



长读长测序解码转录组复杂性

徐楚雯^{1#}, 李佳^{1#}, 尹晨熙^{2#}, 李珂莹^{2#}, 王甜², 高石文¹, 余越¹, 常乘³, 王也⁴, 童国庆², 龚亮⁵, 胡雯⁶, 柴国梁^{7,8,9},
韩怡来⁸, 秦谦¹⁰, 金锁钦¹, 邹秀芬¹, 李博¹¹, 张天缘¹², Ana Conesa¹³, 王运浩^{2*}, 王定杰^{1,14*}

¹武汉大学数学与统计学院

²西安交通大学第一附属医院生殖医学科

³国家蛋白质科学中心（北京）

⁴四川省濒危野生动物保护生物学重点实验室，成都大熊猫繁育研究基地

⁵浙江大学医学院附属第一医院泌尿外科和良渚实验室

⁶中国科学院深圳先进技术研究院脑认知与脑疾病研究所

⁷首都医学科学创新中心（CIMR）

⁸首都医科大学宣武医院神经内科，国家神经疾病医学中心

⁹北京市老年医学研究中心

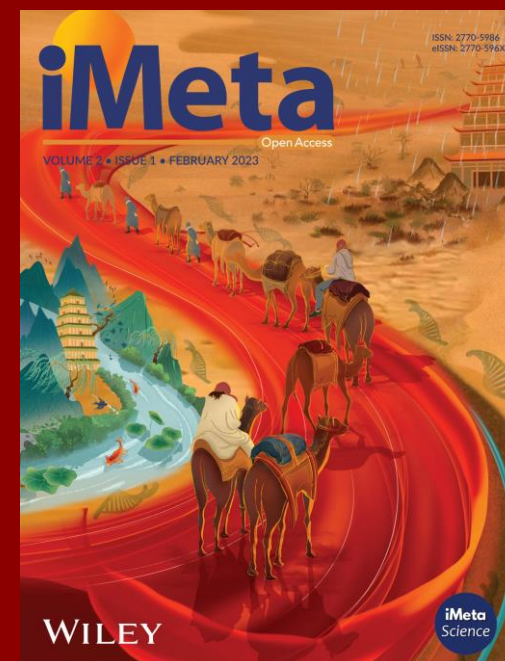
¹⁰武汉大学人工智能学院

¹¹密歇根大学医学中心Frankel心脏与脑健康研究所

¹²贝纳基因研究院

¹³西班牙国家研究委员会整合系统生物学研究所

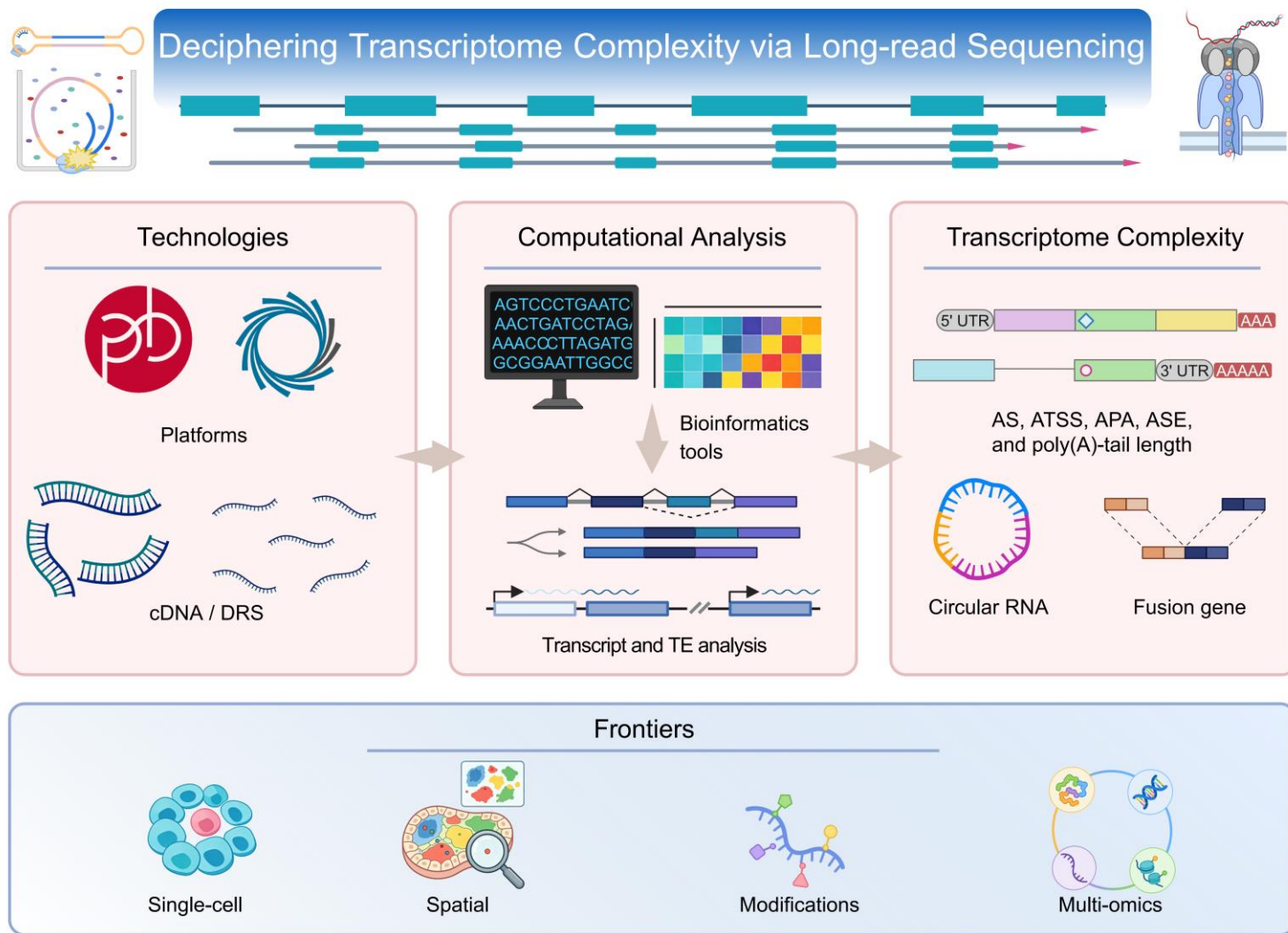
¹⁴武汉大学武汉数学与智能研究院



Chuwen Xu, Jia Li, Chenxi Yin, Keying Li, Tian Wang, Shiwen Gao, Yue Yu, *et al.* 2026. Deciphering Transcriptome Complexity via Long-read Sequencing. *iMeta* 5: e70171. <https://doi.org/10.1002/imt2.70171>



概述



- 系统梳理长读长测序解析转录组复杂性的技术优势：总结 PacBio、ONT 及国产长读长测序平台在全长转录本、可变剪接及重复序列来源转录解析中的应用，阐明其如何突破短读长局限，推动研究迈向异构体水平。
- 构建贯穿实验设计、计算分析与生物学解释的实践指南：围绕平台与建库策略选择、核心计算处理及下游生物学解释，归纳关键决策与实施要点。
- 总结前沿进展与最佳实践，促进分析的标准化和可重复性：提炼可复用的分析框架与实践检查清单，梳理长读长转录组学的新兴方向，为研究者开展高分辨率、异构体层面的发现与解释提供指导。

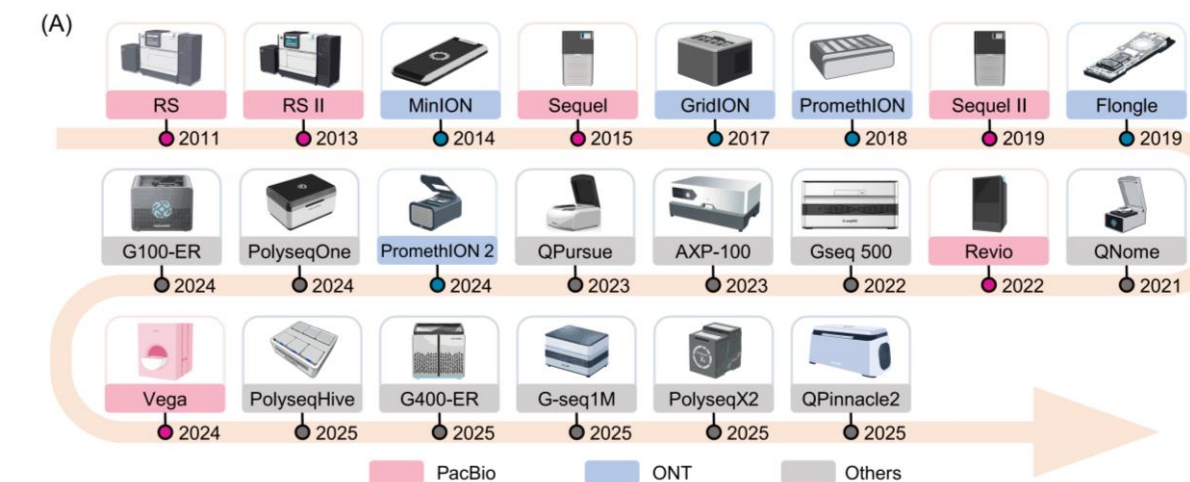


亮点

- **解析转录组复杂性：** 本文阐述长读长测序在解析全长转录本及关联调控特征中的优势，推动转录组研究迈向异构体水平。
- **指导平台与建库选择：** 本文系统梳理测序平台与建库策略对转录本完整性、准确性及生物学解读的影响，为实验设计提供依据。
- **构建标准化分析流程：** 本文整合核心数据处理与下游分析方法，提供可复用的计算框架，促进研究的标准化与可重复性。
- **展望新兴研究前沿：** 本文总结单细胞、空间、直接 RNA 测序及多组学整合的最新进展，展望长读长转录组学的发展方向。



长读长转录组学分析流程全景概览



(A) 长读长测序平台发展时间线：PacBio（粉色）、ONT（蓝色）及其他（灰色）。

(B) 基于长读长测序的转录组学端到端分析流程概述。

(C) 长读长转录组学前沿方向：单细胞、空间转录组、直接 RNA 测序及多组学整合应用。

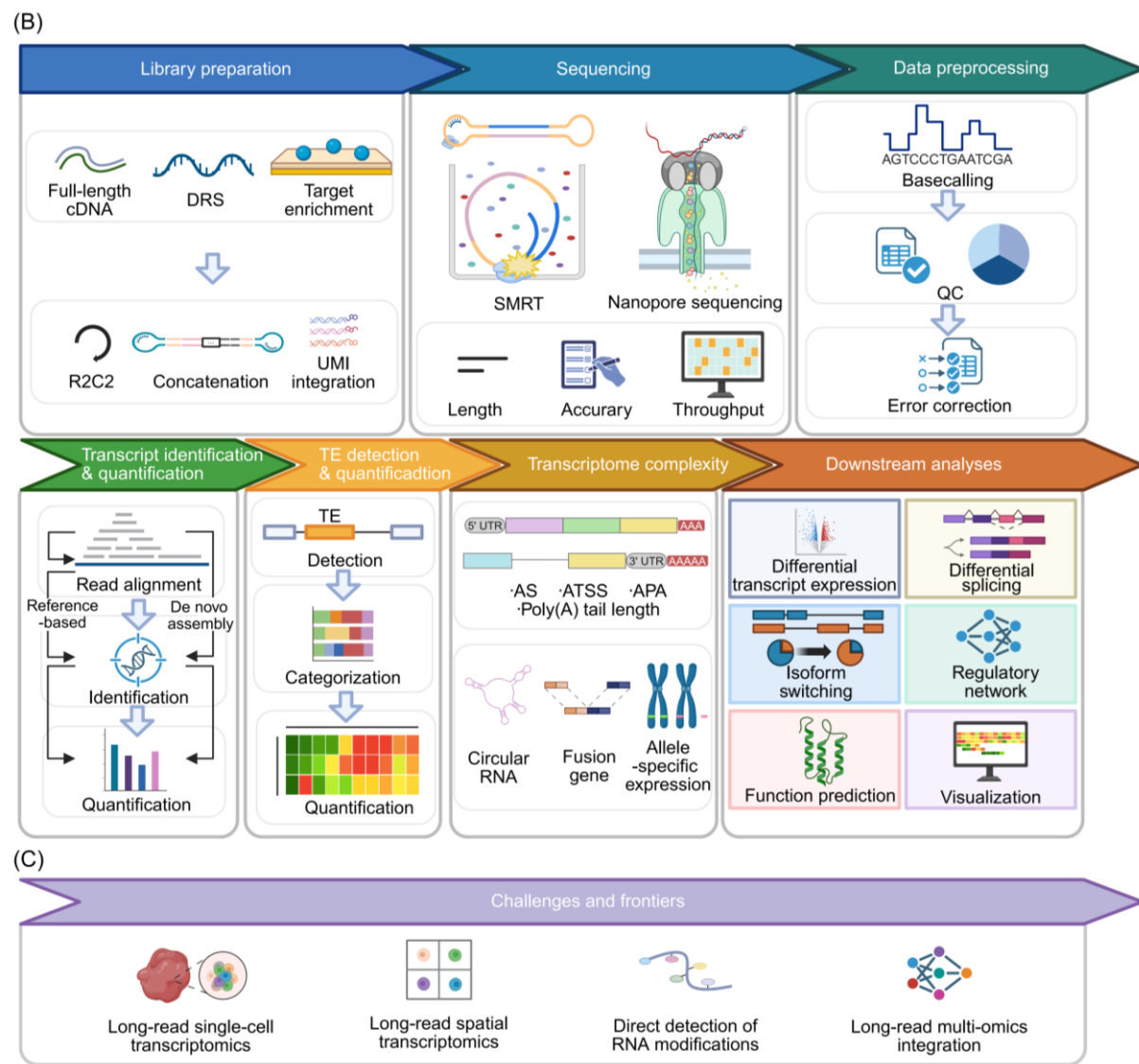
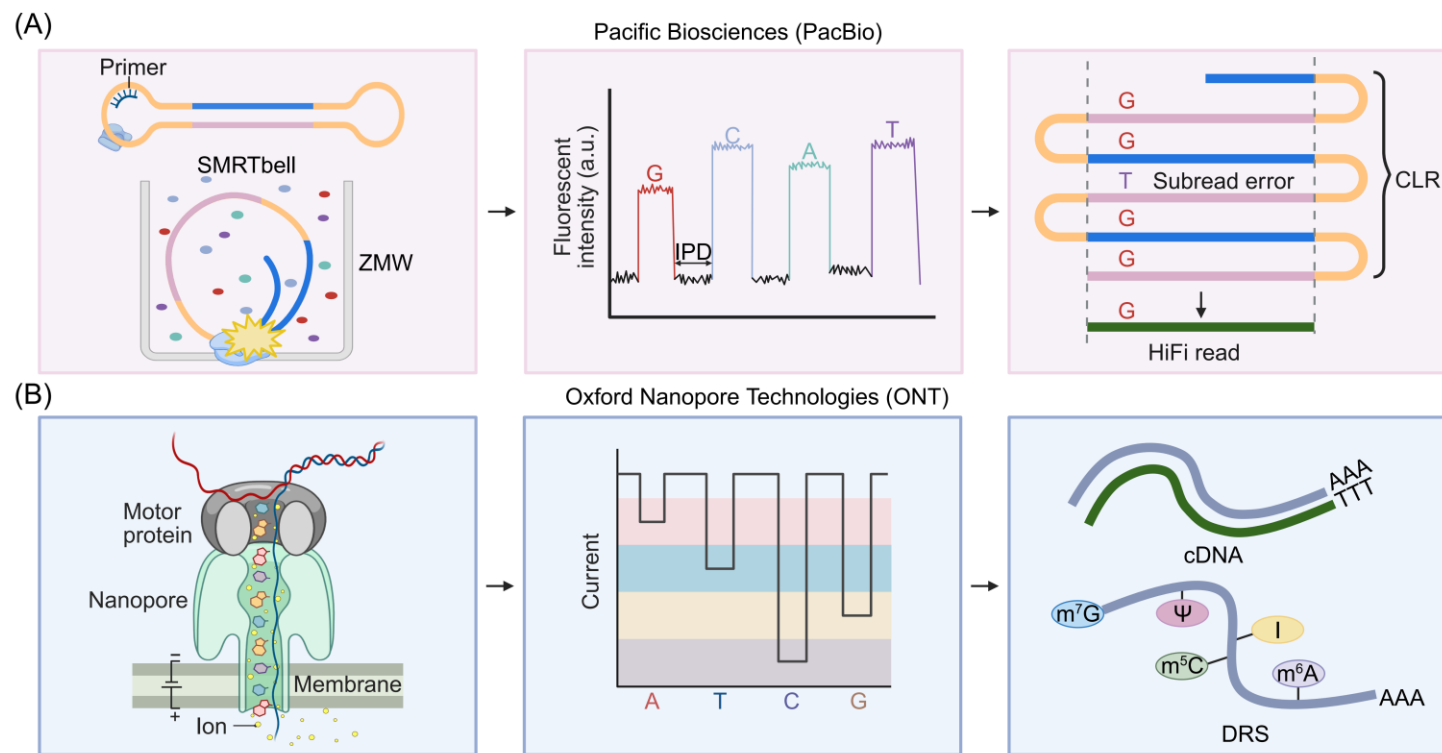


图1. 长读长转录组学分析流程全景概览



长读长转录组测序平台与原理

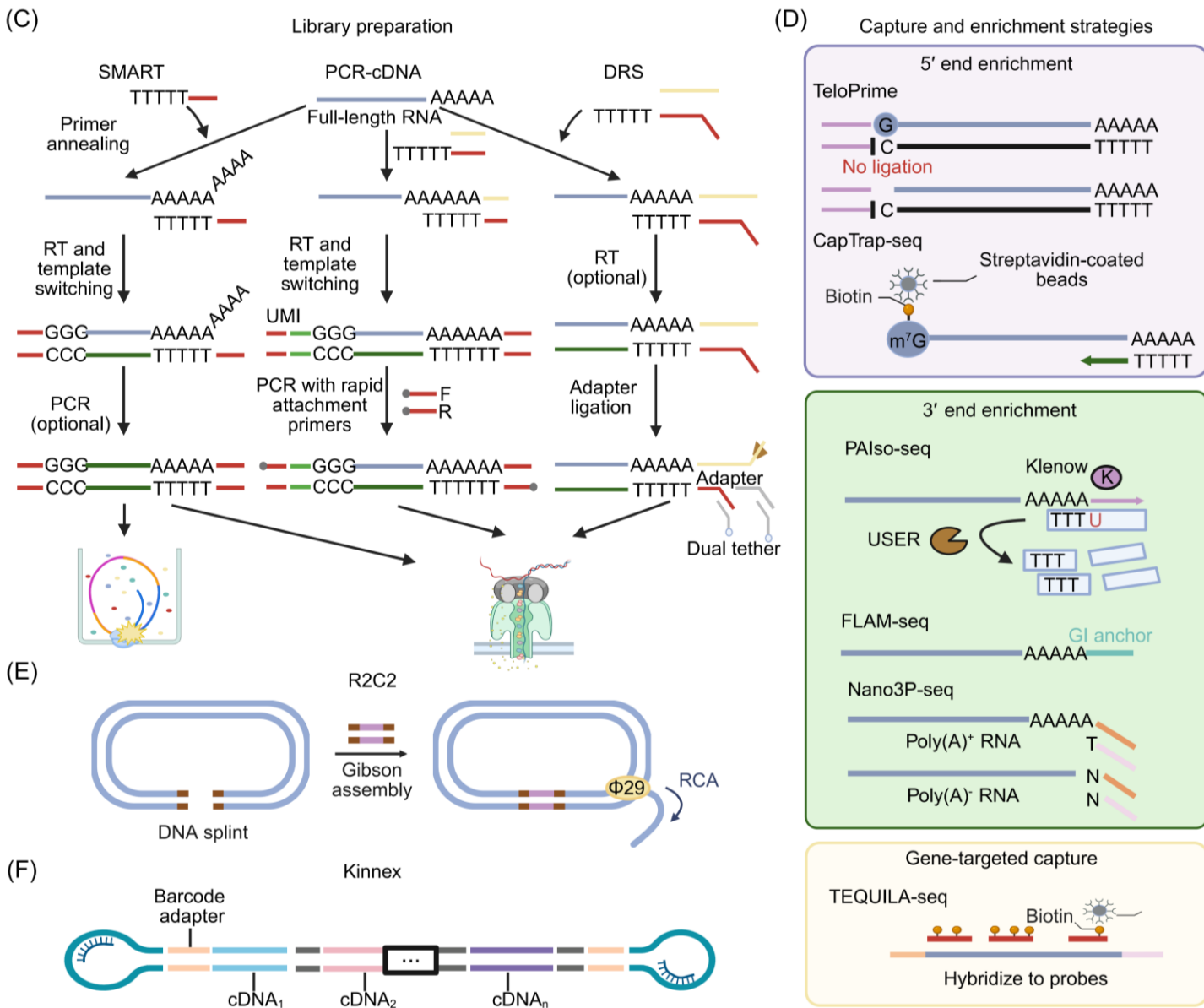


(A) PacBio 测序： SMRTbell 环状模板进入零模波导孔（ZMW），通过聚合酶合成过程中的荧光脉冲读取序列；经多轮测序与环形一致性分析，生成高准确度 HiFi 读段。

(B) ONT 测序： 马达蛋白调控核酸链通过纳米孔，记录随局部序列变化的离子电流，并经碱基识别转换为序列；直接 RNA 测序还可利用原始电流信号解析 RNA 修饰。

图2. 长读长测序平台与原理及转录组文库制备与富集策略概览

长读长文库制备与富集策略概览

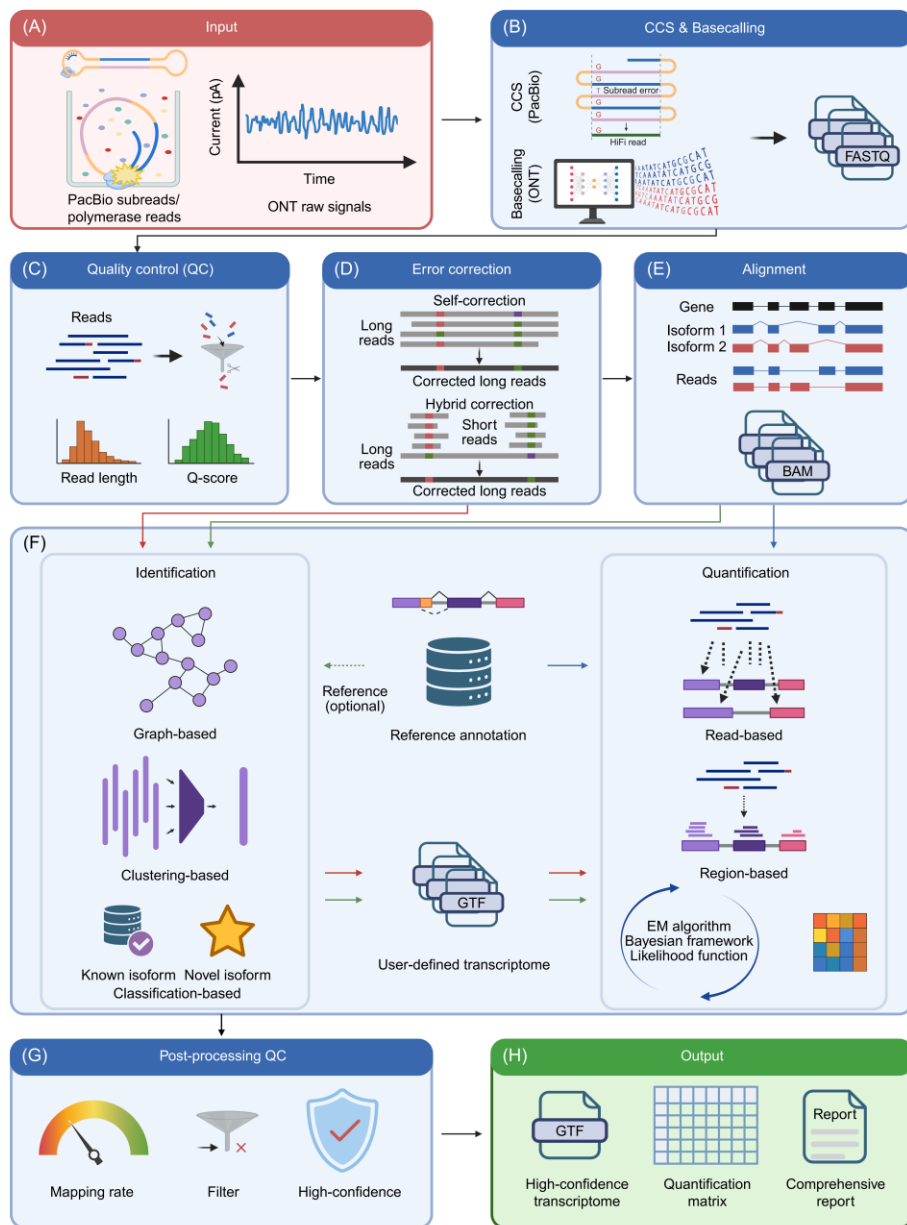


- (C) 文库制备流程：概述 SMART、ONT PCR-cDNA 与直接 RNA 测序（DRS）的主要建库步骤。
- (D) 捕获与富集策略：展示针对 5'端帽结构（CapTrap-seq）、3'端 poly(A) 尾（PAIso-seq）及目标基因（TEQUILA-seq）的代表性策略。
- (E) R2C2 一致性测序：通过滚环扩增生成串联重复单元，经测序与一致性计算提高读段准确度。
- (F) Kinnex 串联测序：将多个 cDNA 分子串联为长模板，提高单次测序的转录本产出。

图2. 长读长测序平台与原理及转录组文库制备与富集策略概览



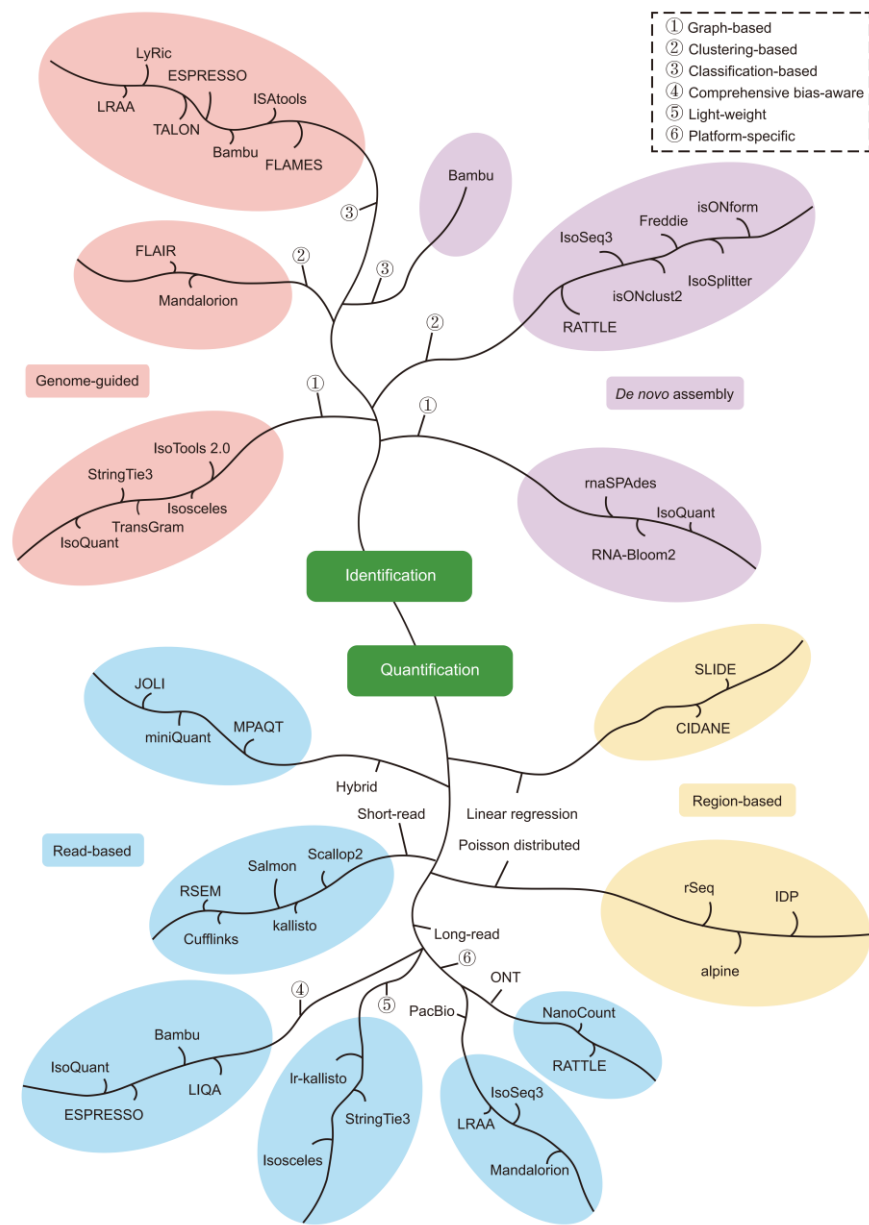
长读长转录组数据分析计算流程



- **预处理与质控:** 对 PacBio 和 ONT 数据进行平台特异性预处理、质量评估及按需纠错。
- **比对与重建:** 将高质量reads比对至参考基因组，或采用无参考策略重建转录本。
- **转录本识别与定量:** 结合参考注释或从数据中推断异构体结构，生成转录本注释并进行丰度估计。
- **结果筛选与输出:** 通过下游质控过滤低置信度结果，形成高置信度转录本集、异构体定量矩阵及分析报告。

图3. 长读长转录组数据分析计算流程

转录本识别与定量计算工具的方法学分类



- **异构体识别**：包括参考基因组引导和从头组装两类方法，并依据核心算法范式进一步分为：（1）基于图的方法、（2）基于聚类的方法和（3）基于分类的方法。
- **异构体定量**：方法分为基于读段和基于区域的模型，适用于短读长、混合及长读长数据。长读长方法进一步分为（4）综合偏倚校正方法、（5）轻量级方法和（6）平台特异性方法。统计假设（如泊松模型和线性回归）在适用处标注。每类方法中展示代表性工具。

图4. 转录本识别与定量计算工具的方法学分类



转座元件衍生转录本的特征分析

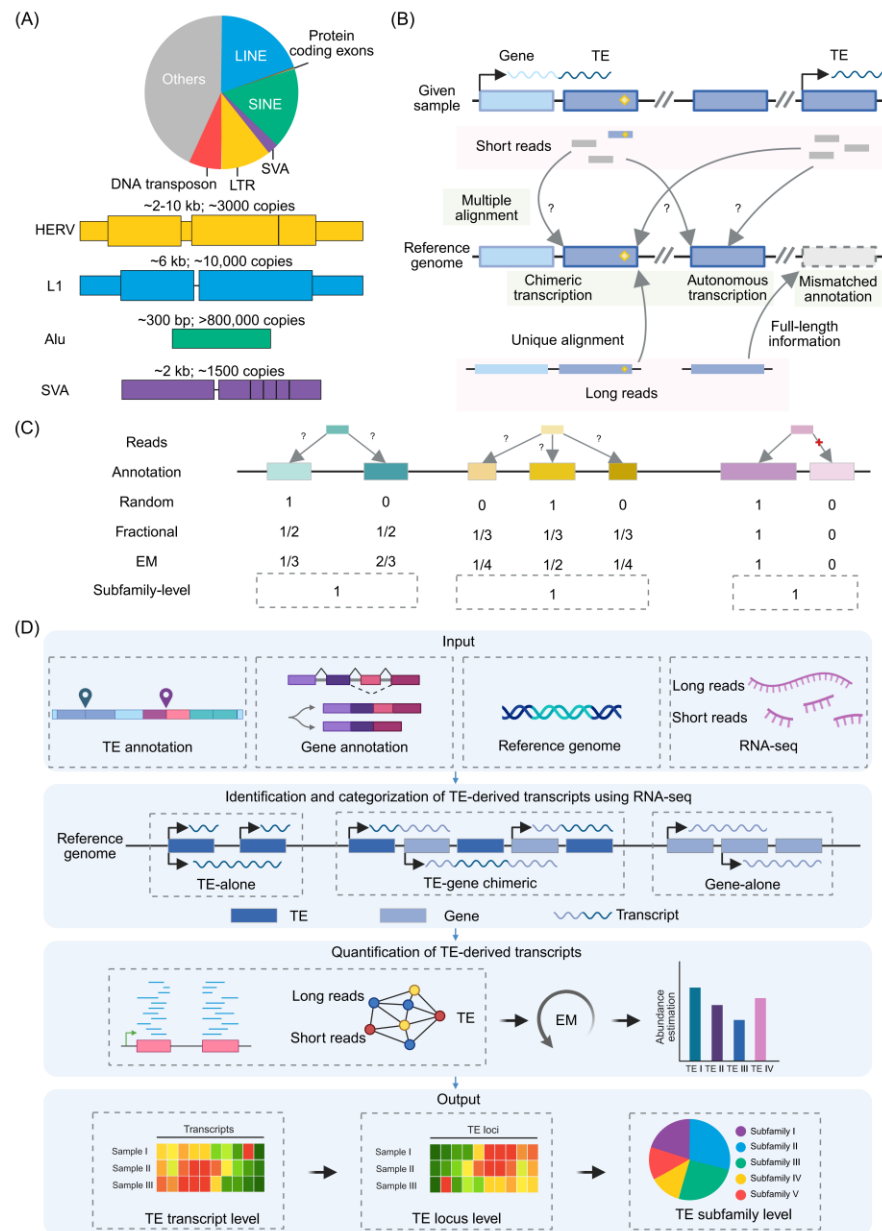
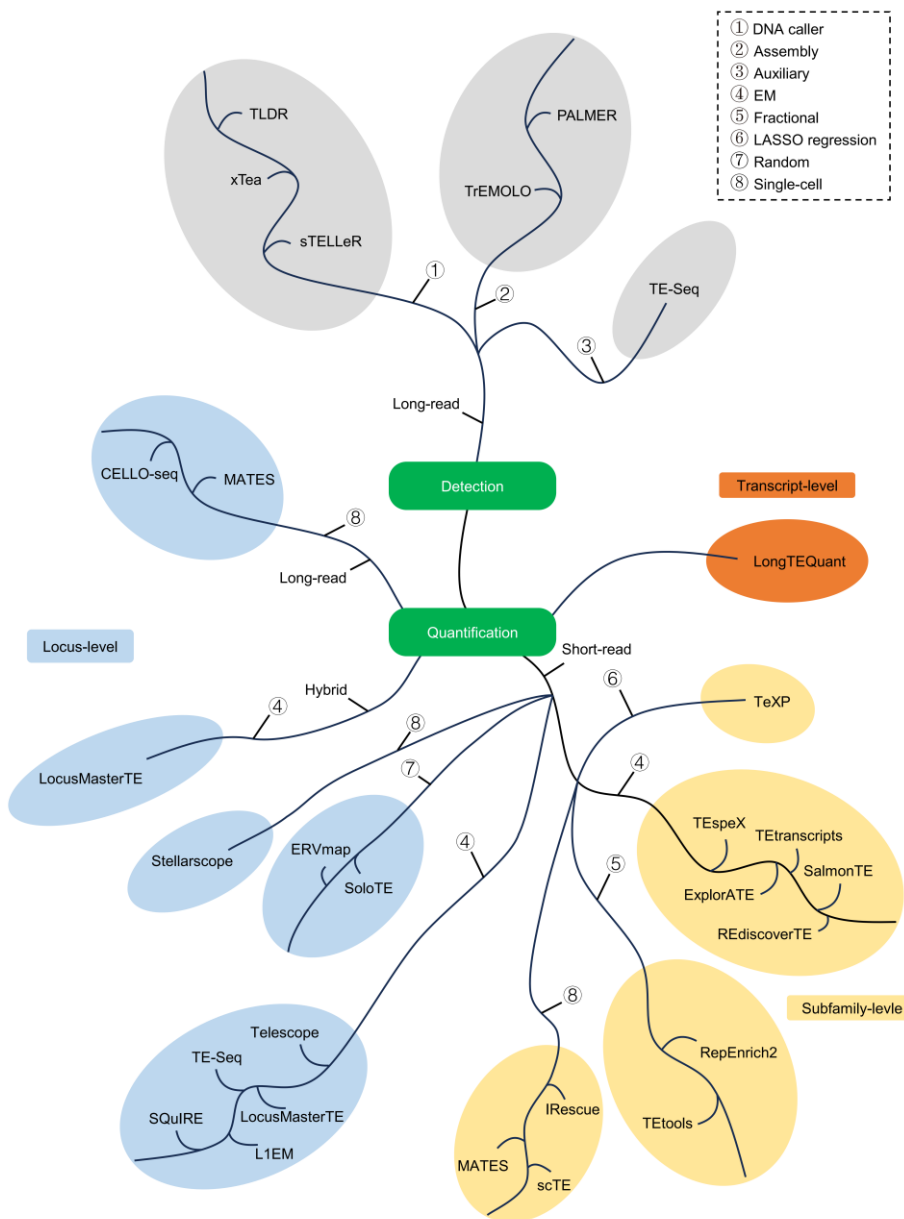


图5. 转座元件衍生转录本的特征分析



转座元件插入检测与衍生转录本定量计算方法概览

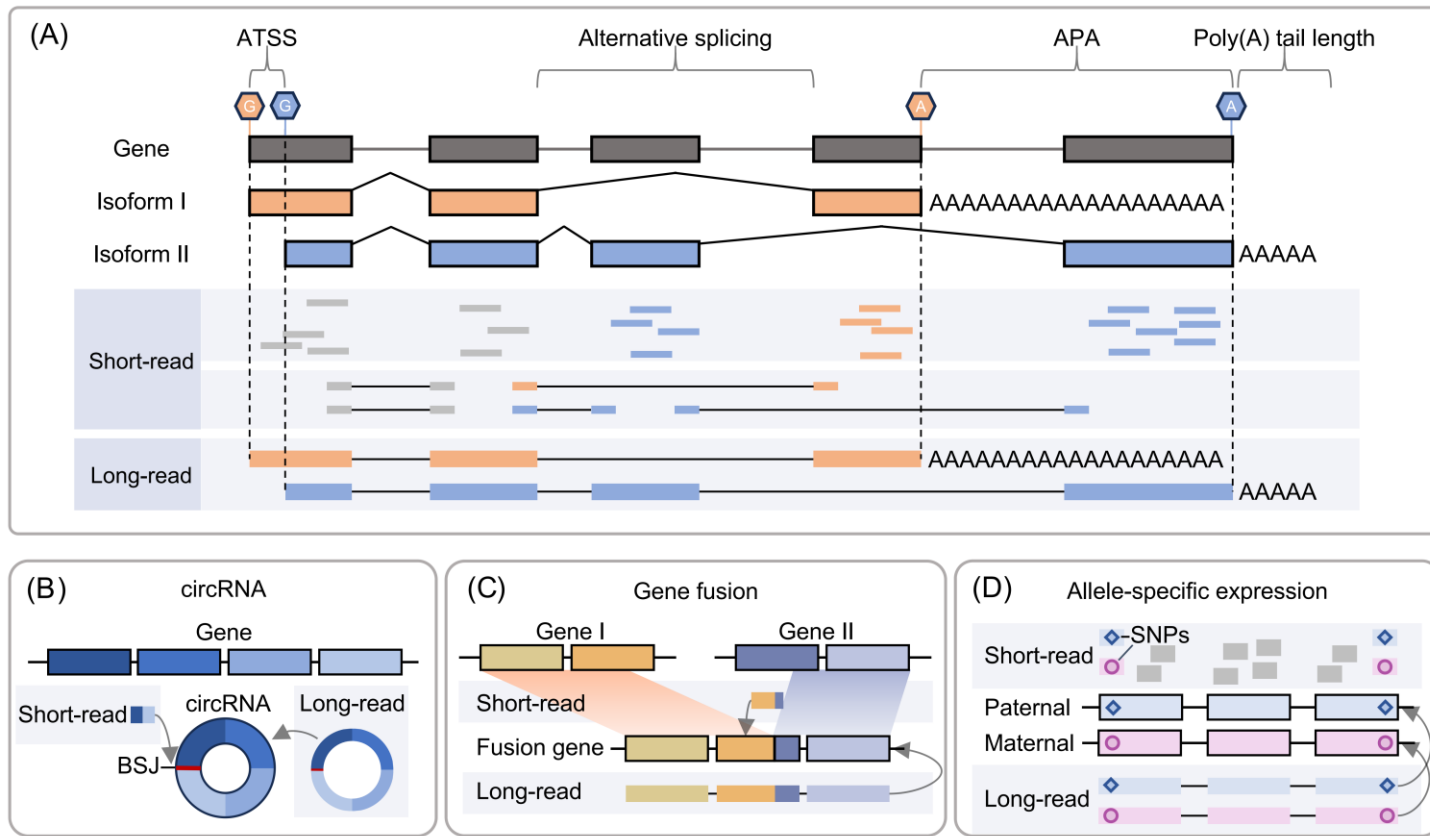


- **TE检测：** 基于长读长的方法分为（1）基于 DNA 序列比对的检测方法、（2）基于组装的方法和（3）辅助分析方法。
- **TE定量：** 方法分为短读长、长读长及混合策略，其中短读长方法进一步按亚家族水平和位点水平分辨率划分。核心策略包括（4）期望最大化（EM）、（5）比例分配、（6）LASSO 回归、（7）随机分配和（8）面向单细胞数据的方法。此外，转录本水平的分析框架利用转录本分辨率信息，直接定量 TE 来源的转录本。各模块中列出的工具名称代表相应方法类别下的典型实现。

图6. TE 插入检测与衍生转录本定量计算方法概览



长读长测序解析转录组复杂性的核心优势

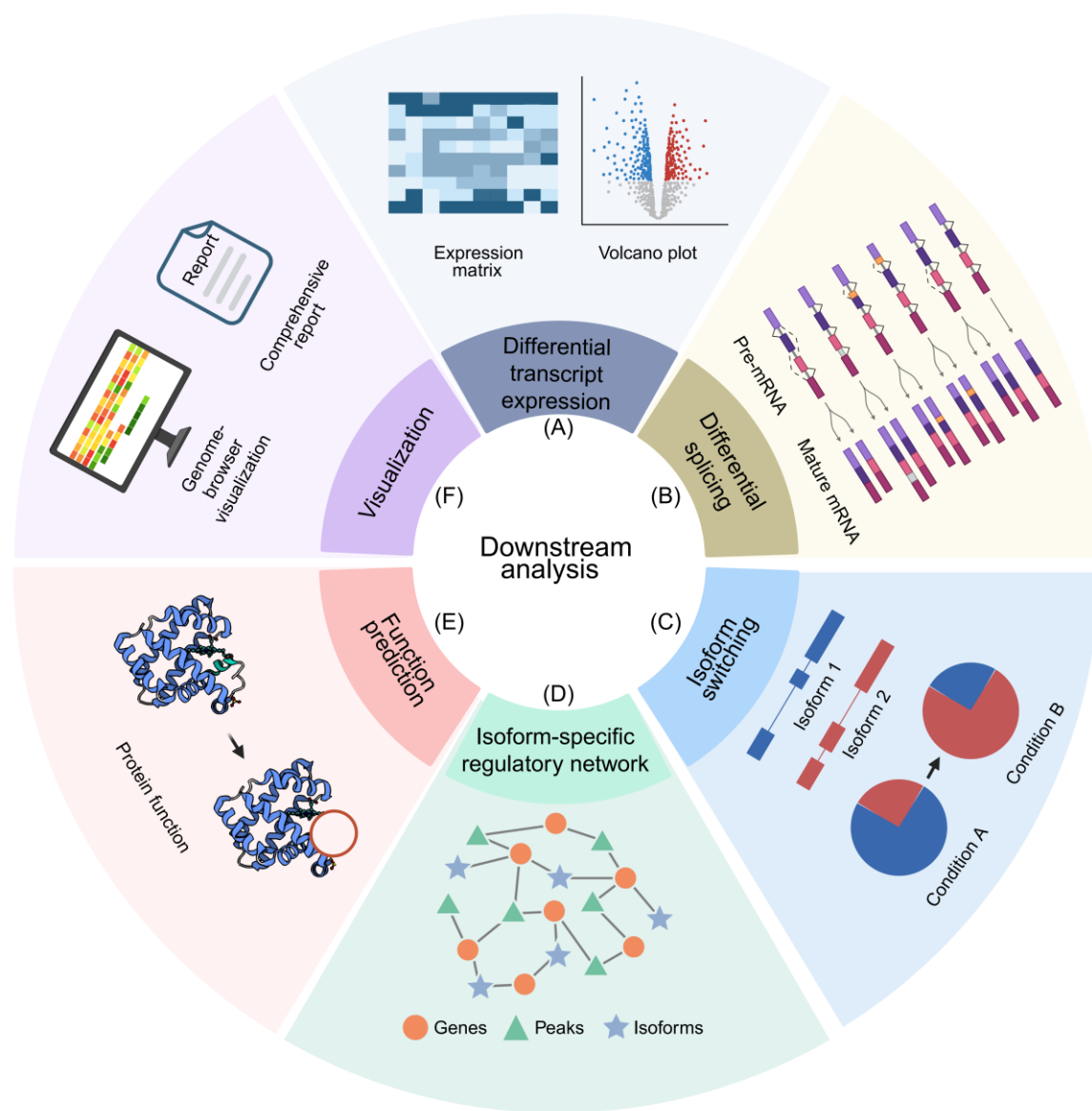


- (A) 长读长测序解析异构体结构、转录本末端及复杂剪接组合；结合适用的测序方案，还可推断poly(A) 尾长度。
- (B) 与短读长相比，长读长显著提升全长环状 RNA 结构的分辨率。
- (C) 长读长可跨越融合断点并解析完整转录本结构，补充短读长的局部片段证据。
- (D) 长读长通过连接跨转录本的多个 SNP，实现可靠的亲本等位基因分配；短读长则受限于稀疏的信息性 SNP 覆盖。

图7. 长读长测序在全长转录本异构体、环状RNA、融合转录本及等位基因分配中的解析优势



长读长测序驱动的转录组学下游分析



(A) 差异转录本表达: 基于表达矩阵, 识别不同条件下丰度显著变化的异构体。

(B) 差异剪接分析: 结合全长转录本结构, 解析与条件相关的可变剪接变化。

(C) 异构体切换: 比较同一基因内异构体的相对使用比例, 识别条件相关的使用偏好转变。

(D) 异构体调控网络: 关联异构体与候选调控因子及靶标, 探索异构体层面的调控关系。

(E) 功能注释与预测: 整合异构体功能特征, 评估其对蛋白结构域及生物学功能的潜在影响。

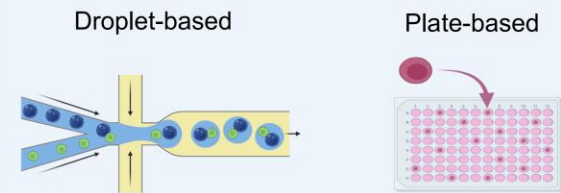
(F) 可视化与报告: 通过基因组浏览器与综合报告展示结构、表达及支持证据, 促进结果解读与交流。

图8. 长读长测序驱动的转录组学下游分析

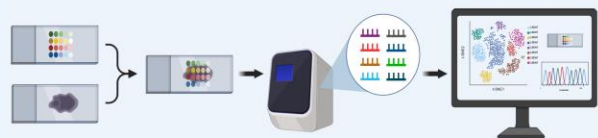


长读长转录组学研究的新兴前沿

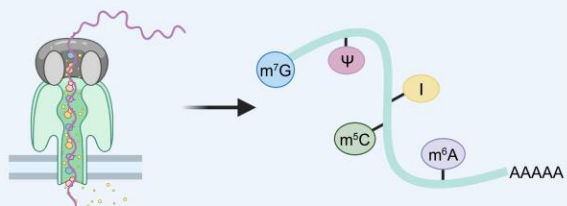
(A) Long-read single-cell transcriptomics



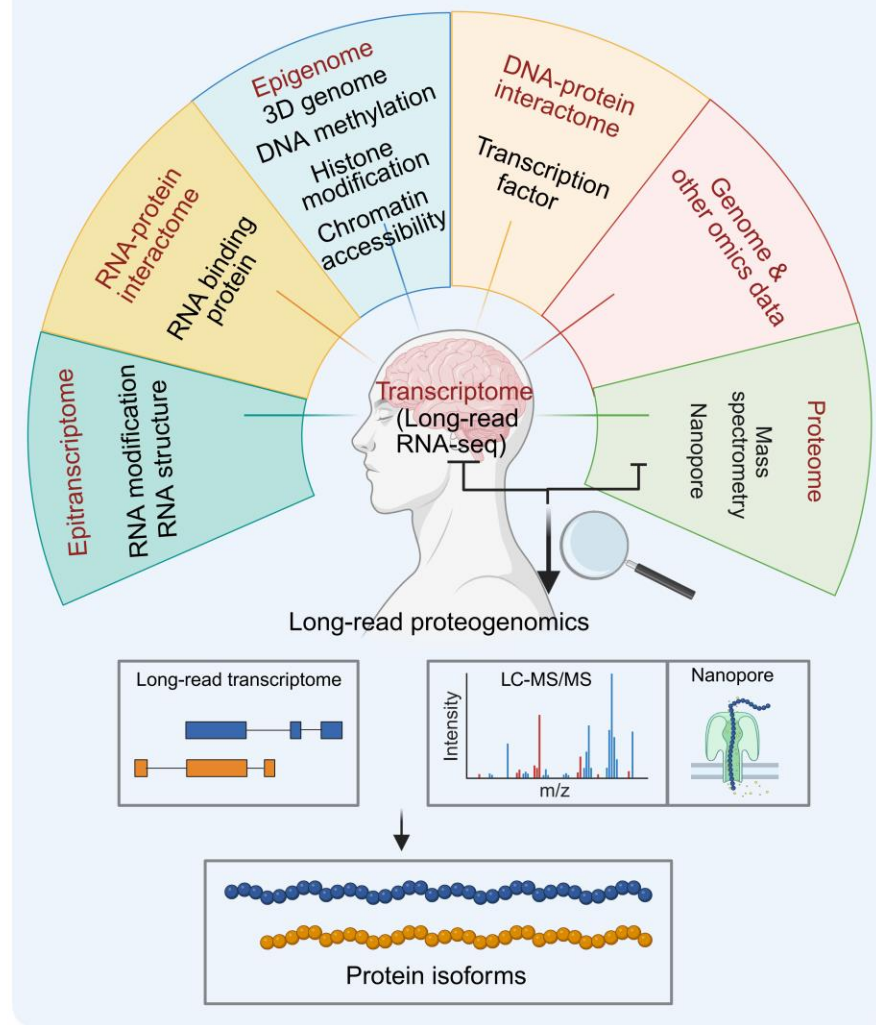
(B) Long-read spatial transcriptomics



(C) Direct detection of RNA modifications



(D) Long-read multi-omics integration



A) 长读长单细胞转录组： 在单细胞水平解析转录本异构体，揭示细胞间的转录多样性。

B) 长读长空间转录组： 关联空间位置与全长转录本结构，刻画异构体的空间分布。

C) 直接 RNA 测序： 读取原生 RNA 分子，结合信号分析识别 RNA 修饰。

D) 长读长多组学整合： 融合转录本结构与多层分子信息，探索异构体相关的调控机制。

图9. 长读长转录组学研究的新兴前沿



总结

- **从基因表达迈向异构体解析：**长读长测序揭示全长转录本、重复区域转录及关联分子特征，深化对转录组复杂性的认识。
- **从技术选择贯通生物学解读：**本综述整合平台、建库与计算分析策略，提供贯穿实验设计和下游应用的全流程指南。
- **从当前挑战展望研究前沿：**聚焦单细胞、空间、直接 RNA 修饰检测及多组学整合，展望长读长转录组学的发展方向。

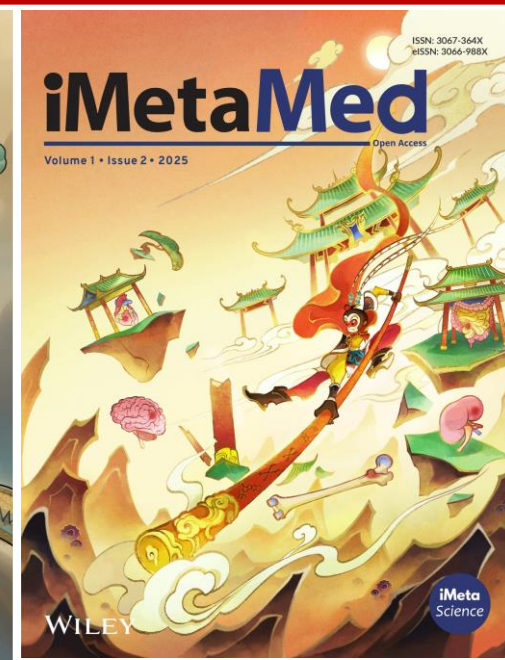
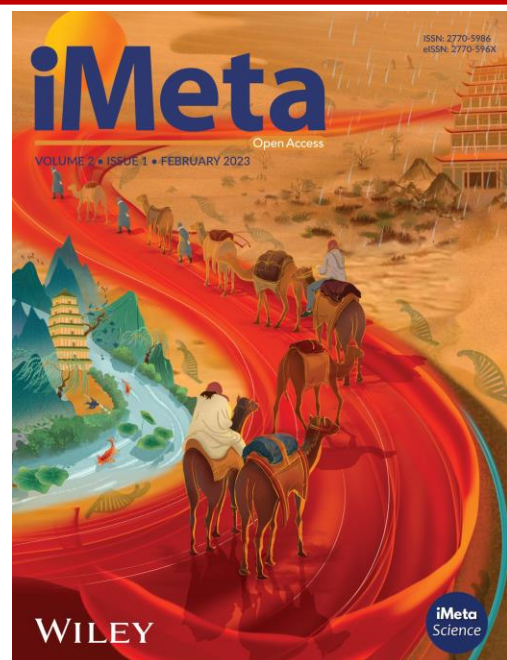
Chuwen Xu, Jia Li, Chenxi Yin, Keying Li, Tian Wang, Shiwen Gao, Yue Yu, *et al.* 2026. Deciphering Transcriptome Complexity via Long-read Sequencing. *iMeta* 5: e70171. <https://doi.org/10.1002/imt2.70171>



论文相关领域专家名单

序号	姓名	所属单位	电子邮箱
1	汤富酬	北京大学生物医学前沿创新中心	tangfuchou@pku.edu.cn
2	付向东	西湖大学生命科学学院	fuxiangdong@westlake.edu.cn
3	陈玲玲	中国科学院分子细胞科学卓越创新中心	linglingchen@sibcb.ac.cn
4	伊成器	北京大学生命科学学院	chengqi.yi@pku.edu.cn
5	杨运桂	中国科学院北京基因组研究所	ygyang@big.ac.cn
6	赵方庆	中国科学院动物研究所	zhfq@ioz.ac.cn
7	叶凯	西安交通大学	kaiye@xjtu.edu.cn
8	郭国骥	浙江大学医学院	ggj@zju.edu.cn
9	杨建华	中山大学生命科学学院	yangjh7@mail.sysu.edu.cn
10	高歌	北京大学生物医学前沿创新中心	gaog@pku.edu.cn
11	肖传乐	中山大学	xiaochuanle@126.com
12	高远	中国科学院北京基因组研究所 (国家生物信息中心)	gaoy@big.ac.cn
13	肖传乐	中山大学	xiaochuanle@126.com
14	Heng Li	Harvard Medical School	hli@jimmy.harvard.edu
15	Kin Fai Au	University of Michigan, Ann Arbor	kinfai@med.umich.edu

序号	姓名	所属单位	电子邮箱
16	Nancy R. Zhang	University of Pennsylvania	nzh@wharton.upenn.edu
17	Ting Wang	Washington University in St. Louis	twang@wustl.edu
18	Ana Conesa	Spanish National Research Council	ana.conesa@csic.es
19	Yi Xing	University of Pennsylvania	xingyi@chop.edu
20	Hagen Tilgner	Weill Cornell Medicine in New York City	hut2006@med.cornell.edu
21	Ali Mortazavi	University of California, Irvine	ali.mortazavi@uci.edu
22	Angela N. Brooks	University of California, Santa Cruz	anbrooks@ucsc.edu
23	Jonathan Göke	Genome Institute of Singapore	gokej@gis.a-star.edu.sg
24	Steven L. Salzberg	Johns Hopkins University	salzberg@jhu.edu
25	Lior Pachter	California Institute of Technology	lpachter@caltech.edu



iMeta(宏)期刊由宏科学创办，对标**Cell**的生物/医学综合期刊，**SCIE**、**PubMed**收录，影响因子(IF)44.4，位列全球第47，中国第4，**分区表生物学1区Top**，外审平均21天，投稿至发表中位数87天(比同水平期刊快3倍)，欢迎高影响力的研究、方法和综述投稿，CNS外审稿件可带意见和回复转投。

iMetaOmics (宏组学)，定位IF>15对标**NC/SA**的生物/医学综合期刊，已被**ESCI**、**PubMed**等收录。

iMetaMed (宏医学)定位IF>15的医学综合期刊，欢迎投稿！



主页: <http://www.imeta.science>

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



iMeta: <https://wiley.atyponrex.com/journal/IMT2>

投稿: iMetaOmics: <https://wiley.atyponrex.com/journal/IMO2>

iMetaMed: <https://wiley.atyponrex.com/journal/IMM3>



office@imeta.science

imetaomics@imeta.science



宣传片



[iMeta](https://www.imeta.science)



更新日期
2026/6/17