Medicinal and molecular profiling of selected tea varieties of Darjeeling and Dooars

Thesis submitted to the University of North Bengal
For the Award of
Doctor of Philosophy
in
Botany

By

Reha Labar

Supervisor
Prof. Arnab Sen

Co- Supervisor
Dr. Malay Bhattacharya

Department of Botany
University of North Bengal
Raja Rammohunpur, Siliguri
2022

This work is dedicated to my Family

Curigional

Document Information

Analyzed document	Reha Labar_Botany.pdf (D129002741)
Submitted	2022-02-28T07:02:00.0000000
Submitted by	University of North Bengal
Submitter email	nbuplg@nbu.ac.in
Similarity	1%
Analysis address	nbuplg.nbu@analysis.urkund.com

Sources included in the report

W	URL: https://www.mdpi.com/2304-8158/9/4/483/pdf Fetched: 2021-03-22T15:36:20.5030000	1
W	URL: https://www.science.gov/topicpages/g/green+tea+oolong.html Fetched: 2020-06-16T09:07:15.5700000	1
W	URL: https://archive.org/stream/in.ernet.dli.2015.228381/2015.228381.Tea-Cultivars_djvu.txt Fetched: 2022-02-28T07:03:09.9800000	7
W	URL: https://crazyteamakerblog.wordpress.com/page/2/ Fetched: 2021-10-07T13:51:50.0400000	1

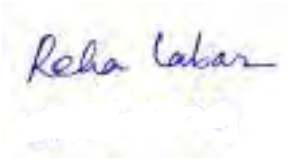
Reha Labar
09.03.2022

Dr Arnab Sen
Professor in Botany
University of North Bengal

Dr. Malay Bhattacharya
Assistant Professor
Department of Tea Science
University of North Bengal

Declaration

I declare that the thesis entitled “Medicinal and molecular profiling of selected tea varieties of Darjeeling and Dooars” has been prepared by me under the supervision of Prof. Arnab Sen, Department of Botany, University of North Bengal and under the co– supervision of Dr. Malay Bhattacharya, Assistant Professor, Department of Tea Science, University of North Bengal. No part of the thesis has formed the basis for the award of any degree or fellowship previously.



[Reha Labar]
Department of Botany,
University of North Bengal
RajaRammohunpur,
Siliguri-734013
Date: 09.03.2022

Certificate

I certify that Miss. Reha Labar has prepared the thesis entitled "Medicinal and molecular profiling of selected tea varieties of Darjeeling and Dooars", for the award of Ph.D. degree of the University of North Bengal, under our guidance. She has carried out the work at the Department of Botany, and at the Department of Tea Science, University of North Bengal. The results incorporated in this thesis have not been submitted for any other degree elsewhere.

[Prof. Arnab Sen]
Supervisor
Department of Botany
University of North Bengal
Raja Rammohunpur,
Siliguri-734013

Date: 09/03/2022

Dr Arnab Sen
Professor in Botany
University of North Bengal

[Dr. Malay Bhattacharya]
Co-Supervisor
Department of Tea Science
University of North Bengal
Raja Rammohunpur,
Siliguri-734013

Date: 09/03/2022

Dr. Malay Bhattacharya
Assistant Professor
Department of Tea Science
University of North Bengal

Abstract

"Tea", the worldwide renowned beverage is not only popular as a drink but also as a healthy beverage. The brain-stimulating drink is a popular refreshing cup in every household, street, or corner of this world. The presence of phytochemicals with pharmaceutical importance and health benefits has rendered its popularity among consumers, industrialists, and researchers. Green tea and black tea are the most popular of different types of tea. Lots of research works related to tea has been compiled for decades and it is still being compiled since "tea' not only has a complex genome or a complex taxonomy, it is also a rich reservoir of ailments to various diseases which has been already explored or is yet to be explored.

The main objective of this thesis was centralized in the medicinal and molecular profiling of some elite tea clones of Darjeeling and Dooars. The research topic was framed to explore the uncharted areas of research in tea with special reference to the elite tea clones established for Darjeeling and Dooars.

Germplasm characterization is critical for plant breeders to efficiently use and manage gene banks to promote cultivar improvement and increase crop productivity. When compared to biochemical and molecular markers, morphological descriptors are a significant preliminary and cost-effective characterization of genetic resources. It undoubtedly has numerous drawbacks when subjected to the effects of environmental factors or parameters, resulting in constant fluctuations and enhanced plasticity. Herein, we made an effort to collect or collate the morphological data of the selected tea clones. The data is recorded based on the parameters like standard, yield, quality, suitability, resistance and susceptibility towards various pests and diseases, flavor, etc.

Environmental factors such as plant age and phenology influence morphological features, making the use of such descriptors problematic in the identification and discerning of genetic variation. As a result, molecular markers such as RAPD and ISSR were used to investigate the genome directly since molecular markers are less affected by environmental factors, removing the limits of phenotyping.

We employed a total of 46 different decamer primers to study the genetic diversity of 33 tea accessions. 45 of the

46 RAPD primers tested, resulted in distinct and scorable bands. RAPD screening yielded a total of 807 bands, 4 were monomorphic and the rest 803 were polymorphic. The 45 random primers had a polymorphism percentage of 99.50% and a PIC value of 0.41 on average. The number of polymorphic bands ranged from nine (OPA16) to twenty-six (OPA26) (OPB10). A total of 15 ISSR primers was also employed to screen the genetic differentiation among the 33 tea clones. All of the primers amplified distinct and scorable bands which rendered a total of 298 bands were created, all of which were polymorphic with a PIC value of 0.42 and showed 100 percent polymorphism. The polymorphic bands varied from 14 (UBC808) to 28(UBC808) (UBC825). The Dice coefficient of similarity created a similarity matrix with values ranging from 0.581 to 0.844. The highest similarity at a nodal value of 0.84) was deduced from the dendrogram where the clones like P-1258 and RR4/5, C28 T-246, and C29 TV-19 shared high relatedness. The ISSR primers yielded similar results, with the clones clustering into two primary groups and subgroups, and the dendrogram deduced the highest similarity at a nodal value of 0.82, between the clones like K1/1 and B-15/263. The results of RAPD and ISSR profiling were mostly coordinated by grouping into two major groups and several subgroups. Regardless of the sample source, sampling area, or point of selection or collection, the pattern of sub-grouping does not disclose a distinct pattern of grouping.

Apart from utilizing molecular markers like RAPD and ISSR, no comprehensive attempt was carried out to explore the genetic variations within the chloroplast region of *Camellia sinensis*. Hence, an initiative was undertaken to explore the variation within the matk region employing the barcode primer (matK). We determined the DNA barcode of the tea clones employing the matK (maturase K) gene and investigated the variation within the chloroplast region. A BLAST search of the NCBI revealed 24 clones to be 100% identical to *C. sinensis*. The other remaining clones showed 99.29–99.89% identity with *C. sinensis*. Clones like Thurbo 3 (11125) and Thurbo 9 (11126) showed more similarity with *Camellia mairei* with a percent identity value being 99.87% and 99.57%.

Clones like Thurbo 3 (11125) and Thurbo 9 (11126), on the other hand, exhibited a higher proportion of similarity to other *Camellia* species, with identity values as 99.87% and 99.57% respectively, when compared to 99.61% and 99.29% similarity with

Camellia sinensis. The unique cluster in the phylogenetic tree demonstrated the relatedness to other *Camellia species*. Our analysis reports a total of 14 variable sites within the matK region, with nine in the high consensus region and five in the low consensus region. This is the first report of barcode analysis of Indian tea clones, in which we successfully used the single locus matK gene to examine variation within the chloroplast region and came to the conclusion that the matK region is not 100% conserved across *Camellia* species.

Phytochemical research in tea has been primarily focused either on processed or packaged tea. Therefore, we aimed at phytochemical extraction from mature fresh tea leaves utilizing extracting solvents with varied polarity. The sole aim of this particular research was not to study the variation of the phytoconstituents among different clones but to screen the potent solvent for extracting phytochemicals from the underutilized mature tea leaves. Different qualitative, quantitative, and antioxidant assays were used to identify potent solvents. Antimicrobial screening and its validation were carried out utilizing GC-MS and *in silico* methods. Acetone extracts were discovered to be the most effective extraction solvent, followed by methanol and ethanol.

The three solvent extracts, i.e., acetone, methanol, and ethanol consistently proved to be the most effective solvents for extraction of phytochemicals like flavonoid, tannin, steroid, diterpenes, terpenoids, coumarin, cardiac glycoside, saponin, and reducing sugar. However, traces of phytochemicals like tannin, cardiac glycoside, and flavonoid were extracted by the less polar solvents benzene, chloroform, and hexane. In-depth antioxidant profiling revealed decent free radical scavenging activity of varied solvent extracts of the mature tea leaf confirming their potent role against oxidative stress. The *in vitro* free radical scavenging assay through DPPH revealed enhanced antioxidant potentiality of acetone extracts with an average inhibition percentage of 79.86% at 200ug/ml followed by methanol (75.31%), and ethanol (73.90%) when compared alongside the standard ascorbic acid (94.82%). Enhance antioxidant activity of acetone extracts was also evident from ferric reducing power assay, total phenol assay with total phenol quantified as 37.77 mg GAE/g. The highest flavonoid content was recorded as 722.94 mg QE/g in methanolic extract although the average value surpassed in acetone extracts.

The potent extracts showed significant inhibition against *S. aureus* with the

Minimum inhibitory concentrations of 4 mg/ml (acetone extracts) and 8 mg/ml, (methanol extracts) respectively. Probable antimicrobial compounds like phenol, 3, 5-bis (1, 1-dimethyl ethyl), caffeine, and Vitamin E were detected by Gas chromatography-mass spectrometry analysis. With a binding affinity of -7.2 kcal/mol to *S. aureus* DNA gyrase, *in silico* findings confirmed phenol as the most powerful antibacterial compound. The potency of a particular solvent in extracting a certain group of chemicals was thus determined through detailed qualitative and quantitative phytochemical profiling. Out of the other bacterial strains tested, acetone and methanol extract showed increased bioactivity against *S.aureus*. The type of chemicals extracted was determined by GC-MS analysis of extracts. Further *in silico* experiments revealed some intriguing details about phenol's ability to bind to *S. aureus* DNA gyrase.

Apart from molecular and phytochemical aspect unchartered areas of tea research was explored in this study. “Tea”, the reservoir of phytochemicals, especially green tea and black tea has been utilized to synthesize metal nanoparticles via the green route of synthesis. However, no comprehensive attempt has been made to explore the underutilized, other forms of tea. Therefore we aimed to synthesize silver and zinc nanoparticles employing the purple tea extract.

The synthesis of silver nanoparticles (AgNPs) was initially monitored from the change in color from colorless to brown and its characterization further affirmed its synthesis. The reaction mixtures containing different molar concentrations of $AgNO_3$ i.e., 1 mM-10 mM produced Surface Plasmon Resonance (SPR) band ranging from 414 nm - 442 nm. The constant increase in SPR band intensity up to 3mM revealed that the increased concentration of $AgNO_3$ also increased the synthesis of AgNPs. The intensity of the SPR band increased up to 3mM following which it decreased. With an increase in $AgNO_3$ concentration, a red sift towards higher wavelengths was noticed from 418 to 442 nm. The photocatalytic effect aided in the rapid synthesis of AgNPs. The sample kept at 30°C (in the sun) produced a sharper SPR band at 430 nm, whereas the other reaction mixture stored at various temperatures showed a blue shift towards lower wavelengths (304nm). XRD analysis clearly showed the crystalline profile of AgNPs showing major peaks of Silver, at three diffraction angles 38.03°, 44.15°, and 64.40° corresponding to the Ag [111], Ag [002], and Ag [220] respectively. The XRD study also determined the size of the synthesized AgNPs as 38.4

nm. The obtained AgNPs were found to be stable only for 24 hours and after that settled at the bottom due to agglomeration.SEM revealed non-spherical and irregular morphology of synthesized AgNPs with a rough surface with the particles size ranging between 10nm to 40nm in size. The synthesized AgNPs showed antimicrobial properties against *Staphylococcus aureus* and *Bacillus subtilis*.

The highly blue-shifted, maximum absorption peak appearing about 273 nm confirmed the nanoscale production of ZnO product although the highest absorption for bulk ZnO usually occurs around 385 nm. The decrease in absorption could be due to nanoparticle agglomeration and settling in the cuvette. XRD analysis of the sample annealed at 400°C revealed the major peaks [100], [002], and [101] of ZnO occurring at the diffraction angles 31.69º and 34.77º and 36.1º respectively. These diffraction angle values for the ZnO peaks are quite similar to the PDF card number 01-080 -0074, showing that pure ZnO nanoparticles were formed. The samples annealed at 400°C clearly show the creation of ZnO nanoparticles, whereas the samples annealed at 100°C exhibit just the [002] peak with low intensity. The nanoparticle's size was estimated to be on the order of 22.4 nm using XRD analysis. The synthesized ZnO was found to be spherical, with an approximate size of 15-25nm, as indicated by SEM. The synthesized ZnO NPs showed antibacterial activity against all the strains studied with predominant inhibition against *Staphylococcus aureus* and *Bacillus subtilis*.

Bioinformatics study on the nuclear genome revealed *C. sinensis* to be AT-rich, where the preference of AT-rich codons over the GC ones was evident. The total codon usage pattern revealed that AT-rich codons with RSCU > 1 coded for 15 amino acids. 29 of the 61 codons were AT-rich, with 22 codons ending in either Adenine or Thymine. Our research revealed the use of AT-rich codons such as AAG, GAA, GAT, AAT, AAA, GTT, GCT, TTG, CTT, TCT, and ATG.

The AT nucleotide compositional constraint was the predominant factor governing the codon usage pattern of *C. sinensis* as evident from a significantly high and positive correlation of Axis 1 of Relative synonymous codon usage (RSCU) with A3 and T3 indices. The pronounced effect of nucleotide compositional constraint was also evident from GC3 versus Nc plots. Apart from that the GC3 versus Nc plot also deduced the role of mutational bias and

translational selection in dictating the codon usage. The impact of translational selection was further intensified by the RSCU based scatter plots where the potentially highly expressed genes (PHX) and the potentially lowly expressed genes (PLX) clustered separately on the extreme opposite ends on the axis of separation. Gene expression level was the other determinant affecting the codon usage signatures as evident from the correlation. The complementary strand-associated genes clustered together in RSCU-based scatter plots of leading and lagging strand-associated genes, indicating the absence of replication-associated mutational pressure. Amino acid usage revealed the highest use of Leucine (L), Serine (S) and the other amino acids used in higher frequencies were Alanine (A), Lysine (K), Glutamic acid (E), Valine (V), and Glycine (G). Strong and significant correlations of the two major axes of RAAU with hydropathicity index GRAVY and aromaticity of the gene products revealed hydrophobicity and aromaticity to govern the amino acid usage signatures of the *C. sinensis* genome. The role of gene expression level in dictating amino acid usage in amino acid usage in all the *Camellia sinensis* was apparent from the high correlation between expression level (CAI) and the two major axes of separation of genes based on RAAU. A significant positive correlation noted between (CAI) and energy cost in the tea genome ($r = .271$, $p<0.01$) indicated the utilization of less energetically costly amino acids in highly expressed genes.

Lifestyle disorders are gradually taking center stage in the global clinical landscape. Increased stress and poor eating habits have increased the risk of diseases such as type 2 diabetes, anxiety, obesity, cardiovascular issues, and so on. Every disorder leads to a chain of disorders. For instance, obesity could lead to diabetes, polycystic ovary syndrome, a higher level of triglycerides which can further add up to hypertension, etc. This chain of linked disorders could act vice versa. Some of these disorders could be treated naturally with functional meals and drinks. Tea use is rapidly increasing in popularity due to its numerous health advantages. Many works have been accomplished in tea as such with specific reference to green tea and black tea. So, the sole aim was to explore the unchartered area of polypharmacology. Purple tea (*Camellia sinensis* var. kitamura) has recently sparked interest as a functional beverage due to its high antioxidant content. To fully escalate the potential of Purple tea (PT), reverse

pharmacology and a target-fishing technique were used. The findings from target-fishing studies have shown PT to be effective against numerous lifestyles disorders such as type 2 diabetes, hyperlipidemia, and coronary heart disease, as well as being important in the treatment of cancer-related complications. Through gene ontology and pathway enrichment analysis, an attempt has also been made to elucidate the mechanism of action of PT against various disorders. It was found to target pathways such as ErbB signaling, Natural killer cell-mediated cytotoxicity, Neurotrophin signaling, PPAR signaling, Insulin signaling, Ras signaling, and others. As a result, we can speculate that PT consumption may protect us from a variety of lifestyle diseases. Additionally, similar studies conducted in under-utilized mature leaf (ML) of tea were mainly found to target the PPAR signaling pathway. ML probably could be effective against various diseases like Type 2 diabetes, obesity, polycystic ovary syndrome, higher level of plasma HDL cholesterol (hypercholesterolemia), high level of triglycerides, Alzheimer's disease, atherosclerosis, insulin resistance, metabolic syndrome, cardiovascular and coronary heart diseases, ovarian cancer, fatty liver, anxiety disorder, hypertension.

Conclusively, this work touches the unexplored multidimensional areas of tea research with special reference to the clones established for Darjeeling and Dooars in a hope that this present endeavor would pave its path for future research for the management, development, and utilization of the most popular health drink.

Preface

“A journey, after all, neither begins in the instant we set out nor ends when we have reached our doorstep once again. It starts much earlier and is never over because the film of memory continues running on inside of us long after we have come to a physical standstill. Indeed, there exists something like a contagion of travel, and the disease is essentially incurable” ― *Ryszard Kapuscinski.*

The thought of pursuing research had already kindled within me when I was doing my master's in Visva Bharati. The thought of wanting to do something for my birthplace and my people made me keen on “Tea Research” since tea is what has given a worldwide identity to my Darjeeling and my people. Such a naive thought and it manifested into reality before I could even realize that it happened. It is said that Life is indeed a journey with different paths and every path I took, every door I knocked just kicked me somehow to the path where I started a new journey of Ph.D. in the Department of Botany, University of North Bengal. Everything seemed surreal when I joined as an RGNF fellow under the project of Tea Research.

“At times, our light goes out and is rekindled by a spark from another person. Each of us has cause to think with deep gratitude of those who have lighted the flame within us.” – Albert Schweitzer

We should always develop an attitude of gratitude and be grateful for everything and every step we take with the help of others. Today, I take this opportunity to thank my supervisor Prof. Arnab Sen, Molecular Cytogenetics Laboratory, Department of Botany, University of North Bengal. Prof. Arnab Sen was the one who opened the door, welcomed me, and showed me the way when I had nothing in me that could contribute to his research endeavor. He neither tested me nor examined me and without hesitation or any reluctance asked me to join his lab from the very next day. The very first day he mesmerized me with his humility and the next moment I knew I belonged here. He had no idea about my capabilities but he allowed me to explore research, being in his team, and

always pushed me to find myself and do things on my own. He allowed me to fall and at the same time felt proud when I stood up on my own. Words are certainly not enough to express my gratitude to you. Not many people have the privilege to work under someone like you. I shall be forever grateful to you for giving me that privilege.

Albert Einstein once quoted, “In the middle of difficulty lies opportunity” and I am much privileged since my opportunity was knocked along with my co-supervisor, Dr. Malay Bhattacharya, Molecular and Tissue Culture Laboratory, Department of Tea Science, University of North Bengal. He made every path look easier when I was even reluctant to walk. He never allowed me to be still, instead encouraged me to keep moving. I shall always abide by your words, “First do it and we shall see what happens.” I do not have enough words to express my gratitude for the efforts you made for me and I shall always be indebted to you.

I am very much grateful to all the teachers of the Department of Botany as well as the Department of Tea Science. Their valuable suggestion and ideologies inspired me and helped me in shaping this thesis. I extend my gratitude to all the non-teaching staff of the Department of Botany who continuously worked round the clock to make our work easier. Some knew more about our deadline of paperwork and like a savior, they walked from one lab to the other and reminded us beforehand while few of them constantly worked under the scorching heat taking care of the gardens and even the plants planted by the scholars. I am very much grateful for all the unconditional service you have provided to us. I would also like to thank the non-teaching staff of the Department of Tea Science who waited for us selflessly beyond their working hours.

“God sends help in disguise”, and I can't think of a better way to thank them than now. I owe my gratitude to Dr. Laxmikanta Padhi and his wife Dr. Tanmayee Mishra, Dr. Ritu Rai, and Dr. Pranay Bantawa, Mr. Jolen Tamang, and Mrs. Saroja Chettri, Dr. Tilak Saha, and his wife Mrs. Mosumi Sarkar Saha, and my little companion Manyata Saha aka Tinni. They were the ones that accepted me into their family, listened to me, cared for me, and even reprimanded me when I did something wrong. They made me feel at home when I was away from my home during my tenure.

Some left footprints to follow which made my journey easier and with due respect, I would like to thank all my senior Research colleagues, Dr. Ritu Rai, Dr. Ayan Roy, Dr. Sanghati

Bhattacharya, Dr. Indrani Sarkar, and Dr. Pallab Kar, and Mr. Arnab Chakraborty. I especially want to acknowledge the guidance, suggestion, and mentoring of my senior lab mates, Dr. Indrani Sarkar and Dr. Pallab Kar, and Mr. Arnab Chakraborty whose valuable guidance helped me in shaping my thesis.

There is a famous saying of Izaak Walton, "Good company in the journey makes the way seem shorter." When you find a bunch of people walking down the same lane, the journey becomes way much adventurous, experiencing everything together, be it the low and high tides, the uphill and downhill roads, twists and turns, cool breeze, or the scorching heat. My journey of a Ph.D. would certainly be dreary without them. Today, I take this opportunity to thank each one of them. I appreciate the good company of my lab mates, Indrani Sarkar, Mousikha Lala, Soumita Bhattacharjee, Saroja Chettri, and Sandipan Ghosh. Things would have been much complicated, uninteresting, and tiring without my lab associates. Thank you very much for just being there for me and helping me in every possible way. I would also like to acknowledge my budding juniors Sutapa Dutta and Swarnendra Banerjee. I would feel regretful if I fail to mention the name of Mr. Basudev Singha whose presence in our lab makes our work way much easier.

Ryunosuke Satoro quoted, "Individually, we are one drop. Together, we are an ocean."

As we navigate through research, we should always be open to collaboration. Other people and their ideas can always be better than others. We should collaborate since as individuals we can do a lot, but as members of a team, we can achieve more. I am sincerely grateful since I got an opportunity to collaborate with such brilliant minds and I barely have any words to express my gratitude to them. I am very much indebted to the guidance and assistance of Prof. Tapas Kumar Kundu and his inquisitive and hard working scholar Miss Rini Labar, Nanotechnology Lab, Department of Physics, Visva Bharati. Working with them in collaboration not only helped me in shaping my thesis or my research work but also helped me to grow as an individual.

One should always be grateful even for the smallest help or a kind courtesy and generosity. Even a smile can make a difference to this world, forget about kind gestures. I would like to thank a bunch of people who were directly or indirectly helped me in every way possible during my research tenure. I am thankful to Smt. Dipali Sen, Mrs. Gargi Sen for their support, and homely feeling in Siliguri. I am

grateful to my landlord Shyam Ghosh and family whose presence made me feel like having a guardian away from home. I would also like to acknowledge the constant support provided by Mr. Angad Thapa and his family during a crucial time.

I especially want to thank my friends Aishwarya Ramchandran, Yanki Lama, Mridula Singh, Upasana Banerjee, Moumita Das, and Laxmi Tamang for being there for me and encouraged me during my hard times and celebrated with me during my smooth sailing.

I would also want to acknowledge the kind help provided by the scholars of Molecular and Tissue Culture Laboratory, Department of Tea Science, University of North Bengal. I very much appreciate the kind help and gesture of Mr. Prashant Barailey, Fulbari Tea Estate for enlightening and helping me with some crucial facets of tea. I am also gratified by all the necessary help provided by the staff of the administrative section, especially the Ph.D. cell, student's cell, finance cell, and the registrar branch of the University of North Bengal. I owe a debt of gratitude to the persons whom I failed to name or enlist here but who, intentionally or unknowingly, contributed significantly to the successful completion of this work.

I am obliged to University Grants Commission for providing me the funding support through UGC- RGNF-SC fellowship during the tenure of five years.

I also appreciate the kind help provided by Mrityunjay Chaube, in charge, Darjeeling Tea Research, and Development Center (DTRDC), Kurseong.

"When I started counting my blessings, my whole life turned around." – Willie Nelson. I have received the highest form of blessing in my family. I am blessed and feel lucky to be born as a daughter to Mr. Suresh Kumar Labar and Mrs. Nikita Labar. I would certainly not have arrived here today if they had not encouraged and supported me in every aspect of life. I have always idolized my father and mother and no one else. They taught me through their actions to never take things for granted but instead keep on moving ahead. They gave me my life, the ladder of dreams, steps of freedom, and zeal to pursue and move forward no matter what. They also gave me my dear siblings Miss. Rini Labar and Mr. Riaz Labar who has always been a constant support during my hard and smooth sailing of life. My sister Rini Labar is my friend and someone I have always looked up to. She has protected me in every aspect and she is the key to my inspiration and I have no regrets about following in her footsteps. My brother Riaz Labar is in awe of

inspiration and my only reason to keep me going since I have to show him the way my sister did. He has been the energy that keeps me going. I have no words for how grateful and blessed I feel being part of you and I can express my gratitude only through my unconditional love towards you all.

“Family isn't always by blood. It is the people in your life who want you in theirs; the ones who accept you as you are. The ones who would do anything to see your smile”- Nishan Panwar.

My in-laws Mr. Balaram Sundas and Mrs. Geeta Sundas became my parents as soon as I wedded their son. I feel blessed and I am grateful to them for all the things they have done for me and I respect their understanding towards my ambition and responsibility as a daughter. I feel extremely happy, lucky, and grateful to have them beside me as my support. I am also thankful to my dear sisters Prerna Subba, Asmita Sundas, and Aditi Sundas for their moral support and understanding.

I have no words to express my gratitude towards my better half, Anupam Sundas. You certainly are my best friend, my guidance, and my inspiration. I am very lucky to have you as my life partner. I am indebted to you for all your support and understanding. You made my dream as yours and you fostered my dreams with your dedication, devotion, unconditional love, and moral support. You never allowed mere responsibilities of a wife towards a husband to overburden me but instead took the burden of my responsibilities so that I can complete my work and my thesis. Thank you for always believing in me more than myself.

I extend my gratitude to my family in Kalimpong, Teesta, and Tindharia. I will be always grateful for the admiration you showed towards me and my dreams.

I have saved the dear one and the merciful one for the last since without the glory and blessings of the Lord I would not have come across such a challenge to overcome my weakness. He gave me challenges, showed my weakness, and also sent help in all possible sources. His blessings touched me through various people I came across during this endeavor.

"I give thanks to you, O Lord my God, with my whole heart, and I will glorify your name forever." Psalm 86:12

[REHA LABAR]

Table of Contents

Contd….to next page.

Contd….from previous page..

List of Tables

Contd….to next page.

Contd….from previous page..

List of Figures

Contd….to next page.

Contd….from previous page..

Contd….to next page.

Contd….from previous page..

Contd….to next page.

Contd….from previous page..

List of Appendices

Chapter 1
Introduction

"The real voyage of discovery consists not in seeking new landscapes but in having new eyes."
- Proust

Tea, a popular non-alcoholic beverage is consumed worldwide where the beverage is made from the infusion of the manufactured tea leaves. The world's tea consumption is highest for black tea, followed by green tea, oolong tea, and white tea (Khan and Mukhtar, 2019). The history of tea consumption dated back to the 5th century and by the end of the 6th century AD, tea was consumed not only for a therapeutic purpose but it became popular as a refreshing drink (Wilson and Clifford, 2012). Kakuzo Okakura (1863-1913), a Japanese cultural philosopher and the author of "The book of tea" (Okakura, 1964) stated, "Tea began as a medicine and grew into a beverage". In ancient China and Japan, tea was considered a ritual element and an essential part of daily life. The effective pharmaceutical activity of tea was very much known from the beginning thus ancient China consumed it as a medicinal beverage (Hilal, 2017). Tea in today's era is the most commonly consumed drink next to water (Hoffmann and Manning, 2005).

The tea plant is most suitable to cultivate in areas with increased humidity, moderate temperature, and acidic soils (Dufresne and Farnworth, 2001). *Camellia* is a genus with about 350 species available (Mukhopadhyay *et al.,* 2016). However, only two species, i.e., *Camellia sinensis* (L.) O. Kuntz and *Camellia assamica*, produce tea, for consumption in a variety of

forms, including black, green, and oolong tea (Mondal *et al.,* 2004). However, in some parts of China, a different species, *C. taliensis*, is used to make tea (Rawal *et al.,* 2021). The variation in leaf size i.e., the small leaf tea plant and the large leaf tea plant has raised much speculation regarding the origin of tea hence providing many theories.

There are many theories and speculations about the origin of tea. Some report states that tea originated from Yunan and Sichuan in China based on studies in numerical taxonomy and multivariate analysis (Hasimoto, 2001). The history of tea dates back to China which states that"Tea" was first discovered around 2737 B.C. by the Emperor of China, Shen Nong when the tea leaves secretly blew into the pot of boiling water by a matter of chance (Patel, 2005). Tea was first originated in South Asia and slowly grown widely throughout Asia, Africa, and various parts of the Middle East (Chopade *et al.,* 2008).

Tea was originally introduced in India when the tea seeds were first brought from China and planted at the Botanical Garden of Calcutta in 1780 (Bezbaruah, 1999). Tea was planted in northeast India in 1836. However, reports say that Major Bruce in 1823 discovered wild tea plants growing in some parts of Sibsagar hills, which were then known as Rangpur hills, the capital of "Assam" (Ukers, 1935). Robert Bruce was a Scottish gentleman who introduced the tea plantations in Assam in the early 19th century. The first Indian tea from Assam was shipped to England for public sale in May 1823. After establishing a profitable industry in Assam's Brahmaputra valley, the viability of cultivating tea in the Himalayan foothills and elsewhere in India was investigated. In Kumaon, Dehra Dun, Garhwal, Kangra Valley, and Kulu, 78 plantations had been developed by 1863.

While attempting to introduce tea to India, British colonists discovered that tea plants with thicker leaves flourished in Assam as well and that these, when planted, thrived. The Singphos tribe of Assam had long grown the same plants, using tea chests provided by tribal chief Ningroola. In the past, botanists categorized the Assamese and Chinese types as separate species, but botanists now classify both as the same species, *Camellia sinensis*. The British East India Company began large-scale tea manufacturing in Assam, India, in the early 1820s, using a tea type traditionally prepared by the Singpho people.

Beginning in the 1850s, the tea business quickly grew, consuming vast tracts of land for tea plantations. By the turn of the century, Assam had surpassed China as the world's leading tea-producing region. Robert Fortune, who spent around two and a half years in China working for the Royal Horticultural Society of London from 1848 to 1851, is often credited with introducing Chinese tea plants, which are distinct from Indian tea. Fortune transported 20,000 tea plants and seedlings to the Darjeeling region of India, which is located on steep slopes in the Himalayan foothills and has the acid soil that *Camellia* plants prefer. He also brought a crew of Chinese tea workers who were taught to help in the production of tea leaves. Most of the Chinese tea plants failed to survive, except for few which remained in established in Indian gardens.

Tea plantation in the area of around 3.8 million hectares resulted in the worldwide production of 6.1 million metric tons of tea (Zhang *et al.,* 2020). The economically important tree crop is grown in over 52 countries (Mondal *et al.,* 2004; Chen and Chen, 2012). It is mostly cultivated in the region of China, India, Taiwan, Sri Lanka, Indonesia, Japan, and central African countries (Mukhtar and Ahmad, 2000). China and India are globally the two largest tea producers (Meegahakumbura *et al.,* 2018). As per the data given by FAOSTAT(http://www.fao.org/faostat/en/#rankings/countries_by_commodity), mainland China was the largest tea producer (2777200 tonnes) and India was the second-largest tea producer (1390080 tonnes).

Tea belongs to the genus *Camellia* and comes under the Theaceae family. Linnaeus previously named the tea plant *Thea sinensis*, which was published in Species Plantarum in the year 1753. However, Carl Ernst Otto Kuntze shifted the species to the genus *Camellia* later in 1887. It is now widely accepted in combination as *Camellia sinensis* (Linnaeus) O. Kuntze (Das and Ghosh, 2016).

The taxonomy of tea is still very confusing due to its plasticity of morphological traits that distinguishes tea taxa (Meegahakumbura *et al.,* 2018). Due to the environmental plasticity of morphological traits, previous attempts to classify tea germplasm by using morphological markers, gave unsuccessful results in revealing the taxonomic affinities between accessions (Balasaravanan *et al.,* 2003).

Classification of tea plants given by Sealy (1958) was based mainly on morphological and yield characters and subsequently, tea plants were grouped

Table 1.1 APG IV classification of tea updated by Angiosperm Phylogeny Group (Chase *et al.,* 2016)

Clade:	Angiosperms
Clade:	Eudicots
Clade:	Asterids
Order:	Ericales
Family:	Theaceae
Genus:	*Camellia*
Species:	C. *sinensis*

in two different taxa i.e., dwarf, small leaf, and slow-growing variety, *Camellia sinensis* var. *sinensis* from China and tall, large leaf and quickly growing variety, *Camellia sinensis* var. *assamica* (Masters) Kitamura from Assam (India) i.e., Assam type. White (1962) revised the classification relying mainly on morphological features. He assigned a specific status to var. *sinensis* and var. *assamica* and recognized a Southern or Cambod form of *Camellia assamica* thus differentiating cultivated tea into three main species each with specific plant types viz. *Camellia sinensis* (China type), *Camellia assamica* (Assam type), and *Camellia assamica lasiocalyx* (Cambod type). However, according to Ming *et al.,* (2000), tea plant has broadly two varieties, *C.sinensis* var. *sinensis* (Chin a type) and *C. sinensis* var. *assamica* (Assam type). Although previous reports have placed Cambod type tea as an alternative name for *C. sinensis* var. *assamica*, recent studies revealed it to be a hybrid of China and Assam teas (Wambulwa *et al.,* 2016). Nowadays, all tea is placed under *Camellia sinensis* irrespective of taxonomic differences. However, China, Assam, and Cambod varieties are still referred to as *Camellia sinensis*, *Camellia assamica*, and *Camellia assamica* ssp. *lasiocalyx* respectively (Kaundun *et al.,* 2000).

Sealy (1958) reported 82 species of *Camellia* native to southeast India. However, to date, The Plant List has included a total of 464 scientific names of the plant under the species rank for the genus *Camellia* and out of this a total of 248 are accepted species name (http://www.theplantlist.org/browse/A/Theaceae/*Camellia*/).The classification of tea updated by Angiosperm Phylogeny Group (Chase *et al.,* 2016) is provided in Table 1.1.

Tea is both commercially and taxonomically important among all the species of *Camellia* since other *Camellia* species do not harvest the world-famous brew (Banerjee 1992). Tea offers its contribution greatly to the country's economy as well as job opportunities in countries like China, Kenya, India, Sri Lanka, etc.(Chen *et al.,* 2005). Other species of *Camellia* i.e., non-tea producers are used as

ornamental plants. *Camellia sinensis* var. *sinensis* and C. *sinensis* var. *assamica* are the two main commercial taxa that are largely cultivated. However, some species like *C. taliensis, C. sinensis* var. *dehungensis, C. crassi columna*, have been used as tea in local parts of Asia, and particularly in Yunnan province (Chen *et al.,* 2005). The tea taxonomy is still a challenge and question of debate due to its complexity as well as a particular interest in only one taxon due to its commercial and economic demand. Studies related to taxonomic characteristics; genetic variation and biogeography would be of great benefit to identify genotypes with increased productivity, which can further be utilized to improve the cultivars (Banerjee, 1992).

Morphologically, the tea plant is an evergreen shrub or tree with numerous branches with leaves appearing dark, leathery, shiny, or shiny green, elongate ovate, roughly serrate, and alternate with the short petiole (Gruenwald, 2007). The cultivated plant attains a height of 0.6-1.5m whereas the height of wild tea plants goes up to 10 - 15 m. The leaves are usually alternate, lanceolate, and short-stalked; light green and leathery with serrated margin and hairs present beneath. It varies in length from 5 - 30 cm and about 4 cm in breadth. The mature leaves are usually smooth, coriaceous, and bright green, whereas young leaves are pubescent. The flowers of tree plants are white and fragrant having a diameter of 2.5-4cm and are usually found single or clustered in groups of two or four. The flowers produce brownish red capsules (Mahmood *et.al.,* 2005). The flattened fruit has a smooth, rounded, and trigonous three-celled capsule which bears solitary nut size seeds in each (Biswas, 2006).

The economic beverage crop is highly cross-pollinated and heterogeneous. The commercial pressure exerts improvement chances of high-yielding tea plants but it cannot help in broadening of genetic variability since the choices are confined to select elite mother plants within the usual hybrid populaces of tea (Hajra, 2001). The threat of losing numerous significant tea germplasm is expanding due to the quick uprooting of the old plantation. Therefore, it is very important to characterize the existing germplasm so that we can measure the plasticity shown in varying climatic conditions. Over time, the profundity of cross-pollination has led to the establishment of a broad genetic pool. Consecutively to study the genetic variation and maintain the gene pool, the germplasm of Indian tea is kept and maintained in three regions namely the UPASI

(southern), TRA (northeastern), and IHBT (western Himalaya). These centers have not only maintained the germplasm but also developed and released biclonal seed stocks and varieties based on selection and selective breeding. Several accessions maintained in the gene banks may have few accessions with some specific and favorable dormant alleles, which might be very important for the improvement of cultivars. Overall, the preservation of tea germplasm is primarily of great importance since previous sectors of seed-grown tea are enormously being uprooted and their seed sources hardly exist, are no longer found, or are lost eventually (Karthigeyan and Sud, 2010).

The successful collection of tea genetic resources, its preservation, and utilization, at present or future breeding programs depend entirely on the knowledge of its genetic background relationship, diversity, and identification (Chen *et al.,* 2005).

Discrimination between archetypal China, Assam, and Cambod varieties is very difficult because of the excessive natural hybridization in tea (Visser, 1969). Widespread selection and cultivation of clonal tea based on yield, biotic and abiotic stress resistance, and quality, among the existing materials, can reduce genetic diversity. Therefore, utmost care should be taken to utilize clones of different origins and understand the genetic diversity at the molecular level (Mondal, 2002). Knowledge of the genetic variation is important for the breeding program since it provides the basis for the development of desirable genotypes (Kaundun *et al.,* 2000).

The practice of recurrent use of the same tea plants or popular cultivars as a parent along with the introduction of vegetative propagation for better productivity has rendered the tea plants limited genetic variation and also prone to diseases, pests, and abiotic stresses (Gunasekare *et al.,* 2012). The importance of germplasm collection and conservation is highlighted in many research articles. Core collection for Chinese tea germplasm based on cultivated region grouping and phenotypic data was established to enhance the utilization of genetic resources in improvement programs and simplify their management. Wang *et al.* (2011) has highlighted the importance of biotechnological approaches to strengthen germplasm activity for using them in a rational manner, assembling a functional core of germplasm to enhance practical applicability of tea genetic resources in tea breeding program and addressing different needs of growers and consumers (Gunasekare *et al.,* 2012).

Further characterization of germplasm based on different morphological characters and division accordingly based on traits has highlighted the importance of the phenetic study to utilize the germplasm to manage the gene bank effectively and efficiently (Piyasundara *et al.,* 2006). Marker-assisted breeding provides great advantages for varietal improvement by increasing the gene pool (Balasaravanan *et al.,* 2003). Precise assessment of the genetic relationships either among closely related species or among the cultivars belonging to different geographic areas is very much essential to give the source of information about desirable genes to the breeders and also establish core collections for improvement of the cultivar. Studying intraspecific relationships between the endemic varieties is likewise essential to the source of favored or specific genes.

Different molecular techniques like Restriction Fragment Length Polymorphism (RFLP), and Random Amplified Polymorphic DNA Polymerase Chain Reaction (RAPD), provide a useful approach for studying the variability within plant germplasms (Kaundun *et al.,* 2000). The RFLP technique on the other hand requires a huge amount of pure DNA, which is quite time-consuming and robust. Whereas, the RAPD technique overcomes these limitations since this method can generate a large number of polymorphic bands relatively easily with a minimum amount of DNA. We can overcome the technical limitations of RFLP by the RAPD technique but the sensitivity of this technique to experimental conditions raises a question about its reproducibility (Mondal, 2002). The SSR (Simple sequence repeats) method reveals an extreme degree of allelic variation (Schlotterer *et al.,* 1991). Despite this advantage, this technique is laborious due to the requirement of complete sequence information, which is eventually needed for primer designing (Mondal, 2002). The Inter-SSR-PCR (ISSR-PCR) technique or PCR analysis employing anchored simple sequence repeats primers has grabbed the attention as an alternative method to characterize complex genomes. This technique employs SSR based oligonucleotides attached to either the 5’ or 3’ end consisting of 2-4 purine or pyrimidine residues, which aids in the initiation of PCR amplification of DNA segments flanked by microsatellite repeats that are inversely placed, closely spaced (Zietkiewicz *et al.,* 1994). This technique generates a PCR product, which shows multiple polymorphisms when resolved on an agarose gel. This has gained much attention due to its low cost, which

avoids costly cloning and sequencing, unlike the microsatellite-based method.

Additionally, DNA barcoding is a very precise method for taxonomic identification. It uses a short standard sequence with abundant variation, which helps to analyze variation among the species. The various regions from the plastid genome have been used for DNA barcoding of land plants (Kress and Erickson, 2007; Singh *et al.,* 2012). However, the Consortium for the Barcode of Life (CboL) recommends rbcL and matK as standard. It is mostly suggested for DNA barcoding of land plants due to its higher variability between the species and its role for the phylogenetic re-establishment (Bafeel *et al.,* 2011; Kuzmina *et al.,* 2012). Information on variations in the chloroplast DNA (cpDNA) is being used widely in lineage and variation studies among the population. Inter and intra specific polymorphism can result from the variation of chloroplast DNA. Due to maternal transmission and lack of recombination, chances of detectable geographical subdivision of populations in a species increase (Katoh *et al.,* 2003). Previous work on tea clones established for Darjeeling and Dooars lacks a report on a study of genetic diversity using barcode primers. On the molecular front, there is very little knowledge regarding the study of genetic diversity employing various elite tea clones of Darjeeling using techniques like RAPD and ISSR since it is very important to record data regarding genetic fidelity. There is also a need to employ a barcode technique to study the variation within the chloroplast regions of elite tea clones.

Tea extracts are popular as ingredients in many dietary supplements and foods. Abundant scientific reports on the positive effect of tea on human health have aroused lots of interest among the researchers regarding chemical composition analysis of tea. Tea has a complex chemical composition, containing over 2000 components (Yashin *et al.,* 2015).The tea plant is the reservoir of phytochemical compounds of pharmaceutical importance. The chemical constituents of tea mainly include the polyphenols like catechins and flavonoids, alkaloids like caffeine, theobromine, theophylline, etc. It also consists of volatile oils, amino acids, polysaccharides, lipids, vitamins (vitamin C, Vitamin E, etc.), and inorganic elements like aluminum, fluorine, manganese, etc.(Sharangi, 2009). The polyphenols comprise mainly of the catechins and flavonoids which have mainly beneficial properties imparting good health. The flavonoids have antioxidant, antiallergic, anti-inflammatory, and

anti-microbial properties. The primary catechin compounds mainly present in green tea are catechin, epicatechin, gallocatechin, epigallocatechin, epicatechin gallate, and epigallocatechin gallate (EGCG). Out of these, the epigallocatechin gallate (EGCG) is the most active component with enhanced bioactivity. Along with the assessment of tea quality, knowledge of the chemical composition of tea also helps to screen and manage the process of tea growing, manufacturing, and storage conditions. In particular, knowledge about the tea chemical constituents has enabled the establishment of the relationships between the chemical constituents and properties of tea owing to the identification of compounds imparting its aroma and taste. Therefore, assessment of tea quality should also focus on the extra determination of tea components beneficial to human health which will not only enhance the nutritional value of tea but also increase the quality of tea by optimization of the growth, manufacturing, and storage conditions (Yashin *et al.,* 2015).

Owing to the health benefits of polyphenols in tea different epidemiologic observations and laboratory studies have indicated that polyphenolic compounds present in tea may reduce the risk of a variety of illnesses, including cancer, coronary heart disease, atherosclerosis, high blood cholesterol concentrations, high blood pressure where most studies involved green tea limiting black tea evaluation also suggested combining tea polyphenols with other consumer products like food items and vitamin supplements for better utilization. Studies on the animal model have suggested the protective property of tea against certain cancers, cardiovascular diseases, and neurodegenerative diseases (Yang & Koo, 2000; Mandel & Youdim, 2004; Butt *et al.,* 2015).

The extracting solvent usually diffuses into solid plant tissue and solubilizes compounds with similar polarity (Chaitanya *et al.,* 2013). The successful extraction and determination of biologically active compounds from plants are mainly dependent on the type of solvent preferred and extraction techniques used in the pharmaceutical industry (Tiwari *et al.,* 2011), A varied range of research work has been performed on tea with the majority of work being carried on tea especially green tea and its pharmaceutical importance to prevent a various form of diseases like cardiovascular disease, cancer, neurodegenerative diseases, etc. Most of the research work has been focused on made tea or processed tea putting a limitation as such to tea plant (Mukhtar and Ahmad, 2000). Very

little work has been done on phytochemical screening of fresh leaves (especially the mature ones) using a range of organic solvents with research being limited to standard solvents and processed tea.

Tea is a repository of phytochemicals and among different types of tea, green tea has gained much importance for its health benefits.

Various lifestyle diseases are caused by people's daily routines and habits (Tabish, 2017). The rising level of competitiveness, along with an unhealthy diet, has become a major source of stress. A variety of foods and beverages are becoming increasingly popular as stress-relieving foods and beverages. None, however, have been able to match the appeal of tea, an age-old useful beverage. Tea especially green tea and black tea is popular for its health benefits and various work has been accomplished to elucidate its role in reducing the risk of various diseases. However, there is still a lacuna and needs tea research to focus on other forms of tea. For instance, purple tea (PT) is unique of its kind due to its high level of anthocyanin coloring, which contributes to its powerful antioxidant capabilities (Sun *et al.,* 2016; Lv *et al.,* 2015). Apart from antioxidant function, anthocyanins have been shown to have cytotoxicity against cancer cells, induce apoptosis, and have an immunostimulatory effect (Joshi *et al.,* 2017). Although some of PT's health advantages have been investigated, its entire pharmacological potential has yet to be discovered. Almost no research has been done on the Indian PT, and most people are unaware of its existence, popularity, and health advantages. Furthermore, it is unknown how PT may operate against certain diseases or control many pathways in our bodies. PT has only a rudimentary medicinal understanding, and its mechanism of action is unknown. A thorough examination of Indian PT's secondary metabolites and pharmacological qualities is required to maximize its health benefits and appeal among customers.

Research trends in nanotechnology have also opened up in the green synthesis of various metal nanoparticle using green tea extracts (Gottimukkala *et al.,* 2017; Nakhjavani *et al.,* 2017; Senthilkumar *et al.,* 2014; Selvan *et al.,* 2018). Black tea has also been employed in the biogenic synthesis of various metal nanoparticles like Au and Ag (Begum *et al.,* 2009), Zinc Oxide (Taghavi *et al.,* 2017), palladium (Lebaschi *et al.,* 2017), iron copper, and silver (Asghar *et al.,* 2018). Green synthesis of metal nanoparticles has gained much attention over the years

due to its simple yet large-scale production, less expensive, non-hazardous method, unlike the chemical route which is expensive and hazardous releasing various toxic byproducts (Senthilkumar and Sivakumar, 2014). Tea is enriched with polyphenols which act as both reducing and capping agents during the green synthesis of nanoparticles (Huang *et al.,* 2014). Various types of tea have been employed for the green synthesis but still there lack reports on other less popular or lesser-known types of tea.

Biological research has been improvised along with the tremendous progress of genome sequencing technology. The revolutionary advancement in genome sequencing has not only opened up the floodgates for whole-genome profiling but also has provided enormous scope for exciting investigations regarding the codon usage pattern of several genes and genomes and simultaneously address their evolutionary dynamics (Roy *et al.,* 2015). The patterns of codon usage differ significantly between organisms where genes and genomes exhibit diverse patterns of synonymous codon usage (Guo and Yuan, 2009; Sau and Deb, 2009). The three nucleotide bases building up the triplet codons act as basic coding components that code for a particular amino acid or cause protein chain initiation or termination. Except for Met and Trp, the 20 standard amino acids are coded by two to six synonymous codons which are used at frequencies that vary in different organisms (Feng *et al.,* 2013; Sablok *et al.,* 2011). extensive research has been focused on model plant genomes like *Arabidopsis* (Duret and Mouchiroud, 1999), important genes of monocot and dicot species (Murray *et al.,* 1989), monocot plastomes (Sablok *et al.,* 2011), plastomes of *Coffea arabica* (Nair *et al.,*2012), ESTs of *Citrus* species (Xu *et al.,* 2013), etc. However, detailed investigations relating to codon usage patterns in the whole genome of non-model yet, essential plants remain largely unaccomplished. The complexity of the tea genome made whole genome sequencing and assembly quite challenging but with the advancement of sequencing technology, we can now avail whole genome sequence of the tea plant. This opens up the door to a plethora of research and a new trend of analysis at the Bioinformatics level.

Hence, based on the critical appraisal on tea research as mentioned above, we selected the tea clones established for Darjeeling and Dooars region and laid the following objectives.

Objectives

1. Plant selection and maintenance of

germplasm.

2. Morphological documentation.

3. Molecular documentation

To study the genetic diversity using different RAPD (Random Amplified Polymorphic DNA), ISSR (Inter simple sequence repeat), and RFLP (Restriction fragment length polymorphism) techniques.

4. DNA Barcode

Explore the variation within the chloroplast region employing barcode primer, matK.

5. Phytochemical screening

Qualitative and quantitative phytochemical assessment of different tea extracts prepared using various extracting solvents ranging from polar to nonpolar.

Chemical Characterization and bioactivity study of selected tea extracts.

6. Green synthesis of the metal nanoparticles employing the lesser-known purple tea extracts.

Synthesis of Silver nanoparticles

Synthesis of ZnO nanoparticles

7. Tea Bioinformatics

Exploring the different determinants governing the codon and amino acid usage of the tea plant (*Camellia sinensis*).

8. Tea polyphenols and human health

Use of *in silico* polypharmacology to investigate the efficacy of Purple tea and mature tea leaf of clones endemic to the Darjeeling Hills against various lifestyle– diseases.■

Chapter 2
Review of Literature

"The universe is not only queerer than we suppose, but queerer than we can suppose."

-J. B. S. Haldane

2.1. History of tea

Historical evidence from Chinese mythology states that tea was first discovered in the year 2737 B.C., by Emperor Shen Nung (Han dynasty) as mentioned in the Pen T'sao (a Chinese medical book written during the Han Dynasty, circa 25 to 221 A.D). Tea is believed to be first mentioned in the ErhYa, (a Chinese dictionary circa 400 B.C). However, the term ch'a used for tea was not so popular before until the writing of Lu Yu (*The Classic of Tea, Ch'a Ching)* which was published during 780A.D. As cited in *The Story of Tea* (Heiss and Heiss, 2007), the consumption of tea as a medicinal drink began in Yunnan province during the Shang Dynasty (1500 BC–1046 BC). After the Tang dynasty (618-906 A.D.), tea became popular as a refreshing beverage (Benn, 2015). Tea's origin is likewise a point of contention that has to be clarified. Although the Indo-Burma region around the Irrawaddy River is thought to be the genesis point (Mondal, 2009), it is unclear if Assam and China type of tea have the same or separate domestication origins (Rawal *et al.,* 2021).

The famous tea transferred from one culture to another so the brewing, consuming method and chemical aspects also varied among different cultures. However, the modern way of tea brewing technique by infusing tea leaves into hot water arose during the

Ming dynasty (1368 to 1644) and subsequently got popular throughout the world.

During the Tang dynasty tea was believed to be consumed as a vegetable with which they prepared soups mixing with other ingredients. Later Brick tea (Mair and Hoh, 2009) made by steaming and compressing the tea leaves into bricks, got popularized. The powdered form of tea with bright tea green color and low astringency got popularized during the Song Dynasty (960 to 1279 A.D.). In the mid 13th century, the Chinese discovered a new way to process tea. Rather than being steamed, tea leaves were roasted and crushed. Unfermented tea leaves were pan-fried, then rolled and dried throughout the Yuan and Ming dynasties. This inhibits the oxidation process that darkens the leaves, allowing the tea to remain green. Oolong tea was invented in the 15th century when the tea leaves were allowed to partially ferment before being pan-fried (Benn, 2015). Yellow tea was discovered by accident during the Ming dynasty's manufacture of green tea, when seemingly poor techniques caused the leaves to turn yellow, thus yielding a distinctive flavor (Mair and Hoh,2009).

Tea drinking and tea culture transferred to Japan, Korea, Portugal, Italy, Iran, Taiwan, United Kingdom, America, Australia, Africa, Srilanka, India, and so on to many other parts of the world.

While attempting to introduce tea to India, British colonists discovered that tea plants with thicker leaves flourished in Assam as well and that these, when planted in India, thrived. The Singphos tribe of Assam had long grown the same plants, using tea chests provided by tribal chief Ningroola (Acharya, 1998).

2.2. Tea taxonomy

Tea, used as a popular beverage worldwide is distinguished as three taxa, namely *C. sinensis* (L.) O, Kuntze (“China type”), *C. assamica* (Masters) Wight (“Assam type”), and *C. assamica* ssp. *lasiocalyx* (“Southern form” or “Cambod type"). A careful observation, however, indicates an entirely different situation in the field, where one comes across a wide spectrum of tea populations, particularly in leaf form and other morphological characters of taxonomic significance. We have in our collection tea selections varying widely from the small-leaved “China type” to the large-leaved “Assam type”, with innumerable intermediates between these two extreme forms. Apart from variations in leaf size and form, one often recognizes certain other features in cultivated tea populations namely,

pigmented young shoots and leaves, punctuations on the leaf either present or absent, varying number of styles, nature of the stylar arms, glabrous ovary, or pubescent to different degrees, which are not readily visible in the typical forms of the three taxa mentioned above (Sharma and Venkataramani, 1974). The characters of *C. sinensis*, *C. assamica* and *C. assamica* ssp. *lasiocalyx* often overlap and also a few other characters unknown to these taxa appear in some populations of the cultivated tea. One is thus left perplexed with the taxonomy of the cultivated tea plant. Tea reproduces freely among themselves because of their high out-crossing tendencies, resulting in plant types that are midway between the two extreme forms, i.e., Assam type huge leaf and China type small leaf. As a result, tea taxonomic categorization based on morphology remains a muddle, and the species name is also a misnomer. Even though China and Assam teas are classified as distinct species, there is no limit on gene transfer between them (Rawal *et al.,* 2021). Desirable characteristics of Darjeeling tea, such as anthocyanin coloration or particular quality, may have been introduced from two wild species, *C. irrawadiensis* and *C. taliensis* (Wood and Barua, 1958).

2.3. Tea morphology

Camellia sinensis (Figure 2.1) is small and nearly 1-2 m tall with numerous stems and 2-inch leaves resistant to cold weather. The leaves are leathery, hard, and thick with a matt surface. Due to its thickness, it is very difficult to distinguish the veins present in the lamina. The leaf blade is elliptical with an obtuse end and square base with teeth present in the edges. The young leaves are smooth with hairs in the bottom, whereas the older leaves have few hairs or eventually it disappears. The Assam tea plant widely known as *Camellia assamica* (Figure 2.2) belongs to North East India and is considered small even though it has a height ranging between 10-15m. Warm weather environments are best suited for this plant. The branching system is strong with thick-thin, shiny, dependent, and pointed leaves, which are hairless or hairy on the vein below. The marginal veins are distinguishable. The leaf blade is oval with leaves 8-20 cm in length and 3.5-7.5 cm wide. *Camellia assamica* sub sp. *Lasiocalyx* (Figure 2.3) or southern form of tea is a small tree with upright branches between 6–10 m tall (Hilal, 2017).

Fig.2.1. *Camellia sinensis* var *sinensis* with small leaves

Fig.2.2. *Camellia sinensis* var *assamica* with large leaves

2.4. Importance of germplasm collection and characterization

Germplasm characterization is very much essential for plant breeders to effectively utilize and manage the gene banks to enhance the improvement of cultivars for higher crop productivity (Piyasundara *et al.,* 2006). The information gathered from the characterization of accessions will also be beneficial to understand genetic variation and select genetically distant relatives for successful hybridization (Anandappa, 1993). Morphological descriptors are the important preliminary characterization of the genetic resources and it is a cost-effective method when compared to the biochemical and molecular markers (Martinez *et al.,* 2003). It surely has many disadvantages under influence of environmental factors or parameters which results in continuous variations and increased plasticity. Despite such shortcomings or disadvantages, morphological traits are utilized in tea for its characterization and identification as

Fig.2.3.*Camellia assamica* subsp. *Lasiocalyx* with intermediate leaves

well as for the study of genetic diversity, and phylogeny (Ranatunga *et al.,* 2017). Piyasundara *et al.* (2006) characterized 20 accessions of *Camellia sinensis* using 13 standard morphological descriptors (Table 2.1). Out of these 13 morphological descriptors, a total of 11 descriptors contributed to the phenotypic variation where width, leaf shape, leaf, and petiole pigmentation mainly contributed to the variation. Likewise, 89 accessions from tea germplasm have been characterized employing 16 floral traits (Table 2.1) comprising of both qualitative and quantitative traits and the result identified considerable variations among the accessions with the pistil trait being highly variable (Ranatunga *et al.,* 2017).

The role of tea germplasm in crop improvement is well recognized but it still lacks adequate information so it cannot be utilized to its full potential. Sri Lanka has reported a total of 600 accessions but the narrow genetic choices for the breeding program have led to the utilization of only 4% of accessions (Ranatunga *et al.,* 2017). These accessions also lack adequate evaluation in the biochemical, agronomic, and molecular aspects (Gunasekare, 2007; Ranatunga and Gunesekera, 2008). Tea being an allogamous plant shows many overlapping, phenotypic and biochemical characteristics and the excessive hybridization may cease the existence of pure archetypes of tea (Banerjee, 1992). Study regarding the characterization of tea germplasm based on morphological descriptors not only helped in the study of variation but also put some light on the collection gap where the majority of the accessions collected were Cambod followed by Assam, showing more preference and bias towards these two types showing void or less preference in China type collection (Ranatunga *et al.,* 2017). Therefore, morphological characterization can generate valuable information regarding the genotypes that predominate in a particular region and thus help to fill the gap in the core collection of germplasm.

Table 2.1. Morphological descriptors used for characterizing tea accessions

(Piyasundara *et al.,* 2006)	(Ranatunga *et al.,* 2007)
Leaf length	**Qualitative traits:**
Leaf width	Ovary pubescence
Ratio of leaf length to leaf width	Stigma position
Leaf shape	Split pattern of style
Leaf pose	**Quantitative traits:**
Immature leaf color	Bud length
Mature leaf color	Bud width
Length of the mature leaf petiole	Petal length
Leaf pigmentation	Petal width
Petiole pigmentation	Total number of sepals
Appearance of the leaf upper surface	Total number of petals
Young shoot pubescent	Number of whorls of sepals
Leaf apex habit	Number of whorls of petals
	Number of style arms
	Style length
	Style column length
	Style arm lengths
	Corolla diameter

2.5. Genetic diversity of tea using molecular markers

The tea germplasm can be characterized using morphological, phytochemical, and molecular descriptors but among the former descriptors, namely the morphological and phytochemical is affected or influenced by environmental factors (Hyun *et al.,* 2020).

The morphological and geographical characters are also considered to be less beneficial when studying the relationships among the cultivars (Katoh *et al.,* 2003). The genetic diversity of tea also has been studied using amplified fragment length polymorph (AFLP) (Paul *et al.,* 1997) but the requirement of radioactive labeling and comparatively better need of quantitative and qualitative DNA limits the utilization of AFLP in every research laboratory. The RAPD technique has been successfully employed to determine the relationships between the species of *Camellia* (Wachira *et al.,* 1997) along with its successful utilization to differentiate between varieties and clones of *C. sinensis* (Wachira *et al.,* 1995). Additionally, we can develop locus-specific probes and other SSR sequences of interest by isolating and reamplifying individual bands (Wu *et al.,* 1994; Zietkiewicz *et al.,* 1994). Previous work on Darjeeling tea reports the study of genetic diversity using RAPD (Roy and Chakraborty, 2009; Baruah *et al.,*2010), ISSR (Roy

and Chakraborty, 2009), and AFLP (Mishra and Sen-Mandi, 2004) markers.

Genetic diversity study has its importance. Genetic diversity and taxonomic relationships were estimated in 38 clones (Wachira *et al.,* 1995) belonging to three tea varieties through RAPD analysis and reported genetic variability between species partitioned between and within-population with 70 percent detected in the latter. Band sharing analysis separated three populations according to present taxonomy and known pedigree of some clones and RAPD analyses also discriminated against 38 commercial clones including the clone which was morphologically and phenotypically indistinguishable (Wachira *et al.,* 1995). Genetic diversity study of 27 superior tea (*Camellia sinensis* var. *sinensis*) accessions from Korea, Japan, and Taiwan using RAPD markers led to the suggestion that the Taiwan tea studied may have a different origin from that of Korea and Japan and also reported 71 percent of variability residing within groups and rest between groups with the greatest variation within the Korean group followed by Taiwan and Japan (Kaundun *et al.,* 2000). Random amplified polymorphic DNA (RAPD) techniques were applied to assess genetic instability among micro propagated tea [*Camellia sinensis* (L.) O. Kuntze] cultivar 'T-78' and suggested RAPD be used successfully to determine the genetic instability among micropropagated plants which otherwise were morphologically indistinguishable (Nondal and Chand, 2002). RAPD markers provided a practical method not only to estimate the genetic diversity and relationship but also to identify tea genetic resources of the 15 well-known, widely planted traditional Chinese elite tea genetic resources [*Camellia sinensis* (L.) O. Kuntze] (Chen *et al.,* 2005). Inter simple sequence repeats (ISSR) technique was adopted to analyze the genetic variability of somatic embryo-derived tea plants and the study confirmed the existence of wide genetic variation among the somaclones (Thomas *et al.,* 2006). Genetic diversity and differentiation were examined in 10 ancient tea populations (*Camellia sinensis* var. *assamica*) by using ISSR markers and found a moderate level of genetic differentiation among the population (Ji *et al.,* 2011). The genetic fidelity of the micropropagated plants of three tea clones was assessed by analyzing their nuclear, mitochondrial (mt), and chloroplast (cp) genomes using multiple molecular DNA markers like RFLP, RAPD, SSR, and ISSR and revealed that genomic changes in tea

clones are genotype-dependent rather than culture condition-dependent (Devarumath *et al.,* 2002). RFLP analysis with phenylalanine ammonia-lyase *(PAL)*cDNA as a probe was used to evaluate the genetic diversity of the Korean tea plant *(Camellia sinensis* var. *sinensis*) and reported the genetic background of Korean teas, to differ from those of Japanese teas along with different morphological characters (Matsumoto *et al.,* 2004). Genomic fingerprinting in 21 tea genotypes was carried out using 7 ISSR and 12 RAPD primers with China variety showing the largest within-group diversity, followed by Assam tea and Cambod tea (Roy and Chakraborty, 2009). Recovery of standard DNA barcodes was tested for land plants from a large array of commercial tea products their performance was analyzed in identifying tea constituents using existing databases and thus indicated that unlisted ingredients are common in herbal teas (Stoeckle *et al.,* 2011)**.** Use of ITS2 barcode accurately and effectively to distinguish herbal tea ingredient *Plumeria rubra* from its adulterants was reported, which provided a new molecular method to identify *P. rubra* and ensure its safety in use (Shi *et al.,* 2014). Four candidate regions (*rbc*L, *mat*K, ITS2, *psb*A-*trn*H) were amplified by a polymerase chain reaction and DNA barcodes were used for the first time to discriminate the commercial non-*Camellia* tea and their adulterants, and to evaluate their safety and suggested non-limitation of the sequence of original plants in GenBank and thus requires submission of more original plant sequences to the GenBank for evaluating the safety of non-*Camellia* teas (Long *et al.,* 2014). A total of 113 tea plants [*Camellia sinensis* (L.)O. Kuntze] housed at the Tea Research and Extension Substation (Taiwan), was used in the experiment from which 113 internal transcribed spacer 2 (ITS2) fragments, 104 *trnL* intron, and 98 *trnL-trnF* intergenic sequence regions were successfully sequenced and found ITS2 nucleotide sequence variation larger compared to trnL intron and trnL intergenic sequence fragments of chloroplast cpDNA, thus suggesting it more suitable for establishing a DNA barcode database to identify tea plant varieties (Lee *et al.,* 2017).

2.6. Different types of tea

Sharangi (2009), elaborated on different types of tea consumed worldwide which are discussed briefly below. The different forms of tea are green tea, black tea, white tea, oolong tea, pu'erh tea, etc. The process of production of respective teas is provided in Table 2.2.

The consumption of tea imparts many

potential health benefits. The health benefits of a different form of tea are enlisted in Table 2.3. Green tea is loaded with beneficial chemicals with maximum health benefits (Sinija and Mishra, 2008). Due to no fermentation, it has more polyphenols and less caffeine which imparts various positive health effects. White tea on the other hand retains high levels of antioxidants than other types of tea due to its unique process of selection and production. Black tea accounts for the largest production worldwide i.e., 72% of total tea production (Sharangi, 2009). Black tea requires longer fermentation so most of the antioxidants mostly the EGCG antioxidants are oxidized although it retains a higher number of the antioxidant compounds especially the flavonoids which helps in the elimination of various harmful toxins from the body. The partially fermented tea i.e., the oolong tea has the flavor and health attributes of both black tea and green tea. It is rich in antioxidants beneficial for skin cells and slowing down the aging process. Another form of tea i.e., the Pu'erh tea is unique in its way due to its aged consumption since the large tea leaves once plucked are piled and stored for 50-100 years. The consumption of green tea and oolong tea is mainly prominent in Asian countries whereas black tea is popular in Western countries (Sharangi, 2009). The polyphenol content decreases and the caffeine level increases with the increase in fermentation process, the lower the polyphenol content, and the higher the caffeine content (Lin *et al.,* 2003).

2.7. Tea phytochemistry

The chemical composition of tea is highly complex, owing to the abundant presence of different classes of chemical compounds. The primary chemical compounds found in tea are given in Figure 2.4. and the structural formula of some main components is provided in Figure 2.5. The main class

Table 2.2. Different types of tea and its production (Sharangi , 2009)

Types of tea	Process of production
Green tea	Prepared from unfermented leaves
White tea	Buds (before fully opened) and young tea leaves collected followed by steaming and drying of leaves. Minimum amount of processing.
Black tea	Longer fermentation
Oolong tea	Partial fermentation
Pu'erh tea	Proceesing similar to black tea. The only difference is that large leaf variety of tea plant is used which can be picked throughout the year.

of chemical compounds is described below.

2.7.1. Polyphenols

Tea is regarded as the most popular dietary source of polyphenols that accounts for 50–70% of tea water extracts (Yao *et al.,* 2006). Polyphenols provide astringency, the 'drying' sensation that we experience in the mouth after consumption of the tea beverage. A strong cup of tea usually is loaded with 180 to 240 mg of polyphenols.

The polyphenols are flavonoids that are produced biosynthetically in considerable amounts ranging from 0.5% to 1.5% (Vinson *et al.,* 1995). The major abundant flavonoids comprise catechins (flavan-3-ols), (Millin *et al.,*1969) and: (-)-epigallocatechin-3-gallate (EGCG), epigallocatechin (EGC), epicatechin-3-gallate (ECG), and epicatechin (EC) are the four major catechins predominantly found in green tea. These different structures are very important for their pharmacological usage (Zhao *et al.,* 2001). Tannins are the second major polyphenols present

Table 2.3. Different types of tea and is health benefits

Tea type	Health benefits	Reference
Black tea	Antioxidant and anti-inflammatory, Cardiovascular diseases, Diabetes, GIT cancer (Esophageal, gastric, pancreatic, colon, & colorectal cancer), Other cancers (Prostate, lung, breast, liver & skin cancer), Mental alertness/ memory improvement, Bone health, Gastric emptying, Obesity, Rheumatoid arthritis, Neurological effects, Viral infection, Bacterial infection, Asthma & allergy, Mood & cognitive performance, Dental health.	(Naveed *et al.,* 2018)
Green tea	Antioxidant and hepatoprotective activity, anticancer and anti-mutagenic activity, antimicrobial and antiviral activity, anti-schisosomiasis and antiparasitic activity, cardioprotective activity, antidiabetic and anti-obesity activity, gastrointestinal tract problems relieving activity, neuroprotective activity, anti-inflammatory, analgesic, antipyretic,and anti-allergic activity, skeletomuscular system relieving activity.	(Aboulwafa *et al.,*2019)
Oolong tea	Anti-inflammation,antioxidant activity, anti-obesity, anti-cancer, hypoglycemic effect, prevention of heart disease, antimicrobial activity (*Bacillus subtilis*, *Escherichia coli*, *Proteus vulgaris*, *Pseudomonas fluorescens*, Salmonella sp. and *Staphylococcus aureus)*	(Weerawatanakorn*e t al.*, 2015)
White tea	Cardioprotective effects, antidiabetic potential, anticarcinogenic and antimutagenic activities, neuroprotective activity, antimicrobial properties, Anti-obesity potential	(Dias, 2013)

in tea products, which are responsible for astringency (Okuda *et al.,*1985). Additionally, the phenolic acids consist of chlorogenic acid, caffeic acid, gallic acid, coumaric acid, and quinic acid ester. The flavanols comprises of kaempferol, myricetin**,**and quercetin (Aboulwafa *et al.,* 2019).

2.7.2. Purine Alkaloids

The second major constituent of tea is caffeine which belongs to the xanthine bases or the purine alkaloids. Other metabolites like theophylline and bromine, are found in smaller amounts.

2.7.3. Triterpenoid Saponins

The triterpenoid saponins comprise floratheasaponin A-F (Yoshikawa *et al.,* 2008), which are predominantly found in higher concentrations in seeds and flowers (Man *et al.,* 2010).

2.7.4. Amino Acids

Arginine, aspartic acid, glutamic acid, glutamine, and serine**,** as well as theanine, accounts for more than 90% of the whole amino acids (1–4% of dry weight) found in the leaves of *C. sinensis* (Horie *et al.,* 1993). Other amino acids like Tryptophan, valine,

Table 2.4. Primary chemical compounds of tea characterized by chromatic methods (Yashin *et al.,* 2015)

Sl No.	Compound name	Main representative elements
1.	Catechins	Flavanols: 12 catechins are indentified, including 8 occurring in significant quantity, i.e., (+)-catechin, (-)- epicatechin, (-)-gallocatechin, (-)-epigallocatechin, (-)-catechin gallate, (-)-epicatechin gallate, (-)-gallocatechin gallate, (-)-epigallocatechin gallate
2.	Oxyaromatic acids	Gallic, caffeic, quinine, chlorogenic, n-coumaric acids
3.	Flavonols	Quercetin, kaempferol, myricetin
4.	Theaflavins	Theaflavin, theaflavin-3-O-gallate, theaflavin-3'-O-gallate, theaflavin-3-3'-O-gallate
5.	Teagallins	Teagallin
6.	Thearubigins	High-molecular weight polymers of catechin gallates with molecular weight from 1000 to 40000 Da
7.	Pigments	Carotenoids and chlorophyll
8.	Alkaloids	Caffeine, theophylline, theobromine
9.	Sugars	Glucose, fructose, saccharose
10.	Amino acids	Isoleucine, leucine, methionine, threonine, phenylalanine,glutamine, asparagine, alanine, serine, proline, histidine, glutamic acid, aspartic acid, theanine
11.	Vitamins	C, α-, β-, γ-, δ-tocopherols, riboflavin
12.	Dibasic acids	Succinic, malic, tartaric, citric, quinic, aspartic, glutamic, oxalic acids
13.	Cations	K+, Na+,Ca2+, Mg2+, NH4 +, Al3+
14.	Metals	Fe, Zn, Cu, Ni, Al
15.	Lignans and triterpenoid saponins	Mixture of many compounds

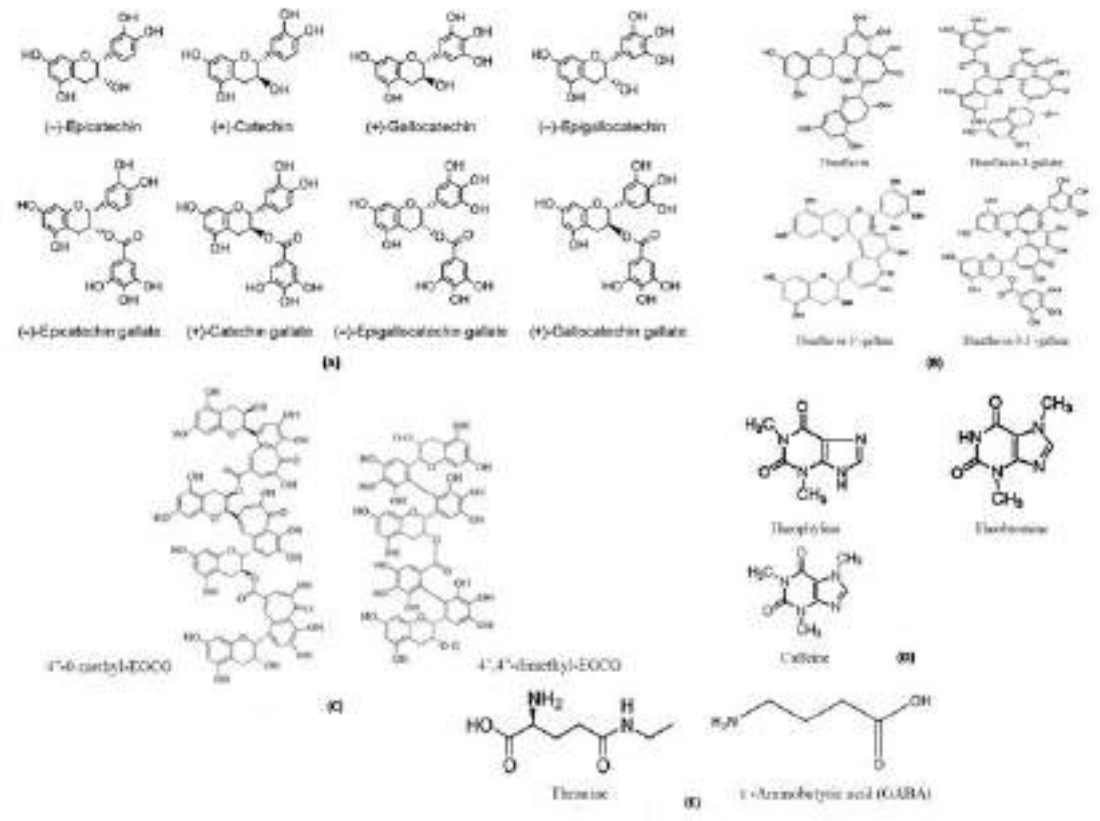


Fig.2.4. Structure of primary chemical compounds of tea (A) Eight major catechins found in tea (B) Tea theoflavins (C) Oligomeric catechins (D) Tea methylxanthins (E) Tea theanine

threonine, lysine, tyrosine, leucine, and lysine are also found. The amino acid attributing to the flavor and exotic taste of green tea i.e., Theanine (5-Nethylglutamine) is the only amino acid existing solely in tea plants and is the major amino acid found in largest amounts in green tea i.e., about 1–2% of the dry weight of leaf (Aboulwafa *et al.,* 2019). The presence of this rare amino acid is also reported (Casimir *et al.,* 1960) in the *Xerocomusbadius* (edible bay boletes mushroom**).**

2.7.5. Other constituents

Tea comprises minerals and traces elements like, Na, Cr, Al, Ca, K, Co, Cu, F, Fe, Mg, Se, Sr, Mo, Ni, P, and Zn, which comprises about 5% of dry weight (Cabrera *et al.,* 2006). About 15 –20% of dry weight is represented by proteins, whose enzymes comprise an important fraction (e.g., carotenoid cleavage enzymes), and 5–7% of dry weight is represented by carbohydrates like sucrose, glucose, fructose, and pectin. The most found lipids are linoleic and α-linolenic acids. Sterols such as stigma sterols are also found in tea. Other important compounds found in tea are vitamins (B, C, and E), chlorophyll pigments (carotenoids),

volatile compounds mainly related to tea aroma (aldehydes, alcohols, long-chain aliphatic alcohols). Green tea is enriched with the bioactive compound policosanol, long-chain aliphatic alcohol which has much health-promoting activity (Choi *et al.,* 2016).

2.8. Health benefits of tea

2.8.1. Antioxidant

Tea is enriched with polyphenols like catechins, theaflavins, and thearubigins which are high in antioxidants and beneficial for health. The antioxidant activity of tea polyphenols is due to their ability to scavenge superoxide along with the increased or enhanced activity of some detoxifying enzymes like catalase, glutathione-S-transferase, glutathione peroxidase, quinine reductase, and glutathione reductase present in the small intestine, liver, and lungs. The free radical scavenging activity of EGCG has been confirmed by previous *in vitro* and *in vivo* studies. Among all the antioxidants like vitamins C and E (tocopherol) and carotene, tea catechins are better. The antioxidant properties of tea are very much crucial for human health which aids in the prevention of atherosclerosis and specifically some coronary artery disease (Miura *et al.,* 2001).

2.8.2. Anticancer

Tea contains various antioxidants and phenolic compounds, where some compounds have shown anti-cancer properties in laboratory conditions as stated in earlier reports (Kris-Etherton *et al.,* 2002). Consumption of green tea decreases the frequency of cancer development and imparts reduced risk of several human malignancies and diabetes. The molecular mechanisms of green tea polyphenols and their therapeutic use in cancer are studied previously by Shankar *et al.* (2007) and stated epigallocatechin- 3-gallate (EGCG), a major polyphenol found in green tea, to be widely studied for its chemopreventive and potential anticancer activity. A wide range of mechanisms is outlined by which epigallocatechin gallate (ECGC) and other green and black tea polyphenols inhibit cancer cell survival as outlined previously by Beltz *et al.* (2006). The polyphenols of tea especially found in green tea inhibit angiogenesis and metastasis which arrest the growth and induces apoptosis celebration of various signaling pathways. Some form of skin tumor formation which is cost by different chemical carcinogens or exposure to UV radiation is inhibited by oral or local application of green tea extracts.

This research compelled many other cosmetic and pharmaceutical companies to supplement their skincare products with green tea extracts (Katiyar and Elmets, 2001). Million cases of skin cancer are diagnosed annually across the globe where most of the cases are of non-melanoma associated with chronic exposure to ultraviolet light. The non-melanoma form of skin cancer is very common and is increasing in occurrence therefore the potential of green tea polyphenols, epigallocatechin gallate as a chemopreventive agent for non-melanoma skin cancer was examined where a double-blind randomized and placebo-controlled phase II clinical trial of epigallocatechin gallate in the prevention of non-melanoma skin cancer was conducted (Linden *et al.,* 2003).

2.8.3. Cardiovascular diseases

Hypercholesterolaemia, a condition when the cholesterol level increases in the circulating blood where the concentration of cholesterol rises above 200 mg/ dL and the level of the high-density cholesterol are below 35 mg/ dL. This condition of high cholesterol level is associated with increased risk of cardiovascular disease and the causative factors for this condition are stress or high intake of refined sugar and animal fats. So along with lifestyle changes, some dietary changes need to be made by replacing animal fats with industrially processed vegetable oils along with increased consumption of green tea which is beneficial for the cardiovascular system which ultimately lowers the cholesterol level and prevents clumping of the platelet.

Coronary artery disease is a result of increased oxidative stress and dysfunction of endothelium and black tea antioxidants can reverse the endothelial dysfunction (Duffy *et al.,* 2001). Prior clinical trial suggests a total of five servings of black tea per day which could help in reducing low-density lipoprotein by 11.1% and total cholesterol by 6.5% in moderately hypercholesterolemic adults (Davies *et al.,* 2003).

Tea is believed to have the potential ability to limit cholesterol absorption in the intestine. Therefore, a study was conducted where the effect of tea on blood lipids was studied keeping all other components of the diet constant. The continuous ingestion of green tea extract which is high in catechin aided in the reduction of body fat, SBP, and LDL cholesterol. The study emphasized the role of such catechin enriched extract to decrease obesity and risk of cardiovascular disease (Nagao *et al.,* 2007). Similarly, the other study also emphasized the effects

of a beverage rich in catechin to reduce the cardiovascular disease risk in obese children and thus concluded that consumption of tea enriched with catechin decreases obesity and related risk factors of cardiovascular disease without having any side effects (Matsuyama *et al.,* 2008).

2.8.4. Respiratory disorders

Some respiratory diseases like wheezing, shortness of breath, and difficulty breathing arise as a result of some health conditions like chronic bronchitis, emphysema, asthma, and other lung diseases. Theophylline in tea is used to prevent such alleviating symptoms where it relaxes the lungs and opens air passages thus making it easier to breathe (Sharangi, 2008).

2.8.5. Skin Problems

Tea is used as an age-old home remedy for burns, wounds, and swelling. Green tea extract reduces inflammation, stops or slower bleeding, relieves itchy rashes caused by insect bites.

Tannins present in tea have both antiseptic and anti-inflammatory properties. The important antioxidant compound like flavonoids also possesses antiseptic properties. Green tea extract is considered by researchers as a natural sunblock agent. At certain concentrations, the EGCG present in green tea and the other green tea polyphenols showed the possibility for renewed cell division (Hsu *et al.,* 2003).

2.8.6. Digestion

Tea loaded with phytochemicals having bioactivity like antibacterial, antioxidant, antiseptic, and detoxifying properties are very much effective in treating infectious dysentery as well as easing inflammatory bowel disease. Green tea is very much useful as a traditional home remedy for various digestive problems like inflammatory bowel disease (Rahman *et al.,* 2018).

2.8.7. Anti-diabetic

Studies on animal models revealed that green tea may slow or prevent the progression of Type 1 diabetes. The role of polyphenolic components of tea to inhibit a-amylase from human saliva was investigated using *in vitro* studies by Hara and Honda, (1990). The polyphenols present in tea lower the serum glucose by inhibiting the activity of both salivary and intestinal amylase (starch digesting enzyme). This results in the slow breakdown of starch and minimization in the sudden rise of serum glucose.

2.8.8. Arthritis

Tea especially the green tea has anti-inflammatory properties where the compound like EGCG inhibits the production of several molecules in the

immune system that leads to inflammation and damage of joints in arthritis patients (http:// www.nerve.in/ news, 25350069280). The green tea antioxidant compounds will thus help to prevent developing arthritis and consumption of at least four cups of green tea can reduce inflammation in patients with arthritis.

2.8.9. Anti-obesity

A metabolic disorder resulting in obesity is caused by an imbalance between intake and expenditure of energy thus putting a person at stake for lifestyle-related disease. Consumption of a tea rich in catechins along with regular exercise helps to activate whole-body energy metabolism and thus reduce obesity induced by diet (Murase *et al.,* 2006).

2.8.10. Cognitive functions with tea consumption

Tea loaded with compounds of pharmaceutical importance protects the brain corresponding to the aging process and is also inversely associated with the occurrence of dementia, Alzheimer's, and Parkinson's diseases. The main catechin polyphenols compound i.e., the EGCG, has shown neuroprotective properties in a wide range of cellular as well as animal models with neurological disorders (Mandel *et al.,* 2008). A cognitive test examined by studying the relation between intake of flavonoid-rich foods and its effect on cognitive function led to the findings of enhanced cognitive abilities in the elderly (Nurk *et al.,* 2009).

2.9. Solvent extraction and bioactivity

Polar solvents like water, ethanol, and acetone were found to extract major phytochemicals groups than non-polar ethyl acetate and chloroform from *Camellia sinensis* and *Psidium guajava* leaves (Geoffrey *et al.,* 2014). Extraction method should ensure complete extraction of the desired compounds of interest without any chemical modification (Zuo *et al.,* 2002). Extraction and determination of biologically active compounds depend upon the type of solvent used where solvents will diffuse into solid plant tissue and solubilize compounds with the same polarity (Tiwari *et al.,* 2011).Aqueous mixtures of ethanol, methanol, and acetone, water are commonly exploited to extract plants (Sun and Ho, 2005). Researchers have reported the use of aqueous methanol, acetone and ethanol, absolute methanol, absolute ethanol, and boiling water for the extraction of polyphenols from green, black and other types of manufactured tea (Labar *et al.,* 2019). Different solvents like water; aqueous ethanol in different extracting time has

been employed to extract phenolics from green and white tea (Rusak *et al.,* 2008). Tea components have been analyzed using high-performance liquid chromatography and high-performance capillary electrophoresis (Horie and Kohata, 2000). A comparative analysis of tea catechins and theaflavins has been performed using high-performance liquid chromatography and capillary electrophoresis (Lee and Ong, 2000). Metabolite Profiling and Quality Assessment of Green Tea or *Camellia sinensis* (L.) was done using ^{1}H NMR Spectroscopy and reported the presence of theanine, gallic acid, caffeine, epigallocatechin gallate, and epicatechin gallate in higher levels than epigallocatechin in Longjing teas compared to other teas (Le Gall *et al.,* 2004). Phenolic ingredients were separated by multilayer countercurrent chromatography (MLCCC) and preparative high-performance liquid chromatography (HPLC) was then applied to obtain pure flavonoids and the purity and identity of isolated compounds was confirmed by different NMR experiments, HPLC-diode array detector (DAD), or gas chromatography−mass spectrometry (GC-MS) analysis (Krafczyk and Glomb, 2008).

Studies have been performed to check the antimicrobial properties of polyphenol fractions of tea and also shown purified catechin fractions from green and black tea, and especially ECG and EGCG inhibited growth of many bacterial species and possessed anticarcinogenic properties, proved to be powerful antagonists of human immunodeficiency virus reverse transcriptase, causing 50% inhibition at a concentration ranging from 10 to 20 ng/ml (Hamilton-Miller, 1995). The flavonols, quercetin, kaempferol, and myricetin showed activity against gram -positive bacteria and phytopathogenic fungi (El-Gammal and Mansour, 1986).

2.10. Tea in nanotechnology

Nanostructured noble metals are widely used in a variety of technological applications and a variety of synthetic approaches have been used to suit these demands but nowadays green chemistry concepts have recently sparked renewed interest in the production of noble metal nanoparticles (Moulton *et al.,* 2010). The development of biological nanoparticle synthesis utilizing microorganisms or plant extracts is essential in nanotechnology since it is not harmful to the environment, unlike chemical synthesis which releases toxic chemicals (Loo *et al.,* 2012). Green synthesis methods, which use biological microbes or plant extracts

instead of chemicals, have evolved as a simple and effective alternative to chemical synthesis. Green synthesis outperforms chemical approaches in terms of environmental friendliness, cost-effectiveness, and ease of scaling up for large-scale synthesis. Plant extracts can be used to synthesize nanoparticles, which has the advantage over other biological procedures in that it eliminates the time-consuming process of maintaining cell cultures and aseptic environments and can be scaled up for large-scale production (Loo *et al.,* 2012). There are three main steps during the synthesis of metal nanoparticles, (i) solvent selection, (ii) reducing agent selection, and (iii) capping agent selection (Moulton *et al.,* 2010).

There has been a growing interest in finding environmentally friendly, multifunctional materials for the green synthesis of metal nanoparticles. Tea contains polyphenols, which act as both reducing and capping agents during the synthesis of nanoparticles (Huang *et al.,* 2014). Catechins, a flavanol category of polyphenols found in fresh tea leaves, are extremely abundant (approximately 30 percent of the dry leaf weight). Flavonoids and their glycosides, chlorogenic acid, gallic acid, coumarylquinic acid, and theogallin are among the other polyphenols found. Phenolic compounds have a higher antioxidant potential. Hence, they are excellent reducers of metal ions which highly favor the green synthesis of nanoparticles (Senthilkumar and Sivakumar, 2014).

The simple tea leaf broth, as well as the one containing the ethyl acetate extract of tea leaves, was shown to be highly effective in forming stable Au and Ag nanoparticles quickly (Begum *et al.,* 2009). Three different tea extracts, including green, oolong, and black teas, were used to make iron nanoparticles (Fe NPs) among which FeNPs synthesized using green tea extracts was reported to be the best method for decomposing malachite green (MG). This was due to a higher concentration of caffeine/polyphenols in green tea that function as both reducing and capping agents during the formation of Fe NPs (Huang *et al.,* 2014).

The green synthesis of nanoparticles of the noble metals gold (Au) and silver (Ag) and their antimicrobial applications have been studied extensively. However, very few works have been reported on other metal oxide nanoparticles like FeO_2, TiO_2, CuO, AlO, MgO, ZnO (Senthilkumar and Sivakumar, 2014). Silver nanoparticles (AgNPs) are widely used in biomedical sectors because of their

powerful antibacterial effect. The biogenic nanoparticles antibacterial activity was demonstrated against *Staphylococcus aureus*, *Escherichia coli*, *Klebsiella pneumoniae*, *Salmonella enterica,* and *Pseudomonas aeruginosa* out of which *Salmonella* enterica was found to be most sensitive against the synthesized AgNPs *(Rolim et al., 2019).*

Because of its antibacterial, UV-blocking, high catalytic, and photochemical capabilities, zinc oxide (ZnO) nanoparticles have gotten a lot of attention (Meruvu *et al.,* 2011). According to Sharma *et al.* (2010), ZnO nanoparticles (ZnO NPs) have antibacterial and antifungal properties even at low concentrations, making them appropriate for thin coating applications. Raghupathi *et al.* (2011) examined the antibacterial action of ZnO NPs and developed antibacterial agents to suppress bacterial infections against a wide spectrum of pathogens. In comparison to manufactured medicines, the antibacterial activity of ZnO NPs synthesized using green tea was better and comparable (Senthilkumar and Sivakumar, 2014).

2.11. Tea in Bioinformatics

Comprehensive codon and amino acid usage profiling have been executed proficiently with the accessibility of a huge number of plant genomes. In plants, an overall length of the coding sequences, level of gene expression, etc., have been reported to be decisive in governing codon usage (Sablok *et al.,* 2011). In *Arabidopsis thaliana* the correlation between gene expression and codon usage signified a distinct impact of tissue-specific gene expressivity to be an imperative factor influencing codon usage bias (Duret and Mouchiroud, 1999). Analyzing expressed sequence tags of *Citrus* species, translational selection, mutational bias, and gene length were found to influence codon usage variation and it was also concluded that positive selection governs codon usage (Xu *et al.,* 2013). Sablok and colleagues (Sablok *et al.,* 2011) in their study of complete chloroplast genomes of species of pooid grass family reported mutational bias to play a major role in codon biology and also showed hydrophobicity and aromaticity of encoded proteins of each species to be other factors of codon bias. Examination of 207 plant genes subdivided into 53 monocot and 154 dicot genes, suggested codon usage to vary between taxonomic groups mainly due to varying use of G+C content at the degenerate third base (Murray *et al.,* 1989).

In a previous study (Zhao *et al.,* 2011), the coding sequences of 134 proteins from the tea plant (*C. sinensis*) were

analyzed using the CodonW and CUSP programs, and the frequency of codon usage that encoded amino acids was calculated and compared to four classes of representative organisms: *Homo sapiens* (human), *Saccharomyces cerevisiae* (yeast), *Drosophila melanogaster* (fruit fly) and *Escherichia coli* (bacteria). The protein-coding sequences of cp genes were investigated (Yengkhom *et al.,* 2019) using bioinformatic techniques to identify the patterns of codon use among the cp genes of three tea groups i.e., *Camellia sinensis* var. *sinensis* (Chinese type), Camellia *sinensis* var. *assamica* (Assam type), *and Camellia pubicosta* (wild tea species).■

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			

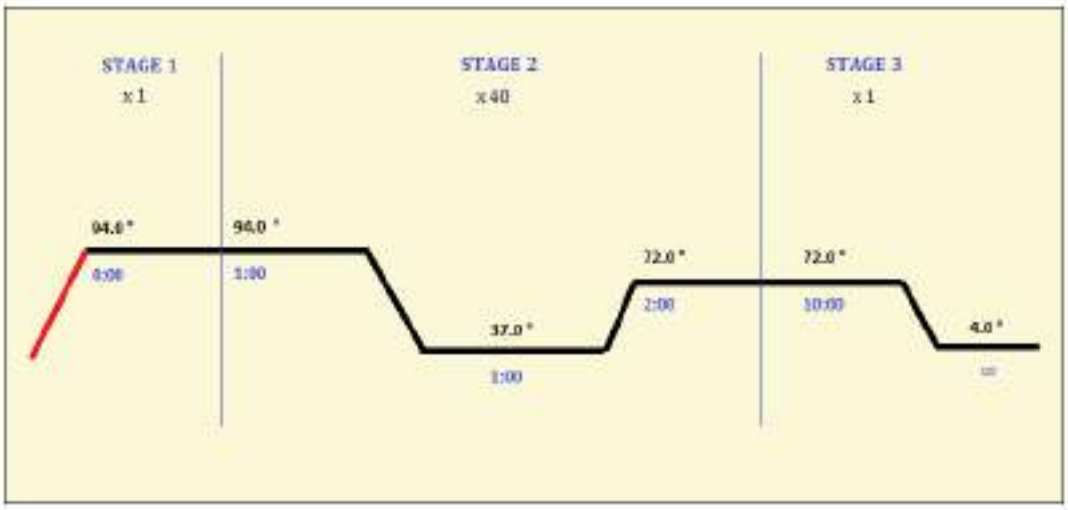


Fig.3.2. PCR cycle of RAPD

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

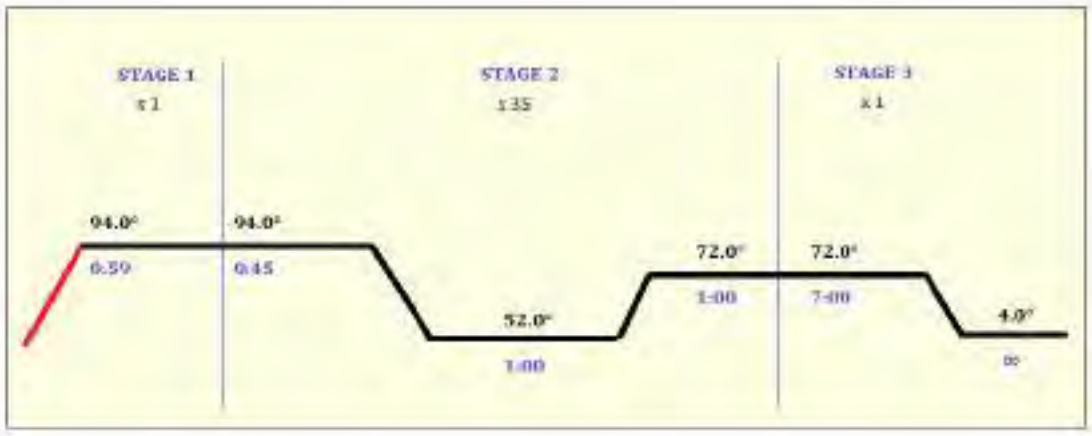


Fig.3.3. PCR cycle of ISSR

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).

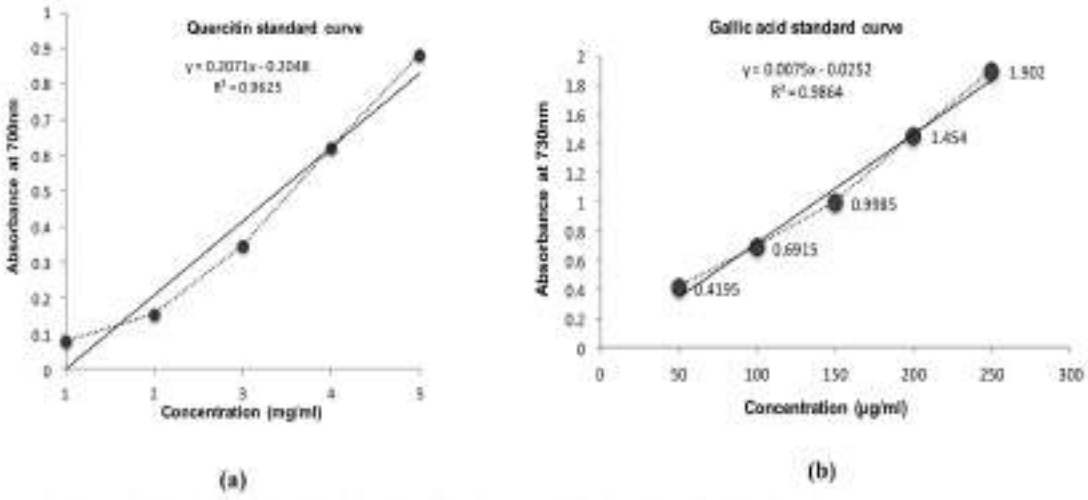
The total flavonoids were thus expressed as mg of quercetin equivalent per g of extract i.e., mg QE/ g. Data were expressed as a mean of triplicates ± standard deviation (Blois, 1958).

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.■

Chapter 4
Results & Discussion

If I have seen further than others, it is by standing upon the shoulder of giants.

-Isaac Newton

4.1. Morphological documentation

The tea clones or planting material for northern regions has been certified by Tocklai Research Association (TRA) based on standard, yield, quality, suitability, etc. Some of the TRA certified clones for Darjeeling based on standard are AV2, B 7/1A/76, B 9/3/76, BB157, Lingia 12, TV 14, T78, T135, T383, P1258, RR17/14, SKM 1, TV 1, Thurbo 3, Thurbo 9, Badamtam 15/263, P1404, T246. The TRA-approved Darjeeling clones with high yield are Happy Valley 39, RR4/5, TV 19, CP1, K1/1, Sundaram, B/5/63, T 253, and the quality clones are B688, B777, and T145.

Fig.4.1. Leaves of *Camellia sinensis* showing alternate and serrate morphology

The clones of Darjeeling and Dooars chosen for our study comprised of all the three main types of tea i.e the China type (small-sized leaf), Assam type (large-sized leaf), and the Cambod type (intermediated sized leaf). Generally, tea (*Camellia sinensis*) is an evergreen shrub or tree with simple, alternate, and serrate leaf (Figure 4.1). It is a diploid plant (2n=30) with 15 chromosomes. It is usually cross-pollinated since it is a self-incompatible plant. It has bisexual and regular flowers usually with five petals and sepals. It has multiple two-celled stamens with anthers. The ovary is superior, 2-4 locular, axile, and rarely solitary. It bears a capsule-like fruit with recalcitrant seeds as shown in Figure 4.2 (the seeds which losses viability on drying).

Tea or *Camellia sinensis* usually share more or less a common morphology among the different types and only differs in leaf morphology (size, shape, pubescence, etc), shoot (large, small, or medium). The leaf and shoot morphology of the studied tea clones is provided in Table 4.1. Different parameters like shoot size, pubescence, yield, flavor, drought resistance, suitability, resistance against pests, etc. were evaluated for the clones under investigation and the detailed information is given in Table 4.1. Among the 33 studied tea clones, the standard clones included clones like T 78, T 246, T 383, AV2, P312, RR 17/144, P 1404, P 1258, TTV 1, SKM 1, BT 15/263, BS 7/1A/ 76, BS 9/3/76, TS 378. The clones with high yield included T 253, RR 4/5, CP 1, HV 39, Sundaram, TV 19, and the quality clones included clones like T 145, K 1/1, B 777, and B 688 (Singh, 2005).

Fig.4.2. Capsule like fruit of *C. sinensis* with recalcitrant seeds

Clones like AV2, B 668, T 383, and K 1/1 comprised of the clones with very good flavor, and the rest are categorized into average (T 145, T 253, T 246, B 777, RR 4/5, CP 1, P1404, P 1258, Sundaram, HV 39) or good flavor (B 157, T 78, T 383, RR 17/144, TV 19, TS 378, TS379, B 777, P 312, SKM 1, BT 15/ 263, BS 7/1A/76, BS 9/3/76). Likewise, the clones are categorized as per their yield into clones with very high yield (RR 17/144, CP 1, P1404, Sundaram, HV 39, TV 19), high yield (P 312, AV 2, T 253, RR 4/5, T 78, T 383, P 1258, TTV 1, BT 15/263, BS 7/1A/76, BS 9/3/76, SKM 1, TS 379), average yield (T 145, B 777, B 688, K1/1, TS 378) and

Table 4.1. Data collected for clones under investigation considering various parameters (Singh, 2005)

	CATEGORY
Standard	T 78, T 246, T383, AV2, P 312, RR 17/144, P 1404, P 1258, TTV 1, SKM 1, BT 15/263, BS 7/1A/76, BS 9/3/76, TB 3, TB 9, TB 3, TS 378, TS 379, T 157
Yield	T 253, RR 4/5, CP 1, HV 39, Sundaram, TV 19
Quality	T 145, K 1/1, B777, B688
	SHOOT SIZE
Small	RR 4/5,TB 3, TS 378
Medium	T 78, AV2, P312, RR 17/144, P1258, TTV 1, SKM 1, BT 15/263, BS 9/3/76,TB 9,T 145, B 777, B 688, K1/1
Large	T 253, T 246, T 383, CP 1, P 1404, Sundaram, BS 7/1A/76, TV 19
	PUBESCENCE
Low	B 157, T 145, TV 19, Sundaram
High	P 312, AV2, RR 4/5, T 78, T 383, RR 17/144, P 1404, K 1/1, TTV 1, HV 39, BS 7/1A/76, BT 15/263,TB 3, TB 9, TS 378, P 1258, BS 9/3/76
Very high	T 253, T 246, B 777, B 688, SKM 1, CP 1
	YIELD
Average	T 145, B 777, B 688, K 1/1, TS 378
Above average	B 157, T 246, TB 3
High	P 312, AV 2, T 253, RR 4/5, T 78, T 383, P 1258, TTV 1, BT 15/263, BS 7/1A/76, BS 9/3/76, SKM 1, TB 9
Very high	RR 17/144, CP 1, P 1404, Sundaram, HV 39, TV 19
	FLAVOUR
Average	T 145, T 253, T 246, B 777, RR 4/5, CP 1, P 1404, P 1258, Sundaram, HV 39
Good	B 157, T 78, T 383, RR 17/144, TV 19, TS 378, B 777, P 312, SKM 1, BT 15/263, BS 7/1A/ 76, BS 9/3/76, TB 3, TB 9
Very good	AV2, B 688, T 383, K 1/1
	DROUGHT RESISTANCE
Poor	T 246
Fair	B 777, K 1/1, TB 9
Good	P 312, B 157, T 145, AV 2, T 253, RR 4/5, B 688, T 78, T 383, P 1404, P 1258, TTV 1, Sundaram, HV 39, SKM 1, BT 15/263, TV 19, BS 7/1A/76, BS 9/3/76, TB 3
Very good	RR 17/144, CP 1, TS 378
	SUITABILITY
Sub-marginal/ droughty conditions	P 312, B 157, T 145, T 253, T 78, T 383, AV 2, RR 4/5, P 1258, TTV 1, TS 378, BS 7/1A/76, BS 9/3/76, TB 3, TB 9
Infilling and Interplanting	T 78, RR 4/5, T 253, CP 1, HV 39, P 1404, Sundaram, SKM 1, BT 15/263, TS 378, TV 19

above-average yield (B 157, T 246). Some abiotic stress like drought were also considered for evaluation and the clones are categorized following their ability to resist drought with clones having poor (T 246), fair (B 777, K1/1), good (P312, B 157, T 145, AV2, T 253, RR4/5, B 688, T 78, T 383, P 1404, P 1258, TTV 1, Sundaram, HV 39, SKM 1, BT 15/263,

Table 4.2. Data collected for resistance or susceptibility against different diseases and pest (Singh, 2005)

RESISTANT TO MITE	AV2, B/5/63***, Sundaram, TS-378***, HV39***, TV 19
RESISTANT TO BLISTER BLIGHT	AV2, P312****, TS 378*, T 383, TS 378***, Nandadevi, TB 9***, T 145***, T 78***
RESISTANT TO RED SPIDER	HV 39*, T 383, B/5/63***, T 246**, B 157****
SUSCEPTIBLE TO BLISTER BLIGHT	HV 39****, T 253, TTV 1, K1/1, B 15/263, B5/1A/76*, B 777, T 135, BS 9/3/76**, CP 1, P 1404*, P 1258**, RR 4/5**, B 157
SUSCEPTIBLE TO RED SPIDER	T 253, TTV 1, B 15/263, B5/1A/76*, BS 9/3/76**, P 1258**, RR 4/5**, TB 3, TB 9, T 78
SUSCEPTIBLE TO MITES	K 1/1, B5/1A/76*, B 777, BS 9/3/76**, P 1404*, P 1258**, RR 4/5**

*Slightly resistant or susceptible; **Moderately resistant or susceptible; *** Fairly resistant or susceptible; ****Highly resistant or susceptible

TV19, BS 7/1A/76, BS 9/3/76)) and very good (RR 17/144, CP 1, TS 378) resistance towards drought.

The studied clones also exhibit resistance against various pests and diseases (Table 4.2.). The tea clones, resistant to mites include AV2, B/5/63, Sundaram, TS 38, HV 39, TV 19 and, the clones resistant to red spiders are HV39, T383, B/5/63, B668, T 246, B 157. Some clones are also resistant to the havoc-causing blister blight disease. The resistant clones towards the blister blight are AV2, P312, TS 378, T 383, Thurbo 9, T 145, T78. However, some clones are susceptible to the mite, red spider, and blister blight, the details of which are provided in Table 4.2. Morphological descriptors do not give us a complete idea about the genetic variation or diversity due to excess cross-breeding. However, prior knowledge about the phenotype is crucial for further breeding programs.

4.2. Molecular Documentation

Crude DNA was isolated from 33 tea samples of clones using the protocol given by Doyle and Doyle, (1987). After the successful CTAB extraction of DNA, agarose gel electrophoresis was performed to visualize the bands and any form of impurities. Further, protein and other contaminants were removed employing the isolated DNA for downstream purification. The Agarose gel electrophoresis revealed distinct and clear bands. The purity or quality of DNA was further checked using DNA quantifying methods where

the DNA was quantified using UV spectrophotometer at 260 nm and 280 nm filters. The ratio of 260/280 equivalent to 1.8 was considered of good quality. DNA of good quality will have an A_{260} /A_{280} ratio of 1.7–2. However, a value of 1.6 does not render the unsuitability of DNA for further analysis or application, but a lower ratio below 1.7 indicates the presence of more contaminants in the DNA. Hence, combining the three steps mentioned above i.e., DNA isolation, DNA purification, and DNA quantification allowed a good amount of pure DNA from the leaves of different clones of tea (*Camellia sinensis*) that can be utilized for further PCR amplification.

Tea is an out-crossing species, where the selected elite tea genotypes are vegetatively propagated and later released in the gardens as clonal varieties. Identification of the clones has been mostly done in a traditional way using morphological descriptors like size and shape of the leaf, fruit, etc. However, tea is an outcrossing and heterozygous crop, the different descriptors like the morphological, biochemical, and physiological descriptors show continuous variation with increased plasticity (Lai *et al.,*2001). Morphological traits are influenced by environmental factors like age of the plant and phenology which makes use of such descriptors difficult in identification and discernment of genetic diversity. Therefore, molecular markers such as RAPD, ISSR, AFLP, etc. can be utilized to study the genome directly since the molecular markers are least affected by environmental influence thus eliminating the limitations of a phenotype observation (Lai *et al.,*2001).

4.2.1. RAPD - Random Amplified Polymorphic DNA

RAPD marker developed by Williams *et al.*(1990) has been used extensively to identify and characterize cultivars and is also used to determine the genetic diversity within and between cultivated tea and other related species of *Camellia* (Lai *et al.,*2001). RAPD is a technique used routinely to evaluate the genetic relationship of variation among the species, varieties, and cultivars. Evaluation of genetic diversity using DNA molecular markers unlike the morpho-anatomical is the easiest when it comes to studying the genetic variation or relatedness among the species or cultivars (Roy and Chakraborty, 2009). Previous reports suggest the use of molecular markers like AFLP (Paul *et al.,*1997), RFLP (Devarumath *et al.,*2002), RAPD (Wachira *et al.,*1995; Kaundun *et al*., 2002; Roy and Chakraborty,2007) and

Table 4.3. Total number of amplified bands, number of polymorphic and monomorphic bands and percentage of polymorphism generated by the RAPD primers

SL no.	PRIMER	TB	PB	MB	POL%
1.	OPA01	24	24	0	100%
2.	OPA02	16	16	0	100%
3.	OPA03	22	22	0	100%
4.	OPA04	18	18	0	100%
5.	OPA05	18	18	0	100%
6.	OPA06	13	13	0	100%
7.	OPA07	21	21	0	100%
8.	OPA08	22	22	0	100%
9.	OPA09	20	20	0	100%
10.	OPA10	22	22	0	100%
11.	OPA11	15	15	0	100%
12.	OPA12	16	16	0	100%
13.	OPA13	15	15	0	100%
14.	OPA14	14	14	0	100%
15.	OPA15	14	13	1	92.90%
16.	OPA16	9	9	0	100%
17.	OPA17	13	11	2	84.61%
18.	OPA18	17	17	0	100%
19.	OPA19	24	24	0	100%
20.	OPA20	21	21	0	100%
21.	OPB01	23	23	0	100%
22.	OPB02	14	14	0	100%
23.	OPB03	15	15	0	100%
24.	OPB04	15	15	0	100%
25.	OPB05	25	25	0	100%
26.	OPB06	13	13	0	100%
27.	OPB07	23	23	0	100%
28.	OPB08	16	16	0	100%
29.	OPB09	22	22	0	100%
30.	OPB10	26	26	0	100%
31.	OPB11	11	11	0	100%
32.	OPB12	16	16	0	100%
33.	OPB13	21	21	0	100%
34.	OPB14	20	20	0	100%
35.	OPB15	23	23	0	100%
36.	OPB17	18	18	0	100%
37.	OPB18	17	17	0	100%
38.	OPB19	19	19	0	100%
39.	OPB20	22	22	0	100%
40.	OPF09	18	18	0	100%
41.	OPG19	14	14	0	100%
42.	OPN13	14	14	0	100%
43.	OPH04	18	18	0	100%
44.	OPN04	17	16	1	94.11%
45.	OPN19	13	13	0	100%
	Total	**807**	**803**	**4**	**99.50%**

TB= total bands, PB = polymorphic bands, MB = monomorphic bands, POL% = polymorphism percentage

ISSR (Mingzhe *et al.,* 2010) to study the genetic variation within the tea germplasm. Random Amplified Polymorphic DNA (RAPD) markers have been efficiently utilized earlier to detect genetic variation or establish a genetic relationship in 25 tea genotypes comprising of garden clones, TRA garden series, Darjeeling clones, and some novel clones (Baruah *et al*., 2010).

In our study, a total of 46 different decamer primers have been employed to study the genetic diversity of 33 tea accessions. (Table 4.3). Out of the 46

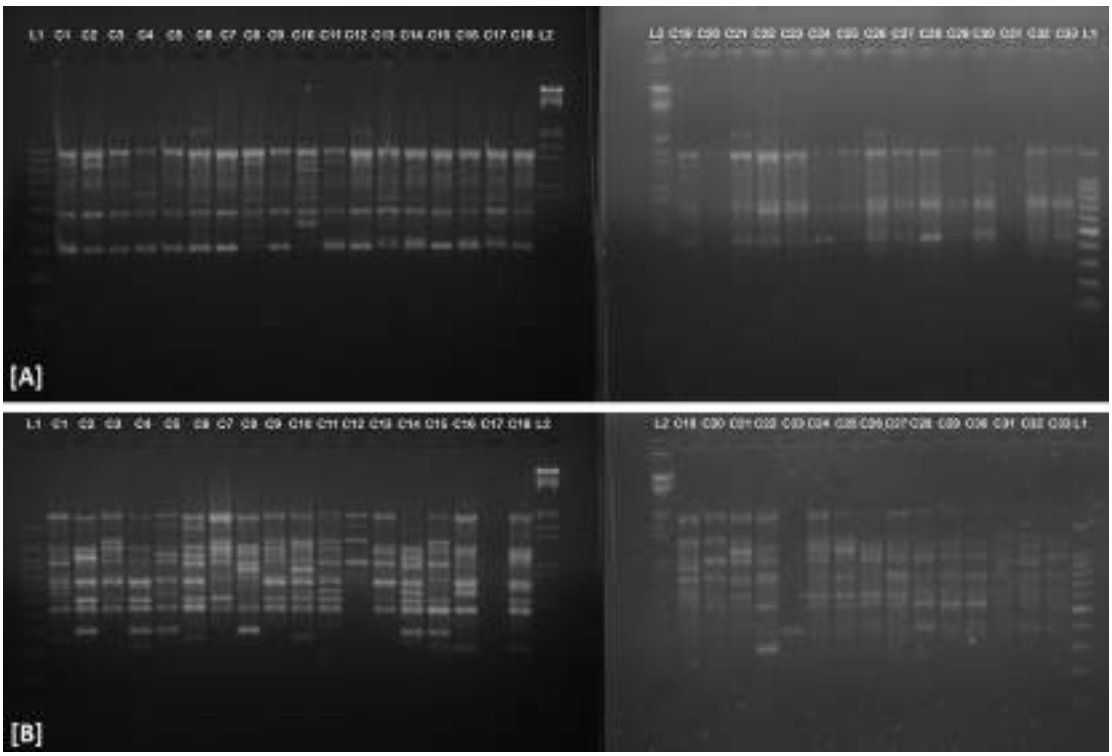


Fig.4.3. RAPD screening: L1- 100 bp DNA ladder, L2- : λ DNA/EcoRI/HindIII double digest DNA ladder, 1-33-tea clones; (A) amplification with OPB06 (B) amplification with OPB15

RAPD primers screened, 45 primers resulted in distinct and scorable bands .

RAPD screening rendered a total of 807 bands out of which 4 bands were monomorphic and the remaining 803 bands showed 100% polymorphism (Table 4.3). The genetic diversity of 27 superior tea germplasms from three different regions i.e., Japan, Korea, and Taiwan were examined using 50 RAPD primers where only 17 primers generated 58 polymorphic bands (Kaundun *et al.,* 2000). The 45 random primers showed a polymorphism percentage of 99.50% and an average PIC value of 0.41. Earlier, 10 random primers used to screen genetic diversity among tea genotypes showed 51.47% polymorphism and a PIC value of 0.356 (Baruah *et al*., 2010). The number of polymorphic bands ranged between 9(OPA16) to 26 (OPB10) (Table 4.3). The RAPD profile of the tea clones generated using primers OPB06 (A) and OPB15 (B) is given in Figure 4.3.

4.2.2. Intersimple Sequence Repeat (ISSR) analysis

ISSR markers developed by Zietkiewicz *et al.* (1994) can also be employed for the same purpose i.e., to study genetic diversity (Lai *et al.,*2001). A total of 15 ISSR primers were further employed to study the genetic variation among the 33 tea clones. All the primers amplified

Table 4.4. Total number of amplified bands, number of polymorphic and monomorphic bands and percentage of polymorphism generated by the ISSR primers

Sl No.	PRIMER	TB	PB	MB
1	UBC807	16	16	0
2	UBC808	14	14	0
3	UBC810	20	20	0
4	UBC811	17	17	0
5	UBC818	15	15	0
6	UBC822	17	17	0
7	UBC824	17	17	0
8	UBC825	28	28	0
9	UBC834	23	23	0
10	UBC836	19	19	0
11	UBC841	16	16	0
12	UBC856	26	26	0
13	UBC873	27	27	0
14	UBC813	21	21	0
15	UBC815	22	22	0
		298	**298**	**0**

TB= total bands, PB = polymorphic bands, MB = monomorphic bands, *100% polymorphism

distinct and scorable bands. Previously twenty 3'-anchored ISSR primers were screened among tea, out of which six primers produced clear and reproducible bands (Lai *et al.,*2001). A total of 298 bands were generated out of which all the bands were polymorphic showing 100% polymorphism (Table 4.4) with an average PIC value of 0.42. The polymorphic bands ranged between 14 (UBC808) to 28(UBC825). The ISSR profile of the studied clones amplified using the primer UBC818 (A), UBC10 (B) is given in Figure 4.4.

4.2.3. DNA fingerprinting analysis

The genetic relatedness is depicted in the dendrogram given (Figure 4.5) where we can see the tea samples clustered into two main groups and many other subgroups showing genetic diversity. The highest similarity was observed between C22 (P-1258) and C23 (RR4/5) with a nodal value of 0.84. The same nodal value of similarity was also shared by the clones C28(T-246) and C29(TV-19). The dendrogram revealed genetic variation within the studied sample. The results were further supported or validated by the 2D (Supplementary Figure 4.1) and 3D (Supplementary Figure 4.2) plot which revealed two major clusters thus pointing towards genetic diversity within the studied clones. A similarity matrix was further drawn using the Dice coefficient of similarity (Nei and Li, 1979) ranging from 0.581 to 0.844 (Figure 4.6).

The ISSR primers also revealed similar kinds of results where the clones studied clustered into two major groups and other subgroups as depicted by the dendrogram (Figure 4.7).

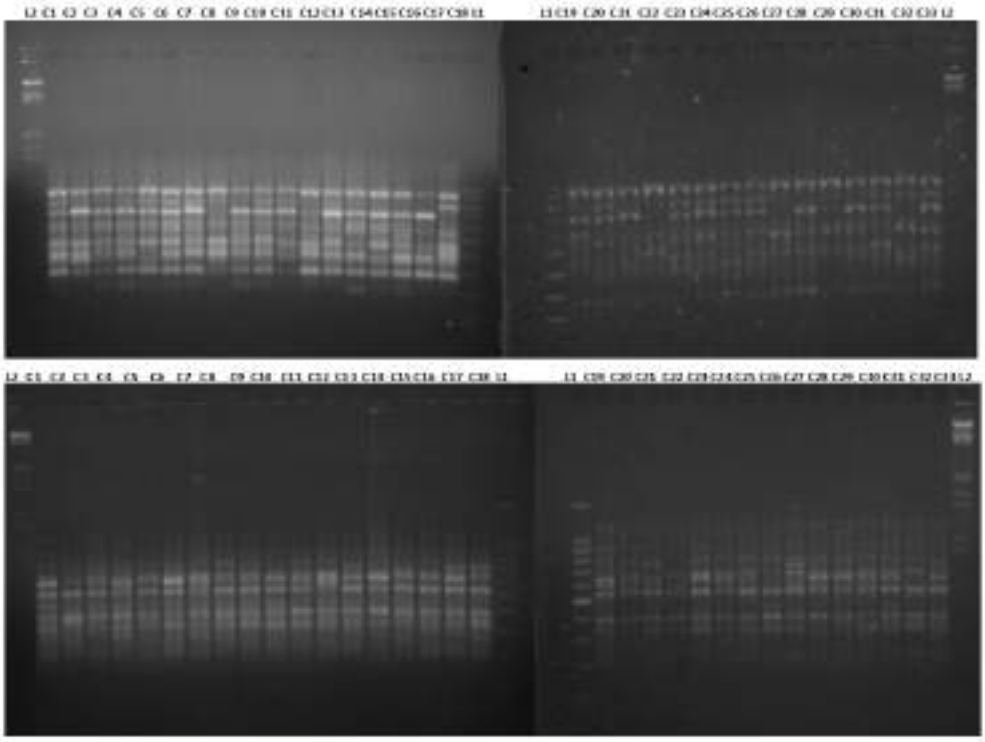


Fig.4.4. ISSR screening: L1- 100 bp DNA ladder, L2- : λ DNA/EcoRI/HindIII double digest DNA ladder, 1-33-tea variety; (A) amplification with UBC818 (B) amplification with UBC10

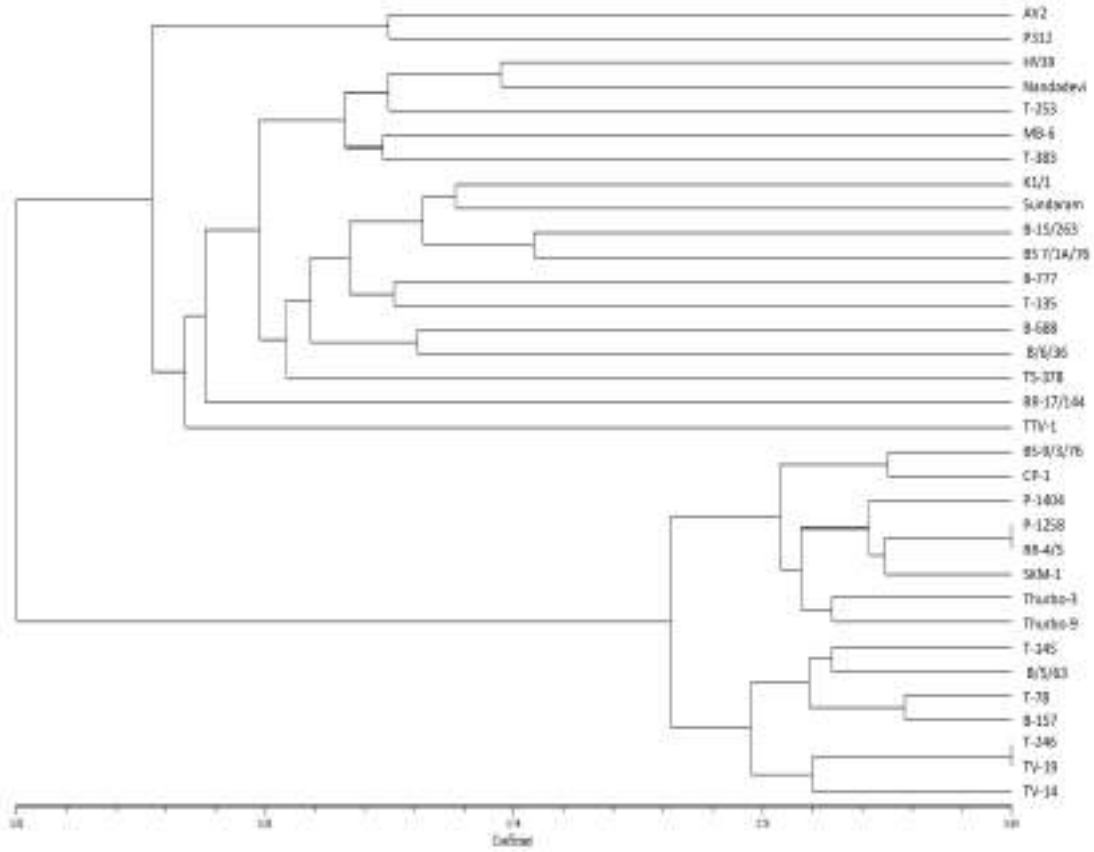


Fig.4.5. Dendrogram obtained from UPGMA cluster analysis of RAPD markers illustrating the genetic relationships among the 33 tea clones

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30	C31	C32	C33
C1	1.00																																
C2	0.71	1.00																															
C3	0.69	0.69	1.00																														
C4	0.68	0.69	0.71	1.00																													
C5	0.66	0.70	0.73	0.70	1.00																												
C6	0.67	0.69	0.70	0.69	0.71	1.00																											
C7	0.64	0.68	0.69	0.65	0.66	0.68	1.00																										
C8	0.63	0.65	0.70	0.67	0.72	0.71	0.65	1.00																									
C9	0.67	0.67	0.72	0.66	0.69	0.67	0.68	0.71	1.00																								
C10	0.63	0.67	0.70	0.67	0.66	0.70	0.67	0.66	0.71	1.00																							
C11	0.64	0.66	0.71	0.70	0.71	0.70	0.68	0.71	0.72	0.74	1.00																						
C12	0.64	0.66	0.67	0.67	0.67	0.68	0.66	0.68	0.71	0.68	0.70	1.00																					
C13	0.64	0.64	0.66	0.67	0.66	0.68	0.68	0.68	0.72	0.72	0.71	0.71	1.00																				
C14	0.65	0.66	0.69	0.69	0.67	0.68	0.66	0.69	0.68	0.70	0.69	0.71	0.71	1.00																			
C15	0.63	0.65	0.65	0.66	0.69	0.67	0.64	0.65	0.69	0.67	0.69	0.68	0.70	0.69	1.00																		
C16	0.62	0.65	0.69	0.68	0.66	0.68	0.66	0.68	0.67	0.70	0.67	0.69	0.70	0.68	0.68	1.00																	
C17	0.65	0.66	0.69	0.69	0.65	0.68	0.65	0.65	0.69	0.68	0.71	0.69	0.72	0.70	0.68	0.71	1.00																
C18	0.61	0.63	0.65	0.66	0.66	0.65	0.63	0.65	0.65	0.68	0.66	0.68	0.69	0.68	0.68	0.66	0.70	1.00															
C19	0.61	0.64	0.66	0.60	0.67	0.62	0.64	0.61	0.63	0.64	0.66	0.62	0.61	0.62	0.60	0.62	0.63	0.64	1.00														
C20	0.64	0.65	0.66	0.64	0.66	0.63	0.63	0.61	0.64	0.66	0.65	0.62	0.61	0.62	0.61	0.61	0.62	0.63	0.82	1.00													
C21	0.63	0.62	0.64	0.63	0.65	0.63	0.63	0.61	0.63	0.64	0.65	0.64	0.63	0.63	0.60	0.63	0.63	0.63	0.79	0.82	1.00												
C22	0.62	0.64	0.64	0.65	0.66	0.64	0.64	0.61	0.64	0.67	0.66	0.64	0.62	0.64	0.63	0.62	0.64	0.65	0.80	0.82	0.82	1.00											
C23	0.61	0.61	0.62	0.61	0.65	0.61	0.60	0.61	0.61	0.62	0.63	0.62	0.59	0.59	0.58	0.61	0.60	0.61	0.79	0.81	0.81	0.84	1.00										
C24	0.63	0.63	0.64	0.62	0.64	0.63	0.64	0.62	0.62	0.64	0.65	0.63	0.60	0.62	0.61	0.61	0.61	0.63	0.78	0.78	0.81	0.82	0.82	1.00									
C25	0.60	0.63	0.63	0.62	0.66	0.63	0.62	0.62	0.61	0.65	0.65	0.64	0.63	0.63	0.61	0.62	0.61	0.63	0.78	0.81	0.80	0.80	0.80	0.80	1.00								
C26	0.61	0.60	0.62	0.61	0.63	0.62	0.59	0.61	0.61	0.62	0.63	0.62	0.61	0.62	0.59	0.58	0.61	0.61	0.76	0.80	0.78	0.79	0.79	0.81	0.80	1.00							
C27	0.62	0.60	0.64	0.62	0.66	0.65	0.60	0.62	0.63	0.65	0.65	0.63	0.62	0.62	0.61	0.61	0.63	0.62	0.74	0.78	0.76	0.76	0.79	0.79	0.79	0.80	1.00						
C28	0.62	0.62	0.64	0.63	0.62	0.60	0.61	0.60	0.62	0.64	0.65	0.63	0.63	0.63	0.59	0.61	0.62	0.62	0.76	0.76	0.77	0.78	0.78	0.75	0.78	0.78	0.79	1.00					
C29	0.64	0.63	0.65	0.63	0.66	0.64	0.62	0.64	0.63	0.64	0.65	0.64	0.62	0.63	0.60	0.61	0.63	0.63	0.76	0.78	0.78	0.79	0.80	0.79	0.80	0.77	0.81	0.84	1.00				
C30	0.61	0.61	0.67	0.64	0.66	0.63	0.63	0.62	0.61	0.63	0.64	0.62	0.60	0.64	0.58	0.62	0.62	0.61	0.74	0.77	0.75	0.77	0.80	0.76	0.77	0.74	0.79	0.78	0.82	1.00			
C31	0.61	0.63	0.64	0.62	0.65	0.62	0.61	0.60	0.61	0.66	0.63	0.63	0.61	0.62	0.59	0.62	0.62	0.62	0.74	0.75	0.76	0.77	0.77	0.79	0.80	0.76	0.79	0.76	0.80	0.79	1.00		
C32	0.62	0.64	0.66	0.64	0.67	0.62	0.62	0.64	0.63	0.64	0.65	0.64	0.62	0.65	0.61	0.64	0.64	0.63	0.73	0.78	0.75	0.75	0.77	0.77	0.79	0.77	0.80	0.78	0.77	0.78	0.82	1.00	
C33	0.62	0.62	0.65	0.64	0.66	0.63	0.63	0.60	0.62	0.62	0.64	0.64	0.63	0.62	0.61	0.63	0.63	0.61	0.74	0.75	0.76	0.76	0.77	0.77	0.79	0.75	0.80	0.77	0.81	0.80	0.80	0.80	1.00

Fig.4.6. Similarity matrix of RAPD analysis calculated using Dice coefficient (Nei and Li, 1979) based on band-sharing from the binary data

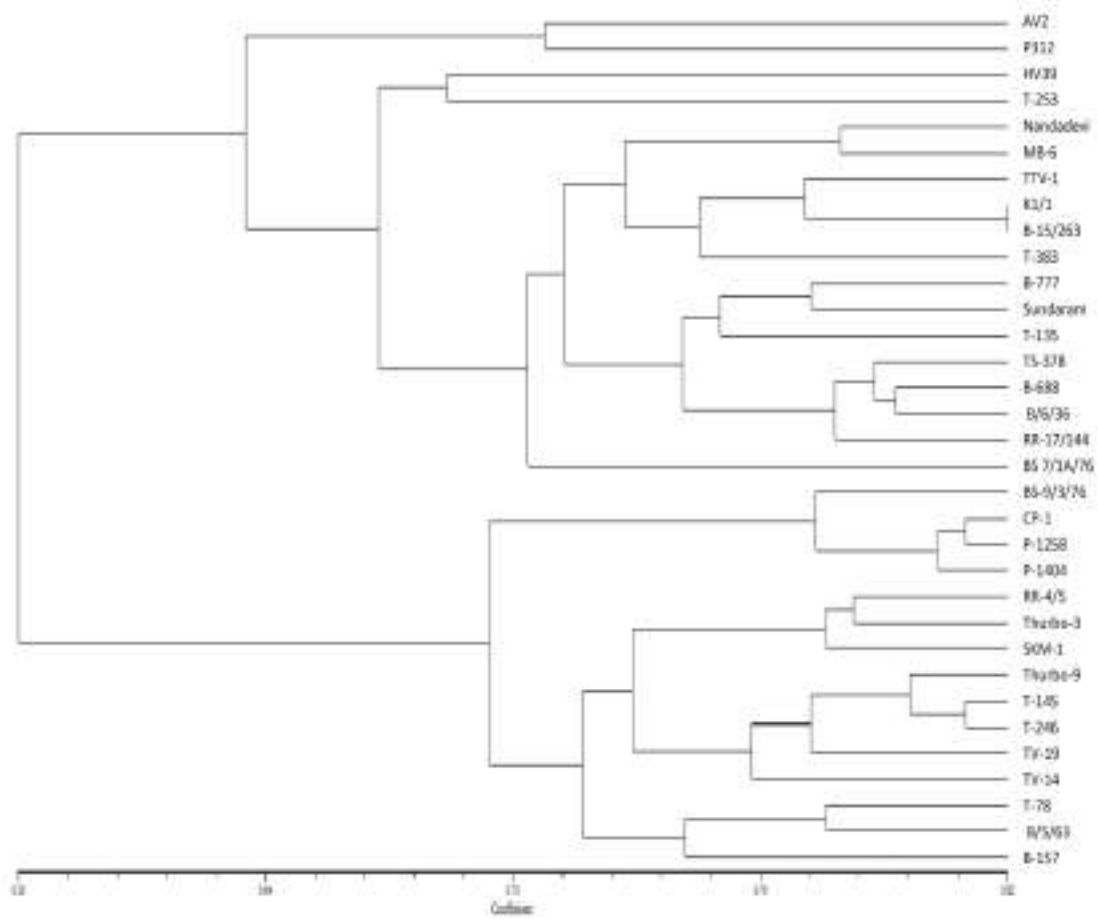
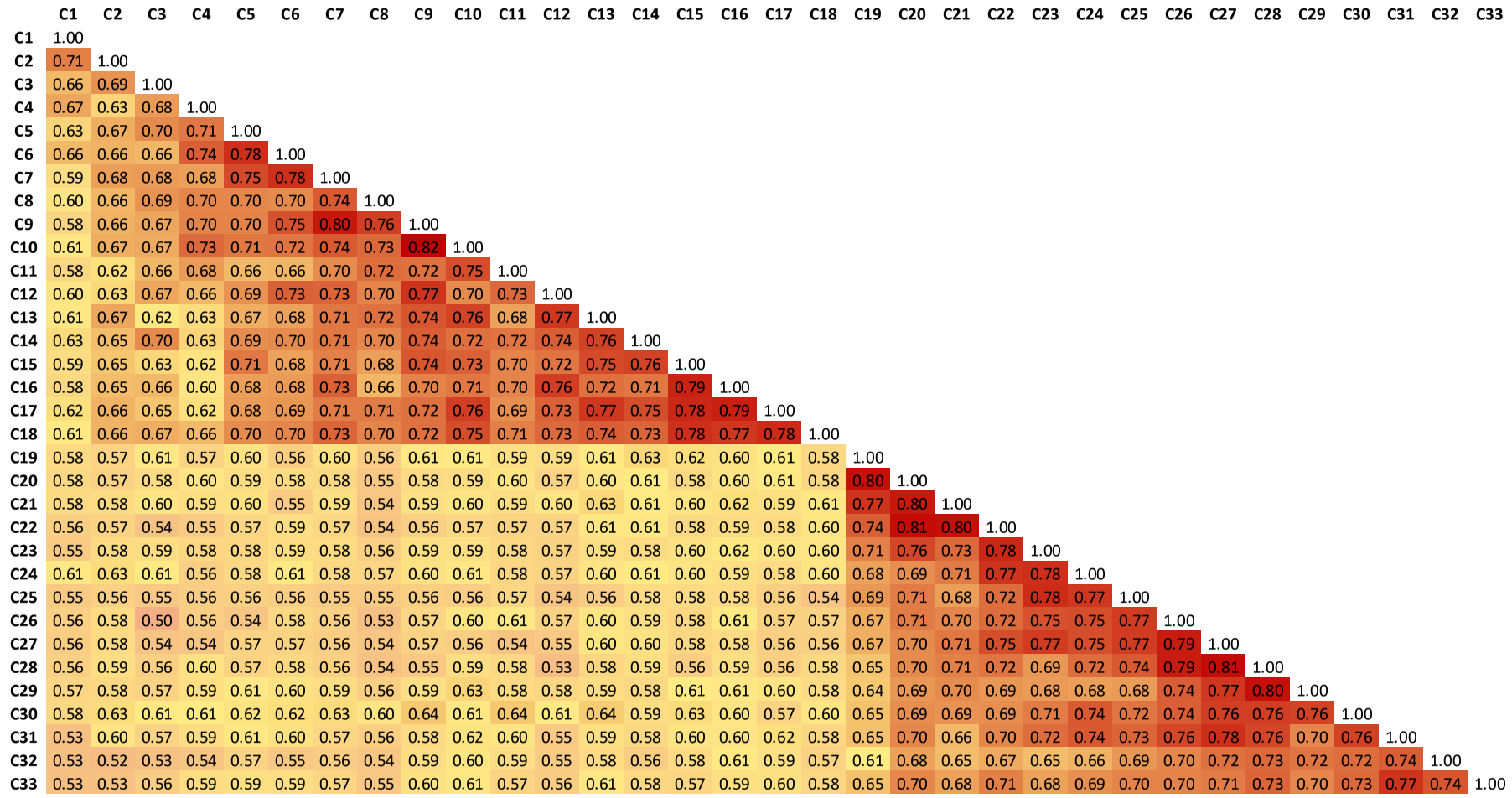


Fig.4.7. Dendrogram obtained from UPGMA cluster analysis of ISSR markers illustrating the genetic relationships among the 33 tea clones

The highest similarity was observed between C9 (K1/1) and C10 (B-15/263) with a nodal value of 0.82. Our findings were corroborated by 2D (Supplementary Figure 4.3) and 3D plots (Supplementary Figure 4.4). A similarity value ranging from 0.500 to 0.817 was observed from the similarity matrix drawn (Figure 4.8).

The results of RAPD and ISSR profiling more or less corroborated by clustering into broad two groups and many other subgroups. However, the pattern of sub-grouping does not reveal as such a clear pattern of grouping irrespective of the sample source, sampling area, or the place of selection or collection. Our findings thus suggest an extensive exchange of genetic resources between the sampling regions. A minimum genetic diversity was recorded previously among the tea populations irrespective of the place of selection or collection (Baruah *et al*., 2010). In our study ISSR fingerprinting revealed more polymorphism than RAPD fingerprinting. A similar kind of result was reported earlier (Devarumath *et al.,* 2002) where ISSR profiling detected more polymorphic loci than RAPD profiling.



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30	C31	C32	C33
C1	1.00																																
C2	0.71	1.00																															
C3	0.66	0.69	1.00																														
C4	0.67	0.63	0.68	1.00																													
C5	0.63	0.67	0.70	0.71	1.00																												
C6	0.66	0.66	0.66	0.74	0.78	1.00																											
C7	0.59	0.68	0.68	0.68	0.75	0.78	1.00																										
C8	0.60	0.66	0.69	0.70	0.70	0.70	0.74	1.00																									
C9	0.58	0.66	0.67	0.70	0.70	0.75	0.80	0.76	1.00																								
C10	0.61	0.67	0.67	0.73	0.71	0.72	0.74	0.73	0.82	1.00																							
C11	0.58	0.62	0.66	0.68	0.66	0.66	0.70	0.72	0.72	0.75	1.00																						
C12	0.60	0.63	0.67	0.66	0.69	0.73	0.73	0.70	0.77	0.70	0.73	1.00																					
C13	0.61	0.67	0.62	0.63	0.67	0.68	0.71	0.72	0.74	0.76	0.68	0.77	1.00																				
C14	0.63	0.65	0.70	0.63	0.69	0.70	0.71	0.70	0.74	0.72	0.72	0.74	0.76	1.00																			
C15	0.59	0.65	0.63	0.62	0.71	0.68	0.71	0.68	0.74	0.73	0.70	0.72	0.75	0.76	1.00																		
C16	0.58	0.65	0.66	0.60	0.68	0.68	0.73	0.66	0.70	0.71	0.70	0.76	0.72	0.71	0.79	1.00																	
C17	0.62	0.66	0.65	0.62	0.68	0.69	0.71	0.71	0.72	0.76	0.69	0.73	0.77	0.75	0.78	0.79	1.00																
C18	0.61	0.66	0.67	0.66	0.70	0.70	0.73	0.70	0.72	0.75	0.71	0.73	0.74	0.73	0.78	0.77	0.78	1.00															
C19	0.58	0.57	0.61	0.57	0.60	0.56	0.60	0.56	0.61	0.61	0.59	0.59	0.61	0.63	0.62	0.60	0.61	0.58	1.00														
C20	0.58	0.57	0.58	0.60	0.59	0.58	0.58	0.55	0.58	0.59	0.60	0.57	0.60	0.61	0.58	0.60	0.61	0.58	0.80	1.00													
C21	0.58	0.58	0.60	0.59	0.60	0.55	0.59	0.54	0.59	0.60	0.59	0.60	0.63	0.61	0.60	0.62	0.59	0.61	0.77	0.80	1.00												
C22	0.56	0.57	0.54	0.55	0.57	0.59	0.57	0.54	0.56	0.57	0.57	0.57	0.61	0.61	0.58	0.59	0.58	0.60	0.74	0.81	0.80	1.00											
C23	0.55	0.58	0.59	0.58	0.58	0.59	0.58	0.56	0.59	0.59	0.58	0.57	0.59	0.58	0.60	0.62	0.60	0.60	0.71	0.76	0.73	0.78	1.00										
C24	0.61	0.63	0.61	0.56	0.58	0.61	0.58	0.57	0.60	0.61	0.58	0.57	0.60	0.61	0.60	0.59	0.58	0.60	0.68	0.69	0.71	0.77	0.78	1.00									
C25	0.55	0.56	0.55	0.56	0.56	0.56	0.55	0.55	0.56	0.56	0.57	0.54	0.56	0.58	0.58	0.58	0.56	0.54	0.69	0.71	0.68	0.72	0.78	0.77	1.00								
C26	0.56	0.58	0.50	0.56	0.54	0.58	0.56	0.53	0.57	0.60	0.61	0.57	0.60	0.59	0.58	0.61	0.57	0.57	0.67	0.71	0.70	0.72	0.75	0.75	0.77	1.00							
C27	0.56	0.58	0.54	0.54	0.57	0.57	0.56	0.54	0.57	0.56	0.54	0.55	0.60	0.60	0.58	0.58	0.56	0.56	0.67	0.70	0.71	0.75	0.77	0.75	0.77	0.79	1.00						
C28	0.56	0.59	0.56	0.60	0.57	0.58	0.56	0.54	0.55	0.59	0.58	0.53	0.58	0.59	0.56	0.59	0.56	0.58	0.65	0.70	0.71	0.72	0.69	0.72	0.74	0.79	0.81	1.00					
C29	0.57	0.58	0.57	0.59	0.61	0.60	0.59	0.56	0.59	0.63	0.58	0.58	0.59	0.58	0.61	0.61	0.60	0.58	0.64	0.69	0.70	0.69	0.68	0.68	0.68	0.74	0.77	0.80	1.00				
C30	0.58	0.63	0.61	0.61	0.62	0.62	0.63	0.60	0.64	0.61	0.64	0.61	0.64	0.59	0.63	0.60	0.57	0.60	0.65	0.69	0.69	0.69	0.71	0.74	0.72	0.74	0.76	0.76	0.76	1.00			
C31	0.53	0.60	0.57	0.59	0.61	0.60	0.57	0.56	0.58	0.62	0.60	0.55	0.59	0.58	0.60	0.60	0.62	0.58	0.65	0.70	0.66	0.70	0.72	0.74	0.73	0.76	0.78	0.76	0.70	0.76	1.00		
C32	0.53	0.52	0.53	0.54	0.57	0.55	0.56	0.54	0.59	0.60	0.59	0.55	0.58	0.56	0.58	0.61	0.59	0.57	0.61	0.68	0.65	0.67	0.65	0.66	0.69	0.70	0.72	0.73	0.72	0.72	0.74	1.00	
C33	0.53	0.53	0.56	0.59	0.59	0.59	0.57	0.55	0.60	0.61	0.57	0.56	0.61	0.58	0.57	0.59	0.60	0.58	0.65	0.70	0.68	0.71	0.68	0.69	0.70	0.70	0.71	0.73	0.70	0.73	0.77	0.74	1.00

Fig.4.8. Similarity matrix of ISSR analysis calculated using Dice coefficient (Nei and Li, 1979) based on band-sharing from the binary data

4.3. DNA barcode

The advancement of sequencing technology has uplifted molecular research, and even a minor difference whether being intraspecific or interspecific can be detected which would otherwise not be easily accomplished using other morphological or robust molecular techniques. DNA barcoding utilizes numerous plastid and nuclear regions to identify species. Seven different plastid regions have been explored in land plants where they have suggested the combination of rbcL+matK to be most precise and appropriate for plant barcode analysis (Group *et al.*, 2009). Successful utilization of a combination of matK+ITS and rbcL+trnH-psbA to study 100% differences between *Cassia* species has already been reported in previous studies (Purushothaman *et al.*, 2014). The matK region solely has been used efficiently to differentiate *Vachellia* species from other *Acacia* species (Steven *et al.*, 2009), where the study highlighted the possibility of utilizing matK for separating taxa at the genus level. Some previous studies have highlighted the successful amplification and utilization of the matK region to explore the phylogeny in both monocots and dicots such as Zingiberaceae (Selvaraj *et al.*, 2008), Erythronium (Allen *et al.*, 2003), *Myristica fragrans* (Tallei and Kolondam, 2015), local tomato (Caprar *et al.*, 2017) and oil-bearing roses (Wang *et al.*, 2012). Apart from species recognition or variation study, nowadays the barcode technique is also used to detect other contaminants. For instance, one study reports the origin of bamboo leaves as well as the presence of adulterants in manufactured bamboo tea products. The adulterants present in bamboo tea have counter-indications for pregnant women in products (Horn and Haser, 2016). The DNA barcodes like rbcL, matK, ITS2, and psbA-trnH have been used previously to differentiate between the commercial non-*Camellia* tea and the adulterants present in it just to ensure its safety. However, a limited number of original plant sequences in GenBank limited the findings of the study (Long *et al.*, 2014).

4.3.1. matK Amplification and Sequencing

The primer successfully amplified the matK (Figure 4.9) region of all the 33 tea clones under investigation. The size of the PCR product amplified approximately ranged between 900 bp –1000 bp. However, the sequencing result deduced the size of the final amplified product ranging from 644 bp to 876 bp. The accession number for the submitted sequences is provided in Table 4.5 & Appendix- C.

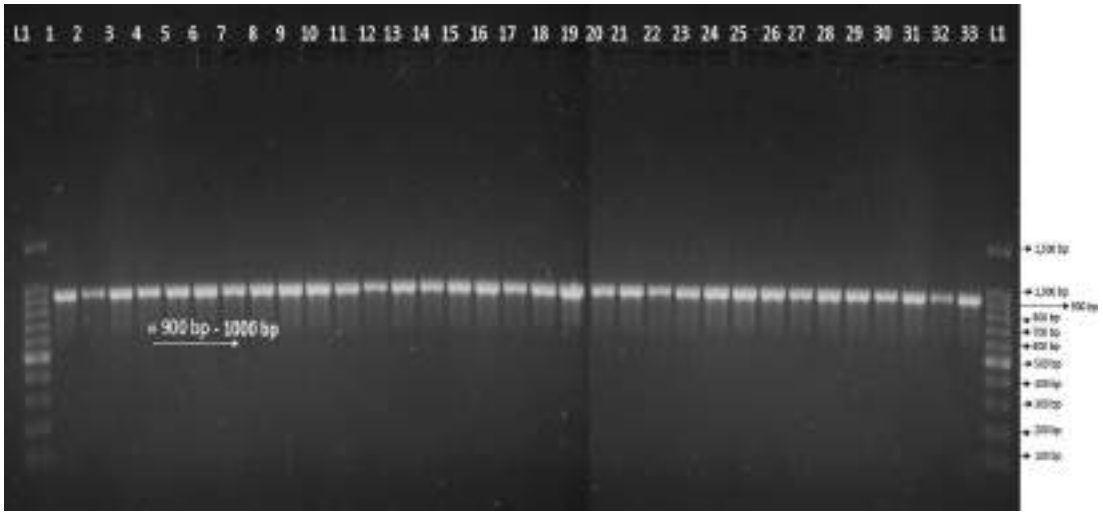


Fig.4.9. Amplification of the matK region. Lane L1- 100 bp DNA ladder; Lane 1-33: 33 tea clones.

Table 4.5. Accession number and details of the submitted matK sequence retrieved from NCBI

Sl. no	Full name	Abbreviation	NCBI accession	Unique Id	Base pairs
1.	Ambari Vallai 2	AV2	MH649284	1111	757bp
2.	Phoobsering 312	P312	MK393394	1112	871bp
3.	Happy valley 39	HV39	MH791417	1113	864bp
4.	Tukdah 253	T-253	MH920315	1114	876bp
5.	Nanda Devi	TS 378	MH920316	1115	816bp
6.	Makaibari-6	MB-6	MH920317	1116	758bp
7.	Teesta Valley 1	TTV-1	MH920318	1117	861bp
8.	Kopati 1/1	K1/1	MH920319	1119	774bp
9.	Balasun 7/1A/76	BS 7/1A/76	MK393393	11111	833bp
10.	Bunnockburn 777	B-777	MK393395	11112	757bp
11.	Sundaram	B/5/63	MK393396	11113	644bp
12.	Tukdah-135	T-135	MK393397	11114	763bp
13.	Bunnockburn 688	B-688	MK393398	11116	833bp
14.	Golconda	Golconda	MK393399	11117	644bp
15.	RungliRungiliot 17/144	RR-17/144	MK393400	11118	826bp
16.	Balasun 9/3/76	BS-9/3/76	MK393401	11119	756bp
17.	Chiradew Parbat1	CP-1	MK393402	11120	761bp
18.	Phoobsering 1404	P-1404	MK393403	11121	763bp
19.	Phoobsering 1258	P-1258	MK393404	11122	751bp
20.	RungliRungiliot 4/5	RR-4/5	MK393405	11123	746bp
21.	Sikkim 1	SKM-1	MK424865	11124	867bp
22.	Thurbo 3	Thurbo-3	MN480321	11125	761bp
23.	Thurbo 9	Thurbo-9	MN480322	11126	707bp
24.	Tukdah 145	T-145	MK424866	11127	761bp
25.	Tukdah 246	T-246	MK424867	11128	756bp
26.	Tocklai variety 19	TV-19	MK424868	11129	750bp
27.	Tocklai variety 14	TV-14	MK424869	11130	735bp
28.	Tukdah 78	T-78	MK424870	11131	755bp
29.	Bannockburn 157	B-157	MK424871	11132	735bp

4.3.2. Blast Result

As per the blast results, a total of 24 clones out of 29 clones was found to be 100% identical with *Camellia sinensis*. The percentage of similarity with *Camellia sinensis* as shown by the remaining five clones were 99.29% (Thurbo 9), 99.61% (Thurbo 3), 99.64% (RR-4/5), 99.88% (SKM-1), and 99.89% (P312). However, the clones like Thurbo 3 and Thurbo 9 showed a relatively higher percentage of similarity with other species of *Camellia*, i.e., 99.87% (Thurbo 3) and 99.57% (Thurbo 9) similarity with *Camellia mairei* (KJ197933.1). The clones were therefore placed under different species of *Camellia* with a percentage similarity value below 99.64%. The matK locus placed two genera *Myristica* and *Knena* differently at a sequence similarity of 99.43% while genus *Virola* differed with 99.25% (Tallei and Kolondam, 2015). With a similarity percentage of 99.64%, tomatoes were placed within the same species contradicting the assumption of required percent identity as 99.74 –100% to place the organisms within the same species (Lawodi, 2013). Our study has differentiated two species at percent identity below 99.64% where the clones like Thurbo 3 showed 99.61% and Thurbo 9 showed 99.29% identity with *C. sinensis*. However, our study showed 99.87% (Thurbo 3) and 99.57% (Thurbo 9) identity with *C. mairei* which is comparatively higher than the identity shared with *C. sinensis*.

4.3.3. Sequence Alignment and Phylogenetic Tree Construction

The phylogenetic tree constructed using neighbor-joining (Figure 4.10) and UPGMA (Supplementary Figure 4.5) method depicted variation among the sequences. The combined nucleotide sequences of almost all the clones clustered together except for clones like Thurbo 3 (11125), Thurbo 9 (11126), P312 (1112), and RR-17/144 (11118). The clones like Thurbo 3 and Thurbo 9 clustered together whereas the clones P312 and RR-17/144 diverged from the main group. To validate our results, a phylogenetic tree was further constructed considering a reference matK sequence of different *Camellia* species taken from NCBI (KJ197933.1). The clones like Thurbo 3 (11125) and Thurbo 9 (11126) are now clustered with the reference sequence of *C. mairei* (KJ197933.1) as shown in the neighbor-joining (Figure 4.11) and UPGMA tree (Supplementary Figure 4.6).

4.3.4. Sequence Analysis

The genetic distances for the matK sequence as calculated using Nucleotide: Maximum composite likelihood method and Nucleotide

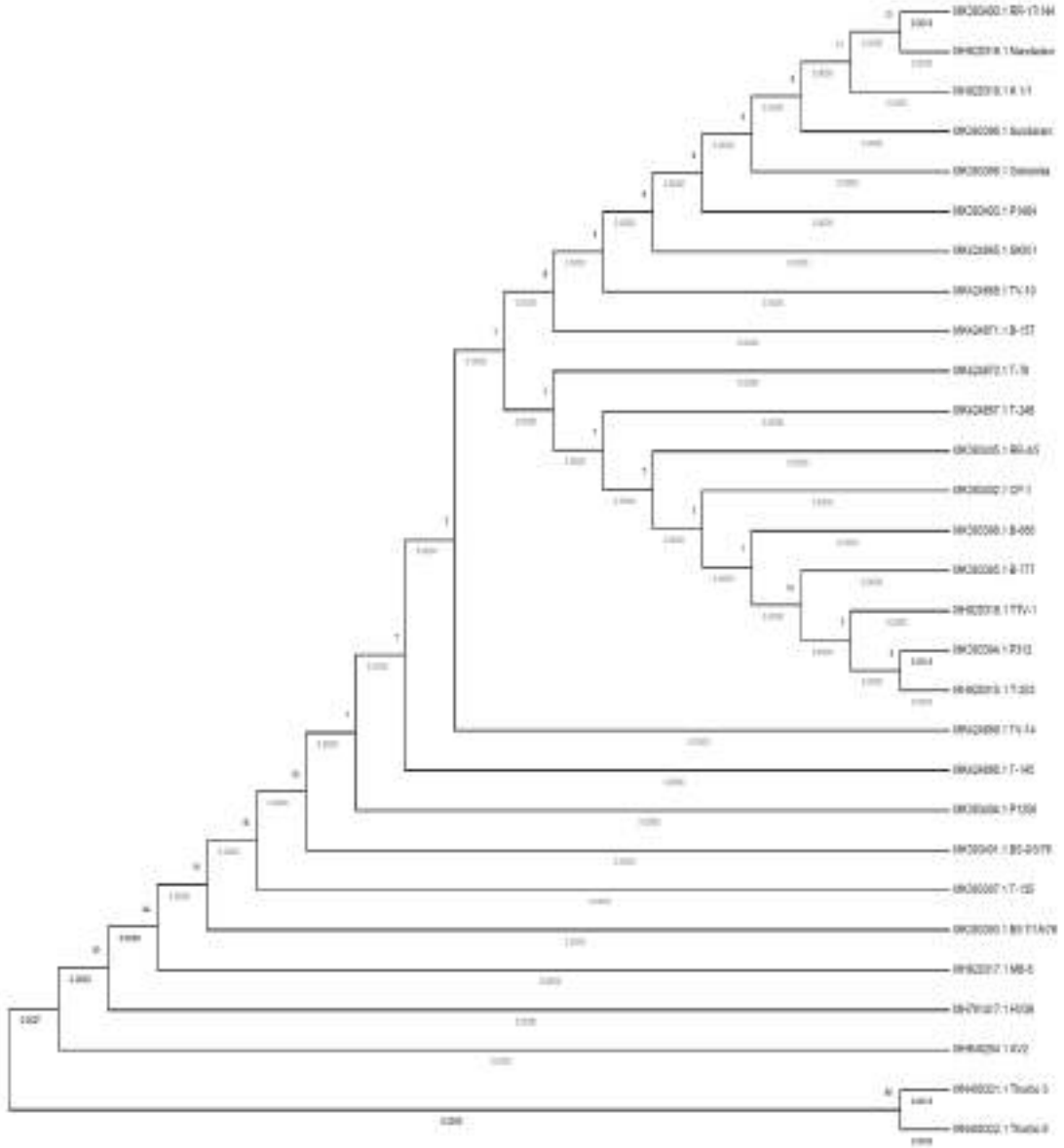


Fig.4.10. Neighbor joining tree method showing the genetic relationship of matK region between 29 tea clones

Kimura 2-parameter method ranged between 0 to 0.0090 (Figure 4.12) and 0 to 0.0089 (Supplementary Figure 4.7) respectively. The overall calculated mean distance was 0.0013. The results show the number of base substitutions per site and are based on an analysis of a total of 29 sequences with all codon positions included and with a total of 563 positions in the final dataset excluding the eliminated positions containing gaps and missing data. The matK sequence of Thurbo 3, Thurbo 9, and *C. mairei* shared two unique variable sites that differed from the rest of the sequences. It was further validated by analyzing a high consensus sequence of 563 bp prepared using Multalin software. The high consensus region showed multiple

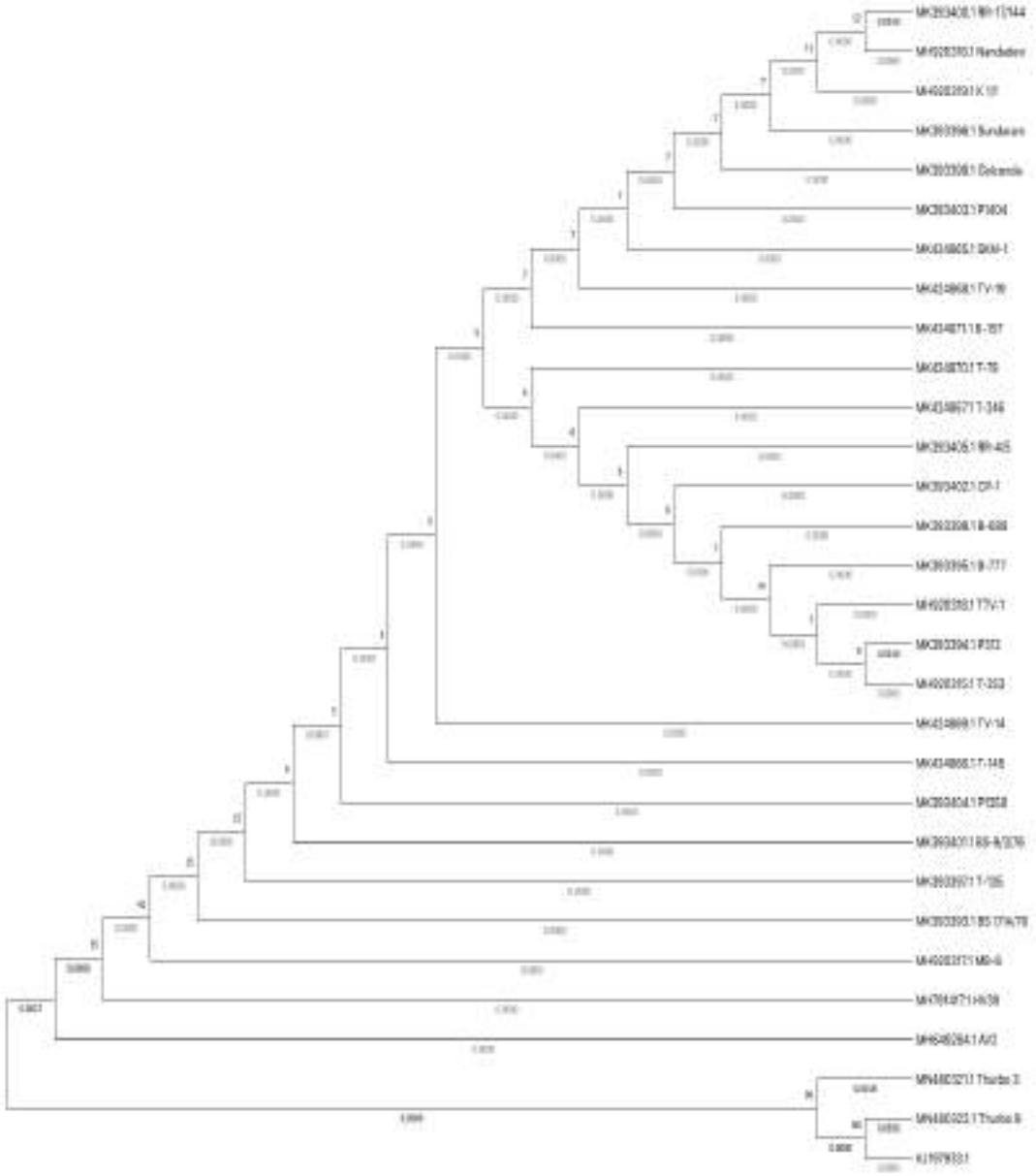


Fig.4.11. Neighbour joining tree method showing the genetic relationship of matK region between 29 tea clones along with sequence of *Camellia mairei* (KJ197933.1) taken from NCBI

variable sites, where Thurbo 3 (11125) showed 3 substitutions, Thurbo 9 (11126) showed 4 substitutions, and some single substitutions were found in P312 (1112) and RR-17/144 (11118) as shown in Figure 4.13 and Supplementary material 1 (SM1). Similarly, the low consensus region also revealed a total of five nucleotide substitutions or variations in clones like SKM-1 (3 substitutions), Thurbo 9 (1 substitution), and P1258 (1 substitution) as shown in Supplementary material 2 (SM2). Earlier, two species *T. cope* and

T. wightii (Ragupathy *et al.*, 2009) did not show variation at rbcl locus but

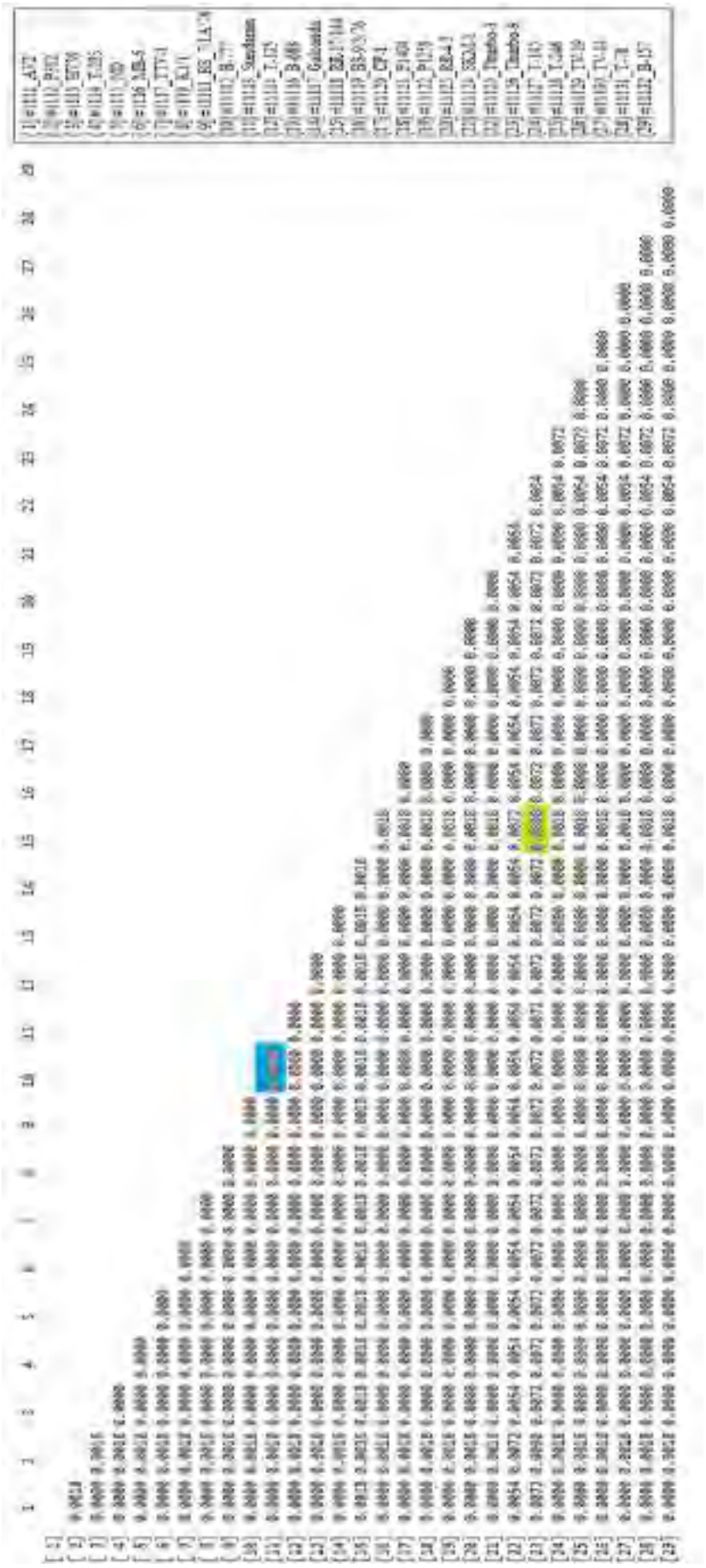


Fig.4.12. Genetic distances of the matK sequence calculated using Nucleotide: Maximum Composite Likelihood method.

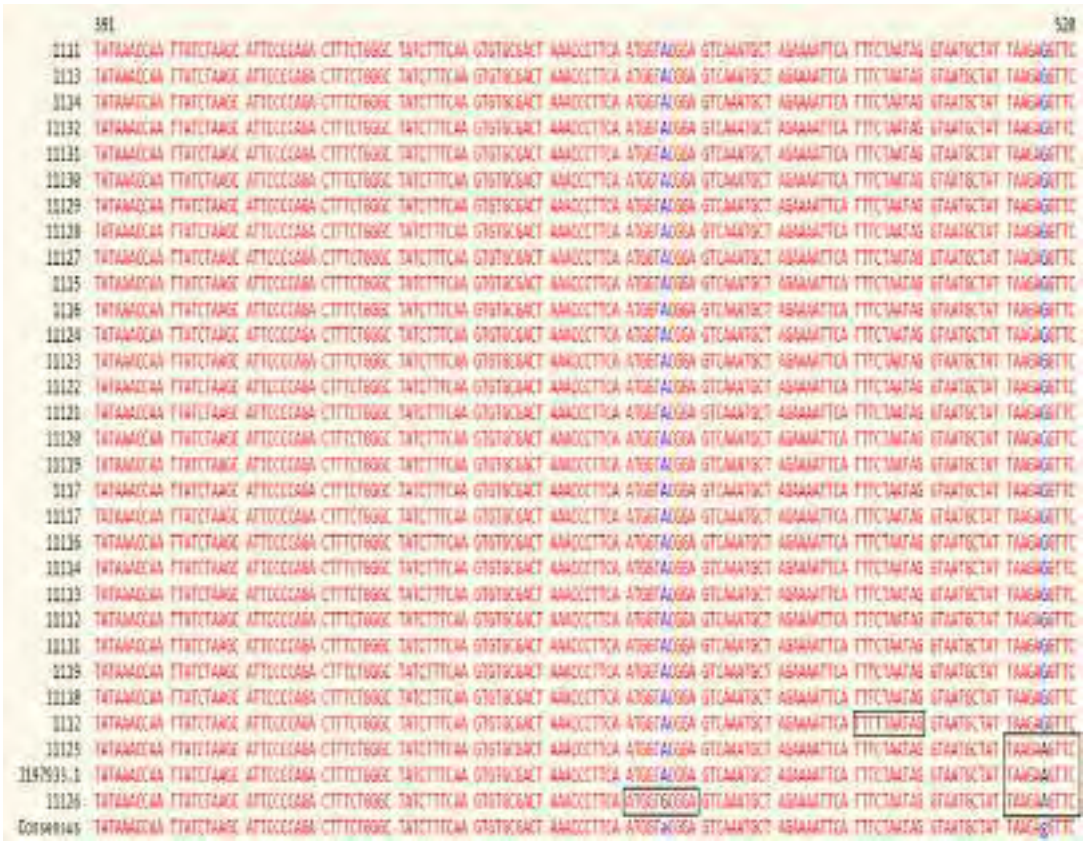


Fig.4.13. Consensus region of the aligned sequences showing nucleotide substitution as highlighted by black box

showed a variable site in the matK region where it showed 2 nucleotide variation and trnHpsbA showed 1 nucleotide variation. The interspecific variation could be broadened if the two loci are considered as a two-gene approach and therefore they report interspecific variation at p-distance 0.002– 0.003 but with the p-distance value of 0.00, they report the absence of intraspecific variation. A total of three variable sites was reported in trnH-psbA sequences of seven tomato varieties with the genetic distance ranging from 0 to 0.004 (Caprar *et al.*, 2017). On the other hand, the study reported no variable sites in the rbcL, rpoC1, and rpoB sequences thus suggesting it to be 100% conserved within the species. The ability of matK locus to differentiate only at the genus level within the family of Myristicaceae has been reported earlier where the matK locus failed to differentiate *Myristica* at species level showing 100% similarity in blast analysis with other three species of *Myristica* sharing three nucleotide differences with *Rivola sebifera* and four nucleotide differences with *Knema laurina*. However, our study reports a total of nine variable sites in the high consensus region and a total of five variable sites in the low consensus

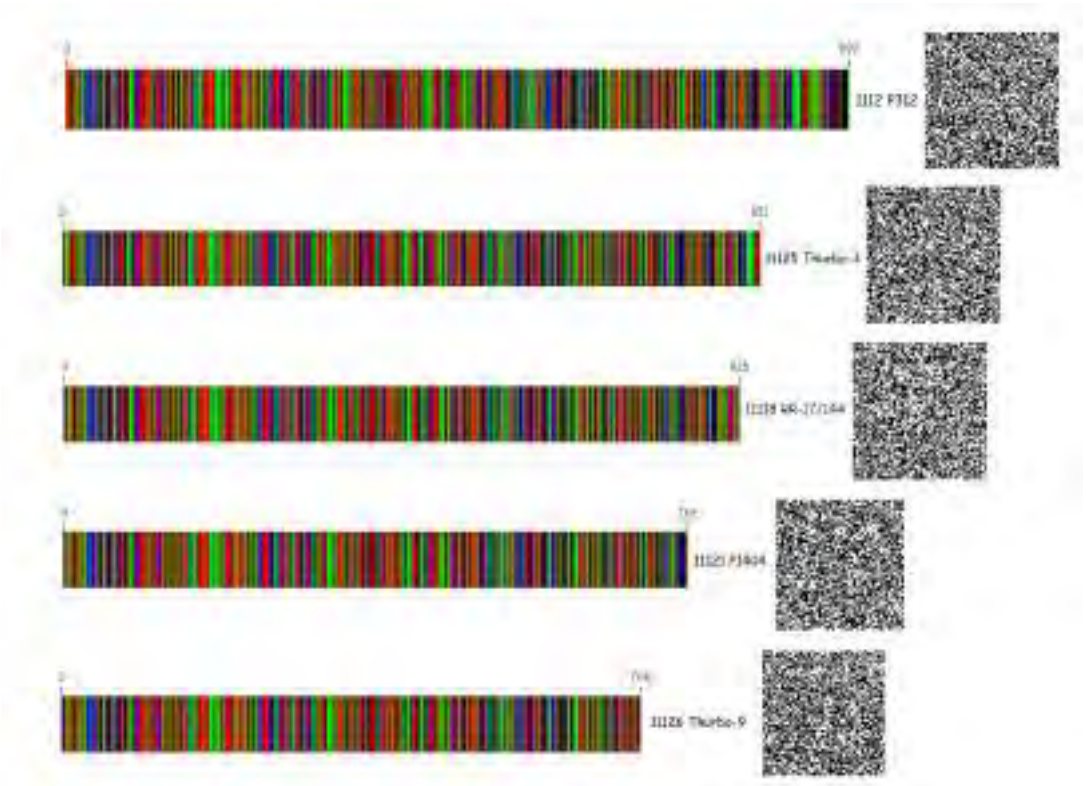


Fig.4.14. Illustrative representation of matK sequences as barcode and QR code

region of matK sequences within the same species of *Camellia sinensis*. Therefore, we report intraspecific variation and conclude with a fact of matK sequence not being 100% conserved within the same species of *Camellia*.

The matK sequence under investigation is represented illustratively as barcode and QR code [Figure 4.14 and Supplementary material 3 (SM3)]. The four different color code in the barcode depicts the different nucleotide composition (A, T, G or C) of the matk sequence. The QR code generated can be decoded as DNA sequences which make data storage and retrieval easy.

Fig 4.15. Herbarium of Thurbo 3 (*C. sinensis*) with the partial sequence information of its matk region enscrypted as QR code

The sequence information can also be enscrypted as QR code along with the herbarium of the specimen as shown in Figure 4.15 .

4.4. Phytochemical screening

4.4.1. Qualititative tests

The three polar solvents i.e., acetone, methanol, and ethanol persistently gave better results in almost all the qualitative tests like tests for cardiac glycoside, steroid, flavonoid, diterpenes, terpenoids, saponin, tannin, and reducing sugar as shown in Figure 4.16. These three solvents showed the highest potency to extract important phytochemicals from tea leaves. Extracting solvent and its nature (polar/

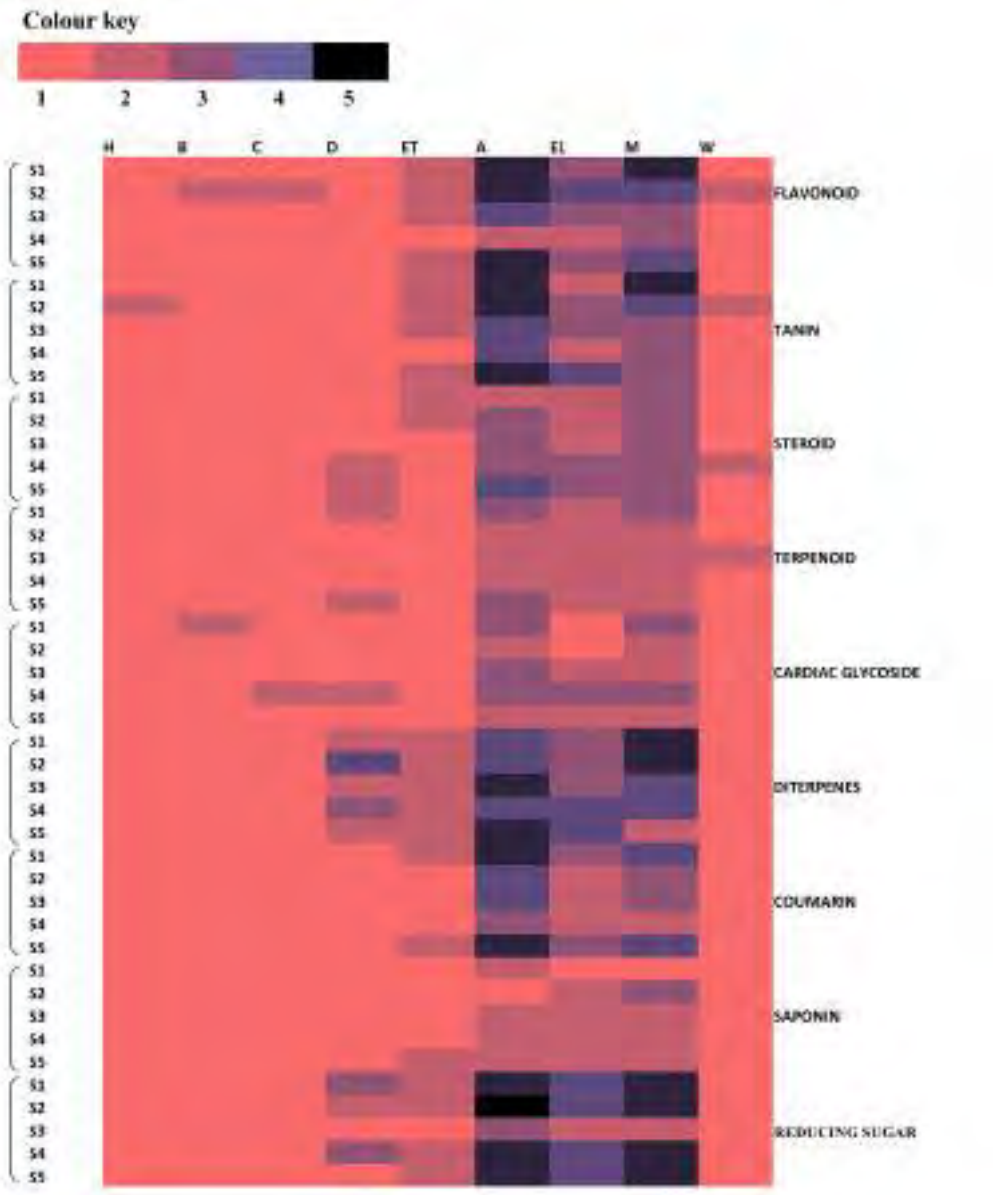


Fig.4.16. Heatmap representing the qualitative phytochemical profiling of different extracts of selected tea varieties. Colour key provided from pink to black represents the intensity in increasing order

non-polar) is very crucial in the extraction of phytochemicals since the compounds have different polarities, chemical characteristics, and solubilities (Ozarkar, 2005). An abundant number of phytochemicals was present in the extracts prepared using acetone, methanol, and ethanol, and traces of phytochemicals like flavonoid, tannin, and cardiac glycoside was found in the extracts prepared using less polar or non-polar solvents like hexane, benzene, and chloroform. High polar solvents like water, ethanol, and acetone extracted major phytochemicals groups like flavonoids, alkaloids, terpenoids, saponins, and phenols from plant extracts of *C. sinensis* (purple tea) when compared with non-polar ethyl acetate and chloroform (Geoffrey *et al.,* 2014). One of the reports suggests methanol as the most suitable solvent than acetone and ethyl acetate to extract phytochemicals from black packaged tea. Methanol showed major extraction properties for extracting phytochemicals like flavonoid, reducing sugar triterpenes, lipid and tannin while other solvents showed minimum activity (Patil *et al.,* 2016). Along with the use of varying solvents for extraction, the time of extraction, and the procedure followed (cold extraction / hot extraction), the state of the sample also plays a crucial role during phytochemical extraction since the phytochemicals slowly degenerate from the time of plucking up to manufacturing. Qualitative screening of phytochemicals very much plays an important role since the presence of a phytochemical of interest or importance may lead to further isolation, purification, and characterization of desired phytochemical of pharmaceutical importance (Ugochukwu *et al.,* 2013).

Since the antioxidant ability of a particular sample or extract cannot be determined using only qualitative method, so methods like total phenol and flavonoid estimation, DPPH free radical scavenging assay, and FRP assay provided some insight into the potent capability of extracting solvents to extract crucial phytochemical compounds with the antioxidant or bioactive property.

4.4.2. DPPH assay

At a single concentration of 200 µg/ml, the percentage of inhibition or the highest radical scavenging percentage in the DPPH assay was observed in acetone extracts followed by methanol extract and ethanol extract (Figure 4.17). The lowest radical scavenging activity was shown by benzene extracts.

The antioxidant potential of the two

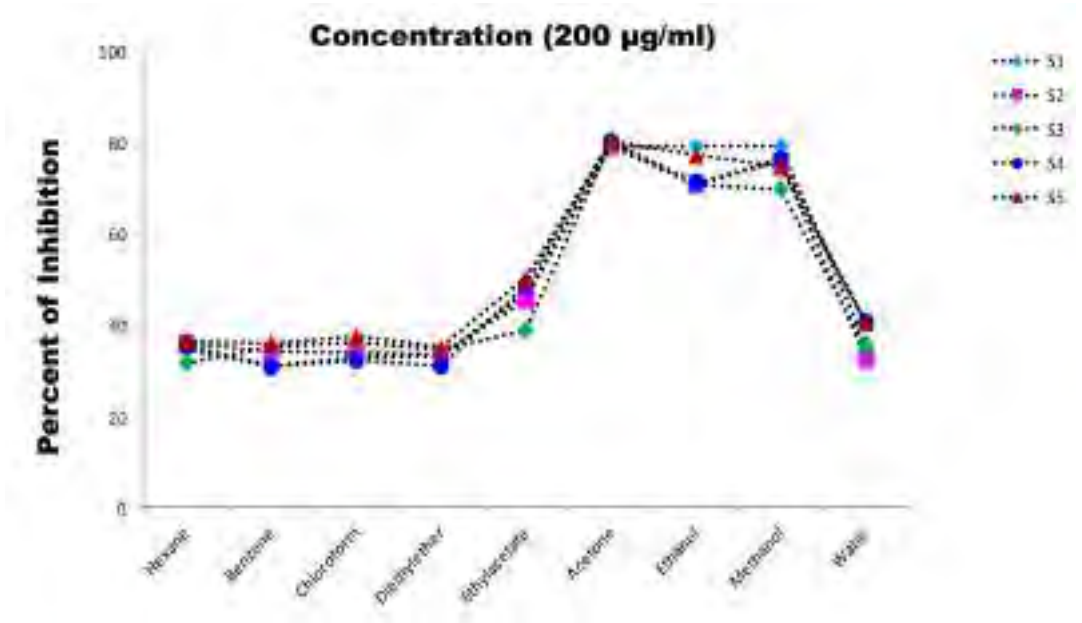


Fig.4.17. DPPH scavenging activity of different solvent leaf extracts of tea clones at single concentration (200µg/ml)

best solvent extracts i.e., the acetone extracts and the methanol extracts were further compared alongside the standard ascorbic acid using various concentrations ranging from 1 mg/ml up to 5 mg/ml. Acetone (Figure 4.18.a) and methanol extracts (Figure 4.18.b) showed antioxidant activity or free radical scavenging property almost similar to the standard used. When compared among all the five clones, acetone gave promising results than methanol. The sample concentration providing 50% of inhibition is the IC_{50} value. The lower the IC_{50} value the higher the antioxidant activity. It was determined by plotting a curve of inhibition using various concentrations. Most of the time the IC_{50} value increases with increasing concentration

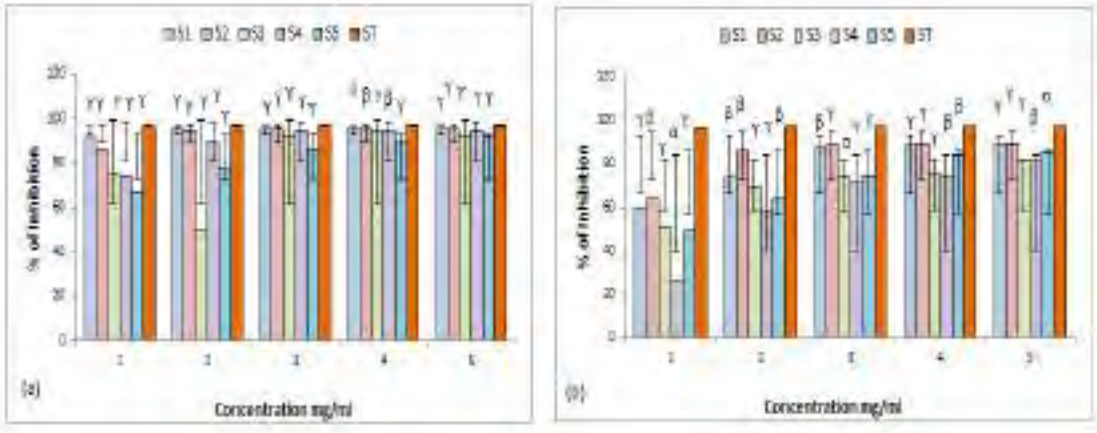


Fig.4.18.DPPH scavenging activity of (a) acetone extracts; methanol extracts; $p<0.05$; β $p<0.01$; ϒ$p<0.001$; € non-significant

Table 4.6. IC_{50} value of acetone and methanol extracts prepared in different concentrations (mg/ml) during DPPH assay

Accession	IC_{50} (mean ± sd) of acetone extracts(mg/ml)	IC_{50} (mean ± sd) of methanol extracts(mg/ml)
S1 (TS569)	0.111±0.001	0.635±0.028
S2 (China)	0.159±0.002	0.478±0.028
S3 (AV2)	0.600±0.002	1.030±0.061
S4 (P312)	0.294±0.008	1.595±0.297
S5 (Assam)	0.515±0.002	1.016±0.039
ST (Ascorbic acid)	0.057±0.000	0.056±0.000

Data expressed as means of triplicates ±sd

(Labar *et al*., 2019). Here, the lowest IC_{50} value was recorded as 0.111 ± 0.001 mg/ml (S1 acetone extracts), and 0.478 ± 0.028 mg/ml (S2 methanol extracts) when compared against the IC50 value of the standard ascorbic acid i.e., 0.057 ± 0.000 mg/ml (Table 4.6).

A strong antioxidant property of methanol (Turkmen *et al.,* 2006), ethanol, and acetone extract of black tea has been previously reported. Varying percent concentration of the solvent resulted in varying antioxidant activity with 50% ethanol and 50% acetone showing maximum antioxidant activity in a mate and black tea where they inferred that the solvent potential in phytochemical extraction can be enhanced or reduced just by altering the concentration percentage of the solvent. The result also highlighted the importance of hot water in extraction where the extracts showed moderate (black tea) and higher (mate tea) antioxidant potential.

4.4.3. Ferric reducing power assay

Acetone extracts (Figure 4.19.a) comparatively showed higher ferric reducing power than methanol (Figure 4.19.b) and the results were almost at par with the standard ascorbic acid. Previously methanol extract of fresh tea leaf has shown the highest reducing power based on the sample state. The reducing power also varied among the sample type i.e., greater ferric reducing power was showed by shoot extract which was followed by young leaves and mature leaves (Chan *et al.,* 2007).

4.4.4. Total Phenol

The highest value of total phenol was recorded in acetone extracts of seed clone (S2) as 37.77 mg GAE/g followed by methanol and ethanol as mentioned in Table 4.7. On the other hand, the lowest value was recorded for chloroform extracts of S2 (0.25 mg GAE/g of total phenols). The minimum

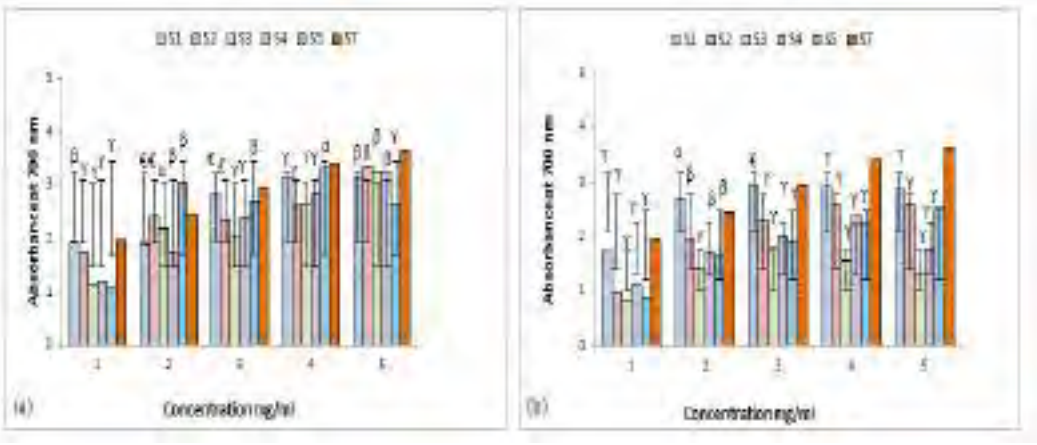


Fig.4.19. Ferric reducing power of acetone extracts (a); methanol extracts (b); p<0.05; β p<0.01; ϒp<0.001; € non-significant

level of phenol was recorded in extracts of non-polar solvents like hexane, benzene, and chloroform. However, this result did not just limit to the varying range of polarity. It varied among the different samples along with varying solvents. Previous works suggest green tea to be a richer source of phenolics than white tea where they reported the potential of 40% aqueous ethanol to extract catechin (Rusak *et al.,* 2008). The estimation of total phenol content using the Folin-Ciocalteu reagent method showed solvents like acetone (50%), N, N-dimethylformamide (DMF) (50%), ethanol (50%), and methanol (50%), to be suitable for extracting total phenols (Turkmen *et al.*, 2006).

4.4.5. Total flavonoid

The highest total flavonoid content

Table 4.7. Determination of total phenol content (TPC) expressed as mg/g GAE

Solvents	S1 (mgGAE/g)	S2 (mgGAE/g)	S3 (mgGAE/g)	S4 (mgGAE/g)	S5 (mgGAE/g)
Hexane	1.61	0.60	3.66	1.00	1.08
Benzene	2.14	0.38	2.86	1.21	1.50
Chloroform	1.77	0.25	1.50	1.45	1.85
Diehylether	3.42	7.18	2.49	2.14	3.34
Ethylacetate	6.41	5.40	3.77	3.69	5.42
Acetone	32.44	37.77	32.06	14.57	18.65
Ethanol	18.25	16.92	17.88	14.25	13.72
Methanol	32.89	31.50	23.82	10.73	11.50
Water	4.04	1.53	2.36	2.52	2.25

Table 4.8. Determination of total flavonoid content (TFC) expressed as mg QE/g

Solvents	S1 (mgQE/g)	S2 (mgQE/g)	S3 (mgQE/g)	S4 (mgQE/g)	S5 (mgQE/g)
Hexane	412.94	446.74	399.42	418.73	537.52
Benzene	406.18	441.91	488.27	401.35	463.16
Chloroform	428.39	482.47	412.94	467.02	477.64
Di ethyl ether	467.99	473.78	433.22	448.67	561.66
Ethyl acetate	470.88	507.58	433.22	424.53	469.92
Acetone	513.76	675.62	617.67	579.04	669.82
Ethanol	528.83	557.80	570.35	506.62	611.88
Methanol	722.94	581.94	545.24	519.17	574.22
Water	424.53	406.18	432.25	423.56	431.29

(Table 4.8) was found in the methanol extracts of S1(722.94 mg QE/g) and the lowest value was recorded that for hexane extracts of S3 (399.42 mg QE/ g).

4.5. Chemical Characterization and bioactivity study

4.5.1. Gas chromatography-mass spectrometry

GC-MS identified crucial bioactive compounds (Appendix D), which could impart antimicrobial properties. The compounds under scrutiny were Phenol, 3,5bis (1,1-dimethyl ethyl), caffeine, and Vitamin E, which could impart antimicrobial properties. GC-MS analysis detected these compounds in the acetone extracts of S3 (Figure 4.20) and methanol extracts of S1 (Supplementary Figure 4.8). The GC-MS data is provided in Appendix D (Table 1– 4). A previous study reports the caffeine-independent antimicrobial activity of coffee extracts against pathogenic strains like *Staphylococcus epidermidis* and *Enterococcus faecalis* (Runti *et al.,* 2015). Contrary to this result, caffeine was also found to enhance or increase the antimicrobial activity of amoxicillin when used against the gram-positive strain i.e., *Staphylococcus aureus* (Esimone *et al.,* 2008). Vitamin E, a known compound with powerful antioxidant activity and the anti-inflammatory property was found to inhibit the growth of mostly gram-positive bacteria rather than gram -negative bacteria due to the presence of lipopolysaccharides in their outer

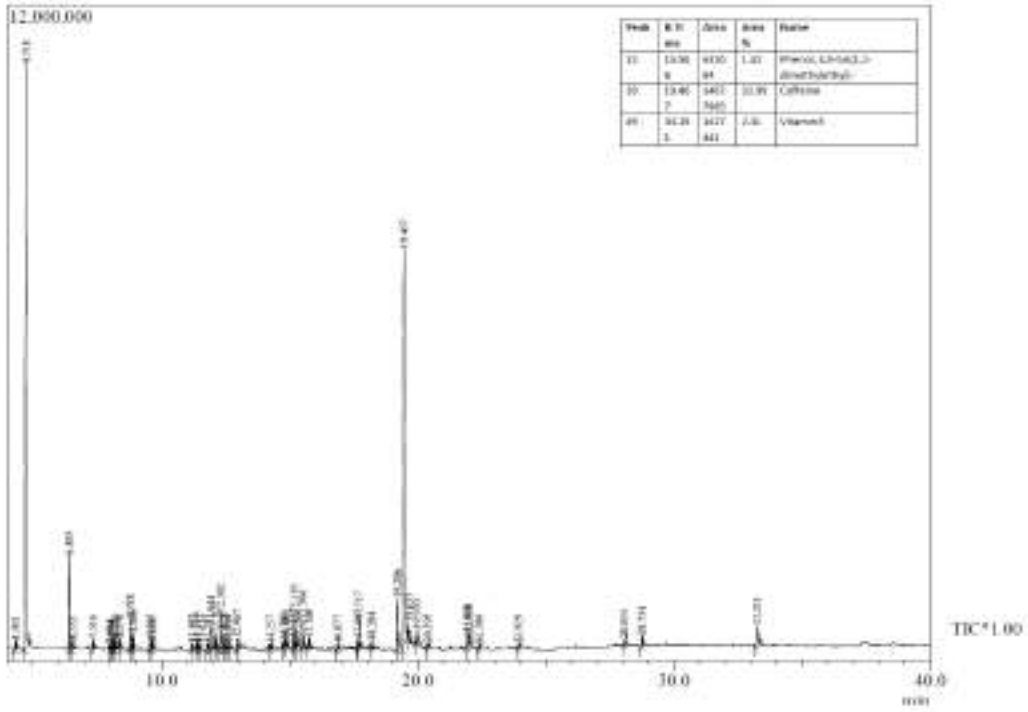


Fig.4.20. Chromatogram of the acetone extracts of AV2 (S3)

membrane (Jehad *et al.,* 2013). The presence of phenol, 3, 5-bis (1, 1-dimethyl ethyl) in GCMS data of *Ulva reticulata* indicated its probable role for imparting anti-microbial effect against *Staphylococcus aureus* and *Escherichia coli* (Dhanya *et al.,* 2016).

4.5.2. Antimicrobial screening

Out of the four bacterial strains studied, acetone and methanol extracts were more effective or bactericidal against the gram-positive *S. aureus* (Figure 4.21). Both the extracts showed minimum or no activity towards the other bacterial strains employed. A broad-spectrum antibiotic i.e., Streptomycin sulfate was used as a standard and the plant extracts were pipetted into each well with volumetric difference. The volumetric increase in the sample ultimately increased the inhibition, which was at par with 100 µl of the standard (Supplementary Figure 4.9). The minimum Inhibition concentration or the MIC value was found to be 4mg/ml for acetone extracts (Supplementary Figure 4.10) and 8mg/ml for the methanol extracts (Figure 4.22).

Previous reports suggest tea extract to be effective against various bacterial strains with bactericidal activity against *Staphylococcus aureus* and *Yersinia enterocolitica* (Yam *et al.,* 1997). However, our study reported acetone and methanol extract to show potent antimicrobial and bactericidal activity against the studied strain, *S. aureus*.

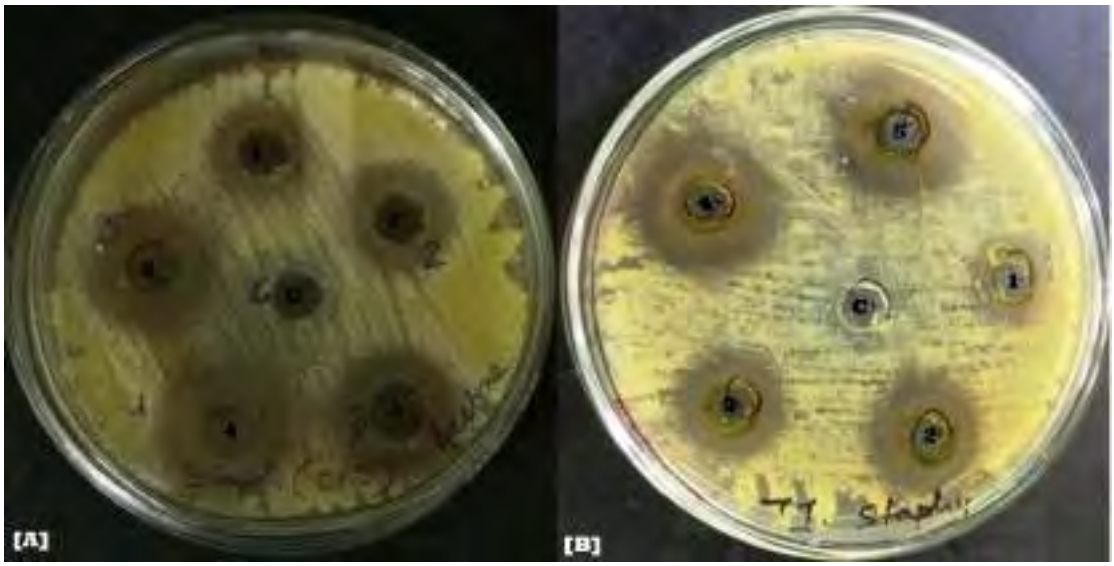


Fig.4.21. Antimicrobial activity of tea extracts against *S. aureus* [A] Acetone [B] Methanol; 1-5 different tea clones; C sterility control (DMSO)

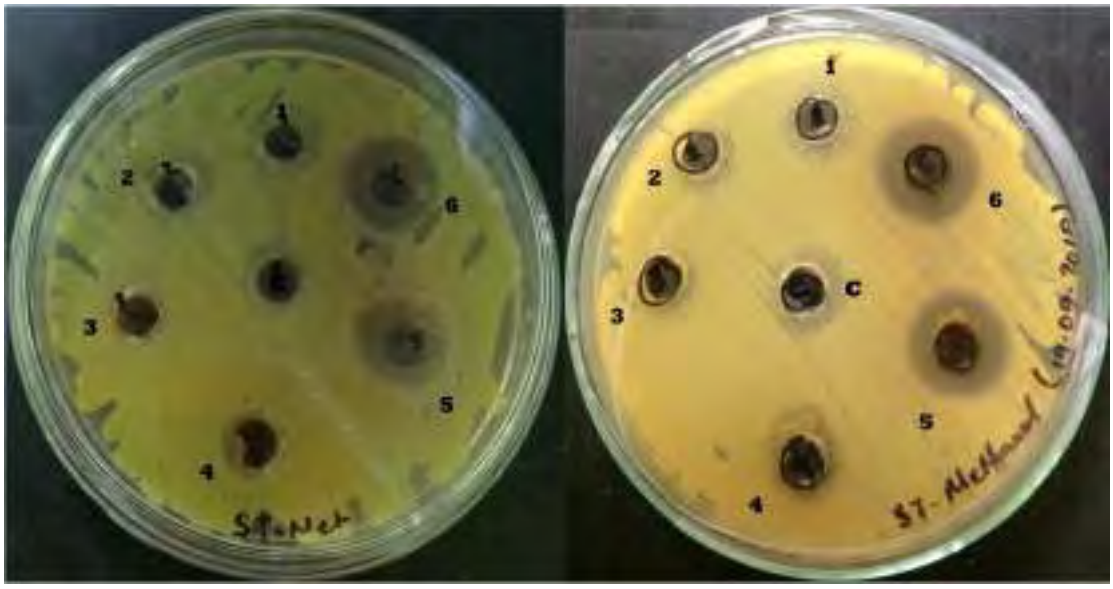


Fig.4.22.MIC value determination from antimicrobial activity of tea extracts against *S. aureus* using agar well diffusion method (C) Methanol extracts of S4(D) Methanol extracts of S7; 1-5 different concentrations 1-0.25mg/ml, 2– 05.mg/ml, 3– 1mg/ml, 4– 4mg/ml, 5– 8 mg/m,6– 10mg/ml; C sterility control DMSO

Both the extracts showed minimum or negligible activity against other strains examined.

4.5.3. In silico analysis

Compounds like caffeine and Vitamin E did not show effective binding with the gyrase protein. The compound Phenol, 3,5bis (1,1-dimethylethyl) showed significant binding to 5CDN with the binding affinity of -7.2 kcal/ mol (Figure 4.23). This result showed the significant effect of phenol with the gyrase protein.

Phenol is a well-known antimicrobial

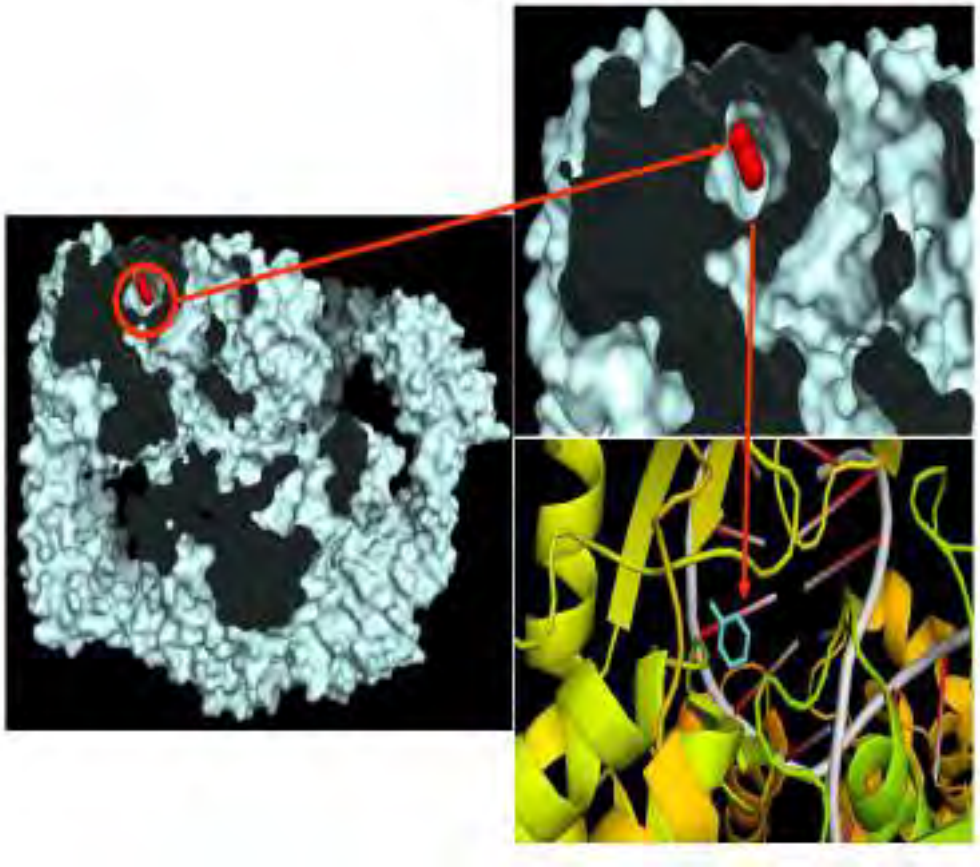


Fig.4.23. *In silico* docking of *S.aureus* DNA gyrase (pdb id: 5CDN) showing significant binding affinity -7.2 kcal/mol with phenol

agent with both bacteriostatic and bactericidal effects which targets the bacterial cell wall, cytoplasmic membrane, organelles, cytosol, and spores (Sabbinenj, 2016). The high penetrating power of the phenol into any organic matter may render its easy access or entry inside the cytoplasmic membrane of bacteria via passive diffusion and active transport mechanisms. Along with its easy permeability, phenol is also a strong oxidizing agent which converts akin to oxide and gas, eventually leading to destabilization of the spores (Sabbinenj, 2016). Therefore, we can infer that the binding of phenol to gyrase A protein may exert a strong antimicrobial activity.

4.6. Purple Tea and its utility in green synthesis of nanoparticles

4.6.1. Silver nanoparticles (AgNPs)

A green route for the rapid synthesis of silver nanoparticles employing extract of purple tea (*C. sinensis*) has been established in our study. The phytochemicals present in the purple tea acts as both reducing as well as a capping agent since earlier reports suggest that the plant extracts contain different phytochemicals like

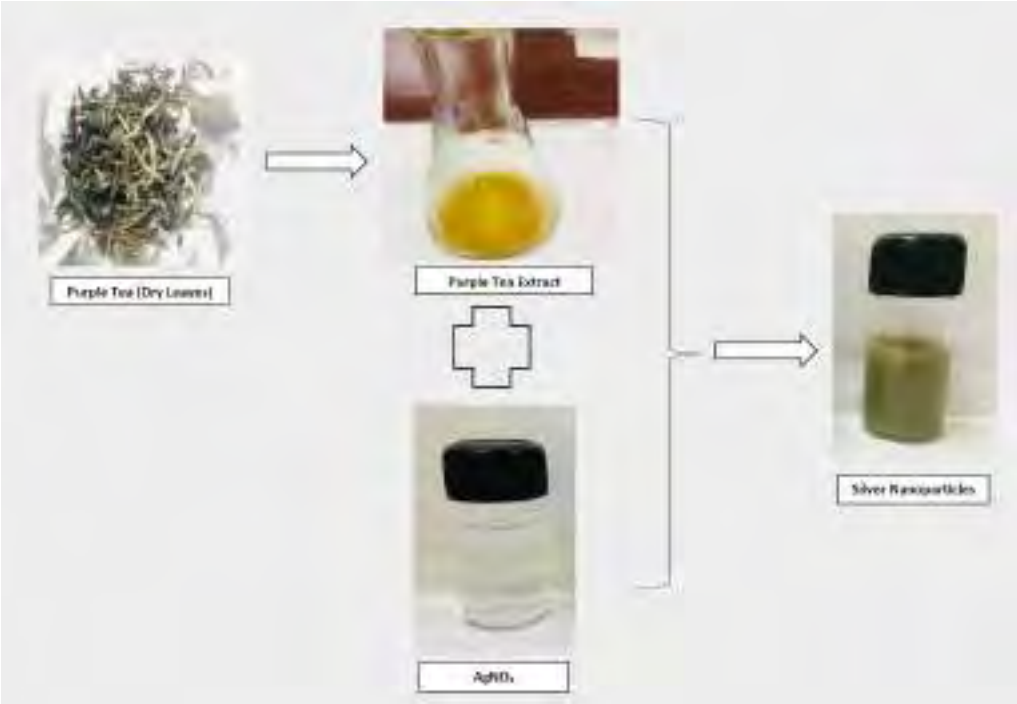


Fig.4.24. Schematic diagram of synthesis of silver nanoparticles from purple tea extract

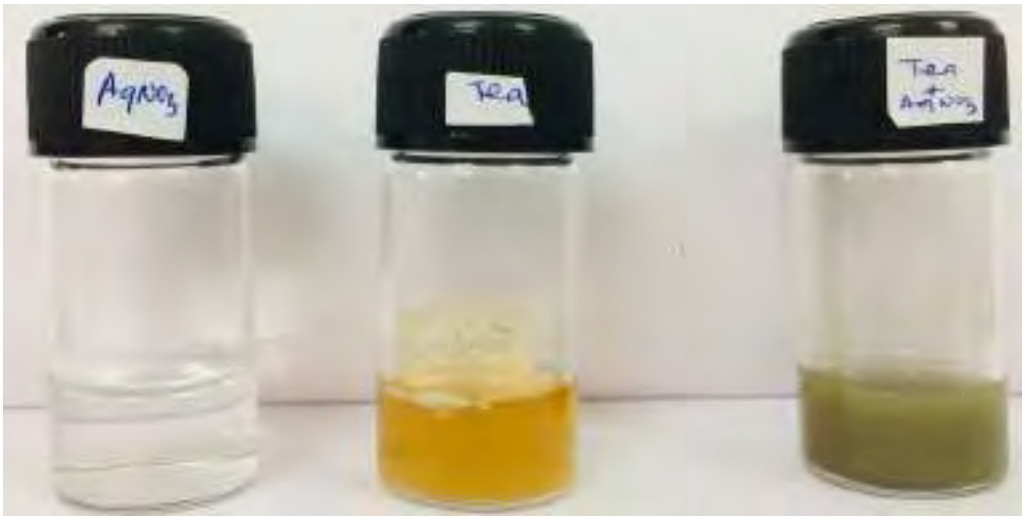


Fig.4.25. Synthesis of silver nanoparticles characterized by change in color

flavonoid, terpenoid, phenol derivatives, etc. which shows the bioreduction property in the presence of metal salts (Kumar *et al.*, 2017). The formation of silver nanoparticles synthesized from purple tea extract was primarily screened by the change in color of the reaction mixture (Figure 4.24). The color of the solution changed from yellow to green after 5 minutes of incubation (Figure 4.25). A previous study on the synthesis of silver nanoparticles using pu-erh tea extract also reports the formation of silver nanoparticles by a change in color of the solution to dark brown

(Loo *et al.*, 2012). The formation or synthesis of AgNPs was further confirmed by its characterization using UV-visible spectroscopy (UV-vis), X-ray diffraction (XRD), and Scanning electron microscopy (SEM).

4.6.1.1. Characterization of silver nanoparticles

The reaction parameters like reaction temperature, the concentration of reactants used in the solution, time of reaction, pH, etc. have to be controlled and optimized to obtain the desired monodispersed nanoparticles with tunable size and morphology (Tripathy *et al.,* 2010).

4.6.1.1.1. Effect of varying molar concentration of $AgNO_3$

Firstly, the effect of different molar concentrations was studied to explore the optimum synthesis of silver nanoparticles. A varying molar concentration of silver nitrate was screened in the range of 1 mM to 10mM (Figure 4.26). The reaction mixture was exposed to bright sunlight for 15 minutes. The reaction mixtures containing different metal ion ($AgNO_3$) concentrations; 1 mM, 2 mM, 3 mM, 4 mM, 5 mM, 6 mM, 7 mM, 8 mM, 9 mM and 10 mM produced (Surface Plasmon Resonance) SPR band at 414 nm, 417 nm, 418 nm, 431 nm, 430 nm, 428 nm, 437 nm, 436 nm, 439 nm and 442 nm respectively. It was observed that the SPR band of 10mM was the least intense as compared to others (Figure 4.26). The steady increase in SPR band intensity up to 3mM revealed that the synthesis of silver

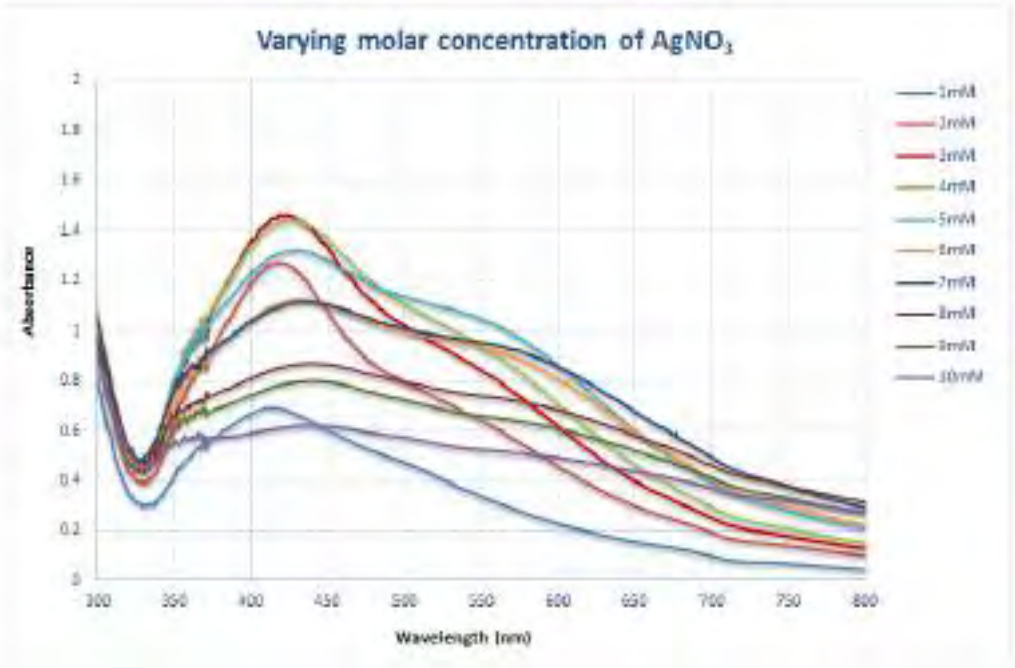


Fig.4.26.A varying molar concentration of silver nitrate was screened for AgNPs synthesis in the range of 1 mM -10mM

nanoparticles increased with the increasing concentration of $AgNO_3$. Synthesis of AgNPs increased while increasing the $AgNO_3$ concentration up to 2 mmol L^{-1} that was evident from a regular increase in SPR band intensity (Philip, 2010; Bar *et al.*, 2009). A redshift in the SPR band from 453 nm to 459 nm is directed towards the increase in the size of the nanoparticle (Jagtap and Bapat, 2013). However, the intensity of the SPR band (431nm) of 4mM decreased when compared against 3mM (418nm). The intensity of the SPR band of 3mM increased but following that the intensity of the band decreased.

A redshift towards a higher wavelength was observed from 418 to 442 nm with an increase in the concentration of $AgNO_3$. The redshift towards a higher wavelength with an increase in the concentration of $AgNO_3$ directly relates to an increase in the size of the silver nanoparticles (Mock *et al.*, 2002). Thus, 3mM was chosen as the optimal molar concentration of $AgNO_3$ for this current study. The obtained AgNPs were found to be stable only for 24 hours and after that settled at the bottom due to agglomeration. The previous study reports such agglomeration during green synthesis of nanoparticles, which was also a case, reported in the previous study (Kumar *et al.,* 2017).

4.6.1.1.2. Effect of sunlight exposure

When the reaction mixture containing $AgNO_3$ and plant extract was exposed to bright sunlight, the color of the solution changed from colorless to brown rapidly within 10 minutes of incubation, and as time advanced the color became darker. However, the

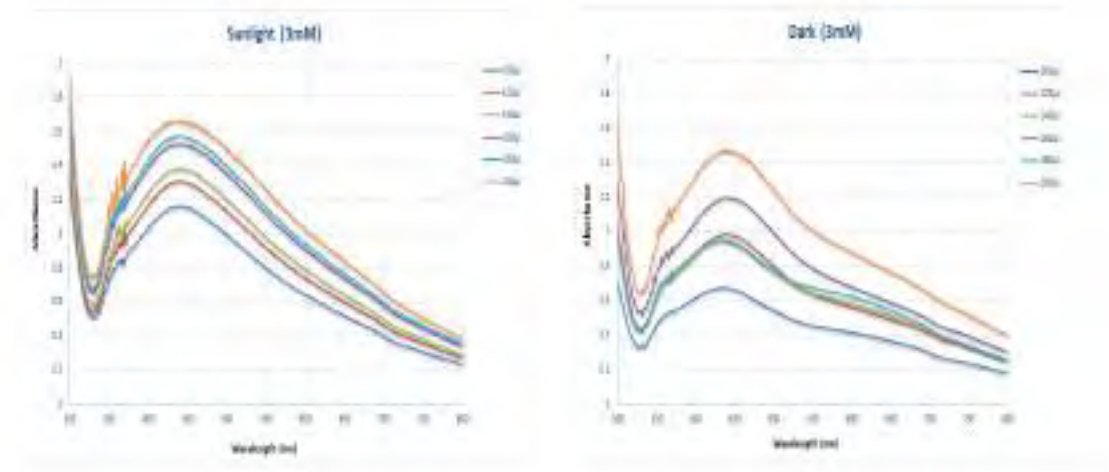


Fig.4.27.UV-vis absorption spectrum of purple tea extract mediated AgNPs synthesized using varying conditions of light

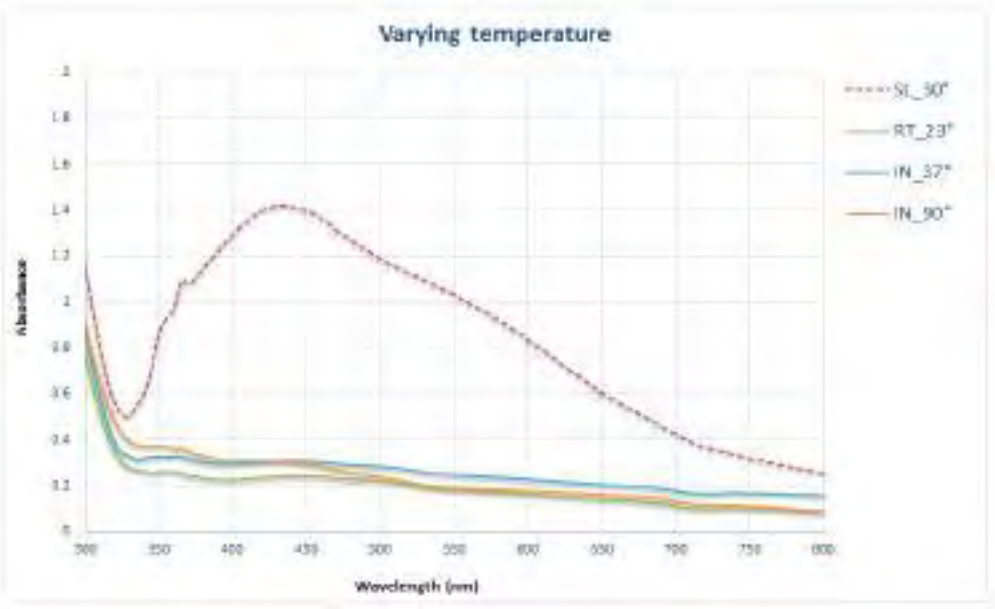


Fig.4.28.UV-vis absorption spectrum of purple tea extract mediated AgNPs synthesized using varying conditions of temperature

reaction mixture kept in the dark also rendered the same color but the change was not rapid as compared to the sample exposed to the bright sunlight. This result highlights the importance of the photocatalytic action of sunlight on the synthesis of AgNPs (Kumar *et al.,* 2017). UV–vis spectroscopy further monitored the reaction mixture kept at varying conditions of light and darkness (Figure 4.27).

4.6.1.1.3. Effect of varying temperature

A different range of temperature was used to study the effect of temperature on AgNPs synthesis. The sample or the reaction mixture was exposed to bright sunlight (30°C), room temperature of 23°C and incubator temperatures of 37°C and 90°C. The reaction mixture exposed to varying temperature conditions for 1 hour was monitored using UV Vis spectroscopy. The sample kept at 30°C (under sunlight) produced the sharper SPR band at 430 nm whereas the other reaction mixture showed a blue shift towards a lower wavelength (304nm) with a decrease in the SPR band intensity (Figure 4.28).

4.6.1.2. Structural Analysis

The structural properties of the sample were studied by carrying out the (X-Ray Diffraction) XRD analysis employing Rigaku Ultima IV. The crystalline profile of our sample clearly shows the peaks of Silver, the major peaks (Figure 4.29) occurring at the diffraction angle 38.03°, 44.15°, and 64.40° correspond to the Ag [111], Ag [002], and Ag [220] respectively. These values of the diffraction angles for the Ag peaks closely match with the PDF card number 00-004-0783

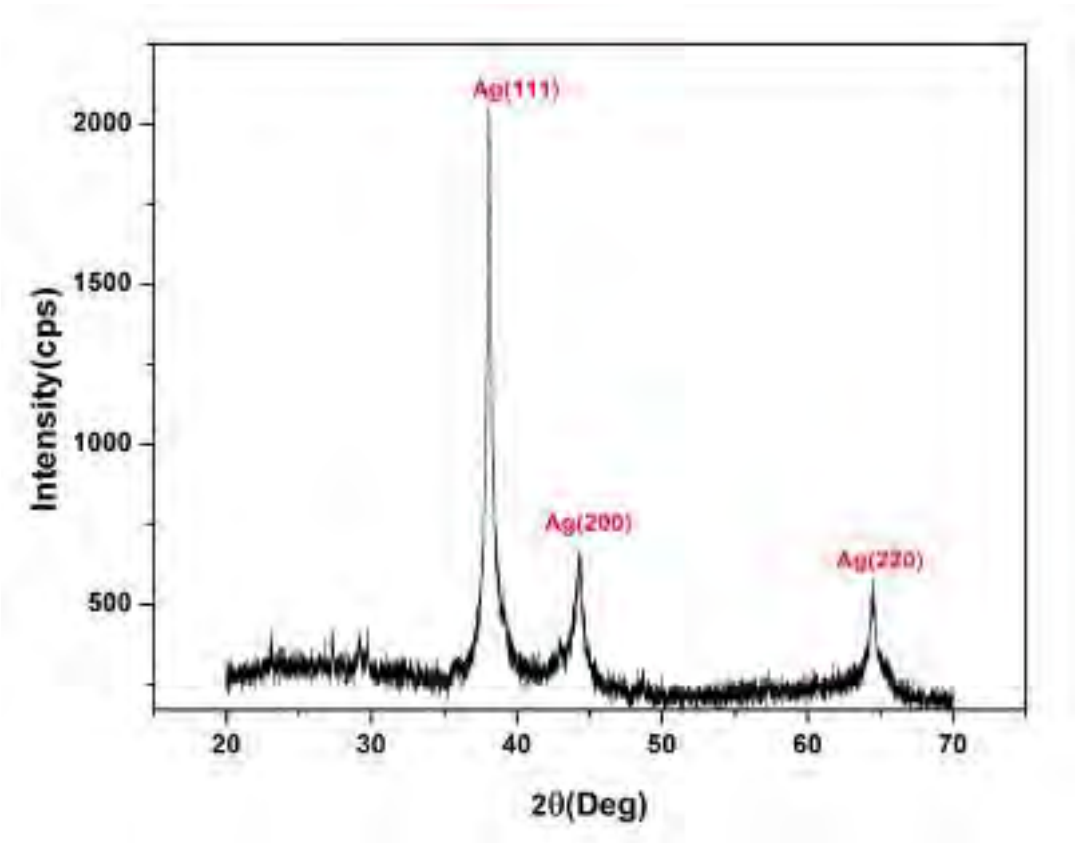


Fig.4.29.XRD pattern of synthesized silver nanoparticle. The crystalline profile of the sample clearly shows the peaks of Silver, the major peaks occurring at the diffraction angle 38.03°, 44.15° and 64.40° corresponding to the Ag [111], Ag [002] and Ag [220]

indicating the formation of silver nanoparticles. The size of the nanoparticle was determined using the Debye Scherrer's formula given as,

$$D = \frac{K\lambda}{\beta\cos(\theta)}$$

Where,

K- Boltzmann Constant

λ – Wavelength of radiation

β – FWHM value of the corresponding peak

θ- Diffraction angle

The size of the nanoparticle hence determined from our analysis was of the order of 38.4 nm. Thus, we can infer from the XRD study, that pure Ag nanoparticles are formed with the particle size of the order of 38 nm.

4.6.1.3. Scanning electron microscopy (SEM)

The SEM imaging revealed varying morphology with the majority of silver nanoparticles to be non-spherical or either irregular in shape with a rough surface, and few were rod-shaped (Supplementary Figure 4.11). The

particle size ranged from 10nm to 40 nm. The particle size of a few exceeded the range, which is probably due to the overlapping of the synthesized nanoparticles. The particles were agglomerated due to crosslinking (Shankar *et al.*, 2017) or solvent evaporation of solvent during the preparation of the sample (Jagtap and Bapat 2013).

4.6.1.4. Antimicrobial property

The synthesized AgNPs were found to show antimicrobial properties against *Staphylococcus aureus* and *Bacillus subtilis* (Figure 4.30) out of the four strains studied namely *Staphylococcus aureus*, *Bacillus subtilis, Escherichia coli*, and *Klebsiella pneumonia.* The plant extract itself showed antimicrobial activity against the gram-positive strains and showed minimum or no activity against the gram-negative ones. However, the antimicrobial activity enhanced with the formation of silver nanoparticles (AgNPs) using the purple tea extract.

4.6.2. Zinc oxide nanoparticles (ZnO NPs)

The schematic diagram of ZnO nanoparticle synthesis is given in Figure 4.31. The brown color of the ZnO NPs arises due to the capping action of purple tea biomolecules on the surface of the nanoparticles. The formation of ZnO nanoparticles synthesized using leaf extract of purple tea was initially monitored by the change in color occurring during the reaction period (Figure 4.32). Upon addition of tea extract to the reaction mixture, the color slowly changes to pale white which then progresses to yellowish-green, and finally, the color of the end product was brown. The

Fig.4.30. Antimicrobial activity of purple tea extract mediated AgNPs:
(a) *S. aureus*, (b) *E. coli*

Fig.4.31. Schematic diagram of synthesis of Zinc Oxide nanoparticles (ZnO NPs) from purple tea extract

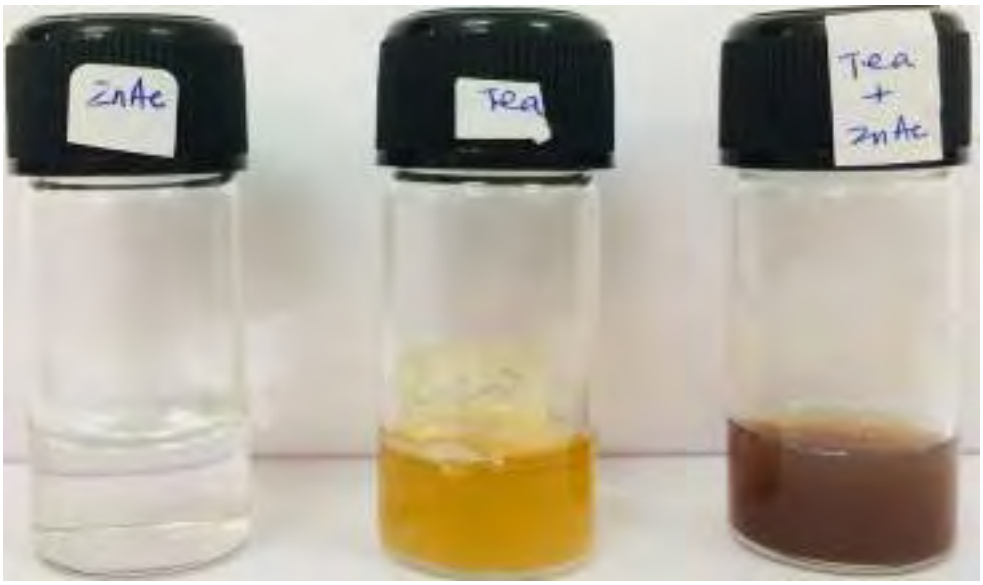


Fig.4.32. Synthesis of Zinc oxide nanoparticles characterized by change in color

color change arises due to the excitation of surface plasmon resonance in the metal nanoparticles, indicating the formation of ZnO nanoparticles.

4.6.2.1. ZnONPs characterization

4.6.2.1.1.UV–characterization

The UV-Vis spectrum of the ZnO nanoparticle is provided in Figure 4.33. The synthesis of ZnO product in nanoscale was confirmed by the highly blue-shifted, maximum absorption peak occurring around 273 nm. The zinc oxide synthesized using leaf extract of *Sesbania grandiflora* (Rajendran and Sengodan, 2017) exhibited UV-visible absorption peaks at 235 nm which is nearly similar to our results. Another work related to

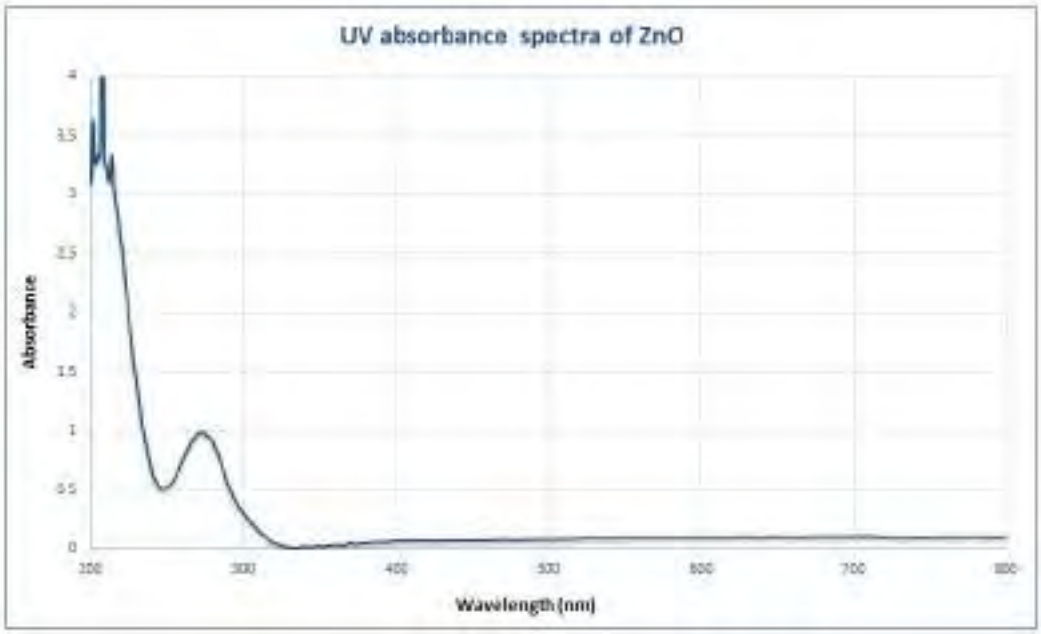


Fig.4.33.UV-vis absorption spectrum of Zinc oxide nanoparticle synthesized from purple tea extract

ZnO synthesis using the microemulsion method also reports a similar kind of result (Kumar and Rani, 2013). Usually, maximum absorption for bulk ZnO approximately occurs around 385 nm. The decrease in absorption may result from agglomeration and settling down of nanoparticles in the cuvette. However, we conclude from our analysis that the synthesized zinc nanoparticles were photosensitive in the UV region since they showed an intense absorbance peak of ~270 (Savi *et al.*, 2014). When the concentration of the plant extract was increased, the absorbance value also increased where the peak intensified at 400nm. The increased concentration of the phytochemicals in the extract increases the efficacy of the extract to reduce the precursor present in the solution rapidly which eventually also enhances the synthesis of nanoparticles as well as increases the absorbance value (Isaac *et al.*, 2013).

4.6.2.1.2. Structural Analysis using XRD (X-ray diffraction)

To study the crystalline properties of the sample, we performed the (X-Ray Diffraction) XRD analysis of the sample, employing Rigaku Ultima IV. The X-ray diffraction profile of the sample shows a crystalline nature. The detailed analysis of the peaks indicates the presence of ZnO, the occurrence of which is visibly clear in the sample annealed at temperature 400°C (Figure 4.34). The major peaks [100], [002], and [101] of ZnO occur at the diffraction angles 31.69º and 34.77º and 36.1º respectively. These values of

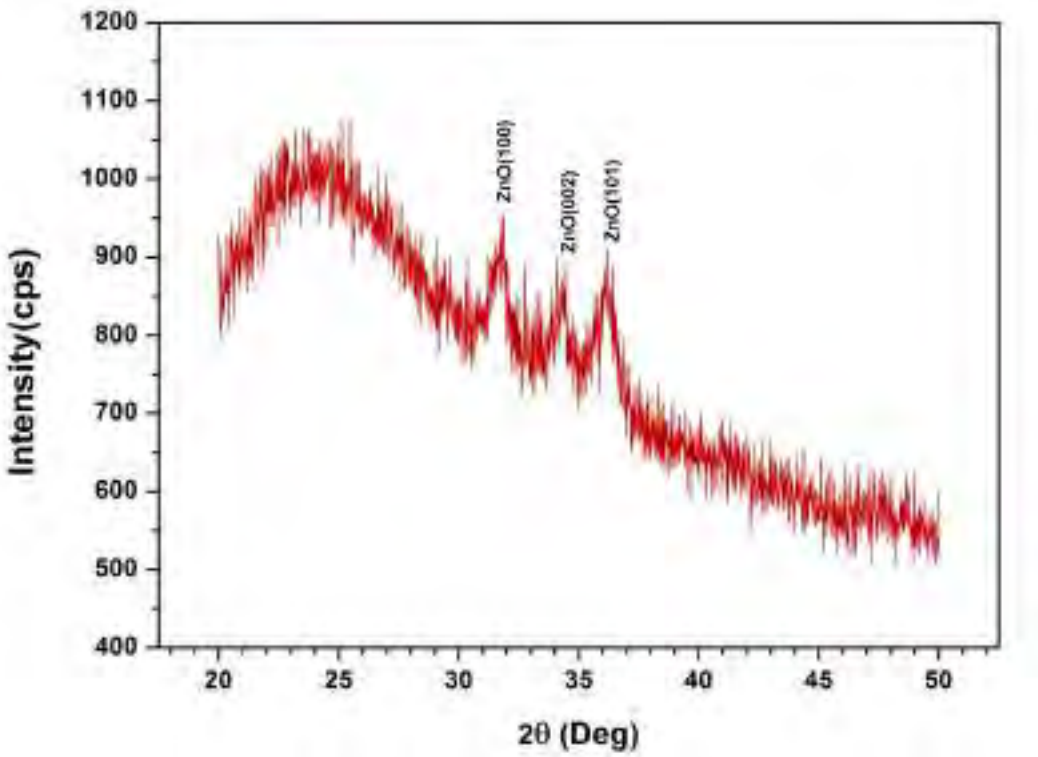


Fig.4.34.XRD pattern of synthesized Zno NPs (Calcination temperature - 400° C). The major peaks [100], [002] and [101] of ZnO occur at the diffraction angles

the diffraction angles for the ZnO peaks closely match with the PDF card number 01-080-0074 indicating the formation of pure ZnO nanoparticles. The size of the nanoparticle was determined using the Debye Scherrer's formula given as,

$$D = \frac{K\lambda}{\beta\cos(\theta)}$$

Where,

K- Boltzmann Constant

λ – Wavelength of radiation

β – FWHM value of the corresponding peak

θ- Diffraction angle

The size of the nanoparticle hence determined from our analysis was of the order of 22.4 nm. Though the samples annealed at 400°C show a clear formation of ZnO nanoparticles, the samples annealed at 100°C (Figure 4.35) shows the occurrence of only the [002] peak with a minimum intensity. The profile also shows a broad background, somewhat like a hump, along with the crystalline peaks, the possible reason for which could be due to the presence of the organic compound in the material. The broad background is minimized slightly at higher temperatures probably due to the removal of the organic compound at a higher temperature.

The occurrence of this arch-like peak

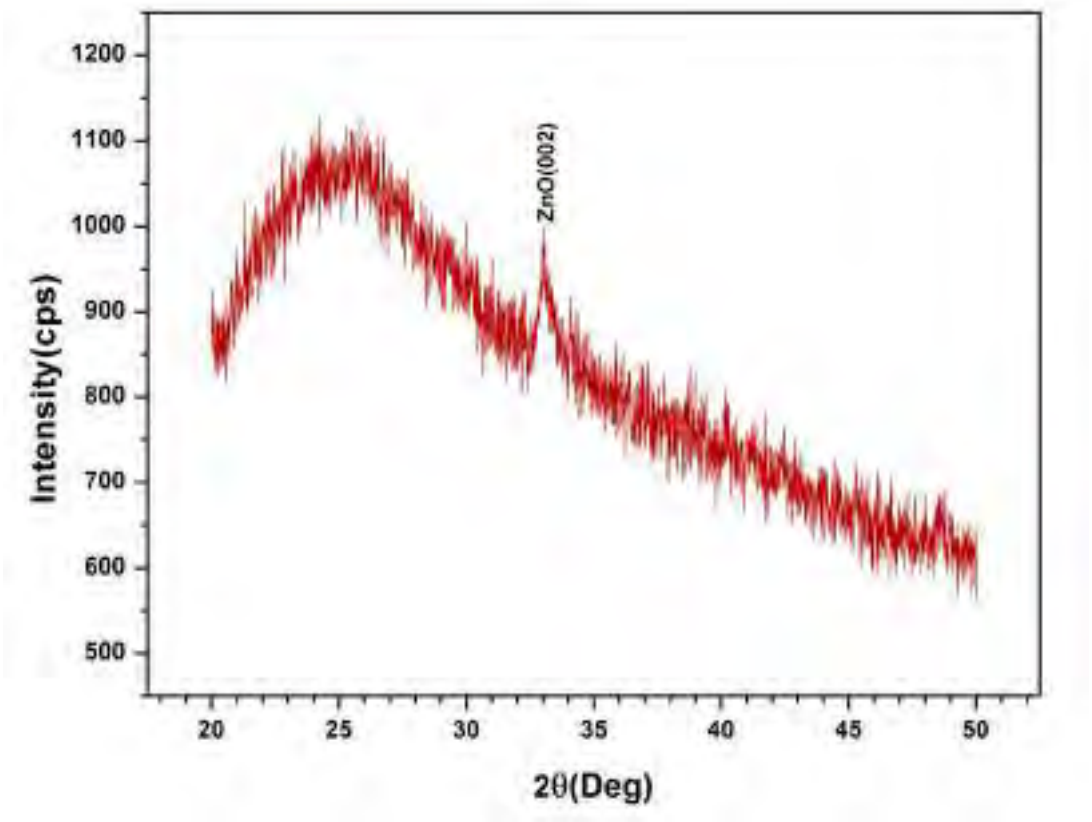


Fig.4.35.XRD pattern of synthesized Zno NPs (Calcination temperature - 100° C). Single peak [002] of ZnO occur at the diffraction angle 34.77º

in the profile also indicates the presence of tea in the sample as reported in various works.

4.6.2.1.3. Scanning electron microscopy (SEM)

SEM was carried out to determine the surface morphology of the synthesized ZnO. The micrographs (Figure 4.36) showed network formation or cross-linking during the synthesis of ZnO. This indicates agglomeration (Shankar *et al.*, 2017) during the process of synthesis. Moreover, the synthesized ZnO was found to be spherical with an approximate size of around 15-25nm.

4.6.2.2. Antimicrobial property

The antimicrobial activity of the plant extract (purple tea extract), zinc acetate, and zinc nanoparticle (ZnO) was screened against bacterial strains including both the gram-positive (*Staphylococcus aureus*, *Bacillus subtilis*) and gram-negative strains (*Escherichia coli*, *Klebsiella pneumoniae*). We found that zinc acetate itself has an antimicrobial property towards all the investigated strains while our plant extract was effective against the strains like *Staphylococcus aureus* and *Bacillus subtilis* and we solely aimed to see the

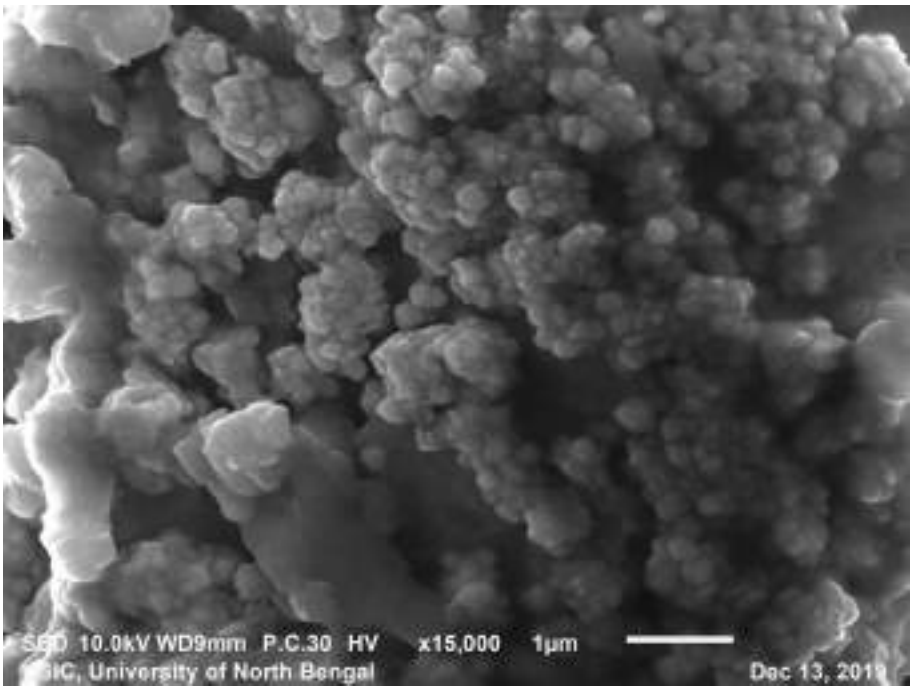


Fig.4.36.SEM image of ZnONPs synthesized using aqueous purple tea extract

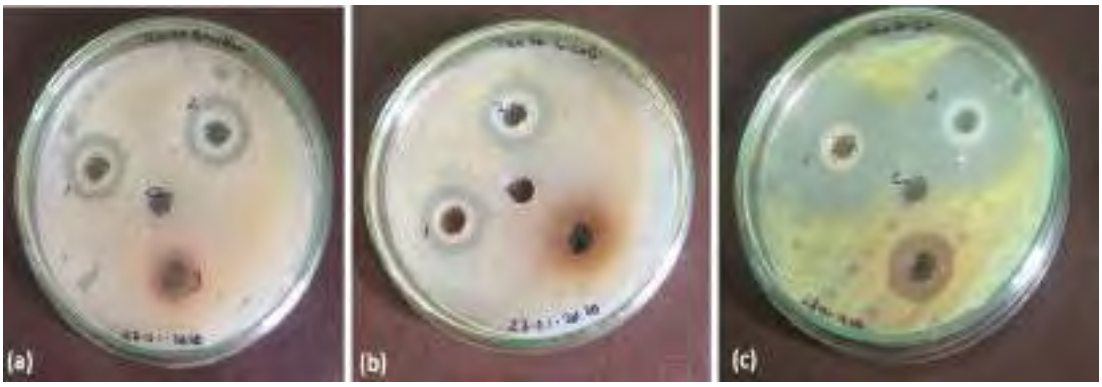


Fig.4.37. Antimicrobial activity of purple tea extract mediated ZnONPs: (a) *Bacillus subtilis*, (b) *Escherichia coli*, (c) *Staphylococcus aureus*

antimicrobial activity of the synthesized ZnO. The inhibition zone of the ZnO was bigger when compared solely to Zinc acetate and plant extract. Thus, we found that the antimicrobial property was enhanced in the synthesized zinc nanoparticles (Figure 4.37). The synthesized ZnO showed antimicrobial activity against all the strains but the inhibition property towards *Escherichia coli* was due to the antimicrobial property of zinc acetate alone since the plant extract did not show significant antimicrobial property towards *E. coli*.

4.7. Tea Bioinformatics

4.7.1. Factors governing codon usage behavior

Availability of whole-genome sequence of the reference genome of

Camellia sinensis var. *sinensis* cv. Shuchazao (Xia *et al*., 2020) allowed a detailed study of its codon usage and amino acid usage patterns. *Camellia sinensis* was found to be AT-rich (AT 55% and GC 45%). The tea genome preferred AT-rich codons over the GC ones. The overall codon usage pattern revealed that 15 amino acids were coded by AT-rich codons with RSCU > 1 as marked in bold (Table 4.9). Out of 27 optimal codons (RSCU>1), 19 optimal codons were AT-rich. Out of the 61 codons, 29 codons were AT-rich out of which 22 codons were found to end with either Adenine or Thymine. Usage of AT-rich codons like AAG, GAA, GAT, AAT, AAA, GTT, GCT, TTG, CTT, TCT, and ATG was quite perceptible from the given Table 4.9. Preference for AT-rich codons was evident from codon usage and amino acid usage analysis. Axis 1 of Relative synonymous codon usage (RSCU) was found to exhibit a significantly high and positive correlation with A3 and T3 indices (Table 4.10) which pointed towards a mighty effect of the AT nucleotide compositional constraint in shaping the codon usage pattern of *Camellia sinensis*. Various important correlation patterns between the codon usage parameters pointed out the importance of nucleotide compositional constraint in directing codon usage variation. Cases of nucleotide composition governing the codon usage have been reported in unicellular as well as multicellular eukaryotes like *Entamoeba histolytica* (Romero *et al*., 2000), different human genes (Karlin and Mrazek, 1996), and nuclear genes of *Oryza sativa*, *Zea mays*, and *Tuber aestivum* (Liu and Xue, 2005).

GC3 versus Nc plots were plotted where we considered total protein-coding genes (TPC) along with potentially highly expressed genes (PHX) and potentially lowly expressed genes (PLX) as shown in Figure 4.38. We also plotted a GC3 versus Nc plot (Figure 4.39) considering total protein-coding genes along with genes of different metabolic pathways rendering synthesis of important metabolites like flavonoid, phenylpropanoid, flavone and flavonol, caffeine biosynthesis, and streptomycin biosynthesis. GC3 versus Nc plot has been suggested to be effective in the proper assessment of probable factors that govern codon usage patterns among genes and genomes (Wright 1990). Previous studies have stated that all the genes would fall on a continuous Nc plot if the codon usage is solely governed by nucleotide compositional constraint (Wright 1990). The fact that nucleotide compositional constraint was having a pronounced effect on the codon usage behavior of *C. sinensis* was evident

Table 4.9. Overall codon usage table of *C. sinensis*

AA	CODON	N	RSCU	AA	CODON	N	RSCU
Phe	**UUU**	367773	**1.13**	**Ser**	**UCU**	343826	**1.51**
	UUC	284965	0.87		UCC	190052	0.84
Leu	UUA	181203	0.72		**UCA**	308178	**1.36**
	UUG	405410	**1.6**		UCG	119393	0.53
	CUU	351724	**1.39**		**AGU**	228392	**1.01**
	CUC	241117	0.95		AGC	172260	0.76
	CUA	158489	0.63	**Thr**	**ACU**	259137	**1.38**
	CUG	181791	0.72		ACC	173905	0.93
Ile	**AUU**	393437	**1.43**		**ACA**	245262	**1.31**
	AUC	229411	0.83		ACG	73037	0.39
	AUA	202753	0.74	Ala	GCU	374391	1.53
Val	**GUU**	385936	**1.54**		GCC	191336	0.78
	GUC	174705	0.7		GCA	319405	1.31
	GUA	153571	0.61		GCG	92867	0.38
	GUG	290487	1.16	**Arg**	CGU	92103	0.71
Pro	CCU	256840	1.38		CGC	69350	0.53
	CCC	124824	0.67		CGA	97952	0.75
	CCA	271364	1.46		CGG	8314	0.6
	CCG	88950	0.48		**AGA**	237445	**1.83**
Gly	GGU	308462	1.22		AGG	4585	1.57
	GGC	178355	0.7	**His**	**CAU**	225773	**1.22**
	GGA	309165	1.22		CAC	144199	0.78
	GGG	16733	0.86	**Gln**	**CAA**	319649	**1.16**
Asn	**AAU**	424258	**1.22**		CAG	232013	0.84
	AAC	271869	0.78	**Asp**	**GAU**	537901	**1.37**
Lys	AAA	425374	0.95		GAC	249058	0.63
	AAG	66503	**1.05**	**Glu**	**GAA**	511574	**1.06**
Tyr	**UAU**	239990	**1.15**		GAG	451716	0.94
	UAC	176468	0.85				
Cys	**UGU**	165325	**1.12**				
	UGC	130754	0.88				

Bold letters represent the amino acids preferentially coded by AT rich codons
RSCU cumulative RSCU values by summation of RSCU of all genes, *N* number of codons, *AA* amino acids

from the GC3 versus Nc plots. However, many genes were found to position well below the continuous Nc curve suggesting the influence of factors other than compositional bias. Apart from the role of nucleotide compositional constraint, the impact of translational selection on codon usage patterns was apparent from the RSCU based scatter plots and GC3 versus NC plots. It has already been reported that if translational selection acts on a gene, the aforesaid gene lies distantly below the continuous plot (Nair *et al*., 2012). Efficient translation, under the deep influence of natural selection, was most prevalent on PHX genes. Codon usage variation due to the influence of translational selection has been observed previously in organisms like *Caenorhabditis* (Moriyama and Powell, 1997).

Multivariate statistical analysis (correspondence analysis) of RSCU

Table 4.10. Correlation analysis between different parameters of codon and amino acid usage obtained from SPSS

	T3s	C3s	A3s	G3s	CAI	CBI	Fop	GC3s	GC	L_sym	L_aa	Gravy	Aromo	MeanCost	Axis1_RSCU	Axis2_RSCU	Axis3_RSCU	Axis4_RSCU	Axis1_RAAU	Axis2_RAAU	Axis3_RAAU	Axis4_RAAU
T3s	1.000	-.758**	.293**	-.327**	.814**	-.139**	-.061**	-.832**	-.665**	.212**	.212**	-.039**	.034**	.037**	.789**	-.061**	.329**	.108**	.084**	.138**	-0.009	-.190**
C3s	-.758**	1.000	-.523**	.033**	-.702**	.353**	.332**	.825**	.608**	-.217**	-.217**	.077**	.150**	-.052**	-.867**	-.336**	-.198**	.014**	-.214**	-.103**	.114**	.121**
A3s	.293**	-.523**	1.000	-.461**	.567**	-.199**	-.104**	-.730**	-.675**	.151**	.152**	-.192**	-.048**	.178**	.685**	-.117**	-.218**	-.342**	.234**	.064**	.157**	-.138**
G3s	-.327**	.033**	-.461**	1.000	-.286**	-.220**	-.211**	.561**	.353**	-.095**	-.094**	-.209**	-.132**	.184**	-.344**	.696**	.018**	.226**	.384**	.102**	-.190**	.043**
CAI	.814**	-.702**	.567**	-.286**	1.000	-.032**	.107**	-.809**	-.837**	.183**	.183**	-.171**	.058**	.271**	.846**	-.080**	.013*	.015**	.269**	.228**	.130**	-.262**
CBI	-.139**	.353**	-.199**	-.220**	-.032**	1.000	.943**	.180**	.146**	-.093**	-.094**	.080**	0.003	-.102**	-.213**	-.320**	-.197**	-0.010	-.139**	-.118**	.061**	.071**
Fop	-.061**	.332**	-.104**	-.211**	.107**	.943**	1.000	.120**	-.036**	-.089**	-.090**	.047**	.066**	.037**	-.138**	-.331**	-.238**	-.012*	-.039**	.019**	.106**	.055**
GC3s	-.832**	.825**	-.730**	.561**	-.809**	.180**	.120**	1.000	.780**	-.232**	-.232**	0.003	0.008	-.036**	-.923**	.125**	-.125**	.134**	-.038**	-.095**	-.053**	.163**
GC	-.665**	.608**	-.675**	.353**	-.837**	.146**	-.036**	.780**	1.000	-.205**	-.207**	-.071**	-.184**	-.485**	-.792**	.113**	.075**	.102**	-.154**	-.452**	-.369**	.234**
L_sym	.212**	-.217**	.151**	-.095**	.183**	-.093**	-.089**	-.232**	-.205**	1.000	1.000**	-.016**	-.053**	0.010	.217**	0.001	.081**	-.040**	.035**	.027**	.087**	-.131**
L_aa	.212**	-.217**	.152**	-.094**	.183**	-.094**	-.090**	-.232**	-.207**	1.000**	1.000	-.012*	-.047**	.015**	.218**	0.002	.080**	-.039**	.033**	.033**	.086**	-.128**
Gravy	-.039**	.077**	-.192**	-.209**	-.171**	.080**	.047**	0.003	-.071**	-.016**	-.012*	1.000	.379**	.136**	-.028**	-.134**	.023**	-.024**	-.680**	.536**	.059**	.278**
Aromo	.034**	.150**	-.048**	-.132**	.058**	0.003	.066**	0.008	-.184**	-.053**	-.047**	.379**	1.000	.417**	-.016**	-.109**	-.043**	-.014**	-.505**	.358**	.097**	.133**
MeanCost	.037**	-.052**	.178**	.184**	.271**	-.102**	.037**	-.036**	-.485**	0.010	.015**	.136**	.417**	1.000	.130**	-.024**	-.084**	-0.009	.279**	.683**	.355**	.179**
Axis1_RSCU	.789**	-.867**	.685**	-.344**	.846**	-.213**	-.138**	-.923**	-.792**	.217**	.218**	-.028**	-.016**	.130**	1.000	0.000	0.000	0.000	.134**	.185**	.021**	-.179**
Axis2_RSCU	-.061**	-.336**	-.117**	.696**	-.080**	-.320**	-.331**	.125**	.113**	0.001	0.002	-.134**	-.109**	-.024**	0.000	1.000	0.000	0.000	.181**	-0.009	-.188**	-.045**
Axis3_RSCU	.329**	-.198**	-.218**	.018**	.013*	-.197**	-.238**	-.125**	.075**	.081**	.080**	.023**	-.043**	-.084**	0.000	0.000	1.000	0.000	-.019**	-.051**	-.089**	.079**
Axis4_RSCU	.108**	.014**	-.342**	.226**	.015**	-0.010	-.012*	.134**	.102**	-.040**	-.039**	-.024**	-.014**	-0.009	0.000	0.000	0.000	1.000	.046**	.012*	-.105**	.038**
Axis1_RAAU	.084**	-.214**	.234**	.384**	.269**	-.139**	-.039**	-.038**	-.154**	.035**	.033**	-.680**	-.505**	.279**	.134**	.181**	-.019**	.046**	1.000	.023**	-.035**	-.042**
Axis2_RAAU	.138**	-.103**	.064**	.102**	.228**	-.118**	.019**	-.095**	-.452**	.027**	.033**	.536**	.358**	.683**	.185**	-0.009	-.051**	.012*	.023**	1.000	-.048**	-.028**
Axis3_RAAU	-0.009	.114**	.157**	-.190**	.130**	.061**	.106**	-.053**	-.369**	.087**	.086**	.059**	.097**	.355**	.021**	-.188**	-.089**	-.105**	-.035**	-.048**	1.000	-.012*
Axis4_RAAU	-.190**	.121**	-.138**	.043**	-.262**	.071**	.055**	.163**	.234**	-.131**	-.128**	.278**	.133**	.179**	-.179**	-.045**	.079**	.038**	-.042**	-.028**	-.012*	1.000

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

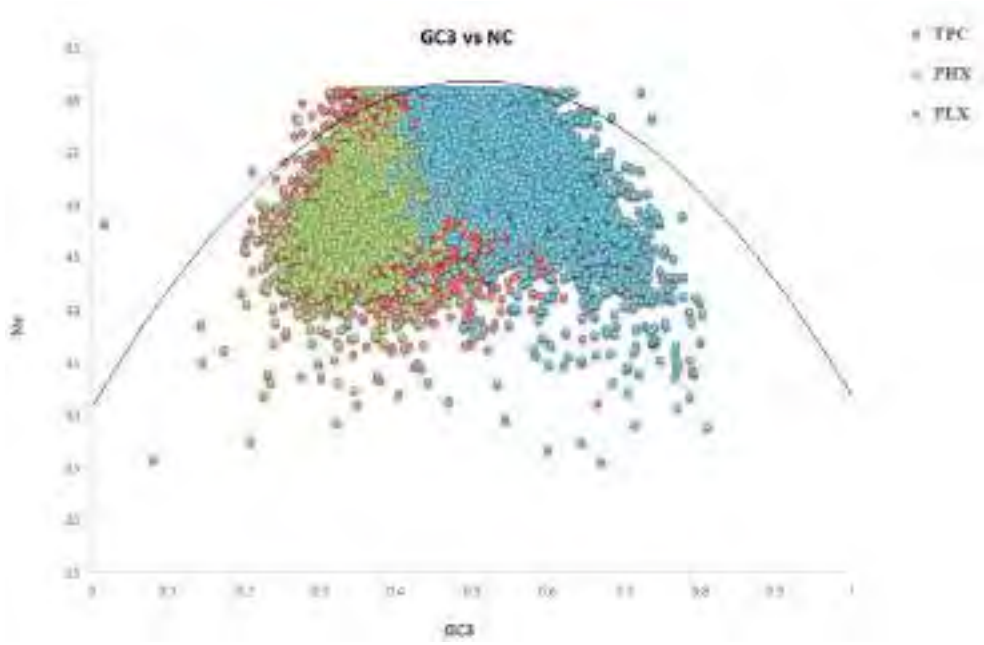


Fig.4.38.GC3 versus Nc plot for *Camellia sinensis* genome. Orange colored circles represent total protein coding genes (TPC), blue colored circles represent potentially lowly expressed genes (PLX), greencolored circles represent potentially highly expressed genes (PHX)

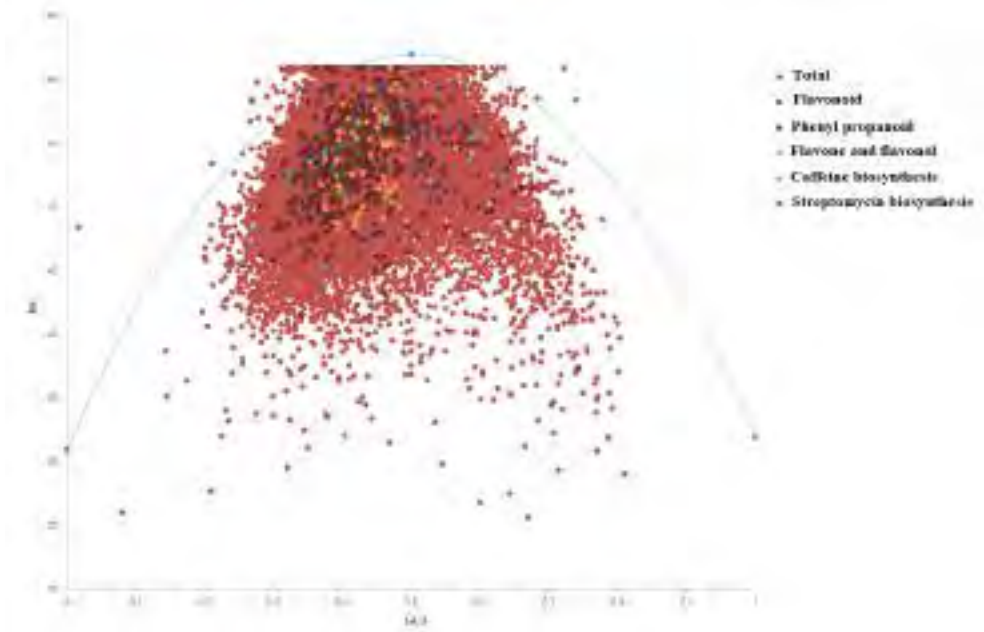


Fig.4.39.GC3 versus Nc plot for *Camellia sinensis* genome highlighting some important metabolic genes

data was executed to gain a better insight into the tentative factors dictating codon usage signatures in the tea genome. Correspondence analysis (CoA) based on RSCU data revealed that Axes 1 mainly contributed to total inertia (3.78) in the 59- dimensional hyperspace. Such an observation signified that Axis 1 has been the principal axis governing the variation

among genes. RSCU based scatter plots revealed an interesting fact that PHX and PLX genes clustered separately at the extreme opposite ends of the major axis of separation of genes (Figure 4.40). Such differential trends of codon usage among the potential highly expressed and lowly expressed genes signified a marked impact of translational selection to be influential on the *C. sinensis* genome.

Gene expression level was found to be another important determinant influencing codon usage pattern in genome of *C. sinensis* as obvious from the CoA plots. Such cases of gene expressivity affecting codon usage have been reported in *Caenorhabditis elegans*, *Drosophilla melanogaster*, and *Arabidopsis thaliana* (Duret and Mouchiroud, 1999), *Corynebacterium glutamicum* (Liu *et al*. 2010), etc. It was interesting to note that there was a strong significant and positive correlation between Axis 1 of RSCU and CAI (Table 4.10). Meaningful correlation analysis and RSCU based scatter plots clearly emphasized the influence of gene expression level to be a major determinant in framing the codon usage pattern of the *C. sinensis* genome.

The PHX genes were the ones to encode for proteins like disease resistance family protein, nucleic acid-binding protein, plant calmodulin-binding protein, SAP domain-containing protein, galactose oxidase, leucine-rich repeat, thioesterase superfamily protein, etc. On the other

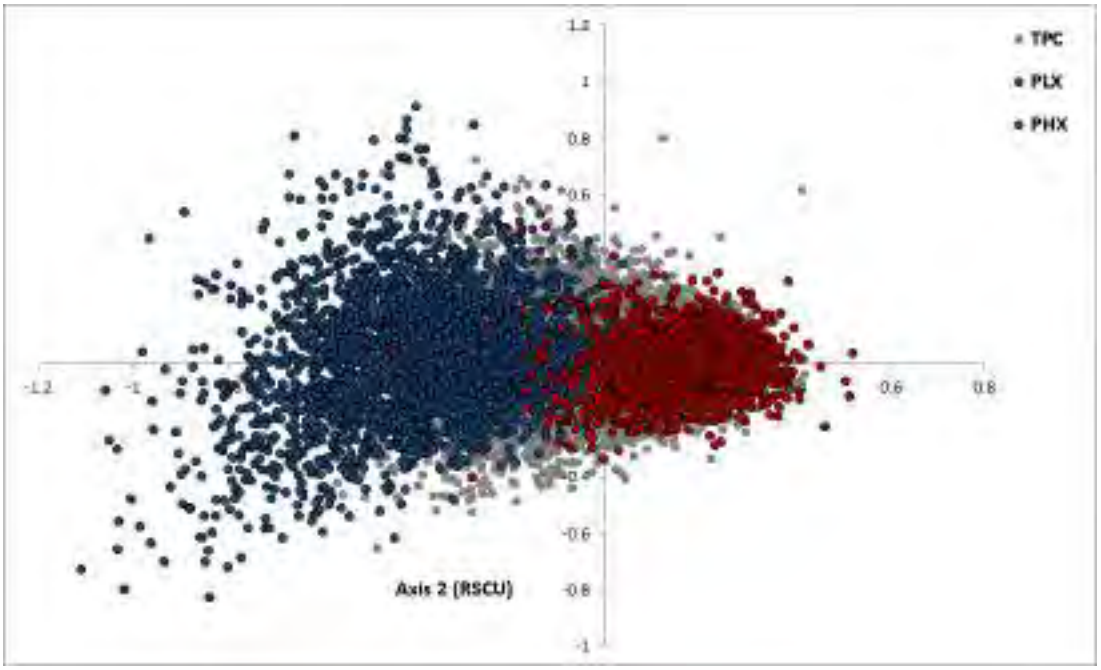


Fig.4.40.RSCU scatter plot for *Camellia sinensis*. Gray colored circles represent total protein coding genes (TPC), blue colored circles represent potentially lowly expressed genes (PLX), red colored circles represent potentially highly expressed genes (PHX)

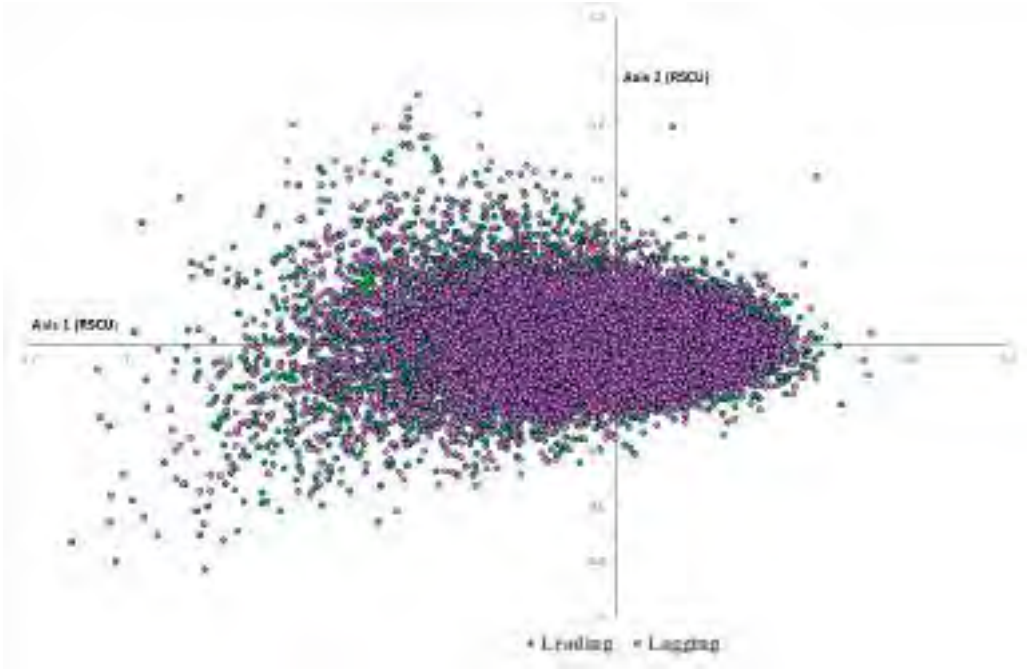


Fig.4.41. RSCU scatter plot for the leading and lagging strand-specific genes of *C. sinensis* showing absence of replicational selection. The green colored circle represents genes on the leading strand and the purple colored represents genes on the lagging strand

hand, the PLX genes were the ones to encode proteins like calcium-binding EF-hand family protein, cupredoxin superfamily, carbohydrate-binding X8 domain superfamily, etc. The list of PHX and PLX genes is given in Supplementary material 4 (SM4) .

The length of the coding sequence was found to correlate strongly with Axis 1 of RSCU (Table 4.10) suggesting the role of length of coding sequences in codon usage variation. The fact that replicational and transcriptional selection dictates the codon usage signatures in organisms, experiencing strand (leading/lagging strand of replication) asymmetry, has been well established (Das *et al*., 2006; Guo and Yu, 2007; Guo and Yuan, 2009; Lafay *et al*., 1999; McInerney, 1998; Wei and Guo, 2010). RSCU based scatter plots of leading and lagging strand-associated genes of the tea genome revealed that the complimentary strand -associated genes clustered together signifying the absence of replication-associated mutational pressure (Figure 4.41). Correlation studies of codon usage parameters directed towards the role of length of coding sequences in influencing codon usage behavior.

A significant positive correlation between gene expression level and gene length in the tea genome pointed towards the tendency of the potential highly expressed genes to be longer in stretch than the lowly expressed genes. However, *C. sinensis* being a

eukaryotic organism has deviated from the trends reported earlier in eukaryotes like *D. melanogaster* (Miyasaka, 2002), *C. elegans* (Marais and Duret, 2001) where the length of coding sequences was found to correlate negatively with the level of gene expression. Though PHX genes in eukaryotes may be shorter in length to minimize energy cost (Moriyama and Powell, 1998), the behavioral pattern of the genome prefers a longer stretch of highly expressed genes over shorter ones. Such an observation of the positive correlation between gene expression levels and gene length to minimize misincorporation and missense errors to increase translational accuracy has been formerly reported in prokaryotic organisms like *E. coli* (Eyre-Walker, 1996), *Pseudomonas aeruginosa* (Gupta and Ghosh, 2001), *Streptococcus pneumonia* (Hou and Yang, 2001), etc.

4.7.2. Exploring the sources of amino acid usage heterogeneity

Amino acids like Leucine (L), Serine (S) were used in the highest frequencies as evident from the amino acid usage (Figure 4.42). The other amino acids used in higher frequencies were Alanine (A), Lysine (K), Glutamic acid (E), Valine (V), and Glycine (G). The first and the second principal axes i.e., Axis 1 and Axis 2 were mainly found to govern the amino acid usage variation in the tea genome). The two major axes of RAAU data were found to exhibit significant correlations with hydropathicity index GRAVY and aromaticity of the gene products of *C. sinensis* genome (Table 4.10).

It was perceptible from RAAU data-based scatter plot that the protein-coding genes separated based on higher or lower values of Gravy and Aromo value. The genes with high GRAVY (high values of hydrophobicity) represented by red color and the genes with low GRAVY (low values of hydrophobicity) marked by blue color (Figure 4.43.a) clustered separately along the axis of separation. Similarly, the genes with higher aromo (purple color) and lower aromo (green color) clustered separately in the RAAU data-based scatter plot (Figure 4.43.b). Detailed analysis of amino acid usage-based parameters revealed the possible sources of amino acid usage variation in *C. sinensis*. Hydrophobicity and aromaticity were found to govern amino acid usage signatures. The role of gene expression level in amino acid usage in all the *C. sinensis* was obvious from the high correlation of CAI and the two major axes of separation of genes based on RAAU i.e., Axis 1 and Axis 2. Axis 1 and Axis 2 of RAAU

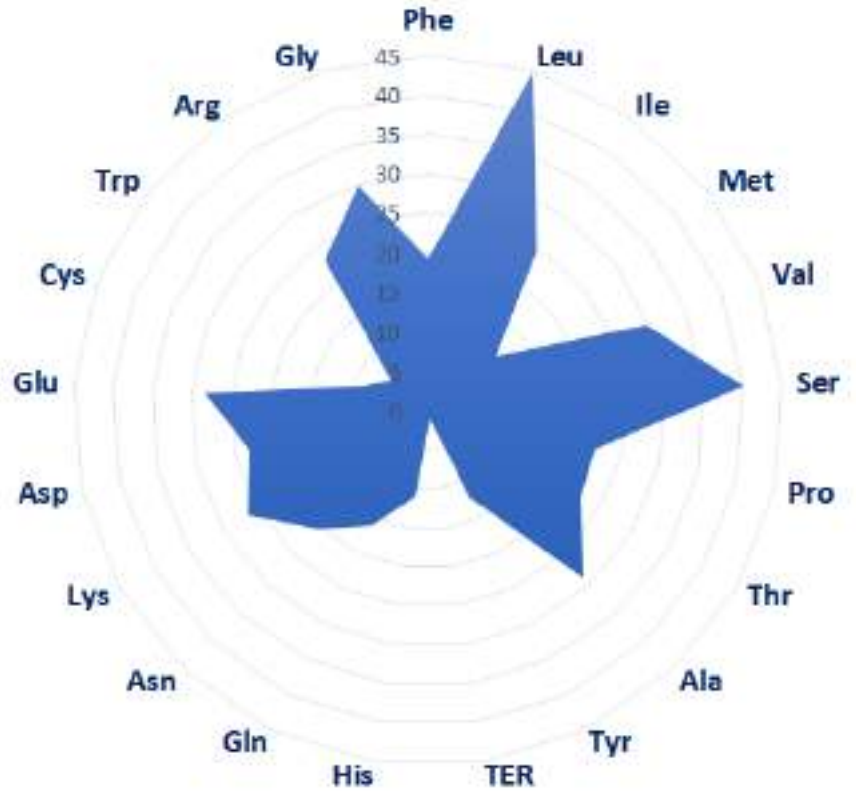


Fig.4.42.A radar plot showing amino acid usage in *C. sinensis* genome

were found to correlate significantly with gene expressivity, as estimated by CAI (Table 4.10). Thus, gene expression level not only influenced the codon usage signatures but also contributed substantially to producing the observed amino acid usage variations in the genome.

4.7.3. Protein-energy cost

It has been reported in prokaryotic organisms that the highly expressed gene products tend to be less expensive in terms of biosynthetic energy (protein -energy cost) and judiciously abide by the policy of cost-minimization (Roy *et al*., 2015; Seligmann, 2003). A significant positive correlation (Table 4.10) was noted between expression level (CAI) and energy cost in the tea genome ($r = 0.271$, $p<0.01$). Significant positive correlation of gene expression level with protein-energy cost indicated towards the utilization of less energetically costly amino acids in highly expressed genes which have also been reported in *Saccharomyces cerevisiae* (Kahali *et al*., 2007; Raiford *et al*., 2008). Increased utilization of inexpensive amino acids in highly expressed genes to lessen the use of expensive amino acids is true for prokaryotes (Akashi and Gojobori, 2002; Heizer *et al*., 2006).

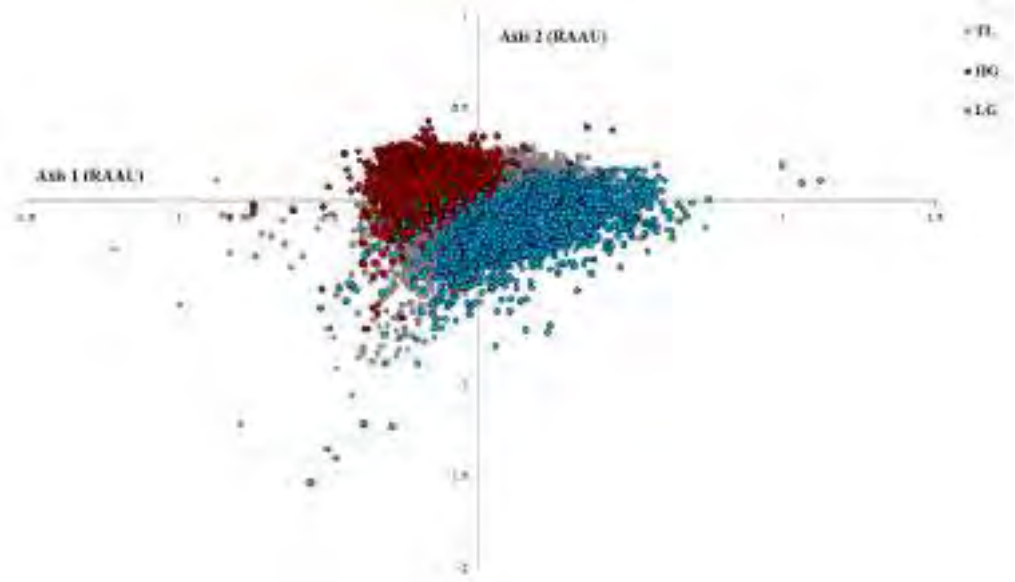


Fig.4.43.a.RAAU scatter plot showing genes with high hdrophobicity (red circle) and genes with low hydrophobicity (blue circle) of *C. sinensis* cluster separately . Grey coloured regions represents total genes

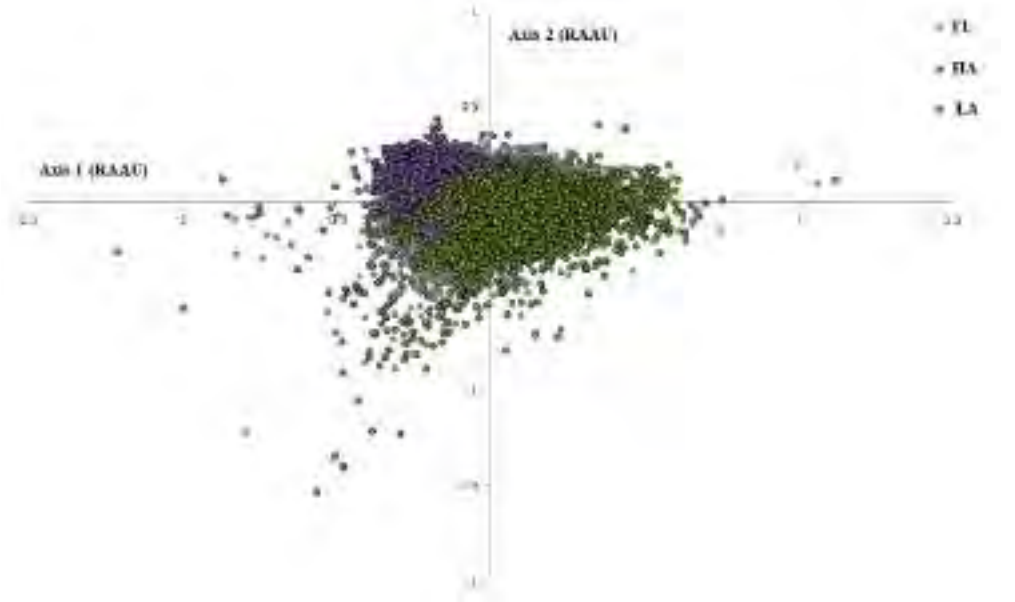


Fig.4.43.b.RAAU scatter plot showing genes with high aromo (green circle) and genes with low aromo (purple circle) of *C. sinensis* cluster separately .Grey coloured regions represents total genes

4.7.4. Eukaryotic Orthologous Groups (KOG)

Extensive profiling of Eukaryotic Orthologous Groups (KOG) revealed that the tea genome was mostly enriched with the proteins representing the KOG categories T (Signal transduction mechanisms), O (Posttranslational modification, protein turnover, chaperons), Q (Secondary metabolites biosynthesis, transport, and catabolism), G (Carbohydrate transport and metabolism), K (Transcription), C (Energy production and conversion),

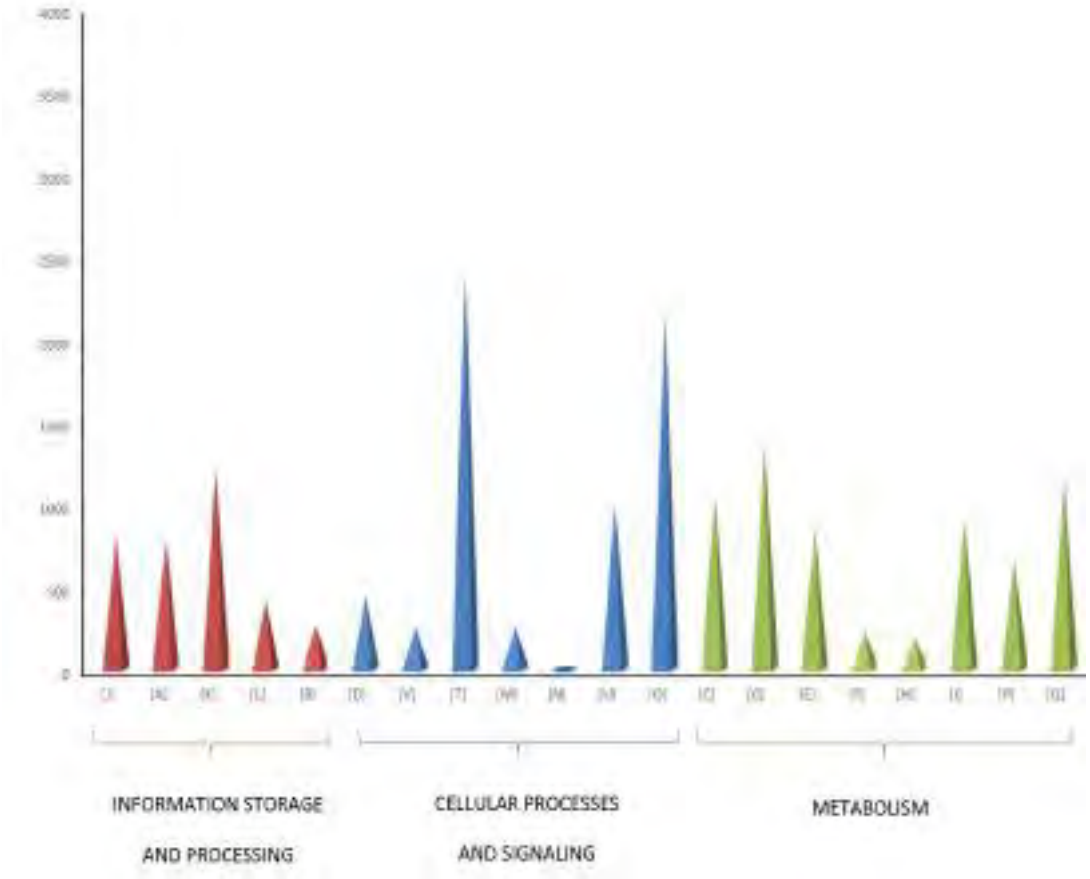


Fig.4.44. Characterization of Eukaryotic orthologous groups (KOGs) of the *C. sinensis* genome

U(Intracellular trafficking, secretion, and vesicular transport (Figure 4.44). However we did not consider the poorly characterized categories R (general function prediction) and S (unknown function).

Thus, the KOG categories in the majority were under the group like metabolism-related proteins (T, O, Q, G, and C), cellular processes and signaling (T, U, O), and information storage and processing (J, A, K).

4.8. *In silico* Polypharmacology of tea

4.8.1. Polypharmacological studies of Purple tea

4.8.1.1. Phytocompounds and their respective target proteins

GC/MS analysis of Purple tea (PT) extract yielded a list of 29 component chemicals (Table 5- Apendix D). Four phytocompounds, namely Hexadecanoic acid, 1,2,3-benzenetriol, Caffeine, and γ -Sitosterol, have more than 80% interaction probability with 19 target proteins, according to target prediction.

4.8.1.2. Phytocompounds and their Drug likeliness assessment

The details of this analysis are provided in Supplementary material 5 (SM5) and the results are discussed below briefly. Hexadecanoic acid and 1,2,3-benzenetriol have permeability across the blood-brain barrier (BBB), whereas caffeine and γ -Sitosterol do not. All of the substances, except for γ -Sitosterol, have a high GI absorption rate. None of the chemicals are permeability glycoprotein substrates (P -gp). One of the most important properties that influence drug absorption is solubility. Three topological approaches are employed in SwissADME to predict water solubility (Gfeller *et al.,* 2014). Hexadecanoic acid was found to be 'moderately soluble', 1,2,3-benzenetriol and Caffeine are 'very soluble,' while γ -Sitosterol is 'poorly soluble,'. The Lipinski rule-of-five is one of the filters used in this study to determine a compound's 'drug-likeliness.' The rule is followed by 1,2,3-benzenetriol and Caffeine, while Hexadecanoic acid and γ -Sitosterol each have one violation. However, Hexadecanoic acid has the greatest Bioavailability score (chance of a chemical having at least 10% oral bioavailability in rats or detectable Caco-2 permeability) among the substances (0.85).

4.8.1.3. Interaction with human proteins

The purple tea-derived compounds interacted with a total of 19 human proteins (Table 4.11, Figure 4.45). The network showing interaction of target proteins of Purple tea (PT) derived compounds and other proteins of *Homo sapiens* is shown in Figure 4.46. The compound 1, 2, 3-Benzenetriol interacted with epidermal growth factor (EFGR), with proto-oncogene tyrosine-protein kinase Fyn, and Carbonic anhydrase (CA) 1, 2, and 6. The EFGR activates numerous signaling pathways to transform extracellular inputs into suitable cellular responses. Fyn regulates cell growth and survival, cell adhesion, integrin-mediated signaling, cytoskeletal remodeling, cell motility, immunological response, and axon guidance, among other biological activities. The regulation of fluid secretion into the anterior chamber of the eye is linked to carbonic anhydrase (CA) proteins. These CA proteins contribute to intracellular pH regulation in the upper villous epithelium of the duodenum during proton-coupled peptide absorption and boost chloride-bicarbonate exchange activity, which aids CO2 transport into the blood.

Table 4.11. Target proteins associated with PT and their probability of interaction

Compound	Target	Common Name	Uniport ID	Probability
1,2,3-Benzenetriol	Tyrosine-protein kinase FYN	FYN	P06241	0.91
	Epidermal growth factor receptor erbB1	EGFR	P00533	0.91
	Carbonic anhydrase II	CA2	P00918	0.91
	Carbonic anhydrase I	CA1	P00915	0.91
	Carbonic anhydrase VI	CA6	P23280	0.91
Coffeine	Acetylcholinesterase	ACHE	P22303	0.85
	HERG	KCNH2	Q12809	0.85
	Adenosine A2a receptor	ADORA2A	P29274	0.85
γ-Sitosterol	Nuclear receptor ROR-gamma	RORC	P51449	0.92
	Niemann-Pick C1-like protein 1	NPC1L1	Q9UHC9	0.92
	LXR-alpha	NR1H3	Q13133	0.92
	Sterol regulatory element-binding protein 2	SREBF2	Q12772	0.83
	HMG-CoA reductase	HMGCR	P04035	0.83
Hexadecanoic Acid	Fatty acid-binding protein adipocyte	FABP4	P15090	0.94
	Peroxisome proliferator-activated receptor alpha	PPARA	Q07869	0.94
	Fatty acid-binding protein muscle	FABP3	P05413	0.94
	Fatty acid-binding protein epidermal	FABP5	Q01469	0.94
	Peroxisome proliferator-activated receptor delta	PPARD	Q03181	0.94
	Fatty acid-binding protein intestinal	FABP2	P12104	0.94

Acetylcholinesterase (ACE), potassium voltage-gated channel subfamily H member 2 (KCNH2), and adenosine A2a receptor (ADORA2A) are all associated with caffeine . Through fast hydrolysis of acetylcholine delivered into the synaptic cleft, ACE plays a key role in signal transduction at the neuromuscular junction. Additionally, this protein was linked to neuronal death.

Proteins like Fatty acid-binding protein adipocyte (FABP4), Peroxisome proliferator-activated receptor alpha (PPARA), Fatty acid-binding protein muscle (FABP3), Fatty acid-binding protein epidermal (FABP5), Peroxisome proliferator-activated receptor delta (PPARD), and Fatty acid -binding protein intestinal (FABP2) were found to form a direct network with Hexadecanoic Acid. In adipocytes, these proteins are linked to the Lipid Transport Protein. PPARD

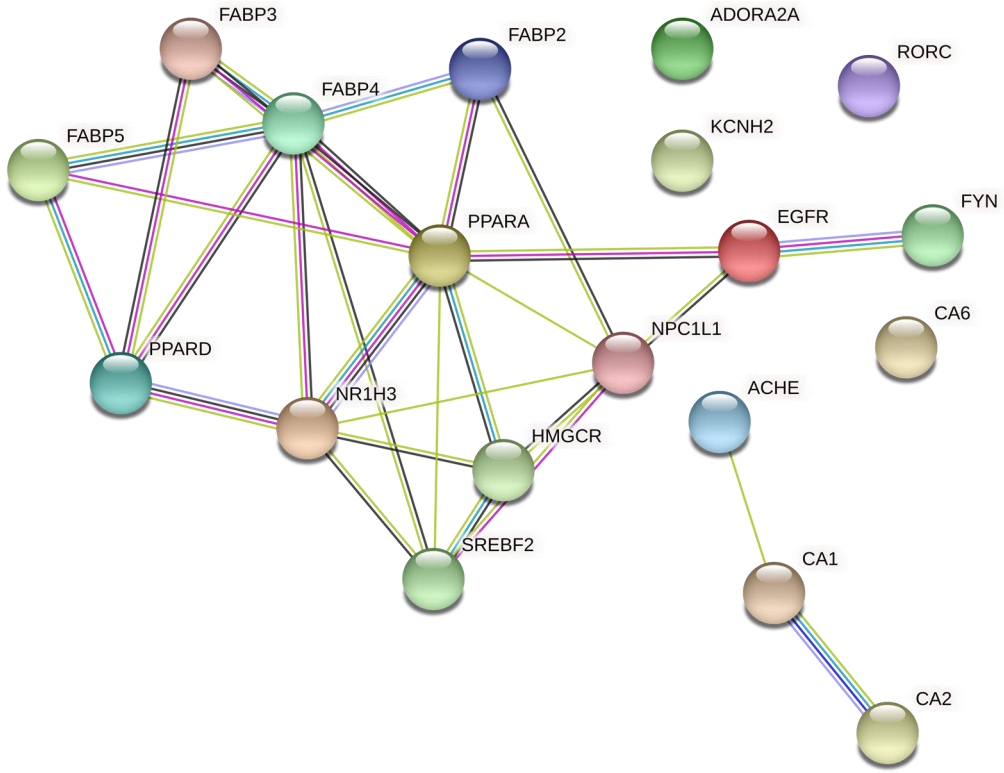


Fig 4.45. Interaction between the target proteins of Purple tea (PT) derived compounds

also functions as a Ligand-activated transcription factor that prefers polyunsaturated fatty acids like gamma -linoleic acid and eicosapentaenoic acid. They regulate the peroxisomal beta-oxidation pathway of fatty acids after activation.

γ-Sitosterol was found to interact with nuclear receptor ROR-gamma (RORC), Niemann-Pick C1-like protein 1 (NPC1L1), LXR-alpha (NR1H3), Sterol regulatory element-binding protein 2 (SREBF2), and HMG -CoA reductase (HMGCR). These interacting proteins are essential for the body's homeostasis to be maintained. RORC controls cellular differentiation, immunity, peripheral circadian rhythm, lipid, steroid, xenobiotic, and glucose metabolism, among other things. Through cholesterol uptake across the plasma membrane of the intestinal enterocyte, NPC1L1 plays a critical function in cholesterol homeostasis. Multiple lipid transport abnormalities result from a loss-of-function mutation in this protein.

With a p-value of less than 0.05, GO analysis of the target genes reveals that they are related with 152 BP, 62 MF, and 32 CC (Supplementary Material 6: ST 2, 3 and 4) The top 15 GO results based on gene count are shown in Figure 4.47. The majority of the genes

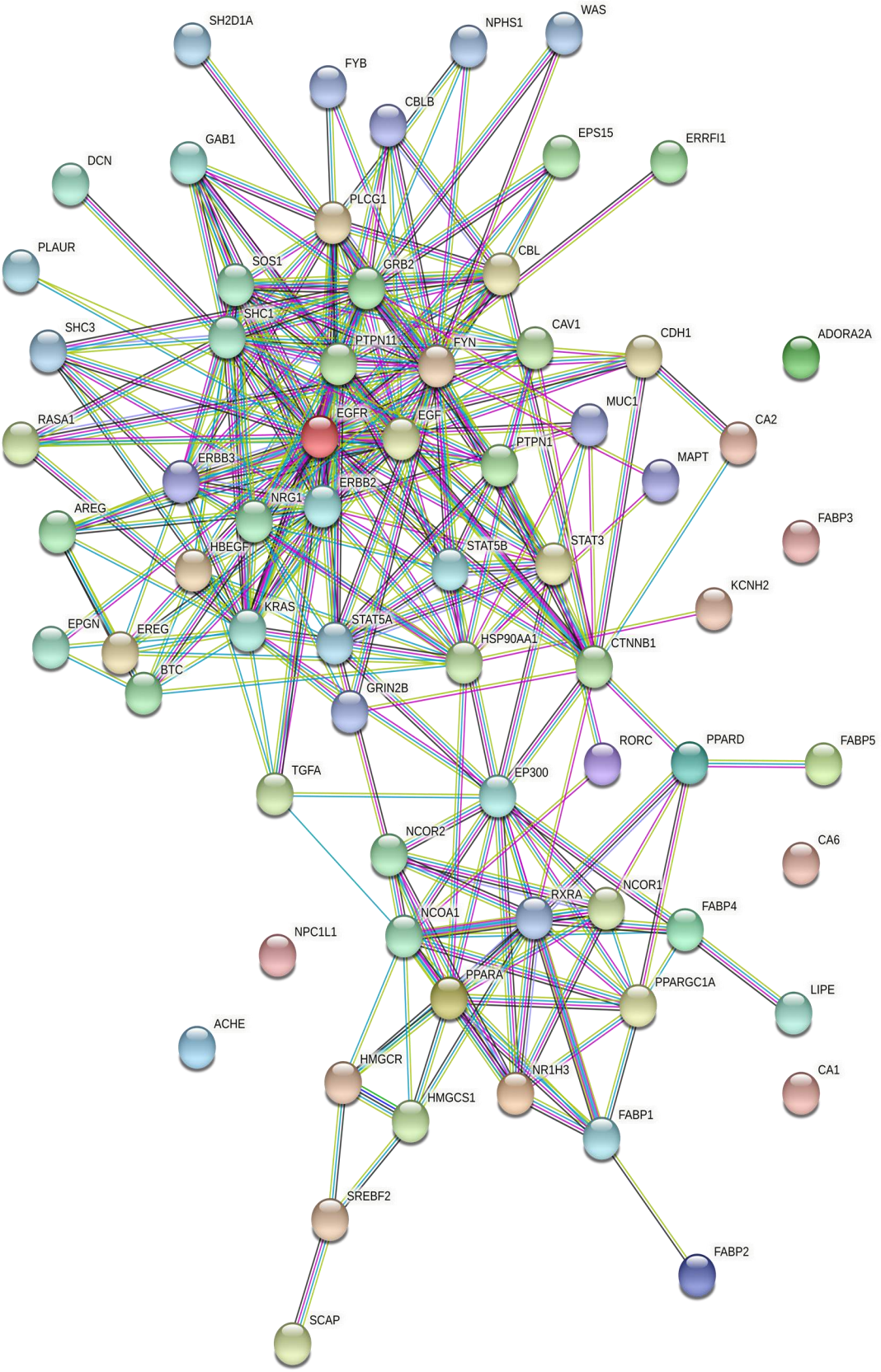


Fig 4.46. PPI network showing interaction of target proteins of Purple tea (PT) derived compounds and other proteins of *Homo sapiens*

(58) are involved in the molecular function of protein binding. The biggest number of genes (32), according to cell component analysis, are found in the cytosol.

As a result of this investigation, it was discovered that human proteins that interact directly with the selected compounds are functionally important and are primarily linked to cell growth and survival, axon guidance, neural transport, immune response, cell motility, lipid metabolism, and cholesterol metabolism. Furthermore, KEGG pathway enrichment analysis shows that PT targets are associated with pathways like ErbB signaling (21 proteins), PPAR signaling (8 proteins), Insulin signaling (10 proteins), Ras signaling (12 proteins), Natural killer cell-mediated cytotoxicity (9 proteins), Neurotrophin signaling (8 proteins), and others (Supplementary Material 7 : ST 5, SF 1,2,3,4,5 and 6).

4.8.1.4. Role in Diseases

4.8.1.4.1. Gut-Brain Axis

Compounds produced from PT were discovered to interact with many genes linked to various illnesses like Hypercholesterolemia (increased levels of HDL cholesterol in the blood), Type 2 diabetes, edema, cardiovascular illnesses, colorectal cancer, Alzheimer's disease, atherosclerosis, insulin resistance, obesity, hyperlipidemias, and several other metabolic syndromes (Supplementary Material 6: ST 1, Figure 4.48). As a result, these proteins were divided into two categories: (a) gut-related disorders and (b) neurological diseases, indicating the impact of PT on the Gut-Brain Axis (GBA). The GBA is a two-way communication system that connects the enteric and neurological systems. This is the link between the brain's emotional and cognitive centers and the functioning of the peripheral intestine.

We discovered 34 proteins that interacted with PT-derived chemicals were linked to Type 2 Diabetes and Edema while evaluating the data (Supplementary Material 6: ST 1 TGFA, EGFR, ERBB3, PLCG1, STAT5B, EGF, GAB1, NRG1, GRB2, SOS1, and HBEGF were discovered to be linked to the ErbB signaling pathway among the 34 proteins studied (Supplementary Material 7 :ST 5, SF 2). This pathway is important for directing intracellular signaling cascades, either directly or indirectly, in a variety of hereditary diseases (Holbro and Hynes, 2004). Peptides linked to epidermal growth factor (EGF) connect to ErbB receptors, activating intrinsic kinases (Fiske *et al.,* 2009). These Fyn-kinases phosphorylate or dephosphorylate a

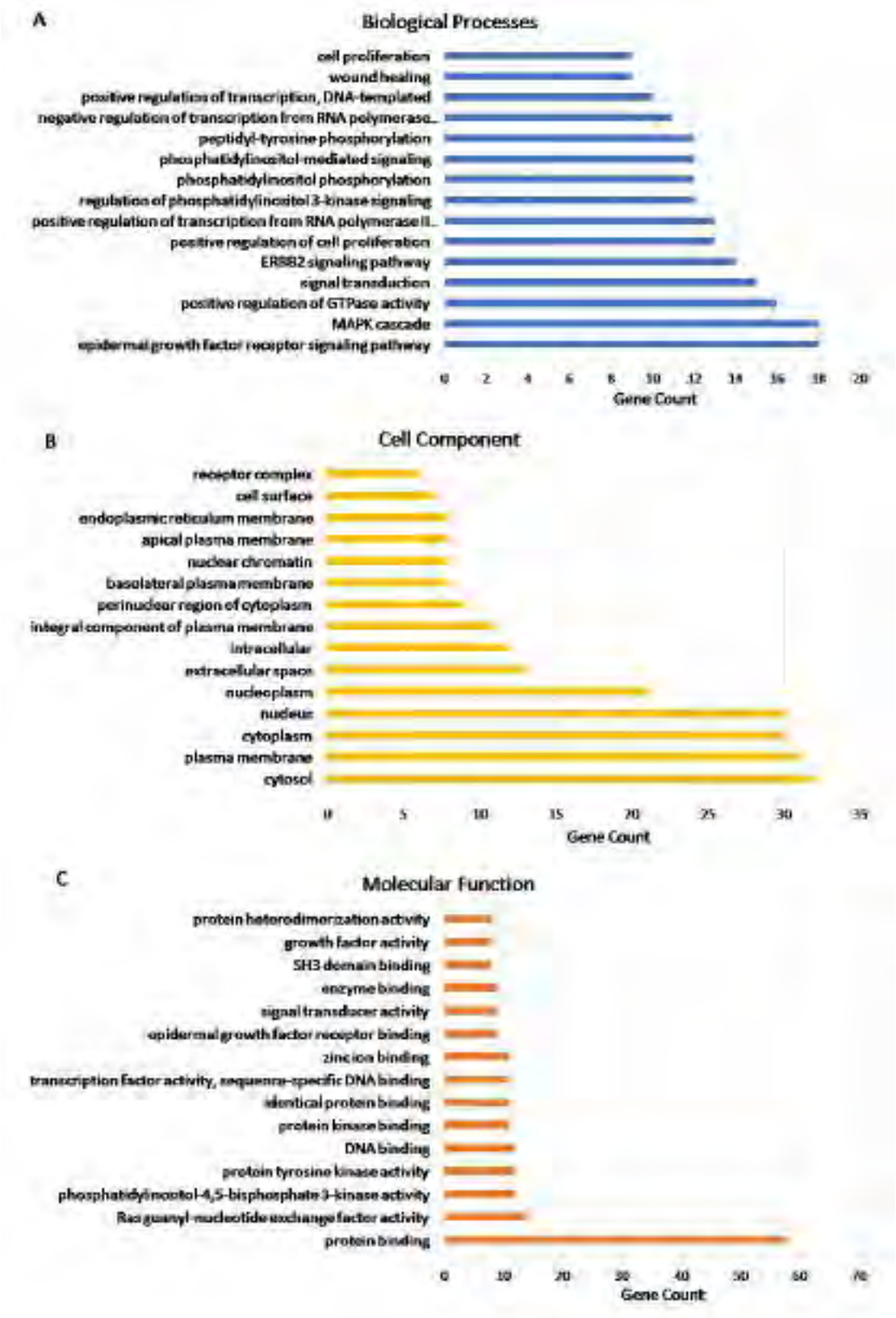


Fig 4.47. Gene Ontology Enrichment Analysis. The top 15 results (with p-value < 0.05) of Gene ontology enrichment analysis of the target genes show them to be associated with different (A) Biological Processes, (B) Cell Components and (C) Molecular Functions

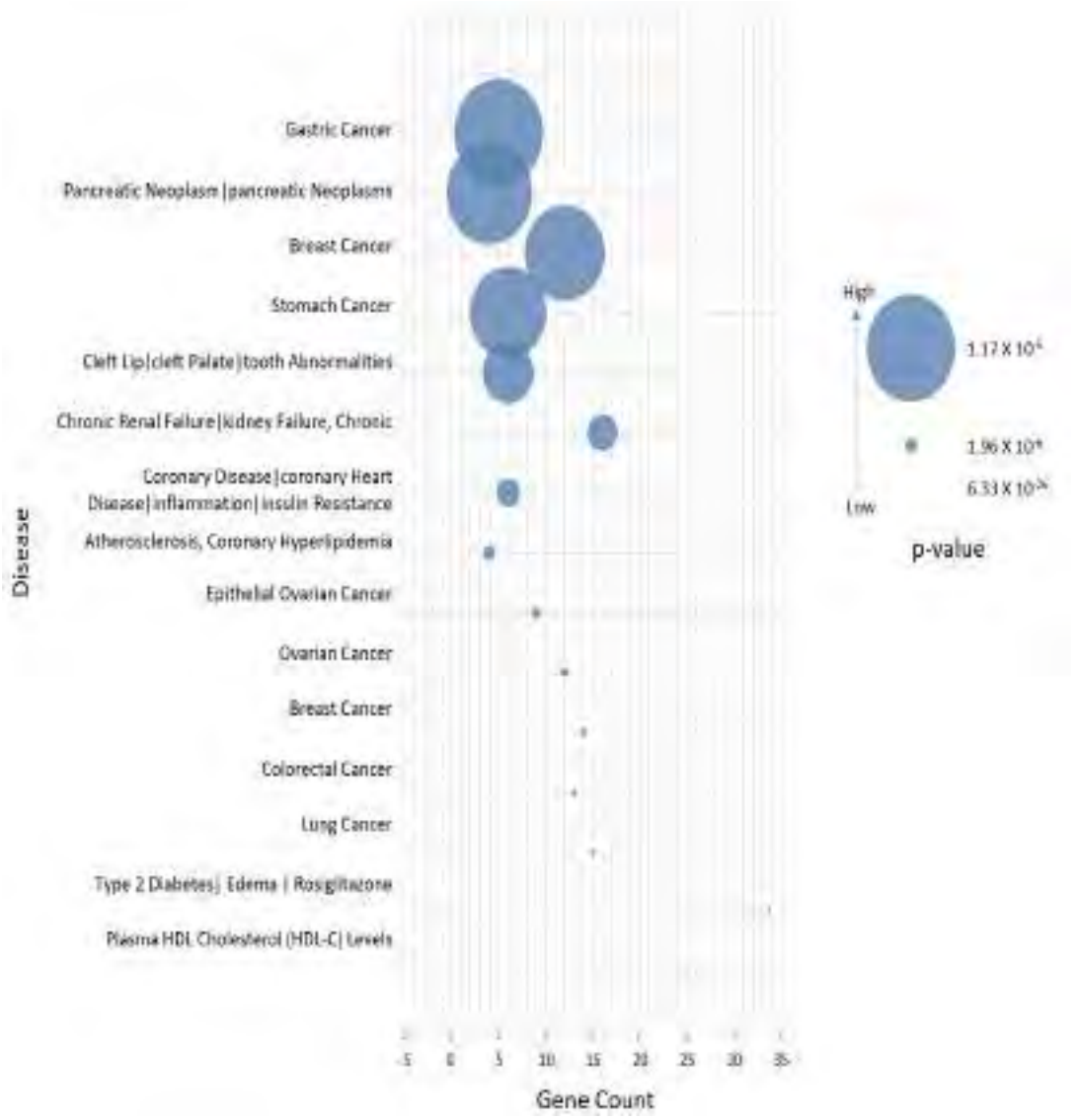


Fig 4.48. Functional Annotation (Disease). The top 15 results (with p-value < 0.05) of Functional annotation of target genes show their association with diseases such as plasma HDL cholesterol (Hypercholesterolemia), Type 2 diabetes, edema, coronary diseases, etc. Size of the bubbles indicate the p-value

variety of target proteins, regulating the cell-cell contact cascade (Schreier *et al.,* 2021). Thus, the intrinsic signaling cascades broadly linked with the Gut-Brain axis were revealed to be modulated by both EGFR related to ErbB signaling and Fyn kinase, which was discovered to be the principal target of PT-derived components.

4.8.1.4.2. Immune Booster

A total of 9 targets namely SHC3, SHC1, SH2D1A, KRAS, PTPN11, GRB2, FYN, PLCG1, and SOS1 were found to be related with 'Natural killer cell-mediated cytotoxicity in our body' in a KEGG pathway enrichment study of the PT targets (Supplementary Maetrial 7: ST 5, SF 6). Some of the

key immune cells linked with the innate immune system include natural killer (NK) cells, macrophages, and other phagocytes. These cells serve a key part in cell-mediated cytotoxicity that is mediated by antibodies (ADCC). Furthermore, NK cells can kill tumor cells without the need for immunization or immune activation (Waldhauer and Steinle, 2008). In the presence of increased reactive oxygen species (ROS), these NK cells are especially vulnerable to and can lose their ability to destroy antigens (Nakamura and Matsunaga, 1998). PT can also protect NK cells from ROS-induced damage, enhancing NK cell-mediated immunity. Furthermore, anti-oxidant properties are invariably linked to apoptosis induction and cancer inhibition (Harris and De Nicola, 2020). The PT interacts with proteins involved in the PI-3 kinase and MAP kinase pathways (Downward, 2004). These pathways play an important role in cell growth, survival, aging, and apoptosis regulation. As a result, PT chemicals aid in the general preservation of the normal cell cycle, avoiding the start of uncontrolled cellular proliferation, which leads to cancer.

4.8.1.4.3. Alzheimer's disease management

Alzheimer's disease (AD) is a progressive neurological illness that causes brain atrophy and cell death. This is the most prevalent form of dementia, and it causes patients' intellectual abilities to deteriorate over time, as well as causes significant damage to their behavioral and social skills. All of these things have an impact on a person's ability to think and function independently in the long run. The tau protein has been linked to the development of Alzheimer's disease. This protein is mostly found in brain cells and is linked to the production of microtubules (Wenk, 2003). These microtubules aid in the transportation of vital elements such as nutrients from one section of the nerve cell to another. Tau can diffuse throughout the brain as an oligomer, allowing chemical and electrical messages to pass via synapse (Fiske *et al.,* 2009). When tau accumulates, however, clinical consequences similar to those seen in Alzheimer's disease develop. In Alzheimer's disease, two primary accumulations have been identified: (a) β -amyloid (A) plaques and (b) tau neurofibrillary tangles (NFTs). These disrupt the brain's and neurons' homeostasis, resulting in Alzheimer's disease (Wenk, 2003). Tyrosine-protein kinases have been discovered in several studies on this disease. Fyn is crucial in the treatment of Alzheimer's disease. They regulate

cell growth and survival, cell adhesion, integrin-mediated signaling, cytoskeletal remodeling, cell motility, immunological response, and axon guidance, among other biological activities (Shirazi and Wood, 1993). Furthermore, Fyn regulates the nervous system's functionality by phosphorylating proteins involved in synaptic signal transmission. Recent research has also demonstrated the role of Fyn in preventing the onset of Alzheimer's disease via regulating tau aggregation (Briner *et al.,* 2020). Fyn knockout mice also lacked normal myelination, according to previous research. The tau protein interacts with the microtubules, and Fyn interacts with it. Both Fyn and tau are found in oligodendrocyte rafts, generating a Fyn -tau complex. Tau is attached to the SH3 domain of Fyn (Fyn has both SH2 and SH3 domains) on one side and the microtubules on the other in this complex. The interaction of Fyn-tau microtubules permits oligodendrocytes to link with nearby neurons throughout their extension process (Rojo *et al.,* 2006).

The phytocompounds in PT have been discovered to interact with Fyn kinase through the ErbB signaling pathway. Furthermore, Functional Annotation reveals that five of PT drugs' key targets (ACHE, FYN, HMGCR, PPARA, PPARD) are linked to Alzheimer's disease (Supplementary Material 6: ST 1). These chemicals can also control axonal outgrowth, neurotransmitter retrograde transmission, and the neuro-myelination process. As a result, we might speculate that PT can maintain neuronal homeostasis by interacting with the Fyn kinase protein, preventing or at least delaying the onset of AD and associated neurological diseases.

4.8.1.4.4. Hyperglycemia

People with diabetes might be affected by high blood sugar (hyperglycemia). Hyperglycemia is caused by several causes that might aggravate Type 2 Diabetes (T2D). T2D is a chronic condition that affects glucose metabolism and renders the body insulin resistant. This means that the cells will not employ insulin to utilize glucose. The insulin signaling pathway is linked to Fyn-kinase and other Src-kinase family proteins. They work with the help of the plasma membrane's lipid raft microdomains (PM). Post-transcriptional changes shift Fyn kinase from the cytoplasm to the plasma membrane (Dai *et al.,* 2017). They are specifically found within the lipid rafts. Fyn can phosphorylate caveolin protein in lipid rafts in response to insulin stimulation. Fyn also interacts with flotillin and other proteins found in lipid rafts (Dai *et al.,* 2017). Fyn aids in

fatty acid absorption and oxidation, which is directly linked to insulin signaling, through these interactions (Dai *et al.,* 2017). Furthermore, constitutively active Src family proteins can block pyruvate kinase, the enzyme that catalyzes the final step of glycolysis (phosphor-enol pyruvate →). As a result, the use of glucose as a source of energy is inhibited. Furthermore, inhibiting Fyn and other Src kinases can stop adipocyte differentiation from happening. These findings confirmed Fyn kinase's role in insulin signaling. As previously established, PT drugs interact with the Fyn kinase via the ErbB signaling pathway to keep it active in adipocyte development and insulin signaling. Accurate insulin signaling will keep glucose uptake and utilization in check, reducing the risk of insulin resistance and type 2 diabetes. LIPE, PTPN1, SHC3, SHC1, CBLB, KRAS, GRB2, CBL, SOS1, and PPARGC1A are among the Insulin signaling pathways ten targets (Supplementary Material 7 , -ST 5, SF 4.Furthermore, PT drugs interact with eight proteins of the peroxisome-proliferator-activated-receptors (PPAR) signaling pathway (Supplementary Material 7: ST 5, SF 3) that play an essential role in adaptive thermogenesis (FABP1, FABP2, RXRA, FABP4, FABP5, NR1H3, PPARA, and PPARD). This is the production of body heat in stressful situations (Polvani *et al.,* 2016). When people are stressed, their rate of gluconeogenesis increases, which boosts blood glucose levels. Insulin is secreted to generate heat (or energy) from glucose. In the absence of stress, the entire situation is 'cooled down.' In the presence and absence of stress, adaptive thermogenesis mediated by PPAR aids in maintaining insulin homeostasis (Polvani *et al.,* 2016). PPAR interacts with the RAS signaling system and is involved in insulin signaling, as well as glucose and lipid metabolism (Polvani *et al.,* 2016). As a result, regular PT consumption can help to maintain a healthy blood glucose level in the body, reducing the risk of type 2 diabetes and insulin resistance.

4.8.1.4.5. Hypercholesterolemia, obesity, and coronary heart disease management

The prevalence of hypercholesterolemia, obesity, and coronary heart disease is quickly rising over the world. Excess fat buildup in tissues other than adipose tissues is sometimes referred to as obesity. Fat is stored mostly in the liver and skeletal muscles, which can result in obesity-related inflammation, high blood cholesterol, and coronary heart disease. Obesity and related pathophysiological

problems are linked to several biological pathways. One of them is PPAR, which is still linked to lipid metabolism, adipocyte development, fatty acid breakdown, and bile acid production (Polvani *et al.,* 2016). As a result, PPAR is linked to maintaining cholesterol homeostasis and fat metabolism (Polvani *et al.,* 2016). Furthermore, Fyn-significance kinases in obesity and associated illnesses have been confirmed in recent investigations (Jung et al., 2008). Furthermore, functional annotation and GO enrichment analysis revealed that PT compounds influenced 'plasma HDL cholesterol (HDL-C) levels and 'coronary heart disease' linked genes (25 genes and 6 genes, respectively) (Supplementary Material 6: ST 1). As a result, we can speculate that PT drugs that interact with PPAR and Fyn play a direct role in reducing obesity and coronary heart disease.

4.8.2. Polypharmacological studies of the mature leaf of *C. sinensis*

4.8.2.1. Identification of target proteins and physiochemical analysis of the derived Phyto-compounds

Analysis of the phytocompounds from mature leaf (ML) extract of *C. sinensis* (Table 1-4: Appendix D), revealed five compounds with interaction probability of more than 60% i.e., Caffeine, Hexadecanoic acid, Octadecanoic acid, Olean-12-en- 3-one, and Stigmasta-7,22-dien-3-ol (Table 4.12). Out of the five compounds (Supplementary Material 8), only hexadecenoic acid showed blood-brain barrier (BBB) permeability. Gastro-intestinal (GI) absorption was higher in compounds like Caffeine, Hexadecanoic acid, and Octadecanoic acid. No compounds among five serve as the substrate of permeability glycoprotein (P-gp). Here, only Caffeine abides by the rule of drug likeliness implemented using Lipinski rule-of-five. Whereas, the other compounds showed one violation each. The compounds like hexadecenoic acid and Octadecanoic acid showed the highest Bioavailability score (0.85).

4.8.2.2. Phyto compounds and their interaction with human proteins

A total of 12 human proteins was found to directly interact with the phytocompounds found in mature tea leaf (Figure 4.49). The Protein-Protein interaction data obtained from STRING Database with minimum required interaction score (0.400) and maximum number of interactors set to none is provided in Supplementary Material 9. The network showing interaction of target proteins of Mature Tea Leaf (ML) derived compounds and other proteins of *Homo sapiens* with minimum required interaction score

Table 4.12: Target proteins associated with ML and their proability of interaction

Compound	Target	Common name	Uniprot ID	Probability*
Caffeine	Acetylcholinesterase	ACHE	P22303	0.85
	HERG	KCNH2	Q12809	0.85
	Adenosine A2a receptor	ADORA2A	P29274	0.85
Hexadecanoic acid	Fatty acid binding protein adipocyte	FABP4	P15090	0.94
	Peroxisome proliferator-activated receptor alpha	PPARA	Q07869	0.94
	Fatty acid binding protein muscle	FABP3	P05413	0.94
	Fatty acid binding protein epidermal	FABP5	Q01469	0.94
	Peroxisome proliferator-activated receptor delta	PPARD	Q03181	0.94
	Fatty acid binding protein intestinal	FABP2	P12104	0.94
Octadecanoic acid	Peroxisome proliferator-activated receptor alpha	PPARA	Q07869	0.93
	Peroxisome proliferator-activated receptor delta	PPARD	Q03181	0.93
	Fatty acid binding protein adipocyte	FABP4	P15090	0.71
	Fatty acid binding protein epidermal	FABP5	Q01469	0.71
Olean-12-en- 3-one	Cytochrome P450 19A1	CYP19A1	P11511	0.61
Stigmasta-7,22-dien-3-ol	Androgen Receptor	AR	P10275	0.68

(0.900) and maximum number of interactors set to 50 is shown in Figure 4.50. The interaction of compounds like caffeine and hexadecenoic acid is already mentioned in the above section (4.8.1.3). Octadecanoic acid was also found to interact with Peroxisome proliferator-activated receptor alpha (PPARA), Peroxisome proliferator-activated receptor delta (PPARD), Fatty acid-binding protein adipocyte (FABP4), Fatty acid-binding protein, and epidermal (FABP5). Both hexadecenoic acid and Octadecanoic acid showed a direct network with proteins directly associated with Lipid transport protein in adipocytes.

Olean-12-en- 3-one was directly

related to Cytochrome P450 family 19 (CYP19A1) which catalyzes the formation of aromatic C18 estrogens from C19 androgens.

Stigmasta-7,22-dien-3-ol on the other hand showed direct relation with Androgen Receptor (AR). These receptors are ligand-activated transcription factors that control cellular proliferation and differentiation in target tissues by regulating eukaryotic gene expression. The activity of transcription factors is influenced by the presence of binding coactivator and corepressor proteins. NR0B2 inhibits the activation of transcription. HIPK3 and ZIPK/ DAPK3 activate it but do not phosphorylate it.

The Gene Ontology results based on gene count showed a majority of the genes to be involved in the molecular function (Supplementary Materia10) of lipid binding and Biological processes (Supplementary Materia 11) like the triglyceride catabolic process (Figure 4.51).

4.8.2.3. Compounds derived from the mature leaf

The compounds were found to interact with several genes linked to various diseases with (Supplementary Materia 13) like Type 2 Diabetes, Obesity, Polycystic ovary syndrome, higher level of plasma HDL cholesterol (Hypercholesterolemia), high level of triglycerides, Alzheimer's disease,

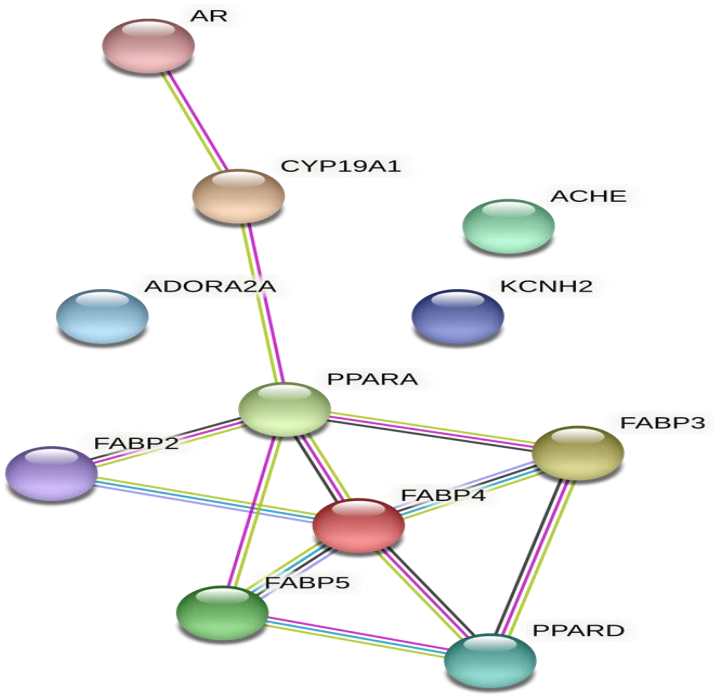


Fig 4.49. Phytocompound-target interaction and PPI network of *C. sinensis* Mature leaf (ML) associated proteins

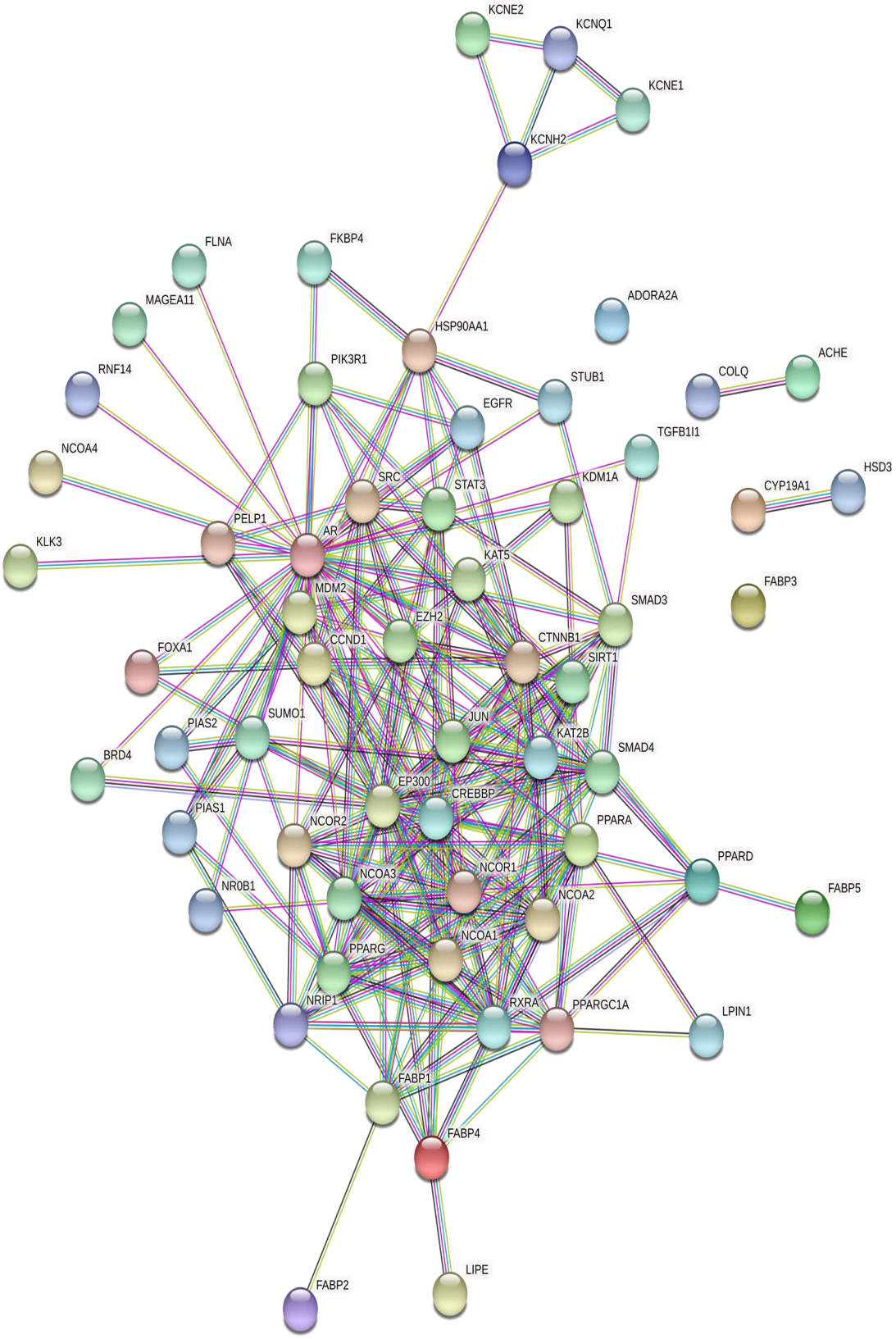


Fig 4.50. PPI network showing interaction of target proteins of Mature Tea Leaf (ML) derived compounds and other proteins of *Homo sapiens*

atherosclerosis, insulin resistance, metabolic syndrome, cardiovascular and coronary heart diseases, ovarian cancer, fatty liver, anxiety disorder, hypertension, and several other diseases. The top 15 results with p-value < 0.05 of Functional annotation of target genes is provided in Figure 4.52 which shows their association with diseases such as Type 2 diabetes, edema, Polycystic ovary syndrome, plasma HDL cholesterol (Hypercholesterolemia), coronary diseases, etc. Decreasing the size of the bubbles as well as the color gradiation from light pink to deep red indicates decrease in the the p-Value.

Among the 12 studied proteins, 11 proteins interacted with compounds associated with Type 2 diabetes and edema, 7 proteins were linked with Obesity, 4 proteins were linked with polycystic ovary syndrome, 5 proteins were linked with plasma HDL cholesterol, Alzheimer's disease, and hypertension (Supplementary Material 12).

Diabetes, cardiovascular disease, and cancer are the most frequent and complex diseases that are currently having a significant impact on global health issues that lead to death. Diabetic individuals are at a greater risk of heart problems, and cardiovascular complications are the primary cause of mortality in diabetic patients (Mirza *et al.,* 2019). Furthermore, type 2 diabetes can lead to an increase in cholesterol issues, blood pressure, and obesity, all of which can lead to cardiovascular difficulties. Aside from the increased risk of heart disease in diabetic people, researchers are looking into a possible link between diabetes and cancer. Millions of people have died as a result of cancer, which has now become a major global threat. Diabetes is thought to increase the incidence of certain malignancies, which may be attributed to hyperinsulinemia (Volkers, 2000). Diabetic people are at a much-increased risk for several forms of cancer. Type 2 diabetes has been linked to a variety of malignancies, and it has been claimed that treatments for type 2 diabetes may directly or indirectly affect cancer cells (Giovannucci *et al.,* 2010).

PCOS (polycystic ovarian syndrome) is a prevalent and complex endocrine disorder in women that is regarded as one of the leading causes of infertility (Cheshmeh *et al.,* 2021). PCOS is linked to hyperandrogenism, hyperinsulinemia, and hypothalamic–pituitary–adrenal axis dysfunction. Changes in the pituitary–ovarian axis, ovulation difficulties, and irregular menstruation affect these women, and they also experience mood swings,

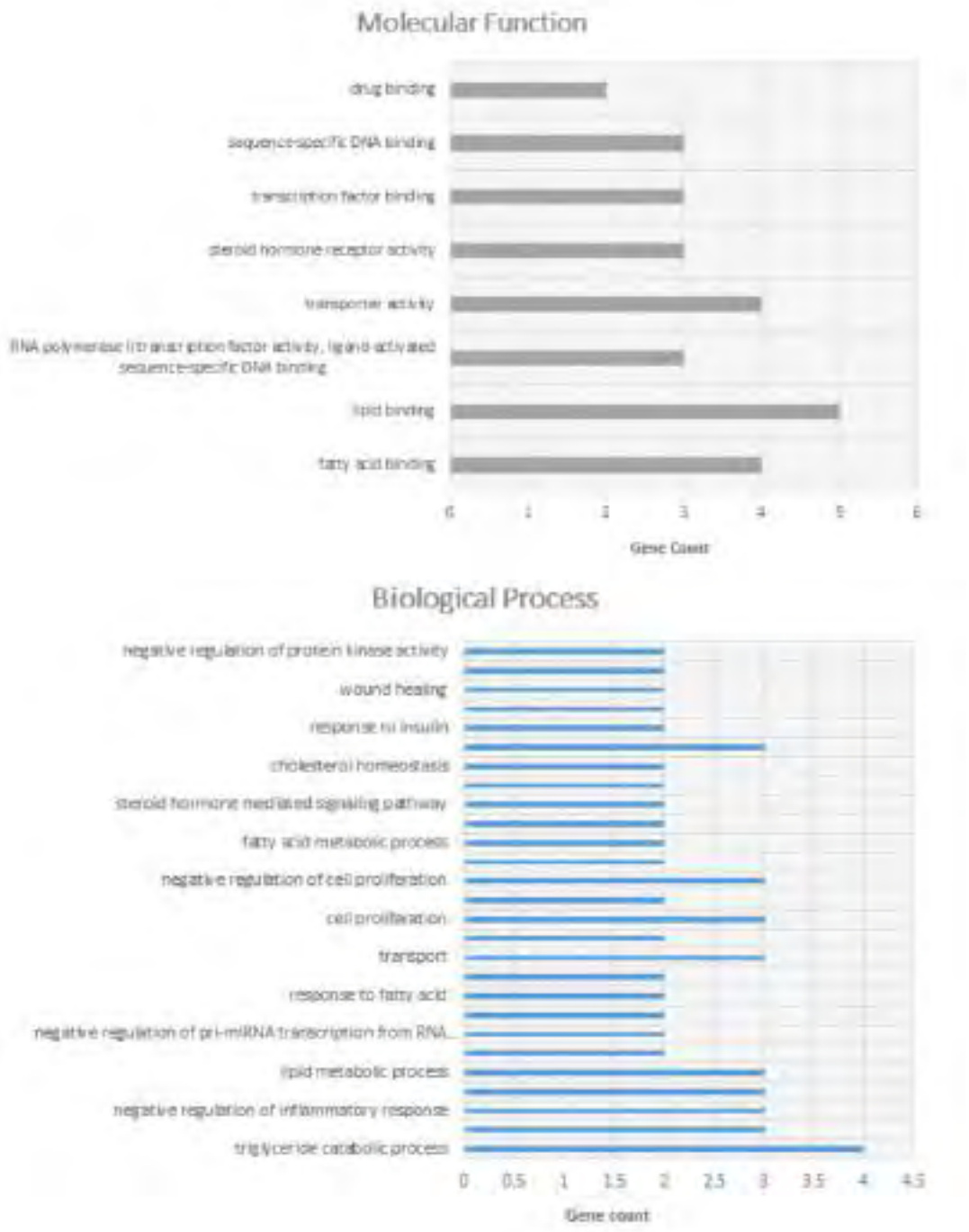


Fig 4.51. Gene Ontology Enrichment Analysis. The top 15 results (with p-value < 0.05) of Gene ontology enrichment analysis of the target genes show them to be associated with different Molecular Functions and Biological Processes

anxiety, despair, and mood changes (Machado *et al.,* 2020). Inter-linked metabolic disorders, such as weight gain and obesity, Insulin resistance, type 2 diabetes, and cardiovascular disease (Zhang *et al.,* 2020) are common in patients with PCOS. According to recent studies, genetic factors have a major role in the development of obesity and insulin

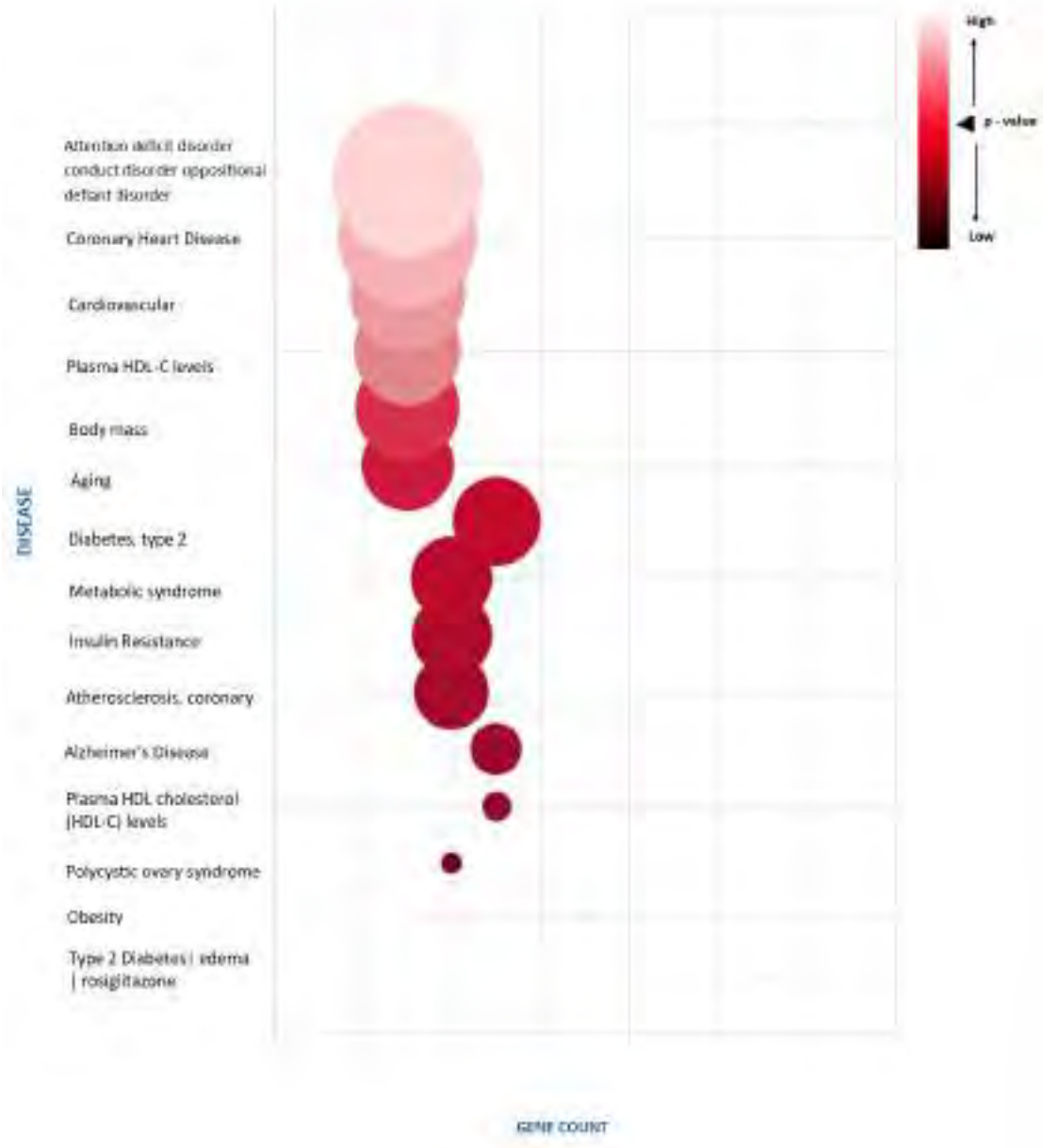


Fig 4.52. Functional Annotation (Disease). The top 15 results (with p-value < 0.05) of Functional annotation of target genes show their association with diseases such as plasma HDL cholesterol (Hypercholesterolemia), Type 2 diabetes, edema, Polycystic ovary syndrome, Obesity etc. Decreasing size of the bubbles as well as the color gradiation from light pink to deep red indicates decrease in the the p-value

resistance in PCOS patients (Khorshidi *et al.,* 2018). Furthermore, oxidative stress and an increase in inflammatory cytokines have been linked to the development of PCOS (Machado *et al.,* 2020).

4.8.2.4. Proteins related with PPAR signaling pathway and their role in several diseases

A total of 6 proteins (FABP2, FABP3, FABP4, FABP5, PPARA, PPARD), among the, studied 12 proteins were

found to be connected with PPAR (peroxisome proliferator-activated receptors) signaling pathway (Supplementary Material 13). The nuclear receptor superfamily includes the peroxisome proliferator-activated receptors (PPARs) which are activated by fatty acids and their derivatives. Invertebrates, PPAR has three subtypes namely PPAR alpha (PPARA), beta/ delta (PPARB/PPARD), and gamma (PPARG) with distinct expression patterns. Each of them binds fatty acids and eicosanoids and is encoded by a different gene. PPARalpha regulates gene expression associated with lipid metabolism in the liver and skeletal muscle, which aids in the clearance of circulating or cellular lipids. PPARbeta/delta is implicated in cell growth and lipid oxidation. PPARgamma enhances blood glucose absorption by promoting adipocyte differentiation (https:// www.genome.jp/entry/hsa03320).

The peroxisome proliferator-activated receptors (PPARA, PPARB/D, and PPARG), serve as ligand-inducible transcription factors and play important roles in glucose and lipid metabolism. These are well-known diabetes treatment receptor that not only affects the cardiovascular system but is also found in many human solid tumors. Peroxisome proliferator-activated receptors (PPARs) have a well-established role in a variety of chronic illnesses, including diabetes, cancer, inflammation, and atherosclerosis (Mirza *et al.,* 2019). It has a crucial function in maintaining lipid, glucose, and energy homeostasis (Towfighi and Ovbiagele, 2008). The PPARs are key therapeutic targets for atherosclerosis, inflammation, and hypertension and the careful and advanced design of a partial agonists of PPAR may also increase the therapeutic importance for several diseases like cardiovascular, cancer, and inflammation (Mirza *et al.,* 2019).

PPARA and other members of this family of receptors, such as PPARB/D and PPARG, are found in the brain and other organs and have a role in oxidative stress, energy balance, mitochondrial fatty acid metabolism, and inflammation. In the brain, PPAR- plays a role in the regulation of genes that code for proteins involved in glutamate homeostasis and cholinergic/ dopaminergic transmission. Downregulation of PPARA may reduce anti-oxidative and anti-inflammatory activities, and may be responsible for changes in fatty acid transport, lipid metabolism, and mitochondrial function in Alzheimer's disease (AD) patients' brains. In neurodegenerative and neurodevelopmental illnesses, specific PPAR-activators may help improve

brain cell metabolism and cognitive performance (Wojtowicz *et al.,* 2020).

PPARG is extensively expressed in adipose tissue and has a variety of activities, including adipocyte development, lipid storage, mitochondrial uncoupling protein expression, leptin downregulation, and insulin sensitivity modulation (Towfighi and Ovbiagele, 2008). They play a role in inflammatory processes and are involved in cell cycle regulation and insulin sensitivity development (Lehrke and Lazar, 2005). PPAR-agonists are thought to serve as negative regulators in T cell development, activating inflammatory responses and hence having a role in adaptive immunity (Sun *et al.,* 2017). PPARG regulates glucose homeostasis, lipid metabolism, and is a key therapeutic target for the treatment of type 2 diabetes as well as metabolic syndrome (Mirza *et al.,* 2019). It increases endothelial cell function by reducing inflammation in diabetes and atherosclerosis (Hsueh and Law, 2003). PPARG also plays an important function in cardiovascular disease. PPARG is abundantly expressed in atherosclerotic lesions, and its activation increases inflammatory effects in cardiovascular cells (Hamblin *et al.,* 2009). PPARG, in addition to its role in cardiovascular systems, has a strong affinity for tumor -related issues and plays an important role in cancer cell growth regulation. Many cancer cells have been found to contain it. Potent ligands that activate PPARG activation can inhibit cancer cell proliferation and differentiation. It plays a role in lipid metabolism as well as cancer cell growth. As a result, it could be used as a therapeutic target for cancer treatment (Mirza *et al.,* 2019). Activators of the PPARA and PPARG receptors can lower atherogenic triglycerides, enhance cardio-protective HDL levels, and improve insulin resistance simultaneously. Since PPARA and PPARD have significant fat-burning activity (Kersten *et al.,*2000) dual, selective, or triple agonists for PPAR-A/D/G are being investigated, since they can overcome the negative effects of PPARG agonists (Towfighi and Ovbiagele, 2008).■

Conclusion

"It is wiser to find out than to suppose."
– Mark Twain

Having been born and brought up among the beautiful and lusty tea gardens of the magnificent Darjeeling district, I always wanted to contribute my knowledge to my paradise. Since time immemorial, the local people have been taking so much effort in nurturing and maintaining the essence of Darjeeling tea. I have always wondered how they can handpick and trap the most mystical flavor because of which our Darjeeling tea is renowned worldwide.

Subsequently, I made an effort to contribute towards tea research so that my work could put some light and benefit to the tea industry in any way possible. I initiated the present endeavor in midst of 2017 taking elite tea clones established for Darjeeling and Dooars, to explore different areas of research in tea. Comprehensive morphological documentation of the clones was collated referring to the previous records from different sources so that it could benefit the readers or give necessary information.

The mature tea leaves are not used for tea manufacturing and it is usually discarded at the time of pruning. Therefore, we screened the phytochemicals in mature tea leaves employing a different range of solvents. Polar solvents like acetone, methanol, and ethanol were found to be the most potent solvents for extraction of various phytochemicals like flavonoid, coumarin, cardiac glycoside, diterpenes, terpenoids, steroid, saponin, tannin, and reducing sugar. The potency of a particular solvent in extracting a specific compound of interest was thus determined through detailed qualitative phytochemical

profiling. Acetone extracts outperformed by giving the best results in radical scavenging assay, ferric reducing power assay, total phenol, and flavonoid estimation. The extracts also showed potent antimicrobial and bactericidal activity against *Staphylococcus* sp. The powerful antibacterial component identified by GC-MS and *in silico* investigation was phenol 3, 5-bis (1, 1-dimethyl ethyl). The information gathered in this study could help future research to be more efficient in its use of resources.

While considering molecular study, a good deal of genetic diversity was observed among the studied tea clones. A total of 803 and 298 polymorphic bands were generated using several RAPD primers and ISSR primers with polymorphism percentages of 99.50% and 100% respectively.

DNA barcode analysis through matK revealed variation within the matK gene among the elite tea clones. The matK region was used to explore intraspecific variation because of its known enhanced rate of substitution, low transition/transversion ratio, and speedy evolving rate. Our findings show that the matK region has evolved within the same species, since we report varied sites within the consensus region, and some of our studied tea clones showed more similarity with the other species of *Camellia* and therefore, we conclude with the fact of matK gene not being 100% conserved within *Camellia sinensis*.

Tea, on the other hand has various forms and types and the most popular being green tea and black tea. Lots of research work has been carried in these two forms of tea. However, there are still some other forms of tea like white tea, oolong tea, yellow tea, purple tea, etc. which are yet to be explored. Therefore, upon survey and analysis, I picked up purple tea manufactured in Darjeeling and tried to explore the unchartered areas of nanoscience and went for the biogenic synthesis of silver and zinc nanoparticles. My findings suggested purple tea as a good candidate for synthesis of pure metal nanoparticles with antimicrobial activity but also conclude the absence of stability over 24 hours due to agglomeration and settling down of nanoparticles. Since I aimed to go for green synthesis of nanoparticles, no chemical surfactant was added to avoid agglomeration. Therefore, I conclude purple tea with high antioxidant activity to be a good reducing agent but not a good capping agent during the synthesis of nanoparticles.

With the advancement of genome sequence technology and with the availability of genome sequence of

C. sinensis, I further studied the factors governing its codon and amino acid usage. The factors dictating the codon usage signature in the nuclear genome of tea were mainly the nucleotide compositional constraint (AT richness) along with the role of other factors like mutational bias, translational selection, and gene expression level. The tea genome indicated the absence of replication-associated mutational pressure. The amino acid usage signatures on the other hand were dictated mainly by the factors like hydrophobicity, aromaticity, and gene expression level. Leucine (L) and Serine (S) were found to be the amino acids used in higher frequencies and in both cases Axis 1 was found to be the principal governing axis of codon and amino acid usage signatures.

Further, the phytochemistry and polypharmacological studies was employed to investigate role of purple tea as well as mature tea leaf against several diseases. Due to the ever-increasing demands of competitiveness and workload in the modern period, stress, sedentary habits, and poor diet may eventually lead to the urban population succumbing to different lifestyle-ailments such as obesity, diabetes, heart disease, and cancer as a result of elevated ROS generation. People consume numerous health-related foods such as tea, coffee, wine, and other beverages to reduce stress as well as the risk of such diseases. Purple tea consumption is the most recent addition to this trend. By opting for this healthy option, you will be able to reduce the bad health effects. Due to its significant antioxidant characteristics, which makes this anthocyanin-rich tea more efficient and unique, the beverage is steadily attracting commercial attention. The polypharmacology investigation via network pharmacology and target fishing studies provided insight into the mechanism of action of purple tea as well as mature tea leaf in managing the aforementioned disorders.■

Bibliography

Achaya, K. T. (1998). Indian Food Tradition: A Historical Companion. Oxford University Press.

Aiyegoro, O. A., & Okoh, A. I. (2009). Phytochemical screening and polyphenolic antioxidant activity of aqueous crude leaf extract of *Helichrysum pedunculatum*. International Journal of Molecular Sciences, 10(11), 4990-5001.

Akashi, H., & Gojobori, T. (2002). Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. Proceedings of the National Academy of Sciences, 99(6), 3695-3700.

Allen, G. A., Soltis, D. E., & Soltis, P. S. (2003). Phylogeny and biogeography of *Erythronium* (Liliaceae) inferred from chloroplast matK and nuclear rDNA ITS sequences. Systematic Botany, 28(3), 512-523.

Al-Salih, D. A. A. K., Aziz, F. M., Mshimesh, B. A. R., & Jehad, M. T. (2013). Antibacterial effects of vitamin E: In vitro study. Jornal of Biotechnology Research Center, 7(2).

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403-410.

Anandappa, T. I. (1993). The breeding selection and testing of tea clones.

Angiosperm Phylogeny Group, Chase, M. W., Christenhusz, M. J., Fay, M. F., Byng, J. W., Judd, W. S., ... & Stevens, P. F. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Botanical Journal of the Linnean Society, 181(1), 1-20.

Asghar, M. A., Zahir, E., Shahid, S. M., Khan, M. N., Asghar, M. A., Iqbal, J., & Walker, G. (2018). Iron, copper and silver nanoparticles: Green synthesis using green and black tea leaves extracts and evaluation of antibacterial, antifungal and aflatoxin B1 adsorption activity. Lwt, 90, 98-107.

B. Banerjee, “Botanical classification of tea, in Tea cultivation to consumption,” in Chapman Hall publication, K. C. Wilson and and M. N. Clifford, Eds., pp. 555–601, London, 1992

Bafeel, S. O., Arif, I. A., Bakir, M. A., Khan, H. A., Al Farhan, A. H., Al Homaidan, A. A., ... & Thomas, J. (2011). Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose -1, 5-bisphosphate carboxylase oxygenase large subunit (rbcL) for barcoding of some arid plants. Plant Omics, 4(4), and 195.

Balasaravanan, T., Pius, P. K., Kumar, R. R., Muraleedharan, N., & Shasany, A. K. (2003). Genetic diversity among south Indian tea germplasm (*Camellia sinensis*, *C. assamica* and *C. assamica* spp. *lasiocalyx*) using AFLP markers. Plant Science, 165(2), 365-372.

Banerjee, B. (1992). Botanical classification of tea. In Tea (pp. 25-51). Springer, Dordrecht.

Bar, H., Bhui, D. K., Sahoo, G. P., Sarkar, P.,

Pyne, S., & Misra, A. (2009). Green synthesis of silver nanoparticles using seed extract of *Jatropha curcas*. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 348(1-3), 212-216.

Baruah, A. R., Hemanta, S., & Biswajit, B. (2010). Detection of close genetic relatedness in some tea genotypes of Assam and Darjeeling using RAPD markers. Journal of Plantation Crops, 38 (1), 11-15.

Begum, N. A., Mondal, S., Basu, S., Laskar, R. A., & Mandal, D. (2009). Biogenic synthesis of Au and Ag nanoparticles using aqueous solutions of Black Tea leaf extracts. Colloids and surfaces B: Biointerfaces, 71(1), 113-118.

Benn, J. A. (2015). Tea in China: A religious and cultural history. University of Hawaii Press.

Bezbaruah, H. P. (1999). Origin and history of development of tea. Global advance in tea science.

Biswas, K. P. (2006). Description of tea plant. Encyclopaedia of Medicinal Plants, 964-966.

Blois, M. S. (1958). Antioxidant determinations by the use of a stable free radical. Nature, 181(4617), 1199-1200.

Book of Tea by Kakuzo Okakura (pp. 5–6). Published 1964. Courier Dover Publications. Sociology. 94 pages. ISBN 0 -486-20070-1

Brain, K. R., & Turner, T. D. (1975). The practical evaluation of phytopharmaceuticals (Vol. 1). Bristol: Wright-Scientechnica.

Briner, A., Gotz, J., & Polanco, J. C. (2020). Fyn kinase controls tau aggregation in vivo. Cell Reports, 32(7), 108045.

Caprar, M., Copaci, C. M., Chende, D. M., Sicora, O., Sumalan, R., & Sicora, C. (2017). Evaluation of Genetic Diversity by DNA Barcoding of Local Tomato Populations from North-Western Romania. Notulae Botanicae Horti Agrobotanici Cluj-Napoca, 45(1), 276-279.

Casimir, J., Jadot, J., & Renard, M. (1960). Separation et caracterisation de la N-ethyl-γ-glutamine a partir de *Xerocomus badius*. Biochimica et biophysica acta, 39 (3), 462-468.

Chan, E. W. C., Lim, Y. Y., & Chew, Y. L. (2007). Antioxidant activity of *Camellia sinensis* leaves and tea from a lowland plantation in Malaysia. Food Chemistry, 102(4), 1214-1222.

Chen, J., Wang, P., Xia, Y., Xu, M., & Pei, S. (2005). Genetic diversity and differentiation of *Camellia sinensis* L. (cultivated tea) and its wild relatives in Yunnan province of China, revealed by morphology, biochemistry and allozyme studies. Genetic Resources and Crop Evolution, 52(1), 41-52.

Chen, L., Gao, Q. K., Chen, D. M., & Xu, C. J. (2005). The use of RAPD markers for detecting genetic diversity, relationship and molecular identification of Chinese elite tea genetic resources [*Camellia sinensis* (L.) O. Kuntze] preserved in a tea germplasm repository. Biodiversity & Conservation, 14(6), 1433-1444.

Chen, Z. M., & Chen, L. (2012). Delicious and healthy tea: an overview. In Global tea breeding (pp. 1-11). Springer, Berlin, Heidelberg.

Cheshmeh, S., Elahi, N., Ghayyem, M., Mosayebi, E., Moradi, S., Pasdar, Y., & Tahmasebi, S. (2021). Effects of green cardamom supplementation on obesity and diabetes gene expression among obese women with polycystic ovary syndrome; A double blind randomized controlled trial.

Chopade, V., Phatak, A., Upaganlawar, A., & Tankar, A. (2008). Green tea (*Camellia sinensis*): Chemistry, traditional, medicinal uses and its pharmacological activities-a review. Pharmacognosy Reviews, 2(3), 157.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Research, 16(22), 10881-10890.

Dai, X., Yan, X., Zeng, J., Chen, J., Wang, Y., Chen, J., ... & Tan, Y. (2017). Elevating CXCR7 improves angiogenic function of EPCs via Akt/GSK-3β/Fyn-mediated Nrf2 activation in diabetic limb ischemia. Circulation Research, 120(5), e7-e23.

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness, and medicinal chemistry friendliness of small molecules. Scientific Reports, 7(1), 1-13.

Das, A. P. *Camellia sinensis* var. *lasiocalyx* (G. Watt) AP Das & C. Ghosh–new combination name for the Cambod variety of tea.

Das, S., Paul, S., & Dutta, C. (2006). Evolutionary constraints on codon and amino acid usage in two strains of human pathogenic actinobacteria *Tropheryma* whipplei. Journal of Molecular Evolution, 62(5), 645-658.

Das, S., Vasudeva, N., & Sharma, S. (2014). Chemical composition of ethanol extract of *Macrotyloma uniflorum* (Lam.) Verdc. using GC-MS spectroscopy. Organic and Medicinal Chemistry Letters, 4(1), 1-4.

Devarumath, R., Nandy, S., Rani, V., Marimuthu, S., Muraleedharan, N., & Raina, S. (2002). RAPD, ISSR and RFLP fingerprints as useful markers to evaluate genetic integrity of micropropagated plants of three diploid and triploid elite tea clones representing *Camellia sinensis* (China type) and *C. assamica* ssp. *assamica* (Assam-India type). Plant Cell Reports, 21 (2), 166-173.

Dey, P., & Chaudhuri, T. K. (2016). Phytochemical Characterization of *D ioscorea alata* Leaf and Stem By Silylation Followed by GC-MS Analysis. Journal of Food Biochemistry, 40(4), 630-635.

Dhanya, K. I., Swati, V. I., Vanka, K. S., & Osborne, W. J. (2016). Antimicrobial activity of *Ulva reticulata* and its endophytes. Journal of Ocean University of China, 15(2), 363-369.

Downward, J. (2004, April). PI 3-kinase, Akt and cell survival. In Seminars in cell & developmental biology (Vol. 15, No. 2, pp. 177-182). Academic Press.

Doyle J. J., and Doyle J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11-15.

Dufresne, C. J., & Farnworth, E. R. (2001). A review of latest research findings on the health promotion properties of tea. The Journal of Nutritional Biochemistry, 12(7), 404-421.

Duret, L., & Mouchiroud, D. (1999). Expression pattern and, surprisingly, gene length shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*. Proceedings of the National Academy of Sciences, 96(8), 4482-4487.

Esimone, C. O., Okoye, F. B., Nworu, C. S., & Agubata, C. O. (2008). In vitro interaction between caffeine and some penicillin antibiotics against *Staphylococcus aureus*. Tropical Journal of Pharmaceutical Research, 7(2), 969-974.

Eyre-Walker, A. (1996). Synonymous codon bias is related to gene length in *Escherichia coli*: selection for translational accuracy? Molecular biology and evolution, 13(6), 864-872.

FAOSTAT (2015) FAO database. Food Agric Organ United Nations. URL http:// faostat3.fao.org/download/Q/QC/E(2015). Accessed 19 November 2015

Fiske, W. H., Threadgill, D., & Coffey, R. J. (2009). ERBBs in the gastrointestinal tract: recent progress and new perspectives. Experimental Cell Research, 315(4), 583-601.

Flindt, M. L. H. (1969). Pulmonary disease due to inhalation of derivatives of *Bacillus subtilis* containing proteolytic enzyme. The Lancet, 293(7607), 1177-1181.

Folin, O., & Ciocalteu, V. (1927). On tyrosine and tryptophane determinations in proteins. Journal of Biological Chemistry, 73(2), 627-650.

Futcher, B., Latter, G. I., Monardo, P., McLaughlin, C. S., & Garrels, J. I. (1999). A sampling of the yeast proteome. Molecular and Cellular Biology, 19(11), 7357.

Geoffrey, K. K., John, K. M., Naomi, M., & Simon, K. M. (2014). Qualitative phytochemical screening of *Camellia sinensis* and *Psidium guajava* leaves extracts from Kericho and Baringo counties. International Journal of Advanced Biotechnology and Research, 5 (3), 506-512.

Gfeller, D., Grosdidier, A., Wirth, M., Daina, A., Michielin, O., & Zoete, V. (2014). SwissTargetPrediction: a web server for target prediction of bioactive small molecules. Nucleic Acids Research, 42 (W1), W32-W38.

Giovannucci, E., Harlan, D. M., Archer, M. C., Bergenstal, R. M., Gapstur, S. M., Habel, L. A., ... & Yee, D. (2010). Diabetes and cancer: a consensus report. Diabetes Care, 33(7), 1674-1685.

Gottimukkala, K. S. V., Harika, R. P., & Zamare, D. (2017). Green synthesis of iron nanoparticles using green tea leaves extract. Journal of Nanomedicine and Biotherapeutic Discovery, 7, 151.

Group, C. P. W., Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., ... & Little, D. P. (2009). A DNA barcode for land plants. Proceedings of the National Academy of Sciences, 106(31), 12794-12797.

Gruenwald, J., Brendler, T., & Jaenicke, C. (2007). PDR for Herbal Medicines. Montvale, NJ: Thomson Healthcare.

Gunasekare, M. T. K. (2007). Applications of molecular markers to the genetic improvement of *Camellia sinensis* L.(tea)– a review. The Journal of Horticultural Science and Biotechnology, 82(2), 161-169.

Guo, F. B., & Yuan, J. B. (2009). Codon usages of genes on chromosome, and surprisingly, genes in plasmid are primarily affected by strand-specific mutational biases in *Lawsonia intracellularis*. DNA research, 16(2), 91-104.

Gupta, S. K., & Ghosh, T. C. (2001). Gene expressivity is the main factor in dictating the codon usage variation among the genes in *Pseudomonas aeruginosa*. Gene, 273 (1), 63-70.

Hajra, N. G. (2001). Advances in selection and breeding of tea-a review. Journal of Plantation Crops, 29(3), 1-17.

Hamblin, M., Chang, L., Fan, Y., Zhang, J., & Chen, Y. E. (2009). PPARs and the cardiovascular system. Antioxidants & Redox Signaling, 11(6), 1415-1452.

Harris, I. S., & DeNicola, G. M. (2020). The complex interplay between antioxidants and ROS in cancer. Trends in Cell

Biology, 30(6), 440-451.

Heiss, M. L., & Heiss, R. J. (2007). The story of tea: a cultural history and drinking guide. Random House Digital, Inc.

Heizer Jr, E. M., Raiford, D. W., Raymer, M. L., Doom, T. E., Miller, R. V., & Krane, D. E. (2006). Amino acid cost and codon-usage biases in 6 prokaryotic genomes: a whole-genome analysis. Molecular biology and evolution, 23(9), 1670-1680.

Hilal, Y. (2017). Morphology, manufacturing, types, composition and medicinal properties of tea (*Camellia sinensis*). Journal of Basic and Applied Plant Sciences, 1(2), 107.

Hoffmann, F. W., & Manning, M. (2002). Herbal medicine and botanical medical fads. Psychology Press.

Holbro, T., & Hynes, N. E. (2004). ErbB receptors: directing key signaling networks throughout life. Annual Review of Pharmacology & Toxicology, 44, 195-217.

Horn, T., & Haser, A. (2016). Bamboo tea: reduction of taxonomic complexity and application of DNA diagnostics based on rbcL and matK sequence data. PeerJ, 4, e2781.

Hou, Z. C., & Yang, N. (2002). Analysis of factors shaping *S. pneumoniae* codon usage. Yi chuan xue bao= Acta genetica Sinica, 29(8), 747-752.

Hsueh, W. A., & Law, R. (2003). The central role of fat and effect of peroxisome proliferator-activated receptor–γ on progression of insulin resistance and cardiovascular disease. The American Journal of Cardiology, 92(4), 3-9.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research, 37(1), 1-13.

Huang, L., Weng, X., Chen, Z., Megharaj, M., & Naidu, R. (2014). Green synthesis of iron nanoparticles by various tea extracts: comparative study of the reactivity. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 130, 295-301.

Hyun, D. Y., Gi, G. Y., Sebastin, R., Cho, G. T., Kim, S. H., Yoo, E., & Lee, K. J. (2020). Utilization of Phytochemical and Molecular Diversity to Develop a Target-Oriented Core Collection in Tea Germplasm. Agronomy, 10(11), 1667.

Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. Molecular Biology and Evolution, 2(1), 13-34.

Isaac, R. S., Sakthivel, G., & Murthy, C. H. (2013). Green synthesis of gold and silver nanoparticles using *Averrhoa bilimbi* fruit extract. Journal of Nanotechnology, 2013.

Jagtap, U. B., & Bapat, V. A. (2013). Green synthesis of silver nanoparticles using *Artocarpus heterophyllus* Lam. seed extract and its antibacterial activity. Industrial crops and products, 46, 132-137.

Joshi, R., Rana, A., Kumar, V., Kumar, D., Padwad, Y. S., Yadav, S. K., & Gulati, A. (2017). Anthocyanins enriched purple tea exhibits antioxidant, immunostimulatory and anticancer activities. Journal of Food Science and Technology, 54(7), 1953-1963.

Jung, M. Y., Kim, B. S., Kim, Y. J., Koh, I. S., & Chung, J. H. (2008). Assessment of relationship between Fyn-related kinase gene polymorphisms and overweight/ obesity in Korean population. The Korean Journal of Physiology & Pharmacology, 12(2), 83-87.

Kahali, B., Basak, S., & Ghosh, T. C. (2007). Reinvestigating the codon and amino acid

usage of *S. cerevisiae* genome: a new insight from protein secondary structure analysis. Biochemical and biophysical research communications, 354(3), 693-699.

Kang, H., Zheng, W., Kong, Z., Jiang, F., Gu, B., Ma, P., & Ma, X. (2020). Disease burden and molecular epidemiology of carbapenem-resistant *Klebsiella pneumonia* infection in a tertiary hospital in China. Annals of translational medicine, 8(9).

Karlin, S., & Mrázek, J. (1996). What drives codon choices in human genes? Journal of molecular biology, 262(4), 459-472.

Karthigeyan, S., & Sud, R. K. (2010). Genetic diversity of Indian tea (*Camellia* sinensis (L.O) Kuntze) germplasm detected using morphological characteristics. Journal Cell and Plant Science, 1(1), 13-22.

Katoh, Y., Katoh, M., Takeda, Y., & Omori, M. (2003). Genetic diversity within cultivated teas based on nucleotide sequence comparison of ribosomal RNA maturase in chloroplast DNA. Euphytica, 134(3), 287-295.

Kaundun, S. S., & Park, Y. G. (2002). Genetic Structure of Six Korean Tea Populations as Revealed by RAPD-PCR Markers. Crop Science, 42(2), 594-601.

Kaundun, S. S., Zhyvoloup, A., & Park, Y. G. (2000). Evaluation of the genetic diversity among elite tea (*Camellia sinensis* var. *sinensis*) accessions using RAPD markers. Euphytica, 115(1), 7-16.

Kersten, S., Desvergne, B., & Wahli, W. (2000). Roles of PPARs in health and disease. Nature, 405(6785), 421-424.

Khorshidi, M., Moini, A., Alipoor, E., Rezvan, N., Gorgani-Firuzjaee, S., Yaseri, M., & Hosseinzadeh-Attar, M. J. (2018). The effects of quercetin supplementation on metabolic and hormonal parameters as well as plasma concentration and gene expression of resistin in overweight or obese women with polycystic ovary syndrome. Phytotherapy Research, 32(11), 2282-2289.

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., ... & Bolton, E. E. (2021). PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Research, 49(D1), D1388-D1395.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution, 16(2), 111-120.

Koonin, E. V., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Krylov, D. M., Makarova, K. S., ... & Natale, D. A. (2004). A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. Genome Biology, 5(2), 1-28.

Kumar, A., Ilavarasan, R., Jayachandran, T., Decaraman, M., Aravindhan, P., Padmanabhan, N., & Krishnan, M. R. V. (2009). Phytochemicals investigation on a tropical plant, *Syzygium cumini* from Kattuppalayam, Erode district, Tamil Nadu, South India. Pakistan Journal of Nutrition, 8(1), 83-85.

Kumar, H., & Rani, R. (2013). Structural and optical characterization of ZnO nanoparticles synthesized by microemulsion route. International Letters of Chemistry, Physics and Astronomy, 14, 26-36.

Kumar, V., Singh, D. K., Mohan, S., Gundampati, R. K., & Hasan, S. H. (2017). Photoinduced green synthesis of silver nanoparticles using aqueous extract of *Physalis angulata* and its antibacterial and antioxidant activity. Journal of Environmental Chemical Engineering, 5

(1), 744-756.

Kuzmina, M. L., Johnson, K. L., Barron, H. R., & Hebert, P. D. (2012). Identification of the vascular plants of Churchill, Manitoba, using a DNA barcode library. BMC Ecology, 12(1), 1-11.

Labar, R., Sarkar, I., Sen, A., & Bhattacharya, M. (2019). Effect of solvent with varying polarities on phytochemical extraction from mature tea leaves and its evaluation using biochemical, antimicrobial and *in-silico* approaches. International Research Journal of Pharmacy, 10(8), 59-67.

Lafay, B., Sharp, P. M., Lloyd, A. T., McLean, M. J., Devine, K. M., & Wolfe, K. H. (1999). Proteome composition and codon usage in spirochaetes: species-specific and DNA strand-specific mutational biases. Nucleic Acids Research, 27(7), 1642-1649.

Lai, J. A., Yang, W. C., & Hsiao, J. Y. (2001). An assessment of genetic relationships in cultivated tea clones and native wild tea in Taiwan using RAPD and ISSR markers. Botanical Bulletin of Academia Sinica, 42.

Lawodi, E. N. (2013). 2. Variasi genetik tanaman tomat dari beberapa tempat di sulawesi utara berdasarkan gen matK. Pharmacon, 2(4).

Lebaschi, S., Hekmati, M., & Veisi, H. (2017). Green synthesis of palladium nanoparticles mediated by black tea leaves (*Camellia sinensis*) extract: catalytic activity in the reduction of 4-nitrophenol and Suzuki-Miyaura coupling reaction under ligand-free conditions. Journal of Colloid and Interface Science, 485, 223-231.

Lehrke, M., & Lazar, M. A. (2005). The many faces of PPARγ. Cell, 123(6), 993-999.

Lin, Y. S., Tsai, Y. J., Tsay, J. S., & Lin, J. K. (2003). Factors affecting the levels of tea polyphenols and caffeine in tea leaves. Journal of Agricultural and food Chemistry, 51(7), 1864-1873.

Lithwick, G., & Margalit, H. (2005). Relative predicted protein levels of functionally associated proteins are conserved across organisms. Nucleic Acids Research, 33(3), 1051-1057.

Liu, G., Wu, J., Yang, H., & Bao, Q. (2010). Codon usage patterns in *Corynebacterium glutamicum*: mutational bias, natural selection and amino acid conservation. Comparative and Functional Genomics, 2010.

Liu, Q., & Xue, Q. (2005). Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species. Journal of Genetics, 84 (1), 55-62.

Long, P., Cui, Z., Wang, Y., Zhang, C., Zhang, N., Li, M., & Xiao, P. (2014). Commercialized non- *Camellia* tea: traditional function and molecular identification. Acta Pharmaceutica Sinica B, 4(3), 227-237.

Loo, Y. Y., Chieng, B. W., Nishibuchi, M., & Radu, S. (2012). Synthesis of silver nanoparticles by using tea leaf extract from *Camellia sinensis*. International Journal of Nanomedicine, 7, 4263.

Lv, H. P., Dai, W. D., Tan, J. F., Guo, L., Zhu, Y., & Lin, Z. (2015). Identification of the anthocyanins from the purple leaf coloured tea cultivar Zijuan (*Camellia sinensis* var. *assamica*) and characterization of their antioxidant activities. Journal of Functional Foods, 17, 449-458.

Machado, V., Escalda, C., Proenca, L., Mendes, J. J., & Botelho, J. (2020). Is there a bidirectional association between polycystic ovarian syndrome and periodontitis? A systematic review and meta-analysis. Journal of Clinical

Medicine, 9(6), 1961.

Mahmood, T., Akhtar, N., & Khan, B. A. (2010). The morphology, characteristics, and medicinal properties of *Camellia sinensis* tea. Journal of Medicinal Plants Research, 4(19), 2028-2033.

Mair, Victor H.; Hoh, Erling (2009). The True History of Tea. Thames & Hudson. ISBN 978-0-500-25146-1.

Marais, G., & Duret, L. (2001). Synonymous codon usage, accuracy of translation, and gene length in *Caenorhabditis elegans*. Journal of Molecular Evolution, 52(3), 275-280.

Martinez, L., Cavagnaro, P., Masuelli, R., & Rodriguez, J. (2003). Evaluation of diversity among Argentine grapevine (*Vitis vinifera* L.) varieties using morphological data and AFLP markers. Electronic Journal of Biotechnology, 6(3), 244-253.

McInerney, J. O. (1998). Replicational and transcriptional selection on codon usage in *Borrelia burgdorferi*. Proceedings of the National Academy of Sciences, 95(18), 10698-10703.

Meegahakumbura, M. K., Wambulwa, M. C., Li, M. M., Thapa, K. K., Sun, Y. S., Moller, M., ... & Li, D. Z. (2018). Domestication origin and breeding history of the tea plant (*Camellia sinensis*) in China and India based on nuclear microsatellites and cpDNA sequence data. Frontiers in Plant Science, 8, 2270.

Meruvu, S., Hugendubler, L., & Mueller, E. (2011). Regulation of adipocyte differentiation by the zinc finger protein ZNF638. Journal of Biological Chemistry, 286(30), 26516-26523.

Ming, T. L., Gu, Z. J., Zhang, W. J., & Xie, L. S. (2000). Monograph of the Genus *Camellia*. Yunnan Science and Technology Press. Kunming, China, 128-134.

Mingzhe, Y., Liang, C., Xinchao, W., Liping, Z., & Yajun, Y. (2010). Genetic diversity and relationship of clonal tea cultivars in China revealed by ISSR markers.

Mirsepasi-Lauridsen, H. C., Vallance, B. A., Krogfelt, K. A., & Petersen, A. M. (2019). *Escherichia coli* pathobionts associated with inflammatory bowel disease. Clinical Microbiology Reviews, 32(2), e00060-18.

Mirza, A. Z., Althagafi, I. I., & Shamshad, H. (2019). Role of PPAR receptor in different diseases and their ligands: Physiological importance and clinical implications. European Journal of Medicinal Chemistry, 166, 502-513.

Mishra, R. K., & Sen-Mandi, S. (2004). Genetic diversity estimates for Darjeeling tea clones based on amplified fragment length polymorphism markers. Journal of Tea Science, 24(2), 86-92.

Miyasaka, H. (2002). Translation initiation AUG context varies with codon usage bias and gene length in *Drosophila melanogaster*. Journal of Molecular Evolution, 55(1).

Mock, J. J., Barbic, M., Smith, D. R., Schultz, D. A., & Schultz, S. (2002). Shape effects in plasmon resonance of individual colloidal silver nanoparticles. The Journal of Chemical Physics, 116(15), 6755-6759.

Mondal, T. K. (2002). Assessment of genetic diversity of tea (*Camellia sinensis* (L.) O. Kuntze) by inter-simple sequence repeat polymerase chain reaction. Euphytica, 128 (3), 307-315.

Mondal, T. K., Bhattacharya, A., Laxmikumaran, M., & Ahuja, P. S. (2004). Recent advances of tea (*Camellia sinensis*) biotechnology. Plant Cell, Tissue and Organ Culture, 76(3), 195-254.

Mondal, T.K. Tea. In Breeding Plantation Tree

crops Tropical Species (Eds. M. Prydarsini & S.M. Jain) (ISBN: 978-0-387-71199-7) 545–587 (Springer USA, New York, 2009).

More, R. P., Mane, R. C., & Purohit, H. J. (2016). matK-QR classifier: a patterns-based approach for plant species identification. BioData mining, 9(1), 1-15.

Moriyama, E. N., & Powell, J. R. (1997). Codon usage bias and tRNA abundance in *Drosophila*. Journal of Molecular Evolution, 45(5), 514-523.

Moriyama, E. N., & Powell, J. R. (1998). Gene length and codon usage bias in *Drosophila melanogaster*, *Saccharomyces cerevisiae* and *Escherichia coli*. Nucleic Acids Research, 26(13), 3188-3193.

Moulton, M. C., Braydich-Stolle, L. K., Nadagouda, M. N., Kunzelman, S., Hussain, S. M., & Varma, R. S. (2010). Synthesis, characterization and biocompatibility of "green" synthesized silver nanoparticles using tea polyphenols. Nanoscale, 2(5), 763-770.

Mukhopadhyay, M., Mondal, T. K., & Chand, P. K. (2016). Biotechnological advances in tea (*Camellia sinensis* [L.] O. Kuntze): a review. Plant Cell Reports, 35(2), 255-287.

Mukhtar, H., & Ahmad, N. (2000). Tea polyphenols: prevention of cancer and optimizing health. The American Journal of Clinical Nutrition, 71(6), 1698S-1702S.

Nair, R. R., Nandhini, M. B., Monalisha, E., Murugan, K., Sethuraman, T., Nagarajan, S., & Ganesh, D. (2012). Synonymous codon usage in chloroplast genome of *Coffea arabica*. Bioinformation, 8(22), 1096.

Nakamura, K., & Matsunaga, K. I. (1998). Susceptibility of natural killer (NK) cells to reactive oxygen species (ROS) and their restoration by the mimics of superoxide dismutase (SOD). Cancer Biotherapy & Radiopharmaceuticals, 13(4), 275-290.

Nakhjavani, M., Nikkhah, V., Sarafraz, M. M., Shoja, S., & Sarafraz, M. (2017). Green synthesis of silver nanoparticles using green tea leaves: Experimental study on the morphological, rheological and antibacterial behaviour. Heat and Mass Transfer, 53(10), 3201-3209.

Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences, 76(10), 5269-5273.

Ngbede, J., Yakubu, R. A., & Nyam, D. A. (2008). Phytochemical screening for active compounds in *Canarium schweinfurthii* (Atile) leaves from Jos North, Plateau State, Nigeria. Res J Biol Sci, 3(9), 1076-1078.

Ozarkar, K. R. (2005). Studies on anti-inflammatory effects of two herbs *Cissus quadrangularis* Linn. and *Valeriana wallichi* DC using mouse model. University of Mumbai, Mumbai.

Patel, Salil H. "*Camellia sinensis*: historical perspectives and future prospects." Journal of Agromedicine 10, no. 2 (2005): 57-64.

Patil, M. P., Patil, K. T., Ngabire, D., Seo, Y. B., & Kim, G. D. (2016). Phytochemical, antioxidant and antibacterial activity of black tea (*Camellia sinensis*). International Journal of Pharmacognosy and Phytochemical Research, 8(2), 341-346.

Paul, S., Wachira, F. N., Powell, W., & Waugh, R. (1997). Diversity and genetic differentiation among populations of Indian and Kenyan tea (*Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. Theoretical and Applied Genetics, 94(2), 255-263.

Perriere, G., & Gouy, M. (1996). WWW-

query: an on-line retrieval system for biological sequence banks. Biochimie, 78 (5), 364-369.

Philip, D. (2010). Rapid green synthesis of spherical gold nanoparticles using *Mangifera indica* leaf. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 77(4), 807-810.

Piyasundara, J. H. N., Gunasekare, M. T. K., Peiris, T. U. S., & Wickramasinghe, I. P. (2006). Phenotypic diversity of Sri Lankan tea (*Camellia sinensis* L.) germplasm based on morphological descriptors.

Polvani, S., Tarocchi, M., Tempesti, S., Bencini, L., & Galli, A. (2016). Peroxisome proliferator activated receptors at the crossroad of obesity, diabetes, and pancreatic cancer. World Journal of Gastroenterology, 22(8), 2441.

Purushothaman, N., Newmaster, S. G., Ragupathy, S., Stalin, N., Suresh, D., Arunraj, D. R.,& Parani, M. (2014). A tiered barcode authentication tool to differentiate medicinal *Cassia* species in India. Genetics and Molecular Research, 13(2), 2959-68.

Raghupathi, K. R., Koodali, R. T., & Manna, A. C. (2011). Size-dependent bacterial growth inhibition and mechanism of antibacterial activity of zinc oxide nanoparticles. Langmuir, 27(7), 4020-4028.

Ragupathy, S., Newmaster, S. G., Murugesan, M., & Balasubramaniam, V. (2009). DNA barcoding discriminates a new cryptic grass species revealed in an ethnobotany study by the hill tribes of the Western Ghats in southern India. Molecular Ecology Resources, 9, 164-171.

Rahman, S. U., Li, Y., Huang, Y., Zhu, L., Feng, S., Wu, J., & Wang, X. (2018). Treatment of inflammatory bowel disease via green tea polyphenols: Possible application and protective approaches. Inflammopharmacology, 26 (2), 319-330.

Raiford, D. W., Heizer, E. M., Miller, R. V., Akashi, H., Raymer, M. L., & Krane, D. E. (2008). Do amino acid biosynthetic costs constrain protein evolution in *Saccharomyces cerevisiae*? Journal of molecular evolution, 67(6), 621-630.

Rajendran, S. P., & Sengodan, K. (2017). Synthesis and characterization of zinc oxide and iron oxide nanoparticles using *Sesbania grandiflora* leaf extract as reducing agent. Journal of Nanoscience, 2017.

Ranatunga, M. A. B., & Gunesekera, M. T. K. (2008). Assembling of preliminary core collection of tea (*Camellia sinensis* (L.) O. Kuntze) genetic resources in Sri Lanka. Plant Genetic Resources Newsletter, 155, 41-45.

Ranatunga, M. A. B., Arachchi, J. D. K., Gunasekare, K., & Yakandawala, D. (2017). Floral diversity and genetic structure of tea germplasm of Sri Lanka. International Journal of Biodiversity, 2017, 1-11.

Rawal, H. C., Borchetia, S., Bera, B., Soundararajan, S., Ilango, R. V. J., Barooah, A. K., & Mondal, T. K. (2021). Comparative analysis of chloroplast genomes indicated different origin for Indian tea (*Camellia assamica* cv TV1) as compared to Chinese tea. Scientific Reports, 11(1), 1-6.

Rohlf, F. J. (1998). NTSYS-pc version 2.0. Numerical taxonomy and multivariate analysis system. Exeter software, Setauket, New York.

Rojo, L., Sjöberg, M. K., Hernández, P., Zambrano, C., & Maccioni, R. B. (2006). Roles of cholesterol and lipids in the etiopathogenesis of Alzheimer's disease.

Journal of Biomedicine and Biotechnology, 2006.

Rolim, W. R., Pelegrino, M. T., de Araujo Lima, B., Ferraz, L. S., Costa, F. N., Bernardes, J. S., ... & Seabra, A. B. (2019). Green tea extract mediated biogenic synthesis of silver nanoparticles: characterization, cytotoxicity evaluation and antibacterial activity. Applied Surface Science, 463, 66-74.

Romero, H., Zavala, A., & Musto, H. (2000). Codon usage in *Chlamydia trachomatis* is the result of strand-specific mutational biases and a complex pattern of selective forces. Nucleic acids research, 28(10), 2084-2090.

Roy, A., Mukhopadhyay, S., Sarkar, I., & Sen, A. (2015). Comparative investigation of the various determinants that influence the codon and amino acid usage patterns in the genus *Bifidobacterium*. World Journal of Microbiology and Biotechnology, 31(6), 959-981.

Roy, S. C., & Chakraborty, B. N. (2007). Evaluation of genetic diversity in tea of the Darjeeling foot hills, India using RAPD and ISSR markers. Journal of Hill Research, 20, 13-19.

Roy, S. C., & Chakraborty, B. N. (2009). Genetic diversity and relationships among tea (*Camellia sinensis*) cultivars as revealed by RAPD and ISSR based fingerprinting.

Ruiz, L. M., Armengol, G., Habeych, E., & Orduz, S. (2006). A theoretical analysis of codon adaptation index of the *Boophilus microplus* bm86 gene directed to the optimization of a DNA vaccine. Journal of Theoretical Biology, 239(4), 445-449.

Runti, G., Pacor, S., Colomban, S., Gennaro, R., Navarini, L., & Scocchi, M. (2015). *Arabica coffee* extract shows antibacterial activity against *Staphylococcus epidermidis* and *Enterococcus faecalis* and low toxicity towards a human cell line. LWT-Food Science and Technology, 62(1), 108-114.

Rusak, G., Komes, D., Likic, S., Horzic, D., & Kovac, M. (2008). Phenolic content and antioxidative capacity of green and white tea extracts depending on extraction conditions and the solvent used. Food Chemistry, 110(4), 852-858.

Sabbineni, J. (2016). Phenol-an effective antibacterial agent. Research & Reviews: Journal of Medicinal & Organic Chemistry, 3(2), 182-191.

Sablok, G., Nayak, K. C., Vazquez, F., & Tatarinova, T. V. (2011). Synonymous codon usage, GC 3, and evolutionary patterns across plastomes of three pooid model species: Emerging grass genome models for monocots. Molecular Biotechnology, 49(2), 116-128.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution, 4(4), 406 -425.

Savi, M., Rossi, S., Bocchi, L., Gennaccaro, L., Cacciani, F., Perotti, A., & Zaniboni, M. (2014). Titanium dioxide nanoparticles promote arrhythmias via a direct interaction with rat cardiac tissue. Particle and Fibre Toxicology, 11(1), 1-16.

Schlotteroer, C., Amos, B., & Tautz, D. (1991). Conservation of polymorphic simple sequence loci in cetacean species. Nature, 354(6348), 63-65.

Schreier, B., Stern, C., Dubourg, V., Nolze, A., Rabe, S., Mildenberger, S., ... & Gekle, M. (2021). Endothelial epidermal growth factor receptor is of minor importance for vascular and renal function and obesity-induced dysfunction in mice. Scientific Reports, 11(1), 1-11.

Scudiero, O., Brancaccio, M., Mennitti, C., Laneri, S., Lombardo, B., De Biasi, M. G., ... & Pero, R. (2020). Human defensins: a novel approach in the fight against skin colonizing *Staphylococcus aureus*. Antibiotics, 9(4), 198.

Sealy, J. R. (1958). A revision of the genus *Camellia*. A Revision of the Genus *Camellia*.

Seligmann, H. (2003). Cost-minimization of amino acid usage. Journal of Molecular Evolution, 56(2), 151-161.

Selvan, D. A., Mahendiran, D., Kumar, R. S., & Rahiman, A. K. (2018). Garlic, green tea and turmeric extracts-mediated green synthesis of silver nanoparticles: Phytochemical, antioxidant and in vitro cytotoxicity studies. Journal of Photochemistry and Photobiology B: Biology, 180, 243-252.

Selvaraj, D., Sarma, R. K., & Sathishkumar, R. (2008). Phylogenetic analysis of chloroplast matK gene from Zingiberaceae for plant DNA barcoding. Bioinformation, 3(1), 24.

Sen, A., Sur, S., Bothra, A. K., Benson, D. R., Normand, P., & Tisa, L. S. (2008). The implication of life style on codon usage patterns and predicted highly expressed genes for three *Frankia* genomes. Antonie Van Leeuwenhoek, 93(4), 335-346.

Senthilkumar, S. R., & Sivakumar, T. (2014). Green tea (*Camellia sinensis*) mediated synthesis of zinc oxide (ZnO) nanoparticles and studies on their antimicrobial activities. International Journal of Pharmacy and Pharmaceutical Science, 6(6), 461-465.

Shankar, T., Karthiga, P., Swarnalatha, K., & Rajkumar, K. (2017). Green synthesis of silver nanoparticles using *Capsicum frutescence* and its intensified activity against *E. coli*. Resource-Efficient Technologies, 3(3), 303-308.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research, 13(11), 2498-2504.

Sharma, D., Rajput, J., Kaith, B. S., Kaur, M., & Sharma, S. (2010). Synthesis of ZnO nanoparticles and study of their antibacterial and antifungal properties. Thin Solid films, 519(3), 1224-1229.

Sharma, V. S., & Venkataramani, K. S. (1974, October). The tea complex. In Proceedings of the Indian Academy of Sciences-Section B (Vol. 80, No. 4, pp. 178-187). Springer India.

Sharp, P. M., & Li, W. H. (1986). An evolutionary perspective on synonymous codon usage in unicellular organisms. Journal of Molecular Evolution, 24(1-2), 28-38.

Sharp, P. M., & Li, W. H. (1987). The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Research, 15(3), 1281-1295.

Shirazi, S. K., & Wood, J. G. (1993). The protein tyrosine kinase, fyn, in Alzheimer's disease pathology. Neuroreport, 4(4), 435-437.

Singh, I.D., (2005). The Planter's Guide to Tea Culture and Manufacture. NB Modern Agencies.

Sinija, V. R., & Mishra, H. N. (2008). Green tea: Health benefits. Journal of Nutritional & Environmental Medicine, 17(4), 232-242.

Sneath, P. H., & Sokal, R. R. (1973). Numerical taxonomy. The principles and

practice of numerical classification.

Steven G, New master. & Subramanyam, R. (2009). Testing plant barcoding in a sister species complex of pantropical *Acacia* (Mimosoideae, Fabaceae). Molecular Ecology Resources, 9, 172-180.

Sun, B., Zhu, Z., Cao, P., Chen, H., Chen, C., Zhou, X., ... & Liu, S. (2016). Purple foliage coloration in tea (*Camellia sinensis* L.) arises from activation of the R2R3-MYB transcription factor CsAN1. Scientific Reports, 6(1), 1-15.

Sun, H., Zhu, X., Cai, W., & Qiu, L. (2017). Hypaphorine attenuates lipopolysaccharide -induced endothelial inflammation via regulation of TLR4 and PPAR-γ dependent on PI3K/Akt/mTOR signal pathway. International Journal of Molecular Sciences, 18(4), 844.

Sutormin, D., Rubanova, N., Logacheva, M., Ghilarov, D., & Severinov, K. (2019). Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the *Escherichia coli* genome. Nucleic Acids Research, 47(3), 1373-1388.

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., ... & von Mering, C. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Research, 49(D1), D605-D612.

Tabish, S. A. (2017). Lifestyle diseases: consequences, characteristics, causes and control. Journal of Cardiology & Current Research, 9(3), 00326.

Taghavi Fardood, S., Ramazani, A., & Woo Joo, S. (2017). Sol-gel synthesis and characterization of zinc oxide nanoparticles using black tea extract. Journal of Applied Chemical Research, 11(4), 8-17.

Tallei, T. E., & Kolondam, B. J. (2015). DNA barcoding of Sangihe Nutmeg *(Myristica fragrans*) using matK gene. HAYATI Journal of Biosciences, 22(1), 41-47.

Tamura, K., Nei, M., & Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences, 101(30), 11030-11035.

Towfighi, A., & Ovbiagele, B. (2008). Partial peroxisome proliferator-activated receptor agonist angiotensin receptor blockers. Cerebrovascular Diseases, 26(2), 106-112.

Tripathy, A., Raichur, A. M., Chandrasekaran, N., Prathna, T. C., & Mukherjee, A. (2010). Process variables in biomimetic synthesis of silver nanoparticles by aqueous extract of *Azadirachta indica* (Neem) leaves. Journal of Nanoparticle Research, 12(1), 237-246.

Turkmen, N., Sari, F., & Velioglu, Y. S. (2006). Effects of extraction solvents on concentration and antioxidant activity of black and black mate tea polyphenols determined by ferrous tartrate and Folin–Ciocalteu methods. Food chemistry, 99(4), 835-841.

Ugochukwu, S. C., Uche, A., & Ifeanyi, O. (2013). Preliminary phytochemical screening of different solvent extracts of stem bark and roots of *Dennetia tripetala* G. Baker. Asian Journal of Plant Science and Research, 3(3), 10-13.

Ukers, W. H. (1935). All about tea (Vol. 1). Tea and coffee trade journal Company.

Vaghasiya, Y., & Chanda, S. (2009). Screening of some traditionally used Indian plants for antibacterial activity against *Klebsiella pneumonia*. Journal of Herbal Medicine and Toxicology, 3(2), 161-164.

Vila, J., Sáez-Lopez, E., Johnson, J. R., Romling, U., Dobrindt, U., Canton, R., ... & Soto, S. M. (2016). *Escherichia coli*: an old friend with new tidings. FEMS Microbiology Reviews, 40(4), 437-463.

Visser, T. (1969). *Camellia sinensis* (L.) O. Kuntze. Outlines of perennial crop breeding in the tropics, (4), 459.

Volkers, N. (2000). Diabetes and cancer: scientists search for a possible link. Journal of the National Cancer Institute, 92 (3), 192-194.

Wachira, F. N., Powell, W., & Waugh, R. (1997). An assessment of genetic diversity among *Camellia sinensis* L. (cultivated tea) and its wild relatives based on randomly amplified polymorphic DNA and organelle-specific STS. Heredity, 78 (6), 603-611.

Wachira, F. N., Waugh, R., Powell, W., & Hackett, C. A. (1995). Detection of genetic diversity in tea (*Camellia sinensis*) using RAPD markers. Genome, 38(2), 201-210.

Waldhauer, I., & Steinle, A. (2008). NK cells and cancer immunosurveillance. Oncogene, 27(45), 5932-5943.

Wambulwa, M. C., Meegahakumbura, M. K., Chalo, R., Kamunya, S., Muchugi, A., Xu, J. C., & Gao, L. M. (2016). Nuclear microsatellites reveal the genetic architecture and breeding history of tea germplasm of East Africa. Tree Genetics & Genomes, 12(1), 11.

Wang, H., Yao, L., Cai, R., Pan, J., & Chen, X. (2012). Genetic relationship analyses of oil -bearing roses in China using matK sequences. Scientia Horticulturae, 137, 121-124.

Wei, W., & Guo, F. B. (2010). Strong strand composition bias in the genome of *Ehrlichia canis* revealed by multiple methods. The Open Microbiology Journal, 4, 98.

Wenk, G. L. (2003). Neuropathologic changes in Alzheimer's disease. Journal of Clinical Psychiatry, 64, 7-10.

Wight W (1962) Tea classification revised. Current Science 31:298-299

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research, 18(22), 6531-6535.

Wojtowicz, S., Strosznajder, A. K., Jeżyna, M., & Strosznajder, J. B. (2020). The novel role of PPAR alpha in the brain: promising target in therapy of Alzheimer's disease and other neurodegenerative disorders. Neurochemical Research, 45(5), 972-988.

Wood, D. J., & Barua, P. K. (1958). Species hybrids of tea. Nature, 181(4624), 1674-1675.

Wright, F. (1990). The 'effective number of codons' used in a gene. Gene, 87(1), 23-29.

Wu, K. S., Jones, R., Danneberger, L., & Scolnik, P. A. (1994). Detection of microsatellite polymorphisms without cloning. Nucleic Acids Research, 22(15), 3257.

Xia, E. H., Li, F. D., Tong, W., Li, P. H., Wu, Q., Zhao, H. J., ... & Wan, X. C. (2019). Tea Plant Information Archive: a comprehensive genomics and bioinformatics platform for tea plant. Plant Biotechnology Journal, 17(10), 1938-1953.

Xia, E., Tong, W., Hou, Y., An, Y., Chen, L., Wu, Q., & Wan, X. (2020). The reference genome of tea plant and resequencing of 81 diverse accessions provide insights into its genome evolution and adaptation. Molecular Plant, 13(7), 1013-

1026.

Xia, X. (2007). An improved implementation of codon adaptation index. Evolutionary Bioinformatics, 3, 117693430700300028.

Xia, X. (2013). DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Molecular Biology and Evolution, 30(7), 1720-1728.

Yam, T. S., Shah, S., & Hamilton-Miller, J. M. T. (1997). Microbiological activity of whole and fractionated crude extracts of tea (*Camellia sinensis*), and of tea components. FEMS Microbiology Letters, 152(1), 169-174.

Yengkhom, S., Uddin, A., & Chakraborty, S. (2019). Deciphering codon usage patterns and evolutionary forces in chloroplast genes of *Camellia sinensis* var. *assamica* and *Camellia sinensis* var. *sinensis* in comparison to *Camellia pubicosta*. Journal of Integrative Agriculture, 18(12), 2771-2785.

Zhang, J., Xu, J. H., Qu, Q. Q., & Zhong, G. Q. (2020). Risk of cardiovascular and cerebrovascular events in polycystic ovarian syndrome women: a meta-analysis of cohort studies. Frontiers in Cardiovascular Medicine, 184.

Zhao, Y., Yang, Y., Liu, Z., & Yang, P. (2011). Analysis of codon usage in tea plant (*Camellia sinensis*). Journal of Tea Science, 31(4), 319-325.

Zietkiewicz, E., Rafalski, A., & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics, 20(2), 176-183.

Zou, Y., Lu, Y., & Wei, D. (2004). Antioxidant activity of a flavonoid-rich extract of *Hypericum perforatum* L. in vitro. Journal of Agricultural and Food Chemistry, 52 (16), 5032-5039.

Index

Appendix – A

List of Publication

RESEARCH ARTICLE

- Bose, D., Sarkar, I., **Labar, R**., Oshone, R., Ghazal, S., Morris, K., ... & Sen, A. **(2016)**. Comparative genomics of Prauserella sp. Am3, an actinobacterium isolated from root nodules of Alnus nepalensis in India. ***Symbiosis***, *70*(1), 49-58.

- **Labar, R.,** Sen, A., & Bhattacharya, M. **(20)**. Effect of solvents on qualitative and quantitative phytochemical constituent profiles of fresh leaves of TV26. ***NBU J Plant Sc***. *Vol*, *11*, 115-123.

- **Labar, R.,** Sarkar, I., Sen, A., & Bhattacharya, M. **(2019)**. Effect of solvent with varying polarities on phytochemical extraction from mature tea leaves and its evaluation using biochemical, antimicrobial and in-silico approaches. ***Int Res J Pharm***, *10*(8), 59-67.
- **Labar, R.,** Kar, P., Biswas, P., Sen, A., & Bhattacharya, M. **(2021)**. Evolution of matK Gene among the Elite Tea Clones (*Camellia sinensis*) Revealed by Nucleotide Substitution within the Consensus Region**.** ***J Appl Biol Biotechnol***, *Vol*, *9*(1), 32-40.

REVIEW ARTICLE

- **Labar, R.,** and Sen, A. **(2015)**. Efficacy of Anthocyanin in Production of Remedial Tea. *NBU J Plant Sc*. *Vol*, *9*(1), 18-32.

AWARDS AND HONOUR

- Outstanding paper award: 3rd Regional Science Congress
- Outstanding paper award: 27th West Bengal state Science Congress

Appendix – B

Buffers and chemicals used for DNA fingerprinting studies

CTAB- buffer

100mM Trizma Base (Sigma, Cat# T1503) (pH-8.0)

20mM EDTA (Merck India, Cat# 60841801001730) (pH-8.0)

1.4 M NaCl (Merck India, Cat#60640405001730)

2% (w/v) CTAB (Hexadecyl cetyl trimethyl ammonium bromide) (Sigma, Cat# H6269)

12.11g of molecular grade Trizma base was dissolved in 400 ml double distilled water, pH was adjusted to 8.0 and was divided into two parts of equal volume. To one part 7.44g EDTA was added and to the other part 81.8g NaCl and 20g CTAB. Both the parts were than mixed and the final volume was made up to 1000ml with double distilled water prior to autoclaving. The buffer was autoclaved at 121ºC and 15 psi for 20 mins and stored at room temperature for further use.

Note: Add 1% PVP (Polyvinylpyrrolidone) (Sigma, Cat #P5288) and 0.3% β-mercaptoethanol (Sigma, Cat# M3148) just before use.

CTAB/NaCl

CTAB (Sigma, Cat# H6269) =10gm

NaCl (Merck India, Cat#60640405001730) = 4.1gm

Final volume=100ml

In 80ml double distilled water, 4.1gm of NaCl was dissolved. Slowly 10gm of CTAB was added with continuous heating and stirring until the solution becomes clear. The final volume

of the solution was adjusted to 100ml by adding ddH2O. The solution was autoclaved at 121ºC and 15 psi for 20 mins and stored at room temperature for further use.

5X TBE (Tris-borate-EDTA) buffer

Trizma base (Sigma, Cat# T1503) = 27 gm

Boric acid (Sigma, Cat# 15663) = 13.75 gm

0.5M EDTA (pH 8.0) =1.86 gm

All the reagents were dissolved separately and finally mixed together and the final volume was made up to 1000ml with double distilled water prior to autoclaving. The buffer was autoclaved at 121ºC and 15 psi for 20 mins and stored at room temperature for further use.

1X TE

Tris- Cl (pH 8.0) (i.e., 10Mm) =0.6055gm

EDTA (pH 8.0) (i.e., 1mM) =0.186 gm

Both the reagents were dissolved separately and finally mixed together and the final volume was made up to 1000ml with double distilled water prior to autoclaving. The buffer was autoclaved at 121ºC and 15 psi for 20 mins and stored at room temperature for further use.

3M Sodium Acetate (Sigma, Cat# S9513)

The required amount of sodium acetate i.e.,12.31 g was dissolved in 50ml double distilled water prior to autoclaving. The solution was autoclaved at 121ºC and 15 psi for 20 mins and stored at room temperature for further use.

6X gel loading buffer

TYPE 3:

0.25% Bromophenol blue (Sigma, Cat# B0126)

0.25% Xylene cyanol FF (Sigma, Cat# X4126)

30% Glycerol (Merck India, Cat#61756005001730) in water

Store at 4ºC.

RNase A

The RNase A enzyme (Sigma, Cat# R4875) was dissolved at a concentration of 10mg/ml in 0.01M sodium acetate (Sigma, Cat# S9513) (pH 5.2). The solution was heated at 100ºC for 15 minutes in a water bath and allowed to cool slowly to room temperature. The pH was adjusted by adding 1/10 volume of 1M Tris- Cl (pH 7.4) and stored at -20ºC for further use.

Note: Both 0.01M sodium acetate and 1M Tris-Cl were prepared and autoclaved at 121ºC and 15 psi for 20 mins prior to use.

SDS (10%)

SDS (Sigma, USA. Cat #L4390) =5gm

Final volume=50ml

In 20ml of ddH2O, 5gm of SDS was added and heated to dissolve. The final volume was made upto 50ml and autoclaved for future use.

Appendix – C

SEQUENCE SUBMITTED IN NCBI

1. MH649284 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Ambari Vallai 2 maturase K (matK) gene, partial cds; chloroplast.
2. MH791417 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Happy Valley 39 maturase K (matK) gene, partial cds; chloroplast
3. MH920315 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Takdah 253 maturase K (matK) gene, partial cds; chloroplast.
4. MH920316 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Nandadevi maturase K (matK) gene, partial cds; chloroplast.
5. MH920317 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone MB - 6 maturase K (matK) gene, partial cds; chloroplast.
6. MH920318 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Teesta Valley 1 maturase K (matK) gene, partial cds; chloroplast.
7. MH920319 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Kopati 1/1 maturase K (matK) gene, partial cds; chloroplast.
8. MH393392- **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Badamtam 15/263 maturase K (matK) gene, partial cds; chloroplast.

9. MH393393 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Balasun 7/1A/76 maturase K (matK) gene, partial cds; chloroplast.
10. MH393394 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Phoobsering 312 maturase K (matK) gene, partial cds; chloroplast.
11. MK393395 - Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Bannockburn 777 maturase K (matK) gene, partial cds; chloroplast.
12. MK393396 - Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Sundaram maturase K (matK) gene, partial cds; chloroplast.
13. MK393397- Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Takdah 135 maturase K (matK) gene, partial cds; chloroplast.
14. MK393398 - Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Bannockburn 688 (matK) gene, partial cds; chloroplast.
15. MK393399 - Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Golconda maturase K (matK) gene, partial cds; chloroplast.
16. MK393400 - Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Rungli Rungiliot 17/144 maturase K (matK) gene, partial cds; chloroplast.
17. MK393401 - Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Balasun 9/3/76 maturase K (matK) gene, partial cds; chloroplast.
18. MK393402- Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Chiradew Parbat1 maturase K (matK) gene, partial cds; chloroplast.
19. MK393403 - Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Phoobsering 1404 maturase K (matK) gene, partial cds; chloroplast.

20. MK393404 - Kar, P., **Labar, R.,** Bhattacharya, M., Biswas, P. and Sen, A. *Camellia sinensis* clone Phoobsering 1258 maturase K (matK) gene, partial cds; chloroplast.
21. MK393405 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Rungli Rungiliot 4/5 maturase K (matK) gene, partial cds; chloroplast.
22. MK424865 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Sikkim 1 maturase K (matK) gene, partial cds; chloroplast.
23. MK424866 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Takdah 145 maturase K (matK) gene, partial cds; chloroplast.
24. MK424867 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Takdah 246 K (matK) gene, partial cds; chloroplast.
25. MK424868 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Tocklai Variety 19 maturase K (matK) gene, partial cds; chloroplast.
26. MK424869 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Tocklai Variety 14 maturase K (matK) gene, partial cds; chloroplast.
27. MK424870 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Takdah 78 maturase K (matK) gene, partial cds; chloroplast.
28. MK424871 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Bannockburn 157 maturase K (matK) gene, partial cds; chloroplast.
29. MK424872 - Bhattacharya, M., **Labar, R.,** Kar, P., Biswas, P. and Sen, A. *Camellia sinensis* clone Takdah 383 maturase K (matK) gene, partial cds; chloroplast.
30. MN480321 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Thurbo 3 maturase K (matK) gene, partial cds; chloroplast.
31. MN480322 - Labar**, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Thurbo 9 maturase K (matK) gene, partial cds; chloroplast.

32. MZ130930 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Thurbo 3 ribosomal protein S16 (rps16) gene, partial cds; chloroplast.
33. MZ130931 -- **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone Thurbo 9 ribosomal protein S16 (rps16) gene, partial cds; chloroplast.
34. MZ130932 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone BS 7/1A/76 ribosomal protein S16 (rps16) gene, partial cds; chloroplast.
35. MZ130933 - **Labar, R.,** Kar, P., Biswas, P., Bhattacharya, M. and Sen, A. 2019. *Camellia sinensis* clone clone B-777 ribosomal protein S16 (rps16) gene, partial cds; chloroplast.

Appendix - D

Table 1: Phytoconstituents of tea extract (Assam) characterized by GC-MS.

Peak Report TIC

Peak#	R.Time	Area	Area%	Name
1	4.683	941660	0.08	2,4-Dimethyl-1-heptene
2	4.740	847411	0.08	2-HYDROXY-2-METHYL-4-PENTANONE (DIACETONE
3	5.086	1310017	0.12	4,4-Dimethoxy-2-methyl-2-butanol
4	5.451	13969189	1.26	BUTANOIC ACID, 3-METHYL-
5	5.528	14008026	1.26	1-BUTANOL, 3-METHYL-, ACETATE
6	6.346	7918761	0.71	HYDROPEROXIDE, 1,1-DIMETHYLETHYL
7	7.490	1467464	0.13	BUTANE, 1,1-DIETHOXY-3-METHYL-
8	8.043	2803378	0.25	PENTANE, 1-(1-ETHOXYETHOXY)-
9	8.253	1463283	0.13	ETHANE, PENTACHLORO-
10	8.364	5782140	0.52	TETRAHYDRO-2-(TETRAHYDRO-2-FURFURYL)-FURA
11	9.064	3123505	0.28	ISOBUTYLISOVALERATE
12	9.224	8266587	0.75	Propanoic acid, 2-methyl-, 3-methylbutyl ester
13	9.318	1191227	0.11	PROPANOIC ACID, 2-METHYL-, 3-METHYLBUTYL EST
14	9.835	1393664	0.13	TETRAHYDRO-2-(TETRAHYDRO-2-FURFURYL)-FURA
15	10.232	3538133	0.32	BUTANOIC ACID, 3-METHYL-, BUTYL ESTER
16	10.511	4426347	0.40	ISO AMYL N-BUTYRATE
17	11.044	7472070	0.67	Ethane, hexachloro-
18	11.177	5464583	0.49	1,3-Dioxan-4-one, 2-heptyl-6-methyl
19	11.699	3961589	0.36	2-Furanol, tetrahydro-2,3-dimethyl-, trans-
20	11.828	48697428	4.39	ISOPENTYL 2-METHYLBUTANOATE
21	11.919	7053237	0.64	BUTANOIC ACID, 2-METHYL-, 2-METHYLBUTYL EST
22	12.081	158324513	14.27	PENTANOIC ACID, PENTYL ESTER
23	12.121	19795080	1.78	PENTYL 3-METHYLBUTANOATE
24	14.169	29317951	2.64	2-Butanol, 3-(2,2-dimethylpropoxy)-
25	14.305	14439683	1.30	NAPHTHALENE
26	14.406	1156236	0.10	PROPANOIC ACID, 2-METHYL-, 1-METHYLBUTYL EST
27	14.534	1283502	0.12	1-DODECANOL
28	14.900	19681178	1.77	PROPANOIC ACID, 2-METHYL-, 3-METHYLBUTYL EST
29	15.279	2070734	0.19	2,4-OXAZOLIDINEDIONE, 5,5-DIMETHYL-
30	15.422	9573134	0.86	Butane, 2,2'-[methylenebis(oxy)]bis[2-methyl-
31	16.463	6206086	0.56	Pentan-2-yl 2-methylbutanoate
32	16.924	12677893	1.14	TERT-PENTYL BUTYRATE #
33	17.108	75090507	6.77	PENTANE, 1,1'-OXYBIS-
34	17.234	1854308	0.17	TERT-PENTYL BUTYRATE #
35	17.609	4045864	0.36	Oxalic acid, 2-ethylhexyl pentyl ester
36	17.832	1932328	0.17	Sulfurous acid, octadecyl 2-propyl ester
37	18.057	1586207	0.14	2-Isopropyl-5-methyl-1-heptanol
38	18.851	2168485	0.20	Butanoic acid, 2,2-dimethylpropyl ester
39	19.988	2690410	0.24	2-METHYLPROPIONYL PHENYLSELENO SULFIDE
40	20.100	45543516	4.11	diethyl 2-hydroxy-3-(tetrahydrofuran-2-yl)succinate
41	20.204	77426697	6.98	Tripentyl orthoformate
42	20.279	1435480	0.13	TETRADECANE
43	20.813	4148818	0.37	Propanoic acid, 2-methyl-, pentyl ester

Peak#	R.Time	Area	Area%	Name
44	22.903	6216186	0.56	PHENOL, 2,4-BIS(1,1-DIMETHYLETHYL)-
45	23.595	1908744	0.17	Eicosane
46	24.993	3799147	0.34	1-HEXADECENE
47	25.174	2182662	0.20	Hexadecane
48	28.380	1314281	0.12	Eicosane
49	28.674	1972073	0.18	OCTADECANOIC ACID
50	29.442	3860426	0.35	1-OCTADECENE
51	29.593	1621985	0.15	EICOSANE
52	30.324	5566676	0.50	Caffeine
53	31.647	1983574	0.18	7,9-DITERT-BUTYL-1-OXASPIRO[4.5]DECA-6,9-DIENE-
54	32.107	1222258	0.11	HEXADECANOIC ACID, METHYL ESTER
55	32.634	1579085	0.14	Dibutyl phthalate
56	32.827	10709536	0.97	HEXADECANOIC ACID
57	33.475	3716006	0.33	n-Tetracosanol-1
58	33.605	1813258	0.16	EICOSANE
59	34.422	1686006	0.15	Hexadecanoic acid, trimethylsilyl ester
60	35.476	4245849	0.38	Heneicosane
61	36.573	2258057	0.20	Octadecanoic acid
62	37.254	2544252	0.23	n-Tetracosanol-1
63	37.373	15575865	1.40	DOCOSANE
64	38.066	1292782	0.12	3-METHYLBUTYL DECANOATE
65	38.840	27213621	2.45	DOCOSANE
66	39.913	35284249	3.18	PENTACOSANE
67	40.776	30259090	2.73	PENTACOSANE
68	41.013	1230102	0.11	OCTADECANAL
69	41.254	2688734	0.24	Tetratetracontane
70	41.551	32340484	2.92	TETRACONTANE
71	42.054	4042684	0.36	Tetratetracontane
72	42.139	2857543	0.26	Octadecane, 9-ethyl-9-heptyl-
73	42.380	34998234	3.16	TETRACONTANE
74	42.890	1511823	0.14	TRICOSANE
75	42.947	5818265	0.52	Tetratetracontane
76	43.049	6031542	0.54	Hexacosane, 9-octyl-
77	43.241	2605687	0.23	1-Cyclopentyleicosane
78	43.323	36774495	3.32	TETRACONTANE
79	43.462	4625992	0.42	Squalene
80	43.691	1302174	0.12	2-Cyclohexylnonadecane
81	43.825	1489030	0.13	CELIDONIOL, DEOXY-
82	43.916	2976271	0.27	Tetratetracontane
83	43.981	4512312	0.41	Tetratetracontane
84	44.104	7465125	0.67	11,15-Dimethylpentatriacontane
85	44.422	45056493	4.06	TETRACONTANE
86	44.529	3354502	0.30	n-Heptadecylcyclohexane
87	44.817	1569506	0.14	Z-14-Nonacosane

Peak#	R.Time	Area	Area%	Name
88	44.890	1705818	0.15	CYCLOHEXANE, EICOSYL-
89	44.959	1585120	0.14	TRITETRACONTANE
90	45.137	2752709	0.25	TRICOSANE
91	45.209	2842808	0.26	Tetratriacontane
92	45.359	3311273	0.30	13,17,21-Trimethylheptatriacontane
93	45.735	29429447	2.65	TETRACONTANE
94	45.913	7238643	0.65	n-Heptadecylcyclohexane
95	46.250	4077668	0.37	1-Triacontanol
96	46.337	9403903	0.85	4,4,6a,6b,8a,11,11,14b-Octamethyl-1,4,4a,5,6,6a,6b,7,8,8a,9
97	46.571	4837342	0.44	NOROLEAN-12-ENE
98	47.337	14507781	1.31	PENTATRIACONTANE
99	47.516	1645312	0.15	a-Neooleana-3,12-diene
100	47.596	2164856	0.20	n-Heptadecylcyclohexane
101	48.010	5579470	0.50	Urs-12-ene
102	49.316	7571589	0.68	PENTATRIACONTANE
		1109278714	100.00	

Table 2: Phytoconstituents of tea extract (China) characterized by GC-MS.

Peak Report TIC

Peak#	R.Time	Area	Area%	Name
1	4.737	951562	0.08	2-HYDROXY-2-METHYL-4-PENTANONE (DIACETONE
2	5.081	1407061	0.12	4,4-Dimethoxy-2-methyl-2-butanol
3	5.457	11929558	1.01	BUTANOIC ACID, 3-METHYL-
4	5.525	12949100	1.09	1-BUTANOL, 3-METHYL-, ACETATE
5	6.345	7891633	0.67	HYDROPEROXIDE, 1,1-DIMETHYLETHYL
6	7.490	1644077	0.14	BUTANE, 1,1-DIETHOXY-3-METHYL-
7	8.041	2829773	0.24	PENTANE, 1-(1-ETHOXYETHOXY)-
8	8.256	1794026	0.15	ETHANE, PENTACHLORO-
9	8.361	4581997	0.39	TETRAHYDRO-2-(TETRAHYDRO-2-FURFURYL)-FURA
10	9.059	2711807	0.23	ISOBUTYLISOVALERATE
11	9.220	7808394	0.66	Propanoic acid, 2-methyl-, 3-methylbutyl ester
12	9.313	1183931	0.10	PROPANOIC ACID, 2-METHYL-, 3-METHYLBUTYL EST
13	9.831	1130599	0.10	TETRAHYDRO-2-(TETRAHYDRO-2-FURFURYL)-FURA
14	10.229	3503651	0.30	BUTANOIC ACID, 3-METHYL-, BUTYL ESTER
15	10.506	4385203	0.37	ISO AMYL N-BUTYRATE
16	11.012	8973287	0.76	4,5-Octanediol, 2,7-dimethyl-
17	11.172	6044376	0.51	1,3-Dioxan-4-one, 2-heptyl-6-methyl
18	11.691	4173155	0.35	2-Furanol, tetrahydro-2,3-dimethyl-, trans-
19	11.826	48749178	4.12	ISOPENTYL 2-METHYLBUTANOATE
20	11.917	7123689	0.60	BUTANOIC ACID, 2-METHYL-, 2-METHYLBUTYL EST
21	12.076	160496607	13.55	PENTANOIC ACID, PENTYL ESTER
22	12.115	18650969	1.57	PENTYL 3-METHYLBUTANOATE
23	12.447	975139	0.08	2H-PYRAN, TETRAHYDRO-2-[(TETRAHYDRO-2-FURA
24	14.167	30003187	2.53	2-Butanol, 3-(2,2-dimethylpropoxy)-
25	14.299	8631565	0.73	NAPHTHALENE
26	14.403	1324213	0.11	PROPANOIC ACID, 2-METHYL-, 1-METHYLBUTYL EST
27	14.898	20052190	1.69	PROPANOIC ACID, 2-METHYL-, 3-METHYLBUTYL EST
28	15.276	2049766	0.17	2,4-OXAZOLIDINEDIONE, 5,5-DIMETHYL-
29	15.421	9262473	0.78	Butane, 2,2'-[methylenebis(oxy)]bis[2-methyl-
30	16.464	6127497	0.52	Pentan-2-yl 2-methylbutanoate
31	16.924	14366096	1.21	TERT-PENTYL BUTYRATE #
32	17.113	84538200	7.14	PENTANE, 1,1'-OXYBIS-
33	17.234	2146469	0.18	TERT-PENTYL BUTYRATE #
34	17.612	3790373	0.32	Oxalic acid, neopentyl octyl ester
35	17.829	1719951	0.15	2-Hexyldecyl acetate
36	17.884	1396155	0.12	Butanoic acid, anhydride
37	18.055	1285506	0.11	2-Isopropyl-5-methyl-1-heptanol
38	18.849	2187833	0.18	Butyric acid, 1-propylpentyl ester
39	19.087	1050635	0.09	Propanoic acid, 2-methyl-, 1,2,3-propanetriyl ester
40	19.315	1745230	0.15	1,3-Dioxane, 5-ethyl-2,2-dimethyl-
41	19.987	2808445	0.24	2-METHYLPROPIONYL PHENYLSELENO SULFIDE
42	20.098	45528177	3.84	diethyl 2-hydroxy-3-(tetrahydrofuran-2-yl)succinate
43	20.203	79417711	6.70	Tripentyl orthoformate

Peak#	R.Time	Area	Area%	Name
44	20.276	1725891	0.15	TETRADECANE
45	20.649	1618868	0.14	2-Butenoic acid, 1-methylethyl ester
46	20.811	4663100	0.39	Propanoic acid, 2-methyl-, pentyl ester
47	22.902	6012571	0.51	PHENOL, 2,4-BIS(1,1-DIMETHYLETHYL)-
48	23.594	1837488	0.16	Eicosane
49	24.992	3680355	0.31	1-HEXADECENE
50	25.174	2449935	0.21	Hexadecane
51	28.672	2018213	0.17	OCTADECANOIC ACID
52	29.440	3768722	0.32	1-OCTADECENE
53	29.593	1439797	0.12	EICOSANE
54	30.381	25396950	2.14	Caffeine
55	31.646	2027066	0.17	7,9-Di-tert-butyl-1-oxaspiro(4,5)deca-6,9-diene-2,8-dione
56	32.632	2488888	0.21	Dibutyl phthalate
57	32.830	12529090	1.06	HEXADECANOIC ACID
58	33.473	3524058	0.30	1-OCTADECENE
59	34.421	1706418	0.14	Hexadecanoic acid, trimethylsilyl ester
60	35.473	3728859	0.31	Heneicosane
61	36.574	7045354	0.59	Octadecanoic acid
62	37.252	2483297	0.21	n-Tetracosanol-1
63	37.369	14612691	1.23	DOCOSANE
64	38.838	26715832	2.26	DOCOSANE
65	39.912	36944873	3.12	PENTACOSANE
66	40.777	33438898	2.82	PENTACOSANE
67	41.253	2765827	0.23	Tetratetracontane
68	41.552	36692995	3.10	TETRACONTANE
69	42.053	4499247	0.38	Tetratetracontane
70	42.140	3081651	0.26	Octadecane, 9-ethyl-9-heptyl-
71	42.381	40946772	3.46	TETRACONTANE
72	42.890	1646893	0.14	TRICOSANE
73	42.947	6857133	0.58	Tetratetracontane
74	43.048	6760985	0.57	Hexacosane, 9-octyl-
75	43.240	2835215	0.24	1-Cyclopentyleicosane
76	43.323	43898109	3.71	TETRACONTANE
77	43.689	1318998	0.11	CYCLOHEXANE, EICOSYL-
78	43.825	1591084	0.13	CELIDONIOL, DEOXY-
79	43.916	4923851	0.42	Tetratetracontane
80	43.977	7288682	0.62	Tetratetracontane
81	44.103	8358966	0.71	TETRACONTANE
82	44.422	49583281	4.19	TETRACONTANE
83	44.528	3591126	0.30	n-Heptadecylcyclohexane
84	44.817	1776119	0.15	1-EICOSANOL
85	44.885	2053745	0.17	CYCLOHEXANE, EICOSYL-
86	44.956	1972063	0.17	TRITETRACONTANE
87	45.132	3391172	0.29	TRICOSANE

Peak#	R.Time	Area	Area%	Name
88	45.206	3208297	0.27	Hexatriacontane
89	45.355	3915551	0.33	13,17,21-Trimethylheptatriacontane
90	45.732	44696244	3.77	Docosanoic acid, docosyl ester
91	45.911	7634386	0.64	n-Heptadecylcyclohexane
92	46.252	3125807	0.26	1-EICOSANOL
93	46.688	4739472	0.40	Hexatriacontane
94	47.331	20061001	1.69	PENTATRIACONTANE
95	47.595	5853300	0.49	n-Heptadecylcyclohexane
96	48.001	3543885	0.30	Heptacosane, 1-chloro-
97	49.313	8767795	0.74	PENTATRIACONTANE
		1184557310	100.00	

Table 3: Phytoconstituents of tea extract (AV2) characterized by GC-MS

Peak Report TIC

Peak#	R.Time	Area	Area%	Name
1	4.305	279010	0.52	Heptane, 2,4-dimethyl-
2	4.711	15208680	28.58	2-HYDROXY-2-METHYL-4-PENTANONE (DIACETONE
3	5.131	75414	0.14	OCTANE, 4-METHYL-
4	6.403	2431765	4.57	Bicyclo[3.1.0]hex-2-ene, 2-methyl-5-(1-methylethyl)-
5	6.554	248137	0.47	BICYCLO[3.1.1]HEPT-2-ENE, 2,6,6-TRIMETHYL-
6	7.125	63520	0.12	NONANE, 2-METHYL-
7	7.313	273593	0.51	BICYCLO[3.1.0]HEXANE, 4-METHYLENE-1-(1-METHY
8	7.414	63413	0.12	BICYCLO[3.1.1]HEPTANE, 6,6-DIMETHYL-2-METHYLE
9	7.941	90963	0.17	1,3-CYCLOHEXADIENE, 2-METHYL-5-(1-METHYLETH
10	7.993	64009	0.12	BICYCLO[4.1.0]HEPT-3-ENE, 3,7,7-TRIMETHYL-
11	8.033	85066	0.16	Undecane, 5,6-dimethyl-
12	8.133	69327	0.13	HEXANE, 2,2,3,3-TETRAMETHYL-
13	8.199	213758	0.40	Octane, 3,3-dimethyl-
14	8.281	47689	0.09	BENZENE, 1-METHYL-4-(1-METHYLETHYL)-
15	8.369	191218	0.36	CYCLOHEXENE, 1-METHYL-4-(1-METHYLETHENYL)-
16	8.423	143138	0.27	2-OXABICYCLO[2.2.2]OCTANE, 1,3,3-TRIMETHYL-
17	8.786	570544	1.07	Decane, 3,7-dimethyl-
18	8.885	281466	0.53	Decane, 3,7-dimethyl-
19	9.563	186001	0.35	Decane, 3,7-dimethyl-
20	9.657	146365	0.28	Hexane, 3,3,4-trimethyl-
21	10.618	62843	0.12	Decane, 1-iodo-
22	10.774	54187	0.10	NONANE, 3,7-DIMETHYL-
23	11.196	112822	0.21	HEPTADECANE
24	11.326	34735	0.07	Heptane, 3,4,5-trimethyl-
25	11.377	105264	0.20	Undecane, 2,6-dimethyl-
26	11.510	104600	0.20	Dodecane, 4-methyl-
27	11.795	91320	0.17	Sulfurous acid, 2-ethylhexyl isohexyl ester
28	11.880	55358	0.10	NONANE, 3,7-DIMETHYL-
29	11.946	561494	1.06	Benzene, 1,3-bis(1,1-dimethylethyl)-
30	12.112	191592	0.36	DECANE, 2,3,7-TRIMETHYL-
31	12.300	574484	1.08	Nonane, 5-(2-methylpropyl)-
32	12.424	86957	0.16	Undecane, 3,8-dimethyl-
33	12.510	58713	0.11	Undecane, 3,8-dimethyl-
34	12.618	68803	0.13	Decane, 3,7-dimethyl-
35	12.760	56353	0.11	Hexane, 3,3-dimethyl-
36	12.967	180836	0.34	Decane, 3,7-dimethyl-
37	13.559	47673	0.09	Decane, 1-iodo-
38	14.153	29629	0.06	3-OXO-1-CYCLOHEXENYL 2-METHYLPROPIONAT
39	14.196	46705	0.09	Decane, 2,5,6-trimethyl-
40	14.256	82061	0.15	Sulfurous acid, 2-ethylhexyl isohexyl ester
41	14.325	51059	0.10	Hexane, 3,3,4-trimethyl-
42	14.703	58862	0.11	PENTADECANE
43	14.779	78123	0.15	Sulfurous acid, 2-ethylhexyl isohexyl ester

Peak#	R.Time	Area	Area%	Name
44	14.837	43386	0.08	HEPTANE, 3,3,5-TRIMETHYL-
45	14.867	54626	0.10	Undecane, 3,4-dimethyl-
46	14.953	40840	0.08	Undecane, 3,8-dimethyl-
47	15.136	117326	0.22	Nonane, 5-(2-methylpropyl)-
48	15.193	546331	1.03	Heptadecane
49	15.280	44950	0.08	Nonane, 2,5-dimethyl-
50	15.315	32980	0.06	Nonane, 5-methyl-5-propyl-
51	15.505	648502	1.22	Phenol, 3,5-bis(1,1-dimethylethyl)-
52	15.624	41840	0.08	Nonane, 5-(2-methylpropyl)-
53	15.759	200674	0.38	Nonane, 5-butyl-
54	15.885	42477	0.08	Decane, 3,7-dimethyl-
55	16.480	55612	0.10	Diethyl Phthalate
56	16.653	71483	0.13	Octane, 2-methyl-
57	16.877	80747	0.15	Heptadecane, 2,6,10,15-tetramethyl-
58	17.113	50039	0.09	Octane, 3,5-dimethyl-
59	17.230	39182	0.07	1-PIPERAZINEETHANAMINE, 4-METHYL-
60	17.333	21828	0.04	Undecane, 3,8-dimethyl-
61	17.659	128321	0.24	Dodecane, 2-methyl-
62	17.716	388524	0.73	OCTADECANE
63	17.765	45848	0.09	Nonane, 5-(2-methylpropyl)-
64	17.820	80651	0.15	Nonane, 5-methyl-5-propyl-
65	18.031	117184	0.22	Dodecyl nonyl ether
66	18.153	33571	0.06	1-METHYL-1,2,4-TRIAZOLE-4-15N
67	18.203	90801	0.17	Nonane, 5-methyl-5-propyl-
68	18.257	29316	0.06	NONANE, 4,5-DIMETHYL-
69	19.205	1399143	2.63	Neophytadiene
70	19.477	18352952	34.49	Caffeine
71	19.655	460801	0.87	3,7,11,15-Tetramethyl-2-hexadecen-1-ol
72	19.805	50137	0.09	Nonadecane
73	19.888	67777	0.13	2-Methyltetracosane
74	19.968	211675	0.40	Eicosane
75	20.335	39057	0.07	Nonane, 5-(1-methylpropyl)-
76	20.395	71431	0.13	Nonane, 5-methyl-5-propyl-
77	21.294	43598	0.08	OCTADECANE
78	21.660	24894	0.05	HEXADECANE
79	21.701	28941	0.05	Octane, 2-methyl-
80	21.934	200468	0.38	Phytol
81	21.998	166408	0.31	2-Methylhexacosane
82	22.385	75117	0.14	Dodecane, 2-methyl-
83	23.929	99635	0.19	2-methyloctacosane
84	28.753	628222	1.18	Squalene
85	29.727	326299	0.61	2-methyloctacosane
86	32.693	410095	0.77	2-methyloctacosane
87	33.255	1962515	3.69	Vitamin E

Peak#	R.Time	Area	Area%	Name
88	37.354	780592	1.47	STIGMASTA-7,22-DIEN-3-OL, (3.BETA.,5.ALPHA.,22E,2
89	38.532	1230945	2.31	24-Noroleana-3,12-diene
		53205290	100.00	

Table 4: Phytoconstituents of tea extract (TS569) characterized by GC-MS.

Peak Report TIC

Peak#	R.Time	Area	Area%	Name
1	4.305	200603	0.45	Heptane, 2,4-dimethyl-
2	4.718	17134202	38.51	2-HYDROXY-2-METHYL-4-PENTANONE (DIACETONE
3	6.405	1591110	3.58	Bicyclo[3.1.0]hex-2-ene, 2-methyl-5-(1-methylethyl)-
4	6.553	148471	0.33	BICYCLO[3.1.1]HEPT-2-ENE, 2,6,6-TRIMETHYL-
5	7.316	188820	0.42	BICYCLO[3.1.0]HEXANE, 4-METHYLENE-1-(1-METHY
6	7.994	65491	0.15	BICYCLO[4.1.0]HEPT-3-ENE, 3,7,7-TRIMETHYL-
7	8.034	78393	0.18	UNDECANE, 5,6-DIMETHYL-
8	8.135	57915	0.13	OCTANE, 2,4,6-TRIMETHYL-
9	8.201	206359	0.46	Octane, 3,3-dimethyl-
10	8.370	156799	0.35	CYCLOHEXENE, 1-METHYL-4-(1-METHYLETHENYL)-
11	8.788	548475	1.23	Decane, 3,7-dimethyl-
12	8.886	269234	0.61	Decane, 3,7-dimethyl-
13	9.565	187744	0.42	Decane, 3,7-dimethyl-
14	9.660	129537	0.29	Hexane, 3,3-dimethyl-
15	11.197	126743	0.28	PENTADECANE
16	11.379	116268	0.26	Octane, 3,5-dimethyl-
17	11.512	113475	0.26	Dodecane, 4-methyl-
18	11.797	102577	0.23	PENTADECANE
19	11.948	556730	1.25	Benzene, 1,3-bis(1,1-dimethylethyl)-
20	12.113	208916	0.47	NONADECANE
21	12.302	616648	1.39	Nonane, 5-butyl-
22	12.428	93012	0.21	Nonane, 5-methyl-5-propyl-
23	12.512	64515	0.15	3-Ethyl-3-methylheptane
24	12.619	77773	0.17	Nonane, 5-(2-methylpropyl)-
25	12.967	199358	0.45	Decane, 3,7-dimethyl-
26	14.257	84577	0.19	Sulfurous acid, 2-ethylhexyl isohexyl ester
27	14.781	82594	0.19	Sulfurous acid, 2-ethylhexyl isohexyl ester
28	14.869	239257	0.54	Undecane, 3,4-dimethyl-
29	15.137	130360	0.29	Nonane, 5-(2-methylpropyl)-
30	15.193	594150	1.34	Heptadecane
31	15.284	152185	0.34	HEXADECANE
32	15.506	632004	1.42	Phenol, 3,5-bis(1,1-dimethylethyl)-
33	15.760	214860	0.48	Nonane, 5-(2-methylpropyl)-
34	16.877	82707	0.19	Undecane, 3,8-dimethyl-
35	17.661	139622	0.31	OCTADECANE
36	17.717	422825	0.95	Eicosane
37	18.204	114323	0.26	Nonane, 5-methyl-5-propyl-
38	19.206	960408	2.16	Neophytadiene
39	19.467	14677005	32.99	Caffeine
40	19.655	327184	0.74	3,7,11,15-Tetramethyl-2-hexadecen-1-ol
41	19.970	196535	0.44	Eicosane
42	20.395	83780	0.19	Nonane, 5-methyl-5-propyl-
43	21.938	316892	0.71	Phytol

Peak#	R.Time	Area	Area%	Name
44	22.000	188275	0.42	2-Methylhexacosane
45	22.386	104357	0.23	Decane, 1-iodo-
46	23.929	88508	0.20	2-methyloctacosane
47	28.054	133644	0.30	Carbonic acid, 2-ethylhexyl nonyl ester
48	28.754	255553	0.57	Squalene
49	33.251	1027441	2.31	Vitamin E
		44488214	100.00	

Table 5: Phytoconstituents of tea extract (Purple Tea) characterized by GC-MS

Peak Report TIC

Peak#	R.Time	Area	Area%	Name
1	13.108	1707360	5.01	1,2,3-BENZENETRIOL
2	16.968	154335	0.45	Neophytadiene
3	17.224	37841	0.11	3,7,11,15-Tetramethyl-2-hexadecen-1-ol
4	17.375	20770300	60.99	COFFEINE
5	17.883	681456	2.00	CYCLOPENTANETRIDECANOIC ACID, METHYL EST
6	19.518	42029	0.12	10,12-HEXADECADIEN-1-OL
7	19.578	291133	0.85	9-Octadecenoic acid (Z)-, methyl ester
8	19.693	75125	0.22	Phytol
9	21.220	62679	0.18	Fumaric acid, 2-dimethylaminoethyl octadecyl ester
10	21.337	164596	0.48	CHLOROMETHYL 2-CHLORODODECANOATE
11	21.498	60047	0.18	HEXADECANOIC ACID
12	21.554	47262	0.14	OXAZOLE, 2,2'-(1,4-BUTANEDIYL)BIS[4,5-DIHYDRO-
13	22.627	29738	0.09	Petroselinic acid, TBDMS derivative
14	22.677	92796	0.27	Bis(2-(Dimethylamino)ethyl) ether
15	22.820	229857	0.67	Glycidyl oleate
16	23.020	240975	0.71	Hexanoic acid, octadecyl ester
17	23.272	273553	0.80	Di-n-octyl phthalate
18	24.367	461379	1.35	(14Z)-14-TRICOSENYL FORMATE #
19	25.543	42003	0.12	Squalene
20	25.620	1627200	4.78	8-Methoxy-11-methyl-11H-indolo[3,2-c]quinoline, 5-oxide
21	26.140	247977	0.73	2-(3-METHOXY-5-METHYL-BENZYLIDENE)-7-METH
22	27.102	75379	0.22	1,3-benzenedicarboxaldehyde, 4-hydroxy-5,6-dimethoxy-
23	27.836	178786	0.52	3-Dimethylamino-6-nitro-4-phenyl-quinolin-2-ol
24	28.130	4560247	13.39	7-METHOXY-2,3,6-TRIAZAPHENOTHIAZIN-1(2H)-ON
25	28.497	136001	0.40	.gamma.-Tocopherol
26	28.755	257605	0.76	PYRAZINE, TETRAKIS(1-METHYLETHYL)-
27	29.069	410371	1.20	1-(4-Methoxybenzyl)isoquinoline
28	29.739	391353	1.15	Vitamin E
29	33.661	707104	2.08	.gamma.-Sitosterol
		34056487	100.00	

Journal of Applied Biology & Biotechnology Vol. 9(1), pp. 32-40, Jan-Feb, 2021
Available online at http://www.jabonline.in
DOI: 10.7324/JABB.2021.9105

Evolution of matK Gene among the Elite Tea Clones (*Camellia sinensis*) Revealed by Nucleotide Substitution within the Consensus Region

Reha Labar[1], Pallab Kar[1], Prosenjit Biswas[2], Arnab Sen[1], Malay Bhattacharya[3]*

[1]Department of Botany, University of North Bengal, Raja Rammohunpur, West Bengal, India.
[2]Department of Chemistry, University of North Bengal, Raja Rammohunpur, West Bengal, India.
[3]Department of Tea Science, University of North Bengal, Raja Rammohunpur, West Bengal, India

ARTICLE INFO

Article history:
Received on: July 17, 2020
Accepted on: October 01, 2020
Available online: January 17, 2021

Key words:
Camellia sinensis,
matK,
Blast,
Phylogenetic,
Consensus

ABSTRACT

The medicinally and economically important tea plant of India lacks a report on barcode study. Thus, we aimed to establish the DNA barcode of some elite tea clones of Darjeeling and Dooars along with the study of variation within the chloroplast region. A thorough investigation of 29 tea clones based on the matK (maturase K) gene has been carried out in our study. The laid objectives were fulfilled following DNA isolation, purification, amplification of the matK region, and sequencing. The sequences were further analyzed using BLAST analysis and phylogenetic tree construction along with the study of the aligned consensus region among all the clones. A BLAST search of NCBI revealed 24 clones to share 100% identity with *Camellia sinensis*. The remaining 5 clones showed 99.29–99.89% identity with *Camellia sinensis*. However, clones such as 11125 and 11126 showed a higher percentage of similarity, that is, 99.87% and 99.57% with other species of *Camellia* when compared, respectively, to 99.61% and 99.29% with *Camellia sinensis*. The relatedness to other *Camellia* species was also evident from the distinct cluster in the phylogenetic tree. This study reports a total of 14 variable sites within the matK region where the high consensus region revealed a total of nine variable sites and the low consensus region revealed a total of five variable sites. Therefore, this study is the first report of barcode analysis of Indian tea clones, wherein we successfully utilized the single locus matK gene to study variation within the chloroplast region and also conclude that the matK region is not 100% conserved with the same species of *Camellia*.

1. INTRODUCTION

The medicinally important health beverage consumed worldwide as tea, belongs to the genus *Camellia* under the Theaceae family. The world-famous tea originated in China with five reported subspecies and two varieties among which *Camellia sinensis* L.O Kuntze is mostly cultivated worldwide to make the famous tea [1]. The predominant tea varieties cultivated worldwide are the China variety with small leaves (*Camellia sinensis* L.), large leaf Assam variety (*Camellia sinensis* var. *assamica*), and intermediate leaf Cambod varieties (*Camellia assamica* var. *lasiocalyx*) [2].

Apart from its medicinal importance, tea is one of the most important cash crops of India. The world-famous Darjeeling tea serves its purpose for the Indian economy due to its unique flavor and aroma. Darjeeling tea gardens have established several elite tea clones. Environmental influences and plant age make it harder to study genetic diversity based on morphological traits unlike the molecular traits [1].

A previous study on Darjeeling tea clones reports the study of genetic diversity using a robust technique such as the RAPD, ISSR, and AFLP markers [3-5]. However, no study has been reported in the genetic diversity of Darjeeling tea clones using matK primers.

DNA barcoding, a concise method for taxonomic identification, uses a standard short sequence with ample variation to differentiate among species. Many regions from the plastid genome such as the rbcL, rpoC1, rpoB, and trnH-psbA intergenic spacer apart from the matK region have been suggested and exploited for DNA barcoding of land plants [6,7]. However, the Consortium for the Barcode of Life (CboL) has recommended rbcL and matK as standard DNA barcode for plants because of its increased variation between species and the important role it plays in the phylogenetic restoration of terrestrial plants [8,9]. The matK gene around 1500 base pairs (bp) also known as orfK is utilized in the study of molecular systematics and evolution since the matK gene contains high substitution rates within species [10]. The matK gene, coding maturase protein, is responsible for splicing of Group II intron. It is located within the intron of the trnK gene

*Corresponding Author:
Malay Bhattacharya,
Molecular Biology and Tissue Culture Laboratory,
Department of Tea Science, University of North Bengal,
Raja Rammohunpur, West Bengal, 734013, India,
Email: malaytsnbu@gmail.com

whose two flanking exons were lost, thus leaving the gene intact for splicing [11]. Due to its high degree of substitution and variation than other genes, the matK gene is considered to evolve quickly [12]. The matK gene has ideal size and also a mutational conserved region along with a greater rate of substitution and low transition/transversion ratio. The sequence varies at the nucleic acid level at first and second codon positions. All of these features of matK have a profound impact on relationship study at family and species level [11].

Based on previous reports, we found that the Indian tea clones have no report on barcode analysis. Thus, the main objective of the work is to study the genetic diversity and variation within the chloroplast region of the tea clones grown in Darjeeling and Dooars. We collected 33 elite tea clones from the Darjeeling region, which is famous worldwide for tea. We have chosen matK, one of the standard DNA barcodes recommended for plants and also because of the fact of the matK gene having high substitution rates within species, which makes it ideal for our study. Therefore, we aim to perform barcode analysis of collected tea clones by fulfilling steps such as genomic DNA isolation, DNA purification, and quantification, PCR amplification of the matK region, sequencing of the amplified fragments, and sequence analysis using different tools of bioinformatics. This work highlights the importance of the utilization of a single-locus matK region to study intraspecific variation and also infers the evolution of the matK gene within the same species of *Camellia*.

2. MATERIALS AND METHODS

2.1. Sample Selection and Collection

A total of 33 tea clones were collected for this study [Table 1]. The samples were collected from Darjeeling hills.

Table 1: List of tea clones chosen for study.

SI. No	Clone	Abbreviation/Alternative name
1	Ambari Vallai 2	AV2
2	Phoobsering 312	P312
3	Happy valley 39	HV39
4	Tukdah 253	T-253
5	Nanda Devi	TS 378
6	Makaibari-6	MB-6
7	Teesta Valley 1	TTV-1
8	Tukdah 383	T-383
9	Kopati 1/1	K1/1
10	B-15/263	Badamtam -15/263
11	Balasun 7/1A/76	BS 7/1A/76
12	Bunnockburn 777	B-777
13	Sundaram	B/5/63
14	Tukdah-135	T-135
15	Tocklai seed 378	TS 378 or Nanda Devi
16	Bunnockburn 688	B-688
17	Golconda	B/6/36
18	Rungli Rungiliot 17/144	RR-17/144
19	Balasun 9/3/76	BS-9/3/76
20	Chiradew Parbat1	CP-1
21	Phoobsering 1404	P-1404
22	Phoobsering 1258	P-1258
23	Rungli Rungiliot 4/5	RR-4/5
24	Sikkim 1	SKM-1
25	Thurbo 3	Thurbo-3
26	Thurbo 9	Thurbo-9
27	Tukdah 145	T-145
28	Tukdah 246	T-246
29	Tocklai variety 19	TV-19
30	Tocklai variety 14	TV-14
31	Tukdah 78	T-78
32	Bannockburn 157	B-157
33	B/5/63	Sundaram

2.2. DNA Barcoding

2.2.1. DNA Isolation

DNA was isolated from a 5 g fresh leaf sample of tea clones (*Camellia sinensis*) using the CTAB extraction method with slight modification [13]. For the CTAB DNA extraction method, 5 g of fresh and tender leaves of *Camellia* was taken and pulverized using a mortar and pestle using liquid nitrogen. The pulverized material was quickly transferred and mixed into an Oakridge tube containing pre-warmed CTAB extraction buffer. It was then incubated for 1 h at 65°C with occasional mixing in-between. An equal volume of chloroform and isoamyl alcohol (24:1) was added and gently mixed. It was then centrifuged at 6000 rpm for 10 min at room temperature. An equal volume of ice-cold isopropanol was added to the supernatant and mixed by inversion. Following incubation for 2 h at −20°C, the mixture was centrifuged at 6500 rpm for 30 min at 4°C. The supernatant was discarded and the pellet was washed thoroughly with 70% ice-cold ethyl alcohol and allowed to dry for about an hour. The pellet was dissolved in 500 µl of 1X TE buffer and to it, and an equal volume of equilibrated phenol was added and mixed properly followed by centrifugation at 13,000 rpm for 20 min. The upper aqueous layer was transferred into a fresh tube followed by the addition of an equal volume of chloroform:isoamyl alcohol (24:1) and centrifuged at 10,000 rpm for 15 min at room temperature. The supernatant was again transferred to a new tube and treated with 1/10th volume of 3 M sodium acetate and double volume of ice-cold absolute ethyl alcohol. It was mixed gently and then centrifuged at 13,000 rpm for 30 min at 4°C. The supernatant was discarded and the pellet was washed using 70% ethanol, air-dried, and finally dissolved in 500 µl of 1X TE buffer.

2.2.2. DNA purification

RNA, protein, and polysaccharides being the main contaminants in crude DNA, hamper the isolation process and it is, therefore, very important to remove such impurities. CTAB was used to eliminate polysaccharides from DNA along with subsequent use of phenol:chloroform and RNase to further eliminate proteins and RNA to a large extent from crude DNA.

For the purification process, freshly prepared RNaseA was added into the buffered solution of DNA and incubated at 37°C for 1 h in a dry bath. An equal volume of chloroform:isoamyl alcohol (24:1) was added and mixed properly and centrifuged at 10,000 rpm for 15 min at room temperature. The supernatant was transferred to another tube and 1/10th volume of 3 M sodium acetate and double volume of ice-cold absolute ethyl alcohol was added followed by centrifugation at 13,000 rpm for 30 min at 4°C. Finally, the DNA pellet was washed using ice-cold 70% ethyl alcohol and air-dried and finally dissolved in 100 µl of 1X TE buffer.

Table 2: Details of matK primer.

Primer used	Tm of primer	Annealing temperature	Conc. of primer (pm/µl)	Length of primer with sequence
matK forward (F)	46	48°C	161.83	22 (CGATCTATTCATTCAATATTTC)
matK reverse (R)	53	48°C	208.38	22(TCTAGCACACGAAAGTCGAAGT)

2.2.3. DNA quantification

The isolated DNA was quantified using a UV spectrophotometer (Agilent Technologies Cary 60 UV–Vis) at 260 and 280 nm filters. The samples providing the ratio of absorbance at 260 nm to absorbance at 280 nm equivalent to 1.8 was only considered of good quality.

2.2.4. PCR amplification and sequencing of the matK region

The matK region was amplified in a 25 µl reaction volume comprising of 12.5 µl of GoTaq PCR master mix, 1.25 µl of matK-F and matK-R, 2 µl of DNA, and 8 µl of pyrogen-free water. The working concentrations taken for primers were 1.236 µM (matK-F) and 0.958 µM (matK-R), respectively. The details of the primers are given in Table 2. The PCR reactions were performed on a thermocycler (Applied Biosystems Veritti 96-well Thermal Cycle) using the following conditions: Denaturation of template DNA at 94°C for 4 min followed by 35 cycles of reactions: 94°C for 1 min, primer annealing at 48°C for 30 s, and primer extension at 72°C for 1 min with the final extension cycle at 72°C for 7 min. The success of the PCR was verified by agarose gel electrophoresis. The PCR product (5 µl) was run in an agarose gel (1%) and visualized under UV transilluminator. DNA sequencing was done using (ABI 3730 XL) from Bioserve Biotechnologies Pvt. Ltd.

2.2.5. Sequence submission

The obtained sequences were edited as per the NCBI guidelines (www.ncbi.nlm.nih.gov) and further compared by querying against existing global sequences in the GenBank database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using Nucleotide BLAST algorithm. Blastn and blastx were performed consecutively. The sequences were then further submitted in NCBI using the BankIt submission gateway (https://www.ncbi.nlm.nih.gov/WebSub/). The accessions number provided was retrieved from the database [Table 3].

2.3. Sequence Analysis

Complete 29 experimental sequences of *Camellia sinensis* clones were taken for final sequence analysis. A total of four sequences were excluded from the final dataset based on its short size (<600 base pairs) or sequencing error. The barcode sequences of the matK region of tea clones were further analyzed using MEGA X (Molecular Evolutionary Genetics Analysis) software (https://www.megasoftware.net/).

The DNA sequences were first aligned (pairwise alignment and multiple sequence alignment) using ClustalW and an evolutionary history was inferred using clustering methods such as neighbor-joining [14] and Unweighted Pair Group Mean Average method [15] with arithmetic mean (UPGMA) of MEGA X [16]. The bootstrap value was set to 1000 replicates. The phylogenetic tree drawn to scale was inferred with branch lengths in the same units as evolutionary distances. The evolutionary distances were computed using the Kimura 2- parameter method [17] with the units given as the number of base substitutions per site. The first, second, and third codon positions were included. Small sequences were not considered for analysis as well as all positions containing gaps and missing data were eliminated from the dataset. There were a total of 563 positions in the final dataset.

Table 3: Accession number and details of the submitted matK sequence retrieved from NCBI.

Full name	Abbreviation	NCBI accession	Unique Id	Base pairs
Ambari Vallai 2	AV2	MH649284	1111	757 bp
Phoobsering 312	P312	MK393394	1112	871 bp
Happy valley 39	HV39	MH791417	1113	864 bp
Tukdah 253	T-253	MH920315	1114	876 bp
Nanda Devi	TS 378	MH920316	1115	816 bp
Makaibari-6	MB-6	MH920317	1116	758 bp
Teesta Valley 1	TTV-1	MH920318	1117	861 bp
Kopati 1/1	K1/1	MH920319	1119	774 bp
Balasun 7/1A/76	BS 7/1A/76	MK393393	11111	833 bp
Bunnockburn 777	B-777	MK393395	11112	757 bp
Sundaram	B/5/63	MK393396	11113	644 bp
Tukdah-135	T-135	MK393397	11114	763 bp
Bunnockburn 688	B-688	MK393398	11116	833 bp
Golconda	B/6/36	MK393399	11117	644 bp
Rungli Rungiliot 17/144	RR-17/144	MK393400	11118	826 bp
Balasun 9/3/76	BS-9/3/76	MK393401	11119	756 bp
Chiradew Parbat1	CP-1	MK393402	11120	761 bp
Phoobsering 1404	P-1404	MK393403	11121	763 bp
Phoobsering 1258	P-1258	MK393404	11122	751 bp
Rungli Rungiliot 4/5	RR-4/5	MK393405	11123	746 bp
Sikkim 1	SKM-1	MK424865	11124	867 bp
Thurbo 3	Thurbo-3	MN480321	11125	761 bp
Thurbo 9	Thurbo-9	MN480322	11126	707 bp
Tukdah 145	T-145	MK424866	11127	761 bp
Tukdah 246	T-246	MK424867	11128	756 bp
Tocklai variety 19	TV-19	MK424868	11129	750 bp
Tocklai variety 14	TV-14	MK424869	11130	735 bp
Tukdah 78	T-78	MK424870	11131	755 bp
Bannockburn 157	B-157	MK424871	11132	735 bp

Further genetic pairwise distance for matK was calculated using the Kimura 2-parameter model [17] and maximum composite likelihood model [18] as given in MEGA X. Sequences were further aligned using Multalin V.5.4.1 (http://multalin.toulouse.inra.fr/ multalin/) and analysis of both high consensus and low consensus region was done. A different number of nucleotide frequencies and position of nucleotide along the consensus region were studied using Multalin (http://multalin.toulouse.inra.fr/multalin/) software V.5.4.1 [19]. Illustrative representation of the DNA barcode was done employing the sequences in matK QR code generator [20] and Biorad barcode generator (http://biorad-ads.com/DNABarcodeWeb/).

3. RESULTS

3.1. matK Amplification and Sequencing

The primer used for analysis showed successful amplification of matK [Figure 1] in all the studied clones. The size of the amplified PCR products was approximately within the range of 900 bp–1000 bp. However, the final sequencing result provided matK sequences of size ranging from 644 bp to 876 bp. The accession number for the submitted sequences and the details is provided in Table 3.

3.2. Blast Result

Blast analysis revealed 24 clones out of 29 to be 100% identical with *Camellia sinensis*. The percentage of similarity with *Camellia sinensis* recorded for other clones was 99.29% (Thurbo 9), 99.61% (Thurbo 3), 99.64% (RR-4/5), 99.88% (SKM-1), and 99.89% (P312). Despite showing 99.61% (Thurbo 3) and 99.29% (Thurbo 9) similarity with *Camellia sinensis*, both the clones showed a higher percentage of similarity with other species of *Camellia*, that is, Thurbo 3 showed 99.87% and Thurbo 9 showed 99.57% similarity with *Camellia mairei* (KJ197933.1). Thus, a percentage similarity value below 99.64% placed the clones under different species of *Camellia*.

3.3. Sequence Alignment and Phylogenetic Tree Construction

Both neighbor-joining [Online Resource 1(SM1)] and UPGMA [Figure 2] tree revealed variation among the sequences. All the combined nucleotide sequences clustered together with exceptions such as Thurbo 3 (11125), Thurbo 9 (11126) clustering together, and P312 (1112) and RR-17/144 (11118) differing from the main group. To validate our results, we also constructed a phylogenetic tree adding a sequence of different *Camellia* species (KJ197933.1) taken from the NCBI database. Thurbo 3 (11125) and Thurbo 9 (11126) now clustered with the reference sequence of *Camellia mairei* (KJ197933.1) as depicted by the neighbor- joining [Figure 3] and UPGMA tree [Online Resource 2(SM1)]).

3.4. Sequence Analysis

The genetic distances for the matK sequence ranged from 0 to 0.0090 (Nucleotide: Maximum composite likelihood method) given in Figure 4 and from 0 to 0.0089 (Nucleotide: Kimura 2-parameter method) given as Online Resource 3 (SM1). The overall mean distance was recorded as 0.0013. The results show the number of base substitutions per site and are based on an analysis of a total of 29 sequences (all codon positions included) with a total of 563 positions in the final dataset excluding the eliminated positions containing gaps and missing data.

The matK sequence showed two unique variable sites in Thurbo 3, Thurbo 9, and *Camellia mairei* that differed from the rest of the sequences. This was validated by a high consensus sequence of 563 bp prepared using Multalin software. A total of nine substitutions were observed in high consensus region where Thurbo 3 (11125) showed a total of three variable sites, Thurbo 9 (11126) showed a total of four variable sites and some single variable site was seen in P312 (1112) and RR-17/144 (11118), as shown in Figure 5 and Online Resource 4a (SM2). Study of low consensus region also revealed a total of five nucleotide substitution or variation with SKM-1 (11124) showing three variable sites, and Thurbo 9 (11126) and P1258 (11122) sowing one variable site each [Online Resource 4b (SM2)].

The sequences are represented illustratively as barcode and QR code [Figure 6 and Online Resource 5 (SM3)]. The QR code generated can be decoded as DNA sequences that make data storage and retrieval comparatively easy.

4. DISCUSSION

With the advancement of technology, sequencing analysis has uplifted the research in the molecular field, and thus, a small difference or rather variation (intraspecific or interspecific) can be studied which could not be accomplished easily using morphological means or other robust molecular techniques.

DNA barcoding is used for species identification and it utilizes many plastid and nuclear regions. A total of the seven-plastid region were explored in land plants and the combination of rbcL+matK was considered as the best combination for plant barcode [21]. Some other scientific studies reported the successful use of a combination of matK+ITS and rbcL+trnH-psbA to study 100% differences between *Cassia* species [22]. However, the efficiency of only a single matK region to differentiate *Vachellia* species from other *Acacia* species was reported earlier with concluding remarks about the possibility of utilizing matK for separating taxa at the genus level [23].

Some previous reports have suggested successful amplification and

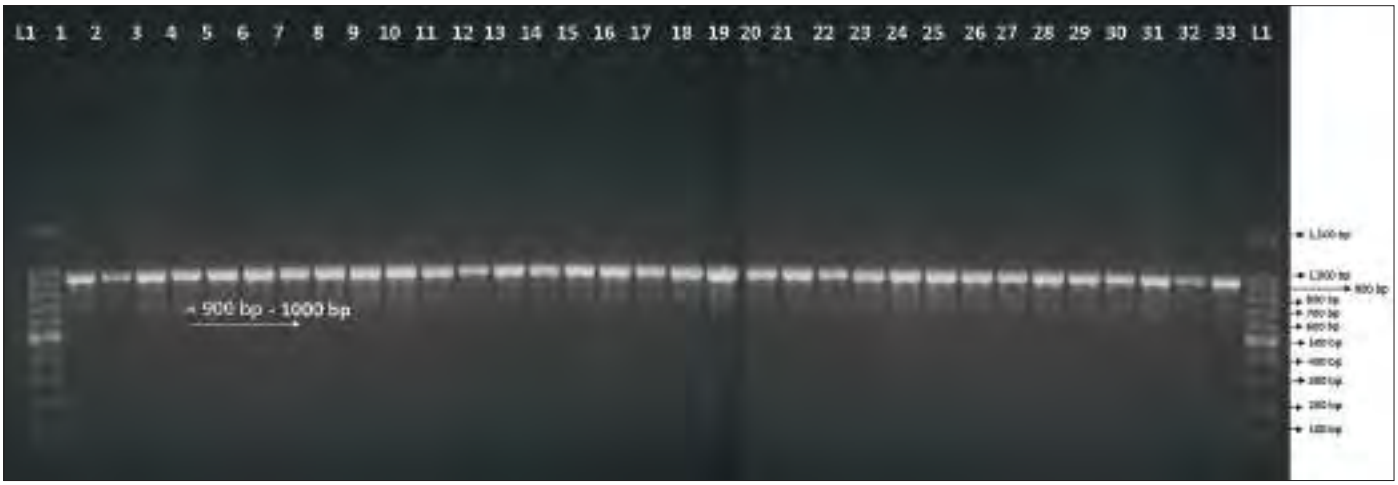


Figure 1: Amplification of the matK region. Lane L1: 100 bp ladder; lane 1–33: 33 tea clones.

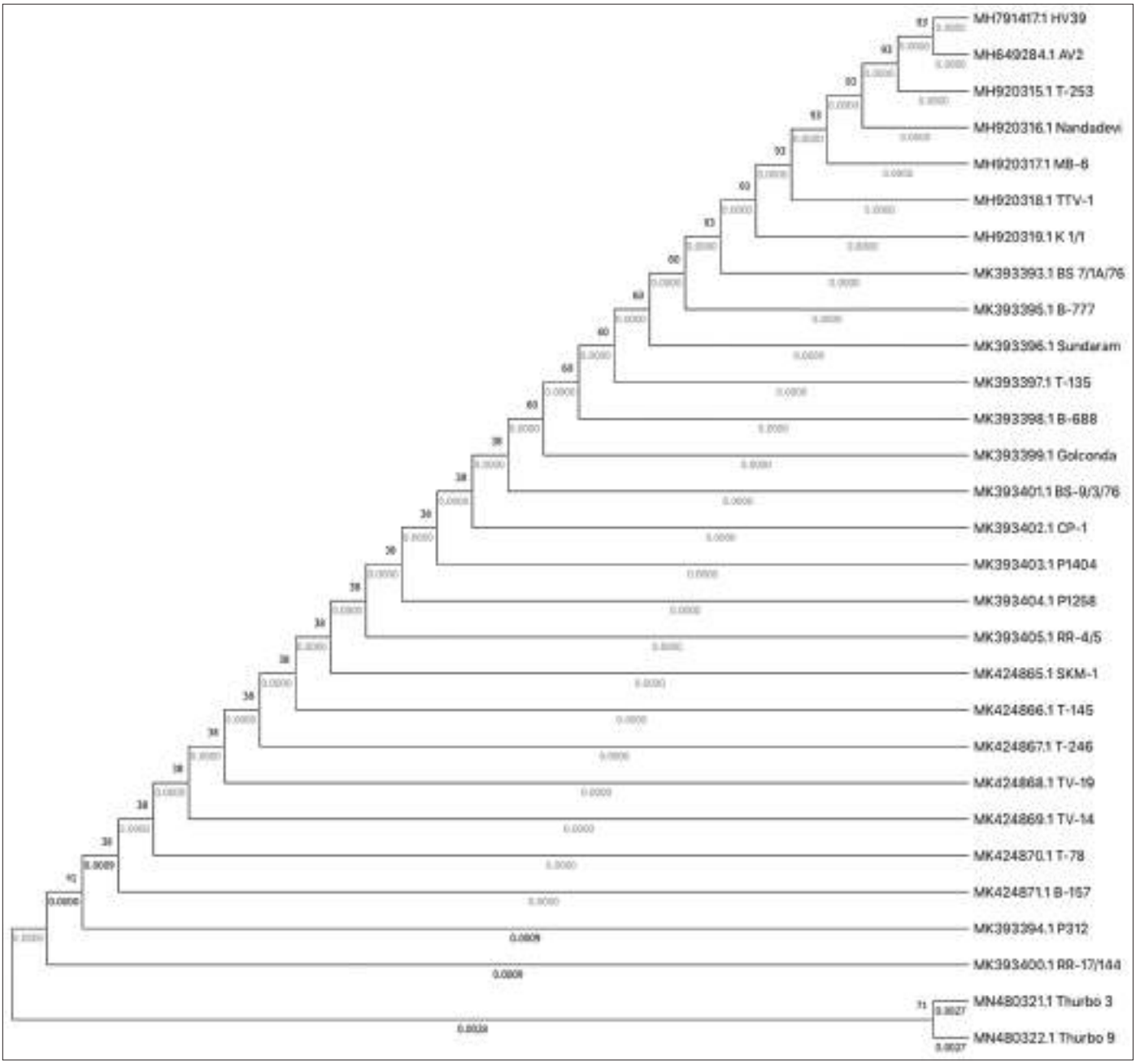


Figure 2: UPGMA tree method showing the genetic relationship of matK region between 29 tea clones.

use of the matK region to investigate phylogeny in both monocots and dicots such as *Zingiberaceae* [11], *Erythronium* [24], *Myristica fragrans* [25], local tomato [26], and oil-bearing roses [27].

The barcode technique is also used nowadays to detect any kind of contaminants. Researchers have reported the presence of adulterant with counter indications for pregnant women in bamboo tea products and also detected the origin of bamboo leaves that were used in the product [28]. Researchers have also used DNA barcodes (rbcL, matK, ITS2, and psbA-trnH) to distinguish between the commercial non-*Camellia* tea and the adulterants present in it, to assess their safety, although a limited number of original plant sequences in GenBank limited the findings [29]. There are reports of matK locus placing two genera *Myristica* and *Knena* differently at a sequence similarity of 99.43% while genus *Virola* differed with 99.25% [25]. The tomatoes were placed within the same species even with 99.64% similarity, thus limiting the assumption of percentage identity required as 99.74–100% to place organisms within the same species [30]. Whereas our study has differentiated two species at percentage identity below 99.64% with clones such as 11125 (Thurbo 3) and 11126 (Thurbo 9) showing 99.61% and 99.29% identity with *Camellia sinensis* when compared, respectively, to 99.87% and 99.57% identity with *Camellia mairei*.

Two species *T. cope* and *T. wightii* [31] did not differ at rbcl locus but showed a difference in matK (2 nucleotide difference) and trnH-psbA (1 nucleotide difference). This could broaden the interspecific variation if the two loci are considered as two-gene approach and thus they reported interspecific variation (p-distance 0.002–0.003) but

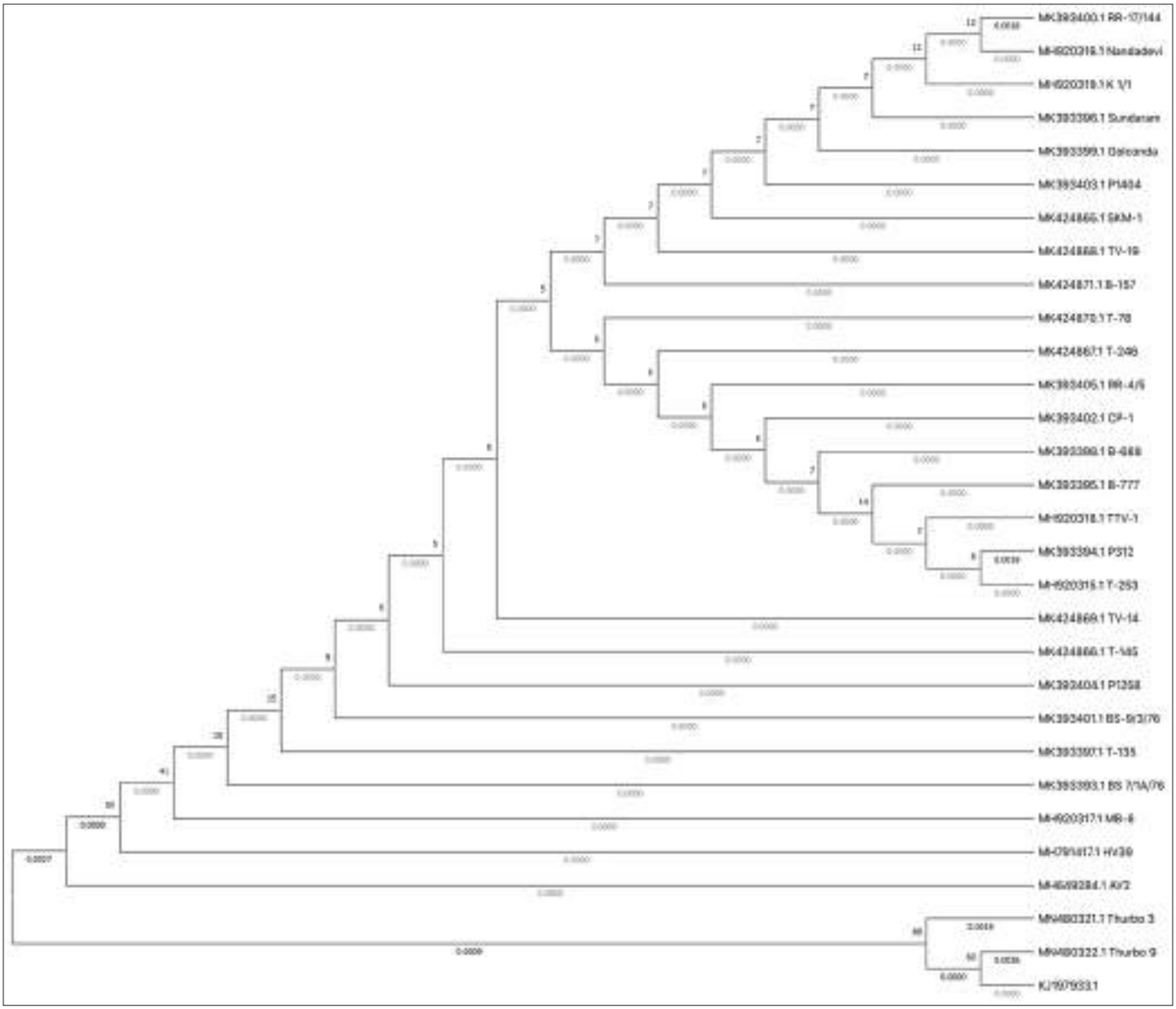


Figure 3: Neighbor-joining tree method showing the genetic relationship of matK region between 29 tea clones along with sequence of *Camellia mairei* (KJ197933.1) taken from NCBI.

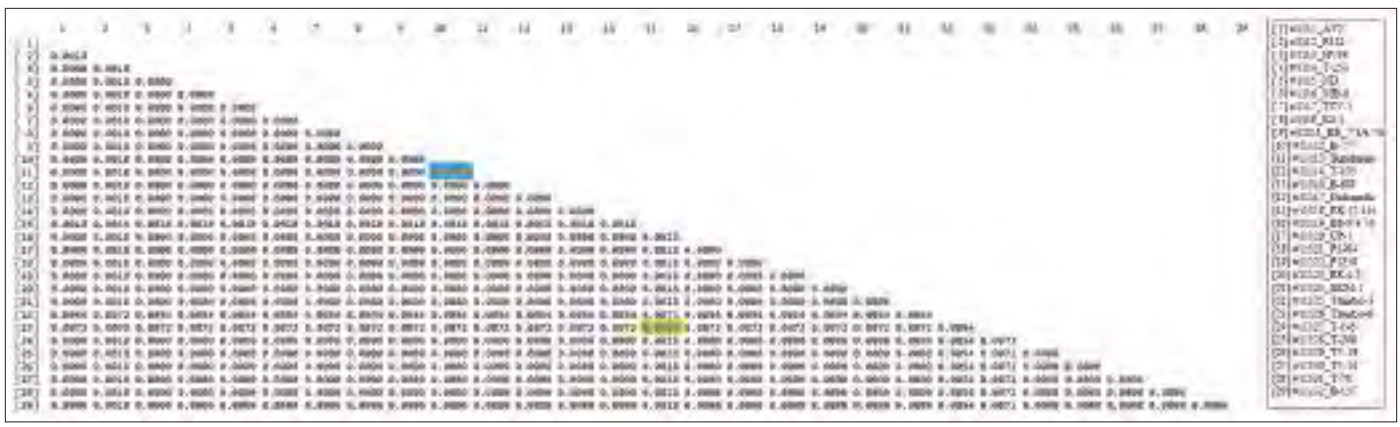


Figure 4: Genetic distances of the matK sequence calculated using nucleotide: maximum composite likelihood method.

Figure 5: Consensus region of the aligned sequences showing nucleotide substitution as highlighted by black box.

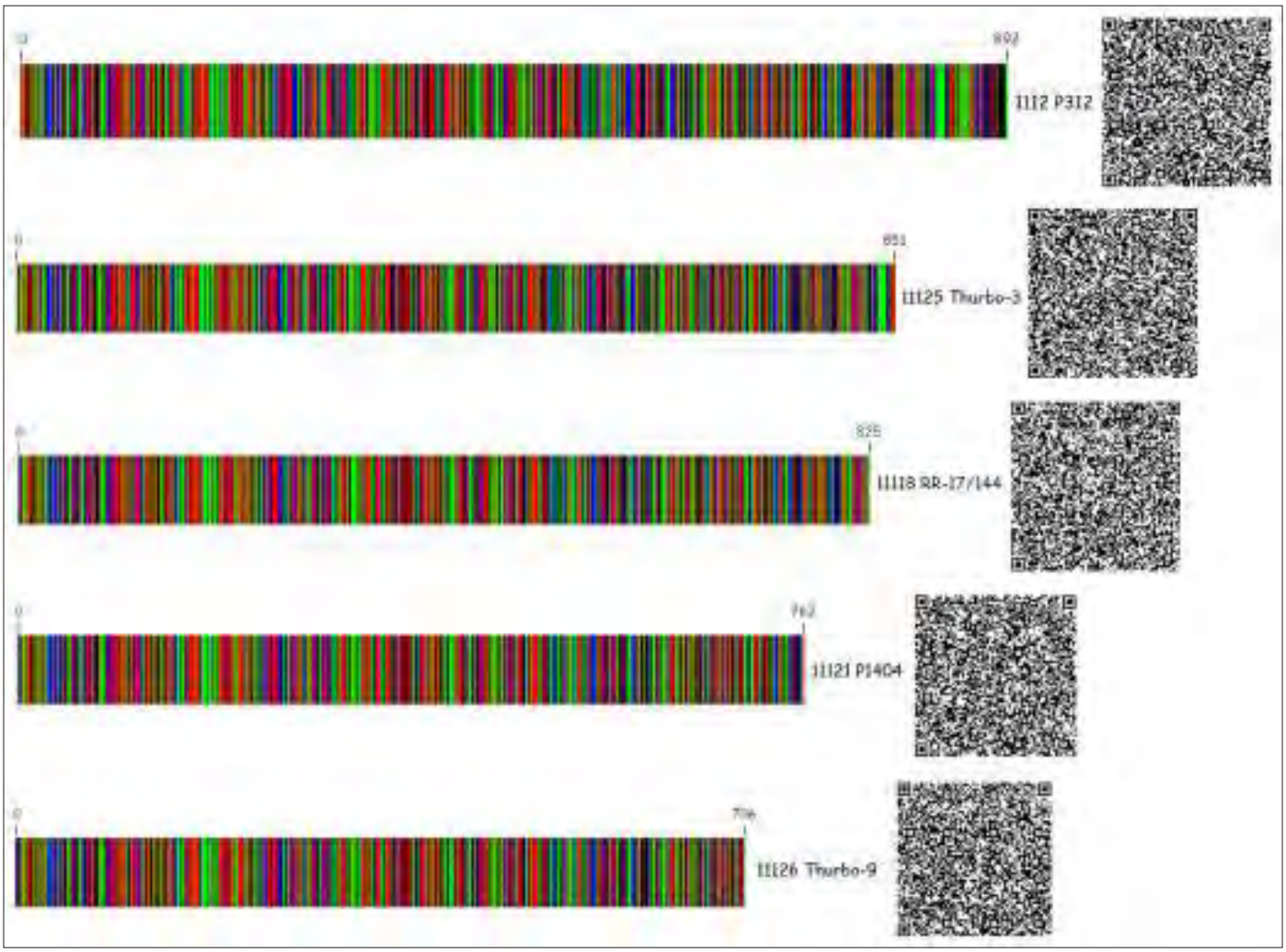


Figure 6: matK sequences represented illustratively as barcode and QR code.

no intraspecific variation (p-distance 0.00). Another work reported having three variable sites in trnH-psbA sequences among seven tomato varieties with genetic distance ranging from 0 to 0.004 [26]. On the contrary, rbcL, rpoC1, and rpoB sequences did not show any variable sites, thus suggesting it to be 100% conserved within the species. The matK locus failed to differentiate *Myristica* at species level since the blast results showed 100% similarity with other three species of *Myristica* and also reported three nucleotide differences with *Rivola sebifera* and four nucleotide difference with *Knema laurina*, thus concluding the ability of matK locus to differentiate only at the genus level within the family of Myristicaceae. However, in our study, we report a total of nine variable sites in the high consensus region and a total of five variable sites in the low consensus region of matK sequences within the same species of *Camellia sinensis*. Therefore, we report intraspecific variation and conclude with a fact of matK sequence not being 100% conserved within the same species of *Camellia*.

5. CONCLUSION

This work reports the successful use of the matK region to explore the genetic diversity and variation within the matK gene of chloroplast region among the elite clones of Darjeeling and Dooars. The employment of the matK region with its known increased rate of substitution, low transition/transversion ratio, and quick evolving rate aided to study the intraspecific variation due to probable contamination from other tea plants. The evolution of the matK region within the same species of *Camellia sinensis* was evident from our results where we report variable sites within the consensus region and conclude with the fact of the matK gene not being 100% conserved among *Camellia sinensis*. The DNA barcode of elite tea clones of Darjeeling and Dooars was thus established, wherein we conclude with the remark of matK being a good candidate for DNA barcoding of *Camellia sinensis* as well as for rapid detection of variation and molecular evidence of clones at a minimal cost, thus avoiding other robust molecular techniques.

6. ACKNOWLEDGMENTS

Reha Labar is thankful to the University Grants Commission (UGC-RGNF-SC) for providing the fellowship. We also acknowledge the prompt service provided by the National Centre of Biotechnology Information and Bioserve Sequencing Services.

7. FUNDING

University Grants Commission (UGC-RGNF-SC).

8. AUTHOR CONTRIBUTIONS

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

9. CONFLICTS OF INTEREST

The authors report no conflicts of interest in this work.

10. ETHICAL APPROVALS

This study does not involve the use of animals or human subjects.

REFERENCES

1. Lee SC, Wang CH, Yen CE, Chang C. DNA barcode and identification of the varieties and provenances of Taiwan's domestic and imported made teas using ribosomal internal transcribed spacer 2 sequences. J Food Drug Anal 2017;25:260-74.
2. Sharma RK, Bhardwaj P, Negi R, Mohapatra T, Ahuja PS. Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.). BMC Plant Biol 2009;9:53.
3. Baruah AR, Hemanta S, Biswajit B. Detection of close genetic relatedness in some tea genotypes of Assam and Darjeeling using RAPD markers. J Plant Crops 2010;38:11-5.
4. Mishra RK, Sen-Mandi S. Genetic diversity estimates for Darjeeling tea clones based on amplified fragment length polymorphism markers. J Tea Sci 2004;24:86-92.
5. Roy SC, Chakraborty BN. Genetic diversity and relationships among tea (*Camellia sinensis*) cultivars as revealed by RAPD and ISSR based fingerprinting. Indian J Biotechnol 2009;8:370-6.
6. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS One 2007;2:e508.
7. Singh HK, Parveen I, Raghuvanshi S, Babbar SB. The loci recommended as universal barcodes for plants on the basis of floristic studies may not work with congeneric species as exemplified by DNA barcoding of *Dendrobium* species. BMC Res Notes 2012;5:42.
8. Bafeel SO, Arif IA, Bakir MA, Khan HA, Al Farhan AH, Al Homaidan AA, *et al*. Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (rbcL) for barcoding of some arid plants. Plant Omics 2011;4:195.
9. Kuzmina ML, Johnson KL, Barron HR, Hebert PD. Identification of the vascular plants of Churchill, Manitoba, using a DNA barcode library. BMC Ecol 2012;12:1-11.
10. Hilu KW, Liang G. The matK gene: Sequence variation and application in plant systematics. Am J Bot 1997;84:830-9.
11. Selvaraj D, Sarma RK, Sathishkumar R. Phylogenetic analysis of chloroplast matK gene from *Zingiberaceae* for plant DNA barcoding. Bioinformation 2008;3:24.
12. Barthet MM. Expression and Function of the Chloroplast-encoded Gene matK, Doctoral Dissertation. Virginia: Virginia Polytechnic Institute and State University; 2006.
13. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 1987;19:11–15
14. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406-25.
15. Sneath PH. Numerical Taxonomy: The Principles and Practice of Numerical Classification. United States: W.H. Freeman & Co Ltd.; 1973. p. 573.
16. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 2018;35:1547-9.
17. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980;16:111-20.
18. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci 2004;101:11030-5.
19. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 1988;16:10881-90.
20. More RP, Mane RC, Purohit HJ. MatK-QR classifier: Patterns based approach for plant species identification. BioData Mining 2016;9:39.
21. Group CP, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, *et al*. A DNA barcode for land plants. Proc Natl Acad Sci 2009;106:12794-7.

22. Purushothaman N, Newmaster SG, Ragupathy S, Stalin N, Suresh D, Arunraj DR, *et al*. A tiered barcode authentication tool to differentiate medicinal *Cassia* species in India. Genet Mol Res 2014;13:2959-68.
23. Steven GN, Subramanyam R. Testing plant barcoding in a sister species complex of pantropical *Acacia* (*Mimosoideae*, *Fabaceae*). Mol Ecol Resour 2009;9:172-80.
24. Allen GA, Soltis DE, Soltis PS. Phylogeny and biogeography of *Erythronium* (*Liliaceae*) inferred from chloroplast matK and nuclear rDNA ITS sequences. Syst Bot 2003;28:512-23.
25. Tallei TE, Kolondam BJ. DNA barcoding of Sangihe Nutmeg (*Myristica fragrans*) using matK gene. Hayati J Biosci 2015;22:41-7.
26. Caprar M, Copaci CM, Chende DM, Sicora O, Sumalan R, Sicora C. Evaluation of genetic diversity by DNA barcoding of local tomato populations from North-Western Romania. Not Bot Horti Agrob Cluj Napoca 2017;45:276-9.
27. Wang H, Yao L, Cai R, Pan J, Chen X. Genetic relationship analyses of oil-bearing roses in China using matK sequences. Sci Horticult 2012;137:121-4.
28. Horn T, Haser A. Bamboo tea: Reduction of taxonomic complexity and application of DNA diagnostics based on rbcL and matK sequence data. PeerJ 2016;4:e2781.
29. Long P, Cui Z, Wang Y, Zhang C, Zhang N, Li M, *et al*. Commercialized non-*Camellia* tea: Traditional function and molecular identification. Acta Pharm Sin B 2014;4:227-37.
30. Lawodi EN. Variasi genetik tanaman tomat dari beberapa tempat di sulawesi utara berdasarkan gen matk. Pharmacon 2013;2:3098.
31. Ragupathy S, Newmaster SG, Murugesan M, Balasubramaniam V. DNA barcoding discriminates a new cryptic grass species revealed in an ethno botany study by the hill tribes of the Western Ghats in Southern India. Mol Ecol Resour 2009;9:164-71.

How to cite this article:
Labar R, Kar P, Biswas P, Sen A, Bhattacharya M. Evolution of matK Gene among the Elite Tea Clones (*Camellia sinensis*) Revealed by Nucleotide Substitution within the Consensus Region. J App Biol Biotech. 2021;9(1):32-40. DOI: 10.7324/JABB.2021.9105

INTERNATIONAL RESEARCH JOURNAL OF PHARMACY
www.irjponline.com
ISSN 2230 – 8407

Research Article

EFFECT OF SOLVENT WITH VARYING POLARITIES ON PHYTOCHEMICAL EXTRACTION FROM MATURE TEA LEAVES AND ITS EVALUATION USING BIOCHEMICAL, ANTIMICROBIAL AND *IN-SILICO* APPROACHES

Reha Labar [1], Indrani Sarkar [1], Arnab Sen [1], Malay Bhattacharya [2]*
[1] Molecular Cytogenetics Lab, Department of Botany, University of North Bengal, Siliguri, West Bengal, India
[2] Molecular Biology and Tissue Culture Laboratory, Department of Tea Science, University of North Bengal, Siliguri, West Bengal, India
*Corresponding Author Email: malaytsnbu@gmail.com

Article Received on: 06/05/19 Approved for publication: 10/06/19

DOI: 10.7897/2230-8407.1008247

ABSTRACT

Research in tea has been limited to processed or packaged tea. Owing to its economical as well as medicinal importance, our work is focused on phytochemical extraction from mature fresh tea leaves based on extracting solvents with varying polarities with the prime focus on bio activity study using biochemical, microbial and *in-silico* approaches. Potent solvents were screened from different qualitative, quantitative and antioxidant tests. Antimicrobial screening was done along with its validation using GC-MS and *in-silico* approaches. Acetone extracts was found to be the most potent solvent for extraction followed by methanol and ethanol. Acetone and methanol extracts showed antagonist activity against *Staphylococcus* sp. with Minimum Inhibitory Concentration of 4 mg/ml and 8 mg/ml respectively. Gas chromatography - mass spectrometry identified bioactive compounds like Phenol, 3, 5-bis (1, 1-dimethylethyl), caffeine and Vitamin E as the probable compounds as antibacterial agents. Further *in-silico* results validated phenol as the most potent antimicrobial compound with its binding affinity of -7.2 kcal/mol to *S. aureus* DNA gyrase. Elaborate qualitative and quantitative phytochemical profiling thus gave an idea about the potency of particular solvent in extracting specific group of compounds. Bioactivity of potent extract against *Staphylococcus* sp. was recorded out of other bacterial strains investigated. GC-MS analysis of extracts gave insight into the type of compounds extracted. Further *in- silico* results provided interesting insights into the ability of phenol to bind against *S. aureus* DNA gyrase.

Keywords: Polarities, Phytochemicals, Phenol, docking, gyrase

INTRODUCTION

Tea the most popular health beverage enriched with important phytochemicals belongs to the genus *Camellia* under Theaceae family. The important phytochemicals in tea includes the polyphenols (catechins and flavonoids), alkaloids (caffeine, theobromine, theophylline etc.), volatile oils, polysaccharides, amino acids, lipids, vitamins (e.g., vitamin C), inorganic elements (e.g., aluminum, fluorine and manganese etc.) with polyphenols being the most important compound of pharmacological importance[1]. It has already been proved in series of experiments that abundant polyphenols in tea imparts many health protecting activities[2]. These compounds have a wide range of pharmaceutical properties which includes anti oxidative, anti carcinogenic and anti arteriosclerotic property[3-6]. Polyphenolic compounds present in tea may reduce the risk of a variety of illnesses, including cancer, coronary heart disease, atherosclerosis, high blood cholesterol concentrations and high blood pressure. Most of the research work has been focused on made tea or processed tea putting a limitation as such to tea plant[7].

The objective of this work mainly focuses on the qualitative and quantitative phytochemical screening as well as studying the antioxidant and antimicrobial activity of fresh mature leaves of five different clones of *Camellia sinensis* based on extracting solvents of different polarities ranging from non polar to polar.

Very little work has been done on phytochemical screening of fresh leaves of tea using range of organic solvents with research being limited to standard solvents and processed tea. Extraction method should ensure complete extraction of the desired compounds of interest without any chemical modification[8]. Extraction and determination of biologically active compounds depends upon the type of solvent used where solvents will diffuse into solid plant tissue and solubilize compound with same polarity[9]. Aqueous mixtures of ethanol, methanol and acetone, water, are commonly exploited to extract plants[10]. Researchers have reported use of aqueous methanol, acetone and ethanol[6,11], absolute methanol[12], absolute ethanol[13] and boiling water for the extraction of polyphenols from green, black and mate teas[14]. Different solvents like water, aqueous ethanol in different extracting time has been employed to extract phenolics from green and white tea[15].

MATERIAL AND METHODS

Plant selection

Five tea elite clones were selected as experimental material, namely TS569 (S1), China (S2), AV2 (S3), P312 (S4) and Assam (S5), the details of which are provided in Supplementary Table 1. The samples were collected from Mirik hills of Darjeeling district and was identified by Taxonomy lab, of Botany Department, University of North Bengal (Accession no- 10339).

Solvent selection

Nine different solvents in increasing order of polarity namely hexane, benzene, chloroform, diethylether, ethyl acetate, acetone, ethanol, methanol and cold water were chosen as extracting solvents.

Sample extraction

Fresh leaves of the samples were washed thoroughly under running tap water, air-dried and then pulverized using a grinder. The sample was weighed and 3 g each was distributed equally and immersed in 30 ml each of nine different solvents ranging from non polar to polar. After 48 hours the aqueous cold extracts was centrifuged and the supernatant thus collected was dried and stored in 4°C for future use.

Qualitative screening of phytochemicals

Qualitative test for phytochemicals included test for flavonoid, tannin, steroid, terpenoid, cardiac glycoside, diterpenes, coumarin, reducing sugar, protein, and saponin with slight modification[16-18]. The method is given in Supplementary Table 2.

2, 2-diphenyl-1-picrylhydrazyl (DPPH) free radical scavenging assay

As mentioned previously[19], 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 atom and scavenges the unpaired electron of the stable free radical of DPPH. To 100 µl of plant extracts (5 mg/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 spectrophotometer. Ascorbic acid was taken as a standard.
The total scavenging activity was calculated using the following equation:

$$\text{DPPH scavenging (\%)} = \{A_{control} - A_{sample}/A_{control} \times 100\}$$

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

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

Ferric reducing power assay

Ferric reducing power assay was done as per the protocol[20] 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 700 nm. In a test tube 250 µl of leaf extract was taken with addition of 625 µl of sodium phosphate buffer (0.2M, pH 6.6), 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 addition of 625 µl of TCA (10%). The upper layer of the solution or supernatant (625 µl) was mixed with equal volume of distilled water and 125 µl of $FeCl_3$ (0.1% w/v). The absorbance was finally recorded at 700 nm. Higher absorbance value indicated higher reducing power.

Quantitative estimation of total flavonoids

The total flavonoids were estimated quantitatively using $AlCl_3$ method with some modifications[21]. A total of 250 µl of sample was taken in a test tube to which 750 µl of deionized water and 75 µl of 5% $NaNO_2$ was added. Following incubation for five minutes at room temperature, 150 µl of 10% $AlCl_3$ was added. It was then incubated for six minutes at room temperature followed by addition of 500µl of 1 mM NaOH and 275 µl of deionized water. It was then incubated for 30 minutes at room temperature. The absorbance of the yellowish orange color produced by interaction of flavonoid with $AlCl_3$ was recorded at 510 nm using 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 taken at different concentrations (1-5 mg/ml) [Supplementary Figure 1(a)]. The total flavonoids were thus expressed as mg of quercetin equivalent per g of extract i.e., mg QE/ g. Data was expressed as mean of triplicates ± standard deviation.

Quantitative estimation of total phenol

The total phenolic content of the sample was determined using the Folin – Ciocalteu method[22] with slight modification. 100 µl samples were taken in a test tube and to it 400 µl of 10 % Folin reagent was added (1 ml Folin + 9 ml distilled water). The mixture was incubated in dark for 5 minutes at room temperature followed by addition of 1 ml of 5% Na_2CO_3. After incubating it for 2 h in dark at room temperature, absorbance was recorded at 730 nm using 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 [Supplementary Figure 1(b)] taken at different concentrations (50-300 µg/ml). The total phenols in the extract were thus expressed as mg of gallic acid equivalent per gram of extract (mg GAE/g). Data was expressed as mean of triplicates ± standard deviation.

Antimicrobial activity

Antimicrobial activity of the acetone and methanol extracts of different tea clones was carried by agar well diffusion method[23-25]. Four different bacterial strains were employed out of which two were gram positive bacteria (*Staphylococcus* sp. and *Bacillus* sp.) and the other two were gram negative bacteria (*Escherichia coli* and *Klebsiella* sp.).

Agar well diffusion method

Sterile petri plates 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 a 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 set up was 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 with reference to appearance of clear halo or inhibition zone on the plates.

MIC (Minimal Inhibitory Concentration) determination

The MIC value was further determined only for *Staphylococcus* sp. using agar well diffusion technique as mentioned above. Different concentrations were taken i.e., 0.25 mg/ml, 0.5 mg/ml, 1 mg/ml, 4 mg/ml, 8 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 *Staphylococcus* sp. where DMSO was used as a control. The least concentration of the acetone or methanol extract of sample, showing a clear zone of inhibition was considered as its MIC[25].

GC-MS analysis

The acetone extracts of the sample S1 (TS-569) and S3 (AV2) was further employed for GC-MS analysis. The GC-MS analysis was performed using a GCMS-QP2010 Plus (Shimadzu Corporation, Kyoto Japan) the details of which are discussed in previous reports[26].

***In-silico* analysis**

DNA gyrase is one of the Type II topoisomerase relieving strain while ds DNA is unwounded by RNA polymerase or by helicase specifically at replication fork. It either introduces negative supercoiling or relaxes the positive supercoiling and thus help in the initiation and progression of DNA replication[27]. Hence gyrase may be regarded as one of the major house-keeping proteins among organisms. This fact has been used in the docking study. Crystal structure of *S. aureus* DNA gyrase (pdb id: 5CDN) was downloaded from PDB database (http://www.rcsb.org/structure/5CDN). The resolution of the structure was 2.79 Å and was obtained through X-ray diffraction method. This crude structure was polished by removal of water and addition of polar hydrogen bonds. The dimensions of grid box were taken in such a way that the whole protein could be considered for searching the best binding site and finally the protein structure was ready for docking. The structure of Phenol was downloaded from NCBI Pubchem database. The physiochemical properties of phenol were taken into consideration. Molecular docking was done in Auto dock Vina software and was visualized via Pymol software.

Statistics

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 level of significance $p \leq 0.001$.

RESULTS

Qualitative test

Acetone, methanol, and ethanol persistently proved to be the most potent solvent for various phytochemical extractions as it gave positive result for tests like flavonoid, tannin, steroid, diterpenes, terpenoids, coumarin, cardiac glycoside, saponin and reducing sugar (Figure 1). The result varied to some extent in different clones and solvents since traces of phytochemicals named above were also found in extracts prepared from, diethylether, ethylacetate, and cold water as shown in the heatmap (Figure 1). However traces of phytochemicals like cardiac glycoside, flavonoid and tannin were found in extracts of less polar solvents like benzene, chloroform and hexane of seed clones S1 and S2.

The antioxidant ability of the sample cannot be concluded only by one method[28], so we used methods to estimate total phenols and flavonoids quantitatively, and also performed DPPH free radical scavenging assay and FRP assay.

DPPH assay

The highest percentage of radical scavenging (Supplementary Figure 2) at single concentration 200 µg/ml was recorded in acetone extracts (200 µg/ml) i.e. 71.90% followed by methanol (75.31%) and ethanol (73.90%) against ascorbic acid used as standard (94.82%). Overall, the lowest scavenging activity was seen in benzene extracts. Acetone and methanol being the best solvents to extract compounds with greater antioxidant potential was further compared with ascorbic acid (standard) at various concentrations ranging from 1 mg/ml to 5 mg/ml. Both the solvent extracts i.e. acetone Figure 2(a) and methanol Figure 2(b) showed antioxidant potential or rather free radical scavenging activity at par with the standard (ascorbic acid). However, acetone gave more promising result comparatively in almost all clones. Regarding samples, acetone extracts of S1 and methanol extracts of S2 showed better antioxidant potential compared to other clones with lowest IC_{50} value being 0.111 ± 0.001 mg/ml and 0.478 ± 0.028 mg/ml respectively. The IC_{50} value of standard (ascorbic acid) was recorded as 0.057 ± 0.000 mg/ml (Table 1).

Ferric reducing power assay

Similarly, acetone extracts Figure 2(c) showed higher ferric reducing power in almost all the clones than in methanol Figure 2(d) with S1 and S2 mostly giving better results. The ferric reducing power was almost at par with the ferric reducing power of ascorbic acid used as standard.

Total Phenol

Acetone extracts gave the best result overall along with the highest value of phenol recorded in acetone extracts of China seed variety (S2) as 37.77 ± 1.28 mg GAE/g. Next to acetone, methanol and ethanol gave persistently better results (Table 2). The lowest value recorded was that of chloroform extracts of S2 i.e. 0.25 mg GAE/g of total phenols.

Total flavonoid

The highest value of 722.94 ± 127.01 mg QE/g was recorded for methanolic extracts of S1 (Table 3). The lowest value was recorded for hexane extracts of S3 as 399.42 ± 2.73 mg QE/g.

Antimicrobial screening

From the antimicrobial activity studied, acetone and methanol extracts were found to be more effective and bactericidal against *Staphylococcus* sp. Figure 3 (1) and on the other hand minimum or negligible activity was observed against other bacterial strains. When compared with the standard, pipetting more extracts into the well proved beneficial as the volumetric increase of plant extracts gave results almost at par with the 100 µl of the standard Figure 3(2). The MIC of the acetone extracts of tea clones was found to 4 mg/ml Figure 4 (1) whereas for methanol extracts the MIC was recorded as 8 mg/ml Figure 4 (2).

Gas chromatography- mass spectrometry analysis

Further analysis of the extracts using Gas chromatography- mass spectrometry identified bioactive compounds like Phenol, 3,5-bis(1,1-dimethylethyl), caffeine and Vitamin E as the probable compounds as antibacterial agents in both the extracts of S1

(Supplementary Figure 3 ; Supplementary Table 3) and S3 (Supplementary Figure 4; Supplementary Table 4).

***In-silico* results**

The binding affinity of phenol to 5CDN was found to be -7.2 kcal/mol which showed the significant effect of phenol with the gyrase protein (Figure 5). The other two compounds did not show any effective binding.

DISCUSSION

The nature of extracting solvent plays an important role in extraction of potential compounds of antioxidant activity since the compounds differ in chemical characteristics, polarities and solubilities[29]. Presence of alkaloids, flavonoids, saponins, terpenoids and phenols were reported in plant extracts of *Camellia sinensis* (purple tea) and the solvents with higher polarity i.e. water, ethanol and acetone were found to extract major phytochemicals groups than non-polar ethyl acetate and chloroform[30]. Methanol showed better extraction properties than acetone and ethyl acetate for extracting few phytochemicals like flavonoid, tannin, triterpenes, lipid and reducing sugar in black packaged tea where other solvents showed minimum activity[28]. The polar solvents and in some cases even the least polar solvents showed best result in extracting phytochemicals from fresh leaves of seed clones like S1 and S2 and we could therefore infer that in addition to extraction of samples using solvent with different polarities, extraction time and procedure, the state of sample also matters in phytochemical extraction since the phytochemical constituent slowly degenerates from the time of plucking up to manufacturing. Qualitative screening of phytochemicals is important to the pharmaceutical industry since the presence of a phytochemical of interest may lead to its further isolation, purification and characterization[31].

Table 1: **IC50 value of acetone and methanol extracts prepared in different concentrations (mg/ml) during DPPH Assay**

Accession	IC_{50} (mean ± sd) of acetone extracts (mg/ml)	IC_{50} (mean ± sd) of methanol extracts (mg/ml)
S1 (TS569)	0.111 ± 0.001	0.635 ± 0.028
S2 (China)	0.159 ± 0.002	0.478 ± 0.028
S3 (AV2)	0.600 ± 0.002	1.030 ± 0.061
S4 (PS312)	0.294 ± 0.008	1.595 ± 0.297
S5 (Assam)	0.515 ± 0.002	1.016 ± 0.039
ST (Ascorbic acid)	0.057 ± 0.000	0.056 ± 0.000

Data expressed as means of triplicates ± sd

Table 2: Determination of total phenol content (TPC) expressed as mg Gallic acid equivalent (GAE) /g in fresh leaves tea clones extracted by different solvents

Solvents	S1 Mean ± sd (mgGAE/g)	S2 Mean ± sd (mgGAE/g)	S3 Mean ± sd (mgGAE/g)	S4 Mean ± sd (mgGAE/g)	S5 Mean ± sd (mgGAE/g)
Hexane	1.61 ± 0.08	0.60 ± 0.75	3.66 ± 3.51	1.00 ± 0.04	1.08 ± 0.38
Benzene	2.14 ± 0.98	0.38 ± 0.00	2.86 ± 0.79	1.21 ± 0.04	1.50 ± 0.08
Chloroform	1.77 ± 0.08	0.25 ± 0.19	1.50 ± 0.00	1.45 ± 0.08	1.85 ± 0.26
Diethylether	3.42 ± 0.15	7.18 ± 0.72	2.49 ± 0.26	2.14 ± 0.00	3.34 ± 0.11
Ethylacetate	6.41 ± 0.15	5.40 ± 0.68	3.77 ± 0.11	3.69 ± 0.08	5.42 ± 0.26
Acetone	32.44 ± 0.98	37.77 ± 1.28	32.06 ± 0.75	14.57 ± 1.66	18.65 ± 2.07
Ethanol	18.25 ± 0.15	16.92 ± 0.98	17.88 ± 0.60	14.25 ± 4.75	13.72 ± 0.08
Methanol	32.89 ± 7.96	31.50 ± 9.99	23.82 ± 6.45	10.73 ± 0.15	11.50 ± 1.24
Water	4.04 ± 0.19	1.53 ± 0.11	2.36 ± 0.53	2.52 ± 0.38	2.25 ± 0.30

$y = 0.0075x-0.0252$, $R^2 = 0.9864$. Data expressed as means of triplicates ± sd

Table 3: Determination of total flavonoid content (TFC) expressed as mg Quercetin equivalent (QE)/g in fresh leaves of tea clones extracted by different solvents

Solvents	S1 Mean ± sd (mgQE/g)	S2 Mean ± sd (mgQE/g)	S3 Mean ± sd (mgQE/g)	S4 Mean ± sd (mgQE/g)	S5 Mean ± sd (mgQE/g)
Hexane	412.94 ± 8.19	446.74 ± 23.22	399.42 ± 2.73	418.73 ± 13.66	537.52 ± 28.68
Benzene	406.18 ± 1.37	441.91 ± 24.58	488.27 ± 73.75	401.35 ± 0.00	463.16 ± 5.46
Chloroform	428.39 ± 2.73	482.47 ± 19.12	412.94 ± 2.73	467.02 ± 10.93	477.64 ± 15.02
Diethyl ether	467.99 ± 12.29	473.78 ± 36.87	433.22 ± 1.37	448.67 ± 4.10	561.66 ± 16.39
Ethyl acetate	470.88 ± 5.46	507.58 ± 10.93	433.22 ± 4.10	424.53 ± 8.19	469.92 ± 23.22
Acetone	513.76 ± 136.03	675.62 ± 51.90	617.67 ± 32.78	579.04 ± 8.19	669.82 ± 87.41
Ethanol	528.83 ± 19.12	557.80 ± 8.19	570.35 ± 6.83	506.62 ± 9.56	611.88 ± 16.39
Methanol	722.94 ± 127.01	581.94 ± 17.75	545.24 ± 45.07	519.17 ± 35.51	574.22 ± 1.37
Water	424.53 ± 10.93	406.18 ± 6.83	432.25 ± 10.933	423.56 ± 34.14	431.29 ± 4.10

$y = 0.2071x - 0.2048$, $R^2 = 0.9625$. Data expressed as means of triplicates ± sd

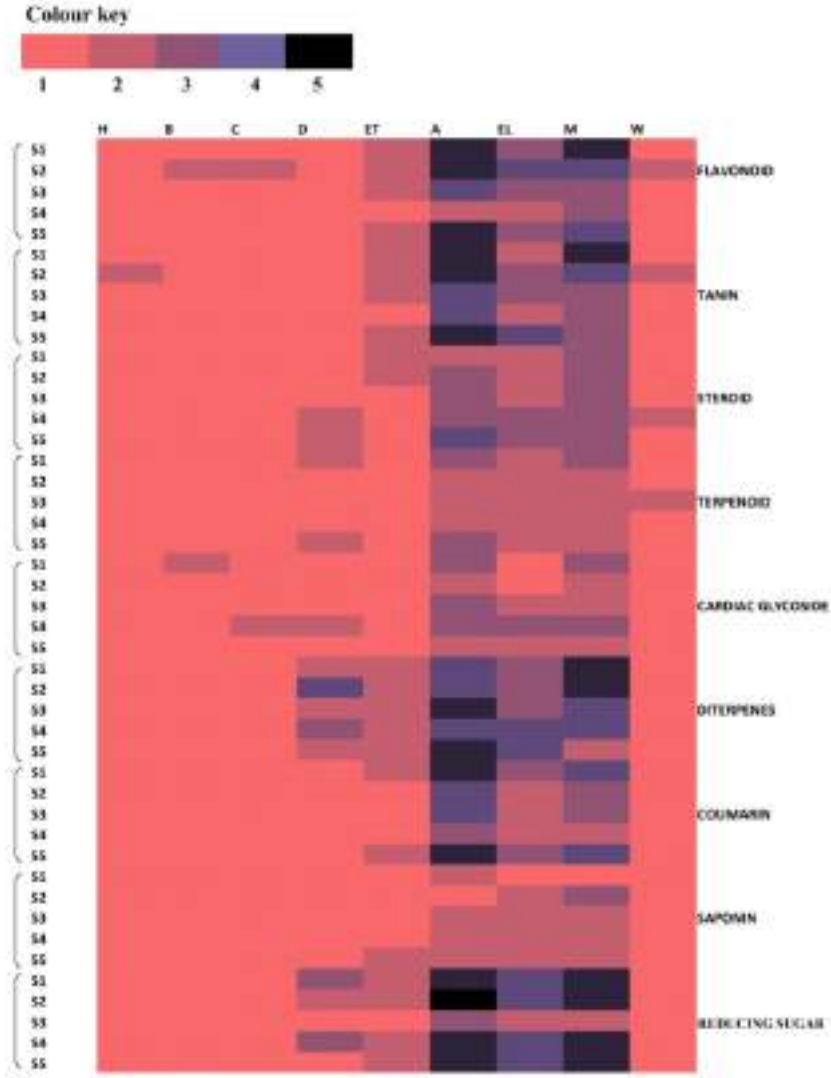


Figure 1: Heat map representing the qualitative phytochemical profiling of different extracts of selected tea clones
Colour key provided from pink to black represents the intensity in increasing order. H-Hexane, B-Benzene, C-Chloroform, D-Diethyl ether, ETEthyl acetate, A-Acetone. EL-Ethanol, M-Methanol, W-Water

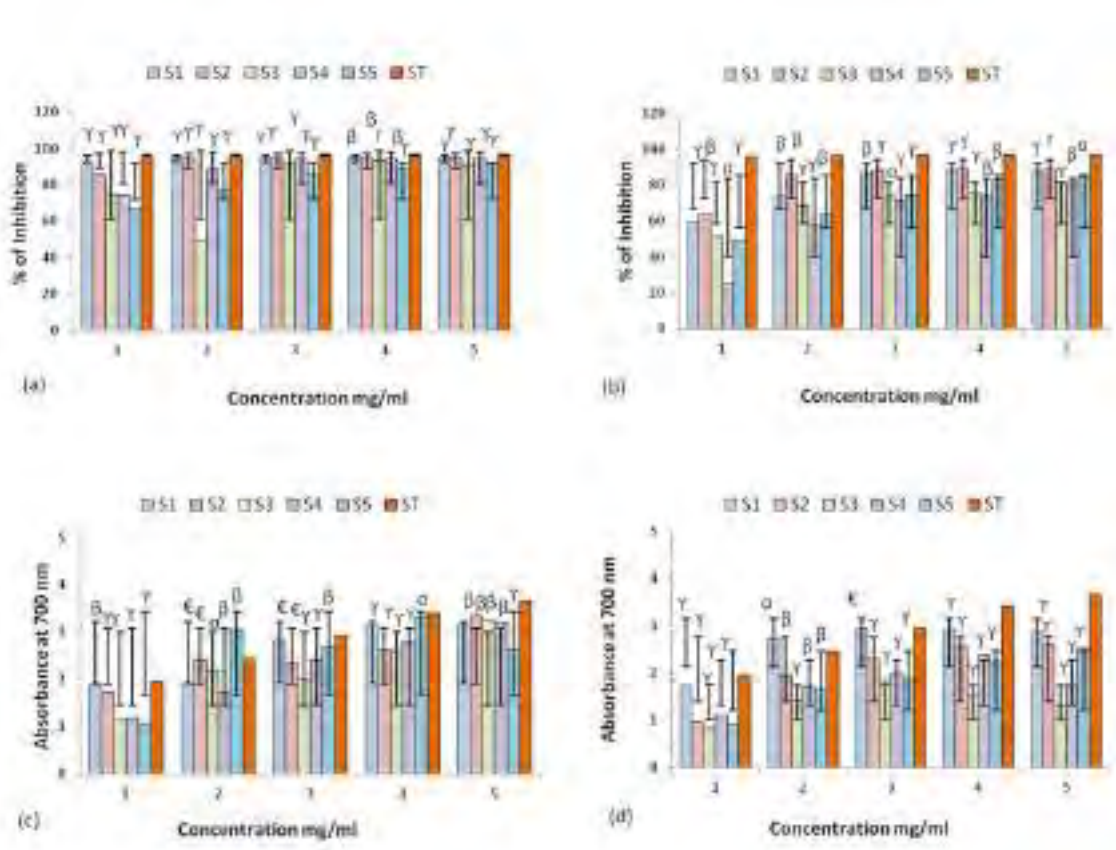


Figure 2(a): DPPH scavenging activity of acetone extracts; (b): DPPH scavenging activity of methanol extracts; (c): Ferric reducing power of acetone extracts (d): Ferric reducing power of methanol extracts;
$^{\alpha}p < 0.05$; $^{\beta}p < 0.01$; $^{\Upsilon}p < 0.001$; $^{\epsilon}$non significant

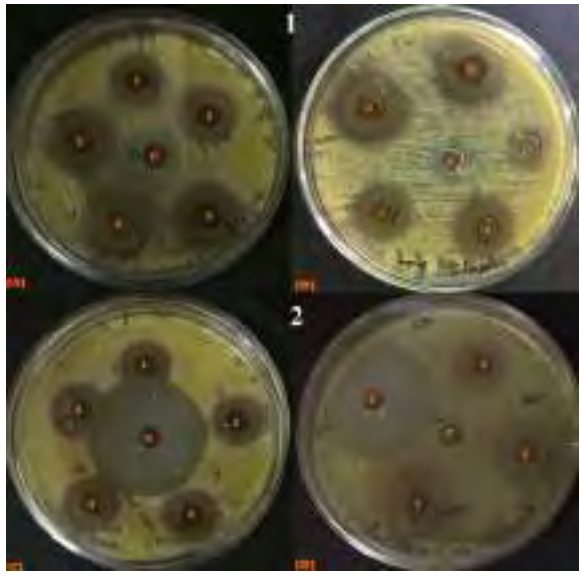


Figure 3(1): Antimicrobial activity of tea extracts against *Staphylococcus* sp.
(A) Acetone (B) Methanol, 1-5 different tea clones; C sterility control (DMSO); (2) Antimicrobial activity of tea extracts against *Staphylococcus* sp. (C) Acetone extract (100 µl of 10 mg/ml) of five different tea clones with standard; 1-5 different five clones; S positive control (Streptomycin sulphate), (D) Acetone extract of S1 with standard; 1-3 different volumetric dose; 1– 100 µl , 2– 200 µl, 3– 300 µl, S positive control (Streptomycin sulphate); C sterility control DMSO

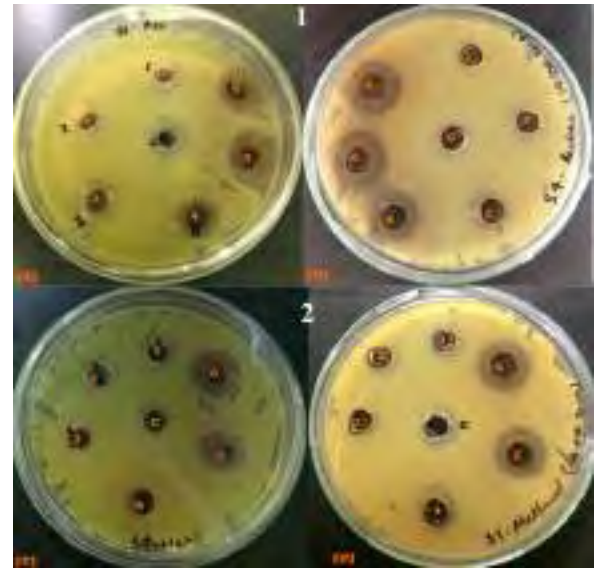


Figure 4(1): MIC value determination from antimicrobial activity of tea extracts against *Staphylococcus* sp. using agar well diffusion method (A) Acetone extracts of S2 (B) Acetone extracts of S4; 1-5 different concentrations 1-0.25 mg/ml, 2– 05 mg/ml, 3– 1 mg/ml, 4– 4 mg/ml, 5– 8 mg/m, 6– 10 mg/ml; C sterility control (DMSO); (2) MIC value determination from antimicrobial activity of tea extracts against *Staphylococcus* sp. using agar well diffusion method (C) Methanol extracts of S4(D) Methanol extracts of S7; 1-5 different concentrations 1-0.25 mg/ml, 2– 05 mg/ml, 3– 1 mg/ml, 4– 4 mg/ml, 5– 8 mg/m, 6– 10 mg/ml; C sterility control DMSO

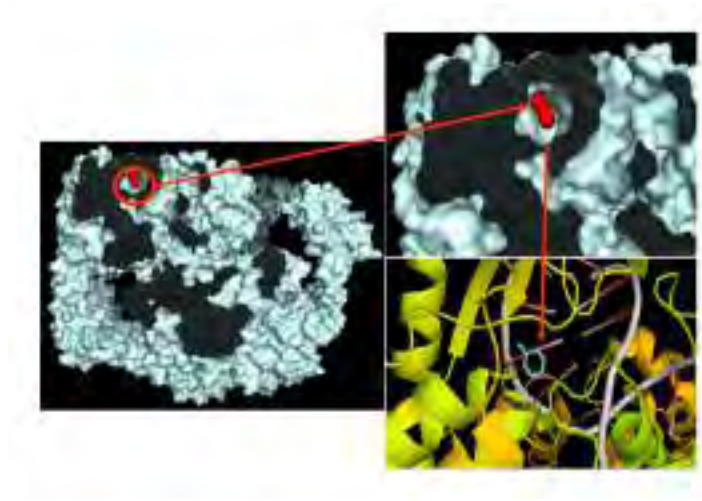


Figure 5: *In-silico* docking of *S. aureus* DNA gyrase (pdb id: 5CDN) showing significant binding affinity -7.2 kcal/mol with phenol

Methanol[28] ethanol and acetone extract[14] has already been reported to have strong antioxidant property in black tea. Previous report[14] highlights the varying antioxidant activity with changes in percent concentration of the solvent thus reporting 50% ethanol and 50% acetone with maximum antioxidant potential in mate tea and black tea respectively and also showed the capability of hot water showing moderate antioxidant potential in black tea and higher antioxidant potential in mate tea when compared to other 100% solvent. Thus it could be inferred that solvent potential may be enhanced or reduced with altering the percent concentration of the solvent.

The IC_{50} value calculated represents the exact concentration for 0.5 absorbance at 700 nm and the reducing power mostly increases with increasing concentration of antioxidant compounds[28,32,33]. Previous report on fresh tea leaf sample[34] has stated methanol extract showing highest reducing power based on state of sample where greater ferric reducing power was showed by shoot extract followed by young leaves and mature leaves.

The minimum level of phenols was recorded in less polar solvent extracts like hexane, benzene, and chloroform. However, the consistency of better result did not limit to the range of polarity in increasing or decreasing order. It varied among the different sample as well as solvents. It has already been reported in previous works that green tea was a richer source of phenolics than white tea and found 40% aqueous ethanol to be useful for extracting catechin[15]. Total phenol content estimation following Folin-Ciocalteu reagent method showed 50% acetone, 50% N, N-dimethylformamide (DMF), 50% ethanol and 50% methanol to be suitable for extracting total phenols[14].

Previously aqueous tea extracts were found to be effective against range of bacterial strains including *Staphylococcus aureus* with bactericidal activity against *Staphylococcus aureus* and *Yersinia enterocolitica*[35]. However in our study significant antimicrobial activity was found for acetone and methanol extract against *Staphylococcus* sp. with minimum or negligible activity against other strains.

Vitamin E being a powerful antioxidant showing scavenging as well as anti-inflammatory properties was also found to be effective against gram positive bacteria than gram negative ones owing to the fact of lipopolysaccharides present in their outer membrane[36]. On the other hand, the antimicrobial activity of *Arabica coffee* extracts against *Staphylococcus epidermidis* and *Enterococcus faecalis* was found to be independent from caffeine content[37]. However, caffeine was found to potentiate or enhance the antibacterial activity of amoxicillin against *Staphylococcus aureus*[38]. Significant inhibition of hexane extract of *Ulva reticulata* against *Staphylococcus aureus* and *Escherichia coli* reported presence of phenol, 3, 5-bis (1, 1-dimethylethyl) in GC-MS analysis as the probable compound giving antibacterial activity[39].

Phenol is a well-known antimicrobial agent with both bacteriostatic and bactericidal effect. Phenol may exert its effect on external membrane of bacterial cell wall, cytoplasmic membrane, organelles, cytosol as well as on spores[40]. The penetrating power of phenol into organic matter is quite high and active phenol may enter through bacterial cytoplasmic membrane via passive diffusion and active transport. Moreover, phenol acts as strong oxidizing agent and converts akin to oxide and gas, which ultimately destabilize the spores[40]. Thus, the binding of phenol to gyrase a protein may also exert as strong antimicrobial effect.

CONCLUSION

Acetone, methanol, and ethanol showed persistent and better results to extract various phytochemicals like flavonoid, tannin, steroid, diterpenes, terpenoids, coumarin, cardiac glycoside, saponin, and reducing sugar. Elaborate qualitative phytochemical profiling thus gave an idea about the potency of particular solvent in extracting specific compound of interest. Acetone extracts gave the best result overall during radical scavenging assay, ferric reducing power assay, total phenol and flavonoid estimation followed by methanol. Out of the four bacterial strain studied the acetone extracts along with methanol extracts was found effective against the *Staphylococcus* sp. with the MIC value 4 mg/ml for acetone extracts and 8 mg/ml for methanol extracts respectively. GC-MS and *in- silico* analysis suggested phenol 3, 5-bis (1, 1-dimethylethyl) as the potent antimicrobial compound. Thus, it can be concluded that abundancy of phytochemicals may vary from one clone to the other or its state (either fresh, dry or manufactured), beside varying time and procedure and the type of solvent utilized for extraction. We could also infer from our results that there isn't any pattern of potentiality of solvents in increasing or decreasing order. The collective information of this work could promote future research to utilize resources in an efficient manner.

ACKNOWLEDGEMENTS

We are very much grateful for financial assistance provided by University Grants Commission (UGC- RGNF). We also acknowledge AIRF, JNU, Delhi, for instrumentation facility for GC-MS analysis.

REFERENCES

1. Sharangi AB.Medicinal and therapeutic potentialities of tea (*Camellia sinensis* L.)- A review. Food Research International. 2009; 42(5-6):529-535.
2. Manzocco L, Anese M, Nicoli MC.Antioxidant properties of tea extracts as affected by processing. LWT-Food Science and Technology. 1998; 31(7-8):694-698.
3. Atoui AK, Mansouri A, Boskou G, Kefalas P .Tea and herbal infusions: their antioxidant activity and phenolic profile. Food Chemistry. 2005; 89(1):27-36.
4. Dufresne CJ, Farnworth ER. A review of latest research findings on the health promotion properties of tea. The Journal of Nutritional Biochemistry.2001; 12(7):404-421.
5. Filip R, Ferraro GE.Researching on new species of "Mate": *Ilex brevicuspis*. European Journal of Nutrition. 2003; 42(1):50-54.
6. Wang H, Helliwell K.Determination of flavonols in green and black tea leaves and green tea infusions by high-performance liquid chromatography. Food Research International. 2001; 34(2-3):223-227.
7. Mukhtar H, Ahmad N. Tea polyphenols: prevention of cancer and optimizing health. The American Journal of Clinical Nutrition. 2000; 71(6):1698S-1702S.
8. Zuo Y, Chen H, Deng Y.Simultaneous determination of catechins, caffeine and gallic acids in green, Oolong, black and pu-erh teas using HPLC with a photodiode array detector. Talanta. 2002; 57 (2):307-316.
9. Tiwari P, Kumar B, Kaur M, Kaur G, Kaur H. Phytochemical screening and extraction: a review Internationale Pharmaceutica Sciencia. 2011; 1(1):98-106.
10. Sun T, Ho CT. Antioxidant activities of buckwheat extracts. Food Chemistry 2005; 90(4):743-749.
11. Martinez MADP, Pelotto JP, Basualdo N. Distribution of flavonoid aglycones in *Ilex* species (Aquifoliaceae). Biochemical Systematics and Ecology.1997; 25(7):619-622.
12. Yao L, Jiang Y, Datta N, Singanusong R, Liu X, Duan J, *et al.* HPLC analyses of flavanols and phenolic acids in the fresh young shoots of tea (*Camellia sinensis*) grown in Australia. Food Chemistry. 2004; 84(2):253-263.
13. Opie SC, Robertson A, Clifford MN. Black tea thearubigins-their HPLC separation and preparation during in-vitro oxidation. Journal of the Science of Food and Agriculture. 1990; 50(4):547-561.
14. Turkmen N, Sari F, Velioglu YS. Effects of extraction solvents on concentration and antioxidant activity of black and black mate tea polyphenols determined by ferrous tartrate and Folin-Ciocalteu methods. Food Chemistry. 2006; 99(4):835-841.
15. Rusak G, Komes D, Likic S, Horzic D, Kovac M. Phenolic content and antioxidative capacity of green and white tea extracts depending on extraction conditions and the solvent used. Food Chemistry. 2008; 110(4):852-858.
16. Brain KR, & Turner TD. The Practical Evaluation of Phytopharmaceuticals. Bristol: Wright-Scientechnica; 1975.
17. Kumar A, Ilavarasan R, Jayachandran T, Decaraman M, Aravindhan P, Padmanabhan N, *et al.* Phytochemicals investigation on a tropical plant, *Syzygium cumini* from Kattuppalayam, Erode district, Tamil Nadu, South India. Pakistan Journal of Nutrition. 2009; 8(1):83-85.
18. Ngbede J, Yakubu RA, Nyam DA.Phytochemical screening for active compounds in *Canarium schweinfurthii* (Atile) leaves from Jos North, Plateau State, Nigeria. Research Journal of Biological Sciences. 2008; 3(9):1076-1078.
19. Blois MS. Antioxidant determinations by the use of a stable free radical. Nature. 1958; 181(4617):1199-1200.
20. Aiyegoro OA, Okoh AI .Phytochemical screening and polyphenolic antioxidant activity of aqueous crude leaf extract of *Helichrysum pedunculatum*. International Journal of Molecular Sciences. 2009; 10(11):4990-5001.
21. Zou Y, Lu Y, Wei D. Antioxidant activity of a flavonoid-rich extract of *Hypericum perforatum* L. in vitro. Journal of Agricultural and Food Chemistry.2004; 52(16):5032-5039.
22. Folin O, Ciocalteu V.On tyrosine and tryptophane determinations in proteins. Journal of Biological Chemistry. 1927; 73(2):627-650.
23. Boyanova L, Gergova G, Nikolov R, Derejian S, Lazarova E, Katsarov N, *et al.* Activity of Bulgarian propolis against 94 *Helicobacter pylori* strains in vitro by agar-well diffusion, agar dilution and disc diffusion methods. Journal of Medical Microbiology. 2005; 54(5):481-483.
24. Holder IA, Boyce ST.Agar well diffusion assay testing of bacterial susceptibility to various antimicrobials in concentrations non-toxic for human cells in culture. Burns. 1994; 20(5):426-429.
25. Okeke MI, Iroegbu CU, Eze EN, Okoli AS, Esimone CO.Evaluation of extracts of the root of *Landolphia owerrience* for antibacterial activity. Journal of Ethnopharmacology. 2001; 78(2-3):119-127.
26. Das S, Vasudeva N, Sharma S. Chemical composition of ethanol extract of *Macrotyloma uniflorum* (Lam.) Verdc. using GC-MS spectroscopy.Organic and Medicinal Chemistry Letters. 2014;4(1): 13.
27. Sutormin D, Rubanova N, Logacheva M, Ghilarov D, Severinov K. "Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the *Escherichia coli* genome". Nucleic Acids Research. 2018; 47(3):1373-88.
28. Patil MP, Patil KT, Ngabire D, Seo YB, Kim GD. Phytochemical, Antioxidant and Antibacterial Activity of Black Tea (*Camellia sinensis*).International Journal of Pharmacognosy and Phytochemical Research. 2016; 8(2):341-346.
29. Ozarkar KR. Studies on anti-inflammatory effects of two herbs *Cissus quadrangularis* Linn. and *Valeriana wallichi* DC using mouse model. Ph. D. Thesis, University of Mumbai, Mumbai; 2005.
30. Geoffrey KK, John KM, Naomi M, Simon KM.Qualitative phytochemical screening of *Camellia sinensis* and *Psidium guajava* leave extracts from Kericho and Baringo counties. International Journal of Advanced Biotechnology and Research. 2014; 5(3):506-512.
31. Ugochukwu SC, Uche A, Ifeanyi O. Preliminary phytochemical screening of different solvent extracts of stem bark and roots of *Dennetia tripetala* G. Baker. Asian Journal of Plant Science and Research. 2013; 3(3):10-13.
32. Ferreira IC, Baptista P, Vilas-Boas M, Barros L. Free-radical scavenging capacity and reducing power of wild edible mushrooms from northeast Portugal: Individual cap and stipe activity. Food Chemistry. 2007; 100(4):1511-1516.
33. Manivasagan P, Alam MS, Kang KH, Kwak M, Kim SK. Extracellular synthesis of gold bionanoparticles by *Nocardiopsis* sp. and evaluation of its antimicrobial, antioxidant and cytotoxic activities. Bioprocess and Biosystems Engineering. 2015; 38(6):1167-1177.
34. Chan EW, Lim YY, Chew YL. Antioxidant activity of *Camellia sinensis* leaves and tea from a lowland plantation in Malaysia. Food Chemistry. 2007; 102(4):1214-1222.

35. Yam TS, Shah S, Hamilton-Miller JM. Microbiological activity of whole and fractionated crude extracts of tea (*Camellia sinensis*), and of tea components. FEMS Microbiology Letters. 1997; 152(1):169-174.
36. Jehad MT, Mshimesh BA, Aziz FM, Al-Salih DA. Antibacterial effects of vitamin E: In vitro study. Jornal of Biotechnology Research Center. 2013; 7(2):17-23.
37. Runti G, Pacor S, Colomban S, Gennaro R, Navarini L, Scocchi M. Arabica coffee extract shows antibacterial activity against *Staphylococcus epidermidis* and *Enterococcus faecalis* and low toxicity towards a human cell line. LWT-Food Science and Technology. 2015; 62(1):108-114.
38. Esimone CO, Okoye FB, Nworu CS, Agubata CO.*In vitro* interaction between caffeine and some penicillin antibiotics against *Staphylococcus aureus*. Tropical Journal of Pharmaceutical Research. 2008; 7(2):969-974.
39. Dhanya KI, Swati VI, Vanka KS, Osborne WJ. Antimicrobial activity of *Ulva reticulata* and its endophytes Journal of Ocean University of China. 2016; 15(2):363-369.
40. Sabbineni, J. Phenol an effective antibacterial agent. Research and Review. Journal of Medicinal & Organic Chemistry. 2016; 3(2): 182-191.

Cite this article as:

Reha Labar *et al*. Effect of solvent with varying polarities on phytochemical extraction from mature tea leaves and its evaluation using biochemical, antimicrobial and *in-silico* approaches. Int. Res. J. Pharm. 2019;10(8):59-67 http://dx.doi.org/10.7897/2230-8407.1008247

Source of support: University Grants Commission- Rajiv Gandhi National Fellowship for SC (UGC- RGNF), Conflict of interest: None Declared

NBU Journal of
Plant Sciences

Vol. 11, March 2019 p. 115-123
ISSN No. 09746927

Effect of solvents on qualitative and quantitative phytochemical constituent profiles of fresh leaves of TV26

Reha Labar[1], Arnab Sen[1] and Malay Bhattacharya[2]*
[1]Molecular Cytogenetics Lab, Department of Botany, University of North Bengal, Siliguri, West Bengal, 734013
[2]Molecular Biology and Tissue Culture Laboratory, Department of Tea Science, University of North Bengal, Siliguri, West Bengal, 734013
Email: malaytsnbu@gmail.com

Abstract
The aim of this study was to qualitatively and quantitatively screen and identify major phytochemical groups from leaves extract of TV26 extracted by nine different solvents having different polarities. Qualitative screening suggested acetone, methanol, ethanol and ethyl acetate to be the most potent solvent for various phytochemical extractions like flavonoid, tannin, steroid, diterpenes, terpenoids, coumarin, cardiac glycoside, saponin, protein and reducing sugar. The highest percent of radical scavenging was recorded in cold water extracts (5mg/ml) i.e. 91.10% and was at par with 93.40% percent scavenging activity of ascorbic acid taken as standard (5mg/ml). Methanol, acetone, ethanol and ethyl acetate gave the best results with the total phenol content value (GAE) recorded as 100.60mg/g, 87.07mg/g, 58.73 mg/g and 51.47mg/g respectively with methanol giving the best result. Acetone extracts (5mg/ml) showed higher ferric reducing power with IC_{50} value 426.45±1.12 µg/ml compared to the standard (ascorbic acid) 270.35±0.66 µg/ml. Our findings suggest that the polar solvents were more beneficial and potent against the other non polar counterparts during phytochemical extraction but the polarity of solvents need not be in increasing order since we can assume from our results that acetone being less polar than ethanol, methanol and water showed better results. In addition to different polarities, state of the sample and extraction technique is also crucial for better extraction.

Keywords: TV26, Phytochemicals, Solvent, Qualitative, Quantitative.

Tea, the most popular health beverage next to water in the world, has aroused great interest among the world of scientific research due to its beneficial health effects. It belongs to the genus *Camellia* under Theaceae family with, *Camellia sinensis* (L.)O. Kuntze being the mostly used species for making the health beverage (Kaundun *et al*. 2000). The important phytochemicals in tea includes the polyphenols (catechins and flavonoids), alkaloids (caffeine, theobromine, theophylline etc.), volatile oils, polysaccharides, amino acids, lipids, vitamins (e.g., vitamin C), inorganic elements (e.g., aluminium,fluorine and manganese etc.) with polyphenols being the most important compound of pharmacological importance (Sharangi 2009). It has already been proved in series of experiments that abundant polyphenols in tea imparts many health protecting activities (Manzocco *et al*. 1998). These compounds have a wide range of pharmaceutical properties which includes antioxidative, anticarcinogenic and antiarteriosclerotic (Atoui *et al*. 2005; Dufresne and Farnworth 2001; Filip and Ferraro 2003; Wang and Helliwell 2001). Polyphenolic compounds present in tea may reduce the risk of a variety of illnesses, including cancer, coronary heart disease, atherosclerosis, high blood cholesterol concentrations and high blood pressure. Most of the research work has been focused on manufactured tea or processed tea putting a limitation as such to tea plant (Mukhtar and Ahmad 2000).

Varied range of solvents has been used for extracting polyphenols from plants (Chavan *et al*. 2001) and the role of extracting solvents and the method of extraction on total extraction yield has already been highlighted in several articles (Goli *et al*. 2005). Extraction method should ensure complete extraction of the desired compounds of interest without any chemical modification (Zuo *et*

al. 2002). Extraction and determination of biologically active compounds depends upon the type of solvent used where solvents will diffuse into solid plant tissue and solubilize compound with same polarity (Tiwari *et al*. 2011).

Aqueous mixtures of ethanol, methanol and acetone, water, are commonly exploited to extract plants constituents (Sun and Ho 2005). Researchers have reported use of aqueous methanol, acetone and ethanol (Martinez *et al*. 1997; Wang and Helliwell 2001), absolute methanol (Yao *et al*. 2004), absolute ethanol (Opie *et al*. 1990) and boiling water for the extraction of polyphenols from green, black and mate teas (Turkmen *et al*. 2006). Different solvents like water, aqueous ethanol in different extracting time has been employed to extract phenolics from green and white tea (Rusak *et al*. 2008).

Very little research has been done on phytochemical screening of tea using range of organic solvents. Research has been mostly limited to standard solvents and processed tea thus limiting the study of different phytochemicals in fresh leaves of tea extracted by range of organic solvents. Therefore our study mostly focused on preliminary area of research covering the qualitative and quantitative phytochemical screening as well as studying the antioxidant activity of fresh leaves of TV26 based on extracting solvents of different polarities. The range of organic solvent system with different polarities included absolute hexane, benzene, chloroform, diethylether, ethylacetate, acetone, ethanol, methanol and water.

Materials and methods

Chemicals

10% $FeCl_3$, 1% $FeCl_3$, chloroform, concentrated H_2SO_4, glacial acetic acid, 5% copper acetate, 10%NaOH, Benedict reagent, concentrated HNO_3, distilled water, 2,2-diphenyl-1-picrylhydrazyl, Hexane, Benzene, Diethylether, Ethylacetate, Acetone, Ethanol, Methanol and Water, sodium phosphate buffer (0.2M, pH 6.6), K2 $[Fe(CN)_6]$ 1(% w/v), TCA (10%), $FeCl_3$ (0.1% w/v), Folin reagent, 5% Na_2CO_3, gallic acid, ascorbic acid.

Variety selection

TV26 (Tocklai variety) was chosen as the experimental variety and leaves were collected from tea garden of University of North Bengal. TV26 is a variety of Cambod origin with average quality, high yield and is moderately tolerant to drought (http://www.tocklai.net/activities/tea-cultivation/tra-garden-series-clones/).

Sample extraction

Fresh leaves of TV26 collected were washed thoroughly under running tap water, air dried and then pulverized using a grinder. The sample was weighed and 3g each was distributed equally and immersed in 30 ml each of nine different solvents ranging from polar to non polar. Nine different solvents in increasing order of polarity namely hexane, benzene, chloroform, diethylether, ethylacetate, acetone, ethanol, methanol and water were chosen as extracting solvents. After 48 hours the aqueous cold extracts was centrifuged and the supernatant thus collected was dried and dissolved in methanol. The aliquots of different aqueous extract were thus stored at room temperature.

Qualitative screening of phytochemicals

Qualitative test for phytochemicals included test for flavonoid, tannin, steroid, terpenoid, cardiac glycoside, diterpenes, coumarin, reducing sugar, protein, and saponin with slight modification (Brain and Turner 1975; Kumar *et al*. 2009; Ngbede *et al*. 2008).

Flavonoid

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

Tanin

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

Steroid

About 250 µl of 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. Appearance of reddish brown colour ring confirmed the presence of steroids.

Terpenoid

The evaporated extract (250 µl) was dissolved in chloroform and concentrated H_2SO_4 was added from the sidewall of test tubes and then shaken.

Formation of red to reddish brown coloration at the base confirmed the presence of terpenoids.

Cardiac glycoside

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

Diterpenes

Copper acetate was performed to confirm the presence of diterpenes wherein addition of few drops of 5 % copper acetate to 250 µl of extract dissolved in equal volume of distilled water formed emerald green color.

Coumarin

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

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 sample gave green coloration (color varies from green to red depending upon the percentage of reducing sugar present).

Protein

Xanthoproteic test was performed where 1ml concentrated HNO_3 was added to about 250 µl sample thus giving a yellow precipitate.

Saponin

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

Quantitative screening of Phytochemicals

2, 2-diphenyl-1-picrylhydrazyl (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 atom 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 spectrophotometer. Ascorbic acid was taken as a standard.

The total scavenging activity was calculated using the following equation:

$$\text{DPPH scavenging (\%)} = [A_{control} - A_{sample}/A_{control} \times 100]$$

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

Folin-Ciocalteu reagent method for 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 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 spectrophotometer. Gallic acid was taken as a standard and the total phenol content was expressed in mg of GAE per g of extract or GAE mg/g.

Ferric reducing power assay

Ferric reducing power assay was done as per the protocol (Aiyegoro and Okoh 2009) with slight modification. In a test tube 250 µl of leaf extract was taken with addition of 625 µl of sodium phosphate buffer (0.2M, pH 6.6), 625 µl of K2 $[Fe(CN)_6]$ 1(% w/v) and incubated for 20 minutes at 50 °C. The tubes were then cooled and centrifuged at 3000 rpm after addition of 625 µl of TCA (10%). The upper layer of the solution or supernatant (625 µl) was mixed with equal volume of distilled water and 125 µl of $FeCl_3$ (0.1% w/v). The absorbance was finally recorded at 700nm. Higher absorbance value indicated higher reducing power.

Results and Discussion

Acetone, methanol, ethanol and ethyl acetate persistently proved to be the most potent solvent for various phytochemical extractions as it gave

positive result for various test like flavonoid, tannin, steroid, diterpenes, terpenoids, coumarin,cardiac glycoside, saponin, protein and reducing sugar (Table1, Fig. 1). However, extraction of phytochemicals varied in some solvents. Methanol, cold water, ethanol, and ethylacetate proved to be the best solvents for extracting cardiac glycosides. Traces of steroid, diterpenes, reducing sugar, saponin was found in diethylether extracts. Traces of phytochemicals were also seen in less polar solvents like chloroform (steroid and saponin) and benzene (steroid and protein). The nature of extracting solvent plays an important role in extraction of potential compounds of antioxidant activity since the compounds differ in chemical characteristics, polarities and solubilities (Ozarkar 2005). Presence of alkaloids, flavonoids, saponins, terpenoids and phenols were reported in plant extracts of *Camellia sinensis* (purple tea) and the solvents with higher polarity i.e. water, ethanol and acetone were found to extract major phytochemicals groups than non-polar ethyl acetate and chloroform (Geoffrey *et al*. 2014). Methnaol showed better extraction properties than acetone and ethylacetate for extracting few phytochemicals like flavonid, tannin, triterpenes, and lipid and reducing sugar in black packaged tea. Other solvents showed minimum activity (Patil *et al*. 2016). The polar solvents and in some cases even the least polar solvents showed best result in extracting phytochemicals from fresh leaves of TV26 and we could therefore infer that in addition to extraction of samples using solvent with different polarities, extraction time and procedure, the state of sample also matters in phytochemical extraction. The phytochemical constituent slowly degenerates from the time of plucking upto manufacturing. Qualitative screening of phytochemicals is important to the pharmaceutical industry since the presence of a phytochemical of interest may lead to its further isolation, purification and characterization (Ugochukwu and Arukweuche 2013).

Table 1. Qualitative phytochemical screening of TV26 extracted by nine different solvents.

	Hexane	Benzene	Chloroform	Diethylether	Ethylacetate	Acetone	Ethanol	Methanol	Water
Flavonoid	-	-	-	-	++	++++	+	+++	-
Tanin	-	-	-	-	+	++++	++	+++	-
Steroid	-	+	+	++	+++	++++	+++	+++++	+
Terpenoid	-	-	-	-	+	++++	++	+++	-
Cardiac glycoside	-	-	-	-	++	-	++	+++++	+++
Diterpenes	-	-	-	+	++	+++++	+++	++++	-
Coumarin	-	-	-	-	++	++++	++	+++	-
Saponin	-	-	++	-	++	+	+	+++	++
Reducing sugar	-	-	-	+	+	+++	++	++++	-
Protein	-	+	-	+	++	+++	++	++	-

+ Positive test; - negative test; the number of + or - indicates the higher or lower intensity.

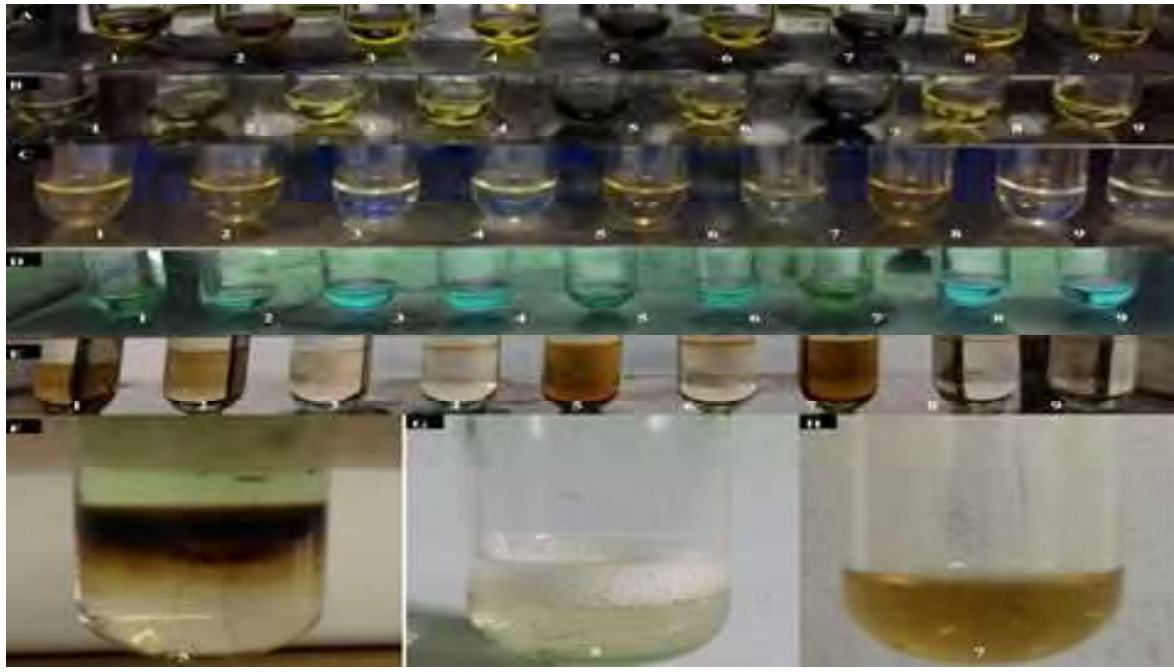


Fig. 1: Qualitative phytochemical screening of TV26 in nine different extracting solvents **1**–Ethanol,**2**–Ethylacetate, **3**-Water, **4**– Benzene, **5**– Methanol, **6**– Diethylether, **7**-Acetone, **8**-Chloroform, **9**-Hexane **[A]** Flavonoid; **[B]** Tanin; **[C]** Coumarin; **[D]** Reducing sugar; **[E]** Terpenoid; **[F]** Cardiac glycoside; **[G]** Saponin; **[H]** Protein.

The antioxidant ability of the sample cannot be concluded only by one method (Patil *et al*. 2016), so we used methods to estimate total phenols quantitatively, DPPH free radical scavenging assay and FRP assay. During DPPH assay, antioxidants acts as a proton donor where the free radical is scavenged and absorbance is decreased thereby rendering change in color from purple to yellow (Liu *et al*. 2010; Manivasagan *et al*. 2015). The highest percent of radical scavenging was recorded in cold water extracts (5mg/ml) i.e. 91.10% and was at par with 93.40% percent scavenging activity of ascorbic acid taken as standard (5mg/ml). Diethylether, ethylacetate, acetone, ethanol, and methanol also gave better results with percent scavenging activity above 50% (Fig. 2). The lowest scavenging activity was seen in hexane extracts. Methanol (Patil *et al*. 2016) ethanol and acetone extract (Turkmen *et al*. 2006b) has already been reported to have strong antioxidant property in black tea. Turkmen (2006) reported varying antioxidant activity with changes in percent concentration of the solvent thus reporting 50% ethanol and 50% acetone with maximum antioxidant potential in mate tea and black tea respectively and also showed the capability of hot water showing moderate antioxidant potential in black tea and higher antioxidant potential in mate tea when compared to other 100% solvent. Thus we could infer that solvent potential may be enhanced or reduced with altering the percent concentration of the solvent.

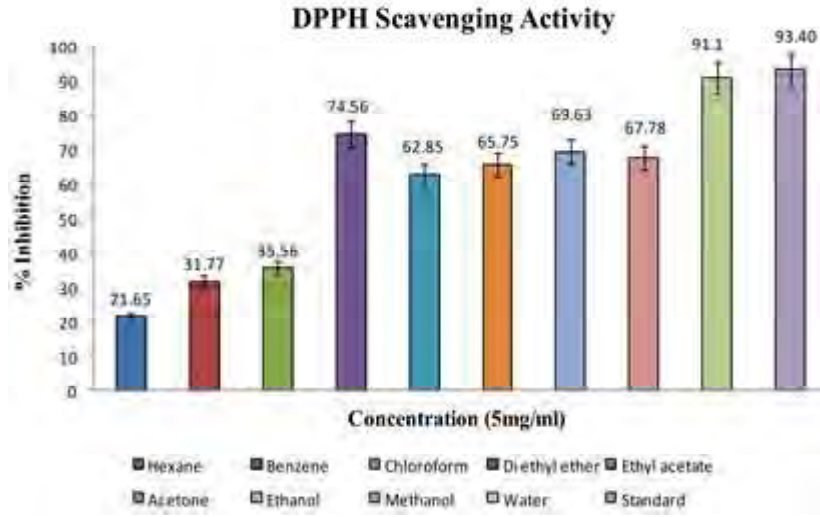


Fig. 2: Percent DPPH free radical scavenging activity of all the nine extracts and ascorbic acid used as standard at concentration 5mg/ml.

The total phenol content was expressed as GAE (mg/g). Methanol, acetone, ethanol and ethyl acetate gave the best results with the total phenol content value recorded as 100.60mg/g, 87.07mg/g, 58.73 mg/g and 51.47mg/g respectively with methanol giving the best result followed by water, diethylether, benzene, chloroform and hexane (Table 2). The lowest value recorded was that of hexane i.e. 4.30mg/g of total phenols. Green tea was found to be a richer source of phenolics than white tea and found 40% aqueous ethanol to be useful for extracting catechin (Rusak *et al*. 2008). Total phenol content estimation following Folin-Ciocalteu reagent method showed 50% acetone, 50% N,N-dimethylformamide (DMF), 50% ethanol and 50% methanol to be suitable for extracting total phenols (Turkmen *et al*. 2006).

Table 2. Determination of total phenolic content (TPC) from fresh leaves of TV26 extracted by different organic solvents.

Solvents	TPC(µg/ml)	GAE (mg/g)	Total phenolic content (mean±SD)
Hexane	2.2	4.3	4.3±0.2
Benzene	3.7	6.9	6.9±0.6
Chloroform	3.0	5.8	5.8±0.2
Diehylether	7.4	12.8	12.8±1.9
Ethylacetate	27.1	51.5	51.5±2.6
Acetone	43.7	87.1	87.1±0.3
Ethanol	31.1	58.7	58.7±3.5
Methanol	54.1	100.6	100.6±7.7
Water	37.5	36.8	36.8±38.2

$y= 0.0075x-0.0252$; $R^2=0.9864$

Acetone extracts showed higher ferric reducing power than other extracts as par with ascorbic acid taken as standard (Fig. 3). The antioxidant compounds acts 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 .The IC_{50} value calculated represents the exact concentration for 0.5 absorbance at 700 nm and the reducing power mostly increases with increasing concentration of antioxidant compounds (Ferreira *et al*. 2007; Manivasagan *et al*. 2015; Patil *et al*. 2016). Acetone extracts showed higher ferric reducing power with IC_{50} value 426.45±1.12 µg/ml compared to the standard (ascorbic acid) 270.35±0.66 µg/ml.

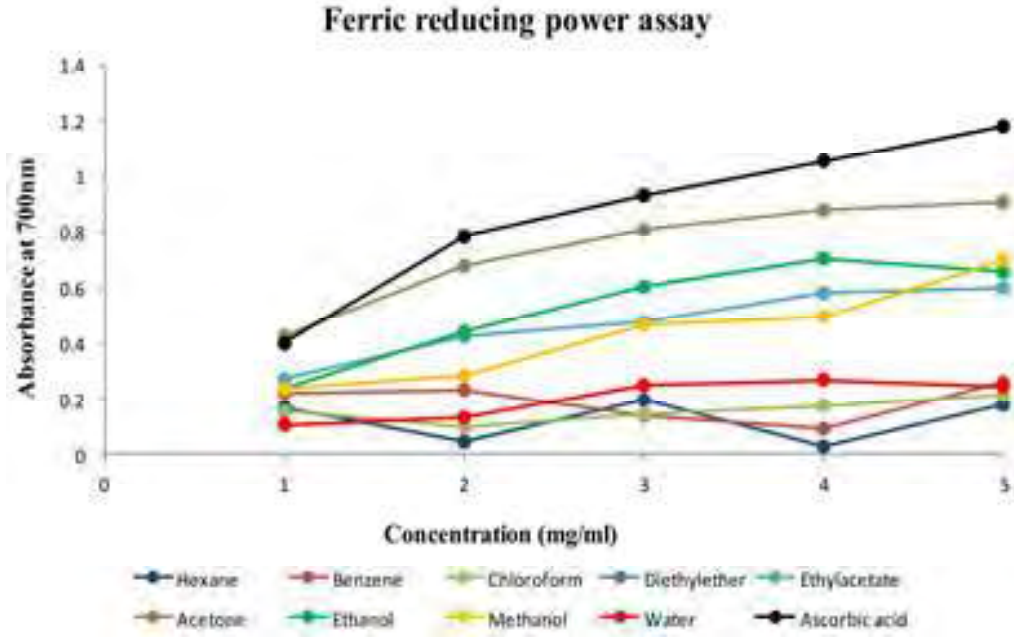


Fig. 3: Ferric reducing power of all the extracts and ascorbic acid (standard) at different concentration (mg/ml).

Conclusion

The qualitative and quantitative phytochemical screening as well as studying the antioxidant activity based on extracting solvents of different polarities was studied for fresh leaves of TV26. The ranges of organic solvent used with differing polarities included absolute hexane, benzene, chloroform diethylether, ethylacetate, acetone, ethanol, methanol and water. Solvents like acetone, methanol, ethanol and ethyl acetate was found to be the most potent solvent for various phytochemical extractions like flavonoid, tannin, steroid, diterpenes, terpenoids, coumarin, cardiac glycoside, saponin, protein and reducing sugar. The highest percent of radical scavenging was recorded in cold water extracts (5mg/ml) i.e. 91.10% and was at par with 93.40% percent scavenging activity of ascorbic acid taken as standard (5mg/ml). Methanol, acetone, ethanol and ethyl acetate gave the best results with the total phenol content value recorded as 100.60mg/g, 87.07mg/g, 58.73 mg/g and 51.47mg/g respectively. Acetone extracts showed higher ferric reducing power with IC_{50} value 426.45±1.12 µg/ml compared to the standard 270.35±0.66 µg/ml. The polar solvents and in some cases even the least polar solvents showed best result in extracting phytochemicals from fresh leaves of TV26 and we could therefore infer that

in addition to extraction of samples using solvent with different polarities, extraction time and procedure, the state of sample also matters in phytochemical extraction.

The collective information of this work could promote future research work in broader ways and to check whether the result varies or remain same among other varieties. Elaborate qualitative phytochemical profiling thus gave an idea about the potency of particular solvent in extracting specific compound of interest.

Acknowledgement

We acknowledge University Grants Commission (UGC) for providing necessary fund (UGC-RGNF).

References

Aiyegoro, O.A., and Okoh, A.I. 2009. Phytochemical screening and polyphenolic antioxidant activity of aqueous crude leaf extract of *Helichrysum pedunculatum*. International Journal of Molecular Sciences. 10:4990-5001.

Atoui, A.K., Mansouri, A., Boskou, G. and Kefalas, P. 2005. Tea and herbal infusions: their antioxidant activity and phenolic profile. Food Chemistry. 89:27-36.

Blois, M.S., 1958. Antioxidant determinations by the use of a stable free radical. Nature. 181:1199-1200.

Brain, K.R. and Turner, T.D., 1975. The practical evaluation of phytopharmaceuticals. Wright-Scientechnica Bristol.

Chavan, U.D., Shahidi, F. and Naczk, M. 2001. Extraction of condensed tannins from beach pea (*Lathyrus maritimus* L.) as affected by different solvents. Food Chemistry. 75:509-512.

Dufresne, C.J and Farnworth, E.R 2001. A review of latest research findings on the health promotion properties of tea. The Journal of Nutritional Biochemistry. 12:404-421.

Ferreira, I.C.F.R, Baptista, P., Vilas-Boas, M. and Barros, L. 2007. Free-radical scavenging capacity and reducing power of wild edible mushrooms from northeast Portugal: Individual cap and stipe activity. Food Chemistry. 100:1511-1516.

Filip, R. and Ferraro, G.E., 2003. Researching on new species of "Mate" : *Ilex brevicuspis*. European Journal of Nutrition. 42:50-54.

Folin, O. and Ciocalteu, V. 1927. On tyrosine and tryptophan determinations in proteins. Journal of Biological Chemistry. 73:627-650.

Geoffrey, K.K., John, K.M., Naomi, M. and Simon, K.M. 2014. Qualitative phytochemical screening of *Camellia sinensis* and *Psidium guajava* leave extracts from Kericho and Baringo countries. International Journal Of Advanced Biotechnology And Research (IJBR). Vol5, Issue3:506-512.

Goli, A.H., Barzegar, M. and Sahari, M.A. 2005. Antioxidant activity and total phenolic compounds of pistachio (*Pistachia vera*) hull extracts. Food Chemistry. 92:521-525.

Kaundun, S.S., Zhyvoloup, A. and Park, Y.G. 2000.Evaluation of the genetic diversity among elite tea (*Camellia sinensis* var. *sinensis*) accessions using RAPD markers. Euphytica. 115:7-16.

Kumar, A., Ilavarasan, R., Jayachandran, T., Decaraman, M., Aravindhan, P., Padmanabhan, N. and Krishnan, M.R.V. 2009. Phytochemicals investigation on a tropical plant, *Syzygium cumini* from Kattuppalayam, Erode district, Tamil Nadu, South India. Pakistan Journal of Nutrition. 8:83-85.

Liu, Q., Kong, B., Xiong, Y.L. and Xia, X. 2010. Antioxidant activity and functional properties of porcine plasma protein hydrolysate as influenced by the degree of hydrolysis. Food Chemistry. 118:403-410.

Manivasagan, P., Alam, M.S., Kang, K.H., Kwak, M. and Kim, S.K. 2015. Extracellular synthesis of gold bionanoparticles by *Nocardiopsis* sp. and evaluation of its antimicrobial, antioxidant and cytotoxic activities. Bioprocess and Biosystems Engineering. 38:1167-1177.

Manzocco, L., Anese, M. and Nicoli, M.C. 1998. Antioxidant properties of tea extracts as affected by processing. LWT-Food Science and Technology. 31:694-698.

MartiNez, M.A.D.P., Pelotto, J.P. and Basualdo, N. 1997. Distribution of flavonoid aglycones in *Ilex* species (Aquifoliaceae). Biochemical Systematics and Ecolog. 25:619-622.

Mukhtar, H. and Ahmad, N. 2000. Tea polyphenols: prevention of cancer and optimizing health. The American Journal of Clinical Nutrition. 71:1698-1702.

Ngbede, J., Yakubu, R.A. and Nyam, D.A. 2008. Phytochemical screening for active compounds in *Canarium schweinfurthii (*Atile) leaves from Jos North, Plateau State, Nigeria. Research Journal of Biological Sciences. 3:1076-1078.

Opie, S.C., Robertson, A. and Clifford, M.N. 1990. Black tea thearubigins-their HPLC separation and preparation during in-vitro oxidation. Journal of the Science of Food and Agriculture. 50:547-561.

Ozarkar, K.R. 2005. Studies on anti-inflammatory effects of two herbs *Cissus quadrangularis* Linn. and *Valeriana wallichi* DC using mouse model. Ph. D. Thesis, University of Mumbai, Mumbai.

Patil, M.P., Patil, K.T., Ngabire, D., Seo, Y.B. and Kim, G.D. 2016. Phytochemical, antioxidant and antibacterial Activity of Black Tea (*Camellia sinensis*). International Journal of Pharmacognosy and Phytochemical Research. 8:341-346.

Rusak, G., Komes, D., Likic, S., Horzic, D. and Kovac, M. 2008. Phenolic content and antioxidative capacity of green and white tea extracts depending on extraction conditions and the solvent used. Food Chemistry. 110:852-858.

Sharangi, A.B. 2009. Medicinal and therapeutic potentialities of tea (*Camellia sinensis* L.)-A review. Food Research International. 42:529-535.

Sun, T., Ho, C.T. 2005. Antioxidant activities of buckwheat extracts. Food Chemistry. 90:743-749.

Tiwari, P., Kumar, B., Kaur, M., Kaur, G. and Kaur, H. 2011. Phytochemical screening and extraction: A review. Internationale Pharmaceutica Sciencia. 1:98-106.

Turkmen, N., Sari, F., Velioglu, Y.S. 2006. Effects of extraction solvents on concentration and antioxidant activity of black and black mate tea polyphenols determined by ferrous tartrate and Folin-Ciocalteu methods. Food Chemistry. 99:835-841.

Ugochukwu, S.C. and Arukwe, U.I., and Onuoha, I. 2013. Preliminary phytochemical screening of different solvent extracts of stem bark and roots of *Dennetia tripetala* G. Baker. Asian Journal of Plant Science and Research. 3:10-13.

Wang, H., Helliwell, K. 2001. Determination of flavonols in green and black tea leaves and green tea infusions by high-performance liquid chromatography. Food Research International. 34:223-227.

Yao, L., Jiang, Y., Datta, N., Singanusong, R., Liu, X., Duan, J., Raymont, K., Lisle, A. and Xu, Y. 2004. HPLC analyses of flavanols and phenolic acids in the fresh young shoots of tea (*Camellia sinensis*) grown in Australia. Food Chemistry. 84:253-263.

Zuo, Y., Chen, H. and Deng, Y. 2002. Simultaneous determination of catechins, caffeine and gallic acids in green, oolong, black and pu-erh teas using HPLC with a photodiode array detector. Talanta. 57:307-316.

NBU Journal of Plant Sciences

Vol. 9, No. 1 March 2015 p. 18-32
ISSN No. 09746927

EFFICACY OF ANTHOCYANIN IN PRODUCTION OF REMEDIAL TEA

Reha Labar and Arnab Sen*

Department of Botany, University of North Bengal, Raja Rammohunpur, Siliguri 734013, West Bengal, India

ABSTRACT

Tea is consumed as a medicinal beverage from centuries, as the medicinal component includes polyphenols, caffeine, and amino acids. Apart from that it also contains flavonoids; compounds reported to have antioxidant properties with many beneficial effects. Anthocyanins belonging to the flavonoid group are naturally occurring compounds that imparts colour to fruits, vegetables and plants. Apart from that it has an array of health promoting benefits. This article has been reviewed to highlight the importance of anthocyanin as well as to motivate research in exploring tea varieties with abundant anthocyanin so that consumption of anthocyanin rich tea or beverage and also use of natural dye made using anthocyanin would replace harmful effects of chemical drugs and also improve country's economy by flourishing the tea industry with increased consumption.

Keywords: *Camellia sinensis*, anthocyanin, pharmacological activity, industrial use

INTRODUCTION

Cultivated tea belongs to genus Camellia, consisting of three species each with specific plant types viz. *Camellia sinensis* China type, *Camellia assamica* Assam type and *Camellia assamica lasiocalyx* Cambod type (Wight, 1962). Tea is an important agricultural and commercial crop consumed worldwide, mainly as a beverage made from processed tea leaves and it has also been used for medicinal purposes for several centuries. (Friedman et al., 2007; Wang et al., 2012)According to varying processing procedures, tea currently made in the world can be classified into six main types including black, green, white, yellow, oolong and reprocessed tea (P12). Of all the types, green tea is mostly favored as a medicinal tea as many of its medicinal properties like antioxidative, anti-mutagenic, anticarcinogenic, antihypersensitive, antibacterial, antiviral and also weight reducing property have been already reported (Saito et al., 2011). Besides green tea, purple tea is gaining much importance since, interest in anthocyanins is growing among researchers owing to their potential health benefits(Kong et al., 2003).

Anthocyanins are the another most important plant pigment besides chlorophyll visible to the human eye belonging to the widespread class of phenolic compounds collectively named flavonoids (Kong et al., 2003). They occur in different colour basically red, blue or purple depending upon their pH. Synthesized through the phenylpropanoid pathway, the water soluble vacuolar pigment has many important roles to play besides imparting colour and contributing to astringent sensation.

It is found in many plant species including red grapes (Rivero-Perez et al., 2008), berries (blueberry, strawberry, raspberry, blackcurrant, bil- berry, cranberry, elderberry) (Nicoue et al., 2007), eggplant (Azuma et al., 2008), purple fleshed sweet potatoes (Oki et al., 2002) and flowers like Hibiscus (Lo et al., 2007a). These pigments have been found to be the largest and most important group of water soluble pigments found in nature and they contribute to the attractive colours of fruits, vegetables and flowers imparting red, orange, purple, violet and blue colours (Feild et al., 2001). Anthocyanins in plants normally accumulate in the vacuoles of the epidermal and sub epidermal cells (Steyn et al., 2002). The colours of these pigments are pH-dependent (Mazza and Miniati, 1993). Interest in anthocyanins has recently increased owing to their potential health benefits (Kong, Chia, Goh, Chia, & Brouillard, 2003) and their use as an alternative source of synthetic colourants/dyes. (Jackman et al., 1987; Kerio et al., 2012).

The anthocyanin pigments that create the

*Corresponding Author:
Email: senarnab_nbu@hotmail.com

color (Tsai and Ou, 1996) are responsible for the wide range of coloring in many foods. Recently, the biological activities of anthocyanin, such as antioxidant activity, protection from atherosclerosis and anti-carcinogenic activity have been investigated, and shown to have some beneficial effects in the treatment of diseases.

For instance, there have been reports on the antioxidant activity in grape anthocyanin (Igarashi et al., 1989) and the biological effect of anthocyanin in low density lipoprotein and lecithin-liposome systems (Meyer et al., 1997). Anthocyanins were also found to have many times more activity than common antioxidants such as ascorbate (Wang et al., 1997). Overall, there is now increasing evidence that antioxidants in the human diet are of major benefit for health and well-being. (Tsai et al., 2002)

However, no tea has been demonstrated to contain an abundant amount of anthocyanins. Therefore, 'Sunrouge' was developed, a red leaf tea cultivar that is rich in anthocyanins by natural crossing in 2009. An anthocyanin-rich parental line, 'Cha Chuukanbohon Nou 6', which was derived from *C. taliensis x C. sinensis* in 2004, was previously developed. *C. taliensis* is closely related to *C. sinensis*. However, the anthocyanin content of 'Cha Chuukanbohon Nou 6'suddenly diminishes as the leaf matures. Therefore, anthocyanin-rich tea was developed the cultivar which was higher in anthocyanin content than 'Cha chuukanbohon Nou 6', and in which the anthocyanin content did not diminish after leaf maturation. 'Sunrouge' is an offspring of 'Cha Chuukanbohon Nou 6'. Saito *et al*. isolated six anthocyanins from 'Sunrouge' leaves.(Maeda-Yamamoto et al., 2012) Because of the sedentary nature of plants, they are prone to UVirradiation which can cause oxidative stress. Anthocyanins protect plants against such irradiation. Their biosynthesis has been demonstrated to be upregulated when the plant is exposed to UV-B irradiation (Merzlyak et al., 2008). Although anthocyanin is not found abundantly in Camellia sinensis, there are many more purple leaf coloured tea plants that are yet to be explored to make promising anthocyanin rich tea.

BIOCHEMISTRY OF ANTHOCYANIN

Anthocyanin belongs to the widespread class of phenolic compounds collectively named flavonoids. They are glycosides of poly-hydroxy and poly-methoxy derivatives of 2-phenylbenzopyrylium or flavylium salts (Kong et al., 2003). Chemically, anthocyanins are glycoside moieties of anthocyanidins derived from the flavylium (2-phenylbenzopyrilium) cation shown in Fig 1 (Kerio et al., 2012). The differences between individual anthocyanins relate to the number of hydroxyl groups, the nature and number of sugars attached to the molecule, the position of this attachment, and the nature and number of aliphatic or aromatic acids attached to sugars in the molecule (Kong et al., 2003).

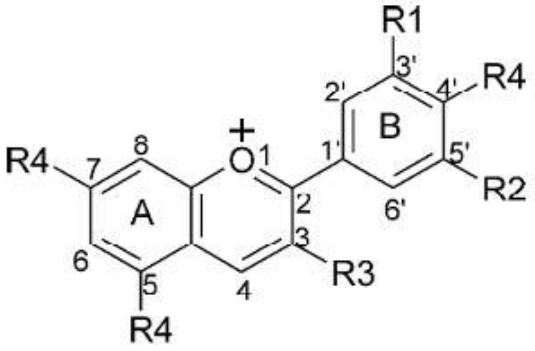


Fig 1: Basic structure of anthocyanidin pigment, the flavylium cation where R1 and R2 are H, OH, or OCH; R3 is a glycosyl or H; and R4 is OH or a glycosyl. (Kong et al., 2003)

There are several anthocyanidins described in nature but among these, six are widespread in fruits and vegetables namely; pelargonidin, cyanidin, peonidin, delphinidin, petunidin and malvidin as shown in Table 1 (Kerio et al., 2012). The various anthocyanidins differs in number and position of the hydroxyl and /or methyl ether groups attached on 3, 5, 6, 7, 3', 4' and/or 5' positions (Table 1). Despite the fact that 31 different monomeric anthocyanidins have been identified (including 3-deoxy-anthocyanidins, pyranoanthocyanidins and sphagnorubins), 90% of the naturally occurring anthocyanins are based on only six structures (30% on cyanidin, 22% on delphinidin,18% on pelargonidin and in summary 20% on peonidin, malvidin and petunidin). Those six anthocyanidins are usually known as common

Table 1: Common anthocyanidins occurring in the nature. *Excerpted from Ananga et al. (2013)*

ANTHOCYANIN COMMON IN HIGHER PLANTS									
Sl. no	Name	Abbreviation	Substitution Pattern						
			3	5	6	7	3´	4´	5´
1	Cyanidin	Cy	OH	OH	H	OH	OH	OH	H
2	Delphinidin	Dp	OH	OH	H	OH	OH	OH	OH
3	Malvidin	Mv	OH	OH	H	OH	OMe	OH	OMe
4	Pelargonidin	Pg	OH	OH	H	OH	H	OH	H
5	Peonidin	Pn	OH	OH	H	OH	OMe	OH	H
6	Petunidin	Pt	OH	OH	H	OH	OMe	OH	OH

anthocyanidins (Ananga et al., 2013).The glycosides of the three non-methylated anthocyanidins (Cy, Dp and Pg) are the most widespread in nature, being present in 80% of pigmented leaves, 69% of fruits and 50% of flowers. The distribution of the six most common anthocyanidins in the edible parts of plants is cyanidin (50%), pelargonidin (12%), peonidin (12%), delphinidin (12%), petunidin (7%), and malvidin (7%). The following four classes of anthocyanidin glycosides are common: 3-monosides, 3-biosides, 3,5-diglycosides and 3,7-diglycosides. 3-glycosides occur about two and half times more frequently than 3,5-diglycosides. So, the most widespread anthocyanin is cyanidin 3-glucoside. (Kong et al., 2003)

Because of their polar nature, anthocyanins are soluble in polar solvents, such as methanol (MEOH), ethanol and water. The initial step in their isolation therefore involves solvent extraction, which includes the use of acidified methanol or ethanol. The use of acid stabilizes anthocyanins in the flavylium cation form, which is red at low pH (Mc. Ghie and Walton, 2007).

The color of anthocyanidins differs with the number of hydroxyl groups, attached on their molecules (especially those substituted in ring B). With the increase of attached hydroxyl groups, the visible color of entire molecule shift from orange to violet (Andersen and Jordheim, 2006; Delgado-Vargas et al., 2000; Delgado-Vargas and Paredes-Lopez, 2002; Tanaka et al., 2008).

Glycosylation of anthocyanidins results to additional reddening of obtained anthocyanins, whereas the presence of aliphatic or aromatic acyl moieties causes no color change or slight blue shift and has significant effect on their stability and solubility (Tanaka et al., 2008). Changes in pH can also cause reversible structural transformations in anthocyanins molecules, which has a dramatic effect on their color Delgado-Vargas and Paredes-Lopez, 2002; Wrolstad, 2004). Most of the anthocyanins are *O*-glycosylated at 3 (except those based on 3-deoxyanthocyani-dins and sphagnorubinns), 5 or 7 positions and in some cases at 3´, 4´ and 5´ positions(Ananga et al., 2013). However, 8-*C*-glycosylanthocyanins have been found only in *Tricyrtis formosana* Baker (Saito et al., 2003; Tatsuzawa et al., 2004). Anthocyanins contain two, one or tree monosaccharide units in their molecules. The usual monosaccharide residues are glucose, galactose, arabinose, ramnose, xylose and glucuronic acid. However, anthocyanins containing disaccharides and trisaccharides were also found in nature but no tetrasaccharides have been discovered yet (Ananga et al., 2013).

BIOSYNTHESIS OF ANTHOCYANIN IN TEA

Anthocyanin molecules are produced via flavonoid pathway. Anthocyanin pigments are assembled like all other flavonoids from two

different streams of chemical raw materials in the cell where one stream involves the shikimate pathway to produce the amino acid phenylalanine and the other stream produces three molecules of malonyl-CoA, a C3 unit from a C2 unit (acetyl-CoA). The Shikimate pathway leads to the formation of chorismate, which is the precursor of the aromatic amino acids phenylalanine, tyrosine and tryptophan. Phenylalanine (and in some cases tyrosine, but not in the case of tea plants) is the primary precursor of catechins (Tounekti et al., 2013).The flavonoid pathway starts with phenylalanine, produced via shikimate pathway and trans- formed to 4 coumaroyl-CoA. The key enzyme, chalcone synthase (CHS) produce a naringe-nin chalcone by condensing one molecule of 4-coumaroyl-CoA and three malonyl-CoA molecules (derived from citrate produced by The Krebs cycle). In this case, the rings A and C are derived from the acetate pathway, whereas the ring B is derived from shikimate pathway (Ananga et al., 2013).Currently there are three isoforms of chalcone synthase (Park et al., 2004).The three genes act to synthesize naringenin chalcone, which is used in the formation of anthocyanins, proanthocyanidins, and other phenolic compounds. According to (Ageorges et al., 2006), the three different CHSs may act in three different pathways to produce different secondary metabolites. In the next step, chalcone isomerase (CHI) converts stereospecifically the naringenin chalcone to its isomer naringenin. Ring B of the naringenin undergoes further hydroxylation by the enzymes flavonoid 3´-hydroxylase (F3´H), flavonoid 3´5´-hydroxylase (F3´5´H) or flavanon 3ß-hydroxylase (F3H) (He et al., 2010). Then, the obtained dihydroflavonols are reduced by the enzyme dihydroflavonol 4-reductase (DFR) to the corresponding leucoanthocyanidins. After this reduction, anthocyanidin synthase (ANS) oxidize leucoanthocyanidins to their corresponding anthocyanidins. Anthocyanidins are inherently unstable under physiological conditions and are immediately glycosylated to anthocyanins by UDP-glucose: Anthocianidin: Flavonoid glucosyltransferase (UFGT) (He et al., 2010). Anthocyanins, containing methylated anthocyani-dins (peonidin 4, petunidin 5 and malvidin 6) as aglycone can be obtained by methylation of hydroxyl groups on the ring B of the cyanidin-3-*O*-glucoside 7, delphinidin-3-*O*-glucoside and petunidin-3-*O*-glucoside by the enzyme *O*-methyltransferase (OMT). Future acylation of produced anthocyanins is possible by the action of different anthocyanin acyltransferases (ACT).

EXTRACTION OF ANTHOCYANIN

Anthocyanins are soluble in polar solvents, and they are normally extracted from plant materials by using methanol that contains small amounts of hydrochloric acid or formic acid. The acid lowers the solution's pH value and prevents the degradation of the non-acylated anthocyanin pigments. However, as hydrochloric acid or formic acid is concentrated during the evaporation of the methanol-hydrochloric acid or methanol-formic acid solvent, pigment degradation occurs (e.g. in the extract of Azalea cv. Alice Erauw, the cyanidin-3monosides are converted into unstable aglycone). Small amounts of acid may also cause partial or total hydrolysis of the acyl moieties of acylated anthocyanins that are present in some plants. One report compared various techniques for the extraction of anthocyanins from red grapes and demonstrated that solvents containing up to 0.12 mol/l hydrochloric acid can cause partial hydrolysis of acylated anthocyanins (Revilla et al., 1998). Acetone has also been used to extract anthocya nins from several plant sources (GarciaViguera et al., 1998; Giusti et al., 1994). In comparison to acidified methanol, this technique allows an efficient and more reproducible extraction, avoids problems with pectins, and permits a much lower temperature for sample concentration (GarciaViguera et al., 1998). Solid-phase extraction (SPE) on C_{18} (SPE) cartridges or Sephadex is commonly used for the initial purification of the crude anthocyanin extracts. The anthocyanins are bound strongly to these adsorbents through their unsubstituted hydroxyl groups and are separated from unrelated compounds by using a series of solvents of increasing polarity. Sunrouge' tea leaves i.e., anthocyanin rich tea leaves (87.2 g) were extracted with 15% acetic acid (600 mL 3), and additionally extracted with 15% acetic acid-

containing 50% EtOH (600 mL 4) (Saito et al., 2011). The anthocyanins were extracted successfully from tea products processed from a number of newly bred purple leaf coloured Kenyan tea cultivars (*Camellia sinensis*) using acidified methanol/HCl (99:1 v/v). (Kerio et al., 2012).

CHARACTERIZATION OF ANTHOCYANIN

The characterization of a mixture of anthocyanins usually involves the separation and collection of each compound, and subsequent analysis by nuclear magnetic resonance (NMR) and fast atom bombardment mass spectroscopy (FAB-MS). For the separation and structural analysis, the use of liquid chromatography–mass spectrometry (LC–MS) technique, which combines the separation of LC with the selectivity and sensitivity of the MS detector, permits the identification of individual compounds in a mixture of compounds. Recently liquid chromatography-electron impact ionization mass spectrometry (LC-EI-MS) was also used to identify the anthocyanins of Catharanthus roseus extracts (Piovan et al., 1998). LC–MS with an atmospheric pressure-ionization ion-spray interface to analyze the anthocyanins contained in the grape skins (Vitis vinifera L.). Nineteen derivatives of cyanidin, delphinidin, petunidin, malvidin and peonidin were identified by this ionization technique. The individual mass spectra showed peaks for the molecular ions, together with a fragment corresponding to aglycone; when acylation was present, an additional fragment was detected at mass/charge values corresponding to the loss of acyl moiety from the molecular ion (Baldi et al., 1995). Many new acylated anthocyanins have been found with the help FAB-MS (Saito et al., 1983).

Atmospheric-pressure ionization (API) techniques have several advantages over other MS detection methods. In API-MS the ion source is located outside the MS; the ions are formed at atmospheric pressure, and then sampled into the mass spectrometer. These are soft ionization techniques (only the molecular ion is formed), although the application of a potential at the entrance of the mass spectrometer (fragment voltage) creates suitable conditions for CID, and the production of fragment ions. Two API interfaces are available commercially, namely, the atmospheric pressure chemical ionization interface (APCI) and the ESI interface. LC–MS system equipped with an ESI interface was used to analyze anhocyanins present in extracts of grape skins and red wine (Revilla et al., 1999) Another technique recently used for anthocyanin analysis is capillary electrophoresis (CE) which has excellent mass sensitivity, high resolution, low sample consumption and minimal generation of solvent waste. The separation of a mixture of standards, as well as strawberry and elderberry anthocyanins, by capillary zone electrophoresis (CZE) has already been reported by Bridle et al. (1997). Matrix-assisted laser desorption/ionization mass spectrometry (MALDI–MS) to perform both qualitative and quantitative analyses of anthocyanins in wine and fruit juice, and matrix-assisted laser desorption/ionization time-offlight mass spectrometry (MALDI–TOF–MS) to analyze the content of anthocyanins in various foods (Wang and Sporns, 1999; Wang and Lin, 2000). Purification and isolation of 6 anthocyanins from 'Sunrouge' by chromatography, and identification of them by LC/MS/MS and NMR analysis has already been reported.(Saito et al., 2011).

NMR, an important technique is useful for structural elucidation of anthocyanin. Some examples of its use are given where it has been reported that the separation of anthocyanins from red radish (Raphanus sativus) and their structural elucidation by one- and two-dimensional NMR.

Four anthocyanins were obtained: pelargonidin 3-O- [2-O-(b-glucopyranosyl) -6-O-(trans-p-coumaroyl) -b-glucopyranoside] 5O-(6-O-malonyl-b-glucopyranoside); pelargonidin 3-O[2-O -(b-glucopyranosyl)-6-O- (trans-feruloyl)-b-glucopyranoside] 5-O-(6-O- malonyl-b-glucopyranoside); pelargonidin 3-O-[2-O-(b-glucopyranosyl)-6-O- (trans-pcoumaroyl)-b-d-glucopyranoside] 5-O-(b-glucopyranoside) and pelargonidin 3-O-[2-O-(b-glucopyranosyl)-6O-(trans-feruloyl)-b- gluco-pyranoside] 5-O-(b-glucopyr-anoside). They also investigated the three-dimensional conformation of the molecule by using NOESY techniques, which showed

proximity between the hydrogen from the cinnamic acid acylating group and the C-4 of the pelargonidin (Giusti et al., 1998). Similarly, anthocyanin trisaccharides of *Vaccinium padifolium* were identified by using NMR and other techniques (Cabrita et al., 2000). Anthocyanins from Kenyan teas were purified by C_{18} solid phase extraction (SPE) catridges and characterised by HPLC-UV–Visible. They were identified according to their HPLC retention times, elution order and comparison with authentic standards that were available. Total monomeric anthocyanins were determined by the pH-differential method.(Kerio et al., 2012). Anthocyanins content in purple tea was measured by single-pH and the content was much more than that measured by pH-Differential spectrophotometry. The two methods showed good linear relation(R2=0.9927).The effects of interfering substances on quantitative analysis of anthocyanins could be eliminated by using pH-Differential spectrophotometry. The results provided a reliable basis for the measurement of Anthocyanins (Chen and Lu, 2011).

ANTHOCYANIN RICH TEA

Sunrouge is the red tea cultivar made by naturally crossing *Camellia taliensis* and *Camellia sinensis* for which an application for registration was made in 2009. An anthocyanin-rich parental line, 'Cha Chuukanbohon Nou 6', was previously developed by crossing *C. taliensis x C. sinensis* in 2004 where *C. taliensis* is closely related to *C. sinensis*. However, the anthocyanin content of 'Cha Chuukanbohon Nou 6 suddenly diminished after maturation of leaves. Therefore Sunrouge was introduced as an offspring of 'Cha Chuukanbohon Nou 6' whose anthocyanin content did not diminish after leaf maturation (Maeda-Yamamoto et al., 2012).The total anthocyanin content of four tea cultivars: 'Sunrouge', 'Cha Chuukanbohon Nou 6', 'Benibana-cha' (C. sinensis), the anthocyanin-rich tea cultivar, and 'Yabukita' (C. sinensis), the common tea cultivar in Japan was quantified and it has been already reported that the anthocyanin content of 'Sunrouge' was the highest among 4 tea cultivars, and was 8.4 times higher than that of 'Yabukita'. Purification and isolation of 6 anthocyanins from 'Sunrouge' has already been done using chromatography, and also has been identified using LC/MS/MS and NMR analysis, where the four anthocyanins were identified as delphinidin-3-O- ßD-(6-(E)-pcoumaroyl) galactopyranoside(2),delphinidin-3-O-ßD-(6-(E)-p-coumaroyl) glucopyranoside (3), cyanidin-3-O-ß-D-(6-(E)-pcoumaroyl) galacto-pyranoside (4), and cyanidin-3-O-ßD-(6-(E)-p-coumaroyl)glucopyranoside(5), and the other two were estimated respectively as delphinidin-(Z)-p-coumaroylgalactopyranoside (1), petunidin-(E)-p-coumaroylgalactopyranoside (6). Compound 3 was found in tea for the first time. In general, anthocyanins has been reported to have various bioactivities, including relieving eyestrain and antioxidative effects, so it is expected that drinking 'Sunrouge' tea would bring in the above bioactivities.(Saito et al., 2011).

Another anthocyanin rich tea was extracted from tea products processed from a number of newly bred purple leaf coloured Kenyan tea cultivars (Camellia sinensis) using acidified methanol/HCl (99:1 v/v). Extracted anthocyanins were purified by C_{18} solid phase extraction (SPE) catridges and characterised by HPLC-UV–Visible. Of the six most common natural anthocyanidins, five were identified in the purified extracts from purple leaf coloured tea, in both aerated (black) and unaerated (green) teas namely; delphinidin, cyanidin, pelargonidin, peonidin and malvidin. The most predominant anthocyanidin was malvidin in both tea products. In addition, two anthocyanins namely, cyanidin-3-O-galactoside and cyanidin-3-O-glucoside were also identified. (Kerio et al., 2012).

ANTHOCYANIN AND ITS UTILITY IN PLANTS

Apart from imparting colour to the plants, anthocyanin play a definite role in the attraction of animals for pollination and seed dispersal, and hence they are of considerable value in the co-evolution of these plant-animal interactions. Anthocyanins and 3-deoxyanthocyanidins however have roles in flowering plants other than as attractants. They can act as antioxidants, phytoalexins or as antibacterial

agents. Anthocyanins may be important factors along with other flavonoids in the resistance of plants to insect attack(Harborne, 1988). For example, cyanidin 3-glucoside was shown to protect cotton leaves against the tobacco budworm (Hedin and Hedin, 1983).. In photosynthetic tissues (such as leaves and sometimes stems), anthocyanins have been shown to act as a "sunscreen", protecting cells from high-light damage by absorbing blue-green and ultraviolet light, thereby protecting the tissues from photo-inhibition, or high-light stress. This has been shown to occur in red juvenile leaves, autumn leaves, and broad-leaf evergreen leaves that turn red during the winter. The red coloration of leaves has been proposed to possibly camouflage leaves from herbivores blind to red wavelengths, or signal unpalatability, since anthocyanin synthesis often coincides with synthesis of unpalatable phenolic compounds (Sullivan, 1998). Some roles of anthocyanin have been tabulated in Table 2.

PHARMACOLOGICAL ACTIVITY

Anthocyanins possess known pharmacological properties and are used by humans for therapeutic purposes. Following the recognition that pigment extracts are more effective than O-(b-hydroxyethyl) rutin in decreasing capillary permeability and fragility and in their anti-inflammatory and anti-oedema activitiest is possible that anthocyanins may replace rutin and its derivatives in the treatment of illnesses involving tissue inflammation or capillary fragility. The crude anthocyanin extracts of *Vaccinium myrtillus* have been given orally and by intravenal or intramuscular injection to reduce capillary permeability and fragility. (Kong et al., 2003)

Anthocyanins were not found effective in suppressing tumor growth (Ghiselli et al., 1998). However, an antioxidant activity study of anthocyanin fractions from Italian red wine showed that the anthocyanin fraction was the most effective both in scavenging reactive oxygen species and in inhibiting lipoprotein oxidation and platelet aggregation (Ghiselli et al., 1998). This result suggests that anthocyanins could be the key component in red wine that protects against cardiovascular disease. Another report on the anti-tumor activity of anthocyanins was reported where they found that the anthocyanin fraction from red wine suppressed the growth of HCT-15 cells, which are derived from human colon cancer or

Table 2: The role of anthocyanins and 3-deoxyanthocyanidins in plants. *Excerpted from Kong et al. (2003)*

Plant	Compound	Origin	Function
	Angiosperms		
Senecio cruentus	Cinerarin	Petals	Pollination
Sorghum	Apigeninidin	Leaf sheath	Phytoalexin anti-microbial antioxidants
	Gymnosperms		
Abies concolor	Petunidin-3- glucoside Cyanidin-3- glucoside	Cone	-
Pinus contorta	Anthocyanin	Leaves	Cold tolerance
Pinus banksiana	-	Seedlings	Photoinhibition tolerance
	Ferns		
Davallia divaricata	Pelargonidin-3-p-coumaryl-glc-5-glc (monardein)	Young leaves	-
Ferns species	Apigenidin	Leaves	-
	Mosses		
Bryum, Splachunm	Luteolinidin-5-glc	Leaves	-
	Liverwort		
Cephaloziella exilifolia	Anthocyanin like	Thallus	-

AGS cells from human gastric cancer. The suppression rate by the anthocyanin fraction was significantly higher than that of the other fractions (Kamei et al., 1998). The ability of anthocyanin obtained from the petals of *H. rosa-sinensis* was examined which prevented carbon tetrachloride-induced acute liver damage in rats. The results showed that those rats treated with anthocyanin and carbon tetrachloride had significantly less hepatotoxicity ($P<0.05$) than those given carbon tetrachloride alone. This was assessed by measuring the levels of serum aspartate and alanine aminotransferase activities 18 hours after carbon tetrachloride was given. This result suggested that *H. rosa-sinensis* anthocyanin may be protective against carbon tetrachloride-induced liver injury (Obi et al., 1998).

The antimutagenicity of water extracts prepared from the storage roots of four varieties of sweet potato with different flesh colors, using Salmonella typhimurium TA 98. They found that two anthocyanin pigments purified from the purple colored sweet potato 3-(6,6´-caffeylferulylsophoroside) 5-glucoside of cyanidin (YGM-3) and peonidin (YGM6), effectively inhibited the reverse mutation induced by heterocyclic amines-mutagen, Trp-P-1, Trp-P-2, and IQ in the presence of rat liver microsomal activation systems (Yoshimoto et al., 1999).

It has been reported that the the administration of anthocyanin dyes from Aronia melanocarpa to rats before the intraperitoneal injections of PlateletActivating Factor (PAF) and ceruleine had a beneficial effect on the development of acute experimental pancreatitis in rats (Jankowski et al., 2000). It was revealed that this was due to the reduction of pancreatic swelling and a decrease in lipid peroxidation and adenosine deaminase activity.

They also examined the effect of anthocyanins from Cabernet red wine on the course and intensity of symptoms of experimental diabetes in rats (Jankowski et al., 1999). The results showed that a simultaneous daily administration of anthocyanins obtained from Cabernet red wine and streptozotocin substantially decreased sugar concentrations in the urine and blood serum. These anthocyanins also inhibited the loss of body mass caused by the injection of streptozotocin. Simultaneously, the anthocyanin pigment prevented the generation of free oxygen radicals, and decreased the peroxidation of lipids. The influence of anthocyanins was determined from chokeberries on the generation of autoantibodies to oxidize low density lipoproteins (oLAB) in pregnancies complicated by intrauterine growth retardation (IUGR). An experiment was conducted with a study group of 105 pregnant women (on the turn of trimester two according to LMP) with IUGR (sonographic examination results below the 5th percentile for real gestational age) who were randomly divided into 2 groups. Fifty women were administered anthocyanins and 55 women were given a placebo. There was a control group of 60 healthy pregnant women. They then examined the level of oxidative stress measured by the serum concentration of autoantibodies required to oxidize low density lipoproteins (oLAB). In the anthocyanin group, the oLAB titres decreased from 1104_41 mU/ml before treatment to 752_36 mU/ml in the first month and 726_35 mU/ml in the second month, at $P<0.01$. In the placebo group, the oLAB titres showed a slightly increasing trend: 1089_37 mU/ml before treatment, 1092_42 mU/ml in the first month and 1115_43 mU/ml in the second month, at $P>0.05$. The oLAB titres in the control group were 601_49 mU/ml before treatment, 606_45 mU/ml in the first month, and 614_43 mU/ml in the second month, at $P>0.05$. The results indicated that natural antioxidants (anthocyanins) can be useful in controlling oxidative stress during pregnancies complicated by IUGR (Pawlowicz et al., 2000).

Hibiscus anthocyanins (HAs), a group of natural pigments occurring in the dried flowers of *Hibiscus sabdariffa L.*, are used in soft drinks and herbal medicines. Their antioxidant bioactivity has been studied and it appears that HAs can significantly decrease the leakage of lactate dehydrogenase and the formation of malondialdehyde induced by a treatment of tert-butyl hydroperoxide (t-BHP). The in vivo investigation showed that the oral pretreatment of HAs before a single dose of t-BHP significantly lowered the serum levels of hepatic enzyme markers (alanine and aspartate aminotransferase) and reduced oxidative liver

damage. The histopathological evaluation of the liver revealed that hibiscus pigments reduce the incidence of liver lesions including inflammation, leucocyte infiltration, and necrosis induced by t-BHP in rats (Wang et al., 2000).

Their pharmaceutical value has been additionally increased due to their high bioavailability. However, the administration and metabolism of Anthocyanins *in vivo* have been investigated in details mostly in rats, whereas the detailed studies on humans still are scantly presented in scientific literature (He and Giusti, 2010; Yue et al., 2011).

The colorful anthocyanins are the most recognized, visible members of the bioflavonoid phytochemicals. The free-radical scavenging and antioxidant capacities of anthocyanin pigments are the most highly publicized of the modus operandi used by these pigments to intervene with human therapeutic targets, but, in fact, research clearly suggests that other mechanisms of action are also responsible for observed health benefits (Lila, 2004). Anthocyanin isolates and anthocyanin-rich mixtures of bioflavonoids may provide protection from DNA cleavage, estrogenic activity (altering development of hormone-dependent disease symptoms), enzyme inhibition, boosting production of cytokines (thus regulating immune responses), anti-inflammatory activity, lipid peroxidation, decreasing capillary permeability and fragility, and membrane strengthening (Lila, 2004).

The roles of anthocyanin pigments as medicinal agents have been well-accepted dogma in folk medicine throughout the world, and, in fact, these pigments are linked to an amazingly broad-based range of health benefits. For example, anthocyanins from *Hibiscus sp* have historically been used in remedies for liver disfunction and hypertension; and bilberry (*Vaccinium*) anthocyanins have an anecdotal history of use for vision disorders, microbial infections, diarrhea, and diverse other health disorders (Rice-Evans and Packer, 2003; Smith et al., 2000; Wang et al., 2000).

But while the use of anthocyanins for therapeutic purposes has long been supported by both anecdotal and epidemiological evidence, it is only in recent years that some of the specific, measurable pharmacological properties of isolated anthocyanin pigments have been conclusively verified by rigorously controlled in vitro, in vivo, or clinical research trials (Tsuda et al., 2003). For example, visual acuity can be markedly improved through administration of anthocyanin pigments to animal and human subjects, and the role of these pigments in enhancing night vision or overall vision has been particularly well documented (Matsumoto et al., 2001). Oral intake of anthocyanosides from black currants resulted in significantly improved night vision adaptation in human subjects (Nakaishi et al., 2000) and similar benefits were gained after administration of anthocyanins from bilberries (Muth et al., 2000). Three anthocyanins from black currant stimulated regeneration of rhodopsin (a G-protein-coupled receptor localized in the retina of the eye), and formation of a regeneration intermediate was accelerated by cyanidin 3-rutinoside (Matsumoto et al., 2003). These studies strongly suggest that enhancement of rhodopsin regeneration is at least one mechanism by which anthocyanins enhance visual acuity. In both in vitro and in vivo research trials, anthocyanins have demonstrated marked ability to reduce cancer cell proliferation and to inhibit tumor formation (Lila, 2004). Anthocyanins inhibit tumorigenesis by blocking activation of a mitogen-activated protein kinase pathway. This report provided the first indication of a molecular basis for why anthocyanins demonstrate anticarcinogenic properties.

The role of anthocyanins in cardiovascular disease protection is strongly linked to oxidative stress protection. Since endothelial dysfunction is involved in initiation and development of vascular disease, four anthocyanins isolated from elderberries were incorporated into the plasma. Crude anthocyanin extracts from bilberry have been administered both orally and via injection to reduce capillary permeability (Kong et al., 2003). Protection from heart attacks through administration of grape juice or wine was strongly tied to the ability of the anthocyaninrich products to reduce inflammation and enhance capillary strength and permeability, and to inhibit platelet formation and enhance nitric oxide (NO) release (Folts, 1998).

Their important function in cognitive decline and neural dysfunction has been investigated and found that fruit extracts (from blueberry)

including anthocyanins were effective in reversing age-related deficits in several neural and behavioral parameters, e.g. oxotremorine enhancement of a K1 evoked release of dopamine from striatal slices, carbachol-stimulated GTPase activity, striatal Ca buffering in striatal synaptosomes, motor behavioral performance on the rod walking and accelerod tasks, and Morris water maze performance and thus proved to improve neural and behavioral parameters (memory and motor functions. (Joseph et al., 1999).

It has already been reported that the Anthocyanins extracted from purple corn, when provided to mice in tandem with a high-fat diet, effectively inhibited both body weight and adipose tissue increases. Typical symptoms of hyperglycemia, hyperinsulinemia, and hyperleptinemia provoked by a high-fat diet did not occur when mice also ingested isolated anthocyanins. The experiments suggest that anthocyanins, as a functional food component, can aid in the prevention of obesity and diabetes(Tsuda et al., 2003).

MEDICINAL BENEFITS OF ANTHOCYANIN RICH FOODS

Consumption of anthocyanin adds as a beneficiary source of neutraceuticals and therapeutics and it has various application in the pharmaceutical industry. Recurrent consumption of anthocyanins could provide various health benefits including reduced risk of coronary heart diseases, anti-carcinogenic activity, antioxidant activity, reduced risk of stroke, anti-inflammatory effects etc. (Davies, 2009; Lila, 2004; Stintzing and Carle, 2004; Wrolstad, 2004).

Biological activity of anthocyanin has already been discussed in the above section. Anthocyanin in diet inhibits body weight and adipose tissue increase that could prevent symptoms of hyperglycemia, hyperinsulinemia, and hyperleptinemia provoked by a high-fat diet and thus can aid in the prevention of diabetes and obesity (Tsuda et al., 2003). It improves and cures vision disorders, microbial infections, diarrhea, and diverse other health disorders (Rice-Evans and Packer, 2003; Smith et al., 2000; Wang et al., 2000). It has the capacity to modulate cognitive and motor function, to enhance memory, and to have a role in preventing age-related declines in neural function (Lila 2004). It also reduces inflammation and capillary fragility (Kong et al., 2003), scavenges reactive oxygen species and in inhibiting lipoprotein oxidation and platelet aggregation (Ghiselli et al., 1998), provides protection from DNA cleavage, estrogenic activity (altering development of hormone-dependent disease symptoms), enzyme inhibition, boosting production of cytokines (thus regulating immune responses).

Above all anthocyanin has many other health beneficial properties, which include antioxidant (Bae and Suh, 2007), anticarcinogenic (Lee et al., 2009), anti-angiogenic (Bagchi et al., 2004), antimicrobial (Viskelis et al., 2009) antiapoptotic (Elisia and Kitts, 2008) and pro-apoptotic (Lo et al., 2007b) properties. Despite their health enhancing properties, no work had been carried out to determine the presence of these coloured polyphenols in newly bred purple leaf coloured tea clone (Kerio et al., 2012).

UTILITY IN VARIOUS INDUSTRIES

The world market of natural food colorants expands with the annual growth rate of 4-6% (Cormier et al., 1996). In USA 4 of the 26 colorants approved by the food administration, that are exempt from certification, are based on anthocyanin pigments (Wrolstad, 2004). In European Union, all anthocyanin-containing colorants are classified as natural colorants under the classification E163 (Socaciu, 2007).

Currently most of the worldwide anthocyanins supply comes from processing of grape pomace, which is a waste product from wine making. But in European Union other plant sources such as red cabbage, elderberry, black currant, purple carrot, sweet potato, and red radish are also allowed (Mortensen, 2006). Anthocyanins, produced by grape cell suspensions can be a promising alternative supply of natural colorants. It has already been demonstrated that the produced pigments by the grape cell suspensions undergo significant structural modifications. Grape cell suspensions accumulates higher levels of metabolically more

evolved structures (methylated and acylated anthocyanins). Acylated anthocyanins are suitable for application in food products, mainly because of the improved color stability compared to non-acylated structures (B[1]kowska-Barczak, 2005). Moreover, the grape cell suspensions can also produce elevated levels of beneficial phenolic compounds such as flavonoids, stilbenes, phenolics, etc., which are capable of increasing the added value of the final additive. The overall metabolite profile of grape cells in combination with the lack of microbial and toxic contaminations will give the potential for development of new types of food additives if the entire cell suspension biomass are utilized.

The commercial interest of cosmetic companies to apply plant additives, derived by biotechno-logical cultivation of plant cells to their products has increased remarkably in the last few years (Schurch et al., 2008). The addition of plant cell derived extracts in cosmetic products has been considered as a powerful approach used to increase their health benefits. Several plant extracts have been added to various cosmetic products as moisturizers, antioxidants, whitening agents, colorants, sunscreens, preservatives. With the advancement of plant cell biotechnology, more and more cosmetic companies have been attracted for application of additives, based on plant cell suspensions. Recently the application of so-called plant "steam" cells attracts industry's attention (Schurch et al., 2008). In the last few years, the French company "Sederma" launched the product "Re- sistem™" based on application of *in vitro* cultivated plant cells (www.sederma.fr). The other company, "Mibelle Biochemistry", situated in Switzerland, developed a "PhytoCellTec" prod- uct, based on grape cell suspension of *V. vinifera* L. *cv.* Gamay Fréaux, which was processed by high-pressure homogenizer to produce liposomes for application in cream products (www.mi-bellebiochemistry.com). According to the company, the grape cell derived liposomes contained higher amounts of anthocyanins and when applied on skins serve as strong UV protectors and fight photoaging. The presented examples clearly demonstrate the commercial interest to application of grape cell suspension derived products. However, it is a matter of time for the scientists to develop the biotechnological approach of producing anthocyanins by grape cell suspensions from the frame of experimental scale to commercially applicable products.

Anthocyanins have also been employed to produce juices and red wine whose natural colour as well as high antioxidant property adds to the quality of product.

Anthocyanins can also be used as pH indicators because their color changes with pH; they are pink in acidic solutions (pH < 7), purple in neutral solutions (pH ~ 7), greenish-yellow in alkaline solutions (pH > 7), and colourless in very alkaline solutions, where the pigment is completely reduced (Michaelis et al., 1936) and thus, it is employed in many chemical or pharmaceutical industry as well as in the fied of research.

Nowadays anthocyanins are being used widely in organic solar cells because of their ability to convert light energy into electrical energy (Cherepy et al., 1997). he many benefits to using dye-sensitized solar cells instead of traditional pn junction silicon cells include lower purity requirements and abundance of component materials, such as titania, as well as the fact they can be produced on flexible substrates, making them amenable to roll-to-roll printing processes (Gratzel, 2003).

CONCLUSION AND FUTURE PROSPECTS

Anthocyanins represent a class of important antioxidants, as they are so common in human foods. In recent years, many papers have been published on the in vitro antioxidant activity of anthocyanins and their other functions. However, there are still fewer studies on anthocyanin compared to the studies of other flavonoids. On the other hand tea is a pleasant, popular, socially accepted, economical and safe drink that is initially taken as medicine and later as beverage and now, it has proven well as future potential of becoming an important industrial and pharmaceutical raw material. As green tea, the purple coloured anthocyanin rich tea may also be a popular health drink since anthocyanins has many medicinal properties and is particularly known to be beneficial against cardiovascular diseases, for providing anticancer

benefits, improving vision, cholesterol and blood sugar metabolism as discussed above in the article. Most importantly it sports much lower caffeine content than black or green tea which is beneficial in beverage. Anthocyanin content in tea in addition to other polyphenols and other medicinal compounds would add a splendid color as well as an enigmatic healing property.

Based on these facts, this review is directed to highlight the importance of anthocyanins in order to improve further research in this field, discovering tea cultivars or wild tea plant, rich in anthocyanin.

REFERENCES

Ageorges, A., Fernandez, L., Vialet, S., Merdinoglu, D., Terrier, N. and Romieu, C. (2006). Four specific isogenes of the anthocyanin metabolic pathway are systematically co-expressed with the red colour of grape berries. *Pl. Sci.* 170: 372-383.

Ananga, A., Phills, B., Ochieng, J., Georgiev, V. and Tsolova, V. (2013). Production of anthocyanins in grape cell cultures: A potential source of raw material for pharmaceutical, food, and cosmetic industries. INTECH Open Access Publisher.

Andersen, Q.M. and Jordheim, M. (2006). Anthocyanins. eLS, 471-553

Azuma, K., Ohyama, A., Ippoushi, K., Ichiyanagi, T., Takeuchi, A., Saito, T. and Fukuoka, H. (2008). Structures and antioxidant activity of anthocyanins in many accessions of eggplant and its related species. *J. Agric food chemi.* 56: 10154-10159.

B¹kowska-Barczak, A. (2005). Acylated anthocyanins as stable, natural food colorantsâea review. *Polish J. Food Nutri. Sci.* 14(2): 107-116.

Bae, S.-H. and Suh, H.-J., (2007). Antioxidant activities of five different mulberry cultivars in Korea. *LWT- Food Sci. Technol.,* 40: 955-962.

Bagchi, D., Sen, C. K., Bagchi, M. and Atalay, M., (2004). Anti-angiogenic, antioxidant, and anti-carcinogenic properties of a novel anthocyanin-rich berry extract formula. *Biochem*. (Moscow) 69: 75-80.

Baldi, A., Romani, A., Mulinacci, N., Vincieri, F.F. and Casetta, B. (1995). HPLC/MS Application to Anthocyanins of *Vitis vinifera* L. *J. Agric. and Food Chem.* 43: 2104-2109.

Cabrita, L., Frøystein, N.A., Andersen, Ø.M., 2000. Anthocyanin trisaccharides in blue berries of *Vaccinium padifolium*. Food Chem 69: 33-36.

Chen, Q. and Lu, R.Q. (2011). Method for quantitative determination of anthocyanins of purple tea leaves [J]. J. Beijing Technol. Business University (Natural Science Edition) 2, 010.

Cherepy, N. J., Smestad, G. P., Gratzel, M. and Zhang, J. Z., (1997). Ultrafast electron injection: implications for a photoelectrochemical cell utilizing an anthocyanin dye-sensitized TiO2 nanocrystalline electrode. *J. Phy. Chem. B* 101: 9342-9351.

Cormier, F., Brion, F., Do, C.B. and Moresoli, C. (1996). Development of process strategies for anthocyanin-based food colorant production using *Vitis vinifera* cell cultures. CRC Press LLC: Boca Raton, FL, USA.

Davies, K. (2009). Annual Plant Reviews, Plant Pigments and their Manipulation. John Wiley & Sons.

Delgado-Vargas, F., Jimenez, A. R. and Paredes-Lopez, O. (2000). Natural pigments: carotenoids, anthocyanins, and betalainsâe"characteristics, biosynthesis, processing, and stability. *Crit. Rev. Food Sci. Nutri*. 40: 173-289.

Delgado-Vargas, F. and Paredes-Lopez, O., (2002). Natural colorants for food and nutraceutical uses. CRC Press.

Elisia, I. and Kitts, D.D. (2008). Anthocyanins inhibit peroxyl radical-induced apoptosis in Caco-2 cells. *Mol. Cell. Biochem.* 312: 139-145.

Feild, T.S., Lee, D.W. and Holbrook, N. M. (2001). Why leaves turn red in autumn. The role of anthocyanins in senescing leaves of red-osier dogwood. *Pl. Physiol.* 127: 566-574.

Folts, J. D. (1998). Antithrombotic potential of grape juice and red wine for preventing heart attacks. *Pharm. Biol* 36, 21-27.

Friedman, M., Mackey, B. E., Kim, H.J., Lee, I.S., Lee, K.R., Lee, S.U., Kozukue, E. and Kozukue, N. (2007). Structure-activity relationships of tea compounds against human cancer cells. . *Agric. Food Chem.,* 55: 243-253.

GarciaViguera, C., Zafrilla, P. and Tomas Barberan, F.A. (1998). The use of acetone as an extraction solvent for anthocyanins from strawberry fruit. *Phytochem. Anal.*, 9: 274-277.

Ghiselli, A., Nardini, M., Baldi, A. and Scaccini, C., (1998). Antioxidant activity of different phenolic fractions separated from an Italian red wine. *J. Agric. Food Chem.* 46: 361-367.

Giusti, M.M., Ghanadan, H. and Wrolstad, R.E. (1998). Elucidation of the structure and conformation of red radish (*Raphanus sativus*) anthocyanins using one-and two-dimensional nuclear magnetic resonance techniques. *J Agric. Food Chem.* 46: 4858-4863.

Giusti, M.M.N., Rodriguez-Saona, L.E., Baggett, J.R., Reed, G.L., Durst, R.W. and Wrolstad, R.E., (1994). Anthocyanin pigment composition of red radish cultivars as potential food colorants. red 23.

Gratzel, M., (2003). Dye-sensitized solar cells. Journal of Photochemistry and Photobiology C: *Photochem. Rev.* 4:145-153.

Harborne, J.B. (1988). Introduction to ecological biochemistry. Academic Press.

He, F., Mu, L., Yan, G.L., Liang, N.N., Pan, Q.H., Wang, J., Reeves, M.J., and Duan, C.Q. (2010). Biosynthesis of anthocyanins and their regulation in colored grapes. *Mol*. 15: 9057-9091.

He, J. and Giusti, M.M. (2010). Anthocyanins: natural colorants with health-promoting properties. *Ann*. *Rev*. *Food Sci. Technol.* 1: 163-187.

Hedin, P.A. and Hedin, P.A. (1983). Plant resistence to insects. *American Chemical Society Symposium Series*.

Igarashi, K., Takanashi, K., Makino, M., Yasui, T., (1989). Antioxidative activity of major anthocyanin isolated from wild grapes (*Vitis coignetiae*). *J. Japanese Soc. Food Sci. Technol*.

Jackman, R.L., Yada, R.Y., Tung, M.A., Speers, R., 1987. Anthocyanins as food colorantsâe"a review. *J. Food Biochem.* 11: 201-247.

Jankowski, A., Jankowska, B., Niedworok, J., 1999. The effect of anthocyanin dye from grapes on experimental diabetes. *Folia medica Cracoviensia* 41: 5-15.

Jankowski, A., Jankowska, B. and Niedworok, J. (2000). The influence of *Aronia melanocapra* in experimental pancreatitis. Polski merkuriusz lekarski: organ Polskiego Towarzystwa Lekarskiego 8: 395-398.

Joseph, J.A., Shukitt-Hale, B., Denisova, N.A., Bielinski, D., Martin, A., McEwen, J.J. and Bickford, P.C. (1999). Reversals of age-related declines in neuronal signal transduction, cognitive, and motor behavioral deficits with blueberry, spinach, or strawberry dietary supplementation. *J. Neurosci.* 19: 8114-8121.

Kamei, H., Hashimoto, Y., Koide, T., Kojima, T. and Hasegawa, M. (1998). Anti-tumor effect of methanol extracts from red and white wines. *Cancer Biother. Radiopharma.* 13: 447-452.

Kerio, L.C., Wachira, F.N., Wanyoko, J.K. and Rotich, M.K., (2012). Characterization of anthocyanins in Kenyan teas: Extraction and identification. *Food Chem*. 131: 31-38.

Kong, J.M., Chia, L.S., Goh, N.K., Chia, T.F. and Brouillard, R. (2003). Analysis and biological activities of anthocyanins. *Phytochem.* 64: 923-933.

Lee, S.H., Park, S.M., Park, S.M., Park, J.H., Shin, D.Y., Kim, G.Y., Ryu, C.H., Shin, S.C., Jung, J.M. and Kang, H.S. (2009). Induction of apoptosis in human leukemia U937 cells by anthocyanins through down-regulation of Bcl-2 and activation of caspases. *Int. J. Oncol.* 34: 1077-1083.

Lila, M.A. (2004). Anthocyanins and human health: an *in vitro* investigative approach. *Bio. Med Res. Int.* 306-313.

Lo, C.W., Huang, H., Lin, H., Chien, C. and Wang, C. (2007a). Effect of Hibiscus anthocyanins • rich extract induces apoptosis of proliferating smooth muscle cell via activation of P38 MAPK and p53 pathway. *Mol. Nutri. Food Res.* 51:1452-1460.

Lo, C.W., Huang, H.P., Lin, H.M., Chien, C.T. and Wang, C.J. (2007b). Effect of *Hibiscus* anthocyanins • rich extract induces apoptosis of proliferating smooth muscle cell via activation of P38 MAPK and p53 pathway. *Mol. Nutri. Food Res.* 51: 1452-1460.

Maeda-Yamamoto, M., Saito, T., Nesumi, A., Tokuda, Y., Ema, K., Honma, D., Ogino, A., Monobe, M., Murakami, A. and Tachibana, H. (2012). Chemical analysis and acetylcholinesterase inhibitory effect of anthocyanin • rich red leaf tea (cv. Sunrouge). *J. Sci. Food Agric.* 92: 2379-2386.

Matsumoto, H., Inaba, H., Kishi, M., Tominaga, S., Hirayama, M. and Tsuda, T. (2001). Orally administered delphinidin 3-rutinoside and cyanidin 3-rutinoside are directly absorbed in rats and humans and appear in the blood as the intact forms. *J. Agric. Food Chem.* 49: 1546-1551.

Matsumoto, H., Nakamura, Y., Tachibanaki, S., Kawamura, S. and Hirayama, M. (2003). Stimulatory effect of cyanidin 3- glycosides on the regeneration of rhodopsin. *J. Agric. Food Chem*. 51, 3560-3563.

Mazza, G. and Miniati, E., (1993). Anthocyanins in fruits, vegetables, and grains. CRC press.

McGhie, T.K. and Walton, M.C. (2007). The bioavailability and absorption of anthocyanins: towards a better understanding. *Mol. Nutri. Food Res.* 51: 702-713.

Merzlyak, M.N., Chivkunova, O.B., Solovchenko, A.E. and Naqvi, K.R. (2008). Light absorption by anthocyanins in juvenile, stressed, and senescing leaves. *J. Exp. Bot.* 59: 3903-3911.

Meyer, A.S., Yi, O.S., Pearson, D.A., Waterhouse, A.L. and Frankel, E.N. (1997). Inhibition of human low-density lipoprotein oxidation in relation to composition of phenolic antioxidants in grapes (*Vitis vinifera*). *J. Agric. Food Chem.* 45: 1638-1643.

Michaelis, L., Schubert, M.P. and Smythe, C.V. (1936). Potentiometric study of the flavins. *J. Biol. Chem.* 116: 587-607.

Mortensen, A. (2006). Carotenoids and other pigments as natural colorants. *Pure Appl. Chem.* 78: 1477-1491.

Muth, E.R., Laurent, J.M. and Jasper, P. (2000). The effect of bilberry nutritional supplementation on night visual acuity and contrast sensitivity. *Alt.*

Medi. Rev. 5: 164-173.

Nakaishi, H., Matsumoto, H., Tominaga, S. and Hirayama, M. (2000). Effects of black currant anthocyanoside intake on dark adaptation and VDT work-induced transient refractive alteration in healthy humans. *Alt. Med. Rev.* 5: 553-562.

Nicoue, E.E., Savard, S., Belkacemi, K. (2007). Anthocyanins in wild blueberries of Quebec: extraction and identification. *J. Agric. Food Chem.* 55: 5626-5635.

Obi, F.O., Usenu, I.A. and Osayande, J.O. (1998). Prevention of carbon tetrachloride-induced hepatotoxicity in the rat by H. rosasinensis anthocyanin extract administered in ethanol. *Toxicol*. 131: 93-98.

Oki, T., Masuda, M., Furuta, S., Nishiba, Y., Terahara, N. and Suda, I. (2002). Involvement of Anthocyanins and other Phenolic Compounds in Radical • Scavenging Activity of Purple • Fleshed Sweet Potato Cultivars. *J. Food Sci.,* 67:1752-1756.

Park, J.S., Kim, J.B., Hahn, B.S., Kim, K.H., Ha, S.H., Kim, J.B. and Kim, Y.H. (2004). EST analysis of genes involved in secondary metabolism in *Camellia sinensis* (tea), using suppression subtractive hybridization. *Pl. Sci.* 166:953-961.

Pawlowicz, P., Wilczynski, J., Stachowiak, G. and Hincz, P. (2000). Administration of natural anthocyanins derived from chokeberry (*Aronia melanocarpa*) juice in the treatment of fetal intrauterine growth retardation of idiopathic and preeclamptic origin. Influence on metabolism of plasma oxidized lipoproteins-the role of autoantibodies to oxidized low density lipoproteins. *Ginekologia Polska* 71: 848-853.

Piovan, A., Filippini, R., Favretto, D., 1998. Characterization of the anthocyanins of *Catharanthus roseus* (L.) G. Don in vivo and in vitro by electrospray ionization ion trap mass spectrometry. *Rapid communications in mass spectrometry* 12: 361-367.

Revilla, E., Ryan, J.M. and Martin-Ortega, G. (1998). Comparison of several procedures used for the extraction of anthocyanins from red grapes. *J. Agric. Food Chem.* 46: 4592-4597.

Revilla, I., Perez-Magarino, S., Gonzalez-SanJose, M. L. and Beltran, S. (1999). Identification of anthocyanin derivatives in grape skin extracts and red wines by liquid chromatography with diode array and mass spectrometric detection. *J. Chromatogra.* 847: 83-90.

Rice-Evans, C.A. and Packer, L. (2003). Flavonoids in health and disease. CRC Press.

Rivero-Perez, M.D., Muniz, P. and Gonzalez-Sanjose, M.L., (2008). Contribution of anthocyanin fraction to the antioxidant properties of wine. *Food Chem. Toxicol.* 46: 2815-2822.

Saito, N., Tatsuzawa, F., Miyoshi, K., Shigihara, A., Honda, T. (2003). The first isolation of C-glycosylanthocyanin from the flowers of *Tricyrtis formosana*. *Tetrahedron Letters.* 44: 6821-6823.

Saito, N., Timberlake, C.F. Tucknott, O.G. and Lewist, I.A.S. (1983). Fast atom bombardment mass spectrometry of the anthocyanins violanin and platyconin. *Phytochem.* 22: 1007-1009.

Saito, T., Honma, D., Tagashira, M., Kanda, T., Nesumi, A. and Maeda-Yamamoto, M., (2011). Anthocyanins from new red leaf tea 'Sunrouge'. *J. Agric. Food Chem.* 59: 4779-4782.

Schurch, C., Blum, P. and Zulli, F. (2008). Potential of plant cells in culture for cosmetic application. *Phytochem. Rev.* 7:599-605.

Smith, M.A., Marley, K.A., Seigler, D., Singletary, K. W. and Meline, B. (2000). Bioactive properties of wild blueberry fruits. *J. Food Sci. Chicago* 65: 352-356.

Socaciu, C. (2007). Food colorants: chemical and functional properties. CRC Press.

Steyn, W.J., Wand, S.J.E., Holcroft, D.M. and Jacobs, G. (2002). Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. *New Phytol.* 155: 349-361.

Stintzing, F.C. and Carle, R. (2004). Functional properties of anthocyanins and betalains in plants, food, and in human nutrition. Trends *Food Sci. Technol*. 15: 19-38.

Sullivan, J. (1998). Anthocyanin. *Carnivor. Pl. Newslett.* 27: 26-28.

Tanaka, Y., Sasaki, N. and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Pl. J.* 54: 733-749.

Tatsuzawa, F., Saito, N., Miyoshi, K., Shinoda, K., Shigihara, A. and Honda, T. (2004). Diacylated 8-C-glucosylcyanidin 3-glucoside from the flowers of *Tricyrtis formosana*. *Chemi. Pharma Bull.* 52: 631-633.

Tounekti, T., Joubert, E., Hern-Ã¡ndez, I. and Munn Bosch, S. (2013). Improving the polyphenol content of tea. *Crit. Rev. Pl. Sci.* 32:192-215.

Tsai, P.J., McIntosh, J., Pearce, P., Camden, B. and Jordan, B.R., (2002). Anthocyanin and antioxidant capacity in Roselle (*Hibiscus sabdariffa* L.) extract. *Food Res. Int.* 35:351-356.

Tsai, P.J. and Ou, A.S.M. (1996). Colour degradation of dried roselle during storage. *Food Sci.* 23: 629-640.

Tsuda, T., Horio, F., Uchida, K., Aoki, H. and Osawa, T. (2003). Dietary cyanidin 3-O-Î²-D-glucoside-rich purple corn color prevents obesity and ameliorates hyperglycemia in mice. *J. Nutri.* 133: 2125-2130.

Viskelis, P., Rubinskiene, M., Jasutiene, I., Sarkinas, A., Daubaras, R. and Cesoniene, L. (2009). Anthocyanins, antioxidative, and antimicrobial

properties of American cranberry (*Vaccinium macrocarpon* Ait.) and their press cakes. *J. Food Sci.* 74: C157-C161.

Wang, C.J., Wang, J.M., Lin, W.L., Chu, C.Y., Chou, F.P. and Tseng, T.H. (2000). Protective effect of Hibiscus anthocyanins against tert-butyl hydroperoxide-induced hepatic toxicity in rats. *Food Chem. Toxicol*. 38: 411-416.

Wang, H., Cao, G., Prior, R.L., (1997). Oxygen radical absorbing capacity of anthocyanins. *J. Agric. Food Chem.* 45: 304-309.

Wang, J., Sporns, P., (1999). Analysis of anthocyanins in red wine and fruit juice using MALDI-MS. *J. Agric. Food Chem.,* 47: 2009-2015.

Wang, S. Y., Lin, H.-S., (2000). Antioxidant activity in fruits and leaves of blackberry, raspberry, and strawberry varies with cultivar and developmental stage. *J. Agric. Food Chem.,* 48: 140-146.

Wang, Y., Gao, L., Shan, Y., Liu, Y., Tian, Y., Xia, T., (2012). Influence of shade on flavonoid biosynthesis in tea (*Camellia sinensis* (L.) O. Kuntze). *Scientia Horticulturae.* 141: 7-16.

Wight, W., (1962). Tea classification revised. 31: 298.

Wrolstad, R.E., (2004). Anthocyanin pigments - ″Bioactivity and coloring properties''. *J. Food Sci.* 69: C419-C425.

Yoshimoto, M., Okuno, S., Yoshinaga, M., Yamakawa, O., Yamaguchi, M., Yamada, J., (1999). Antimutagenicity of sweetpotato (Ipomoea batatas) roots. Bioscience, biotechnology, and biochemistry 63, 537-541.

Yue, X., Zhang, W., Deng, M., (2011). Hyper-production of 13 C-labeled trans-resveratrol in Vitis vinifera suspension cell culture by elicitation and *in situ* adsorption. *Biochem. Eng.J.* 53:292-296.

Symbiosis
DOI 10.1007/s13199-016-0401-3

Comparative genomics of *Prauserella* sp. Am3, an actinobacterium isolated from root nodules of *Alnus nepalensis* in India

Debadin Bose[1] · Indrani Sarkar[1] · Reha Labar[1] · Rediet Oshone[2] · Shimaa Ghazal[2] · Krystalynne Morris[2] · Feseha Abebe-Akele[2] · W. Kelley Thomas[2] · Louis S. Tisa[2] · Arnab Sen[1]

Received: 30 November 2015 / Accepted: 11 March 2016

Abstract A novel actinomycete strain, assigned as Am3, was isolated from the root nodules of *Alnus nepalensis* at Mirik hills, India. Analysis of the 16s rRNA gene sequence placed this new strain within the genus *Prauserella*. The genome was sequenced by Illumina sequencing and resulting 5.33-Mbp high quality draft genome sequenced with a G + C content of 70.0 % and 4828 candidate protein-encoding genes. Phylogenetically, *Prauserella* clusters very close to *Amycolatopsis* and was previously placed under the genus *Amycolatopsis*. Our main focus was to reveal the genomic similarities and dissimilarities of the newly sequenced *Prauserella* sp. Am3 with the type strain, *Prauserella rugosa* DSM 43194 T, and to determine its relationship with *Amycolatopsis*, which is happened to be the closest genus of *Prauserella*. Taking an in silico approach, bioinformatic analysis revealed that the core genome of *Amycolatopsis* and *Prauserella* contained 1589 genes. The two *Prauserella* genomes shared approximately 4224 genes, and 237 and 245 unique genes were found in the *P. rugosa* and *Prauserella* sp. Am3 genomes, respectively. Analysis of various phylogenetic trees including a 16s rRNA gene tree, MLSA protein-based tree and concatenated core-genome-based tree, placed both *Prauserella* genomes together with *Amycolatopsis halophila* YIM 93233 as its closest neighbor. Blast Matrix analysis of the predicted proteomes revealed about 86 % homology between the two *Prauserella* genomes. Analysis of the strand variation property revealed the absence of replication-transcriptional selection. Overall, a high degree of similarity was found between the two *Prauserella* genomes and a high percentage of similarity occurred among the *Prauserella* genomes and *Amycolatopsis halophila*.

Keywords Sequencing · MLSA · ANI score · Genome plasticity · Pan-core plot

Presented at the 18th International Meeting on Frankia and Actinorhizal Plants (ACTINO2015), August 24–27, 2015, Montpellier, France

Electronic supplementary material The online version of this article (doi:10.1007/s13199-016-0401-3) contains supplementary material, which is available to authorized users.

✉ Arnab Sen
senarnab_nbu@hotmail.com

[1] NBU Bioinformatics Facility, Department of Botany, University of North Bengal, Siliguri 734013, India

[2] Department of Molecular, Cellular & Biomedical Sciences, University of New Hampshire, 46 College Road, Durham, NH 03824-2617, USA

1 Introduction

Actinorhizal plants are a group of dicotyledonous plants that form a symbiotic association with the nitrogen-fixing actinobacterium *Frankia* which results in root nodule structures. These plants are distributed among eight orders, eight families, 24 genera and 194 species of angiosperms (Benson and Silvester 1993; Normand and Fernandez 2009). Actinorhizal plants are found worldwide and in diverse environments ranging from arctic tundra (*Dryas* spp.) and alpine forest (*Alnus* sp., *Coriaria* sp. etc) to coastal and xeric condition (*Casuarina* sp.). *Frankia* in symbiosis with actinorhizal plants fixed sustainable amount of molecular nitrogen (ranges from 22 to 300 kg of N_2 per ha per year) and are comparable to those of leguminous plants (Hibbs and Cromack 1990; Wall 2000; Wheeler and Miller 1990). As early visitors of marginal soils, actinorhizal plants are considered pioneer species in the landslides and other threatened areas.

Beside *Frankia*, different actinobacteria genera have also been isolated from the rhizosphere and root nodules of actinorhizal plants (Ghodhbane-Gtari and Tisa 2014; Ghodhbane-Gtari et al. 2010; Liu et al. 2009; Carro et al. 2013; Trujillo et al. 2006; Valdes et al. 2005). Many of these diverse actinobacteria have been characterized up to the genus level and include *Streptomyces*, *Nocardia*, *Micromonospora*, and *Actinoplanes*. These non-*Frankia* actinobacteria are filamentous, and mostly fall into the Actinomycetales category. Colonies and cultures of the non-*Frankia* isolates showed a wide range of colors and physiological traits. These pigmentations ranged in color from white, orange, yellow to maroon (Ghodhbane-Gtari and Tisa 2014), pink to brownish red (Liu et al. 2009), intense orange (Trujillo et al. 2006), and yellow-white (Carro et al. 2013; Valdes et al. 2005). Many of them produce highly branched filaments with some exceptions including *Nocardia* and *Micromonospora* (Maldonado and Quintana 2015).

An isolate from cattle rumen was first designated as *Nocardia rugosa* by Di Marco and Spalla (1957) and latter classified as *Amycolatopsis rugosa* (Lechevalier et al. 1986). However, *A. rugosa* differed in physiological and chemical properties from many of the known *Amycolatopsis* strains and also clustered differently in 16s rDNA sequence based phylogenetic tree (i.e. *A. rugosa* clustered with the *Sacharomonospora* genus instead of being grouped together with other *Amycolatopsis* strains) (Kim and Goodfellow 1999). Based on these physiological and genotypical differences, a new genus *Prauserella* gen nov. was proposed and *Amycolatopsis rugosa* was renamed as *Prauserella rugosa* (Kim and Goodfellow 1999). Since then, many different species have been identified including *Prauserella halophila* (Li et al. 2003), *Prauserella alba* (Li et al. 2003), *Prauserella salsuginis* (Li et al. 2009), *Prauserella flava* (Li et al. 2009), *Prauserella aidingensis* (Li et al. 2009), *Prauserella sediminis* (Li et al. 2009), *Prauserella manina* (Wang et al. 2010) and *Prauserella muralis* (Schafer et al. 2010; Solanki and Kothari 2011). Members of this genus are aerobic, Gram-positive, non-acid-alcohol-fast and non-motile filamentous actinomycetes (Kim and Goodfellow 1999). These actinobacteria produce branched substrate mycelium having a diameter of 0.6–0.8 μm. After 24–48 h growth on rich medium, the hyphae fragment into irregular rods. One of the major characteristics of *Prauserella* is the absence of aerial hyphae. The temperature range for their growth is 10–45 °C with optimum temperature of 34 °C and the optimum pH range is 6.8–7.2 (Kim and Goodfellow 1999).

During the process of isolating *Frankia* from root nodules of *Alnus nepalensis*, an actinobacterium was isolated and 16s rRNA gene analysis identified it as a member of the genus *Prauserella.* Although physiological studies are available on this genus, *Prauserella rugosa* DSM 43194 T, the type strain of this genus, is the only genome (Genbank Accession Number- NR_024889) from this genus that has been sequenced. In this study, we have taken a genomic approach and have sequenced the genome of this new *Prauserella* strain Am3 isolated from the root nodules of *Alnus nepalensis.* The genome of *Prauserella* strain Am3 was analyzed and compared to the *P. rugosa* DSM 43194 T and some selected *Amycolatopsis* genomes, which are very close to *Prauserella* genus.

2 Materials and methods

2.1 Isolation procedures

Nodules from *Alnus nepalensis* were collected along the Siliguri –Darjeeling road via Mirik-Pasupati-Ghoom-Sonada (26.8870° N, 88.1870° E) of West Bengal, India. As a road side tree, *A. nepalensis* is commonly found in abundance at this location. Fresh, young and light-brown-colored root nodules from young *A. nepalensis* were collected in the Mirik hills during the summer 2007. The nodule lobes were surface sterilized with a series of mild surface-sterilizing agents as previously described (Bose and Sen 2006). The nodule lobes were incubated in Qmod medium (Lalonde and Calvert 1979) supplemented with 3 % activated charcoal for 48 h at 30 °C. Activated charcoal was used to absorb out phenolic extracts from the nodule lobes which may hinder the proper growth of the endophytes (Thomas 2008). After incubation, the nodule lobes were washed twice with sterile distilled water to remove charcoal. Epidermal layers were removed and nodule lobes were chopped into small pieces. Each nodule piece was incubated with 10 mL of Qmod (Lalonde and Calvert 1979) medium in 30 mL screw cap test tube at 28 °C for around one month in dark without shaking.

After one-month incubation in dark, bacterial colonies formed minute cottony-white dots at the bottom of the tube. Each colony was collected and subcultured in separate tubes containing Qmod medium.

2.2 Culture growth conditions

After successful isolation of the bacteria in pure culture, the isolate was maintained in Qmod medium in 250 ml flask in dark with occasional shaking.

2.3 Field emission scanning electron microscopy

For field emission scanning electron microscopy, one-month-old *Prauserella*, cultures were used. The bacterial cells were harvested in a micro-centrifuge tube and centrifuged at 1000 g for 5 min. The cells were washed with distilled water and treated with 4 % (*w*/*v*) glutaraldehyde solution in phosphate

buffer (pH 7.2) for 4 h at 30 °C. The glutaraldehyde solution was decanted and the samples were washed with the same buffer for three times. The samples were treated with 1 % Osmium tetra oxide solution in phosphate buffer (pH 7.2) for 1 h and washed with the phosphate buffer (pH 7.2) for 3 h. The *Prauserella* samples were treated bypassing through with different graduated ethanol solutions from 50 % to absolute for 10 min each. The 95 % ethanol wash was performed twice. After dehydration, the samples were stored in absolute alcohol for microscopy (Bajwa 2004).

The samples were critical point dried for 2 h. Each sample was mounted on copper tape and platinum coated (by JFC 1600 platinum caster, JEOL, Japan) and viewed on JEOL JSM-6700F field emission scanning electron microscope at an accelerating voltage of 2KV.

2.4 Plant infectivity test

The seeds of healthy *A. nepalensis* were collected in the month of March from Darjeeling hills. The seeds were soaked overnight with double distilled water. The soaked seeds were surface sterilized with 30 % H_2O_2 for 10 min followed by washing several times with sterile double distilled water. The seeds were placed on a sterile moist filter paper on a Petri plate with 1X Hoagland solution (Hoagland and Arnon 1950) and kept in a seed germinator at 30 ± 1 C for germination. Fifteen days old seedlings were transferred to a magenta box containing sterile N_2 free Hoagland solution on sterile blotting paper supported by stainless steel supporter. The following three sets were prepared for each isolates: (1) seedlings inoculated with crushed nodule suspension (positive control); (2) seedlings inoculated with 100 μl of 30 day old culture; (3) un-inoculated seedlings (negative control). Twenty seedlings were used in each set. After an incubation period of 20–30 days, all the seedlings were carefully examined to detect any root hair deformation or nodule formation.

2.5 Genome sequencing and assembly

The draft genome of *Prauserella* strain Am3 was generated at the Hubbard Genome Center (University of New Hampshire, Durham, NH) using Illumina technology (Kuhn et al. 2004) techniques. A standard Illumina shotgun library was constructed and sequenced using the Illumina HiSeq2000 platform, which generated 29,139,048 reads (260 bp insert size) totaling 4370.9 Mbp. The Illumina sequence data were assembled using CLC Genomics Workbench (6.5.1) and AllPaths-LG (Version r41043) (Gnerre et al. 2011). The final draft assembly contained 61 contigs with an N_{50} of 370 kb. The total size of the genome is 5.3 Mbp, and the final assembly is based on 2057.3 Mb of Illumina draft data, provided an average 680 x coverage of the genome.

Genome annotation was performed within the Integrated Microbial Genomes (IMG) platform developed by the Joint Genome Institute, Walnut Creek, CA, USA (Markowitz et al. 2006). Genes were identified using Prodigal (Hyatt et al. 2010). The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) non redundant database, UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAScanSE tool (Lowe and Eddy 1997) was used to find tRNA genes, whereas ribosomal RNA genes were found by searches against models of the ribosomal RNA genes built from SILVA (Pruesse et al. 2007). The metabolic potential of the genome was analyzed by the use of the antiSMASH program (Blin et al. 2013; Medema et al. 2011) and IMG platform.

2.6 Nucleotide sequence accession numbers

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JTJI00000000. The version described in this paper is version JTJI01000000.

2.7 Bioinformatics study

2.7.1 Retrieval of genome sequences

The nucleotide sequences along with the translated protein sequences of 21 *Amycolatopsis* and 2 *Prauserella* genomes were retrieved from JGI-IMG database (http://img.jgi.doe.gov/) (Mavromatis et al. 2009). A list of the genomes and general properties are presented in supplementary Table 1. Major information regarding the strand specificity and functional Cluster of Orthologs (COG) category (Tatusov et al. 2000) for each gene were also obtained from the above mentioned database.

2.7.2 16s rRNA gene and MLSA phylogeny tree

The 16s rRNA genes for the 21 *Amycolatopsis* strains and two *Prauserella* strains were collected from IMG database. The partial 16s rRNA sequence for remaining *Prauserella* strains were collected from NCBI database. The 16s rRNA sequences were truncated to match the sequence length. This drastically reduces the unaligned regions too. The truncated 16s rRNA gene sequence alignment was done by ClustalW (Li 2003). Mega ver. 6 software (Tamura et al. 2007) was used to build the 16s rRNA gene based tree using neighbor-joining model and a bootstrap value of 1000. The Multi-Locus Sequence Analysis (MLSA) tree was generated with Random Axelerated Maximum Likelihood (RAxML) (Stamatakis 2006) software. For MLSA analysis, AtpI, GyrA, FtsZ, SecA, DnaK proteins were used. Although there are several

housekeeping proteins present in actinobacteria group, the five proteins were chosen because they represent single proteins universally found among all of the studied strains. The amino acid sequences of the five housekeeping genes were aligned by the use of ClustalW software (Li 2003) and alignment files were concatenated by Seaview software (Gouy et al. 2010). The concatenated file was used to make the tree in RAxML software with PROTGAMMAWAG (de Mendoza et al. 2010) substitution model and 1000 bootstrap value. *Pseudonocardia* was used as outgroup in these phylogenetic study.

2.8 ANI score and genome-to- genome distance calculation

Average nucleotide Identity (ANI) of several genomes was performed to reveal the genomic closeness. The ANI score calculates the similarity index between a given pair of genomes. If the score is >95 %, they are considered the same species (Kim et al. 2014). The ANI score was calculated via the ANI calculator tool developed by Kostas Lab (http://enve-omics.ce.gatech.edu/ani/). The Genome-to Genome distance calculator was also used to show the genomic distances among the four strains (2 *Prauserella* and 2 *Amycolatopsis*). The distance was calculated by GGDC 2.1 software (http://ggdc.dsmz.de/distcalc2.php). This tool calculates the genomic distance and transforms these values analogous to DNA-DNA Hybridization (DDH) values (Meier-Kolthoff et al. 2013). Generalized Linear Model (GLM) is used to estimate the confidence level of the DDH values. If the DDH value > = 70 %, the target and query genomes are consider to be same species. If the DDH value > = 79 %, they are considered to be the same sub-species (Meier-Kolthoff et al. 2013).

2.8.1 Amino acid and codon usage count

The codon and amino acid usage patterns were determined and compared among the studied genomes. A simple BioPerl module implemented in CMG Biotools (Vesth et al. 2013) was used to evaluate the codon and amino acid usage patterns. From that module, the third position codon bias (bT) was calculated by the following formula:

$$bT = \frac{No.\text{of each base on each position of each codon}}{\text{total number of codons}}$$

Two separate heat maps were generated on amino acid usage and codon usage by R package (Ihaka and Gentleman 1996). Rose plots were also generated to have a deeper look on the amino acid and codon usage preference of each genome under investigation (Vesth et al. 2013).

2.8.2 Identification of putative horizontally transferred (pHGT), pseudogenes and insertion sequence

Comparative analysis of the *Prauserella* genomes was performed using the IMG system (Markowitz et al. 2006). Putative horizontally transferred genes (pHGT) were defined as genes that have the best hits to genes that are not in the phylogenetic group of the query sequence. To determine a pHGT, not only was the best hit (highest bit score) used, but all of the hits that have bit score equal or greater than 95 % of the best hit were taken under consideration. Pseudogenes were identified on JGI-IMG platform via "Pseudo genes" option under the tab "Genes total number" (Markowitz et al. 2012). IMG database actually calculate the Pseudogene number via GenePRIMP (Pati et al. 2010). Insert sequences (IS) were identified via IS Finder server (Siguier et al. 2006).

2.8.3 Dot-plot generation

Dot Plots were generated via Mummer (Armengol et al. 2003) on the IMG site. The DNA sequences for two genomes were chosen and the nucmer (NUCleotide MUMmer) pipeline was used to generate the plots. The whole genome alignment of two *Prauserella* and *A. halophila* genomes was done by Mauve software. The circular genome view of *Prauserella* genomes were generated by GC view (Stothard and Wishart 2005) and have been given in supplementary files (ESM 1,2,3).

2.8.4 Creation of BLAST matrix and Pan Core plot

The BLAST matrix is a visual representation of a pairwise comparison of targeted proteomes using the BLAST algorithm (Vesth et al. 2013). The whole process of creating a blast matrix was performed on the CMG-Biotool platform using "matrix_createconfig" and "matrix" programs.

The pan and core genome analyses were also done by pancoreplot_createconfig (Vesth et al. 2013) program implemented in the CMG platform. The concatenated core gene family tree with Manhattan distance was built by pancoreplot tree program.

3 Results and discussion

3.1 Isolation of an actinobacterium from *Alnus* root nodules

The organism was isolated from root nodules of the *Alnus nepalensis* and it was identified as *Prauserella* by 16s rRNA gene analysis. An electron micrograph (Fig. 1) shows that the bacterium has fragmented mycelium. The mean diameter of the hyphae was found to be ±0.3 μm.

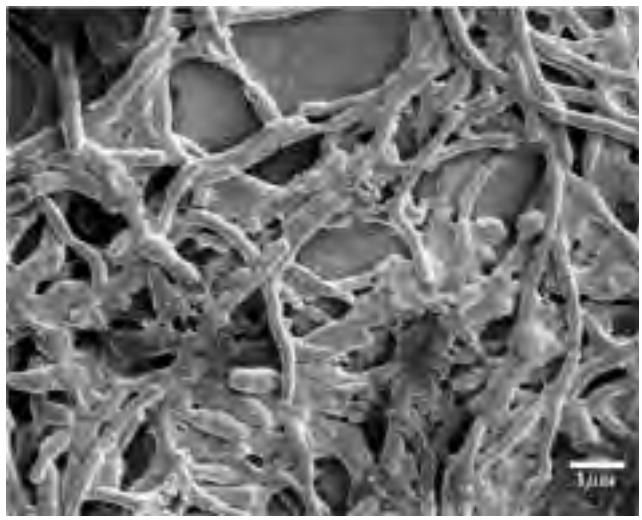


Fig 1 Scanning electron micrograph of *Prauserella* sp. Am3. The size bar represent 1 μm

3.2 Plant infectivity

We attempted to re-infect the host plant *A. nepalensis*, but failed to produce either root hair deformation or nodules on these root. The positive controls produced nodulated plants suggest that our test was okay. However, this result is not surprising. Most of the non-*Frankia* actinobacteria isolated from actinorhizal nodules are unable to re-infect their host plant.

3.3 Phylogenetic analysis of the isolates

A multi-tiered approach was used for phylogenetic analysis. First, a neighbor-joining (NJ) 16s rRNA gene phylogenetic tree grouped *P. rugosa* DSM 43194 and *Prauserella* sp. Am3 together with *A. halophila* YIM 93233 and *A. orientalis* DSM 43388. All of the partially sequenced 16s rRNA from different *Prauserella* strains formed one clade. However, all of the other *Amycolatopsis* strains were placed at a considerable distance from *P. rugosa* and *Prauserella* sp. Am3 (Fig. 2). Next, we determined the ANI scores among *A. halophila* YIM 93233, *A. orientalis* DSM 43388, *P. rugosa* and *Prauserella* sp. Am3 (Supplementary Table 2). *A. halophila* was closer to both *Prauserella* genomes than to *A. orientalis* DSM 43388. An in silico DNA-DNA hybridization via the Genome-to- Genome Distance calculator indicates that *Prauserella* sp. Am3 and *P. rugosa* may be the same species [DDH = 96.22 %].

Because of this close clustering pattern found with *A. halophila* YIM 93233, *A. orientalis* DSM 43388, *P. rugosa* and *Prauserella* sp. Am3, a phylogenomic study was performed by the use of a MLSA tree as described by Sen et al. (2014). The MLSA amino-acid-sequence-based tree also placed *A. halophila* YIM 93233 with both *P. rugosa* and *Prauserella* sp. Am3 in one clade with a good bootstrap value (Fig. 3). However, *A. orientalis* DSM 43388 was distantly placed with the other *Amycolatopsis* strains. This phylogenomic evidence supports the idea that *A. halophila* is somewhat closer to *Prauserella* than other *Amycolatopsis* strains. A comparative genomic summary of these three strains has been given in Table 1.

3.4 Amino acid and codon usage analysis

The codon and amino acid usage patterns of the genomes were determined and are shown in heatmaps (ESM4). All of the selected genomes showed a distinct GC bias. The most widely used codons were: GCG, GCC (alanine); GTC, GTG (valine); CGG, CGC (Arginine); CTC, CTG (Leucine); GAG, GGC (Glycine). Bioinformatics analysis of newly sequenced *Prauserella* sp. Am3 indicates the use of aliphatic amino acids with low energy cost and a preference of either G or C at the 3rd position of the codon. Thus, GC percentage and GC3 percentage are obviously governing the codon usage pattern in *Prauserella* sp. Am3.

3.5 Major shifts and realignments among *Prauserella* genomes

To understand the presence of major shifts and realignments, the *Prauserella* genomes were analyzed by the use of a DNA-DNA dot plot (Fig. 4). Both *Prauserella* genomes were very similar and showed no signs of translocation. These results along with results from the ANI and GGDC scores described above indicating that *Prauserella* sp. Am3 strain may be a member of *P. rugosa* species.

3.6 Genome plasticity

The two *Prauserella* and *A. halophila* genomes were analyzed for the presence of pHGT, IS elements and pseudogene (ESM5). pHGTs were approximately 0.68 % and 1.03 % of gene content of the *P. rugosa* and *Prauserella* sp. Am3 genomes, respectively. However, no pHGTs were identified in the *A. halophila* genome which might be due to nature of minimum draft sequence data. The percentage of IS elements present in *P. rugosa*, *Prauserella* sp. Am3 and *A. halophila* genomes were 5.26, 5.53 and 5.36 %, respectively. This result suggests that the selected genomes are not very stable and show high degree of plasticity. Pseudogenes were not identified in either of the two *Prauserella* genomes. Insertion sequences play an important role in evolution as they can promote the genome plasticity and gene inactivation (Touchon and Rocha 2007). pHGT generally is a positive indicator of IS element amount (Touchon and Rocha 2007).

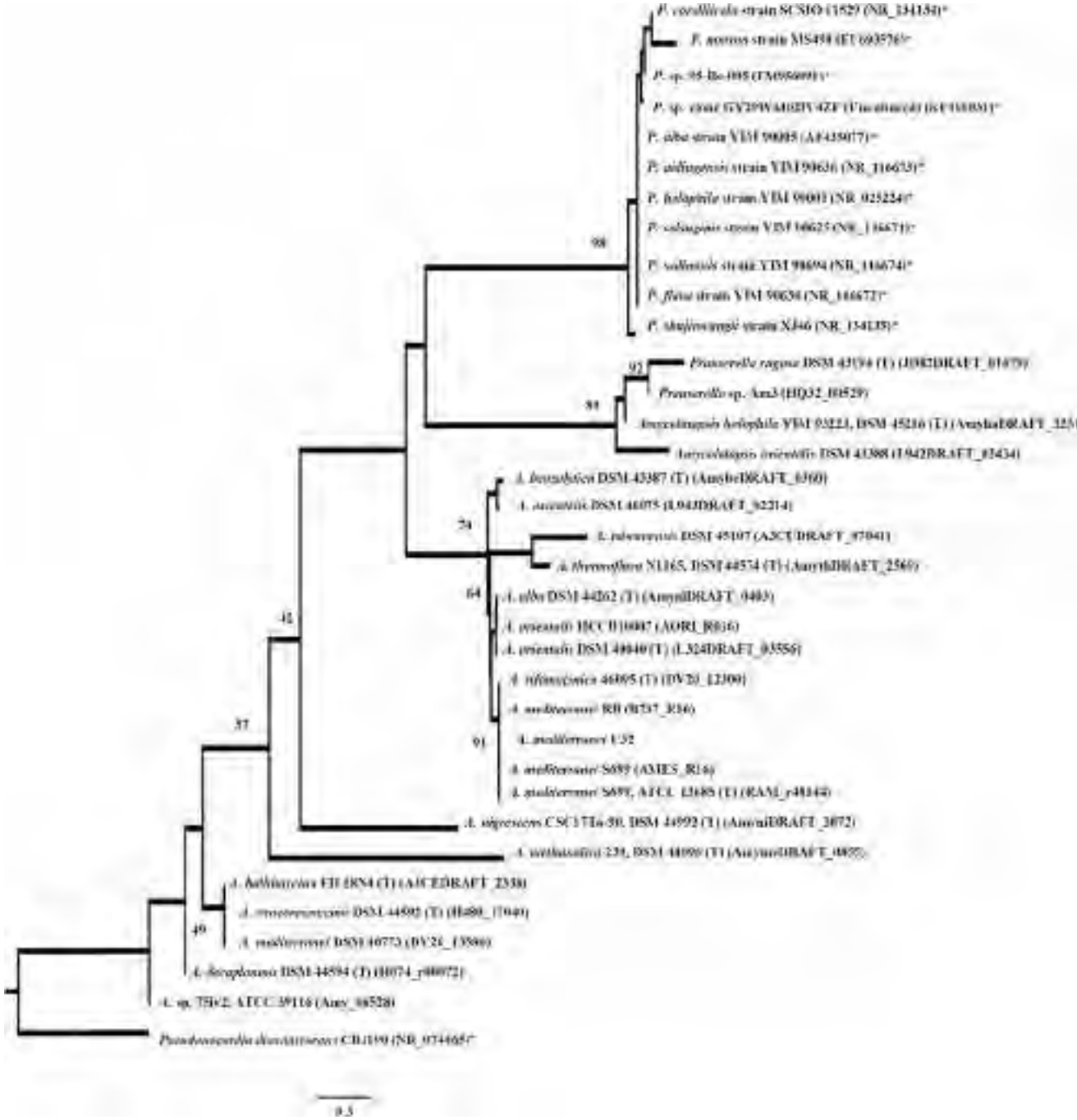


Fig 2 16s rRNA gene based Neighbour joining tree of *Prauserella* and *Amycolatopsis* genomes with1000 bootstrap values represented as percentage. Two strains of *Prauserella* have clustered with *A. halophila.* The (T) designate type strains and * designate genes taken from NCBI. Other genes are taken from IMG database. Corresponding Genbank accession numbers have been given for the sequences obtained from NCBI and for others Locus Tag has been given within parenthesis

3.7 Strand variation property

The strand-biased properties of the two *Prauserella* and *A. halophila* genomes were analyzed and visualized by a bar diagram (ESM 6). Stringent strand asymmetric bias was not observed (regardless of their expression pattern). The differences between the two *Prauserella* strains regarding the strand specificity nature may result from inversion mutations which have the potentiality to convert the leading strand genes to lagging strand genes as proposed by Chen and Zhang

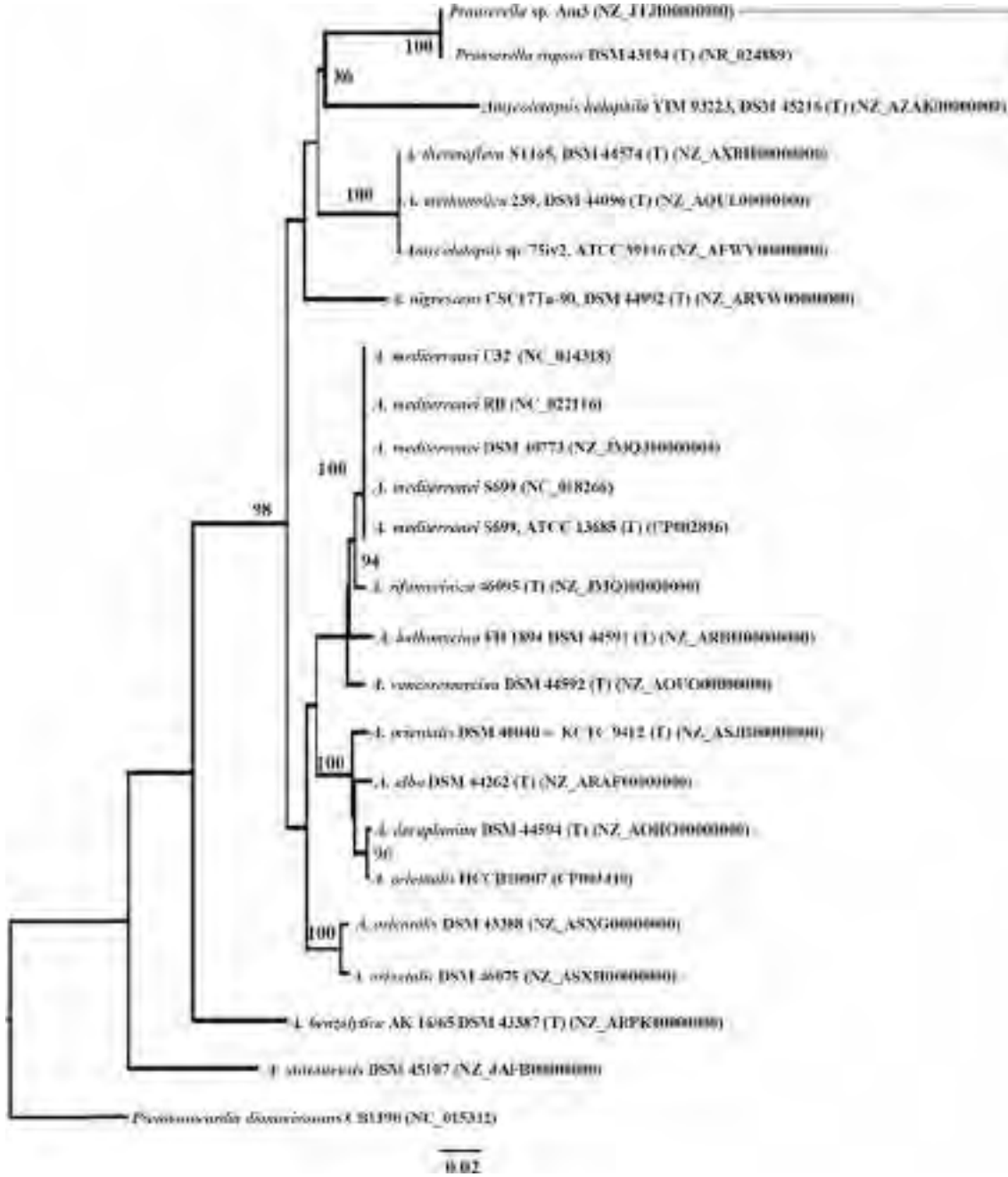


Fig 3 MLSA protein based tree of *Prauserella* and *Amycolatopsis* genomes built by RAxML using PROTGAMMAWAG substitution method and 1000 bootstrap value. Bootstarp values were represented in percentage. Two strains of *Prauserella* have clustered with *A. halophila* indicated by *ash box*. The (T) designate type strains and genbank accession numbers are given inside the parenthesis

(2013). However, it still needs a detailed study. Thus, absence of replication associated mutational pressure is confirmed on these genomes as there was no strand asymmetric nature persisted on those three considered genomes (Das et al. 2006). However, when we considered the potentially highly expressed genes, a considerable number of highly expressed genes were found to be located on the lagging strand in all three strains. These results indicate that transcriptional-associated mutational pressure was not apparent. This type of result was also found in *Bifidobacteria* (Roy et al. 2015), another group of actinobacteria living in the gut of animal and human host. Chen and Zhang (2013) proposed that, some microbial genomes have highly-expressed genes on their lagging strand for attaining proper mutation-selection balance and our data can be supported by this theory.

Table 1 Summary of genome properties

Characterization	*Prauserella rugosa* DSM 43194	*Prauserella* sp. Am3	*Amycolatopsis halophila* DSM 45216
Accession Number	NR_024889	NZ_JTJI00000000	NZ_AZAK00000000
Genome Size(bp)	5,220,823	5,321,061	5,550,897
Scaffolds	2	46	1
GC content (%)	70	70	68
Gene content	4985	4969	5187
No. of genes for:			
rRNA	9	8	6
tRNA	48	47	49
CDS	4923	4907	5187
Biosynthetic cluster count	62	53	54
Biosynthetic cluster gene count	689	593	817
NCBI Taxon ID	43,354	1,515,610	592,678
No. of CRISPR	2	1	-

3.8 Pan- and core- genome plot

The pan and core-genomes for *Amycolatopsis* and *Prauserella* were determined as described in the Methods. As expected, with the addition of the further genomes, the pan-genome (blue circles) increased due to the accumulation of accessory genes whereas the conserved core genome (red circles) decreased. The core-gene pool contained genes which are essential for transcription, translation, replication and metabolism. A total of 1589 genes were found to be in the core genome for the 21 *Amycolatopsis* and 2 *Prauserella* strains and the pan genome contained 35,441 genes (ESM 7). Manhattan-distance-based concatenated phylogenetic tree based on the core genome (1589 genes) was build (ESM 8). The *Prauserella* genomes clustered with *A. halophila* and were separated from the rest of *Amycolatopsis* genomes. To get a clear picture of these genomes, we focused on pan and core-genomes of the two *Prauserella* and *A. halophila* strains. The distribution of commonly shared genes is shown in Venn diagrams (Fig. 5). The two *Prauserella* genomes shared a total of 4224 genes (Fig. 5a core gene indicated by black box). When the *A. halophila* genome was included in the analysis, this core-genome was reduced to 2124 genes (Fig. 5b, core gene indicated by black box).

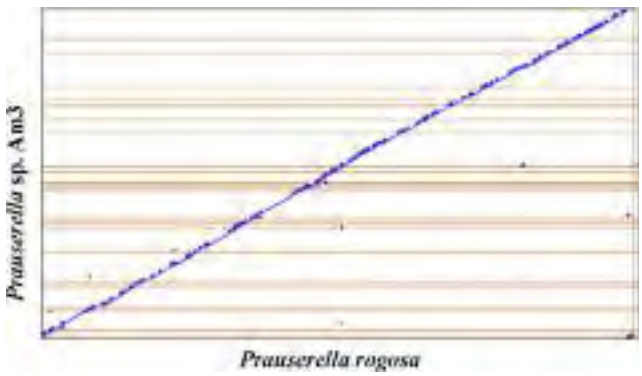


Fig 4 DNA-DNA dot plot of *Prauserella rugosa*, *Prauserella* sp. Am3 showing a high similarity between these two genomes

3.9 Blast matrix

The blast matrix of these genomes was generated by CMG biotools. The percentage of similarity among different genomes was represented with green color and the homology within a strain was visualized by red color (ESM 9). Two

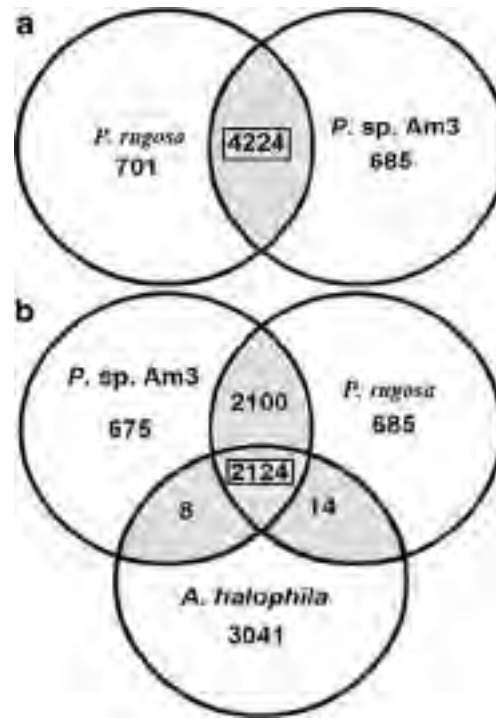


Fig 5 Venn diagrams of core and unique genes (**a**) between two *Prauserella* genomes and (**b**) among two *Prauserella* genomes and *A. halophila*

Prauserella genomes shared 85.7 % homology with each other. Though, blast matrix did not show a very high similarity between the *Prauserella* genomes and *A. halophila*, the percentage of similarity were more than the similarities obtained between *A. halophila* and other *Amycolatopsis* strains.

3.10 Proteins of Major Functional categories among investigated organisms

The predicted proteins were classified according to their Cluster of Orthologs (COG) category (Tatusov et al. 2000) and the results are presented in supplementary figure ESM 10 showing their distribution among the genomes. Most of the proteins were classified into the following COG categories: E (Amino acid transport and metabolism), G (Carbohydrate transport and metabolism), C (Energy production and conversion), and K (Transcription), which are mainly associated with the essential functions for an organism to survive in nature.

4 Conclusion

In this study, we have sequenced a new strain of *Prauserella* and have compared its genome with the type strain of the genus *P. rugosa* and other selected *Amycolatopsis* genomes, which are were reported as the closest genus of *Prauserella*. From both the 16s rRNA gene sequence based phylogenetic and MLSA amino acid sequence based phylogenomic analyses, *P. rugosa* and *Prauserella* sp. Am3 were found to be closer to *A. halophila* rather than any other *Amycolatopsis* strains. There were 4224 shared genes (core genome) between the two *Prauserella* genomes, which reduced to 2124 genes when the *A. halophila* genome was included. Analysis of the DNA-DNA dot plot revealed a high degree of similarity between the two *Prauserella* genomes.

Acknowledgments This work was supported by the USDA National Institute of Food and Agriculture Hatch 022821 (LST), and Department of Biotechnology, Govt. of West Bengal, India through grant no. 206/Bt (Estd.)/RD-22/2014 (AS). Authors acknowledge Department of Biotechnology, Govt. of India for the creation of Bioinformatics Facility at North Bengal University. IS acknowledges the receipt of BSR, UGC fellowship. Partial funding was provided by the New Hampshire Agricultural Experiment Station. This is Scientific Contribution Number 2643. Sequencing was performed on an Illumina HiSeq2500 purchased with an NSF MRI Grant: DBI-1229361to WKThomas.

Compliance with ethical standards This article does not contain any studies with human participants or animals performed by any of the authors.

Funding This work is also partially supported by Department of Biotechnology, Govt. of West Bengal, India through grant no. 206/Bt(Estd.)/RD-22/2014. LST is supported in part by a USDA National Institute of Food and Agriculture Hatch 022821.

Conflict of interest The authors declare that they have no conflict of interest.

References

Armengol ,L, Pujana MA, Cheung J, Scherer SW, Estivill X (2003) Enrichment of segmental duplications in regions of breaks of synteny between the human and mouse genomes suggest their involvement in evolutionary rearrangements. Hum Mol Genet 12: 2201–2208

Bajwa BS (2004) Molecular characterization of *Frankia* and *Alder-Frankia* symbiosis in Eastern India. Ph. D Thesis, NBU, Siliguri

Benson DR, Silvester WB (1993) Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. Microbiol Rev 57:293

Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T (2013) antiSMASH 2.0–a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res 41: W204–W212. doi:10.1093/nar/gkt449

Bose D, Sen A (2006) Isolation and heavy metal resistance pattern of *Frankia* from *Casuarina equisetifolia* nodules. Indian J Med Microbiol 46:9

Carro L et al. (2013) Micromonospora is a normal occupant of actinorhizal nodules. J Biosci 38:685–693

Chen X, Zhang J (2013) Why are genes encoded on the lagging strand of the bacterial genome? GBE 5:2436–2439

Das S, Paul S, Dutta C (2006) Evolutionary constraints on codon and amino acid usage in two strains of human pathogenic actinobacteria *Tropheryma whipplei*. J Mol Evol 62:645–658

De Mendoza A, Suga H, Ruiz-Trillo I (2010) Evolution of the MAGUK protein gene family in premetazoan lineages. BMC Evol Biol 10:93

Di Marco A, Spalla C (1957) La produzione di cobalamine da fermentazione con una nuova specie di *Nocardia*: *Nocardia rugosa*. Giorn Microbiol 4:24–30

Ghodhbane-Gtari F, Tisa LS (2014) Ecology and physiology of non-*frankia* actinobacteria from actinorhizal plants. In: Plasticity in plant-growth-promoting and phytopathogenic bacteria. Springer, New York, p 27–42

Ghodhbane-Gtari F, Essoussi I, Chattaoui M, et al. (2010) Isolation and characterization of non-*Frankia* actinobacteria from root nodules of *Alnus glutinosa*, *Casuarina glauca* and *Elaeagnus angustofolia*. Symbiosis 50:51–57

Gnerre S et al. (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. PNAS 108: 1513–1518

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27:221–224

Hibbs DE, Cromack K Jr. (1990) Actinorhizal plants in Pacific Northwest forests. In: Schwintzer CR, Tjepkema JD (eds) The biology of frankia and actinorhizal plants. Academic Press, San Diego, p 343–363

Hoagland DR, Arnon DI (1950) The water-culture method for growing plants without soil. Calif Agr Expt Sta Circ 347(2):32

Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. doi:10.1186/1471-2105-11-119

Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comput Graph Stat 5:299–314

Kim SB, Goodfellow M (1999) Reclassification of *Amycolatopsis rugosa* lechevalier et al. 1986 as *Prauserella rugosa* gen. nov., comb. nov. Int J Syst Evol Microbiol 49:507–512

Kim M, Oh H-S, Park S-C, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351

Kuhn K et al. (2004) A novel, high-performance random array platform for quantitative gene expression profiling. Genome Res 14: 2347–2356

Lalonde M, Calvert HE (1979) Production of *Frankia* hyphae and spores as an infective inoculant for *Alnus* species. Symbiotic Nitrogen Fixation in the Management of Temperate Forests 95:110

Lechevalier MP, Prauser H, Labeda DP, Ruan JS (1986) Two new genera of nocardioform actinomycetes: *Amycolata* gen. nov. and *Amycolatopsis* gen. nov. Int J Syst Evol Microbiol 36:29–37

Li K-B (2003) ClustalW-MPI: ClustalW analysis using distributed and parallel computing. Bioinformatics 19:1585–1586

Li W-J, Xu P, Tang S-K, Xu L-H, Kroppenstedt RM, Stackebrandt E, Jiang C-L (2003) *Prauserella halophila* sp. nov. and *Prauserella alba* sp. nov., moderately halophilic actinomycetes from saline soil. Int J Syst Evol Microbiol 53:1545–1549

Li Y, Tang S-K, Chen Y-G, Wu J-Y, Zhi X-Y, Zhang Y-Q, Li W-J (2009) *Prauserella salsuginis* sp. nov., *Prauserella flava* sp. nov., *Prauserella aidingensis* sp. nov. and *Prauserella sediminis* sp. nov., isolated from a salt lake. Int J Syst Evol Microbiol 59: 2923–2928

Liu N, Wang H, Liu M, Gu Q, Zheng W, Huang Y (2009) *Streptomyces alni* sp. nov., a daidzein-producing endophyte isolated from a root of *Alnus nepalensis*. Int J Syst Evol Microbiol 59:254–258

Lowe TM, Eddy SR (1997) tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Maldonado LA, Quintana ET (2015) Unexpected properties of *Micromonosporae* from marine origin. Adv Microbiol 5:452–456

Markowitz VM et al. (2006) The integrated microbial genomes(IMG) system. Nucleic Acids Res 34:D344–D348. doi:10.1093/Nar/Gkj024

Markowitz VM et al. (2012) IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res 40: D115–D122

Mavromatis K, Ivanova NN, Chen IMA, Szeto E, Markowitz VM, Kyrpides NC (2009) The DOE-JGI Standard operating procedure for the annotations of microbial genomes. Stand Genomic Sci 1:63

Medema MH et al. (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res 39: W339–W346. doi:10.1093/Nar/Gkr466

Meier-Kolthoff JP, Auch AF, Klenk H-P, Goker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60

Normand P, Fernandez MP (2008) Evolution and diversity of *Frankia*. In: Prokaryotic symbionts in plants. Springer Berlin Heidelberg, pp 103–125

Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC (2010) GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat Methods 7: 455–457

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196. doi:10.1093/nar/gkm864

Roy A, Mukhopadhyay S, Sarkar I, Sen A (2015) Comparative investigation of the various determinants that influence the codon and amino acid usage patterns in the genus *Bifidobacterium*.World J Microbiol Biotechnol 31(6):959–981. doi:10.1007/s11274-015-1850-1

Schafer J, Martin K, Kamfer P (2010) *Prauserella muralis* sp. nov., from the indoor environment. Int J Syst Evol Microbiol 60:287–290

Sen et al. (2014) The phylogeny of actinobacteria revisited in the light of complete genomes, the orders frankiales and micrococcales should be split into coherent entities. proposal of frankiales ord. nov., geodermatophilales ord. nov., acidothermales ord. nov. and nakamurellales ord. nov. Int J Syst Evol Microbiol 64:3821–3832

Siguier P, Parochon J, Lestrade L, Mahillon J, Chandler M (2006) ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 34:D32–D36

Solanki P, Kothari V (2011) Halophilic actinomycetes: salt-loving filaments. Int J Life Sci Technol 4:7–13

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690

Stothard P, Wishart DS (2005) Circular genome visualization and exploration using CGView. Bioinformatics 21:537–539

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Tatusov RL, Galperin MY, Natale DA, Koonin EV (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28:33–36

Thomas TD (2008) The role of activated charcoal in plant tissue culture. Biotechnol Adv 26:618–631

Touchon M, Rocha EPC (2007) Causes of insertion sequences abundance in prokaryotic genomes. Mol Biol Evol 24:969–981

Trujillo ME, Kroppenstedt RM, Schumann P, Carro L, Martanez-Molina E (2006) *Micromonospora coriariae* sp. nov., isolated from root nodules of *Coriaria myrtifolia*. Int J Syst Evol Microbiol 56: 2381–2385

Valdes M, Parez N-O, Estrada-de Los Santos P, Caballero-Mellado J, Pe-Cabriales JJ, Normand P, Hirsch AM (2005) Non-frankia actinomycetes isolated from surface-sterilized roots of *Casuarina equisetifolia* fix nitrogen. Appl Environ Microbiol 71:460–466

Vesth T, Lagesen K, Ãn A, Ussery D (2013) CMG-biotools, a free workbench for basic comparative microbial genomics. PLoS One 8: e60120

Wall LG (2000) The actinorhizal symbiosis. J Plant Growth Regul 19: 167–182

Wang J et al. (2010) *Prauserella marina* sp. nov., isolated from ocean sediment of the South China Sea. Int J Syst Evol Microbiol 60: 985–989

Wheeler CT, Miller IM (1990) Current and potential uses of actinorhizal plants in Europe. The Biology of Frankia and Actinorhizal Plants 365:389

Nucleotide

GenBank

Camellia sp. RL-2020 isolate Thurbo 9 maturase K (matK) gene, partial cds; chloroplast

GenBank: MN480322.1

FASTA Graphics

Go to:

```
LOCUS       MN480322                707 bp    DNA     linear   PLN 19-MAY-2020
DEFINITION  Camellia sp. RL-2020 isolate Thurbo 9 maturase K (matK) gene,
            partial cds; chloroplast.
ACCESSION   MN480322
VERSION     MN480322.1
KEYWORDS    .
SOURCE      chloroplast Camellia sp. RL-2020
  ORGANISM  Camellia sp. RL-2020
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
            Pentapetalae; asterids; Ericales; Theaceae; Camellia; unclassified
            Camellia.
REFERENCE   1  (bases 1 to 707)
  AUTHORS   Labar,R., Bhattacharya,M., Kar,P., Biswas,P. and Sen,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2019) Molecular Cytogenetics Laboratory,
            Department of Botany, University of North Bengal, Rajarammohunpur,
            Siliguri, West Bengal 734013, India
FEATURES             Location/Qualifiers
     source          1..707
                     /organism="Camellia sp. RL-2020"
                     /organelle="plastid:chloroplast"
                     /mol_type="genomic DNA"
                     /isolate="Thurbo 9"
                     /db_xref="taxon:2734848"
                     /tissue_type="leaf"
     gene            <1..>707
                     /gene="matK"
     CDS             <1..>707
                     /gene="matK"
                     /codon_start=2
                     /transl_table=11
                     /product="maturase K"
                     /protein_id="QJT93349.1"
                     /translation="VLNYVLDILIPHSIHLEISVQTLRYWVKDASSLYLLRFFLHMYW
                     NWNSLITPKKSSFYFSKRNQRLFLFLYNFHICEYESIFVFLRKQSSHLRSISSGTFLE
                     RRYFYGKIEHFLEVFTKDFQVILWLFKDPFIHYVRYQGKCILASKGTSLLMNKWKSYL
                     VNFWQCYFYMWSQPGRIHINQLSKHSPDFLGYLSSVRLNPSMVRSQMLENSFLIGNAI
                     KKFDTIVPIIPMIGSLS"
ORIGIN
        1 cgtgctaaat tatgtgttag atatactaat accccactcc atccatctgg aaatctcggt
       61 tcaaacgctt cgttactggg taaaagatgc ctcttctttg tatttattac gattctttct
      121 ccatatgtat tggaattgga atagtcttat tactccaaag aaatctagtt tttatttttc
      181 aaaaagaaat caaagattat tcttgtttct atataatttt catatatgtg aatacgaatc
      241 catcttcgtt tttctccgca agcaatcttc tcatttacga tcaatatctt ctggaacctt
      301 tctcgaacga agatatttct atggaaaaat agaacatttt ttagaagtct ttactaagga
      361 ttttcaggtc attctgtggt tgttcaagga tccttttatt cattatgtta ggtatcaagg
      421 aaaatgcatt ttggcttcaa aaggtacgtc tcttttgatg aataaatgga aatcttacct
      481 tgtcaatttc tggcaatgtt atttttacat gtggtctcaa ccaggaagga tccatataaa
      541 ccaattatct aagcattccc cagactttct gggctatctt tcaagtgtgc gactaaaccc
      601 ttcaatggtg cggagtcaaa tgctagaaaa ttcatttcta ataggtaatg ctattaagaa
      661 gttcgatacc atagttccga ttattcctat gattggatca ttgtcta
//
```

Nucleotide

GenBank

Camellia sp. RL-2020 isolate Thurbo 3 maturase K (matK) gene, partial cds; chloroplast

GenBank: MN480321.1

FASTA Graphics

Go to:

```
LOCUS       MN480321                761 bp    DNA     linear   PLN 19-MAY-2020
DEFINITION  Camellia sp. RL-2020 isolate Thurbo 3 maturase K (matK) gene,
            partial cds; chloroplast.
ACCESSION   MN480321
VERSION     MN480321.1
KEYWORDS    .
SOURCE      chloroplast Camellia sp. RL-2020
  ORGANISM  Camellia sp. RL-2020
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
            Pentapetalae; asterids; Ericales; Theaceae; Camellia; unclassified
            Camellia.
REFERENCE   1  (bases 1 to 761)
  AUTHORS   Labar,R., Bhattacharya,M., Kar,P., Biswas,P. and Sen,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2019) Molecular Cytogenetics Laboratory,
            Department of Botany, University of North Bengal, Rajarammohunpur,
            Siliguri, West Bengal 734013, India
FEATURES             Location/Qualifiers
     source          1..761
                     /organism="Camellia sp. RL-2020"
                     /organelle="plastid:chloroplast"
                     /mol_type="genomic DNA"
                     /isolate="Thurbo 3"
                     /db_xref="taxon:2734848"
                     /tissue_type="leaf"
     gene            <1..>761
                     /gene="matK"
     CDS             <1..>761
                     /gene="matK"
                     /codon_start=3
                     /transl_table=11
                     /product="maturase K"
                     /protein_id="QJT93348.1"
                     /translation="LNYVLDILIPHSIHLEISVQTLRYWVKDASSLYLLRFFLHMYWN
                     WNSLITPKKSSFYFSKRNQRLFLFLYNFHICEYESIFVFLRKQSSHLRSISSGTFLER
                     RYFYGKIEHFLEVFTKDFRVILWLFKDPFIHYVRYQGKYILASKGTSLLMNKWKSYLV
                     NFWQCYFYMWSQPGRIHINQLSKHSPDFLGYLSSVRLNPSMVRSQMLENSFLIGNAIK
                     KFDTIVPIIPMIGSLSKAKFCNVLGHPISKPVWAD"
ORIGIN
        1 atctaaatta tgtgttagat atactaatac cccactccat ccatctggaa atctcggttc
       61 aaactcttcg ttactgggta aaagatgcct cttctttgta tttattacga ttctttctcc
      121 atatgtattg gaattggaat agtcttatta ctccaaagaa atctagtttt tatttttcaa
      181 aaagaaatca aagattattc ttgtttctat ataatttttca tatatgtgaa tacgaatcca
      241 tcttcgtttt tctccgcaag caatcttctc atttacgatc aatatcttct ggaacctttc
      301 tcgaacgaag atatttctat ggaaaaatag aacatttttt agaagtcttt actaaggatt
      361 ttcgggtcat tctgtggttg ttcaaggatc cttttattca ttatgttagg tatcaaggaa
      421 aatacatttt ggcttcaaaa ggtacgtctc ttttgatgaa taaatggaaa tcttaccttg
      481 tcaatttctg gcaatgttat ttttacatgt ggtctcaacc aggaaggatc catataaacc
      541 aattatctaa gcattcccca gactttctgg gctatctttc aagtgtgcga ctaaaccctt
      601 caatggtacg gagtcaaatg ctagaaaatt catttctaat aggtaatgct attaagaagt
      661 tcgataccat agttccgatt attcctatga ttggatcatt gtctaaagcg aaattttgta
      721 acgtgttagg acatcccatt agtaaaccag tctgggccga t
//
```

Nucleotide

GenBank

Camellia sinensis clone Takdah 383 maturase K (matK) gene, partial cds; chloroplast

GenBank: MK424872.1

FASTA Graphics

Go to:

```
LOCUS       MK424872                659 bp    DNA     linear   PLN 21-AUG-2019
DEFINITION  Camellia sinensis clone Takdah 383 maturase K (matK) gene, partial
            cds; chloroplast.
ACCESSION   MK424872
VERSION     MK424872.1
KEYWORDS    .
SOURCE      chloroplast Camellia sinensis
  ORGANISM  Camellia sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
            Pentapetalae; asterids; Ericales; Theaceae; Camellia.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Labar,R., Bhattacharya,M., Kar,P., Biswas,P. and Sen,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-2019) Molecular Cytogenetics Laboratory,
            Department of Botany, University of North Bengal, Rajarammohunpur,
            Siliguri, West Bengal 734013, India
FEATURES             Location/Qualifiers
     source          1..659
                     /organism="Camellia sinensis"
                     /organelle="plastid:chloroplast"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4442"
                     /clone="Takdah 383"
                     /tissue_type="leaf"
     gene            <1..>659
                     /gene="matK"
     CDS             <1..>659
                     /gene="matK"
                     /codon_start=2
                     /transl_table=11
                     /product="maturase K"
                     /protein_id="QEE04476.1"
                     /translation="KRNQRLFLFLYNFHICEYESIFVFLRKQSSHLRSISSGTFLERR
                     YFYGKIEHFLEVFTKDFQVILWLFKDPFIHYVRYQGKYILASKGTSLLMNKWKSYLVN
                     FWQCYFYMWSQPGRIHINQLSKHSPDFLGYLSSVRLNPSMVRSQMLENSFLIGNAIKR
                     FDTIVPIIPMIGSLSKAKFCNVLGHPISKPVWADLSDSDIIDRFGRIYRNLSHYHSGS
                     S"
ORIGIN
        1 aaaaagaaat caaagattat tcttgtttct atataatttt catatatgtg aatacgaatc
       61 catcttcgtt tttctccgca agcaatcttc tcatttacga tcaatatctt ctggaacctt
      121 tctcgaacga agatatttct atggaaaaat agaacatttt ttagaagtct ttactaagga
      181 ttttcaggtc attctgtggt tgttcaagga tccttttatt cattatgtta ggtatcaagg
      241 aaaatacatt ttggcttcaa aaggtacgtc tcttttgatg aataaatgga aatcttacct
      301 tgtcaatttc tggcaatgtt atttttacat gtggtctcaa ccaggaagga tccatataaa
      361 ccaattatct aagcattccc cagactttct gggctatctt tcaagtgtgc gactaaaccc
      421 ttcaatggta cggagtcaaa tgctagaaaa ttcatttcta ataggtaatg ctattaagag
      481 gttcgatacc atagttccta ttattcctat gattggatca ttgtctaaag cgaaattttg
      541 taacgtgtta ggacatccca ttagtaaacc agtctgggcc gatttatcag attctgatat
      601 tattgaccga ttcgggcgta tatatagaaa tctttctcat tatcatagcg gatcctcaa
//
```

Nucleotide

GenBank

Camellia sinensis clone Bannockburn 157 maturase K (matK) gene, partial cds; chloroplast

GenBank: MK424871.1

FASTA Graphics

Go to:

```
LOCUS       MK424871                735 bp    DNA     linear   PLN 21-AUG-2019
DEFINITION  Camellia sinensis clone Bannockburn 157 maturase K (matK) gene,
            partial cds; chloroplast.
ACCESSION   MK424871
VERSION     MK424871.1
KEYWORDS    .
SOURCE      chloroplast Camellia sinensis
  ORGANISM  Camellia sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
            Pentapetalae; asterids; Ericales; Theaceae; Camellia.
REFERENCE   1  (bases 1 to 735)
  AUTHORS   Bhattacharya,M., Labar,R., Kar,P., Biswas,P. and Sen,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-2019) Molecular Cytogenetics Laboratory,
            Department of Botany, University of North Bengal, Rajarammohunpur,
            Siliguri, West Bengal 734013, India
FEATURES             Location/Qualifiers
     source          1..735
                     /organism="Camellia sinensis"
                     /organelle="plastid:chloroplast"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4442"
                     /clone="Bannockburn 157"
                     /tissue_type="leaf"
     gene            <1..>735
                     /gene="matK"
     CDS             <1..>735
                     /gene="matK"
                     /codon_start=2
                     /transl_table=11
                     /product="maturase K"
                     /protein_id="QEE04475.1"
                     /translation="LITPKKSSFYFSKRNQRLFLFLYNFHICEYESIFVFLRKQSSHL
                     RSISSGTFLERRYFYGKIEHFLEVFTKDFQVILWLFKDPFIHYVRYQGKYILASKGTS
                     LLMNKWKSYLVNFWQCYFYMWSQPGRIHINQLSKHSPDFLGYLSSVRLNPSMVRSQML
                     ENSFLIGNAIKRFDTIVPIIPMIGSLSKAKFCNVLGHPISKPVWADLSDSDIIDRFGR
                     IYRNLSHYHSGSSKKTSLYRIKYILRL"
ORIGIN
        1 tcttattact ccaaagaaat ctagttttta tttttcaaaa agaaatcaaa gattattctt
       61 gtttctatat aattttcata tatgtgaata cgaatccatc ttcgtttttc tccgcaagca
      121 atcttctcat ttacgatcaa tatcttctgg aacctttctc gaacgaagat atttctatgg
      181 aaaaatagaa catttttag aagtctttac taaggatttt caggtcattc tgtggttgtt
      241 caaggatcct tttattcatt atgttaggta tcaaggaaaa tacattttgg cttcaaaagg
      301 tacgtctctt ttgatgaata aatggaaatc ttaccttgtc aatttctggc aatgttattt
      361 ttacatgtgg tctcaaccag gaaggatcca tataaaccaa ttatctaagc attccccaga
      421 ctttctgggc tatctttcaa gtgtgcgact aaacccttca atggtacgga gtcaaatgct
      481 agaaaattca tttctaatag gtaatgctat taagaggttc gataccatag ttcctattat
      541 tcctatgatt ggatcattgt ctaaagcgaa attttgtaac gtgttaggac atcccattag
      601 taaaccagtc tgggccgatt tatcagattc tgatattatt gaccgattcg ggcgtatata
      661 tagaaatctt tctcattatc atagcggatc ctcaaaaaaa acgagtttgt atagaataaa
      721 atatatactt cgact
//
```

Nucleotide

GenBank

Camellia sinensis clone Takdah 78 maturase K (matK) gene, partial cds; chloroplast

GenBank: MK424870.1

FASTA Graphics

Go to:

```
LOCUS       MK424870                755 bp    DNA     linear   PLN 21-AUG-2019
DEFINITION  Camellia sinensis clone Takdah 78 maturase K (matK) gene, partial
            cds; chloroplast.
ACCESSION   MK424870
VERSION     MK424870.1
KEYWORDS    .
SOURCE      chloroplast Camellia sinensis
  ORGANISM  Camellia sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
            Pentapetalae; asterids; Ericales; Theaceae; Camellia.
REFERENCE   1  (bases 1 to 755)
  AUTHORS   Bhattacharya,M., Labar,R., Kar,P., Biswas,P. and Sen,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-2019) Molecular Cytogenetics Laboratory,
            Department of Botany, University of North Bengal, Rajarammohunpur,
            Siliguri, West Bengal 734013, India
FEATURES             Location/Qualifiers
     source          1..755
                     /organism="Camellia sinensis"
                     /organelle="plastid:chloroplast"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4442"
                     /clone="Takdah 78"
                     /tissue_type="leaf"
     gene            <1..>755
                     /gene="matK"
     CDS             <1..>755
                     /gene="matK"
                     /codon_start=2
                     /transl_table=11
                     /product="maturase K"
                     /protein_id="QEE04474.1"
                     /translation="YVLDILIPHSIHLEISVQTLRYWVKDASSLYLLRFFLHMYWNWN
                     SLITPKKSSFYFSKRNQRLFLFLYNFHICEYESIFVFLRKQSSHLRSISSGTFLERRY
                     FYGKIEHFLEVFTKDFQVILWLFKDPFIHYVRYQGKYILASKGTSLLMNKWKSYLVNF
                     WQCYFYMWSQPGRIHINQLSKHSPDFLGYLSSVRLNPSMVRSQMLENSFLIGNAIKRF
                     DTIVPIIPMIGSLSKAKFCNVLGHPISKPVWAD"
ORIGIN
        1 ttatgtgtta gatatactaa taccccactc catccatctg gaaatctcgg ttcaaactct
       61 tcgttactgg gtaaaagatg cctcttcttt gtatttatta cgattctttc tccatatgta
      121 ttggaattgg aatagtctta ttactccaaa gaaatctagt ttttatttt caaaaagaaa
      181 tcaaagatta ttcttgtttc tatataattt tcatatatgt gaatacgaat ccatcttcgt
      241 ttttctccgc aagcaatctt ctcatttacg atcaatatct tctggaacct ttctcgaacg
      301 aagatatttc tatggaaaaa tagaacattt tttagaagtc tttactaagg attttcaggt
      361 cattctgtgg ttgttcaagg atccttttat tcattatgtt aggtatcaag gaaaatacat
      421 tttggcttca aaaggtacgt ctcttttgat gaataaatgg aaatcttacc ttgtcaattt
      481 ctggcaatgt tatttttaca tgtggtctca accaggaagg atccatataa accaattatc
      541 taagcattcc ccagactttc tgggctatct ttcaagtgtg cgactaaacc cttcaatggt
      601 acggagtcaa atgctagaaa attcatttct aataggtaat gctattaaga ggttcgatac
      661 catagttcct attattccta tgattggatc attgtctaaa gcgaaatttt gtaacgtgtt
      721 aggacatccc attagtaaac cagtctgggc cgatt
//
```

Nucleotide

GenBank

Camellia sinensis clone Tocklai Variety 19 maturase K (matK) gene, partial cds; chloroplast

GenBank: MK424868.1

FASTA Graphics

Go to:

```
LOCUS       MK424868                750 bp    DNA     linear   PLN 21-AUG-2019
DEFINITION  Camellia sinensis clone Tocklai Variety 19 maturase K (matK) gene,
            partial cds; chloroplast.
ACCESSION   MK424868
VERSION     MK424868.1
KEYWORDS    .
SOURCE      chloroplast Camellia sinensis
  ORGANISM  Camellia sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
            Pentapetalae; asterids; Ericales; Theaceae; Camellia.
REFERENCE   1  (bases 1 to 750)
  AUTHORS   Bhattacharya,M., Labar,R., Kar,P., Biswas,P. and Sen,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-2019) Molecular Cytogenetics Laboratory,
            Department of Botany, University of North Bengal, Rajarammohunpur,
            Siliguri, West Bengal 734013, India
FEATURES             Location/Qualifiers
     source          1..750
                     /organism="Camellia sinensis"
                     /organelle="plastid:chloroplast"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4442"
                     /clone="Tocklai Variety 19"
                     /tissue_type="leaf"
     gene            <1..>750
                     /gene="matK"
     CDS             <1..>750
                     /gene="matK"
                     /codon_start=1
                     /transl_table=11
                     /product="maturase K"
                     /protein_id="QEE04472.1"
                     /translation="VLDILIPHSIHLEISVQTLRYWVKDASSLYLLRFFLHMYWNWNS
                     LITPKKSSFYFSKRNQRLFLFLYNFHICEYESIFVFLRKQSSHLRSISSGTFLERRYF
                     YGKIEHFLEVFTKDFQVILWLFKDPFIHYVRYQGKYILASKGTSLLMNKWKSYLVNFW
                     QCYFYMWSQPGRIHINQLSKHSPDFLGYLSSVRLNPSMVRSQMLENSFLIGNAIKRFD
                     TIVPIIPMIGSLSKAKFCNVLGHPISKPVWAD"
ORIGIN
        1 gtgttagata tactaatacc ccactccatc catctggaaa tctcggttca aactcttcgt
       61 tactgggtaa aagatgcctc ttctttgtat ttattacgat tctttctcca tatgtattgg
      121 aattggaata gtcttattac tccaaagaaa tctagttttt atttttcaaa aagaaatcaa
      181 agattattct tgtttctata taattttcat atatgtgaat acgaatccat cttcgttttt
      241 ctccgcaagc aatcttctca tttacgatca atatcttctg gaacctttct cgaacgaaga
      301 tatttctatg gaaaaataga acatttttta gaagtcttta ctaaggattt tcaggtcatt
      361 ctgtggttgt tcaaggatcc ttttattcat tatgttaggt atcaaggaaa atacattttg
      421 gcttcaaaag gtacgtctct tttgatgaat aaatggaaat cttacctttgt caatttctgg
      481 caatgttatt tttacatgtg gtctcaacca ggaaggatcc atataaacca attatctaag
      541 cattccccag actttctggg ctatctttca agtgtgcgac taaacccttc aatggtacgg
      601 agtcaaatgc tagaaaattc atttctaata ggtaatgcta ttaagaggtt cgataccata
      661 gttcctatta ttcctatgat tggatcattg tctaaagcga aattttgtaa cgtgttagga
      721 catcccatta gtaaaccagt ctgggccgat
//
```

Nucleotide

GenBank

Camellia sinensis clone Phoobsering 1258 maturase K (matK) gene, partial cds; chloroplast

GenBank: MK393404.1

FASTA Graphics

Go to:

```
LOCUS       MK393404                751 bp    DNA     linear   PLN 21-AUG-2019
DEFINITION  Camellia sinensis clone Phoobsering 1258 maturase K (matK) gene,
            partial cds; chloroplast.
ACCESSION   MK393404
VERSION     MK393404.1
KEYWORDS    .
SOURCE      chloroplast Camellia sinensis
  ORGANISM  Camellia sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
            Pentapetalae; asterids; Ericales; Theaceae; Camellia.
REFERENCE   1  (bases 1 to 751)
  AUTHORS   Kar,P., Labar,R., Bhattacharya,M., Biswas,P. and Sen,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2019) Molecular Cytogenetics Laboratory,
            Department of Botany, University of North Bengal, Rajarammohunpur,
            Siliguri, West Bengal 734013, India
FEATURES             Location/Qualifiers
     source          1..751
                     /organism="Camellia sinensis"
                     /organelle="plastid:chloroplast"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4442"
                     /clone="Phoobsering 1258"
                     /tissue_type="leaf"
     gene            <1..>751
                     /gene="matK"
     CDS             <1..>751
                     /gene="matK"
                     /codon_start=1
                     /transl_table=11
                     /product="maturase K"
                     /protein_id="QEE04467.1"
                     /translation="VLDILIPHSIHLEISVQTLRYWVKDASSLYLLRFFLHMYWNWNS
                     LITPKKSSFYFSKRNQRLFLFLYNFHICEYESIFVFLRKQSSHLRSISSGTFLERRYF
                     YGKIEHFLEVFTKDFQVILWLFKDPFIHYVRYQGKYILASKGTSLLMNKWKSYLVNFW
                     QCYFYMWSQPGRIHINQLSKHSPDFLGYLSSVRLNPSMVRSQMLENSFLIGNAIKRFD
                     TIVPIIPMIGSLSKAKFCNVLGHPISKPVWAD"
ORIGIN
        1 gtgttagata tactaatacc ccactccatc catctggaaa tctcggttca aactcttcgt
       61 tactgggtaa aagatgcctc ttctttgtat ttattacgat tctttctcca tatgtattgg
      121 aattggaata gtcttattac tccaaagaaa tctagttttt atttttcaaa aagaaatcaa
      181 agattattct tgtttctata taattttcat atatgtgaat acgaatccat cttcgttttt
      241 ctccgcaagc aatcttctca tttacgatca atatcttctg gaacctttct cgaacgaaga
      301 tatttctatg gaaaaataga acatttttta gaagtcttta ctaaggattt tcaggtcatt
      361 ctgtggttgt tcaaggatcc ttttattcat tatgttaggt atcaaggaaa atacattttg
      421 gcttcaaaag gtacgtctct tttgatgaat aaatggaaat cttacctttgt caatttctgg
      481 caatgttatt tttacatgtg gtctcaacca ggaaggatcc atataaacca attatctaag
      541 cattccccag actttctggg ctatctttca agtgtgcgac taaacccttc aatggtacgg
      601 agtcaaatgc tagaaaattc atttctaata ggtaatgcta ttaagaggtt cgataccata
      661 gttcctatta ttcctatgat tggatcattg tctaaagcga aattttgtaa cgtgttagga
      721 catcccatta gtaaaccagt ctgggccgat a
//
```

26th West Bengal State Science & Technology Congress, 2019

Organised by

Department of Science & Technology and Biotechnology,
Government of West Bengal

The Paper titled Phytochemical and antimicrobial screening of selected Darjeeling based On varying Polarities of extracting solvents

authored by Reha Labar, Malay Bhattacharya, Annab Sen

presented by Prof./Dr./Smt./Sri Reha Labar.

is awarded as 'Outstanding Paper' in the 26th West Bengal State Science and Technology Congress, 2019 held on 28th February / 1st March, 2019, at Science City, Kolkata.

Dr. A. R. Ghosh
Senior Scientist
& Organizing Secretary

Dr. S. Kishore, IAS
Principal Secretary

Bratya Basu
Minister-In-Charge

UNIVERSITY OF NORTH BENGAL

Accredited by NAAC with Grade A

International Seminar on "Frontiers in Tea Research-2020"

March 6, 2020

Organized by

DEPARTMENT OF TEA SCIENCE

CERTIFICATE

This is to certify that Prof./Dr./Mr./Mrs/Miss. REHA LABAR

of Dept. of Botany, NBU has delivered an invited Lecture/Oral presentation/Poster presentation/participated in the one day International Seminar on "Frontiers in Tea Research-2020" held at the University of North Bengal on 6th March, 2020.

He/She has presented a paper titled "DNA barcode profilling..... consensus region"

Prof. Pranab Ghosh
Chairman
Organizing Committee

Dr. Chandra Ghosh
Head, Department of Tea Science
Convener, Organizing Committee

Dr. M. Bhattacharya & Dr. Sonali Ray
Conveners
Organizing Committee

3rd Regional Science & Technology Congress - 2018

West Bengal, (Northern Region)

Organised jointly by

Jalpaiguri Government Engineering College

and

Department of Science and Technology and Biotechnology, Government of West Bengal

Certificate of Outstanding Paper

This is to certify that the Paper titled Phytochemical and antimicrobial Screening of selected Darjeeling tea clones based on varying polarities of extracting Solvents.

Authored by Neha Labar has been adjudged as Outstanding Paper in the session of Botany presented in the 3rd Regional Science & Technology Congress - 2018, West Bengal (Northern Region) held on 12th & 13th December, 2018 at Jalpaiguri Government Engineering College, Jalpaiguri, West Bengal, India.

Dr. Bimalesh Samanta
Senior Scientist
Dept. of Science and Technology
and Biotechnology, Government of West Bengal
and Nodal Officer & Joint Organising Secretary
3rd Regional Science and
Technology Congress
[Northern Region]

Dr. Amitava Ray
Principal
Jalpaiguri Govt. Engg. College

Prof. Goutam Kumar Panda
Organising Secretary
3rd Regional Science and
Technology Congress
[Northern Region]

ENLIGHTENMENT TO PERFECTION

NATIONAL SEMINAR ON
NEW HORIZON IN
BOTANICAL RESEARCH

Botsem 2020

FEBRUARY 20-21, 2020

ORGANISED BY

DEPARTMENT OF BOTANY
UNIVERSITY OF NORTH BENGAL
Accredited by NAAC with Grade A

CERTIFICATE

This is to certify that Prof. / Dr. / Mr. / Ms Reha Lahar

of Botany *participated /chaired a session/*

presented a paper in oral/~~poster session~~/entitled Effect of solvent with varying polarities on phytochemical extraction ..... in - silico approaches *in the above seminar.*

Prof. Aniruddha Saha
Organizing Secretary
BOTSEM- 2020

Prof. Subhas Ch. Roy
Chairman
BOTSEM-2020