



## 北京基因组所（国家生物信息中心）发布更新版表观基因组关联研究开放平台EWAS Open Platform

作者： 发布时间：2022-01-05 | 【大 中 小】 | 【打印】 【关闭】



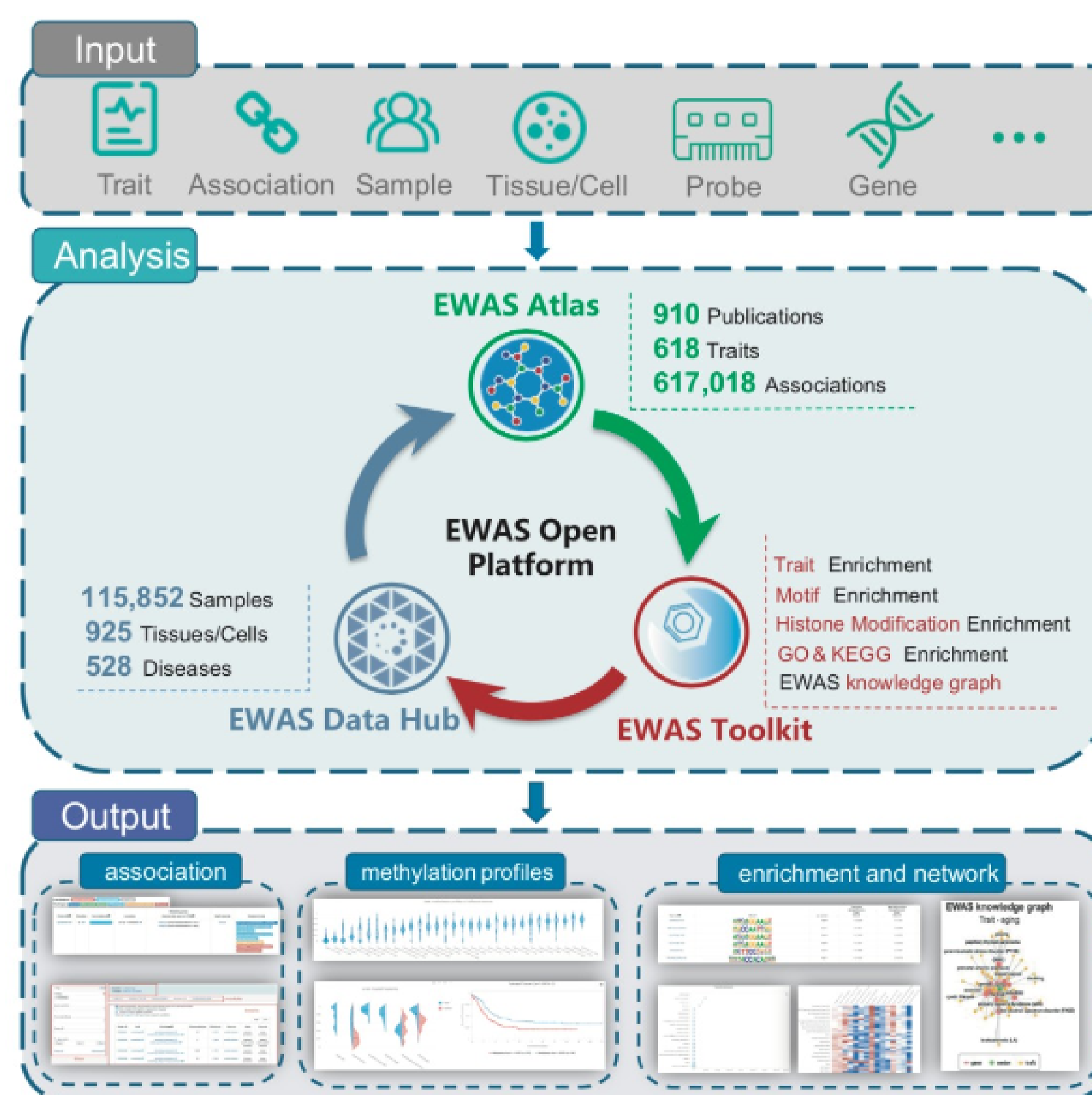
近日，由中国科学院北京基因组研究所（国家生物信息中心）国家基因组科学数据中心（NGDC）开发的表观基因组关联研究资源开放平台EWAS Open Platform正式上线。该研究成果以“EWAS Open Platform: integrated data, knowledge and toolkit for epigenome-wide association study”为题在国际学术期刊*Nucleic Acids Research* 在线发表。

随着表观基因组关联研究（EWAS）的爆炸式增长，发表了大量EWAS学术论文，积累了海量EWAS相关的数据。对这些数据进行标准化整合，并从已发表论文中提取和挖掘表观关联知识，对于系统的表征和研究不同实验条件下的甲基化状态、探索与各种性状相关的表观遗传分子机制具有重要意义。NGDC在2019年和2020年先后开发了基于高质量的人工审编EWAS知识库（EWAS Atlas）和存储了海量标准化的DNA甲基化芯片数据的EWAS数据库（EWAS Data Hub），得到了业内的广泛使用和同行的高度评价。

为了提供从数据浏览与下载、在线分析与可视化到知识解释与验证的全面系统的资源和服务，NGDC研究团队在不断整合和更新中心已有EWAS资源基础上，构建了表观基因组关联研究资源开放平台（EWAS Open Platform）。EWAS Open Platform包括标准化的数据信息库（EWAS Data Hub）、人工信息提取的知识库（EWAS Atlas）和表观-特征关联在线工具（EWAS Toolkit）三部分。EWAS Data Hub整合了115,852个样本的DNA甲基化芯片数据和对应的元数据，并统一采用GMQN方法进行标准化。同时，EWAS Data Hub利用海量高质量DNA甲基化芯片数据和标准化元数据的优势，为485,512个探针和36,397个基因提供了一系列重要的评估值（包括组织特异性、年龄相关性、性别差异和种族特异性）和不同背景下的参考DNA甲基化图谱；EWAS Atlas共整合了910篇文献中报道的617,018个高质量的甲基化与表型关联，涉及到618种表型和3,385个队列；EWAS Toolkit利用EWAS Atlas和EWAS Data Hub提供的高质量甲基化与表型关联知识和标准化的DNA甲基化芯片数据，为用户提供多种在线分析和可视化工具，包括富集分析、注释、知识图谱可视化等。

北京基因组所（国家生物信息中心）博士研究生熊壮、杨飞以及博士毕业生李萌伟为本文共同第一作者，鲍一明研究员、章张研究员及李茹姣高级工程师为共同通讯作者。该研究得到了中科院战略性先导科技专项、国家重点研发计划、中科院关键技术人才等项目资助。

[论文链接](#)



EWAS Open Platform数据处理流程示意图

