25	F: TGTAACGACGGCCAGTATGCGCAGATCGCCTACTCACAAA	(GAT) ₆
	R: CGTTGACGAGAAGCATCAAG	

To calculate genetic diversity parameters, including the number of alleles per locus (N_a), observed heterozygosity (H_o), and expected heterozygosity (H_e), SpaGeDi v.1.2 software was employed (Hardy and Vekemans, 2002). The polymorphic information content (PIC) was calculated using Cervus v.3.0.7 (Kalinowski *et al.*, 2007).

Genetic distance among accessions was visualized using the Jaccard distance and clustered by the UPGMA method (Unweighted Pair Group Method with Arithmetic Mean) in the MEGA 6 software (Tamura *et al.*, 2013). All input data for statistical analysis were prepared using MADC v.2.0 software (Grahić and Grahić, 2017).

RESULTS AND DISCUSSION

A total of 26 microsatellite loci were successfully amplified from 21 local common bean accessions collected from various regions of Bosnia and Herzegovina (B&H), covering a wide range of agroecological zones, including continental lowlands, hilly regions, and mountainous rural areas. This geographic diversity ensures that the genetic variation assessed in this study accurately reflects the traditional cultivation range of common beans within the country. The SSR markers demonstrated significant genetic variability among the accessions, highlighting the effectiveness of SSRs for studying diversity in *Phaseolus vulgaris*.

Table 2. Characterization of the 26 microsatellite loci used on 21 common bean accessions

Analyzed loci	N_a	Range (bp)	H_o	H_e	PIC
BM157	4.0	96/114	0.048	0.603	0.546
BM172	4.0	81/109	0.000	0.562	0.495
BM143	7.0	112/152	0.000	0.832	0.790
GATS91	9.0	216/258	0.143	0.816	0.778
Pvctt001	4.0	153/163	0.095	0.548	0.436
BM210	5.0	162/188	0.048	0.710	0.645
BMd20	2.0	116/120	0.000	0.316	0.261
BMd12	3.0	151/163	0.048	0.554	0.465
BM151	4.0	139/149	0.000	0.660	0.581
PVag001	5.0	150/160	0.048	0.717	0.650
PVat007	11.0	192/224	0.095	0.893	0.859
BMd42	6.0	105/161	0.095	0.683	0.632
BMd45	2.0	88/126	0.048	0.494	0.366

BMb247	4.0	131/139	0.053	0.622	0.548
BMb96	5.0	114/197	0.048	0.758	0.699
BMd22	2.0	116/118	0.048	0.470	0.354
BMb508	3.0	244/250	0.050	0.432	0.365
BMd47	3.0	150/154	0.000	0.455	0.398
BMb469	2.0	141/147	0.053	0.235	0.202
BMd46	3.0	322/328	0.000	0.400	0.355
BM197	2.0	200/202	0.095	0.251	0.215
BMb267	2.0	223/226	0.000	0.193	0.171
BMd53	2.0	103/106	0.048	0.494	0.366
PVag003	2.0	163/165	0.000	0.455	0.346
BMb174	5.0	89/197	0.053	0.292	0.274
BMd25	2.0	115/118	0.000	0.483	0.360
Mean	4.0		0.043	0.536	0.468

Na - number of alleles; H_O - observed heterozygosity; H_E - expected heterozygosity; PIC - polymorphic information content.

The number of alleles per locus (Na) ranged from 2 to 11, with an average of 4.0 alleles per locus. The most polymorphic loci included PVat007 (Na = 11), GATS91 (Na = 9), and BM143 (Na = 7), which aligns with their previously reported informativeness in common bean research (Carvalho *et al.*, 2020; Gioia *et al.*, 2019; Carović-Stanko *et al.*, 2017) (Table 2.). Observed heterozygosity (H_O) was low (mean = 0.043), which is typical for a predominantly self-pollinating species like the common bean. In contrast, the expected heterozygosity (H_E) was higher (mean = 0.536), reflecting the underlying allelic diversity present within the collection. Polymorphic information content (PIC) values varied significantly, ranging from 0.171 (BMb267) to 0.859 (PVat007), with an average of 0.468. According to Botstein's classification, markers with PIC values greater than 0.5 are deemed highly informative. In this study, 11 loci exceeded this threshold, indicating that nearly half of the SSR markers utilized are suitable for diversity and differentiation analyses within local common bean germplasm. These findings are consistent with previous SSR-based studies on common bean landraces in the Western Balkans (Maras *et al.*, 2015; Savić *et al.*, 2021) and align with recent research on Bosnian accessions, where similar PIC values and allele ranges have been reported (Bajrić *et al.*, 2021). The presence of multiple alleles per locus reaffirms that the gene bank collection conserves a broad genetic base, highlighting its importance for conservation and breeding efforts.

The UPGMA dendrogram, which is based on Jaccard's genetic distance, categorized the 21 accessions into several clusters. Notably, no two accessions displayed identical allele profiles (Figure 1.). This finding strongly indicates the absence of genetic duplicates within the analyzed set, affirming the authenticity and uniqueness of each accession. Although a few accessions, such as BH04 and BH09, as well as BH11 and BH12, formed closely related pairs or subclusters, their profiles still exhibited sufficient allelic differences to classify them as genetically distinct. This clustering pattern aligns

with the observations of Grahić *et al.* (2013), who noted that landraces demonstrated gradual genetic differentiation rather than strict organization based on geographic origin. This pattern may reflect historical seed exchange practices among farming communities within B&H, where traditional landraces were frequently shared between regions regardless of geographic proximity. Such informal exchange systems could explain the presence of genetically similar but geographically dispersed accessions, contributing to a mosaic-like distribution of genetic variation. While specific accessions were found to be more closely related than others, no highly uniform subgroups emerged. Similar patterns of diversity have been documented in other studies focusing on landraces, wherein clustering reflects intricate genetic backgrounds rather than clear-cut substructuring (Maras *et al.*, 2015).

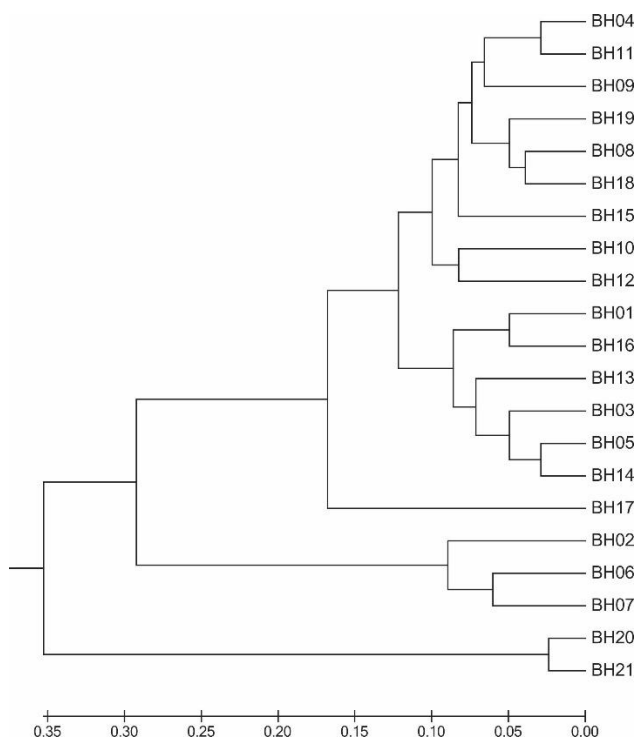


Figure 1. UPGMA cluster analysis based on polymorphisms of 26 SSR loci in 21 common bean accessions using Jaccards distance generated in MEGA 6 software (Tamura *et al.*, 2013). The findings indicate the absence of duplicates

CONCLUSIONS

The results of this study confirm the significant genetic diversity present among local common bean (*Phaseolus vulgaris* L.) accessions stored in the Gene Bank of the Faculty of Agriculture and Food Sciences in Sarajevo. No genetically redundant accessions were identified within the analyzed subset, indicating that each represents a unique genotype. These findings underscore the critical importance of preserving and further characterizing this germplasm, which serves as a valuable genetic resource for future breeding efforts aimed at improving stress resilience, yield, and adaptability.

In addition to molecular characterization, further evaluation of these accessions should include detailed phenotypic assessments, particularly for traits related to drought tolerance, disease resistance, and agronomic performance under low-input conditions. Such integrated characterization would enhance their practical value and facilitate their inclusion in targeted breeding programs.

Molecular tools remain essential for prioritizing unique accessions for conservation. Furthermore, the highly informative SSR loci identified in this study, especially PVat007, BM143, and GATS91, may prove effective tools for future germplasm screening and marker-assisted selection programs.

REFERENCES

- Agency for Statistics of B&H (2023). Harvested Area and Yield of Major Crops – 2023. Sarajevo: Agency for Statistics of Bosnia and Herzegovina. Available at: <https://bhas.gov.ba> (Accessed: June 2025).
- Bajrić, M., Gaši, F., Okić, A., Hadžidedić, M., Konjić, A., Grahić, J. (2021). Genetic diversity of the common bean (*Phaseolus vulgaris* L.) landraces from Bosnia and Herzegovina assessed using microsatellite markers. Works of the Faculty of Agriculture and Food Sciences, University of Sarajevo, Vol. 66, No. 71/2; pp. 20-29.
- Carović-Stanko, K., Liber, Z., Vidak, M., Barešić, A., Grdiša, M., Lazarević, B., Šatović, Z. (2017). Genetic diversity of Croatian common bean landraces. *Frontiers in Plant Science*, Vol. 8: 604; doi: 10.3389/fpls.2017.00604.
- Carvalho, M.S., de Oliveira Moulin Carias, C.M., Silva, M.A., da Silva Ferreira, M.F., de Souza, T.L.P.O., Posse, S.C.P., Ferreira, A. (2020). Genetic diversity and structure of landrace accessions, elite lineages and cultivars of common bean estimated with SSR and SNP markers. *Molecular Biology Reports*, Vol. 47, No. 9; pp. 6705-6715; doi: 10.1007/s11033-020-05726-7.
- Gioia, T., Logozzo, G., Marzario, S., Spagnoletti Zeuli, P., Gepts, P. (2019). Evolution of SSR diversity from wild types to U.S. advanced cultivars in the Andean and Mesoamerican domestications of common bean (*Phaseolus vulgaris*). *PLoS ONE*, Vol. 14, No. 1; doi: 10.1371/journal.pone.0211342.g005.

- Grahić, A., Grahić, J. (2017): MADC – Marker Analysis Data Compiler user's Manual.(www.divisionagro.ba/apps/docs/madc-marker-analysis-data-compiler/user-manual).
- Grahić, J., Ahmetović, A., Gaši, F., Karić, L., Šimon, S., Lazarević, B., Okić, A., Kurtović, M. (2018). Ispitivanje pouzdanosti morfoloških markera pri određivanju gen-centra porijekla običnog graha kolekcionisanog na području Bosne i Hercegovine. Radovi Poljoprivredno-prehrambenog fakulteta Univerziteta u Sarajevu, Vol. 68, No. 2; pp. 18-24.
- Grahić, J., Gaši, F., Kurtović, M., Karić, L., Đikić, M., Gadžo, D. (2013). Morphological evaluation of common bean diversity in Bosnia and Herzegovina using the discriminant analysis of principal components (DAPC) multivariate method. Genetika, Vol. 45, No. 3; pp. 963-977.
- Hardy, O.J., Vekemans, X. (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology, Vol. 2; pp. 618-620
- Kalinowski, S.T., Taper, M.L., Marshall, T.C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology, Vol. 16, No. 5; pp. 1099–1106.
- Maras, M., Pipan, B., Sustar-Vozlic, J., Todorović, V., Đurić, G., Vasić, M., Jordanovska, S., Ibusoska, A., Agić, R., Matotan, Z., Cupic, T., Meglič, V. (2015). Examination of genetic diversity of common bean from the Western Balkans. Journal of the American Society for Horticultural Science, Vol. 140, No. 4; pp. 308-3016; doi: 10.21273/JASHS.140.4.308.
- Pipan, B., Meglič, V. (2019). Diversification and genetic structure of the western-to-eastern progression of European *Phaseolus vulgaris* L. germplasm. BMC Plant Biol, Vol. 19, No. 442; doi: 10.1186/s12870-019-2051-0.
- Savić, A., Pipan, B., Vasić, M., Meglič, V. (2021). Genetic diversity of common bean (*Phaseolus vulgaris* L.) germplasm from Serbia, as revealed by single sequence repeats (SSR). Scientia Horticulturæ, Vol. 288; doi: 10.1016/j.scienta.2021.110405.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol, Vol. 30, No. 12; pp. 2725-2729; doi: 10.1093/molbev/mst197.

MARKERIMA POTPOMOĞNUTI SKRINING POTENCIJALNIH GENETIČKIH DUPLIKATA U KOLEKCIJI OBIČNOG GRAHA POHRANJENOJ U GEN-BANCI

Sažetak

Detaljna genetička karakterizacija lokalnih sorti graha od ključne je važnosti za njihovo očuvanje i buduću upotrebu u oplemenjivanju i selekciji. Ova studija imala je za cilj procijeniti genetičku raznolikost i identificirati moguće duplikate između 21 lokalne primke običnog graha (*Phaseolus vulgaris* L.) pohranjene u gen banci Poljoprivredno-prehrambenog fakulteta Univerziteta u Sarajevu. Za genotipizaciju je korišteno ukupno 26 SSR markera, pri čemu je prosječan broj alela po lokusu iznosio 4,0. Ustanovljena heterozigotnost (H_o) iznosila je, u prosjeku, 0,043, dok je prosječna očekivana heterozigotnost (H_E) imala vrijednost od 0,536. Vrijednosti polimorfnog informativnog sadržaja (PIC) kretale su se od 0,171 do 0,859, pri čemu je 11 lokusa klasificirano kao visoko informativno ($PIC > 0,5$). Klaster analiza zasnovana na Jaccardovoj genetičkoj distanci i UPGMA metodi pokazala je da nijedna od analiziranih primki ne dijeli identičan alelni profil, što potvrđuje odsustvo genetičkih duplikata u okviru ispitivane kolekcije. Dobijeni rezultati ističu značaj očuvanja i daljnje karakterizacije lokalnog sortimenta graha kao vrijednog genetičkog resursa za oplemenjivačke programe usmjerene na poboljšanje prinosa, otpornosti na stres i adaptabilnosti. Također, nekoliko SSR lokusa, naročito PVat007, BM143 i GATS91, izdvojilo se kao visoko informativno i pogodno za buduće studije diverziteta i selekcije.

Ključne riječi: *obični grah, SSR markeri, genetički diverzitet, Gen banka, lokalne sorte, identifikacija duplikata*