ANNUAL REPORT OF THE

BEAN IMPROVEMENT COOPERATIVE

A VOLUNTARY AND INFORMAL ORGANIZATION
TO EFFECT THE EXCHANGE OF INFORMATION AND MATERIALS

VOLUME 62
2019

GENOTYPING THE *EX SITU* GENETIC RESOURCES OF WILD AND CULTIVATED TEPARY BEAN

John P. Hart[1], Ana G. Vargas[2], James S. Beaver[2], Daniel G. Debouck[3], and Timothy G. Porch[1]

[1]USDA-ARS-TARS, 2200 P.A. Campos Ave, Suite 201, Mayagüez, PR 00680
[2]Dept. of Agroenvironmental Sciences. University of Puerto Rico, Mayagüez, PR 00680
[3]Genetic Resources Program, CIAT, Km 17 Recta Cali - Palmira, Cali, Colombia

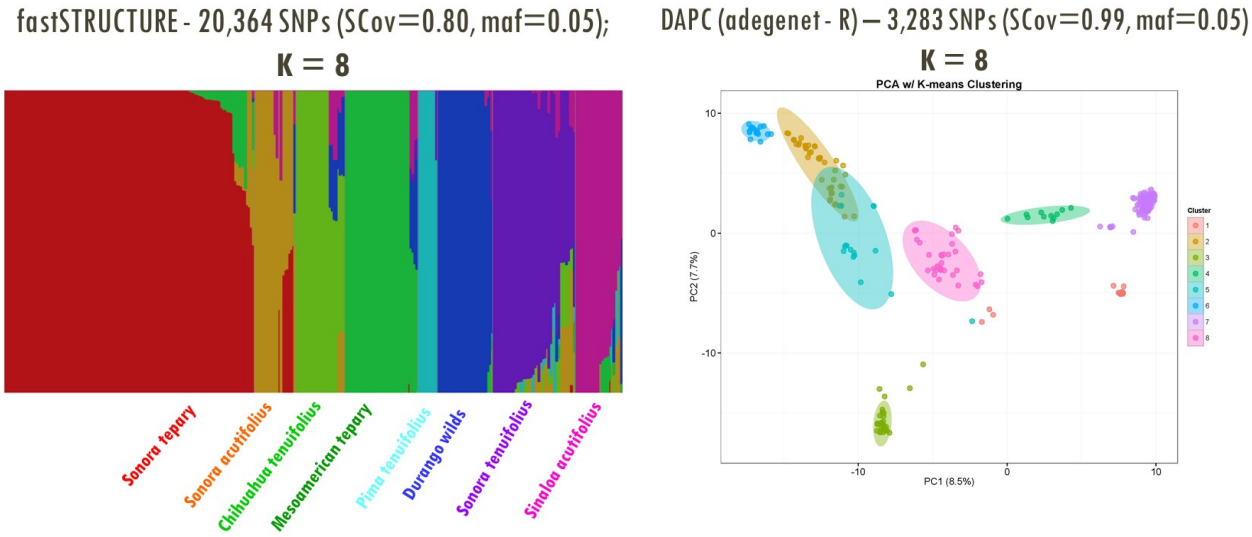
INTRODUCTION: The tepary bean (*Phaseolus acutifolius* A. Gray) is a relatively untapped source of novel allelic diversity both as a donor for common bean improvement, and as an intrinsically stress-tolerant and nutritious food crop. Given the present and future socioeconomic and climatic scenarios for Phaseolus bean production and the immense potential that tepary beans may offer for adaptation to these scenarios, further characterization of the *ex situ* genetic resources is timely. The objectives of this research were to assemble, increase, and genotype all of the available wild and cultivated tepary bean accessions held by the USDA, CIAT, and TARS germplasm collections, and to investigate the genetic diversity and population structure of this germplasm as it relates to domestication status, morphological classification, and geographical distribution. These objectives are part of an overall objective to advance the conservation, utilization, and deployment of the genetic diversity of tepary beans.

MATERIALS AND METHODS: An initial total of 314 accessions (158 wild, 156 cultivated) were obtained and increased in the screenhouse at TARS as single plants. The 156 cultivated accessions were further increased in the field and have been phenotyped extensively. These 314 accessions were genotyped as part of a 384-plex *Ape*KI genotyping-by-sequencing (GBS) library (Elshire et al., 2011; Hart and Griffiths, 2015) by submitting this library to 4 lanes of 101-cycle sequencing on an Illumina HighSeq 2500 at the Weill Cornell Genomics Resources Core. The resulting GBS tags were processed with the GBS Discovery Pipeline for species with a reference genome in TASSEL v3.0 (Bradbury et al., 2007), aligned to the *P. vulgaris* v1.0 reference genome (Schmutz et al., 2014), and SNPs were called and filtered with the TASSEL v3.0 Discovery SNP Caller (Glaubitz et al., 2014).

RESULTS AND DISCUSSION: This genotyping effort resulted in 3.2 million unique GBS tags of which 50% could be aligned to the reference genome with the Burrows-Wheeler Alignment (BWA) tool (Li and Durbin, 2009), or 64% with Bowtie 2 (Langmead and Salzberg, 2012). After alignment with BWA we were able to discover and genotype 20,364 SNPs (with MAF $\geq$ 0.05) that were present in at least 80% of the accessions. When the wild germplasm was considered exclusively, this number changed to 23,070 SNPs and was in sharp contrast with the cultivated germplasm where only 7,642 SNPs were discovered. This is another indicator of the severely reduced diversity in cultivated tepary germplasm. We used the dataset for all of the accessions to investigate population structure with fastSTRUCTURE (Raj et al., 2014) and 3,283 SNPs that were called in all accessions for discriminant analysis of principal components (DAPC) (Jombart et al., 2010). The results of both analyses suggested that the number of subpopulations present in the germplasm is equal to eight (Figs. 1 and 2), and that the membership of each accession in the eight subpopulations is structured based on domestication status, geographic origin, and

morphological variation. These preliminary results confirm the strong bottleneck caused by tepary domestication, identify subpopulations of tepary germplasm in both the wild and cultivated genepools according to geographical origin, and present extensive opportunities for further research into the evolution, domestication, diversity, and improvement of tepary bean.

Figures 1 & 2: Results of fastSTRUCTURE analysis (left) and DAPC (right) with 314 wild and cultivated tepary germplasm accessions.



REFERENCES:

Bradbury, P.J., et al.2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinforma. 23:2633-2635.

Elshire, R.J., et al. 2011. A Robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6(5): e19379. doi:10.1371/journal.pone.0019379

Glaubitz, J.C., et al. 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis platform. PLoS ONE 9(2):e90346. doi:10.1371/journal.pone.0090346

Hart, J.P. and P.D. Griffiths. 2015. Genotyping-by-sequencing enabled mapping and marker development for the *By-2* potyvirus resistance allele in common bean. Plant Genome 8:1-14. doi:10.3835/plantgenome2014.09.0058

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11:94. doi: 10.1186/1471-2156-11-94

Langmead, B., and S.L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods. 9:357-359. doi:10.138/nmeth.1923

Li, H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinform. 25:1754-60. doi: 10.1093/bioinformatics/btp324

Raj, A., M. Stephens, and J.K. Pritchard. 2014. fastSTRUCTURE: Variational inference of population structure in large SNP datasets. Genet. 197:573-589. doi:10.1534/genetics.114.164350

Schmutz, J., et al. 2014. A reference genome for common bean and genome-wide analysis of dual domestications. Nat. Genet. 46: 707–713. doi:10.1038/ng.3008