



过表达 PDI 对毕赤酵母分泌淀粉样蛋白的影响——以胱抑素为例

牛婷婷[#], 周雪洁[#], 余小波, 蔡宜生, 陈静, 王瑞, 付爽, 栾闯, 何剑为^{*}

辽宁大学生命科学院, 辽宁 沈阳 110036

摘要: 【目的】本研究旨在结合酵母菌蛋白质二硫键异构酶(protein disulfide isomerase, PDI)与其底物蛋白鸡胱抑素 C (chicken cystatin C, cC)在酵母中的共表达, 理解 PDI 影响外源蛋白合成与表达的调控规律。运用转录组深度测序技术(RNA-Seq)筛选差异基因, 调取并鉴定影响 cC 表达的关键基因, 为解析外源蛋白高效表达机制, 改造工程菌株提供理论支撑。【方法】以巴斯德毕赤酵母 GS115、GS115-cC 为出发菌株, 采用电转的方法将携带 PDI 编码基因的载体 pPIC3.5K 转入到 GS115/GS115-cC 菌株, 使其在菌株中过表达, 研究过表达 PDI 对 cC 表达的影响。采用 RNA-Seq 深度测序方法, 研究重组毕赤酵母基因表达差异情况。并结合 KEGG 注释结果对数据进行分析, 挑选差异显著表达基因进行验证, 初步明确其在蛋白表达调控方面的功能。【结果】本研究通过构建过表达 PDI 重组毕赤酵母菌株, 使得外源蛋白 cC 的表达量显著增加。利用 RNA-seq 技术分析过表达 PDI 菌株与正常菌株的差异, 最终筛选了 373 个差异表达基因, 其中有 122 个差异基因注释到 KEGG 生物通路, 包括 12 个基因注释到蛋白质转运和分解代谢途径, 21 个基因注释到蛋白质折叠分选和降解途径, 以及 24 个基因参与蛋白质的翻译途径等。【结论】在毕赤酵母中过表达 PDI 能显著增加外源蛋白 cC 的表达量。通过对过表达与正常表达 PDI 的毕赤酵母基因的表达谱分析, 初步确定了其中一些转录情况变化显著的基因, 明确了它们参与的细胞途径和信号通路, 为改造具有高效率表达淀粉样蛋白的酵母菌株奠定基础。

关键词: 毕赤酵母, 鸡胱抑素 C, 蛋白质二硫键异构酶, 转录组测序

巴斯德毕赤酵母(*Komagataella phaffii*, 原命名为 *Pichia pastoris*) 在表达外源蛋白时可将甲醇作为唯一的碳源和能源来诱导和调控外源基因

的表达。毕赤酵母表达系统具有大肠杆菌表达系统的诸多优势, 同时又具有高等真核表达系统的许多优点: 如蛋白加工、折叠、翻译后修饰等^[1];

基金项目: 国家自然科学基金(31670103)

[#]共同第一作者。

^{*}通信作者。Tel/Fax: +86-24-62202232; E-mail: jwhe@lnu.edu

收稿日期: 2021-01-23; 修回日期: 2021-03-30; 网络出版日期: 2021-08-06

其胞外蛋白分泌的非常少,有利于后续对外源蛋白的分离和纯化;通过使用工业生物反应器,可以从较小的培养体积大规模地生产目的蛋白质。因此,毕赤酵母已成为比较理想的外源蛋白表达宿主之一,且已被成功地应用于工业、生物制药及疫苗研究领域^[2]。

蛋白质二硫键异构酶 (protein disulfide isomerases, PDI) 是内质网中与折叠相关的负责二硫键形成的多功能蛋白,它既可以催化新生肽链形成二硫键,又能够帮助蛋白质折叠形成其正确的天然构象^[3]。酵母的 PDI 与哺乳动物 PDI 结构相似,全长 PDI 分子结构在空间上呈“U”型,具有 a、b、b'、a'、c 功能结构域,“U”型结构的底部由 b 和 b' 结构域组成,通过其疏水表面结合底物。“U”型结构的两个侧臂由 a 和 a' 结构域组成,主要起催化功能^[4]。由于“U”型结构的内表面富集疏水残基,有利于其与错误折叠蛋白之间发生相互作用。

鸡胱抑素 C (chicken cystatin C, cC), 即鸡半胱氨酸蛋白酶抑制剂, 是人胱抑素 C (human cystatin C, HCC) 的同源蛋白, 二者的氨基酸序列、结构、生理特性等都具有高度的同源性^[5]。HCC 在体内易形成二聚体或寡聚体, 导致其半胱氨酸蛋白酶抑制剂的活性丧失, 而进一步的寡聚化能形成多聚体最终形成成熟的淀粉样纤维, 高度稳定的淀粉样纤维沉积在脑动脉血管壁直接导致了脑淀粉样血管病的发生^[6-7]。cC 大小约为 14 kDa, 由 116 个氨基酸组成, 是一种低分子量蛋白。cC 在生理条件下以单体形式存在, 更易于分离出蛋白单体, 并具有更高的热力学稳定性,

因此在进行实验操作时具有更高的可行性和便利性。在科研中也常用 cC 来替代 HCC 作为淀粉样蛋白的模型蛋白进行科学研究^[8-9]。

有关 PDI 调控外源新生蛋白表达的相关研究显示: 在敲除 PDI 同源蛋白 Eps1 的酿酒酵母中表达外源淀粉样蛋白时, 外源蛋白的胞外分泌量和胞内驻留量出现显著变化^[10]; 同时, 毕赤酵母在表达外源淀粉样蛋白时存在二硫键未正确形成的问题, 干扰细胞自身蛋白的正确折叠, 误折叠蛋白的积累会导致折叠胁迫^[11]。折叠胁迫会改变细胞的正常代谢, 降低细胞的抗逆性, 甚至造成细胞的死亡^[12]。大量研究表明, 毕赤酵母中 PDI 的过表达可以促进毕赤酵母分泌外源重组蛋白, 如 PDI 的过表达使得抗体蛋白 2F5 Fab 在 *Pichia pastoris* 的表达水平提高了 2 倍^[13]。虽然通过与 PDI 的共表达提高了外源蛋白的表达水平和活性, 但在真核细胞中, 由于蛋白质折叠是一个复杂的多步骤过程, PDI 帮助新生蛋白质合成与折叠的调控机制以及在蛋白分泌网络中与 PDI 相关的调控因子仍不明确。

本研究通过构建 PDI 过表达菌株, 研究 PDI 过表达对外源淀粉样蛋白 cC 在毕赤酵母中合成、折叠及分泌的影响。结合 RNA-Seq 组学方法与高通量分析手段比较 PDI 过表达菌株与对照菌株, 揭示 RNA 转录本在特定时刻的存在情况和数量的多少。根据基因在不同菌株中的不同表达量进行差异基因表达分析、筛选、功能注释等生物学信息分析, 以期识别显著上调/下调的基因, 在表达宿主全基因组范围内挖掘改造宿主分泌途径的新靶点。

1 材料和方法

1.1 菌种和质粒

毕赤酵母 GS115 (*his4*)菌株由福州大学刘树滔教授惠赠。酵母表达载体 pPICZαA 购自 Invitrogen 公司。酵母胞内表达载体 pPIC3.5K 购自武汉森灵生物科技有限公司。

1.2 培养基和菌体的培养条件

PDI 相关重组 DNA 操作在大肠杆菌菌株 (DH5α, Invitrogen)中进行, 该菌株在 37 °C 于 LB 培养基中培养。在我们先前的工作中, 已经构建了质粒 pPICZαA-cC 并将其转化到大肠杆菌菌株 (Top10, Invitrogen)中。酵母细胞在 YPD (1% 酵母提取物, 2% 蛋白胨和 2% 葡萄糖)中于 30 °C 培养至对数生长期。携带外源基因的重组酵母细胞在培养基 YPM [1% 酵母提取物, 2% 蛋白胨, 0.5% 甲醇(V/V)]中诱导表达 72 h。

1.3 表达载体和过表达菌株的构建

根据已经发表的 GenBank 序列号 EU805807 的酵母 PDI 基因序列设计引物 5'-TACGTAACCA TGCAATTCAACTGGGATATT-3' (正向, 下划线为 *Sna*B I 位点)和引物 5'-GCGGCCGCTTAAAGC TCGTCGTGAGCGTC-3' (反向, 下划线为 *Not* I 位点)从 *Pichia pastoris* X33 (Invitrogen)的基因组 DNA 扩增出来。扩增的 PDI 片段被 *Sna*B I/*Not* I (TaKaRa)酶切后与载体 pPIC3.5K 连接, 构建 pPIC3.5K-PDI 重组载体。将含有 cC 编码基因的重组载体 pT-cC (前期构建)用 *Xho* I/*Xba* I (Thermo Fisher)酶切后与载体 pPICZαA 连接, 构建 pPICZαA-cC 重组载体。通过电穿孔将 *Sac* I 线性化的质粒 pPICZαA-cC 转化到毕赤酵母 GS115 细

胞中, 并在含有 100 μg/mL Zeocin (Invitrogen)的 YPDS 平板(1% 酵母提取物, 2% 蛋白胨, 2% 葡萄糖, 1 mol/L 山梨醇和 2% 琼脂粉)中筛选重组子。同时, 采用同样方法将线性化的 pPIC3.5K-PDI 也转化到毕赤酵母 GS115 细胞中, 并在 MD 平板上选择重组子。选择后的酵母细胞的阳性转化子即为 Mut⁺ (EasySelect™ *Pichia* 表达试剂盒)。

1.4 蛋白表达分析

将在 YPM 培养基诱导 72 h 的酵母细胞于 4 °C、12000 r/min 离心 30 min 后收集上清即为胞外蛋白样品。应用酵母蛋白质提取试剂(TaKaRa)从酵母细胞中提取胞内蛋白样品。在变性条件下用 15% SDS 聚丙烯酰胺凝胶电泳(PAGE)检测 cC 蛋白, 用 12% PAGE 观察检测 PDI 蛋白表达情况。将凝胶用考马斯亮蓝显示蛋白条带, 然后将其转移到 PVDF 膜(Millipore)上进行蛋白质印迹分析。1:2000 稀释度的 PDI 一抗(ABclonal)用于检测 PDI, 1:2000 稀释度的 cC 一抗(自制)检测 cC。二抗为 1:5000 稀释度的 HRP-conjugated Goat Anti-Rabbit IgG。使用 Image J 软件分析 PVDF 膜上的条带, 并使用 Graphpad Prism 5.0 处理生成的数据。

1.5 Total RNA 样品抽提和检测

当酵母生长至 $OD_{600}=2.5$, 处于对数生长期时, 依据酵母 RNA 提取试剂盒 Yeast RNAiso Kit (TaKaRa)操作说明提取酵母 Total RNA。

1.6 高通量 RNA-Seq 测序

使用 NanoDrop 2000 测量 RNA 浓度, 之后应用 Agilent Bioanalyzer 2100 系统的 RNA Nano 6000 测定试剂盒评估 RNA 的完整性, 后进行转录组测序的文库制备。根据厂家说明, 将文库制

备物在 Illumina HiSeq Xten 平台上测序, 并产生配对末端读数。此实验委托北京百迈克生物科技有限公司操作。

1.7 实时荧光定量 qRT-PCR

将提取的总 RNA 作为模板, 使用反转录试剂盒 (RT Master Mix, TaKaRa) 合成第一条链 cDNA, 具体操作步骤如产品说明书所描述。按照荧光定量试剂盒 (SYBR Premix Ex Taq II) 说明书配置荧光定量 PCR 体系的混合液, 并且依据酶的特性设定荧光定量 PCR 循环条件, 进行荧光定量 PCR 实验。

2 结果和分析

2.1 GS115 毕赤酵母菌株 PDI 过表达系统的建立

已有文献报道, 存在于内质网并且与蛋白折叠相关的分子伴侣中, PDI 的过表达可以增加毕赤酵母的分泌外源蛋白的表达量^[13]。因此, 我们研究使用过表达 PDI 的方法来提高外源淀粉样蛋白 cC 的表达量。为了建立用于表达分泌型 cC 的 PDI 过表达酵母菌株, 我们首先选择 pPIC3.5K 作为胞内表达载体用于胞内表达 PDI, 构建了过表达 PDI 的重组毕赤酵母菌株 GS115-PDI, 然后选择 pPICZαA 载体进行 cC 的胞外分泌蛋白表达, 构建了过表达 PDI 状态下的重组毕赤酵母菌株 GS115-PDI-cC。

2.2 PDI 的过表达增加了外源淀粉样蛋白 cC 的表达

由于醇氧化酶启动子 (AOX1) 受甲醇诱导, 利用甲醇可诱导 AOX1 连接的外源蛋白编码基因的表达, 通过蛋白质免疫印迹法检测诱导后 PDI 在

各个菌株中的表达情况, 结果见图 1, 过表达酵母菌株中 PDI 蛋白的含量显著高于对照菌株, 表明成功构建了 PDI 过表达的毕赤酵母菌株。然后, 我们使用 SDS-PAGE 方法分析了每个菌株中胞内和胞外蛋白的表达情况, 发现胞内和胞外蛋白都显著增加 (图 2)。同时, 使用蛋白质免疫印迹法检测每个菌株中表达的 cC 蛋白量, 结果表明, 过表达 PDI 菌株中胞内和胞外的 cC 的表达量显著增加, 并且胞内 cC 二聚体也明显增加 (图 3)。以上结果表明, PDI 的过表达显著增加了毕赤酵母 GS115 中外源淀粉样蛋白 cC 的表达量。

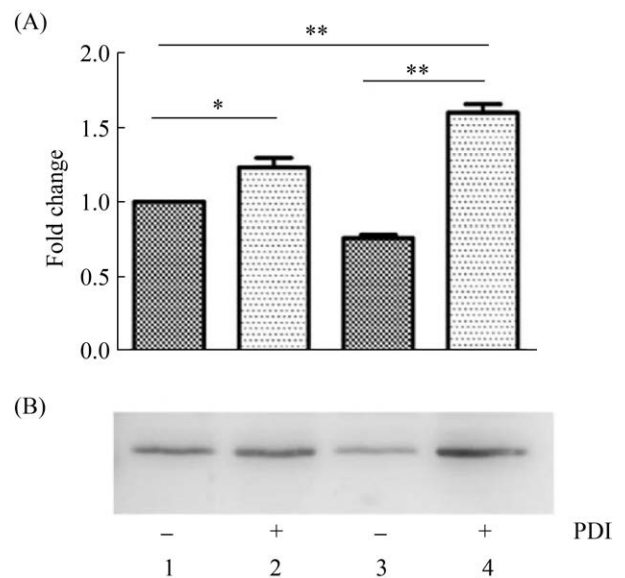


图 1. Western blotting 验证毕赤酵母 GS115 中 PDI 的表达水平

Figure 1. Western blotting verifies the expression level of PDI in *Pichia pastoris* GS115. +: the plasmid carrying PDI gene is transformed into *Pichia pastoris* GS115 and the PDI protein is successfully overexpressed; -: the PDI gene is not overexpressed in *Pichia pastoris* GS115. The abscissa of the Figure 1-A corresponds to the band of every lane in Figure 1-B.

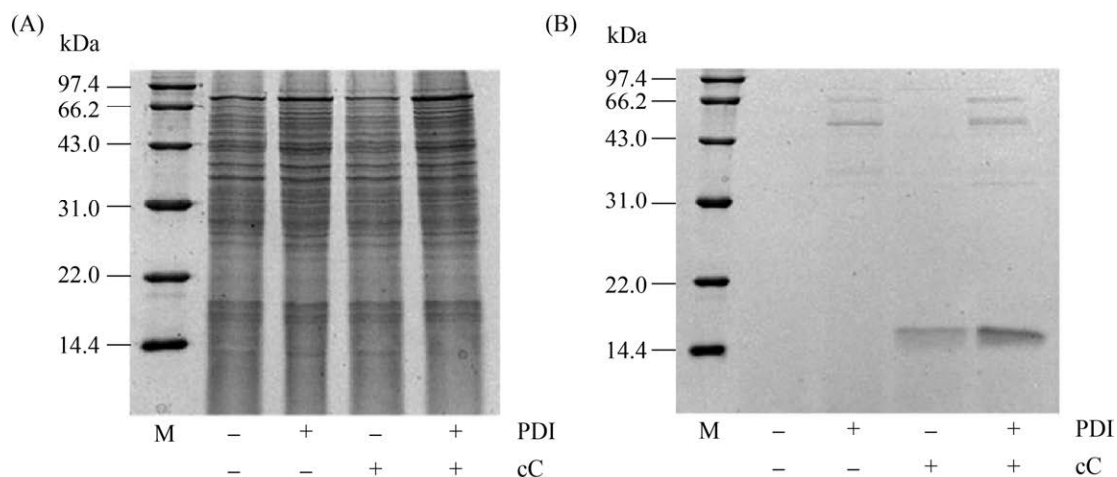


图 2. 各菌株细胞内总蛋白(A)与细胞外分泌蛋白(B)的 SDS-PAGE 凝胶成像图

Figure 2. SDS-PAGE gel image of total intracellular protein (A) and extracellular secretion protein (B) of each strain. +: the plasmid carrying this gene is transformed into *Pichia pastoris* GS115; -: the vector carrying the target gene is not transfected into *Pichia pastoris* GS115 relatively. PDI: disulfide bond isomerase; cC: chicken cystatin.

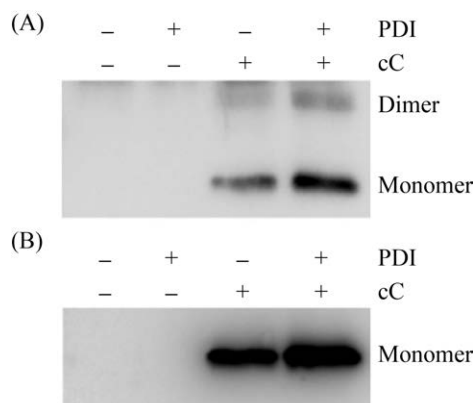


图 3. Western blotting 分析诱导 72 h 后 cC 蛋白表达情况

Figure 3. Western blotting analysis of the expression of cC protein 72 h after induction. +: the plasmid carrying this gene is transformed into the host (*Pichia pastoris* GS115); -: the target gene is not transfected into the host strain.

2.3 高通量 RNA-Seq 测序

已有研究表明基因表达存在着生物变异性^[14-15]。为了寻找真正差异表达的基因, 需要考虑并解决由于生物变异引起的表达差异。目前最

常用且最有效的方法是在实验设计中设立生物学重复。重复条件限制越严格, 重复样品数目越多, 寻找到的差异表达基因越可靠^[16]。研究表明, 皮尔逊相关系数 r 常作为生物学重复相关性的评估指标^[17]。 r^2 越接近 1, 说明 2 个重复样品相关性越强^[18]。因此, 我们对菌株 GS115、GS115-PDI、GS115-cC、GS115-PDI-cC 每个菌株进行 3 次重复共完成了 12 个样品的转录组测序分析。分析样本之间的相关系数如图 4 所示, 从图中可以看到 2 个相同样品的相关性 r^2 值都接近 1, 说明样品的重复性达到了生物学统计的标准, 可以用于筛选差异表达基因。

2.4 利用 RNA 测序技术筛选并注释 GS115-cC 与 GS115-PDI-cC 比较之后的差异表达基因

如表 1、图 5 所示, 通过比较不同菌株, 我们鉴定出了数千个差异表达基因。但是对于只属于在过表达 PDI 时影响毕赤酵母 GS115 宿主菌株增加 cC 表达量的基因, 我们最终筛选出 373 个

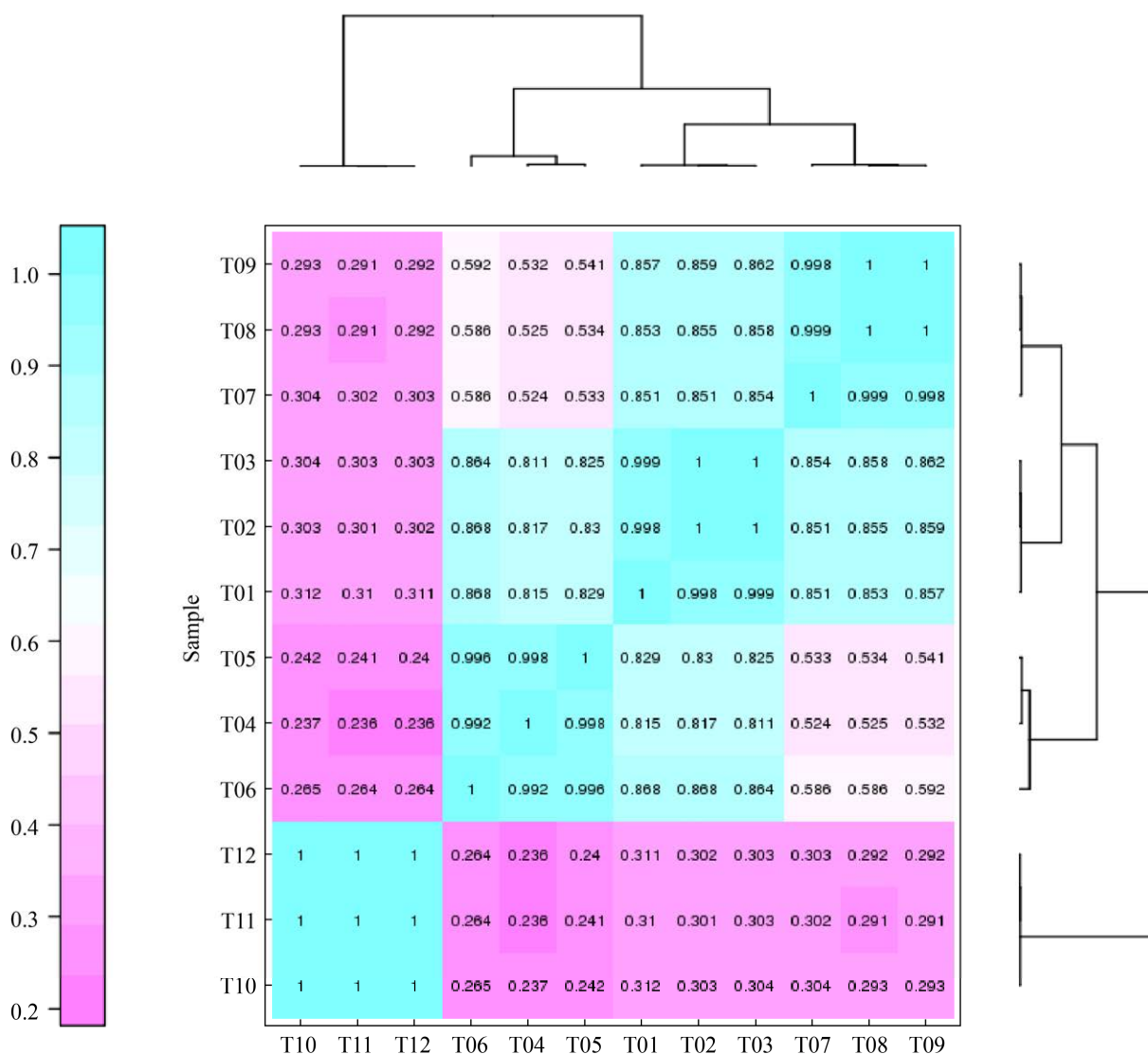


图 4. 两两样品的表达量相关性热图

Figure 4. Heat map of expression level correlation of pairwise samples.

表 1. 不同菌株的两两比较表

Table 1. Pairwise comparison table of different strains

Group	Strains comparison	Samples in RNA-seq
Group0(G0)	GS115 VS GS115-PDI	T1,T2,T3 VS T4,T5,T6
Group1(G1)	GS115 VS GS115-cC	T1,T2,T3 VS T7,T8,T9
Group2(G2)	GS115 VS GS115-PDI-cC	T1,T2,T3 VS T10,T11,T12
Group3(G3)	GS115-PDI VS GS115-PDI-cC	T4,T5,T6 VS T10,T11,T12
Group4(G4)	GS115-cC VS GS115-PDI-cC	T7,T8,T9 VS T10,T11,T12

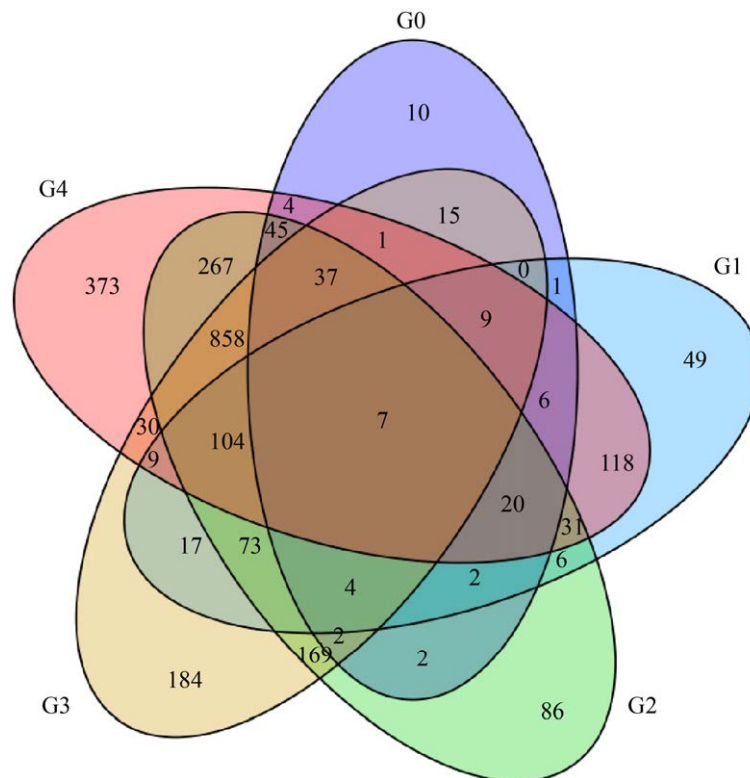


图 5. 不同菌株比较组差异表达基因(DEG)维恩图

Figure 5. Differentially expressed genes (DEG) Venn diagram of different strains comparison group.

差异表达基因, 其中上调和下调的基因分别有 191 和 182 个。注释到 KEGG 数据库的差异表达基因 (DEG) 的结果^[19]如图 6 所示, 有 122 个基因成功注释到 KEGG 数据库中, 其中上、下调的基因分别有 52 和 70 个。然而, 并不是所有这些差异表达基因都参与蛋白质的运输和分泌以及内质网质量控制系统。通过对图 6 的分析, 我们发现 24 个 DEG 参与蛋白质翻译通路, 12 个 DEG 参与蛋白质的运输和分解代谢, 21 个 DEG 参与蛋白质折叠、分选和降解的途径。基因差异表达结果显示, 大部分参与翻译通路以及大部分参与蛋

白质运输和分解代谢的 DEG 表达下调, 暗示着蛋白质过量表达会消耗宿主细胞可利用的碳源和能量, 因此会给宿主带来代谢负担。而参与蛋白质折叠、分选和降解通路的 DEG 表达上调。这暗示着由于 PDI 本身是一种存在于内质网中的分子伴侣, 它能够帮助蛋白质正确折叠防止那些错误折叠蛋白被蛋白质降解系统降解, 从而来增加外源蛋白的表达量。在这些 DEG 中, 我们进一步在蛋白质折叠及内质网应激途径进行差异基因的再次筛选, 如表 2 所示, 我们选取到了参与上述两途径的 8 个 DEG。



图 6. 差异表达基因 KEGG 分类图

Figure 6. Differentially expressed genes KEGG classification map. The ordinate is the name of the KEGG metabolic pathway, the abscissa is the ratio of the number of genes annotated to the pathway and their number to the total number of annotated genes.

表 2. 选择与蛋白质折叠及内质网品质管理系统相关的 DEG

Table 2. Select DEG related to protein folding and endoplasmic reticulum quality management system

Gene ID	Gene name	GenBank	Regulation	KEGG pathway annotation
Gene 263	PAS_chr1-3_0261	XM_002489561.1	Up	Endocytosis (ko04144)
Gene 3092	PAS_chr3_1165	XM_002492390.1	Up	Endocytosis (ko04144)
Gene 3874	PAS_chr3_0974	XM_002493172.1	Up	Endocytosis (ko04144)
Gene 3122	PAS_chr3_0249	XM_002492420.1	Up	Protein processing in endoplasmic reticulum (ko04141)
Gene 4006	PAS_chr3_1104	XM_002493304.1	Up	Protein processing in endoplasmic reticulum (ko04141)
Gene 463	PAS_chr1-1_0160	XM_002489761.1	Down	Protein processing in endoplasmic reticulum (ko04141)
Gene 160	PAS_chr1-3_0161	XM_002489458.1	Up	SNARE interactions in vesicular transport (ko04130)
Gene 3814	PAS_chr3_1238	XM_002493112.1	Up	SNARE interactions in vesicular transport (ko04130)

2.5 差异基因的再次验证

我们使用 qRT-PCR 的方法对表 2 中的 DEG 进行再次验证,证实这些基因在分子水平上是否出现了相同的上下调的变化情况。表 3 列出目的基因引物序列。如图 7-A 所示,在内吞通路中的基因(PAS_chr1-3_0261, PAS_chr3_1165, PAS_chr3_0974)均上调。图 7-B 所示功能未知的基因

(PAS_chr3_0249)和参与内质网蛋白加工的基因(PAS_chr3_1104)上调。从 qRT-PCR 结果可知,这与我们表 4 的 RNA-seq 结果一致。因此,我们再一次在分子水平上验证了差异表达基因的变化情况。我们认为对差异表达基因的再次鉴定是基于客观的实验数据,鉴定出的差异表达基因具有一定的研究价值,可以作进一步分析。

表 3. 基因的引物序列

Table 3. Gene primer sequence

Gene ID	Gene name	GenBank	Sequences of primer
Gene 263	PAS_chr1-3_0261	XM_002489561.1	F: GGTAGAGGGGCAAGAACATGAG R: TGGGCTGAGTAGCAGATGGTG
Gene 3092	PAS_chr3_1165	XM_002492390.1	F: ATCGGTTTCGTCCTGTTTCTC R: ACTGGATTTCCTGCTTGTTC
Gene 3874	PAS_chr3_0974	XM_002493172.1	F: GGAGAAGGGTGGTAATGAACG R: GAGACTGTCCCGAATGGTCAG
Gene 3122	PAS_chr3_0249	XM_002492420.1	F: TTCGCCCTCCTGAACTGATAG R: CCTGGGGAACCTACGGTCTTG
Gene 4006	PAS_chr3_1104	XM_002493304.1	F: GACCAATCGTCGTTTAGAGCTAC R: CTGCATTCTGTCTTGTCTGTC
Gene 463	PAS_chr1-1_0160	XM_002489761.1	F: AGTTTTATGCCGATTGGTGTTC R: TGAGTATTTGTTAGCATCAATGGC
Gene 160	PAS_chr1-3_0161	XM_002489458.1	F: TAGAATCGCAAATGAGGACG R: TTTCTTAGCCATTACCATCATCC
Gene 3814	PAS_chr3_1238	XM_002493112.1	F: GGCGAACCTACCCTCATAACAC R: TGGCTTCATTTCTATTGCATCTC

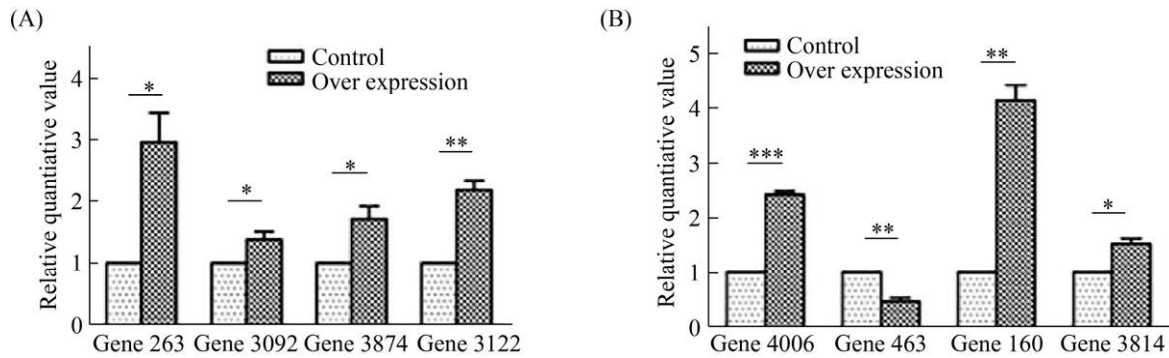


图 7. 关键差异基因相对荧光定量柱状图

Figure 7. Relative fluorescence quantitative histogram of key differential genes. The control group is the GS115-cC strain. The overexpression group represents the GS115-PDI-cC strain.

表 4. 关键差异基因 RNA-Seq 与 qRT-PCR 结果比较

Table 4. Comparison of RNA-Seq and qRT-PCR results of key differential genes

Gene ID	Gene name	Log2FC	FDR	Regulation	ΔC_t / Over-expression	ΔC_t /Control	$2^{-\Delta\Delta C_t}$	Regulation
Gene 263	PAS_chr1-3_0261	1.907035171	2.17E-85	Up	-0.657	0.893	2.954	Up
Gene 3092	PAS_chr3_1165	1.038464554	2.21E-21	Up	-0.107	0.353	1.380	Up
Gene 3874	PAS_chr3_0974	1.054612732	2.46E-15	Up	1.230	2.000	1.713	Up
Gene 3122	PAS_chr3_0249	1.056165135	1.68E-22	Up	-3.257	-2.137	2.177	Up
Gene 4006	PAS_chr3_1104	1.032749024	1.48E-99	Up	1.793	3.063	2.412	Up
Gene 463	PAS_chr1-1_0160	-1.181488455	7.85E-46	Down	0.59	-0.503	0.472	Down
Gene 160	PAS_chr1-3_0161	1.047977247	1.42E-07	Up	1.227	3.273	4.138	Up
Gene 3814	PAS_chr3_1238	1.046077293	2.14E-42	Up	1.140	1.743	1.522	Up

FC: fold change; FDR: false discovery rate.

3 讨论

毕赤酵母在表达外源蛋白时可将甲醇作为唯一的碳源来严格控制外源基因的表达^[20]。但是对淀粉样蛋白进行表达时，由于其易于聚集，所以此系统对其表达的结果不是很理想^[21]。内质网是合成、加工蛋白质的主要场所，当内质网中存在大量未正确折叠的蛋白时，会引起 UPR 反应，

启动蛋白质降解途径使其降解^[22]。PDI 是存在于内质网中的氧化还原酶，它既能够帮助蛋白质形成二硫键又可以促使蛋白正确折叠^[23]。本研究通过构建过表达 PDI 的毕赤酵母菌株，在蛋白水平上证明了过表达 PDI 的菌株在表达蛋白 cC 时，无论在胞内还是胞外，cC 表达量都显著高于 PDI 正常水平的毕赤酵母菌株。这表明过表达 PDI 的毕赤酵母菌株能显著提高 cC 的表达量。

基于以上蛋白水平的表达结果, 应用高通量测序平台并采用转录组深度测序分析手段, 初步挖掘出部分在蛋白折叠、修饰、转运与分泌过程中的潜在参与基因。通过与正常 PDI 水平的毕赤酵母表达 cC 比较, 当过表达 PDI 时, 我们筛选到了 373 个发生了上调或下调的基因。进一步将这些基因注释到 KEGG 途径中后发现它们主要集中在蛋白质翻译修饰、内质网加工、未折叠蛋白反应 (unfolded protein response, UPR) 等途径。这个结果表明当过表达 PDI 时, 酵母细胞会通过调节这些差异基因的表达, 对新生的大量未正确折叠的蛋白和错误折叠的蛋白进行必要的辅助折叠和修饰, 并启动 UPR 对错误折叠的蛋白质进行降解, 以适应折叠胁迫对于细胞表达外源蛋白的微环境造成不利影响^[11]。这些连锁反应不仅提高了蛋白的正确折叠效率, 也提高了外源蛋白的表达量和分泌量。另一方面, 折叠胁迫会改变细胞的正常代谢, 降低细胞的抗逆性, 甚至造成细胞的死亡^[12]。因此本研究中筛选到的候选基因也为折叠胁迫相关细胞信号通路、途径的深入分析提供了实验数据上的有力支撑。

综上, 对差异表达基因的进一步挖掘分析可能为寻找降低淀粉样蛋白在胞内的聚集、提高毕赤酵母中外源淀粉样蛋白表达效率的方法提供新思路, 同时也为发现新的有关淀粉样蛋白沉积疾病的药物提供线索, 为改造高效过表达淀粉样蛋白的工程酵母菌株奠定基础。

参考文献

- [1] Karbalaeei M, Rezaee SA, Farsiani H. *Pichia pastoris*: a highly successful expression system for optimal synthesis of heterologous proteins. *Journal of Cellular Physiology*, 2020, 235(9): 5867–5881.
- [2] Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*, 2005, 22(4): 249–270.
- [3] Ellgaard L, Ruddock LW. The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Reports*, 2005, 6(1): 28–32.
- [4] Tian G, Xiang S, Noiva R, Lennarz WJ, Schindelin H. The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. *Cell*, 2006, 124(1): 61–73.
- [5] Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *The EMBO Journal*, 1988, 7(8): 2593–2599.
- [6] Janowski R, Abrahamson M, Grubb A, Jaskolski M. Domain swapping in N-truncated human cystatin C. *Journal of Molecular Biology*, 2004, 341(1): 151–160.
- [7] Janowski R, Kozak M, Abrahamson M, Grubb A, Jaskolski M. 3D domain-swapped human cystatin C with amyloidlike intermolecular beta-sheets. *Proteins*, 2005, 61(3): 570–578.
- [8] Mori K. Tripartite management of unfolded proteins in the endoplasmic Reticulum. *Cell*, 2000, 101(5): 451–454.
- [9] Higgins DR, Cregg JM. Introduction to *Pichia pastoris*. *Methods in Molecular Biology: Clifton, N J*, 1998, 103: 1–15.
- [10] He JW, Sakamoto T, Song YT, Saito A, Harada A, Azakami H, Kato A. Effect of EPS₁ gene deletion in *Saccharomyces cerevisiae* on the secretion of foreign proteins which have disulfide bridges. *FEBS Letters*, 2005, 579(11): 2277–2283.
- [11] Gasser B, Saloheimo M, Rinas U, Dragosits M, Rodríguez-Carmona E, Baumann K, Giuliani M, Parrilli E,

- Branduardi P, Lang C, Porro D, Ferrer P, Tutino ML, Mattanovich D, Villaverde A. Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microbial Cell Factories*, 2008, 7: 11.
- [12] Zhu TC, Guo MJ, Zhuang YP, Chu J, Zhang SL. Understanding the effect of foreign gene dosage on the physiology of *Pichia pastoris* by transcriptional analysis of key genes. *Applied Microbiology and Biotechnology*, 2011, 89(4): 1127–1135.
- [13] Gasser B, Maurer M, Gach J, Kunert R, Mattanovich D. Engineering of *Pichia pastoris* for improved production of antibody fragments. *Biotechnology and Bioengineering*, 2006, 94(2): 353–361.
- [14] Inan M, Aryasomayajula D, Sinha J, Meagher MM. Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. *Biotechnology and Bioengineering*, 2006, 93(4): 771–778.
- [15] Elowitz MB. Stochastic gene expression in a single cell. *Science*, 2002, 297(5584): 1183–1186.
- [16] Hansen KD, Wu ZJ, Irizarry RA, Leek JT. Sequencing technology does not eliminate biological variability. *Nature Biotechnology*, 2011, 29(7): 572–573.
- [17] Robasky K, Lewis NE, Church GM. The role of replicates for error mitigation in next-generation sequencing. *Nature Reviews Genetics*, 2014, 15(1): 56–62.
- [18] Schulze SK, Kanwar R, Gölzenleuchter M, Therneau TM, Beutler AS. SERE: single-parameter quality control and sample comparison for RNA-Seq. *BMC Genomics*, 2012, 13: 524.
- [19] Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y. KEGG for linking genomes to life and the environment. *Nucleic Acids Research*, 2008, 36(suppl_1): D480–D484.
- [20] Daly R, Hearn MTW. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *Journal of Molecular Recognition*, 2005, 18(2): 119–138.
- [21] Montoliu-Gaya L, Esquerda-Canals G, Bronsoms S, Villegas S. Production of an anti-A β antibody fragment in *Pichia pastoris* and *in vitro* and *in vivo* validation of its therapeutic effect. *PLoS ONE*, 2017, 12(8): e0181480.
- [22] Guan B, Chen FX, Su S, Duan ZY, Chen Y, Li HZ, Jin J. Effects of co-overexpression of secretion helper factors on the secretion of a HSA fusion protein (IL2-HSA) in *Pichia pastoris*. *Yeast*, 2016, 33(11): 587–600.
- [23] Cheng H, Wang L, Wang CC. Domain a' of protein disulfide isomerase plays key role in inhibiting α -synuclein fibril formation. *Cell Stress and Chaperones*, 2010, 15(4): 415–421.

Effect of PDI overexpressing on secretion of exogenous amyloid protein in *Pichia pastoris* —take cystatin as an example

Tingting Niu[#], Xuejie Zhou[#], Xiaobo Yu, Yisheng Cai, Jing Chen, Rui Wang, Shuang Fu, Chuang Yi, Jianwei He^{*}

School of Life Sciences, Liaoning University, Shenyang 110036, Liaoning Province, China

Abstract: [Objective] The aim of this study was to investigate the underlying mechanism for the regulation of protein disulfide isomerase (PDI) in protein folding and expression of exogenous proteins in yeast by co-expression of yeast PDI and its substrate protein chicken cystatin (cC). RNA-seq was carried out to screen for genes that are differentially expressed in *Pichia pastoris* (*P. pastoris*) and identify the key genes that could affect the expression of cC. This might provide theoretical support for the mechanistic analysis of high-level expression of exogenous protein and the construction of high-yield heterologous protein-expression strains. [Methods] PDI coding gene was transferred into GS115 and GS115 cC strains for the overexpression of PDI. RNA-seq was carried out to study the transcription of gene differences of two recombinant *P. pastoris* strains. Combined with the result of KEGG annotation, the data was analyzed to identify the differentially expressed genes between control and PDI/cC co-overexpressed cells, the result were verified by qRT-PCR and the functions in protein expression regulation were also clarified. [Results] The expression of cC was significantly increased in the PDI-overexpressing strain. A total of 373 differentially expressed genes were screened by RNA-seq analysis. Among them, 122 differentially expressed genes were assigned to the KEGG biological pathways, including 12 genes annotated into protein transport and catabolism pathway, 21 genes related to the protein folding sorting and degradation pathway, and 24 genes involved in the protein translation pathway. [Conclusion] Overexpression of PDI in *P. pastoris* significantly increased the expression of the heterologous amyloidogenic protein cC. By analyzing the expression profiles of cC-overexpressed normal *P. pastoris* cells or PDI-overexpressed strain that co-overexpressed cC, some of the genes with significant transcriptional changes were preliminarily identified, laying a foundation for the transformation of yeast strains with high expression of amyloid protein.

Keywords: *Pichia pastoris*, chicken cystatin C, protein disulfide isomerase, transcriptome sequencing

(本文责编: 李磊)

Supported by the National Natural Science Foundation of China (31670103)

[#]These authors contributed equally to this work.

^{*}Corresponding author. Tel/Fax: +86-24-62202232; E-mail: jwhe@lnu.edu.cn

Received: 23 January 2021; Revised: 30 March 2021; Published online: 6 August 2021