

Article

# Mining and Characterizing the SSR Markers for Black Rice Using the Illumina Sequencing Platform

Dhananjay Kumar<sup>1,3†</sup>, Prasanta Kumar Das<sup>1,2†</sup>, Chandra Singha<sup>1</sup> and Bidyut Kumar Sarmah<sup>1,2\*</sup>

<sup>1</sup> Department of Agricultural Biotechnology, Assam Agricultural University Jorhat-785013, India; dhananjayphdresearch@gmail.com (D.K.); prasantabio@gmail.com (P.-K.D.); chandrasingha95@gmail.com (C.S.)

<sup>2</sup> DBT-AAU Centre, Assam Agricultural University Jorhat-785013, India

<sup>3</sup> Present Address: Swedish University of Agricultural Sciences, Box 101, SE-230 53 Alnarp, Sweden

\* Corresponding author: bidyutsarmah@aaau.ac.in; Tel.: +91-9435091484

† These authors have contributed equally to this work

**Abstract:** Study in black rice has gain prominence in recent times due to its high nutritive value, curative effect, and anti-oxidant properties. However, its poor agronomic traits, including low yield necessitates the incorporation of the colour-grain trait into elite varieties through plant breeding techniques. SSR markers play an important role in plant identification and breeding. Here, the generation of reference-based transcriptome, annotation of transcriptome datasets, and a large set of simple sequence repeat (SSR) markers derived from Black rice have been described. In all 28664 SSRs were predicted in 34978 (48.59%) expressed transcripts. However, 7068 (20.20%) transcripts were found to have more than one SSR. The identified SSRs were dominated by tri-nucleotide and tetra-nucleotide repeats representing about 54.11% and 33.31% respectively, of total SSRs. Validation of selected markers associated with anthocyanin trait performed across different black rice accessions established the reliability of the process used for mining SSR markers. The SSR markers identified in this study could be used to select varieties with desired traits, and to investigate the genetic mechanism underlying anthocyanin accumulation in the pericarps of black rice. Furthermore, the findings from this study may prove beneficial in future genetic diversity studies, primer development, and selective breeding programs.

**Keywords:** black rice; transcriptome sequencing; genic SSRs; genetic diversity

## 1. Introduction

Black rice constitutes a group of speciality rice varieties which are glutinous, packed with a high level of nutrients, anthocyanin, and is cultivated in many countries across Asia. The black colouration of the pericarp is attributed to a class of flavonoid pigment known as anthocyanin [1]. Anthocyanins are the primary pigments in grains, flowers, fruits, and vegetables showing blue, purple and red pigmentations. It is estimated that there are more than 400 naturally occurring anthocyanins in nature [2,3] due to different glycosylation and acylation in which cyanidin-3-O-glucoside is the most abundant, followed by peonidin-3-O-glucosid. Black rice is also considered by many as a panacea of culinary diseases because of its high nutritive value, curative effect coupled with beneficial properties of flavonoids that not only act as an antioxidant [4,5] and anti-inflammatory [6,7] but has also been linked to anti-carcinogenic properties [8]. These black rice may also be helpful for the prevention or cure of diseases caused by vitamin A and B deficiencies [9,10]. There are more than 200 types of black rice varieties in the world [11] with highest production reported in China followed by Sri Lanka, Indonesia, India, and the Philippines.

In spite of the demand for black rice, its availability is limited due to factors like low yield, lodging susceptibility and long duration of this crop that makes farmers less interested in the cultivation of black rice. Thus, its improvement through the incorporation of coloured-pericarp trait

into elite varieties, use of genotypic markers would be more beneficial than conventional approaches [12]. The SSR markers are an ideal choice for the large scope of the markers availability and have been used for genetic variability assessment, molecular characterization, genotypic identification, and population structure estimations in rice [13,14]. Besides diversity studies [15], the SSR markers have also been used extensively for the construction of linkage maps [16-18] and QTL analysis [19] in rice. Genic SSRs are located in the coding region of the genome relatively easy and inexpensive, and highly transferable to related taxa [20,21]. Although SSRs spanning the linkage groups in different rice varieties, as well as transcriptome studies have been reported in black rice, there are no reports till date in our knowledge, on the development of genic SSR markers from expressed gene sequences which could be used for trait-specific selection of population during a breeding programme.

In this present study, we have generated a large number of gene-based SSR markers from the transcriptome sequences of black rice variety Chakhao poreitin using the Illumina HiSeq sequencing platform. The distribution of the SSR motifs in the sequences was functionally characterized and annotated according to gene ontology terms, the cluster of orthologous genes (KOGs) and KEGG pathways. Subsequently, a large set of genic SSR markers were developed for Black rice from transcriptome sequences. This molecular markers might prove beneficial in genetic diversity studies, primer development, and selective breeding programs.

## 2. Materials and Methods

### 2.1. Plant Materials

Black rice variety (Chakhao poreitin), and white rice variety (IR-64) were sampled from Regional Agricultural Research Station (RARS), Shillongoni, Assam Agricultural University, Jorhat India. The plants were maintained in the polyhouse under irrigated conditions. The immature seeds were collected 6 days after heading (6 DAH). The immature seeds with intact pericarp were dehusked and stored in RNA-later (Catalog no R0901-500ML) until further processing. For the validation of the candidate SSR markers, four additional black rice landraces were collected from different regions of Assam, India.

### 2.2. Illumina Sequencing, Quality Control, and Assembly

The reference-based transcriptome was performed on a total of four rice samples with two biological replicates each from both black rice as well as white rice [36]. The raw data generated were checked for the quality using FastQC and preprocessed, which include removing the adapter sequences and the low-quality bases (<q30). Pre-processing of the data was done with Cutadapt. HISAT is a fast and sensitive spliced alignment program, was used to align the high-quality data to the reference genome with the default parameters. Reads were classified into aligned reads (which align to the reference genome) and unaligned reads. Cufflinks were used to estimate and calculate transcript abundance. It resulted in a normalized read count in the form of FPKM values. FPKM is a unit of measuring gene/transcript expression. Cuffdiff was used to calculate the differentially expressed transcripts and categorize them into up, down and neutrally regulated based on the log<sub>2</sub>fold change values. Enrichment analysis for differentially expressed genes was performed using PlantRegMap tool.

### 2.3. SSR Prediction search, annotation, and primer design

MISA (MicroSATellite) (<http://pgrc.ipk-gatersleben.de/misa>) and SAMtools were employed for SSR mining and identification of SSRs and the sequence that contained SSR. These sequences had SSRs were aligned with BLASTX against protein databases, including NCBI nr database. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Eukaryotic Ortholog Groups (KOG) with an E-threshold value of 1.0E-5. Blast2GO software was used to obtain GO annotation. Primers for the SSRs were designed using Batchprimer3 tool (<https://probes.pw.usda.gov/batchprimer3/>). BatchPrimer3 is a comprehensive web-based primer design program in which we had primer lengths of 16–22 bases,

GC content of 40%–60%, annealing temperature of 40 °C–60 °C, and PCR product size of 100 to 600 bp.

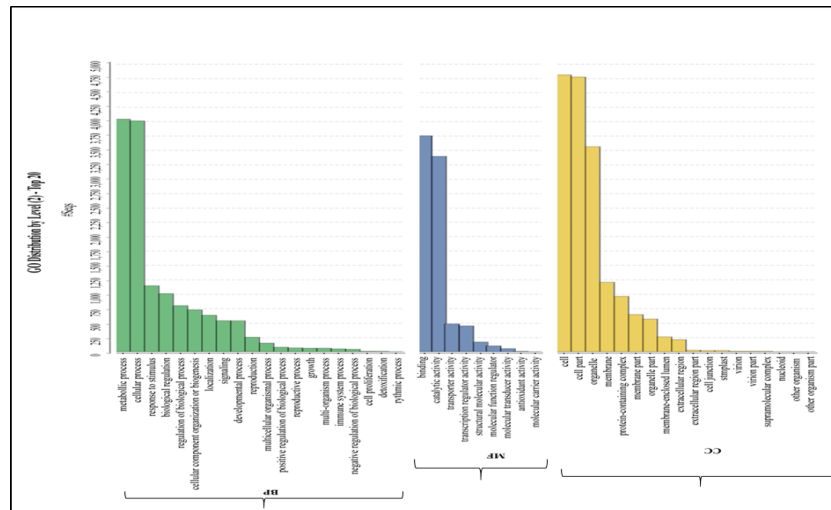
#### 2.4. Validation of SSR Markers

SSR markers, except mononucleotide repeats, were selected randomly for the validation in the laboratory. The primers were designed using the primer blast tool, and DNA was extracted by the CTAB method. The quality of extracted DNA was assessed through agarose gel electrophoresis on 0.7% of agarose gel. The isolated DNA was quantified by measuring the absorbance at 260 nm in a UV-visible spectrophotometer (Thermo scientific). Polymerase chain reaction (PCR) was carried out in a reaction volume of 10 µL with 5 µL 2<sup>+</sup>Taq PCR Master Mix, 0.2 µM of each primer, 1 µL template DNA and 3.6 µL ddH<sub>2</sub>O. All amplifications were carried out in SimpliAmp™ Thermal Cycler (Applied Biosystems, Carlsbad, CA, USA) as follow: denaturation at 94 °C for 5 min, followed by 30 cycles of 94 °C for 50s, at specific annealing temperature (T<sub>m</sub>) for 30s, 72 °C for 40s and 72 °C for 5min as final extension. PCR amplified products were electrophoresed on 3.5% Agarose. A 1000 bp DNA ladder was used as a molecular marker to determine the approximate size of the fragments.

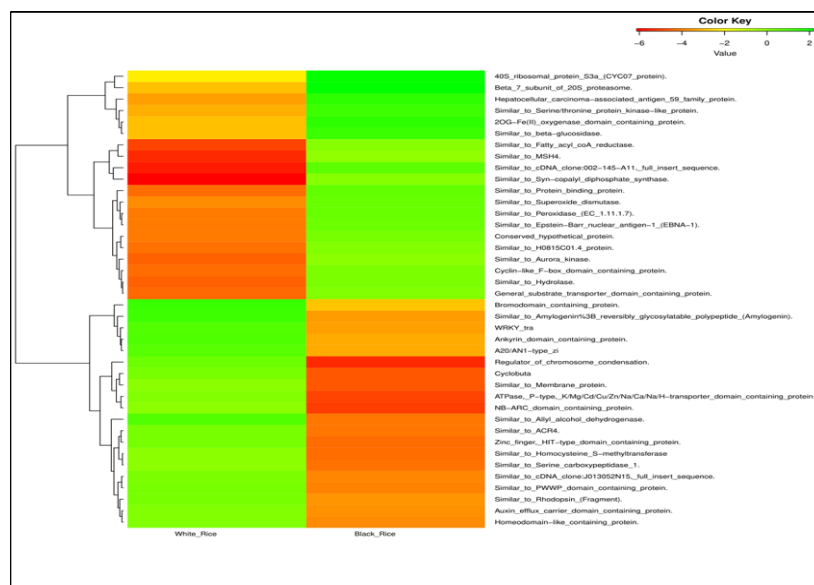
### 3. Results

#### 3.1. Sequence Annotation of black rice containing SSR

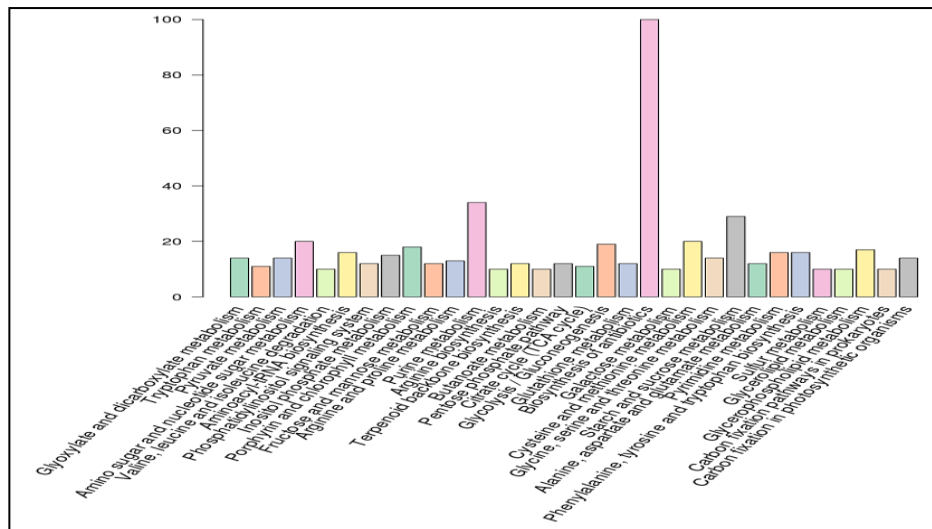
A total of 16,996 SSR containing sequences were obtained, of which 10,034 were found to be having significant BLAST hits. These SSR sequences were annotated by KOG classification, KEGG (The Kyoto Encyclopedia of Genes and Genomes) classification and GO annotation that further categorized them by molecular function, cellular component, and biological process. Using KOG functional classification, 5,384 genes were aligned to the KOG database, which revealed that other than general function prediction, which showed the highest number of repeats, 10.04 % of the repeats are involved in signal transduction mechanisms, followed by posttranslational modification, protein turnover, chaperones (8.43 %), intracellular trafficking, secretion and vesicular transport (6.09 %). A total of 8,473 (84.0 %) SSR sequences were assigned to gene ontology (GO) terms on the basis of the annotation against the NR (non-redundant) database. Within the cellular component category that represented the major class, 4,800 SSR sequences were assigned to cell component (24.3%) while 3,620 were assigned to organelle component (18.9%). Only a few repeats were assigned to the other categories with the least distribution in virion (0.013%) and symplast (0.012%) region. Within the biological process category, the highest sub-category with 20,281 SSRs were assigned to metabolic process (21.0%), followed by a cellular process with 18,363 SSR (19.0%). Within molecular function category, 15,837 and 15,475 SSR were predominantly associated with the binding activity (42.0%) and catalytic activity (41.0%) respectively, followed by 499 SSR for transporter activity (4.73%), while only a few sequences were distributed in the other molecular functional categories (Figure 1). Furthermore, a general overview of the expression pattern was visualized in a heat map (Figure 2), which provided an overall understanding of the change in gene expression. The KEGG was used to analyze the pathway annotations of SSR sequences and to characterize the functional genes in black rice based on active biological pathways. In all, 901 SSR sequences were annotated to 130 pathways based on the KEGG pathway database. As per KEGG classification, it was observed that the highest number of repeats are involved with biosynthesis of antibiotics (11.09 %) followed by purine metabolism (3.77%), while the other repeats are more or less equally distributed in the various KEGG pathways (Figure 3). A total of 2,168 (7.83 %) sequences were also given enzyme codes according to KEGG, and it was observed that the majority of the sequences coded for hydrolases (39.25%) followed by transferases (37.03%) whereas the least distribution was shown in lygases with a distribution of 2.35 % (Figure 4).



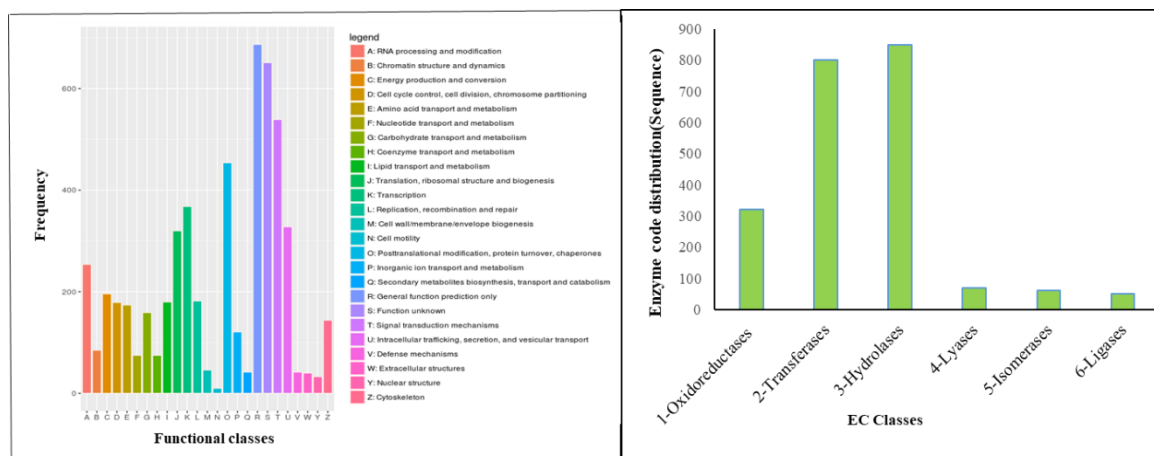
**Figure 1.** Histogram presentation of GO classification of genes from Black rice transcriptome sequences. Histogram representing gene ontology (GO) classification of the assembled unigenes. The results are summarized in three main GO categories: cellular component, molecular functions, and biological process. The x-axis depicts the sub-categories of the unigenes, and y-axes represent the percentage and the number of unigenes, respectively



**Figure 2.** Heat map showing top 20 up and down-regulated transcripts in white and black rice. The heat map generation log<sub>2</sub>fold change was sorted in descending order, ranging from according to the colored key. Expression values of both control and treated samples were considered along with the corresponding protein names. Clustvis (<http://biit.cs.ut.ee/clustvis/>) tool was used to prepare the heat map in this study.



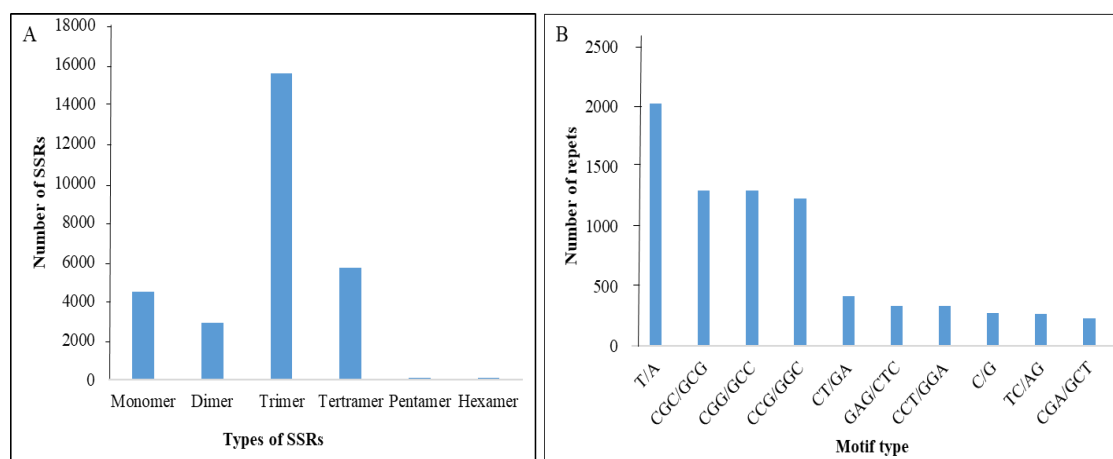
**Figure 3.** Histogram representations for the top twenty pathways of unigenes assigned based on the Kyoto Encyclopedia of Genes and Genomes (KEGG). The columns are the number of unigenes in each subcategory and Y-axis represents no of unigenes in each category.



**Figure 4.** COG function classification and EC class functions A. COG functional classification of consensus sequences (x-axis represented the 26 different groups, y-axis referred to the frequency of genes). B. EC class function classification of consensus sequences (x-axis represented the EC class groups, y-axis referred enzyme code distribution to the of genes)

### 3.2. Frequency and distribution of SSRs in the black rice genome

A total of 28,664 SSRs were found out of the 34,978 transcripts expressed in immature black rice pericarp. In all, 16,996 sequences contained SSR of which 10,034 sequences were found to have significant blast results, 7,068 sequences had more than one SSR and 4044 were compound SSRs. (Table 1). The proportion of SSRs was not evenly distributed; the trinucleotide repeat motifs were the most abundant (15,512 or 62.89 %), followed by tetra- (5,662 or 22.95 %), mono- (4446 or 18.02 %), di- (2879 or 11.67 %), penta- (91 or 0.36 %) and hexanucleotide (755 or 0.30 %) repeat motifs. Among the nucleotide repeats, A/T with 2,026 occurrences was the most abundant. The other three major motifs were CGC/GCG with 1302 occurrences, CGG/GCC with 1299 occurrences and CCG/GGC with 1231 occurrences (Figure 5a and 5b).

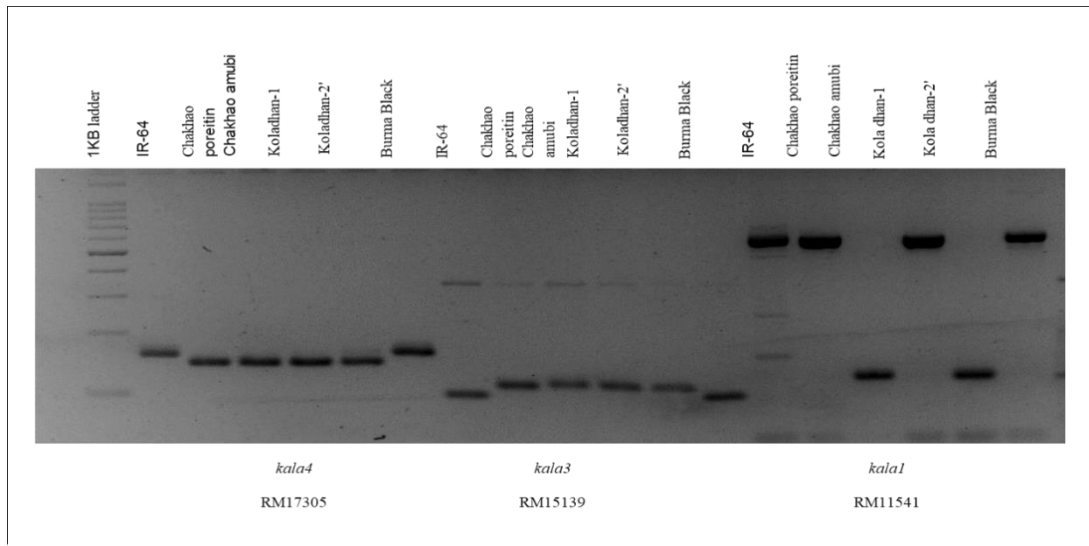


**Figure 5.** Frequency distribution of the SSRs identified in Black rice transcriptome. (A) Frequency distribution of the Black rice SSRs; (B) Total numbers of different SSR motifs in Black rice are represented.

### 3.3. Development of polymorphic genic SSR markers and Validation of trait-linked functional SSR markers

In this study, a large number of high-quality unigene sequences have been developed from black rice via next-generation sequencing from immature black rice seeds and of genic SSRs have been identified. A whole-genome survey of the introgressed black rice line using DNA markers suggested three regions on chromosomes 1, 3 and 4 were associated with black pigmentation namely *Kala1*, *Kala3*, and *Kala4* [22]. In all, 61 SSR primer pairs from the unigene sequences corresponding to the genic region on linkage groups 1, 3 and 4 that bears the loci *Kala1*, *Kala3*, and *Kala4* were identified. Of the 61 tested markers, 18 (29.5 %) amplified non-specific products, 34 (55.7%) failed to amplify the DNA product, and only 6 (9.8 %) showed polymorphism. Although, all 6 markers showed polymorphism, of which 3 could exhibit distinct polymorphism for the *Kala1*, *Kala3* and *Kala4* locus across all the black rice accession. Our results demonstrate that in-silico SSR marker development by transcriptome sequencing using NGS can be used to detect (or allow detecting) genic SSRs for rice and could supplement the high genomic resources available depending on the trait of interest (Figure 5).

In order to validate genic SSR, we chose DNA markers associated with black pigmentation to correspond to the three loci, namely *Kala1*, *Kala3* and *Kala4* [22]. Furthermore, *Kala1* has been found to encode the DFR enzyme and mapped to chromosome 1 while *Kala3* and *Kala4* correspond to bHLH domain-containing and MYB domain-containing TF, respectively. As these three loci reportedly acted as a major determinant of pigment formation in black rice pericarp [22], they were chosen as the candidate for validation as well as determining the genic SSR markers that would show polymorphism in a genotype independent manner. The genic SSRs were tested for validation in 5 black rice landraces, namely, Chakhao poreitin, Chakhao amubi, kola-Dhan1, Kola-Dhan-2, and Burma Black. We were able to successfully validate the polymorphism for the 6 SSR markers, i.e. RM 11343, RM11441 and RM11541 present in the gene *Kala1* locus, RM15139 present in *Kala3* locus and RM17305 and RM5714 present in *Kala4* locus (Figure 6).



**Figure 6.** Validation of polymorphic genic SSR. The three major genes responsible for anthocyanin accumulation in the black rice such as *Kala1*, *Kala3*, and *Kala4* were validated in white rice (IR-64) and five black rice land races. RM17305, RM 15139 and RM11541 are polymorphic for 2,3,4, and 5 while not showing any polymorphism for 6.

#### 4. Discussion

SSR regions are more abundant in non-coding regions of the genomes, however expressed regions may also harbour such repeat motifs [23-25]. Genic-SSR markers are considered to have strong potential for genetic analysis and linkage map construction in crop species due to their specificity and a high degree of conservation [26,27]. Polymorphism in the genic regions, though less, has a greater likelihood of identifying functional variability among genotypes. Black or purple rice is known for its high anthocyanin content. Genetic improvement of this particular trait through marker-assisted breeding needs identification of SSR markers that would exhibit greater redundancy for polymorphism across different varieties of black rice. Our study showed that transcriptome sequencing is a useful tool for unigene discovery and marker development in Black rice. We have developed a total of 10034 genic SSR markers from the 16996 SSR-containing unigene sequences in this study.

For the perfect repeat motifs types, tri-nucleotide repeats are observed to have the highest frequency in many crops, including cotton, barley, wheat, maize, sorghum, rice and peanut [28]. However, we observed that the proportion of SSRs was not evenly distributed; the trinucleotide repeat motifs were the most abundant (15,512 or 62.89 %), followed by tetra- (5,662 or 22.95 %), mono- (4446 or 18.02 %), di- (2879 or 11.67 %), penta- (91 or 0.36 %) and hexanucleotide (755 or 0.30 %) repeat motifs. It has been emphasized that the frequency of SSRs is correlated with many factors, such as SSR detection criteria, dataset size, database-mining tools, different species and different materials [26,29,30]. The genic SSR markers with di-, tri-, tetra-, penta-, and hexanucleotide repeats were 8.7%, 60.9%, 26.1%, and 4.3%, respectively. In other crops like wheat also, not more than 6.25% of primers exhibit polymorphisms between the parents of any individual mapping population, although 81.25% of detected EST-SSRs have been reported to exhibit polymorphisms in 18 alien species [31]. In another report in peanut, 10.3% EST-SSRs exhibited polymorphisms between 22 cultivated peanut accessions and 88% were polymorphic between 16 wild peanut species [32]. The low specificity in amplification rate and low polymorphism rate in our study could be due to several factors like large intronic region between primers, unrecognized intron splice sites that can disrupt priming sites, and or those that could be simply for sequencing error. Among the nucleotide repeats, A/T with 2,026 occurrences was the most abundant. The other three major motifs were CGC/GCG with 1302 occurrences, CGG/GCC with 1299 occurrences and CCG/GGC with 1231 occurrences. These data suggest that the AT motifs are abundant marker resources in the transcript sequences of Black rice. It was reported that among

hundreds of types of repeat motifs, the (AG/CT)<sub>n</sub> di-nucleotide motifs showed the highest frequency in other species, The lowest frequent motif was CG/CG which was also rare in studies using the plants such as coffee, wheat and rosaceae [28,29,33], but differs from some cereal species, in which (TA/TA)<sub>n</sub> motif was the most abundant of the repeat motifs in black rice.

To determine the level of polymorphism among our set of new genic-SSR markers, we validated primer pairs. However, the universality of these markers for use as foreground markers in other black rice in the breeding program for the trait of interest, which in this case purple colouration in pericarp was found to be lacking. Similar reports indicated in sesame that the low polymorphism of SSR markers is likely due to its narrow genetic basis [34,35]. Our results indicate that large numbers of polymorphic SSR markers from transcriptome data sets are available to use in marker-assisted gene selection and QTL analysis gene mapping, comparative genome analysis.

## 5. Conclusions

We performed a de novo transcriptome sequencing analysis of black rice seed (Chakhao) tissues before maturation using an Illumina HiSeq technology. This paper reports on the transcriptome characterizations of black rice and provides a large number of SSR markers to elucidate the functional annotation of black rice. To our knowledge, this is the first report to develop SSR markers for black rice using a transcriptome sequencing method, and these SSR markers will significantly be useful to study the genetic diversity, QTL mapping, and marker-assisted breeding for black rice. These developed SSR markers from transcriptome sequencing may be an efficient tool to accelerate molecular breeding in black rice.

**Author Contributions:** Conceptualization, D.K., P.K.D., and B.K.S.; methodology, D.K.; software, B.K.S.; validation, D.K., P.K.D.; formal analysis, D.K. and C.S.; investigation, D.K., P.K.D.; resources, B.K.S.; data curation, D.K.; writing—original draft preparation, D.K. and C.S.; writing—review and editing, P.K.D., and B.K.S.; visualization D.K., P.K.D., and B.K.S.; supervision, P.K.D., and B.K.S.; project administration, B.K.S.; funding acquisition, B.K.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Department of biotechnology government of India. Project: DBT-ER/AGRI/29/2015 (Group-4) dtd. Oct. 19, 2016.

**Acknowledgments:** We acknowledge Dr. P K Debchoudhury, Plant Breeder at the Regional Agricultural Research Stations, Shillongani under Assam Agricultural University, Jorhat, for providing the seed material and Jumi Borah for validation of SSR marker.

**Conflicts of Interest:** “The authors declare no competing financial interests”.

## References

1. Kushwaha, U. Black rice anthocyanin content increases with increase in altitude of its plantation. *Advances in Plants and Agriculture Research* 2016, 5, 00170.
2. Kong, J.-M.; Chia, L.-S.; Goh, N.-K.; Chia, T.-F.; Brouillard, R. Analysis and biological activities of anthocyanins. *Phytochemistry* 2003, 64, 923-933.
3. Prior, R.L.; Wu, X. Anthocyanins: structural characteristics that result in unique metabolic patterns and biological activities. *Free radical research* 2006, 40, 1014-1028.
4. Reddy, A.M.; Reddy, V.S.; Scheffler, B.E.; Wienand, U.; Reddy, A.R. Novel transgenic rice overexpressing anthocyanidin synthase accumulates a mixture of flavonoids leading to an increased antioxidant potential. *Metabolic engineering* 2007, 9, 95-111.
5. Sun, L.; Zhang, J.; Lu, X.; Zhang, L.; Zhang, Y. Evaluation to the antioxidant activity of total flavonoids extract from persimmon (*Diospyros kaki* L.) leaves. *Food and chemical toxicology* 2011, 49, 2689-2696.
6. Sakamoto, W.; Ohmori, T.; Kageyama, K.; Miyazaki, C.; Saito, A.; Murata, M.; Noda, K.; Maekawa, M. The Purple leaf (Pl) locus of rice: the Pl w allele has a complex organization and includes two genes encoding basic helix-loop-helix proteins involved in anthocyanin biosynthesis. *Plant and cell physiology* 2001, 42, 982-991.
7. Kim, B.G.; Kim, J.H.; Min, S.Y.; Shin, K.-H.; Kim, J.H.; Kim, H.Y.; Ryu, S.N. Anthocyanin content in rice is related to expression levels of anthocyanin biosynthetic genes. *Journal of Plant Biology* 2007, 50, 156-160.



8. Bowen-Forbes, C.S.; Zhang, Y.; Nair, M.G. Anthocyanin content, antioxidant, anti-inflammatory and anticancer properties of blackberry and raspberry fruits. *Journal of food composition and analysis* 2010, 23, 554-560.
9. Dipti, S.S.; Bergman, C.; Indrasari, S.D.; Herath, T.; Hall, R.; Lee, H.; Habibi, F.; Bassinello, P.Z.; Graterol, E.; Ferraz, J.P. The potential of rice to offer solutions for malnutrition and chronic diseases. *Rice* 2012, 5, 16.
10. Panhwar, Q.; Naher, U.; Radziah, O.; Shamshuddin, J.; Razi, I.M.; Dipti, S.; Aghamolki, M.K. Quality and antioxidant activity of rice grown on alluvial soil amended with Zn, Cu and Mo. *South African Journal of Botany* 2015, 98, 77-83.
11. Kushwaha, U. Black rice. In *Black Rice*, Springer: 2016; pp. 21-47.
12. Oikawa, T.; Maeda, H.; Oguchi, T.; Yamaguchi, T.; Tanabe, N.; Ebana, K.; Yano, M.; Ebitani, T.; Izawa, T. The birth of a black rice gene and its local spread by introgression. *The Plant Cell* 2015, 27, 2401-2414.
13. D'Hont, A.; Denoeud, F.; Aury, J.-M.; Baurens, F.-C.; Carreel, F.; Garsmeur, O.; Noel, B.; Bocs, S.; Droc, G.; Rouard, M. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 2012, 488, 213.
14. Olufowote, J.O.; Xu, Y.; Chen, X.; Goto, M.; McCouch, S.R.; Park, W.D.; Beachell, H.M.; Dilday, R.H. Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLP markers. *Genome* 1997, 40, 370-378.
15. Salgotra, R.; Gupta, B.; Bhat, J.A.; Sharma, S. Genetic diversity and population structure of Basmati rice (*Oryza sativa* L.) germplasm collected from North Western Himalayas using trait linked SSR markers. *Plos one* 2015, 10, e0131858.
16. Wei, L.-B.; Zhang, H.-Y.; Zheng, Y.-Z.; Miao, H.-M.; Zhang, T.-Z.; Guo, W.-Z. A genetic linkage map construction for sesame (*Sesamum indicum* L.). *Genes & genomics* 2009, 31, 199-208.
17. Zhao, X.; Huang, L.; Zhang, X.; Wang, J.; Yan, D.; Li, J.; Tang, L.; Li, X.; Shi, T. Construction of high-density genetic linkage map and identification of flowering-time QTLs in orchardgrass using SSRs and SLAF-seq. *Scientific reports* 2016, 6, 29345.
18. Shabir, G.; Aslam, K.; Khan, A.R.; Shahid, M.; Manzoor, H.; Noreen, S.; Khan, M.A.; Baber, M.; Sabar, M.; Shah, S.M. Rice molecular markers and genetic mapping: Current status and prospects. *Journal of integrative agriculture* 2017, 16, 1879-1891.
19. Zhao, X.; Daygon, V.D.; McNally, K.L.; Hamilton, R.S.; Xie, F.; Reinke, R.F.; Fitzgerald, M.A. Identification of stable QTLs causing chalk in rice grains in nine environments. *Theoretical and applied genetics* 2016, 129, 141-153.
20. Varshney, R.; Bertoli, D.; Moretzsohn, M.d.C.; Vadez, V.; Krishnamurthy, L.; Aruna, R.; Nigam, S.; Moss, B.; Seetha, K.; Ravi, K. The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics* 2009, 118, 729-739.
21. Varshney, R.K.; Marcel, T.C.; Ramsay, L.; Russell, J.; Röder, M.S.; Stein, N.; Waugh, R.; Langridge, P.; Niks, R.E.; Graner, A. A high density barley microsatellite consensus map with 775 SSR loci. *Theoretical and Applied Genetics* 2007, 114, 1091-1103.
22. Maeda, H.; Yamaguchi, T.; Omoteno, M.; Takarada, T.; Fujita, K.; Murata, K.; Iyama, Y.; Kojima, Y.; Morikawa, M.; Ozaki, H. Genetic dissection of black grain rice by the development of a near isogenic line. *Breeding science* 2014, 64, 134-141.
23. Cho, Y.G.; Ishii, T.; Temnykh, S.; Chen, X.; Lipovich, L.; McCOUCH, S.R.; Park, W.D.; Ayres, N.; Cartinhour, S. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 2000, 100, 713-722.
24. Thumilan, B.M.; Sajeevan, R.; Biradar, J.; Madhuri, T.; Nataraja, K.N.; Sreeman, S.M. Development and characterization of genic SSR markers from Indian mulberry transcriptome and their transferability to related species of Moraceae. *PloS one* 2016, 11, e0162909.
25. McCouch, S.R.; Chen, X.; Panaud, O.; Temnykh, S.; Xu, Y.; Cho, Y.G.; Huang, N.; Ishii, T.; Blair, M. Microsatellite marker development, mapping and applications in rice genetics and breeding. In *Oryza: From Molecule to Plant*, Springer: 1997; pp. 89-99.
26. Zhang, H.; Wei, L.; Miao, H.; Zhang, T.; Wang, C. Development and validation of genic-SSR markers in sesame by RNA-seq. *BMC genomics* 2012, 13, 316.
27. Zhai, L.; Xu, L.; Wang, Y.; Cheng, H.; Chen, Y.; Gong, Y.; Liu, L. Novel and useful genic-SSR markers from de novo transcriptome sequencing of radish (*Raphanus sativus* L.). *Molecular Breeding* 2014, 33, 611-624.

28. Nicot, N.; Chiquet, V.; Gandon, B.; Amilhat, L.; Legeai, F.; Leroy, P.; Bernard, M.; Sourdille, P. Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). *Theoretical and Applied Genetics* 2004, 109, 800-805.
29. Aggarwal, R.K.; Hendre, P.S.; Varshney, R.K.; Bhat, P.R.; Krishnakumar, V.; Singh, L. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theoretical and Applied Genetics* 2007, 114, 359.
30. Wang, P.; Yang, L.; Zhang, E.; Qin, Z.; Wang, H.; Liao, Y.; Wang, X.; Gao, L. Characterization and development of EST-SSR markers from a cold-stressed transcriptome of Centipedegrass by Illumina paired-end sequencing. *Plant molecular biology reporter* 2017, 35, 215-223.
31. Gupta, P.K.; Rustgi, S.; Sharma, S.; Singh, R.; Kumar, N.; Balyan, H. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Molecular genetics and genomics* 2003, 270, 315-323.
32. Liang, X.; Chen, X.; Hong, Y.; Liu, H.; Zhou, G.; Li, S.; Guo, B. Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species. *BMC Plant Biology* 2009, 9, 35.
33. Jung, S.; Abbott, A.; Jesudurai, C.; Tomkins, J.; Main, D. Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs. *Functional & integrative genomics* 2005, 5, 136-143.
34. Dixit, A.; JIN, M.H.; CHUNG, J.W.; YU, J.W.; CHUNG, H.K.; MA, K.H.; PARK, Y.J.; CHO, E.G. Development of polymorphic microsatellite markers in sesame (*Sesamum indicum* L.). *Molecular Ecology Notes* 2005, 5, 736-738.
35. Li-Bin, W.; Zhang, H.-Y.; Zheng, Y.-Z.; Wang-Zhen, G.; Zhang, T.-Z. Developing EST-derived microsatellites in sesame (*Sesamum indicum* L.). *Acta agronomica sinica* 2008, 34, 2077-2084.
36. Oh, J.-H.; Lee, Y.-J.; Byeon, E.-J.; Kang, B.-C.; Kyeoung, D.-S.; Kim, C.-K. Whole-genome resequencing and transcriptomic analysis of genes regulating anthocyanin biosynthesis in black rice plants. *3 Biotech* 2018, 8, 115.