**题目**

培养及非培养宏基因组技术揭示微生物群落降解复杂微生物聚合物的遗传和酶活潜力

**翻译：张琳琦 赵会成**

**期刊：Microbiome**

**IF：10.465**

**发表时间：2020.06.01**

**第一作者：Ohana Y.A.Costa**

**通信作者：刘彬彬、Eiko E.Kuramae**

**通讯作者单位：中科院遗传所农业资源研究中心，荷兰皇家生态研究所(NIOO-KNAW)**

**摘要**

目前不依赖于微生物纯培养的方法，包括宏基因组学，是探索和发现自然环境中微生物产生的生物技术化合物的工具。糖苷水解酶类(GHs)在工业化产品和生物燃料的生产以及去除生物膜(biofilms)和胞外多聚物(EPS)的工业流程中十分重要。生物膜和胞外多糖成分复杂，需要种类繁多的酶才能完全降解。本研究利用Acidobacteria Granulicellas菌株WH15的胞外多糖复合体结构(WH15EPS)作为富集因子，使用不依赖于微生物纯培养和依赖于微生物纯培养的方法，鉴定潜在的GH微生物产生菌和具有生物技术应用潜力的GH基因。作者使用稳定性同位素标记技术(SIP)结合宏基因组技术研究了使用WH15EPS改良的表层土壤，并通过宏基因组数据与固体纯培养-EPS改良培养基进行耦合研究。

**引言**

宏基因组学方法能够解析依赖于传统微生物纯培养技术无法得到的微生物基因库。因此，长期以来不依赖于纯培养的方法一直被用作探究和发现自然环境中微生物产生的生物技术化合物的工具，特别是检测潜在的酶和其他具有经济意义的生物技术产品。不依赖于微生物纯培养的方法可以解析潜在的微生物作用；然而，仍然需要基于纯培养的研究来理解微生物的特征和表型。宏基因组学的使用促进了工业生产系统和生物酶的研究，特别是在动物肠道中的应用，但是在其他类型的生态系统，如森林凋落物，仍未得到充分利用。

糖苷水解酶类(GHS)是宏基因组学中广泛研究的工业应用酶之一，因为它们在食品生产和其他工业生产活动中十分重要。这些酶被用于酿造、烘焙、糖浆生产、食品加工、调味以及乳制品和发酵食品的生产。微生物通过将纤维素和木质生物质转化为可以被发酵成生物乙醇的糖。

GHs的另一个应用是降解多糖以去除生物膜。胞外多糖是胞外多聚物(EPS)的主要成分，也是目前研究最多的成分，胞外多聚物是由多种微生物菌株合成的生物聚合物。EPS是保存生物膜三维结构、维持内部凝聚力和促进表面粘附的成分。一般来说，去除生物膜对人类健康十分重要，因为这些结构与多种疾病相关联，在医院和食品加工业中会造成许多问题。此外，生物酶去除生物膜优于传统的清洗剂，因为传统的清洗剂不环保，会产生有毒残留物，还会腐蚀设备。酶类是一种环境友好的替代品，因为它们具有可生物降解的特性。EPS和生物膜十分复杂，需要多种酶参与才能形成；酶类如溶菌酶、淀粉酶、β-N-乙酰氨基葡萄糖苷酶和藻酸盐裂解酶已经用于食品和制药中生物膜的去除或抑制。目前超过50%的工业用酶是由微生物产生的，例如芽孢杆菌和曲霉属真菌，大约15%来自植物。此外，具有应用潜力的微生物酶是从诸如火山热液口、北极苔原，牛瘤胃及百白蚁内脏中获得。

本研究的主要目的是利用温带森林凋落物微宇宙培养和培养基纯培养实验，来研究靶向降解EPS的微生物和功能。植物凋落物主要由难降解的生物聚合物组成，它们是微生物群落的碳、能量和养分的来源。纤维素、半纤维素和果胶是植物细胞壁的主要成分。纤维素是最丰富的植物细胞壁成分，由右旋葡萄糖组成。半纤维素主要由木聚糖、β-葡聚糖、甘露聚糖和其他低聚糖组成。果胶含有高半乳糖醛酸、木糖半乳糖醛酸和鼠李半乳糖醛酸。由于其复杂性，植物细胞壁成分需要由微生物产生的多种酶在凋落物分解过程中降解。因此，这是一个探究复杂的多糖降解酶的特殊环境。另一方面，森林生态系统中的微生物群落主要由酸杆菌门组成，该门成员通常与碳降解有关。酸杆菌中的*Granulicella*sp.可以产生大量的EPS。*Granulicella*属不是人类病原体，其胞外多糖的独特组成对于探究复杂多样的糖苷水解酶基因十分重要，这些基因可被应用于工业生产活动中。*AcidobacteriaGranulicella*sp.菌株WH15(WH15EPS)的EPS组成上比大多数商业上可用的微生物聚合物都要复杂，它由7种单糖(甘露糖、葡萄糖、半乳糖、木糖、鼠李糖、葡萄糖醛酸和半乳糖醛酸)组成，而其他已知的EPS最多由4种不同的单糖组成。与其他EPS相比，WH15EPS的降解需要更复杂的酶；因此，将WH15EPS添加到表层土壤样品能够富集更多加样化的GHs。稳定性同位素标记技术(SIP)能够通过超高速离心分离物中的重层组分有效鉴定被标记基质的利用情况，例如13C、18O和15N能够进入代谢特定底物的微生物的细胞成分。因此，SIP能够鉴定参与特定标记基质代谢的活性微生物。目前已被成功地应用于研究含有多种化合物的微生物，如甲醇、酚类和其他物质。

**材料和方法**

**土壤样品**

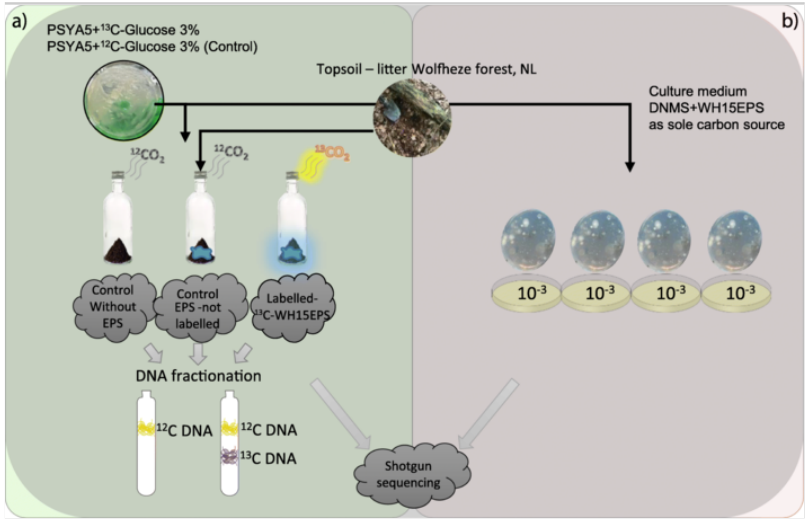
于2017年春季从荷兰的Wolfheze森林采集了四个表层土壤凋落物样本。样本取自倒下树干附近的表层土壤(0-5厘米)。将收集的样品混合过筛，并立即用于SIP培养，并添加来自于WH15EPS的EPS。同时测定表层土壤-凋落物样品的理化性质。实验流程如图7所示。

图7：实验设计流程图。A.在PSYL5培养基中分别添加13C-葡萄糖和12C-葡萄糖用于*Granulicella*spWH15.生产及纯化，13C-和12C-WH15EPS用于培养试验，对照处理不添加WH15EPS；每个处理六个重复。经过35天的培养并测定CO2呼吸测量，并提取DNA进行超高速离心。13C-WH15EPS培养的重层DNA及12C-WH15EPS的DNA和对照进行高通量测序。B.与此同时，纯化的12C-WH15EPS用于纯培养培养基DNMS的碳源。将森林凋落物表层土壤样品稀释至10-3并接种，室温培养30天。每个平板2个重复。接下来，从培养板上刮取细胞；提取总DNA并进行高通量测序。

**SIP宏基因组**

**[13C]标记和未标记的EPS**

*Granulicella*spWH15.使用含有3%(重量/体积)的PSY5固体培养基，分别包括13C标记的葡萄糖作为唯一碳源或未标记的葡萄糖作为碳源的对照实验。在20℃环境下培养30天后，提取EPS并纯化。向每个样品中加入60微升36.5%的甲醛，在4℃下培养1小时。接下来，加入4毫升1M氢氧化钠，并在4℃培养3小时。在9000转离心40分钟后，在室温下过滤除去上清液中的细胞碎片，单糖在4℃温度下通过透析管去除，持续48小时。并测定EPS溶液中的DNA及蛋白质的浓度。总碳水化合物含量使用改良的苯酚-硫酸法估算。将多糖溶液在−80℃下冷冻干燥72h，为后续实验做准备。纯化后的胞外多糖含碳水化合物约400 mg/ml，蛋白质约1%，未检出DNA。

**稳定性同位素标记实验**

在接种表土凋落物之前，使用1毫升的灭菌水对冻干的多糖进行稀释以形成均匀的分布。将含有0.05%的WH15EPS(标记和未标记对照)及不含EPS的5克表土凋落物样本加入120毫升的瓶子，用丁基橡胶塞密封，在室温黑暗条件下培养。为了维持有氧条件并防止13CO2交叉利用，所有的瓶子每4天打开盖子一次。微生物群落对WH15EPS的利用通过气相色谱法检测CO2。并分析标记实验中CO2的13C/12C比值。

**DNA提取和分离**

提取不同处理中的土壤DNA并通过分光光度计进行定量。SIP超高速分离过程具体为：将两微克的DNA与氯化铯溶液和梯度缓冲液混合后，置于超速离心管中以125,395×g在真空下20℃下超高速离心65小时。高速离心后进行DNA梯度分离，用折光仪测量浮力密度确定重层及轻层样品。DNA从氯化铯溶液中沉淀并使用聚乙二醇溶液和乙醇纯化及洗脱，并进行文库制备和高通量鸟枪法测序。

**WH15EPS为唯一碳源的培养基中培养微生物的宏基因组**

为了研究能够在WH15EPS为唯一碳源的培养基中生长的微生物的宏基因组，将10g新鲜表土凋落物样品与100ml的MES缓冲液混合，在室温下涡旋30分钟。使用无菌MES缓冲液制备梯度稀释液。并将冷冻干燥纯化的WH15EPS用灭菌水混合过滤灭菌后加入培养基中。接种土壤悬浮液的平板在室温下培养1个月。选择稀释度为10-3的样品进行测序。培养后刮取菌落，用试剂盒提取总DNA及宏基因组测序。

**宏基因组数据的生物信息学和统计分析**

**SIP宏基因组**

使用EBIMGnify和SqueezeMeta流程处理SIP宏基因组序列。首先对序列进行质量控制，使用Megahit对每个样品分别进行组装；Prodigal用于ORF预测,barrnap用于rRNA基因序列检测，使用RDP分类器进行物种分类。Diamond比对Genbanknr数据库对ORFs进行分类，并使用eggNOG数据库进行功能注释获得KO和COG编号。eggNOG-mapper用于碳水化合物酶类注释。SqueezeMeta脚本SQM2tables.py计算基因相对丰度。处理之间差异分析使用参数和非参数检验进行评估，包括方差分析、KruskalWallis和Tukey的HSD检验。通过多变量方差分析(PERMANOVA)检验组间的差异。并使用“随机森林”算法执行特征选择。

**纯培养微生物的宏基因组分析**

对纯培养微生物的DNA进行宏基因组测序，并用EBIMGnify和ATLAS流程进行数据处理。对于ATLAS，用BBDuk2进行质量控制，用MegaHit进行混合组装；在ORF水平上对组装的contigs进行功能和分类分析。Prodigal用于ORF预测，eggNOG数据库进行功能注释获得KO和COG编号。eggNOG-mapper用于碳水化合物酶类注释。Kaiju用于NCBIRefSeq数据库的ORF分类分配。使用COCORT、MAXBIN和Metabat将质量控制后的contigs>1000的序列用于分箱。使用CheckM检查组装基因组的完整性和污染，并进行分类分配。Prodigal预测基因组的ORF，eggNOG统计COG和KO号。eggNOG-mapper用于碳水化合物酶类注释。

**结果**

**SIP宏基因组**

经过质量控制过滤后，共得到了18,762,958条序列，平均每个样本1,563,580条序列，预测出1,209,745个具有功能的ORFs。

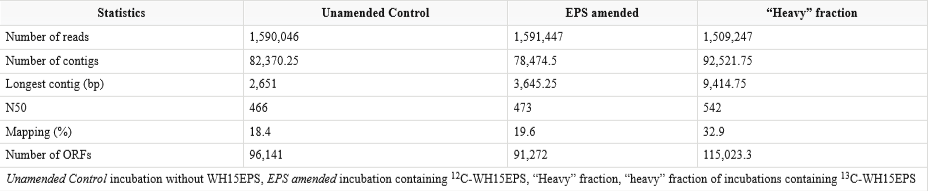


表1：SIP实验的宏基因组数据，数据为四个重复的平均值。

**基于16s与宏基因组数据SIP实验的物种组成**

基于SSUrRNA注释的分类注释表明细菌、真菌和古菌约占序列的84%、4%和2%,在门水平上，在所有样品中观察到17个细菌群，5个真菌群和3个古菌群。*Proteobacteria*在所有处理中是丰度最高的，约占序列数的26.4–28%,其次是*Actinobacteria*，约占序列数的14.5–17.5%。在对照和12C标记样品中*Acidobacteria*是丰度第三高的门，约占序列数的14.5–15.8%。在13C标记样品中,*Planctomycetes*是丰度第三高的门，约占序列数的16.45%。a)在属水平上，我们在所有样本中观察到167个属，其中110个属是未分类的。在对照组中中，“未分类细菌”是丰度最高的属。在未标记样品中，“未分类*Acidobacteriaceae*”是丰度最高的属.在添加13C标记的样品中,“未分类*Planctomycetes*”是丰度最高的属。在丰度最高的前十个属中，只观察到2个已知的属：*Acidothermus*和*Singulisphaera*。b)在属水平上，我们发现了1541个属，其中667个属未分类。两个对照处理中丰度最高的三个属是“未分类微生物”，“未分类*Bacteria*”(12.7–16%oftheORFs),and“未分类*Acidobacteriaceae*”，而在添加13C标记的样品中的丰度最高的属为“未分类微生物”，“未分类*Bacteria*”和“未分类*Planctomycetes*”。

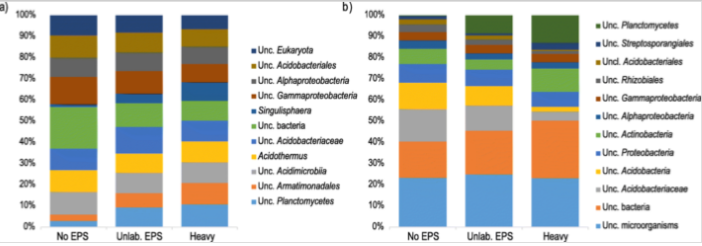


图1：基于SIP处理的属级微生物种群组成及相对丰度：a)SSUrRNA基因分类b)ORF分类。SIP实验分为三个处理：NoEPS:没有添加EPS的样品；Unlab.EPS:含有12C和EPS的样品；Heavy:含有13C和EPS的样品。

PERMANOVA(p值<0.001)表明，对于SSUrRNA数据和基于ORF数据的分析，两种处理之间的微生物群落不同，但两种对照处理相似。在PCOA图中，两种对照处理与13C标记样本分离。对于SSUrRNA群落，PCOA的前两个轴解释了43.3%的变异，而对于ORF的数据，解释了90.6%的变异。两个数据集的RDA分析(p=0.002)主要差异的组是*Planctomycetes*，如“未分类*Planctomycetes*”，“未分类*Planctomycetales*”，“未分类*Planctomycetia*”和*Singulisphaera*造成13C-EPS标记样品与两种对照组之间的微生物群落差异，与13C标记样品中含有较高丰度的“未分类*Planctomycetes*”结果一致。

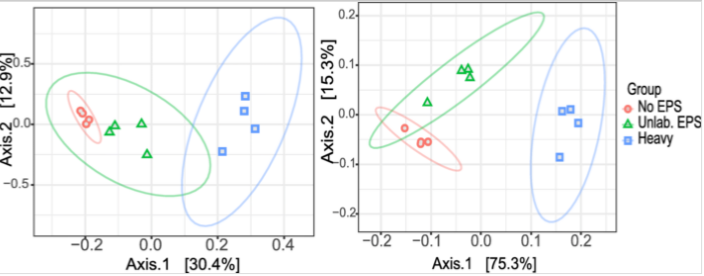


图2：基于Bray-Curtis距离的主坐标分析：a)SSUrRNA基因分类b)ORF分类。

**SIP宏基因组的功能分析**

用KEGG、COG和CAZy数据库进行功能基因注释，探讨微生物群落的功能特征。大约60%的ORFs由COGs注释，注释出20,644个COGs，占总ORFs的10.8–11.6%。用Boruta算法进行“随机森林”分析(p<0.05)，识别在处理之间显著分离的注释。a）采用Boruta算法共选取32个COGs.这其中13个在对照样品中丰度更高，19个在标记样品中丰度更高。大多数COG属于功能未知。b）采用Boruta算法共选取40个KEGGs.这其中14个KEGGs在对照样品中丰度更高，其中8个KEGGs与代谢途径有关；26个KEGGs在标记样品中丰度更高，其中13个KEGGs主要与“代谢途径”和“微生物在不同环境中的代谢”有关。c）Boruta特征选择确定了27个影响样品处理分散的CAZY家族，其中绝大多数属于糖苷水解酶类(GH).15个在标记样品中丰度更高，12个在对照样品中丰度更高.在标记样品中丰度高的种群涉及木聚糖和木聚糖模块、木聚糖酶、甘露糖转移酶和琼脂酶，而在对照样品中丰度高的种群为α和β-半乳糖酶和葡萄糖苷酶。PERMANOVA(p值<0.001)分析表明，对于KEGG、COG和dbCAN数据库，不同处理间的功能基因组成不同与但注释分类相似，对照处理相似，两种对照处理与13C标记样本分离。

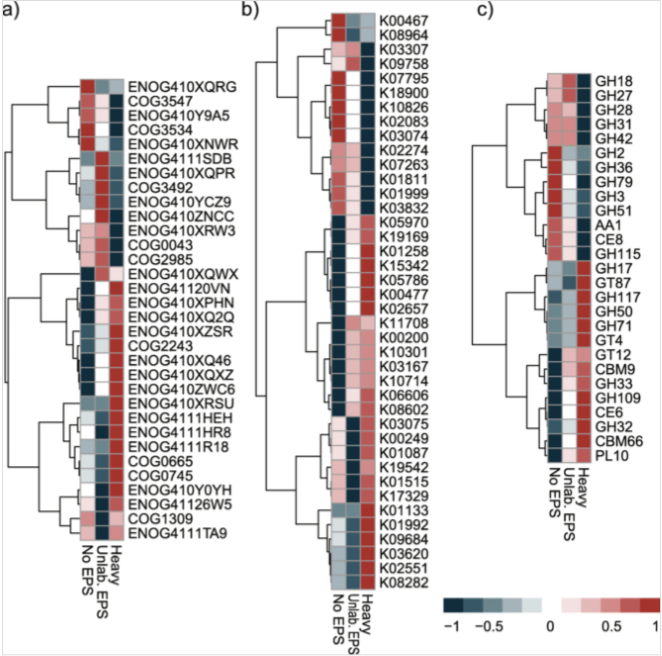


图3：Boruta随机森林选择基于1000排列的不同处理间显著分离的特征：a）COG注释；b）KEGG注释；c）dbCAN注释。

**纯培养微生物的宏基因组分析**

纯培养微生物的宏基因组数据经过质量过滤后，共获得422,735,048条序列，平均80%的ORFs可通过KEGG和COG数据库注释。

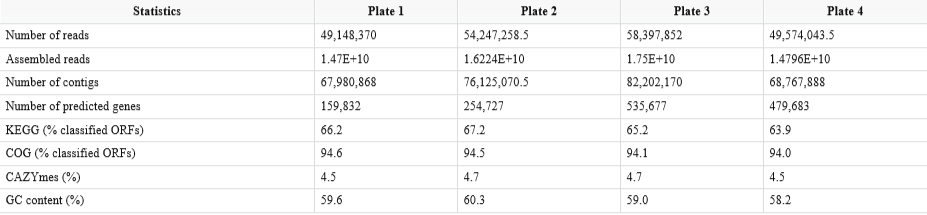


表2：纯培养实验每个板的宏基因组数据，数据为两个重复的平均值。

**基于16s与宏基因组数据纯培养微生物实验的物种组成**

基于SSUrRNA的分类组成分析表明，平均73%的序列属于细菌，20%属于真菌，7%来自真核生物。在门水平，共鉴定出17个细菌群，7个真菌群，14个真核生物群。丰度最高的的细菌门是*Proteobacteria*，占序列数的47.9%。其次最高的是真菌门*Ascomycota*，占序列数的14.5%。a)在属水平，共观察到450个属，其中丰度最高的种群为细菌.优势种群为“未分类Bacteria”和*Dyella*、*Silvimonas*和*Burkholderia*也是丰度最高的10个属之一。b)同样，对于基于ORF的数据，属级丰度最高的种群是细菌，揭示了在属级存在1930个种群。“未分类microbes”就是丰度最高的种群，其次是*Caballeronia*和*Paraburkholderia*.其他属，如*Burkholderia*、*Rhodanobacter*和*Dyella*也属于优势种群。

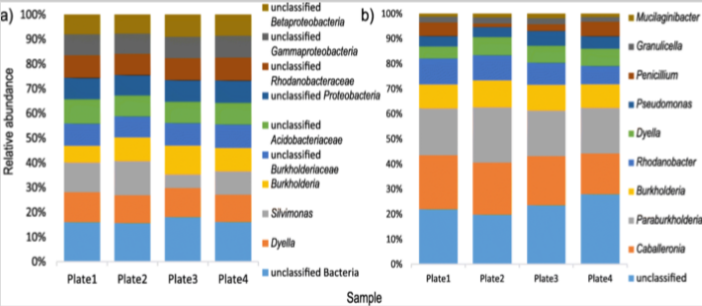


图4：纯培养微生物的属级微生物种群组成及相对丰度：a)SSUrRNA基因分类;b)ORF分类。

**纯培养微生物宏基因组功能分析**

通过KEGG、COG和dbCAN数据库进行注释，探讨微生物培养宏基因组功能概况。COG分析表明，约20.6%的COGs功能未知。a）在COG注释结果中，与SIP宏基因组结果相似，主要类别包括“E-氨基酸的转运和代谢”、“G-碳水化合物的运输和代谢”和“C-能源生产和转换”。b）20个丰度最高的KEGGs属于“基因信息处理”，“新陈代谢”和“信号和细胞过程”等类别。其中13个KEGGs被注释为转运蛋白。c）用dbCAN对碳水化合物活性酶进行分析，发现有298个CAZyme家族。23个家族占主导地位，丰度在1%以上，其中包含2个AA家族(占CAZymes的7.75%)、1个CBM、4个CE、10个GH和6个GT。最丰富的是GT-41，其中包括UDP-GlcNAc（β-N-乙酰氨基葡萄糖转移酶）和UDP-GLC（N-β-葡萄糖转移酶），属于参与蛋白质糖基化的酶。在GH家族中，最丰富的是GH13。

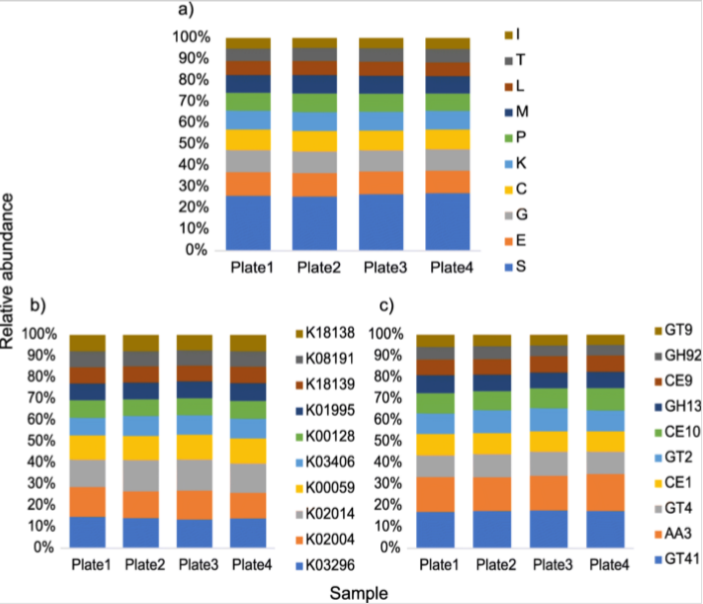


图5：微生物培养样品丰度前十的功能类别的相对丰度分布。a）COG注释b）KEGG注释.c）dbCAN注释。（E-氨基酸的转运和代谢;G-碳水化合物的运输和代谢;C-能源生产和转换;I-脂质转运和代谢;M-细胞壁/膜/包膜生物发生;P-无机离子的转运和代谢;T-信号转导机制;L-复制、重组和修复;K-转录;S-功能未知）

在这两个宏基因组数据集中共发现127个GH家族，有114个家族在这两个数据集中都存在，而5个家族仅存在于SIP数据集(GH112、GH48、GH52、GH86、GH98)，8个家族仅存在于纯培养微生物数据集中(GH111、GH131、GH132、GH134、GH45、GH7、GH80、GH85)。

在两个宏基因组数据集中GH13丰度最高，纯培养微生物数据集中的GH13大多数序列来源于*Proteobacteria*(占占GH序列的66.8%)和*Acidobacteria*(占占GH序列的21.8%)；而在SIP数据集中，GH13最丰富的门是*Actinobacteria*(占占GH序列的20.4%-45.7%)、*Acidobacteria*（占占GH序列的4-24.7%）和其他门(占占GH序列的27-34%)。

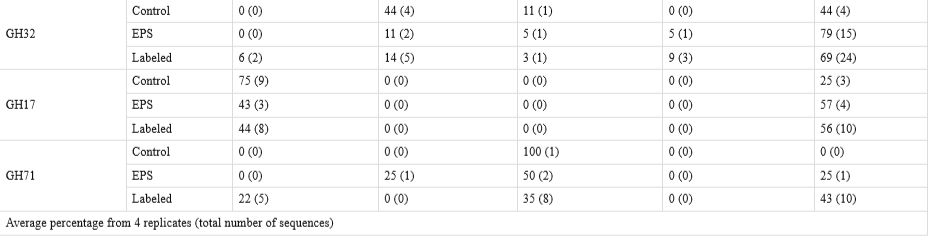
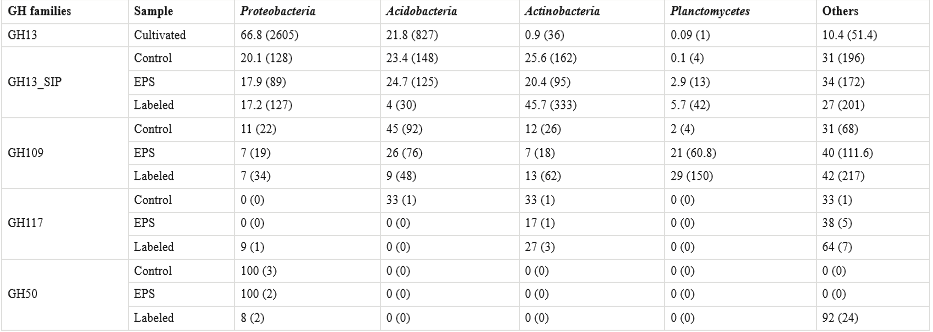


表3：SIP实验和纯培养试验中富集的GH家族的相关序列的分类。

在GH家族中，SIP实验标记样品的丰度较高。GH109的对应序列属于*Acidobacteria*(占GH序列的45%)、其他门(占GH序列的31–42%)和*Planctomycetes*(占GH序列的2–29%)。GH117的对应序列属于*Actinobacteria*(占GH序列的17–33%)、*Acidobacteria*(占GH序列的0–33%)和其他门(占GH序列的33–64%)。GH50的对应序列属于*Proteobacteria*(占GH序列的8–100%)和其他门(占GH序列的0–92%)。GH32的对应序列属于*Acidobacteria*(占GH序列的11–44%)和其他门(占GH序列的44–79%)。GH17的对应序列属于*Proteobacteria*(占GH序列的44–75%)和其他门(占GH序列的25–57%)。GH71的对应序列属于*Actinobacteria*(占GH序列的35–100%)、*Proteobacteria*(占GH序列的0-25%),*Acidobacteria*(占GH序列的0–25%),和其他门(占GH序列的0–43%)。

**纯培养微生物基于宏基因组数据组装出的MAGs**

使用大于5kb的Contigs进行binning，经过质量过滤，产生了4个基因组草图。基因组长度为3.0-6.3Mb，GC含量为57-62%。所有的MAGs都属于*Proteobacteria*门。没有一个MAGS被分类到属水平；最接近MAG1的分类是*Paraburkholderia*，最接MAG2和MAG4的分类是*Amantichitinum*(MAG2和MAG4)，最接近MAG3的分类是*Rhodanobacteraceae*。

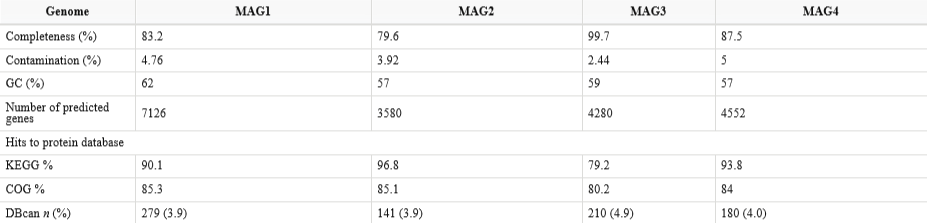
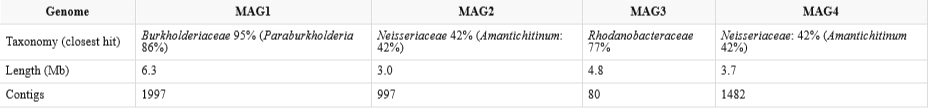


表4：4个MAGs的基因组特征。

MAGs中大约有83.7%的ORF可以通过COG注释出。分析表明，大部分COGs相对应的ORF功能未知(占ORFs的16.4–18.4%)。a)在分类的COGs中，丰度最高的种类是“K-转录”、“E-氨基酸代谢”、“G-碳水化合物代谢”。“C-能量产生”、“P-无机离子代谢”和“M细胞壁/膜生物发生”。b)KEGG通路分析表明，约90%的ORFs预测克归到KEGG同源物。在所有的MAGs中，大多数丰度最高的KEGGs具有几个类型的转运蛋白功能。c)在MAG1中，观察到分布在90个家族中的279个CAZymes，其中丰度最高的是CE1、GT4、GT42、CE10和AA3；在MAG2中，观察到分布在65个家族中的141个CAZymes，其中GT41、GT2和CE1最为丰富；在MAG3中，观察到分布在81个家族中的210个CAZymes，其中GT41、GT2、CE1和CE10最为丰富；在MAG4中，观察到分布在73个家族中的180个CAZymes，其中丰度最高的是CE1、GT2和GT41。

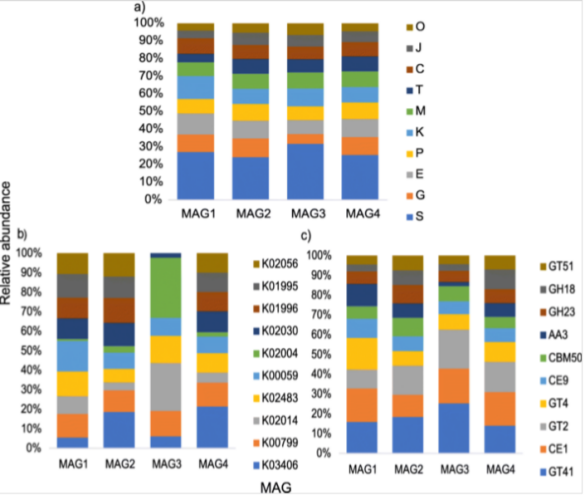


图6：组装的基因组草图中功能类别的相对丰度分布a)COG注释b)KEGG注释c)dbCAN注释。（E-氨基酸的转运和代谢;G-碳水化合物的运输和代谢;C-能源生产和转换;M-细胞壁/膜/包膜生物发生;P无机离子的转运和代谢;O-post translational modification翻译后的校正;T-信号转导机制;K-转录;J-翻译;S-未知功能;）

**讨论**

在本研究中，我们应用依赖于纯培养和不依赖于纯培养的技术来评估微生物多样性和降解微生物生物聚合物WH15EPS所涉及的功能，重点关注具有生物技术应用潜力的酶。首先，我们比较了添加和不添加WH15EPS的环境的功能潜力，采用稳定同位素探测技术评价添加生物聚合物所造成的物种和功能的富集。其次，我们利用宏基因组学技术评价以WH15EPS为单一碳源的培养基中所生长的微生物的功能潜力。

SIP实验证明，无论是16SrRNA数据集还是ORF数据集，无论WH15EPS是标记还是未标记处理，*Proteobacteria*，*Actinobacteria*，*Acidobacteria*和*Planctomycetes*都是丰度最高的门。然而，在凋落物样品中添加WH15EPS可促进*Planctomycetes*门的丰度增加，这在标记样品中更为明显，表明*Planctomycetes*门在WH15EPS的降解中起着积极的作用。此外，在16SrRNA分析中，属于同一门的“*未分类Planctomycetes*”和*Singulisphaera*是标记样品中丰度最高的种群，而“*未分类Planctomycetes*”同样也是ORF分析中丰度最高的种群。*Proteobacteria*，*Actinobacteria*和*Acidobacteria*被广泛认为参与了碳降解过程，例如葡萄糖、木聚糖和纤维素同化。Ivanovaetal.最近证明了*Planctomycetes*门有糖酵解潜力。例如，*Singulisphaera*属对果胶和木聚糖有显著的反应。

正如预期的那样，依赖于纯培养的方法得到的物种多样性较低，其中*Proteobacteria*门是丰度最高的。在依赖于纯培养和不依赖于纯培养的方法中观察到物种多样性之间的差异，特别是丰度最高的种群的差异，被认为是平板计数异常。微生物在实验室的可培养性取决于许多因素，如营养物质、氧水平、温度、*pH*值和生长因子，这些因素限制了在培养基中可以实际恢复的种群的总数。然而，使用*WH15EPS*作为替代碳源可以证明，不寻常的化合物可以使一些仍然未知的微生物在实验室条件下生长。培养基平板的较低多样性促成丰度最高*Proteobacteria*门组装出四个基因组草图，属于属级未知菌种，再次证明了未知微生物有可富集和分离的潜力。

为了寻找具有生物技术应用潜力的酶，我们研究了CAZymes在依赖于纯培养和不依赖于纯培养的方法中生成的数据集的多样性，因为它们几乎在所有的工业部门，如化学、制药、食品工业，洗涤剂生产、纺织品、皮革、造纸业和生物能源的生产方面有非常重要的意义。此外，我们还研究出可用于生物膜去除的酶。

在所有数据集中得到的CAZymes中，我们发现无论是在依赖于纯培养和不依赖于纯培养的数据集中，丰度最高的CAZymes是糖苷转移酶家族，如GT41、GT2和GT4。众所周知，GTs通过将碳水化合物、蛋白质、脂类、DNA和其他分子等从供体转移到受体，从而催化糖苷键的形成。尽管微生物基因组的很大一部分基因编码GTs（约占基因总数的1-2%），但这些酶仍然不如GHs那样广为人知。糖苷化合物发挥着广泛的作用，如能量储存、细胞完整性和信号传递等，天然产物的糖基化在生物活性化合物的探索中具有重要意义。GTs参与抗生素的生产，如氯雷莫霉素、万古霉素和红霉素。因此，制药工业可能特别对GTs感兴趣。

在糖苷水解酶中，这两个宏基因组数据集中丰度最高的家族是来自*Proteobacteria*的GH13，其中包括淀粉和普鲁兰修饰酶，即α-淀粉酶、普鲁兰酶、α-1,6-葡萄糖苷酶、分支酶、麦芽原酰胺酶、新普鲁兰酶和环糊精酶。淀粉酶是食品工业中最重要的酶，可用于生产葡萄糖和麦芽糖糖浆、降低糖浆粘度、生产果汁，淀粉在酿造过程中的增溶，以及生产烘焙产品。此外，研究了α-淀粉酶在抑制生物膜形成中的作用。Fleming等人研究发现，淀粉酶和纤维素溶液显著降低金黄色葡萄球菌和铜绿假单胞菌生物膜的生物量，提高抗生素治疗的有效性。

Craigen等人的研究中也发现了类似的作用，在商业上可用的α-淀粉酶分离金黄色葡萄球菌产生的聚集体并抑制其生物膜的产生。

尽管如此，Boruta包的特征选择揭示了GH家族在SIP实验重层样品中的丰度差异，且来源于能够降解WH15EPS的微生物。这些微生物主要属于*Proteobacteria*,*Acidobacteria*,*Actinobacteria*,*Planctomycetes*以及高比例的未知微生物。GH109含有α-N-乙酰半乳糖胺酶，可通过酶去除红细胞膜中的单糖从而开发通用的红细胞，改善医院的血液供应。此外，由于这些酶包含木糖、葡萄糖和阿拉伯糖，还可以参与WH15EPS的降解。GH117和GH50含有琼脂酶，可用于生产具有抗氧化活性的低聚糖，主要应用于食品、制药和化妆品行业。GH32包括转化酶和菊粉酶，可用于食品和发酵过程。GH17是由具有抗β-葡聚糖和层粘连蛋白活性的内切葡聚糖酶组成，是动物饲料中降解多糖的有效添加剂。GH71的突变体已经表现出对牙菌斑中的葡聚糖的活性。

有趣的是，在不依赖纯培养的数据集丰度最高的GHs中，有16个在依赖于纯培养数据集中的起主导作用，SIP实验标记样品中所有丰度较高的GHs也在依赖于纯培养的数据集中出现。此外，MAGs也包含不同丰度的GH相关家族。MAG1中有8种属于GH92家族的ORFs，其中包括在食品和制药工业广泛应用的α-甘露糖酶，常用于生产果汁、植物材料的降解和提取咖啡。在MAG2中，有5种ORFs属于GH23，其中含有溶菌酶，常被用作多糖水解剂破碎生物膜。MAG3含有GH92、GH23和GH2家族的ORFs。GH2包括几种酶，例如β-半乳糖酶，主要用于生产无乳糖乳制品和其他半乳糖产品。MAG4含有丰富的GH18，例如几丁质酶，在控制真菌生物生长和生物修复过程的重要作用。我们认识到即使MAGs具有较低的污染水平（<5%），它们并不是分离微生物实验的代表性基因组。因此，应用纯培养试验培养出对应微生物来充分验证我们的MAGs。

我们的研究表明，结合SIP和EPS培养，我们可以检测到能够结合生物聚合物的总微生物群落的子集。结合生物信息学、基因合成和酶筛选方法，我们发现*Planctomycetes*是生物技术研究和异源表达的一个有趣的靶点。我们还证明了添加WH15EPS在依赖于纯培养和不依赖于纯培养的方法中，导致GHs编码基因得到了富集，例如淀粉酶、几丁质酶、琼脂酶和内切葡聚糖酶等，主要应用于化工、制药和食品工业。此外，使用WH15EPS进行培养可研究和分离尚未知的微生物，如未分类的*Proteobacteria*和*Planctomycetes*，增加了可培养微生物数量。

**结论**

我们观察到，在依赖于纯培养和不依赖于纯培养的实验方法中，添加WH15EPS所引起的功能多样性中，共发现310个CAZyme家族，其中38.4%(119个)属于GH家族。几乎所有工业部门，如化学、制药和食品工业，以及洗涤剂、纺织品的生产都可能使用GHs。此外，我们还观察到酶的存在，可用于生物膜去除。即使这些潜在的酶可能属于在实验条件下生长缓慢的微生物，如*Acidobacteria*,*Planctomycetes*和*Verrucomicrobia*，但序列仍可用于进一步的异源表达和表征。此外，基于培养的宏基因组学数据集组装出4个基因组草图(MAGs)，这些基因组可能属于未分类*Proteobacteria*.我们发现WH15EPS可用于分离已知和未知的微生物，以及靶向广泛的CAZyme家族的序列.

我们证明，WH15EPS在依赖于纯培养和不依赖于纯培养的方法中的存在所引起的功能多样性中GHs中得到了富集，如淀粉酶和内切葡聚糖酶，可应用于化工、制药和食品工业部门。此外，WH15EPS可用于调查和分离尚未未知的种群，如未分类*Proteobacteria*和*Planctomycetes*，增加目前培养的具有潜在生物技术特征的细菌代表的数量。