# 沙地云杉体细胞胚胎发生全长转录组分析

代金玲，闫艳秋，曹佳璐，白玉娥*

（内蒙古农业大学林学院，内蒙古 呼和浩特，100019）

**摘要：**沙地云杉（*Picea mongolica*）是松科云杉属常绿乔木，具有耐沙埋、耐寒、耐旱等特性，为我国特有珍稀树种。由于天然林更新慢，种子收获量小，病虫害严重等原因，严重影响沙地云杉种群的稳定性和遗传多样性，阻碍了沙地云杉的推广应用。本研究以课题组建立的沙地云杉体细胞胚胎发生体系为基础，以胚性细胞系 A23 为材料，分别收集胚性愈伤组织、球形体细胞胚，晚期体细胞胚和成熟体细胞胚，利用 SMRT 测序技术进行全长转录组测序，并将获得的转录本在 NR、SwissProt、KEGG 和 KOG 四大数据库中进行功能注释，对 AS、SSR、LncRNA、TF 等基因结构进行分析。SMAT 测序最终获得的数据量为 33.88 G，沙地云杉全长转录本序列 12232 条，平均长度 2331.98 bp，N50 为 2486 bp，Reads 长度主要分布在 1000-4000 bp 之间。全长转录本序列在四大数据库中共有 12092 条被成功注释。其中 NR 数据库注释到 12092 条、SwissProt 数据库注释到 10751 条、KEGG 数据库注释到 11895 条、KOG 数据库注释到 8444 条，未被注释的全长转录本有 140 条。比对发现最相似物种为加西云杉。通过结构分析发现 AS 事件 83 个，SSR 位点 108 个，LncRNA35 条和转录因子家族 46 个。因沙地云杉无公布的基因组序列，研究中常以挪威云杉（*P. abies*）基因组为参考，但由于裸子植物基因组庞大，二代测序读长较短，导致组装后质量难以满足目前研究需求。三代全长转录组采用单分子实时测序技术，具有精度高、读长长、能够检测碱基修饰等优点。因可获得全长基因和基因结构，在研究无基因组物种时有很大优势。本研究获得的沙地云杉全长转录序列有助于弥补缺少基因组序列对相关分子生物学研究的影响，为沙地云杉体细胞胚胎发生关键调控基因挖掘及机制研究、高效稳定体胚发生体系的建立奠定基础。

**关键词：**沙地云杉；体细胞胚胎发生；全长转录组；功能注释；结构分析