Chapter 3

Materials & Methods

"Everything is theoretically impossible until it is done"
– Robert A Heinlein

3.1. Study area

"Darjeeling", the queen of hills renowned worldwide for "tea" was selected as our study area. Darjeeling is situated in the northern parts of India with latitude 27.03° N, and longitude 88.18° E extending from the tropical Tarai plains (300 feet above sea level) to the Sandakphu-Phalut ridge (12,000 feet). It is the most beautiful hill station situated at the foothills of the Himalayas surrounded by the borders of Sikkim (north,) Bhutan (east), and Nepal (West) with a total area of 3,149 sq. km. Darjeeling is famous for its spectacular tea gardens where tea is not only the major cash crop but also the main source of income for many local people residing there. There are about 80+ tea gardens in the Darjeeling area (Figure3.1). The Dooars(26.82° N - latitude, 89.62° E - longitude) also known as duars meaning door since it opens up 18 gateways between the hills of Bhutan and Indian plains. The eastern dooars (Assam dooars) and the western dooars (Bengal Dooars) consist of a total area of 880 km^2 and are divided by the Sankosh river. Many rivers flow in this region from the hills of Bhutan thus making this region a very fertile land where tea is one of the major economic crops.

3.2. Plant selection and collection

A total of 33 standard tea clones (*Camellia sinensis*) established mainly for Darjeeling and Dooars, were collected from the Darjeeling region. Standard clones were maintained in the Department of Tea Science, University of North Bengal. The core collection

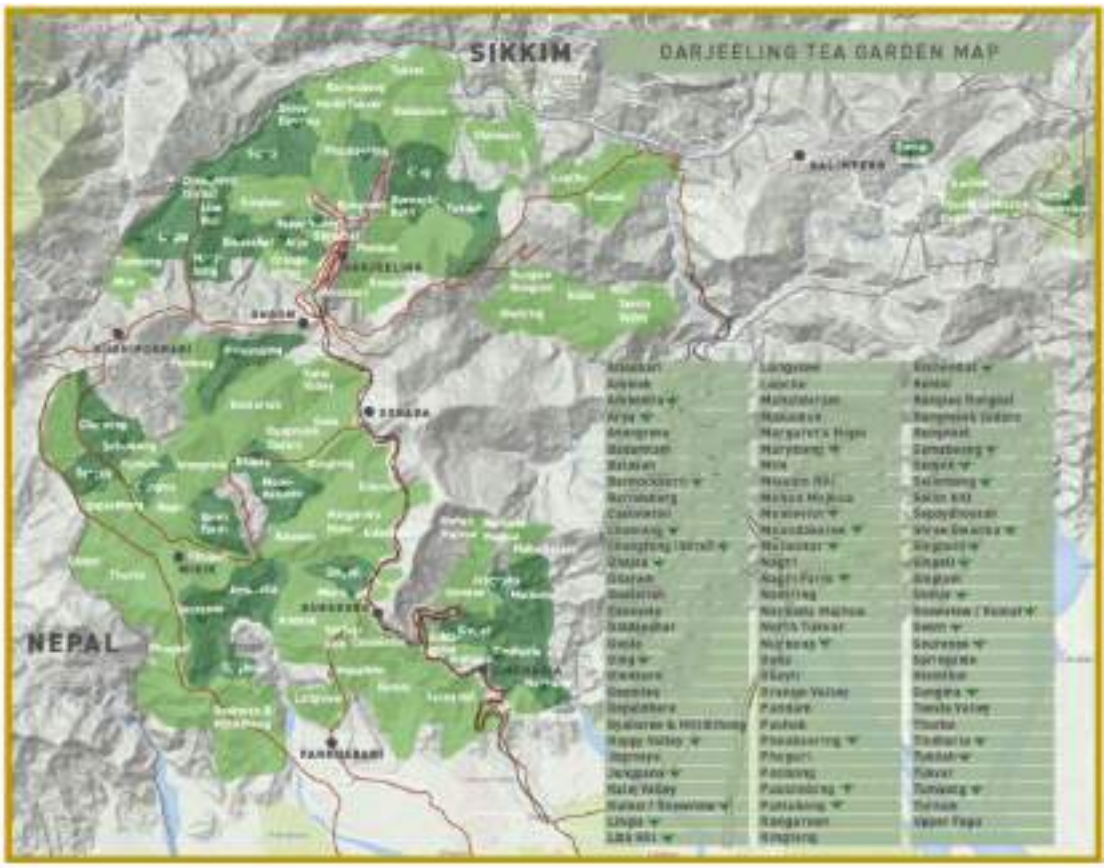


Fig.3.1. Tea gardens of Darjeeling (https://www.teacupsfull.com)

comprised of all the China, Assam, and the Cambod type. The name of the clones along with their specific details is enlisted in Table 3.1. The plant taxonomist of the Department of Botany, University of North Bengal authenticated the voucher specimen and later deposited it in the herbarium of the Botany Department (Accession no- 10339).

3.3. Morphological documentation

Although tea morphology is more or less the same within the same species, there is variation in leaf size, leaf type, plant size, etc. among the China type, Assam type, and the Cambod type. Morphological documentation was done using the data recorded by the Tea Board of India.

3.4. Molecular documentation

3.4.1. Sample selection

All 33 collected clones were employed for molecular studies. Tender leaves of the tea plant (*C. sinensis*), were collected for this study, and only fresh leaves preserved and contained well during the collection were further processed for the downstream analysis.

3.4.2. Chemicals and reagents

The chemicals and reagents used in the present study for varied experiments

Table 3.1. List of tea clones chosen for study

Sl.no	Full name	Abbreviation	Type
1.	Ambari Vallai 2	AV2	China hybrid
2.	Phoobsering 312	P312	China hybrid
3.	Happy valley 39	HV39	China hybrid
4.	Tukdah 253	T-253	Assam-China hybrid
5.	Nandadevi	ND or TS- 378	China hybrid
6.	MB-6	MB-6	China hybrid
7.	Teesta Valley 1	TTV-1	China hybrid
8.	Tukdah 383	T-383	China hybrid
9.	Kopati 1/1	K1/1	Assam Hybrid
10.	B-15/263	B-15/263	Assam hybrid
11.	Balasun 7/1A/76	BS 7/1A/76	China hybrid
12.	Bunnockburn 777	B-777	China hybrid
13.	Sundaram	B/5/63	Assam hybrid (triploid clone)
14.	Tukdah-135	T-135	China hybrid
15.	Tocklai seed-378	TS-378, Nandadevi	China hybrid
16.	Bunnockburn 688	B-688	Assam hybrid
17.	Golconda	UPASI-8 (B/6/36)	Cambod
18.	Rungli Rungiliot 17/144	RR-17/144	China hybrid
19.	Balasun 9/3/76	BS-9/3/76	China hybrid
20.	Chiradew Parbat1	CP-1	Assam hybrid
21.	Phoobsering 1404	P-1404	Assam hybrid
22.	Phoobsering 1258	P-1258	China hybrid
23.	RungliRungiliot 4/5	RR-4/5	China hybrid
24.	Sikkim 1	SKM-1	China Hybrid
25.	Thurbo 3	Thurbo-3	China hybrid
26.	Thurbo 9	Thurbo-9	China - Cambod hybrid
27.	Tukdah 145	T-145	Assam hybrid
28.	Tukdah 246	T-246	Assam-China hybrid
29.	Tocklai variety 19	TV-19	Indochina/ Cambod
30.	Tocklai variety 14	TV-14	Assam hybrid
31.	Tukdah 78	T-78	China hybrid
32.	Bannockburn 157	B-157	China hybrid
33.	B/5/63	Sundaram or B/5/63	Assam hybrid triploid clone

were of analytical grade and purchased either from HiMedia Laboratories Pvt. Ltd. (Mumbai, India), or Merck (Mumbai, India), or Sigma-Aldrich (USA), Promega unless mentioned otherwise. Milli-Q ultrapure water (grade >1.83) was used for all the experiments.

The list of chemicals used for this study is provided in Appendix B.

3.4.3. DNA extraction

The DNA from fresh leaves of all of the tea clones was isolated using the standard protocol given by Doyle and Doyle (1987) with slight modifications.

- Approximately 5 grams of fresh and tender tea leaves were taken in a mortar and pestle and pulverized using liquid nitrogen. It was then quickly transferred into an Oakridge tube containing 15 ml of pre-warmed CTAB extraction buffer (Refer Appendix- B for composition) and mixed properly.
- The tube was then incubated in a water bath (65ºC) for 1h with occasional mixing during intervals.
- An equal volume of chloroform: isoamyl alcohol (24:1) was added subsequently and gently mixed. The mixture was then centrifuged at room temperature (25ºC) for 20 min at 6,500 rpm (5,000xg). The supernatant was carefully transferred to a fresh tube without disturbing the middle layer of cell debris.
- An equal volume of ice-cold isopropanol was added to the supernatant and mixed gently. The mixture was then incubated at -20ºC for almost 2h.
- Following incubation, the mixture was again centrifuged at about 6,500 rpm (5,000xg) for 30 minutes at 4ºC. Afterward, the supernatant was discarded and the pellet was washed gently with chilled 70% ethyl alcohol. It was then and air-dried at least for an hour.
- The pellet was dissolved in 500µl of 1X TE buffer. The pH of the buffer was maintained at pH7.4.
- The buffered DNA solution was now extracted with an equal volume of equilibrated phenol (pH 8.0) and mixed properly. It was then centrifuged at 13,000 rpm (16,000xg) for 20 minutes.
- The upper aqueous layer was then transferred into a fresh tube and extracted with an equal volume of chloroform: isoamyl alcohol (24:1) followed by centrifugation at 10,000 rpm (10,000xg) at room temperature for 15 minutes.
- The upper aqueous layer was again transferred to a fresh tube. To it, $1/10^{th}$volume of 3M sodium acetate (pH 5.2) and double volume of ice-cold absolute ethyl alcohol was added and mixed gently for 10 minutes.
- Afterward, it was then centrifuged at 13,000 rpm (16,000xg) for 30 min at 4ºC.

- Finally, the pellet thus collected was washed gently with 70% ethanol (chilled). It was then air-dried and finally dissolved in 500µl of 1X TE buffer with pH 7.4 (Refer Appendix B for composition).

- Agarose gel analysis was done where the isolated DNA exhibited distinct and clear bands.

- Crude DNA is mixed with many contaminants including RNA, protein, and polysaccharides, etc., which leads to an enzymatic reaction with DNA. The inclusion of the CTAB method in the DNA extraction process helps to eliminate polysaccharides from DNA precipitations to a large extent. Subsequently extraction with phenol: chloroform: isoamyl alcohol indicates the removal of protein impurities from the DNA samples. After the proper removal of impurities, the DNA-CTAB complex provided a network of whitish precipitate of nucleic acids, which was further used for downstream processing. Therefore, DNA purification is a prerequisite step before performing downstream analyses like PCR amplification, DNA restriction, and gene cloning.

3.4.4. Purification

As stated earlier, the main contaminants found in crude DNA are mainly RNA, protein, and polysaccharides, which usually hampers the further downstream process. Therefore, the removal of such impurities is the main step in the purification of crude DNA. The use of CTAB helps in the elimination of polysaccharides from DNA. The subsequent treatment with phenol: chloroform and RNase facilitate the elimination of most of the proteins and RNA respectively from the crude DNA. The purification step of DNA is given as follows:

- 50µg/ml of freshly prepared RNase A was added into the genomic DNA, dissolved in 500µl of 1X TE buffer with pH~7.4 (Refer Appendix B for buffer composition). It was then incubated at 37ºC for 1h in a dry water bath.

- Afterward, an equal volume of chloroform: isoamyl alcohol (24:1) was added and mixed properly. It was then centrifuged at room temperature for 15 minutes at 10,000 rpm (10,000xg).

- The upper aqueous phase was then pipetted into a fresh microcentrifuge tube and to it, $1/10^{th}$ volume of 3M sodium acetate (pH 5.2) was added followed by the addition of a double volume of absolute ethyl alcohol, which helps

in precipitating the DNA.

- It was then centrifuged at 4ºC for 30 min at 13,000 rpm (16,000xg). Finally, the DNA pellet was collected and washed gently with chilled 70% ethyl alcohol. It was air-dried and finally dissolved in 100µl of 1X TE buffer with pH 7.4 (Refer Appendix B for composition).

3.4.5. Quantification

DNA quantification is carried out by spectrophotometric measurements or by agarose gel analysis. Both methods were used in the present study. The crude DNA was quantified using UV Spectrophotometric method using 260 nm and 280nm filters. The results were calculated as the ratio of A_{260}/A_{280} after performing the UV scanning in triplicates and only those samples were considered which showed a ratio of A_{260}/A_{280}equivalent to 1.8. Proper DNA quantification is required in molecular biology for the amplification of target DNA by a polymerase chain reaction and complete digestion of DNA by restriction enzymes.

At first, the UV spectrophotometer (Agilent Technologies Cary 60 UV-Vis) was calibrated at both filters of 260nm and280nm by taking 600µl 1X TE buffer in a cuvette. The purified DNA (6µl diluted in594µl of 1X TE) was taken in a cuvette, mixed properly and the optical density (OD) was recorded at both 260nm and 280nm.

DNA concentration was estimated by using the following formula:

$$\text{DNA} = \frac{\text{OD}_{260} \times 50 \times \text{DF}}{1000}$$

Here DNA=Amount of DNA (ng/µl), DF= Dilution factor. The quality of DNA was considered from the OD values recorded at 260nm and 280nm. The DNA showing an A_{260}/A_{280} ratio of around 1.8 was chosen for further downstream process like RAPD, ISSR, and DNA barcoding techniques.

3.4.6. Electrophoresis separation

For electrophoresis separation, 0.8% of agarose gel (gelling temperature 36°C) was cast in 0.5X TBE (Tris-Borate-EDTA) buffer (Refer Appendix B for composition) containing 0.5µg/ml Ethidium bromide on gel platform (100×70mm). Before loading the DNA into the agarose gel, 5µl of sample DNA was mixed with 3µl of 6X gel loading dye (Refer Appendix B fore composition) and pipetted into the well. A molecular marker(1µl) was run along with the sample to determine the molecular size of the adjacent genomic DNA. Here we used lambda DNA/ EcoRI/HindIII double digest as our molecular marker. The same buffer i.e., the 0.5X TBE was used as a running

buffer. The gel was run at 50 Volt (V) and 100 mA for 1hr in a Mini Submarine Gel Electrophoresis using Electrophoresis Power Supply Unit. At the end of the run time, the gel was visualized under UV light on a UV Transilluminator. The good quality of DNA was considered by the presence of a single compact band at the corresponding position to λ DNA/ EcoRI/HindIII double digest indicating the high molecular weight of the DNA. The quantity of the DNA was estimated by comparing the sample DNA with the control by visualizing the intensity of the band under UV light. The pure and good quality of DNA thus obtained was used for various fingerprinting studies. Gel photographs were captured using a gel documentation system fitted with a Cannon SLR camera (EOS350D) and Marumi orange filter (58mm YA2, Marumi, Japan). Besides, the EOS utility software was used for this purpose.

3.4.7. RAPD analysis

A total of 45 random primers were tested for 33 different tea clones (Table 3.2).

A total of 25µl of PCR reaction mixture was taken in an autoclaved and sterile PCR tube. The reaction mixture (25µl) for PCR amplification consisted of the following solutions:

- PCR master Mix - 12.5µl
- Primer - 1.25µl (0.25µM)
- Template DNA - 2µl (25ng)
- Pyrogen free water- 9.25µl

One negative control PCR mix tube was prepared by omitting DNA from the mixture. The PCR amplification was performed on Applied BiosystemsVeritti 96 well Thermal Cycler and the cycle (Figure 3.2) consisted of the following specifications:

- Cycle 1 - Denaturation at 94ºC for 4min, Primer annealing at 37ºC for 1min, Primer extension at 72ºC for 2min.
- Cycle 2-44 - Denaturation at 94ºC for1 min, Primer annealing at 37ºC for 1min, Primer extension at 72ºC for 2min.
- Cycle 45 - Denaturation at 94ºC for 1min, Primer annealing at 37ºC for 1min., Primer extension at 72ºC for 10 min.

After the completion of the total PCR cycle, the PCR products were separated on 1.8% (w/v) agarose gel containing Ethidium bromide (EtBr) solution (0.5µg/ml) run in 0.5X TBE buffer. Next, the PCR product (12µl) was mixed with 4µl of 6X Gel loading dye and the samples were loaded and

Table 3.2. List of RAPD primers and its sequences

SL no.	PRIMER	SEQUENCE	SL no.	PRIMER	SEQUENCE
1.	OPA01	CAGGCCCTTC	24.	OPB04	GGACTGGAGT
2.	OPA02	TGCCGAGCTG	25.	OPB05	TGCGCCCTTC
3.	OPA03	AGTCAGCCAC	26.	OPB06	TGCTCTGCCC
4.	OPA04	AATCGGGCTG	27.	OPB07	GGTGACGCAG
5.	OPA05	AGGGGTCTTG	28.	OPB08	GTCCACACGG
6.	OPA06	GGTCCCTGAC	29.	OPB09	TGGGGGACTC
7.	OPA07	GAAACGGGTG	30.	OPB10	CTGCTGGGAC
8.	OPA08	GTGACGTAGG	31.	OPB11	GTAGACCCGT
9.	OPA09	GGGTAACGCC	32.	OPB12	CCTTGACGCA
10.	OPA10	GTGATCGCAG	33.	OPB13	TTCCCCCGCT
11.	OPA11	CAATCGCCGT	34.	OPB14	TCCGCTCTGG
12.	OPA12	TCGGCGATAG	35.	OPB15	GGAGGGTGTT
13.	OPA13	CAGCACCCAC	36.	OPB17	AGGGAACGAG
14.	OPA14	TCTGTGCTGG	37.	OPB18	CCACAGCAGT
15.	OPA15	TTCCGAACCC	38.	OPB19	ACCCCCGAAG
16.	OPA16	AGCCAGCGAA	39.	OPB20	GGACCCTTAC
17.	OPA17	GACCGCTTGT	40.	OPF09	CCAAGCTTCC
18.	OPA18	AGGTGACCGT	41.	OPG19	GTCAGGGCAA
19.	OPA19	CAAACGTCGG	42.	OPN13	AGCGTCACTC
20.	OPA20	GTTGCGATCC	43.	OPH04	GGAAGTCGCC
21.	OPB01	GTTTCGCTCC	44.	OPN04	GACCGACCCA
22.	OPB02	TGATCCCTGG	45.	OPN19	GTCCGTACTG
23.	OPB03	CATCCCCCTG			

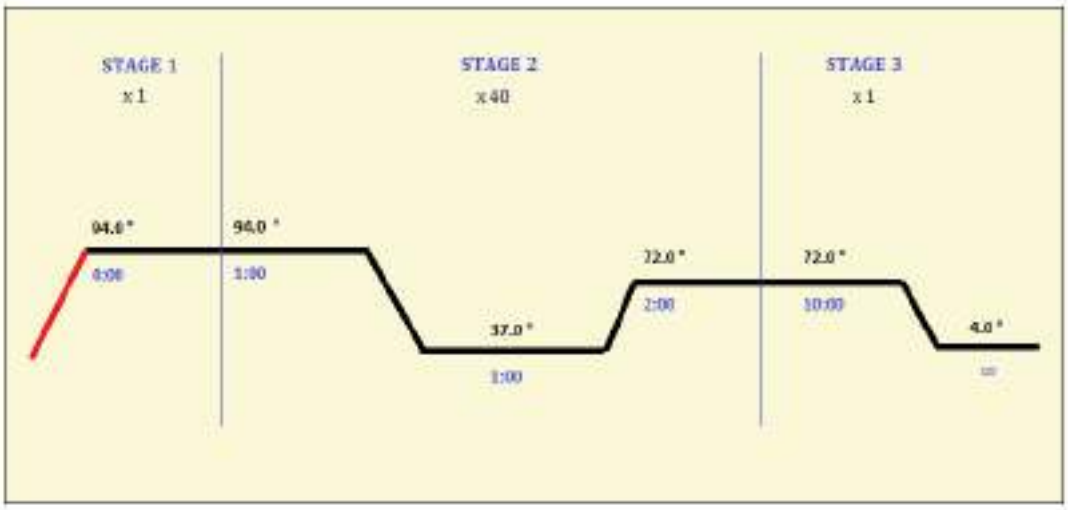


Fig.3.2. PCR cycle of RAPD

electrophoresis was carried out at 50V for 1 hour 40 minutes.

Finally, the gels were visualized with a UV transilluminator and photographed with Gel Documentation System as mentioned earlier. One DNA ladder (λDNA/EcoRI/HindIII double digest) and a 100bp DNA ladder were used as molecular size markers. All PCR reactions were run at least thrice and only reproducible and clear bands were scored using PyElph software (version 1.4). It was then aligned by diversity database software NTSYSpc (version 2.0).

3.4.8. ISSR analysis

A total of 15 ISSR primers were screened for 33 tea clones (Table 3.3).

An autoclaved sterile PCR tube was taken containing 25 µl of PCR reaction mixture for PCR amplification. The mixture consisted of the following solutions;

- PCR master Mix - 12.5µl
- Primer - 1.25µl (0.25µM)
- Template DNA - 2µl (25ng)
- Pyrogen free water- 9.25µl

The PCR reaction was performed on Applied BiosystemsVeritti 96 well Thermal Cycler and the cycle (Figure3.3) consisted of the following specifications:

- Cycle 1 - Denaturation at 94 ºC for 5 mins, primer annealing at 52ºC for 1 min, primer extension at 72ºC for 1 min.
- Cycle 2-34 - Denaturation at 94ºC for 45sec, Primer annealing at 52ºC for 1 minute, Primer extension at 72ºC for 1 min.
- Cycle 35 - Denaturation at 94ºC for 45 sec. Primer annealing for 1 minute. Primer extension at 72ºC for 7 minutes.

After the end of the cycle or run time the PCR products were separated on agarose (1.8% w/v) gel. The gel contained Ethidium bromide solution (0.5µg/ml). TBE buffer (0.5X) was used as a running buffer. While loading the sample, 12µl of PCR product was mixed with 4µl of 6X Gel loading dye (Refer Appendix - B for composition) and the samples were loaded into the

Table 3.3. List of ISSR primers and sequences

PRIMER	SEQUENCE
UBC807	(AG)8T
UBC808	(AG)8C
UBC810	(GA)8T
UBC811	(GA)8C
UBC818	(CA)8G
UBC822	(TC)8A
UBC824	(TC)8G
UBC825	(AC)8T
UBC834	(AG)8YT
UBC836	(AG)8YA
UBC841	(GA)8YC
UBC856	(AC)8YA
UBC873	(GACA)4
UBC813	(CT)8T
UBC815	(CT)8G

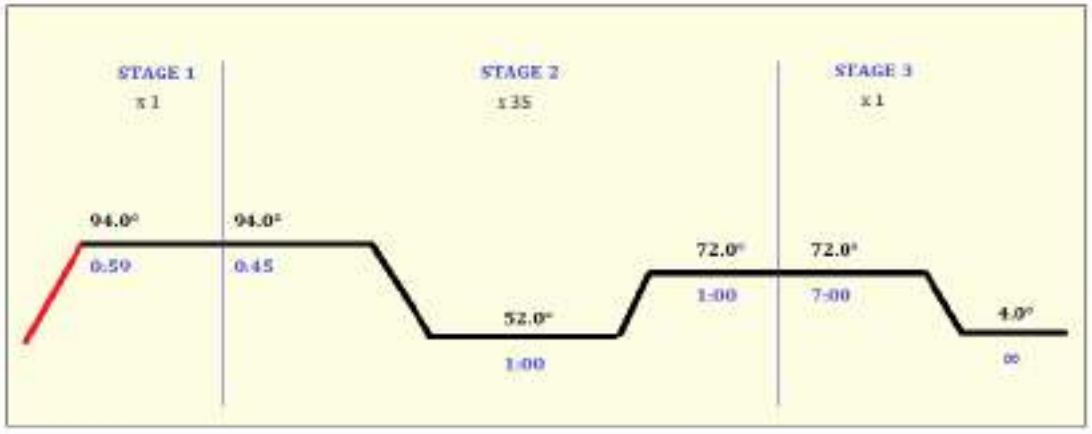


Fig.3.3. PCR cycle of ISSR

well of agarose gel. Electrophoresis was carried out at 50V and 100mA for 1 hour and 40 minutes.

Finally, the bands were visualized under a UV transilluminator and photographed with Gel Documentation System as mentioned earlier (section 3.2.6.). λDNA/EcoRI/HindIII double digest and 100bp DNA ladder were used as a molecular size marker. All PCR reactions were run at least thrice and only reproducible and clear bands were scored using PyElph software (version 1.4) and aligned by diversity database software NTSYSpc (version 2.0).

3.4.9. Data analysis

In the present study, each polymorphic band was considered a binary character and was scored either as 1 or 0 for presence or absence of bands and was later accumulated in a data matrix. Further, a similarity matrix was calculated based on band sharing from the binary data using the Dice coefficient (Nei and Li, 1979) while a dendrogram of similarities was generated using the group average agglomerative clustering tool. All the analysis was done employing the software package NTSYSpc (version 2.0) givewn by Rohlf (1998). In addition, correspondence analysis (2D and 3D plot) of right vectors from the binary data was performed using NTSYSpc (version 2.0), to exhibit a graphical representation among the clones.

3.5. DNA Barcode

3.5.1. Primers used

The matk primer was chosen for the present study. The sequences of the primer are given in Table 3.4.

3.5.2. PCR amplification of the matK region

The study was undertaken in a 25µl volume containing the following

components:

- PCR master Mix– 12.5µl
- Primer – 1.25µl each (forward and reverse)
- Template DNA - 2µl (25ng)
- Pyrogen free water- 8µl

The PCR reactions were performed on Applied BiosystemsVeritti 96 well Thermal Cycler. The amplification cycle (Figure 3.4) consists of the following specifications:

- Cycle 1 - Denaturation at 94ºC for 4 min, Primer annealing at 48ºC for 30 sec, Primer extension at 72ºC for 1 min.
- Cycle 2-34 - Denaturation at 94ºC for1 min, Primer annealing at 48ºC for 30 sec, Primer extension at 72ºC for 1 min.
- Cycle 35 - Denaturation at 94ºC for 1 min, Primer annealing at 48ºC for 30 sec, Primer extension at 72ºC for 7 min.

3.5.3. Agarose gel electrophoresis

After the compliance of total PCR cycle, the PCR products were separated on 2% (w/v) agarose gel containing Ethidium bromide solution (0.5µg/ml) run in 0.5X TBE buffer (pH- 8.0).

Table3.4. Details of matK primer

SL.no	Primer Used	Tm of primer	Annealing temperature	Conc. Of Primer (pm/µl)	Length of Primer with sequence
1.	matK forward (F)	46	48°C	161.83	22 (CGATCTATTCATTCAATATTTC)
2.	matK reverse (R)	53	48°C	208.38	22(TCTAGCACACGAAAGTCGAAGT)

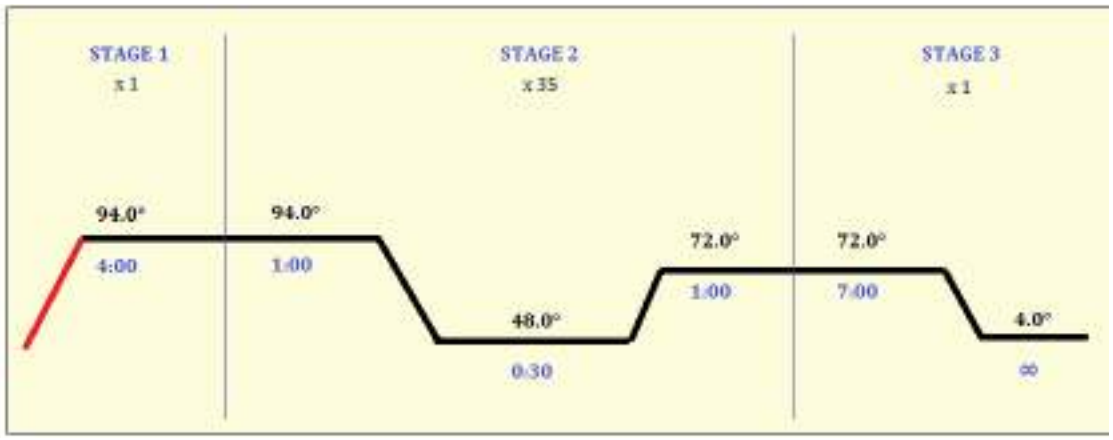


Fig.3.4. PCR cycle of matK

Next, the PCR product (12µl) was mixed with 4µl of 6X Gel loading dye and the samples were loaded and electrophoresis was carried out at 50V for 1h 40 min. Finally, the gels were visualized with a UV transilluminator and photographed with Gel Documentation System as mentioned earlier. One DNA ladder (100 bp) DNA ladder was used to determine the band size.

3.5.4. Sequencing of amplified products

A total of 33 tea clones were subjected for DNA sequencing purposes using matK region of the chloroplast genome. The DNA sequencing was done from Bioserve Biotechnologies Pvt. Ltd.

3.5.5. Sequence analysis and construction of a phylogenetic tree

The obtained sequences were edited or corrected if required (correction of chimeric sequence) as per the guidelines given by NCBI (http:// www.ncbi.nlm.nih.gov) and it was further compared by querying against existing global sequences submitted in the GenBank database (https:// blast.ncbi.nlm.nih.gov/Blast.cgi) using Nucleotide BLAST algorithm. Both Blastn and Blastx were performed simultaneously.

The final and corrected sequences were then further submitted in NCBI using the BankIt submission portal (https:// www.ncbi.nlm.nih.gov/WebSub/). The accessions number provided was later retrieved from the database.

After the successful submission, a total of 29 experimental sequences of the *Camellia sinensis* matK region were taken for final sequence analysis using MEGA 4.0 (Molecular Evolutionary Genetics Analysis) software. Pairwise alignment and multiple sequence alignment using ClustalW were first performed during the analysis. Subsequently, evolutionary history was deduced using various clustering methods of MEGA 4.0 (Tamura *et al.,* 2007). Two methods were chosen here namely the Neighbor Joining method given by Saitou and Nei (1987) and the Unweighted Pair Group Mean Average (UPGMA) method with arithmetic mean given by Sneath and Sokal (1973). The bootstrap value was preferred for 1000 replicates. The phylogenetic tree was deduced with the branch lengths in the same units as evolutionary distances.

The evolutionary distances were calculated using the Maximum Composite Likelihood method of MEGA 4.0 (Tamura *et al.,* 2007). Here the units are given several base substitutions per site. All the three codon positions i.e., the first, second

and third were included. Sequences with small size were not taken for further analysis and all the positions which contained gaps as well as missing data were removed from the dataset. A total of 563 positions were taken in the final dataset.

Further genetic pair-wise distance for matK was calculated using the Maximum Composite Likelihood model (Tamura *et al.,* 2007) and Kimura 2 parameter model (Kimura, 1980) given in MEGA 4.0. Sequences were further aligned using web based Multalin software V.5.4.1 (http:// multalin.toulouse.inra.fr/ multalin/) and both ends were trimmed to avoid any error and the total base pair characters were kept as 563bp. Different number of nucleotide frequencies and position of nucleotide along the consensus region was also studied using the same Multalin software V.5.4.1 (http:// multalin.toulouse.inra.fr/multalin/) (Corpet, 1988).

3.5.6. Illustrative representation of DNA barcode

Illustrative representation of the DNA barcode was done representing the sequences either in the form of QR code or barcodes. Specific QR code was generated using matK QR code generator (More *et al.,*2016), encoding the sequence for each matK sequence of respective tea clones. Any QR code scanner would decode the encoded sequence. On the other hand, barcode with color codes for specific bases or nucleotides was generated using the Biorad barcode generator (http://biorad -ads.com/DNABarcodeWeb/).

3.6. Phytochemical screening

3.6.1. Sample collection

Five tea clones were selected as experimental clones, the details of which are provided in Table 3.5 as given below. The fresh mature tea leaves were collected from the Mirik hills of Darjeeling district.

3.6.2. Extract Preparation

Fresh leaves of the samples were washed thoroughly under running tap water followed by rinsing it twice with distilled water. It was then air-dried under shade (Figure 3.5) at room temperature. The dried leaves were finely powdered using a grinder. A total of 3g of each sample was taken and dissolved in 30 ml of solvent and incubated for 48 hours under room temperature. Nine different solvents were chosen for this study in the order of varying polarity i.e., varying from polar to non-polar. The extracting solvents chosen in increasing order of polarity were hexane; benzene, chloroform, diethyl ether, ethyl acetate, acetone, ethanol, methanol, and water. After 48 hours of incubation, the cold

extracts were centrifuged and the supernatant was collected and preserved at room temperature for future use.

3.6.3. Qualitative screening of phytochemicals

The qualitative test included a qualitative screening of various phytochemicals like flavonoid, tannin, cardiac glycoside, steroid, terpenoid, diterpenes, saponin, coumarin, reducing sugar, and protein. The test was done using the methods given previously (Brain and Turner, 1975; Kumar *et al.,* 2009; Ngbede *et al.,* 2008).

3.6.3.1. Flavonoid

To 250µl of sample few drops of 10%, $FeCl_3$ were added to which a blue or green coloration confirmed the presence of flavonoids.

3.6.3.2. Tanin

The appearance of blue or green color formation on the addition of few drops of 1% $FeCl_3$ to 250 µl of extract confirmed the presence of tannins.

3.6.3.3. Steroid

About 250 µl of the extract was evaporated and dissolved in 2ml of chloroform and about 2ml of concentrated H_2SO_4 was added from the sidewall of the test tube. The appearance of a reddish-brown color ring confirmed the presence of steroids.

3.6.3.4. Terpenoid

The evaporated extract (250 µl) was dissolved in chloroform and concentrated H_2 SO_4 was added from the sidewall of test tubes and then shaken. The formation of red to reddish -brown coloration at the base confirmed the presence of terpenoids.

3.6.3.5. Cardiac glycoside

250 µl of the extract was evaporated and to it, 1ml of glacial acetic acid, one drop of 10% $FeCl_3$, and 1ml of concentrated H_2SO_4 were added. A brown ring at the interface indicated the presence of cardiac glycosides.

3.6.3.6. Diterpenes

Copper acetate was used to confirm the presence of diterpenes wherein the

Table 3.5 Tea clones chosen for phytochemical screening based on varying polarities of extracting solvents

Accession	Clone	Type
S1	TS569	Seed
S2	China	Seed
S3	AV2	Clonal
S4	P312	Clonal
S5	Assam	Seed

Fig.3.5. Mature and fresh leaves of *C. sinensis* washed and air dried under shade

addition of few drops of 5 % copper acetate to 250 µl of extract dissolved in an equal volume of distilled water formed an emerald green color.

3.6.3.7. Coumarin

The yellow coloration on the addition of about 500 µl of 10% NaOH to 250 µl of the sample confirmed the presence of Coumarins.

3.6.3.8. Reducing sugar

Benedict test was performed to estimate the presence of reducing sugar in the extracts wherein 1ml of Benedict reagent added to 250 µl of the sample gave green coloration (color varies from green to red depending upon the percentage of reducing sugar present).

3.6.3.9. Saponin

Froth test was conducted to confirm the presence of saponin with appearance as well as the persistence of froth while shaking the sample diluted with distilled water.

3.6.4. Quantitative screening of Phytochemicals

3.6.4.1. DPPH free radical scavenging assay

As mentioned by (Blois, 1958), DPPH was used to determine the antioxidant activity of the mixture of compounds extracted employing different solvents. The decrease in absorbance is marked by the free radical scavenging property of the compound, which donates hydrogen atoms and scavenges the unpaired electron of the stable free radical of DPPH. To 100 µl of plant extracts (5mg/ml) prepared from different solvents, 1900 µl of methanol was added and shaken. The mixture was incubated at room temperature for 30 minutes in dark. The absorbance was then recorded at 520 nm using a spectrophotometer. Ascorbic acid was taken as a standard.

The total scavenging activity was calculated using the following equation:

$$\text{DPPH scavenging (\%)} = \{Acont - Asamp/AcontX100\}$$

Where, A_{cont} denotes absorbance of only methanol and DPPH and A_{samp} denotes absorbance of sample dissolved in methanol (Plant extract/ standard) along with DPPH.

The IC_{50} value was calculated using the software KyPlot (version 5.0), which gives an idea about the amount of extract needed for 50% inhibition. The IC_{50} value (µg/ml) is the concentration required to inhibit 50% of the initial DPPH free radical, and it was calculated from the graph of the inhibition curve. All the reactions were monitored in triplicate and the value was expressed as the mean ± standard deviation (S.D.).

3.6.4.2. Ferric reducing power assay

Ferric reducing power assay was done as per the protocol given (Aiyegoro and Okoh, 2009) with slight modification. The antioxidant compounds act as reducers causing the reduction of $Fe^{3+/}$ferricyanide complex to the ferrous form, which can be monitored by determining the formation of Perl's Prussian blue at 700nm. In a test tube 250 µl of leaf extract was taken with the addition of 625 µl of sodium phosphate buffer (0.2M, pH 6.6), and a total of 625 µl of K_2 [Fe $(CN)_6$] 1(% w/v) and incubated for 20 minutes at 50 °C. The tubes were then cooled and centrifuged at 3000 rpm after the addition of 625 µl of TCA (10%). The upper layer of the solution or supernatant (625 µl) was mixed with an equal volume of distilled water and 125 µl of $FeCl_3$ (0.1% w/v). The absorbance was finally recorded at 700nm. A higher absorbance value indicated higher reducing power.

3.6.4.3. Quantitative estimation of total flavonoids

The total flavonoids were estimated quantitatively using the $AlCl_3$ method with some modifications (Zou *et al.,* 2004). A total of 250µl of the sample was taken in a test tube to which 750µl of deionized water and 75µl of 5% $NaNO_2$ were added. Following incubation for five minutes at room temperature, 150µl of 10% $AlCl_3$ was added. It was then incubated for six minutes at room temperature followed by the addition of 500µl of 1mM NaOH and 275µl of deionized water. It was then incubated for 30 minutes at room temperature. The absorbance of the yellowish-orange color produced by the interaction of flavonoid with $AlCl_3$ was recorded at 510 nm using a spectrophotometer. Quercetin was taken as a standard and the total flavonoid content was calculated by taking reference from a calibration curve (y = 0.2071x - 0.2048) of quercetin (Figure 3.6.a.) taken at different concentrations (1-5 mg/ml).

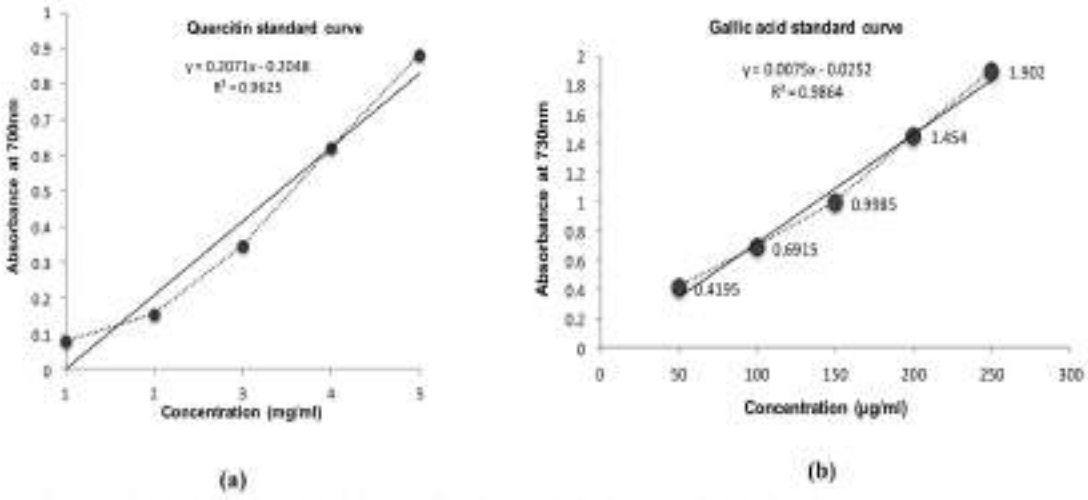
The total flavonoids were thus expressed as mg of quercetin equivalent per g of extract i.e., mg QE/ g. Data were expressed as a mean of triplicates ± standard deviation (Blois, 1958).

3.6.4.4. Determination of phenol

The total phenolic content of the sample was determined using the Folin – Ciocalteu method (Folin and Ciocalteu, 1927) with slight modification. 100 µl samples were taken in a test tube and to it, 400 µl of 10 % Folin reagent was added (1ml Folin + 9 ml distilled water). The mixture was incubated in dark for 5 minutes at room temperature followed by the addition of 1 ml of 5% Na_2CO_3. After incubating it for 2 hrs in dark at room temperature, absorbance was recorded at 730 nm using a spectrophotometer. Gallic acid was taken as a standard and the total phenol content was calculated by taking reference from a calibration curve (y = 0.0075x - 0.0252) of gallic acid (Figure 3.6.b.) taken at different concentrations (50-300µg/ml). The total phenols in the extract were thus expressed as mg of gallic acid per gram of extract (mg GAE/g). Data were expressed as a mean of triplicates ± standard deviation (Blois, 1958).

3.6.4.5. Statistical tests

Statistical tests were performed using the KyPlot software (version 5.0). The readings were taken in triplicates and data was represented as mean ±SD. Statistical analysis relating to t-test for paired comparison of mean was performed employing KyPlot software (version 5.0). Results were considered significant at a level of significance $p \leq 0.001$.



(a) (b)

Fig.3.6. Standard curve (a) Quercetin standard curve for total flavonoid estimation; (b) Gallic acid standard curve for total phenol estimation

3.7. Chemical Characterizations of selected plant extracts

Characterizations of plant extracts provide major opportunities for new drug discovery as well as for the pharmaceutical industry. In essence, natural products or phytochemicals are constantly being screened due to their safe and fewer adverse effects. Several modern techniques of isolation are being included in the characterization of bioactive molecules from plant extracts.

3.7.1. GC-MS analysis

3.7.1.1. Sample preparation

The leaf sample was collected and washed thoroughly under running tap water followed by distilled water. It was then left in the room for shade dry. The sample was now pulverized using a grinder. The plant extract was prepared by mixing the powder (50mg) with 50 ml of acetone with occasional shaking in a magnetic stirrer placed at room temperature (20-25°C). After 48hrs of incubation, the mixture was centrifuged at 15,000 rpm for 20 min at 25°C. The supernatants were collected and passed through the anhydrous Na_2SO_4 and activated charcoal (2:1; w/ w) in a 10 cm mini-column $1/10^{th}$ packed with non-absorbent cotton to remove moisture and color (Dey and Chaudhuri, 2015). The eluted liquid was re -centrifuged at 15,000 rpm for 20 min and passed through Whatman filter paper no. 1(11µm). The resultant clear solutions were used for further analysis.

3.7.1.2. Sample analysis

GC-MS analysis was performed in JNU. The acetone extracts of the sample S1 (TS569) and S3 (AV2) were further employed for GC-MS analysis. The extract was directly used for the analysis. GC-MS was done using a GCMS-QP2010 Plus (Shimadzu, Kyoto, Japan). Das *et al.* (2014) has discussed elaborately the instrument and the process involved in GC-MS analysis which is stated below briefly. The system consists of a headspace sampler (AOC-20s) and auto-injector (AOC-20i), equipped with a mass selective detector with ion source temperature and interface temperature being 230°C and 260°C respectively. The solvent cut time is 2.50 min threshold of 1,000 eV and mass range of 40 - 650 m/z. The compounds present in the solvent extracts separated employing an Rtx 5 MS capillary column (Restek Company, Bellefonte, USA). The split mode is used at a ratio of 10:1. The injector with a split injection mode temperature is initialized to 250°C. The temperature is first programmed from 100°C for 3 minutes which then further increases to

280°C at a ramp rate of 10°C/min. (19 min hold). Helium is used as a carrier gas with a linear flow velocity of 40.9 cm/s. A total of 1.0 µL of the solvent extract is injected and the components present in the extracts were identified by comparing their retention indices (RI) alongside homologous alkane series and by comparing their mass spectral fragmentation patterns against the data provided in different libraries like NIST.LIB, NIST08.LIB, NIST08s.LIB, and WILEY8.LIB. A good match of the mass spectrum with the RI leads to the assumption or the identification of the compounds.

3.8. Bioactivity of the extracts

3.8.1. Antimicrobial activity

The acetone and methanol tea extracts was first evaporated and then dissolved in DMSO to carry out the antimicrobial activity owing to the fact that polar solvents like acetone and methanol imparts its own antimicrobial property. The antimicrobial activity was carried out using agar well diffusion method. Four different bacterial strains were employed (Table 3.6) out of which two were gram-positive bacteria (*Staphylococcus aureus* and *Bacillus subtilis*) and the other two were gram-negative bacteria (*Escherichia coli* and *Klebsiella pneumonia*).

3.8.2. Agar well diffusion method

Sterile Petri dishes of 90 mm were taken to which 25 ml of MHA (Muller Hinton agar) media was poured under aseptic conditions. After the solidification of the media, 45 µl of bacterial strains was pipetted into each plate and swabbed or spread uniformly using a sterile cotton swab. A well was punctured into the solidified agar using a sterile micro tip. To each agar wells, 100 µl of plant extracts (10 mg/ml) was pipetted along with a sterility control (DMSO) and positive control (Streptomycin Sulphate: a broad-spectrum antibiotic used as a standard). The concentration of the standard taken was 10 folds lower than the acetone and methanol extracts. The experimental setup was the same for all the samples, which were performed in duplicates. The plates were incubated at 37°C for 24 hours or 48 hours if needed. The antimicrobial activity was recorded regarding the appearance of a clear halo of inhibition zone on the plates.

3.8.3. MIC (Minimal Inhibitory Concentration) determination

The MIC value was further determined only for *S. aureus* using the agar well diffusion technique as mentioned above elaborately. Different concentrations were taken i.e., 0.25 mg/ml, 0.5 mg/ml,1 mg/ml, 4 mg/ml, 8

Table 3.6 Bacterial strains chosen for study and its history of pathogenecity

BACTERIAL STRAIN	PATHOGENECITY	REFERENCE	TYPE
Staphylococcus aureus	* Sepsis * Mastitis * Septicaemia * urinary tract infection * osteomyelitis * endocarditis * food poisoning * meningitis * biofilm associated infections * skin infections	(Scudiero *et al.,* 2020)	Gram-positive
Bacillus subtilis	Pulmonary disease	(Flindt *et al.,* 1969)	Gram- positive
Escherechia coli	* Inflammatory bowel disease * Extraintestinal infections (urinary tract infection, sepsis, intra-amniotic and puerperal infection, neonatal infections, neonatal sepsis)	(Mirsepasi *et al.,* 2019)	Gram-negative
Klebsiella pneumonia	* Intra- abdominal infections * Urinary tract infection * Lower respiratory tract infection * Baacteremia * Meningitis * bacterial pneumonia	(Kang *et al.,* 2020) (Vaghasiya and Chanda, 2009) (Ashurst and Dawson, 2018)	Gram-negative

mg/ml and 10 mg/ml respectively. 100 µl of each dilution rendering different concentrations was introduced into the wells of MHA (Muller Hinton Agar) plates in duplicates which were pre-inoculated with *S.aureus* and DMSO was used as a control. The least concentration of the acetone or methanol extract of the sample, showing a clear zone of inhibition was considered as its MIC.

3.8.4. Docking

DNA gyrase, one of the Type II topoisomerases relieving strain helps in the progression of DNA replication by either introducing negative supercoiling or relaxing the positive supercoiling (Sutormin *et al.,* 2018). Hence gyrase may be regarded as one of the major housekeeping proteins among organisms. The fact of gyrase being one of the major housekeeping

proteins among the other organisms led us to further progress in the docking study.

The crystal structure of DNA gyrase (PDB id: 5CDN) of *S. aureus* was downloaded from the PDB database (http://www.rcsb.org/structure/5CDN). The resolution of the structure was 2.79Å as obtained through the X-ray diffraction method. The crude structure of DNA gyrase was polished by removing water and adding polar hydrogen bonds to it. The grid box dimensions were considered in such a way as to accommodate the whole protein so that the whole protein could be taken to find the best binding site. After the final protein structure was ready for docking, we downloaded the structure of Phenol from the NCBI Pubchem database. The physicochemical properties of phenol were taken into consideration. Molecular docking was performed in Auto dock Vina software considering the physicochemical properties of phenol. The final results were then visualized via Pymol software.

3.9. Purple Tea and synthesis of nanoparticles

3.9.1. Silver nanoparticles (AgNPs)

3.9.1.1. Extract preparation

5g of plant sample were pulverized and the extract was prepared by suspending the powdered plant sample into 100 ml of distilled water. The aqueous extract was placed in a hot magnetic stirrer (50°C) for two hours. The sample was then filtered using a Whatman filter paper. The filtrate was stored at 4°C for future use (used until 1 week).

3.9.1.2.Optimization

3.9.1.2.1. Varying concentration of $AgNO_3$

In this reaction, the concentration of silver nitrate ($AgNO_3$) was optimized using different concentrations of silver nitrate solution, where the reaction was maintained at a concentration of 1 mM, 2 mM, 3 mM, 4 mM, 5mM, 6 mM, 7 mM, 8 mM, 9 mM, and 10mM. The absorbance of the resulting solution was measured spectrophotometrically.

3.9.1.2.2. Temperature and Light

Similarly, keeping the reaction mixture in different conditions of temperature and light optimized the right temperature or favorable condition.

3.9.1.3. Synthesis

3mM of silver nitrate was dissolved in 200 ml of distilled water and kept at room temperature. 5 ml of plant extract was added to 50 ml of prepared silver nitrate solution and kept under room temperature and natural sunlight. A change in the color of the solution was noted.

3.9.1.4. Characterization

3.9.1.4.1. UV–Vis's spectroscopy

The characterization of synthesized nanoparticles was done by measuring the optical density of reduced zinc ions using UV–Vis's spectroscopy (Agilent Technologies Cary 60 UV-Vis). Small aliquots of the reaction mixture were diluted ten times with double distilled water and 1 ml of the diluted solution was transferred to the cuvette following the analysis using UV–Vis spectrophotometer (Ocean Optics, USA).

3.9.1.4.2. X-ray diffraction (XRD)

X-ray diffraction technique was used to check the formation, quality, and purity of the compounds. The synthesized AgNPs were first centrifuged at 10000 rpm for 15 minutes and then the supernatant was discarded. The pellet was washed using ethanol repeatedly three times and finally, it was rinsed with sterile water.

3.9.1.4.3.Scanning electron microscopy (SEM)

SEM or scanning electron microscopy was done to determine the surface morphology of synthesized AgNPs. A small amount of the synthesized zinc oxide nanoparticle was adhered to the grid of carbon-coated tape and dried under a mercury lamp for 5 minutes. Gold coating (3nm) was done using a gold sputtering unit. The prepared sample was visualized using Scanning electron microscopy (JEOLJSM-IT100InTouchScope™ Scanning Electron Microscope, JEOL Solutions for Innovation, Tokyo, Japan).

3.9.1.4.4. Antimicrobial activity

In this study, bacteria like gram-positive bacteria (*Staphylococcus aureus* and *Bacillus subtilis*) and gram -negative bacteria (*Escherichia coli* and *Klebsiella pneumonia*.) were used as test organisms and were cultured in Muller Hinton Agar medium (MHA). Well, a diffusion assay was performed to determine the antibacterial activity of biosynthesized silver nanoparticles. Different concentration of silver nanoparticles was loaded into each well of the MHA plates. The well without any treatment (only distilled water) was served as a control throughout the experiment. After inoculation, the plates were incubated at 37°C. The zone of inhibition was measured in terms of centimeters. The elaborate protocol is mentioned in the previous section (3.8.1.).

3.9.2. Zinc oxide nanoparticles (ZnO NPs)

3.9.2.1. Preparation of Plant extract

5g of plant sample were pulverized and the extract was prepared by suspending the powdered plant sample into 100 ml

of distilled water. The aqueous extract was placed in a hot magnetic stirrer (50°C) for two hours. The sample was then filtered using a Whatman filter paper. The filtrate was stored at 4°C for future use. The filtrate if stored under given condition can be used until for 1 week

3.9.2.2. Synthesis

0.2 M of zinc acetate dihydrate was dissolved in 40mL of distilled water and stirred for half an hour using a magnetic stirrer. 30 mL of aqueous extract of purple tea was mixed homogenously with 40ml of zinc acetate solution. The mixture was left to dry overnight at 60 °C to yield pale-white ZnO nanoparticles (pale-white or brown). The precipitate of pure ZnO nanoparticle was left to dry overnight at 60°C. The dried sample was pulverized using ceramic mortar and pestle.

3.9.2.3. Calcination of sample using two different temperatures

Calcination was done at 100 °C and 400°C for 1 h. The calcined sample was preserved in an airtight container until further analysis. X-ray diffractometer was employed to analyze the calcined powdered sample and confirm the size of ZnO nanoparticles synthesized.

3.9.2.4. Characterization

The synthesized ZnO NPs were characterized using UV–visible spectrophotometry, scanning electron microscopy (SEM), and X-ray diffraction (XRD), which has been elaborately mentioned previously (Refer section 3.9.1.4)

3.9.2.5. Antimicrobial activity

This study too consisted of antimicrobial activity of synthesized ZnO NPs against gram-positive bacteria (*Staphylococcus aureus* and *Bacillus subtilis*) and gram-negative bacteria (*Escherichia coli* and *Klebsiella pneumonia*). The same protocol was followed as mentioned in the previous section (Refer section 3.9.1.4.4.)

3.10. Tea Bioinformatics

3.10.1. Data collection

The coding sequence of the reference genome of *Camellia sinensis* var. *sinensis* cv. Shuchazao (Xia *et al*., 2020) was downloaded from the Tea Plant Information Archive (http:// tpia.teaplant.org/download.html). The nucleotide, as well as the amino acid sequence, was downloaded from TPIA (Xia *et al*., 2019) for further codon usage and amino acid usage analysis (Figure 3.7).

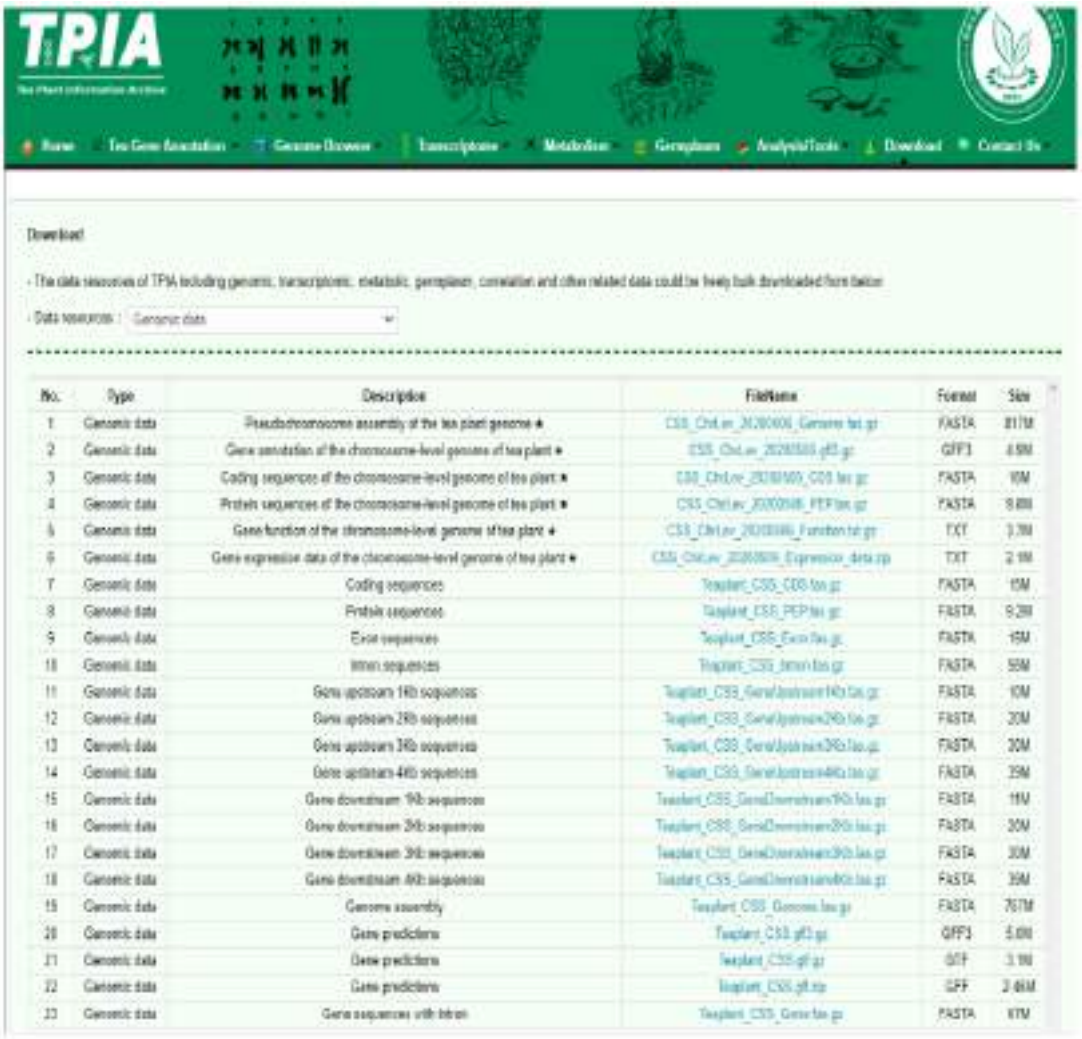


Fig.3.7. Snapshot of data retrieved from Tea Plant Information Archive for codon and amino acid usage of *Camellia sinensis* genome

3.10.2.Codon usage

Different codon usage parameters like Nc, GC, GC3, A3s, G3s, T3s, C3s, Fop, and relative synonymous codon usage (RSCU) described below were calculated using Codon W (version 1.4.2) program (http:// www.molbiol.ox.ac.uk/cu).

The effective number of codons (Nc) is a simple measure of bias in codon usage. It is not dependent on the length of a gene and serves as a direct estimate of synonymous codon usage in the given set of genes or genome (Wright, 1990). The extremity of codon usage bias is seen if the value of Nc ranges 20, one codon codes only for a single amino acid whereas for genes showing no codon bias, the Nc value ranges 61. The value of Nc is calculated employing the formula;

$$Nc = 2 + s + \left(\frac{29}{s^2 + [1 - s]^2}\right)$$

where 's' symbolizes the GC content at

the third position i.e., the GC3s frequency (Wright, 1990).

GC3 is the frequency of codons ending with Guanine [G] or Cytosine [C] in the third synonymous position of the codon which also gives a good indication of the degree of bias in nucleotide composition without including Met, Trp, and termination codons (Roy *et al.,*2015; Sablok *et al.,* 2011).

A3s, G3s, T3s, C3s indicate the composition of the four nucleotide bases i.e., Adenine [A], Guanine [G], Thymine [T], and Cytosine [C] at the third synonymous codon position. The frequency of optimal codons (Fop) signifies the synonymous codons used optimally following the abundance of tRNAs (Ikemura, 1985). RSCU gives a measure of the heterogeneity in the synonymous codon usage pattern (Sharp and Li, 1986) which is represented as;

$$RSCU = \frac{\text{Frequency of codon}}{\text{Expected frequency of codon (in uniform codon usage)}}$$

The value of RSCU>1 indicates that the observed frequency of synonymous codon is greater than one whereas RSCU<1 indicates vice versa.

3.10.3.Amino acid usage

Relative amino acid usage (RAAU) and other amino acid usage parameters like GRAVY (hydrophobicity) and aromaticity were also calculated using the Codon W (version 1.4.2) program (http://codonw.sourceforge.net/). The total protein energetic cost (biosynthetic cost) of the genomes was calculated using the DAMBE5 software (http://dambe.bio.uottawa.ca/) (Akashi and Gojobori, 2002).

3.10.4.Codon adaptation index (CAI)

Codon adaptation index (CAI) gives a measure of the extent of bias towards a preferred set of optimal codons in a gene. CAI gives an accurate estimate of the extent of bias among the codons in a gene of interest or a set of genes under scrutiny. Other than its principal use for calculating the efficiency of translation precision, it has been used to study functional conservation of gene expression across different microbial species, to envisage protein production, and to optimize DNA vaccines (Futcher *et al.,* 1999; Lithwick and Margalit, 2005; Ruiz *et al.,* 2006; Xia, 2007). The value of CAI ranges from 0 to 1. CAI value around 1 reveals the extent of a strong bias towards the preferred codons or rather a codon usage and thus, reflects the high expression level of the measured gene (Sharp and Li, 1987). CAI of a coding sequence (CDS) is calculated from the codon frequencies of the coding sequence and the codon frequencies of a set of known highly

expressed genes (reference set) using the following formula:

$$W_{ij} = \frac{f_{ij.ref}}{Max f_{i.ref}}$$

Here, $f_{ij.ref}$ denotes the frequency of codon j in synonymous codon family i, and $Maxf_{i.ref}$ denotes the maximum codon frequency in synonymous codon family i (Xia 2007).

The CAI value is computed as;

$$CAI = \exp\left(\frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}[f_{ij}\ln(W_{ij})]}{\sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}}\right)$$

Where m refers to the number of synonymous codon families, n_i refers to the number of synonymous codons in codon family i, and f_{ij} indicates the frequency of codon j in codon family i. The exponent is a weighted average of ln(w). The utmost CAI value is 1 (Xia, 2007).

The CAI values for the protein-coding sequences of the tea genome 82 were calculated using the DAMBE5 software (http://dambe.bio.uottawa.ca/) (Xia, 2013). After the computation of CAI values, the top 10% of the genes with the highest CAI values was considered as the potentially highly expressed (PHX) genes and similarly, the 10% genes with the lowest CAI values were considered as potentially lowly expressed (PLX) genes (Sen *et al.,* 2008).

3.10.5.Correspondence analysis (CoA)

It is a multivariate statistical tool that is employed to check the variation in codon usage or amino acid usage in a genome and is thus used to analyze multidimensional data (Perriere and Gouy, 1996; Roy *et al.,*2015). Following the variation in the data, CoA places the data in a continuous axis of separation. CoA analyzes RSCU or RAAU for 59 useful codons excluding Met, Trp and other three stop codons and plots the data along 59 orthogonal axes with 41 degrees of freedom according to the variation in the data of codon usage or amino acid usage (Perriere and Gouy, 1996; Roy *et al.,* 2015; Sablok *et al.,*2011). Most of the variation in codon usage is captured by axis 1 with a gradually diminishing amount of variance along the succeeding axis. RSCU or RAAU value around to 1.0 reflects the equal use of synonymous codon usage without any bias (Sablok *et al.,* 2011). The same Codon W (version 1.4.2) program (http://www.molbiol.ox.ac.uk/cu) was used to calculate CoA based on RSCU and RAAU.

3.10.6. Cluster of Eukaryotic Orthologous Groups (KOG)

Analysis of Eukaryotic Orthologous Groups (KOG) helps in exploring the functional and evolutionary patterns in proteins of a eukaryotic organism.

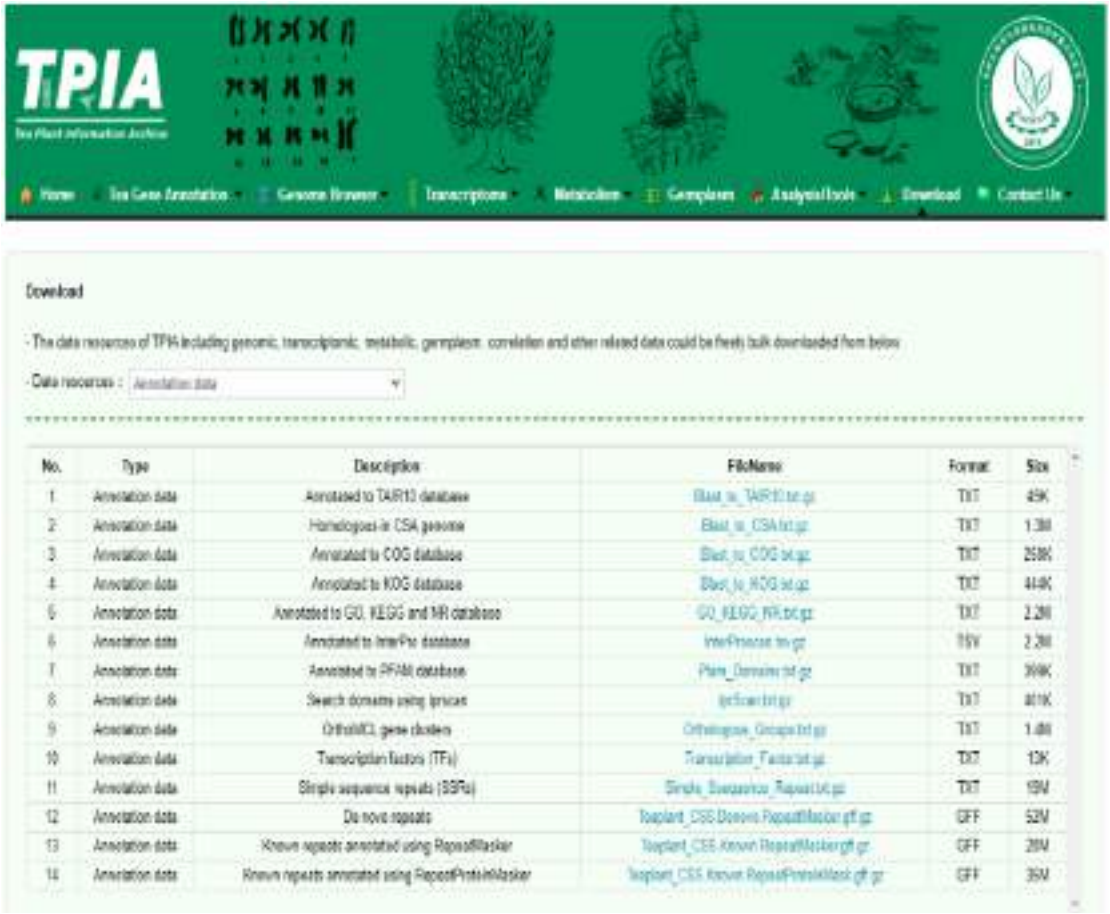


Fig.3.8. Snapshot of KOG data retrieved from Tea Plant Information Archive

During KOG scrutiny the phyletic patterns of KOGs are analyzed in connection with known and predicted protein functions (Koonin *et al.,* 2004). The KOG data (Figure 3.8) was downloaded from TPIA- Tea Plant Information Archive (http:// tpia.teaplant.org/download.html).

3.10.7. Statistical analysis

SPSS version 21 was implemented to interpret all correlations (based on Spearman's rank correlation method) and statistical tests of significance (at levels of significance $P < 0.01$ $P < 0.05$).

3.11. *In silico* polypharmacology of tea

3.11.1. Target Prediction

The SwissTarget Prediction web tool (http://www.swisstargetprediction.ch/, Gfeller *et al.,* 2014) was used to predict the constituent phytomolecules of the tea extract using protein homology-based target prediction. The chemicals' structure and information were obtained from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/ search/search.cgi) (Kim *et al.,* 2021). The accessible compounds' SMILES formulas were acquired and uploaded to SwissTarget Prediction. As our

target organism, we choose "*Homo sapiens*." For further analysis, predicted targets with a likelihood score of 60% or higher were selected.

3.11.2. ADME analysis

The ADME (absorption, distribution, metabolism, and excretion) examination of phytocompounds is used to find candidates that have the best possibility of becoming a medicine. As a result, we used SwissADME (http:// www.swissadme.ch/, Daina *et al.,* 2017). to estimate the drug-like character and ADME parameters of bioactive substances by submitting the list of SMILES formula of compounds with higher probability rates.

3.11.3. Protein-protein interaction (PPI) and network construction

To obtain Target-Target interaction networks as well as their interaction with other human proteins, the list of likely targets acquired from SwissTargetPrediction was uploaded to STRING Database v11.0 (https://string -db.org/) (Szklarczyk *et al.,* 2021). To limit the chance of getting erroneous findings, we set the interaction confidence level to 0.9 and the maximum number of interactions to 50.

3.11.4. Gene enrichment and functional annotation analysis

Gene enrichment and functional annotation analysis were performed using DAVID Bioinformatics Resources v6.8 (https:// david.ncifcrf.gov/tools.jsp, Huang *et al.,* 2009), a free online bioinformatics server. Disease annotation along with Gene Ontology (GO) enrichment analyses was conducted. Biological Process (BP), Cell Component (CC), and Molecular Function were all included in the GO enrichment analysis (MF). The KEGG Pathway was used to conduct the pathway enrichment study. Only results with a significant p-Value ($p \leq 0.05$) were considered.■