**Perspective in Plant taxonomy through conventional to contemporary biotechnological approach**

**Sapna Tamang & Saurav Moktan\***

Department of Botany

University of Calcutta

35, B.C. Road, Kolkata, West Bengal, India

\*Email: smbot@caluniv.ac.in

**ABSTRACT**

Taxonomy, the foundation of conservation efforts, has evolved over centuries, aiming to identify, name, and categorize species. However, the complexity of morphology in plant species has challenged taxonomists, leading to an increased reliance on modern technologies. Contemporary plant taxonomy benefits significantly from DNA barcoding, next-generation sequencing, and AI-based plant identification tools, enhancing biosystematics and ecological surveillance. Molecular techniques, such as DNA markers, barcodes, and sequencing, play a crucial role in studying phylogenetics, reconstructing evolutionary histories, and enriching plant taxonomy. Conventional markers like RFLP, AFLP, RAPD, SSR and SNP, as well as modern markers, offer complementary insights, enabling a more comprehensive understanding of plant diversity and evolutionary relationships. The integration of data from conventional and contemporary biotechnological tools provides a powerful approach to address and enhance the challenges of plant taxonomy.

**Keywords-** Taxonomy; DNA barcoding; Phylogenetics; Molecular markers; Next-generation sequencing; Artificial Intelligence

1. **INTRODUCTION**

Taxonomy is the foundation for all conservation. For more than 250 years, the objective of taxonomy has been to identify, name, and categorize species [1].The taxonomists' ability to make decisions may be hampered by the fact that a small number of species from various populations have been found to exhibit complicated morphology. Thus, the majority of the time spent by botanists is in manually examining and determining the characteristics of various plant species. So, the modern techniques for studying the plants have been substantially enriched by a number of technology used in contemporary plant taxonomy. For the study of biosystematics, the DNA barcoding, next-generation sequencing, AI related plant identification are essential technological tools. Similarly, different artificial systems have been used to identify plants, and it has been discovered that the performances of automated plant identification systems are highly encouraging and may pave the way for a new generation of ecological surveillance systems [2]. At the same time, the molecular techniques in the study of phylogenetics of plants have a significant role in plant taxonomy. Thereby, molecular approaches could be used to reconstruct the evolution of organisms and improve their taxonomy [3].

The use of molecular taxonomic techniques for vegetation surveys has the potential to reduce the effects of the taxonomic obstacle and boost the effectiveness of conservation initiative [4]. To solve the morphological complexity problem, the genetic data are also extremely encouraging to be employed in higher plant systematics [5]. Recent advancement of the different varieties of molecular markers, DNA markers, DNA barcodes and different DNA sequencing techniques plays a pivotal role in the study of plants systematics. Every botanist, including molecular biologists and plant hunters, who still have much to offer, find molecular systematics to be extremely relevant for the study of plants [6]. Markers like RFLP, AFLP, RAPD, SSR, and SNP are considered traditional markers because they have been widely used in plant taxonomy and genetics for several decades. Modern markers offer advantages in terms of throughput, genome-wide coverage, and accuracy, making them essential tools for advancing our understanding of plant diversity, evolution, and taxonomy. Combining data from both conventional and modern markers can provide a more comprehensive frameowrk on plant taxonomy and evolutionary relationships.

1. **DNA BARCODING AND SPECIES AUTHENTICATION**

Emile Zuckerkandl and Linus Pauling presented the idea of using molecular data for phylogenetic inference during early sixties. They proposed that DNA and protein sequences may be utilized as markers for evolutionary history [7]. A year later, Carl Woese, compared the 18S rRNA sequences from various organisms and discovered that they exhibit significant differences. This led him to propose a new classification of life, with three domain concept: Bacteria, Archeae, and Eukarya [8]. This work revolutionized our understanding of the tree of life and the evolutionary relationships between organisms. Woese's research laid the groundwork for DNA barcoding. Phylogenetic analysis and species identification frequently employ the technique of DNA barcoding. It is predicted on amplifying of brief, conserved genomic regions with sufficient variance to distinguish between species with little intraspecific variation.

Previously, [9], an international group of plant systematists made the initial attempt to rearrange the orders and families of flowering plants into a phylogenetic system based on the molecular analysis of *rbcL*, *atpB*, and 18S rDNA genes. Later, the term DNA barcoding was coined and popularized by Paul Hebert and his colleagues with their foundational study on mitochondrial gene Cytochrome C Oxidase subunit I (COI) [10]. The majority of modern computational barcoding techniques have made an effort to include known modelling strategies from molecular phylogenetic research. Conventional barcoding techniques are essentially tree-based evolutionary systems in which identification choices are determined using the tree-induced distances [10, 11]. A Consortium for the Barcode of Life (CBOL), which currently includes more than 120 organizations from 45 countries, was founded in May 2004 to advocate the use of DNA barcoding for all eukaryotic life on this planet [12]. DNA barcoding technique is a useful tool for analyzing small amounts of plant data to identify the species and genus of a given plant [13]. Over 5000 angiospermic taxa have sequenced *rbcL,* the major subunit of ribulose-bisphosphate carboxylase-oxygenase, which is encoded by a plastid gene. The number of species included in published analyses has reached 2230 numbers [14]. These relatively short sequences (650 symbols in the case of mtDNA) serve as identifiers for determining the species identification through the use of mtDNA [15]. A useful tool for conducting vegetation surveys, the multi-marker DNA barcoding method using *rbcL, matK*, and *trnH-psbA* may drastically cut down the time and expense required to identify different species [16]. Reference [17] expands the use of DNA barcoding in the field of medicinal plants and helps phylogenetic research by investigating the use of the DNA barcode ITS2 to identify medicinal plants for the first time. Polymerase Chain Reaction (PCR) is used to amplify a highly variable region, such as the DNA barcode region of the nuclear, chloroplast, or mitochondrial genome. Nuclear DNA, chloroplast DNA (Figure 1) (e.g. *rbcL, trnL-F, matK, psbA, trnH, psbK*), and mitochondrial DNA (e.g. COI) are regions that are frequently utilized for DNA barcoding [18].Similarly, [19] discusses that the DNA barcoding technique with ITS2 region is a potential DNA marker for authentication of selected plants. The phylogenetic study proposed by [20] reveals that the barcode sequences *psbK-psbI, atpF-atpH,* andITS2 had a greater resolution at the species level.

1. **DNA Metabarcoding**

Metabarcoding relies on particular locus selection criteria for isolating distinct taxa from mixed data. The inclusion of numerous loci, each optimized to capture species-level resolution in different taxa (e.g., *matK/rbcL* for plants, ITS for fungi), would be necessary for metabarcoding to be genuinely universal. In contrast, a single bulk sample utilized in metabarcoding precludes such treatment, but several locus sequencing data can be merged for analysis of individual specimens in barcoding to improve taxonomic resolution [21]. Reference [22] demonstrated that DNA metabarcoding with nrITS2 marker improved taxonomic resolution for aerobiological pollen samples thereby showing better correlation and spatiotemporal patterns for airborne pollen trends using nrITS2, making it the preferred marker for molecular airborne pollen monitoring. The enormous potential of ultra-barcoding is tackling difficult plant taxonomy problems and for discovering cryptic species in taxonomically difficult plant taxa [23].

1. **Microfluidic enrichment barcoding (MEBarcoding)**

DNA barcoding called Microfluidic Enrichment Barcoding (MEBarcoding) is an effective substitute for conventional PCR and Sanger sequencing for producing huge numbers of plant DNA barcodes and creating more complete barcode databases. During a single thermal cycling technique, it simultaneously amplifies specific regions for 48 DNA samples and hundreds of PCR primer pairs using the Fluidigm Access Array (generating up to 23,040 PCR products) [24]. Using microfluidic PCR and high-throughput sequencing (HTS), the researchers sequenced 576 samples from plant species across 96 target locations to produce a significant amount of sequence data for phylogenetic studies. The research was done on South American lineage of the genus *Bartsia* under familyOrobanchaceae [25].



**Figure 1: Barcoding loci in ITS region of rRNA, cpDNA and mtDNA [18]**

**III. BIOINFORMATIC DATABASES**

One of the earliest and most notable attempts to launch bioinformatic databases for plant DNA barcoding was made by the Consortium for the Barcode of Life (CBOL). CBOL was formed in 2004 with the aim of promoting the use of DNA barcoding in global standard for identification and biodiversity research [12]. Online databases, like GenBank, NCBI, and BOLD stores vast amounts of genetic and taxonomic information, allowing researchers to access and analyze data for their taxonomic studies. Similarly, TIGR Plant Repeat Databases provide a resource for locating, categorizing, and analyzing repetitive sequences in 12 plant genera and four plant families, despite the fact that repetitive sequences in plants can obstruct genome annotation and sequencing efforts [26]. Current computational approaches to barcoding are more scalable and interpretable as a result of newly created alignment-free methods for DNA barcoding that can quickly and accurately identify specimens by analyzing only a small number of barcode features [27].

The composition vector (CV) technique has been demonstrated to be a trustworthy and quick alignment-free way to analyze big COI (cytochrome C oxidase) barcoding datasets. The CV method is also effective for analyzing huge multi-gene datasets for plant DNA barcoding [28]. Methods that most directly address the issue of barcode-based identifications have lately come into existence. Table 1 depicts the details about the available DNA barcoding tools.

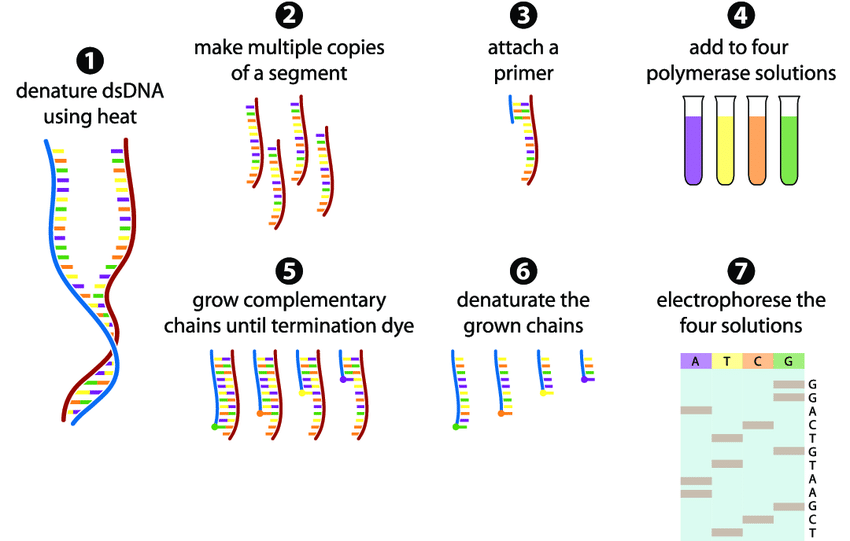
**Table 1: Bioinformatic database tools and their web address [18]**

|  |  |  |  |
| --- | --- | --- | --- |
| **Tools** | **Launch** | **Method** | **Available at** |
| Taxl | 2005 | Distance based | [Axei.meyer@uni-konstanz.de](mailto:Axei.meyer@uni-konstanz.de) |
| CBCAnalyzer | 2005 | Phylogenies based on CBC | http://cbcanalyzer.bioapps.biozentrum.uni  -wuerzburg.de/cgi-bin/index. |
| 4SALE | 2006 | RNA alignment and editing | http://4sale.bioapps.biozentrum.uniwuerzburg.  de/ |
| CodonCode Aligner | 2007 | Codon based | <http://www.codoncode.com/index.html> |
| BPSI | 2008 | Back Propagation neural networks | [zhangab2008@yahoo.com.cn](mailto:zhangab2008@yahoo.com.cn) |
| SAP | 2008 | Bayesian phylogenetics | http://ib.berkeley.edu/labs/slatkin/munch/  StatisticalAssignmentPackage.html |
| CAOS | 2008 | Character Based | <http://sarkarlab.mbl.edu/CAOS> |
| TaxonGap | 2008 | Operational Taxonomic Unit (OTU) based | http://www.kermit.ugent.be/software.php?  navigatieId=37&categorie  Id=17 |
| BioBarcode | 2009 | Sequence based | <http://www.asianbarcode.org> |
| BLOG | 2009 | Data mining approach | <http://dmb.iasi.cnr.it/blog-downloads.php> |
| B | 2010 | Sequence quality and contig overlap | http://www.nybg.org/files/scientists/dlittle  /B.html |
| OFBG | 2010 | Spp. Discrimination using oligonucleotide frequencies | <http://www.nbri.res.in/ofbg.php> |
| OTUBase | 2011 | Operational Taxonomic Unit based | http://www.bioconductor.org/packages/rel  ease/bioc/html/OTUbase.html |
| jMOTU | 2011 | Multiple Operational taxonomic Unit (MOTU) based | http://www.jmotu.com-about.com/ |
| TAxonerator | 2011 | OTU and taxonomy data based | http://www.taxonnerator.com-about.com/ |
| CLOTU | 2011 | Amplicon and taxa based | http://www.mn.uio.no/ibv/bioportal/ |
| Eco Primers | 2011 | Barcode markers and primer based | http://www.grenoble.prabi.fr/trac/ecoPrim  ers |
| PTIGS- ldit | 2011 | psbA-trnHintergenic Spacer (PTIGS) based | <http://psba-trnh-plantidit.dnsalias.org> |
| BRONX | 2011 | Sequence Identification Incorporating Taxonomic Hierchy | http://www.nybg.org/files/scientists/dlittle  /BRONX.html |
| Spider | 2012 | Analysis of species identity and evolution | http://spider.r-forge.rproject.  org/SpiderWebSite/spider.html |
| ISHAM | 2013 | Mycological classification | http://www.isham.org/ |
| LV barcoding | 2013 | Locality sensitive hashing-based | http://msl.sls.cuhk.edu.hk/vipbarcoding/ |
| Excali BAR | 2014 | Calculate intra- and interspecific distances | http://datadryad.com/resource/doi:10.5061  /dryad.r458n |
| VIP Barcoding | 2014 | Vector-based software | http://msl.sls.cuhk.edu.hk/vipbarcoding/ |
| Q-Bank | 2015 | Identification and detection reference database | http://www.q-bank.eu/ |
| Obitools package | 2015 | NGS data based | <http://metabarcoding.org/obitools> |

1. **CONVENTIONAL DNA SEQUENCING**

Reference [29] introduced the dideoxynucleotide sequencing technique also known as Sanger method of DNA sequencing or first generation sequencing. The technique was simpler and quick and it replaced other DNA sequencing techniques in the vast majority of applications (Figure 2). Subsequently, other enzymatic sequencing methods were devised including partial ribosubstitution [30] the plus and minus method of Sanger [31] and the chemical cleavage method end-radio-labeled DNA fragments [32].

Numerous nuclear, mitochondrial, and chloroplast genes have been used to examine sequence variation at the genus level. The success of species-level assignment of plants using Basic Local Alignment Search Tool (BLAST) [33] with individual barcodes was obtained with *matK* (99%), followed by *trnH-psbA* (95%) and then *rbcL* (75%). Use of these three-locus DNA barcode resulted in >98% correct identifications of 296 species of woody trees, shrubs and palms [34]. Recently, a group of plant DNA barcode researchers proposed two genes, *rbcL* and *matK,* taken together, as appropriate for barcoding of plants (CBOL). Gene mapping, QTL analysis, map construction, and marker-assisted selection all relied increasingly on the use of molecular markers [35].



**Figure 2:  Sanger sequencing for application in phylogenetic analysis [36]**

1. **Next Generation Sequencing**

Next-generation sequencing (NGS) also known as 2nd generation sequencing is set to transform plant systematics, just as Sanger sequencing did more than 20 years ago [37]. In comparison to the conventional PCR and Sanger sequencing procedures currently employed in plant systematic research, the next-generation and targeted sequencing approaches offers significant time and cost benefits, especially in situations involving large numbers of taxa and phylogenetic markers [38]. In order to assure the discovery of variations that are clinically relevant, it is advised to use multiple analysis tools in conjunction with next-generation sequencing, which offers time and money saving methodology for evaluating multiple targets across several modalities [39].

Several recent developments in next-generation sequencing (NGS) technology have fuelled the development of the promising field of herbarium genomics, which enables testing of historical biological ideas in plant research [40, 41]. The function of PCR in library preparation allows commercial 2nd generation sequencing technologies to be distinguished from one another (Figure 3). Mostly, PyrosequencingTM, and Illumina® sequencing are the two NGS methods most frequently employed [42]. These NGS technologies have significantly advanced the field of plant taxonomy by providing high-throughput and cost-effective methods by providing large-scale genomic data, higher resolution, and comprehensive insights into plant diversity, evolutionary relationships, and species identification. NGS have proven to be an an indispensable tool for taxonomists, facilitating more accurate and efficient classification and understanding of the complex relationships among plant species.



**Figure 3: Comparison of steps involved in DNA fingerprinting, barcoding and NGS [42].**

1. **Oxford Nanopore Technology**

Oxford Nanopore Technology comes under third generation sequencing that has been working on developing a single-molecule, electrical, label-free DNA sequencing technique. This method aims to eliminate the requirement for amplification or labelling by sensing a straight electrical signal instead [43]. The use of Oxford Nanopore technology, along with complementary sequencing and analysis methods, significantly enhanced the understanding of *Atriplex hortensis*, its genetic variation, and phylogenetic positioning [44].

1. **High-throughput metagenomic shotgun sequencing**

High-throughput metagenomic shotgun sequencing is a powerful and advanced method used to analyze the collective genetic material of microbial communities present in a given environment. It provides a comprehensive and unbiased snapshot of all the DNA sequences (including both host and microbial DNA) present in a sample, without the need for prior knowledge or specific target sequences [45]. High-throughput metagenomic shotgun sequencing is very helpful for generating more complete genetic data from taxonomically significant decade old isotype herbarium specimens [46]. Cncurrently, the msGBS methodology, aids in plant taxonomy by quantifying multiple plant species in belowground interactions offering an advanced and scalable tool for studying complex root communities [47].

1. **Metagenomics**

Metagenomics enables the study of entire genetic material from environmental samples, providing insights into the diversity and distribution of plants in specific habitats. It harnesses the power of next-generation sequencing and bioinformatics technologies to explore the genetic diversity, abundance, composition, and metabolic pathways. Metagenomics can be a valuable tool in plant taxonomy, providing data that complements traditional morphological and molecular methods. The analysis of plant-associated microbial communities can enhance our understanding of plant diversity, evolutionary relationships, and ecological interactions, ultimately contributing to the advancement of plant taxonomy [48]. Metagenomics is known to enhance plant taxonomy by analyzing the diverse microbial communities in the rhizosphere of *Paspalum scrobiculatum* [49].

E. **Transcriptome sequencing**

Transcriptome sequencing, also known as RNA-Seq, is a powerful biotechnological tool to study and analyze the transcriptome of an organism. The transcriptome represents all the RNA molecules, including messenger RNA (mRNA), non-coding RNA (ncRNA), and other functional RNA molecules, that are transcribed from the DNA of a particular cell or tissue at a specific time point. Complete plastid genome sequencing has facilitated analyses of hundreds of taxa at deep levels and allowed phylogeographic studies at the population level. Gene capture methods show promise for rapid and inexpensive analyses of plastid genomes and targeted nuclear loci [50].

The transcriptome sequencing of *Dendrocalamus sinicus* study identified 8,553 simple sequence repeats (SSRs) and 81,534 single-nucleotide polymorphisms (SNPs). These molecular markers are valuable for population studies, genetic diversity assessment, and breeding programs [51]. The insights gained from transcriptome analysis can enhance the accuracy of plant classification, identify diagnostic markers for species discrimination, and shed light on the evolutionary relationships between different plant taxa.

1. **Plastome sequencing**

The process of determining and analyzing the complete DNA sequence of the plastid genome (plastome) of an organism. Next-generation sequencing (NGS) technologies are commonly used to sequence the plastome. Reference [52] used 93 specimens from 12 different Angiosperm families, 73 of which came from herbarium samples that were up to 146 years old. A sufficient number of paired-end reads were produced for 84 specimens resulting in successful plastome assemblies for 74 specimens. This shows that outline plastome sequencing from herbarium specimens is feasible and affordable and can be carried out with little sample destruction.

1. **Genotyping by sequencing**

A revolutionary technique called Genotyping by sequencing (GBS) combines genotyping and next-generation sequencing. It has a variety of uses, from general marker discovery to genome selection, making it a promising strategy that is likely to offer fresh insights into plant biology [53]. In 94 Amaranth accessions, GBS was used to identify 10,668 SNPs, the majority of which were species-specific, and these SNPs can be used for marker creation during further Amaranth research [54]. GBS uses genome-wide SNP markers to characterize *Lens culinaris* germplasm and identify gene pools in wild relatives. It aids in building relationships and detecting misclassified samples, making it a valuable tool for plant breeders interested in crop wild relatives [55].

1. **Phylogenetics and Phylogenomics**

Unlike traditional phylogenetics, which often focuses on a few specific genes or traits, phylogenomics involves analyzing whole genomes or a significant portion of the genome of multiple species. They use genetic and genomic data to infer the branching patterns of evolutionary history, showing how different species are related to each other through common ancestry [56]. The current phylogenomic research on *Oryza* serves as an illustration of how phylogenomics has proven its strength and enormous potential in resolving challenging phylogenetic questions [57].

1. **Genome skimming**

Genome skimming is a next-generation sequencing (NGS) approach used to obtain a broad overview of genomic information from an organism without performing whole-genome sequencing [58]. Genome skimming of Core Goodeniaceae samples allowed researchers to analyze plastome coding regions (CDS), nuclear ribosomal repeats (NRR), and nuclear G3PDH gene, significantly contributing to plant taxonomy by providing extensive genetic data that aids in resolving deep phylogenetic nodes and make informed taxonomic decisions [59]. Genome skimming was applied to milkweed plants (*Asclepias syriaca*) and related genera to demonstrate its effectiveness in generating genome-scale data sets for phylogenomics and has proven to be highly valuable in plant systematics and evolution studies [60].

**V. CONVENTIONAL MOLECULAR MARKERS**

Genetic diversity in conventional plant breeding was identified through observational selection. Simple sequence repeat variations (SSRs, or microsatellite polymorphisms), single nucleotide polymorphisms (SNPs), and copy number variations (CNVs) are the three main kinds of genetic variants found in biological genomes [61]. DNA polymorphisms are useful for analysis and are frequently utilized in molecular genetic investigations because they act as a genetic marker [62].

These DNA-based markers can be divided into two categories: PCR-based markers (RAPD, AFLP, SSR, SNP, etc.) and non-PCR-based markers (RFLP) (Table 3). The microsatellite DNA marker, among others, has been the one that is most frequently employed because it is straight forward to utilize by PCR, followed by a denaturing gel electrophoresis for determining allele size, and because of the high level of information offered by its numerous alleles per locus [63]. Study on *G. hirsutum* (cv. CCRI36) and *G. barbadense* (cv. H7124) as the plant species for the development and application of the ISAP (Intron-based Sequence Amplification Polymorphism) marker system suggested that these are PCR-based marker that targets gene sequences, providing functional molecular markers with high polymorphism and efficient amplification of adjacent expressed sequences. It offers valuable applications in map construction, QTL analysis, and gene mapping for plant breeding and selection [35].

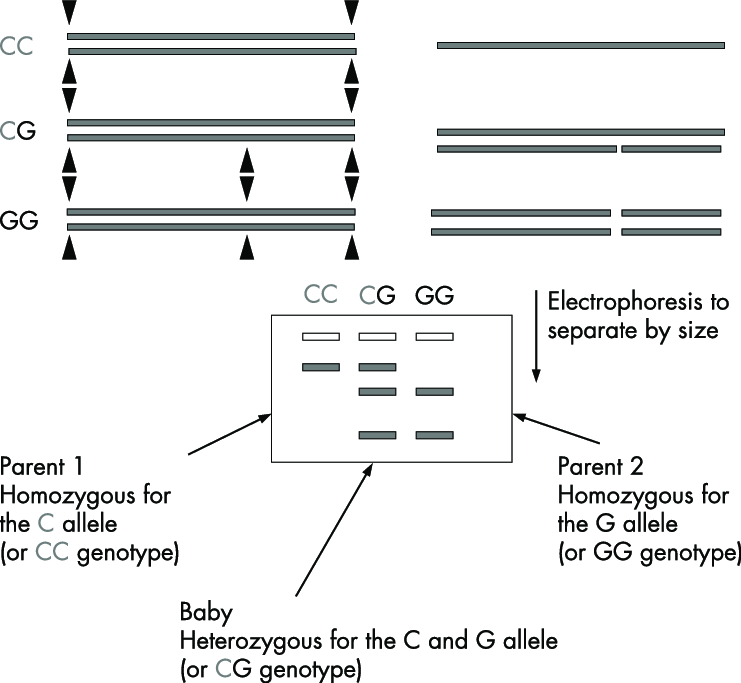
1. **DNA FINGERPRINTING IN PLANTS**

DNA fingerprinting in plants was developed later, building upon the principles and methods established by Alec Jeffreys and other researchers in the field of human genetics. The application of DNA fingerprinting to plants emerged as a powerful tool for studying plant genetics, biodiversity, and conservation. The use of DNA fingerprinting as a taxonomic tool in finding the variation in the species proved to be a helpful addition to morphology, particularly in plant groups with low rates of genetic recombination [64]. A rapid, dependable, and highly informative technique for DNA fingerprinting is provided by bulk analyses of RAPD and ISSR PCR markers [65]. Reference [66], evaluated the use of several DNA marker methods for fingerprinting 39 potato cultivars. RAPDs (20 primers), ISSRs (6 primers), AFLPs (2 primers), and SSRs (5 primer pairs) were the four methodologies that were looked into.

In addition to traditional phenotypic techniques, RFLP, RAPD, AFLP, microsatellites (SSRs), and SNPs, are used in plant taxonomy to identify and characterize plant species, evaluate genetic diversity and address evolutionary and taxonomic questions, complementing traditional phenotypic methods [67].

1. **Restriction Fragment Length Polymorphism**

DNA sequence polymorphisms in genes or other DNA regions of interest can be found using the RFLP study, which makes use of restriction endonuclease digestion (Figure 4). The nuclear genome of *Arabidopsis thaliana* has been assembled using an integrated genetic/restriction fragment length polymorphism (RFLP) linkage map, which offers a foundation for creating a more precise and practical map [68]. The use of RFLP analysis to investigate genetic linkages and variation within the tomato genus *Lycopersicon*, yielding important information about how different species within the genus are classified, behave when mating, and produce varied colours of fruit [69]. RFLP analysis is utilized in plant taxonomy to investigate the origin and evolution of cultivated *Brassica* species [70].

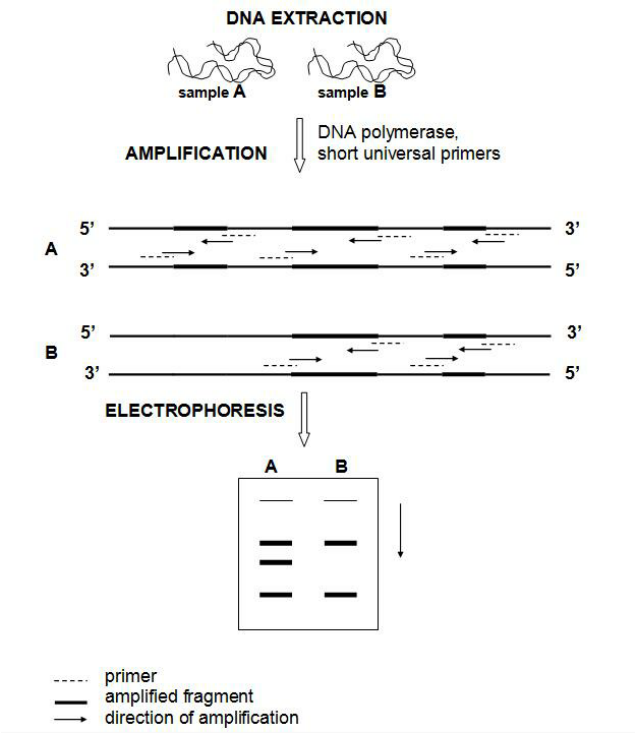


.

**Figure 4: RFLP and detection of alleles [71]**

1. **Random Amplified Polymorphic DNA**

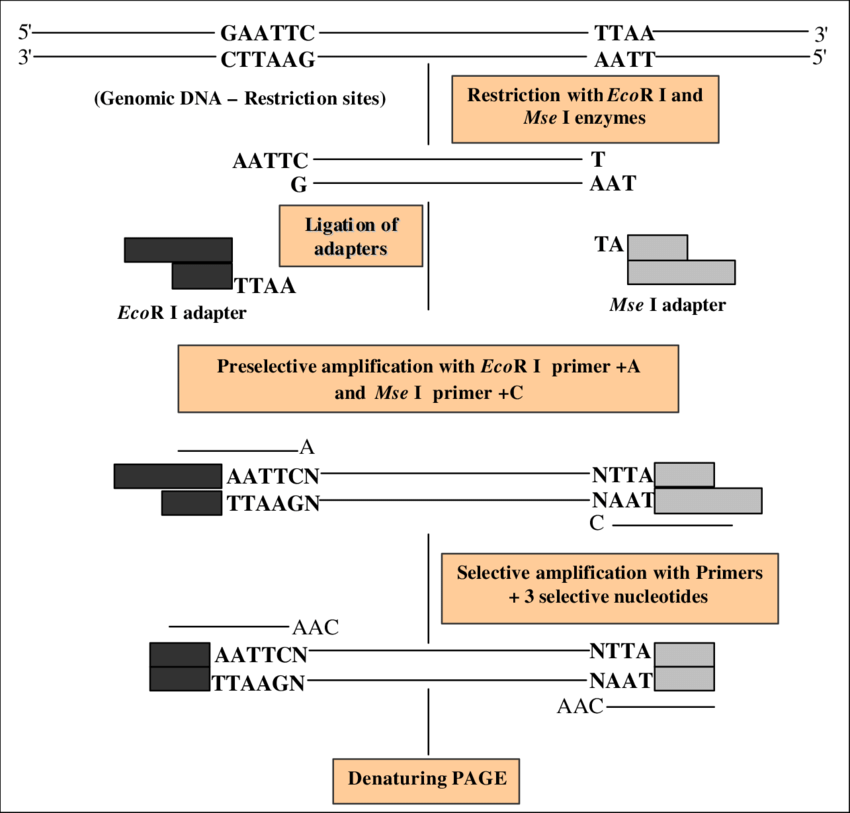
RAPD is a PCR- based technique that uses short, random primers to amplify genomic DNA at multiple sites (Figure 5). The amplified fragments are then visualized using agarose gel electrophoresis. RAPDs are able to identify significant genetic polymorphism in complete genomes [72]. RAPD analysis can be used to create genome-specific markers that can identify between cultivars of wheat, wild *Triticum* and *Aegilops* species, and other plant species. It has also been used to find particular markers for the D and U genomes [73]. Using RAPD analysis, it was possible to determine the species and relationships of the *Brassica, Sinapis,* and *Raphanus*. In general, RAPD analysis has the potential to be used for taxonomic studies at several levels, such as populations, species, and perhaps genera.



**Figure 5: Principle of RAPD-PCR technique [74]**

1. **Amplified Fragment Length Polymorphism**

The selective PCR amplification of DNA restriction fragments under exacting PCR conditions is the foundation of the AFLP technology (Figure 6). Utilizing two restriction endonucleases in tandem, the method entails digesting genomic DNA [75]. AFLP is a promising tool for evolutionary investigations because it is an effective and trustworthy method for producing biosystematic data [75]. AFLP is a new molecular marker technology that is a straightforward and reliable method that might be highly beneficial in a wide range of conservation studies [76]. As part of a study by reference [77], 87 taxa of the *Citrus, Fortunella,* and *Poncirus* families were examined using the AFLP method with two chosen primer pairs. The evolutionary relationships among the species in these genera were revealed by constructing a molecular systematic tree based on Nei's genetic distance. AFLP is used to study taxonomic relationships in *Vicia* and it distinguishes closely related taxa within the *Vicia sativa* aggregate. AFLP helps clarify the taxonomy and detect potential hybridization events [78].



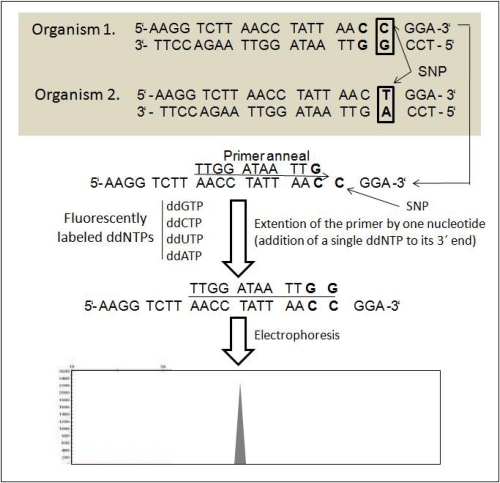
**Figure 6: Principle of the AFLP method [79]**

1. **Microsatellites**

Microsatellites or Single Sequence Repeats (SSRs) are frequently used in plant genetics investigations, employing both low and high throughput genotyping methods. Because of its co-dominance and stability of results, SSR is a more accurate molecular marker type than other marker types [80]. As SSR-based markers exhibit high levels of allelic variation, codominant inheritance, and simplicity of analysis, they will be an effective tool for taxonomic, phylogenetic, genome mapping, and population genetic studies [81]. Reference [82], adopted the SSR markers to study *Rosa* hybrid, in order to enhance flower trait development, breeding, and taxonomy, genomic and floral transcriptome sequencing. These markers allowed the examination of genetic links across contemporary rose accessions and other *Rosa* species.

1. **Single Nucleotide Polymorphism**

SNPs are the most prevalent marker system in both plant and animal genomes, and they have lately become the new generation of molecular markers for a variety of uses (Figure 7). Furthermore, unlike microsatellites, their potency comes from the vast number of loci that may be evaluated rather than the number of alleles [[83]. The study on Litchi cultivars suggested that the SNP markers could be used to identify and characterize more precisely, clearing up confusion in cultivar nomenclature and improving knowledge of the genetic connections between different Litchi accessions [84]. SNP markers help assess the molecular classification of Melon cultivars. It also highlights the limitation of using horticultural groups as botanical taxa [85].



**Figure 7: A flow-chart showing the basic principle of SNP method [67]**

**Table 2: Comparison of the five widely used DNA markers in plants [86]**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Criteria** | **RFLP** | **RAPD** | **AFLP** | **SSR** | **SNP** |
| Genomic coverage | Low copy coding region | Whole genome | Whole genome | Whole genome | Whole genome |
| Amount of DNA required | 10µg-50 | 100ng-1 | 100ng-1 | 120ng-50 | ≥50ng |
| Quality of DNA required | High | Low | High | Medium High | High |
| Type of polymorphism | Single base changes, indels | Single base changes, indels | Single base changes, indels | Changes in length of repeats | Single base changes indels |
| Level of polymorphism | Medium | High | High | High | High |
| Effective multiplex ratio | Low | Medium | High | High | Medium to high |
| Inheritance | Co-dominant | Dominant | Dominant / Codominant | Co-dominant | Co-dominant |
| Types of probes/ primers | Low copy DNA cDNA clone or | Usually 10 bp random nucleotides | Specific sequence | Specific sequence | Allele- specific PCR primers |
| Technically demanding | High | Low | Medium | Low | High |
| Radioactive detection | Usually yes | No | Usually yes | Usually no | No |
| Reproducibility | High | Low to medium | High | High | High |
| Time demanding | High | Low | Medium | High | High |
| Automation | Low | Medium | High | High | High |
| Development start up cost | High | Low | Medium | High | High |
| Proprietary rights required | No | Yes and licensed | Yes and licensed | Yes and some licensed | Yes and some  Licensed |
| Suitable utility in diversity, genetics and breeding | Genetics | Diversity | Diversity and genetics | All purposes | All purposes |

1. **Fluorescence in situ hybridization**

Complete nuclear genome sequencing is becoming commonplace, providing new opportunities for systematic comparative genomics research. Nevertheless, despite falling sequencing prices and technological breakthroughs, genome assembly continues to be a significant difficulty [50]. Fluorescence in situ hybridization (FISH) has proven to be a valuable tool in plant taxonomy, as demonstrated in the study of three *Larix* species (*L. sibirica, L. gmelinii,* and *L. cajanderi*). FISH was used to analyze the karyotypes of these taxa and identify specific ribosomal RNA gene loci [87]. FISH mapping of 35S rDNA in wild *Lilium* species helped to understand their taxonomic status, evolution, and karyotype diversity [88].

1. **CRISPR-Cas9 Technology**

Clustered Regularly Interspaced Short Palindromic Repeats-Cas9 is a revolutionary gene-editing tool that enables precise modification of plant genomes. It has been used in plant taxonomy to study specific genes and genetic markers, helping to resolve phylogenetic relationships between closely related species. Study on the Orchid *Dendrobium officinale* successfully applied the CRISPR/Cas9 system for editing endogenous genes in the genome[89]. These edits can potentially serve as DNA markers for studying genetic variation and evolutionary relationships within *D. officinale* and related orchid taxa. CRISPR/Cas9 has the potential to significantly advance the study of plant taxonomy, allowing for precise genetic modifications and molecular research in various plant species, including medicinal plants [90, 91, 92]. Its versatility allows researchers to precisely edit the DNA sequences of various plant species [93].

1. **Short Interspersed Nuclear Elements**

SINEs are repetitive DNA non-coding elements that can be found in plants and are capable of retrotransposition and can move within the genome, making them potentially useful markers for phylogenetic studies [94, 95]. The first report of a plant SINE family occurring in many lineages was made by [96], and they explored the distribution and evolution of Au SINE in plants. They also discussed the evolution of Au SINE in the plant kingdom, particularly in the Gramineae and Fabaceae. The 'Angio-domain' is present and conserved in SINEs across a variety of plant species, which raises the possibility that it could be used as an important identifier in plant taxonomy [97].

1. **Proteomics**

Proteomics is an essential aspect of plant biology that aids in understanding the phylogenetic relationships among plant taxa, characterizing individual lines, deciphering gene functions, and studying plant development and responses to the environment [98]. In the case of Holm Oak (*Quercus ilex* subsp. *ballota*) populations, proteomic analyses help catalog and understand the protein profiles, contributing to the study of plant taxonomy and the relationships between different populations of the species [99]. Similarly, proteomics in plant taxonomy involves comparing proteomes of various Brassicaceae species and genera to establish genetic relationships [100].

1. **ARTIFICIAL INTELLIGENCE AND PLANT TAXONOMY**

The proposed Artificial Intelligence (AI) system, which employs portrait and aerial photos for plant and weed identification, enhances accuracy and is appropriate and accurate in every class of comparison, making it a useful tool for farmers in obtaining the highest possible return on vegetable plantations [101]. Artificial neural networks (ANNs), more specifically a multilayer perceptron (MLP), can be used to identify higher plants using morphological traits gathered through conventional methods. ANNs outperform the DELTA (DEscription Language for TAxonomy) key generator [102]. Based on an Android application created as part of the Pl@ntNet project, these apps offer Android users a useful method for plant identification and have acceptable identification accuracy [103] (Figure 8). The semi-automatic graphical tool and the automatic plant identification method based on leaf images are the two main components of the computer-based plant identification system [104]. ApLeaf, an Android-based mobile application created to automatically identify plant species based on images of tree leaves, performs well with state-of-the-art identification performance and gives user access to several species that best match the query leaf image [105]. The Tchebichef Moment Invariant (TMI) feature and General Regression Neural Network (GRNN) classifier, which obtained a 100% classification rate in identifying plant species based on leaf photos, can also be used to create automated plant classification tools [106]. Automatic plant identification crowd sourcing systems based on images for botanical data collection have been accepted by a large number of users [107]. The accuracy of an automated plant identification system that uses a deep convolutional neural network to identify plant species through their leaves is 97% [108]. Reference [109], reports a new CNN (Convolutional Neural Network) based technique called D-Leaf can be a useful automated system for identifying plant species. The efficiency and potential capabilities of methods and software to image-based plant research are demonstrated by the Maize-IAS's simplicity of use, which also highlights the viability and promise of AI technology used in agriculture and plant science [110].

**

**Figure 8: Pl@ntNet (Android based application)**

1. **CONCLUSION AND FUTURE PERSPECTIVE**

The modern biotechnological tools brings an insight into the techniques that are decisive in taxonomical investigations for delineation between the taxa, and for understanding the unresolved morphological complexity with species level resolution. In particular, next generation sequencing and bioinformatic analysis, high-throughput and plastome sequencing have proven to be outstanding taxonomic tool. Most herbarium specimens are now potentially accessible to phylogenetic, population genetic, and barcoding studies because of recent advancements in non-destructive genetic sampling and working with very small amounts of genomic DNA, particularly in next-generation sequencing and bioinformatic analysis of ancient DNA.

A better way to disseminate taxonomic alterations and provide important information about plants is through online-based resources that offer a unique community platform for plant genomics research, which includes studies in evolution, genetics, plant breeding, molecular biology, biochemistry, and system biology to gather plant sequencing data that will advance plant molecular taxonomic research. The effectiveness of automated plant identification systems has been found to be highly encouraging after several artificial systems have been utilized to identify plants. Researchers and scientists have been able to advance science and knowledge through the application of these contemporary approaches in the study of plant taxonomy and systematics. Overall, the potential of biotechnological tools utilized in plant taxonomy to enhance our knowledge of plant diversity, evolution, and conservation is propitious. These technologies will remain as the leading competitors for taxonomic study, enabling more precise and effective species identification and categorization, ultimately supporting our efforts to protect and sustain the plant life on Earth.

1. **REFERENCES**

[1] B.P. Sheth and V.S. Thaker, “DNA barcoding and traditional taxonomy: an integrated approach for biodiversity conservation,” Gen. vol. 60, pp. 618-628, 2017.

[2] P. Bonnet, A. Joly, H. Goëau, J. Champ, C. Vignau, J. Molino, D. Barthélémy, and N. Boujemaa, “Plant identification: man vs. machine. Multimedia Tools and Applications,” vol. 75, pp. 1647-1665, Feb 2016.

[3] A. Wilmotte, “Molecular Evolution and Taxonomy of the Cyanobacteria,” in: Bryant, D.A. (eds) The Molecular Biology of Cyanobacteria. Advances in Photosynthesis, vol 1. Springer, Dordrecht, 1994, pp. 1-25.

[4] K. Thompson and S. Newmaster, “Molecular taxonomic tools provide more accurate estimates of species richness at less cost than traditional morphology-based taxonomic practices in a vegetation survey,” Biodiver. Cons. vol. 23, pp. 1411-1424, June 2014.

[5] S. Suranto, “Studies on Ranunculus Population: Isozymic Pattern,” Biodiv. J. Bio. Div. vol 2, 2001.

[6] C. Stace, “Classification by molecules: what's in it for field botanists?,” Wats. vol. 28, pp.103, 2010.

[7]E. Zuckerkandl and L Pauling, “Evolutionary divergence and convergence in proteins,” in InEvolving genes and proteins, Academic Press, 1965, pp. 97-166.

[8] C.R. Woese and G.E. Fox, “Phylogenetic structure of the prokaryotic domain: the primary kingdoms,” Pro. Nat. Aca. Sci. vol. 74, pp.5088-5090, Nov 1977.

[9] A.P.G, “An ordinal classification for the families of flowering plants,” An. Mis. Bot. Gar. pp.531-153, Jan 1998.

[10] P.D. Hebert, A. Cywinska, S. Ball and J. deWaard “ Biological identifications through DNA barcodes,” Pro. Roy. Soc. London, vol. 270, pp.313-321, Feb 2003.

[11] D. Steinke, M. Vences, W. Salzburger and A. Meyer “ TaxI a software tool for DNA barcoding using distance methods,” Phi. Tran. Roy. Soc. B: Bio. Sci. vol. 360, pp.1975-1980, Oct 2005.

[12] S. Ratnasingham, and P.D. Hebert, “BOLD: The Barcode of Life Data System (http://www. barcodinglife. org),” Mole. Eco. Not. vol. 7, pp. 355-364, May 2007.

[13] E. Park, J. Kim and H. Lee, “Plant DNA barcoding system for forensic application,” For. Sci. Int. Gen. Sup. Ser. vol.6, pp.e282-e283, Dec 2017.

[14] M. Källersjö , J.S Farris, M.W. Chase, B. Bremer, M.F Fay, C.J. Humphries, G. Petersen, O. Seberg and K. Bremer, “Simultaneous parsimony jackknife analysis of 2538 rbcL DNA sequences reveals support for major clades of green plants, land plants, seed plants and flowering plants,” Pl. Syst. Evo. vol. 213, pp.259-287, Sep 1998.

[15] M.Y. Stoeckle and P.D. Hebert. “Barcode of life,” Sci. Am. vol. 299, pp. 82-89 ,Oct 2008.

[16] F. Mattia, R. Gentili, I. Bruni, A. Galimberti, S. Sgorbati, M. Casiraghi, and M. Labra, “A multi-marker DNA barcoding approach to save time and resources in vegetation surveys,” Bot. J. Linn. Soc. vol. 30, pp.1207-1216, Sep 2012.

[17] W. Gu, J. Song, Y. Cao, Q. Sun, H. Yao, Q. Wu, J. Chao, J. Zhou, W. Xue and J. Duan, “Application of the ITS2 Region for Barcoding Medicinal Plants of Selaginellaceae in Pteridophyta,” PL. ONE.vol. 8, p.e67818, Jun 2013.

[18] R.S. Purty, and S. Chatterjee, “DNA barcoding: an effective technique in molecular taxonomy” Austin J Biotechnol Bioeng, vol 3, pp.1059, 2016.

[19] C. Husin, D. Naim and M. Ahmad, “Authentication of a selected medicinal plants using DNA barcoding technique” In SHS Web of Conferences, vol. 45,2018, EDP Sci. pp. 05004.

[20] P. Soledispa, E. Santos-Ordóñez, M. Miranda, R. Pacheco, Y.I. Gaiten and R. Scull, “ Molecular barcode and morphological analysis of Smilax purhampuy Ruiz,” Ecuador. PeerJ, vol. 9, pp.e11028, Mar 2021.

[21] Q. Li, J. Wu, Y. Wang, X. Lian, F. Wu, L. Zhou, Z. Huang and S. Zhu, “The phylogenetic analysis of *Dalbergia* (Fabaceae: Papilionaceae) based on different DNA barcodes,” Holzfors.vol. 71, pp. 939-949, Nov 2017.

[22] M. Polling, M. Sin, L.A. de Weger , A.G Speksnijder, M.J. Koenders, H. de Boer and B. Gravendeel, “DNA metabarcoding using nrITS2 provides highly qualitative and quantitative results for airborne pollen monitoring,” Sci. Tot. Env. vol. 806, pp.150468, Feb 2022.

[23] Y. Ji, C. Liu, J. Yang, L. Jin, L., Z. Yang and J. Yang, “ Ultra-Barcoding Discovers a Cryptic Species in *Paris yunnanensis* (Melanthiaceae), a Medicinally Important Plant,” Front. Pl. Sci, vol. 11, pp.411, April 2020.

[24] M. Gostel, J. Zúñiga, W. Kress, V. Funk, and C. Puente-Lelièvre, “Microfluidic Enrichment Barcoding (MEBarcoding): a new method for high throughput plant DNA barcoding,” Sci. Rep. vol.10, pp.8701, May 2020.

[25]S. Uribe-Convers, M.L. Settles and D.C. Tank, “A targeted subgenomic approach for phylogenomics based on microfluidic PCR and high throughput sequencing,” bioRxiv. pp. 021246, Jun 2015.

[26] S. Ouyang and C. Buell, “The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants,” Nuc. Ac. Res, vol 32, pp.D360-D363, Jan 2004.

[27] P. Kuksa and V. Pavlovic, “Efficient alignment-free DNA barcode analytics,” BMC Bioinfor, vol. 10, pp. 1-8, Nov 2009.

[28] C. Li, Z. Yu, G. Han, and K. Chu, “Analyzing Multi-locus Plant Barcoding Datasets with a Composition Vector Method Based on Adjustable Weighted Distance,” PL. ONE. vol. 7, pp. e42154, July 2012.

[29] F. Sanger, S. Nicklen and A.R. Coulson, “DNA sequencing with chain-terminating inhibitors,” Proc. Natl. Acad. Sci, vol.74, pp.5463-5467, Dec. 1977.

[30] W.M Barnes, “DNA sequencing by partial ribo-substitution,” J. Mol. Bio. vol. 119, pp. 83-99, Feb 1978.

[31] F. Sanger, A.R. Coulson, G.F. Hong, D.F. Hill and G.B Petersen, “Nucleotide sequence of bacteriophage λDNA,” J. Mol. Biol. vol. 162, pp. 729-773, Dec 1982.

[32] A.M. Maxam and W. Gilbert, “A new method for sequencing DNA,” Proc. Natl. Acad. Sci. vol. 74, pp.560-564, Feb 1977.

[33] S.F. Altschul, W. Gish, W. Miller, E.W. Myers and D.J. Lipman,” Basic local alignment search tool,” J. Mol. Biol, vol 215, pp. 403-410, Oct 1990.

[34] W.J Kress, D.L. Erickson, F.A. Jones, N.G. Swenson, R. Perez, O. Sanjur and E. Bermingham, “Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama,” Proc. Natl. Acad. Sci. USA, vol. 106, pp. 18621-18626, Nov. 2009.

[35] C. Lu, S. Yu, J. Yu, S. Fan, M. Song, W. Wang and S. Ma. “Development and appraisement of functional molecular marker: intron sequence amplified polymorphism (ISAP)]” Yi chuan = Hereditas, vol.30, pp. 1207-1216, Sep 2008.

[36] M.G. Gauthier, “Simulation of polymer translocation through small channels,” PhD Dissertation, University of Ottawa, 2007.

[37] S. Straub, M. Parks, K. Weitemier, M. Fishbein, R. Cronn, and A. Liston, “ Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematic,” Amer. J. Bot, vol. 99, pp.349-64. Feb 2012.

[38] G. Godden, I. Jordon-Thaden, S. Chamala, A. Crowl, N. García, C. Germain-Aubrey, J. Heaney, M. Latvis, X. Qi, and M. Gitzendanner, “Making next-generation sequencing work for you: approaches and practical considerations for marker development and phylogenetics,” Pl. Eco. Div. vol. 5, pp. 427-450, Dec 2012.

[39] H. Williams, K. Walsh, A. Diamond, A. Oniscu, and Z. Deans, “ Validation of the Oncomine™ focus panel for next-generation sequencing of clinical tumour samples,” Vir. Archiv, vol. 473, pp. 489-503, Oct 2018.

[40] R.M Gutaker and H.A Burbano, “Reinforcing plant evolutionary genomics using ancient DNA. Current opinion in plant biology,” vol. 36, pp. 38-45, April 2017.

[41] F.T. Bakker, “Herbarium genomics: plant archival DNA explored,” Paleogen. Gen. Sc. Ana. An. DNA, vol. 117, pp 33-43, Jan 2016.

[42] N.J. Sucher, J.R. Hennell, and M.C. Carles, “DNA fingerprinting, DNA barcoding, and next generation sequencing technology in plants.” (eds) Plant DNA fingerprinting and barcoding: methods and protocols, vol. 862, Humana Press, 2012, pp.13-22.

[43] J. Clarke, H.C. Wu, L. Jayasinghe, A. Patel, S. Reid and H. Bayley, “Continuous base identification for single-molecule nanopore DNA sequencing,” Nat. Nanotech, vol.4, pp. 265-270, April 2009.

[44] S.P. Hunt , D.E. Jarvis ,D.J. Larsen, S.I. Mosyakin, B.A. Kolano, E.W. Jackson, S.L. Martin, E.N Jellen and P.J. Maughan, “ A chromosome-scale assembly of the garden orach (*Atriplex hortensis* L.) genome using Oxford Nanopore sequencing,” Front. Pl. Sci. vol.11, pp. 624, May 2020.

[45] T.J. Sharpton, S.J Riesenfeld, S.W Kembel, J. Ladau, J.P. O'Dwyer, J.L. Green, J.A. Eisen and K.S. Pollard, “ PhylOTU: a high-throughput procedure quantifies microbial community diversity and resolves novel taxa from metagenomic data,” PLoS Comp. bio. vol.7, pp. e1001061, Jan 2011.

[46] S. Leavitt, R. Keuler, C. Newberry, R. Rosentreter, and L. Clair, “Shotgun sequencing decades-old lichen specimens to resolve phylogenomic placement of type material,” Pt. F. Sys. vol. 64, pp. 237-247, Oct 2019.

[47] C.A. Wagemaker, L. Mommer, E.J Visser, A.Weigelt, T.P. van Gurp, M. Postuma, A.E. Smit‐Tiekstra and H. de Kroon, “ msGBS: a new high‐throughput approach to quantify the relative species abundance in root samples of multispecies plant communities,” Mole. Eco. Res. vol 21, pp.1021-1036, May 2021.

[48] A.E. Fadiji and O.O. Babalola, “Metagenomics methods for the study of plant-associated microbial communities: a review,” J. Microbio. Met. Vol.170, pp. 105860, Mar 2020.

[49] R. Prabha, D.P. Singh, S. Gupta , V.K. Gupta, H.A. El-Enshasy and M.K. Verma, “ Rhizosphere metagenomics of *Paspalum scrobiculatum* l.(kodo millet) reveals rhizobiome multifunctionalities,” Microorg. vol.7, pp. 608, Nov 2019.

[50] D.E. Soltis, M.A. Gitzendanner, G. Stull, M.Chester, A. Chanderbali, S. Chamala, I. Jordon-Thaden, P.S. Soltis, P.S. Schnable and W.B. Barbazuk, “The potential of genomics in plant systematic,” Taxon. vol. 62, pp.886-898, Oct 2013.

[51] K. Cui, H. Wang, S. Liao , Q. Tang, L. Li, Y. Cui and Y. He, “Transcriptome sequencing and analysis for culm elongation of the world’s largest bamboo (*Dendrocalamus sinicus*),” Pl. one, vol. 11, pp. e0157362, June 2016.

[52] F.T. Bakker, D. Lei, J. Yu, S. Mohammadin, Z. Wei, S. van de Kerke, B. Gravendeel, M. Nieuwenhuis, M. Staats, D. E. A-Planas and R. Holmer, “Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an Iterative Organelle Genome Assembly pipeline,” Bio. J. Linn. Soc. vol. 117, pp.33-43, Jan 2016.

[53] A.A. Golicz, P.E. Bayer and D Edwards, “Skim-based genotyping by sequencing,” in Plant Genotyping: Methods and Protocols. 2015, pp.257-270.

[54] X. Wu, and M. Blair, “Diversity in Grain Amaranths and Relatives Distinguished by Genotyping by Sequencing (GBS),” Front. Pl. Sc. vol. 8. 2017.

[55] M.M. Wong, N. Gujaria-Verma, L. Ramsay, H.Y. Yuan, C. Caron, M. Diapari, A. Vandenberg and K.E. Bett, “Classification and characterization of species within the genus Lens using genotyping-by-sequencing (GBS),” PL. One, vol.10, pp. e0122025, Mar 2015.

[56] M.R. McKain, M.G. Johnson, S. Uribe‐Convers, D. Eaton and Y. Yang , “ Practical considerations for plant phylogenomics,” App. Pl. Sci, vol.6, pp.e1038, Mar 2018.

[57] X. Zou and S. Ge, “Conflicting gene trees and phylogenomics,” J. Syst. Ev. vol. 46, pp.795-807, Nov 2008.

[58] C.X. Zeng, P.M. Hollingsworth, J. Yang, Z.S. He, Z.R. Zhang, D.Z. Li and J.B. Yang, “Genome skimming herbarium specimens for DNA barcoding and phylogenomics,” Pl. Met. vol.14, pp.1-4, Dec 2018.

[59] A.G. Gardner, E.B. Sessa, P. Michener, E. Johnson, K.A. Shepherd , D.G. Howarth and R.S. Jabaily, “Utilizing next-generation sequencing to resolve the backbone of the Core Goodeniaceae and inform future taxonomic and floral form studies,” Mole. Phylogen. Evo. 2016, vol. 94, pp. 605-617, Jan 2016.

[60] K. Weitemier, S.C. Straub, R.C. Cronn, M. Fishbein , R. Schmickl, A. McDonnell and A. Liston,“Hyb‐Seq:Combining target enrichment and genome skimming for plant phylogenomics,” App. Pl. Sci, vol. 2, pp. 1400042, Sep 2014.

[61] R. Bernardo, “Molecular markers and selection for complex traits in plants: learning from the last 20 years,” Cr. Sci. vol. 48, pp. 1649-1664, Sep 2008.

[62] S. Teama, “DNA Polymorphisms: DNA-Based Molecular Markers and Their Application in Medicine,” Gen. Diver. Dis. Suscep. Oct 2018.

[63] P. Kumar, V.K. Gupta, A.K. Misra, D.R. Modi, and B.K. Pandey, “Potential of molecular markers in plant biotechnology,” Pl. om. vol.2, pp.141-162, July 2009.

[64] T. Kraft, H. Nybom, and G. Werlemark, “DNA fingerprint variation in some blackberry species (*Rubus* subg.*Rubus*, Rosaceae),” Pl. Syst. Evo. vol. 199, pp. 93-108, Mar 1996.

[65] M. Fernández, A. Figueiras, and C. Benito, “The use of ISSR and RAPD markers for detecting DNA polymorphism, genotype identification and genetic diversity among barley cultivars with known origin,” Theo. App. Gen, vol. 104, pp. 845-851, April 2002.

[66] C.E. McGregor, C.A. Lambert, M.W. Greyling, J.H. Louw and L. Warnich, “A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum* *tuberosum* L.) germplasm,” Euphy. vol. 113, pp.135-144, May 2000.

[67] I.A. Arif, M.A. Bakir, H.A. Khan, A.H. Al Farhan, A.A. Al Homaidan, A.H. Bahkali, M.A. Sadoon and M. Shobrak, “A brief review of molecular techniques to assess plant diversity,” Int. J. Mole. Sci. vol. 11, pp. 2079-2096. May 2010.

[68] B. Hauge, S. Hanley, S Cartinhour, J. Cherry, H. Goodman,M. Koornneef, P. Stam, C. Chang, S. Kempin, L. Medrano, and E. Meyerowitz, “An integrated genetic/RFLP map of the Arabidopsis thaliana genome,” Pl. J. vol. 3, pp.745-54,May 1993.

[69] J.C. Miller and S. D. Tanksley, “RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*,” Theo. App. Gen. vol. 80, pp. 437-448, Oct 1990.

[70] K. M. Song, T.C. Osborn and P.H. Williams, “ Brassica taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 1. Genome evolution of diploid and amphidiploid species,” Theo. App. Gen. vol. 75, pp.784-94, May 1988.

[71] D. Harding, “Impact of common genetic variation on neonatal disease and outcome,” in Archives of Disease in Childhood-Fetal and Neonatal Edition. vol.92, pp. 2007, F408-13.

[72] K.J. Edwards, “Randomly amplified polymorphic DNAs (RAPDs). In: Molecular Tools for Screening Biodiversity,” (eds Karp A, Isaac PG, Ingram DS), Chapman & Hall, New York, 1998 ,pp. 171–175.

[73] S. Cenkci, M. Yıldız, M. Konuk and Y. Eren, “RAPD analyses of some wild Triticum l. and Aegilops l. species and wheat cultivars in Turkey”. Ac. Biolog. Cracov. Ser. Bot, vol. 50, pp. 35-42, Jan 2008.

[74] A. Pasqualone, “Cultivar identification and varietal traceability in processed foods: A molecular approach,” in Cultivars: Chemical Properties, Antioxidant Activities and Health Benefits, 2013, pp. 83-105.

[75] J. Kardolus, H. Eck and R. Berg, “The potential of AFLPs in biosystematics: A first application in *Solanum* taxonomy (Solanaceae),” Pl. Syst. Evol, vol. 210, pp. 87-103, March 1998.

[76] V. Lucchini, “AFLP: a useful tool for biodiversity conservation and management.” Comp. Ren. Bio. vol.326, pp. 43-48, Aug 2003.

[77] G. Liang, G. Xiong, Q. Guo, Q. He and X. Li, “AFLP analysis and the taxonomy of Citrus,” In XXVII Int. Hort. Cong -IHC2006, II Inter. Sym. Pl. Gen. Res. Hort. vol. 760 pp. 137-142, Aug 2006.

[78] M. van de Wouw, N. Maxted,K. Chabane and B.V. Ford-Lloyd, “Molecular taxonomy of *Vicia* ser. *Vicia* based on amplified fragment length polymorphisms,” Pl. Sys. Evo. vol. 229, pp. 91-105, Aug 2001.

[79] M.C. Yadav, “DNA markers in crop improvement. Molecular Genetics Laboratory,” Nat. Res. Cen. Mush. Ind. Coun. Agri. Res. Solan-173, pp. 213, 2008.

[80] K. Abdellatif, and Y. Khidr, “Genetic diversity of new maize hybrids based on SSR markers as compared with other molecular and biochemical markers,” J. Cr. Sci. Biotech. vol.13, pp.139-145, Sep 2010.

[81] R. Singh, S. Sheoran, P. Sharma and R. Chatrath, “Analysis of simple sequence repeats (SSRs)dynamics in fungus *Fusarium graminearum*,” Bioinfor. vol.5, pp. 402 – 404, Feb 2011.

[82] W. Qi, X. Chen, P. Fang , S. Shi , J. Li , X. Liu ,X. Cao, N. Zhao, H. Hao, Y. Li and Y. Han, “Genomic and transcriptomic sequencing of Rosa hybrida provides microsatellite markers for breeding, flower trait improvement and taxonomy studies,” BMC Pl. Bio. Vol. 18, pp. 1-1, Dec 2018.

[83] J.T. Foster, G.J. Allan, A.P. Chan, P.D. Rabinowicz, J. Ravel, P.J. Jackson, and P. Keim, “Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis*),” BMC Plant Biol,vol 10, pp. 13–23,Dec 2010.

[84] W. Liu, Z. Xiao, X. Bao , X. Yang, J. Fang and X. Xiang, “ Identifying litchi (*Litchi chinensis* Sonn.) cultivars and their genetic relationships using single nucleotide polymorphism (SNP) markers,” PL. ONE, vol. 10, pp.e0135390, Aug 2015.

[85] J. Jung, G. Park, J. Oh, J.K. Jung, E.J. Shim, S.M. Chung, G.P. Lee and Y. Park, “ Assessment of the current infraspecific classification scheme in melon (*Cucumis melo* L.) based on genome-wide single nucleotide polymorphisms,” Hort. Env. Biotech. vol. 61, pp. 537-547, June 2020.

[86] M. Kordrostami and M. kers Rahimi, “Molecular marin plants: concepts and applications,” Genet. 3rd Millenn,vol. 13, pp.4024-4031, 2015.

[87] O.V. Goryachkina, E.D. Badaeva, E.N. Muratova and A.V. Zelenin, “Molecular cytogenetic analysis of Siberian Larix species by fluorescence in situ hybridization,” Pl. Syst. Evo.vol.299, pp.471-479, Feb 2013.

[88] S. Zhou, X. Chai, Z. Yang, H. Wang,C. Yang, and T. Sun, “Maize-IAS: a maize image analysis software using deep learning for high-throughput plant phenotyping,” Pl. Met.vol.17, p.48 , April 2021.

[89] L. Kui, H. Chen, W. Zhang , S. He, Z. Xiong, Y. Zhang, L. Yan, C. Zhong, F. He, J. Chen and P. Zeng ,“ Building a genetic manipulation tool box for orchid biology: identification of constitutive promoters and application of CRISPR/Cas9 in the orchid, *Dendrobium officinale*,” Front. In pl. Sci. vol. 7, pp 2036, Jan 2017.

[90] L. Zeng, Q. Zhang, C. Jiang, Y. Zheng, Y. Zuo, J. Qin, Z. Liao and H. Deng, “Development of *Atropa belladonna* L. plants with high-yield hyoscyamine and without its derivatives using the CRISPR/Cas9 system,” Int. J. Mole Sci. vol. 22, pp. 1731, Feb 2021.

[91] X. Li, X. Zuo, M. Li, X. Yang, J. Zhi, H. Sun, C. Xie, Z. Zhang and F. Wang, “ Efficient CRISPR/Cas9-mediated genome editing in *Rehmannia glutinosa*,” Pl. Ce. Rep. vol. 40, pp.1695-1707, Sep 2021.

[92] H.F. Wan, W.J. Han, L. Zhou, S. Wang and C. Sui, “New Advances of CRISPR/Cas9 Technique and its Application in Disease Treatment and Medicinal Plants Research,” Cur. Phar. Biotech. vol. 23, pp.1678-1690, Nov 2022.

[93] X. Liu, S. Wu, J. Xu, C. Sui and J. Wei, “ Application of CRISPR/Cas9 in plant biology.,” Ac. Phar. Sin. B vol.7, pp. 292-302, May 2017.

[94] A. Lenoir, B. Cournoyer, S. Warwick, G. Picard and J.M. Deragon, “ Evolution of SINE S1 retroposons in Cruciferae plant species,” Mole. Bio. Evo, vol. 14, pp. 934-941, Sep 1997.

[95] J.M. Deragon and X. Zhang. “Short interspersed elements (SINEs) in plants: origin, classification, and use as phylogenetic markers,” Syst Bio. vol.55, pp.949-956, Dec 2006.

[96] J.A. Fawcett, T. Kawahara, H. Watanabe and Y. Yasui, “ A SINE family widely distributed in the plant kingdom and its evolutionary history,” Pl. Mole. Bio.vol. 61, pp.505-514, June 2006.

[97] K.M. Seibt, T. Schmidt and T. Heitkam, “The conserved 3′ Angio‐domain defines a superfamily of short interspersed nuclear elements (SINEs) in higher plants,” Pl. J. 2020.

[98] C.M. Ishtiaq, Q. He, J.P. Huang, Y. Wang, P.G. Xiao and Y.C. Yi, “Biosystematics and plant proteomics: role of proteomics in plant phylogenetic analysis,” Pak. J. Bio. Sci. vol.10, pp. 3487-3496, Oct 2007.

[99] J.V. Galván, R.G. Fernández, L. Valledor, R.M. Cerrillo and J.V. Jorrin-Novo, “ Proteotyping of Holm oak (*Quercus ilex* subsp. *ballota*) provenances through proteomic analysis of acorn flour,” Pl. Pro: Met. Prot. pp.709-23, 2014.

[100] K. Marquès, B. Sarazin, L. Chané‐Favre, M. Zivy and H. Thiellement, “ Comparative proteomics to establish genetic relationships in the Brassicaceae family,” PRO: Int. Ed. Vol.1, pp. 1457-1462, Nov 2001.

[101] A. Samraj, R. Kumarasamy, N.T, K.T and R.P, “Plant Variety and Weed Growth Identification: Trending Towards Machine Learning,” Inter. J. Rec. Tech. Eng. Vol. 8, July 2019.

[102] J.Y. Clark, “ Artificial neural networks for species identification by taxonomists,” Biosys. vol. 72, pp.131-147, Nov 2003.

[103] S. Mouine, I. Yahiaoui, A. Verroust-Blondet, L. Joyeux,, S. Selmi, and H. Goëau, “An android application for leaf-based plant identification,” In Proceedings of the 3rd ACM conference on International conference on multimedia retrieval, 2013, pp. 309-310.

[104] N. Pham, T. Le, P. Grard and V. Nguyen, “Computer aided plant identification system,” Inter. Con. Comp. Manage. Telecom. (ComManTel). pp. 134-139, Jan 2013.

[105] Z. Zhao, L. Ma, Y. Cheung, X. Wu, Y. Tang and C. Chen, “ApLeaf: An efficient android-based plant leaf identification system,” Neurocom. Vol. 151, pp 1112-1119, March 2015.

[106] Z. Zulkifli, P. Saad, and I. Mohtar, “ Plant leaf identification using moment invariants & General Regression Neural Networ,” in 11th International Conference on Hybrid Intelligent Systems (HIS), pp. 430-435, Dec 2011.

[107] T.T. Nguyen, T.L. Le ,H.Vu, V.S Hoang and T.H. Tran, “ Crowdsourcing for botanical data collection towards to automatic plant identification: a review,” Comp. Elec. Agri, vol. 1, pp. 412-425, Dec 2018.

[108] D. Bisen, “Deep convolutional neural network based plant species recognition through features of leaf,” Multi. Tool App. vol. 80, pp. 6443-64456, Feb 2020.

[109] J. Tan, S. Chang, S. Abdul-kareem, H. Yap and K. Yong, “ Deep Learning for Plant Species Classification Using Leaf Vein Morphometric” IEEE/ACM Transac. Comput. Bio. Bioinfor. Vol. 17, pp. 82-90, June 2020.

[110] Y.P. Zhou, Z.X. Wang, Y.P. Du, J.W. Li, H.B. He and G.X. Jia, “ Fluorescence in situ hybridization of 35S rDNA sites and karyotype of wild *Lilium* (Liliaceae) species from China: taxonomic and phylogenetic implications,” Gen. Res. Cr. Evo. vol. 67, pp. 1601-1617, Aug 2020.