CORRECTION

Vol. 173: 2041–2059, 2017

Schläpfer P., Zhang P., Wang C., Kim T., Banf M., Chae L., Dreher K., Chavali A.K., Nilo-Poyanco R., Bernard T., Kahn D., and Rhee S.Y. Genome-Wide Prediction of Metabolic Enzymes, Pathways, and Gene Clusters in Plants.

In one of the input files used for predicting metabolic gene clusters for potato (*Solanum tuberosum*), a gene location file downloaded using the biomart tool at Phytozome, protein coordinates were used instead of gene coordinates. Consequently, isoforms of a gene were treated as separate genes. In addition, 24 genes in corn (*Zea mays*) were misidentified as splice variants of other genes. This led to an underprediction of clusters in potato (379 versus 623) and corn (284 versus 306).

The final step in PlantClusterFinder Version 1.0 was to filter all the predicted gene clusters for the enrichment of metabolic genes and retain only those whose enzyme density is above the top 5% of the theoretical clusters (any region of the genome containing the same no. of genes). This filter was unintentionally turned off, and all predicted clusters were reported. Although not all of them are above the top 5%, they all were above the top 27%. Predicted clusters below the top 5% cut can encode functional metabolic gene clusters. Using the published boundaries of the 12 known clusters included in the prediction validation leads to a ranking among the top 6.7%.

The authors have corrected the two errors and reperformed all the analyses for the 18 genomes in the study. While there was a reduction (as expected) in the overall numbers of predicted clusters, all the major conclusions remain unchanged. The changed numbers include the following:

- The total number of predicted clusters across the 18 analyzed genomes decreased by 34% (7,926 versus previously 11,969), with 426 (previously 665) clusters on average per species.

- The total number of high-confidence clusters in four species analyzed with coexpression data are 195 (previously 233).

- More than 1,200 (previously 1,700) clusters contained signature enzymes that could generate specialized metabolite scaffolds and tailoring enzymes that could modify the scaffolds.

- A total of 42 (previously 43) high-confidence clusters contained signature and tailoring enzymes. The cluster that didn't make the top 5% cut is rice (*Oryza sativa*) cluster C42_4 (belongs to the top 6.31%).

- Finally, 51% (previously 50%) of all clusters and 30% (previously 24%) of high-confidence clusters can be considered to encode primary metabolism.

In addition, there were a few minor errors in the figures. Revised versions of the figures have been provided below.

In the Figure 3C legend, "carboxylic ester hydrolase" should be "thromboxane synthase." In the figure, "polysaccharide-lyase" should be placed under "Others," not "Tailoring enzymes."

In Figure 5A, to see if specialized and all other pathways are differently clustered, the authors replaced the two-sided Kolmogorov-Smirnov test with two-sample *t* test (resulting in a *P* value of 3.3570e-04) since there was a statement on the difference of the mean and not the distributions themselves.

The authors regenerated all relevant figures and tables showing results for the top 5% clusters (Figs. 2, 3, and 5; Supplemental Figs. S4–S9; Supplemental Tables S4, S7, S9, and S12). Revised versions of the figures included in the article follow this statement. Revised versions of the supplemental data can be found in the supplemental data area of the article. An updated version of the code (PlantClusterFinder Version 1.2) is available at https://dpb.carnegiescience.edu/labs/rhee-lab/software. This updated code can process genomes where gene identifier to protein identifier mappings are complex.

www.plantphysiol.org/cgi/doi/10.1104/pp.18.00079