# Genetic Diversity Analysis of *Vitex negundo* L. Using SNP Markers

**R. Mary Sujin**

Research Department of Botany, Pasumpon Thiru Muthuramalinga Thevar Memorial College,

Kamuthi, Ramnad District-623 604, Tamilnadu, India; e-mail: sujiphd@gmail.com

## Abstract

**Introduction:***Vitexnegundo*L.is one of the most important medicinal plants because it contains a set of clinically important compounds.

**Objective:**In the present study, the genetic diversity of*V. negundo*from five different geographical origins of Western Ghats was assessed through Single Nucleotide Polymorphism (SNPs).

**Methods:** To assess genetic diversity we conducted sequencing of the genomes from five populations of *V. negundo* and compared the data to a reference genome assembly.

**Results:**14 SNPs were discovered when compared the sequence of five different populations by direct partial sequencing of genes (trnL). UPGMA clustering indicated two main groups.Our genome-wide assessment of SNP variation in *V. negundo* revealed high levels of genetic variation.

**Conclusions:**This approach determining population genetic structure using SNPs from genome-wide comparison constitute a framework for high-throughput analyses genetic diversity in plants, particularly in species with genetic diversity.

**Keywords:**Genetic diversity, *Vitex negundo* L. SNP markers.

## INTRODUCTION

Determining the extent and distribution of genetic diversity is an essential component for conservation strategies. (Estoup*et al.,* 1998). Assessment of genetic diversity prevalent in the germplasm needs immediate attention for the improvement of a species (Lakhanpaul*et al.,* 2003). To understand the effective management of plant genetic diversity from a conservation point of view it is essential to consider the variation richness and distribution at intra and interspecific levels (Padmalatha and Prasad 2007). Conservation of plant genetic diversity has recently generated a lot of interest in the tropics as a result of mismanagement, adverse environment as well as socio-economic changes. Population genetic theory predicts that the decrease in the genetic diversity limits a species ability to keep pace with the changing selection pressure (Young and Merriam, 1992). Characterization of the genetic diversity and examination of

the genetic relationship among *V. negundo* are important for the sustainable conservation and increased use of plant genetic resources.

*Vitex negundo* L. (Verbenaceae), a large, aromatic shrub found throughout the greater part of India. The whole plant is used in anticancer, inflammations, antiseptic, antipyretic, diuretic, antihistamine, antioxidant, antibacterial, antinociceptive, antiandrogenic, hepatoprotective, antifertility, skin aging inhibitor and anti dopaminergic effects (Vishal, 2005). Constituents previously isolated from the plant include eight lignans, flavonones, non diterpene, pentacyclictriterpenoidsand flavonoid glycoside (Azhar*et al.*, 2006). Vitexin (Flavonoid) has potent and broad antitumor efficacy in preclinical models of ectopic growth of breast, prostate, liver and cervical cancer cells (Ying *et al.*, 2009).The extensive development of molecular techniques for genetic analysis in the past decade has led to the increase of the knowledge of plant genetic diversity. The molecular markers best suited for detecting genetic diversity should be relatively easy and inexpensive to use and should evolve rapidly enough to be variable within populations. DNA makers are valid tools for the evaluation of biological materials, both for genetic diversity studies and for the discrimination of samples (Angiocillo*et al.* (1999); Bandely*et al.* (2002); Garant and Kruuk, (2005); Morgante and Salamini, (2003). It is therefore, a natural extension for this molecules approach to be applied for medicinal plants for identifying the superior genotypes (Dal Marso*et al.*,2004; Hernandez *et al.*, 2005).

SNPs are nucleotide variations in the DNA sequence of individuals in a populationand constitute the most abundant molecular markers in the genome. Analysis of variations in plant genomes is increasingly focused on single-nucleotide polymorphism (SNP) analysis, increasing the need for fast yet reliable, simple, and cost-effective techniques to handle the large number of these polymorphisms within large plant genomes (Toni and Robert, 2003). They aresuitable for automation and can be used for a range of purposes, including rapididentification of species, construction of ultra high-density genetic maps,and association studies related to genetic disorders (Douabin-Gicquel et al.,2001). SNPs are widely distributed throughout genomes (Halushka et al., 1999),although studies have shown that the occurrence and distribution of SNPs variesamong species.

The aim of this work was to discover new SNPs and genetic diversity among five population of *V. negundo* from Western Ghats of India among SNPs markers.

## METERIALS AND METHODS

### Plant Material

Five accessions of *V. negundo* originating from different growing areas of Western Ghats of India were used in this study. Young leaves of each sample uses take in plastic bags for transport from the field to the laboratory. The samples were maintained in deep freezer at.70%.

**DNA extraction from leaves**

Genomic DNA was extracted from fresh leaves of the *V. negundo* by using a CTAB method UtaPich and Ingo Schubert (1993). After extraction, the samples were treated with RNase A for 3o min at 37 °C and run in 1 % agarose gel in TAE 1X buffer in presence of ethidium bromide (1µg/ml).

**Primer desing**

PCR primers were designed using primer 3 program (Rozen and Skaletsky, 2000) and the sequence of partial cDNA of trnL genes available from GenBank at NCBI website (http://www.ncbi.nlm.nih.gov). The sequences were amplified by polymerase chain reaction (PCR). Sequence of the trnL genes were PCR amplified for the five different accessions of *V.negundo*using the primer: trnL F: 5´ ATGTCACCACAAACAGAGACTAAAGC 3´ and trnL R: 5´ GTAAAATCAAGTCCACCRCG3´.

**PCR and Sequencing**

PCR reaction mixes were prepared for each sample by mixing 5µl of 10X PCR buffer, 2µl 10 mMdNTPs, 5 µl of 10 mM $MgCl_2$, 1µl of primer (at 10µM), 0.5 U of Taq DNA polymerase, 10 µl of diluted genomic DNA (25 ng/µl) and 25.5 µl of water. Cycling parameters were as follows: 94°C for 5 min followed by 40 cycles of 94°C for 30 second, hybridizing step for 1 min. Products were separated by agarose gel electrophoresis to check for efficiency of amplification and to ensure that only a single product of the expected size was present. PCR products were then purified by passage through Wizard$^R$ 'SV gel and PCR clean-up system purification columns and sequenced tree times from either end using the same forward primers are used in initial terminator cycle sequencing kit version 3.1 (Applied parameters) was used according to manufacturer's instructions.

**Marker discover strategies**

The most direct approach to the discovery of DNA polymorphisms is direct sequencing of PCR products from five populations of *V. negundo*. Polymorphisms between the sequences were identified by sequence alignment. Furthermore, the trace outputs from the sequences were evaluated by eye to identify possible areas of heterozygous sequence.

**Genetic diversity and Data analyzes**

For each SNP marker two alleles are genetically present yielding five populations of *V. negundo*. Allelic and genotype frequencies for each marker were estimated by simple counting. Power of discrimination (PD) was calculated using following formula (Kloosterman*et al.*, 1993).

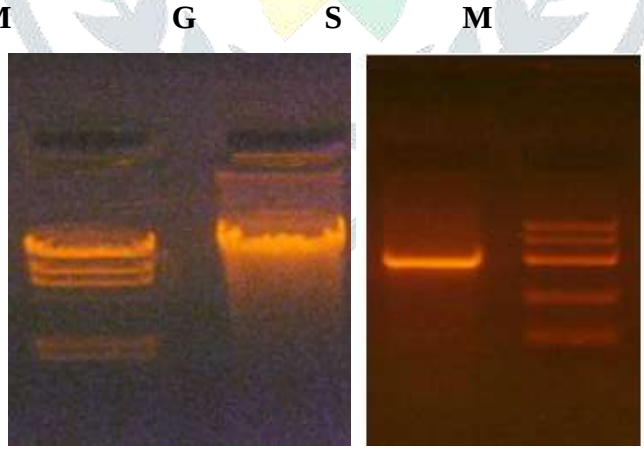
$$PD=1-\sum_{i=1}^{g} f_i^2$$

Where $f_i$ is the frequency of the $i$th genotype and the sum is overall genotypes. The SNP genotypes were then recorded in a 0/1 data matrix, where 1 indicates the presence of a given allele and o its absence. So each SNP was transformed into two columns with three possible states: 1/1 for heterozygote and 0/1 or 1/0 for the two homozygous genotypes.Jaccard's similarity coefficient was calculated to measure the genetic variability between five populations (Jaccard, 1908). Based on the genetic similarity matrix, the populations were clustered by the unweighed pair group method with arithmetic averaging (UPGMA) using the program NTSYS –pc version 2.1 (Rohlf, 1999).

## RESULTS

### Targeted sampling

The total length of high quality sequences analyzed obtained for the trnL genes was 495 bp. Comparison of these sequences to the cDNA original sequences from which the primers were designed by BLAST2 seq (Tatusova and Madden, 1999). SNP discovery was validated using direct sequencing of PCR products and comparative analysis of sequences of five different populations of *V. negundo* (Fig 1).Itshaving SNPs or insertions/deletions were detected in an allele-specific manner.Fourteen SNPs were finally identified within the sequences of the genes of five different accessions (Table 1& Fig 2).




**M- Marker; G- Genomic DNA; S- PCR product**

**Figure 1.**

Genomic DNA and PCR products of *V. trifolia* using SNP primer

**Table 1.**

Position and types of the three SNPs discovered

| Gene | SNP position | Type | Accessions |
|---|---|---|---|
| trnL 1 | 2 | A/G | $V_4$, $V_5$ |
| trnL 2 | 4 | C/G | $V_5$ |
| trnL 3 | 7 | T/G | $V_5$ |
| trnL 4 | 10 | T/G | $V_4$ |
| trnL 5 | 11 | A/G | $V_4$ |
| trnL 6 | 13 | G/T | $V_4$ |
| trnL 7 | 15 | A/G | $V_4$ |
| trnL 8 | 159 | C/T | $V_2$ |
| trnL9 | 417 | G/T | $V_4$ |
| trnL 10 | 479 | A/G | $V_4$ $V_5$ |
| trnL 11 | 485 | T/C | $V_4$ $V_5$ |
| trnL 12 | 487 | G/C | $V_1$ |
| trnL 13 | 494 | T/A | $V_3$ |
| trnL 14 | 494 | T/G | $V_1$ |

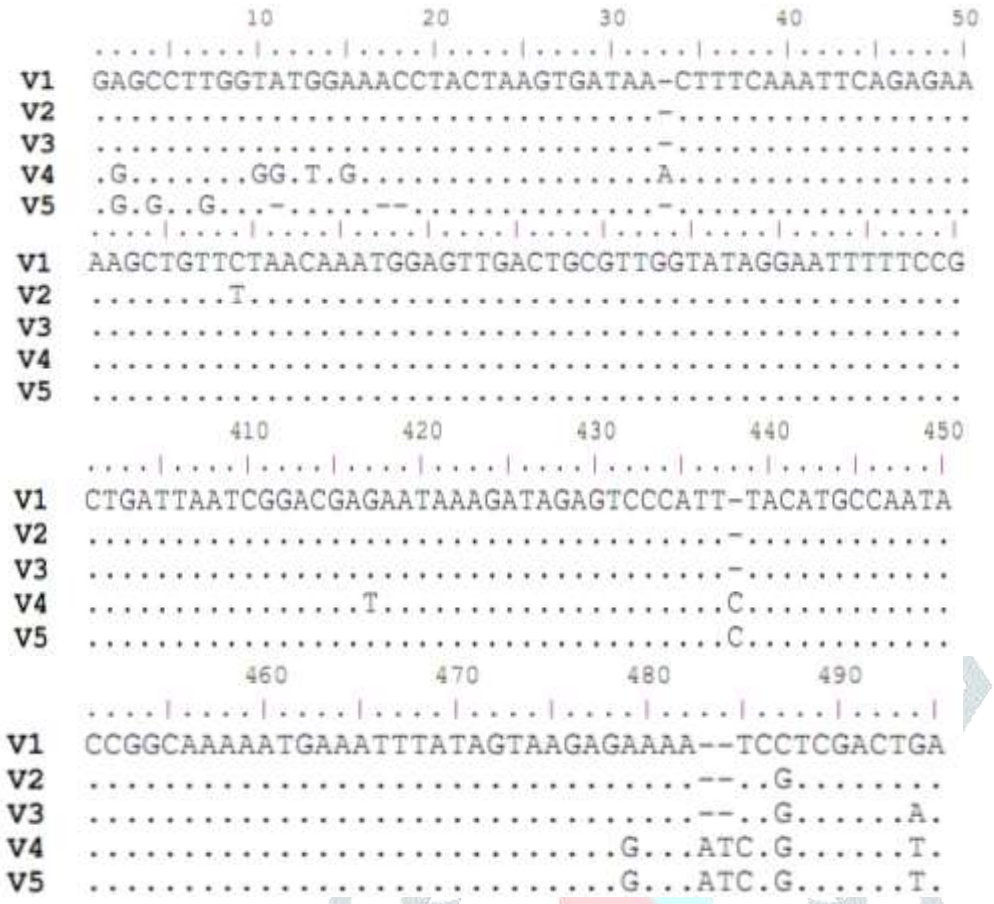



**Figure 2.**

SNP position in the sequence comparison using the program

**Marker characteristics**

The genotype and allelic frequencies of the three SNPs and their power of discrimination (PD) varies from 0.45 to 0.82 for the marker rbcL with an average value of 0.66.The observed heterozygoisty (*Ho*) ranged from 0.21 to 0.54. The genotypic frequencies and allelic frequencies of each SNPs were varied (Table 2).

**Table 2.**

Features and frequencies of new SNPs

| SNP | *Ho* | PD (%) | Genotypic Frequencies (%) | | | Allelic Frequencies (%) | |
|---|---|---|---|---|---|---|---|
| trnL 2 | 0.35 | 62 | 28.57A/A | 57.14G/G | A/G | 56.4:A | 47.9:G |
| trnL 4 | 0.28 | 73 | 14.28C/C | 64.28G/G | C/G | 43.5:C | 78.3:G |
| trnL 7 | 0.40 | 82 | 35.71T/T | 42.85G/G | T/G | 58.9:T | 53.6:G |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| trnL 10 | 0.29 | 74 | 28.57T/T | 35.71/GG | T/G | 42.7:T | 68.9:G |
| trnL 11 | 0.56 | 79 | 21.42A/A | 50.00G/G | A/G | 62.4:A | 73.1:G |
| trnL 13 | 0.34 | 45 | 21.42G/G | 42.85T/T | G/T | 43.8:G | 48.2:T |
| trnL 15 | 0.23 | 57 | 57.14A/A | 42.85G/G | A/G | 72.4:A | 68.1:G |
| trnL 159 | 0.38 | 63 | 57.14C/C | 28.57T/T | C/T | 54.3:C | 54.1:T |
| trnL 417 | 0.21 | 72 | 35.71G/G | 21.42T/T | G/T | 68.9:G | 59.2:T |
| trnL 479 | 0.54 | 64 | 57.14A/A | 64.28G/G | A/G | 57.8:A | 81.2:G |
| trnL 485 | 0.48 | 73 | 42.85T/T | 28.57C/C | T/C | 79.3:T | 78.2:C |
| trnL 487 | 0.47 | 81 | 78.57G/G | 50.00C/C | G/C | 65.8:G | 63.4:C |
| trnL 494 | 0.39 | 75 | 14.28T/T | 42.85A/A | T/A | 35.9:T | 63.7:A |
| trnL 494 | 0.32 | 78 | 35.71A/A | 21.42G/G | T/G | 48.9:T | 38.9:G |

**Genetic diversity level**

The three SNP markers in *V. negundo* from five populations and assess their potential in studying genetic diversity. The lower coefficient of similarity (GS=0.20) was observed between the populations of $V_1$ and $V_4$ and the highest similarity (GS=1) occurred between $V_2$& $V_3$ (Table 3).

**Table 3.**

Similarity indices of five population of *V. negundo* using SNP markers

| **Pop** | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|---|
| $V_1$ | 1.00 | | | | |
| $V_2$ | 0.66 | 1.00 | | | |
| $V_3$ | 0.66 | 1.00 | 1.00 | | |
| $V_4$ | 0.20 | 0.22 | 0.22 | 1.00 | |
| $V_5$ | 0.28 | 0.33 | 0.33 | 0.23 | 1.00 |

**Patterns of genetic diversity**

The present study revealed a relatively low level of genetic diversity in *V. negundo*based on SNP markers. Using UPGMA algorithm a dendrogram was constructed to infer phylogenetic relationships between the five populations (Fig**.** 3). A Jaccard's matrix was used to produce a dendrogram based on SI, which showed distinct separation of the five accessions into two major cluster having 23 % similarity. Among the two major clusters, the accessions belonging to the lowest cluster (LC) were collected from $V_4$while accessions $V_1$, $V_2$, $V_3$and $V_5$ belongs to the upper cluster (UC) had 31 % similarity. Further, the accessions of the UC were grouped in to two major sub clusters USC1 from $V_1$,$V_2$ and $V_3$ having 65 % similarity and USC2 from $V_5$. The lower cluster (USC1) was divided into two sub cluster LSC1 and LSC2. The upper sub cluster LSC1 was further sub divided into two (USC1A) from $V_1$and (USC1B) from $V_2$ and $V_3$ with a similarity of 100 %

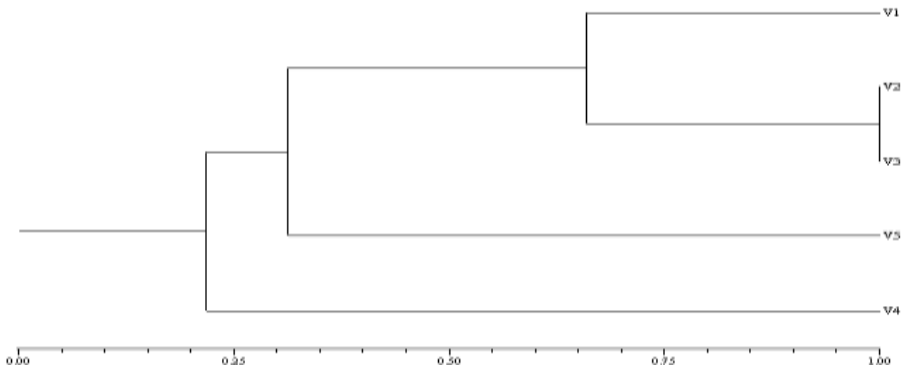



**Figure 3.**
Dendrogram of *V. trifolia* using Jaccard similarity coefficients from SNP markers and UPGMA method

## DISCUSSION

In general, the present study shows that populations of *V. negundo* possess a high level of genetic diversity. The level of genetic diversity was determined by the mean of heterozygosity and percentage of polymorphic loci. According to electrophoretic surveys by Nei and Roychoudhury (1974), Nei (1978), and Gorman and Renzi (1979), a large number of loci should be examined if the number of individuals per locus is small. In fact, a few individuals are sufficient for estimating heterozygosity if a sufficiently large number of loci are examined (Gorman and Renzi, 1979). Although it is difficult to compare the data from electrophoretic and SNP-based studies, preliminary data from recent electrophoretic assays (Norfiza et al., 2001) support the conclusion that SNP markers provide more sensitive assays of genetic diversity.

Cluster analysis between populations of *V. negundo* revealed two major groups, with the Karaiyar population being closely related to that from Maruthamalai population could be explained by its geographical isolation which is expected to limit gene exchange between populations. This observation suggests that SNPs could be employed as molecular markers for predicting whether a particular accession of *V. negundo*

Our assessment of genome wide diversity of *V. negundo*suggests that it has high level of genetic diversity and structure for all populations. Furthermore, our data suggest that SNPs may be necessary for better resolution of relationships of samples among populations.The power of SNP discovery should not be misconstrued as an indication of diversity in a species that shows overall genetic diversity; Although chloroplast markers have been effectively used for studying plant distributions, low effective population size in chloroplast DNA and reduced genetic diversity, compared with nuclear DNA. Chlorplast SNPs are more variable and amenable to high throughput genotyping and will likely be the marker of choice for population-level analyses of species with sequences genomes (Estoup*et al.*, 2002).

From this study, we have identified 14 SNPs that can be used as genetic markers in *V. negundo*. Cluster analysis showed that diversity among the different accessions of *V. trifolia* corresponds well with the geographic origins of each population. Similar work was done in *Eurycomalongifolia* by Asiah*et al.*(2003); in *Ricinuscommunis* by Jeffrey *et al.* (2010); in Oat by Shiaomon*et al.* (2011); *Panaz ginseng* by Hongtao*et al.*(2010). These markers should prove useful in preserving genetic diversity among populations. The distinction between the populations could be explained by its geographical isolation, which is expected to limit gene exchange between populations (Asiah*et al.*, 2003).

## CONCLUSIONS

Our study demonstrates the utility of a SNP-based approach for assessing the population genetics of medicinal plants. From this study, we have identified 14 SNPs that can be used as genetic markers in *V. negundo*. Cluster analysis showed that diversity among the different accessions of *V. negundo* corresponds well with the geographic origins of each population. These markers should prove useful in preserving genetic diversity among domesticated populations of *V. negundo*. These SNPS may also be developed as predictive markers for useful phenotypes.As new sequencing technologies emerge and more genome become more available, our approach promised to be particularly useful for plant population studies due to the resolving power of SNPs and the ability to rapidly assess diversity in a large number of samples. This study represents one of the most extensive genomic studies of worldwide SNP variation in this medicinal plant. With rapidly increasing capabilities in genome sequencing, this work provides a template for assessing population structure. These results indicate that SNP markers would provide a rapid and easy way to establish a fingerprint of each population for genetic variability.

## ACKNOWLEDGEMENT

The authors are grateful to the Indian Council of Medical Research, Government of India, New Delhi (Ref: 59/12/2006/BMS/TRM dt. 26.03.2009) for financial support.

## REFERENCES

1. Angiolillo, A., Mencuccini, M. and Baldoni, L. (1999). Olive genetic diversity assessed using amplified fragment length polymorphisms. *Theoret.Appl.Genet.* 10:411-421.
2. Anonymous (2003). The wealth of India – Raw Materials, Vol. X, Council for Scientific and Industrial Research. New Delhi. 525.
3. Asiah Osman, Barbara Jordan, Philip, A.,Lessard, Norwati Muhammad, M.,RosliHaron, Norifiza Mat Riffin, Anthony, J.,Sinskey, ChoKyunRha, and David Housman, E. (2003). Genetic Diversity of *Eurycomalongifolia* Inferred from Single Nucleotide Polymorphisms. *Plant Physiol.*131(3):1294–1301.

4. Bandelj, D., Jakse, J. and Javornik, B. (2002). DNA fingerprinting of olive varieties by microsatellite markers. *Food Technol. Biotechnol.* 5:185-190.
5. Brumfield, R.T., Beerli, P., Nickerson, D.A. and Edwards, S.V. (2003). The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology & Evolution.* 18:249-256.
6. Consolandi, C., Palmieri, L., Doveri, S., Maestri, E., Marmiroli, N., Reale, S., Lee, D., Baldoni, L., Tosti, N., Severgnini, M., De Bellis, G. and Castiglioi, B. (2007). Olive variety identification by ligation detection reaction in a universal array format. *Afr. J. Biotechnol*. 9:565-574.
7. Dalmasso, A., Fontanella, E., Piatti, P., Civera, T., Rosati, S. and Bottero, M.T. (2004). A multiplex PCR assay for the identification of animal species in feedstuffs. *Mol.Cell.Probes.* 6:81-87.
8. Estoup, A. And Angers, B. (1998). Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. *Advances in Molecular Ecology.*Amsterdam:IOS Press Carvalho GR. 55-86.
9. Estoup, A., Jame, P. and Comuet, J.M. (2002). Homoplasy and muatation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*. 11:1591-1604.
10. Garant, D. and Kruuk, L.E. (2005). How to use molecular marker data to measure evolutionary parameters in wild populations. *Mol. Ecol.* 15:1843-1859.
11. Hernandez, M., Esteve, T. and Pla, M. (2005). Real-time polymerase chain reaction based assays for quantitative detection of barley, rice, sunflower and wheat. *J.Agric.Food* Chem. 6:7003-7009.
12. Hongtao Wang, Hua Sun, Woo-Saeng Kwon, Haizhu Jin and Deok-Chun Yang (201). A PCR based SNP marker for specific authentication of Korean gingeng (Panaxgingeng) cultivar “Chunpoong”. Molecular Biology Reports. 37(2):1053-1057.
13. ImenRekik Hakim, NazhihaGratiKammoun, EmnaMakhloufi and Ahmed Rebai (2010). Discovery and potential of SNP markers in characterization of Tunisian olive germplasm. *Diversity.* 2:17-27.
14. Jaccard, P. (1908). Nouvelles recherché sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat.* 44:223-270.
15. Jeffrey Foster, T., Gerard Allan, J., Agnes Chan, P., Pablo Rabinowicz, D., Jacques Ravel., Pai; Jackson, J. and Paul Keim (2010). Single nucleotide polymorphisms for assessing genetic diversity in castor bean (Ricinuscommunis). *BMC Plant Biology*. 10 (13):1-11.
16. Kannathasan, K., Senthilkumar, A., Chandrasekaran, M. and Venkatesalu, V. (2007). Differential larvicidal efficacy of four species of *Vitex*against *Culexquinquefasciatus* larvae. *Parasitol. Res.* 101:1721-1723.
17. Kloosterman, A.D., Budowle, B. And Daselaar, P. (1993). PCR- amplification and detection of the human DIS80 VNTR locus. Amplification conditions, population genetics and application in forensic analysis. *Int.J.Legal Med.* 7:257-264.

18. Manjunatha, B.K., Vidya, S.M., Krishna, V., Mankani K.L., Singh, S.D. andManohara, Y.N. (2007). Comparative evaluation of wound healing potency of *Vitextrifolia*L. and *Vitexaltissima*L. *Phytother. Res.* 21:457-461.
19. Morgante, M. and Salmini, F. (2003). From plant genomics to breeding practice. *Curr.Opin.Biotechnol.* 5:214-219.
20. Palmieri, L., Vietina, M., Bottura, C., Donini, P. and Nelson, M. (2004). SNP development in olive and their application on varietal determination in olive oil. In Proceedings of the XLVIII It Joint Meeting, Lecce, Italy, September 15-18.
21. Rabinowicz, P.D., Schutz, K., Dedhia, N., Yordan, C., Parnell, L.D., Stein, L., McCombie, W.R. and Martienssen, R.A. (1999). Differential methylation of genes and retrotransposons facilities shotgun sequencing of the maize genome. *Nature Genetics*. 23:305-308.
22. Reale, S., Doveri, S., Diaz, A., Angiolillo, A., Lucentini, L., Pilla, F., Martin, A., Donini, P. and Lee, D. (2006). SNP-based markers for discriminating olive (*Oleaeuropaea* L.) cultivars. *Genome*. 15:1193-1209.
23. Rohlf, F.J. (1999). NTSYS-PC: Numerical Taxonomy and Multivariate Analysis System (Version 2.1), Exeter Software, Setauket, NY, USA.
24. Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 21:365-386.
25. Shiaoman Chao, Rebekah Oliver, Gerard LAzo, Nick Tinker, Jean-Luc Jannink, Rachel Redman and Eric Jackson (2011). Analysis of genetic diversity using SNP markers in Oat. Plant and Animal XIX conference. Jan.15-19. Town & country convention center. San Diego, CA.
26. Tatusova, T.A. and Madden, L.T. (1999). BLAST 2 sequences. A new tool for comparing protein and nucleotide sequences. *FEMSMicrobiolLett.* 3:247-250.
27. Tsuchihashi, Z. And Dracp[p;o, N.C. (2002). Progress in high-throughput SNP genotyping methods. *The Pharmacogenomics Journal.* 2:103-110.
28. UtaPichand Ingo Schubert (1993).Midiprep method for isolation of DNA from plants with a high content of polyphenolics. *Nucl. Acids Res.* 21(14):3328.
29. Vishal Tandon R. Medicinal uses and biological activities of *Vitex negundo*. Nat. Prod. Rad., 2005;4:162-165.
30. AzharUlHaq, Malik A, Khan MT H, Anwar UlHaq MTH, Khan SB, Ahmad Aand Choudhary MI.Tyrosinaseinhibitorylignans from the methanol extract of the roots of *Vitex negundo* L. and their structure-activity relationship. Phytomed.,2006;13(4):255-260.
31. Ying Jun Zhou, Yiliang Ellie Liu, JianGuo Cao, GuangYaoZeng, Cui Shen, YanLan Li, MeiChenZhou,Yiding Chen,WeipingPu, Louis Potters and Eric Shi Y.Vitexins, nature-derived lignan compounds, induce apoptosis and suppress tumor growth. Clin Cancer Res., 2009;15(16):5161-5169.