

Genetic diversity and population structure of diploid potato genotypes

Deissy Juyó¹, María Fernanda Álvarez¹, Helena Brochero¹, Andrés Cortes², Christiane Gebhardt³, Teresa Mosquera¹. Universidad Nacional de Colombia, Facultad de Agronomía, Bogotá-Colombia
 2. University of Uppsala-Department of Plant Ecology and Evolution, Uppsala-Sweeden
 3. Max Planck Institute for Planck Breeding Research, Cologne-Germany

General objective

To perform a genetic characterization of the diploid genotypes of *Solanum tuberosum* group Andigenum from the Colombian Core Collection (Col) with SSR markers to determine its genetic diversity and population structure, with the intention of generating basis for association mapping.

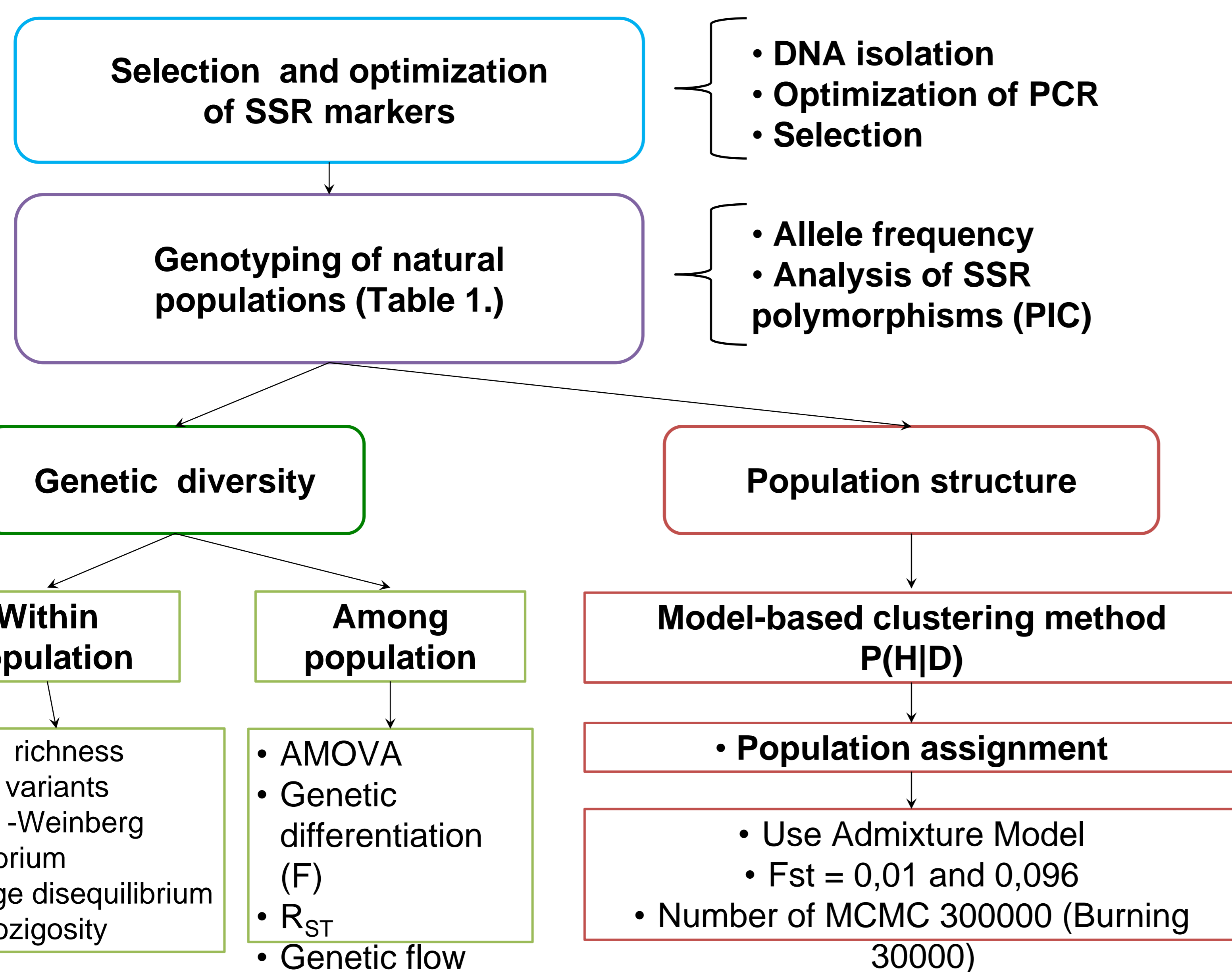
Introduction

Colombia has a diversity center for diploid genotypes of *Solanum tuberosum* group Andigenum, these genotypes are important resources for genetic studies and they have been used in potato breeding programs. In Andean countries these potatoes are staple food and represent important incomes for peasants. We studied the genetic diversity and population structure of 110 diploid potato genotypes by the analysis of allele frequencies using a set of 42 microsatellites (SSRs). Genotypes used in this study came from natural populations: from Col and diploid accessions from the German germplasm bank (Gatersleben Bank-IPK).

Material and methods

Table 1. Natural populations of diploid genotypes of *Solanum tuberosum* group Andigenum

Populations	ID	Genotypes	Origen
Colombian Core Collection	Col	97	Collected in different provinces of Colombia
Gatersleben Bank	IPK	13	Different countries



Conclusion

The diploid population has high genetic diversity, possibly due to the high gene flow and the differentiation among individuals within populations. The lack population structure found in this study allows to use this population for genetic association studies

Genetic diversity

The greater degree of population differentiation was found among individuals within populations, $F_{IT} = 0,251$. There was a low rate of population differentiation $F_{ST} = 0,096$ (Table 2). Genetic diversity within populations was expected heterozygosity or $H_s = 0,555$. The differentiation index for multiple alleles between subpopulations (G_{st}) indicated that only 4.5% of the total detected genetic diversity was explained by differences among populations. The R_{st} statistic reveals significant differences between the two populations ($R_{st} = 0,23635$, $P = 0,0000$). The estimation of gene flow (Slatkin, 1995) between Col and IPK populations was highly significant ($N_m = 1,6154$, $P = 0,00000^*$).

Population structure

Individuals were assigned to a the same specific K populations using Structure (Pritchard *et al.*, 2000) assuming two F_{ST} differentiation values among populations (0.001 default settings and calculated F_{ST} 0.096). These two analyzes showed a continuous increase in the statistic goodness of fit, $\ln [P (D)]$, therefore no population structure was found. According to Evanno *et al.* (2005) (maximum likelihood of $\ln [P (D)]$) the optimum number of populations was two ($K = 2$) which indicated a split between the majority of Col genotypes and IPK genotypes (Figure 1).

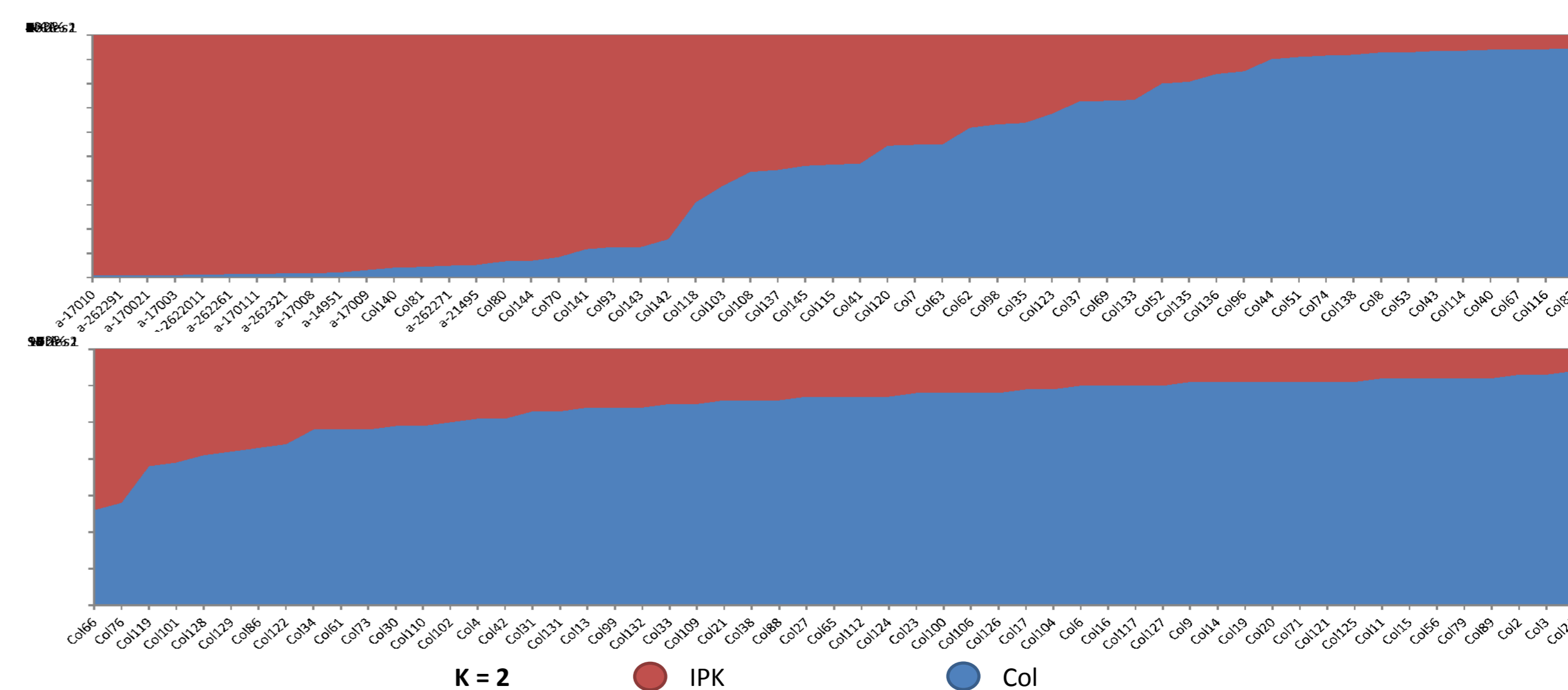


Figure 1. Depiction of Col and IPK population structure. Bar plot of diploid potato genotypes generated by STRUCTURE 2.3.2.1 using the admixture model with independent allele frequencies. Groups are represented by colors, as indicated in the legend. Each column (110 in total) represents an individual, and its genotype is divided into segments indicating its likely genetic origin.

Results

We selected a set of 42 SSRs distributed in the genome which allowed the characterization of the populations. Among the populations, the markers that showed the highest number of alleles were STI023 for Col population (11 alleles) and STI021 for IPK (nine alleles).

PIC values found in this study (0,458) are within the ranges obtained for other analyses of diversity in potato, from 0.25 to 0.884 (Ghislain *et al.*, 2006, 2009). The presence of rare alleles was demonstrated in 22 SSRs from Col population (i.e. STI031, STM1049, STI020, STI002, STI032) and in one SSR from IPK population (STI043).

Table 2. Coefficients of genetic differentiation between Col and IPK populations

Population	F_{IS}	F_{ST}	F_{IT}
Natural	0,17115	0,09636	0,25103
P-value	0,00000*	0,00000*	0,00000*

F_{IS} : Inbreeding coefficient within populations. F_{ST} : Coefficient of genetic differentiation between populations. F_{IT} : Total inbreeding coefficient.

Discussion

The high genetic diversity and lack of population structure in Col and IPK populations can be attributed to a strong self-incompatibility due to the presence of the S alleles system, high intervarietal hybridization (2X x 2X and 2X x 4X) and high genetic flow between diploid genotypes.

Reference

- Evanno *et al.* 2005. Ecology, 14: 2611-2620.
- Pritchard *et al.* 2000. Genetics, 155: 945-959.
- Slatkin M. 1995. Genetics, 139: 457-462.
- Ghislain *et al.* 2009. Molecular Breeding, 23: 377-388.
- Ghislain *et al.* 2006. Theoretical and Applied Genetics, 113 (8): 1515-1527.

Diversidad genética y estructura poblacional de genotipos diploides

Juyó, D.¹, M.F. Álvarez¹, H. Brochero¹, A. Cortes², C. Gebhardt³ y T. Mosquera¹. Universidad Nacional de Colombia 2. Universidad de Uppsala 3. Max Planck Institute for Planck Breeding Research

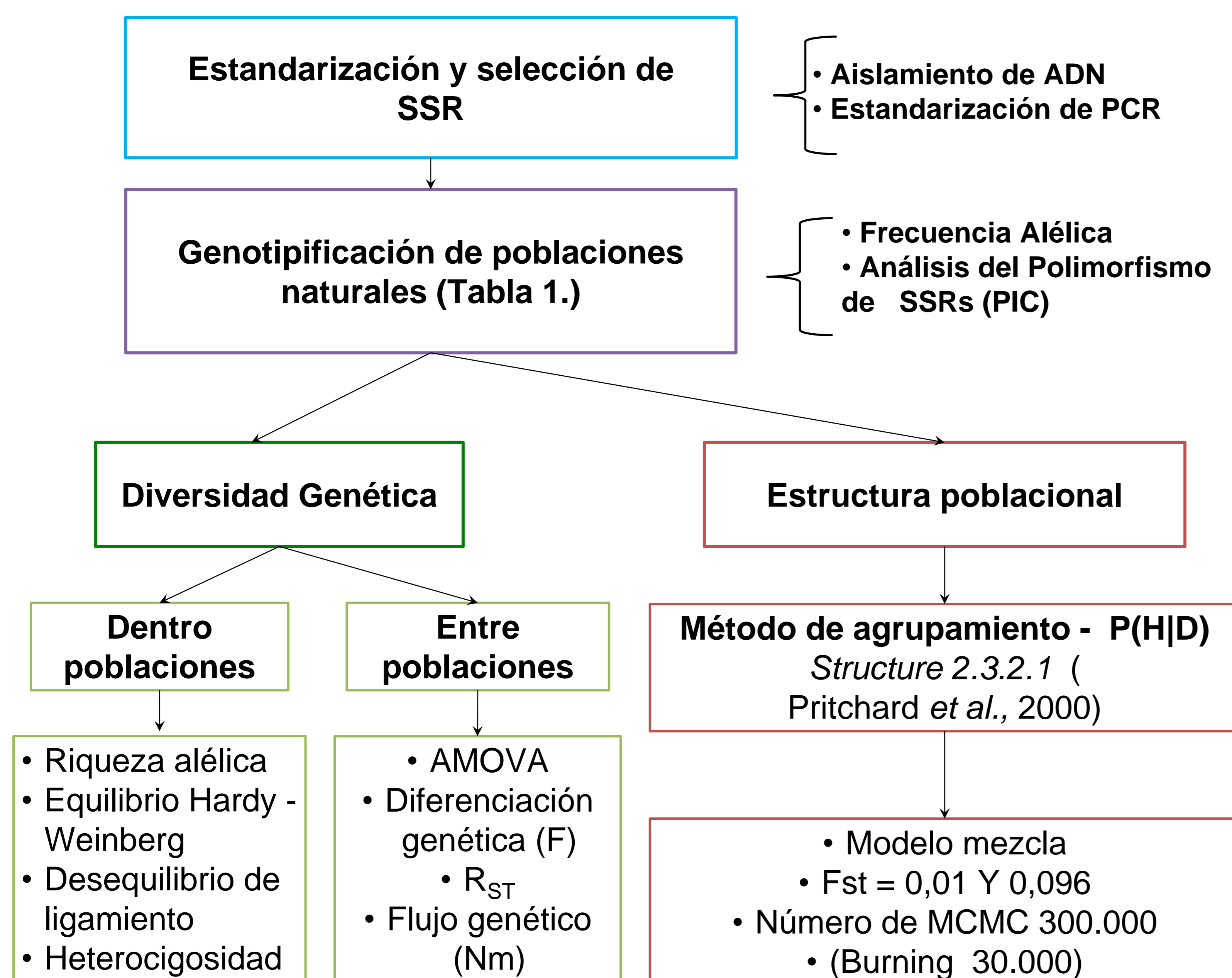
Introducción

Colombia posee un centro de diversidad de genotipos diploides de *S. tuberosum* grupo Andigenum, estos genotipos son un recurso importante para estudios genéticos y los programas de mejoramiento de papa. En este estudio se evaluó la diversidad genética y la estructura poblacional de 110 genotipos diploides de papa por medio del análisis de frecuencias alélicas encontradas con un grupo de 42 microsatélites (SSR), con el propósito de generar conocimiento base para estudios de mapeo por asociación. Los genotipos utilizados provenían de poblaciones naturales: accesiones diploides de la Colección Central Colombiana CCC y del Gatersleben Bank – IPK (Banco de germoplasma Alemán).

Metodología

Tabla 1. Poblaciones naturales de genotipos diploides de *S. tuberosum* grupo Andigenum

Poblaciones	ID	Genotipos	Origen
Colección Central Colombiana	CCC	97	Colectadas en diferentes regiones de Colombia
Gatersleben Bank	IPK	13	Diferentes países



Resultados

Entre las poblaciones el marcador que mostro un mayor número de alelos fué STI023 para Col (11 alelos) y STI021 para IPK (nueve alelos). Los valores de PIC revelados (0,458) se encontraron dentro del rango obtenido para otros análisis de diversidad en papa, 0,25 a 0,884 (Ghislain *et al.*, 2006, 2009). La presencia de alelos raros fué desmostrada para 22 SSRs de la población Col (ej. STI031, STM1049, STI020, STI002, STI032) y solo en un SSR de la población IPK (STI043).

Tabla 2. Coeficientes de diferenciación genética entre las poblaciones naturales Col y IPK

Poblaciones	F_{IS}	F_{ST}	F_{IT}
Naturales	0,17115*	0,09636	0,25103

F_{IS} : Coeficiente de endogamia dentro de poblaciones F_{ST} : Coeficiente de diferenciación genética entre poblaciones F_{IT} : Coeficiente total de endogamia
* P-value = 0,0000

Diversidad genética

El mayor grado de diferenciación genética se encontró entre los individuos dentro de las poblaciones ($F_{IT}=0,251$) y la tasa de diferenciación entre poblaciones $F_{ST} = 0,096$ fué baja (Tabla 2). La diversidad genética en términos de heterocigosidad fué alta (heterocigosidad esperada de 0.555). El estadístico de diferenciación genética (R_{st}) para SSR, reveló diferencias significativas entre las dos poblaciones ($R_{st} = 0.23635$, $P = 0.0000$). La estimación del flujo genético (Slatkin, 1995) entre Col e IPK fué altamente significativa ($N_m = 1,6154$, $P = 0,000$).

Estructura poblacional

Los individuos fueron asignados a la misma población K , asumiendo dos valores de diferenciación poblacional F_{ST} (0,001 valor predeterminado y 0,096 valor calculado). Estos dos análisis mostraron un continuo incremento de la bondad de ajuste estadístico para el logaritmo de la verosimilitud de los datos $\ln[P(D)]$, indicando la falta de estructura poblacional entre los genotipos estudiados.

De acuerdo con Evanno *et al.* (2005) el número óptimo de poblaciones fué dos ($K = 2$) indicando una división entre la mayoría de genotipos Col y genotipos IPK (Figura 1).

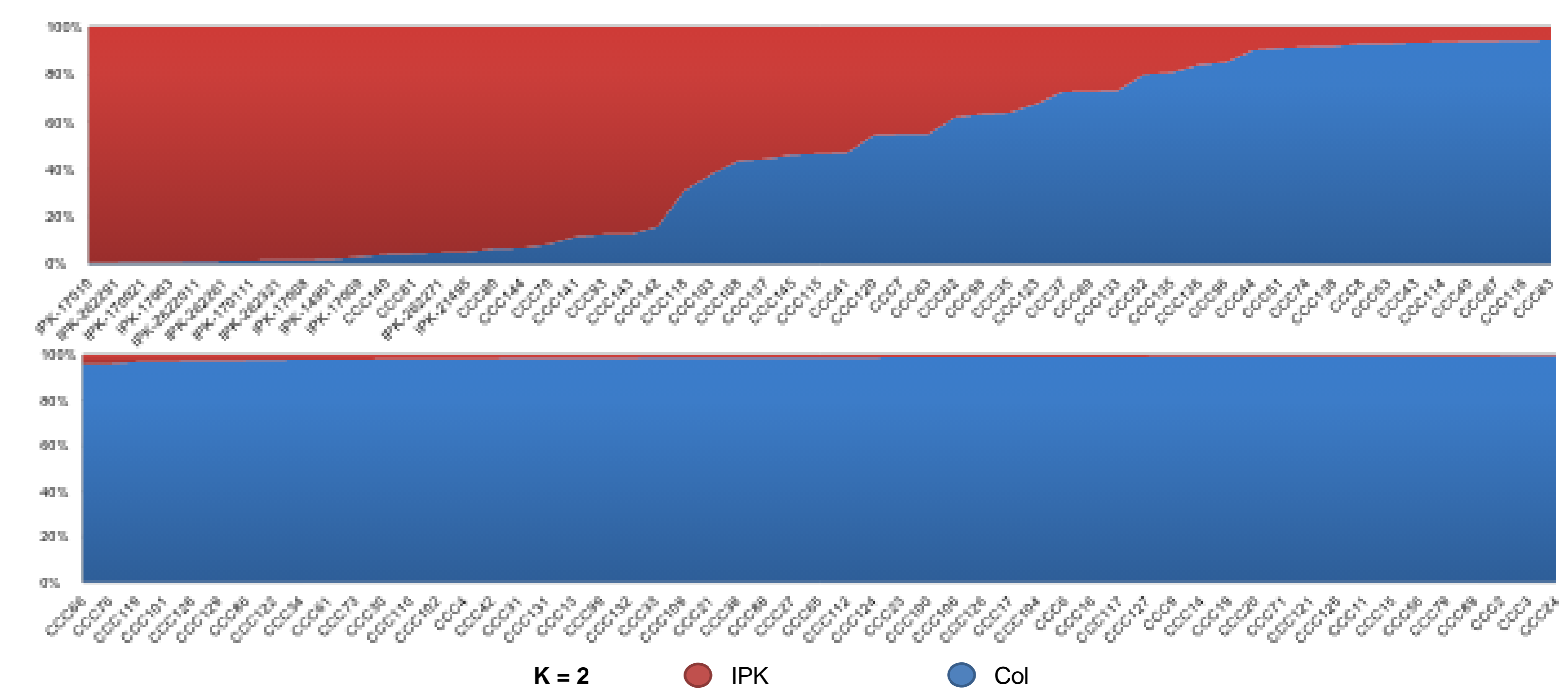


Figura 1. Representación de la asignación de genotipos Col y IPK a una población K , generados por *Structure* 2.3.2.1, bajo un modelo mezcla con frecuencias alélicas independientes. Los grupos son representados por colores, como se indica en la leyenda. Cada columna (110 en total) representa un individuo y su genotipo es dividido entre segmentos que indican su probabilidad de origen genético.

Discusión

El alto grado de diversidad genética y la falta de estructura poblacional en las poblaciones Col e IPK puede ser atribuida a la fuerte autoincompatibilidad entre genotipos de papa causada por la presencia de alelos S (esterilidad) en los gametos. femeninos y masculinos; a la alta hibridación intervarietal (2X x 2X and 2X x 4X) y flujo genético entre genotipos diploides.

Conclusiones

Las poblaciones diploides de papa tienen una alta diversidad genética posiblemente debido al alto flujo genético y la diferenciación entre individuos dentro de las poblaciones.

La falta de estructura poblacional encontrada en este estudio permite usar estas poblaciones en estudios de asociación genética.

Referencias

- Evanno, *et al.* 2005. *Ecology*, 14: 2611-2620.
- Pritchard, *et al.* 2000. *Genetics*, 155: 945-959.
- Slatkin M. 1995. *Genetics*, 139: 457-462.
- Ghislain *et al.* 2009. *Molecular Breeding*, 23: 377-388.
- Ghislain *et al.* 2006. *Theoretical and Applied Genetics* 113 (8): 1515-1527.