Identification of Novel QTL for Black Tea Quality Traits and Drought Tolerance in Tea Plants (*Camellia sinensis*)

Robert. K. Koech[1, 2], Pelly M. Malebe[1], Christopher Nyarukowa[1], Richard Mose[3], Samson M. Kamunya[2] and Zeno Apostolides[1*]

[1]Department of Biochemistry, University of Pretoria, Pretoria 0002, South Africa
[2]Kenya Agriculture and Livestock Research Organization, Tea Research Institute, P.O. Box 820, Kericho 20200, Kenya
[3]James Finlay (Kenya) Limited, P.O. Box 223, Kericho 20200, Kenya.
[*]Corresponding author: za@up.ac.za

Abstract

Tea (*Camellia sinensis*) contain polyphenols and caffeine which have been found to be of popular interest in tea quality. Tea production relies on well distributed rainfall which influence tea quality. Phenotypic data for two segregating tea populations TRFK St 504 and TRFK St 524, were collected and used to identify the quantitative trait loci (QTL) influencing tea biochemical and drought stress traits based on a consensus genetic map constructed using the DArTseq platform. The populations comprised 261 F_1 clonal progeny. The map consisted of 15 linkage groups which corresponds to chromosome haploid number of tea plant (2n = 2x = 30) and spanned 1260.1 cM with a mean interval of 1.1 cM between markers. A total of 16 phenotypic traits were assessed in the two populations. Both interval and multiple QTL mapping revealed a total of 47 putative QTL in the 15 LGs associated with tea quality and percent relative water content at a significant genome-wide threshold of 5%. In total, six caffeine QTL, 25 catechins QTL, three theaflavins QTL, nine QTL for tea taster score and three QTL for percent relative water contents were detected. Out of these 47 QTL, 19 QTL were identified for ten traits in three main regions on LG01, LG02, LG04, LG12, LG13 and LG14. The QTL associated with caffeine, individual catechins, and theaflavins were clustered mostly in LG02 and LG04 but in different regions on the map. The explained variance by each QTL in the population ranged from 5.5 to 56.6%, with an average of 9.9%. Identification of QTL that are tightly linked to markers associated with black tea quality coupled with UPLC assay may greatly accelerate development of novel tea cultivars owing to its amenability at seedling stage. In addition, validated molecular markers will contribute greatly to adoption of marker-assisted selection (MAS) for drought tolerance and tea quality improvement.

Keywords: *Camellia sinensis*; Tea quality; Drought tolerance; UPLC; DArTseq; Linkage map

Abbreviations

DArT-	Diversity Arrays Technology
MAS-	Marker-assisted selection
QTL-	Quantitative trait loci
LOD-	Logarithm of odds
LGs-	Linkage groups
UPLC-	Ultra-performance liquid chromatography
CAFF-	Caffeine
CAT-	(+)-catechin
EC-	(-)-epicatechin
ECG-	(-)-epicatechin gallate

EGC- (-)-epigallocatechin
EGCG- (-)-epigallocatechin gallate
TF1- Simple theaflavin
TF2- Theaflavin -3-monogallate
TF3- Theaflavin-3′-monogallate
TF4- Theaflavin-3, 3′-digallate
CL- Liquor colour
BRT- Liquor brightness
AST- Astringency
BRK- Liquor briskness
AR- Liquor aroma
RWC- Relative water content
St- Stock
TRFK- Tea Research Foundation of Kenya

Introduction

Tea plant (*Camellia sinensis* (L.) O. Kuntze) is the second most popular non-alcoholic beverage consumed in the world after water. The main cultivated taxa comprise of three natural hybrids, viz., *C. sinensis* (L.) O. Kuntze ("China" type), *C. assamica* ssp. Assamica (Masters) Wight ("Assam" jat) and the "Cambod" variety, *C. assamica* ssp. lasiocalyx (Planchon ex Watt) Wight (Banerjee, 1992). The China type has small semi-erect leaves and the Assam type has relatively large, horizontal leaves. The Cambod type is currently considered as a hybrid of *C. sinensis* var. *sinensis* type and *C. sinensis* var. *assamica* type (Meegahakumbura *et al.*, 2016; Wambulwa *et al.*, 2016). Tea is cultivated widely in many Asian and African countries, and it contributes significantly to the local economy. In Kenya, tea is the single largest agribusiness and contributes over 26% of all foreign exchange earnings and over 4% of the gross domestic product (Kenya National Bureau of Statistics, 2012; AFFA, 2013).

Tea being a popular beverage, has attracted the attention of both the consumers and the scientific community because of its multiple health-promoting effects (Preedy, 2012). Tea is known to contain a bioactive compounds, such as the flavonoids, caffeine, L-theanine and γ-aminobutyric acid (GABA) (Lin *et al.*, 2012; da Silva Pinto, 2013; Ma *et al.*, 2014). Catechins, a group of flavonoids, are major constituents in tea and contribute up to 30% of the dry weight of the tender tea shoots (Pang *et al.*, 2013; Ma *et al.*, 2014). In green tea leaves, catechins consist mainly of five flavan-3-ols; (+)-catechin (CAT), (-)-epicatechin (EC), (-)-epicatechin gallate (ECG), (-)-epigallocatechin (EGC) and (-)-epigallocatechin gallate (EGCG) (Ananingsih *et al.*, 2013; Ma *et al.*, 2014). The most abundant catechin is the EGCG which accounts for up to 50-80% of the total catechins (Sang *et al.*, 2011). Theaflavins and thearubigins are the major polyphenols found in black tea formed by the enzymatic oxidation and polymerization of catechins during the so-called "fermentation stage" (Owuor *et al.*, 2006b). The major individual theaflavins in black tea include, simple theaflavin (TF), theaflavin -3-monogallate (TF-3-g), theaflavin-3′-monogallate (TF-3′-g) and theaflavin-3, 3′-digallate (TFdg). Catechins in green tea and theaflavins in black tea have demonstrated to have a wide variety of pharmacological activities, which include; antioxidant, anti-inflammatory, anticancer, antidiabetic and neuroprotective effects (Preedy, 2012).

Tea leaves also contain 2-5% of caffeine and much smaller amounts of theobromine and theophylline (Khan and Mukhtar, 2007). Caffeine is not only known for its stimulatory effect (Sharangi, 2009), but contributes to tea briskness and other taste characteristics, which is considered important parameter for the evaluation of tea quality (Obanda *et al.,* 1997; Nitin Seetohul *et al.,* 2006). Caffeine has also been reported to suppress body weight gain by suppressing food intake and reducing adipose tissue (Kobayashi-Hattori *et al.,* 2005). These health benefits associated with tea intake have led to high consumption of tea products as tea extracts in pharmaceutical, food and beverage and cosmetic industries. In addition, to these health benefits, tea quality is a complex phenomenon and depends mainly on the biochemical components that determine colour, aroma and taste, attached to the tea infusions. As such, tea prices vary greatly depending on the tea quality, which has traditionally been assessed by tea tasters by organoleptic evaluation, to describe various quality attributes of a tea infusion. Furthermore, Kenya tea industry has lately been facing challenges of ever increasing temperatures, prolonged drought periods and changing precipitation patterns especially in the tea growing areas (FAO, 2015).Therefore, this calls for concerted efforts in tea breeding programs to develop tea cultivars with improved biochemical content and drought tolerance.

In a previous study on a method for Short-time Withering Assessment of Probability for Drought Tolerance (SWAPDT) in *Camellia sinensis* validated by targeted metabolomics (Nyarukowa *et al.,* 2016), was developed for predicting the drought tolerance of tea cultivars. The metabolite profiles obtained using the SWAPDT method showed that drought tolerant tea cultivars differed from drought susceptible tea cultivars. The SWAPDT method relied on the percent relative water content (%RWC) of tea leaves after the five-hour withering method. Therefore, the SWAPDT method which was used in the current study provides a basis for selection of new drought tolerant tea cultivars that may lead to improvement of crop productivity, amidst challenges imposed by drought due to climate change.

The tea plant is a woody long-lived perennial tree characterized by a large diploid genome of approximately 3Gb and chromosome number (2n= 2x= 30) that is self-incompatible, highly heterozygous and has a long juvenile period of 4-5 years and about 22-25 years to breed a new cultivar (Chen *et al.,* 2007; Shi *et al.,* 2011). Presently, *Camellia* genus is believed to comprise more than 300 species with *Camellia sinensis* (L.) O. Kuntze as the most important agronomic species (Mondal *et al.,* 2004). The recent discovery of *C. cherryana* in 2012 (Orel and Wilson, 2012) indicates the instability and high out-crossing character or property of the genus. These attributes have led to slow progress in tea breeding; however, tremendous progress has been made regarding the availability of good genetic resources for tea breeding. Recently, Bali *et al.* (2015) constructed a genetic linkage map using the two-way pseudo-testcross approach for mapping drought-tolerant trait of Indian tea cultivars. Construction of a genetic linkage map for tea using pseudo-testcross approach has been used previously (Wachira *et al.,* 1995; Hackett *et al.,* 2000; Kamunya *et al.,* 2010; Taniguchi *et al.,* 2012; Hu *et al.,* 2013; Ma *et al.,* 2014) due to high out-crossing and heterozygous nature of the tea. In spite of all these advances, these genetic linkage maps are mainly based on dominant markers such as random amplified polymorphic DNA (RAPD) or amplified fragment length polymorphism (AFLP) (Wachira *et al.,* 1995; Hackett *et al.,* 2000; Tanaka, 2000; Kamunya *et al.,* 2009; Kamunya *et al.,* 2010) and co-dominant markers such as simple sequence repeats (SSRs) (Hu *et al.,* 2013; Ma *et al.,* 2014; Tan *et al.,* 2016). Also, the mapping populations used in these previous studies are relatively small (52 to 92 F1 progenies),

and a number of gaps that are in between the markers in some of the linkage groups, may decrease the accuracy of detected QTL. Thus, these marker technologies have not fully provided a widely applicable tool that can be used to link genotypes to phenotypes. However, SNP discovery and genotyping based on specific-locus amplified fragment sequencing (SLAF-seq) generated a saturated genetic linkage map length of 3,965 cM, with an average inter-locus distance of 1.0 cM (Ma *et al.,* 2015).

Diversity Arrays Technology (DArT) which involves the isolation and cloning of a random set of DNA fragments from complexity-reduced DNA sample has experienced increased interest in the recent years (Gupta *et al.,* 2013). DArT has provided a standardized high-throughput genotyping, whereby thousands of markers can be readily assayed in parallel in thousands of samples without prior sequence information (Jaccoud *et al.,* 2001; Wittenberg *et al.,* 2005). The DArT array technology has demonstrated excellent performance for phylogenetic and diversity analyses (Steane *et al.,* 2011), genomic selection (Poland *et al.,* 2012) and genetic linkage mapping (Schouten *et al.,* 2012). DArT platform has been applied successfully to construct more dense genetic linkage maps in plants with more complex genomes such as apple (Schouten *et al.,* 2012), strawberry (Sánchez-Sevilla *et al.,* 2015), sugarcane (Racedo *et al.,* 2016), wheat (Zou *et al.,* 2016), citrus (Liu *et al.,* 2016) and trees such as Eucalyptus (Sansaloni *et al.,* 2010; Steane *et al.,* 2011). Therefore, a genetic linkage map based on DArTseq platform could provide a vital tool for relating phenotype to genotype, hence facilitating identification and selection of recombinant parents with desirable attributes. The genetic linkage map will also increase the speed and precision of tea breeding and improvement programmes. The present study constructed a high-density linkage map for tea plant by integrating DArTseq platform and ultra performance liquid chromatography (UPLC) technique for QTL linked to catechins, theaflavins, caffeine content, tea taster scores and percent relative water content for future MAS breeding.

Materials and Methods

Plant Material

Two pseudo-testcross populations consisting of 109 progeny from TRFK St. 504 (TRFK 303/577 x GW Ejulu) and reciprocal cross consisting of 152 progeny coded TRFK St. 524 (GW Ejulu x TRFK 303/577), were developed by crossing two heterozygous parental clones TRFK 303/577 and GW Ejulu. The two parental clones, TRFK 303/577 and GW Ejulu, are of Assamica and China varieties, respectively. The two clones were chosen on the basis of their contrasting attributes. GW Ejulu is a low yielding, high black tea quality and moderate levels of caffeine, but high in total catechins and individual catechins contents. Clone TRFK 303/577, which is an open-pollinated progeny of clone TRFK 6/8, is high yielding, drought tolerant, medium in black tea quality, caffeine and individual catechins.

Description of study site

The cross comprising all above clonal progenies and their parents were established in 2007 at Tea Research Institute (TRI), in Timbilil Estate (Kenya, Kericho county, latitude 0° 22’S, longitude 35° 21’E, altitude 2180 m AMSL). The trial populations are set up in a complete randomized block design with three replications in plots of 15 plants spaced at 0.61 m within rows and 1.22 m between rows. The populations had been receiving 150 Kg N per hectare per year in the form of NPKS 25:5:5:5 compound fertilizer. The trial populations were brought into bearing following the recommended management practices (Anon, 2002; Kamunya *et al.,* 2010).

Sample collection and processing

About 1 Kg of fresh shoots in the form of two leaves and a bud was plucked from heterozygous parents and three independent F_1 progeny in a complete randomized block design between September 2013 and February 2014. The fresh-plucked shoots were placed in appropriately labelled sampling bags for processing at the TRI miniature factory. Five hundred grams of fresh green tea leaves were used to produce green tea by immediately drying using a microwave for 5 minutes, which also deactivated the oxidizing enzyme polyphenol oxidase. The remaining 500 g was used to manufacture black tea using a standard black tea manufacturing procedure by physically withering the tea leaves up to a moisture content of 50-65% for 18 hours. The leaf was passed through the CTC rollers four times to achieve maceration equivalent to that obtained using commercial CTC rollers in black tea manufacture (Owuor and Othieno, 1991). Upon maceration, the tea (dhool) was then aerated for 90 minutes at ambient temperature (22-26 °C) and 100% relative humidity to achieve enzymatic oxidation or so-called "fermentation". A bench top fluid-bed drier system (TeaCraft Ltd) was utilized in firing the tea, initially at 120 °C for 20-25 minutes then brought to 100 °C for 10 minutes. Green and black tea samples were grounded separately using a coffee grinder and stored in sealed silver lined aluminium sachets for UPLC analysis and organoleptic assessment for processed black tea.

Extraction of catechins, caffeine and theaflavins

International Organization for Standardization (ISO) extraction method was used as is described in the ISO14502-2, 2005. Briefly, 0.200 ± 0.001 g of green and black tea samples were weighed out on a Mettler Toledo analytical balance (Microsep, South Africa) and transferred into a 20 ml glass test tube. A five ml volume of 70:30 methanol (Merck, South Africa): water (v/v) at 70 °C was added to each extraction tube, stoppered and vortex mixed for approximately five seconds before placing into a water bath set at 70 °C. The extraction mixture was vortex mixed after five minutes and after ten minutes. The extracts were cooled and centrifuged at 2000 x g using Thermo Scientific Heraeus Labofuge (Sepsci, South Africa) 300 centrifuge for ten minutes, with the resultant supernatant decanted into a ten ml volumetric flask. The extraction step was repeated twice. Both extracts were pooled, and the volume adjusted to ten ml with cold 70:30 methanol: water (v/v). One ml volume of the extracts was then diluted to five mL with stabilizing solution (10 % v/v acetonitrile, 500 mg/mL EDTA and 10 mg/ml ascorbic acid, Sigma-Aldrich, South Africa). The dilution was filtered through 0.2 µm Minisart®RC4 syringe filters (Sartorius, South Africa) with hydrophilic, solvent-resistant regenerated cellulose membranes and ready for UPLC analysis.

UPLC analysis of catechins, caffeine and theaflavins

The UPLC analysis was performed on Waters ACQUITY UPLC H-Class system (Waters, Milford, MA, USA) equipped with a binary solvent delivery pump, an autosampler and a photodiode array detector (PDA) and controlled by the Empower-3 software. Separation was achieved on Waters Acquity HSS T3 column (1.8 µm, 2.1 × 150 mm). The mobile phase was composed of solvent A(2% acetic acid and 9% acetonitrile in water, pH 2.8) and solvent B (2% acetic acid and 80% of acetonitrile in water) with gradient elution: 0 min (5% A), 0-21 min (5-20% A), 21-30 min (20-25% A), 30-32 min (25-100% A), 32-39 min (100-100% A), 39-40 min (100-5% A), and 40-45 min (5-5% A) was used in this study. The mobile phase was filtered through a 0.2 µm cellulose acetate membrane filter and degassed using a Neuberger Laboport (Labotech, South Africa) vacuum pump. The sample injection volume was five µL, and the flow-rate was set at 0.2 mL/min.

The standards consisted of catechins (CAT, EC, ECG, EGC and EGCG), caffeine and gallic acid (Sigma-Aldrich, South Africa). Tryptamine, sulfanilamide and mycophenolic acid (Sigma-Aldrich, South Africa) were used as internal standards for quality control. Three mixed standards; mixture A (catechins, caffeine, gallic acid and three internal standards), mixture B and mixture C which both had all the standards except the three internal standards were used. The identification and quantification of individual catechins, caffeine and individual theaflavins were performed at 278 nm. The individual catechins and caffeine peaks in the tea samples were identified based on retention times of the individual catechins and caffeine standards. The individual theaflavins in black tea samples were quantified as EGCG equivalents, based on the response factor of EGCG.

Determination of plain black tea quality parameters

Tea taster scores

The tea taster's score on black teas was performed on the 261 progenies and their parents. A sample of black tea (2.5 g) was extracted with 150 ml of boiling water for about 5 minutes. The tea sample was then subjected to organoleptic evaluation made by professional tea tasters (WM Cahn, Johannesburg, South Africa) with scores based on liquor colour (CL), liquor brightness (BRT), astringency (AST), liquor briskness (BRK) and liquor aroma (AR). The scores were in the range of 1-5 with the low score representing poor quality and a high score representing high quality. Black teas that are known to be of high quality were used as a control to determine the quality of progenies regarding liquor characteristics and taste.

Determination of percent relative water content.

The % RWC in the tea two populations was determined using the SWAPDT method as described by Nyarukowa *et al.* (2016). Three shoots with two leaves and a bud collected from heterozygous parents and three independent F_1 progeny in a complete randomized block design were immersed in water at room temperature and weighed after 24 hours. The leaves were then oven dried at 37 °C and weighed after five hours. The leaves were again placed in water and weighed after 24 hours, and then oven dried at 105 °C for 24 hours to obtain each leaf's dry weight.

The % RWC was calculated using the formula:

$$\%\ RWC = [(F_{wt} - D_{wt})/(FT_{wt} - D_{wt})]\ x100$$

where F_{wt} is the fresh weight, D_{wt} is the dry weight after 24 h in 105 °C oven, and FT_{wt} is the weight after 24 h rehydration.

Extraction, purification and quantification of DNA

Two tender young leaves and a terminal bud were plucked from a tea bush planted in a complete randomized block design experiments as a representative sample and preserved in zip-lock plastic bags containing dry silica gel (Malebe, 2011). The silica gel was dried in the oven at 70°C for 48 hours to enable it absorb any surface moisture from the preserved leaves. The green tea leaf samples were stored at -20 °C before DNA extraction using a modified method of (Gawal and Jarret, 1991). The NanoDrop spectrophotometer (NanoDropTechnologies, South Africa) which gives the concentration of DNA in ng/µl was used to quantify the DNA. The integrity of DNA was also checked using the agarose gel electrophoresis method (Adkins *et al.,* 2007).

DArTseq assay

The DNA samples were shipped to Diversity Array Technology Pty Ltd (Canberra, Australia) for DArTseq analysis. DNA quality and digestibility were tested using restriction enzyme *Pst*I (Fermentas, Burlington, Canada) and *Eco*RI (Promega, Madison, USA). The DArTseq procedure was carried out as described by (Sansaloni *et al.,* 2010) using *Pst*I and *Mse*I restriction enzymes. Markers were scored 0 or 1 for absence or presence of a polymorphism in the genomic representation of the sample.

Linkage map construction

Six thousand five hundred and eighty eight DArTseq markers derived from C. *sinensis* sequences were tested for segregation using genomic DNA from the two parents and 261 F_1 individuals. The obtained genotyping data was subjected to linkage analyses using JoinMap 4.0 (Van Ooijen, 2006). Markers were grouped at a logarithm of odds (LOD) of (3-12) and a recombination frequency of 0.4. The genetic distances between markers were calculated with the Kosambi mapping function. The LOD thresholds at the genome-wide level were determined by running 1,000 permutations ($P<0.05$).

Phenotypic trait analysis

A total of 16 phenotypic traits were assessed for a normal distribution using the JMP 12 (Sall *et al.,* 2012). The mean and the standard deviation for each phenotypic trait in each parent and the progenies were calculated. The significance difference between parental values and progenies was analyzed using Student's t-test.

QTL analyses

The QTL mapping was carried out using the MapQTL 6.0 software (Van Ooijen, 2006). The first set of markers were selected as cofactors from IM results, and the significant markers were selected using backwards elimination method. The significant markers at $P<0.05$ were only used as cofactors in the multiple MQM method to refine QTL detection. The permutation analysis (1,000 permutations) for significance thresholds was used for accepting the presence of a potential QTL (Churchill and Doerge, 1994). The maximum LOD score on a linkage group was used to estimate the positions of QTL. The location and position of the QTL for individual phenotypic traits was indicated using MapChart software (Voorrips, 2002).

Results

Phenotypic trait analysis

A wide range of variation in individual catechins, theaflavins and caffeine contents was observed in the two parents and F_1 progeny. A chromatogram representing the profiles of black tea and green tea from the two parent lines is presented in Supplementary Fig. 1 and Supplementary Fig. 2, respectively. Besides the five peaks identified, there were other several peaks that were detected but unidentified. This is an indication that there are other several biochemical compounds present in the tea extracts. The mean, standard deviation and coefficient of variation for individual catechins and caffeine content quantified in F1 progeny and in both parents are shown in Table 1 and Table 2. Tea liqour characteristics such as colour, brightness, briskness as well as astringency and aroma were also analyzed in black tea samples.

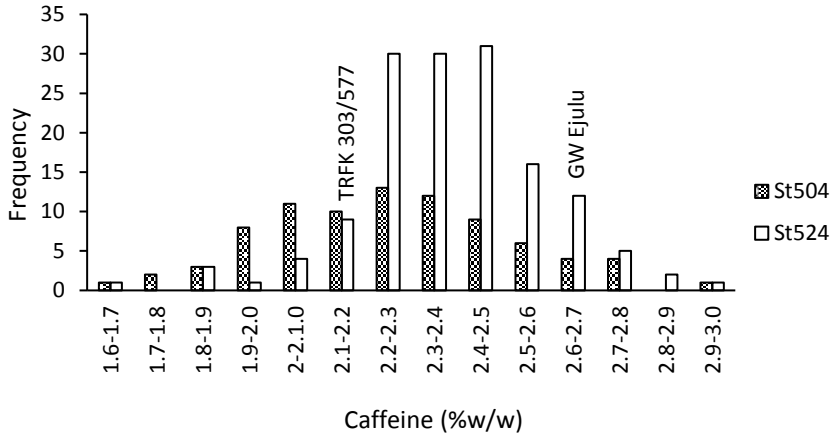
Table 1. Phenotypic variation of caffeine and individual catechins content in parental lines (TRFK 303/577 and GW Ejulu) and 109 F_1 progeny St 504

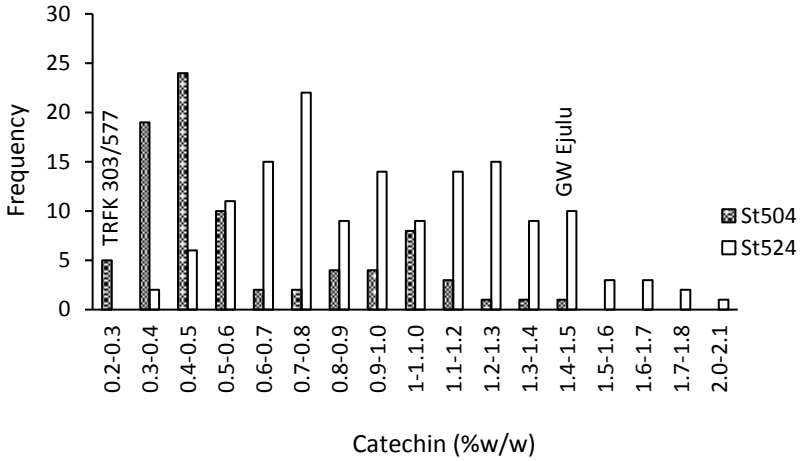
Trait (%w/w)	TRFK 303/577 Mean±SD	GW Ejulu Mean±SD	F_1 Mean±SD	F_1 Range	F_1 CV (%)
Caffeine	2.12±0.11	2.69±0.13	2.25±0.27	1.57-3.02	11.9
CAT	0.44±0.11	1.51±0.17	0.62±0.32	0.17-1.64	53.1
EC	1.22±0.21	0.90±0.07	1.19±0.41	0.51-2.63	34.4
ECG	2.33±0.10	5.21±0.03	3.31±0.83	1.54-6.58	25.6
EGC	3.72±0.72	1.02±0.40	3.25±1.07	0.60-5.79	33.7
EGCG	7.15±0.65	6.50±0.40	6.03±1.13	1.78-12.33	18.9

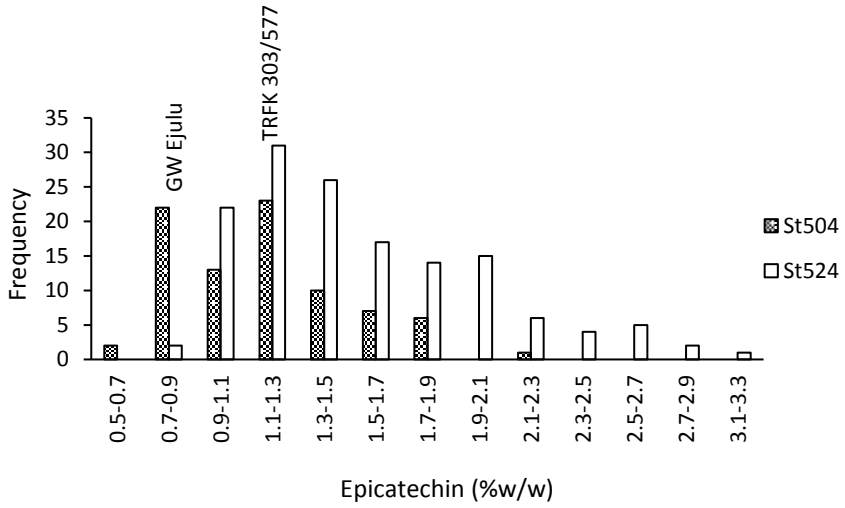
Table 2. Phenotypic variation of caffeine and individual catechins content in parental lines (GW Ejulu and TRFK 303/577) and 152 F_1 progeny St 524

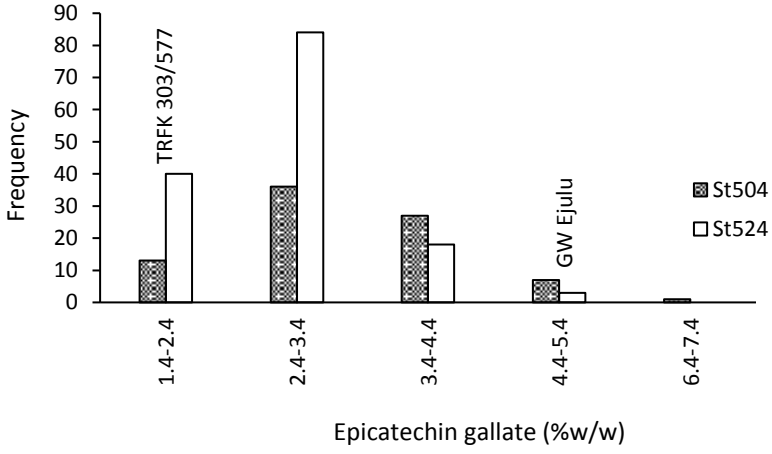
Trait (%w/w)	GW Ejulu Mean±SD	TRFK 303/577 Mean±SD	F_1 Mean±SD	F_1 Range	F_1 CV (%)
Caffeine	2.69±0.13	2.12±0.11	2.38±0.30	1.53-3.69	12.8
EC	0.90±0.07	1.22±0.21	1.57±0.51	0.66-3.40	32.6
ECG	5.21±0.03	2.33±0.10	2.82±0.66	1.06-5.24	24.2
EGC	1.02±0.40	3.72±0.72	3.34±1.05	0.60-7.61	32.3
EGCG	6.50±0.40	7.15±0.65	6.14±1.24	1.78-12.33	20.4

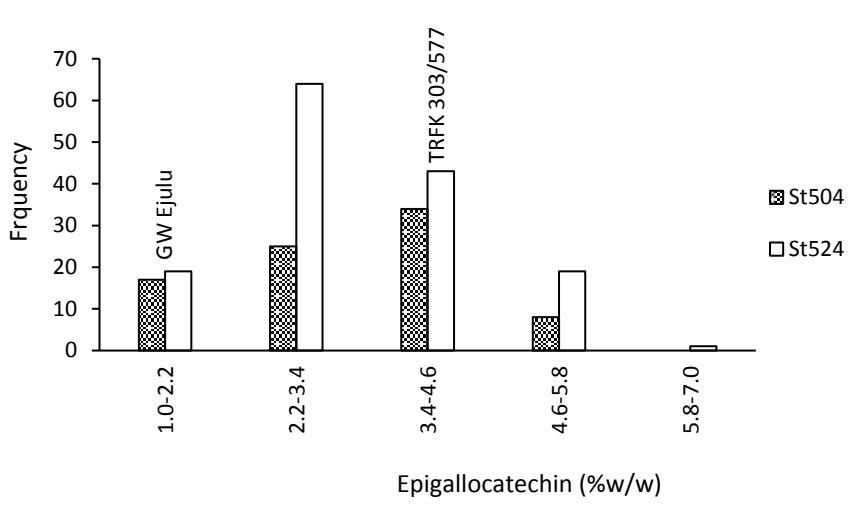
The frequency distributions of phenotype datasets are shown in Fig. 3. The values of the two parents are indicated in the graphs. The levels of the individual catechins analyzed in green teas ranged from 0.1 to 12.3%, with low content of CAT and EC as compared EGCG and EGC, which were the abundant catechins. Caffeine content ranged from 1.5 to 3.7%, while the levels of individual theaflavins in processed black teas ranged from 0.2 to 1.3% with TF3 being the most abundant theaflavin (Fig. 3). A variation in the levels of individual theaflavins content was observed in parental lines and progenies. Parental clone TRFK 303/577 had high levels of all four theaflavins as compared to parental clone GW Ejulu in both tea populations. A number of traits namely caffeine, EGC, TF1, TF2, TF3 and TF4 score as measured in the two populations (St 504 and St 524) had their F_1 means falling in between the two parental means for respective traits (Fig. 3). However, the F_1 means for EC, ECG, EGCG and tea taster scores were marginally lower than the parental means. There were variations in tea taster scores in the parental clones and F_1 progenies. Most of the black tea samples had an average tea quality regarding tea taster's score except for aroma, which scored poorly. This poor aroma score might have been due to the time difference (12 months) between the black tea processing and the tea taster scoring. Tea liqour colour (CL) scored highly than the other tea taster score traits in parental clone GW Ejulu as compared to parental clone TRFK 303/577. There was also variations the percentage relative water content in F_1 progenies and parental clones in the two tea populations (Fig. 3). The % rwc ranged from 49 to 99% which show that there were tea plants that were either tolerant or susceptible to

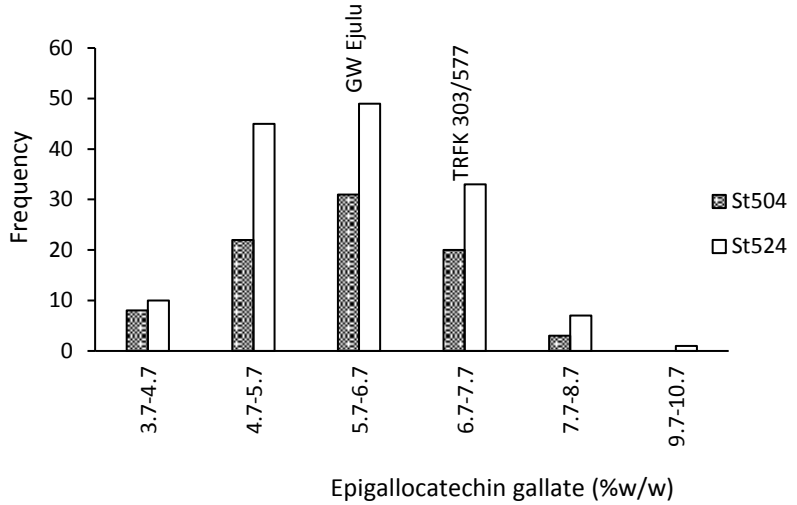

Frequency
35
30
25
20
15
10
5
0
TRFK 303/577
GW Ejulu
St504
St524
1.6-1.7
1.7-1.8
1.8-1.9
1.9-2.0
2-2.1.0
2.1-2.2
2.2-2.3
2.3-2.4
2.4-2.5
2.5-2.6
2.6-2.7
2.7-2.8
2.8-2.9
2.9-3.0
Caffeine (%w/w)

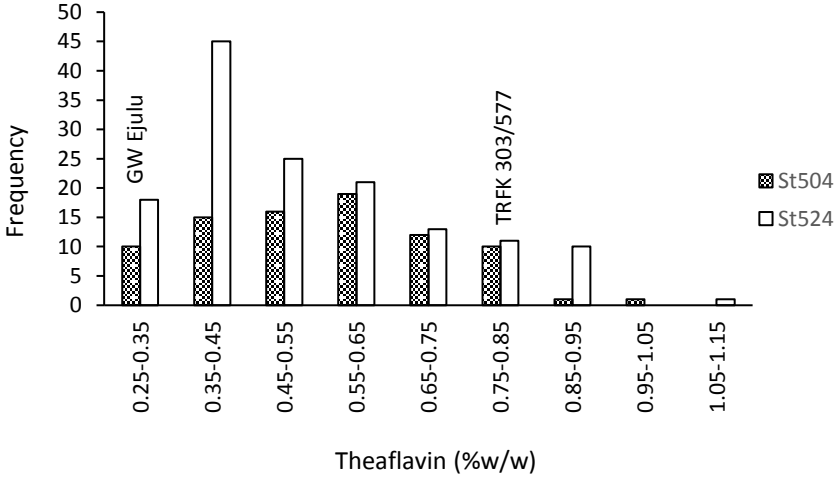

Frequency
30
25
20
15
10
5
0
TRFK 303/577
GW Ejulu
St504
St524
0.2-0.3
0.3-0.4
0.4-0.5
0.5-0.6
0.6-0.7
0.7-0.8
0.8-0.9
0.9-1.0
1-1.1.0
1.1-1.2
1.2-1.3
1.3-1.4
1.4-1.5
1.5-1.6
1.6-1.7
1.7-1.8
2.0-2.1
Catechin (%w/w)

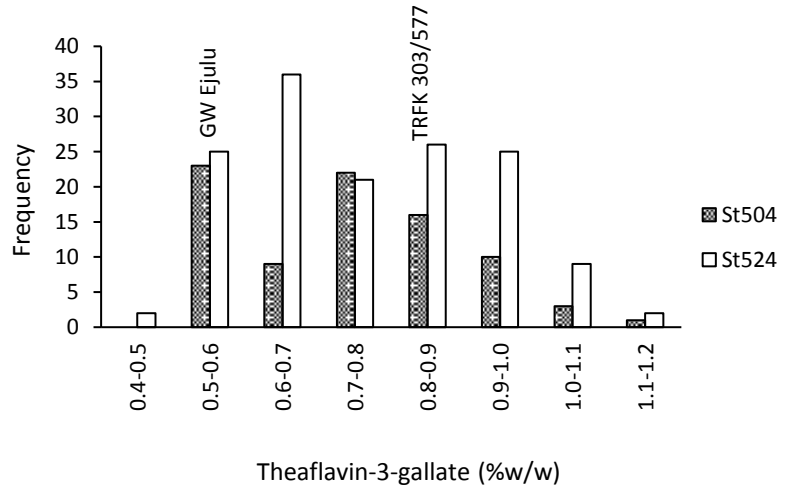

Frequency
35
30
25
20
15
10
5
0
GW Ejulu
TRFK 303/577
St504
St524
0.5-0.7
0.7-0.9
0.9-1.1
1.1-1.3
1.3-1.5
1.5-1.7
1.7-1.9
1.9-2.1
2.1-2.3
2.3-2.5
2.5-2.7
2.7-2.9
3.1-3.3
Epicatechin (%w/w)

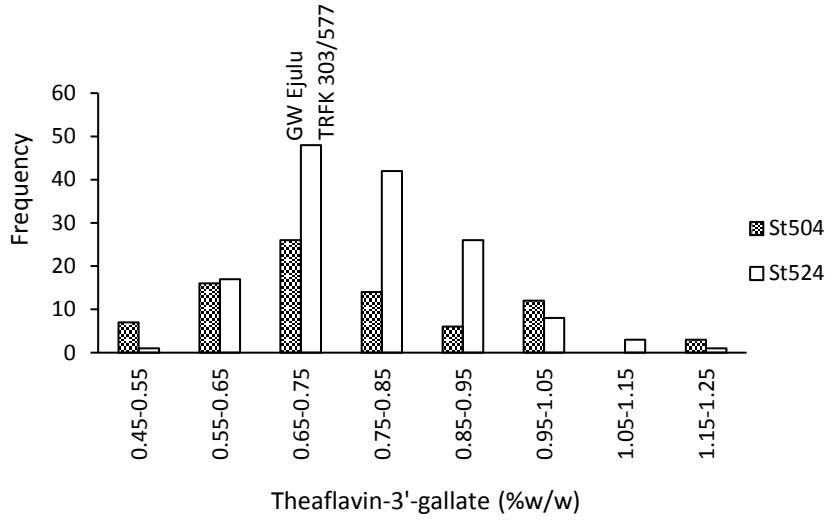

Frequency
90
80
70
60
50
40
30
20
10
0
TRFK 303/577
GW Ejulu
St504
St524
1.4-2.4
2.4-3.4
3.4-4.4
4.4-5.4
6.4-7.4
Epicatechin gallate (%w/w)

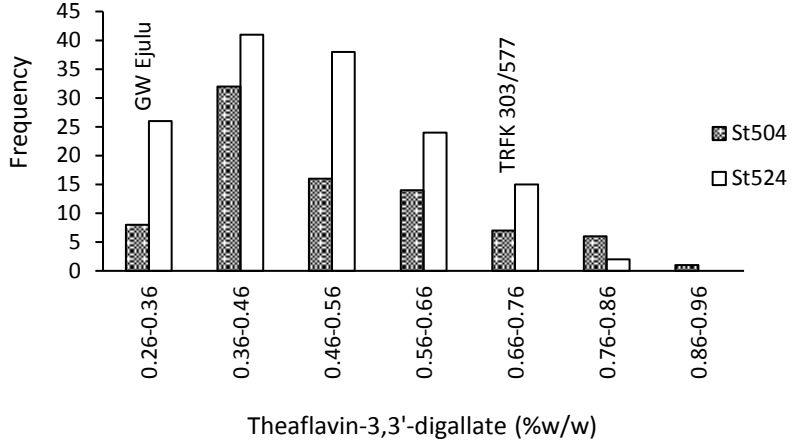

Frquency
70
60
50
40
30
20
10
0
GW Ejulu
TRFK 303/577
St504
St524
1.0-2.2
2.2-3.4
3.4-4.6
4.6-5.8
5.8-7.0
Epigallocatechin (%w/w)

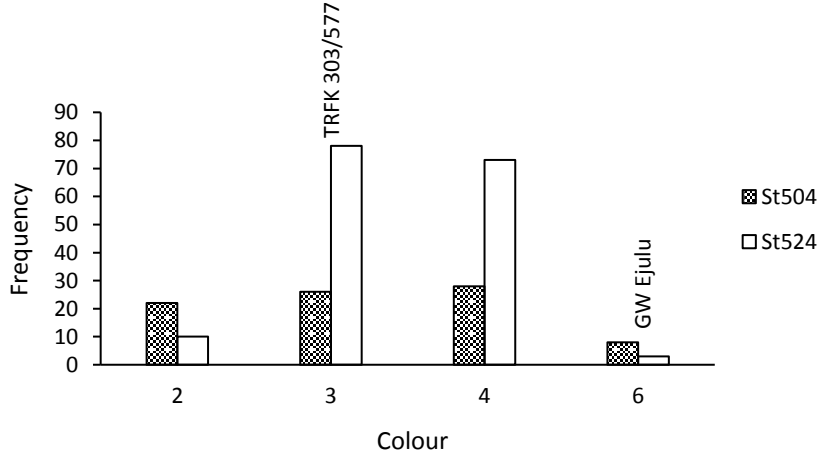

Frequency
60
50
40
30
20
10
0
GW Ejulu
TRFK 303/577
St504
St524
3.7-4.7
4.7-5.7
5.7-6.7
6.7-7.7
7.7-8.7
9.7-10.7
Epigallocatechin gallate (%w/w)

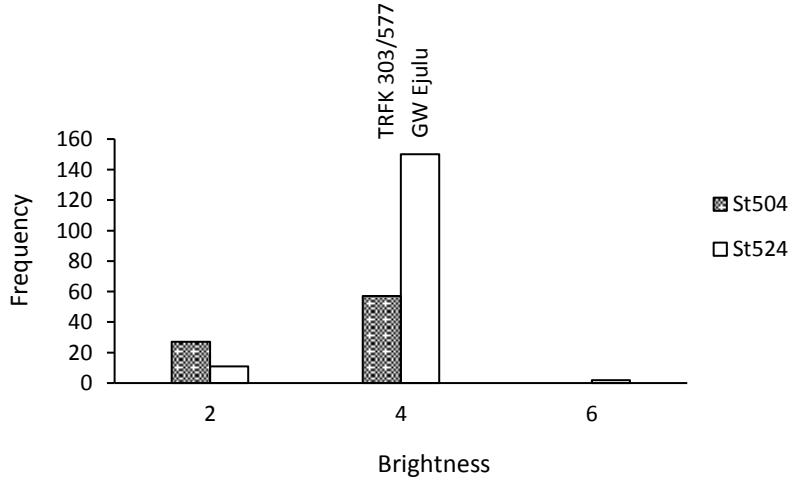

Frequency
50
45
40
35
30
25
20
15
10
5
0
GW Ejulu
TRFK 303/577
St504
St524
0.25-0.35
0.35-0.45
0.45-0.55
0.55-0.65
0.65-0.75
0.75-0.85
0.85-0.95
0.95-1.05
1.05-1.15
Theaflavin (%w/w)


Frequency
40
35
30
25
20
15
10
5
0
GW Ejulu
TRFK 303/577
St504
St524
0.4-0.5
0.5-0.6
0.6-0.7
0.7-0.8
0.8-0.9
0.9-1.0
1.0-1.1
1.1-1.2
Theaflavin-3-gallate (%w/w)


Frequency
60
50
40
30
20
10
0
GW Ejulu
TRFK 303/577
St504
St524
0.45-0.55
0.55-0.65
0.65-0.75
0.75-0.85
0.85-0.95
0.95-1.05
1.05-1.15
1.15-1.25
Theaflavin-3'-gallate (%w/w)


Frequency
45
40
35
30
25
20
15
10
5
0
GW Ejulu
TRFK 303/577
St504
St524
0.26-0.36
0.36-0.46
0.46-0.56
0.56-0.66
0.66-0.76
0.76-0.86
0.86-0.96
Theaflavin-3,3'-digallate (%w/w)


Frequency
90
80
70
60
50
40
30
20
10
0
TRFK 303/577
GW Ejulu
St504
St524
2
3
4
6
Colour


Frequency
160
140
120
100
80
60
40
20
0
TRFK 303/577
GW Ejulu
St504
St524
2
4
6
Brightness

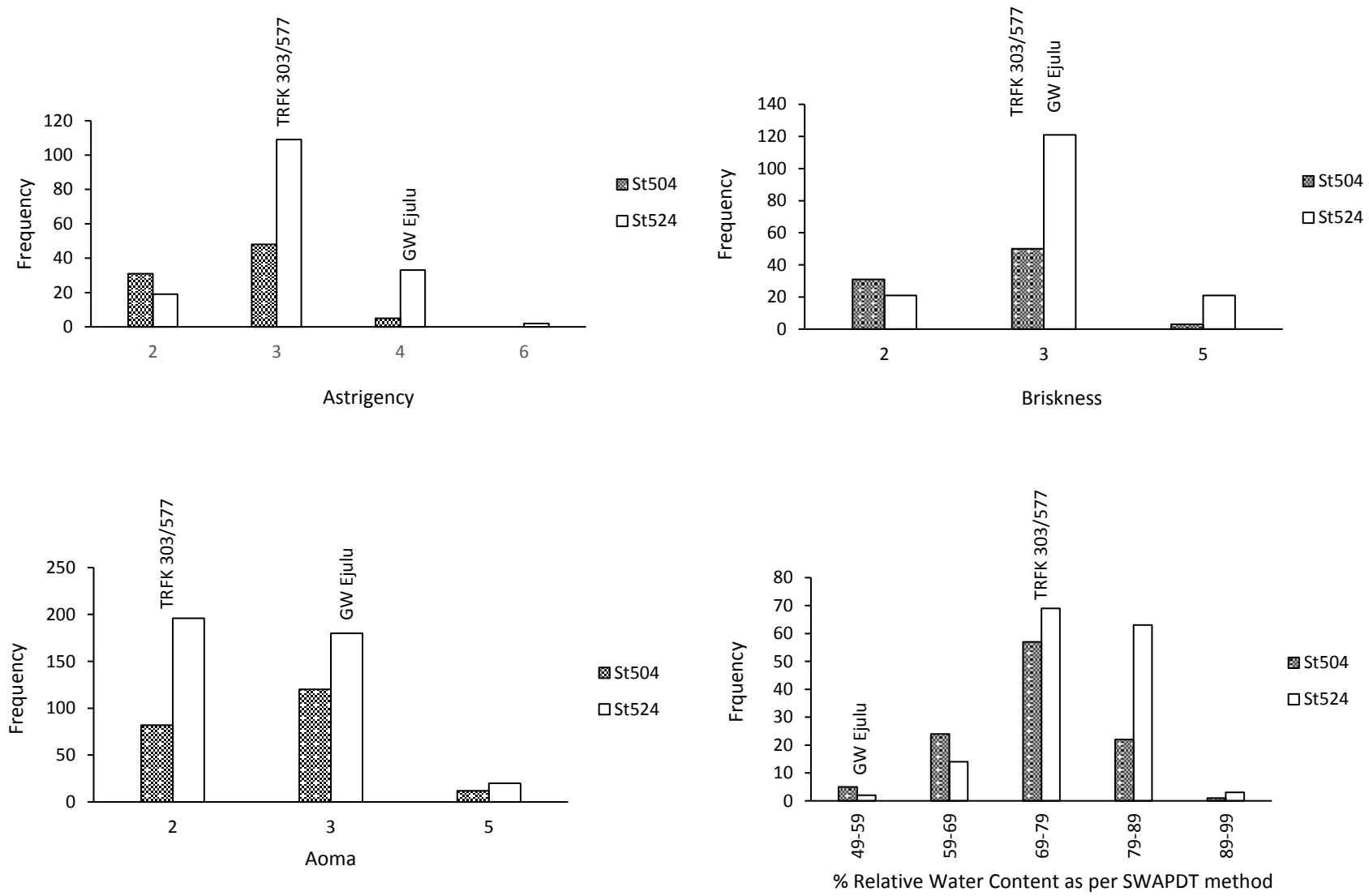


Fig. 1. Frequency distribution pattern of phenotypic traits of both green and black tea from two parental clones and F_1 population derived from a reciprocal cross between TRFK 303/577 and GW Ejulu (i.e. St 504 and St 524), respectively. The phenotypic values of the two parental clones are indicated in the graphs.

drought. Parental clone GW Ejulu which is drought susceptible tea cultivar had low % rwc of 55.1% while parental clone TRFK 303/577 which is drought tolerant cultivar had high % rwc of 74.1%. In overall, a large number of phenotypic traits for the progeny fell outside to those of the parents indicating transgressive segregation, hence polygenic inheritance of the traits in question (Fig. 3). Further, the observed normal distribution for the phenotypic traits shows involvement of several genes in the genetic control of phenotypic traits.

Correlation analysis

Pearson's correlation coefficients between different observations of each trait in the two tea population are shown in Table 3 and Table 4. The correlations among the phenotypic traits were grouped into: catechins (CAT, EC, ECG, EGC and EGCG), caffeine, theaflavins (TF1, TF2, TF3, and TF4), and tea taster's scores (CL, BRT, AST, BRK and AR) and % RWC. In general, the significant correlation coefficients mostly observed in St 504 were between EC and EGC ($r=0.62$, $P<0.01$), TF1 and TF2 ($r=0.89$, $P<0.001$), TF1 and TF4 ($r=0.90$, $P<0.001$), TF2 and TF4 ($r=0.89$, $P<0.001$). For tea taster's score, a high correlation was observed between individual AST and BRK ($r=0.96$, $P<0.001$), and it ranged from $r=0.62$-0.96. There were also significant correlations between AR with CL ($r=0.39$-0.42, $P<0.05$), BRT ($r=0.54$-0.57, $P<0.01$), AST ($r= 0.47$-0.51, $P<0.01$) and BRT ($r=0.56$-0.58, $P<0.01$) (Table 3 and 4). The correlation coefficient for some of the phenotypic traits in St 524, had a similar trend as in St 504 for the tea taster's scores except for catechins and theaflavins. The individual catechins correlated significantly with other individual catechins and theaflavins. There was significant correlation between EGC and EGCG ($r=0.51$, $P<0.01$), EC and EGC ($r=0.73$, $P<0.01$), TF1 and EGC ($r=0.62$, $P<0.01$), TF2 and EC ($r=0.44$, $P<0.05$), TF1 and EGCG ($r=0.42$, $P<0.05$), TF2 and EGC ($r=0.52$, $P<0.01$), TF4 and EGC ($r=0.54$, $P<0.01$) (Table 4). However, there was no correlation between % RWC with all biochemical compounds and tea taster scores in the two tea populations (Table 3 and 4).

Genetic map construction

Regression-based integrated map (RG) resulted into 15 linkage groups (Fig. 2; Table 5). The integrated genetic linkage map covered a total genetic distance of 1,260.1 cM, with an average locus distance of 1.1 cM. The marker density was not always uniform among linkage groups. The individual linkage groups ranged in size from 64.8 cM (LG02) to 160.1 cM (LG09), and the number of markers on each group varied from 50 for LG01 to 219 for LG09 (Table 5). Linkage group 9 had the highest marker density with an average locus distance of 1.4 cM, while the other LGs had relatively lower marker density (0.5-1.7 cM/DArT). However, DArT markers for all the phenotypic traits were absent on the LG03, LG05 and LG11. The constructed maps contained 15 linkage groups which corresponds to haploid chromosome number in *Camellia* spp. ($n = 15$) indicating saturation of the genome.The clustering of markers was observed, and some linkage groups were more densely saturated with markers as compared to others.

Table 3. Correlation coefficients (r) between phenotypic traits in the mapping population St 504

	CAFF	CAT	EC	ECG	EGC	EGCG	TF1	TF2	TF3	TF4	CL	BRT	AST	BRK	AR	RWC
CAFF	1															
CAT	-0.06	1														
EC	-0.07	-0.40	1													
ECG	0.44*	0.22	0.26	1												
EGC	-0.13	-0.78**	0.62**	-0.34	1											
EGCG	0.39	-0.57**	0.08	0.12	0.41*	1										
TF1	-0.24	-0.39	0.14	-0.33	0.42*	0.04	1									
TF2	-0.04	-0.36	0.10	-0.09	0.26	0.12	0.89***	1								
TF3	0.03	-0.08	-0.30	-0.01	-0.13	0.11	0.64**	0.83***	1							
TF4	-0.20	-0.28	-0.03	-0.25	0.23	0.08	0.90***	0.89***	0.80***	1						
CL	0.32	-0.08	0.20	0.26	0.12	0.17	-0.03	0.09	-0.13	-0.10	1					
BRT	0.24	-0.18	0.24	0.24	0.16	0.15	0.16	0.32	0.07	0.11	0.73**	1				
AST	0.20	-0.09	0.21	0.24	0.11	0.18	0.15	0.26	0.08	0.16	0.66**	0.83***	1			
BRK	0.22	-0.12	0.23	0.25	0.12	0.16	0.19	0.31	0.13	0.21	0.62**	0.80**	0.96***	1		
AR	0.18	-0.17	0.12	0.19	0.17	0.15	0.18	0.20	0.01	0.15	0.42*	0.54**	0.51**	0.58**	1	
RWC	0.12	0.19	0.16	-0.19	0.07	-0.03	0.06	0.06	0.02	0.06	-0.05	0.05	0.10	0.05	-0.06	1

CAFF- Caffeine; CAT- Catechin; EC- Epicatechin; ECG- Epicatechin gallate; EGC- Epigallocatechin gallate; EGCG- Epigallocatechin gallate; TF1- Theaflavin; TF2- Theaflavin-3-gallate; TF3- Theaflavin-3'-gallate; TF4- Theaflavin-3,3'-digallate; CL- Colour; BRT- Brightness; AST- Astringency; BRK- Briskness; AR- Aroma; RWC- percent relative water content

**Correlation significant at p≤ 0.05 level*

***Correlation significant at p≤ 0.01 level*

****Correlation significant at p≤ 0.001 level*

Table 4. Correlation coefficients (r) between phenotypic traits in the mapping population St 524

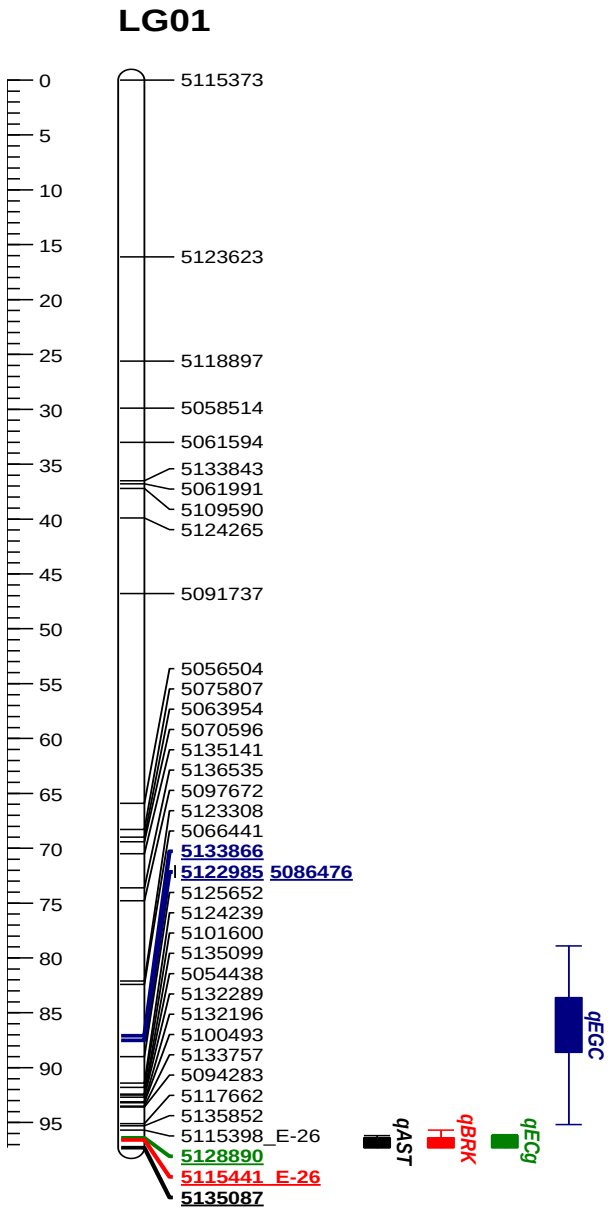
	CAFF	CAT	EC	ECG	EGC	EGCG	TF1	TF2	TF3	TF4	CL	BRT	AST	BRK	AR	RWC
CAFF	1															
CAT	0.12	1														
EC	0.07	-0.41*	1													
ECG	0.39	0.10	0.38	1												
EGC	-0.05	-0.74**	0.73**	-0.02	1											
EGCG	0.34	-0.60**	0.19	0.33	0.51**	1										
TF1	-0.01	-0.44*	0.51**	-0.01	0.62**	0.33	1									
TF2	0.13	-0.46*	0.44*	0.11	0.52**	0.42*	0.77**	1								
TF3	0.12	-0.16	-0.14	-0.03	0.02	0.31	0.44*	0.72**	1							
TF4	0.06	-0.51	0.34	-0.08	0.54**	0.42*	0.82**	0.91**	0.71**	1						
CL	0.11	-0.22	0.20	0.26	0.23	0.31	0.24	0.29	0.05	0.20	1					
BRT	0.09	-0.18	0.12	0.05	0.24	0.21	0.29	0.21	0.03	0.26	0.48*	1				
AST	0.09	-0.17	0.16	0.11	0.22	0.24	0.21	0.12	-0.01	0.17	0.57**	0.59**	1			
BRK	0.10	-0.17	0.14	0.10	0.21	0.22	0.19	0.17	0.01	0.25	0.47*	0.69**	0.76**	1		
AR	0.17	-0.15	0.21	0.10	0.26	0.16	0.22	0.21	-0.03	0.21	0.39*	0.57**	0.47*	0.56**	1	
RWC	0.09	0.07	0.06	-0.05	0.06	-0.03	0.12	0.08	0.06	0.12	-0.05	0.00	0.09	0.02	-0.02	1

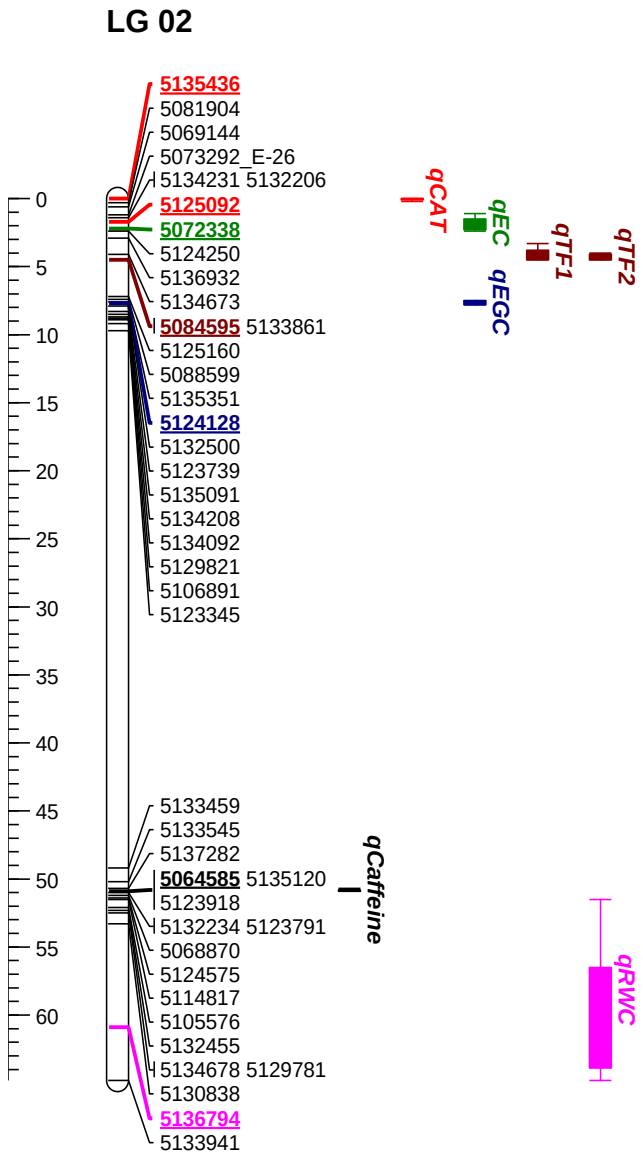
CAFF- Caffeine; CAT- Catechin; EC- Epicatechin; ECG- Epicatechin gallate; EGC- Epigallocatechin gallate; EGCG- Epigallocatechin gallate; TF1- Theaflavin; TF2- Theaflavin-3-gallate; TF3- Theaflavin-3'-gallate; TF4- Theaflavin-3,3'-digallate; CL- Colour; BRT- Brightness; AST- Astringency; BRK- Briskness; AR- Aroma; RWC- percent relative water content

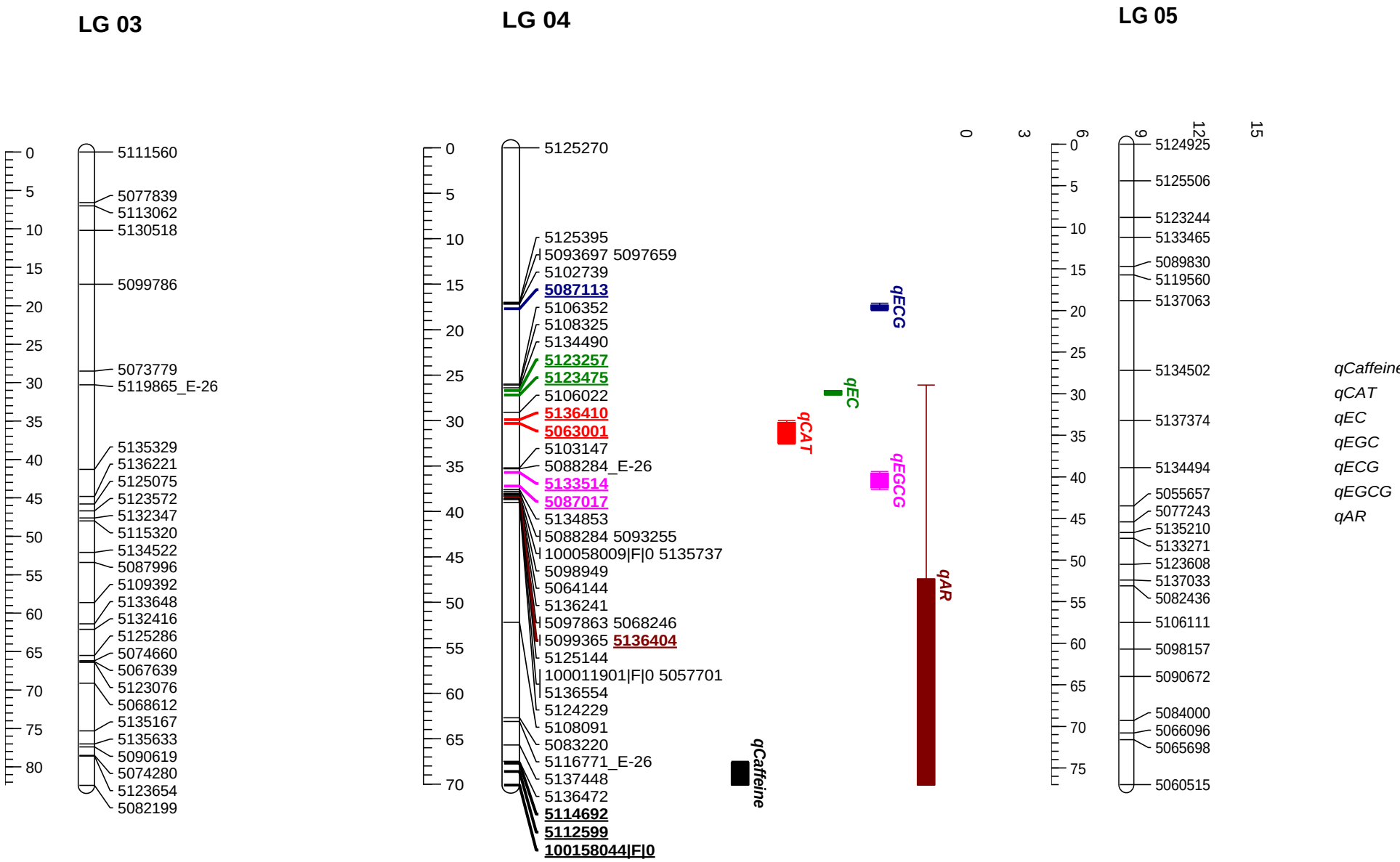
**Correlation significant at p≤ 0.05 level*

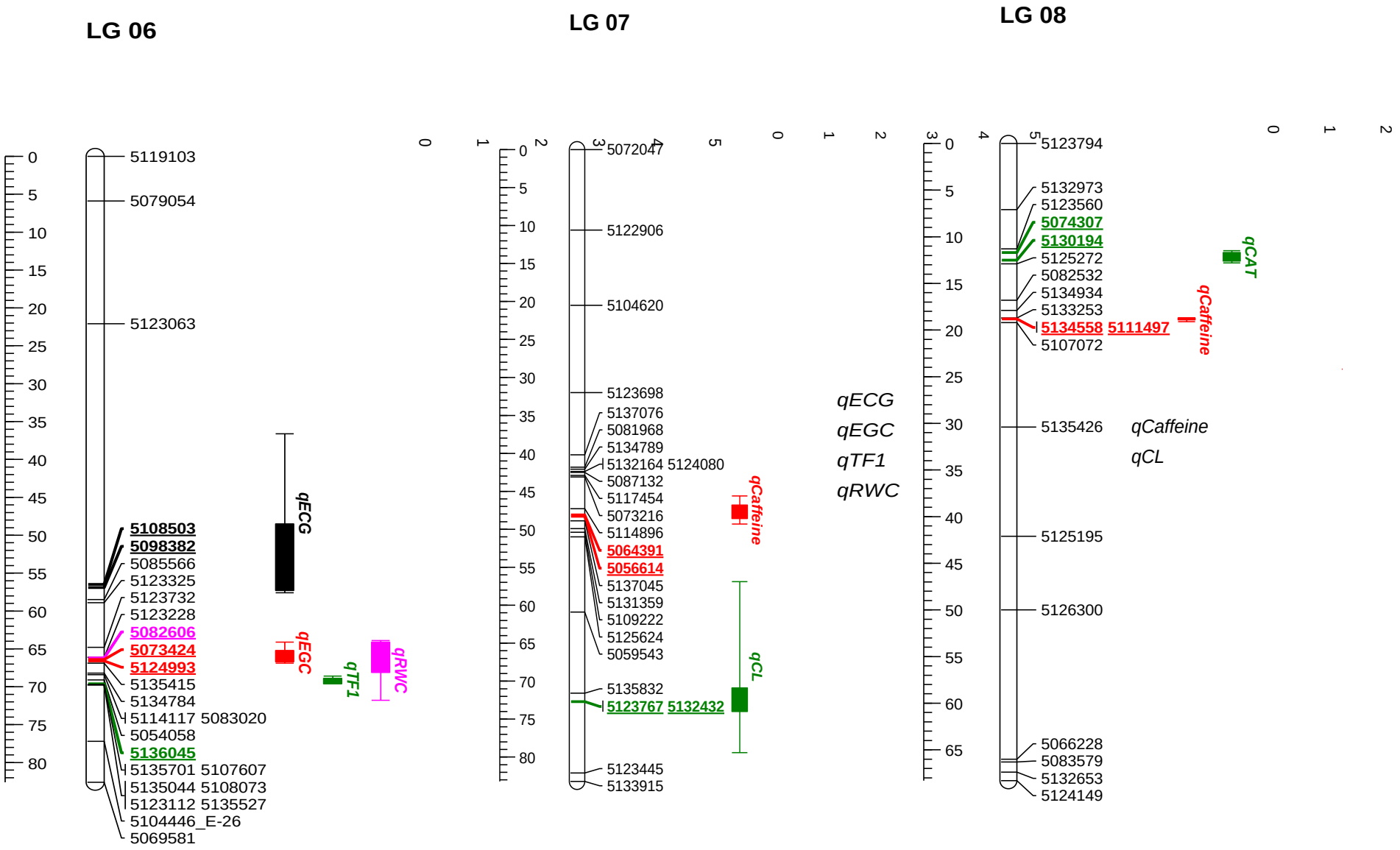
***Correlation significant at p≤ 0.01 level*

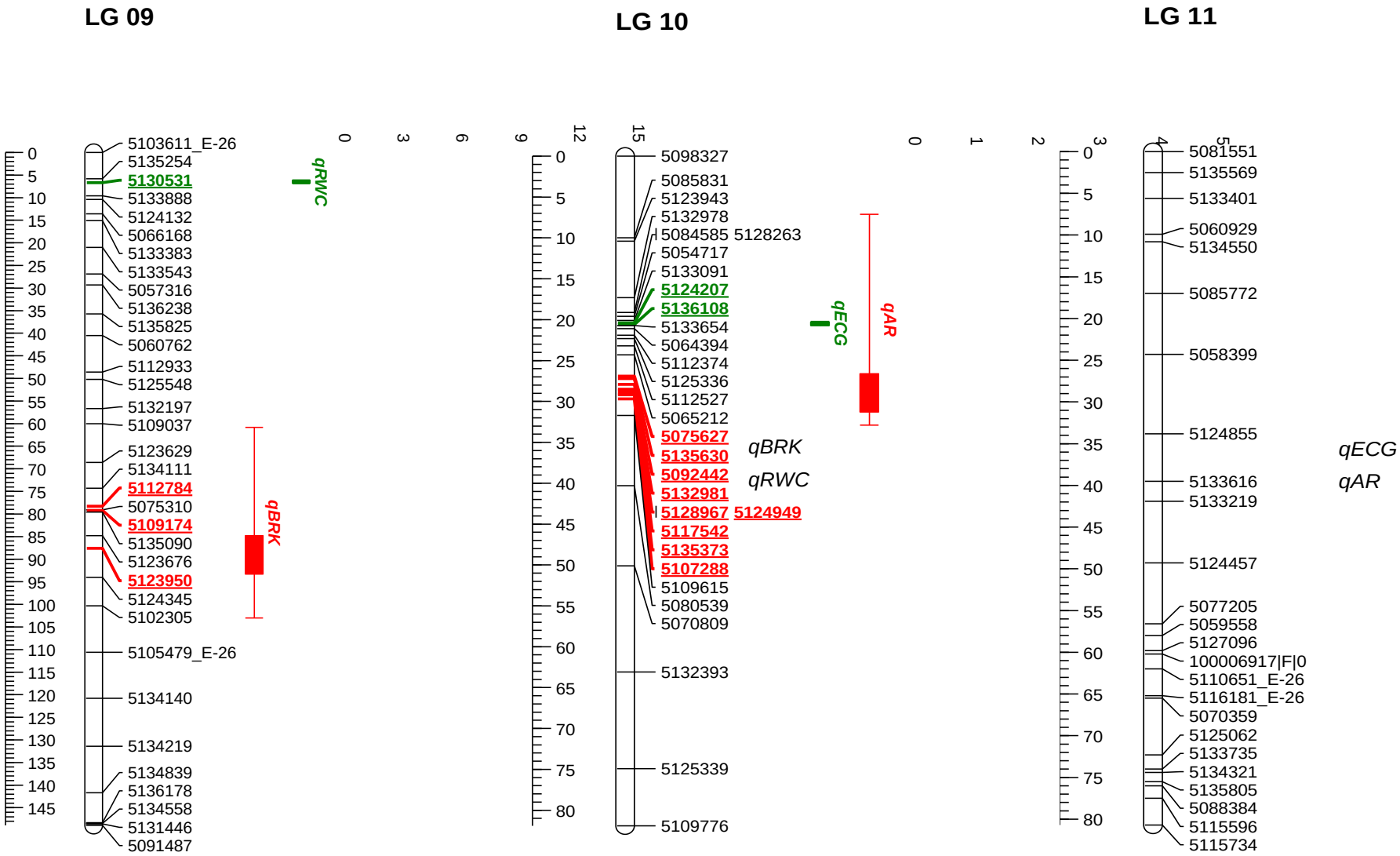
****Correlation significant at p≤ 0.001 leve*

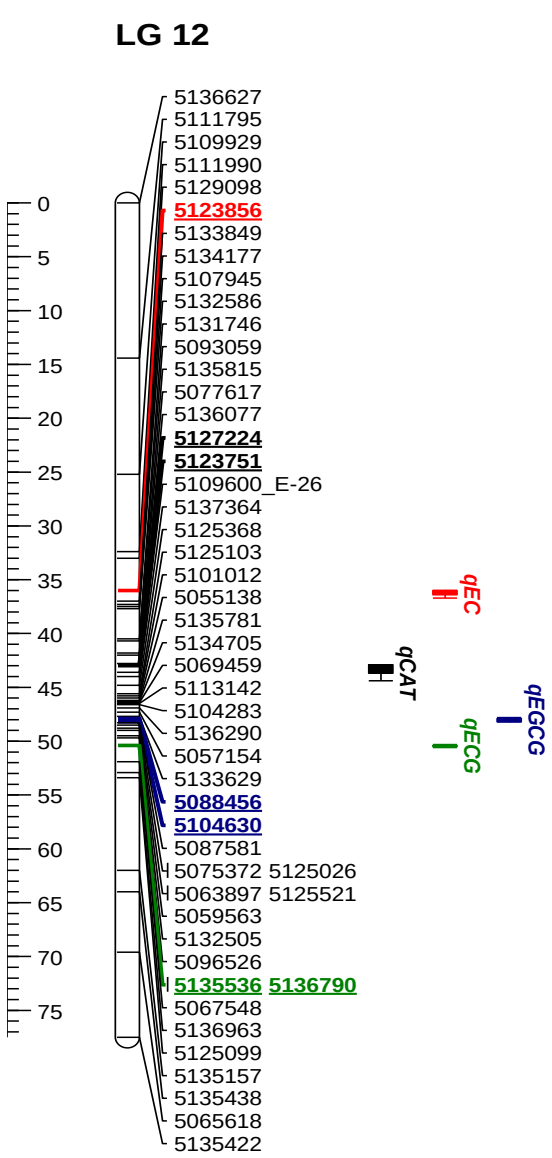

LG01
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
5115373
5123623
5118897
5058514
5061594
5133843
5061991
5109590
5124265
5091737
5056504
5075807
5063954
5070596
5135141
5136535
5097672
5123308
5066441
5133866
5122985 5086476
5125652
5124239
5101600
5135099
5054438
5132289
5132196
5100493
5133757
5094283
5117662
5135852
5115398_E-26
5128890
5115441_E-26
5135087
qAST
qBRK
qECg
qEGC

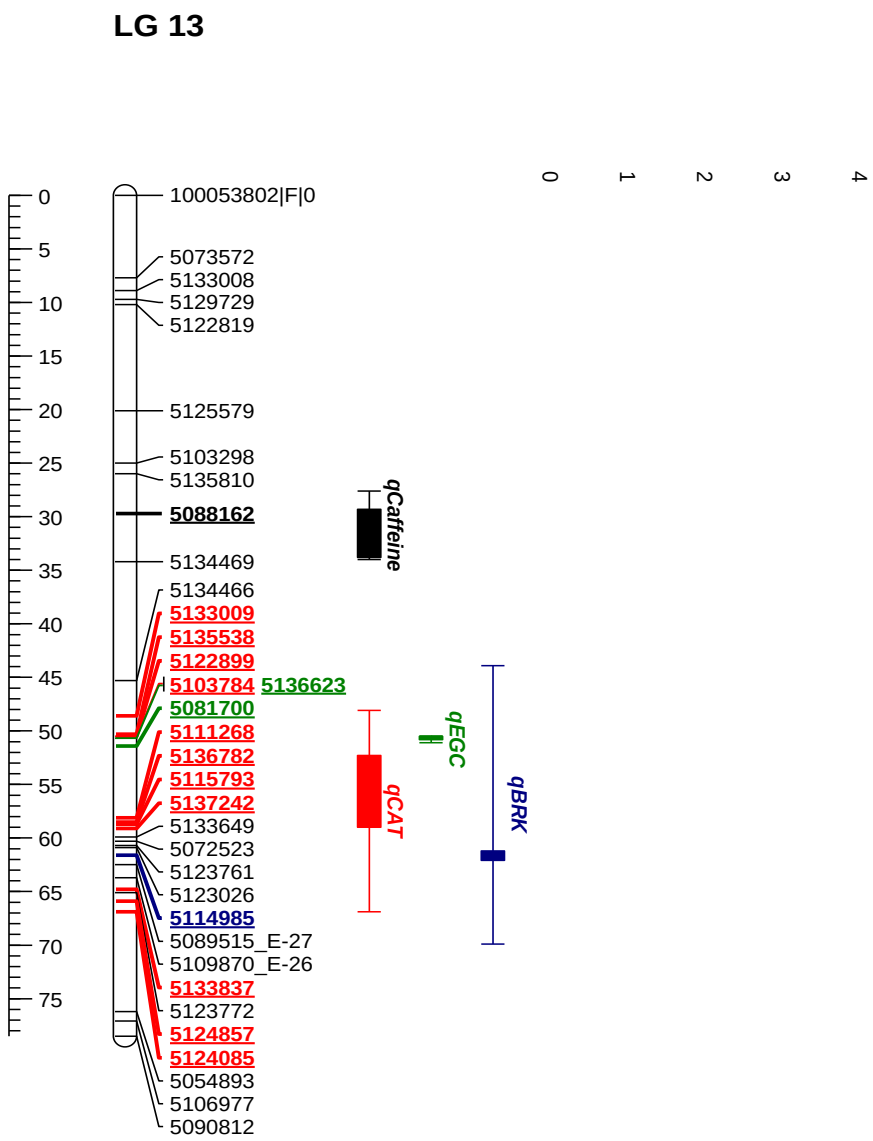

LG 02
0
5
10
15
20
25
30
35
40
45
50
55
60
5135436
5081904
5069144
5073292_E-26
5134231 5132206
5125092
5072338
5124250
5136932
5134673
5084595 5133861
5125160
5088599
5135351
5124128
5132500
5123739
5135091
5134208
5134092
5129821
5106891
5123345
5133459
5133545
5137282
5064585 5135120
5123918
5132234 5123791
5068870
5124575
5114817
5105576
5132455
5134678 5129781
5130838
5136794
5133941
qCAT
qEC
qTF1
qTF2
qEGC
qCaffeine
qRWC


LG 03
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
5111560
5077839
5113062
5130518
5099786
5073779
5119865_E-26
5135329
5136221
5125075
5123572
5132347
5115320
5134522
5087996
5109392
5133648
5132416
5125286
5074660
5067639
5123076
5068612
5135167
5135633
5090619
5074280
5123654
5082199
LG 04
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
5125270
5125395
5093697 5097659
5102739
5087113
5106352
5108325
5134490
5123257
5123475
5106022
5136410
5063001
5103147
5088284_E-26
5133514
5087017
5134853
5088284 5093255
100058009|F|0 5135737
5098949
5064144
5136241
5097863 5068246
5099365 5136404
5125144
100011901|F|0 5057701
5136554
5124229
5108091
5083220
5116771_E-26
5137448
5136472
5114692
5112599
100158044|F|0
qECG
qEC
qCAT
qEGCG
qAR
qCaffeine
0
3
6
9
12
15
LG 05
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
5124925
5125506
5123244
5133465
5089830
5119560
5137063
5134502
5137374
5134494
5055657
5077243
5135210
5133271
5123608
5137033
5082436
5106111
5098157
5090672
5084000
5066096
5065698
5060515
qCaffeine
qCAT
qEC
qEGC
qECG
qEGCG
qAR

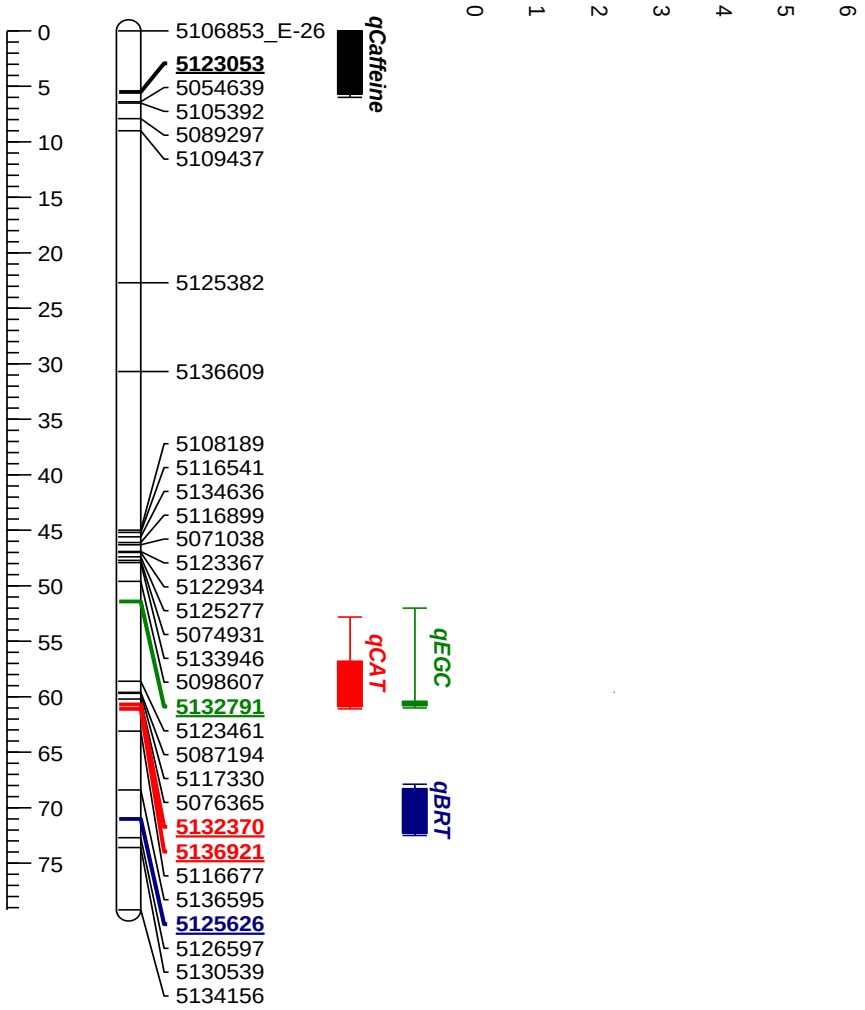

LG 06
LG 07
LG 08
5119103
5079054
5123063
5108503
5098382
5085566
5123325
5123732
5123228
5082606
5073424
5124993
5135415
5134784
5114117 5083020
5054058
5136045
5135701 5107607
5135044 5108073
5123112 5135527
5104446_E-26
5069581
qECG
qEGC
qTF1
qRWC
5072047
5122906
5104620
5123698
5137076
5081968
5134789
5132164 5124080
5087132
5117454
5073216
5114896
5064391
5056614
5137045
5131359
5109222
5125624
5059543
5135832
5123767 5132432
5123445
5133915
qCaffeine
qCL
qECG
qEGC
qTF1
qRWC
5123794
5132973
5123560
5074307
5130194
5125272
5082532
5134934
5133253
5134558 5111497
5107072
5135426
5125195
5126300
5066228
5083579
5132653
5124149
qCAT
qCaffeine
qCaffeine
qCL


LG 09
LG 10
LG 11
5103611_E-26
5135254
5130531
5133888
5124132
5066168
5133383
5133543
5057316
5136238
5135825
5060762
5112933
5125548
5132197
5109037
5123629
5134111
5112784
5075310
5109174
5135090
5123676
5123950
5124345
5102305
5105479_E-26
5134140
5134219
5134839
5136178
5134558
5131446
5091487
qRWC
qBRK
5098327
5085831
5123943
5132978
5084585 5128263
5054717
5133091
5124207
5136108
5133654
5064394
5112374
5125336
5112527
5065212
5075627
5135630
5092442
5132981
5128967 5124949
5117542
5135373
5107288
5109615
5080539
5070809
5132393
5125339
5109776
qBRK
qRWC
qECG
qAR
5081551
5135569
5133401
5060929
5134550
5085772
5058399
5124855
5133616
5133219
5124457
5077205
5059558
5127096
100006917|F|0
5110651_E-26
5116181_E-26
5070359
5125062
5133735
5134321
5135805
5088384
5115596
5115734
qECG
qAR


LG 12
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
5136627
5111795
5109929
5111990
5129098
5123856
5133849
5134177
5107945
5132586
5131746
5093059
5135815
5077617
5136077
5127224
5123751
5109600_E-26
5137364
5125368
5125103
5101012
5055138
5135781
5134705
5069459
5113142
5104283
5136290
5057154
5133629
5088456
5104630
5087581
5075372 5125026
5063897 5125521
5059563
5132505
5096526
5135536 5136790
5067548
5136963
5125099
5135157
5135438
5065618
5135422
qEC
qCAT
qEGCG
qECG


LG 13
0
1
2
3
4
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
100053802|F|0
5073572
5133008
5129729
5122819
5125579
5103298
5135810
5088162
5134469
5134466
5133009
5135538
5122899
5103784 5136623
5081700
5111268
5136782
5115793
5137242
5133649
5072523
5123761
5123026
5114985
5089515_E-27
5109870_E-26
5133837
5123772
5124857
5124085
5054893
5106977
5090812
qCaffeine
qCAT
qEGC
qBRK

LG 14


0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
5106853_E-26
5123053
5054639
5105392
5089297
5109437
5125382
5136609
5108189
5116541
5134636
5116899
5071038
5123367
5122934
5125277
5074931
5133946
5098607
5132791
5123461
5087194
5117330
5076365
5132370
5136921
5116677
5136595
5125626
5126597
5130539
5134156
qCaffeine
qCAT
qEGC
qBRT
0
1
2
3
4
5
6

LG 15

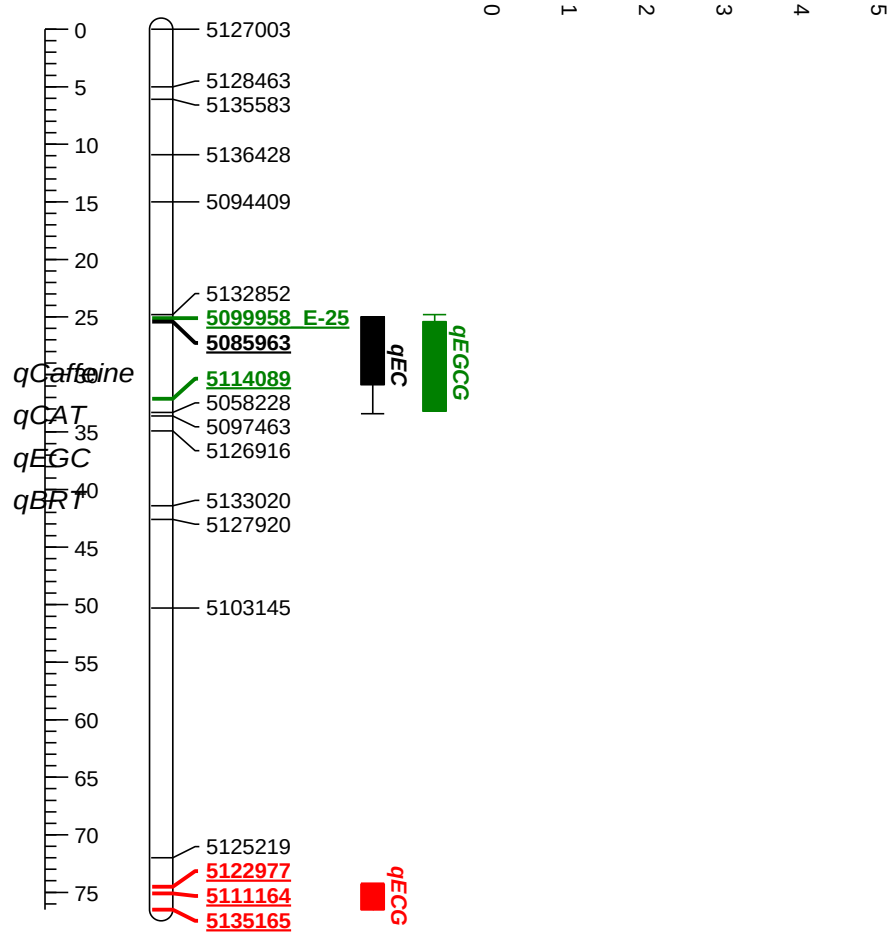

0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
5127003
5128463
5135583
5136428
5094409
5132852
5099958_E-25
5085963
5114089
5058228
5097463
5126916
5133020
5127920
5103145
5125219
5122977
5111164
5135165
qCaffeine
qCAT
qEGC
qBRT
qEC
qEGCG
qECG
0
1
2
3
4
5

Fig. 2. DArT-based genetic map of tea plant showing location of QTLs for caffeine, catechins, theaflavins content, liquor brightness and % rwc identified in the two mapping population. Map distance scales in cM are placed at left margin. Significant QTL for caffeine, individual catechin and individual theaflavin content and liquor brightness is represented by different bars. Bars and lines indicate 1-LOD and 2-LOD support intervals.
qCaffeine- Caffeine; qCAT- Catechin; qEC- Epicatechin; qECG- Epicatechin gallate; qEGC- Epigallocatechin gallate; qEGCG- Epigallocatechin gallate; qTF1- Theaflavin; qTF2- Theaflavin-3-gallate; qCL- Colour; qBRT- Brightness; qAST- Astringency; qBRK- Briskness; qAR- Aroma; qRWC- percent relative water content.

Table 5. DArT markers distribution among the linkage groups

Linkage group	Number of markers	Total length covered (cM)	Average distance between markers (cM)
LG1	50	97.3	0.5
LG2	107	64.8	1.7
LG3	61	82.4	0.7
LG4	98	70.1	1.4
LG5	77	77.0	1.0
LG6	65	82.6	0.8
LG7	139	83.2	1.7
LG8	100	68.3	1.5
LG9	219	160.1	1.4
LG10	114	81.9	1.4
LG11	67	80.7	0.8
LG12	83	77.5	1.1
LG13	68	78.5	0.9
LG14	83	79.2	1.0
LG15	90	76.5	1.2
Total	1421	1260.1	
Average		84	1.1

Phenotype segregation and QTL mapping

The analysis of 16 phenotypic traits revealed a total of 47 putative QTL in 15 LGs associated with tea quality and % RWC by IM and MQM at a significance genome-wide threshold of 5% level of confidence. Significant LOD scores ranged from 3.1 - 47.3 with an average of 6.5, while LOD thresholds of the detected QTL were between 2.9 and 4.6 with an average of 3.2 (Table 6). The interval lengths of the QTL ranged from 0.14 cM to 23.5 cM. QTL associated with individual catechins, caffeine and individual theaflavins were mostly clustered in LG02 and LG04 but in different regions on the map (Fig. 2). Most QTL for tea taster scores (CL, BRT, AST, BRK and AR) were detected at a lower significance level of confidence of 90% as compared to traits for biochemical molecules (Table 6). The QTL for each trait varied from a total of two (EC) to a total of seven (caffeine) in number. A total of nineteen QTL were identified for ten traits and were clustered in three main regions on LG01, LG02, LG04, LG12, LG13 and LG14 (Table 6). The variability explained by each QTL in the population was at average to high levels and ranged from

Table 6. QTLs for caffeine, individual catechins and theaflavins content, tea taster's scores and percent relative water content detected in St 504 and St 524 tea plant population

Traits	QTL	LG	Position (cM)	LOD threshold	LOD score	PVE (%)	Nearest Marker
Caffeine	qCaffeine	2	50.905 (50.247-51.548)	3.3	3.4	6.0	5064585
		4	68.574 (63.137-70.142)	3.2	3.8	6.7	5112599
		7	48.071 (34.695-50.409)	3.2	3.7	6.6	5064391
		8	18.762 (16.783-26.008)	3.3	3.9	6.7	5134558
		13	29.748 (23.653-36.538)	3.1	4.7	8.1	5088162
		14	5.465 (0.000-5.465)	3.3	3.7	6.4	5123053
CAT	qCAT	2	0.000 (0.000-2.421)	3.2	3.7	6.4	5135436
		4	30.322(29.886-30.322)	3.3	46.0	55.6	5063001
		8	12.506 (12.506-12.882)	3.3	3.3	5.6	5130194
		12	42.998 (36.019-54.085)	3.1	6.1	10.3	5123751
		13	50.602 (48.559-58.674)	3.0	4.1	6.9	5111268
		14	60.713 (47.731-61.117)	3.2	4.1	6.9	5132370
EC	qEC	2	2.242 (0.000-8.309)	3.3	4.1	7.0	5072338
		15	25.381 (20.639-25.381)	3.3	4.5	7.6	5085963
ECG	qECG	1	96.36 (96.36-97.284)	3.0	6.8	11.7	5128890
		4	17.658 (13.976-37.226)	3.4	4.7	8.0	5087113

		6	56.948 (56.504-56.948)	3.1	4.0	6.7	5098382
		10	20.631 (20.090-25.539)	3.3	4.8	8.0	5136108
		12	50.441 (50.148-50.441)	2.9	14.9	23.1	5136790
		13	50.602 (48.559-58.674)	3.1	3.7	6.4	5088162
		15	75.067 (75.067-76.498)	3.3	4.2	7.2	5111164
EGC	qEGC	1	87.059 (84.375-88.562)	3.0	3.3	5.8	5133866
		2	7.719 (7.719-7.855)	3.2	3.2	5.5	5124128
		4	27.194 (27.194-28.879)	3.3	47.3	56.6	5123475
		6	66.351 (56.504-72.354)	3.1	4.2	7.2	5073424
		12	42.998 (42.839-54.085)	3.0	14.9	8.6	5123751
		13	50.602 (48.559-58.674)	3.1	3.7	6.4	5136623
		14	60.713 (60.713-63.050)	3.1	4.0	6.9	5132791
EGCG	qEGCG	4	37.226 (28.879-38.559)	3.4	15.6	24.1	5087017
		12	48.179 (36.963-49.744)	3.0	4.2	7.2	5104630
		15	32.140 (17.119-33.578)	3.3	3.8	6.5	5114089
TF1	qTF1	2	4.455 (4.455-5.391)	3.1	3.2	5.8	5084595
		6	69.564 (69.564-75.112)	4.1	4.5	7.9	5136045
TF2	qTF2	2	4.455 (1.220-5.391)	3.1	4.0	7.0	5084595

CL	qCL	7	72.684 (72.677-72.684)	3.0	3.2	6.5	5132432
BRT	qBRT	14	71.046 (65.383-71.046)	3.4	3.3	6.8	5125626
AST	qAST	1	96.577 (88.83-97.284)	3.0	4.5	9.1	5135087
		9	87.614 (87.614)	3.2	3.2	6.6	5123950
BRK	qBRK	1	96.577 (89.022-97.284)	3.0	3.6	7.3	5115441_E-26
		9	87.614 (87.614)	3.6	3.6	7.4	5123950
		13	61.582 (61.582)	3.0	3.1	6.3	5114985
AR	qAR	4	68.574 (68.574)	3.0	3.1	6.4	5112599
		10	28.656 (26.854-29.038)	3.2	3.5	7.0	5128967
% RWC	qRWC	2	60.868 (56.912-60.868)	3.3	4.3	7.3	5136794
		6	66.217 (66.052-66.217)	3.2	3.3	5.7	5082606
		9	6.684 (6.167-6.684)	3.9	4.0	6.9	5130531
Average				3.2	6.5	9.9	

CAT Catechin, EC epicatechin, ECG epicatechin gallate, EGC epigallocatechin, EGCG epigallocatechin gallate, TF1 theaflavin, TF2 theaflavin-3-gallate, CL colour, BRT brightness, AST astringency, BRK briskiness, % RWC percent relative water content
LOD- logarithm of odds ratio
PVE- phenotypic variation explained
The genome-wide LOD significance thresholds ($P<0.05$) based on permutation testing (n=1,000)

5.5 to 56.6%, with an average of 9.9%. The population variability explained (PVE) by all QTL for all the traits ranged between 6.5 (CL) to 97% (EGC).

Seven QTL for caffeine levels were detected and accounted for 40.5% of the total population explained variability (Table 6). Of these, one QTL, qCaffeine, was mapped on LG13 close to marker 5088162, with and LOD score of 4.7 and phenotypic variation explained (PVE) of 8.1% (Table 6). An additional five putative QTL were identified on LG02, LG04, LG07, LG08 and LG14 (Table 6). The qCaffeine LOD score ranged from 3.2 to 4.7, and the population variability explained ranged between 6.0% and 8.1%. Six QTL were detected for CAT content, accounted for a total of 91.1% of the total population variability. The two major QTL, qCAT, were located on LG04 and LG12 close to marker 5063001 and 5123751, respectively. The two QTL, had the largest effect on CAT content, accounting for 55.6 and 10.3% of the population variability, respectively (Table 6). The QTL were distributed across the six LGs, with a LOD score ranging from 3.7 to 46. Two QTL, qEC, were identified for EC content, which accounted for a total of 14.6% of the population variability (Table 6). The qEC was located on LG02 close to marker 5072338 with PVE of 7.0%, and on LG15 close to marker 5085963 with 7.6% of the population variability. Seven QTL, qECG, were detected for ECG content, which collectively explained 71.1% of the total population variability with LOD score range of 3.7 to 14.9 (Table 6). The QTL for ECG, qECG, was located on LG12 close to marker 5136790, had the largest effect on ECG content and accounted for of 23.1% of the population variability. A total of seven QTL were detected for EGC content, which accounted for a total of 97% the population variability (Table 6). Of these, two QTL, qEGC, were mapped on LG04 and LG12, respectively (Table 6). The qEGC on LG04 was located close to marker 5123475 with PVE of 56.6%, while the qEGC on LG12 was located close to marker 5123751 with PVE of 8.6% with LOD score of 47.3 and 14.9, respectively. An additional five putative QTL were identified on LG1, LG02, LG06, LG13 and LG14 (Table 6). The LOD score for the five loci varied from 3.2 to 4.2, and the PVE ranged between 5.5 and 7.2%. The QTL, qEGCG, was identified for EGCG content on LG04, close to marker 5087017, contributing to 24.1% of the population variability (Table 6). In addition, two putative QTL was mapped on LG12 and LG15, and explained between 6.5 and 7.2% of the population variability (Table 6). The three QTL detected for EGCG content accounted for a total of 37.8% the population variability (Table 6). Two QTL for theaflavin, qTF1, were identified and it accounted for a total of 13.7% of the population variability (Table 6). The qTF1 was located on LG02 and LG06 close to markers 5084595 and 5136045, respectively. The nucleotide sequences for these markers are presented in supplementary Table 1. Only one QTL for qTF2 and another one for qBRT, were identified on LG02 and LG14, close to marker 5084595 and 5125626, contributing to 7.0 and 6.8% of the population variability, respectively (Table 6).

The QTL for astringency, qAST, and briskness, qBRK were mapped on two linkage groups (LG01 and LG09). The population variability for the tea taster score varied with aroma, qAR (6.4%) and briskness, qBRK (7.3%), respectively. Tea liqour colour and tea brightness, had one QTL each while tea briskness, tea aroma had two QTL each and tea astringency had three QTL, respectively. Three QTL were detected for % RWC, which accounted for a total of 19.9% of the population variability (Table 6). The three stable QTL, qRWC, were mapped on LG02, LG06 and LG09, respectively. Three QTL were identified for percent rwc. The first QTL (qRWC) was located on LG02, the second QTL was located on chromosome LG06, and the third QTL was located on

LG09. These QTL spanned 60.7 cM, 66.2 cM, and 6.7 cM and have peak LOD scores of 4.3, 3.3 and 4.0, respectively (Table 6).

Discussion

Construction of a genetic linkage map is an important prerequisite for the identification of QTL(s) as well as mapping of agronomically important genes, which will be very useful for developing improved cultivars. Green tea catechins, black tea theaflavins and caffeine are important biochemical traits that determine tea quality given that tea price at the auction is based on tea quality. The two principal gallated catechins present were EGCG and EGC which corroborate with results reported by (Karori *et al.,* 2014). The two gallated catechins (EGCG and EGC) have been shown to contribute significantly to the formation of theaflavins. The B-ring in tri-hydroxylated catechins is oxidized high rate than the B-ring in di-hydroxylated catechins, which include: CAT, EC and EGC because of their lower oxidation ability (Owuor *et al.,* 2006a). Therefore, tea cultivars with high levels of EGCG and EGC can be developed through tea breeding programs to improve black tea quality.

It was demonstrated previously that the ratio of di-hydroxyl flavan-3-ols to tri-hydroxyl flavan-3-ols influenced black tea quality in that, high proportion of simple catechins (CAT, EC and ECG) to gallocatechins (EGC and EGCG), lead to a higher proportion of theaflavin (Ellis and Nyirenda, 1995). The strong correlation of the di-gallate fractions in the present study is an important in factor that can be utilized to enhance the quality of black tea. The changes in theaflavins and residual catechin compositions in relation to sensory characteristics of total CL, BRT, and BRK has also been reported in black tea (Obanda *et al.,* 2001). Further, (Obanda *et al.,* 2001), reported that theaflavin digallate is the most astringent, estimated to be 6.4 times more astringent than simple theaflavin, and 2.88 times more astringent than either theaflavin-3-monogallate or theaflavin-3'-monogallate. The formation of a single theaflavin molecule requires a dihydroxy and a trihydroxy flavan-3-ol (Wright *et al.,* 2002), as follows:

Epicatechin (EC) + Epigallocatechin (EGC) = simple theaflavin (TF1);
Epicatechin (EC) + Epigallocatechin gallate (EGCg) = Theaflavin-3-gallate (TF2);
Epicatechin gallate (ECG) + Epigallocatechin (EGC) = Theaflavin-3'-gallate (TF3);
Epicatechin gallate (ECG) + Epigallocatechin gallate (EGCg) = Theaflavin-3, 3'-digallate (TF4)

The relationship between catechins and TFs content with the quality of black tea has been investigated with reports showing a significant correlation between sensory parameters with catechins and TFs content. Among all the sensory, parameters investigated such as colour, brightness, astringency, briskness and aroma; the results in this present study indicate that neither caffeine, individual catechins nor individual TFs content had any correlation with sensory parameters. This could be due to loss in aromatic compounds between storage and the time of organoleptic evaluation was done which led to flat tea (Sedaghathoor et al., 2013). Although, there was a significant correlation between the individual sensory parameters (CL, BRT, AST, BRK and AR). These results are similar to those reported by (Owuor *et al.,* 2006b) that there was no significant correlation between simple (non-gallated) theaflavin with sensory parameters except for tea liquor brightness. The significant correlations shows that the phenotypic traits in are either controlled by linked genes or genes that have pleiotropic effects on the traits. Hence, improving one phenotypic trait may lead to improvement of other phenotypic traits.

The map spanned a total length of 1260.1 cM with 1421 markers, which is close to tea genome obtained by (Ota and Tanaka, 1999) (1,640 cM), (Hackett *et al.,* 2000) (1,349.7 cM), and (Taniguchi *et al.,* 2012) (1,298 and 1,305 cM), (Ma *et al.,* 2014) (1,143.5 cM), but smaller in length than those reported by (Huang *et al.,* 2006) (2,457.7 and 2,545.3 cM), and (Hu *et al.,* 2013) (4,482.9 cM). The differences in the total map length can result from variation in the number of recombination events in the two parents as well as variations in the numbers and locations of mapped loci. There were 15 linkage groups found in this study, indicating saturation of the genome (n= 15). However, there was gap greater than 20 cM between the adjacent markers on LG06 and LG15. Even though the combination of *PstI* (CTGCAG) and *MseI* (CCGG) restriction enzymes performed best with 16,382 DArTseq achieved, there was still one gap of larger than 20 cM between adjacent markers on LG06 and LG15. The presence of a large gap may be as a result of genome regions corresponding to the gap regions of the genetic map or homozygous could be a result of gaps in both mapping parents which leads no recombination events. Therefore, further studies need to be carried out to fill in the gaps in the current genetic map. However, the mapped markers in the current study were distributed in all 15 LGs with marker density ranging from 0.5 to 1.7 cM. The average interval between two markers was 1.1 cM, which is an improvement to previous genetic maps which ranged from 1.9 to 19.0 cM with the latest by Ma *et al.,* (2014) spacing 2.9 cM. The marker density for genome-wide QTL mapping which is recommended should be less than 10 cM (Doerge, 2002). The genetic map constructed in the present study is thus suitable for identification of QTL.

In total, six caffeine QTL, 25 catechins QTL, three theaflavins QTL, nine QTL for tea taster score and three QTL for percent rwc were detected, and the population variability explained by each QTL varied from 5.5 to 56.6%, with an average of 9.9%. The high PVE exhibited by these QTL suggests that caffeine, catechins, theaflavins content, tea taster scores and percent rwc may be controlled by few critical genes. However, the small size of 261 mapping population used in this study may lead to overestimation of QTL although the population size is slightly smaller than from the previous study of 300 individuals by Ma *et al.* (2014). The individual catechins (ECG, EGC and EGCG) were mostly detected in LG01, LG04, LG12 and LG13. In addition to QTL for catechins, the present study incorporated QTL for caffeine, black tea theaflavins and tea taster's scores, which are major determinants of tea quality. The major QTL, qCAT, qEC, qEGC, qTF1 and qTF2, detected showed a tendency to co-localize and cluster on LG02 while QTL, qCAT, qECG, qEGC and qEGCG clustered in LG12. The QTL were particularly prominent in the region between 0.00 and 7.72 cM of LG02 and 42.9 to 50.4 cM of LG12, where more than four QTL are located, for each trait.

Linkage group one, apart from the QTL for individual catechins content, QTL for individual theaflavins, qTF1 and qTF2 were also detected. These regions of the two chromosomes probably contain genes that are multifunctional which are associated with production and accumulation of catechins and theaflavins and deserve further investigation. More interestingly, QTL associated with caffeine, qCaffeine, and liquor brightness, qBRT, were not clustered with and/or located within the same region on the chromosome with the QTL for catechins and theaflavins in the linkage groups. It is an indication that the genes associated with caffeine production and accumulation as well as liquor brightness are located in different regions on the chromosome from those of catechins and theaflavins. The few loci that were associated with tea catechins and theaflavins will need to be confirmed for MAS. For example, markers 5123475 and 5084595 that

were consistently associated with individual catechins (EC and EGC) and individual theaflavins (TF1 and TF2), indicated their potential as candidate markers for MAS.

Conclusion

DArTseq markers are capable of detecting even very small polymorphisms, are cost-effective and are efficient for whole-genome scans and population structure. The main objective of tea breeding is to improve the quality and quantity of the product. The investigation of biochemical traits such as caffeine, catechins and theaflavins content, which are related to tea quality, presents an alternative approach in tea breeding. The SWAPDT method can also be used since it represents an easy method to measure water deficit in drought stress plants. The use of DArTseq markers in the present study will be useful for tea breeding and will provide valuable information for selecting parents that would give rise to progenies with the desired genotype. Although in the present study, linkage maps might be enhanced by further saturation, they constitute an additional tool that can be used to locate QTL linked with some important agronomic traits. These will contribute to the development of MAS for tea quality improvement, considering the favorable alleles of the markers involved in the QTL for tea quality. The validated molecular markers will contribute greatly to adoption of marker-assisted selection (MAS) for drought tolerance and tea quality improvement. In addition, with recent availability of a complete genome sequence of the tea plant, this will allow identification genes more precisely in relation to the QTL detected in this study.

Author contributions

ZA, SK and RM were involved with the design of the experiment and plant material used. RK, PM and CN were involved in collection of plant material. RK performed the experiments. RK, PM, CN, SK and ZA analyzed samples and interpreted the data. RK wrote the manuscript and revised by PM, CM, RM, SK and ZA. The final manuscript was reviewed and approved by all the authors.

Conflict of interest statement

All the authors declare that there is no commercial or financial relationships that can precedence to conflict of interest in research conducted.

Acknowledgement

The authors acknowledge the financial support to conduct this research, and study grants for RK and PM from James Finlay (Kenya) Ltd., George Williamson (Kenya) Ltd., Sotik Tea Company (Kenya) Ltd., Mcleod Russell (Uganda) Ltd., the TRI of Kenya and *Southern African Biochemistry and Informatics for Natural Products (*SABINA). The *C. sinensis* cultivars used in this study were provided by the TRI of Kenya. Supplementary funding was provided by, the Technology and Human Resources for Industry Programme (THRIP), an initiative of the Department of Trade and Industries of South Africa (dti), the National Research Foundation (NRF) of South Africa, and the University of Pretoria (South Africa).

Data archiving statement

The DArT sequences have been submitted to NCBI (http://www.ncbi.nlm.nih.gov/). BioProject PRJNA398959, Supplementary Table 2.

References

Adkins NL, Hall JA, Georgel PT (2007) The use of quantitative agarose gel electrophoresis for rapid analysis of the integrity of protein-DNA complexes. J Biochem Biophys Methods 70: 721-726.

AFFA (2013) Tea market development in Kenya.

Ananingsih VK, Sharma A, Zhou W (2013) Green tea catechins during food processing and storage: a review on stability and detection. Food Res. Int. 50: 469-479.

Bali S, Mamgain A, Raina SN, Yadava SK, Bhat V, Das S, Pradhan AK, Goel S (2015) Construction of a genetic linkage map and mapping of drought tolerance trait in Indian beveragial tea. Mol Breed 35: 1-20.

Banerjee B (1992) Botanical classification of tea. In: Willson K.C. and Clifford M.N. (Eds), Tea: Cultivation to consumption. Chapman and Hall, London, pp 25-51.

Chen L, Zhou ZX, Yang YJ (2007) Genetic improvement and breeding of tea plant (*Camellia sinensis*) in China: from individual selection to hybridization and molecular breeding. Euphytica 154: 239-248.

Churchill GA Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138: 963-971.

da Silva Pinto M (2013) Tea: A new perspective on health benefits. Food Res Int 53: 558-567.

Doerge RW (2002) Mapping and analysis of quantitative trait loci in experimental populations. Nat Rev Genet 3: 43-52.

Ellis R, Nyirenda H (1995) A successful plant improvement programme on tea (*Camellia sinensis*). J Exp Agric 31: 307-323.

Gawal N, Jarret R (1991) A modified CTAB DNA extraction procedure for Musa and Ipomea. Plant Mol Biol Rep 9: 262-266.

Gupta PK, Rustgi S, Mir RR (2013) Array-based high-throughput DNA markers and genotyping platforms for cereal genetics and genomics. In: Cereal Genomics II, Springer, pp. 11-55.

Hackett CA, Wachira FN, Paul S, Powell W, Waugh R (2000) Construction of a genetic linkage map for *Camellia sinensis* (tea). Heredity 85: 346-355.

Hu CY, Lee TC, Tsai HT, Tsai YZ, Lin SF (2013) Construction of an integrated genetic map based on maternal and paternal lineages of tea (*Camellia sinensis*). Euphytica 191: 141-152.

Huang F, Liang Y, Lu J, Chen R (2006) Genetic mapping of first generation of backcross in tea by RAPD and ISSR markers. J Tea Sci 26: 171-176.

ISO14502-2 (2005) Determination of substances characteristic of green and black tea-Part 2: Content of catechins in green tea- method using high-performance liquid chromatography. Geneva: International Organization for Standardization.

Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res 29: e25-e25.

Kamunya S, Wachira F, Pathak R, Korir R, Sharma V, Kumar R, Bhardwaj P, Chalo R, Ahuja P, Sharma R. (2010). Genomic mapping and testing for quantitative trait loci in tea (*Camellia sinensis* (L.) O. Kuntze). Tree genet. Genomes 6: 915-929.

Kamunya S, Wachira F, Pathak R, Muoki R, Wanyoko J, Ronno W, Sharma R (2009) Quantitative genetic parameters in tea (*Camellia sinensis* (L.) O. Kuntze), I: combining abilities for yield, drought tolerance and quality traits. Afr J Plant Sci 3: 093-101.

Karori S, Wachira F, Ngure R, Mireji P (2014) Polyphenolic composition and antioxidant activity of Kenyan Tea cultivars. J Pharmacogn Phytochem 3: 105-116.

Kenya National Bureau of Statistics (2012) Kenya facts and Fig.s 2012. Kenya.

Khan N, Mukhtar H (2007) Tea polyphenols for health promotion. Life sciences 81: 519-533.
Kobayashi-Hattori K, Mogi A, Matsumoto Y, Takita T (2005) Effect of caffeine on the body fat and lipid metabolism of rats fed on a high-fat diet. Biosci Biotechnol Biochem 69: 2219-2223.
Lin SD, Mau JL, Hsu CA (2012) Bioactive components and antioxidant properties of γ-aminobutyric acid (GABA) tea leaves. LWT-Food Sci Technol 46: 64-70.
Liu X, Tang L, Wu H, Xi W, Yu J, Zhou Z (2016) Development of DArT markers and evaluation of phylogenetic relationship of key Citrus species. Genet Resour Crop Evol 63: 1307-1318.
Ma JQ, Yao MZ, Ma CL, Wang XC, Jin JQ, Wang XM, Chen L (2014) Construction of a SSR-based genetic map and identification of QTL for catechins content in tea plant (*Camellia sinensis*). PLoS One 9: e93131.
Ma JQ, Huang L, Ma CL, Jin JQ, Li CF, Wang RK, Zheng HK, Yao MZ, Chen L (2015) Large-scale SNP discovery and genotyping for constructing a high-density genetic map of tea plant using specific-locus amplified fragment sequencing (SLAF-seq). PLoS One 10(6): p.e0128798.
Malebe MP (2011) Method for screening plants for drought tolerance, Vol. Msc, University of Pretoria, Pretoria, South Africa.
Meegahakumbura MK, Wambulwa MC, Thapa KK, Li MM, Möller M, Xu JC, Yang JB, Liu BY, Ranjitkar S, Liu J, Li DZ (2016) Indications for three independent domestication events for the tea plant (*Camellia sinensis* (L.) O. Kuntze) and new insights into the origin of tea germplasm in China and India revealed by nuclear microsatellites. *PloS one*, *11*(5), p.e0155369.
Mondal TK (2007) Tea. Compendium of Transgenic Crop Plants.
Mondal TK, Bhattacharya A, Laxmikumaran M, Ahuja PS (2004) Recent advances of tea (*Camellia sinensis*) biotechnology. Plant Cell Tissue Organ Cult 76: 195-254.
Nitin SL, Islam M, O'Hare WT, Ali Z (2006) Discrimination of teas based on total luminescence spectroscopy and pattern recognition. J Sci Food Agr 86: 2092-2098.
Nyarukowa C, Koech R, Loots T, Apostolides Z (2016) SWAPDT: A method for short-time withering assessment of probability for drought tolerance in *Camellia sinensis* validated by targeted metabolomics. J Plant Physiol 198: 39-48.
Obanda M, Owuor PO, Mang'oka R (2001) Changes in the chemical and sensory quality parameters of black tea due to variations of fermentation time and temperature. Food Chem 75: 395-404.
Obanda M, Owuor PO, Taylor SJ (1997) Flavanol composition and caffeine content of green leaf as quality potential indicators of Kenyan black teas. J Sci Food Agr 74: 209-215.
Orel G, Wilson PG (2012) *Camellia cherryana* (Theaceae), a new species from China. Ann Bot Fenn pp. 248-254.
Ota S, Tanaka J (1999) RAPD-based linkage mapping using F1 segregating populations derived from crossings between tea cultivar 'Sayamakaori'and Strain 'Kana-Ck17'. Breed Res 1: 16.
Owuor P, Obanda M, Apostolides Z, Wright L, Nyirenda H, Mphangwe N (2006a) The relationship between the chemical plain black tea quality parameters and black tea colour, brightness and sensory evaluation. Food Chem 97: 644-653.
Owuor PO, Obanda M, Nyirenda HE, Mphangwe NI, Wright LP, Apostolides Z (2006b) The relationship between some chemical parameters and sensory evaluations for plain black tea

(*Camellia sinensis*) produced in Kenya and comparison with similar teas from Malawi and South Africa. Food Chem 97: 644-653.

Owuor PO, Othieno CO (1991) Response of black tea quality to locality in stakeholder tea farms. Tea 12: 41-45.

Pang Y, Abeysinghe IBS, He J, He X, Huhman D, Mewan KM, Sumner LW, Yun J, Dixon RA (2013) Functional characterization of proanthocyanidin pathway enzymes from tea and their application for metabolic engineering. Plant Physiol 161: 1103-1116.

Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S, Crossa J, Sánchez-Villeda H, Sorrells M (2012) Genomic selection in wheat breeding using genotyping-by-sequencing. The Plant Genome 5: 103-113.

Preedy VR (2012) Tea in health and disease prevention. Academic Press.

Racedo J, Gutiérrez L, Perera MF, Ostengo S, Pardo EM, Cuenya MI, Welin B, Castagnaro AP (2016) Genome-wide association mapping of quantitative traits in a breeding population of sugarcane. BMC Plant Biol 16: 142.

Sall J, Lehman A, Stephens ML, Creighton L (2012) JMP start statistics: a guide to statistics and data analysis using JMP. SAS Institute.

Sánchez-Sevilla JF, Horvath A, Botella MA, Gaston A, Folta K, Kilian A, Denoyes B, Amaya I (2015) Diversity Arrays Technology (DArT) marker platforms for diversity analysis and linkage mapping in a complex crop, the octoploid cultivated strawberry (Fragaria× ananassa). PLoS One 10: e0144960.

Sang S, Lambert JD, Ho CT, Yang CS (2011) The chemistry and biotransformation of tea constituents. Pharmacol Res 64: 87-99.

Sansaloni CP, Petroli CD, Carling J, Hudson CJ, Steane DA, Myburg AA, Grattapaglia D, Vaillancourt RE, Kilian A (2010) A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in Eucalyptus. Plant Methods 6: 1.

Schouten HJ, van de Weg WE, Carling J, Khan SA, McKay SJ, van Kaauwen MP, Wittenberg AH, Koehorst-van Putten HJ, Noordijk Y, Gao Z (2012) Diversity arrays technology (DArT) markers in apple for genetic linkage maps. Mol Breed 29: 645-660.

Sedaghathoor S, Haghighat SR, Shokrgozar SA (2013) Storage period effects on the qualitative characteristics of scented tea. Int J Biosci 3(7): 66-73.

Shi CY, Yang H, Wei CL, Yu O, Zhang ZZ, Jiang CJ, Sun J, Li YY, Chen Q, Xia T (2011) Deep sequencing of the *Camellia sinensis* transcriptome revealed candidate genes for major metabolic pathways of tea-specific compounds. BMC Genom 12: 131.

Steane DA, Nicolle D, Sansaloni CP, Petroli CD, Carling J, Kilian A, Myburg AA, Grattapaglia D, Vaillancourt RE (2011) Population genetic analysis and phylogeny reconstruction in Eucalyptus (Myrtaceae) using high-throughput, genome-wide genotyping. Mol Phylogenet Evo 59: 206-224.

Tan LQ, Wang LY, Xu LY, Wu LY, Peng M, Zhang CC, Wei K, Bai PX, Li HL, Cheng H, Qi GN (2016) SSR-based genetic mapping and QTL analysis for timing of spring bud flush, young shoot color, and mature leaf size in tea plant (*Camellia sinensis*). Tree Genetics and Genomes, 12(3): 1-13.

Tanaka J (2000) Construction of linkage and QTL analysis of tea plant. Proceedings of the International Symposium on Molecular Biology and Tea Breeding, pp. 23-26.

Taniguchi F, Furukawa K, Ota-Metoku S, Yamaguchi N, Ujihara T, Kono I, Fukuoka H, Tanaka, J (2012) Construction of a high-density reference linkage map of tea (*Camellia sinensis*). Breeding Sci 62: 263-273.

Van Ooijen J (2006) JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen 33: 10.1371.

Voorrips R (2002) MapChart: software for the graphical presentation of linkage maps and QTL. J Hered 93: 77-78.

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

Wambulwa MC, Meegahakumbura MK, Kamunya S, Muchugi A, Möller M, Liu J, Xu JC., Ranjitkar S, Li DZ, Gao LM (2016) Insights into the Genetic Relationships and Breeding Patterns of the African Tea Germplasm Based on nSSR Markers and cpDNA Sequences. Front Plant Sci 7.

Wittenberg AH, van der Lee T, Cayla C, Kilian A, Visser RG, Schouten HJ (2005) Validation of the high-throughput marker technology DArT using the model plant *Arabidopsis thaliana*. Mol Genet Genomics 274: 30-39.

Wright LP, Mphangwe NIK, Nyirenda HE, Apostolides Z (2002) Analysis of the theaflavin composition in black tea (*Camellia sinensis*) for predicting the quality of tea produced in Central and Southern Africa. J Sci Food Agr 82: 517-525.

Zou J, Semagn K, Iqbal M, N'Diaye A, Chen H, Asif M, Navabi A, Perez-Lara E, Pozniak C, Yang RC (2016) Mapping QTL Controlling Agronomic Traits in the 'Attila'×'CDC Go'Spring Wheat Population under Organic Management using 90K SNP Array. *Crop Sci*. *57*(1): 365-377.

Supplementary Table 1. DArTseq marker ID with the corresponding individual marker sequence

Marker ID	Marker Sequence (5'-3')
5063001	TGCAGCCTTATTAGTTTCACGTTACGTATGAAAAACTATGCAATTGTCGAATTTGTTTAGGAGCTCCCA
5064391	TGCAGCCATAGTTTGAAGAACAGCAGAGTTTTTGACCCAAGAATCATGCTCTTAGTTTACAATCTGTAT
5064585	TGCAGTTCAACATCATCAGCACTTTTAGTTTTAGACAACACATTTATTGACCAAAAGCTGAGTGCCATG
5072338	TGCAGCTCTGATTCACTTGTTTCTCAATATTTGAACTTCCACGGTATGAAGGTATATATTTCAAATATG
5073424	TGCAGAGAAATTCCTCCACACACCGAGGCTCGTTTGGGAAGTAACGTTTTTACTAAAAAAAAATTACAG
5082606	TGCAGCAAGCAAAAAAAGTGGAATCCCTCGGAACACTGGACTCGCCGAGGCTAGTCAGAGCCTCTTTTG
5084595	TGCAGTTGAATCTGTAAGAGTGAGACACCCATTAGGCACCCAATAACTTTAGAAATTCAGAAGAAAAAT
5085963	TGCAGCTTGTGGAGCTGCTTATAACTTTTTATTTTGAGAATCGGTGTCAGAGTGTTTGGTGCATCTTTT
5087017	TGCAGTAACAAACATTATAATTTTTGTGTTGCATTATAAAAGCAAGACAATACTAATCAAGCCTTGATT
5087113	TGCAGCGGTTGATGCTGTACGTTTACTCACAGCAGCATCTCCTATCCCGACAAGATGGGTTGACTTTTA
5088162	TGCAGATTTTTTAGGACAACAAGATTCTCCAAGGTCTGAATAATTATGAGAACAAAAATATGCATTCAC
5098382	TGCAGGGCATCAAACAGGACATTAGTCACAAGTAATTAGCATAAGCAACGGATACTATAATTACAGATC
5104630	TGCAGCTGTAGATGAGGAGGTACGCTAGTCTAATTTTAGAGGAATAAAGATGGAGAGTTTTACAGATCG
5111164	TGCAGTACCAGATGGACAAGGAGAGAGAAGCGAGTTAGAGGTCTTTACAGATCGGAAGAGCGGTTCAGC
5111268	TGCAGAGGTAATCTACACCTATCAGTTCACTTCCCATATGTCGTTGAGTACAAACAAGGACACAAAGTT
5112599	TGCAGGAAATCCGGCATTTAGAGGCATTTACAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGA
5114089	TGCAGTCCTTTTTGATTGTTTCTGATCCTTCCATCACGAAACATATATTACAGATCGGAAGAGCGGTTC
5114985	TGCAGCAAACACTTGCTAATTTTCACTCTATCTTACCAAAATGAGGCATTACAGATCGGAAGAGCGGTT
5123053	TGCAGGTCATTTTATAATGTGACACATCGGAGATTACAGATCGGAAGAGCGGTTCAGCAGGAATGCCGA

5123475	TGCAGAGCAATTACCACAGTTTATCGCTGCTGGCAACTTACAGATCGGAAGAGCGGTTCAGCAGGAATG
5123751	TGCAGCTGCCTTTCGTCATGTGGCTTGCCATCAAGAAACCAGAGAAGTTTAGCTTGTCTTGGCTTATTA
5123950	TGCAGCCATTGGTGCCTACCTGGAGTTCCATACACATGGAATGAACTTCTGTATGCTCACCTACTGTCA
5124128	TGCAGCTAGAATAATGTCTACCCACAATCAAATATAGATCTTACAGATTATTACAGATCGGAAGAGCGG
5125626	TGCAGTCGGGCCACTTTTTTTTTAGTTGGTGTGAATGTCCCGTAATTTATTACAGATCGGAAGAGCGGTT
5128967	TGCAGAAGAAAGGCACACATGAGCAGCTGTGATGGAATAACCCAGTTGCAATCTACCACATATTACAGA
5130194	TGCAGTGGTGGAAGCTTGTACATCTGGAGTTTTACAGATCCGAAGAGCGGTTCAGCAGGAATGCCGAGA
5130531	TGCAGTCAGCATAACCGGTGCAGCTATCTGCCACATCTTCTTTGATTCATCTCGACATTTTCTGTTACA
5132370	TGCAGCTGTCCTTTCACCAAGTAACGCATATAAGTTTGCATCGACGAGTTGCTATTCCAAATAAGAGCA
5132432	TGCAGGTTTCTTCTTTCACCCAAAAAAAAAGCCCTCATTTTGGTTCAATTTTACAGATCGGAAGAGCGG
5132791	TGCAGCAATTTGCTAATAACCTCCTCATCCTTGCCTTCAAAATCAATGCCTTCAAAAACAAATCTCATA
5134558	TGCAGCTGAAGTGGTAAAGAGGCTTGTTGAAGGGGCTCAACTCCTCATCCATCAAAGATTTACAGATCG
5135436	TGCAGGAAAGGCAAGGGAAATAACAACAAATAATTACAGATCGGAAGAGCGGTTCAGCAGGAATGCCGA
5136045	TGCAGCAATAATTTACTTCTTCTATGTTTGGCTATTACAGAAGCCTTGTTCTGTTCAAACTTCAGCTTT
5136108	TGCAGAACACATGACTCCAATTTGAAGGATACATCCCAAGCACTCCAGATTACAGATCGGAAGAGCGGT
5136623	TGCAGGGACGTTGGCCTATGTTGAGTAGATTACAAATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACC
5136790	TGCAGCTACAAACCAATCCATAGTCACAGCTCAAGGAGTCACTACAGTTACAGACCAAGCCATAAGACC
5136794	TGCAGGTCCAAGCACTTCAACACGTGTGGCTGTTGATTCTGAAGAAGCCATGCCTGATGTTGACAGTGA

Supplementary Table 2. BioProject PRJNA398959

Accession	Sample Name	Organism	Tax ID	Cultivar	Isolate	BioProject
SAMN07542559	1	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 504/1	PRJNA398959
SAMN07542560	2	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/2	PRJNA398959
SAMN07542561	3	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/3	PRJNA398959
SAMN07542562	4	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/4	PRJNA398959
SAMN07542563	5	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/5	PRJNA398959
SAMN07542564	6	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/7	PRJNA398959
SAMN07542565	7	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/8	PRJNA398959
SAMN07542566	8	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/9	PRJNA398959
SAMN07542567	9	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/10	PRJNA398959
SAMN07542568	10	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/11	PRJNA398959
SAMN07542569	11	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/12	PRJNA398959
SAMN07542570	12	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/13	PRJNA398959
SAMN07542571	13	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/15	PRJNA398959
SAMN07542572	14	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/16	PRJNA398959
SAMN07542573	15	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/17	PRJNA398959
SAMN07542574	16	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/18	PRJNA398959
SAMN07542575	17	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/19	PRJNA398959
SAMN07542576	18	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/21	PRJNA398959
SAMN07542577	19	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/22	PRJNA398959
SAMN07542578	20	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/23	PRJNA398959
SAMN07542579	21	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/24	PRJNA398959
SAMN07542580	22	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/27	PRJNA398959
SAMN07542581	23	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/28	PRJNA398959
SAMN07542582	24	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/31	PRJNA398959
SAMN07542583	25	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/32	PRJNA398959
SAMN07542584	26	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/33	PRJNA398959
SAMN07542585	27	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/34	PRJNA398959
SAMN07542586	28	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/35	PRJNA398959
SAMN07542587	29	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/36	PRJNA398959
SAMN07542588	30	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/37	PRJNA398959
SAMN07542589	31	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/38	PRJNA398959
SAMN07542590	32	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/39	PRJNA398959
SAMN07542591	33	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/40	PRJNA398959
SAMN07542592	34	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/41	PRJNA398959
SAMN07542593	35	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/42	PRJNA398959
SAMN07542594	36	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/43	PRJNA398959
SAMN07542595	37	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/44	PRJNA398959

SAMN07542596	38	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/45	PRJNA398959
SAMN07542597	39	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/47	PRJNA398959
SAMN07542598	40	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/48	PRJNA398959
SAMN07542599	41	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/49	PRJNA398959
SAMN07542600	42	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/50	PRJNA398959
SAMN07542601	43	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/51	PRJNA398959
SAMN07542602	44	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/53	PRJNA398959
SAMN07542603	45	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/55	PRJNA398959
SAMN07542604	46	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/57	PRJNA398959
SAMN07542605	47	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/58	PRJNA398959
SAMN07542606	48	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/59	PRJNA398959
SAMN07542607	49	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/60	PRJNA398959
SAMN07542608	50	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/61	PRJNA398959
SAMN07542609	51	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/62	PRJNA398959
SAMN07542610	52	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/63	PRJNA398959
SAMN07542611	53	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/65	PRJNA398959
SAMN07542612	54	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/68	PRJNA398959
SAMN07542613	55	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/69	PRJNA398959
SAMN07542614	56	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/70	PRJNA398959
SAMN07542615	57	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/71	PRJNA398959
SAMN07542616	58	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/73	PRJNA398959
SAMN07542617	59	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/74	PRJNA398959
SAMN07542618	60	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/75	PRJNA398959
SAMN07542619	61	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/77	PRJNA398959
SAMN07542620	62	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/78	PRJNA398959
SAMN07542621	63	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/79	PRJNA398959
SAMN07542622	64	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/80	PRJNA398959
SAMN07542623	65	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/81	PRJNA398959
SAMN07542624	66	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/82	PRJNA398959
SAMN07542625	67	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/83	PRJNA398959
SAMN07542626	68	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/85	PRJNA398959
SAMN07542627	69	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/86	PRJNA398959
SAMN07542628	70	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/88	PRJNA398959
SAMN07542629	71	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/89	PRJNA398959
SAMN07542630	72	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/91	PRJNA398959
SAMN07542631	73	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/92	PRJNA398959
SAMN07542632	74	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/93	PRJNA398959
SAMN07542633	75	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/94	PRJNA398959
SAMN07542634	76	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/95	PRJNA398959
SAMN07542635	77	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/96	PRJNA398959

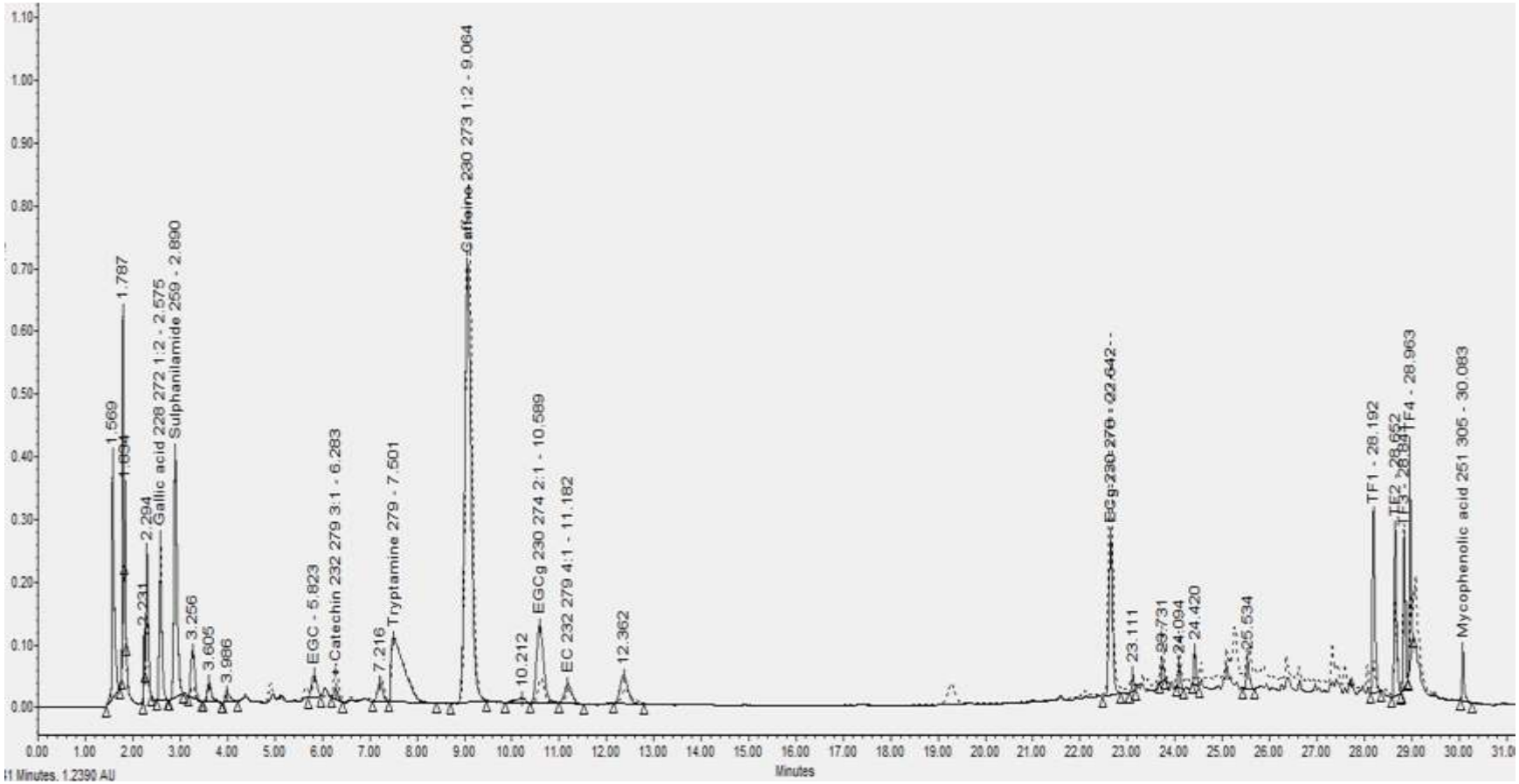
SAMN07542636	78	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/97	PRJNA398959
SAMN07542637	79	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/98	PRJNA398959
SAMN07542638	80	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/99	PRJNA398959
SAMN07542639	81	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/100	PRJNA398959
SAMN07542640	82	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/101	PRJNA398959
SAMN07542641	83	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/102	PRJNA398959
SAMN07542642	84	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/103	PRJNA398959
SAMN07542643	85	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/104	PRJNA398959
SAMN07542644	86	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/105	PRJNA398959
SAMN07542645	87	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/106	PRJNA398959
SAMN07542646	88	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/107	PRJNA398959
SAMN07542647	89	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/108	PRJNA398959
SAMN07542648	90	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/109	PRJNA398959
SAMN07542649	91	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/110	PRJNA398959
SAMN07542650	92	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/111	PRJNA398959
SAMN07542651	93	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/112	PRJNA398959
SAMN07542652	94	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/113	PRJNA398959
SAMN07542653	95	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/115	PRJNA398959
SAMN07542654	96	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/116	PRJNA398959
SAMN07542655	97	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/118	PRJNA398959
SAMN07542656	98	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/119	PRJNA398959
SAMN07542657	99	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/120	PRJNA398959
SAMN07542658	100	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/121	PRJNA398959
SAMN07542659	101	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/122	PRJNA398959
SAMN07542660	102	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/124	PRJNA398959
SAMN07542661	103	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/125	PRJNA398959
SAMN07542662	104	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/126	PRJNA398959
SAMN07542663	105	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/128	PRJNA398959
SAMN07542664	106	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/129	PRJNA398959
SAMN07542665	107	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/131	PRJNA398959
SAMN07542666	108	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/133	PRJNA398959
SAMN07542667	109	*Camellia sinensis*	4442	*Camellia sinensis*	TFRK 504/134	PRJNA398959
SAMN07542668	110	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/1	PRJNA398959
SAMN07542669	111	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/2	PRJNA398959
SAMN07542670	112	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/3	PRJNA398959
SAMN07542671	113	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/4	PRJNA398959
SAMN07542672	114	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/5	PRJNA398959
SAMN07542673	115	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/6	PRJNA398959
SAMN07542674	116	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/7	PRJNA398959
SAMN07542675	117	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/8	PRJNA398959

SAMN07542676	118	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/9	PRJNA398959
SAMN07542677	119	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/10	PRJNA398959
SAMN07542678	120	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/11	PRJNA398959
SAMN07542679	121	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/13	PRJNA398959
SAMN07542680	122	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/15	PRJNA398959
SAMN07542681	123	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/16	PRJNA398959
SAMN07542682	124	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/17	PRJNA398959
SAMN07542683	125	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/18	PRJNA398959
SAMN07542684	126	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/19	PRJNA398959
SAMN07542685	127	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/22	PRJNA398959
SAMN07542686	128	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/23	PRJNA398959
SAMN07542687	129	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/25	PRJNA398959
SAMN07542688	130	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/26	PRJNA398959
SAMN07542689	131	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/27	PRJNA398959
SAMN07542690	132	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/28	PRJNA398959
SAMN07542691	133	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/29	PRJNA398959
SAMN07542692	134	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/30	PRJNA398959
SAMN07542693	135	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/31	PRJNA398959
SAMN07542694	136	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/32	PRJNA398959
SAMN07542695	137	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/34	PRJNA398959
SAMN07542696	138	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/35	PRJNA398959
SAMN07542697	139	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/36	PRJNA398959
SAMN07542698	140	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/38	PRJNA398959
SAMN07542699	141	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/39	PRJNA398959
SAMN07542700	142	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/40	PRJNA398959
SAMN07542701	143	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/41	PRJNA398959
SAMN07542702	144	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/42	PRJNA398959
SAMN07542703	145	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/43	PRJNA398959
SAMN07542704	146	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/44	PRJNA398959
SAMN07542705	147	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/45	PRJNA398959
SAMN07542706	148	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/48	PRJNA398959
SAMN07542707	149	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/49	PRJNA398959
SAMN07542708	150	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/50	PRJNA398959
SAMN07542709	151	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/51	PRJNA398959
SAMN07542710	152	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/52	PRJNA398959
SAMN07542711	153	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/53	PRJNA398959
SAMN07542712	154	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/55	PRJNA398959
SAMN07542713	155	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/56	PRJNA398959
SAMN07542714	156	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/57	PRJNA398959
SAMN07542715	157	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/58	PRJNA398959

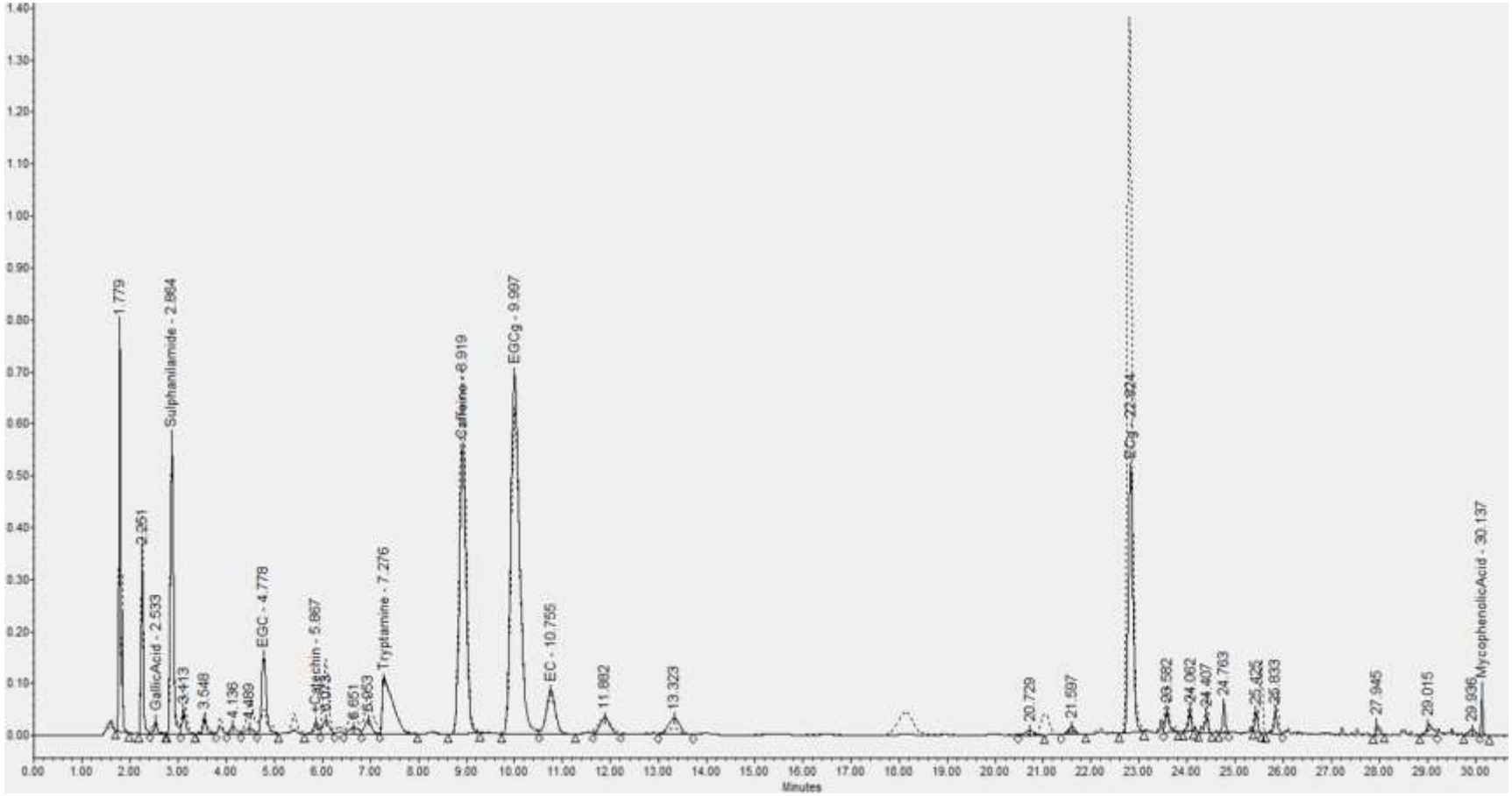
SAMN07542716	158	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/61	PRJNA398959
SAMN07542717	159	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/62	PRJNA398959
SAMN07542718	160	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/63	PRJNA398959
SAMN07542719	161	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/64	PRJNA398959
SAMN07542720	162	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/65	PRJNA398959
SAMN07542721	163	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/66	PRJNA398959
SAMN07542722	164	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/67	PRJNA398959
SAMN07542723	165	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/68	PRJNA398959
SAMN07542724	166	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/69	PRJNA398959
SAMN07542725	167	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/70	PRJNA398959
SAMN07542726	168	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/71	PRJNA398959
SAMN07542727	169	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/72	PRJNA398959
SAMN07542728	170	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/73	PRJNA398959
SAMN07542729	171	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/74	PRJNA398959
SAMN07542730	172	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/75	PRJNA398959
SAMN07542731	173	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/76	PRJNA398959
SAMN07542732	174	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/77	PRJNA398959
SAMN07542733	175	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/78	PRJNA398959
SAMN07542734	176	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/79	PRJNA398959
SAMN07542735	177	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/80	PRJNA398959
SAMN07542736	178	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/81	PRJNA398959
SAMN07542737	179	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/82	PRJNA398959
SAMN07542738	180	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/83	PRJNA398959
SAMN07542739	181	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/84	PRJNA398959
SAMN07542740	182	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/85	PRJNA398959
SAMN07542741	183	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/86	PRJNA398959
SAMN07542742	184	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/87	PRJNA398959
SAMN07542743	185	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/93	PRJNA398959
SAMN07542744	186	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/94	PRJNA398959
SAMN07542745	187	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/96	PRJNA398959
SAMN07542746	188	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/97	PRJNA398959
SAMN07542747	189	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/102	PRJNA398959
SAMN07542748	190	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/104	PRJNA398959
SAMN07542749	191	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/105	PRJNA398959
SAMN07542750	192	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/106	PRJNA398959
SAMN07542751	193	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/108	PRJNA398959
SAMN07542752	194	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/113	PRJNA398959
SAMN07542753	195	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/114	PRJNA398959
SAMN07542754	196	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/115	PRJNA398959
SAMN07542755	197	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/118	PRJNA398959

SAMN07542756	198	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/119	PRJNA398959
SAMN07542757	199	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/121	PRJNA398959
SAMN07542758	200	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/122	PRJNA398959
SAMN07542759	201	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/123	PRJNA398959
SAMN07542760	202	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/124	PRJNA398959
SAMN07542761	203	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/125	PRJNA398959
SAMN07542762	204	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/126	PRJNA398959
SAMN07542763	205	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/127	PRJNA398959
SAMN07542764	206	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/130	PRJNA398959
SAMN07542765	207	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/131	PRJNA398959
SAMN07542766	208	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/132	PRJNA398959
SAMN07542767	209	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/134	PRJNA398959
SAMN07542768	210	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/136	PRJNA398959
SAMN07542769	211	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/137	PRJNA398959
SAMN07542770	212	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/138	PRJNA398959
SAMN07542771	213	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/139	PRJNA398959
SAMN07542772	214	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/140	PRJNA398959
SAMN07542773	215	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/141	PRJNA398959
SAMN07542774	216	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/142	PRJNA398959
SAMN07542775	217	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/143	PRJNA398959
SAMN07542776	218	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/144	PRJNA398959
SAMN07542777	219	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/145	PRJNA398959
SAMN07542778	220	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/146	PRJNA398959
SAMN07542779	221	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/147	PRJNA398959
SAMN07542780	222	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/149	PRJNA398959
SAMN07542781	223	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/150	PRJNA398959
SAMN07542782	224	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/154	PRJNA398959
SAMN07542783	225	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/155	PRJNA398959
SAMN07542784	226	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/156	PRJNA398959
SAMN07542785	227	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/157	PRJNA398959
SAMN07542786	228	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/159	PRJNA398959
SAMN07542787	229	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/160	PRJNA398959
SAMN07542788	230	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/162	PRJNA398959
SAMN07542789	231	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/163	PRJNA398959
SAMN07542790	232	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/165	PRJNA398959
SAMN07542791	233	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/166	PRJNA398959
SAMN07542792	234	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/167	PRJNA398959
SAMN07542793	235	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/168	PRJNA398959
SAMN07542794	236	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/169	PRJNA398959
SAMN07542795	237	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/170	PRJNA398959

SAMN07542796	238	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/172	PRJNA398959
SAMN07542797	239	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/173	PRJNA398959
SAMN07542798	240	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/175	PRJNA398959
SAMN07542799	241	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/176	PRJNA398959
SAMN07542800	242	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/177	PRJNA398959
SAMN07542801	243	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/178	PRJNA398959
SAMN07542802	244	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/179	PRJNA398959
SAMN07542803	245	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/182	PRJNA398959
SAMN07542804	246	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/183	PRJNA398959
SAMN07542805	247	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/184	PRJNA398959
SAMN07542806	248	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/185	PRJNA398959
SAMN07542807	249	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/186	PRJNA398959
SAMN07542808	250	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/187	PRJNA398959
SAMN07542809	251	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/188	PRJNA398959
SAMN07542810	252	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/189	PRJNA398959
SAMN07542811	253	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/190	PRJNA398959
SAMN07542812	254	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/192	PRJNA398959
SAMN07542813	255	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/194	PRJNA398959
SAMN07542814	256	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/195	PRJNA398959
SAMN07542815	257	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/196	PRJNA398959
SAMN07542816	258	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/197	PRJNA398959
SAMN07542817	259	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/198	PRJNA398959
SAMN07542818	260	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/199	PRJNA398959
SAMN07542819	261	*Camellia sinensis*	4442	*Camellia sinensis*	TRFK 524/200	PRJNA398959



Supplementary Fig. 1. A representative UPLC chromatogram of black tea sample from parental lines TRFK 303/577 (solid line) and GW Ejulu (dotted line) with the three internal standards.
KEY: X-axis = Retention time (minutes), Y-axis = Absorbance units (278 nm)



Supplementary Fig. 2. A representative UPLC chromatogram of green tea sample from parental lines TRFK 303/577 (solid line) and GW Ejulu (dotted line) with the three internal standards.
KEY: X-axis = Retention time (minutes), Y-axis = Absorbance units (278 nm)