



ISSN (E): 2277-7695  
 ISSN (P): 2349-8242  
 NAAS Rating: 5.23  
 TPI 2022; 11(7): 394-396  
 © 2022 TPI  
[www.thepharmajournal.com](http://www.thepharmajournal.com)  
 Received: 16-04-2022  
 Accepted: 22-06-2022

**M Rajavel**  
 Department of Crop Physiology,  
 Tamil Nadu Agricultural  
 University, Coimbatore, Tamil  
 Nadu, India

**R Veera Ranjani**  
 Department of Plant  
 Biotechnology, Tamil Nadu  
 Agricultural University,  
 Coimbatore, Tamil Nadu, India

**M Williams**  
 Department of Plant  
 Biotechnology, Tamil Nadu  
 Agricultural University,  
 Coimbatore, Tamil Nadu, India

**A Bharathi**  
 Department of Plant  
 Biotechnology, Tamil Nadu  
 Agricultural University,  
 Coimbatore, Tamil Nadu, India

## Allelic diversity and haplotype analysis of *YLI*, a chloroplast localized protein

M Rajavel, R Veera Ranjani, M Williams and A Bharathi

### Abstract

Rice (*Oryza sativa* L.) is an important food crop species and most widely grown cereal. A key objective of rice breeding is the development of high-yielding, photosynthetically efficient rice varieties. In recent years, many outstanding genes have been used in rice breeding due to the constant development of molecular breeding technique, which is highly important for increasing rice yields. Increasing rice yield necessitates the discovery of novel genes and alleles from the rice gene pool, as well as genetic manipulation of novel traits. To increase the rice yield, it is required to find superior alleles influencing yield component traits. Using publically available sequence data from the IRRI 3K rice genome sequencing project, this work attempted to examine the allelic diversity of a chloroplast localized protein *YLI* (LOC\_Os02g05890) gene across a diverse set of 186 rice accessions. *YLI* has 135 INDELS and 87 SNPs (Single Nucleotide Polymorphisms) according to allelic diversity analysis. In this study, four haplotypes (H1, H2, H3 and H4) contributed by six non synonymous SNPs were identified for *YLI*.

**Keywords:** Rice, photosynthesis, chloroplast development, haplotype analysis, *YLI*

### 1. Introduction

Increasing demands on the global food supply indicate that a 50-70 percent increase in agricultural output is required by 2050. (Jaggard *et al.*, 2010) [5]. With an ever-increasing population, food production is facing a problem and must expand to satisfy demand. Rice, a basic staple food crop consumed by more than half of the world's population, have remained relatively unchanged over the last decade, implying that yield advances achieved by green revolution engineering have been maximized under our existing production systems (Grassini *et al.*, 2013) [3]. Through enhanced photosynthetic efficiencies, the fixing of C4 photosynthetic system into rice has the theoretical potential to double present rice yields (Lee *et al.*, 2021; Von Caemmerer *et al.*, 2012; Hibberd *et al.*, 2008) [6, 9, 4]. Chloroplast ATP synthase has a crucial role play in the light-dependent reactions of photosynthesis. The role of the *YLI* gene in chloroplast development and biosynthesis of the chloroplast ATP synthase in rice is characterized using a leaf-colour mutant in rice (Chen *et al.*, 2016) [2]. The *YLI* mutant showed a reduced chlorophyll content, abnormally formed chloroplast, and reduced photochemical efficiency throughout developmental stage of the crop and also showed deficit in the accumulation cpATPase. Molecular and genetic analyses revealed that *YLI* is a nucleus-encoded protein with a predicted transmembrane domain in its carboxyl-terminus that is conserved to higher plant kingdom. *YLI* localizes to chloroplasts and is preferentially expressed in green tissues containing chloroplasts. Understanding the allelic variation in the chloroplast localized protein *YLI* can aid in haplotype breeding and the development of high yielding photosynthetically efficient rice. The current study was undertaken to find out the allelic variations and haplotype analysis of *YLI* in 3K RG panel RG subset.

### 2. Materials and Methods

#### 2.1 Allelic diversity and haplotype analysis of *YLI* in 3K RG panel subset

A diverse set of 186 lines from 3K RG panel were used as a subset for the present study. The SNP seek database (<https://snp-seek.irri.org/>) was used to retrieve data on a chloroplast localized protein *YLI* of selected 186 rice accessions (Mansueto *et al.*, 2017) [7]. Allelic variation of *YLI* was accessed from the SNP seek database to perform haplotype analysis. The 3K filtered SNP dataset obtainable in the SNP seek database was utilized for the haplotype analysis. The 3K filtered SNP dataset was obtained using the SNP search database with the following criteria: alternative allele frequency of at least 0.01 and 0.2 missing calls per SNP.

#### Corresponding Author:

**M Rajavel**  
 Department of Crop Physiology,  
 Tamil Nadu Agricultural  
 University, Coimbatore, Tamil  
 Nadu, India

By focusing primarily on non-synonymous SNPs, allele mining was done. gPLINK (version 1.07) (Purcell *et al.*, 2007)<sup>[8]</sup> was then used to transform this data into a haploview fileset. Haplo View (version 4.1) was used to analyse the generated haploview fileset to determine the significant SNPs using a cutoff value of 0.001. (Barrett *et al.*, 2005)<sup>[1]</sup>.

### 3. Results and Discussion

#### 3.1 Allelic diversity of *YLI* in 3K RG panel subset

The *YLI* gene had 222 variants (including 135 INDELS and 87 SNPs), although only six of them were nonsynonymous (Table 1), and there were no changes in exons or splice donor/acceptor sites among the 186 rice accessions surveyed. Six non-synonymous variations found in the sequences leads

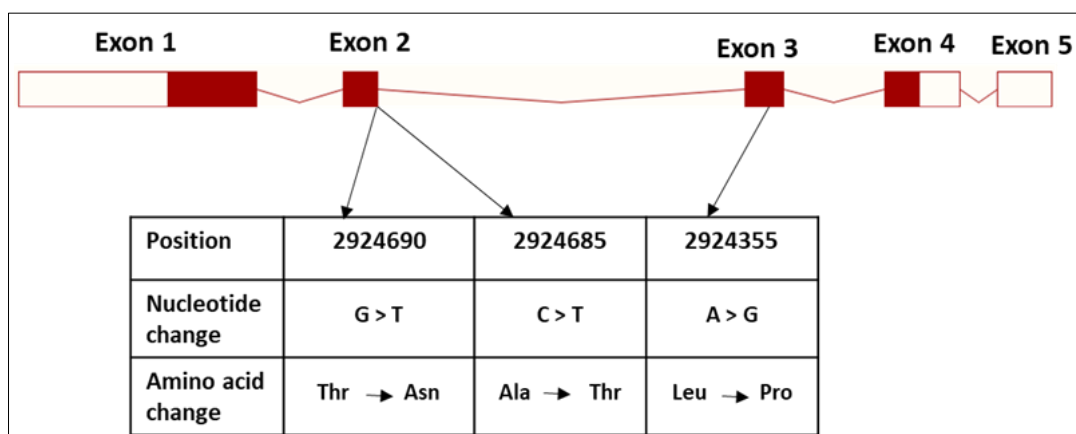
to change of several amino acids (Ala→Thr, Thr→Ala, Leu→Pro, Val→Ala, Ala→Thr, and Thr→Asn) among them only three were significant (Table 2 and Fig. 1). Out of the six Non synonymous SNPs, three were located in the exon 2, two at exon 3 and one at exon 5 of *YLI*.

**Table 1:** Non synonymous SNPs allelic variations in *YLI* gene for the 3K RG panel subset

Position	Position	Alleles	Sequence polymorphisms	Type
2922961	Exon 5	1 bp	C/T	SNP
2924350	Exon 3	1 bp	T/C	SNP
2924355	Exon 3	1 bp	A/G	SNP
2924681	Exon 2	1 bp	A/G	SNP
2924685	Exon 2	1 bp	C/T	SNP
2924690	Exon 2	1 bp	G/T	SNP

**Table 2:** SNP Effect of *YLI* in 3K RG panel subset

POSITION	Alt (ANN)	Effect (ANN)	Putative_impact (ANN)	HGV S.c (ANN)	HGV S.p (ANN)
chr02-2922961	T	Missense variant	Moderate	c.376C>T	p.Ala126Thr
chr02-2924350	C	Missense variant	Moderate	c.136T>C	p.Thr46Ala
chr02-2924355	G	Missense variant	Moderate	c.131A>G	p.Leu44Pro
chr02-2924681	G	Missense variant	Moderate	c.59A>G	p.Val20Ala
chr02-2924685	T	Missense variant	Moderate	c.55C>T	p.Ala19Thr
chr02-2924690	T	Missense variant	Moderate	c.50G>T	p.Thr17Asn



**Fig 1:** Three significant SNPs and their respective amino acid changes. (Solid blocks - exons, lines in between solid blocks – introns)

#### 3.2 Haplotype analysis of 217 genotypes and diversity between haplotypes

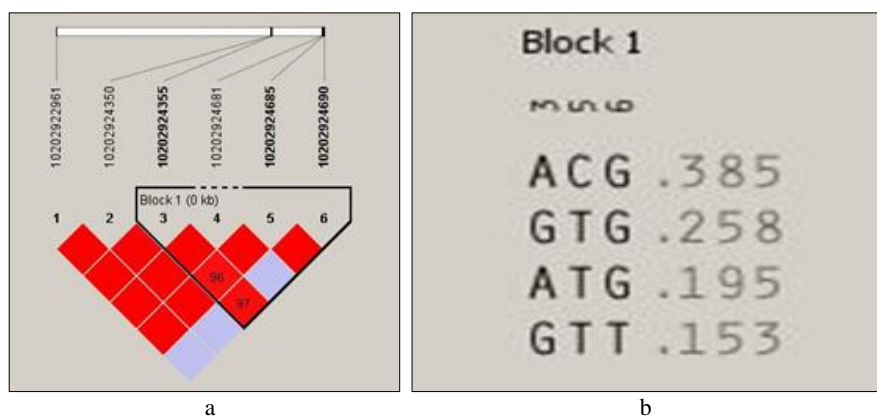
Among the six significant non-synonymous SNPs found in the *YLI* gene, four haplotypes were grouped, comprising 69, 46, 35 and 27 genotypes respectively based on only three SNPs (2924355, 2924685 and 2924690 (Table 3). The LD plot (Fig. 2) revealed a low correlation coefficient ( $r^2 = 5$ ) between three SNPs implicated in haplotype grouping. Previous studies on *YLI*, a nucleus-encoded chloroplast protein indicated that it is involved in chloroplast development and the efficient biogenesis of chloroplast ATP synthase (Chen *et al.*, 2016)<sup>[2]</sup>. The *YLI* mutation resulted in lower chlorophyll concentration, aberrant chloroplast shape, and reduced photochemical efficiency. Furthermore, *YLI* loss impairs the expression of genes involved in chloroplast

formation and photosynthesis (Chen *et al.*, 2016)<sup>[2]</sup>. The present study showed the presence of an allelic polymorphism and haplotype grouping among the 3K RG panel subset in *YLI*, which requires further phenotypic validation.

**Table 3:** Significant SNPs in *YLI* for the 3K RG panel subset

Marker Number	SNP site	Position	HWpval	MAF	Alleles
1	10202922961	2922961	3.1328E-13	0.038	C:T
2	10202924350	2924350	3.1328E-13	0.038	T:C
3	10202924355	2924355	6.6714E-49	0.419	A:G
4	10202924681	2924681	3.1328E-13	0.038	A:G
5	10202924685	2924685	3.9112E-46	0.391	T:C
6	10202924690	2924690	9.567E-29	0.162	G:T

HWpval- Hardy Weinberg p value, MAF- Minor Allele Frequency



**Fig 2:** LD plot of three significant SNPs located in *YLI* (LOC\_Os02g05890) (a) and its haplotype groups (b) using Haploview

#### 4. Conclusion

*YLI* gene has 222 allelic variations in 3K RG subset, only six of which were non-synonymous, and none of them were found to be non-sense. Haplotype analysis revealed four haplotype groups based on three significant non-synonymous SNPs. The identification of novel variations/deletions in *YLI* gene can provide useful information for deploying haplotype-based breeding to develop photosynthetically efficient rice varieties with improved yield potential.

#### 5. References

1. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263-5.
2. Chen F, Dong G, Wu L, Wang F, Yang X, Ma X, *et al.* A Nucleus-Encoded Chloroplast Protein YL1 Is Involved in Chloroplast Development and Efficient Biogenesis of Chloroplast ATP Synthase in Rice. *Scientific reports*. 2016;6:32295. <https://doi.org/10.1038/srep32295>
3. Grassini P, Eskridge KM, Cassman KG. Distinguishing between yield advances and yield plateaus in historical crop production trends. *Nat. Commun*. 2013;4:2918.
4. Hibberd JM, Sheehy JE, Langdale JA. Using C4 photosynthesis to increase the yield of rice-rationale and feasibility. *Curr. Opin. Plant Biol*. 2008;11(2):228-231. [10.1016/j.pbi.2007.11.002](https://doi.org/10.1016/j.pbi.2007.11.002)
5. Jaggard KW, Qi A, Ober ES. Possible changes to arable crop yields by 2050. *Philos. Trans. R. Soc. Lond. B Biol. Sci*. 2010;365:2835-2851.
6. Lee DY, Hua L, Khoshravesh R, Giuliani R, Kumar I, Cousins A, *et al.* Engineering chloroplast development in rice through cell-specific control of endogenous genetic circuits. *Plant biotechnology journal*. 2021;19(11):2291-2303. <https://doi.org/10.1111/pbi.13660>
7. Mansueto L, Fuentes RR, Borja FN, Detras J, Abriol-Santos JM, Chebotarov D, *et al.* Rice SNP-seek database update: new SNPs, indels, and queries. *Nucleic acids research*. 2017;45(D1):D1075-D81.
8. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*. 2007;81(3):559-75.
9. Von Caemmerer S, Quick WP, Furbank RT. The Development of C4 Rice: Current Progress and Future Challenges. *Science*. 2012;336(6089):1671-1672. [10.1126/science.1220177](https://doi.org/10.1126/science.1220177)