新技术助力多种肿瘤内细胞丰度的精准预测

作者：writer 来源：科学网

本文原地址：https://www.iikx.com/news/progress/30295.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

新技术助力多种肿瘤内细胞丰度的精准预测。11月8日，中国科学院深圳先进技术研究院合成生物学研究所副研究员李雪飞团队与香港浸会大学副教授田亮团队展开合作的最新成果发表于《美国科学院院刊》。合作团队开发了一种基于深度学习与公开单细胞数据集的解卷积算法，能够较精确地估算19种实体肿瘤中16种细胞类型的丰度。

肿瘤微环境是一个由多种类型细胞组成的复杂生态系统。实体肿瘤内，除了癌细胞之外，还存在数量可观的血管内皮细胞、成纤维细胞、以及多种免疫细胞。大量研究表明，这些不同类型细胞的比例，在多种癌症类型中与患者的病情进展密切相关。因此，精准量化肿瘤微环境中各类细胞的比例，对于深入理解肿瘤的发生与发展规律、优化癌症诊疗策略等具有重要意义。

在现有的技术中，流式细胞术和单细胞RNA测序等实验方法能够直接测定肿瘤微环境中的细胞比例，但这些方法通常成本较高。

该研究整合了来源于6种实体肿瘤类型的12个单细胞RNA测序数据集，为后续合成虚拟肿瘤批量RNA测序数据（训练集）提供了更为全面的参考数据。该研究提出了一种新的采样方法，从而使得所合成的虚拟肿瘤批量RNA测序数据中包含更多样的细胞比例组合。

此外，在合成批量RNA测序数据时，研究人员通过对基因的过滤来保留与每种细胞类型高度相关的基因，以降低输入数据的维度；再对表达谱进行过滤以保留与真实肿瘤表达谱相似度较高的样本。这些创新点共同有效提升了所合成的虚拟肿瘤批量RNA测序数据的质量。

在深度神经网络结构中，DeSide创新性地采用了两个全连接网络，分别从生物信号通路和基因表达谱中提取特征信息。其中，生物信号通路通过引入粗颗粒度的特征，有效提升了输入数据的多样性。此外，肿瘤细胞的基因表达谱在不同癌症类型之间差异较大，该技术通过改进算法的输出策略，有效减少了直接预测肿瘤细胞比例所带来的误差。

该研究系统对比了DeSide与现有算法在预测肿瘤微环境内细胞比例的准确性，结果表明，新技术能较好地预测多种肿瘤类型内不同细胞类型的比例，即便与依赖参考数据的算法相比较，该技术在特定肿瘤类型上的表现依然出色。同时，DeSide能较准确地预测训练集中未包含的癌症类型，显示了较好的泛化能力。

此外，研究发现，基于DeSide预测的细胞比例能够有效的将患者按照病情进展情况进行评估分型。未来，该技术有望进一步帮助探究不同细胞之间关键的相互作用，从而为寻找潜在的临床治疗靶点提供新的可能。该研究为深入理解肿瘤发生与发展机制、评估患者预后和制定精准治疗策略

提供了有力的方法与数据支持。（来源：中国科学报 刁雯蕙）

相关论文信息：https://doi.org/10.1073/pnas.2407096121

作者：李雪飞等 来源：《美国科学院院刊》

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发