International Journal of Agriculture Innovations and Research
Volume 7, Issue 6, ISSN (Online) 2319-1473
Manuscript Processing Details (dd/mm/yyyy):
Received: 13/05/2019 | Accepted on: 10/06/2019 | Published: 24/06/2019

Application of Phenomics, Genomic Resources and Bioinformatics Tools for Tea Plant Improvement

MRITYUNJAY CHOUBEY[1*], URVASHI LAMA[2], P. CHETRI[3] and B. BERA[4]
[1*]Senior Scientific Officer, Agro-Botany Division, Darjeeling Tea Research & Development Centre, Tea Board, Kurseong - 734203, Darjeeling, West Bengal, India.
[2]Junior Research Fellow, Tea Genome Sequencing Project, Darjeeling Tea Research & Development Centre, Tea Board, Kurseong-734203, Darjeeling, West Bengal, India.
[3]Former Project Director (I/C), Darjeeling Tea Research & Development Centre, Tea Board, Kurseong-734203, Darjeeling, West Bengal, India.
[4]Director (Research), Darjeeling Tea Research & Development Centre, Tea Board, Kurseong-734203, Darjeeling, West Bengal, India.
*Corresponding author email id: choubeyssoabtb@gmail.com

Abstract – **Tea (*Camellia sinensis L. (O) Kuntze*) is the most widely consumed health drink. Due to self incompatibility and highly heterozygous nature along with large genome size of about 4.0 Gigabase, the genomic research on tea is not been explored as much as conducted in other annual crops. Nowadays, advances in genome sequencing technologies are revolutionizing the routine research work of plant genetics. This will pave the way for identification of genes responsible for various agronomically important traits, development of improved tea cultivars, tolerant to biotic and a biotic stress. However, Next generation sequencing (NSG) data analysis in tea is an emerging field of tea research. Hence, the use of bio-informatics tools to evaluate and interpret NGS data is a challenge for the researcher. Analysis of this large amount of heterogeneous genomic data requires effective use and application of bioinformatics tools that must include unified functional annotation, structural search and comprehensive analysis and identification of new genes with fully sequenced genomes. This review discusses momentarily about the applications of phenomics, genomics resources and bio-informatics in next generation sequencing data analysis, with an emphasis on those relevant to tea research. Some fundamental issues related to biological sequence analysis, data basis and software tools are also focussed that facilitate an understanding of the biological processes with the goal to serve as an important information platform for genomic studies which will effectively improve the efficiency of breeding programmes in tea.**

Keywords – **Bioinformatics, *Camellia,* Marker Assisted Selection, Molecular Markers, Next-Generation Sequencing, Phenomics, Tea Genome.**

I. INTRODUCTION

Tea (*Camellia sinensis L. (O) Kuntze*) is one of the most widely consumed beverages having medicinal, anti-oxidative and anti-microbial properties. Apart from this it is an important commercial crop generating employment for a large number of people (Shrivastava et al. 2018). Despite the popularity and economic importance of tea, there have been limited studies on its molecular marker assisted crop improvement programs or genomic research on tea. Most tea clones are diploid with 2n = 2x = 30 chromosomes (Roy 2006) and genome of diploid tea is about 4.0 Gigabase in size (Tanaka and Taniguchi, 2006). Owing to its complex life cycle, due to long gestation period, highly heterozygous and complex chromosomal structure the research on the genome of tea is still in the preliminary stage. Numerous efforts have been made at present to develop and use DNA markers like restriction fragment length polymorphism (RFLP), randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), cleaved and amplified polymorphic sequences (CAPS) and inter simple sequence repeats (ISSR) for characterization of germplasm and genetic enhancement in tea (Sui et al. 2008; Raina et al. 2011). Simple Sequence Repeat (SSR) markers have been frequently used in *Camellia sinensis* for genetic diversity and population structure analysis, evaluation of phylogeny and genetic origin of tea species and construction of genetic linkage maps (Fang et al. 2012; Taniguchi et al. 2012; Yao et al. 2012; Ma et al. 2014; Bali et al. 2015; Tan et al. 2016; Wambulwa et al. 2016).

Nowadays, advances in genome sequencing technologies are revolutionizing the routine research activities of plant genetics. Next generation genome sequencing technologies such as Illumina's genome analyzer, Roche's 454 sequencing platforms have opened the way to tackle sequencing of large genomes like those of barley and wheat which would be impossible to address by Sanger sequencing. Due to the brisk turn down of costs per base pair, next generation sequencing (NGS) is now affordable even for small to midsized laboratories with an overwhelming increase in efficiency and throughput. The combination of conventional breeding techniques with genomic approaches is leading to a new genomics-based plant breeding and accelerating the breeding process in many ways e.g. marker assisted selection, gene pyramiding, etc.

While the complete NGS technologies' data analysis process is complex, which includes multiple steps of analysis dependent on a multitude of programs and databases and involves handling large amounts of heterogeneous data. With the success of NGS base genome experimentation, abundance of tools has been created to support specific parts of the analysis workflow. The appropriate choice of tools is a non-trivial task, especially for inexperienced users. Bioinformatics is the application of computational tools to predict, manage and interpret biological data (Swindells et al., 2002). Bioinformatics has evolved such that the most pressing task now involves the analysis and interpretation of various types of data, including nucleotide, amino acid sequences, protein domains and protein structures (Wales, 2009). Although, there remains a gap between genome data production and next-generation crop improvement strategies, but this is being rapidly closed by far sighted companies and individuals with the ability to combine the ability to mine the genomic data with practical crop-improvement skills (Lai et al., 2012).

The developments in technologies and instrumentation have recently generated many databases for genomic, on a range of platforms and to suite different purposes, like SOL Genomics Network (http: //solgenomics.net/), Auto SNPdb (http: //autosnpdb.appliedbioinformatics.com.), SSR taxonomy tree (http: //appliedbioinformatics.com.au/ projects/ssrtaxono my/php/), Plant mRNA database (PMRD) (http: //bioinformatics.cau.edu.cn/PMRD/) to PMRD integrates the available large information of plant mRNA data consisting of mrNA sequence and their target genes, Arabidopsis small RNA database (http: //mpss.udel.edu/at_sRNA/) to provide sequencing by synthesis based small RNA data, etc (Agarwal and Narayan 2015).

II. PLANT PHENOTYPING AND PHENOMICS FOR PLANT BREEDING

Phenotype which is the outward expression of a particular genotype refers to the physical characteristics of the plant as a result of due interaction of the plant's genes with the environmental factors. As genome is the complete set of genes of an organism, similarly, phenome is the complete set of morphological, anatomical, ontological, physiological, biochemical and organoleptic characteristics or traits expressed by an organism. The study and characterization of these characteristics is called phenomics. The term "phenomics" was coined by Steven A. Garan in the year 1996. Genomics and phenomics are field of studies which are gaining multitude importance in plant breeding. "Plant Breeding is the art and science of changing the traits of plants in order to product desired characteristics" *(Poehlman Sleper, Breeding Field Crops, 1995).* The challenging objective for today's plant breeders is to develop cultivars that can adapt well to the changing global environmental conditions, gives the desired traits or characteristics and also sustains the characteristics over generations. Tea growers, especially of Darjeeling region, face the adversities of climatic changes in hills with change in precipitation pattern and sunlight exposure since the past decade.

Although, genome sequencing has become advanced, fast, easy and cost effective with next-generation sequencing also called as 'high throughput sequencing', plant phenomics still faces challenges in providing similar fast and cost effective high-dimensional phenotypic data and functions also called as 'high - throughput phenotyping'. On the other hand, conventional phenotyping is both labor-intensive and expensive, limiting the number of tests under field conditions to certain periods or seasons. Following the trends in genomics, phenomics platforms are now being developed worldwide for screening purposes as well as to provide new insights into gene function and environmental responses (e.g. Granier et al., 2006; Berger et al., 2010; Furbank and Tester, 2011). Phenomics as a field is growing in importance and technological advancements thereby calling this '*transitional period of change*'as the '*period of phenomics'*. However, more effort is required to map genotype-phenotype relationships for the global development of crop breeding (Tester and Langridge, 2010, Rahaman et al., 2015).

There is a wide difference between the quantity of and time usage for creation of genotypic data and phenotypic data which is a bottleneck in the understanding of the continuous interaction of gene and the environment. The phenotyping bottleneck has therefore aroused the need for robust and vigorous creation of phenotypic data - morphological, physiological and biochemical at high-throughput level and thereby paving way for the development of noble phenotyping technologies.

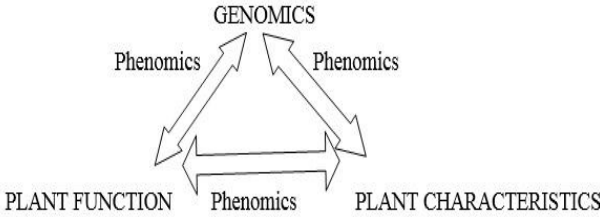


Fig. 1. Phenomics acting as the connecting link between genomics, plant function and plant characteristics

'High-throughput phenotyping' employs advanced robotics, high-tech sensors, imaging systems and computing power to screen the selected phenotypes. Advanced bioinformatics tools further facilitate the analysis of large-scale multi-dimensional, high-resolution data collected through phenotyping. (Furbank and Tester 2011)

Phenomics Ontology Driven Database, abbreviated as PODD, is an ontology - driven data repository for collaborative phenomics research, developed by Australian Plant Phenomics Facility (APPF) that utilizes different softwares for generation of data. International Plant Phenotyping Network (IPPN), an association representing the major plant phenotyping centers, is an International Plant Phenomics Initiative. Other high - throughput phenotyping tools are digital growth analysis, carbon isotope discrimination (CID) and infrared thermography. Sensors used in Infrared thermography detect the plant's temperature from long wavelength infrared radiation according to the relation between plant's body temperature and the light spectrum emitted from the plant's body. Thermography can be effectively used in phenotyping for disease-resistant plant genotypes – one of the most important objectives in tea breeding.

Various imaging techniques like visible light imaging, thermo graphic imaging, hyperspectral imaging, chlorophyll fluorescence, X-Ray, MRI, PET have been applied in crops like rice, wheat, maize, barley to fruits and vegetables like tomatoes and radish respectively, for phenotyping of various growth and developmental processes. Quantitative Phase Imaging (QPI) technique and UAV-enabled phenotyping are some latest high-throughput advancements that can be rendered as very effective phenotyping approaches.

High-throughput phenotyping provides multi-categorical phenotypic traits, and corresponding trait analysis which is essential for the understanding of (a) stress resistance, (b) insect and disease resistance and for the (c) yield and quality improvement (Yang et al., 2013).

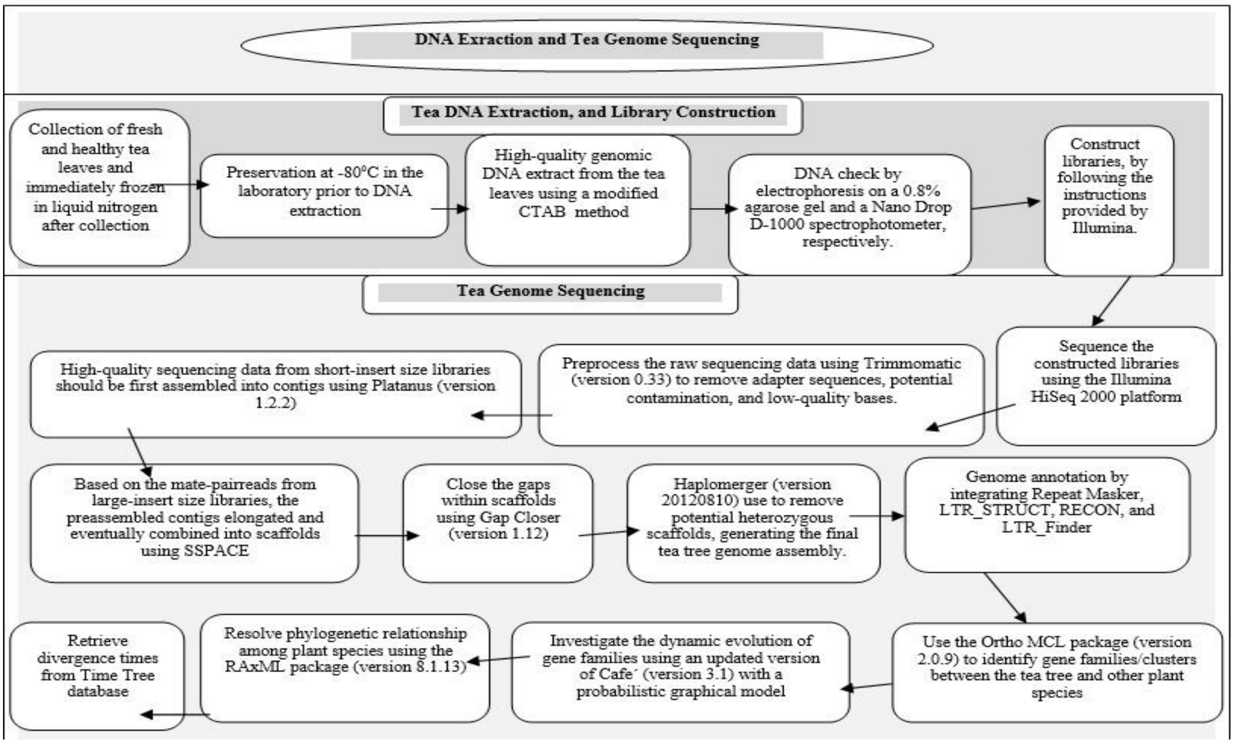
III. Bioinformatics Tools to Retrieve Biological Data

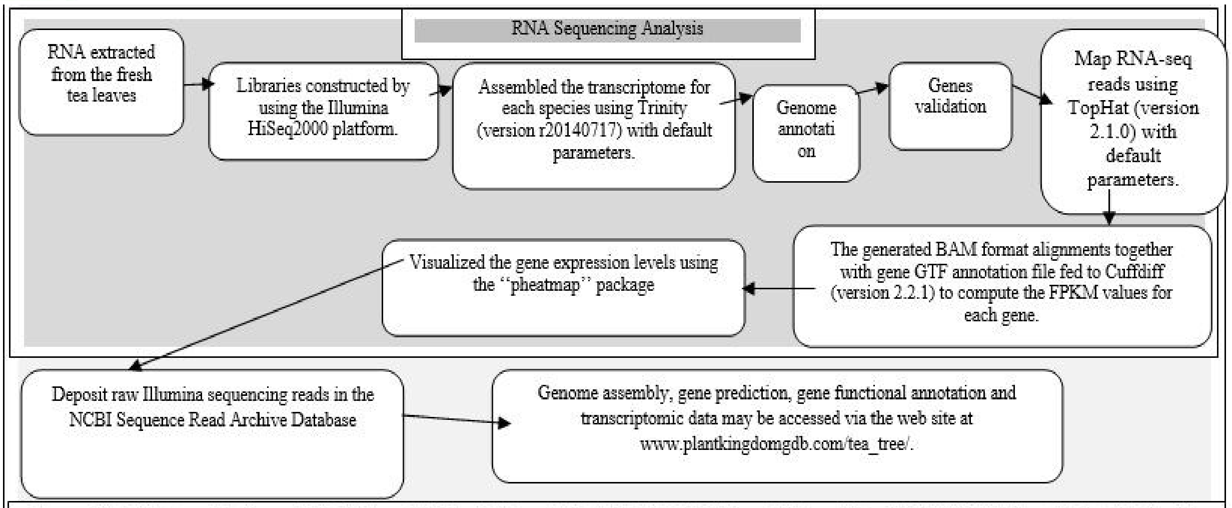
Several bioinformatics tools have been developed to address diverse biological queries. Biological databases are records of consistent data that are stored in a uniform and efficient manner. The main objective of the development of a database is to organize data in a set of structured records to enable simple retrieval of information. To retrieve a particular record from the database, one can specify a particular piece of information, to found in a particular field and expect the computer to retrieve the whole data record. (Kushwaha et al, 2017). Mainly database are divided in two types of database one is the primary while another is secondary database. Primary databases contain information and annotation of DNA and protein sequences, DNA and protein structures and DNA and protein expression profiles. While, secondary databases consist the outcome of analysis on the primary resources including information on sequence patterns, variants, mutations and evolutionary relationships (Rao et al., 2008).

Biologists have a higher level of requirement of data retrieval, which require the identification of connections between pieces of information that were not known when the information was first entered (Xiong, 2009). With the advent of advanced next-generation sequencing technology, the maintenance of this data is becoming an mounting challenge (Batley and Edwards, 2009; Lee et al. 2012). Consequently, the skill to search the huge quantity of data is made viable by the development of custom databases such as TAGdb (http: //flora.acpfg.com.au/ tagdb/) (Marshall et al 2010), but it is increasingly the assembled and annotated genomic data which are applied for crop-improvement applications (Edwards and Batley, 2010). Although genome and transcript sequence information makes up the bulk of genomic data maintained within public databases, which often differ between individuals and varieties that are most helpful for crop-improvement programme. A major emphasis of crop genetic research activities recently is the development of molecular genetic markers associated with important traits. Genetic markers can be assayed with a range of techniques (Duran et al. 2009).

IV. Tea Genome Sequencing and Assembly

As tea plant is self-incompatible little genomic information is currently available. A number of strategies should be adopted to obtain sufficient coverage of expressed transcripts, to improve the accuracy of de novo assembly and to increase the effectiveness of the gene annotations. One of them, first fresh and healthy leaves harvest for DNA / RNA extraction, than construct libraries by following the instructions are provided by Illumina. A paired end library sequencing strategy should be applied to increase the sequencing depth, along with to improve the efficiency of de novo assembly. Finally, all the public databases should be selected for gene annotation comparisons in order to acquire complete functional information (Figure: 2).





Source: Xia EH, Zhang HB, Sheng J, Li1 K, Zhang QJ, Kim C, Zhang Y, Liu Y, Zhu T, Li W, Huang H, Tong Y,Nan H, Shi C, Shi C, Jiang JJ, Mao SY, Jiao JY, Zhang D, Zhao Y, Zhao YJ, Zhang LP, Liu YL,Liu BY,Yu Y, Shao SF, Ni DJ, Eichler EE and Gao LZ; The Tea Tree Genome Provides Insights into Tea Flavor and Independent Evolution of Caffeine Biosynthesis; Molecular Plant 10, 866–877, June 2017.

Fig. 2. Overview of Tea Genome Sequencing and Assembly

The main challenges facing crop bioinformatics researchers is to make the ever increasing volume and types of data available in a suitable format for analysis (Berkman et al. 2012). This includes new high-throughput plant phenotype data as well as the increasing volumes of genetic diversity data. So the task is to associate the huge volume of diversity data with heritable phenotypes which will likely drive genome database development in the coming years (Edwards & Batley 2004, Duran 2010). These databases therefore will require the implementation of appropriate statistical tools for association of high-density genotype and high-throughput phenotype data (Lai et al., 2012). This will cover the approaches for development of improved tea planting materials with the help of identification of genes responsible for various agronomically important traits.

V. GENOMICS BASE MARKER ASSISTED SELECTION

Genomics base marker assisted selection is mainly important in situations, where the accuracy of conventional selection is low for particular agricultural traits, time consuming and often expensive. Marker assisted selection is an approach that has been developed to avoid the problems associated with conventional plant breeding programme, replacing the selection of the phenotype by selection of genes, either directly or indirectly (Schuster 2011, Francia et al 2005). Genomic information tea plantation is limited at present that hinders molecular genetic studies. The modern biological techniques based on genomic sequencing, DNA molecular markers and genetic maps will enable researchers to understand genetic structures of agricultural important traits and improve crops rapidly and precisely (Tan et al 2013, Collard et al 2008). At present, numerous molecular markers have been developed and extensively applied in genetic and genomics analysis in tea for example restriction fragment length polymorphism (RFLP), randomly amplified polymorphic DNA (RAPD), inter simple sequence repeats (ISSR), cleaved and amplified polymorphic sequences (CAPS), amplified fragment length polymorphism to study the genetic diversity of tea, germplasm characterization and genetic enhancement in tea (Liu et al 2017,Yang et al 2016, Bali et al 2013, Sui et al 2008, Raina et at 2011).

The development in plant genomics technology presently, single nucleotide polymorphisms (SNPs) can be easily identified by sequence comparisons of both alleles of a diploid genome, expressed sequence tags (ESTs) and unigenes derived from transcriptome sequences (Yang 2016). The huge classes of polymorphisms in plant genomes SNPs have various advantages that are most important to their use as marker of choice (Fang et al 2014). SNPs have now a day become the preferred marker type for comparative genetic studies due to their low cost, high genotyping efficiency, genome wide coverage and analytical simplicity (Yang 2016). The SNP markers, in combination with a high-throughput genotyping protocol, well establish and verify specific DNA fingerprints for tea varieties. In addition this method provides a powerful tool for variety authentication and quality control for the tea industry. Besides this it is also very useful in the management of tea genetic resources and tea breeding programme, where accurate and efficient genotype identification is essential (Fang et al 2014).

VI. CONCLUSIONS

Use of field based phenotypic information/ data and generation of field data using phenomic technology is becoming very crucial these days. In order to associate these phenotypic/ phenomics data with the NSG technology based genomic resources data application of bioinformatics tools are extremely important for the marker assisted selection in tea for developing superior tea cultivars.

REFERENCES

[1] Agarwal R, Narayan J (2015) Unravelling the impact of bioinformatics and omics in agriculture. Int J Plant Biol Res 3 (2) : 1039.

[2] Bali S, Mamgain A, Raina SN, Yadava SK, Bhat V, Das S, Pradhan AK, Goel S (2015) Construction of a genetic linkage.

[3] Bali S, Raina S N, Bhat V, Aggarwal R K, Goel S (2013), Development of a set of genomic microsatellite markers in tea (*Camellia* L) (Camelliaceae). Mol Breed 32 : 735-741.

[4] Berger, B. et al., High throughput imaging to study drought responses, J. Exp. Bot., 2010, 61, 3519–3528.

[5] Berkman, P.J.; Lai, K.; Lorenc, M.T.; Edwards, D. Next generation sequencing applications for wheat crop improvement. *Amer. J. Bot*. 2012, *99*, 365–371.

[6] Batley, J.; Edwards, D. Genome sequence data: Management, storage, and visualization. *Bio-techniques* 2009, *46*, 333–336.

[7] Collard BCY, Mackill DJ (2008) Marker assisted selection : an approach for precision plant breeding in the twenty first century. Philos T R Soc B 363: 557-572. Doi: 10. 1098/rstb.2007.2170. pubMed:17715053

[8] Duran, C.; Eales, D.; Marshall, D.; Imelfort, M.; Stiller, J.; Berkman, P.J.; Clark, T.; McKenzie, M.; Appleby, N.; Batley, J.; *et al*. Future tools for association mapping in crop plants. *Genome* 2010, *53*, 1017–1023.

[9] Duran, C.; Edwards, D.; Batley, J. Molecular marker discovery and genetic map visualization. In *Applied Bioinformatics*; Edwards, D., Hanson, D., Stajich, J., Eds.; Springer: New York, NY, USA, 2009.

[10] Edwards, D.; Batley, J., Plant Bioinformatics: From genome to phenome. *Trends Biotech.* 2004, *22*, 232–237.

[11] Edwards, D.; Batley, J. Plant genome sequencing: Applications for crop improvement. *Plant Biotechnol. J.* 2010, *7*, 1–8.

[12] Fang W, Li X, Cheng H, Duan Y, Jiang X (2012) Genetic diversity and relationship of clonal Tea (*Camellia sinensis*) cultivars in China as revealed by SSR markers. Plant Syst Evol 298:469-483.

[13] Fang W.P, Meinhardt L.W., Tan H. W., Zhou L., Mischke S and Zhang D (2014) Varietal identification of tea (*Camellia sinensis)* using nano fluidic array of single nucleotide polymorphism (SNP). Horticulture Research (2014) 1, 14035; Nanjing Agricultural University.

[14] Francia E, Tacconi G, Crosatti C, Brarabaschi D, Bulgarelli D, Aglio E D and Vale G (2005); Marker assisted selection in crop plants.; Plant Cell, Tissue and organ Culture (2005) 82 : 317-342.

[15] Furbank R T, Tester M, Phenomics – technologies to relieve the phenotyping bottleneck, Trends in plant science, 2011, 16:12.

[16] Granier C, Aguirrezabal L, Chenu K, Cookson SJ, Dauzat M, Hamard P, Thioux JJ, Rolland G, Bouchier-Combaud S, Lebaudy A, Muller B, Simonneau T, Tardieu F, PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit, New Phytol., 2006, 169(3):623-35.

[17] Kushwaha U.K.S., Deo I., Jaiswal J.P., and Prasad B. Role of Bioinformatics in Crop Improvement. GJSFR: 2017, Vol.17-1; ISSN: 2249-4626.

[18] Lai Kaitao, Michał T. Lorenc and David Edwards (2012) Genomic Databases for Crop Improvement, Agronomy 2012, 2, 62-73; doi:10.3390/agronomy2010062

[19] Lee, H.; Lai, K.; Lorenc, M.T.; Imelfort, M.; Duran, C.; Edwards, D. Bioinformatics tools and databases for analysis of next generation sequence data. *Brief. Funct. Genomics* 2012, *11*, 12–24.

[20] Liu S, Liu H, Wu A, Hou Y, An Y, Wei C (2017); Construction of fingerprinting for tea plant (*Camellia sinensis*) accessions using new genomic SSR markers; Mol Breeding (2017) 37:93

[21] 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 QTLs for catechins content in tea plant (*Camellia sinensis*). PLoS One 9: e93131.

[22] Marshall, D.; Hayward, A.; Eales, D.; Imelfort, M.; Stiller, J.; Berkman, P.; Clark, T.; McKenzie, M.; Lai, K.; Duran, C.; *et al.* Targeted identification of genomic regions using TAGdb. *Plant Methods* 2010, *6*, 19; doi: 10.1186/1746-4811-6-19.

[23] Poehlman, J M, Sleper, D.A. Len, Breeding field crops, Fourth edition, 1995.

[24] Rahman H, Ramanathanm V, Rasi JS, Ramasamy S, Phenomics: Technologies and Applications in Plant and Agriculture, 2015, Plant Omics: The Omics of Plant Science, Edition: first, Chapter, Publisher: Springer, pp. 385-411, DOI: 10.1007/978-81-322-2172-2.

[25] Raina SN, Ahuja PS, Sharma RK, Das SC, Bhardwaj P, Negi R, Sharma V, Singh SS, Sud RK, Kalia RK, Panday V, Banik J, Razdan V, Sehgal D, Dar TH, Kumar A, Bali S, Bhat V, Sharma S, Prasanna BM, Goel S, Negi MS, Vijayan P, Tripathi SB, Bera B, Hazarika M, Mandal AKA, Kumar RR, Vijayan D, Ramkumar S, Chowdhary BR, Mandi SS (2011) Genetic structure and diversity of India hybrid tea. Genet Resour Crop Evol 59 (7):1527-1541.

[26] Rao, V. S., Das, S.K.., Rao, V.J and Srinubabu, G. (2008). Recent developments in life sciences research: Role of Bioinformatics. African Journal of Biotechnology7 (5): 495-503.

[27] Roy SC (2006) Karyotype analysis in three cultivated varieties of tea (*Camellia sinensis (*L) O. Kuntze) for their characterization. J. Phytol Res 19(2): 203-207.

[28] Schuster I (2011), Marker- assisted selection for quantitative traits; Crop Breeding and Applied Biotechnology S1: 50-55, 2011.

[29] Shrivastava R R, Pateriya P, Singh M (2018); Green tea- A short review; International Journal of Indigenous Herbs and Drugs; ISSN-2456-7345; 3 (2).

[30] Sui N, Yao M, Chen L, Zhao L, Wang X (2008) Germplasm and breeding research of tea plant based on DNA marker approaches. Front Agric China 2(2): 200-207.

[31] Swindells, M., Rae, M., Pearce, M., Moodie, S., Miller, R. and Leach, P. (2002) Application of high throughput computing in bioinformatics. Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 360, 1179-1189. Doi: 10.1098/ rsta.2002.0987.

[32] Tan L Q, Wang L Y, Wei K, Zhang C C, Wu L Y, Qi G N, Cheng H, Zhang Q, Cui Q M, Liang J B (2013); Floral Transcriptome Sequencing for SSR Marker Development and Linkage Map Construction in the Tea Plant (*Camellia sinensis*); PLOS ONE, November 2013, Vol 8, issue 11

[33] 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 Genet Genomes 12:52.*

[34] Tanaka J., Taniguchi F. (2006) Estimation of the genome size of tea (*Camellia sinensis*), camellia (C. japonica), and their interspecific hybrids by flow cytometry. *Journal of the Remote Sensing Society of Japan* 101:1-7.

[35] Taniguchi F, Furukawa K, Metoku SO, Yamaguchi N, Ujihara T, Kono I, Fukuoka H, Tanaka J (2012) Construction of a highdensity reference linkage map of Tea (*Camellia sinensis*) Breed Sci 62 : 263-273.

[36] Tester M, Langridge P, Breeding technologies to increase crop production in a changing world, Science, 2010, Feb 12, 327(5967):818-22, doi: 10.1126/science.1183700.

[37] Wales, J. (2009). Bioinformatics. http:// en.wikipedia.org/wiki/ Bioinformatics#Web_services_in_bioinformatics.

[38] Wambulwa MC, Meegahakumbura MK, Chalo R, Kamunya S, Muchugi A, Xu JC, Liu DZ, Gao LM (2016). Nuclear microsatellites reveal the genetic architecture and breeding history of tea germplasm of East Africa. Tree Genet Genomes 12:11

[39] Xia EH, Zhang HB, Sheng J, Li1 K, Zhang QJ, Kim C, Zhang Y, Liu Y, Zhu T, Li W, Huang H, Tong Y,Nan H, Shi C, Shi C,Jiang JJ,Mao SY, Jiao JY, Zhang D,Zhao Y, Zhao YJ, Zhang LP, Liu YL,Liu BY,Yu Y, Shao SF, Ni DJ, Eichler EE and Gao LZ; The Tea Tree Genome Provides Insights into Tea Flavor and Independent Evolution of Caffeine Biosynthesis; Molecular Plant 10, 866–877, June 2017.

[40] Xiong, J. (2009). Essential bioinformatics: introduct ion and biological databases. Cambridge University press, USA. http:// www.cambidge.org /catalogue.

[41] Yang W., Duan L., Chen G., Xiong L., Liu Q., Plant phenomics and high-throughput phenotyping: accelerating rice functional

genomics using multidisciplinary technologies, *Curr. Opin. Plant Biol.*, 2013, 16 180–187. 10.1016/j.pbi.2013.03.005.

[42] Yang H, Wei C L, Liu H W, Wu J L, Li Z G, Zhang L, Jian J B, Li Y Y, Tai Y L, Zhang J, Zhang Z Z, Jiang C J, Xia T, Wan X C (2016), Genetic divergence between *Camellia sinensis* and its wild relatives revealed via genome wide SNPs from RAD sequencing. PLoS One 11:e0151424

[43] Yao MZ, Ma CL, Qiao TT, Jin JQ, Chen L (2012) Diversity distribution and population structure of tea germplasms in China revealed by EST-SSR markers. Tree Genet Genomes 8 : 205-220.

AUTHORS PROFILE’

MRITYUNJAY CHOUBEY

Mr. Mrityunjay Choubey, born in February 1978, received his Master Degree from Dr. Rajendra Prasad Central Agricultural University (RAU), PUSA, Bihar. He is professionally trained in various agricultural aspects at various world reputed institutes like attended training course for agri-clinic & Agri Business, MANA--GE, Hyderabad; training Programme done on “Resource conservation technologies, ICAR, Patna; attended training Programme on “ Resource conservation technologies, BAMETI, Patna; training Programme on ‘Entrepreneurship development in Agriculture’ BAMETI, Patna; attended training Programme on “Organic Farming”, Indian Institute of Soil Science, Bhopal (MP). He has to his credit more than twenty five scientific research papers in national and international repute journal including abstracts besides four popular articles, technical bulletin on various aspects of plant science and one book chapter on Tea Research for Darjeeling Tea Industry - Various Aspects. He has also participated in several national and international conferences. Presently Mr. Choubey is working as Senior Scientific Officer (Agro-Botany Division) in the Darjeeling Tea Research and Development Centre, Tea Board of India.

Urvashi Lama

Address: North Khairbari, Madarihat, Alipurduar, WB – 735220. Date of Birth: 06/02/1994
email id: urviyonzone09@gmail.com.
Contact no. : +919749760541

Educational Qualification:-

- B.Sc. in Botany, Presidency University (2011 - 2014).
- M.Sc. in Botany, Presidency University (2014 – 2016).
- Cell Biology and Molecular Genetics (Special Paper).
- Micro-propagation and some phytochemical aspects of *Gymnema sylvestre* R. Br. (Dissertation thesis).

Work Experience:-

- Volunteer – teacher, Make A Difference, Kolkata (2012 – 2013).
- Principal and Teacher, R.L. Memorial English School, Madarihat (2016 – 2017).
- Junior Research Fellow, Darjeeling Tea Research and Development Centre, Tea Board, India (2017 – present).