



针对 DNA 长片段靶标的锌指筛选体系

刘鹏^{①②}, 朱涛^{①②}, 梁忱^{①②}, 白冰^{①②}, 高冠军^①, 管吉松^{①②*}

① 清华大学生命科学学院分子神经生物学实验室, 北京 100084;

② 清华大学-北京大学生命科学联合中心, 北京 100084

* 联系人, E-mail: jsguan@mail.tsinghua.edu.cn

收稿日期: 2014-04-16; 接受日期: 2014-05-22

国家自然科学基金(批准号: 31371059, 31171008)资助项目

doi: 10.1360/052014-77

摘要 锌指蛋白能够特异性识别目标 DNA 序列, 常被作为分子靶向因子用于定点核酸编辑以及定点转录调控等方面, 具有十分广泛的应用前景. 然而, 目前常规采用的实验方法得到的锌指蛋白通常结合的目标 DNA 序列较短, 结合能力不强, 因而一直难以高效运用在核酸序列与调控蛋白在基因组上的原位互作等方面的研究中. 为了解决这个问题, 本研究构建了一个高通量筛选系统, 利用 N4 噬菌体 gp8 毒蛋白和 LacZ 作为报告基因, 以 300 bp 以上的 DNA 长片段为靶标, 来筛选能够结合多位点的锌指蛋白组合, 提高锌指蛋白应用的精确度以及效率. 该系统针对小鼠 *Nrxn-1*α 启动子区域进行了锌指蛋白文库筛选, 得到了具有序列选择特异性的混合锌指蛋白库, 并对筛选结果进行了初步功能验证. 研究表明, 本系统具有简便快捷的特点, 不仅大幅度缩短了筛选时间, 而且减少了因靶标序列 DNA 片段较长而不得不反复设计多位点结合锌指蛋白造成的成本浪费; 筛选得到的锌指蛋白库具有较高的长片段 DNA 靶标结合能力和一定的序列特异性, 并且能够在真核细胞内特异地结合目标 DNA 序列. 因此, 本研究建立的新型锌指筛选系统不仅可以广泛应用于高通量筛选, 而且在 DNA-蛋白相互作用的研究中也具有重要意义.

关键词

锌指蛋白

单杂交筛选

DNA-蛋白相互作用

锌指蛋白是真核细胞中广泛存在的一类 DNA 结合蛋白, 由可识别特异性 DNA 序列的保守的锌指单元域和发挥调控作用的效应结构域组成. 根据其保守结构域的特点, 可将锌指蛋白分为 C₂H₂, C₄, C₆ 型, 其中最为广泛研究的为 C₂H₂ 型, C₂H₂ 型锌指能够特异性识别一个三联体碱基, 常被用于作为人工锌指设计的模板. 每个锌指单元包含一个 α 螺旋和 β 片层结构, 通过 α 螺旋-1, 3 和 6 位的氨基酸与 DNA 发生作用^[1,2].

利用锌指单元的 DNA 结合能力, 将串联的锌指与转录因子或限制性内切酶相结合, 可分别用作锌指转录因子(zinc finger transcription factors, ZF-TFs)和锌指核酸酶(zinc finger nucleases, ZFNs), 从而进行基因表达控制或基因编辑, 具有重大的科研与医疗价值^[3-8]. 锌指转录因子通过锌指蛋白与转录激活因子或转录抑制因子的融合表达, 从而特异性地调控内源基因的转录水平^[5,8]. 目前, 常用的转录激活因子有单纯性疱疹病毒(herpes simplex virus)的

引用格式: 刘鹏, 朱涛, 梁忱, 等. 针对 DNA 长片段靶标的锌指筛选体系. 中国科学: 生命科学, 2014, 44: 1061-1072

Liu P, Zhu T, Liang C, et al. A one-step screening system for multi-zinc finger proteins targeting a long-DNA sequence. SCIENTIA SINICA Vitae, 2014, 44: 1061-1072, doi: 10.1360/052014-77

VP16(也称 Vmw65, herpes simplex virus protein vmw65)结构域和 NK- κ B 的 P65 亚基等;常用的转录抑制因子有 KOX1(又称 ZNF10, zinc finger protein 10)的 KRAB(Krüppel associated box)亚基和 Mad mSIN3 的结合结构域等^[8]。锌指转录因子在许多内源基因如 *ERBB2*(v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2)与 *ERBB3*(v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3)的表达调控中都取得了较好的成效^[4,5]。锌指核酸酶技术则是将一对靶位点相距 5~7 bp 的锌指蛋白分别与限制性内切酶 *Fok* I 融合表达,当两个锌指蛋白分别与各自的靶位点结合后,*Fok* I 二聚化产生酶切活性继而定点剪切双链 DNA,然后利用细胞内的同源重组(homology-directed repair, HDR)或非同源末端连接(nonhomologous end joining, NHEJ)这两种修复机制,来实现对目标基因的插入、敲除或突变,该技术的产生有效解决了传统的基因靶向技术效率低的问题。在基础科研的许多研究领域,利用锌指核酸酶技术已经能够对果蝇(*Drosophila melanogaster*)、线虫(*Caenorhabditis elegans*)和小鼠(*Mus musculus*)等多种模式生物进行各种成功的基因修饰。在临床应用方面,锌指核酸酶已经成为一种非常有力的分子靶向药物。在抗 HIV(human immunodeficiency virus)的治疗中,通过锌指蛋白设计能够结合 HIV 受体 CCR5(C-C chemokine receptor type 5)并阻断其生物活性的酶,目前已经能够对原始 T 细胞和造血干细胞上 CCR5 进行特异性失活,从而激活 T 细胞抵抗 HIV 的能力,使得此类细胞不再成为 HIV 病毒的攻击对象^[7]。锌指核酸酶也可用于改进 T 细胞免疫疗法的治疗效果,即通过锌指核酸酶技术使 T 细胞受体编码基因失活,提高肿瘤特异性的 T 细胞的获得效率。此外,锌指核酸酶也可与诱导多能干细胞(induced pluripotent stem cell, iPS cell)等技术相结合,从而在自身器官移植等治疗手段中中扮演重要的角色。

但是,制约锌指蛋白应用的一个重要限制因素是“脱靶效应”(off-target effect),即锌指蛋白由与基因组内的其他位点进行非特异性结合,而造成错误切割,DNA 区域产生突变,甚至可能出现较大的细胞毒性等后果。因此,如何在短时间内获得与靶位点结合能力强且特异性强的锌指蛋白成为一个亟待解决的问题。

目前获得针对特定靶序列的锌指单元的途径主要有模块化装配法^[9~12]和文库筛选法(开源法) 2 种。文库筛选法因为能够获得足够数量的候选分子而受到广泛的关注^[9]。迄今为止,已经建立的多种锌指单元的筛选方法,在实际应用中均存在一定的不足。体外筛选方法,如噬菌体展示和指数富集配基的系统进化(systematic evolution of ligands by exponential enrichment, SELEX)都需要做多轮筛选,工作量大且耗时较长^[13~19];芯片技术需要特定的专业知识和分析技能,从而应用范围受到限制^[20~22];其他的基于细菌或者酵母的单杂交或双杂交系统都源于 Hochschild 实验室构建的体系^[23~25]。在这类系统中,锌指结构域与 RNA 聚合酶的一个亚基结合,而报告载体的 DNA 靶序列位于启动子上游。如果锌指单元能够特异性地结合靶序列,则会招募 RNA 聚合酶,激活下游的报告基因表达^[26~30]。然而目前这类系统的靶序列设定主要仍为小片段 DNA,以长片段 DNA 为靶序列的筛选方法并不多见。

近年来,DNA-蛋白相互作用研究领域的成果越来越多地提出了进一步在细胞内进行互作研究的必要性。尤其针对 DNA 特异性区域的染色质纯化研究,目前通常采用的方法是在目标区域插入外源性的多个乳糖操纵子(LacO),再通过亲和层析得到目标区域相关 DNA 与其相互作用的蛋白。这个方法由于需要在原有 DNA 中加入外源性乳糖操纵子,有可能会破坏原有的染色质环境,因此一直受到各种质疑。理想的替代方式是采用 DNA 序列特异性结合蛋白代替乳糖操纵子系统来纯化目标染色质区域。然而,由于所使用的多是单个锌指单元,其结合力有限,同时单个单元的靶序列在细胞中也可能会被蛋白包裹从而导致这种替代方法中的 DNA 序列特异结合蛋白能力不足,无法满足 DNA-蛋白相互作用领域的更多研究要求。所以需要采用较长的靶序列,增加锌指蛋白结合的可能性,也同时增强锌指蛋白的结合力。

一直以来,噬菌体编码蛋白都被广泛用于文库的筛选系统。其中 gp8 是一个由大肠杆菌 N4 噬菌体编码的 71 个氨基酸的肽段^[31]。gp8 能够结合大肠杆菌 DNA 聚合酶 III 的滑动钳装载复合物,并在感染 30 min 后抑制细菌的 DNA 复制^[31]。上述特点使得 gp8 能够作为细菌单杂交系统的筛选报告基因。为了实现长片段靶序列的锌指筛选,本研究以 gp8 作为正筛选报告基因,以 β -半乳糖苷酶(β -galactosidase,

*LacZ*基因作为负筛选报告基因构建了单杂交筛选系统. 本研究采用 Barbas 模块(Barbas Modules)构建人工四锌指蛋白随机文库^[32,33], 并以小鼠 *Nrxn-1α*基因启动子区域为靶序列进行筛选. 通过一轮正筛选和一轮负筛选获得有选择性的锌指蛋白混合文库, 并利用生物化学和细胞生物学等手段验证了其结合能力和特异性. 实验结果表明, 筛选获得的锌指蛋白对靶序列具有较高结合能力和一定的特异性.

1 材料与方法

1.1 锌指随机文库与载体的构建

人工四锌指蛋白随机文库的构建方法同 Gonzalez 等人^[3], 库容约为 10^6 .

用于文库筛选和蓝白斑鉴定的锌指载体 pQE30-Amp-LacZ 由 pQE30(Qiagen, 德国)改造而成, *Lac* 启动子通过 pCR2.1-TOPO(Invitrogen, 美国)亚克隆获得(引物为 Plac-F: 5'-CATCTCGAGGCGCC-CAATACGCAAAC-3', Plac-R: 5'-ATGGAATTC-CTGTGTGAAATTGTTATC-3'), 多克隆位点由 1 对引物退火后经由 *Xma* I 和 *Spe* I 酶切位点连接而成(引物为 Mcs-F: 5'-GATCCGATCCCGGGACCGG-TACTAGTCTCGAGA-3', Mcs-R: 5'-AGCTTCTCG-AGACTAGTACCGGTCCCGGGATCG-3'). 正筛选使用的载体为 TOPO-gp8-LacZ, 由 *Acl* I (NEB, 美国)切除 pCR2.1-TOPO 载体的青霉素抗性基因后获得, 非特异序列位于启动子与 *LacZ* 编码序列之间. 对于负筛选使用的 TOPO-Gen-LacZ 载体, 庆大霉素基因从 pFAST-bac 载体 PCR 得到(引物序列 Gen-F: 5'-AGCCGGCCGCTATGTTACGCAGCAGCAACGAT-3', Gen-R: 5'-AGCCCATGGTTAGGTGGCGGTACTTGGG-3'), 并通过 *Eag* I 与 *Sph* I 酶切位点插入到载体中, 同时也去除了卡那霉素抗性基因. 锌指蛋白表达的载体为 pGEX-4TM, 通过把 pQE30-LacZ 的多克隆位点连入 pGEX-4T-1 载体获得. 正筛选报告载体 TOPO-gp8 是将 *gp8* 人工合成基因(GENEWIZ, 北京)克隆至 TOPO-LacZ 的 *Xho* I 和 *Xba* I 之间获得. 靶序列位于 *Bam*H I 和 *Xho* I 之间. 细胞实验使用载体为 pEGFP-N3, 锌指与 P65 转录活性蛋白激活结构域融合表达(位于 *Xho* I 与 *Kpn* I 之间)^[34], 锌指位于 C 端, 与 P65 之间有一个 SV40 核定位信号.

1.2 蛋白表达与纯化

谷胱甘肽 S-转移酶(glutathione S transferase, GST)与锌指的融合蛋白按如下方法表达: 含有锌指表达载体的 BL21(DE3)感受态在 LB(Luria-Bertani)培养基中(氨苄青霉素 50 μg/mL)经过 37℃培养至 A_{600} 为 0.6, 加入诱导剂异丙基-β-D-硫代半乳糖苷(isopropyl β-D-1-thiogalactopyranoside, IPTG, 购自 Sigma-Aldrich, 美国)至终浓度为 1 mmol/L, 并在 16℃低温诱导 16 h. 细胞收集后在低盐缓冲液(20 mmol/L Tris-HCl, 150 mmol/L NaCl, pH 7.0)中重悬. 超声裂解后, 18516×g 离心 20 min 获得上清, 并按照 glutathione-sepharose(Thermo Scientific, 美国)的标准程序进行纯化.

1.3 EMSA 实验

靶标区域 *Nrxn-1α*(-7)与 *Nrxn-1α*(+923k)的 DNA 探针由 M13 引物(序列为 M13 F: 5'-GTAAAACGACGGCCAGT-3', M13R: 5'-GGAAACAGCTATGACC-ATG-3')从相应负筛选报告载体中扩增获得, 长度为 136 bp, *Nrxn-1α*(-357)靶序列片段与负对照片段(enhanced green fluorescent protein, EGFP)片段分别从小鼠基因组与 pEGFP-N3 载体 PCR 获得(引物分别为 P435F: 5'-AGTGCTAGCACTGTCAGGAAAGC-ACCTTGGGAG-3', P435-R: 5'-GTACTCGAGAGAG-CCTACTGAGCATGCCCAAGC-3', EGFP-F: 5'-ATG-GTGA GCAAGGGCGAGGAGC-3', EGFP-R: 5'-GTT-GTACTCCAGCTTGTGCCCCAG-3'). EMSA 反应体系为 15 μL, 缓冲液含有 50 mmol/L Tris, 150 mmol/L NaCl, pH 8.0, 5% 甘油, 1 mmol/L 二硫苏糖醇(dithiothreitol, DTT)和 0.1% NP-40(NonidetP-40). 锌指蛋白与 30~50 ng DNA 探针 37℃孵育 0.5 h 后, 用 8%非变性聚丙烯酰胺胶在 0.5×Tris-硼酸缓冲液中电泳 2 h, 溴化乙锭(ethidium bromide, EB)染色成像.

1.4 染色质 pull-down 实验

实验材料为 C57B6/L 品系的成年小鼠, 均来自清华大学实验动物中心. 将小鼠用 2%的戊巴比妥钠以 0.01 mL/g 腹腔注射, 待小鼠麻醉后先用心脏灌流, 再用 4%多聚甲醛固定. 取其内脏组织, 匀浆后先用 2%多聚甲醛室温固定 10 min, 再用 0.125 mol/L 甘氨酸室温孵育 5 min 终止反应. 将组织超声裂解后, 离

心并将上清加至结合有筛选锌指蛋白的 glutathione-sepharose 柱, 4℃ 孵育过夜后以 15 mmol/L 还原型谷胱甘肽(Thermo Scientific, 美国)洗脱. 洗脱样品解交联后用 PCR 进行半定量检测(引物为 357-F: 5'-CAGTGCAGGTGTTGAGTCC-3', 357-R: 5'-CGG-GATTTAGCCGTGATTGG-3', 1058K-F: 5'-TCGAG-GCACAACACAAAACC-3', 1058K-R: 5'-AATGCAA-GTTCCAACCCCTG-3').

1.5 正筛选系统的验证

对于 gp8 的抑菌实验, 将 50 ng 正筛选报告载体 TOPO-gp8-Nrxn-1 α (-7)转化至 BL21(DE3)菌株, 复苏后以 1/100 接种于 5 mL 含有 0.05 mg/mL 卡那霉素和 IPTG 的 LB 培养基, 培养 10.5 h 后用分光光度计测定菌液 A_{600} 数值.

对于正筛选系统的验证实验, 将含有相应锌指载体和正筛选报告载体的菌 37℃ 培养过夜后接种于 5 mL LB 培养基(含 0.05 mg/mL 氨苄青霉素, 0.05 mg/mL 卡那霉素以及 IPTG), 17~20 h 后测定 A_{600} , 并用 BX51WI (Olympus, 日本)相差显微镜在明场以 40 $\times$ 物镜拍摄菌体形态.

1.6 负筛选系统的验证

将 Nrxn-1 α (-7)的锌指载体和负筛选报告载体共同转化至 DH5 α 感受态细胞中. 热激复苏后, 将菌液用微量移液器点在 60 mm 培养皿上, 其中含有 0.05 mg/mL 氨苄青霉素, 0.015 mg/mL 庆大霉素, 1 mmol/L IPTG 以及 0.2 mg/mL 5-溴-4-氯-3-吡啶- β -D-半乳糖苷(5-bromo-4chloro-3-indolyl- β -D-galactopyranoside, X-gal, 购自 Sigma-Aldrich, 美国), 室温培养 32 h 左右, 直至菌落颜色发生变化.

1.7 锌指文库的筛选

首先将含有靶序列的正筛选报告载体转化至 BL21(DE3)感受态细胞, 并制成电击感受态.

在正筛选过程中, 将四锌指载体库电击转化至含有报告载体的电击感受态中, 并在 100 mL 含有 0.05 mg/mL 氨苄青霉素, 0.015 mg/mL 庆大霉素以及 0.8 mmol/L IPTG 的 LB 培养基中以 200 r/min 培养 17~20 h. 通常情况下, 100 mL 培养基中有 $10^5 \sim 10^6$ 个转化子. 收集菌体后, 提取载体, 并与含有非特异序列的负筛选报告载体共转至 DH5 α 感受态进行负筛.

在 150 mm $\times$ 150 mm 的 LB 平板上(成分同 1.6 负筛选系统验证)约有 2×10^3 个菌落. 培养 32 h 后, 挑选 400 个蓝色菌落在终浓度 0.05 mg/mL 氨苄青霉素的 LB 中培养, 提取质粒后, 用 *Kpn* I 限制性内切酶去除报告载体. 将酶切产物转化, 并将复苏的菌直接扩大培养, 即为最终的负筛混合锌指文库. 整个筛选过程大约需要 3 天.

1.8 RT-PCR 与 Real time-PCR

将鼠源 NG108-15 细胞系培养至汇合度 40%~60% 时, 用 VigoFect 试剂(Vigorous Biotechnology, 北京)转染 5 μ g 真核表达载体. 30~40 h 后收取细胞, 根据 TRIzol(Invitrogen, 美国)试剂说明书提取总 RNA, 并将其中 1 μ g 用 TransScript® First-Strand cDNA Synthesis SuperMix(TransGen, 北京)试剂进行反转录(42℃, 1 h). Real time-PCR 按照 SsoFast EvaGreen Supermix(Bio-Rad, 美国)说明书进行. 扩增条件为: 95℃ 3 min; 95℃ 15 s, 62℃ 20 s, 72℃ 20 s, 45 个循环. 分别检测 *Nrxn-1 α* , *Nrxn-2*, *Nrxn-3* 和 β -tublin 的表达量, 实验结果用 $2^{-\Delta\Delta C_T}$ 法进行分析, 并用 *gapdh* 作为内参(所用引物分别为 Nrxn-1 α -F: 5'-CCAGCACAACCTGCCAAGAGGATTC-3', Nrxn-1 α -R: 5'-TGGGGCGGTCATTGGGAGGCCAC-3', Nrxn-2-F: 5'-GGCTCACCTGACGTTAAACA-3', Nrxn-2-R: 5'-TCCTTGCCCTTTTGTTCGGCTG-3', Nrxn-3-F: 5'-GTAGAGGAGTGGCTGCAGGA-3', Nrxn-3-R: 5'-CTTCTCCCCTAGCCGGACA-3', β -tublin-F: 5'-TAGACCCCAGCGGCAACTAT-3', β -tublin-R: 5'-GTTCCAGGTTCCAAGTCCACC-3', *gapdh*-F: 5'-GGTGAAGGTTCGGTGTGAACG-3', *gapdh*-R: 5'-CTCGCTCCTGGAAGATGGTG-3').

2 结果

2.1 基于 gp8 和 LacZ 的细菌单杂交筛选系统

本研究首先检测了不同的菌株在表达 gp8 后的生长情况发现, 只有 BL21(DE3)菌株具有较高的 gp8 敏感性(表 1). 对 BL21(DE3)菌株内的抑菌生长实验表明, 通过测量菌液 A_{600} , BL21(DE3)的生长能够受 gp8 的抑制, 在 0.04 mmol/L IPTG 诱导下的 gp8 表达可完全抑制其生长, BL21(DE3)被用作正筛的宿主(图 1).

在正筛选中, 在 gp8 蛋白前构建 DNA 靶标序列, 组成报告载体. 将锌指库转化至含有报告载体的 BL21(DE3)菌株. 当锌指蛋白结合到靶序列上时, gp8 的表达受到抑制, 这样具有靶序列结合能力的锌指蛋白库得到大量扩增(图 2).

在负筛选中, 含有非特异 DNA 序列的负筛选报告载体与正筛选获得的锌指载体同时转入 DH5α 菌株中. 如果锌指蛋白没有结合非特异序列, 表达的 LacZ 蛋白能够降解培养基中的底物 X-gal, 使菌落呈现蓝色; 如果锌指蛋白能够与非特异序列结合, LacZ 的表达将会受到抑制, 菌落呈现白色. 通过负筛选, 进一步挑选具有较高序列特异性的锌指蛋白库(图 2).

整个筛选过程仅需 3 天左右即可完成. 同目前的锌指筛选体系相比, 本系统更为方便快捷.

在正筛选过程中, 将四锌指库(氨苄青霉素抗性)电转化至含有报告载体(卡那霉素)的 BL21(DE3)细胞, 并在含有 IPTG 的培养基中培养 17~20 h. 若锌指结合报告载体中的靶序列, 则会抑制 gp8 的表达, 菌体得以正常生长. 将菌体收集并提取载体, 与含有非特异序列的负筛选报告载体(庆大霉素抗性)共同转化至 DH5α 细胞中进行负筛选. 特异性较好的锌指不与非特异序列结合, LacZ 正常表达. 培养 32 h 后, 收集

表 1 宿主菌的筛选

菌种	是(√) 否(×)受 gp8 抑制
DH5α	×
TOP10	×
JM110	×
Rosetta	×
BL21(DE3)	√

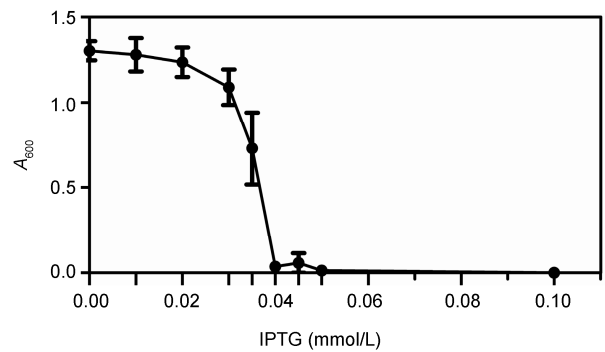


图 1 gp8 蛋白的抑菌作用

表达 gp8 的 BL21(DE3)菌株在含有不同 IPTG 浓度的 LB 培养基中进行培养, A₆₀₀ 在 10.5 h 进行测定. 数据显示为 $\bar{x} \pm \text{SD}$, $n=3$ 个生物学重复

蓝色菌落培养, 提取质粒后, 用 Kpn I 限制性内切酶特异性去除报告载体, 即得到最终的负筛选混合锌指库. 整个筛选过程大约需要 3 天.

2.2 筛选系统的验证

为了验证筛选系统的可行性, 本研究利用 ZiFiT(zinc finger targeter)网络平台进行锌指序列设计, 并采用模块化装配法针对小鼠的 Nr1h3-1α 构建了 2 个六锌指蛋白(表 2), 其中 Nr1h3-1α(-7)结合启动子区域, Nr1h3-1α(+923k)结合 exon20 区域(标记位置相对于转录起始位点, 以靠近起始位点一侧命名). 电泳迁移率实验(electrophoretic mobility shift assay, EMSA)结果表明, Nr1h3-1α(-7)与 Nr1h3-1α(+923k)都能特异性结合各自的靶位点(图 3A 和 B).

在正筛选体系的验证中, 空载体、Nr1h3-1α(-7)锌指载体以及 Nr1h3-1α(+923k)锌指载体分别与 TOPO-gp8-Nr1h3-1α(-7)报告载体转入 BL21(DE3)菌株, 并在含有不同浓度 IPTG 的培养基中进行培养. 在 0.6~1 mmol/L 的 IPTG 浓度范围内, 宿主的繁殖能够被 Nr1h3-1α(-7)锌指拯救. 在转入空载体或 Nr1h3-1α(+923k)的情况下, 细菌的生长受到明显抑制, 说明 Nr1h3-1α(-7)锌指能够结合靶序列, 抑制 gp8 合成(图 3C). 显微成像的结果也进一步证实了这个结论(图 3D).

在负筛选体系的验证中, 空载体, Nr1h3-1α(-7)锌指载体以及 Nr1h3-1α(+923k)锌指载体分别与 TOPO-Gen-LacZ-Nr1h3-1α(-7)报告载体转入 DH5α 菌株, 并在含有 IPTG 和 X-Gal 的 LB 平板上培养. 实验结果表明, 只有转入 Nr1h3-1α(-7)锌指和 TOPO-Gen-LacZ-Nr1h3-1α(-7)报告载体的菌落呈现白色, 说明 Nr1h3-1α(-7)锌指能够与靶序列结合; 而转入空载体或 Nr1h3-1α(+923k)锌指的菌落呈现蓝色, 因此 Nr1h3-1α(+923k)不能与 Nr1h3-1α(-7)的靶序列结合(图 3E). 因此, 如果在负筛选报告载体中插入非特异序列, 即可根据颜色变化去除特异性较低的锌指.

2.3 针对 Nr1h3-1α 启动子特定位点的锌指筛选

为了对小鼠 Nr1h3-1α 启动子区域进行锌指筛选, 本研究构建了一个大小约 10⁶ 的人工四锌指随机库, 靶序列长度为 435 bp(-357 位, 图 4). 在正筛选中, 锌指库被电击转化至含有正筛选报告载体 TOPO-gp8-Nr1h3-1α(-357)的 BL21(DE3)电击感受态, 并在含有

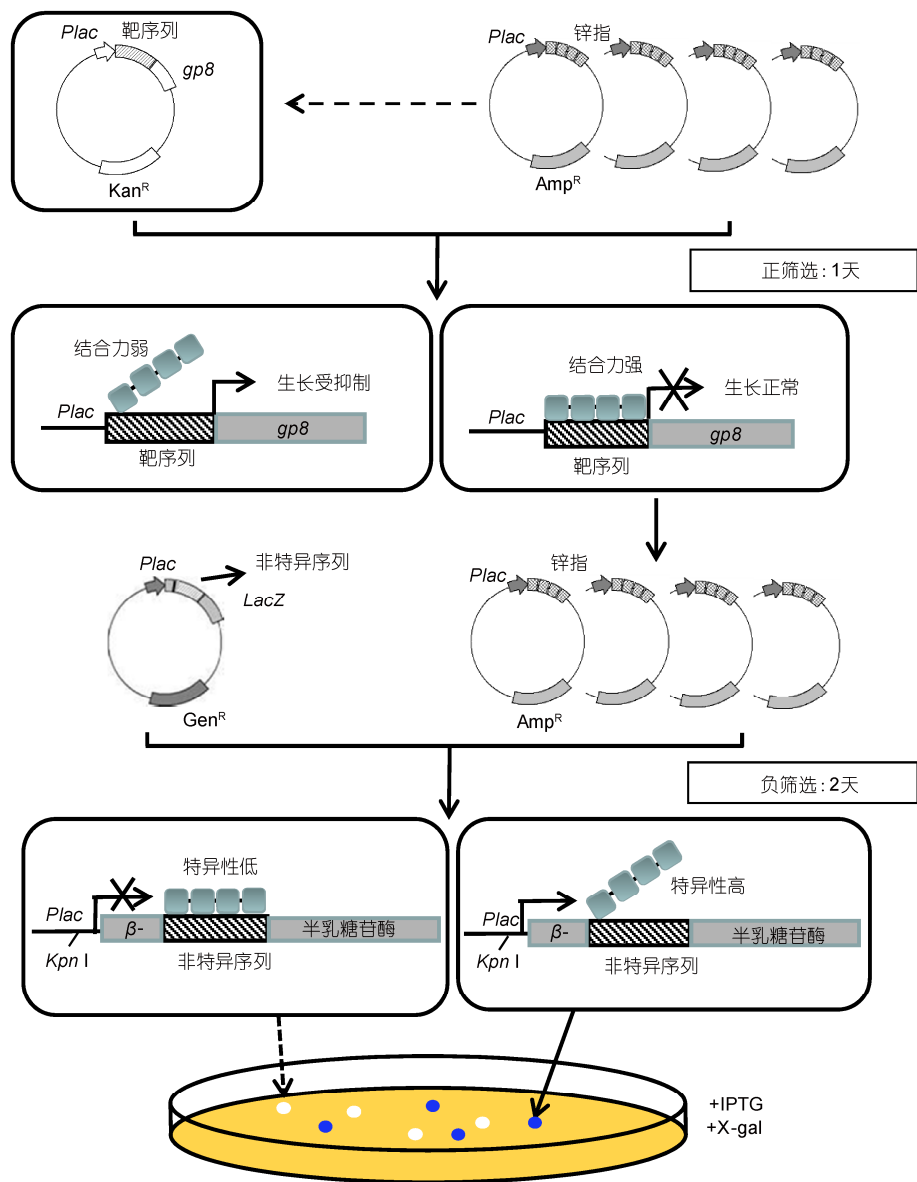


图 2 本研究构建的锌指筛选系统示意图

表 2 Nr₁α(-7)与 Nr₁α(+923k)锌指蛋白的 DNA 识别区

	锌指 1	锌指 2	锌指 3	锌指 4	锌指 5	锌指 6
Nr ₁ α(-7)	DKSCLNR	QKGNLLR	WPSNTLR	WPSNTLR	DPGHLVR	QLANLRA
Nr ₁ α(+923k)	HKNALQN	TTGNLTV	HRTTLTN	ISSNLQR	QSSSLIR	TTGNLTV

IPTG 的培养基中进行培养。培养获得的菌体在提取锌指载体后, 又用化学法转化至含有负筛选报告基因的 DH5α 菌株进行负筛选。在平板上挑取 400 个蓝色菌落培养后, 即获得负筛的混合四锌指库。

同时, 为了探究筛选锌指的多样性, 本研究将 14 个负筛克隆进行测序, 总共获得 7 个不同的四锌指序列, 实验结果表明, 本研究的筛选方法能够在一次筛选中获得种类较多的锌指蛋白(表 3)。

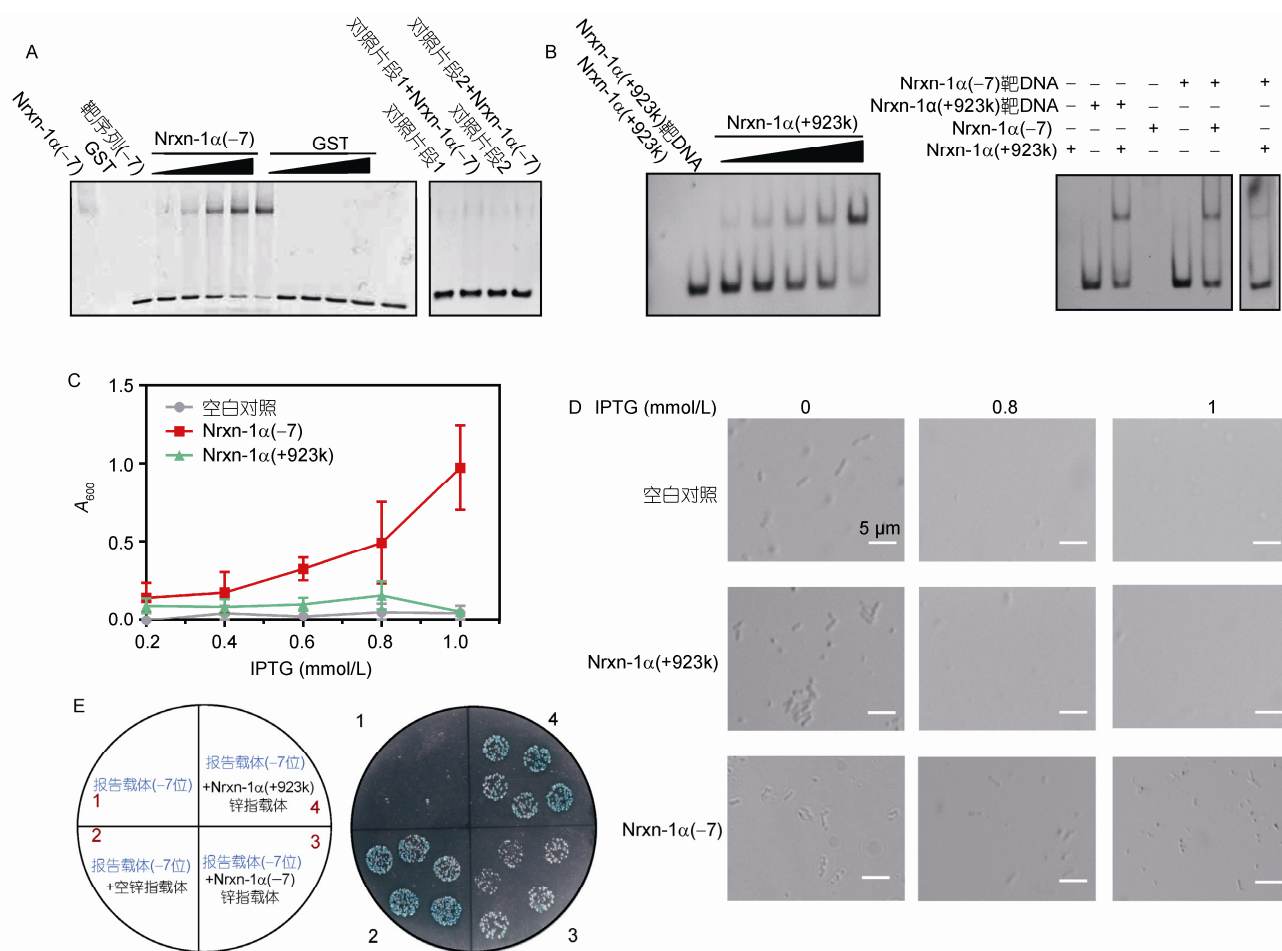


图3 锌指筛选系统的验证

A, B: Nrnx-1α(-7)与 Nrnx-1α(+923k)能够特异性结合自身靶序列。A 中对照片段中的靶序列为 Nrnx-1α的 2 段 18 bp 片段(序列分别为 5'-ATTAACATAGGAATAATT-3', 5'-GAAGTGCCTTGTGGCAAGT-3'), DNA 用量为 30 ng。Nrnx-1α(-7), Nrnx-1α(+923k)与 GST 浓度梯度为 0.04, 0.1, 0.2, 0.3, 0.4 μmol/L。图 A 左图前 2 泳道和图 B 左图第 1 泳道蛋白浓度为 0.4 μmol/L。图 A 和 B 右图的蛋白浓度都为 0.25 μmol/L; C: BL21(DE3)细胞的生长拯救实验。A₆₀₀在 17 h 左右测定, 对照组为锌指空载体。其中对照组为 $n=3$ 个生物学重复; Nrnx-1α(-7)为 $n=5$ 个生物学重复; Nrnx-1α(+923k)为 $n=6$ 个生物学重复。数据显示为 $\bar{x} \pm \text{SD}$; D: C 中相应组的显微成像; E: DH5α细菌的蓝白斑实验。报告载体为含有 Nrnx-1α(-7)靶序列的负筛选载体

2.4 筛选锌指的结合能力与特异性验证

为了检测筛选锌指的结合能力与特异性, 本研究将筛选获得混合四锌指的编码序列连入原核表达载体 pGEX-4TM, 表达并纯化融合蛋白。在电泳迁移率实验(EMSA)中, GST 标签蛋白不能改变靶序列的迁移率, 而锌指融合蛋白能够靶序列条带明显上升, 从而说明筛选获得的锌指具有较高的结合能力(图 5A)。

为了进一步验证筛选系统的有效性, 本研究对 Nrnx-1α的+1058k 附近(3'非翻译区(untranslated region, UTR))一段 51 bp 序列(5'-GGTCAGCT-

GGCACAACTGGAGCAGGGGTTGGAAGTTGCATT TGAGACAAAG-3')进行了锌指筛选, 并通过 pull-down 实验与 Nrnx-1α(-357)筛选锌指进行了对比。在 pull-down 实验中, 带有 GST 标签蛋白的锌指能够与纯化柱结合, 锌指蛋白能够捕获细胞裂解液中的 DNA 片段(图 5B)。将融合蛋白洗脱后, 采用半定量 PCR 和 Real time-PCR 的方法对目的片段的富集进行检测。结果表明, 两组筛选获得的锌指蛋白的 pull-down 产物中, 对自身的靶序列均有 2~3 倍富集(图 5C~E), 然而对对方的靶序列富集虽然较少, 仍有一定程度富集。这种非特异性富集有可能源于体外纯

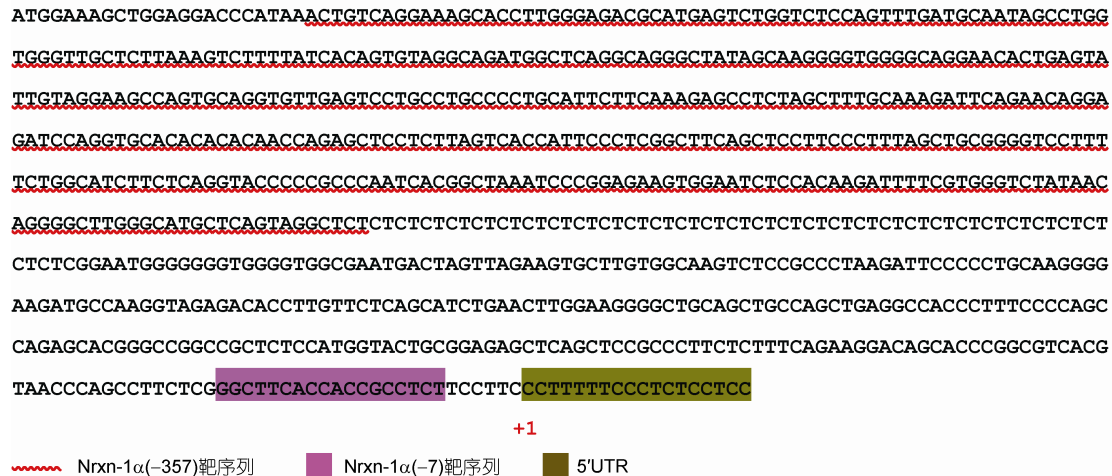


图 4 *Nrnx-1α*的启动子区示意图

表 3 筛选锌指蛋白 *Nrnx-1α*(-357)的 DNA 识别区

	锌指 1	锌指 2	锌指 3	锌指 4	拷贝数
1	RNDALTE	DCRDLAR	SKKALTE	SPADLTR	4
2	SRRTCRA	DCRDLAR	SPADLTR	SPADLTR	4
3	HKNALQN	RSDNLVR	SRRTCRA	DCRDLAR	2
4	DPGHLVR	QSSSLVR	SRRTCRA	RSDHLTN	1
5	HKNALQN	QKSSLIA	TTGNLTV	QKSSLIA	1
6	TTGNLTV	RTDTLRD	QKSSLIA	ERSHLRE	1
7	THLDLIR	TSGSLVR	HRTTLTN	RSDHLTN	1

化体系中的条件没有最优化,当然也可能由于所得到的锌指蛋白本身的选择性不够.

2.5 锌指转录因子的胞内功能检测

为了探究筛选锌指在细胞内的功能,本研究将筛选获得的锌指与 P65 转录激活结构域构建融合表达载体^[34,35],并转染鼠源的 NG108-15 细胞. RT-PCR 与 Real time-PCR 的结果表明,与空载体相比,锌指与 P65 的融合蛋白能够将内源的 *Nrnx-1α* 的表达提高 30% 左右,而单独表达锌指或者 P65 并不能使 *Nrnx-1α* 的表达发生显著变化,说明筛选锌指在胞内对特定定位点具有较强的特异性结合. 此外, β -tublin, *Nrnx-2* 和 *Nrnx-3* 的表达量没有显著升高,说明锌指蛋白与启动子区域的结合具有特异性(图 6).

3 讨论

3.1 基于 gp8 和 LacZ 的单杂交系统的优点

本研究构建的新型细菌单杂交锌指筛选系统具

有显著优点. 目前的筛选系统主要针对小片段 DNA 靶序列进行. 本系统能够针对长片段 DNA 靶序列筛选锌指蛋白,因为长片段 DNA 靶序列具有更多结合位点,筛选得到的锌指蛋白,与以小片段 DNA 为靶序列的筛选系统得到的锌指蛋白相比,具有更强的结合能力.

此外,同 SELEX 与 phage display 相比,本系统仅需一轮筛选即可获得结合能力强、特异性高的锌指,大幅缩短筛选时间,而且筛选是在细胞环境内进行,更好的保证了筛选锌指具有一定的生物活性;同酵母系统相比, *E. coli* 具有生长速度快(代时仅 20 min 左右,而酵母代时约为 90 min)、转化效率高(通常比酵母高 2~4 个数量级)和不需要核定位信号等优点. 与其他细菌单杂交或双杂交系统相比,本系统还有以下优势: (i) 对宿主菌和培养基没有特殊要求. 本系统所需宿主菌为 BL21(DE3)和 DH5 α 这两种常见的菌株,转化效率高,生长周期短. 实验所使用的培养基也是在各个实验室广泛使用的 LB 培养基,使用的药品成本较低. 相对而言,以 *URA3* (orotidine-

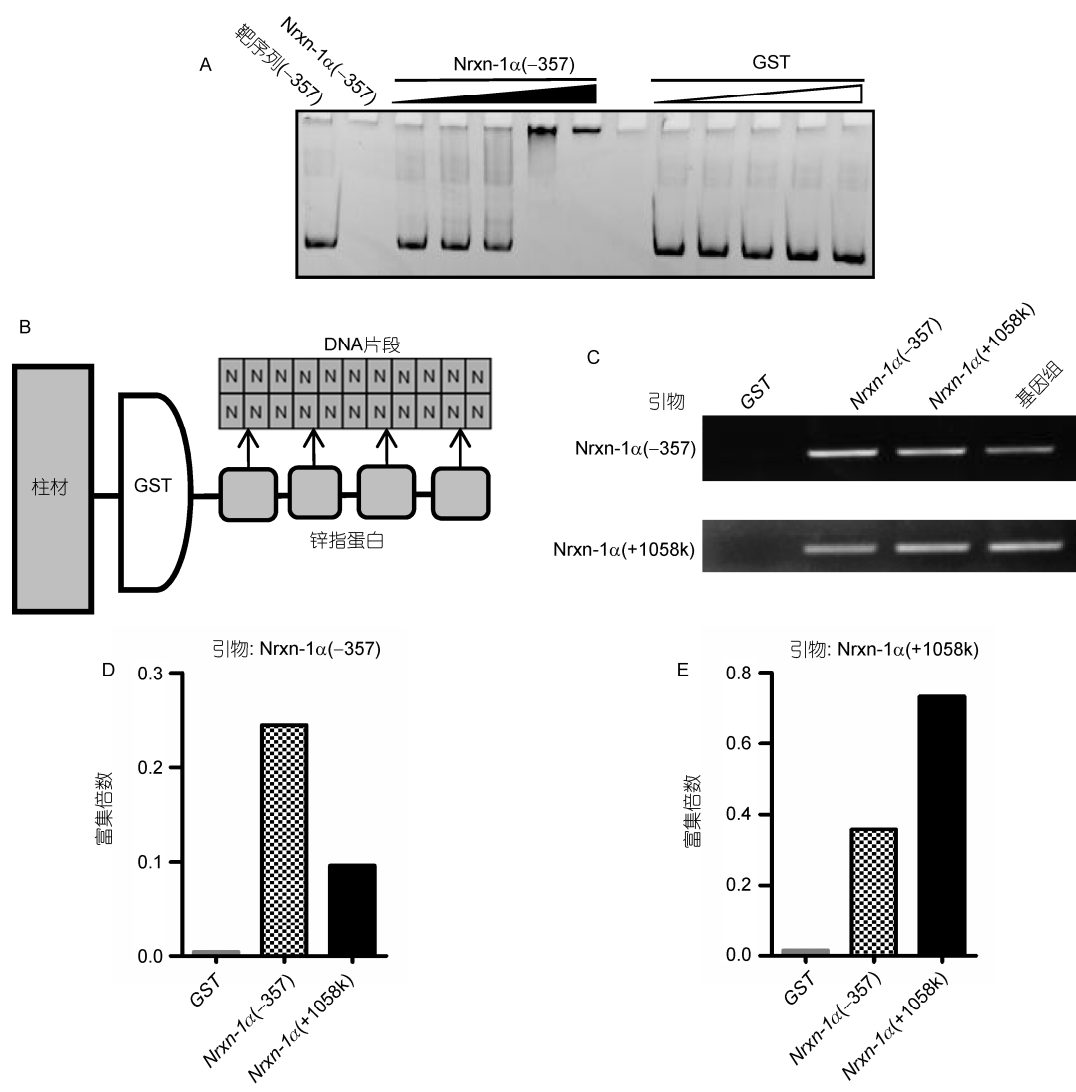


图 5 筛选锌指的结合能力的体外结合实验

A: 筛选锌指 *Nrnx-1α(-357)* 的 EMSA 实验。蛋白浓度梯度为 0.025, 0.0625, 0.125, 0.3125, 0.625 μmol/L; B: pull-down 实验原理示意图; C: *Nrnx-1α(-357)* 混合锌指与 *Nrnx-1α(+1058k)* 混合锌指 pull-down 产物的半定量 PCR 分析; D, E: *Nrnx-1α(-357)* 混合锌指与 *Nrnx-1α(+1058k)* 混合锌指 pull-down 产物的 Real time-PCR 分析

5'-phosphate decarboxylase), *HIS3*(imidazol- eglycerol-phosphate dehydratase)等作为筛选报告基因的筛选系统则需要特殊基因型的菌株(例如营养缺陷性菌株),而且培养基中也需要单独添加或去除某些组分(例如 5-氟尿嘧啶、组氨酸等); (ii) 现存的筛选系统主要针对较短的 DNA 靶序列, 通常不超过 20 bp, 而本系统可以筛选结合 300 bp 以上 DNA 靶序列的锌指。而且, 在筛选报告基因方面, *gp8* 具有很强的抑菌作用, 短时间内即可抑制大肠杆菌的繁殖, 能够有效降低假阳性率, 加之其编码序列十分简短, 在载体构建中操

作简便。

基于以上优点, 本研究构建的这一筛选系统能够成为适用于大多数生物学实验室的有效技术手段。

3.2 筛选系统的优化

本研究筛选获得的锌指蛋白虽然已具有较高的长片段 DNA 靶序列结合能力, 还是存在一定的非特异性, 可以通过以下几点进一步加以优化: (i) 提高蛋白纯度, 可以在 GST-锌指融合蛋白的纯化中, 提

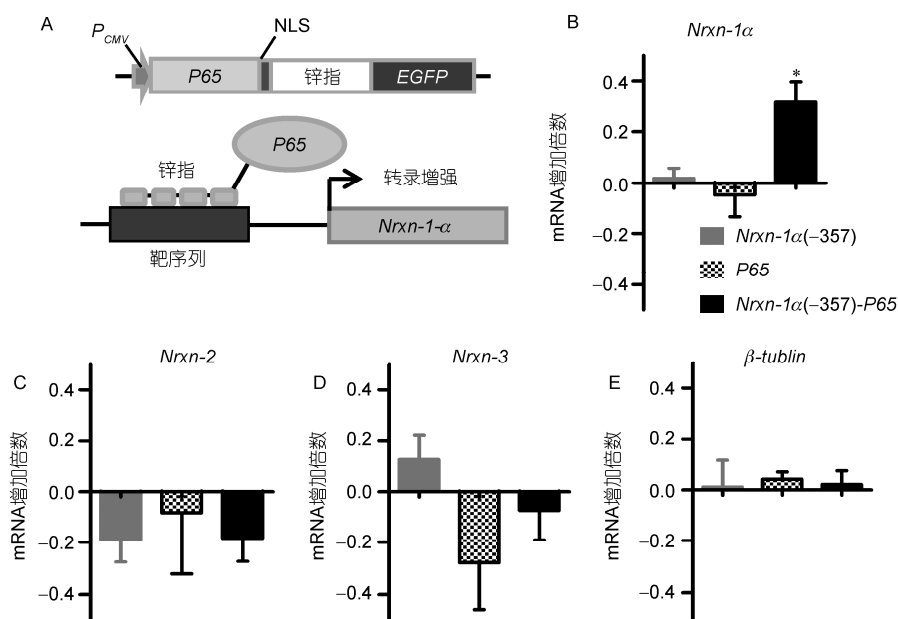


图6 锌指转录因子的胞内功能检测

A: 锌指转录因子的作用示意图; B-E: NG-108-15 细胞转染后的 Real time-PCR 分析. *gapdh* 为内参基因. $n=3$ 个生物学重复; *: $P<0.05$. NLS: nuclear localization signal, SV40 核定位信号. 数据显示为 $\bar{x} \pm SE$

高蛋白纯度并尽量去除特异性较低的锌指蛋白, 例如可以采用串联亲和纯化法(tandem affinity purification)提纯蛋白, 并在 pull-down 实验中通过加入鱼精 DNA 孵育来去除特异性低的锌指蛋白; (ii) 增加库容, 本研究使用的锌指库只包括部分四锌指蛋白, 因此可以通过增加库容提高特异性. 此外, 四锌指在基因组中的靶位点较多, 如果改为构建六锌指库进行筛选, 特异性会大大提高; (iii) 降低假阳性率, 在筛选过程中, 可以通过改变 IPTG 的浓度来调节正筛的严谨性, 从而增强锌指蛋白的结合能力; 在负筛选中, 也可以根据颜色的深浅来判断锌指蛋白结合特异性的高低.

3.3 筛选系统的应用前景

本系统筛选获得的锌指蛋白能够特异的结合长片段靶序列, 具有非常广泛的应用前景. (i) 用于 DNA 结合蛋白研究, 筛选获得的锌指蛋白因为对靶序列结合能力强, 用于 pull-down 实验, 可以富集较

多的目的片段, 有助于进一步研究了解靶序列上的其他结合蛋白; (ii) 用于基因表达调控研究, 筛选获得的锌指在细胞内可以结合转录因子, 调控细胞内源基因的表达. 常规筛选方法得到的单个锌指转录因子, 细胞内的靶位点可能已被其他蛋白占据, 因无法顺利结合靶位点, 单个锌指转录因子在细胞内的表达容易失去其应有的生物功能. 而本系统筛选获得的锌指因为结合位点多, 进入细胞后可以保证与靶序列的结合, 保持其生物功能完整性; (iii) 本系统也可用于其他生物学研究, 例如 DNA 与蛋白相互作用的检测.

4 结论

综上所述, 本研究成功构建了一个细菌单杂交锌指筛选系统. 该系统具有高通量、快速以及筛选在细胞环境进行等优点, 有助于进一步揭示锌指蛋白与 DNA 结合的内在规律, 更为大规模筛选和 DNA-蛋白相互作用的研究提供了一个高效的平台.

致谢 感谢清华大学生命科学学院的徐菁对本文初稿的编辑修改, 丁歆璐、柳三雄、江骏以及中国科学院上海有机化学研究所生命有机化学国家重点实验室的郑庆飞对本研究的实验方面的讨论和建议.

参考文献

- 1 Klug A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu Rev Biochem*, 2010, 79: 213–231
- 2 Miller J, Mclachlan A D, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus oocytes*. *EMBO J*, 1985, 4: 1609–1614
- 3 Gonzalez B, Schwimmer L J, Fuller R P, et al. Modular system for the construction of zinc-finger libraries and proteins. *Nat Protoc*, 2010, 5: 791–810
- 4 Beerli R R, Barbas C F 3rd. Engineering polydactyl zinc-finger transcription factors. *Nat Biotechnol*, 2002, 20: 135–141
- 5 Beerli R R, Segal D J, Dreier B, et al. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc Natl Acad Sci USA*, 1998, 95: 14628–14633
- 6 Isalan M. Zinc-finger nucleases: how to play two good hands. *Nat Methods*, 2012, 9: 32–34
- 7 Gaj T, Gersbach C A, Barbas C F 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*, 2013, 31: 397–405
- 8 Sera T. Zinc-finger-based artificial transcription factors and their applications. *Adv Drug Deliv Rev*, 2009, 61: 513–526
- 9 Sander J D, Dahlborg E J, Goodwin M J, et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods*, 2011, 8: 67–69
- 10 Sander J D, Zaback P, Joung J K, et al. Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic Acids Res*, 2007, 35: W599–W605
- 11 Sander J D, Maeder M L, Reyon D, et al. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res*, 2010, 38: W462–W468
- 12 Mandell J G, Barbas C F 3rd. Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res*, 2006, 34: W516–W523
- 13 Jamieson A C, Kim S H, Wells J A. *In vitro* selection of zinc fingers with altered DNA-binding specificity. *Biochemistry*, 1994, 33: 5689–5695
- 14 Choo Y, Klug A. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc Natl Acad Sci USA*, 1994, 91: 11163–11167
- 15 Rebar E J, Pabo C O. Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science*, 1994, 263: 671–673
- 16 Choo Y, Klug A. Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. *Proc Natl Acad Sci USA*, 1994, 91: 11168–11172
- 17 Klug S J, Famulok M. All you wanted to know about SELEX. *Mol Biol Rep*, 1994, 20: 97–107
- 18 Roulet E, Busso S, Camargo A A, et al. High-throughput SELEX SAGE method for quantitative modeling of transcription-factor binding sites. *Nat Biotechnol*, 2002, 20: 831–835
- 19 李战伟, 王昕, 任刚, 等. 人工锌指蛋白随机库的构建及其在锌指核酸酶筛选中的应用. *西北农业学报*, 2012, 21: 1–10
- 20 Liu X, Noll D M, Lieb J D, et al. DIP-chip: rapid and accurate determination of DNA-binding specificity. *Genome Res*, 2005, 15: 421–427
- 21 Bulyk M L, Huang X, Choo Y, et al. Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. *Proc Natl Acad Sci USA*, 2001, 98: 7158–7163
- 22 Mukherjee S, Berger M F, Jona G, et al. Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nat Genet*, 2004, 36: 1331–1339
- 23 Dove S L, Joung J K, Hochschild A. Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature*, 1997, 386: 627–630
- 24 Dove S L, Hochschild A. Conversion of the omega subunit of *Escherichia coli* RNA polymerase into a transcriptional activator or an activation target. *Genes Dev*, 1998, 12: 745–754
- 25 Hu J C, Kornacker M G, Hochschild A. *Escherichia coli* one- and two-hybrid systems for the analysis and identification of protein-protein interactions. *Methods*, 2000, 20: 80–94
- 26 Meng X, Brodsky M H, Wolfe S A. A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. *Nat Biotechnol*, 2005, 23: 988–994
- 27 Joung J K, Ramm E I, Pabo C O. A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci USA*, 2000, 97: 7382–7387
- 28 Durai S, Bosley A, Abulencia A B, et al. A bacterial one-hybrid selection system for interrogating zinc finger-DNA interactions. *Comb Chem High Throughput Screen*, 2006, 9: 301–311

- 29 Hurt J A, Thibodeau S A, Hirsh A S, et al. Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. *Proc Natl Acad Sci USA*, 2003, 100: 12271–12276
- 30 Zhu T, Tian J, Zhang S, et al. Identification of the transcriptional regulator NcrB in the nickel resistance determinant of *Leptospirillum ferriphilum* UBK03. *PLoS One*, 2011, 6: 0017367
- 31 Yano S T, Rothman-Denes L B. A phage-encoded inhibitor of *Escherichia coli* DNA replication targets the DNA polymerase clamp loader. *Mol Microbiol*, 2011, 79: 1325–1338
- 32 Segal D J, Dreier B, Beerli R R, et al. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc Natl Acad Sci USA*, 1999, 96: 2758–2763
- 33 Dreier B, Fuller R P, Segal D J, et al. Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors. *J Biol Chem*, 2005, 280: 35588–35597
- 34 Seipel K, Georgiev O, Schaffner W. Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO J*, 1992, 11: 4961–4968
- 35 Jouvenot Y, Ginjala V, Zhang L, et al. Targeted regulation of imprinted genes by synthetic zinc-finger transcription factors. *Gene Ther*, 2003, 10: 513–522

A One-Step Screening System for Multi-Zinc Finger Proteins Targeting a Long-DNA Sequence

LIU Peng^{1,2}, ZHU Tao^{1,2}, LIANG Chen^{1,2}, BAI Bing^{1,2}, GAO GuanJun¹ & GUAN JiSong^{1,2}

1 Laboratory of Molecular Neurobiology, School of Life Sciences, Tsinghua University, Beijing 100084, China;

2 Tsinghua-Peking Center for Life Sciences, Beijing 10084, China

Engineered zinc finger proteins have been widely used for gene editing and transcriptional regulation. Previously, various systems have been utilized in zinc finger screening, but only targeting on sequences no more than 18 bp. Here, we present a novel fast screen method, which is a one-hybrid system based on gp8, a phage-encoded inhibitor of DNA polymerase III. The whole screening process takes 3 days to identify a batch of zinc finger proteins, which could specifically and tightly target on a 300 bp sequence of mouse *neurexin1-α* promoter. The specificity and efficiency of zinc finger proteins were further validated by *in vivo* and *in vitro* assays. Our system provides a high-throughput, rapid and *in vivo* method for large-scale screening of zinc finger proteins, which is capable to study the protein-DNA interactions.

zinc finger proteins, one-hybrid system, protein-DNA interactions

doi: 10.1360/052014-77