

**Comparison of selection methods for the establishment of a core collection using SSR
markers for hazelnut (*Corylus avellana* L.) accessions from European germplasm repositories**

Boccacci P.¹⁾, Aramini M.²⁾, Ordidge M.³⁾, van Hintum T.J.L.⁴⁾, Torello Marinoni D.⁵⁾, Valentini N.⁵⁾,
Sarraqigne, J.P.⁶⁾, Solar A.⁷⁾, Rovira M.⁸⁾, Bacchetta L.²⁾, Botta R.⁵⁾

¹⁾ Institute for Sustainable Plant Protection - National Research Council (IPSP-CNR). Strada delle Cacce, 73 -
10135 Torino, Italy

²⁾ ENEA - Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile, Via
Anguillarese 301 - 00123 S.M. di Galeria (RM), Italy

³⁾ School of Agriculture, Policy and Development, University of Reading, Whiteknights, RG6 6AR, Reading,
United Kingdom

⁴⁾ Centre for Genetic Resources, the Netherlands, Wageningen Plant Research, P.O. Box 16, 6700 AA,
Wageningen, the Netherlands

⁵⁾ Department of Agricultural, Forestry and Food Science - University of Torino, Largo Paolo Braccini, 2 - 10095
Grugliasco (TO), Italy

⁶⁾ Association Nationale des Producteurs de Noisette (ANPN), 47290 Cancon, France

⁷⁾ University of Ljubljana, Biotechnical Faculty, Department of Agronomy, Jamnikarjeva 101, 1000 Ljubljana,
Slovenia

⁸⁾ Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Centre Mas Bové, Ctra. Reus-El Morell, km 3.8 -
43120 Constantí (Tarragona), Spain

Corresponding author: Paolo Boccacci, e-mail: paolo.boccacci@ipsp.cnr.it

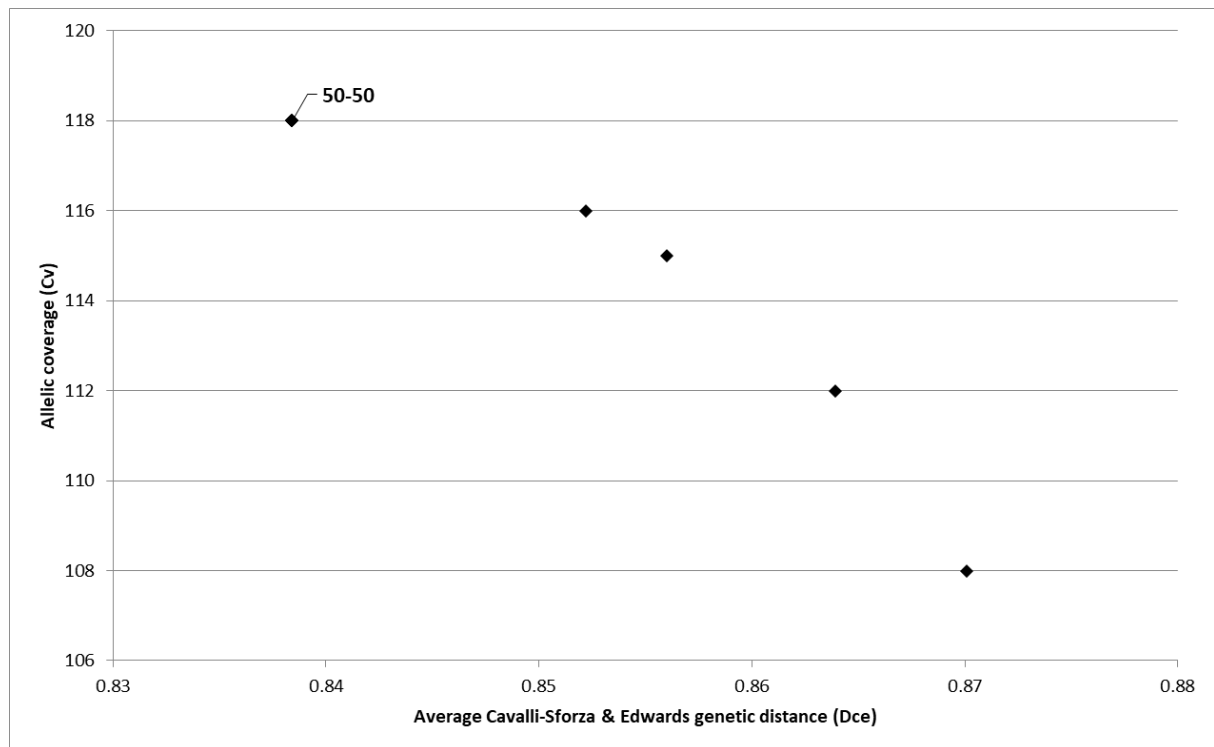
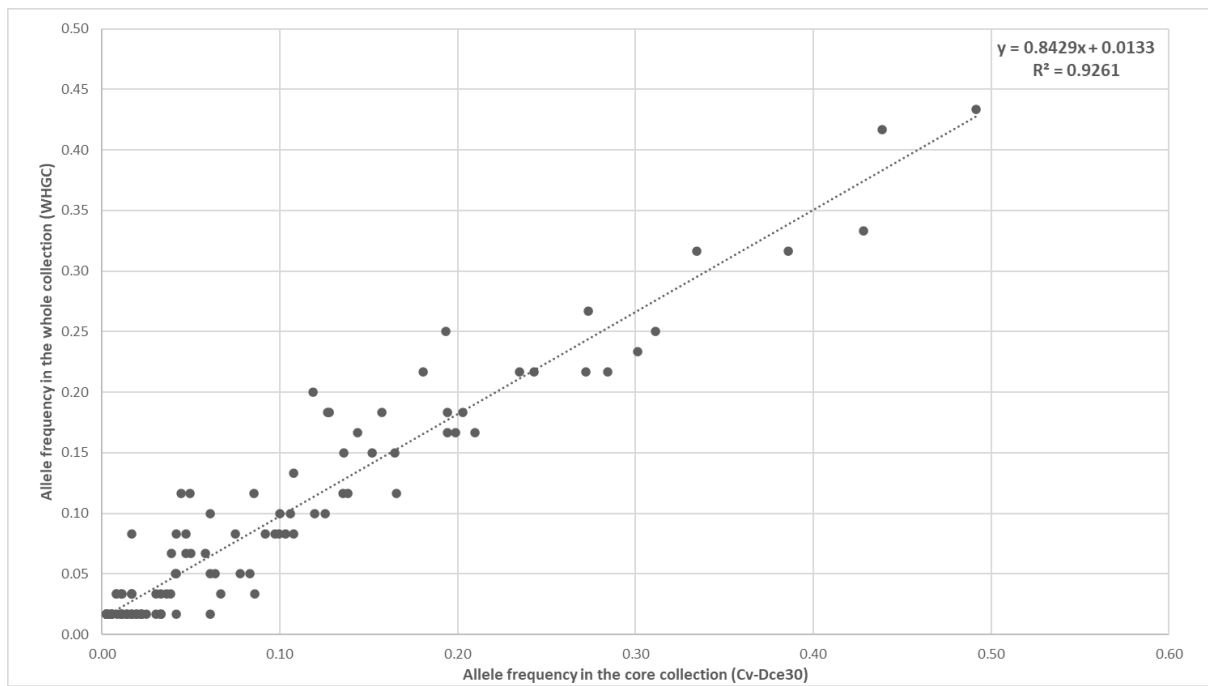


Figure S1 – Maximizing average allelic coverage (Cv) and Cavalli-Sforza and Edwards (Dce) genetic distance. Values of Cv and Dce were maximized simultaneously with respect to a weight assigned to each measure. The Core Hunter software was run independently for 10 different weight values assigned to Cv and Dce measures. When a weight of 50%, 60%, 70%, 80% and 90% was assigned to Cv and 50%, 40%, 30%, 20%, and 10% to Dce, respectively, all observed alleles were captured in the sampled subset (50-50).

1



2

3

4 **Figure S2** – Frequency distribution of the 118 alleles recovered with the final core collection Cv-Dce30
 5 (30 genotypes with Core Hunter multi-strategy) compared with the whole hazelnut germplasm
 6 collection (181 genotypes) after analysing 10 SSR loci.

7

8

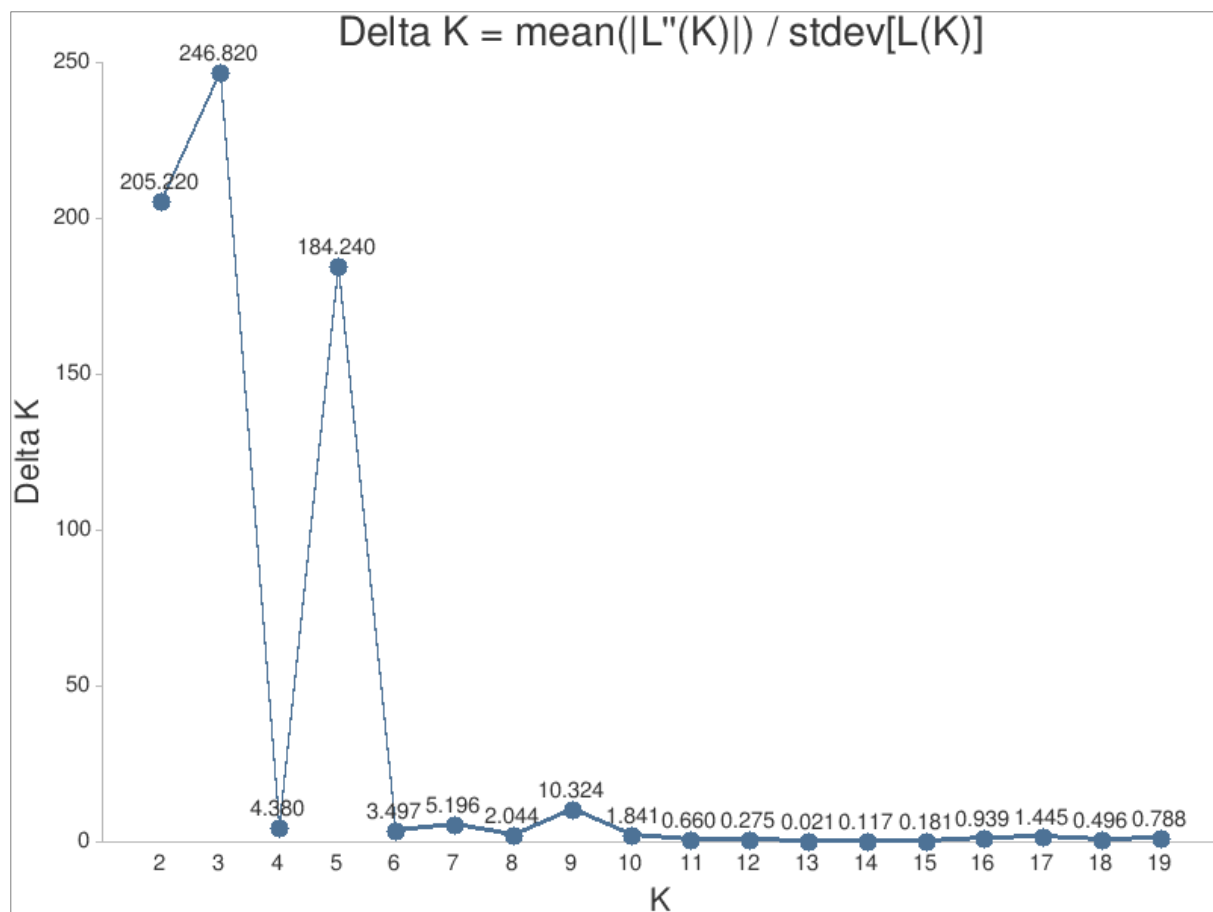


Figure S3 – The most likely K subpopulations according to ΔK method (Evanno et al. 2005). K=3 was the optimal number of genetic clusters in our set of genotypes, but a second most likely subpopulation stratification was observed at K=5.