Madras Agric.J., 2022; https://doi.org/10.29321/MAJ.10.000582 (online first draft)

RESEARCH ARTICLE

Allelic Diversity of *OsGW5.1* Regulating Grain Width in Rice

Bharathi Ayyenar, Ameena Premnath, Sudhakar D* and Raveendran Muthurajan*

Department of Plant Biotechnology, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore - 641 003.

Corresponding author mail id: raveendrantnau@gmail.com; dsudhakar@hotmail.com

ABSTRACT

Rice is the staple food for more than half of the global population and rice production has to be increased by at least 50% during 2030 to meet the requirements of the rapidly increasing population. Rice yield is determined by several quantitative traits *viz*. number of productive tillers, panicle length, number of grains per panicle, grain size, and photosynthetic efficiency. Grain size is a major determinant of rice yield and it is primarily governed by grain length, grain width, and grain thickness. Several candidate genes modulating yield components have been discovered and put into breeding applications. *OsGW5.1* (LOC_Os05g25350), encoding for a receptor-like kinase is reported to alter grain width in rice. Advancements in genomics and re-sequencing of 3000 rice genotypes paved way for measuring allelic diversity of putative candidate genes, identification of elite haplotypes, and developing functional markers for breeding applications. In this study, the allelic diversity of *OsGW5.1* was analyzed in a set of 151 IRRI-3K panel lines. Eight non-synonymous SNPs were observed and haplotype analysis identified four haplogroups of *OsGW5.1*. Differences in grain width of the different haplogroups were tested by Dunnet's test. Results revealed that haplotype 1 (H1) and haplotype 2 (H2) showed a significant difference in grain width. KIKUBA (haplogroup 1) was found to possess greater grain width (3.62 mm) whereas NS1515 (haplotype 2) possessed a slender grain type (1.88 mm). Identified superior donors possessing elite haplotypes of *GW5.1* may be utilized in haplotype-based breeding to develop next generation tailor-made high yielding rice varieties with desired grain size.

Keywords: *Rice; Grain Width; Allelic Diversity; Haplotype Analysis*

INTRODUCTION

Rice, one of the most important staple food crop species, is consumed by half of the world's population, particularly in Asia. By 2050, the world's population is predicted to exceed 9.1 billion which requires a simultaneous increase in rice production (Hibberd *et al*., 2008). Rice productivity has undergone two major leaps *viz.* first during 1960's through the introduction of semi-dwarfism and secondly during the 1980's through the introduction of hybrids. After that, no major breakthrough in rice yields has been achieved. A number of issues, including water scarcity, soil salinity, disease, climate change, and reduced arable land area have indicated exacerbated food shortages over the next 50 years, demanding immediate improvement in crop production (Zhang, 2007). Any further increase in rice production has to overcome challenges *viz.* yield plateau, shrinking natural (land and water) resources, and increased occurrence of abiotic stresses, pests, and diseases due to changing climate (Khush *et al*., 2005). Designing high-yielding rice genotypes adapted to diverse environments depends on the exploitation of genetic resources through phenomics and well knitted genomics-assisted breeding tools. In this study, attempts were made to compile information on various yield genes in rice and survey the allelic diversity of key candidate genes. Literature survey enabled identification of 189 genes related to grain yield attributes and 63 genes regulating grain quality traits. Identification of favorable alleles associated with yield traits may accelerate haplotype based breeding in rice.

An ideal rice cultivar should have high grain yield potential with improved grain shape, nutritional value, disease resistance, and stress tolerance (Rosegrant and Cline, 2003). Rice grain size is a significant trait influencing rice yield and quality.It is also influenced by three geometrical dimensions including grain length, width and thickness with a wide range of variation (Li *et al*., 2019). Domestication indicators include reduced nucleotide diversity and changed allele frequency in domestication loci. Thus, bridging the molecular study of gene function and domestication research requires a thorough understanding of the molecular basis for natural variation in domesticated features (Purugganan, 2019).

Re-sequencing of 3K diverse rice germplasm lines enabled identification and exploring allelic/haplotype variations and thus harnessing genetic diversity (Abbai *et al*., 2019). Ultimately, this paved way for the identification of novel donors and novel alleles associated with the traits of interest, which can in turn be deployed in crop improvement (Varshney *et al*., 2018). In this context, the 3K RG re-sequencing project holds great promise for harnessing genetic diversity in rice (Li *et al*., 2014). The identified superior versions of candidate genes could be combined *via* the recently established, fast and robust 'haplotype assembly' concept (Bevan *et al*., 2017). For grain width, the gene *GW5.1* was encoding a receptor-like kinase (LOC_Os05g25350) was selected (Zhang *et al*., 2021). The mutant *gw5.1* lines showed wider grain width and the lines which harbor the *GW5.1* gene produced narrow grains (slender grain type). Allelic diversity of *OsGW5.1* was measured using appropriate statistical tools and results are discussed.

MATERIAL AND METHODS

A set of 151 diverse rice accessions from IRRI-3K Panel was assembled (Table 1) and evaluated during *kharif* 2020-21. The grain width of 151 accessions was measured using Vernier caliper. Ten grains per genotype were used for measurement of grain width. Measured data were subjected to descriptive statistics analysis using Minitab 19 and frequency distribution and box plot graphs were created (Allen, 2019).

Nucleotide polymorphisms existing in the *GW5.1* (LOC_Os05g25350) were visualized using SNP seek database and haplotype analysis was performed by downloading the allelic variants of the *GW5.1* gene in PLINK format (Weeks, 2010). Haploview 4.2 software was used to create haplotype groups and linkage disequilibrium blocks containing allelic variation, with settings such as a HW p-value (Hardy-Weinberg p value) of 0.001, a minimum genotype percent of 75, and a minimum minor allele frequency of 0.001(Barrett *et al*., 2005). Dunnett's test, performed with Minitab 19 Statistical Software, revealed significant differences between the generated haplotype groups (Allen, 2019).

RESULTS AND DISCUSSION

Descriptive statistics analysis defined the measures of variability among the 151 accessions. The coefficient of variation was found to be moderately high, indicating there is abundant variation in the population and thus selection of suitable donors for the respective trait can be done with the genotypes available (Table 2). The range of grain width varied from 1.88 mm to 3.62 mm, with an average of 2.65 mm, according to the box plot curve (Fig1a). According to the histogram, the population has a normal distribution (Fig1b).

All the 151 germplasm lines were clustered into 4 haplogroups based on the sequence diversity of *OsGW5.1*. Eight non-synonymous SNPs identified in the *OsGW5.1* (LOC_Os05g25350) were identified. The

haplotype group *viz*. H1, H2, H3 and H4 were found to contain 121, 22, 7 and 1 genotype, respectively. The H4 haplotype showed a heterozygous allele and it was not taken for further analysis. Eight SNPs (Table 3) from *GW5.1* gene were employed in the construction of four haplotypes using haploview, as illustrated in Fig. 2a. The linkage disequilibrium block depicts the alleles at nearby positions that can occur on the same haplotype (Fig. 2b). In contrast to Mendel's law of independent assortment, linkage disequilibrium refers to the non-random association of alleles in a locus. The SNP site 10514718976 and 10514720069 had the highest LD r2 value which indicates (Fig 2b-Values inside the blocks) high correlation (98) between these two sites.

From Table 4, it was understood that haplotype H1 contained maximum number of genotypes (121) followed by H2 (22) and H3 (7). H4 was represented by only one genotype that showed heterozygous allele and was not included in the phenotypic analysis. The maximum grain width was observed in haplotype 1 (3.62 mm), followed by H3 (2.97 mm) and H2 (2.83 mm). From the Dunnett's test results, it was observed that there is a significant difference between H2-H1 (p values <0.05) (Table 5 and Figure 3). KIKUBA was identified to possess a wider grain type (3.62 mm) observed in the H1 haplotype group and NS1515 had a narrow grain type (slender, 1.88 mm) present in the haplotype group H2. From this study, it was found that the genotype, NS1515 can be used as a donor for the development of rice genotypes with slender grain type. Similarly, Ramasamy *et al.* (2021) and Sundaramoorthy *et al.* (2022) identified superior haplotypes for *An-1* and *Gn1a* genes, respectively.

CONCLUSION

The study resulted in unravelling the allelic variations in the *GW5.1* gene with the help of 151 genotypes from 3K RG panel, and their grouping based on haplotyping identified a superior haplotype H2 which can be exploited in allele mining and plant breeding programmes. Thus, haplotype-based breeding is expected to aid in the development of high-quality rice varieties with desired grain width to fulfill the growing demands of the rice consuming population.

Funding and Acknowledgment

The authors acknowledge the receipt of seeds.

Ethics statement

No specific permits were required for the described field studies because no human or animal subjects were involved in this research.

Originality and plagiarism

This is original research work and any work and/or words of others, has been appropriately cited.

Consent for publication

All the authors agreed to publish the content.

Competing interests

There were no conflict of interest in the publication of this content.

Data availability

All the data of this manuscript are included in the MS. No separate external data source is required. If anything is required from the MS, certainly, this will be extended by communicating with the corresponding author through corresponding official mail; raveendrantnau@gmail.com; dsudhakar@hotmail.com

Author contributions

Idea conceptualization - MR, Experiments - BA, Guidance – MR, DS, Writing original draft – BA, AP; Writing-reviewing & editing – BA, AP, DS, MR.

REFERENCES

Abbai, R., Singh, V. K., Nachimuthu, V. V., Sinha, P., Selvaraj, R., Vipparla, A. K., Singh, A. K., Singh, U. M., Varshney, R. K. and A. Kumar. 2019. Haplotype analysis of key genes governing grain yield and quality traits across 3K RG panel reveals scope for the development of tailor-made rice with enhanced genetic gains. *Plant biotechnology journal*, **17(8)**: 1612-1622.

Allen, T. T. 2019. Introduction to Engineering Statistics and Lean Six Sigma (London: Springer) Software overview and methods review: Minitab 575-600.

Barrett, J. C., Fry, B., Maller, J. and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**: 263–265.

Bevan, M. W., Uauy, C., Wulff, B. B., Zhou, J., Krasileva, K. and M. D. Clark. 2017. Genomic innovation for crop improvement. *Nature*, **543(7645)**: 346-354.

Hibberd, J. M., Sheehy, J. E. and J. A. Langdale. 2008. Using C4 photosynthesis to increase the yield of rice-rationale and feasibility. *Curr. Opin. Plant Biol*., **11**:228-231.

Khush, G. S. 2005. What it will take to feed 5.0 billion rice consumers in 2030. *Plant Mol. Biol*., **59**: 1-6.

Li, J. Y., Wang, J. and R. S. Zeigler. 2014. The 3,000 rice genomes project: new opportunities and challenges for future rice research. *Gigascience*, **3(1)**: 2047-217X.

Li, N., Xu, R. and Y. Li. 2019. Molecular networks of seed size control in plants. *Annu. Rev. Plant Biol*., **70**: 435-463.

Priya, R. S., Kalaimagal, T., Rajeswari, S., Prasanth, R. A. and M. Raveendran. 2021. Allele mining for the grain number gene I in rice (*Oryza sativa* L.). *Electronic Journal of Plant Breeding*, **12(3)**:772-779.

Purugganan, M. D. 2019. Evolutionary insights into the nature of plant domestication. *Curr. Biol*., **29**: R705-R714.

Rosegrant, M. W. and S. A. Cline. 2003. Global food security: challenges and policies. *Science*, **302**: 1917–1919.

Sundaramoorthy, M., Ramasamy, S. P., Rajagopalan, V. R., Ramalingam, A. P., Ayyenar, B., Mohanavel, V., Narayanan, M. B. and R. Muthurajan. 2022. Pilot scale genome wide association mapping identified novel loci for grain yield traits in rice. *Plant Physiol. Rep*., https://doi.org/10.1007/s40502-021-00641-w.

Varshney, R. K., Singh, V. K., Kumar, A., Powell, W. and M. E. Sorrells. 2018. Can genomics deliver climate-change ready crops?. *Current opinion in plant biology*, **45**: 205-211.

Weeks, J. P. 2010. Plink: An R package for linking mixed-format tests using IRT-based methods. *Journal of Statistical Software*, **35(12)**: 1-33.

Zhang, J., Zhang, D., Fan, Y., Li, C., Xu, P., Li, W., Sun, Q., Huang, X., Zhang, C., Wu, L., Yang, H., Wang, S., Su, X., Li, X., Song, Y., Wu, M., Lian, X. and Y. Li. 2021. The identification of grain size genes by RapMap reveals directional selection during rice domestication. *Nature Communications*, **12(1)**: 1-18.

Zhang, Q. 2007. Strategies for developing Green Super Rice. *Proc. Natl. Acad. Sci. USA*., **104**: 16402-16409.

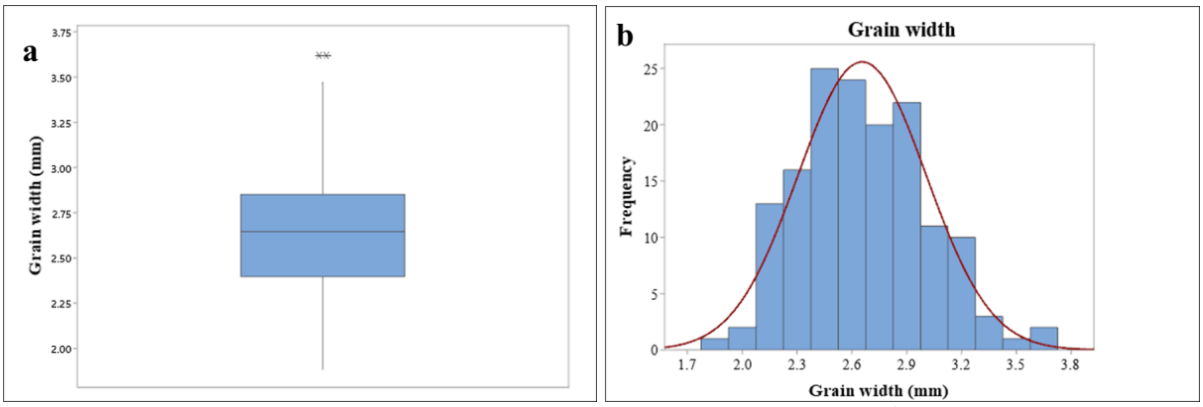

Figure 1. a) Boxplot for grain width; b) Frequency distribution curve for grain width

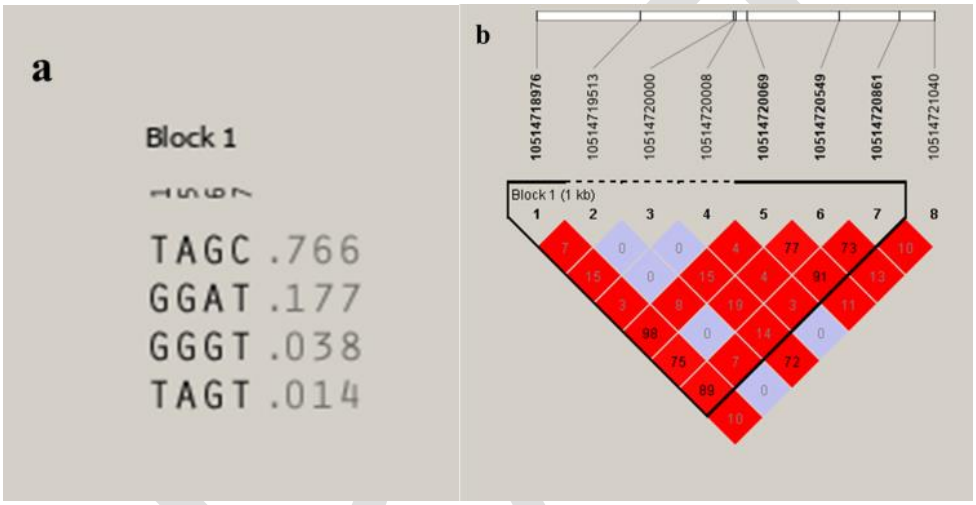

Figure 2. a) Formation of haplotype groups based on allelic variation;
b) Linkage Disequilibrium (LD) block of *GW5.1* gene for 150 genotypes

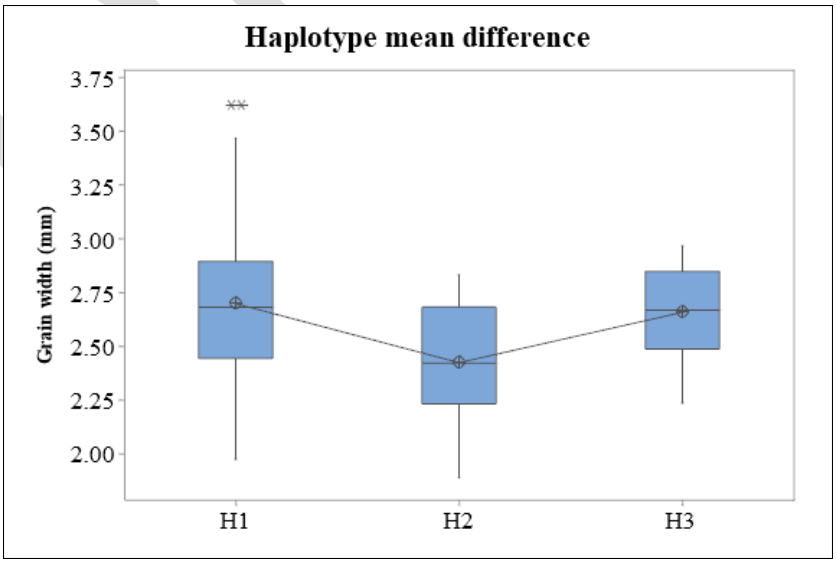

Figure 3. Box plot curve for different haplotypes indicating significant difference by Dunnett's test

Table 1. List of genotypes used in the present study and their haplotype groups

S. No.	Accessions	H	S. No.	Accessions	H	S. No.	Accessions	H
1	CX145	H1	51	IRIS_313-8595	H1	101	IRIS_313-11885	H3
2	CX94	H1	52	IRIS_313-8638	H1	102	IRIS_313-11126	H1
3	IRIS_313-8288	H1	53	IRIS_313-8509	H2	103	IRIS_313-10880	H1
4	IRIS_313-8265	H1	54	IRIS_313-15902	H1	104	IRIS_313-11256	H1
5	IRIS_313-10403	H1	55	IRIS_313-10298	H1	105	IRIS_313-11262	H1
6	IRIS_313-10423	H1	56	IRIS_313-10353	H1	106	IRIS_313-11263	H1
7	IRIS_313-7668	H1	57	IRIS_313-11744	H1	107	IRIS_313-11266	H1
8	IRIS_313-7684	H2	58	IRIS_313-11377	H1	108	IRIS_313-11271	H1
9	IRIS_313-8067	H2	59	IRIS_313-11624	H1	109	IRIS_313-11308	H1
10	IRIS_313-7650	H2	60	IRIS_313-10664	H1	110	IRIS_313-11310	H1
11	IRIS_313-7799	H1	61	IRIS_313-11812	H1	111	IRIS_313-11192	H1
12	IRIS_313-8703	H1	62	IRIS_313-10924	H1	112	IRIS_313-11241	H2
13	IRIS_313-8660	H1	63	IRIS_313-11656	H1	113	IRIS_313-11358	H1
14	IRIS_313-8697	H1	64	IRIS_313-11852	H1	114	IRIS_313-10986	H1
15	IRIS_313-9427	H1	65	IRIS_313-11853	H3	115	IRIS_313-10775	H1
16	IRIS_313-9574	H1	66	IRIS_313-11692	H1	116	IRIS_313-11327	H1
17	IRIS_313-9067	H1	67	IRIS_313-11854	H1	117	IRIS_313-10974	H1
18	IRIS_313-9503	H1	68	IRIS_313-11671	H1	118	IRIS_313-10754	H1
19	IRIS_313-9348	H1	69	IRIS_313-11676	H1	119	IRIS_313-11076	H1
20	IRIS_313-10113	H1	70	IRIS_313-11721	H1	120	IRIS_313-10555	H1
21	IRIS_313-10047	H1	71	IRIS_313-10769	H1	121	IRIS_313-11418	H2
22	IRIS_313-9302	H1	72	IRIS_313-10966	H1	122	IRIS_313-11242	H1
23	IRIS_313-10224	H3	73	IRIS_313-11802	H1	123	IRIS_313-11399	H1
24	IRIS_313-10167	H2	74	IRIS_313-11505	H3	124	IRIS_313-11141	H1
25	IRIS_313-9968	H1	75	IRIS_313-11506	H1	125	IRIS_313-11372	H1
26	IRIS_313-9709	H2	76	IRIS_313-11867	H1	126	IRIS_313-11181	H1
27	IRIS_313-9696	H1	77	IRIS_313-11814	H1	127	IRIS_313-11254	H1
28	IRIS_313-9605	H1	78	IRIS_313-11543	H1	128	IRIS_313-10433	H1
29	IRIS_313-8956	H2	79	IRIS_313-11577	H1	129	IRIS_313-10822	H1
30	IRIS_313-9325	H1	80	IRIS_313-11234	H3	130	IRIS_313-10753	H1
31	IRIS_313-9320	H1	81	IRIS_313-10609	H1	131	IRIS_313-10897	H1
32	IRIS_313-9841	H1	82	IRIS_313-11869	H2	132	IRIS_313-11793	H2
33	IRIS_313-10000	H1	83	IRIS_313-11784	H1	133	IRIS_313-11467	H1
34	IRIS_313-9160	H1	84	IRIS_313-11760	H2	134	IRIS_313-11899	H2
35	IRIS_313-10054	H2	85	IRIS_313-11545	H2	135	IRIS_313-11808	H1
36	IRIS_313-9469	H3	86	IRIS_313-11811	H1	136	IRIS_313-11954	H1

S. No.	Accessions	H	S. No.	Accessions	H	S. No.	Accessions	H
37	IRIS_313-10171	H1	87	IRIS_313-11446	H1	137	IRIS_313-8349	H1
38	IRIS_313-10333	H1	88	IRIS_313-11638	H1	138	IRIS_313-8493	H2
39	IRIS_313-10016	H1	89	IRIS_313-11644	H1	139	IRIS_313-9329	H1
40	IRIS_313-10260	H3	90	IRIS_313-11762	H2	140	IRIS_313-11758	H1
41	IRIS_313-10046	H1	91	IRIS_313-11761	H2	141	IRIS_313-12280	H2
42	IRIS_313-10300	H1	92	IRIS_313-10458	H1	142	IRIS_313-11509	H1
43	IRIS_313-10314	H1	93	IRIS_313-10955	H2	143	IRIS_313-11914	H2
44	IRIS_313-10396	H1	94	IRIS_313-11820	H1	144	IRIS_313-11225	H2
45	IRIS_313-10402	H1	95	IRIS_313-10717	H1	145	IRIS_313-11554	H1
46	IRIS_313-9227	H1	96	IRIS_313-12296	H1	146	IRIS_313-10441	H1
47	IRIS_313-8932	H1	97	IRIS_313-10715	H1	147	IRIS_313-10900	H1
48	IRIS_313-9116	H1	98	IRIS_313-11521	H1	148	IRIS_313-11251	H1
49	IRIS_313-9433	H1	99	IRIS_313-11548	H1	149	IRIS_313-10690	H1
50	IRIS_313-9986	H1	100	IRIS_313-11507	H1	150	IRIS_313-10628	H1
						151	IRIS_313-8293	H4

Table 2. Descriptive statistics for *OsGW5.1* among the study population

Statistics	Total population
Range	1.88 - 3.62 mm
Mean	2.65 mm
Median	2.64 mm
SE Mean	0.03
CV (%)	13.21

Table 3. Details of identified significant alleles

Marker number	SNP site	Position	HW pval	MAF	Alleles
1	**10514718976**	**14718976**	**8.84E-34**	**0.218**	**T:G**
2	10514719513	14719513	1.15E-08	0.022	A:T
3	10514720000	14720000	3.91E-09	0.04	A:T
4	10514720008	14720008	4.38E-05	0.011	G:A
5	**10514720069**	**14720069**	**5.53E-35**	**0.215**	**A:G**
6	**10514720549**	**14720549**	**1.53E-29**	**0.182**	**G:A**
7	**10514720861**	**14720861**	**1.36E-28**	**0.232**	**C:T**
8	10514721040	14721040	1.88E-06	0.03	A:G

Table 4. Statistical difference between haplotype groups

Statistics	H1 (121)	H2 (22)	H3 (7)
Range	1.93 - 3.62 mm	1.88 - 2.83 mm	2.23 - 2.97 mm
Mean	2.7 mm	2.42 mm	2.66 mm
Median	2.68	2.42 mm	2.67 mm
SE Mean	0.03	0.05	0.09
CV (%)	13.12	11.11	9.47

Table 5. Dunnett's Simultaneous Tests for Level Mean-Control Mean

Difference of levels	Mean difference	t value	p value
H2 - H1	-0.2747	-3.49	0.001
H3 - H1	-0.04	-0.31	0.942