

Supplementary Table 1 Comparison between PCMDB and PlantscRNadb on number of marker genes, literature, and cell types.

	Experimentally Validated		Identified by Bulk RNA-seq		Identified by scRNA-seq		Overlap between PCMDB and PlantscRNadb
	PlantscRNadb	PCMDB	PlantscRNadb	PCMDB	PlantscRNadb	PCMDB	
Number of Maker Genes							
Arabidopsis	353	1,718	249	4,004	14,717	17,022	12,750
Rice	72	623	26	3,690	5,362	17,003	5,306
Maize	84	442	0	11,516	5,817	7,015	NA(No GeneID)
Soybean	NA	79	NA	12,314	NA	0	NA(No GeneID)
Tomato	40	169	34	8,056	0	5,875	NA
Tobacco	NA	88	NA	1,045	NA	0	NA
Number of Literatures							
Arabidopsis	239	790	1	15	9	14	/
Rice	48	340	1	14	1	3	/
Maize	63	208	0	9	3	3	/
Soybean	NA	71	NA	7	NA	0	/
Tomato	27	149	2	11	0	1	/
Tobacco	NA	64	NA	2	NA	0	/
Number of Cell Types							
Arabidopsis	121	159	1	54	40	42	/
Rice	39	95	6	25	8	19	/
Maize	25	68	0	25	24	12	/
Soybean	NA	25	NA	14	NA	0	/
Tomato	25	69	8	27	0	9	/
Tobacco	NA	22	NA	6	NA	0	/

Supplementary Table 2 The number of literatures for different source during process of literature.

	Round 1	Round 2	Round 3
	Number of Experimental Literatures		
Arabidopsis	31,967	2,883	790
Rice	17,945	996	340
Maize	16,552	649	208
Soybean	11,593	336	71
Tomato	10,150	277	149
Tobacco	37,283	272	64
	Number of Bulk RNA-seq Literatures		
Arabidopsis	11,011	47	15
Rice	3,169	17	14
Maize	1,593	34	9
Soybean	1,821	22	7
Tomato	609	22	11
Tobacco	1,727	8	2
	Number of scRNA-seq Literatures		
Arabidopsis	28	14	14
Rice	6	3	3
Maize	3	3	3
Soybean	3	0	0
Tomato	2	1	1
Tobacco	3	0	0

Supplementary Figure 1 The number of marker genes for the top 20 cell types identified by bulk RNA-seq in each plant species.



Supplementary Figure 2 The number of marker genes for the top 20 cell types identified by scRNA-seq in each plant species.

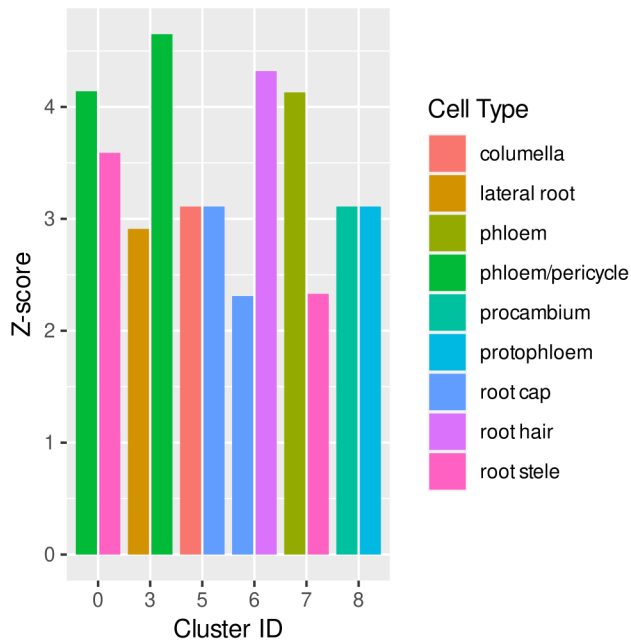
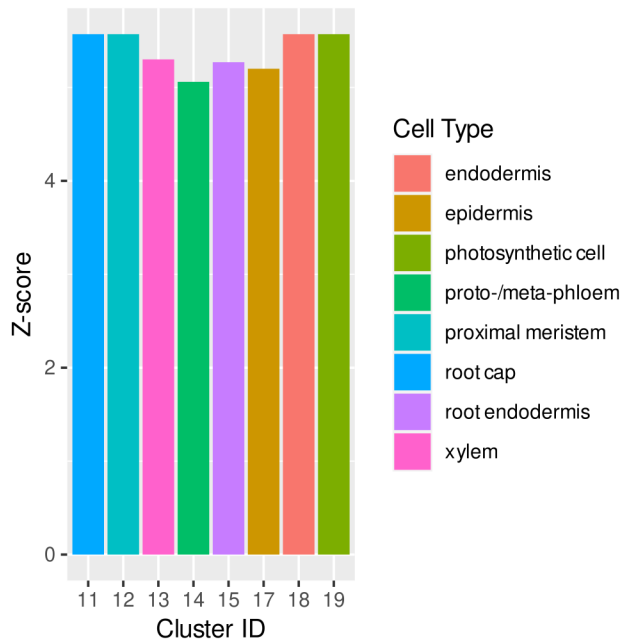
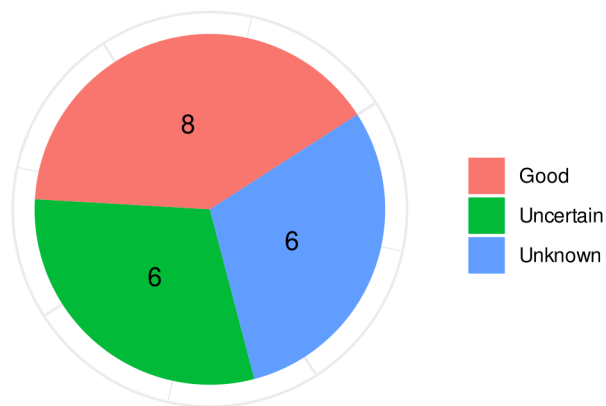
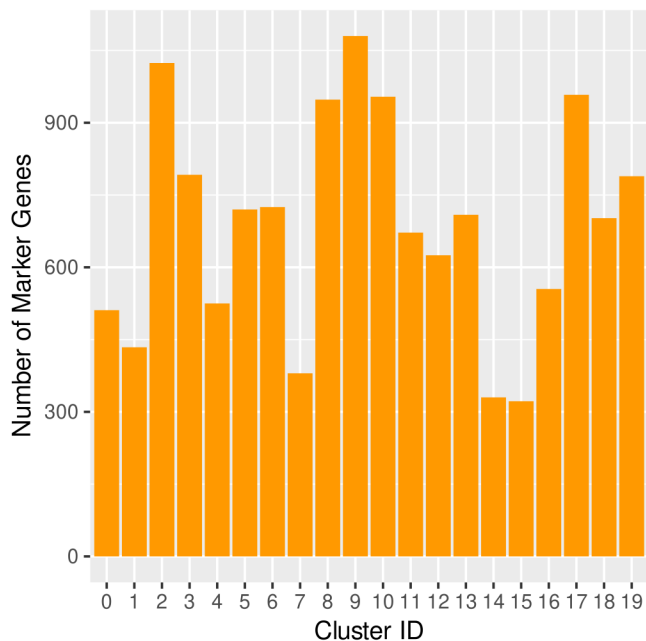


Supplementary Figure 3: The result of SCSA prediction for the scRNA-seq dataset of Zhang et al. (PMID:31004836) using marker genes from different sources.

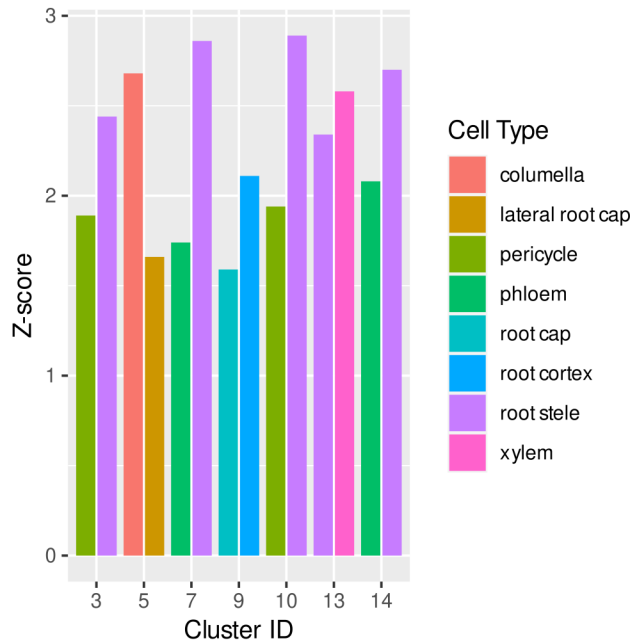
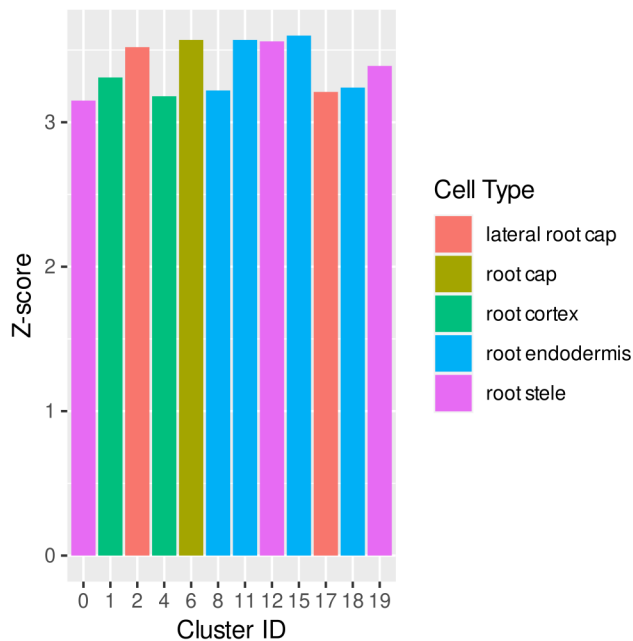
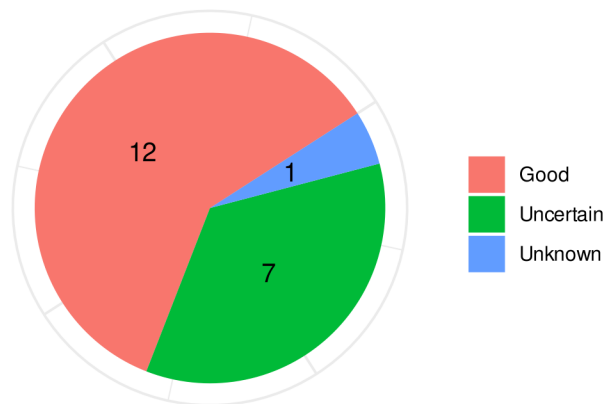
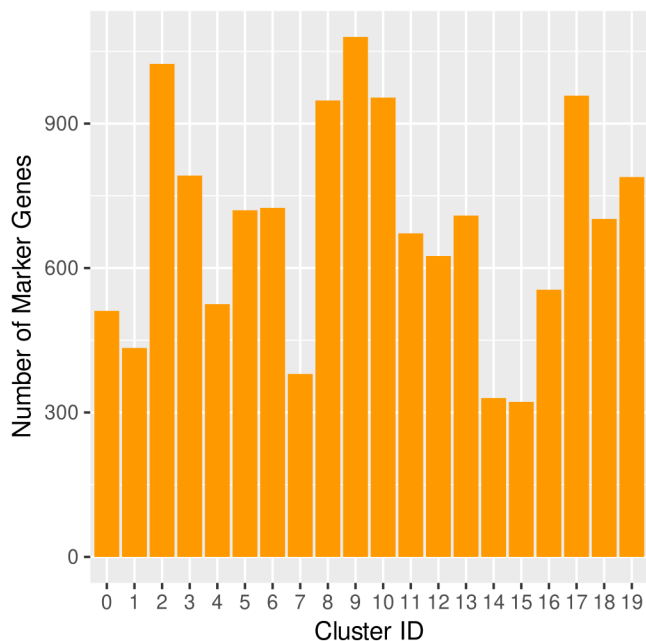
Left Top: The number of marker genes for each cluster of the input data. Right Top: The number of clusters classified into different types (Good, Uncertain, and Unknown) by SCSA. Left Bottom: Z-score of cluster label for clusters classified into the Good type. Right Bottom: Z-score of top 2 cluster labels for clusters classified into the Uncertain type.

- (A) The result of SCSA prediction using 103 known marker genes mentioned in Zhang et al.
- (B) The result of SCSA prediction using 923 experimental related root marker genes in PCMDB.
- (C) The result of SCSA prediction using 162 bulk RNA-seq related root marker genes in PCMDB.
- (D) The result of SCSA prediction using 11,654 scRNA-seq related root marker genes in PCMDB.
- (E) The result of SCSA prediction using all root marker genes in PCMDB.

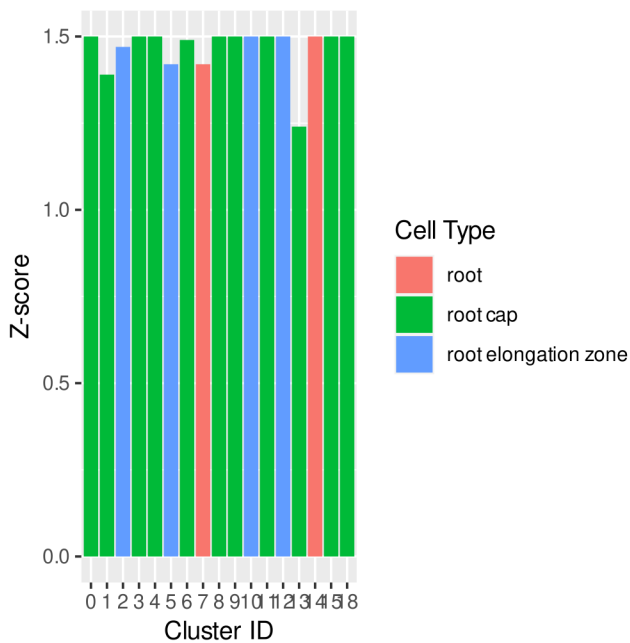
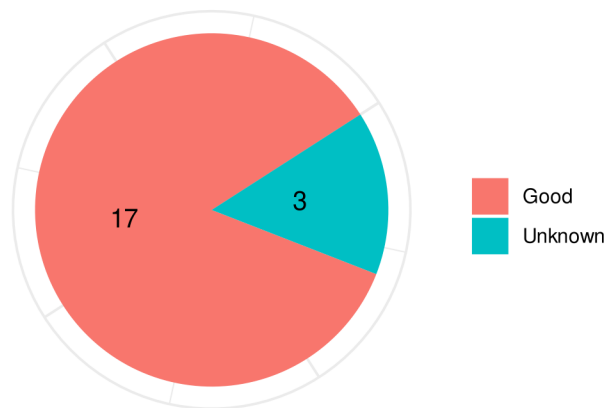
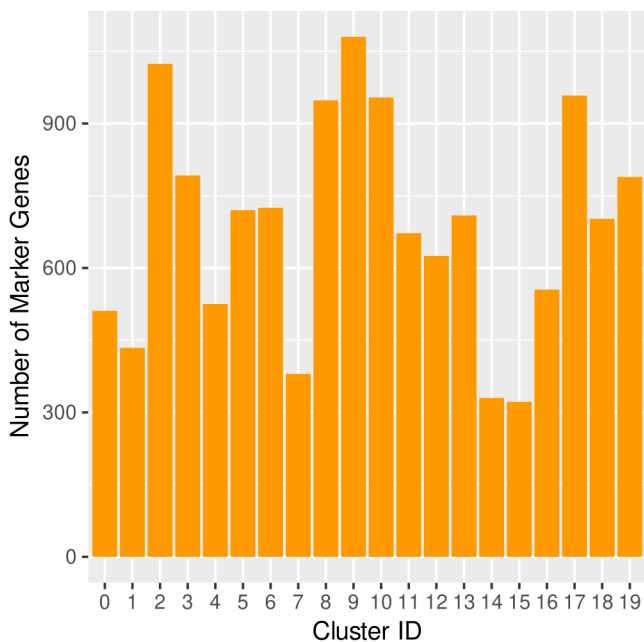
Supplementary Figure 3A. The result of SCSA prediction using 103 known marker genes mentioned in Zhang et al.



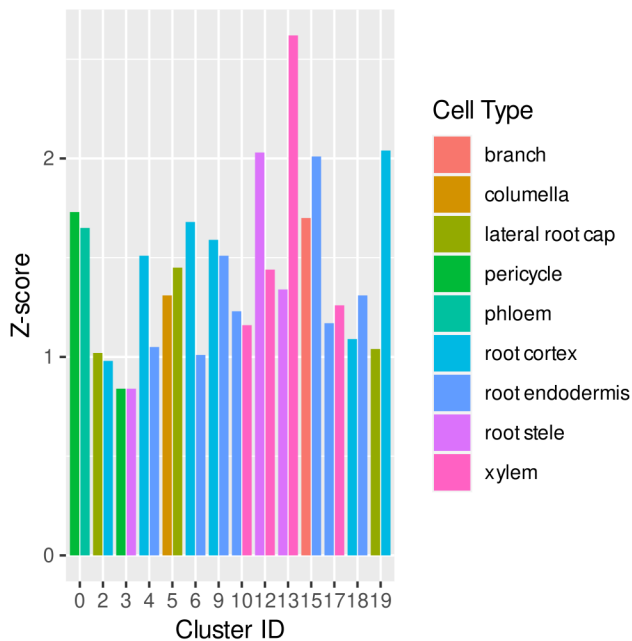
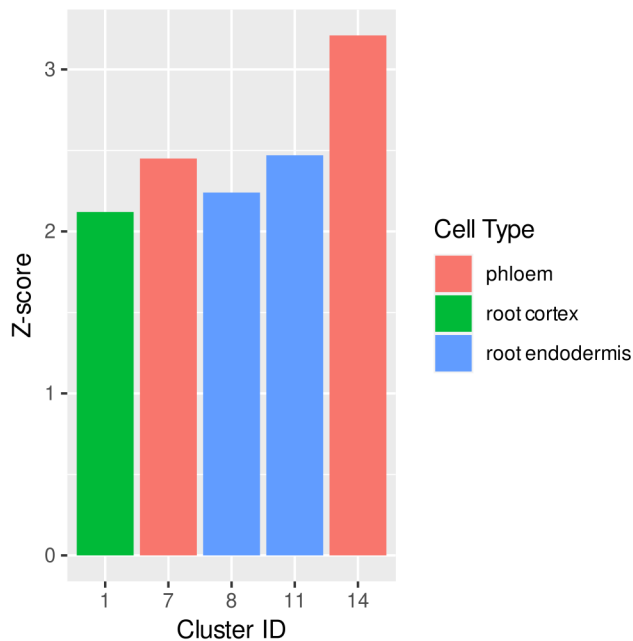
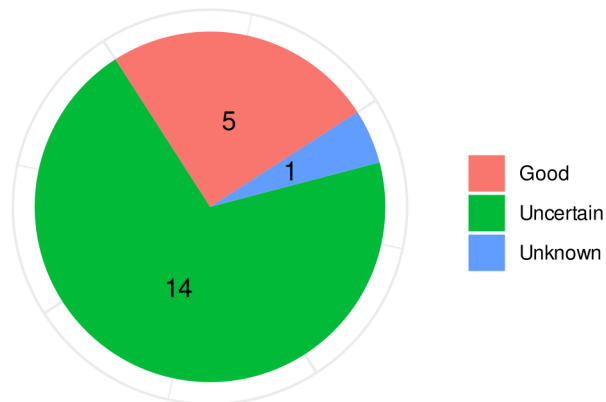
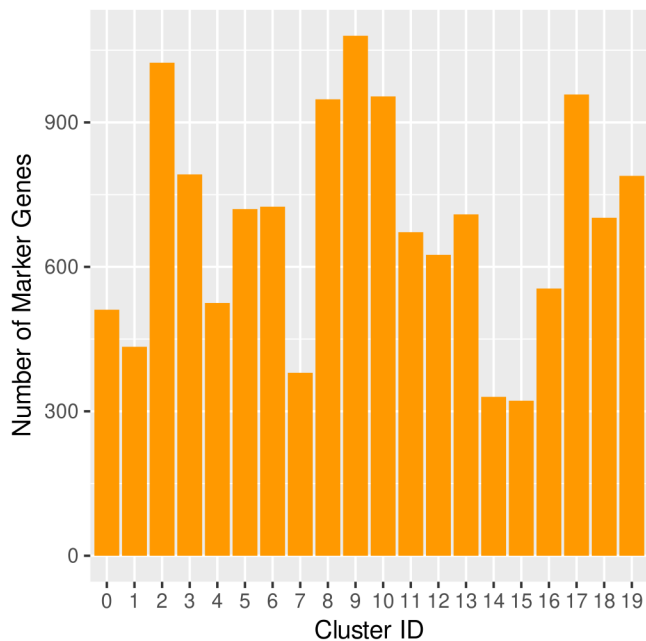
Supplementary Figure 3B. The result of SCSA prediction using 923 experimental related root marker genes in PCMDB.



Supplementary Figure 3C. The result of SCSA prediction using 162 bulk RNA-seq related root marker genes in PCMDB



Supplementary Figure 3D. The result of SCSA prediction using 11,654 scRNA-seq related root marker genes in PCMDB



Supplementary Figure 3E. The result of SCSA prediction using all root marker genes in PCMDB.

