

## **Abstract**

Line 54: remove spp from “Capsicum spp”

Line 59: Capsicum clade accessions are not clear here, clade may not be appropriate to use. Better to say Capsicum accessions

## **Introduction**

Line 71: Put reference

Line 80: remove “an”

Line 81: Remove “to improve crop yield” and Capitalize “t”

Line 111: the sentence seems not completed. Revisit

Line 120: change *Capsicum spp.* with unidentified capsicum spp.

Line 123: In all other parts of the manuscript, authors claimed five candidate genes except here. Change three to five

Line 124-126: Please rewrite the sentence to give meaningful objective of the study to readers.

## **General**

As it is known, there is source-sink, physiological and architectural related issues with respect to having more than one fruit on a single node which directly or indirectly influences the crop yield. Therefore, it will be nice if authors can give their reflection on this issue.

## **Materials and Methods**

Line 130: Change “intraspecific” to “interspecific”

Line 138: The number of plants for each line is not clear. What are biological replicates?

Line 142 – 144: While removing *C. baccatum* and *C. pubescences* for GWAS with the main reason of reducing the population structure, authors included 45 unidentified species. Is this because these species are known to be in any of the three species used, it needs explanation?

Line 150: Since the trait of interest is only one, flower number. Better to put the trait itself on the title instead of “phenotypic data”

Line 151 - 152: I suggest authors to change the sentence into “We completed data collection on the number of flowers per each node after flowers are set or distinguished at the six nodes”.

Line 152 – 154: In calculating the average number of flowers, are you dividing the total number of flowers grown on 18 nodes by the number of nodes for each line? Please make it clear. And why the data collection is different for CSHL population?

I suggest authors to upload phenotype data of all the accessions used in the experiment as supplementary file

Line 159: “all stages” should be qualified.

Line 162: mention the figure number in parenthesis

Line 200: *C. annuum* reference genome v0.1 missed the accession. Maybe it is to refer Dempsey. Please include that

Line 242 – 243: To estimate population structure, PCA is required but the structure is developed not by PCA. Please check this

Line 279 – 283: What authors did is good to compare results with sequences of other accessions as a pre-validation step. In line with it, I recommend to include a figure or modify Supplementary Figure 5a to show significant sequence variation among different lines on the candidate genes.

Line 290: here authors mentioned only cotyledon for RT-qPCR analysis, but in Line 173 – 174, young leaves were also included. Please be consistent

Line 297: put the reference figure number at the end of the sentence

## **Results**

Line 309: I recommend the title to be: “Variation for Number of Flower Per Bud and Shoot Apical Meristem Development”

Line 318 – 321: The average flower number observed is 1.02, 1.33 and 1.87, which logically seems the average flower number ranged from 1 to 2. But what is confusing here is how authors can classify flower numbers into three categories, viz Single, intermediate and multiple?

Line 326: based on fig 1, habanero set a maximum of up to 25 leaves, not 22. Revise it

Line 327: It is okay for the width of TM, but the term “taller” can’t be verified from the figure. Please check this

Line 354 – 355: I recommend to change “multiple flowers-per-node trait” with “flower number per node”

Line 365: Authors should refer to the data while mentioning about the skewed distribution of multi-flowers-per node trait in TH RILs

Line 371: Change the term “standing” with “observed”

Line 375: There is no additional information about the bin markers TH1, TH7 and TH11, authors should explain if they mention it here

Line 402: Besides mentioning about the relatively larger block size than the mean distance between SNPs, authors should provide the significance of this information.

Line 435: Fig 6 has no relevance to be mentioned at this position, please revise it

Line 442: three or four genes in chromosome 2? It confuses with Table 5

Line 447: at the end of the sentence, it is important to mention the reference genome used (based on table 5)

Line 449: Though, it is good to include SP5G candidate gene for validation, why authors could not show expression results of other candidate genes in the co-localized region?

Line 459: why WOX 9 is missed from the list?

Line 469: Please refer the figure number at the end of the sentence

Line 485: WOX9 mentioned here is WUSHEL-related homeobox 9 in Table 5. Can you put WOX9 in parenthesis in Table 5 or otherwise here? It is confusing. Same is true for SP5G gene in Table 5. Be consistent

Line 489- 490: Supplementary Table 1 is wrongly mentioned here. Please check

Generally:

- From the sequence data, it seemed WOX9 and ME12-like 5 are major candidates for flower number per node, as they have amino acid difference between lines. Both these genes are in Chr2
- from the co-localization result, genes in Chr 1, Chr 7 and Chr 11 might have contribution in regulating the trait
- Locus in Chr 5 is detected only by GWAS

Looking into the results (tables, figures and data presentations), much focus has been given to only genes in Chr 2 and Chr 5. Can you please provide any data related to genes associated with the above other regions, or give some concrete reason why you only focused on those ones?

## **Discussion**

Line 526 – 528: Not clear why this is mentioned here. Either substantiate the existence of epistasis in the trait of interest by providing some data from your own, or remove it.

## Tables and Figures

1. The trend of  $R^2$  values of QTLs in Table 3 and Table 4 is different. Why? Please give reason
2. Put Total  $R^2$  value in Table 4 under the  $R^2$  values of others
3. Table 5: For WUSCHEL-related homeobox 9 *include* **WOX9** in parenthesis; and For FLOWERING LOCUS T include "SELF PRUNING LOCUS T"
4. Figure 4. While Y-axis of the box plot is for number of flower per node, X-axis is not mentioned
5. Figure 6. The comparison will give meaning if authors include Chr 7 instead of Chr 5. Check the data
6. Figure 7. It is good that the expression result is presented for SP5G which conform other results, but what about results of other genes?
7. Figure 8. WOX9 is missed from the figure. It should be included, or authors need to give justification why it is not included
8. Supplementary Table 1. In which chromosome is SP5G is located (Chr 05 as it is in Table 5 or Chr 02 as appeared here)?. Check it and make correction
9. Supplementary Figure 1. There is no description given either in the figure itself or in Line 757 about what A, B and C are. Check it