

基于四环素和 Cumate 的共调控序列

技术领域

本公开涉及用于调控目的核酸片段转录的基于四环素和 Cumate 依赖的核酸序列。具体地，本公开涉及用于调控目的核酸片段转录的核酸序列，所述核酸序列包含至少 2 个拷贝的结合四环素依赖的反式激活物(reverse tetracycline controlled transactivator ,rtTA)的 TetO-操作子序列 (Tet operator , TetO)、1 个拷贝的包含 TATA 盒序列的最小启动子序列和至少 1 个拷贝的结合 CymR 转录阻遏物 (transcription repressor) 的 CuO-操作子序列 (Cumate operator , CuO)。进一步，在上述核酸序列 3'端下游和被调控目的核酸片段 5'端上游之间还存在可剪接的内含子序列，其可用于在保持无诱导泄露的情况下，进一步提高诱导后转录活性。再进一步，本公开还涉及含有以上调控序列的载体、病毒、细胞、转基因植物或动物。

背景技术

基因表达可以通过对其中的几个步骤，包括转录，RNA 剪接，翻译和翻译后修饰，进行调控来实现对基因表达的调控，在转录步骤对特定基因或目的核酸片段进行调控是最广泛、最有效的表达调控方式。通过外源因素控制细胞、植物或动物体内特定基因和核酸片段的转录和表达开启或关闭的诱导表达系统在生物基础科研、生物制药、基因治疗、细胞治疗、转基因植物或动物改良等领域起关键作用。通常上述的诱导表达系统在设计上由两个部分组成：1) 在被调控目的核酸片段 5'端上游的一段特定核酸序列，一般可以和转录激活物或阻遏物结合调控目的核酸片段的转录，比如包含操作子 (operator) 的核酸序列 (此核酸序列以下称为响应元件)；2) 在受外部因素调控下，能和上述特定核酸序列结合或分离，一般带有转录激活或阻遏功能的单个或多个调控物，通常是表达后的蛋白。外部因素可以是环境因素，例如温度、光照等；或某种成分物质，例如参与生命活动的代谢物、荷尔蒙、金属离子或人造化合物等。

“四环素诱导表达系统 (Tetracycline inducible system 或 Tet System) ”是目前在科学研究上和商业上最广泛使用的诱导表达系统之一 (参见 ,例如 ,Gossen, M. and H. Bujard (1992). "Tight control of gene expression in mammalian cells by tetracycline-responsive promoters." Proc Natl Acad Sci U S A **89**(12): 5547-5551., 美国专利号5,888,981, 5,814,618, 6,004,941, 5,814,618, 7,541,446, WO96/01313和WO00/75347 (上述文献通过引用结合于此))。其主要由四环素依赖的调控物和含有TetO操作子的响应元件组成，通过使用四环素诱导的方式调控和含有TetO操作子的响应元件相连的目的核酸片段的转录。此含有TetO操作子的响应元件

可以在没有四环素的条件下和转录激活物结合，并在结合后启动下游被调控目的核酸片段转录；也可以在组成型活跃启动子（constitutively active promoter）的下游，被调控目的核酸片段的上游，然后在没有四环素的条件下和四环素依赖的阻遏物结合，抑制下游被调控的目的核酸片段转录。传统上，四环素依赖的调控物在没有四环素的条件下和TetO操作子结合，在有四环素的条件下和TetO操作子分离。比如基于此原理设计的激活物，通过去除环境中的四环素可以启动被调控目的核酸片段的转录(美国专利号5,464,758, 6,914,124, 5,789,156, 6,271,348, WO96/01313和WO00/75347)。之后发展出反式的四环素依赖的调控物，其仅在有四环素的条件下和TetO操作子结合。基于反式的四环素依赖的激活物，在环境中加入四环素可以激活被调控目的核酸片段转录(美国专利号5,654,168, 6,136,954, 5,789,156, 6,271,348, 6,087,166, 6,271,341, 7,541,446, WO96/01313和WO00/75347)，比如在培养基中加入四环素激活被调控目的核酸片段转录或给转基因动物注射四环素激活体内由基因工程转入的被调控目的基因的转录和表达。

被调控核酸片段在诱导条件下的转录活性和非诱导条件下的泄露转录活性是评价诱导表达系统优劣的两个重要指标，好的诱导表达系统需要具有更高的诱导转录活性和更低的泄露转录活性。四环素诱导系统自第一次报道以来被多个实验室优化，其性能不断提升。主要优化包括四环素依赖的调控物和含有TetO操作子的响应元件。在四环素依赖的调控物优化方面主要集中在（1）在尽可能保持诱导后转录活性的情况下，降低无诱导条件下的泄露转录活性，提高诱导/泄露转录活性比例；（2）优化调控物和四环素及其衍生物做为诱导剂结合的特异性和亲和力，在更低的四环素及其衍生物浓度下达到最大诱导转录活性，同时降低调控物在无诱导剂条件和响应元件的亲和力。（参见Zhou, X., et al. (2006). "Optimization of the Tet-On system for regulated gene expression through viral evolution." Gene Ther **13**(19): 1382-1390., 美国专利号8,383,364, 7,541,446）。优化含有TetO操作子的响应元件主要集中在（1）优化TetO操作子的拷贝数目和连接序列。在提高响应元件性能的同时，增强多拷贝连续TetO操作子序列的稳定性尤其是在病毒载体上的稳定性；（2）在不显著影响诱导后转录活性的条件下，降低最小启动子（mini promoter）的无诱导基础泄露转录活性，提高诱导/泄露转录活性比例（参见Loew, R., et al. (2010). "Improved Tet-responsive promoters with minimized background expression." BMC Biotechnol **10**: 81., 美国专利号9,181,556）。目前广泛使用的商业四环素诱导系统主要为Clontech Laboratories, Inc.公司的Tet-On® Advanced Inducible Gene Expression Systems（文件编号：PT3898-1 (102312)，其中的四环素依赖的反式激活物rtTA和含有TetO操作子的响应元件TRE(含有多拷贝连续TetO操作子序列和最小启动子序列的核酸序列，Tet Response Element，下文称为TRE)在下文中分别被称为rtTA_{adv}和

TRE_{adv}), Tet-Off® Advanced Inducible Gene Expression Systems(文件编号 :PT3945-1(101612)) 和Tet-On® 3G Inducible Expression Systems (文件编号 : PT5148-1 (010814), 其中的四环素依赖的调控物rtTA和含有TetO操作子的响应元件TRE在下文中分别被称为rtTA_{3G}和TRE_{3G})。

Cumate诱导系统(Cumate Inducible System)开发自*Pseudomonas putida*的p-cym操纵子, 5 由在组成型活性启动子TATA盒下游的CuO操作子和Cumate及其衍生物依赖的并能和CuO操作子结合的阻遏物CymR蛋白组成。在无Cumate的条件下, CymR蛋白和CuO操作子结合, 抑制下游被调控目的核酸片段的转录; 当CymR蛋白和Cumate结合后, 其和CuO操作子的亲和力下降并分离, 下游被调控目的核酸片段转录不再受抑制。比如将CuO操作子连接在被调控目的核酸片段和组成型活性启动子序列之间, 同时在细胞内组成性活跃表达CymR阻遏物。当培养基中没有Cumate及其衍生物诱导剂时, 表达的CymR蛋白与CuO操作子结合抑制 10 被调控目的核酸片段转录; 当培养基中添加诱导剂时, 与Cumate及其衍生物结合的CymR蛋白和CuO操作子分离, 被调控目的核酸片段的转录不再受抑制(参照WO02088346A2、WO2006037215A1 (上述文献通过引用结合于此))。

Cumate诱导系统的优化主要集中在(1)优化最短的有效CuO操作子序列。(2)筛选可 15 以溶于水的和p-cumate功能相似的Cumate及衍生物诱导剂。(3)构建类似Tet-Off的诱导系统, 由CymR激活物(如构建CymR和HSV VP-16转录激活结构域的融合蛋白)和含有多拷贝连续CuO操作子的响应元件(序列包含多段重复CuO序列和下游最小启动子)组成。在无Cumate诱导剂的条件下, CymR激活物和CuO操作子结合激活被调控目的核酸片段的转录; 在有Cumate诱导剂的条件下, CymR激活物和CuO操作子分离, 被调控目的核酸片段停止转 20 录。(4)构建类似Tet-On的诱导系统, 由反式CymR激活物(通过突变和筛选, 使得反式CymR在有Cumate诱导剂时和CuO操作子结合, 没有Cumate诱导剂时和CuO操作子分离, 并且构建反式CymR和HSV VP-16转录激活结构域的融合蛋白做为反式CymR激活物)和含有多拷贝连续CuO操作子的响应元件组成。在有Cumate诱导剂条件下, 反式CymR激活物和响应元件结合, 激活被调控目的核酸片段转录; 在没有Cumate诱导剂条件下, 反式CymR激活物和响 25 应元件分离, 被调控目的核酸片段停止转录(参照: Mullick, A., et al. (2006). "The cumate gene-switch: a system for regulated expression in mammalian cells." *BMC Biotechnol* 6: 43.; WO02088346A2、WO2006037215A1)。目前广泛使用的Cumate诱导系统主要为SBI System Biosciences公司的SparQ Cumate Switch系统(文件编号: 1-090810)。

在目前的各种诱导表达系统开发和改进中, 主要优化都集中在降低被调控目的核酸片 30 段在无诱导条件下的基础泄露转录活性的控制上, 而在诱导条件下的转录活性不是优化重点, 导致被诱导表达系统调控的目的核酸片段的诱导后转录活性低于常见的组成型活性启

动子 (constitutively active promoter) 控制下的转录活性。而在生物技术开发中, 提高目的核酸片段在诱导后的绝对转录活性和控制其在无诱导条件下的泄露转录活性同样重要。比如在生物制药领域中, 生产带有细胞毒性的重组蛋白或病毒载体, 既需要提高目的核酸片段编码的蛋白在诱导后的表达量, 也需要为保证生产细胞系的培养状态和稳定性而控制目的核酸片段在无诱导条件下的泄露表达量。所以一种既能提高目的核酸片段在诱导条件下的转录活性, 同时又能控制该目的核酸片段在无诱导条件下的泄露转录活性的诱导表达系统可以广泛促进生物技术在诸多领域的研究、开发和产业化。

技术问题

本公开涉及四环素和 Cumate 依赖的, 能调控目的核酸片段转录的核酸序列, 具体地, 所述核酸序列包含至少 2 个拷贝的结合四环素或其衍生物依赖的反式激活物 (reverse tetracycline controlled transactivator, rtTA) 的 TetO 操作子序列 (Tet operator, TetO) 1 个拷贝的包含 TATA 盒序列的最小启动子序列和至少 1 个拷贝的结合 CymR 转录阻遏物 (transcription repressor) 的 CuO-操作子序列 (Cumate operator, CuO)。进一步, 在上述核酸序列 3'端下游和被调控目的核酸片段 5'端上游之间存在的可以剪接的内含子序列可以在保持无诱导泄露的情况下, 进一步提高诱导后转录活性。再进一步, 本公开还涉及含有以上调控序列并以调控基因表达为目的的载体、病毒、细胞、转基因植物或动物。

技术解决方案

术语“目的核酸片段”一般可以指脱氧核糖核酸片段 (DNA 片段), 根据应用目的可以是指基因, 如编码蛋白的核酸序列; 可以为一段核糖核酸 (RNA), 如 RNA 病毒的全部或部分基因组 RNA 片段、微小 RNA (microRNA, 简称 miRNA)、小分子干扰核糖核酸 (small interfering RNA, 简称 siRNA)、长链非编码核糖核酸 (long non-coding RNAs, 简称 LncRNA)、CRISPR 基因编辑系统的导向 RNA (guide RNA, gRNA)、转运核糖核酸 (transfer RNA, 简称 tRNA)、核糖体核糖核酸 (Ribosomal RNA, 简称 rRNA)、适体 (Aptamers)、能和其他核酸结合的核糖核酸片段、能和特定蛋白结合的核糖核酸片段、或任意一段核糖核酸; 也可以是以上一段或多段核酸序列的组合。

术语“载体”是一种包含核酸分子, 通常被用作将外来遗传物质, 如上述目的核酸片段, 人工携带到另一个细胞中的媒介物, 并且在其中复制和/或表达。在功能上, 所有载体一般都可以用于克隆并携带外来目的核酸片段, 同时也有专门设计的用于核酸片段转录和蛋白质表达的表达载体。质粒, 病毒载体, 粘粒 (cosmids) 和人工染色体是载体的四种主要类型。其中, 最常用的载体是质粒。所有工程质粒载体都包含在细菌中复制的复制起点, 用于插入目的核酸片段的多克隆位点和用于选择阳性菌株的标记基因。病毒载体是另一种常用载体, 通常用来将遗传物质, 如目的核酸片段输送到细胞中。此过程可以在生物体内

(体内)或细胞培养物中(体外)进行。基于病毒自身进化出的多种分子机制,如对遗传物质的保护、基于受体对宿主细胞的选择、遗传物质向宿主细胞内的递送、在宿主细胞内的复制和/或表达、对宿主细胞生长、代谢、繁殖复制和防御机制的调整、以及在高等动物体内对免疫系统的抑制和/或逃逸,可以有效地将目的核酸片段转移到感染的靶细胞内。除了被用于分子生物学研究中,病毒载体还常被用于基因治疗、细胞治疗、免疫治疗和疫苗开发。

通常诱导表达系统在增强被调控目的核酸片段的诱导后转录活性的同时也会增加其在无诱导条件下的泄露转录活性;而常用的控制无诱导条件下泄露转录活性的策略比如通过点突变降低反式激活物在无诱导剂的条件下和操作子的亲和力或降低操作子下游的最小启动子的基础活性都会影响诱导后的最大转录活性。这使得绝大多数的单一诱导表达系统都为了平衡在无诱导条件下的泄露转录活性而降低了在诱导条件下对被调控目的核酸片段的转录激活活性。设计含有多调控组件的复合诱导表达系统是解决上述问题的有效方向。在本公开实现的复合诱导表达系统中,一方面优化提高第一诱导表达系统中的激活物和相应响应元件组合的诱导后最大转录活性,另一方面用第二诱导表达系统中的阻遏物和相应响应元件组合控制复合诱导表达系统因此提高的泄露转录活性,所述的复合诱导表达系统比如使用 Tet-On 诱导表达系统做为第一诱导表达系统,Cumate 诱导表达系统做为第二诱导表达系统。上述复合诱导表达系统的复合响应元件序列包含至少 2 个拷贝的结合受四环素或其衍生物调控的反式激活物(reverse tetracycline controlled transactivator,rtTA)的 TetO 操作子序列(tet operator)、1 个拷贝的包含 TATA 盒序列的最小启动子序列和至少 1 个拷贝的结合 CymR 转录阻遏物(transcription repressor)的 CuO 操作子序列(CuO operator)。

本发明人发现,Tet-On 和 Cumate 复合诱导表达系统的诱导转录活性和无诱导泄露转录活性受(1)多拷贝 TetO 操作子及其连接核酸序列,(2)TRE 响应元件中最小启动子序列;(3)CuO 操作子和最小启动子中 TATA 盒相对位置三个条件影响。第三代 Tet-On 系统中 TRE_{3G} 响应元件比第二代 Tet-On 系统中 TRE_{adv} 响应元件控制更严格,有更低的非诱导泄露转录活性;CuO 操作子和 TATA 盒越近对泄露转录活性控制越严格但同时诱导转录活性影响越大,越远对泄露转录活性控制越弱但同时诱导转录活性影响越小。基于以上发现,本公开的用于调控目的核酸片段转录的核酸序列包含:至少 2 个拷贝(例如 2、3、4、5、6、7 或更多个,优选 7 个)的能与受四环素或其衍生物调控的反式激活物 rtTA 结合的 TetO-操作子序列,优选地所述 TetO-操作子序列如 SEQ ID NO: 24 所示;一段包含 TATA 盒序列的最小启动子序列,优选地所述最小启动子序列如 SEQ ID NO:25 或 SEQ ID NO:26 所示,更优选地所述最小启动子序列如 SEQ ID NO:25 所示;以及与受 cumate 调控的转录阻遏物 CymR 结合的

CuO-操作子序列,优选地所述 CuO 操作子序列如 SEQ ID NO:27 所示。优选地,至少一段 CuO 操作子序列在上述启动子 TATA 盒的下游,优选在 TATA 盒下游 10bp 至 100bp 之间,如 14bp, 30bp, 50bp 和 100bp; 进一步优选地,有且仅有一段 CuO 序列在 TATA 盒的下游 30bp 至 50bp; 更近一步,有且仅有一段 CuO 序列在 TATA 盒的下游 30bp 或 50bp 位置。优选地,本公开
5 的用于调控目的核酸片段转录的核酸序列如 SEQ ID NO:23、SEQ ID NO:28、SEQ ID NO:29 或 SEQ ID NO:30 所示,更优选地,如 SEQ ID NO:23 或 SEQ ID NO:28 所示。

在表达质粒中启动子下游的内含子一般可以提高转录信使 RNA (mRNA) 的稳定性并提高 mRNA 转运出细胞核的效率(参照 Akef, A., et al. (2015). "Splicing promotes the nuclear export of beta-globin mRNA by overcoming nuclear retention elements." RNA **21**(11):
10 1908-1920) 并提高目的核酸片段的表达量。本发明人发现在单一诱导表达系统中的响应元件 3'端下游连接可剪接的内含子同样也能提高诱导后被调控目的核酸片段的表达量,但同时会显著增加被调控目的核酸片段的无诱导泄露表达量。本公开所述四环素和 Cumate 依赖的复合响应元件序列 3'端下游以及被调控目的核酸片段 5'端上游之间连接可以剪接的内含子序列可以在不显著增加无诱导泄露表达量的同时显著提高诱导表达量。基于此,在本公
15 开中,优选地,在所述四环素和 Cumate 依赖的复合响应元件序列 3'端下游以及目的核酸片段 5'端上游之间连接可剪接的内含子片段。上述内含子的选择并无特定限制,本领域技术人员可以理解,凡是在哺乳动物细胞中能进行 RNA 剪接 (Splicing) 的序列都能实现上述功能。可以选择的内含子序列包括但不限于常用的克隆载体上的内含子如:兔 β -球蛋白内含子、源自人 β -球蛋白和免疫球蛋白重链内含子的杂合内含子、EF-1 α 内含子 A、SV40 内含
20 子、源自腺病毒和免疫球蛋白重链内含子的杂合内含子、修饰的人巨细胞病毒内含子、源自鸡 β -肌动蛋白(CBA)和小鼠微小病毒(MMV)内含子的杂合内含子、源自鸡 β -肌动蛋白和兔 β -球蛋白内含子的嵌合物和 mP1 内含子;也可以是任意真核生物任意基因的任意内含子;或者是基于内含子剪接规则设计的人造内含子序列。本公开实施例使用的内含子序列可以是如 SEQ ID NO:31 所示。

25 上述 Tet-On 和 Cumate 复合诱导表达系统在无四环素或其衍生物和 Cumate 诱导剂 (Cumate 或其功能类似物)的环境下,rtTA 反式激活物不能和复合响应元件结合,同时 CymR 阻遏物和复合响应元件中 CuO 操作子结合进一步抑制被调控的目的核酸片段因为 rtTA 反式激活物非特异性结合或最小启动子的基础活性引起的泄露转录活性;在仅有 Cumate 诱导剂的环境下,rtTA 反式激活物和结合了 Cumate 诱导剂的 CymR 阻遏物都不和复合响应元件结
30 合,被调控目的核酸片段因为 rtTA 反式激活物非特异性结合或最小启动子的基础活性引起极低量转录;在仅有四环素或其衍生物的环境下,rtTA 反式激活物和复合响应元件中 TetO

操作子结合激活被调控目的核酸片段转录,但 CymR 阻遏物同时也和复合响应元件中的 CuO 操作子结合使得被调控目的核酸片段仅能有限转录;在四环素或其衍生物和 Cumate 诱导剂同时存在的环境下,rtTA 反式激活物和复合响应元件中的 TetO 操作子结合,同时结合了 Cumate 诱导剂的 CymR 阻遏物和复合响应元件中 CuO 操作子分离,使得被调控目的核酸片段最大活性转录。上述 rtTA 反式激活物核酸编码序列优选经过人密码子优化的 SEQ ID NO:18; CymR 优选经过人密码子优化的 SEQ ID NO:15。

上述 Tet-On 和 Cumate 复合诱导表达系统在生物基础科研、生物制药、基因治疗、细胞治疗、转基因植物或动物改良中都能应用,包括但不限于在细胞、动物或植物中,控制编码蛋白质的核酸序列转录的开启或关闭;控制非蛋白编码核酸序列的转录的开启和关闭,如 RNA 病毒的全部或部分基因组 RNA 片段、微小 RNA (microRNA, 简称 miRNA)、小分子干扰核糖核酸 (small interfering RNA, 简称 siRNA)、长链非编码核糖核酸 (long non-coding RNAs, 简称 lncRNA)、CRISPR 基因编辑系统的导向 RNA (guide RNA, gRNA)、转运核糖核酸 (transfer RNA, 简称 tRNA)、核糖体核糖核酸 (Ribosomal RNA, 简称 rRNA)、适体 (Aptamers)、能和其它核酸结合的核糖核酸片段、能和特定蛋白结合的核糖核酸片段、或任意一段核糖核酸;也可以是以上一段或多段核酸序列的组合。在一方面,上述 Tet-On 和 Cumate 复合诱导表达系统可以用于表达和生产在基因治疗、细胞治疗或免疫治疗领域的病毒载体;用于表达和生产有细胞毒性、或不稳定的蛋白质;用于在疫苗领域表达生产减毒疫苗或病毒样颗粒 (viral like particle, 简称 VLP) 疫苗;用于在转基因动物或植物中控制目的核酸片段转录表达的开启或关闭。

在本公开中,可用于 Tet-On 诱导表达系统的四环素及其衍生物包括在结构上与四环素相似的化合物,其能够与本公开所述四环素依赖的反式激活物 rtTA 相结合,其结合常数 K_a 至少达到 $10^{-6}M$;优选地,其结合常数 K_a 达到或强于 $10^{-9}M$ 。四环素衍生物例如可以选自:多西环素(Dox)、脱水四环素(Atc)、氯四环素、土霉素和脱氧四环素。

在本公开中,可用于 Cumate 诱导表达系统的,与阻遏物 CymR 相结合的 Cumate 功能类似物例如可以选自:对乙基苯甲酸、对丙基苯甲酸、对异丙基苯甲酸、对异丁基苯甲酸、对叔丁基苯甲酸,对正二甲基氨基苯甲酸,对正乙基氨基苯甲酸。以及例如在美国专利号 7,745,592 中描述的其他 Cumate 功能类似物。

可以转入包含本公开的四环素和 Cumate 的共调控序列的载体的宿主细胞在原理上可以为任意细胞例如是细菌、真菌、动物细胞或植物细胞。在另一方面,本公开所述的四环素和 Cumate 的共调控序列也可以通过病毒载体使用,例如制备的病毒载体所含核酸片段中含有本公开的四环素和 Cumate 的共调控序列,在 DNA 病毒载体中(如腺病毒、腺相关病毒)

以脱氧核糖核酸存在，在 RNA 病毒载体中(如逆转录病毒)以核糖核酸存在。

如本领域技术人员可以理解的，对于转录受本公开的调控核酸序列控制的目的核酸片段没有特别的限制，只要其转录可以受控于本公开的调控核酸序列即可。

在一个方面，本公开提供以下各项：

5 第 1 项. 核酸序列，所述核酸序列包含至少 2 个拷贝的能与受四环素或其衍生物调控的反式激活物 rtTA 结合的 TetO-操作子序列、1 个拷贝的包含 TATA 盒序列的最小启动子序列和至少 1 个拷贝的与受 cumate 调控的转录阻遏物 CymR 结合的 CuO-操作子序列，其中所述 CuO-操作子序列在所述 TATA 盒序列的 3'端下游，和所述 TATA 盒的距离间隔为 10bp 至 100bp。

10 第 2 项. 根据第 1 项所述的核酸序列，其中所述 CuO-操作子序列和所述 TATA 盒的距离间隔为 30bp 至 50bp。

第 3 项. 根据第 1 项所述的核酸序列，其中所述 CuO-操作子序列和所述 TATA 盒的距离间隔为 50bp。

第 4 项. 根据第 1 项所述的核酸序列，其中所述 TetO-操作子序列如 SEQ ID NO: 24 所示，及/或所述最小启动子序列如 SEQ ID NO: 25 或 SEQ ID NO:26 所示，及/或所述 CuO-操作子序列如 SEQ ID NO:27 所示。

第 5 项. 根据第 4 项所述的核酸序列，所述核酸序列如 SEQ ID NO:23、SEQ ID NO:28、SEQ ID NO:29 或 SEQ ID NO:30 所示。

20 第 6 项. 根据第 1-5 项中任一项所述的核酸序列，所述核酸序列在 3'端还包含可剪接的内含子序列。

第 7 项. 载体，所述载体包含根据第 1-6 项中任一项所述的核酸序列。

第 8 项. 根据第 7 项所述的载体，其中所述载体是表达载体，所述表达载体包含处于根据第 1-6 项中任一项所述的核酸序列的 3'端下游的目的核酸片段，并且所述目的核酸片段的转录受根据第 1-6 项中任一项所述的核酸序列的控制。

25 第 9 项. 宿主细胞，所述宿主细胞包含根据第 1-6 项中任一项所述的核酸序列或根据第 7 项或第 8 项所述的载体。

第 10 项. 在宿主细胞中诱导表达目的核酸片段的方法，所述方法包括以下步骤：

- (1) 将根据第 8 项所述的载体以及 rtTA 的编码序列和 CymR 的编码序列转入所述宿主细胞中；
- 30 (2) 在经过(1)的所述宿主细胞中表达 rtTA 和 CymR；以及
- (3) 向经过(2)的所述宿主细胞提供四环素或其衍生物以及 cumate 或其功能类似物。

第 11 项. 根据第 10 项所述的方法, 其中所述 rtTA 是 rtTA_{adv} 或 rtTA_{3Go}

第 12 项. 根据第 10 项所述的方法, 其中所述 rtTA 的编码序列如 SEQ ID NO. 18 所示。

第 13 项. 根据第 10 项所述的方法, 其中所述 CymR 的编码序列如 SEQ ID NO. 15 所示。

5 附图说明

图 1 显示实施例 2 中 CuO 操作子在复合响应元件中的位置及数量对目的核酸片段诱导表达量和无诱导泄露表达量的影响。图 1A 显示仅加入 DOX 诱导剂和同时加入 DOX 和 Cuamte 诱导剂对应的 Luciferase 检测实验的 RLU 值; 图 1B 显示加入 Cumate 和无加入 Cuamte 诱导剂 RLU 值得比值。

10 图 2 显示实施例 3 中单调调控响应元件、复合调控响应元件及响应元件下游连接内含子对目的核酸片段诱导表达量和无诱导泄露表达量的影响。图 1A 显示有/无 DOX 和 Cumate 诱导剂对应的 Luciferase 检测实验的 RLU 值; 图 1B 显示有/无 DOX 和 Cumate 诱导剂 RLU 值得比值。

图 3 显示实施例 4 中不同内含子对 TRE_{3G}CuO 复合响应元件诱导表达和无诱导泄露表达影响。图 1A 显示有/无 DOX 和 Cumate 诱导剂对应的 Luciferase 检测实验的 RLU 值; 图 1B 显示有/无 DOX 和 Cumate 诱导剂 RLU 值得比值。

图 4 显示实施例 5 中被 TRE_{3G}CuO 和 TRE_{3G}CuO-BGI 调控的目的核酸片段在不同诱导组合下的诱导转录活性。纵坐标显示不同诱导剂条件下 Luciferase 检测实验的 RLU 值; 4 种颜色的柱状图显示四种加入诱导剂的组合; 途中横线上数字代表对应的诱导条件下和无诱导条件下(白色) RLU 比值。

20 本发明的实施方式

提供以下实施例用以对本发明的技术方案进行说明, 以下实施例不应被认为是对本发明的范围和精神的限制。

25 实施例 1: 质粒构建方法

以下实施例所使用的分子克隆技术, 例如, DNA 片段的 PCR 扩增、DNA 片段的限制性内切酶酶切、DNA 片段的凝胶回收、两段或多段 DNA 片段的 T4 DNA 连接酶连接、连接产物感受态细胞的转化、质粒小量制备及鉴定等方法均为本领域熟知技术。以下实施例中涉及以下试剂: PCR 酶(Thermo, F-530S); 限制性内切酶(NEB); T4 DNA 连接酶(Invitrogen, 15224041); DNA 片段凝胶回收试剂盒(Omega, D2500-02); 质粒小提试剂盒(TIANGEN, DP105-03); 感受态细胞(XL-10 Gold, 湖南丰晖生物科技有限公司, JZ011); SEQ ID NO:1 至 SEQ ID NO:22 所示的核酸序列由金斯瑞合成并用于本公开所述质粒构建, 质粒测序鉴定由

Invitrogen 公司完成。表 1 为构建质粒的引物信息；表 2 为附录序列 SEQ ID NO:1 至 SEQ ID NO:31 的元件组成说明；表 3 为质粒中各功能元件说明；表 4 为本公开构建的质粒编号及对应名称。。以下实施例中所涉及各质粒所采用的元件序列信息为实现本公开的示例，本领域技术人员可以预期将以下实施例中所用质粒上各元件序列替换成其它生物学功能类似的元件序列也能达到本公开所述效果，包括但不限于质粒的骨架序列（如复制原点 (replication origin)、抗性基因等）、酶切位点序列、转座子重复序列、诱导系统响应元件序列、隔离子(Insulator)序列、启动子序列、内含子序列、聚腺苷酸信号(PolyA)序列、不同密码子优化的基因序列、以上各功能元件序列和基因序列的突变体以及各功能元件序列和基因序列的克隆位置、克隆顺序和克隆方向。具体的质粒构建方法如下所示：

- 10 1. 构建质粒 18BF007：将合成序列 SEQ ID NO:2 (2900bp)用 NotI 和 AsiSI 酶切并连接在质粒 18BF003 (SEQ ID NO:1)的 NotI 和 AsiSI 酶切位点，从而构建得到质粒 18BF007。
2. 构建质粒 18BF011：将 18BF007 质粒用 MluI 和 SphI 酶切，凝胶回收 1730bp 片段并将其连接在 18BF003 质粒的 MluI 和 SphI 酶切位点从而构建得到质粒 18BF011。
3. 构建质粒 18BF210 将合成序列 SEQ ID NO:3 (1208bp)用 SpeI 和 AgeI 酶切并连接在质粒 18BF011 的 AvrII 和 AgeI 酶切位点，从而构建得到质粒 18BF210。
4. 构建质粒 18BF211，18BF212，18BF213，18BF214，18BF215，18BF216，18BF217 和 18BF218：将合成序列 SEQ ID NO:4 (908bp)、序列 SEQ ID NO:5 (880bp)、序列 SEQ ID NO:6 (890bp)和序列 SEQ ID NO:7 (845bp)分别用 MluI 和 ClaI 酶切并分别连接在质粒 18BF210 的 MluI 和 ClaI 酶切位点，从而分别构建得到质粒 18BF212、18BF211、18BF214 和 18BF213。
- 20 将质粒 18BF211、18BF212、18BF213 和 18BF214 分别用 BstBI 酶切，并分别凝胶回收 DNA 片段 3932bp (18BF211)、3960bp (18BF212)、3897bp (18BF213)和 3942bp (18BF214)，并用 T4 连接酶连接以分别构建质粒 18BF215、18BF216、18BF217 和 18BF218。
5. 构建质粒 18BF229、18BF232、18BF233、18BF234、18BF235、18BF236、18BF237、18BF240 和 18BF241：以 pGL3-Basic (Promega E1751)为模板，Luc-F(SEQ ID NO:32)和 Luc-R(SEQ ID NO:33)为引物 PCR 扩增 luciferase 基因片段(1728bp)，并用 BamHI 和 XhoI 酶切处理，分别连接在质粒 18BF210、18BF217、18BF218、18BF215、18BF216、18BF211、18BF212、18BF213 和 18BF214 的 BamHI 和 XhoI 酶切位点，从而构建得到质粒 18BF229、18BF232、18BF233、18BF234、18BF235、18BF236、18BF237、18BF240 和 18BF241。
- 25 6. 构建质粒 18BF251、18BF252、18BF253、18BF254、18BF255、18BF256 和 18BF257：将合成序列 SEQ ID NO:8 (414bp)、序列 SEQ ID NO:9 (414bp)、序列 SEQ ID NO:10 (415bp)、序列 SEQ ID NO:11 (472bp)、序列 SEQ ID NO:12 (588bp)、序列 SEQ ID NO:13 (704bp)和序列 SEQ ID NO:22 (376bp)分别用 MluI 和 ClaI 酶切并分别连接在质粒 18BF235 的 MluI 和 ClaI 酶切位点，从而分别构建得到质粒 18BF251、18BF252、18BF253、18BF254、18BF255、18BF256 和 18BF257。
- 30

7. 构建质粒 18BF261、18BF262、18BF263 和 18BF264 :以序列 17 为模板 ,BGI(C&R)-F(SEQ ID NO:34)和 BGI(C&R)-R(SEQ ID NO:35)为引物 PCR 扩增 BGI(C&R)内含子序列 (1036bp) ; 以 pEF1alpha-IRES-AcGFP1(Clontech)为模板 , Intron(EF-1a)-F(SEQ ID NO:36)和 Intron(EF-1a)-R(SEQ ID NO:37)为引物 PCR 扩增 Intron(EF-1a)内含子序列 (962bp) ; 以 pSI (Promega#E1721) 质粒为模板 , Intron(pSI)-F(SEQ ID NO:38)和 Intron(pSI)-R(SEQ ID NO:39)为引物 PCR 扩增 Intron(pSI)内含子序列 (152bp) , 并将以上三段 PCR 产物分别用 ClaI 和 BamHI 酶切处理 , 连接在质粒 18BF235 的 ClaI 和 BamHI 酶切位点 , 从而构建得到质粒 18BF261、18BF263 和 18BF264。将合成的序列 14 (210bp)用 ClaI 和 BamHI 酶切 , 并分接在质粒 18BF235 的 ClaI 和 BamHI 酶切位点 , 从而构建得到质粒 18BF262。
8. 构建质粒 19BF075 和 19BF074 : 将合成的序列 SEQ ID NO:15 (633bp)和序列 SEQ ID NO:16 (1496bp)分别用 ClaI 和 XhoI 以及 SpeI 和 AgeI 酶切处理并依次连接在 18BF007 质粒的 ClaI 和 XhoI 酶切位点以及 AvrII 和 AgeI 酶切位点从而构建得到 19BF073 质粒。将合成的序列 SEQ ID NO:17 (1979bp)用 MluI 和 AgeI 酶切并连接在 18BF007 质粒的 MluI 和 AgeI 酶切位点 , 替换 CMV-BGI-MCS-pA 序列 , 从而构建得到 18BF008 质粒。将合成的序列 SEQ ID NO:18 (768bp)和序列 SEQ ID NO:19 (765bp)分别用 ClaI 和 XhoI 酶切并分别连接在 18BF008 质粒的 ClaI 和 XhoI 酶切位点从而分别构建得到 18BF085 和 18BF084 质粒。将质粒 19BF073 用 SpeI 和 AgeI 酶切 , 凝胶回收 3821bp 片段并将其分别连接在 18BF085 和 18BF084 质粒的 AvrII 和 AgeI 酶切位点从而分别构建得到 19BF075 和 19BF074 质粒。
9. 构建质粒 18BF019 : 将合成的序列 SEQ ID NO:21 (1044bp)和序列 SEQ ID NO:20 (1320bp)分别用 BamHI 和 XhoI 以及 XhoI 和 BglII 酶切并连接在 18BF011 质粒的 BamHI 和 BglII 酶切位点从而构建得到 18BF019 质粒。
10. 构建质粒 19BF229、19BF235 和 19BF237 : 将质粒 18BF229、18BF235 和 18BF237 分别用 MluI 和 AgeI 酶切 , 分别凝胶回收 4605bp (18BF229) 3847bp (18BF235) 和 4341bp (18BF237) 片段并将其分别连接在 18BF007 质粒的 MluI 和 AgeI 酶切位点从而分别构建得到 19BF229、19BF235 和 19BF237 质粒。

表 1. 引物信息列表

SEQ ID NO:	引物名称	引物序列(5'-3')
32	Luc-F	TCAGGATCCATCTGCGATCTAAGTAAGCTTG
33	Luc-R	TCAACTCGAGCTAGAATTACACGCGCATC
34	BGI(C&R)-F	GTCAATCGATGGAGTCGCTGCGCGCTG
35	BGI(C&R)-R	GTCGGATCCCTGTAGGAAAAAGAAGAAGG
36	Intron(EF-1a)-F	GTCAATCGATGTAAGTGCCGTGTGTG
37	Intron(EF-1a)-R	GTCGGATCCCTGAAATGGAAGAAAAAACT
38	Intron(pSI)-F	GTCAATCGATGTAAGTATCAAGGTTACAAG
39	Intron(pSI)-R	GTCGGATCCCTGTGGAGAGAAAGGC

表 2. 附录序列元件组成说明

序列编号	序列元件组成描述
------	----------

SEQ ID NO:1	18BF003_pma-MCS 质粒序列 (1893bp)
SEQ ID NO:2	NotI-IR/DR-HS4I-CMV-BGI-MCS-hGHPA-HS4I-IR/DR-AsiSI (2900bp)
SEQ ID NO:3	SpeI-SV40p-EGFP-SV40pA-AgeI (1208bp)
SEQ ID NO:4	MluI-TRE _{3G} CuO-BGI-Clal (908bp)
SEQ ID NO:5	MluI-TRE _{3G} -BGI-Clal (880bp)
SEQ ID NO:6	MluI-TRE _{adv} CuO ₅₂ -BGI-Clal (890bp)
SEQ ID NO:7	MluI-TRE _{adv} -BGI-Clal (845bp)
SEQ ID NO:8	MluI-TRE _{3G} CuO ₁₄ -Clal (414bp)
SEQ ID NO:9	MluI-TRE _{3G} CuO ₃₀ -Clal (414bp)
SEQ ID NO:10	MluI-TRE _{3G} CuO ₁₀₀ -Clal (415bp)
SEQ ID NO:11	MluI-TRE _{3G} CuO _{2x} -Clal (472bp)
SEQ ID NO:12	MluI-TRE _{3G} CuO _{4x} -Clal(588bp)
SEQ ID NO:13	MluI-TRE _{3G} CuO _{6x} -Clal (704bp)
SEQ ID NO:14	Call-Intron(mP1)-BamHI(210bp)
SEQ ID NO:15	Clal-optiCymR-XhoI (633bp)
SEQ ID NO:16	SpeI-SV40p- optiHygroR-SV40pA-AgeI(1496bp)
SEQ ID NO:17	MluI-CAGGS-BGI(C&R)-MCS-SV40pA-AgeI (1979bp)
SEQ ID NO:18	Clal-optirtTA _{3G} -XhoI (768bp)
SEQ ID NO:19	Clal-rtTA _{adv} -XhoI (765bp)
SEQ ID NO:20	XhoI-ires-EGFP-BglIII(1320bp)
SEQ ID NO:21	BamHI-optiSB-XhoI (1044bp)
SEQ ID NO:22	MluI-TRE _{adv} CuO ₃₂ -BGI-Clal (376bp)
SEQ ID NO:23	TRE _{3G} CuO ₅₀ 响应元件序列
SEQ ID NO:24	TetO 操作子序列
SEQ ID NO:25	最小启动子序列#1
SEQ ID NO:26	最小启动子序列#2
SEQ ID NO:27	CuO 操作子序列
SEQ ID NO:28	TRE _{3G} CuO ₃₀ 响应元件序列
SEQ ID NO:29	TRE _{adv} CuO ₃₂ 响应元件序列
SEQ ID NO:30	TRE _{adv} CuO ₅₂ 响应元件序列
SEQ ID NO:31	人β-球蛋白内含子 (BGI) 序列

表 3. 质粒功能元件说明

原件名称	功能描述
IR/DR(L/R)	SB 转座子系统反向重复序列 (IR)和同向重复序列(DR)
HS4I	鸡 β-球蛋白高敏感位置 4 核心隔离子
HygroR	经密码子优化的编码潮霉素抗性基因序列
Sleeping Beauty (optiSB)	经密码子优化的编码 Sleeping Beauty (SB)转座酶的基因序列 (SEQ ID NO:21)
BGI	人 β-球蛋白内含子 (SEQ ID NO:31)
CMV	人类巨细胞病毒的强表达启动子
MCS	限制性内切酶多克隆位点
CAGGS	巨细胞病毒启动子增强子部分与鸡 β-肌动蛋白启动子嵌合启动子
BGI (C&R)	鸡 β-肌动蛋白和兔 β-球蛋白的嵌合内含子
Intron (mP1)	鼠 Prm1 基因内含子 (SEQ ID NO:14 中 7-202bp 与登记号 : FJ411376 中 942-1137bp 一致)
Intron (EF-1a)	pEF1alpha-IRES-AcGFP1(Clontech)质粒 EF-1α 内含子
Intron (pSI)	pSI (Promega#E1721) 质粒人 β 珠蛋白和免疫球蛋白重链基因嵌合体内含子
TRE _{3G} CuO/ TRE _{3G} CuO ₅₀	本申请设计的 Tet-On(基于 TRE _{3G} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, CuO 序列在 TATA 盒之后间隔 50bp 位置 (SEQ ID NO:23)

TRE _{3G} CuO ₁₄	本申请设计的 Tet-On(基于 TRE _{3G} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, CuO 序列在 TATA 盒之后间隔 14bp 位置 (SEQ ID NO:8 中 10-317bp)
TRE _{3G} CuO ₃₀	本申请设计的 Tet-On(基于 TRE _{3G} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, CuO 序列在 TATA 盒之后间隔 30bp 位置 (SEQ ID NO:28)
TRE _{3G} CuO ₁₀₀	本申请设计的 Tet-On(基于 TRE _{3G} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, CuO 序列在 TATA 盒之后间隔 100bp 位置 (SEQ ID NO:10 中 10-403bp)
TRE _{3G} CuO _{2x}	本申请设计的 Tet-On(基于 TRE _{3G} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, 2 段重复 CuO 序列在 TATA 盒之后间隔 50bp 位置 (SEQ ID NO:11 中 10-411bp)
TRE _{3G} CuO _{4x}	本申请设计的 Tet-On(基于 TRE _{3G} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, 4 段重复 CuO 序列在 TATA 盒之后间隔 50bp 位置 (SEQ ID NO:12 中 10-527bp)
TRE _{3G} CuO _{6x}	本申请设计的 Tet-On(基于 TRE _{3G} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, 6 段重复 CuO 序列在 TATA 盒之后间隔 50bp 位置 (SEQ ID NO:13 中 10-643bp)
TRE _{adv} CuO/TRE _{adv} CuO ₅₂	本申请设计的 Tet-On(基于 TRE _{adv} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, CuO 序列在 TATA 盒之后间隔 52bp 位置 (SEQ ID NO:30)
TRE _{adv} CuO ₃₂	本申请设计的 Tet-On(基于 TRE _{adv} 中 7xTetO 序列)和 Cumate 复合诱导表达系统响应元件, CuO 序列在 TATA 盒之后间隔 32bp 位置 (SEQ ID NO:29)
TRE _{3G} /TRE _{adv}	Tet-On 诱导表达系统响应元件: TRE _{3G} 是第三代响应元件/TRE _{adv} 是第二代响应元件 (SEQ ID NO:5/SEQ ID NO:7)
SV40p	猿空泡病毒 40 启动子
IRES	核糖体进入位点
EGFP	编码绿色荧光蛋白的基因序列
ECFP	编码青色荧光蛋白的基因序列
Luciferase (Luc)	编码荧光素酶的基因序列
polyA (hGHpA/SV40pA)	转录终止子聚腺苷酸序列(hGHpA 人生长因子终止子序列/SV40pA 猿空泡病毒 40 终止子)
optiCymR	经密码子优化的编码 Cumate 诱导表达系统阻遏物 CymR 蛋白的编码序列(SEQ ID NO:15 中 13-621bp)
rtTA _{3G}	经密码子优化的编码第三代 Tet-On 调控系统 rtTA _{3G} 反式激活元件的基因序列(SEQ ID NO:18 中 13-756bp)
rtTA _{adv}	编码第二代 Tet-On 调控系统 rtTA _{adv} 反式激活元件的基因序列 (SEQ ID NO:19 中 13-756bp)

表 4. 质粒编号及名称

质粒编号	质粒名称
18BF003	pma-MCS
18BF007	pmaSBT3-2xHS4I-CMV-BGI-MCS
18BF011	pmaCMV-BGI-MCS
18BF210	pmaCMV-BGI-MCS-SV40p-EGFP
18BF211	pmaTRE _{3G} -BGI-MCS-SV40p-EGFP
18BF212	pmaTRE _{3G} CuO ₅₀ -BGI-MCS-SV40p-EGFP
18BF213	pmaTRE _{adv} -BGI-MCS-SV40p-EGFP
18BF214	pmaTRE _{adv} CuO ₅₂ -BGI-MCS-SV40p-EGFP
18BF215	pmaTRE _{3G} -MCS-SV40p-EGFP
18BF216	pmaTRE _{3G} CuO ₅₀ -MCS-SV40p-EGFP
18BF217	pmaTRE _{adv} -MCS-SV40p-EGFP
18BF218	pmaTRE _{adv} CuO ₅₂ -MCS-SV40p-EGFP
18BF229	pmaCMV-BGI-Luciferase-SV40p-EGFP
18BF232	pmaTRE _{adv} -Luciferase-SV40p-EGFP
18BF233	pmaTRE _{adv} CuO ₅₂ -Luciferase-SV40p-EGFP
18BF234	pmaTRE _{3G} -Luciferase-SV40p-EGFP
18BF235	pmaTRE _{3G} CuO ₅₀ -Luciferase-SV40p-EGFP
18BF236	pmaTRE _{3G} -BGI-Luciferase-SV40p-EGFP

18BF237	pmaTRE _{3G} CuO ₅₀ -BGI-Luciferase-SV40p-EGFP
18BF240	pmaTRE _{adv} -BGI-Luciferase-SV40p-EGFP
18BF241	pmaTRE _{adv} CuO ₅₂ -BGI-Luciferase-SV40p-EGFP
18BF251	pmaTRE _{3G} CuO ₁₄ -Luciferase-SV40p-EGFP
18BF252	pmaTRE _{3G} CuO ₃₀ -Luciferase-SV40p-EGFP
18BF253	pmaTRE _{3G} CuO ₁₀₀ -Luciferase-SV40p-EGFP
18BF254	pmaTRE _{3G} CuO _{2x} -Luciferase-SV40p-EGFP
18BF255	pmaTRE _{3G} CuO _{4x} -Luciferase-SV40p-EGFP
18BF256	pmaTRE _{3G} CuO _{6x} -Luciferase-SV40p-EGFP
18BF257	pmaTRE _{adv} CuO ₃₂ -Luciferase-SV40p-EGFP
18BF261	pmaTRE _{3G} CuO-BGI(C&R)-Luciferase-SV40p-EGFP
18BF262	pmaTRE _{3G} CuO-Intron(mPI)-Luciferase-SV40p-EGFP
18BF263	pmaTRE _{3G} CuO-Intron(EF-1a)-Luciferase-SV40p-EGFP
18BF264	pmaTRE _{3G} CuO-Intron(pSI)-Luciferase-SV40p-EGFP
19BF073	pmaSBT3-2xHS4I-CMV-BGI-optiCymR-HygroR
18BF008	pmaSBT3-2xHS4I-CAGGS-BGI (C&R)-MCS
18BF085	pmaSBT3-2xHS4I-CAGGS-BGI (C&R)-optirtTA _{3G}
18BF084	pmaSBT3-2xHS4I-CAGGS-BGI (C&R)-rtTA _{adv}
19BF075	pmaSBT3-2xHS4I-CAGGS-BGI (C&R)-optirtTA _{3G} -CMV-BGI-optiCymR-HygroR
19BF074	pmaSBT3-2xHS4I-CAGGS-BGI (C&R)-rtTA _{adv} -CMV-BGI-optiCymR-HygroR
18BF019	pmaCMV-BGI-optiSB-IRES-ECFP
19BF229	pmaSBT3-2xHS4I-CMV-BGI-Luciferase-SV40p-EGFP
19BF235	pmaSBT3-2xHS4I-TRE _{3G} CUO-Luciferase-SV40p-EGFP
19BF237	pmaSBT3-2xHS4I-TRE _{3G} CUO-BGI-Luciferase-SV40p-EGFP

实施例 2: CuO 操作子位置及数量对诱导表达量和无诱导泄露表达量的影响

本实施例所述实验为研究和验证 CuO 操作子在 Tet-On 和 Cumate 复合响应元件的位置和拷贝数对复合响应元件诱导表达量和无诱导泄露表达量的影响以及优化确认最优的 CuO 操作子位置和拷贝数。本实施例基于 TRE_{3G} 响应元件中 7xTetO 序列和最小启动子序列#1(SEQ ID NO:25), 在 TATA 盒下游间隔 10bp 至 100bp 范围内, 分别在间隔 14bp、30bp、50bp 和 100bp 的位置上连接 CuO 操作子序列设计了 TRE_{3G}CuO₁₄ (SEQ ID NO:8 中 10-317bp) TRE_{3G}CuO₃₀ (SEQ ID NO:28) TRE_{3G}CuO₅₀ (SEQ ID NO:23) 和 TRE_{3G}CuO₁₀₀ (SEQ ID NO:10 中 10-403bp) 响应元件, 并在以上响应元件 3'端下游连接 Luciferase 报告基因序列, 构建质粒 18BF251、18BF252、18BF235 和 18BF253。进一步, 为研究多拷贝 CuO 操作子的影响, 在 TATA 盒下游间隔 50bp 的位置, 分别插入了 2、4 和 6 个拷贝 CuO 操作子序列设计了 TRE_{3G}CuO_{2x} (SEQ ID NO:11 中 10-411bp) TRE_{3G}CuO_{4x} (SEQ ID NO:12 中 10-527bp) 和 TRE_{3G}CuO_{6x} (SEQ ID NO:13 中 10-643bp) 响应元件, 并在以上响应元件下游 3'端连接 Luciferase 报告基因序列, 构建质粒 18BF254、18BF255 和 18BF256。在稳定表达 rtTA_{3G} 和 CymR 基因的 293T-rtTA_{3G}-CymR 细胞内, 瞬时转染上述质粒, 通过测量同时加入 DOX 和 Cumate 诱导剂样品和仅加入 DOX 诱导剂对照的 luciferase 荧光值验证优化 CuO 操作子最优位置和数量。具体实验方法如下:

1. 用 SB 转座子系统构建 293T-rtTA_{adv}-CymR 和 293T-rtTA_{3G}-CymR 细胞:

将 293T 细胞按照 1.5×10^6 个细胞每 60mm 培养皿接种，在 37°C 、5% CO_2 的环境下，培养在添加 10%FBS (ExCell, 11H116) 的 DMEM(Sigam, D6429) 完全培养基中。培养 24 小时后，按照 PEI 方法进行转染，转染时每 60mm 培养皿加入 500 μL 转染试剂，其中含有总质粒量为 5.5 μg ，总质粒量与 PEI MAX (Polysciences, 24765-1) 的质量比为 1:4。其中按照质粒 19BF074:18BF019 摩尔比 10:1 进行转染以获得 293T-rtTA_{adv}-CymR 细胞；按照质粒 19BF075:18BF019 摩尔比 10:1 进行转染以获得 293T-rtTA_{3G}-CymR 细胞。质粒与 PEI MAX 混合均匀，静置 15 分钟后加入培养皿中，转染 3 小时后，更换培养基为 DMEM 完全培养基，转染操作完成。转染 24 小时后，用胰酶消化细胞并全部接种于 100mm 培养皿(Corning, 430167)中，并用 200 $\mu\text{g}/\text{ml}$ 潮霉素 (Hygromycin, 生工 A600230-0001) 药物筛选至少三代后，细胞在药物压力下生长和原始 293T 细胞一致后，进行以下实验。

2. 基于 Luciferase 荧光强度检测在 DOX 和 Cumate 诱导或仅 DOX 诱导条件下各响应元件效能

将 293T-rtTA_{3G}-CymR 细胞按照 2.5×10^4 个细胞每孔接种于 96 孔板(Corning 3916)中，培养基为 100 微升 DMEM 完全培养基。培养 24 小时后，按照 PEI 方法进行转染以下质粒 18BF234 (TRE_{3G})、18BF251 (TRE_{3G}CuO₁₄)、18BF252(TRE_{3G}CuO₃₀)、18BF235 (TRE_{3G}CuO₅₀)、18BF253(TRE_{3G}CuO₁₀₀)、18BF254(TRE_{3G}CuO_{2x})、18BF255(TRE_{3G}CuO_{4x})和 18BF256(TRE_{3G}CuO_{6x})，转染时每孔加入 10 μL 转染试剂，其中含有总质粒量为 0.3 μg ，包含 0.01 μg 以上 8 种待测质粒和 0.29 微克 18BF003 空质粒。总质粒量与 PEI MAX (Polysciences 24765-1) 的质量比为 1:4，每种质粒转染 6 个孔。转染 3 小时后更换 DMEM 完全培养基，在其中 3 孔内加入诱导剂 1 $\mu\text{g}/\text{ml}$ DOX (盐酸多西环素(DOX)，生工生物工程(上海)股份有限公司，A600889)和 200 $\mu\text{g}/\text{ml}$ Cumate (Aladdin, I107765)；另外三个孔内仅加入 1 $\mu\text{g}/\text{ml}$ DOX。转染后 24 小时，使用 Steady-Glo[®] Luciferase Assay System (Promega, E2610)试剂盒，并按照其操作说明(Promega, FB037)检测各孔相对荧光单位 RLU (relative light unit)，检测仪器为荧光酶标仪(Perkin Elmer Victor V)。

结果如图 1 所示：CuO 操作子序列影响 Tet-On 诱导表达系统响应元件的诱导表达量，且 CuO 操作子距离 TRE 响应元件 TATA 盒越近，诱导后的表达量越低。但随着 CuO 操作子与 TATA 盒距离增大，无 Cumate 诱导的泄露表达量越高。在综合诱导后表达量、无 Cumate 诱导泄露表达量和 Cumate 诱导/泄露表达量比例等结果，确定 CuO 在与 TATA 盒间隔 30bp 至 50bp 之间是结果最优，诱导表达量分别为 3.06×10^6 RLU 和 4.53×10^6 RLU，分别为 TRE_{3G} 响应元件诱导表达量的 44.5%和 66.0%；基于 Cumate 诱导表达系统的诱导/泄露表达量比例分别为 TRE_{3G} 响应元件的 4.21 倍和 2.81 倍。增加 CuO 操作子的拷贝至 2, 4 或 6 会进一步

降低诱导后表达量,且不会提高 Cumate 诱导/泄露表达量比例。基于以上结果,在间隔 TRE 响应元件 TATA 盒 30bp 至 50bp 距离插入 CuO 操作子且仅插入 1 个拷贝为最优条件。

基于 TRE_{adv} 响应元件中 7xTetO 序列和最小启动子序列#2(SEQ ID NO:26) 在 TATA 盒下游间隔 32bp 和 52bp 的位置上连接 CuO 操作子序列按照上述方法设计了 TRE_{adv}CuO₃₂ (SEQ ID NO:29)和 TRE_{adv}CuO₅₂(SEQ ID NO:30)响应元件,并在以上响应元件 3'端下游连接 Luciferase 报告基因序列,构建质粒 18BF257 和 18BF233。通过和上述相同实验方法,检测诱导表达量分别为 3.26+06 RLU 和 4.88E+06 RLU, 分别为 TRE_{adv}响应元件诱导表达量的 37.8%和 53.4%; 基于 Cuamte 诱导表达系统的诱导/泄露表达量比例分别为 TRE_{adv}响应元件的 4.97 倍和 3.23 倍。后续实施例以 TRE_{adv}CuO₅₂ 和 TRE_{3G}CuO₅₀ 响应元件进行实验,并分别标记为 TRE_{adv}CuO 和 TRE_{3G}CuO 响应元件。

实施例 3: 单调控/复合调控及内含子对诱导表达量和无诱导泄露表达量的影响

在 Tet-On 诱导表达系统中, TRE 响应元件的 TetO 操作子连接序列和最小启动子序列以及反式激活物 rtTA 的不同突变体综合影响诱导转录活性和泄露转录活性。此外响应元件 3'端下游,被调控目的核酸片段 5'端上游连接可剪切的内含子序列可能增强信使核糖核酸运输和稳定性,但同样可能影响诱导表达系统响应元件的诱导转录活性和泄露转录活性。基于此,本实施例对比(1)基于 TRE_{adv}和 TRE_{3G}响应元件序列设计的 Tet-On 和 Cumate 复合响应元件 TRE_{adv}CuO 和 TRE_{3G}CuO,(2)是否连接内含子,(3)被反式激活物 rtTA_{adv}或 rtTA_{3G}调控,三个条件组合下的 Luciferase 基因诱导表达量和诱导/泄露表达量比例确认最优复合响应元件设计方案。本实施例共对比 8 种设计方案:分别为 TRE_{adv}(18BF232)、TRE_{adv}CuO(18BF233)、TRE_{3G}(18BF234)、TRE_{3G}CuO(18BF235)、TRE_{adv}-BGI(18BF240,响应元件为 TRE_{adv},并且响应元件 3'端和 Luciferase 基因 5'端之间连接人 β -球蛋白内含子)、TRE_{adv}CuO-BGI(18BF241,响应元件为 TRE_{adv}CuO,并且响应元件 3'端和 Luciferase 基因 5'端之间连接人 β -球蛋白内含子)、TRE_{3G}-BGI(18BF236,响应元件为 TRE_{3G},并且响应元件 3'端和 Luciferase 基因 5'端之间连接人 β -球蛋白内含子)和 TRE_{3G}CuO-BGI(18BF237,响应元件为 TRE_{3G}CuO,并且响应元件 3'端和 Luciferase 基因 5'端之间连接人 β -球蛋白内含子)。启动子 CMV-BGI(18BF229,启动子为 CMV,并且 CMV 启动子 3'端和 Luciferase 基因 5'端之间连接人 β -球蛋白内含子)做为阳性对照。在稳定表达 rtTA_{adv}和 CymR 基因的 293T-rtTA_{adv}-CymR 细胞内和稳定表达 rtTA_{3G}和 CymR 基因的 293T-rtTA_{3G}-CymR 细胞内,分别瞬时转染上述质粒,通过测量加入 DOX 和 Cumate 诱导剂样品和无诱导对照的 Luciferase 荧光值验证最优复合响

应元件组合。具体实验方法如下：

将实施例 2 中构建的 293T-rtTA_{adv}-CymR 和 293T-rtTA_{3G}-CymR 细胞分别按照 2.5E+04 个细胞每孔接种于 96 孔板(Corning 3916)中，培养基为 100 微升 DMEM 完全培养基。培养 24 小时后，按照 PEI 方法转染上述 9 种质粒，转染时每孔加入 10 μ L 转染试剂，其中含有总质粒量为 0.3 μ g，包含 0.01 μ g 上述 9 种质粒待测质粒和 0.29 微克 18BF003 空质粒。总质粒量与 PEI MAX (Polysciences, 24765-1)的质量比为 1:4，每种质粒分别在 293T-rtTA_{adv}-CymR 和 293T-rtTA_{3G}-CymR 细胞各转染 6 个孔。转染 3 小时后更换 DMEM 完全培养基，在每种细胞其中 3 孔内加入诱导剂 1 μ g/ml DOX 和 200 μ g/ml Cumate；另外三个孔内加入等量培养基对照。转染后 24 小时，使用 Steady-Glo[®] Luciferase Assay System (Promega, E2610)试剂盒，并按照其操作说明(Promega, FB037)检测各孔相对荧光单位 RLU (relative light unit)，检测仪器为荧光酶标仪(Perkin Elmer Victor V)。

实验结果如图 2 所示：(1) TRE_{CuO} 复合响应元件对比 TRE 单调调控响应元件显著提高了诱导/泄露表达量比例一到两个数量级。TRE_{3G}CuO 复合响应元件对比 TRE_{adv}CuO 复合响应元件能更好的控制泄露表达。在 293T-rtTA_{adv}-CymR 细胞中，TRE_{3G}CuO 和 TRE_{adv}CuO 的诱导/泄露表达量比例分别为 1565 和 1345 倍，提高 16.4%；在 293T-rtTA_{3G}-CymR 细胞中，TRE_{3G}CuO 和 TRE_{adv}CuO 的诱导/泄露表达量比例分别为 2635 和 1915 倍，提高 37.6%。但 TRE_{CuO} 复合响应元件对比 TRE 单调调控响应元件的诱导表达量显著下降。在 293T-rtTA_{adv}-CymR 细胞中，TRE_{3G}CuO 和 TRE_{adv}CuO 的诱导表达量分别为对应的 TRE_{3G} 和 TRE_{adv} 单调调控响应元件的 50.4% 和 36.3%；在 293T-rtTA_{3G}-CymR 细胞中，TRE_{3G}CuO 和 TRE_{adv}CuO 的诱导表达量分别为单调调控响应元件的 61.8%和 40.3%。(2) 在诱导表达量方面，TRE_{3G}CuO 复合响应元件对比 TRE_{adv}CuO 复合响应元件对下游的内含子更加敏感，在有内含子 BGI 的设计中，诱导表达量显著提高。TRE_{3G}CuO 和 TRE_{3G}CuO-BGI 诱导表达量在 293T-rtTA_{adv}-CymR 细胞中分别为 3.4E+06 RLU 和 10.6E+06 RLU，分别是 TRE_{3G} 诱导表达量的 50.4%和 157.9%，连接内含子可以提高 TRE_{3G}CuO 复合响应元件诱导后表达量 2.13 倍；在 293T-rtTA_{3G}-CymR 细胞中分别为 4.24E+06 和 11.3E+06，分别是 TRE_{3G} 诱导表达量的 61.8%和 165.0%，连接内含子可以提高 TRE_{3G}CuO 复合响应元件诱导后表达量 1.67 倍。(3) 对比诱导后表达量，反式激活物 rtTA_{3G} 在 8 种设计方案中和 rtTA_{adv} 没有显著区别，但在对比诱导/泄露表达量比例上 rtTA_{3G} 有显著优势，其 8 种设计方案的诱导/泄露表达量比例平均值为 1110 倍是 rtTA_{adv} 的 1.68 倍 (rtTA_{adv} 诱导/泄露表达量比例平均值为 661 倍)。

基于以上结果，TRE_{3G}CuO 复合响应元件对比 TRE_{adv} 和 TRE_{3G} 以及 TRE_{adv}CuO 复合响应元件拥有更好控制泄露表达的能力；在下游连接内含子 BGI 后可以在维持 TRE_{3G}CuO 复合响应

元件控制泄露表达的同时极大的提高其诱导表达量。在受 rtTA_{3G} 调控条件下, TRE_{3G}CuO-BGI 的诱导表达量分别为 TRE_{adv} 的 80.6%、TRE_{adv}-BGI 的 80.5%、TRE_{adv}CuO 的 199.9%、TRE_{adv}CuO-BGI 的 147.0%、TRE_{3G} 的 165.0%、TRE_{3G}-BGI 的 81.5%和 TRE_{3G}CuO 的 267.2%, 达到组成型活性启动子 CMV-BGI 的 84.7%; TRE_{3G}CuO-BGI 的诱导/无诱导表达量比例分别为

5 TRE_{adv} 24.5 倍、TRE_{adv}-BGI 43.4 倍、TRE_{adv}CuO 1.2 倍、TRE_{adv}CuO-BGI 1.6 倍、TRE_{3G} 5.9 倍、TRE_{3G}-BGI 8.9 倍和 TRE_{3G}CuO 的 0.8 倍。

实施例 4: 不同内含子对 TRE_{3G}CuO 复合响应元件诱导表达和无诱导泄露表达影响

在上述实施例 3 中, 连接内含子 BGI 显著提升了 TRE_{3G}CuO 复合诱导响应元件的诱导转

10 录活性以及 Luciferase 目的基因的表达量。本实施例验证其它内含子是否有相似效果。本实施例将 4 种常用质粒载体上的内含子克隆到 TRE_{3G}CuO 复合响应元件 3'端下游和 luciferase 报告基因 5'端上游之间, 分别构建含有 BGI(C&R)、Intron(mP1)、Intron(EF-1a)和 Intron(pSI) 内含子的质粒 18BF261、18BF262、18BF263 和 18BF264, 质粒构建具体方法见实施例 1 描述。在稳定表达 rtTA_{3G} 和 CymR 基因的 293T-rtTA_{3G}-CymR 细胞内, 分别瞬时转染上述质粒以

15 及 TRE_{3G}CuO (18BF235)、TRE_{3G}CuO-BGI (18BF237) 和 CMV-BGI (18BF229) 质粒, 通过测量加入 DOX 和 Cumate 诱导剂样品和无诱导对照的 luciferase 荧光值验证各内含子对 TRE_{3G}CuO 复合响应元件诱导表达和无诱导泄露表达的影响。具体实验方法如下:

将 293T-rtTA_{3G}-CymR 细胞按照 2.5E+04 个细胞每孔接种于 96 孔板(Corning 3916)中, 培养基为 100 微升 DMEM 完全培养基。培养 24 小时后, 按照 PEI 方法分别转染上述 7 种质粒,

20 转染时每孔加入 10μL 转染试剂, 其中含有总质粒量为 0.3ug, 包含 0.01ug 以上 7 种待测质粒和 0.29 微克 18BF003 空质粒。总质粒量与 PEI MAX (Polysciences 24765-1)的质量比为 1:4, 每种质粒转染 6 个孔。转染 3 小时后更换 DMEM 完全培养基, 在其中 3 孔内加入诱导剂 1ug/ml DOX 和 200ug/ml Cumate; 另外三个孔内加入等量培养基。转染后 24 小时, 使用 Steady-Glo[®] Luciferase Assay System (Promega, E2610)试剂盒, 并按照其操作说明(Promega, FB037)检测各孔相对荧光单位 RLU (relative light unit), 检测仪器为荧光酶标仪(Perkin Elmer Victor V)。

25

实验结果如图 3 所示: 含有四种其它内含子的质粒 18BF261、18BF262、18BF263 和 18BF264 在诱导表达量和诱导/泄露表达量比例上和 TRE_{3G}CuO-BGI (18BF237) 无显著差异, 都能达到实施例 3 所述 TRE_{3G}CuO-BGI 的设计效果。在内含子里 需要有 5' 剪接位点(5' splice

30 site)、3' 剪接位点(3' splice site)及剪接分枝位点(branch point)来进行剪接。本实施例证明符合上述条件, 能剪接的内含子都能达到本公开所述的实验效果。因此本公开所述效果不应

受到特定内含子序列约束，凡是在哺乳动物细胞中能进行 RNA 剪接 (Splicing) 的序列都能实现上述功能。可以选择的内含子序列包括但不限于常用的克隆载体上的内含子如：兔 β -球蛋白内含子、源自人 β -球蛋白和免疫球蛋白重链内含子的杂合内含子、EF-1 α 内含子 A、SV40 内含子、源自腺病毒和免疫球蛋白重链内含子的杂合内含子、修饰的人巨细胞病毒内含子、源自鸡 β -肌动蛋白(CBA)和小鼠微小病毒(MMV)内含子的杂合内含子、源自鸡 β -肌动蛋白和兔 β -球蛋白内含子的嵌合物和 mP1 内含子；也可以是任意真核生物任意基因的任意内含子；或者是基于内含子剪接规则设计的人造内含子序列。

实施例 5: 研究 TRE_{3G}CuO 复合响应元件在不同诱导组合下的诱导转录活性。

本实施例构建受 Tet-ON 和 Cumate 复合调控的，基于 TRE_{3G}CuO 或 TRE_{3G}CuO-内含子设计，以 Luciferase 为报告基因的稳定细胞系，并基于这些细胞系检测无诱导剂、仅加 Cumate、仅加 DOX 或同时加 DOX 和 Cumate 诱导剂 4 种诱导条件组合下，报告基因的不同诱导转录活性。本实施例仅以 Luciferase 报告基因为例，原则上本公开所述方法不受目的核酸片段影响。具体实验方法如下：

1. 利用 SB 转座子系统构建 DOX 和 Cumate 复合调控 Luciferase 稳定细胞系：

将 293T 细胞按照 1.5E+06 个细胞每 60mm 培养皿接种，在 37°C、5% CO₂ 的环境下，培养在添加 10%FBS (ExCell, 11H116) 的 DMEM(Sigam, D6429) 完全培养基中。培养 24 小时后，按照 PEI 方法进行转染，转染时每 60mm 培养皿加入 500 μ L 转染试剂，总质粒量为 5.5ug，总质粒量与 PEI MAX (Polysciences, 24765-1)的质量比为 1:4，共构建 3 种 Luciferase 稳定细胞系。其中转染质粒按照 19BF075:19BF229:18BF019 摩尔比 5:5:1 进行转染以获得 293T(T&C)-CMV-BGI-Luc 细胞；其中转染质粒按照 19BF075:19BF235:18BF019 摩尔比 5:5:1 进行转染以获得 293T(T&C)-TRE_{3G}CuO-Luc 细胞；其中转染质粒按照 19BF075:19BF237:18BF019 摩尔比 5:5:1 进行转染以获得 293T(T&C)-TRE_{3G}CuO-BGI-Luc 细胞。质粒与 PEI MAX 混合均匀后，静置 15 分钟后加入培养皿中，转染 3 小时后，更换培养基为 DMEM 完全培养基，转染操作完成。转染 24 小时后，用胰酶消化细胞并全部接种于 100mm 培养皿(Corning A30167)中，并用 200 μ g/ml 潮霉素 (Hygromycin, 生工 A600230-0001)药物筛选至少三代。当细胞在药物压力下生长和原始 293T 细胞一致后，再用极限稀释法将上述三种细胞分别稀释至 1 个细胞每孔接种于 96 孔板内，待细胞长至大约 50% 孔底面积时，在荧光显微镜下挑选 EGFP 绿色荧光强度较强、发光均匀的孔，每种细胞挑选 3 个独立克隆进行以下实验。

2. 利用不同 DOX 和 Cumate 组合诱导 Luciferase 表达:

将上述 3 种细胞，每种 3 个独立克隆分别按照 2.5E+04 个细胞每孔接种于 96 孔板 (Corning 3916)中，每个克隆接种 8 个孔，培养基为 100 微升 DMEM 完全培养基。培养 24

小时后,更换培养基并在每个细胞克隆 8 个培养孔中取复孔分别加入(1)等量培养基、(2)终浓度 200ug/ml Cumate、(3)终浓度 1ug/ml DOX 或(4)终浓度 1ug/ml DOX 和 200ug/ml Cumate。再培养 24 小时后,使用 Steady-Glo[®] Luciferase Assay System (Promega, E2610)试剂盒,并按照其操作说明(Promega, FB037)检测各孔相对荧光单位 RLU (relative light unit),检测仪器为荧光酶标仪(Perkin Elmer Victor V)。

实验结果如图 4 所示:在仅加入 Cumate 的条件下, TRE_{3G}CuO 和 TRE_{3G}CuO-BGI 的诱导转录活性为 1.36E+04 RLU 和 4.16E+04 RLU,对比无诱导条件分别提升了 4.31 倍和 3.27 倍。在仅加入 DOX 的条件下, TRE_{3G}CuO 和 TRE_{3G}CuO-BGI 的诱导转录活性为 2.14E+05 RLU 和 5.34E+05 RLU,对比无诱导条件分别提升了 67.66 倍和 42.07 倍;对比仅加入 Cumate 诱导条件下提升了 15.71 倍和 12.86 倍。在同时加入 DOX 和 Cumate 的条件下, TRE_{3G}CuO 和 TRE_{3G}CuO-BGI 的诱导转录活性为 7.54E+06 RLU 和 1.29E+07 RLU,对比无诱导条件分别提升了 2381.05 倍和 1011.81 倍;对比仅加入 Cumate 诱导条件下提升了 552.99 倍和 309.27 倍;对比仅加入 DOX 诱导条件下提升了 35.19 倍和 24.05 倍。基于以上结果, TRE_{3G}CuO 和 TRE_{3G}CuO-BGI 复合响应元件按照不同诱导剂调控组合可以在不同转录活性级下调控目的核酸片段的诱导转录活性,其调控范围分别为 4.31 倍至 2381.05 倍和 3.27 倍至 1011.81 倍,在同时加入 DOX 和 Cumate 诱导剂时,最大诱导转录活性分别可以达到 CMV 启动子的 47.84%和 81.54%。再进一步优化 DOX 和/或 Cumate 诱导剂浓度的条件下,可以进一步在不同转录活性级下细致调控目的核酸片段的转录活性。