Methylation Section:

This is an important study in that the authors are trying extend work in a candidate gene previously associated with BMI in EWAS studies. Another strength is that they integrate genetic and epigenetic data into their study. However, I found it hard to follow the logic for why each analysis was done and some simple restructuring of the results and discussion and additional details would improve the clarity of the paper. My detailed comments are below:

Figure 1. This figure would be more informative if more details were given about context of the gene and CpG island locations. Can you add another track showing additional info or at least state some of this in the text?

It would be useful to describe the methylation data. For instance, it looks like methylation levels are on average lower in the 2nd region.

I think the CpG specific results are also valuable. Perhaps, to cut down on number of tables, average methylation can be an additional column in the CpG specific results tables. Another method to summarize methylation at multiple correlated sites is PCA. Then you just includes a few PCs in the regression model.

Table 3 doesn't seem necessary

I'm not sure it makes sense to throw all of these things into the same model. Aren't they all highly correlated? Birth weight and gestational age are probably quite correlated. Does pre-eclampsia also affect gestational age? Have you checked VIF's for evidence of multicollinearity?

Can you compare to publicly available eQTM datasets? There are some studies that have measured 450K data and gene expression and have published which CpG sites are associated with gene expression. Gutierrez-Arcelus et al., eLife, 2013

What is the potential implication of a relationships between genetic variation and cord blood methylation?

It is fine to refer to previous papers to describe the method, but a very brief sentence mentioning that methylation was determined by Sequenom MassARRAY would make sense.

I am not familiar with the methylation quantify method in this study, I am quite interesting the cost for this assay for each sample. And why the author apply this method rather than the traditional pyrosequencing or targeted BSP sequencing.

In this manuscript: “Transcriptomic profiling revealed the role of DNA methylation and histone modifications

in tumor-infiltrating myeloid-derived suppressor cell subsets in colorectal cancer”, Dr. Varun Sasidharan Nair and colleagues conducted a pilot study to the profile of different subtype of myeloid-derived suppressor cells (MDSCs) in colorectal cancer microenvironment and compared the transcriptome (RNA-seq analysis) among I-MDSCs, PMN-MDSCs and APCs in 4 samples. The authors found I-MDSCs and PMN-MDSCs were increased in tumor tissues compared with normal tissues. the authors should find genes associated with immune-suppression and myeloid cell recruitment in CRC tumor tissue are downregulated upon HDAC or IL-6 inhibition. The study was well performed and some findings are quite interesting. However, there are still lot of issues in the data analysis, power evaluation and result interpretation.

1. The sample size is the main problem in this study. First, the sample size for flow cytometric should be evaluated. What’s the power for the conclusion that the levels of I-MDSCs and PMN-MDSCs were increased in tumor tissues compared with normal tissues? Second, the result of the scRNA-seq between two cancer samples and two normal samples will be full of bias which will bring unexpected affection to the conclusion. The result should be validated in another independent dataset. Third, the pathway analysis should be corrected by multiple-test adjustment with FDR or some other approach or else too many false positive result will be obtained (Figure 2B). In the revised manuscript, the authors should be uploaded all the DAVID result as the supplementary so that the readers could check the full picture for the enrichment.
2. The authors should show the tSNE plot including all the four samples. Within tSNE plot, cancer or normal and different cell types should be labelled with clear legends.
3. The authors should compared their result with other single-cell RNA-seq paper to colorectal cancer, such as Zhang Y, Zheng L, Zhang L, Hu X, Ren X, Zhang Z. Deep single-cell RNA sequencing data of individual T cells from treatment-naive colorectal cancer patients. Sci Data 2019;6:131. Full discussion to the comparison should be conducted to interpret the difference between the current manuscript with previous findings.
4. The scRNA-seq data is not public available significantly decreased the contribution to the manuscript. I suggest the authors submit it to SRA and GEO.
5. Figure 1A, the details should be shown with supplementary tables.
6. Figure 2A and 2C: the details should be shown with supplementary tables.
7. Figure 4A and 4C: the details should be shown with supplementary tables.
8. Figure 5A, B, C, D, E, F: the details should be shown with supplementary tables.
9. Figure 2C, the value for the legend bar is RPKM, correct? The authors should provide clear explanation to the legends.
10. Abbreviations should be defined before using, for example CRC and APCs in the abstract. Please check the full manuscript to avoid such issues.
11. In the Figure 6, the authors claimed the results were from 6 patients, however, I found only 4 patients ID were showed #26, #56, #60 and #63?
12. The authors did not show how the dosage and time of the treatment were determined in “Cell Culture” section.

# SNP and Association Study:

In this manuscript, Dr. Wu and colleagues conducted a genetic association study between 8 polymorphisms located in the intron region of CMTM8 and lung cancer risk, in a small (509 case vs 506 control) Chinese Han population. In this study, the inclusion and exclusion criteria was actually not explicit. What is worse, the authors did not provide solid enough logic why the study was conducted. Although some positive association were identified, we do not know whether they are caused by population structure difference between case and control since the author do not have any control to measure the population structure. I have several major concern for the manuscript:

1. In the background section, the authors should introduce the previous GWAS study in Chinese Han population and the author should explain why a small sample size research was conducted when the GWAS study did not identify CMTM8.
2. Although the association significance is very good, I still suggest the author to validate the association study in another independent sample set to make sure the association is not false-positive since small sample size is easily to be influenced by the different population structure between case and control.
3. In the background section, the authors should provide explicit broad-sense heritability and narrow-sense heritability of lung cancer.
4. In the background section, the authors should give certain introduction to genetic background of population. The best solution is to show the current population in the 1st and 2nd components plot in PCA analysis of African, European and the population in this study.
5. Please show the association result (P-values) when you apply 1000 genome or ExAC normal individuals as the control and discuss it in the discussion section.
6. Smoking and drinking information should not only be labelled in Table 1, but also should be adjusted in other tables, especially in logistic regression.
7. Linkage disequilibrium between all the eight SNPs should be provided in the result section.
8. The power analysis should be provided to show the stability of the study.
9. I found typos were still existed in the manuscript, such as “firstiy” in the abstract section (line 42, in page 2)
10. The authors conducted multiple genetic models to evaluate the association between CMTM8 with lung cancer, however, which one is the correct one? The genetic model corresponding to significant association is correct one?

**Association of rs944289, rs965513, and rs1443434 in TITF1/TITF2 with risks of papillary thyroid carcinoma and with nodular goiter in northern Chinese Han populations**

In this manuscript, Dr. Zhang and colleagues conducted a case-control based genetic association study between three polymorphisms located in the TITF1 and TITF2 and lung cancer risk with a medium sample size (861 case, 562 nodular goiter and 896 normal) in a Chinese Han population. Thyroid cancer has been considered as one of highest familial risk carcinomas among all kinds of cancers, therefore, identification of genetic variants associated with thyroid cancer is a very important works. In this study, several positive association were identified which is quite interesting. I have several tiny suggestion for the manuscript:

1. In this study, the inclusion and exclusion criteria was actually not explicit. I suggest the authors provided explicit inclusion and exclusion criteria so that others can repeat this study with more specific guideline. One of the reason why different association cannot be repeated is caused by phenotypes difference in different studies.
2. Although the association significance is very good, I still suggest the author to conduct the association study with 1000 Genome (Chinese) and EXAC as the control to make sure the association is not false-positive since the association is easily to be influenced by the different population structure between case and control.
3. In the background section, the authors should provide explicit broad-sense heritability and narrow-sense heritability of thyroid cancer.
4. Please show the association result (P-values) when you apply 1000 genome or ExAC normal individuals as the control and discuss it in the result section and discuss it in the discussion section.
5. Smoking and drinking information should be considered in the analysis since they are also very significant risk factor for thyroid cancer.
6. Linkage disequilibrium between all the 3 SNPs should be provided in the result section.
7. The power analysis should be provided to show the stability of the study.
8. The authors conducted multiple genetic models to evaluate the association between TITF1/TITF2 with thyroid cancer, however, which one is the correct one? The genetic model corresponding to significant association is correct one?
9. The statistical analysis should be enhanced; I suggest the author refer the analysis strategy in this paper: Huang L, Li Y, Guo S, Sun Y, Zhang C, Bai Y, Li S, et al. Different hereditary contribution of the CFH gene between polypoidal choroidal vasculopathy and age-related macular degeneration in Chinese Han people. Invest Ophthalmol Vis Sci 2014;55:2534-2538.
10. As I mentioned above several important citation were ignored in this manuscript, such as:

Gudmundsson J, Thorleifsson G, Sigurdsson JK, Stefansdottir L, Jonasson JG, Gudjonsson SA, Gudbjartsson DF, et al. A genome-wide association study yields five novel thyroid cancer risk loci. Nat Commun 2017;8:14517.

Guo S, Wang YL, Li Y, Jin L, Xiong M, Ji QH, Wang J. Significant SNPs have limited prediction ability for thyroid cancer. Cancer Med 2014;3:731-735.

In current manuscript: A Possible Association between methylenetetrahydrofolate reductase (MTHFR) polymorphism and Smoking Behavior among Saudi Arabian population, Dr, Almutairi and colleagues conducted a case-control association study in a small sample size from Saudi Arabian population to investigate the possible relationship between MTHFR polymorphism rs1801131 and their interactions with smoking status.

1. In the abstract section, please use OR to describe the result.
2. Another independent validation cohort should be considered to validate current result
3. Why the author select rs1801131 is not clear in the background section?
4. In the result section: Compared with rs1801131 TT genotype carriers, carriers of TG genotype were 0.209-fold decreased in CS patients (TG vs TT: OR=0.209, 95% CI=0.117–0.373, P<0.005). I am not sure how this OR was calculated?

# miRNA Association Section

Dr. Ibrahim and colleagues conducted a pilot functional association between rs11614913 which is a genetic variant located in epigenetic factor, miR-196a2 and human type 1 diabetes in a small (N=300) Egypt population. The study has explicit inclusion and exclusion criteria and the result is interesting. What’s more, the motivation to investigate genetic variation is epigenetic factors is exciting since more and more epigenetic factor have been demonstrated to play important roles in human disease and to be promising drug targets. In this study, the authors find rs11614913 was significantly associated with T1D and miR-196a2 was significantly deregulated in T1D. Although the result is interesting, I still have some tiny suggestions:

1. In the abstract section, the authors should provide explicit population origin of the samples.
2. Although the association significance is very good, I still suggest the author to validate the association study in another independent Egypt sample set to make sure the association is not false-positive since small sample size is easily to be influenced by the different population structure between case and control.
3. In the background section, the authors should provide explicit broad-sense heritability and narrow-sense heritability of T1D.
4. I agree with the authors, genetic variant in epigenetic factors play important roles in susceptibility of human disease, I suggest the author to cite the following research to support the above concepts.

Fan L, Chen L, Ni X, Guo S, Zhou Y, Wang C, Zheng Y, et al. Genetic variant of miR-4293 rs12220909 is associated with susceptibility to non-small cell lung cancer in a Chinese Han population. PLoS One 2017;12:e0175666.

Shen F, Chen J, Guo S, Zhou Y, Zheng Y, Yang Y, Zhang J, et al. Genetic variants in miR-196a2 and miR-499 are associated with susceptibility to esophageal squamous cell carcinoma in Chinese Han population. Tumour Biol 2016;37:4777-4784.

Zhang P, Wang J, Lu T, Wang X, Zheng Y, Guo S, Yang Y, et al. miR-449b rs10061133 and miR-4293 rs12220909 polymorphisms are associated with decreased esophageal squamous cell carcinoma in a Chinese population. Tumour Biol 2015;36:8789-8795.

Ma Y, Wang R, Zhang J, Li W, Gao C, Liu J, Wang J. Identification of miR-423 and miR-499 polymorphisms on affecting the risk of hepatocellular carcinoma in a large-scale population. Genet Test Mol Biomarkers 2014;18:516-524.

1. Meanwhile, in the background section, the previous GWAS studies of T1D should also be cited such as:

Reddy MV, Wang H, Liu S, Bode B, Reed JC, Steed RD, Anderson SW, et al. Association between type 1 diabetes and GWAS SNPs in the southeast US Caucasian population. Genes Immun 2011;12:208-212.

1. All the primers/probes used in the manuscript should be provided as the supplementary or in the main manuscript.
2. The second level miRNA structure should be provided and show the location of the genetic variant as the Figure 1 in Shen F, Tumor Biology, 2016 which I have mentioned in question 4. Meanwhile, prediction model based on miRNA genotyping and RT-PCR can be built to show the AUC which is similar as Figure 2 in Shen F, Tumor Biology, 2016. And the author should discuss the AUC they observed.
3. In the background section, the authors should give certain introduction to genetic background of Egypt Population. The best solution is to show the location of Egypt population in the 1st and 2nd components plot in PCA analysis of African, European and Asian population.
4. Result section should be reformatted with compact format.
5. Please show the association result (P-values) when you apply 1000 genome or ExAC normal individuals as the control and discuss it in the discussion section.

**Predictive and prognostic value of selected microRNAs in luminal breast cancer**

In the manuscript entitled “Predictive and prognostic value of selected microRNAs in luminal breast cancer”, Amorim et al. aim to examine the prognostic and predictive capacity of miRNAs in endocrine resistant luminal breast cancer. The authors claim that "miR-30c-5p, miR-30b-5p, miR-182-5p and miR-200b-3p were found to be independent predictors of clinical benefit from endocrine therapy" and that "miR-182-5p and miR-200b-3p displayed independent prognostic value for disease recurrence in luminal BrC patients after endocrine therapy"

Strength: On the whole well written, good numbers of patients in the validation study and good to see the inclusion of normal samples for comparison

Limitations: Too few samples in the discovery cohort, unlikely that the selected miRNAs reflect endocrine resistance, unclear why specific genes were selected for individual results sections e.g. 3.4, no correction for multiple testing, genes appear to be cherry-picked from discovery cohort for further analysis, use of clinical endpoints not previously defined in a breast cancer setting (ERFS)

Reviewer comments

In the manuscript entitled “Predictive and prognostic value of selected microRNAs in luminal breast cancer”, Amorim et al. aim to examine the prognostic and predictive capacity of miRNAs in endocrine resistant luminal breast cancer. Whilst the mauscript is on the whole well written, the work suffers from a fundamental statistical power flaw at the initial miRNA discovery stage which precludes acceptance with the current sample numbers. This is discussed further below.

Major comments:

1. Three vs. four (Luminal A Rec vs. NRec) or four vs. four (Luminal B Rec vs. NRec) or seven vs. eight (Luminal Rec vs. NRec) is not enough samples to derive a robust gene list given the molecular heterogeneity that exists within Luminal A, B or all Luminal tumours. This initial discovery sample set is statistically underpowered to detect miRNA differences between endocrine resistant vs. non-resistant tumours and this unfortunately renders much of the analyses and validation afterwards inconsequential. With regards to heterogentity, large scale molecular studies including TCGA have demonstrated that Luminal A cancers have a large number of recurrently mutated genes at a >5% frequency including PIK3CA, CDH1, MAP3K1, GATA3, MAP2K4, FOXA1 and TP53 [1], and are mostly diploid but some show copy number changes including 1q gains and 16 q losses [1]. Related to this, Luminal A tumours have been found to be spread out across at least five different IntClust subgroups – a classification scheme mainly based on copy number changes [2]. Similarly, Luminal B tumours are cyclin D1 amplified in 56% of cases and FGFR amplified in 23% [1]. They show mutations in PIK3CA, GATA3, PTEN and TP53 and are also spread out into five different IntClust subgroups [2]. In short, this heterogenetity in gene mutation/amplification and mRNA expression levels simply cannot be controlled for when comparing only four vs four samples. This means that the miRNAs found to be differentially expressed in the discovery cohort between resistant and non-resistant tumours may not at all reflect endocrine resistance but instead underlying molecular differences between tumours. This reference, whilst microarry based, may provide some help in determining a more appropriate sample size for this comparison: [3]. Without a siginificant increase in numbers at the discovery stage this manuscript cannot be accepted.
2. The authors derive a list of miRNAs that specifically try to identify differences between endocrine sensitive vs. resistant tumours. These miRNAs are further tested in a second cohort where about 20% (n = 22) of the tumours are endocrine resistant. The authors see a survival difference in uni- and multi-variable analysis when comparing the lowest quartile (P ≤ 25) of the selected miRNAs to the upper three quartiles (P > 25) (Figs 4 and 5, Tables 5 and 6). The problem here is that in Figure 2 a large overlap is seen in the distribution of these miRNAs when comparing sensitive vs. resistant tumours, particularly as the values approach zero. If you select the samples with the lowest quartile expression (P ≤ 25) for e.g. miR-30c-5p, miR-30b-5p, miR-200b-3p, how many of those approx. 34 tumours are endocrine senitive and how many are resistant? By the looks of the data in Figure 2 there will be quite a bit of overlap, which once again goes back to my point 1 above – it is very unlikely that your chosen miRNAs actually reflect endocrine resistance.

Minor comments:

1. The gene selection from the discovery cohort is very odd – why not pick e.g. the top 5 or top 10 most downregulated genes? It seems like the genes were cherry-picked?
2. Numbers of patients in each group are missing in Figs 4 and 5, Tables 5 and 6
3. Multivariable analyses in Table 5 should contain all variables rather than adjusting for molecular subtype, Her2 and grade separately
4. ERFS is not a standard breast cancer clinical time to event endpoint [4, 5]
5. What was the selection criteria for the miRNAs examined in results section 3.4?
6. How exactly were the luminal A/B defined by ER, PR, HER2 and Ki67 – a better description need to be added. E.g. Luminal A = ER+, PR+, HER2-, Ki67 low, Luminal B = ?
7. There are far too many comparisons being made in uni- and multi-variable analyses, nothing appears to have been adjusted for multiple testing

References 1. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490:61–70. 2. Russnes HG, Lingjærde OC, Børresen-Dale A-L, Caldas C. Breast Cancer Molecular Stratification. Am J Pathol. 2017;187:2152–62. 3. Stretch C, Khan S, Asgarian N, Eisner R, Vaisipour S, Damaraju S, et al. Effects of Sample Size on Differential Gene Expression, Rank Order and Prediction Accuracy of a Gene Signature. PLoS ONE. 2013;8:e65380. 4. Hudis CA, Barlow WE, Costantino JP, Gray RJ, Pritchard KI, Chapman J-AW, et al. Proposal for Standardized Definitions for Efficacy End Points in Adjuvant Breast Cancer Trials: The STEEP System. J Clin Oncol. 2007;25:2127–32. 5. Gourgou-Bourgade S, Cameron D, Poortmans P, Asselain B, Azria D, Cardoso F, et al. Guidelines for time-to-event end point definitions in breast cancer trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials)†. Ann Oncol. 2015;26:873–9.

Dr. Jiang and colleagues conducted an association and function validation study for the SNPs located in the 3’UTR region of lncRNA HOTAIR in a large pancreatic cancer cohort from Chinese population.

Major:

1. In the background section, the authors mentioned overexpression of HOTAIR is associated with the metastasis and poor prognosis, Is there any previous evidence to show the roles of HOTAIR in the diagnosis of pancreatic cancer. If not, how to understand the hypothesis of that high expression of HOTAIR caused by the inhabited binding of miRNA-29a increase the risk to pancreatic cancer. If over-expression HOTAIR could be a diagnostic biomarker for pancreatic cancer, any quantitative evaluation for the diagnosis performance. These information should be mentioned in the background section.
2. In the method section, the author should give explicit description rather than ‘Data was analyzed by MassARRAY Typer software 4.0.3’. What kind of analysis did MassARRAY Typer conduct?
3. Is there any previous evidence to show any of the SNPs in present study were significantly associated with diabetes (type I or II)
4. It would be helpful if the authors should provide the diagram to show the relative location about the four SNPs and HOTAIR in the Figure 1.
5. In the section of “Association of SNP rs200349340 Genotypes with miR-29a binding”, how the author could make the conclusion without confirming the binding alleles (in HOTAIR) for the sw1990 cell line? What would happen if sw1990 have C allele in HOTAIR?
6. In the figure 1 and 2, the significance star should be labelled and with P-value.
7. The authors mentioned Li et .,2018 investigated common variants in HOTAIR,
8. The authors should show the coefficient and P-values for confounders (can be as the supplementary materials)
9. Please add the line number in the manuscript in the further manuscript.
10. I want to know whether this region was nearby any pancreatic cancer GWAS significant regions.
11. Linkage disequilibrium among these four SNPs should be mentioned in the manuscript in Chinese population.
12. In the background section, the authors should provide more details how to select the candidate SNPs. Only 4 miRNA-SNP were found or the authors add more filters to excluded certain miRNA-SNP
13. The authors mentioned batch effect of the samples collected from different hospital, why not take it as one of the confounders in the regression models.
14. How many sample size should be mentioned for Figure 2 and boxplot with scatters, which represents each sample, should be provided. Meanwhile, is there any possibility to show the OR based on gene expression for PC?
15. It would be better to remove 3 non-significant SNPs (from Table 2) to the supplementary and merge Table 2 and Table 3 together to show the conclusion clearly.

# Meta-analysis to Genetic association Study Section

In this systemic review, the authors provided a brief summary of the association between CAG trinucleotides of HTT (Huntington Disease Protein) and Huntington's disease (HD) in different populations. The authors found populations of high HD prevalence have higher frequencies of A1or A2 haplotypes than populations showing low HD prevalence, even with similar average. The review is quite helpful to understand the importance of haplotype research to HTT genes in HD and emphasized the importance of long-region sequencing technique in HD precision medicine. I have several comments to the manuscript:

1. The author should provide exact reference for “Furthermore, several genetic modifiers distinct from the Huntington’s disease locus itself have been identified as being linked to the clinical expression and progression of HD”
2. It will be helpful to provide more basic introduction to A1, A2, B, C, A4, A5, IA haplotypes. They are not common knowledge even for geneticists if they are not familiar with Huntington’s disease.

1. For the haplotype A1/A2 results, the method how to infer the haplotype should be mentioned in the table 1. Is there any difference on the methods used for the haplotype inference in different studies? What kinds of influence will be caused by different methods
2. Prevalence of HD in different countries also significantly affected by the emphasize to the disease, for example, Asian countries usually have decreased emphasize to mental diseases compared with European country. It should be mentioned in the discussion section.
3. I suggest the author to discuss the importance to apply long-fragment sequencing in HD genetic etiology research since which can help to generate accurate haplotype information.
4. Some reference in the background section are out of data, there are lots of recent important research should be mentioned in the background section, for example:

Alexandrov, V., et al., *Large-scale phenome analysis defines a behavioral signature for Huntington's disease genotype in mice.* Nat Biotechnol, 2016. **34**(8): p. 838-44.

Genetic Modifiers of Huntington's Disease, C., *Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease.* Cell, 2015. **162**(3): p. 516-26.

Moss, D.J.H., et al., *Identification of genetic variants associated with Huntington's disease progression: a genome-wide association study.* Lancet Neurol, 2017. **16**(9): p. 701-711.

Ruderfer, D.M. and J.T. Dudley, *Deep phenotyping predicts Huntington's genotype.* Nat Biotechnol, 2016. **34**(8): p. 823-4.

# Review: Aberrant DNA Methylation in Cutaneous Squamous Cell Carcinoma

In this review, the authors provided a summary to several important abnormal epigenetic genes in cutaneous squamous cell carcinoma (CSCC) and discussed the potential diagnostic, prognostic role of these methylation genes and miRNA in CSCC. The review is interesting since current the epigenetic research in CSCC is not very common caused by limited genome-wide DNA methylation data in CSCC. What’s worse, the most important cancer genome project: TCGA project didn’t include CSCC samples which make CSCC methylation research worse. Therefore, this review will help the researcher to have better understanding to CSCC methylation research. However, the review still need more revision before publication:

1. Line 12-15, the incidence and prevalence of CSCC should be introduced with specific region or worldwide. It should be quite explicit.
2. In line 1, “…hypothesized that CpG methylation is associated with closed chromatin configurations” clear reference should be provided.
3. In line 7,” tumor suppressor genes (TSGs) in various types of tumors [5-7]”, the references are not well demonstrated. The recommended way is to list most recent research for most common cancers, such as ESCC, NSCLC, HCC, breast cancer. For example: Wang C, Pu W, Zhao D, Zhou Y, Lu T, Chen S, He Z, et al. Identification of Hyper-Methylated Tumor Suppressor Genes-Based Diagnostic Panel for Esophageal Squamous Cell Carcinoma (ESCC) in a Chinese Han Population. Front Genet 2018;9:356
4. Between line 50 and line 59, LINE-1 methylation research in CSCC is interesting. However, the author would better to compare the methylation level of LINE-1 in cancer (41.09%) and normal (46.05%) with other cancers and normal and try to find whether it is CSCC specific. According to my experience, LINE-1 methylation usually is 60%-80% in normal samples.
5. The authors omitted numbers of methylation research in CSCC, such as FOXE1, MGMT, MTHFR

Tuominen R, Jewell R, van den Oord JJ, Wolter P, Stierner U, Lindholm C, Hertzman Johansson C, et al. MGMT promoter methylation is associated with temozolomide response and prolonged progression-free survival in disseminated cutaneous melanoma. Int J Cancer 2015;136:2844-2853.

Venza I, Visalli M, Tripodo B, De Grazia G, Loddo S, Teti D, Venza M. FOXE1 is a target for aberrant methylation in cutaneous squamous cell carcinoma. Br J Dermatol 2010;162:1093-1097.

Laing ME, Cummins R, O'Grady A, O'Kelly P, Kay EW, Murphy GM. Aberrant DNA methylation associated with MTHFR C677T genetic polymorphism in cutaneous squamous cell carcinoma in renal transplant patients. Br J Dermatol 2010;163:345-352.

1. The authors should mentioned the most recent cfDNA methylation research and tissue-of-origin mapping and non-invasive methylation research in ESCC.
2. The author didn’t provide enough research progress on 5-Aza-CdR and Zebularine in CSCC.
3. The author didn’t explain difference of CSCC methylation research compared with other cancers. Moreover, the author should discuss the challenge and opportunity of CSCC methylation research.

In this manuscript, Dr. Bilinovich conducted hypothesis test to validate the truth autism GWAS signal located in 20p12.1. The authors want to demonstrate the real GWAS risk is caused by lncRNA MSNP1AS rather than previous reported protein-coding gene MACROD2. The authors conducted RNA-seq analysis to show expression level of lncRNA MSNP1AS is corrected with phenotypes (autism and normal), environment (diesel particulate matter exposure) and genotypes of GWAS tag SNP (rs4141463). The authors provided very promising preliminary data to show the interesting relationship between lncRNA MSNP1AS and risk of autism.

How many samples were collected in the RNA-seq assay should be declared in the method section.

It looks all the RNA-seq analysis differential gene analysis and Gene ontology enrichment analysis don’t have multiple-test correction?

In the method section, p<0.05 was taken as significant differential genes. However, in the discussion section, q<0.05 was mentioned. It should be declared clearly.

Data availability statement should be check again. Generated Statement: This manuscript contains previously unpublished data. The name of the repository and accession number(s) are not available. I don’t understand why repository and accession number(s) are not available since the data have been published before. Meanwhile, it will be better to make RNA-seq data generated by the current manuscript available.

It will be better to show as “GTEx consortium” in line 146 and it will be better to check p-value between rs4141463 and MSNP1AS in GTEx database and discuss the consistent between your result with GTEx result.

# Histone modification

Pediatric glioma have decreased BDNF expression and increase histone methylation at the BDNF promoter

In this manuscript, Dr. Thai and colleagues conducted a comprehensive study to investigate the relationship between BDNF and pediatric glioma. several important genes, such as EGFR, IDH1, PDGFRA, PTEN, BDNF, SLC7A11, SXC, were investigated between pediatric and adult glioma. Furthermore, the epigenetic regulation of BDNF was investigated including DNA methylation and histone modification (H3ac and H3K27me3). The authors demonstrated DNA methylation of BDNF was similar between pediatric and adult glioma while H3K27me3 might be the reason to BDNF divergent expression regulation.

# DNA methylation, Virus and human cancer

In this manuscript, Dr. Zheng and colleagues conducted a study to investigate the DNA methylation of Epstein - Barr virus (EBV) DNA in the brushing samples from the patients of nasopharyngeal carcinoma. The authors found significantly increased virus DNA load and hyper-methylation of virus DNA methylation in cancer samples compared with control samples. Although the findings are quite interesting, the limited sample size and the un-solid technique, such as MSP, are main disadvantages for the manuscript. In addition, the manuscript should be modified to make the structure more concise.

1. The background should be significantly shorted. Too many non-related contents are involved. The authors only need to explain the most related 1) background 2) the knowledge gap and 3) the logic of why the authors need to conduct this study with 2-3 paragraphs.
2. The flow-chart of the research was recommended to provided so that the reader can easily understand the whole story as soon as possible, such as sample types, numbers, inclusion and exclusion criterions, experiments procedures.
3. In the discussion section, the authors should have little discussion about why brushing samples were selected in this study since, recently, the tissue-of-origin prediction ability of blood based cell-free DNA methylation have been well confirmed in several studies, such as: 1) Kang S, Li Q, Chen Q, Zhou Y, Park S, Lee G, Grimes B, et al. CancerLocator: non-invasive cancer diagnosis and tissue-of-origin prediction using methylation profiles of cell-free DNA. Genome Biol 2017;18:53. 2) Guo S, Diep D, Plongthongkum N, Fung HL, Zhang K, Zhang K. Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. Nat Genet 2017;49:635-642.
4. The authors are recommended to prepare a Figure to show ROCs for different diagnosis models, such as 1) only EBV load 2) only methylation 3) EBV+methylation
5. The authors should apply golden method, such as BSP to validate the result of MSP to make sure the result are not false positive.

# Diagnostic biomarker + circ-RNA-rejected

In this manuscript, Dr. Zheng and colleagues conducted a study to investigate the DNA methylation of Epstein - Barr virus (EBV) DNA in the brushing samples from the patients of nasopharyngeal carcinoma. The authors found significantly increased virus DNA load and hyper-methylation of virus DNA methylation in cancer samples compared with control samples. Although the findings are quite interesting, the limited sample size and the un-solid technique, such as MSP, are main disadvantages for the manuscript. In addition, the manuscript should be modified to make the structure more concise.

# DNA methylation and pregnancy

In this manuscript, Dr. Robakis and colleagues conducted an epigenetic research to investigate DNA methylation changes during pregnancy and the association between DNA methylation patterns with psychiatric distress and pregnancy complications. The authors identified DNA methylation in CD3 gene strongly associated with anxiety and depression in mid- and late pregnancy, in addition, weakly associated with the presence of complications of pregnancy. The study was performed rigorously and the findings are quite interesting. However, I have several concerns:

**Major Compulsory Revisions**

1. In the method section, the authors mentioned “for the remainder, mean T cell fraction was 93% (range 66-100%)”. How to understand the T-cell fraction could as low as 66% for certain samples considering that the samples were selected by CD3+ criterion?
2. I recommend the author provided a flow-chart or schematic-diagram to show the whole study design and the sample, data relationship since current the samples are under multiple situations that reader will be confuse about the status for these samples. The aim is to let the reader to know which samples are applied in which analysis.
3. Data availability should be noted for DNA methylation data and should be uploaded to GEO.
4. The author mentioned the genome-wide DNA methylation analysis and didn’t identify significant signals between all during-pregnancy to all post-pregnancy samples. Did the authors try to use 5% methylation change or 1% methylation change? 10% methylation actually is a very high threshold. I recommend the authors to prepare the Manhattan plot and qq-plots for the genome-wide DNA methylation assay and provided the most significant different CpGs even the P-value may not pass the multiple test correction. Meanwhile, the authors mentioned in the background section, they want to focus on several T-cell related genes. My question is why not make such assumption? What kinds of evidence make the authors to remove non-T-cell related genes since actually they have genome-wide DNA methylation dataset?

Minor Revisions:

1. How to understand the exclusion criterions in this study: “Women were excluded if they acknowledged smoking or use of recreational drugs, lacked fluency in English or were multiparous or taking medications”. Especially for the language requirement, as we known, it is easy to solve the problem caused by the language.
2. The authors used a term named: “hotspots” should be explained in the method section since it concept is not common used in the methylation data analysis. How to define hotspots, how to convince authors the threshold is correct? Since there are some other hypothesis: DNA methylation in TFBS play more important roles in transcriptional regulation.
3. The authors used a term named “mean DNA methylation density”. Is this to represent “average methylation level”? “methylation density” is not common used in methylation analysis.
4. I recommend the authors to have a brief summary about the changes between gestation and postpartum for all these variables.

In this manuscript, Dr. Kang and colleagues conducted a brief computational analysis to investigate the relationship between SLE gene family and the prognosis of clear cell renal cell carcinoma. The author collected data from several public database including GEPIA, TCGA, cBioPortal, HPA.

**Major Compulsory Revisions**

1. The logic of the study is not very clear. Why only focus on SLC gene family not the whole genome since all the genes are available in these databases?
2. In the method section, the P-value cut-off was set as 0.01, which does not make sense since these P-values are genome-wide based P-value, which should be adjusted by multiple test correction.
3. The authors applied multiple databases to increase the credibility, however, these database enrolled repeated data such as TCGA, however to remove these over-weighted effects should be considered.
4. The authors only collected the result from different website without any innovation on method or algorithm which cannot be published, I think. The corresponding findings must be validated with well-applied method in the clinical application.

In this manuscript, Dr. Yu and colleagues conducted a two-stage association study between genotypes of YWHAZ and biochemical recurrence of prostate cancer in Taiwan population. The author genotyped 11 SNPs located in 11 genes identified from KEGG cell cycle progression item and finally the author identified rs2290291 was associated with biochemical recurrence of prostate cancer. The study was performed rigorously and the findings are quite interesting. However, I have several tiny concerns:

**Major Compulsory Revisions**

1. The authors collected a very valuable biochemical recurrence cohort. I am wondering why the authors do not conducted a GWAS study, but a case-control study to investigate several candidate genes, in which situation; the most interesting signals might be ignored.
2. As Table 1 shown, the P-values were ranged from 0.008 to 0.038. If multiple-test correction was considered these P-values might not be significant. The authors should be careful about it. I suggest the author to conduct a power analysis and false-positive analysis or permutation analysis to check the robust of the conclusion.
3. Figure 2 were derived from public data without any innovation. I suggest the authors to move them to supplementary figures.
4. I suggest the author to validate the eqtl of rs2290291 with in vitro experiments so that the conclusion will be more solid.

Minor Revisions:

1. Please check Figure 2D, the p-value was as significant as Figure 2F, however, the curves looks quite different, please double-check it.
2. The authors should provide explicit cut-off to low and high expression for Figure 2D to Figure 2F or else others cannot repeat the result.

# DNA methylation and cord blood storage

In this manuscript, Dr. Sasaki and colleagues conducted an epigenetic research to investigate Impact of ex-vivo sample handling on DNA methylation profiles in human cord blood and neonatal dried blood spots with high-coverage RRBS technique. The authors found that DNA methylation profiles are resilient to ex-vivo sample handling conditions prior to storage. The study was performed rigorously and the findings are quite interesting and provided fundamental advice that the sample preparation for umbilical cord blood is quite stable for DNA methylation profiling under different approach. Meanwhile, the authors identified several genes whose methylation and expression have consistent change as the storage. I only have several minor concerns for the study:

Major:

1. The sample size is one of most important factors to affect the number of differential methylation loci/regions in the statistical procedure. When the author didn’t identify significant DMCs, one reason might be there is no big difference on DNA methylation while another reason might be the sample size is not large enough to identify tiny methylation changes. It should be discussed in the discussion section.
2. The function of DMCs significantly depend on the genomic location of CpGs, such as CpG Island, promoter region, gene body region. In the table 1, the location of DMRs should be shown. Meanwhile, the meth diff % could be shown as average differential methylation ratio (beta) rather than a range.
3. Please show the qq-plot for all the test as the main figure to show the distribution of the P-values. Meanwhile, Manhattan plots were also recommendation to be showed as the main figure.
4. The authors only conducted analysis to 30x coverage regions; however, it will be quite interesting if the author could conducted similar analysis to 5x, 10x, 20x, 30x to show the stability of the DMC analysis.
5. A RRBS based cluster analysis should be conducted to show the relationship between all these assays.

Minor:

1. The exact sample size should be shown in Figure 1a, 1b, 1c.
2. The authors should explain clearly what’s the aim of complete blood count (CBC) test.
3. The abbreviations should be mentioned before usage, such as nRBC in line 122, page 3.
4. In line 172, page 4, the authors mentioned 5% was taken as threshold since it is the detection limits of DNA methylation pyrosequencing. However, the author did not provide a reference to support this claim.
5. When the author conducted KEGG analysis, I am not sure whether only promoter/enhancer DMCs were applied or DMCs in gene body regions were also applied since the function for these two different regions are quite different.
6. Python script should be provided in 100 line, page 3.
7. Any figure to show the 200, 265 and 330bp peaks in the study?
8. I need the raw fastq or processing beta matrix to have more checking before the publication, therefore, I suggest the author to upload the data to GEO and SRA.

# Genome-wide DNA methylation profiles

In this manuscript: Extensive inter-cyst DNA methylation variation in autosomal dominant polycystic kidney disease revealed by genome scale sequencing, Dr. Bowden and colleagues applied RRBS to generate genome-wide DNA methylation patterns for autosomal dominant polycystic kidney disease (ADPKD) and to infer the intra-variation of DNA methylation within one individual to get rid of genetic variation noise. The study was performed rigorously and the findings are quite interesting and provided useful data for the further research. I only have several concerns for the study:

* 1. In the background section, the basic information about ADPKD is not clear, such as prevalence, heritability
  2. In the 2.1 section, line 69, the detail mutation information should be provided.
  3. It will be better to show a schematic diagram to show the sample collection for section 2.1
  4. How to define or calculate the significance of ICV should be explained again. I can not figure out how to calculate it. It will be better if the author could provide the R or matlab function in specific github page
  5. Line 150, how to define background gene should be explained, overlap with gene body? Promoter? Same problem to n=2,415 and n=4,204. Another solution is to provide the gene list as a supplementary table.
  6. In the line 116, the author should have more explain what exact this non-ADPKD is? it is normal tissue or what?
  7. In the line 166, the author should give explicit information about the data of non-ADPKD and whole tissue, they are RRBS or WGBS
  8. Figure 1A, there is not cluster analysis section in method section. Meanwhile, what will happened if the authors add kidney methylation data from Roadmap project?
  9. The specific genomic region for Figure 1C should be labelled in the Figure.
  10. In the 3.1 section, how to understand the methylation level between among 3 categories? How to explain these difference? Is there any other previous evidence support this result?
  11. In the line 202, how to understand the ICV enrichment to chr4 and chr17?
  12. When the authors calculate ICV, whether sequencing depth was considered in the formula?
  13. Figure 4B, the authors can show the distribution for 3 categories together including 5890, 837 and 1367 regions.
  14. Is there any methylation difference for FBXW10 and SMPD3? I ask this question since there is a report mentioned these two genes are DMRs in kidney cancer. [Abnormal methylation status of FBXW10 and SMPD3, and associations with clinical characteristics in clear cell renal cell carcinoma](javascript:void(0)), J Wang, J Li, J Gu, J Yu, S Guo, Y Zhu, D Ye. Oncology letters 10 (5), 3073-3080
  15. Please show the qq-plot for all the test as the main or supplementary figure to show the distribution of the P-values. Meanwhile, Manhattan plots were also recommendation to be showed as the main or supplementary figure.

# Chemotherapy and mechanism

In this manuscript, the authors conducted an investigation to characterize the synergetic effect of combining ERL with vorinostat (SAHA) on the progression in the A549 human lung cancer cell line. It is a very importance study, however, I have several concerns before the further evaluation:

1. The background section is not well prepared. The authors should provide more information about current status/usage of ERL and SAHA. For example, how often these two drug were used? How often or what’s the prevalence of the drug resistance will happen to ERL and SAHA? What’s the mechanism of the resistance to ERL and SAHA? Whether combination therapy of ERL and SAHA has been used as a regimen? If yes, how often it is used? These information will help reader to understand the ERL and SAHA better and to understand the importance of the study. Although the current manuscript showed clear aim of the study, the authors didn’t showed the importance and the urgency of the study.
2. In the Apoptosis assay method section, what’s the control cell is not clear, please show more details about the control cells.
3. In the Drug preparation and application section, how to determine the cell concentration should be explained? Why the authors set 10, 50, 100uM? Why 5uM and 200 uM were not set? I suggest the authors to add more concentration from very low concertation to very high concentration to show readers why 10, 50, 100uM is an optimized study design.
4. Figure 1 should be re-prepared and the detail numbers should be provided as a supplementary table.
5. In the Figure 1B, how to understand WISH have weaken necrosis for 100E+S compared with 50E+S should be explained. Meanwhile, same question to 100E and 50E.
6. In the Figure 1A and 1B, what’s “C” indicates should be noted.
7. In the Figure 1B, why 100E+S don’t showed higher necrosis should be explained.
8. The layout of Figure 2 (upper panel) should be better prepared so that the reader can compare the different situation easily. Figure 2 bottom panel should be a little larger. How to understand G0-G1 is dominant cells in A549E while pre-G1 is the dominant cells in A549-S? How to understand as the E+S concentration increased, cells in S stage is increasing?
9. RR-PCR should be replaced with RNA-seq since selection to detected several genes is bias to show the result.

# Protein and cancer prognosis

In this manuscript, Dr. Gao and colleagues conducted an association study between WEE1 protein level and liver cancer prognosis. The authors applied immunohistochemically detected the protein level of WEE1 in 325 liver cancer and non-liver cancer samples and found there is no significant association between WEE1 protein level and survival. As a negative result, the authors did not explained the heterogeneity, statistic power and other potential reasons that might cause false negative result and certain statistical analysis are mistakenly applied in the study.

1. Since the author have provide clear functional background of WEE1, therefore, all the analysis should use single-side test rather than two-side test.
2. In the Table 1, how to define high expression and low expression should be explained in the footnote.
3. In the COX regression model, the authors did not show what kinds of confounders have been included in the model.
4. I recommend the authors apply TCGA data together with UALCAN to increase the quality of the manuscript.
5. Figure 1, 2, 3 can be merged into one single Figure to show all the details of the study rather than separating them to three different figures.
6. In the Table 1, the exact P-value should be used rather than “<0.001”
7. The author should check whether the association between WEE1 protein and overall survival showed significant association in certain subgroup.
8. In the discussion section, the authors should discuss why WEE1 was down-regulated in HCC

# Cancer Cell Biology

Dr. Zhu and colleagues conducted a classic molecular and cellular analysis to investigate the roles of EIF3B in pancreatic cancer with multiple approaches including RNA-seq after knocking download, Xenograft tumor mice model, MTT, apoptosis analysis and Western blotting as well immnunohistochemical analysis. the authors identified EIF3B might regulate CDH1, IRS1, DDIT3, PTEN and CDKN1B genes. Frankly speaking, I totally trust the conclusion to this study, however, all these conclusion can be inferred by previous studies and database including eIF3 targets cell proliferation mRNAs for translational activation or repression (2015) Nature.

strengths: None  
limitations:  
  
1. all the Aims for this study have been investigated before, the authors just simple repeated previous findings.  
2. the authors didn't realize they should investigate EIF3 genes rather than EIF3B which is only a subunit  
3. the authors should avoid to repeat known knowledge but to investigate unknown part for the roles of EIF3 in human cancers.

1. the authors knocked down EIF3B and then conducted a RNA-seq analysis which is a novel study and provided very interesting part to the study, however, the authors didn't upload the RNA-seq data to SRA database and then lost all the interesting part of the study.  
  
2. I think the authors could re-write the manuscript and put knocking-down RNA-seq as the main part to be a brief report which will be an interesting paper.  
  
3. another problem is EIF3 paper I mentioned above was absent in the reference list indicating the authors conducted a weak study design.

For IHC, please show adjacent normal and tumor tissue in one image and the boundary between normal and cancer tissue

EIF3B is an essential gene for translation. How would authos convince that this gene is benefit tumor cells specifically, but not a general survival gene even for normal cells?

For EIF3B upregulation, please also show paired normal and tumor (at least three paired) expression by western blot and qPCR

Mechanistically, pleasae show morphology of shControl and shEIF3B (bright field images), western blot of cleaved caspase 3. cell cycle data and apoptosis data are not immpressive. Please not show peak, can authors show dot per cell FACS data? Please double stain Annexin V and PI for apoptosis FACS.

# Methylation meta-analysis

In this manuscript, Dr. Peng and colleagues provided a meta-analysis to evaluate the diagnosis performance of RASSF1A hypermethylation and HBV/HCV induced hepatocarcinoma. DNA methylation have been demonstrated to be the best biomarkers for cancer diagnosis, especially the cell-free DNA methylation. RASSF1A definitely is one of most interesting hyper-methylation markers for cancer diagnosis. It is very important work to quantitative evaluate the performance of RASSF1 in liver cancer. However, I have several major concerns for the manuscript:

Meta-analysis to the RASSF1 and liver cancer have been conducted in previous study: Oncotarget. 2016 Dec 6; 7(49): 81255–81267. What is the difference between the current manuscript and the above one I mentioned?

Diagnosis based meta-analysis is quite different than other meta-analysis, I recommended the authors to check the following study and conducted similar heterogeneity analysis in this study. Guo S, Tan L, Pu W, Wu J, Xu K, Wu J, Li Q, et al. Quantitative assessment of the diagnostic role of APC promoter methylation in non-small cell lung cancer. Clin Epigenetics 2014;6:5.

The methylation detection techniques, primer location and tissue-types (solid tissue or plasma) are significantly affect the study result and these factors should be discussed in the study.

In this manuscript, Dr. Khan and colleagues conducted a bioinformatics analysis to POLG and identified two novel coding sequences/ORF and interesting CUG codon mechanism in POLG which is thought to be uniquely responsible for the replication and repair of mtDNA. The study and the analysis were well prepared and conducted and the finding is very interesting and important for clinical and basic research. The only tiny cons is that the manuscript was not well prepared which could be more concise and clear.

1. The information in the Table 1 should be more clear with genomic version, position, REF and ALT
2. Please provided a detail github project to show the all steps and readme documents for this study to let the reader could repeat the study easily.

In this manuscript: “The utility of new biomarker-based predictive model for clinical outcomes among

ST-elevation myocardial infarction patients”, Dr. Olga V. Petyunina and colleagues conducted a biomarker and prognosis model investigation to predict the outcome of ST-elevation myocardial infarction patient. It is a well-designed study with power estimation and comparison between different models. I enjoy the study, however, there are several tiny concern which will help the manuscript more smoothly.

1, in the abstract the gene symbol and rs id should be provided endothelial No synthase gene and angiotensin receptor-1. In addition, there is an error in the abstract C786C should be T786C.

2, Abbreviations should be defined before using, for example STEMI in the abstract. Please check the full manuscript to avoid such issues.

3, I cannot review table 4,5,6,7 since the right margin are not large enough to show the remaining tables.

4, in the table 4, multiple different genetic model should be applied in the statistical analysis including additive model, recessive model, and over-dominant model. In additional, please show the reference (present or absent ) with an explicit style or please mention which one was recoding to reference when you conduct the association study and prediction model (Table 4).

5, I notice the authors applied SPSS as the analysis tools, It will be more interesting if the author could provide R function or script so that other can easily use his model to their own dataset to check the independent effect or performance of the model. It will great helpful to increase the citation for this manuscript.

1. The sample size is the main problem in this study. First, the sample size for flow cytometric should be evaluated. What’s the power for the conclusion that the levels of I-MDSCs and PMN-MDSCs were increased in tumor tissues compared with normal tissues? Second, the result of the scRNA-seq between two cancer samples and two normal samples will be full of bias which will bring unexpected affection to the conclusion. The result should be validated in another independent dataset. Third, the pathway analysis should be corrected by multiple-test adjustment with FDR or some other approach or else too many false positive result will be obtained (Figure 2B). In the revised manuscript, the authors should be uploaded all the DAVID result as the supplementary so that the readers could check the full picture for the enrichment.
2. The authors should show the tSNE plot including all the four samples. Within tSNE plot, cancer or normal and different cell types should be labelled with clear legends.
3. The authors should compared their result with other single-cell RNA-seq paper to colorectal cancer, such as Zhang Y, Zheng L, Zhang L, Hu X, Ren X, Zhang Z. Deep single-cell RNA sequencing data of individual T cells from treatment-naive colorectal cancer patients. Sci Data 2019;6:131. Full discussion to the comparison should be conducted to interpret the difference between the current manuscript with previous findings.
4. The scRNA-seq data is not public available significantly decreased the contribution to the manuscript. I suggest the authors submit it to SRA and GEO.
5. Figure 1A, the details should be shown with supplementary tables.
6. Figure 2A and 2C: the details should be shown with supplementary tables.
7. Figure 4A and 4C: the details should be shown with supplementary tables.
8. Figure 5A, B, C, D, E, F: the details should be shown with supplementary tables.
9. Figure 2C, the value for the legend bar is RPKM, correct? The authors should provide clear explanation to the legends.
10. In the Figure 6, the authors claimed the results were from 6 patients, however, I found only 4 patients ID were showed #26, #56, #60 and #63?
11. The authors did not show how the dosage and time of the treatment were determined in “Cell Culture” section.

In this manuscript: Modulation of calcium homeostasis may be associated with susceptibility to renal cell carcinoma in diabetic nephropathy rats, Dr. Luo and colleagues apply in vivo rat model to investigate whether diabetic nephrology (DN) increases the risk for the development of renal cell cancer (RCC) and underlying biological pathways. The study is quite interesting, however, there are several major concerns, and especially the data analysis is severely not professional and the statement are not truth for example the PRJNA597461 is not existed at all.

1. The sample size for the mouse model should be discussed. How the sample size were determined? The significant imbalance of the case and control samples were big concerns since it is fundamental for this study and further biological research.
2. In the DATA AVAILABILITY section, PRJNA597461 actually are not existed at all.
3. Figure 2, should be showed as all the images and quantitative dot or bar plot.
4. Figure 3, should be showed as all the images and quantitative dot or bar plot.
5. Figure 4, should be replace with dot plot to show each sample rather than bar plot
6. Figure 5B should be replaced with P-value rather than gene numbers.

In this manuscript, SLC39A4 as a novel prognosis marker promotes tumor progression in esophageal squamous cell carcinoma, Dr. Xia and colleagues conducted a functional and clinical investigation to the roles of SLC39A4 in esophageal squamous cell carcinoma. The authors validated the over-expression of SLC39A4 in a ESCC cohort from Chinese Population. Meanwhile, the author found over-expression of SLC39A4 was positively correlated with clinical stage, T stage and lymph node metastasis. Finally, the functional studies showed SLC39A4 knockdown not only impaired the proliferation and motility capacities of ESCC cells but also enhanced the sensitivity to cisplatin treatment.

Pros:

1, Authors utilized large number of public data to support his hypothesis.

2, multiple phenotypes such as EMT, apoptosis, chemotherapy resistance were measured with SLC39A4 knocking-down.

Cons:

1, Figure legends should be provide more details, for example, which dataset is used in Figure 1A. The authors should make sure readers could figure out majority information in the legends to understand the Figures.

2, Only knocking-down experiments were conducted without over-expression assay which cannot make the solid conclusion that SLC39A4 is an oncogene.

3, How many EMT markers totally? Why the authors only selected these 5 genes? How to avoid the bias caused by the candidate gene selection? Here, RNA-seq data should be applied to make the conclusion.

4, Please provide the regression lines between SLC39A4 expression and clinical stage, T stage as well as lymph node metastasis.

5, Multiple drugs should be tried to check the consistence, meanwhile, over-expression of SLC39A4 should be conducted to check the drug response.

6, Why SLC39A4 is significantly over-expressed in ESCC should be investigated.

Minor:

1, Furthermore, the in vitro experiments showed that SLC39A4 knockdown impaired the proliferation and motility capacities of ESCC cells but also enhanced the sensitivity to cisplatin treatment. I guess the authors lose a “also” after the “but”.

2, Figure 4B, siNC and siSLC39A4 should be provided more details. Please check all the legends and be sure to provided enough information.

# Rheumatoid arthritis

In this manuscript, Dr. Tsai and colleagues hypothesized that exposure to PMs would alter miRNA expression in the FLS, leading to exacerbation of acute joint inflammation and then the authors conducted serials of experiments to determine whether ambient PMs is associated with the expression of specific miRNA, and to investigate the participation of such miRNA in the inflammatory pathways related to RA.

Pros:

Cons:

1. Figure 1A, there are so many cytokines, why the author select IL1, IL6, IL8, IL10, IL13, COX2 and TNF?
2. Figure 1B, did the author tried 100 and 200 ug/cm2