

Misuse of reporter score in microbial enrichment analysis

Lei Liu¹, Ruixin Zhu¹, Dingfeng Wu^{2, *}

¹ The Shanghai Tenth People's Hospital, School of Life Sciences and Technology, Tongji University, Shanghai 200072, P. R. China.

² National Clinical Research Center for Child Health, the Children's Hospital, Zhejiang University School of Medicine, Hangzhou 310058, Zhejiang, P.R. China.

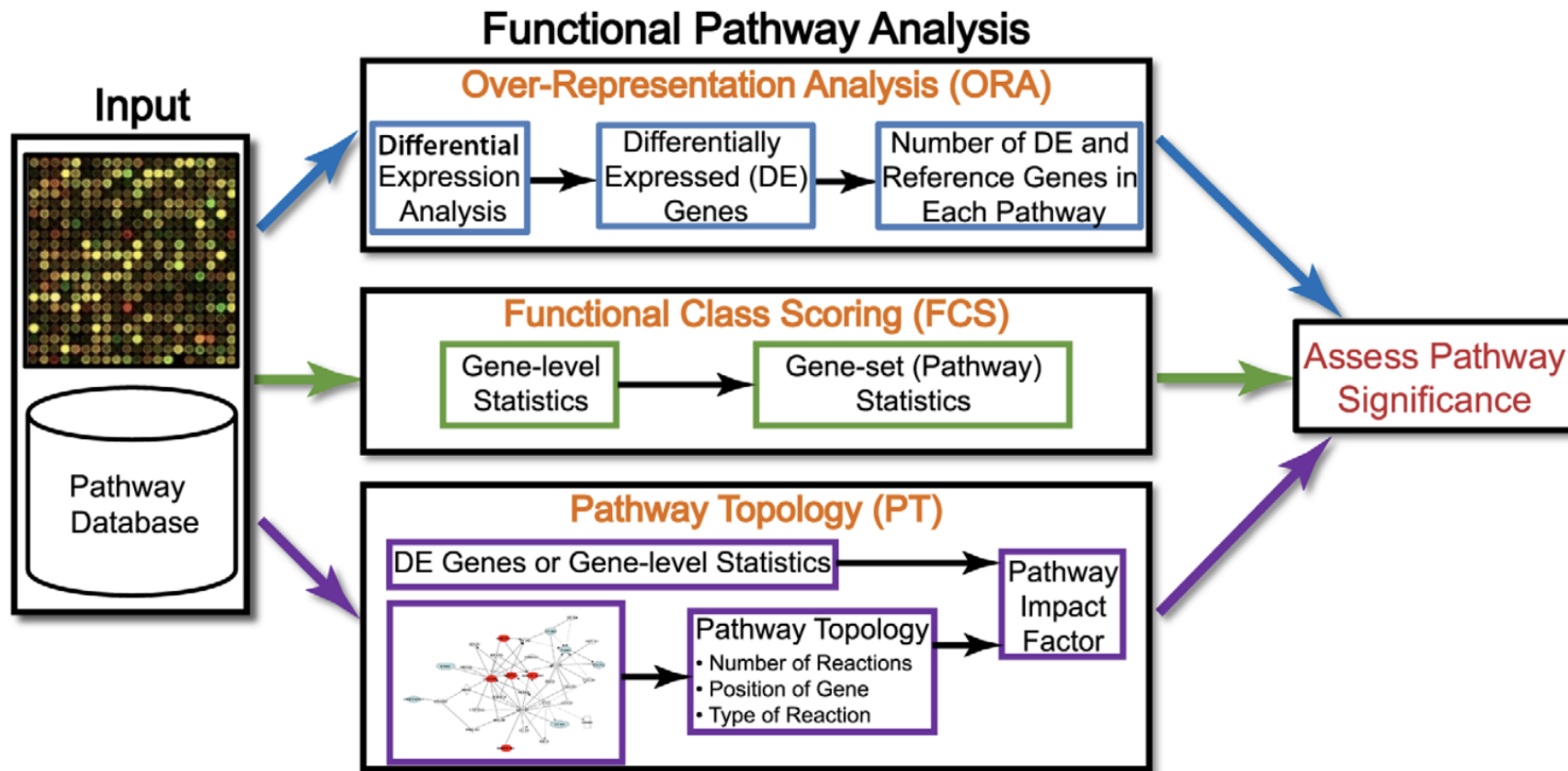


Liu, Lei, Ruixin Zhu, and Dingfeng Wu. 2023. “Misuse of reporter score in microbial enrichment analysis.” *iMeta*. e95. <https://doi.org/10.1002/imt2.95>

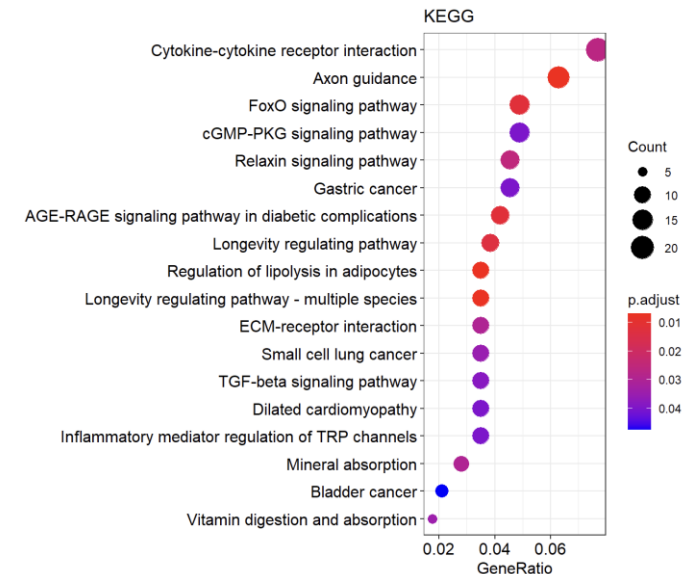


Enrichment analysis

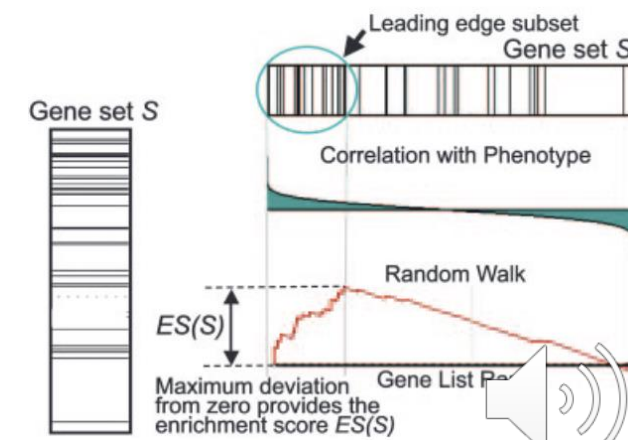
Overview of existing pathway analysis methods



ClusterProfiler

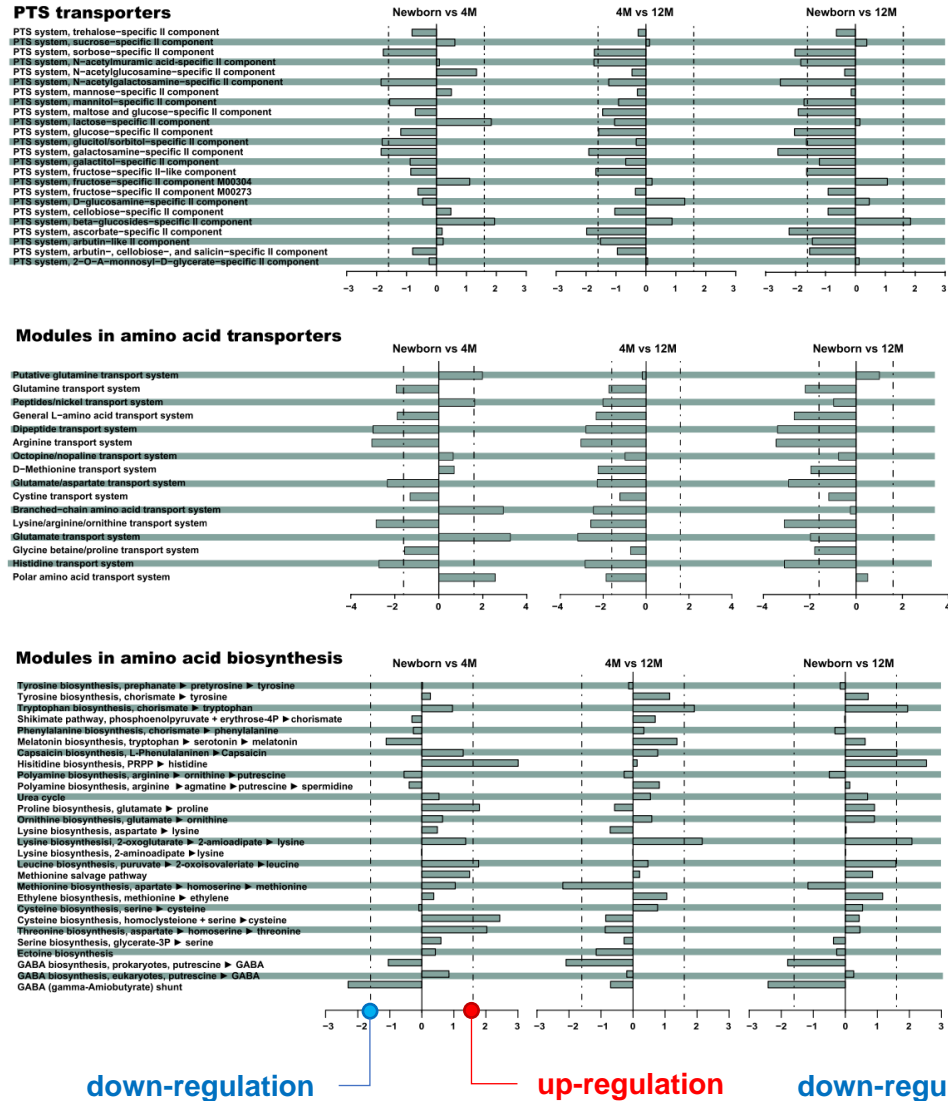


GSEA

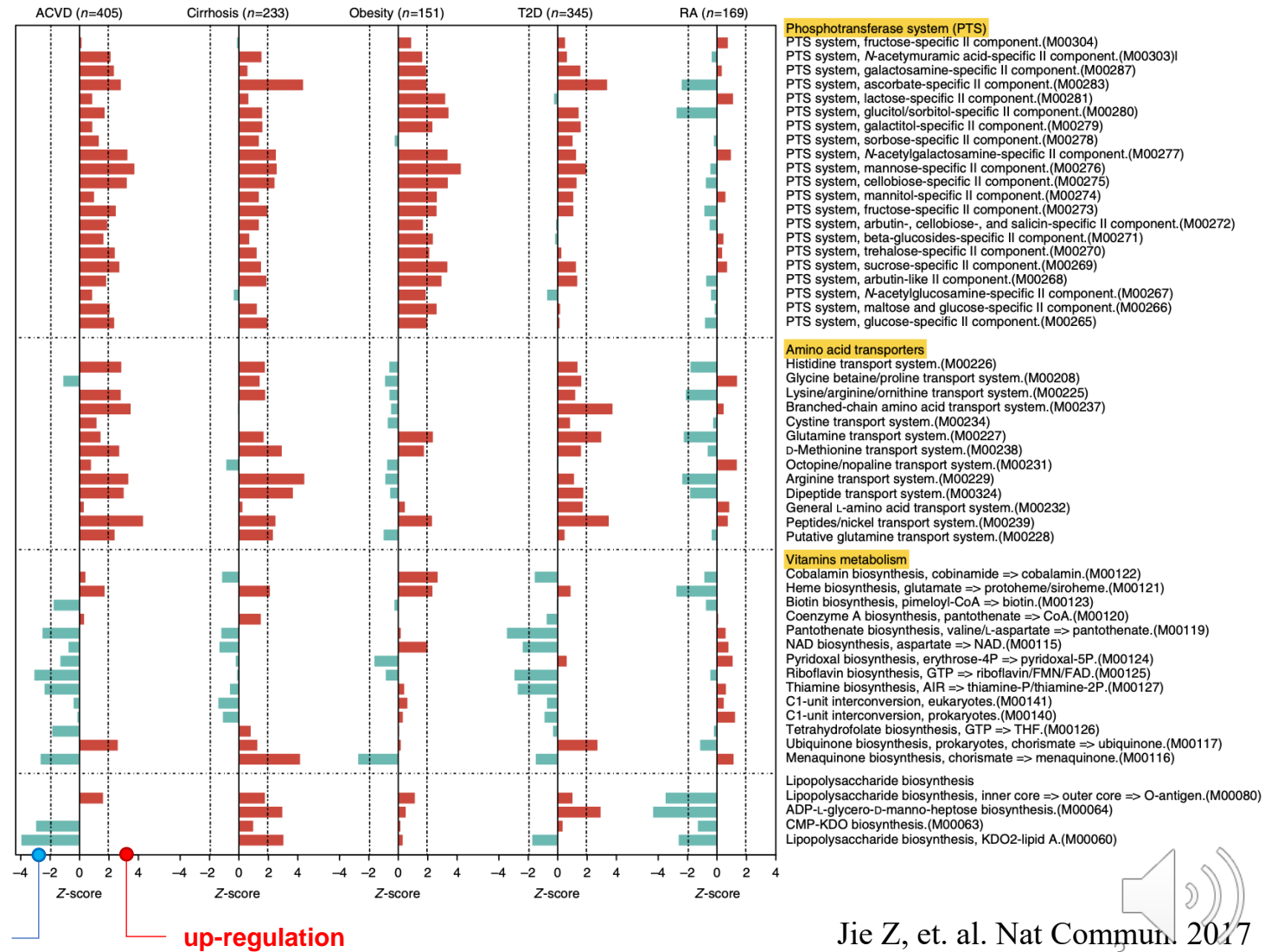


Reporter Score

Reporter score used in gut microbiome of mother and infant

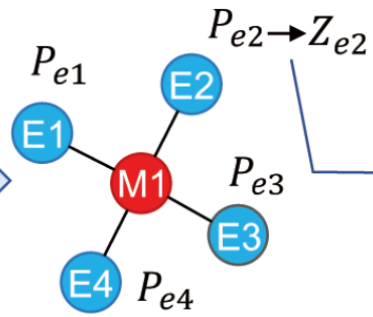
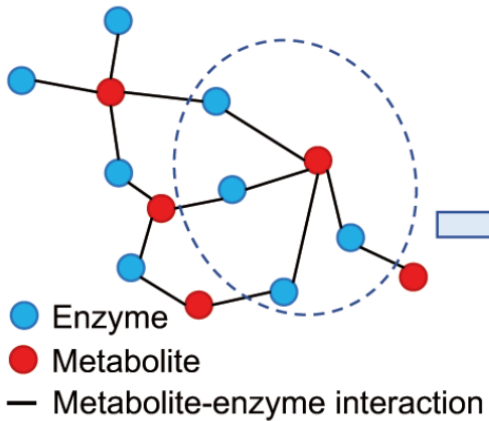


Reporter score used in gut microbiome of atherosclerotic cardiovascular disease



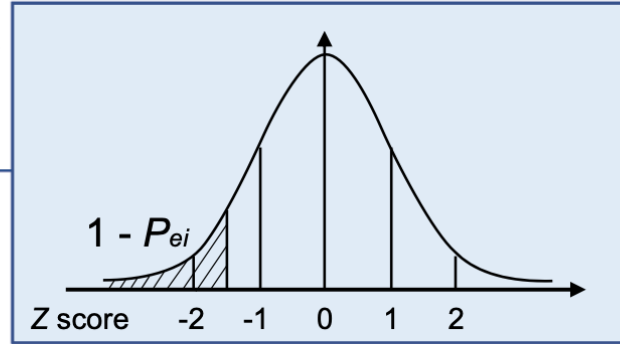
Misuse of Reporter Score Leads to Misleading Conclusions in Microbial Research

1) Patil et al., 2005

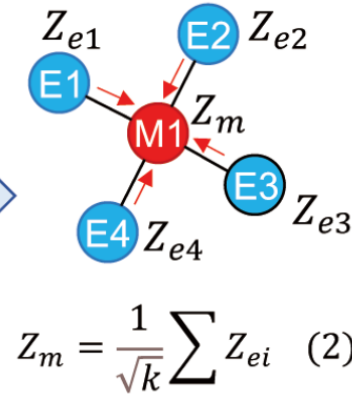


Reporter Score

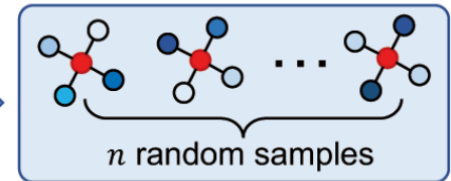
normal distribution



$$Z_{ei} = \theta^{-1}(1 - P_{ei}) \quad (1)$$

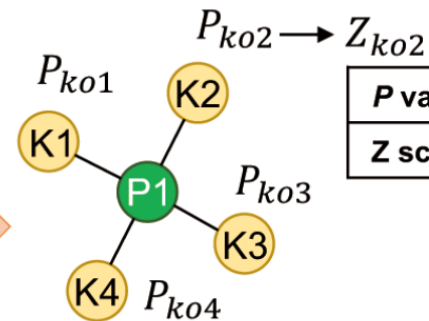
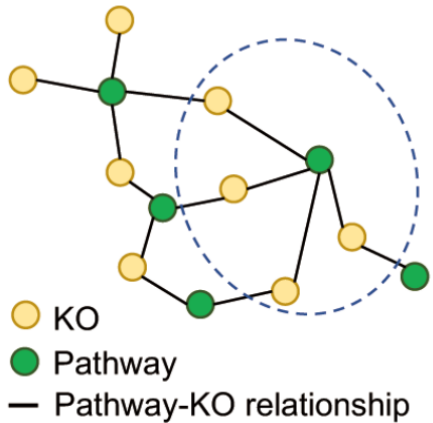


Background distribution



$$Z_m^{corrected} = \frac{Z_m - \mu_k}{\sigma_k} \quad (3)$$

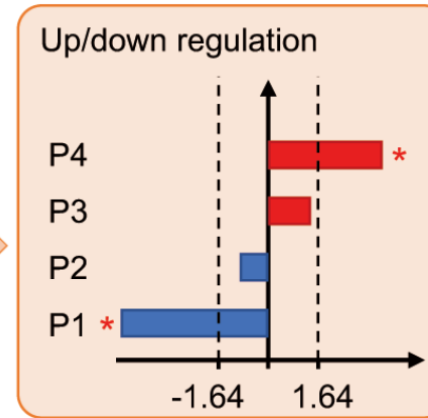
2) Surveyed articles



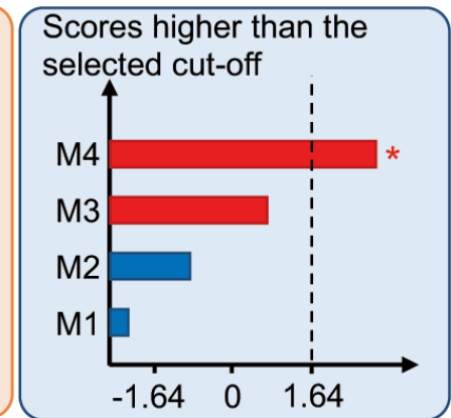
P value	0.01	0.05	0.5	0.95	0.99
Z score	2.33	1.64	0.0	-1.64	-2.33

... $Z_{corrected\ pathway}$

Misuse



Correct use

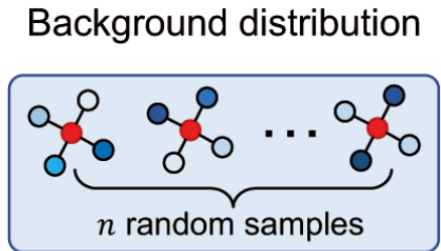
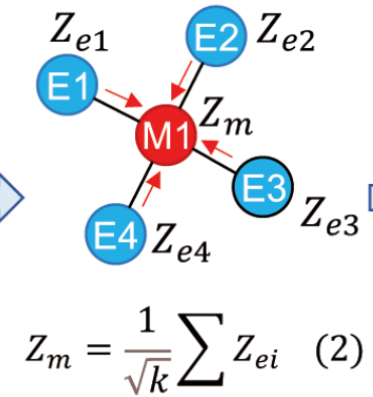
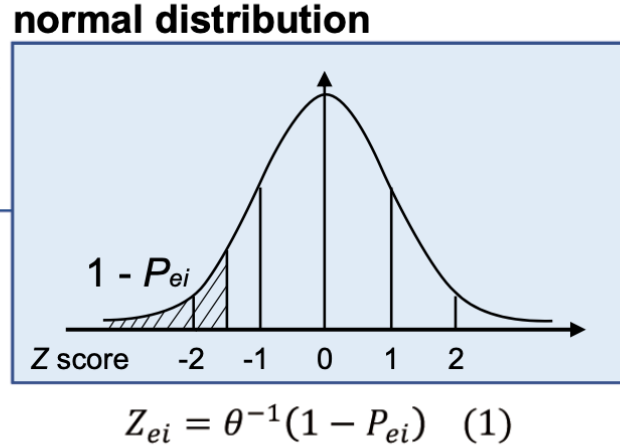
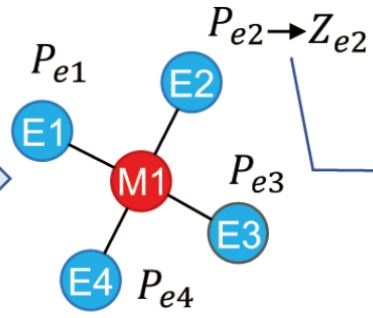
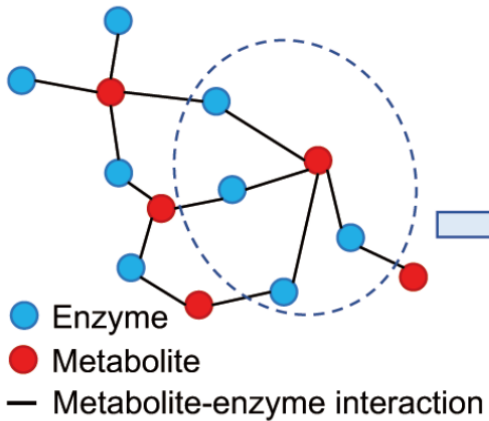


Patil and Nielsen. PNAS. 2005
Feng Q, et. al. Nat Commun. 2013

Misuse of Reporter Score Leads to Misleading Conclusions in Microbial Research

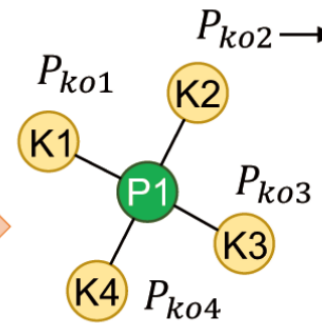
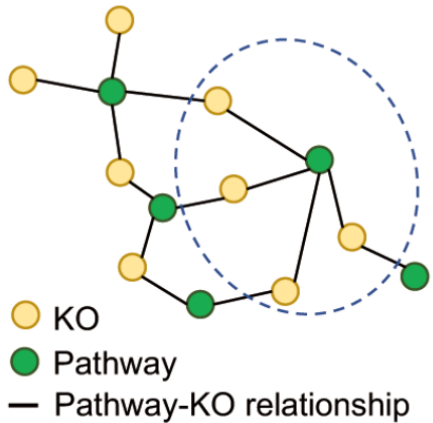
Misuse of Reporter Score

1) Patil et al., 2005

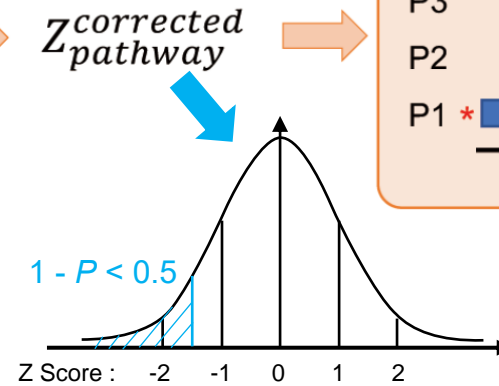
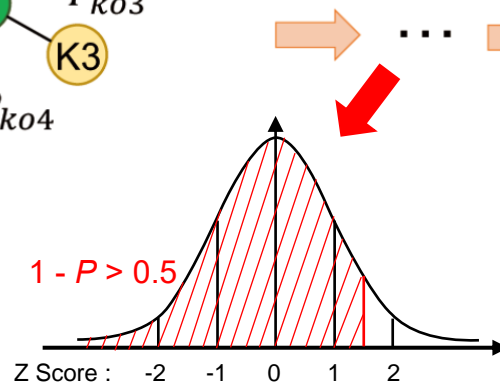


$$Z_m^{corrected} = \frac{Z_m - \mu_k}{\sigma_k} \quad (3)$$

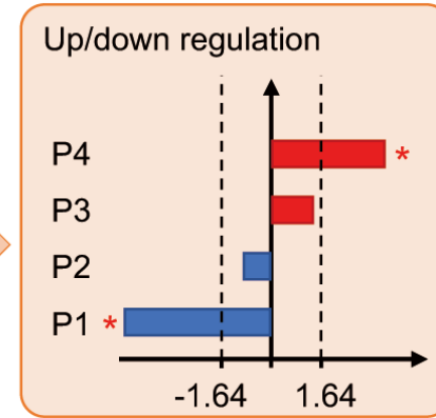
2) Surveyed articles



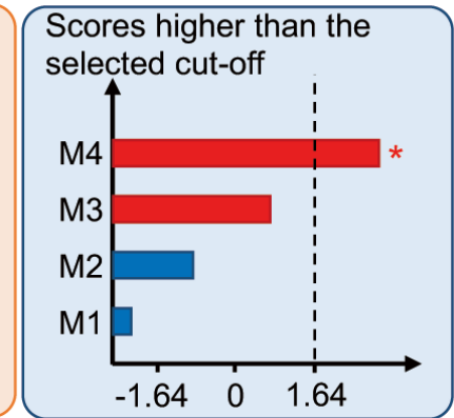
P value	0.01	0.05	0.5	0.95	0.99
Z score	2.33	1.64	0.0	-1.64	-2.33



Misuse



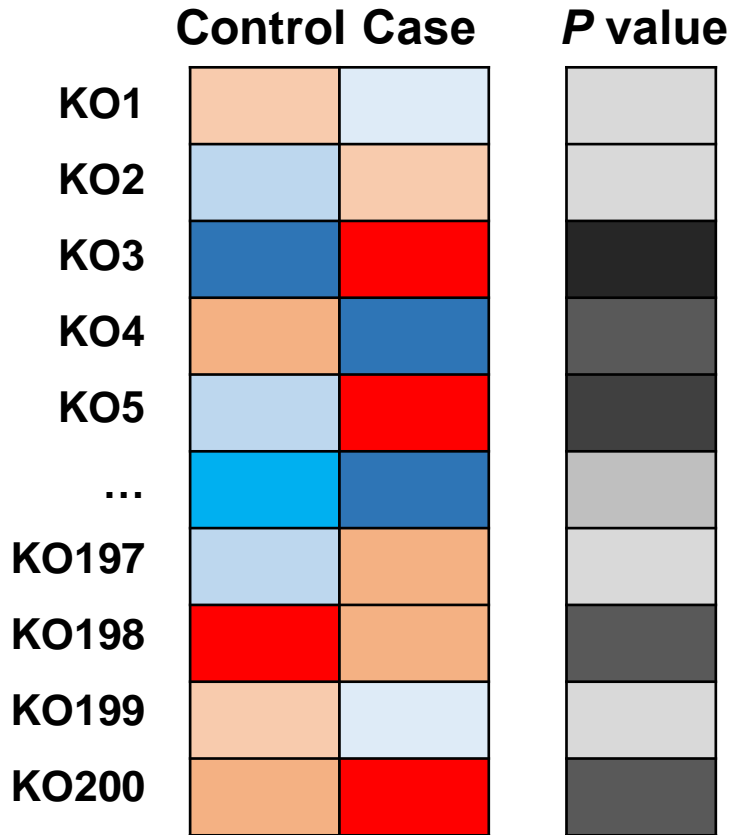
Correct use



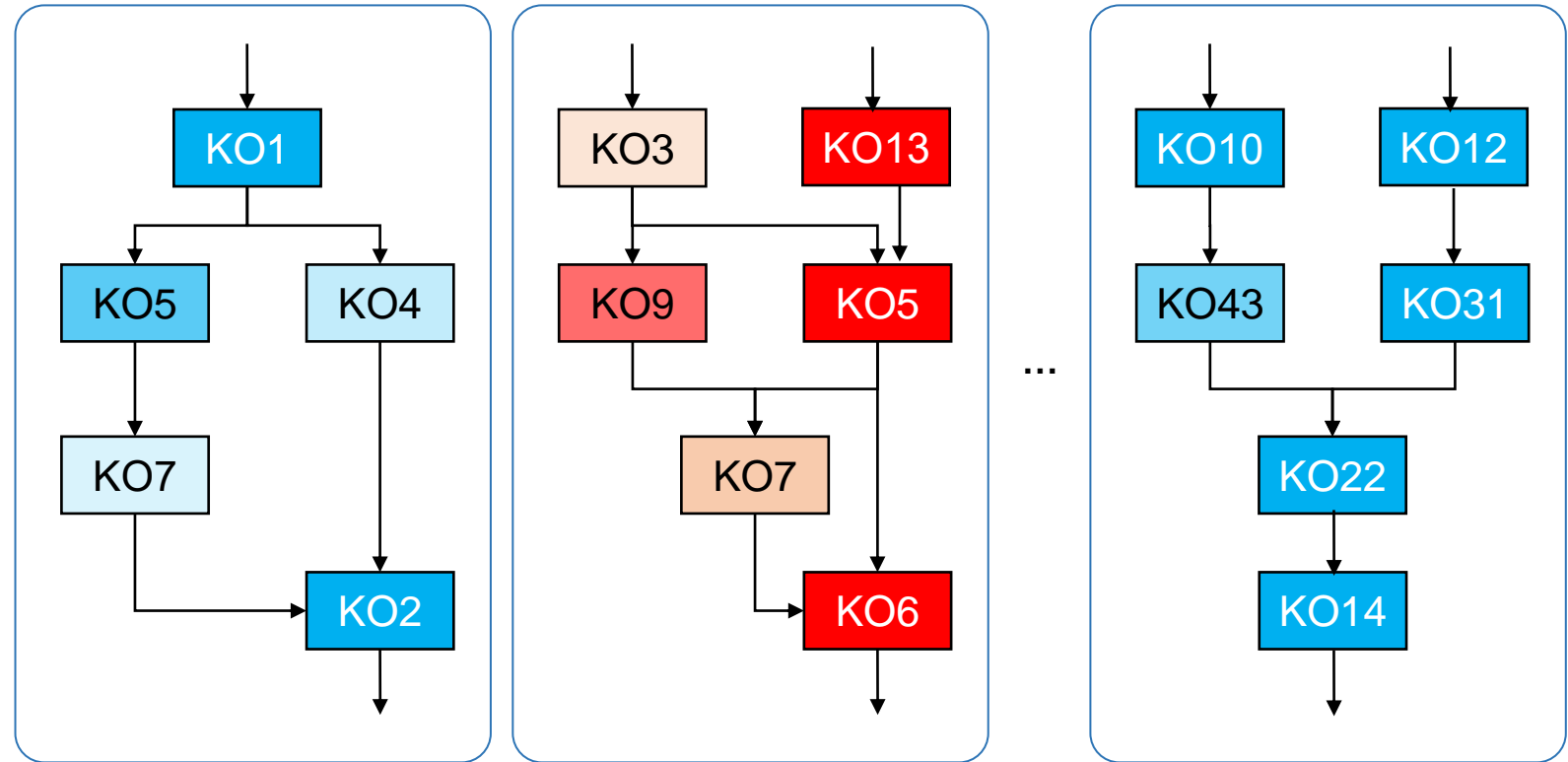
Patil and Nielsen. PNAS. 2005
Feng Q, et. al. Nat Commun. 2015

Misuse of *Reporter Score* Leads to Misleading Conclusions in Microbial Research

Simulated dataset



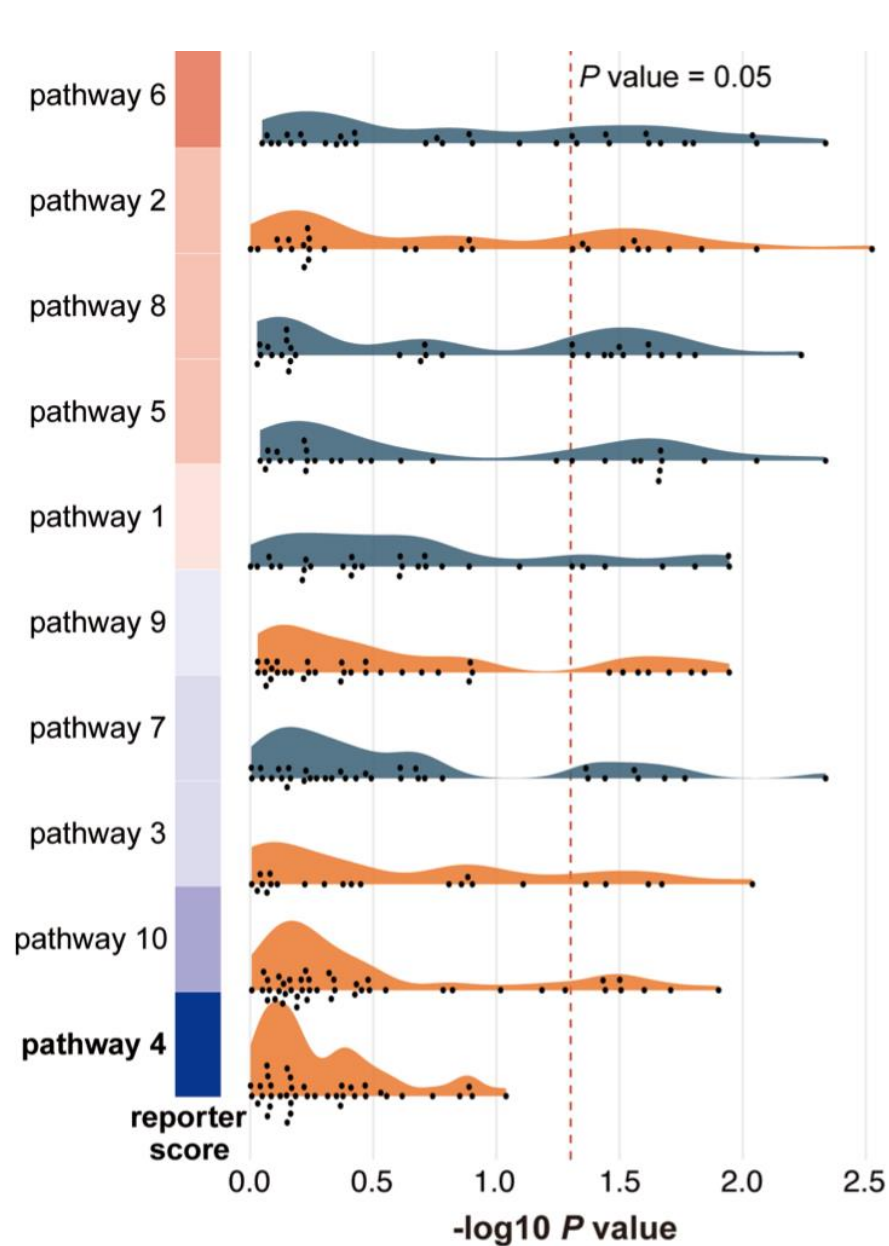
50 KOs were significantly differentially expressed



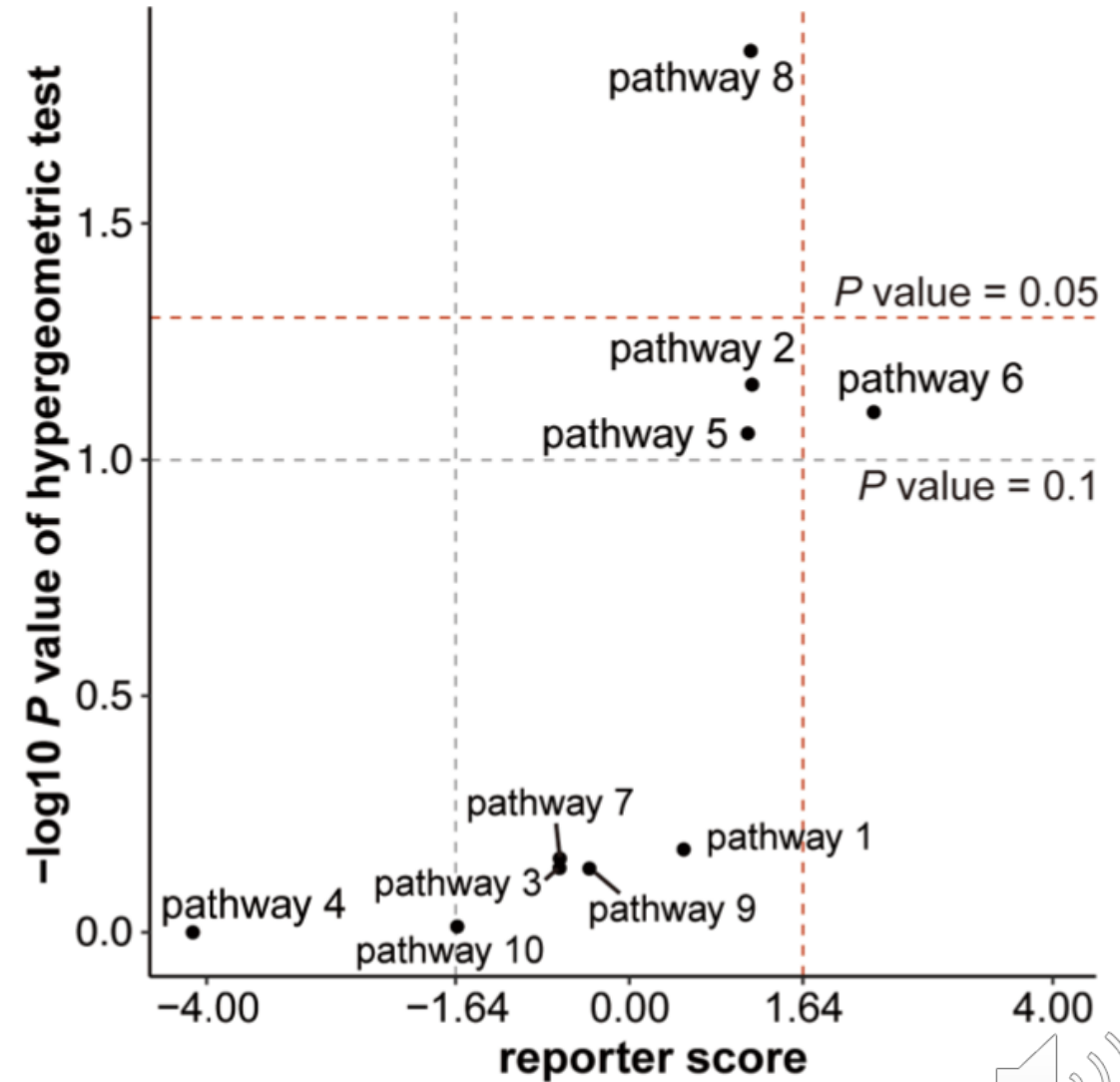
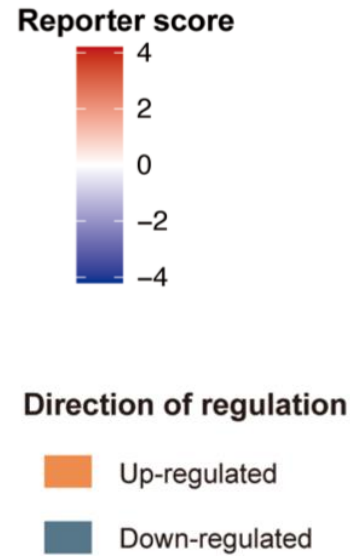
(KOs in the same pathway have consistent directionality of regulation)



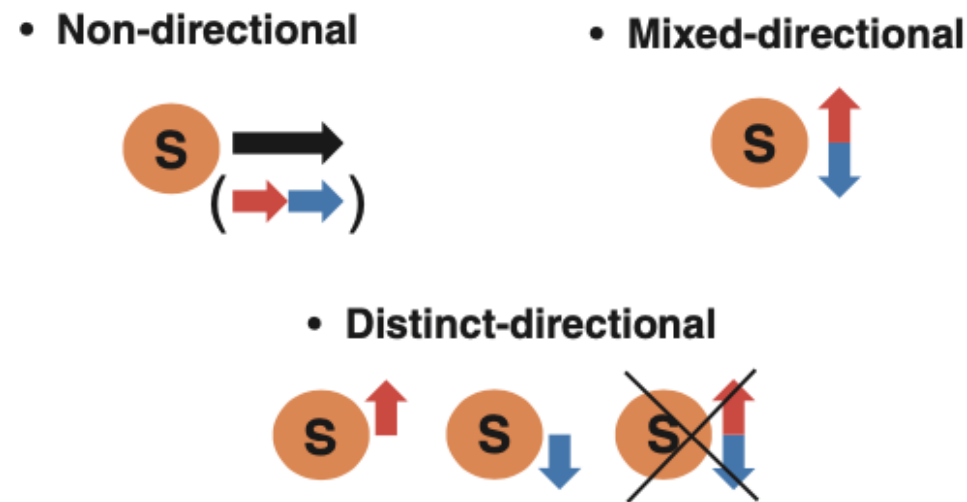
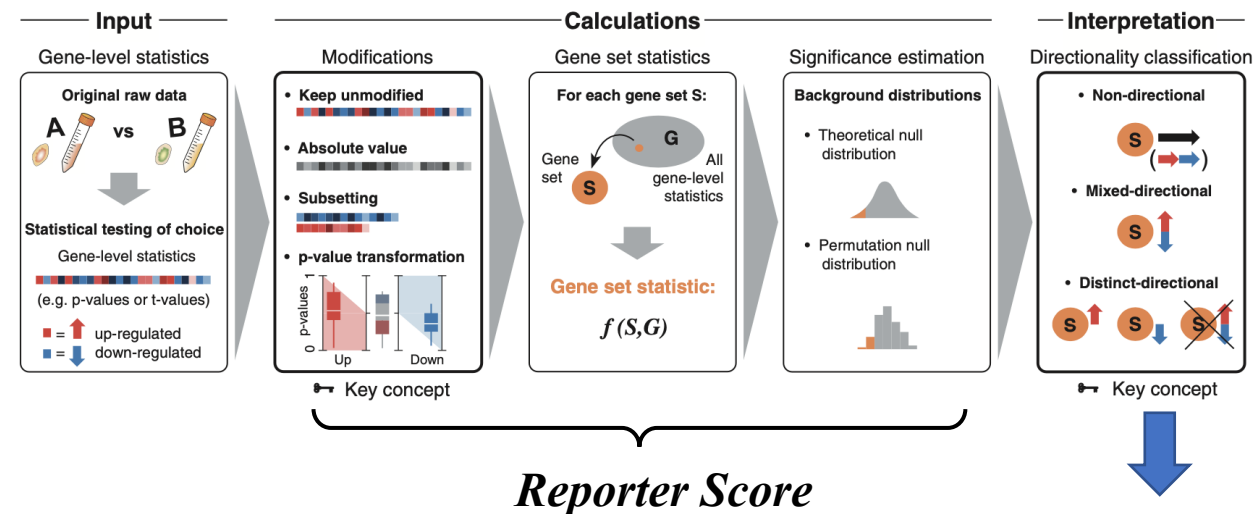
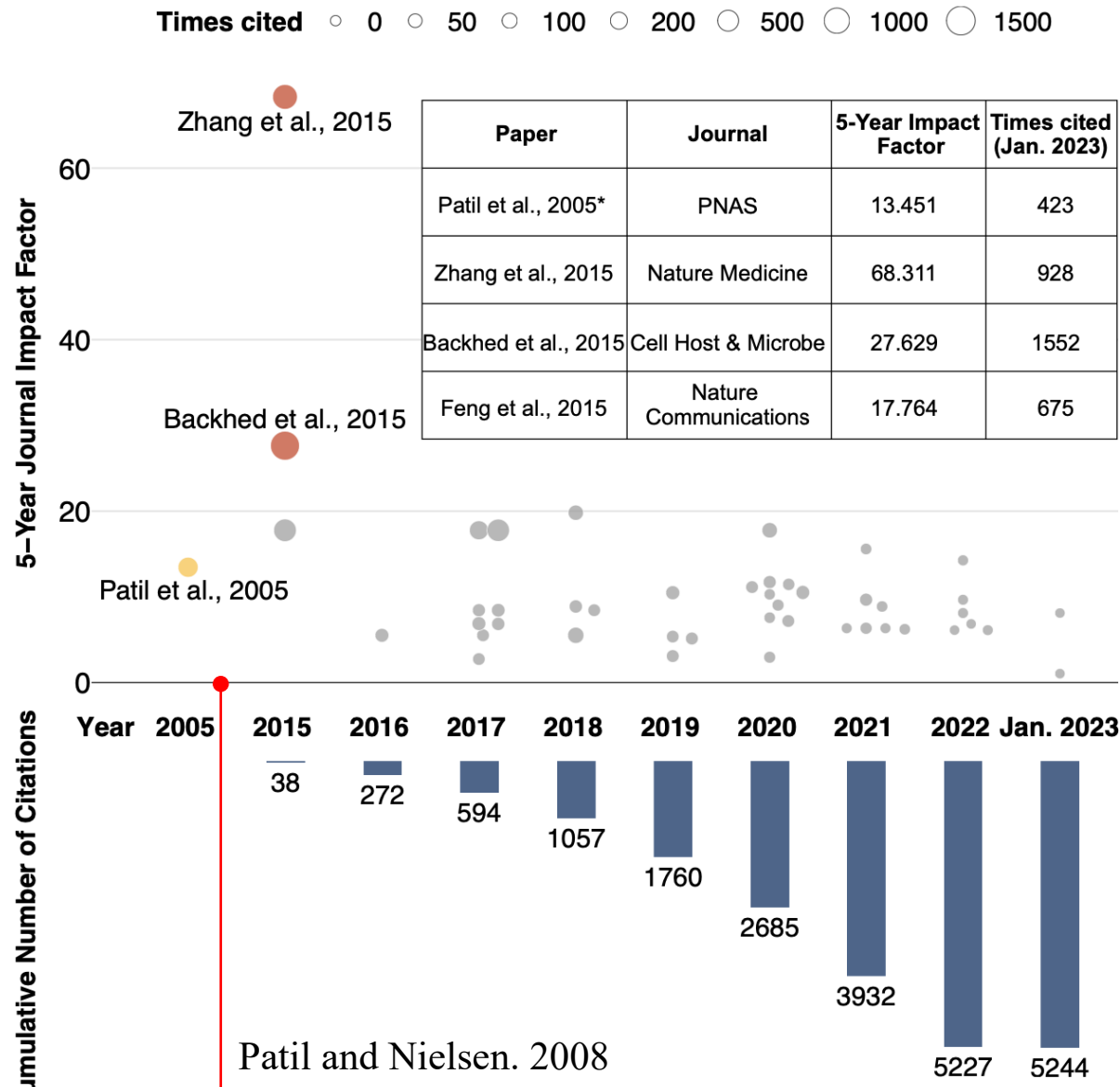
Misuse of *Reporter Score* Leads to Misleading Conclusions in Microbial Research



Simulated dataset



Misusing Reporter Score Has Caused Snowball Effect in Microbiome Field



Reporter Feature analysis (Piano)

Misuse of Reporter Score Leads to Misleading Conclusions in Microbial Research

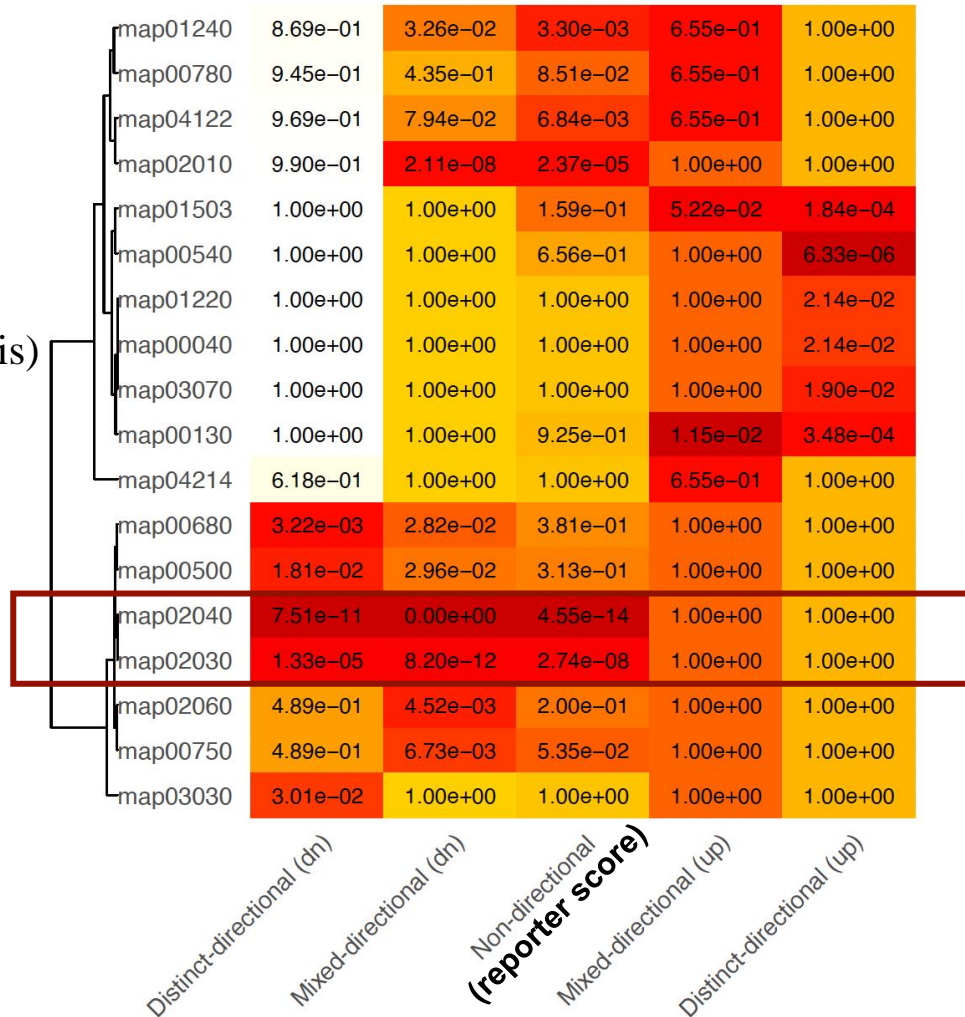
Demo dataset
(NAFLD with
Fibrosis vs. without Fibrosis)
Rohit Loomba, et. al. 2017

7246 KOs

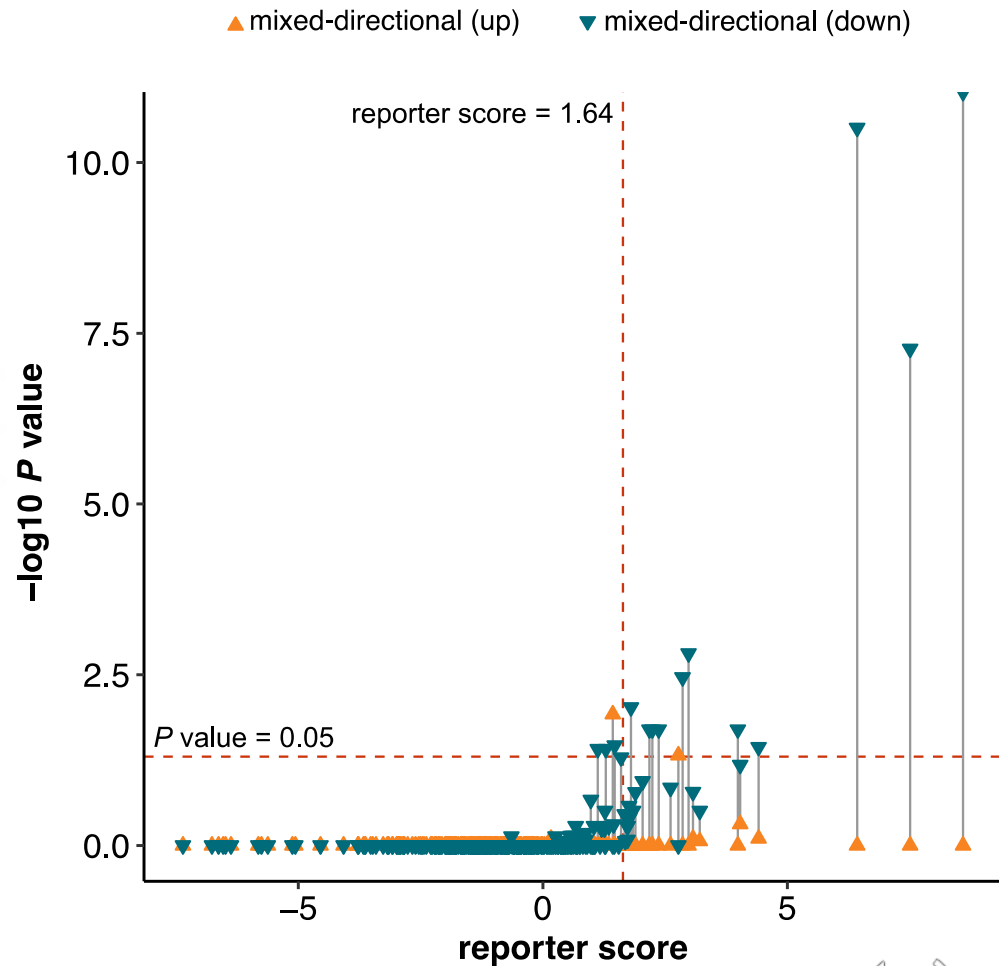


457 Pathways

Reporter feature (Piano R package)



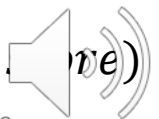
Reporter score vs. Reporter feature



map02040: non-directional reporter score = 8.596
map02030: non-directional reporter score = 6.429

<https://github.com/ddhmed/ReporterScore>

$$P \text{ value} = 1 - \theta(\text{reporter score})$$



Conclusions

- In scientific research, special attention should be paid to the misinterpretation or misuse of statistical techniques, which will cause the problems of unreliability and irreproducibility.
- Interpreting *reporter score* under habitual thinking and adopting a published method on faith without comprehending its statistical meaning are the underlying causes why such a naïve error kept being made.
- A specialized checklist for bioinformatics research is necessary in the review system of journals.
- Multi-disciplinary collaboration, particularly with statisticians, should be encouraged to promote rigorous research.
- Strengthening the training of bioinformaticians with solid basic statistical and mathematical knowledge will guarantee the significant and stable development of bioinformatics, facilitating scientific research in various domains.



Misuse of reporter score in microbial enrichment analysis

Lei Liu¹, Ruixin Zhu¹, Dingfeng Wu^{2, *}

¹ The Shanghai Tenth People's Hospital, School of Life Sciences and Technology, Tongji University, Shanghai 200072, P. R. China.

² National Clinical Research Center for Child Health, the Children's Hospital, Zhejiang University School of Medicine, Hangzhou 310058, Zhejiang, P.R. China.

* Corresponding authors: Dingfeng Wu (dfw_bioinfo@126.com)

Liu, Lei, Ruixin Zhu, and Dingfeng Wu. 2023. “Misuse of reporter score in microbial enrichment analysis.” *iMeta*. e95. <https://doi.org/10.1002/imt2.95>





“iMeta”是由威立、肠菌分会和本领域数百位华人科学家合作出版的开放获取期刊，主编由中科院微生物所刘双江研究员和荷兰格罗宁根大学傅静远教授共同担任。目的是发表原创研究、方法和综述以促进宏基因组学、微生物组和生物信息学发展。目标是发表前10%(IF > 15)的高影响力论文。期刊特色包括视频投稿、可重复分析、图片打磨、青年编委、前3年免出版费、50万用户的社交媒体宣传等。2022年的三月、六月和九月期已正式在线出版发行!



主页: <http://www.imeta.science>

出版社: <https://wileyonlinelibrary.com/journal/imeta>



投稿: <https://mc.manuscriptcentral.com/imeta>



office@imeta.science



[iMeta](#)

[宣传片](#)

