

Using SNP haplotypes to study gene flow in a *P. vulgaris* wild-weedy-crop complex from Colombia: a pilot study

C. Quintero¹, R.I. González-Torres¹, E. Gaitán-Solís², R. Hidalgo⁴, O. Toro³, D.G. Debouck³ and J. Tohme¹

¹CIAT, Biotechnology Research Unit, AA6713, Cali, Colombia; ²Donald Danforth Plant Science Center, Missouri, USA; ³CIAT, Genetic Resources Unit, Cali, Colombia; ⁴Universidad Nacional de Colombia, Palmira, Colombia

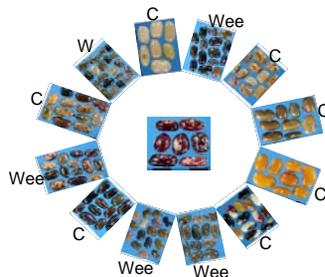
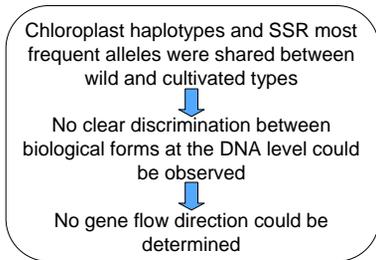
INTRODUCTION

The study of "wild-weedy-crop" complexes along common bean range of distribution in the Americas has shown no preferential gene flow direction in Colombia, Ecuador and Peru (Chacón *et al.*, 2006). We conducted a pilot study in the wild-weedy-crop complex G50879 from Colombia (19, wild; 32 weedy and 34 cultivated), previously evaluated with biochemical markers and SSRs, to test the potential of nuclear SNP markers as an alternate tool for the identification of gene flow events and inference about their direction.

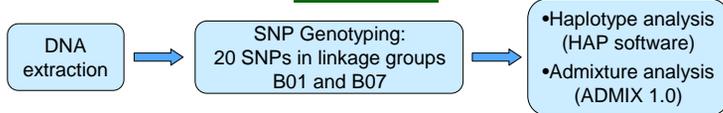
MATERIALS

Wild-weedy crop complex G50879

Seed from the original collection surrounded by wild, weedy and cultivated forms.



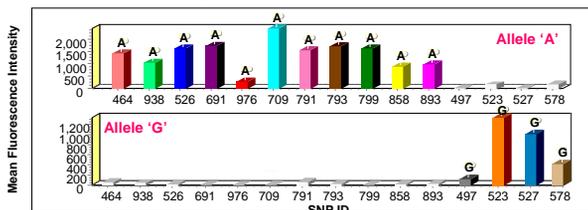
METHODS



RESULTS AND DISCUSSION

SNP genotyping:

Simultaneous genotyping up to 15 alleles in one single base extension reaction was achieved using a Luminex¹⁰⁰ flow cytometer equipped with a Luminex XY Platform plate reader (Quintero *et al.*, 2005).



Haplotype inference:

Block partition and haplotype prediction within each block was done using HAP software (Halperin and Eskin, 2004). Both linkage groups were divided in three blocks of limited diversity. 43 SNP haplotypes were observed in the wild-weedy-crop complex: 84% in weedy forms, 77% in cultivated forms and 70% wild forms.

Linkage group B01

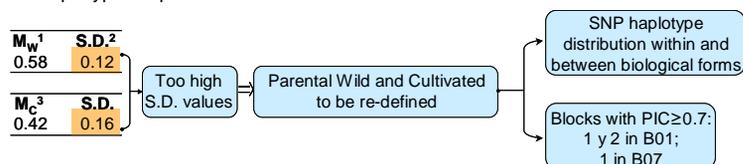
Block	SNPs (#)	Haplotypes (#)	PIC
1	3	8	0,702
2	5	12	0,756
3	2	4	0,310

Linkage Group B07

Block	SNPs (#)	Haplotypes (#)	PIC
1	6	11	0,731
2	3	6	0,498
3	1	2	0,288

Initial admixture analysis

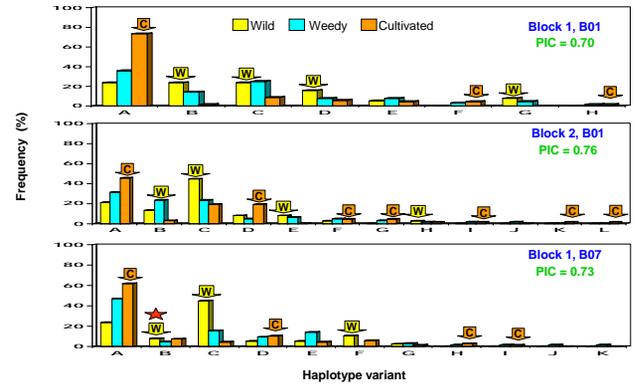
The admixture coefficient M_V (Bertorelle & Excoffier, 1998) was calculated, based on haplotype frequencies and the molecular distances between them.



¹ M_w : relative contribution from the wild parental population
²S.D.: Standard deviation based on 1000 resampling events.
³ M_c : relative contribution from the cultivated parental population

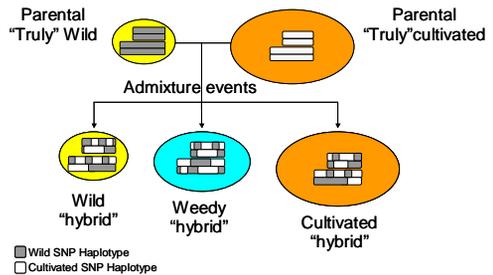
SNP haplotypes frequency in the complex

Differences between haplotype frequencies allowed the identification of those haplotypes that described better each of the parental biological forms. In most cases, haplotypes were classified as wild types when their frequency was at least twice the observed in cultivated forms, and vice versa.



Admixture model

Parental truly-wild and truly-cultivated populations were selected based on the coincidence of their phenotype with their genotype (SNP haplotypes). After admixture events, they could give rise to three hybrid subpopulations in this wild-weedy-crop complex.



A more accurate estimation of admixture coefficients (S.D. values close to zero) in the hybrid subpopulations was calculated.

Cultivated (hybrid)

Mw	SD	Mc	SD	Mc/Mw	Gene flow
0,3237	0,0501	0,6766	0,0495	2,1	C→W

Weedy

Mw	SD	Mc	SD	Mc/Mw	Gene flow
0,4458	0,0436	0,5518	0,043	1,2	C↔W

Wild (hybrid)

Mw	SD	Mc	SD	Mc/Mw	Gene flow
0,4669	0,055	0,5281	0,0588	1,1	C↔W

In the weedy and wild-hybrid subpopulations, the contribution of truly-cultivated was slightly higher than that of truly-wild meaning that both gene flow directions are possible and almost symmetric. For the cultivated-hybrid, the contribution of truly-cultivated was twice that of the truly-wild, suggesting more pollen flow from the cultivated types. Although we have looked only at one wild-weedy-crop complex from Colombia our results show that SNP haplotypes are informative enough to provide evidence about gene flow dynamics in *P. vulgaris*.

REFERENCES

- Bertorelle G, Excoffier L. 1998. Molecular Biology and Evolution. 15:1298-1311.
 Chacón, M.I. González-Torres, R.I. & D. G. Debouck. 2006. Council of Science. CIAT. 2006.
 Halperin E. and E. Eskin. 2004. Bioinformatics. 20(12):1842-9.
 Quintero, C., E. Gaitán-Solís, C. Quigley, P. Cregan, M.C. Duque and J. Tohme. 2005. Annual Report SB02-Project. 21-23p.

ACKNOWLEDGEMENTS

Special thanks go to Drs Perry Cregan and Charles Quigley (Soybean Genomics Lab-USDA) who kindly shared with us the information and primers of the soybean SNPs they discovered. To the "Ginés-Mera memorial fellowship fund" for postgraduate studies in biodiversity for partial funding of this research and to Myriam C. Duque for helpful comments.