A DISSERTATION REPORT ON

Molecular and Phytochemical Diversity analysis in accessions of *Cannabis* sp.

SUBMITTED TO THE
DEPARTMENT OF BIOSCIENCES
INTEGRAL UNIVERSITY

IN PARTIAL FULLFILLMENT
FOR THE
DEGREE OF MASTER OF SCIENCE
IN BIOTECHNOLOGY

SUBMITTED BY
POOJA YADAV
M.Sc. Biotechnology (IV Semester)
Department of Biosciences
INTEGRAL UNIVERSITY, LUCKNOW

UNDER THE SUPERVISION OF

Dr. NARENDRA KUMAR
(Senior Scientist)
Department of Botany and Pharmacognosy,
CSIR- CENTRAL INSTITUTE OF MEDICINAL AND AROMATIC PLANTS
LUCKNOW-226015, UTTAR PRADESH

CSIR-CENTRAL INSTITUTE OF MEDICINAL AND AROMATIC PLANTS

HUMAN RESOURCE DEVELOPMENT PROGRAMME

CERTIFICATE

This certificate is issued to the candidate by *CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow,* on successful completion of her M.Sc. dissertation work.

Name of candidate	: Ms. Pooja Yadav
Institution	: Integral University, Lucknow
Category of training	: Graduate Training
Training Department	: Botany and Pharmacognosy
Title of training	: Molecular and Phytochemical Diversity analysis in accessions of Cannabis sp.L.
Duration of training	: Four months (07/02/2022- 07/06/2022)

The institute wishes the candidate success in her future endeavors.

Supervisor
Dr. Narendra Kumar
(Senior Scientist)

INTEGRAL UNIVERSITY

Established Under U.P. Act No 09 of 2004 by State Legislation
Approved by University Grants Commission

Phone No.: +91 (0552) 2890812, 2890730, 3296117, 6451039, Fax No.: 0522-2890809

Kursi Road, Lucknow-226026, Uttar Pradesh (INDIA)

<u>TO WHOM IT MAY CONCERN</u>

This is to certify that **Ms. Pooja Yadav**, a student of M. Sc. Biotechnology (II Year, IV semester), Integral University has completed her four months dissertation work entitled "***Molecular and Phytochemical Diversity analysis in accessions of Cannabis sp.***" successfully. She has completed this work from Department of Botany and Pharmacognosy, CSIR- central institute of medicinal and aromatic plants, under the   guidance of **Dr. Narendra Kumar (senior scientist).**

The dissertation was a compulsory part of her M. Sc. degree. I wish her good luck and future.

Dr. Snober S. Mir

Head,

Department of Biosciences,

Integral University, Lucknow

DECLARATION

I hereby declare that the dissertation work ***"Molecular and Phytochemical Diversity analysis in accessions of Cannabis sp."*** is being submitted to the partial fulfillment of the award of the degree of Master of Sciences in Biotechnology, Integral University Lucknow.

It is an authentic record of work carried out under the supervision of **Dr Narendra Kumar,** Senior Scientist, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, U.P.

I Pooja Yadav hereby declare that the work embodied that the work is not been submitted to any other University / Institution.

Place: Pooja Yadav

Date:

ACKNOWLEDGEMENT

TABLE OF CONTENT

List of Abbreviation

THC	Tetrahydrocannabinol
CBD	Cannabidiol
CBG	Cannabigerol
ISSR	Inter Simple Sequence Repeats
SCOT	Start Codon Repeats
PCR	Polymerase Chain Reaction
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
HPLC	High Pressure Liquid Chromatography
TLC	Thin Layer Chromatography
µL	Micro Liter
MG	Milligram
ML	Milliliter
RPM	Revolution Per Minute
EDTA	Ethylenediamine Tetra Acetic Acid
ETBR	Ethidium Bromide
DNTPS	Deoxynucleotide Triphosphates
MGCL2	Magnesium Chloride
CTAP	Cetyl Trimethyl Ammonium Bromide
PVP	Polyvinylpyrrolidone
GC	Gas Chromatography
PCA	Principle Component Analysis
OD	Optical Density
DNA	Deoxyribonucleic Acid

LIST OF TABLE

LIST OF FIGURE

1. <u>INTRODUCTION</u>

Cannabis L. is the most ancient legendary plant, principally grown everywhere around the globe. *Cannabis* is a member of the family Cannabacae that is mainly represented by three main species, i.e., *Cannabis sativa*, *Cannabis indica* and *Cannabis ruderalis*. Later on, it was considered a monotypic genus with the species *Cannabis sativa*. It is believed that *Cannabis* has originated in the Central Asian region. However, the plant is now growing in many countries throughout the globe for its wider use as a fibre and oil crop.

Cannabis plants are dioecious with distinct male and female individuals; sometimes, monoecious plants were also found. *Cannabis* is an annual, dicotyledonous flowering plant. The plant is up to 5 m in height, stem slender, much-branched, sometimes with ridges and furrows. Leaves palmately lobed, present 3-11 leaflets, petiole pubescent. Leaflet sessile, lanceolate, narrowed at base, serrate, acuminate at apex, lower surface more densely pubescent than the upper surface, filled with glandular and non-glandular hairs. Male flowers 4 to 6 mm in size, yellowish-green, pedicel 1 to 2 mm long, filiform anther. Tepals entire, acute, elliptic to oblong, pubescent and four in number. Stamen 4 to 5 mm long. Female flowers are covered by persistent bract, bracts 4 to 10 mm long, and are covered with glandular and non-glandular hairs. Ovary globose, sessile, 1 to 1.5 mm long, brown, caduceus, and covered with hairs. Fruits achene, 3 to 6 mm long, persistent bract, yellowish-brown in colour, glabrous, ovate; Seed with small unilateral endosperm. *Cannabis* is an anemophilous plant, i.e., pollination generally occurs by the wind.

Traditionally hemp was used for the high quality of fibre from stem; seeds produce nutrient-rich oil and proteins, essential oils and resins from inflorescence and leaves. *Cannabis* has important biologically active compounds, such as psychoactive compounds, Delta-9-tetrahydrocannabinol (Δ9-THC) and non-psychoactive compounds, i.e. cannabidiol (CBD). THC is used for recreational purposes and substantially impacts the human body, the reason behind the ban of *Cannabis* in many countries. The government allows the cultivation of *Cannabis* specific cultivars with THC content (based on dry weight) of less than 0.2% in Europe and 0.3% in Asian countries. Mainly three different types of *cannabis* products were utilized in India. First is the bhang, prepared from the leaves and tops

of the marijuana plant. It is usually ingested or swallowed as an injection in a beverage. The second is the ganja, which comprises leaves and is generally smoked. The third one is known as "charas" or "hashish", which is the resinous compound and resin which is taken from buds and leaves of marijuana. *Cannabis* has three main and important types of products that are herb, resin, and oil, respectively.

In *Cannabis*, more than 500 specialized metabolites have been reported, such as phytocannabinoids, terpenes and flavonoids. Among them, the significant Cannabinoids are Δ9-tetrahydrocannabinol (Δ9-THC), cannabidiol (CBD), cannabinol (CBN), Δ9-tetrahydrocannabivarin (D9-THCV), cannabigerol (CBG), cannabicyclol (CBL), cannabinol (CBN) and cannabichromene (CBC).

Cannabis plants have broad ranges of therapeutic activity against various disorders like analgesic, anti-inflammatory, appetite stimulant, anti-repulsive properties, nerve disorders, Tourette's syndrome, multiple sclerosis, nausea and vomiting, schizophrenia, Alzheimer's and Parkinson's disease. Besides this, cannabinoids have a very favourable and promising drug for therapeutic uses, especially for cancer and AIDS patients. *Cannabis* has become popular in clinical research due to its non-intoxicating effects. Besides this, it is also effective in treating neuropathic pain, spasms and disorders related to movements. Due to their wide range of pharmaceutical products in humans, cannabinoids are the most studied compounds comprising psychotropic activities. At the starting of the 19th century, the therapeutic use of Cannabis was confirmed in Western medicine and attained its highest point in the last two decades. The earlier research on hemp was limited due to the legal issues and lack of awareness about the plants. The recent trends changed the scenario, the cultivation purposes and medicinal uses of *Cannabis* are increasing daily.

1.2. MOLECULAR MARKER:

A molecular marker is a short sequence of DNA used to determine inheritance patterns that can be observed. A molecular marker is mainly based on naturally occurring DNA polymorphism that confirms the basis for calculating strategies to utilize for applied purposes. DNA markers have numerous applications in plant molecular genetic research (**Gupta et al. 1999; Semagn et al. 2006; Winter and**

Kahl 1995). Two of the most common uses of DNA markers have been assessing genetic diversity within crop germplasm and constructing linkage maps for mapping genes or quantitative trait loci (QTL) controlling agronomically important traits (**Collard et al. 2005; Farooq and Azam 2002**). A marker should exist in various forms that must be polymorphic so that the chromosome with the mutant gene can be differentiated from the chromosomes with the wild gene.

A molecular marker must have some advantages

- It must be polymorphic because polymorphism is the leading characteristic for measuring genetic diversity.
- The molecular marker must be co-dominantly inherited, which means different makers should permit the differentiation of homozygotes and heterozygous.
- It should be uniformly and regularly distributed all over the genome.
- It should be cheap, fast and easy to detect.

1.2.1 Inter simple sequence repeats (ISSR) markers:

ISSR is a PCR-based molecular marker technique. In this technique, single short oligonucleotides primers are inconsistently selected to amplify all over the genome (**Williams et al. 1990**). The charges as polymorphism in the pattern of bands amplified from genetically different individuals acted as Mendelian genetic markers. ISSR amplification is carried out in conditions favouring PCR by using genomic DNA from the species of interest and a single short oligonucleotide. The DNA amplification results are obtained from a region marked by a part of 10 base pair promoting sites in the suitable direction. Genomic DNA from two non-identical individuals frequently produces different amplification patterns. The unique fragments obtained for one individual but not for other represents DNA polymorphism and can be used as a genetic marker. Many random oligonucleotide primers using different nucleotide combinations have been generated and are commercially available. These primers can be synthesized in an oligonucleotide synthesizing facility-based randomly selected sequences.

PCR is based on three simple steps required for any DNA synthesis reaction:

(1) **Denaturation:** denatures the template DNA (break H-bonds b/w bases) into single strands at 94°C.

(2) **Annealing:** cooling of the mixture at 40-60°C to permit the annealing of primers to each original strand for new strand synthesis.

(3) **Extension:** - the extension of the new DNA strands from the primers at 72°C.

The PCR reaction generally requires cycling between the three temperatures. First, to denature or break the template DNA strands; second, to anneal the primers; and third, to extend at temperature optimum for Taq polymerase. This process is generally repeated 25-45 times. For ISSR analysis, the annealing temperature of 40°c- 60°c is used in PCR reactions. The appearance of a band that corresponds to a recessive allele. Therefore, homozygous and heterozygous individual dominants with ISSR markers cannot be differentiated

1.2.1.1 Advantages of ISSR

- The need for a small quantity of DNA (i.e., 25-50mg) makes it possible to work with populations.
- It consists of a simple experimental set-up that needed only an agarose assembly and a thermocycler.
- The work can be managed on a large variety of specie where species-specific gene libraries are unavailable.
- It gives a rapid and efficient screening for the DNA sequence-based polymorphism at numerous loci.
- It does not require blotting and hybridization steps.

1.2.1.2. Limitation of ISSR

- It causes a loss of information-related markers that show co-dominance because ISSR polymorphisms are inherited as dominant, recessive characters.

- A mismatch of even a single nucleotide can frequently prevent the primer from annealing because ISSR primers are generally short, causing a loss of band.
- Changes in some of the amplified fragments have resulted because ISSR is very sensitive to the changes in PCR conditions

1.2.2. Start Codon Targeted (SCoT) marker:

Start Codon Targeted Polymorphism (SCoT) PCR-based technique developed by **Collard and Mackill in 2009**. SCoT is the simple, innovative DNA marker technique that uses an 18-mer single primer in PCR with an annealing temperature of 50°C and PCR products are resolved using standard agarose gel electrophoresis. SCoT primer was designed based on the short conserved start codon (ATG) sequences in plant genes. This method has been confirmed high polymorphic, efficient, and successfully utilized in rice and peanut for cultivar identification and genetic diversity analysis. The SCoT primers are based on conserved regions flanking the initiation codon sequences of genes. It shares the principle of using a single primer like RAPD and ISSR. The following characteristics of SCoT marker:

- SCoT markers are precisely simple for use.
- It is a dominant marker, requires no prior sequence information, and the polymorphism is correlated to functional genes and their analogous characters
- The recombination levels between the SCoT marker and the gene/trait are generally lower when compared with random markers such as RAPDs, ISSRs or SSRs
- Thus it could be used directly in marker-assisted breeding programs for the analysis of point mutations based on polymorphism with SCoT primers
- It has been used to evaluate genetic relationships in plants. It is helpful for plant breeding. It is useful for accessing genetic relationships among the generations.
- This marker system requires no prior knowledge about the sequence under study.

- It is helpful for QTL mapping, which emerged as a superior system compared to RAPD.
- SCoT can give high reproducibility as compared to RAPD.
- The annealing temperature of SCoT is higher than that of RAPD.
- In some cultivars, SCoT has emerged superior to ISSR when used for accessing genetic relationships. The information generated by Start Codon Targeted Polymorphism is valuable. Since it is a gene-targeted marker system, it produces particular and reliable data.

1.2.2.1 Advantages of SCoT marker:

- SCoT markers require very little and not necessarily high-quality DNA.
- It is highly polymorphic.
- It is evenly distributed throughout the genome.
- The result of the SCoT marker is interpreted.
- SCoT markers are easily automated, allowing multiplexing.
- The marker has good analytical resolution and high reproducibility.

1.2.2.2 Limitation of SCoT Marker:

- SCoT marker shows low reproducibility.
- In the SCoT marker, the PCR cycling conditions greatly influence the outcome.
- The SCoT marker is highly sensitive.
- In the SCoT marker, the mismatch between primer and template may result in the total absence of PCR product.

1.3 CHROMATOGRAPHY:

Chromatography is a well-used separation technique. It is used in many areas of study, particularly in chemistry, biology and medicine. It is used in the examination, isolation and purification of biologically active compounds as a

component of small and large-scale production. It often detects one or several features in a complex mixture.

In terms of scale, a minimal amount, i.e., less than a Nanogram, is separated and identified during analysis. While at a large scale, material at a specific time results in a refined product. The versatility of chromatography in its many variants is behind its environmental status in separation science.

Chromatography is based on separating the phytochemical components between two phages, i.e. stationary and mobile phases. The different elements of the mixture travel through the stationary phase at different speeds, passing them to separate from one another. The nature of the specific mobile and stationary phase determines which substance travels more quickly or slowly.

The basic principle in chromatography is separation based on the various components distributed or separated between liquid phases. It involves the use of aqueous solvent held in pores of the stationary phase, whereas the mobile phase travels over the paper/plates. Due to differences in their affinity towards water (in the stationary phase) and mobile phase solvents, the compounds in the mixture get separated by the capillary action of the pores in the paper/plates. The components may also be divided based on the principle of adsorption between solid and liquid phases, where the solid surface of paper serves as the stationary phase, and the mobile step is a liquid solvent. Although the main principle of chromatography is partitioning, this is employed in many pharmaceutical applications (**G. Zweig et al. 1976**)

The term Rf is associated with the migration of the solute relative to the solvent front as:

$$Rf = \frac{Distance\ traveled\ by\ the\ solute\ from\ origin\ line}{Distance\ traveled\ by\ the\ solvent\ from\ origin\ line}$$

The Rf is the function of the partition coefficient. It is constant for the given substance, provided the condition of chromatographic systems are kept consistent concerning temperature, type, duration, the direction of development, nature and the shape and size of the wick. The Rf values of a substance depend on several factors, such as

a) Solvent system

b) Medium of separation, i.e. quality of stationary phase

c) Nature of mixture

d) Temperature

e) Dimensions of the vessel where operation is to be carried out.

1.3.1. Stationary and mobile phase

Chromatography is partition chromatography and various valid combinations of stationary and mobile phases. The two systems don't need to be unmixed. The types of stationary phases used can be classified as aqueous, hydrophilic and hydrophobic systems.

1.3.2 Stationary phases

(1) Aqueous stationary phase is mainly used in paper chromatography. Water is readily detained by paper. Therefore, water equilibrated paper is attached by suspending paper in a closed chamber whose atmosphere is saturated with water. If an aqueous buffer or salt phase is required, the paper is drawn through the respective solution and then exposed to a water-saturated atmosphere in a chamber. This system is particularly suited for separating moderate polar to extremely polar mixtures.

(2) Hydrophilic stationary phase: An organic solvent can be used for the hydrophilic stationary phase. If the solvent is volatile enough, the paper can be equilibrated in a chamber whose atmosphere is saturated with solvent. Alternatively, the stationary phase solvent is dissolved in a very volatile diluent, leaving the stationary phase liquid uniformly distributed throughout the plates. Commonly used hydrophilic solvents include formamide, methanol, glycerol and glycols.

(3) Hydrophobic stationary phase: The plates must be modified previously before they will tend to retain the hydrophobic stationary phase. Equilibration in the vapours of the solvent is the dipping technique in a solution of the solvent, and a volatile

diluent is used for introducing the hydrophobic solvent into the modified paper. Solvents such as dimethylformamide, aromatic and aliphatic hydrocarbons and kerosene are commonly used.

1.3.3. Mobile phase

The mobile phase in various combinations can be used in chromatography. Choosing the optimum eluting condition is a trial and error process. However, specific guidelines can be used to predict eluting conditions. For example, the characteristics of the components in the mixture and the type of stationary phase being employed should be considered. The solvent system used in chromatography is usually a mixture of organic solvent with water (**M. Ghahramani, 2015**). The ionization of analysts can be controlled by adding acids (HCl, HNO3, acetic acid) and bases (NH3). Different solvent systems can be used to identify compounds based on their chemical nature; for example, for amino acids, the solvent system used is acetic acid: water: n-butanol in the ratio of 1:5: 4. For the separation of sugar, a solvent system composed of ethyl acetate: pyridine: water; ethyl acetate is suitable. For inorganic ions, solvents like pyridine: water or HCl: water are more popular (**S. K. Pramod, 2017**).

1.3.4. Types of chromatography

1. Thin-layer chromatography
2. Column chromatography
3. Affinity chromatography
4. Ion exchange chromatography
5. Paper chromatography
6. Gas chromatography
7. Hydrophobic interaction chromatography
8. High-pressure chromatography
9. Pseudo affinity chromatography
10. Dye ligand chromatography
11. Gel permeation chromatography

1.3.4.1 Thin-layer chromatography (TLC)

Like all chromatography techniques, thin-layer chromatography is frequently used for organic chemical reactions and assays organic compounds' purity in photochemistry and biotechnology. TLC takes advantage of different affinities from the stationary and mobile phases to produce a complex mixture of organic molecules. TLC is an easy sensitive, and inexpensive technique used to determine the number of components in the mix and verify and identify the compounds' variety and purity to monitor the reaction's progress.

Thin-layer chromatography uses a thin glass plate coated with either aluminium oxide or silica gel as a solid phage; the mobile phase is a solvent chosen according to the property of the component in the mixture. The principle of TLC is the compound distribution between a solid fixed phase applied to a glass or a plastic plate and a liquid mobile phase, which is moving over a solid phase.

TLC is used to separate unknown mixtures by determining the pigment plant contain to detect the peptide and insecticides in food, to identify the combination present in the sample for the separation of amino acid TLC is used in the separation of vitamins.

1.3.4.2 APPLICATION OF TLC:-

The following application of TLC are mentioned below:;-

1. In monitoring the progress of the reaction.
2. Identify the compounds present in a given mixture
3. Determine the purity of a substance
4. Analyzing the fatty acids
5. Detection of pesticides and insecticides in food and water
6. Assaying the decomposition of fibre in forensics
7. Identification of medicinal plants and their consistence

1.3.4.3 ADVANTAGES OF TLC:-

1. TLC is a simple process with a short development time.
2. Help with the visualization of separated compound spots quickly.
3. It helps in the isolating of most of the compounds.
4. The separation process is faster, and the selectivity of compounds is higher.
5. The purity standard of the given sample can be assessed quickly.
6. It is a cheaper chromatography technique.

1.3.4.4 LIMITATION OF TLC:-

1. To identify specific compounds, the Rf value for the component of interest must be known beforehand
2. TLC plates don't have a prolonged stationary phase; therefore, the separation length is limited compared to other chromatography techniques.
3. It can't tell the difference between antimeres and some isomers.

REVIEW OF LITERATURE

2   <u>REVIEW OF LITERATURE</u>

Molecular marker methods have broad applications in several plant species for diverse purposes such as genetic characterization and marker-assisted selection. Molecular studies using DNA markers (RAPD, ISSR, AFLP, SCoT etc.) are well known in *Cannabis* research (**Kojoma et al., 2011**). ISSR is a technique that is the simplest method with a high level of reproducibility. It involves PCR amplification using a short sequence of primer. Meanwhile, it is simple, fast, cost-effective, highly discriminator and highly reliable and is widely used in plant genetic analysis and shows improved stability and polymorphism (**Qian et al., 2001**).

ISSR markers are a powerful tool for investigating genetic variations within and among species. Similarly, a new random novel DNA marker system called Start Codon Targeted (SCoT) Polymorphism (**Collard and Mackill, 2009**) was developed based on the short conserved flanking sequences of the ATG start codon in plant genes. SCoT markers are highly reproducible, and the primer length and annealing temperature are not the only factors determining reproducibility. These are dominant markers like RAPD and ISSR and could be used for genetic analysis, quantitative trait loci (QTL) mapping and bulk segregation analysis (BSA) (**Collard and Mackill, 2009**).

The genetic diversity using ISSR markers in different crops was easily distinguishable at the species level and based on their geographical distribution and climatic similarity (**Kaivan Bahmani et al., 2012**). The genetic diversity study by ISSR markers in 25 various types of Iranian Fennels population with 7 ISSR primers showed a total of 52 amplified fragments, among which 49 were polymorphic (**Kaivan Bahmani et al., 2012**).

In the same way, the molecular diversity and genetic structure of 29 species of Indian Teak plants were analyzed using ISSR markers (UBC 801, 880, 899, 900). The primers mainly used showed 100% polymorphism, and the highest genetic diversity was led by primer UBC 900 (**S.A Ansari et al. 2012**). They finally concluded that 43 polymorphic bands were amplified among 29 teak species, resulting in approximately 8.6 bars per primer (**S.A Ansari et al. 2012**). To identify the ISSR markers polymorphism in *Lycopersicon* to investigate the polymorphism

level using ISSR. The 5 species of tomatoes, such as *L. esculentum, L. cheesmanii, L. pennellii, L. hirsutum* and *L. humboldtii* by using 14 primers. Among these 14 primers only 9 primers were capable of differentiating all species of tomato. All 9 primers produce different fingerprints bands of all species of tomato. Maximum bands were polymorphic. Therefore, they obtained special fingerprints of ISSR for every species of *Lycopersicon* (**Yu. M. Tikunov et al. 2003**).

According to M.Z. Galvan et al., (2003), the genetic diversity study of *Phaseolus vulgaris* (common bean) using RAPD and ISSR markers, showed the relationship among 10 species of bean that originated in Argentina. The result encompassed the RAPD marker results with the results of ISSR marker and noticed that ISSR markers are best useful tool than RAPD markers in identifying beans by origin (**M.Z. Galvan et al. 2003**).

Similarly, the genetic variation study using RAPD and ISSR markers between wild rice populations of *Oryza granulata* from China shows ISSR is superior to RAPD. They took five rice populations from different places in China and used 20 RAPD primers which gave 199 bands. Among them, 61 were polymorphic. While 12 ISSR primers gave 113 bands, 52 were polymorphic (**W. Qian et al. 2000**). Therefore, RAPD and ISSR both contributed very low genetic diversity level of their wild populations. Later on, further studies showed that the total detected polymorphic bands with the help of ISSR are much higher than those seen with the help of RAPD. So, they concluded that ISSR is much superior to RAPD in detecting polymorphism (**W. Qian et al. 2000**).

Iruela et al., (2001) estimated the phylogenetic analysis using RAPD and ISSR markers in *Cicer* and Chickpea. They mainly analyzed 75 different areas of 14 species of *Cicer* with the help of PCR molecular markers. They analyzed using RAPD that 234 polymorphic fragments were given with the help of 12 primers. They concluded changes among the species. They were only able to find 26 fragments from total of 234 fragments that were monomorphic and specific (**Iruela et al., 2001**).

Martins et al., (2013) spotted the genetic relatedness using RAPD and ISSR markers of Portuguese almond. They used 60 RAPD and 18 ISSR primers. Out of those, only 6 RAPD and 5 ISSR primers were tested for high polymorphism and reproducibility.

However, 124 PCR fragments were estimated, and from that, mainly 120 fragments were polymorphic (**Martins et al. 2013**).

Also, the study by Subramanian et al. showed the genetic diversity by using RAPD and ISSR markers in black gram. They used 25 RAPD and 16 ISSR primers. RAPD primers gave rise to 104 fragments. From those, 44 were polymorphic. ISSR primers gave 101 bands. Total percentage polymorphism that ranges among RAPD primers was maximum up to 66 per cent and that of ISSR primers up to 100 per cent. So, in their study, they revealed that ISSR markers were much more effective than RAPD markers, with an average polymorphism percentage of 57% of ISSR as compared to 42.7% polymorphism of RAPD (**Subramanian et al. 2004**).

It was found that the genetic stability of *Cannabis* species using ISSR markers was reproducible. The evaluation of micro-propagated *Cannabis* with 15 ISSR primers gave 115 prominent and reproducible bands. They also evaluated that all the micro-propagated plants existed in one form similar to mother plants (**Hemant Lata et al., 2010).** The clonal mass propagated cannabis species showed that mainly six cannabinoids were identified. The analysis of some micro-propagated plants using ISSR marker in combination with donor mother plant of Cannabis and some selected primers such as UBC 808, UBC 807 and UBC 811 showed similarity with parent plant (**Hemant Lata et al., 2010**). The plants did not show any differences from the donor plant in morphological features such as the shape of the leaf, height of the plant and pattern of branching (**Hemant Lata et al., 2010**).

Similarly, Hemant Lata et al., (2011) observed the genetic stability of synthetic seeds of *Cannabis* using ISSR fingerprinting under different growth conditions. They used 14 ISSR markers for molecular analysis. A total of 9 primers out of 14 produced 40 different bands (**Hemant Lata et al., 2011**). Genetic stability was confirmed among the mother plants and monomorphic ISSR profiles. The GC analysis shows major differences in cannabinoids content of mainly six cannabinoids which are cannabidiol, cannabigerol, cannabinol, THC, cannabichromene and tetrahydrocannabivarin. Therefore, allowing the pure genetic stability of plants (**Hemant Lata et al., 2011**).

According to Khatak et al., (2016) recognized the specific diversity analysis in various genotypes of *Cannabis* using the ISSR marker. They distinguished between

fibre and drug hemp types of *Cannabis* using molecular fingerprinting using various markers. They mainly collected 14 samples of Cannabis from the edges of roads of Northern India and observed genetic diversity with the help of 28 primers. These 28 primers originated 152 fragments. Of those, 132 were polymorphic, resulting in approximately 10.14 bands per primer (**Khatak et al. 2016**).

Kojoma et al., (2002) identified various samples of *Cannabis sativa* using DNA-based ISSR fingerprinting. They characterize three chemical compounds using HPLC into two types which are tetrahydrocannabinol (THC) and cannabidiol (CBD). They were not able to identify two samples of CBD with the help of HPLC. Then they use ISSR fingerprinting to distinguish DNA patterns among various samples. Differentiation of cannabis samples can be easily achieved by ISSR fingerprinting, which is very typical of the HPLC process (**Kojoma et al. 2002**).

Similarly, a new random novel DNA marker system called Start Codon Targeted (SCoT) Polymorphism (**Collard and Mackill, 2009**) was developed based on the short conserved flanking sequences of the ATG start codon in plant genes. SCoT markers are highly reproducible, and the primer length and annealing temperature are not the only factors determining reproducibility (**Collard and Mackill, 2009**). SCoT markers represent a powerful tool for investigating genetic variations within the species and among the species of several medicinal and aromatic plants.

The Genetic variability study of *Dendrobium nobile* Lind., using a Start Codon Targeted (SCoT) marker system, showed 96.21% polymorphism. Six natural populations were investigated by SCoT primer and produced one hundred and thirty-two (132) bands. Cluster analysis also reveals high genetic variation among the genotypes. The investigation also showed the effectiveness of the SCoT marker system to estimate the genetic diversity of *D. Nobile*, an endangered orchid species of medicinal importance (**Bhattacharya et al. 2013**). Likewise, the genetic fidelity of i*n-vitro* raised plantlets of *Spondias pinnata* (L. f.) Kurz, was confirmed by SCoT and ISSR molecular markers. The genetic fidelity of in vitro regenerated plants showed 12.5% polymorphism (**Jaiswal et al. 2020**).

Silybum marianum, a popular medicinal plant, was used with Start Codon Targeted (SCoT) marker system to investigate the genetic variability. SCoT marker produced 255 amplicons and 84.03% polymorphism a with PIC value was 0.43. The primers

showed resolving power values between 4.18 and 7.84, and the percentage of polymorphic bands was between 33.3 and 100%. Authors suggested that SCoT marker system could effectively evaluate the genetic diversity of milk thistle genotypes (**Rafizadeh et al. 2017**).

Rajesh et al., (2015) evaluated a simple and novel marker system, starting codon targeted polymorphic (SCoT) markers for assessment of genetic diversity in 23 coconut accessions (10 tall and 13 dwarfs) from different geographical regions. A total of 102 reproducible bands were generated by the 15 primers, which showed 88 % polymorphism. Tall and dwarf coconut accessions were separated, and accessions from the same geographical region were grouped. Authors concluded that the potential of SCoT markers to be utilized as molecular markers to detect DNA polymorphism in coconut accessions (**Rajesh et al. 2015**).

A comparative study of start codon targeted (SCoT) and inter-simple sequence repeats (ISSR) markers were used for genetic diversity and relationship analysis of nine *Salvia* species. Some selected ISSR (21) and SCoT (20) primers amplified 350 and 329 reproducible bands, respectively, among them all polymorphic. The dendrogram generated based on SCoT separated the individuals into sub-clusters in accordance with their species and section. The results also showed that the SCoT markers could be a consistent and reproducible marker for estimating genetic diversity and relationships among *Salvia* species (**Etminan et al 2018**).

3   <u>OBJECTIVES:</u>

The detailed literature survey revealed that no scientific reports have been made on molecular marker studies with the help of ISSR and SCoT markers on different accessions of these plants in the Indian population.

In the present study, we used ISSR and SCoT markers to investigate genetic diversity and TLC for phytochemical fingerprinting in *Cannabis*. This work was undertaken to screen Indian *Cannabis* germplasms for genetic variability study; with the following objectives:

1. ISSR and SCoT marker-based molecular characterization of *Cannabis* accessions.
2. Phytochemical evaluation of *Cannabis* accessions using thin-layer chromatography.

4 <u>MATERIAL AND METHODS</u>

The present investigation, "**Molecular and Phytochemical Diversity analysis in accessions of *Cannabis* sp.**" was carried out at the Botany and Pharmacognosy Lab., Plant Breeding and Genetic Resource Conservation Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, from February 2022 to June 2022.

4.1 Plant Material:

Twenty (20) different accessions of *Cannabis* sp*.* plant (CSNK-11 to CSNK-30) were provided from the Botany and Pharmacognosy Lab., Department of Plant Breeding and Genetic Resource Conservation, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow. The plants were studied in detail during their transition from the vegetative to the reproductive stage. Immature leaf materials collected in the reproductive phase were used for DNA extraction. The individual samples were stored at -20°C prior to use.

4.2. Genomic DNA Extraction and PCR Amplification:

Reagents for DNA isolation and purification:

- **Tris-Cl:** pH 8.0 (1.0 M); EDTA pH 8.0 (0.5 M); NaCl (5.0 M); CTAB (20%); Chloroform: Isoamyl alcohol (24:1 v/v); Polyvinylpyrrolidone; and b mercaptoethanol.
- **Extraction buffer:** 100 mMTris-Cl (pH 8.0), 25 mMEDTA, 1.5MNaCl, 2.5% CTAB, 0.2% b-mercaptoethanol (v/v) (added immediately before use) and 1% PVP (w/v) (added immediately before use).
- **High salt TE buffer:** 1 M NaCl, 10 mM Tris-Cl (pH 8.0) and 1 Mm EDTA.
- Liquid Nitrogen for pulverization of leaves.
- Taq assay buffer
- Template DNA
- dNTPs mix
- Forward primer
- Reverse primer
- Taq DNA polymerasE

- Ethidium bromide (ETBr)

4.3 INSTRUMENTS:

- Proflex PCR System
- Gradient cycler
- Nanodrop spectrophotometer
- Water bath
- Milli Q-Biocel system (Millipore)
- Autoclave
- Gel electrophoresis unit
- Incubator (for 1.5ml centrifuge tubes)
- Centrifuge
- Pipette(Eppendorf)
- Mortar and pestle

4.4 PROCESS OF DNA ISOLATION:

The following steps are involved in the process of DNA isolation,

- Grind the plant material in liquid nitrogen (3 g of fresh tissue or 0.5 g of dry tissue).
- Transfer the material to 1.5ml microcentrifuge tubes and add 600ul of freshly prepared extraction buffer, mix by inversion to a slurry.
- Incubate at 60 °C in a shaking water bath (100 rpm) for 1–2 h (dry samples may require overnight incubation at 37°C).
- Centrifuge at 10,000 rpm for 10 min. Transfer the supernatant to new 1.5ml microcentrifuge tubes.
- Add 500ul of chloroform: isoamyl alcohol (24:1) and mix by inversion for about 15 min.
- Spin at 10,000 rpm for 10 min at 25–30°C.
- Carefully transfer the upper transparent aqueous layer to another 1.5ml microcentrifuge tube.
- Add 3ul of RNase solution and incubate at 37 °C for 30 min.
- Then centrifuge it at 10,000rpm for 10 minutes.

- Transfer the supernatant to new 1.5ml microcentrifuge tubes.
- Add 600ul of isopropanol and invert it very gently to make the fibrous thread-like strands of DNA formed, and then let the mixture stand at room temperature for 1h.
- Centrifuge at 13,000rpm for 10 minutes and discard the supernatant carefully, leaving the DNA pellet behind.
- Then add 500ul of 80% Ethanol to wash the DNA. Gently invert the tubes to wash them.
- Then centrifuge it at 13,000rpm for 1minute and remove the ethanol carefully.
- Airs dry the pellet for 30 minute and then adds 100ul of MQ.
- Leave it for overnight at 40°C to dissolve the DNA properly.

4.5 Polymerase chain reaction:

A polymerase chain reaction is a method widely used to rapidly make copies of a specific DNA sample, allowing scientists to take a very small sample of DNA and amplify it to a large enough amount to study in detail.

- Reagents for PCR:
  DNA template, Primer, PCR Master Mixture (DNTPs, Buffer, $MgCl_2$, Taq polymerase), Nuclease free water
- Instrument for PCR:
  Pipettes, Minispin, Centrifuge, Thermal cycler

4.6 ISSR marker analysis

The DNA samples were subjected to ISSR assay, using 10 arbitrary primers (**Table 1**) with GC content of 60%. Different concentrations and combinations of PCR components like dNTPs, MgCl2, and primer concentration were used in different PCR cycles to optimize the best PCR conditions and to obtain the best amplification results. The amplification reaction mixture contained 400 μM each of dNTPs, 10 picomoles primers, 0.9 mM $MgCl_2$, 1U Taq polymerase (Thermo Fishers), 2.5 μL of Taq buffer and 50 ng of genomic DNA (1.0 μL). PCR amplification was carried out at 94°C (5 min), followed by 30 cycles at 94°C (1 min), 48.0 to 60°C (45 sec) and 72°C (1 min), with a final extension for 5 min at 72°C. Reproducibility of the ISSR primers was tested by repeating PCR reactions for at least three times under the same PCR

conditions with the same set of chemicals. All the reactions were carried out in Takara, PCR system. After completion of PCR all the amplified products were stored at -20°C.

Table 1: List of ISSR primers and their details, used in the study

Primer Code	Primer Sequence	BP	T_m (°C)	Annealing temp (°C)
ISSR 806	AGAGAGAGAGAGAGAGT	17	50.36	45
ISSR 807	AGAGAGAGAGAGAGAGC	17	52.77	47
ISSR 809	GAGAGAGAGAGAGAGAT	17	50.36	45
ISSR 814	CTCTCTCTCTCTCTCTG	17	52.77	48
ISSR 821	TCTCTCTCTCTCTCTCA	17	50.36	45
ISSR822	TCTCTCTCTCTCTCTCC	17	52.77	48
ISSR 823	TCTCTCTCTCTCTCTCG	17	52.77	48
ISSR 824	ACACACACACACACT	15	50.36	45

4.7. SCoT marker analysis

Twelve SCoT primers-PCR (**Table 2**) were carried out in a 10 µl reaction volume containing 25−50 ng of template DNA, 10X PCR buffer, 1.5-2.5 mM of each dNTP's, 10 pmol of primer and 2- 3U of Taq DNA polymerase.

The PCR amplifications were performed on a thermal cycler (ProFlex, Applied Biosystems) programmed for initial denaturation at 94 °C for 3 min; 40 cycles of denaturation at 94 °C for 1 min, annealing at 50 °C (for all primers) for 1 min and extension at 72 °C for 2 min; final extension at 72 °C for 5 min and 12 °C forever. The molecular size of the amplified products was determined against a 100 bp −3 kb DNA Ladder (Thermo fisher).

Table 2: List of SCoT primers and their details, used in the study

Primer Code	Primer Sequence	BP	T_m (°C)	Annealing temp (°C)

SCOT1	CAACAATGGCTACCACCA	18	53.69	50
SCOT2	CAACAATGGCTACCACCC	18	55.97	50
SCOT-4	CAACAATGGCTACCACCT	18	53.69	50
SCOT-5	CAACAATGGCTACCACGA	18	53.69	50
SCOT-7	CAACAATGGCTACCACGG	18	55.97	50
SCOT-8	CAACAATGGCTACCACGT	18	53.69	50
SCOT-11	AAGCAATGGCTACCACCA	18	53.69	50
SCOT-12	ACGACATGGCGACCAACG	18	58.24	50

4.8. Detection of the specific fragment(s)

Ten µL of amplified PCR products were mixed with 5 µl of DNA loading dye (0.2% (w/v) bromophenol blue, 0.2% (w/v) xylene cyanol FF, 40% (w/v) sucrose). PCR amplified products were run on 1.2 % (w/v) agarose (Himedia, India) gel at 65 amp at a constant current for 2–4 h in the presence of 1X TBE buffer (40 mM Tris (pH 7.6), 20 mM boric acid, 1 mM EDTA). After agarose gel electrophoreses were completed, DNA fragments on the gels were stained with ethidium bromide (0.5 µg µl^{-1}) and documented by a Gel Documentation System. Molecular weight markers were used to compare duplicate reactions. All experiments were repeated at least two times to ensure reproducibility and consistency.

4.9. Phytochemical analysis (Thin Layer Chromatography)

4.9.1 Sample preparation:

Each dried sample were taken about 100 mg and crushed into coarse powder and soaked in acetonitrile (ACN) (100mg/ml) for overnight at dark then centrifuged. The supernatant was transferred to a new Eppendorf tube and extract was kept at -20°C for further use.

4.9.2 Chromatographic method:

The dried ACN extract was applied to silica gel TLC plate and developed in Hexane: diethyl ether (8:2) in twin trough chamber (CAMAG), pre-saturated with the same

mobile phases for 20 min. Then the plates were air dried and visualized under a UV cabinet (UVP, Germany) at 254nm and 366nm wavelengths.

4.10 Data analysis:

Twenty accessions of *Cannabis* were examined using ISSR amplified bands through the presence (1) and absence (0) matrix. The numbers of monomorphic and polymorphic bands generated by each ISSR primer were noted. The genetic similarity was calculated using the Sorensen-Dice coefficient. The similarity matrix of phytochemical data and ISSR data were used separately to construct the UPGMA (unweighted pair group method with the arithmetic average) dendrograms using the Numerical Taxonomy System, (NTSYS-pc, version 2.0) (Rohlf 1990).

RESULTS

5   <u>**RESULTS**</u>:

The observations of experimental findings obtained during the present investigation were presented in this chapter. Observations were made for various parameters and analyzed.

5.1 Morphological study of *Cannabis* sp.:

The plant accessions are highly polymorphic in growth habit, height, spread, leaf shape, stem width, flower colour, flower diameter and inflorescence, etc.

5.1.1 Habit:
The plants are annual tall and herbaceous—dense primary and secondary branches with pubescent **(Figure 1 a, b)**.

5.1.2. Leaves
Leaves alternate palmate; petiole 2-7 cm long; leaf blade green, strigose, and with dense trichomes; leaflets usually lanceolate to linear, varies in size with longest in middle, margin serrate, apex acuminate.

5.1.3. Male inflorescences
Male inflorescences consist of male flowers that arise on the axils**.** Male flowers: yellowish-green, pedicel 2-4 mm, thin; sepals ovate to lanceolate, membranous with sparse hairs; petals absent; filament 0.5-1 mm; anthers oblong.

5.1.4. Female inflorescences
Female inflorescences crowded in apical leaf axils among leaf-like bracts and bracteoles. Female flowers: green, sessile; calyx sparsely pubescent; ovary globose, enclosed by appressed calyx, surrounded closely by bract and bracteoles. Persistent bracts yellow (**Figure 1 c,d.**

5.1.5. Fruit
Fruits are achene type, oval-shaped, hard with persistent bract and greenish-brown in colour.

5.1.6. Seed
The seeds are oval to round in shape, seed coat hard with surface ornamentation, endospermic with curled embryo and brownish in colour.
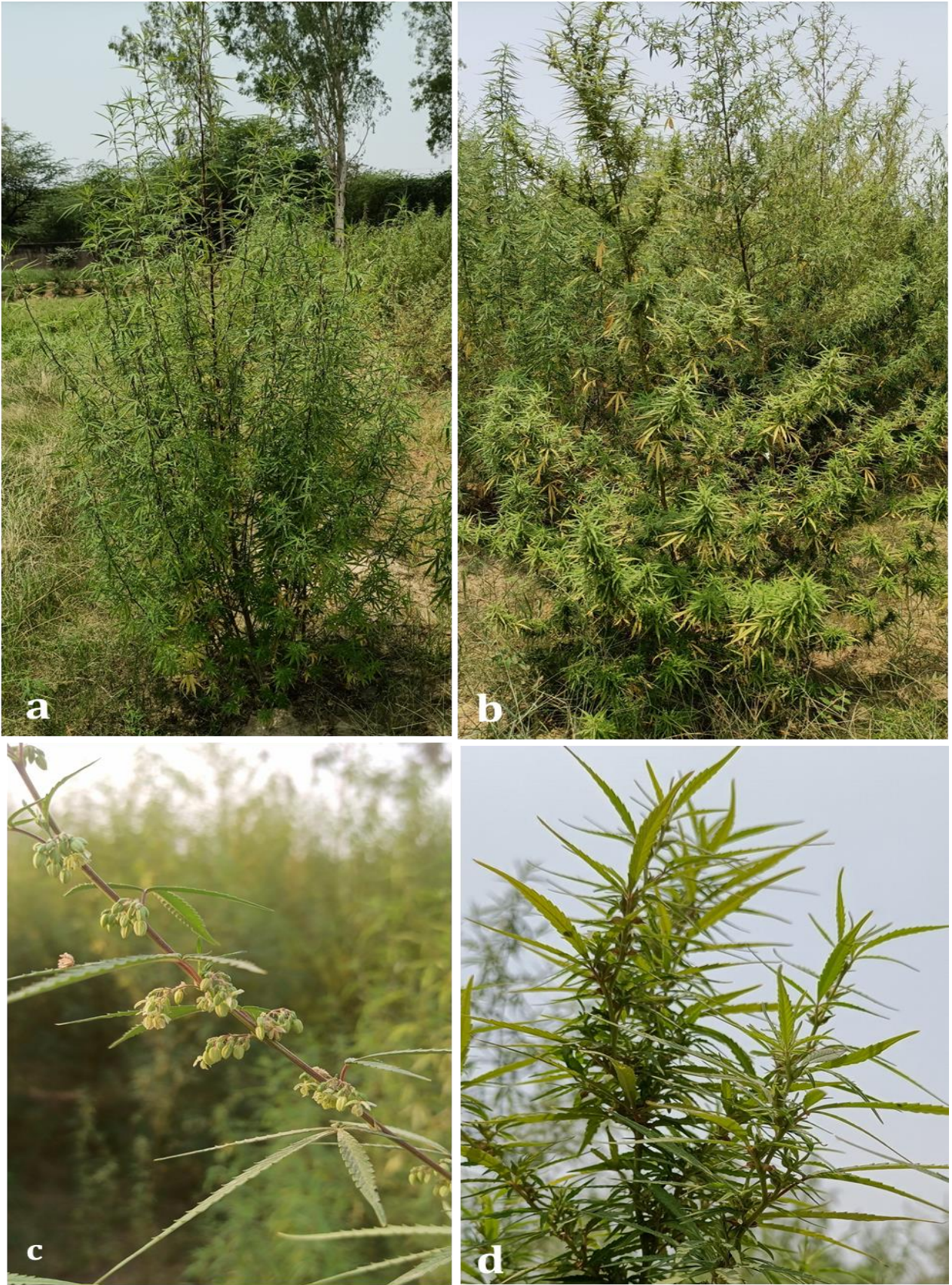
Figure 1: Habit of *Cannabis* sp. **a.** CSNK-15, **b.** CSNK-23, **c-d** inflorescence.

5.2. Molecular study of *Cannabis* accessions

5.2.1 Isolation and purification of genomic DNA

The analysis of molecular diversity of the twenty accessions of *Cannabis* sp. young leaves were used for genomic DNA isolation. Genomic DNA was isolated with the help of modified CTAB method. An optimum DNA isolation was possible from 500 mg of leaf samples with a 2% CTAB buffer solution added with PVP and mercaptoethanol. The adopted method for the isolation of DNA from the leaves shows a good yield of DNA (**Figure 2**). The O.D. value (A260/A280 ratio) of the DNA is found to be in the range of 1.7-1.9, which indicates that the DNA is relatively very pure with the least protein or phenol contamination. The yield and quality of isolated DNA from the leaves are shown in **Table 3.** The quality of isolated genomic DNA of samples was checked on 0.8% agarose gel (**Figure 3**). The isolated DNA sample stock was stored at –20ºC and it was suitably diluted to prepare a working sample of DNA, which was used for further PCR studies. Dilution of stock DNA is made up of PCR-grade water so that the working solution contains about 50ng/µL DNA.

Table 3: Yield and quality of genomic DNA isolated from 20 accessions of *Cannabis* sp.

| Sample	260/280 Ratio	260/230 Ratio	Yield (ng/µl)
CSNK-11	1.69	1.42	1806.4
CSNK-12	1.69	1.49	2529.7
CSNK-13	1.76	1.68	1517.2
CSNK-14	1.71	1.63	2850.5
CSNK-15	1.79	1.52	1548.3
CSNK-16	1.74	1.38	2002.1
CSNK-17	1.79	1.6	1241.4
CSNK-18	1.79	1.63	1649.4
CSNK-19	1.8	1.72	2402.8
CSNK-20	1.78	1.85	1667.9
CSNK-21	1.79	1.79	2232.1
CSNK-22	1.68	1.5	1160.2
CSNK-23	1.72	1.54	1198.7
CSNK-24	1.63	1.26	1317.1
CSNK-25	1.72	1.51	1148.5
CSNK-26	1.66	1.67	1953.8
CSNK-27	1.71	1.32	758.3
CSNK-28	1.75	1.34	1049.7
CSNK-30	1.65	1.08	1338.8
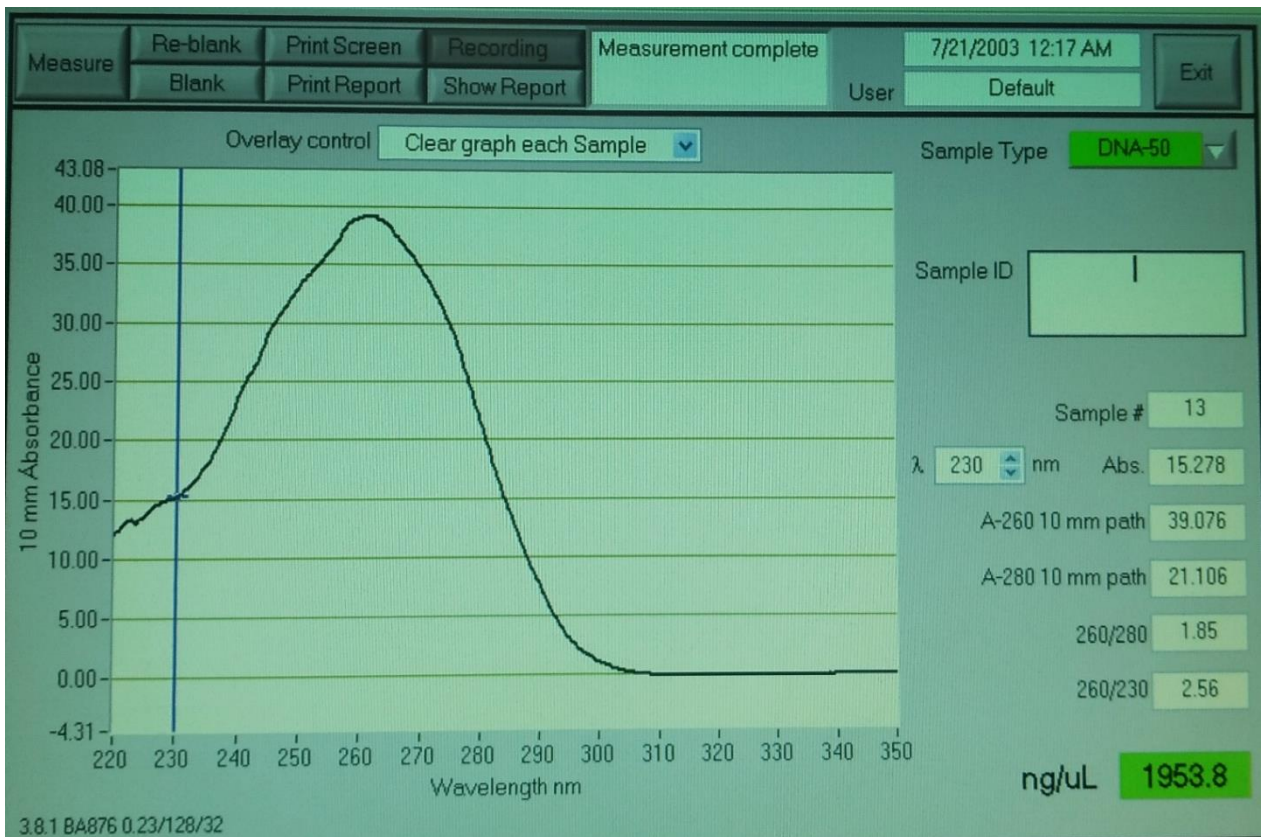
Figure 2: A graph of UV spectrophotometer showing absorbance at 260nm and 280nm of isolated genomic DNA.
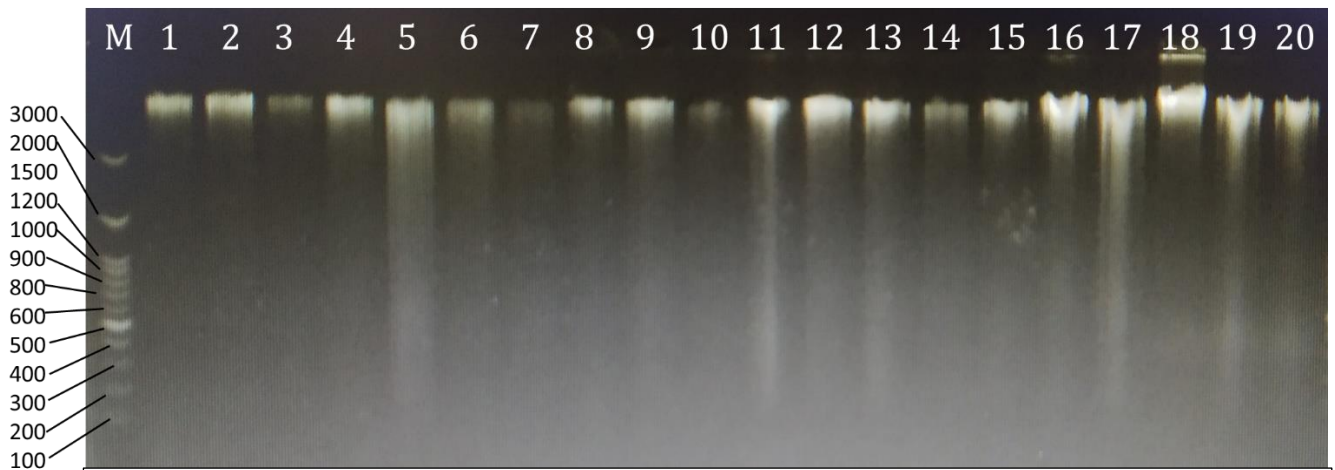

Figure 3: Genomic DNA isolated from leaves loaded on 0.8% agarose gel. L1-L20: Genomic DNA from leaves of *Cannabis* accessions. M: 1kb plus ladder

5.2.2 <u>ISSR Marker analysis of *Cannabis* accessions</u>

5.2.2.1. Optimization of PCR condition for ISSR analysis

The optimization of inter-simple sequence repeats (ISSR-PCR) was done by testing several PCR parameters, including a concentration of $MgCl_2$, annealing temperature, and primer concentration (**Table: 4**). The total volume of the reaction mixture was 20µl. Ten ISSR primers were selected and used for screening. Agarose gel electrophoresis was run using 1.2% gel. The reproducibility of this method was checked by repeating the procedures three times under optimal conditions.

Table 4: Optimized PCR condition for ISSR reaction (20µl)

Components	Concentration
PCR buffer (10X)	1X
Primer concentration	10 picomoles/µl
DNA template	50 ng/µl
MgCl2	0.9 mM
dNTP mix	200 µM
Taq DNA polymerase	1U

5.2.2.2. Generation of the molecular profile of *Cannabis* accessions with ISSR primers

ISSR-PCR system is the simplest method with a high level of reproducibility. It involves PCR amplification using short sequence of primer. It is widely used in plant genetic analysis and show improved stability and polymorphism for diversity analysis (**Qian et al., 2001**).

In the present study, amplification of the genomic DNA of 20 accessions of *Cannabis* species was carried out using 10 selected ISSR markers. The optimum annealing temperature for these ISSR markers varied from 48.7 to 60°C (**Table 1**). Out of 10 selected primer, only 8 markers produced reproducible and scorable bands (**Figure 4**). These primers amplified the fragments across the accessions of the studied species, with several amplified fragments per reaction ranging from 06 to 69. These bands varied from 300 to 2000 bp. The dendrogram based on UPGMA analysis

grouped the accessions of *Cannabis* sp. with Sorensen-Dice coefficient ranging from 0.67 to 0.94 between the 20 accessions of *Cannabis*. (**Figure 5**).
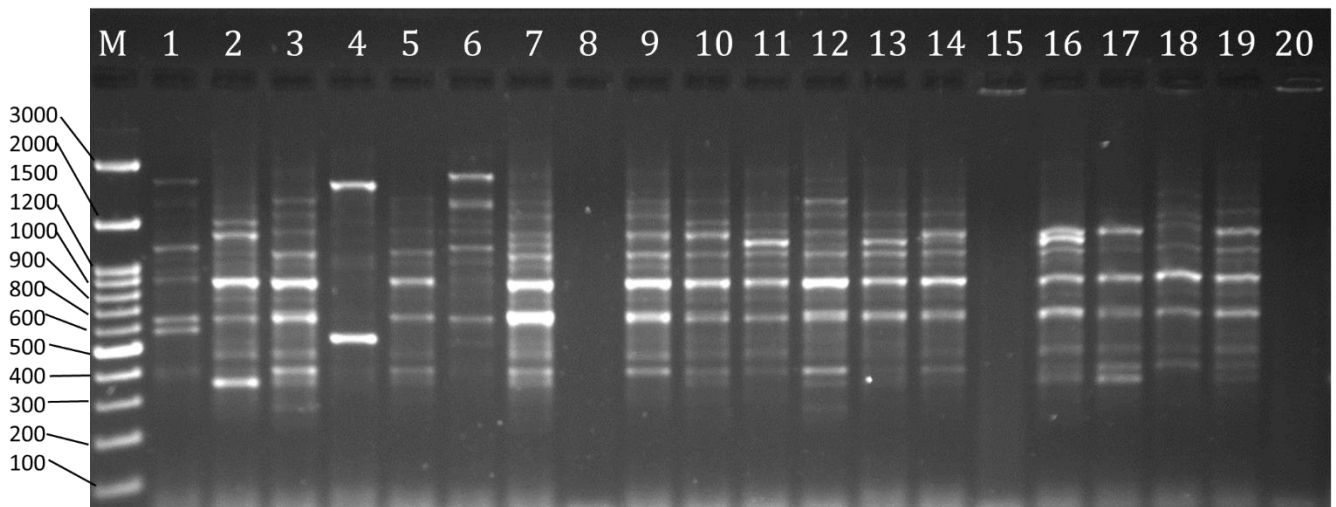
Figure4: PCR-ISSR-809 loaded on 1.2% agarose gel. L1-L20: Genomic DNA from leaves of *Cannabis* accessions. M: 1kb plus ladder
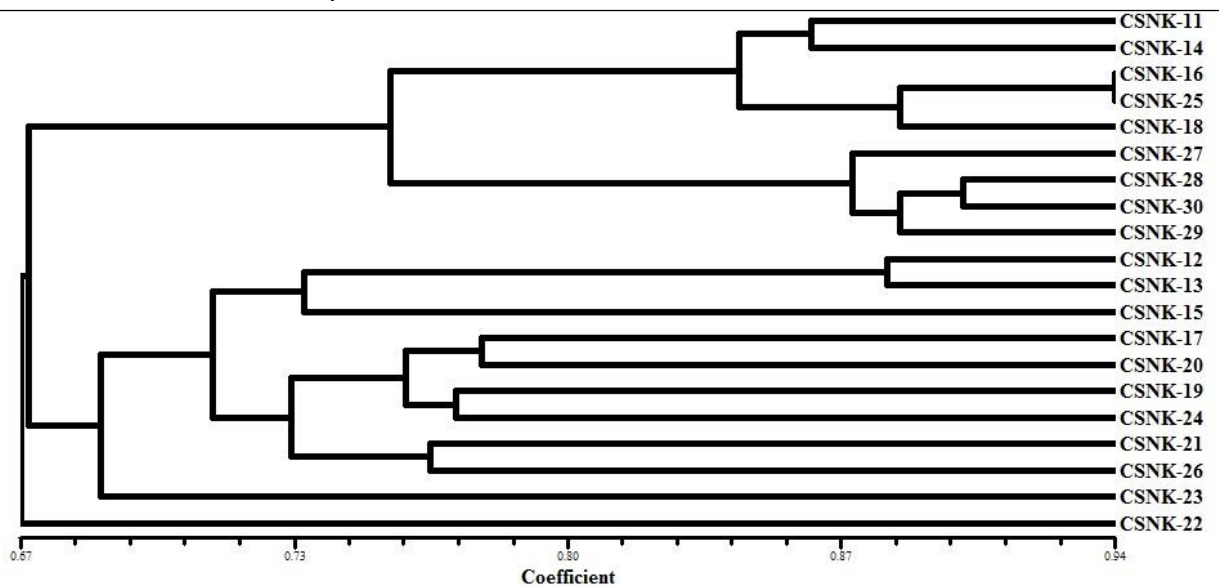
Figure 5: UPGMA dendrogram based on 8 ISSR primer showing similarity relationship between *Cannabis* accessions CSNK-11 to CSNK-30.

5.2.2.3. Cluster analysis of *Cannabis* accessions with ISSR markers

The dendrogram of genetic relationships between the germplasms was constructed on the basis of the similarity matrix by the UPGMA method using binary data of ISSR primers. The dendrogram showed clustering of the 20 accessions into three main clusters (cluster I, II and III) with a similarity coefficient of 0.67 to 0.94. Cluster-I

again separated into two sub-clusters. Sub-cluster-IA consisted of CSNK-11, 14, 16, 25, and 18, and sub-cluster-IB consisted of CSNK-27, 28, 30 and 29. Within sub-cluster-IA, CSNK-16 and 25 are closer, with a 0.94 similarity coefficient. Cluster-II is comprised of three sub-clusters. Likewise, sub-clusters-IIA comprises CSNK-12, 13, 15. Sub-clusters-IIB contained maximum accessions with CSNK-17, 20, 19, 24, 21 and 26. While sub-cluster-IIC included only CSNK-23. A single accessions CSNK-22 made a distinct I in cluster-III, representing a 0.67 similarity coefficient with other accessions (**Table 5).**

Table 5: Clustering analysis through ISSR marker study

Cluster	Sub-cluster	Number of accessions
I	I-A	CSNK-11, 14, 16, 25, 18
	I-B	CSNK-27, 28, 30, 29
II	II-A	CSNK-12, 13, 15
	II-B	CSNK-17, 20, 19, 24, 21, 26
	II-C	CSNK-23
III	--	CSNK-22

5.2.3. <u>SCoT Marker analysis of Cannabis accessions</u>

5.2.3.1 Optimization of PCR condition for SCoT analysis

The optimization of start codon targeted polymorphism (SCoT-PCR) was done by PCR protocol described in Collard & Mackill, (2009). The total volume of reaction mixture was 10µl (**Table 6).** Twelve selected primers were used for screening. A standard PCR cycle was used: an initial denaturation step at 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 2 min; the final extension at 72°C was held for 5 min. Agarose gel electrophoresis was run using 1.2% gel. Reproducibility of this method was checked by repeating the procedures three times under optimal condition.

Table 6: Optimized PCR condition for SCoT reaction (10µl)

Components	Concentration
PCR buffer (10X)	1X
Primer concentration	10 picomole/µl
DNA template	25 ng/µl
MgCl2	1.0 mM
dNTP mix	200 µM
Taq DNA polymerase	1U

5.2.3.2. Generation of molecular profile of *Cannabis* accessions with SCoT primers

Likewise, SCoT marker, a new random novel DNA marker system called Start Codon Targeted (SCoT) Polymorphism (**Collard and Mackill, 2009**) was developed based on the short conserved flanking sequences of the ATG start codon in plant genes. SCoT markers represent a powerful tool for investigating genetic variations within the species and among the species.

In the present study, amplification of the genomic DNA of 20 accessions of *Cannabis* species was carried out using 12 selected SCoT markers. The optimum annealing temperature for these SCoT markers varied from 48 to 50°C (**Table 2**). Out of 12 selected primer, only 8 markers produced reproducible and scorable bands. These primers amplified the fragments across the accessions of the studied species, with several amplified fragments per reaction ranging from 06 to 69 (**Figure 5**). These bands varied from 300 to 2000 bp. The dendrogram based on UPGMA analysis grouped the accessions of *Cannabis* sp. with Sorensen-Dice coefficient ranging from 0.67 to 0.94 between the 20 accessions of *Cannabis* (**Figure 6**).
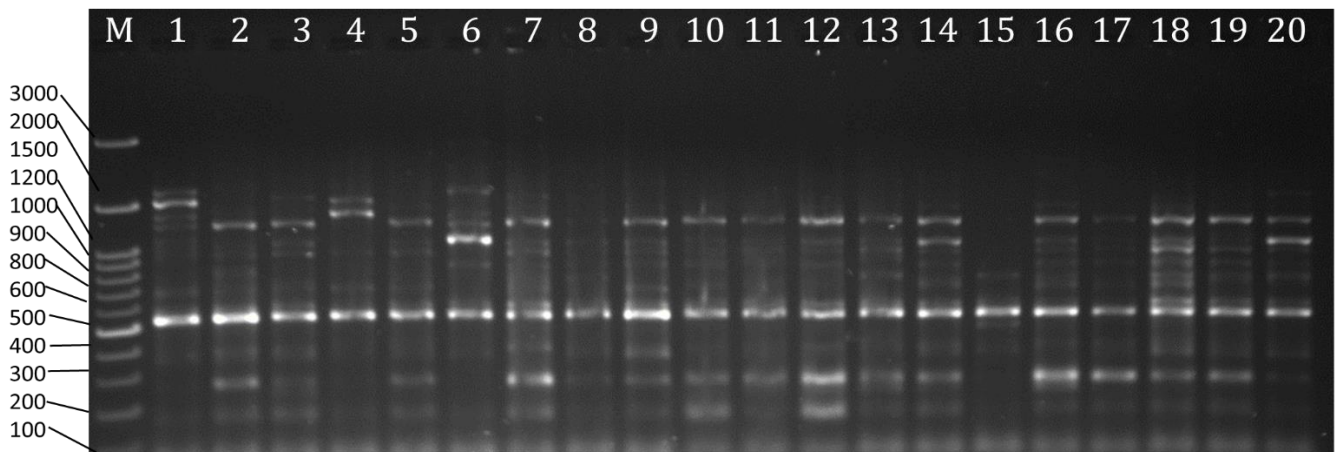
Figure 6: PCR-SCoT-2 loaded on 1.2% agarose gel. L1-L20: Genomic DNA from leaves of *Cannabis* accessions. M: 1kb plus ladder
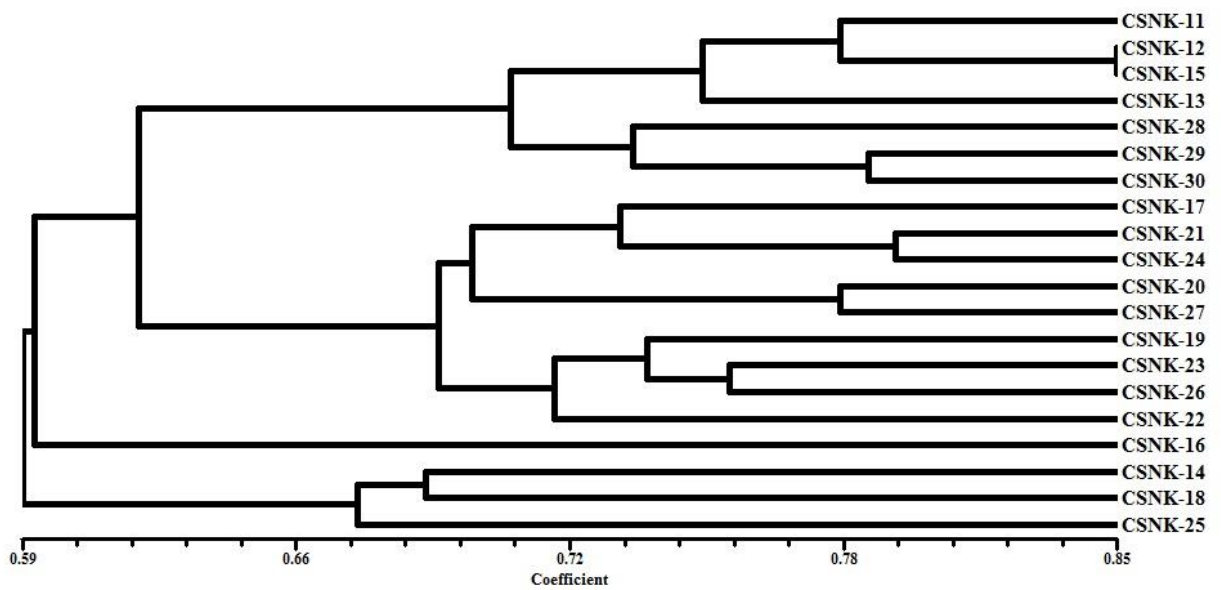
Figure 6: UPGMA dendrogram based on 8 SCoT primer showing similarity relationship between *Cannabis* accessions CSNK-11 to CSNK-30.

5.2.3.3. Cluster analysis of *Cannabis* accessions with SCoT markers

The dendrogram of genetic relationships between the germplasms was constructed on the basis of the similarity matrix by the UPGMA method using binary data of SCoT primers. The dendrogram was showed clustering of the 20 accessions into

four main clusters (cluster I, II, III and IV) with similarity coefficient of 0.59 to 0.85. Cluster-I again separated into two sub-clusters. Sub-cluster-IA consisted of CSNK-11, 12, 15 and 13, and sub-cluster-IB consisted of CSNK-28, 30 and 29. Within sub-cluster-IA, CSNK-12 and 15 closer to each other, with a 0.85 similarity coefficient. Cluster-II comprised with three sub-clusters. Likewise, sub-clusters-IIA comprises CSNK-17, 21 and 24. Sub-clusters-IIB contained maximum accessions with CSNK-20 and 27. While, sub-cluster-IIC included CSNK-19, 23, 26 and 22. A single accession CSNK-16 made a distinct OTU in cluster-III representing a 0.59 similarity coefficient with other accessions. Similarly, cluster IV includes CSNK-14, 18 and 25 with a 0.68 similarity coefficient with other accessions (**Table 7**).

Table 7: Clustering analysis through SCoT marker study

Cluster	Sub-cluster	Number of accessions
I	I-A	CSNK-11, 12, 15, 13
	I-B	CSNK-28, 29,30
II	II-A	CSNK-17, 21, 24
	II-B	CSNK-20, 27
	II-C	CSNK-19, 23, 26, 22
III	--	CSNK-16
IV	--	CSNK-14, 18, 25

5.3. Phytochemical characterization

The chemical fingerprinting of twenty accessions of *Cannabis* species were profiled. TLC experimental condition was optimized by the single solvent system with different ratio of solvents. The presence and absence of major chemical component in the aerial part was examined for differentiation of diversity among the species. Reasonably clear separation of compounds with the mobile phase, hexane: diethyl ether (8:2) at different RF values (**Table 8)**, between them 9 bands were common in all, but varied in intensities (**Figure 7**).   CSNK-12, 17, 25, 26, 27, 28, 29 and 30 shared 1 common bands at Rf 0.25. Similarly, in CSNK-11, 13, 14, 21, 22, 23, 27, 28 and 29, presence of characteristics band at Rf 0.81, which distinguish from the other accessions.
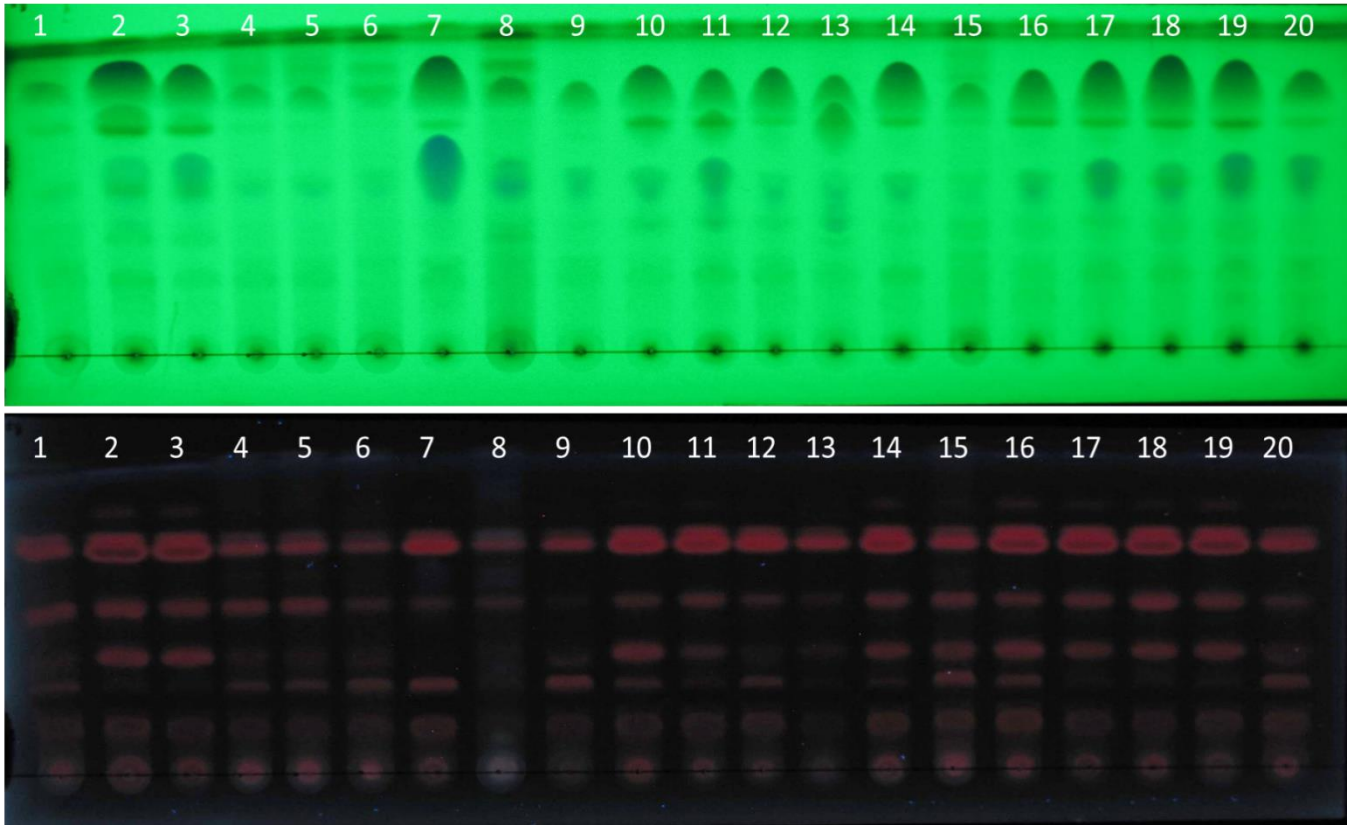
Figure 7: Thin Layer Chromatography of extract of 20 Cannabis accesions. Mobile phase- Hexane diethylether (8:2), visualized at UV-254 nm and UV 366

Table 08: Phyto-Chemical diversity of aerial parts of Cannabis species as determined by Thin Layer Chromatography

Rf	CSNK-11	CSNK-12	CSNK-13	CSNK-14	CSNK-15	CSNK-16	CSNK-17	CSNK-18	CSNK-19	CSNK-20	CSNK-21	CSNK-22	CSNK-23	CSNK-24	CSNK-25	CSNK-26	CSNK-27	CSNK-28	CSNK-29	CSNK-30
The presence/absence matrix of TLC developed in Hexane:diethylether (8:2) and visualized at 254 nm of UV																				
0.98	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.96	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.88	-	-	-	+	+	+	-	+	-	-	-	-	-	-	+	-	-	-	-	-
0.85	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.81	-	+	+	-	-	-	+	-	-	+	+	+	+	+	+	+	+	+	+	+
0.69	-	+	+	-	-	-	+	+	-	-	+	-	-	-	-	-	+	+	+	-
0.65	+	+	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+
0.40	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.25	-	+	-	-	-	-	+	-	-	-	-	-	-	-	+	+	+	+	+	+
0.06	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
The presence/absence matrix of TLC developed in Hexane:diethylether (8:2) and visualized at 366 nm of UV																				
0.81	+	-	+	+	-	-	-	-	-	-	+	+	+	-	-	-	+	+	+	-
0.73	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.67	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.52	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.38	-	-	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+
0.29	+	+	-	-	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+
0.17	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
0.04	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

+ (Present), - (Absent)

6   Discussions

The genus *Cannabis*, belonging to the family Cannabaceae is an important crop mostly grown wild. Hemp is an annual herbaceous, fibre and drug-yielding plant with multipurpose uses. Traditionally *Cannabis* was used for the high quality of fiber from stem; seeds produce nutrient-rich oil and proteins, essential oils and resins from inflorescence and leaves (**Baldini 2018**). The recent attraction of *cannabis* research is mainly on psychoactive compounds Δ9-tetrahydrocannabinol (Δ9-THC) and non-psychoactive compound, cannabidiol (CBD) (**Burggel et al. 2020a**). THC is used for recreational purposes and substantially impacts the human body, the reason behind the ban of hemp in many countries. The government allows the cultivation of hemp-specific cultivar with THC content (based on dry weight) of less than 0.2% in Europe and 0.3% in Asian countries (**Burggel et al. 2020b; Hammami et al. 2021**). *Cannabis* species are reported to identify and taxonomy purely based on morphological characteristics, an attempt was made to identify suitable molecular markers to help in the identification of the crop and to study genetic relationships (**Mukherjee et al. 2008**). The earlier research on hemp was limited due to the legal issues and lack of awareness about the plants (**Bonini et al., 2018**). The recent trends changed the scenario, the cultivation purposes and medicinal uses of Cannabis increasing day by day. Recent studies may be helpful in the increasing economic and pharmacological value of the crop. *Cannabis* is a multipurpose used plant.

The present study has explored diversity analysis using molecular marker aspects of 20 *Cannabis* accessions. Along with phytochemical study for variability analysis of these accessions. To study the molecular variation among the accessions, ISSR and SCoT markers were used. The molecular profile characters indicate that the ISSR primers selected for the study were in determining the diversity among the twenty accessions. Those primers could also give distinct segregation based on the geographical source of the crop **(Rafizadeh et al. 2017)**. The present study showed that selected primers for ISSR profiling could be used to identify the unknown species in *Cannabis*. The ISSR data showed the highest similarity coefficient 0.96 among the 20 accessions. The dendrogram based clustering shows three main cluster in ISSR marker studies. In contrast, the SCoT primer data shows 0.86

similarity coefficient with four distinct cluster of 20 *Cannabis* accessions. Similarly, a comparative study of start codon targeted (SCoT) and inter-simple sequence repeats (ISSR) markers were used for genetic diversity and relationship analysis of nine *Salvia* species (**(Jaiswal et al. 2020**; **Etminan et al 2018)**. The dendrogram generated based on SCoT separated the individuals into sub-clusters in accordance with their species and origin. The results also showed that the SCoT markers could be more consistent and reproducible than ISSR for estimating genetic diversity and relationships among *Salvia* species (**Etminan et al 2018**). SCoT marker was also used to investigate the genetic variability of *Silybum marianum*. SCoT marker produced 255 amplicons and 84.03% polymorphism (**Bhattacharya et al. 2013**). The results show that the SCoT marker system could effectively evaluate milk thistle genotypes genetic diversity (**Rafizadeh et al. 2017**; **Rajesh et al. 2015**). The dendrogram-based cluster analysis of both the primer revealed that the some of the accessions belongs to the same cluster on both the dendrogram-based cluster analysis. CSNK-14, 18 and 25 belong to the same cluster. While CSNK-12, 13 and 15 fits into the same group. On the other hand, CSNK-28, 29 and 30 go to the same group. Moreover, SCoT marker showed high reproducibility with consistent performance for allocation of different accessions of hemp accessions.

The non-targeted phytochemical analysis shows the twenty accessions of *Cannabis* sp. shared a number of common bands. These common bands and differentiating bands could play a more significant role in establishing the pharmacognostic standards for future identification and authentication of genuine plant material (**Khatoon et al. 2005**). In this solvent system with 20 examined, accessions were also clearly differentiated from each other (**De Meijer, E.P., 2014**).

7   Conclusion

The genus *Cannabis* (family Cannabaceae) is an underutilized crop with multipurpose use, mostly grown wild. Lack of awareness of the economic and nutritional utility of the crop and habitat destruction has already resulted in the extinction of several species in India. As the species are reported to identify and taxonomy purely based on morphological characteristics, an attempt was made to identify suitable molecular markers to help in the identification of the crop and to study genetic relationships.

The present aimed to compare the genetic diversity among 20 accessions of *Cannabis* sp. through ISSR and SCoT marker. Along with the phytochemical characterization of the studied accessions. The interpretations of the recent study substantiate, the ISSR marker have a distinct role in the identification of the different accessions of the same species for the evolutionary origin. However, this study involves more specifically with SCoT marker, which implies the 20 accessions more precisely. The same genera in different localities would throw more light on genetic diversity and the adaptive features of the taxa to their microclimatic environments. Similarly, chemical fingerprinting with differentiating bands could play a more significant role in establishing the pharmacognostic standards for future identification and authentication of genuine plant material. Recent studies may be helpful in the increasing economic and pharmacological value of the crop. Hence the conservation of the species in the wild is a must.

8.REFERENCES:

1. Ansari, Shamin Akhtar, et al. "ISSR markers for analysis of molecular diversity and genetic structure of Indian teak (Tectonagrandis Lf) populations." Annals of Forest Research 55.1 (2012): 11-23.

2. Bhattacharyya, P., Kumaria, S., Kumar, S. and Tandon, P., 2013. Start Codon Targeted (SCoT) marker reveals genetic diversity of Dendrobium nobile Lindl., an endangered medicinal orchid species. Gene, 529(1), pp.21-26.

3. Bahmani, Kaivan, et al. "Assessment of genetic diversity in Iranian fennels using ISSR markers." Journal of Agricultural Science 4.9 (2012): 79.

4. Baldini, M., Ferfuia, C., Piani, B., Sepulcri, A., Dorigo, G., Zuliani, F., Danuso, F., Cattivello, C., 2018. The performance and potentiality of monoecious hemp (Cannabis sativa L.) cultivars as a multipurpose crop. Agronomy 8, 162.

5. Baruah, J., et al. "Genetic diversity study amongst Cymbopogon species from NE-India using RAPD and ISSR markers." Industrial Crops and Products 95 (2017): 235-243.

6. Burgel, L., Hartung, J., Pflugfelder, A., Graeff-Hönninger, S., 2020a. Impact of growth stage and biomass fractions on cannabinoid content and yield of different hemp (Cannabis sativa L.) genotypes. Agronomy 10, 372.

7. Burgel, L., Hartung, J., Schibano, D., Graeff-Hönninger, S., 2020b. Impact of Different Phytohormones on Morphology, Yield and Cannabinoid Content of Cannabis sativa L. Plants 9, 725.

8. Galvan, M. Z., et al. "Inter simple sequence repeat (ISSR) markers as a tool for the assessment of both genetic diversity and gene pool origin in common bean (Phaseolus vulgaris L.)." Euphytica 132.3 (2003): 297-301.

9. De Meijer, E.P., 2014. The chemical phenotypes (chemotypes) of Cannabis. Handbook of Cannabis, 89, p.110.

10. Godwin, Ian D., Elisabeth AB Aitken, and Lawrence W. Smith."Application of inter simple sequence repeat (ISSR) markers to plant genetics." Electrophoresis 18.9 (1997): 1524-1528.

11. Iruela, M., et al. "Phylogenetic analysis in the genus Cicer and cultivated chickpea using RAPD and ISSR markers." Theoretical and Applied Genetics 104.4 (2002): 643-651.

12. Khatak, Sunita, MansiGhai, and Swati Dahiya."ISSR marker-based inter and intra-specific diversity analysis in different genotypes of Cannabis sativa."IntConf on Innovative Res in EnggSci and Mang.Vol. 3. 2016.

13. Kojoma, Mareshige, et al. "DNA fingerprinting of Cannabis sativa using inter-simple sequence repeat (ISSR) amplification."PlantaMedica 68.01 (2002): 60-63.

14. Mukherjee, A., Roy, S.C., De Bera, S., Jiang, H.E., Li, X., Li, C.S. and Bera, S., 2008. Results of molecular analysis of an archaeological hemp (Cannabis sativa L.) DNA sample from North West China. Genetic Resources and Crop Evolution, 55(4), pp.481-485.

15. Lata, Hemant, et al. "Assessment of the genetic stability of micropropagated plants of Cannabis sativa by ISSR markers." Plantamedica 76.01 (2010): 97-100.

16. Lata, Hemant, et al. "Molecular analysis of genetic fidelity in Cannabis sativa L. plants grown from synthetic (encapsulated) seeds following in vitro storage." Biotechnology letters 33.12 (2011): 2503-2508.

17. Martins, M., R. Tenreiro, and M. M. Oliveira."Genetic relatedness of Portuguese almond cultivars assessed by RAPD and ISSR markers." Plant cell reports 22.1 (2003): 71-78.

18. Qian, Wei, Song Ge, and D-Y.Hong."Genetic variation within and among populations of a wild rice Oryzagranulata from China detected by RAPD and ISSR markers." Theoretical and Applied Genetics 102.2 (2001): 440-449.

19. Rajesh, M.K., Sabana, A.A., Rachana, K.E., Rahman, S., Jerard, B.A. and Karun, A., 2015. Genetic relationship and diversity among coconut (Cocos nucifera L.) accessions revealed through SCoT analysis. 3 Biotech, 5(6), pp.999-1006.

20. Raman, Vijayasankar, et al. "Morpho-anatomy of marijuana (Cannabis sativa L.)." Cannabis sativa L.-Botany and Biotechnology.Springer, Cham, 2017.123-136.

21. Rafizadeh, A., Koohi-Dehkordi, M. and Sorkheh, K., 2018. Molecular insights of genetic variation in milk thistle (Silybum marianum [L.] Gaertn.) populations collected from southwest Iran. Molecular biology reports, 45(4), pp.601-609.

22. Souframanien, J., and T. Gopalakrishna."A comparative analysis of genetic diversity in blackgram genotypes using RAPD and ISSR markers." Theoretical and Applied Genetics 109.8 (2004): 1687-1693.

23. Strzelczyk, M., Lochynska, M. and Chudy, M., 2021. Systematics and botanical characteristics of industrial hemp Cannabis sativa L. Journal of Natural Fibers, pp.1-23.

24. Srivastava, Ankit, and Vijay Kumar Yadav."Microscopical and chemical study of Cannabis sativa." J Forensic Res 5.1 (2013): 210.

25. Tikunov, Yu M., L. I. Khrustaleva, and G. I. Karlov."Application of ISSR markers in the genus Lycopersicon." Euphytica 131.1 (2003): 71-81.

26. Khatoon, S., Srivastava, M., Rawat, A. and Mehrotra, S., 2005. HPTLC method for chemical standardization of Sida species and estimation of the alkaloid ephedrine. Jpc-Journal of Planar Chromatography-Modern Tlc, 18(105), pp.364-367.