Comments to the Authors,

Dr. Wang conducted a genetic case-control association study between epigenetic modeling gene DNMT1 and Systemic lupus erythematosus (SLE) in a large Chinese population. And further biological analysis shown rs2162560 was in regulatory elements for binding by the transcriptional factors and a DNMT1 cis expression quantitative trait locus (eQTL).

**Major Compulsory Revisions**

1, The authors didn’t mention any result from GWAS result in the background? Is there any signal for DNMT1 in the previous GWAS in SLE?

2, The authors didn’t provide power analysis whether the current sample size is power enough to support the conclusions in the manuscript? Actually, all the P-value in Table 2, Table 3 and Table 4 are weak significant and might be caused by random sampling or stochastic processing. Another independent validation should be conducted to make the conclusion solid.

3, In the function analysis section, the authors could not use prediction result to make conclusion and corresponding prediction should be validated since majority of the prediction is false positive result and therefore cannot be taken as the final evidence.

4, Why only DNMT1 was considered in the current study? Acutally, DNMT gene family should be studied together right?

5, LD analysis shown the SNPs in this region are not strongly linkage disequilibrium, can you explain the reason? Which kind of sample were involved in this analysis, case or control? Did the authors check the LD with hapmap or 1000 Genome data? Suppose it is not a highly linkage disequilibrium, whether is it suitable to do the haplotype analysis?

2, In the association study, as an accurate study design the smoking, BMI between case and control population should be almost same or no significantly difference so that the genetic difference could be estimated. Why the authors didn’t control these confounders? Do you think these effects can be adjusted 100% with the statistic model? What’s worse, I cannot understand why the author show the Chi-square P-value in Table 3, rather than logistic P-value? Finally, although the number of candidate SNPs in the study were not too many, the P-value will turn not to be significant after multiple test correction.

3, It is still very difficult to understand the logic of the study. Such as 1) *YY1* is low expressed in lung cancer cell lines. 2) Expression of *DCBLD1* was not significantly different among different genotypes, even though the authors gave several hypothesis. 3) Another difficult question is that for the SNPs or genotype, the distribution was only little different between cancer and normal population and you can find the proportion of the genotype was almost very close, even the P-value was significantly, in such way, the molecular function interpretation should be carefully. I suggest that the eQTL analysis can be repeated again with large sample size to check whether the difference between different genotype were significant, if so, then everything is clear, or else, the author should be give more comprehensive and reasonable explanation.

**Minor Revisions**

1, Please provided line number so that the comments could be provided to the exactly lines.

2, Is there any possibility to validate the function of DNMT to SLE in cell lines?

3, In Table 3, the authors provided kind of association based on different genetic models, however, for the biology, there should be only one model working for the species, do you which one is right?