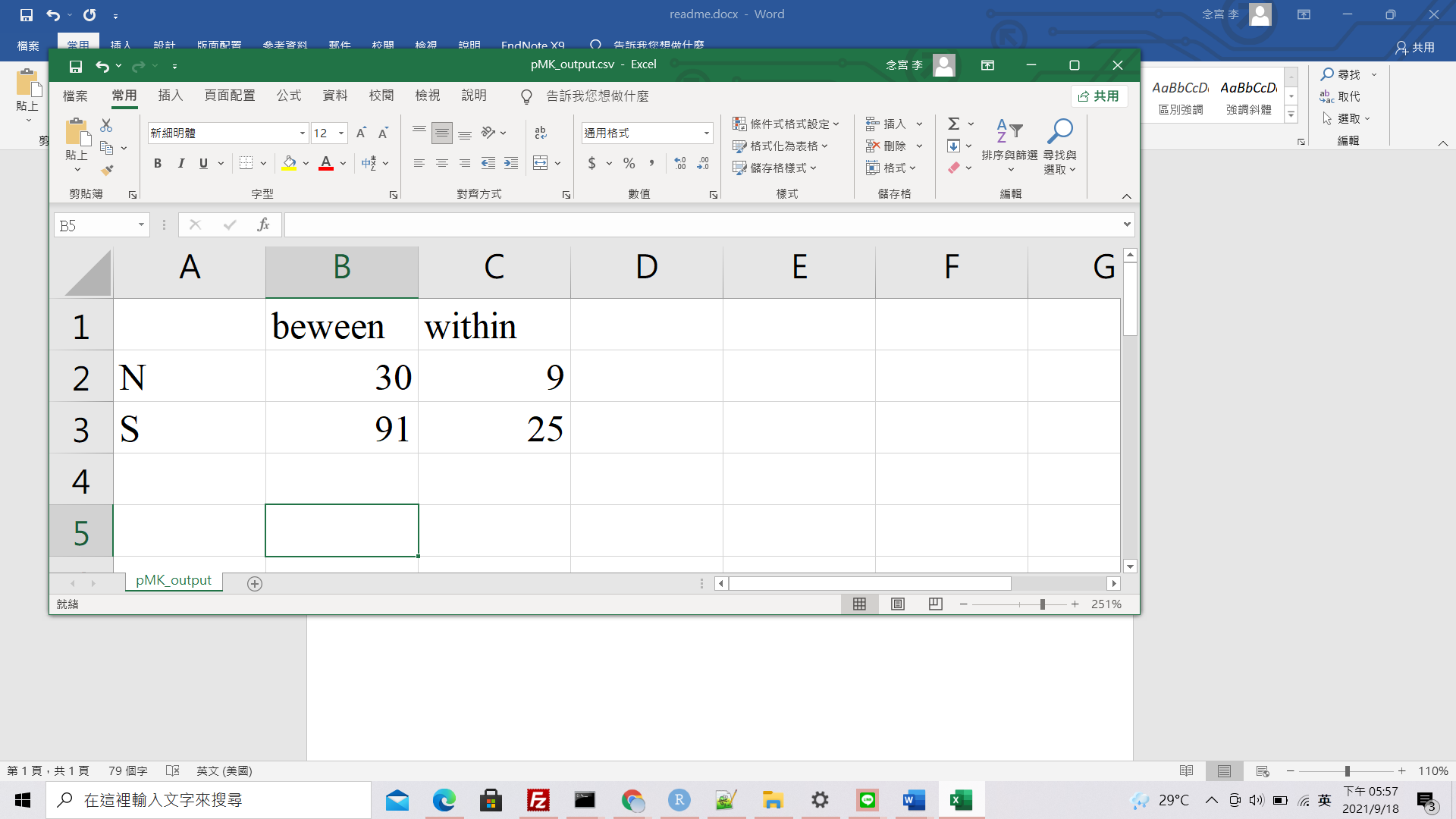
This code runs in R.

Step1: input *PAML* output file (example\_paml\_output.txt) to read the numbers of nonsynonymous and synonymous sites (*N* and *S*) and the nonsynonymous and synonymous substitution rates (*dN* and *dS*) estimated from *PAML*

Step2: input group file (example\_groupfile.csv)

The group file contains the sequence ID and its group and cluster. The program will average the nonsynonymous changes and synonymous changes according to the group, and then summarize the changes base on the cluster. The final output is shown in the figure.



Between: between-cluster level

Within: within-cluster between-group level

N: the number of nonsynonymous changes

S: the number of synonymous changes

\* Input file: example\_paml\_output.txt ; example\_groupfile.csv

Output file: pMK\_output.csv