

复合物及其用途

技术领域

本申请涉及生物医药领域，具体的涉及一种用于基因编辑的复合物及其用途。

背景技术

基因组表观遗传修饰的异常与许多疾病的发生与发展息息相关，如生活中常见的代谢紊乱、心血管疾病和癌症等，而基因表观遗传编辑工具可在不改变基因序列的情况下达到相应基因转录调控目的，该过程不会造成永久性的 DNA 损伤，也不会产生有害突变和脱靶效应；并且表观遗传疗法可以通过同时调节多个基因活性来提供更好的治疗效果，弥补基因治疗存在的不足，以及为治疗这类疾病提供了新的希望。另一方面，基因定点修饰技术的发展更使得在天然染色质环境中有针对性地进行表观基因组编辑和转录调控成为可能，特别是 CRISPR/Cas9 技术的发展。

目前，表观遗传修饰和转录调控的设计主要基于核酸酶的结合运用，其中以工程化缺陷型核酸酶(dCas9)的结合最为有效，其设计原理在于把各种表观调控效应器融合到 dCas9 上，通过该核酸酶的靶向并结合目标 DNA 的特性实现特定基因组位点上表观基因组的编辑。这些反式调控结构域和蛋白是通过向启动子区域的 dCas9 靶向位点阻碍 RNA 聚合酶的结合或募集内源性转录复合物来发挥作用的。近期该领域一项突破性的工作是使用一种 dCas-SunTag 系统融合多个拷贝的转录激活或抑制蛋白实现目标基因转录的调控，例如 Hatada Izuho 等人 (PCT/JP2021/006498) 通过 dCas9-GCN4 融合物来招募包含有表观修饰因子（如甲基转移酶和组蛋白乙酰转移酶）、转录抑制调控因子（如 ZIM3）以及抗体的融合物，从而对其调控的靶基因表达进行抑制。目前利用转录调控和表观遗传定点修饰技术在体内治疗表观遗传异常引起疾病的研究还屈指可数，且现有的编辑工具存在一定问题，例如 SunTag 系统招募效果不理想、转录调控效率低下和对靶基因的修饰（例如甲基化修饰）范围受限等。

发明内容

本申请提供了一种用于基因编辑的复合物以及其编码核酸、载体、组合物、细胞等形式，该复合物相比于现有基于 SunTag 招募策略的基因编辑方法能够取得更高的调控效率和更丰

富的靶基因的修饰和调控范围。并且本申请的复合物可以用于制备抑制靶基因表达的产品和制备药物的用途。

一方面，本申请提供一种复合物，其包含第一融合物和第二融合物，所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域和至少一个招募结构域 A，并且其中另一个融合物包含转录阻遏物结构域和至少一个招募结构域 A'；其中，所述招募结构域 A 和招募结构域 A'能够相互作用，以使所述第一融合物和所述第二融合物其中之一的融合物或其部分能够被招募到另一个融合物附近。

在一些实施方案中，所述第一融合物或所述第二融合物包含核酸结合结构域。

在一些实施方案中，所述第一融合物包含 DNA 甲基化结构域、核酸结合结构域和至少一个招募结构域 A，所述第二融合物包含转录阻遏物结构域和至少一个招募结构域 A'。

在一些实施方案中，所述第一融合物中从 N 端到 C 端依次包含 DNA 甲基化结构域、核酸结合结构域和招募结构域 A。

在一些实施方案中，所述第二融合物中从 N 端到 C 端依次包含转录阻遏物结构域和招募结构域 A'，或者从 N 端到 C 端依次包含招募结构域 A'和转录阻遏物结构域。

在一些实施方案中，所述第一融合物包含转录阻遏物结构域、核酸结合结构域和至少一个招募结构域 A，所述第二融合物包含 DNA 甲基化结构域和至少一个招募结构域 A'。

在一些实施方案中，所述第一融合物中从 N 端到 C 端依次包含招募结构域 A、核酸结合结构域和转录阻遏物结构域。

在一些实施方案中，所述第二融合物中从 N 端到 C 端依次包含 DNA 甲基化结构域和招募结构域 A'，或者从 N 端到 C 端依次包含招募结构域 A'和 DNA 甲基化结构域。

在一些实施方案中，所述复合物的特征在于：1) 所述第一融合物从 N 端到 C 端依次包含 DNA 甲基化结构域、核酸结合结构域和招募结构域 A，所述第二融合物从 N 端到 C 端依次包含转录阻遏物结构域和招募结构域 A'；或 2) 所述第一融合物从 N 端到 C 端依次包含 DNA 甲基化结构域、核酸结合结构域和招募结构域 A，所述第二融合物从 N 端到 C 端依次包含招募结构域 A'和转录阻遏物结构域；或 3) 第一融合物中从 N 端到 C 端依次包含招募结构域 A、核酸结合结构域和转录阻遏物结构域，所述第二融合物中从 N 端到 C 端依次包含 DNA 甲基化结构域和招募结构域 A'；或 4) 第一融合物中从 N 端到 C 端依次包含招募结构域 A、核酸结合结构域和转录阻遏物结构域，所述第二融合物从 N 端到 C 端依次包含招募结构域 A'和 DNA 甲基化结构域。

在一些实施方案中，所述核酸结合结构域为 DNA 结合结构域。

在一些实施方案中，所述 DNA 结合结构域选自：TALE 结构域、锌指结构域、tetR 结构域、大范围核酸酶、Cas 蛋白、Argonaute (Ago) 蛋白，以及其同系物、修饰形式或变体。

在一些实施方案中，所述 DNA 结合结构域能够结合靶基因座的目标序列。

在一些实施方案中，所述 DNA 结合结构域能够与引导 RNA 结合。

在一些实施方案中，所述引导 RNA 能够特异性识别所述靶基因座的目标序列并与其杂交。

在一些实施方案中，所述 DNA 结合结构域为 Cas 蛋白，且所述 Cas 蛋白为 II 类 Cas 核酸酶。

在一些实施方案中，所述 Cas 蛋白选自 II 类 II 型 Cas 核酸酶和 II 类 V 型 Cas 核酸酶。

在一些实施方案中，所述 Cas 蛋白为 Cas9 或 Cas12 蛋白。

在一些实施方案中，所述 Cas 蛋白为失活 Cas9 (dCas9) 蛋白或失活 Cas12 (dCas12) 蛋白。

在一些实施方案中，所述 DNA 结合结构域包含 SEQ ID NO: 1-9、343 和 344 中任一项所示的氨基酸序列。

在一些实施方案中，所述招募结构域 A 选自下列两组结构域其中一组中的任一个，所述招募结构域 A' 选自下列两组结构域中另一组中的任一个：1) 通用控制非去阻遏蛋白 4 (GCN4)、来源于分裂绿色荧光蛋白(GFP)的 GFP11 片段和 GVKESLV 多肽；和 2) 单链抗体(scFv)、来源于分裂绿色荧光蛋白(GFP)的 GFP1-10 片段和 PDZ 蛋白结构域。

在一些实施方案中，所述复合物中：1) 所述招募结构域 A 和所述招募结构域 A' 其中之一的结构域为 GCN4，并且其中另一个结构域为 scFv；或 2) 所述招募结构域 A 和所述招募结构域 A' 其中之一的结构域为 GFP11 片段，并且其中另一个结构域为 GFP1-10；或 3) 所述招募结构域 A 和所述招募结构域 A' 其中之一的结构域为 GVKESLV，并且其中另一个结构域为 PDZ 蛋白结构域。

在一些实施方案中，所述 DNA 甲基化结构域包含至少一种 DNA 甲基转移酶或其功能活性片段。

在一些实施方案中，所述 DNA 甲基转移酶选自 DNMT3A、DNMT3B、DNMT3c、DNMT1、DNMT2 和 DNMT3L。

在一些实施方案中，所述 DNA 甲基化结构域包含至少一个 DNMT3A 和至少一个 DNMT3L。

在一些实施方案中，所述 DNA 甲基转移酶包含 SEQ ID NO: 19-24 任一项所示的氨基酸序列。

在一些实施方案中, 所述 DNA 甲基化结构域包含 DNMT3A-DNMT3L 结构域或 DNMT3L-DNMT3A 结构域; 其中, -表示其两端的结构域按照从 N 端到 C 端的顺序直接或间接地连接。

在一些实施方案中, 所述转录阻遏物选自下列所示结构域中的一种或多种: KRAB, ZIM3, ZNF680, ZNF554, ZNF264, ZNF582, ZNF324, ZNF669, ZNF354A, ZNF82, ZNF595, ZNF419, ZNF566, ZIM2, EHMT2, SUV39H1, ZFPM1, TRIM28, EZH2, MXD1, SID, LSD1, HP1a, HDAC3, HDAC1, PRMT1, SETDB1, hSIRT1, ZNF436, ZNF257, ZNF675, ZNF490, ZNF320, ZNF331, ZNF816, ZNF41, ZNF189, ZNF528, ZNF543, ZNF140, ZNF610, ZNF350, ZNF8, ZNF30, ZNF98, ZNF677, ZNF596, ZNF214, ZNF37A, ZNF34, ZNF250, ZNF547, ZNF273, ZFP82, ZNF224, ZNF33A, ZNF45, ZNF175, ZNF184, ZFP28-1, ZFP28-2, ZNF18, ZNF213, ZNF394, ZFP1, ZFP14, ZNF416, ZNF557, ZNF729, ZNF254, ZNF764, ZNF785, ZNF10, CBX5, RYBP, YAF2, MGA, CBX1, SCM1, MPP8, SUMO3, HERC2, BIN1, PCGF2, TOX, FOXA1, FOXA2, IRF2BP1, IRF2BP2, IRF2BPL IRF-2BP1_2 N-terminal domain, HOXA13, HOXB13, HOXC13, HOXA11, HOXC11, HOXC10, HOXA10, HOXB9, HOXA9, ZFP28, ZN334, ZN568, ZN37A, ZN181, ZN510, ZN862, ZN140, ZN208, ZN248, ZN571, ZN699, ZN726, ZIK1, ZNF2, Z705F, ZNF14, ZN471, ZN624, ZNF84, ZNF7, ZN891, ZN337, Z705G, ZN529, ZN729, ZN419, Z705A, ZN302, ZN486, ZN621, ZN688, ZN33A, ZN554, ZN878, ZN772, ZN224, ZN184, ZN544, ZNF57, ZN283, ZN549, ZN211, ZN615, ZN253, ZN226, ZN730, Z585A, ZN732, ZN681, ZN667, ZN649, ZN470, ZN484, ZN431, ZN382, ZN254, ZN124, ZN607, ZN317, ZN620, ZN141, ZN584, ZN540, ZN75D, ZN555, ZN658, ZN684, RBAK, ZN829, ZN582, ZN112, ZN716, HKR1, ZN350, ZN480, ZN416, ZNF92, ZN100, ZN736, ZNF74, ZN443, ZN195, ZN530, ZN782, ZN791, ZN331, Z354C, ZN157, ZN727, ZN550, ZN793, ZN235, ZN724, ZN573, ZN577, ZN789, ZN718, ZN300, ZN383, ZN429, ZN677, ZN850, ZN454, ZN257, ZN264, ZN485, ZN737, ZNF44, ZN596, ZN565, ZN543, ZFP69, SUMO1, ZNF12, ZN169, ZN433, ZN175, ZN347, ZNF25, ZN519, Z585B, ZN517, ZN846, ZN230, ZNF66, ZN713, ZN816, ZN426, ZN674, ZN627, ZNF20, Z587B, ZN316, ZN233, ZN611, ZN556, ZN234, ZN560, ZNF77, ZN682, ZN614, ZN785, ZN445, ZFP30, ZN225, ZN551, ZN610, ZN528, ZN284, ZN418, ZN490, ZN805, Z780B, ZN763, ZN285, ZNF85, ZN223, ZNF90, ZN557, ZN425, ZN229, ZN606, ZN155, ZN222, ZN442, ZNF91, ZN135, ZN778, ZN534, ZN586, ZN567, ZN440, ZN583, ZN441, ZNF43, ZN589,

ZN563, ZN561, ZN136, ZN630, ZN527, ZN333, Z324B, ZN786, ZN709, ZN792, ZN599, ZN613, ZF69B, ZN799, ZN569, ZN564, ZN546, ZFP92, ZN723, ZN439, ZFP57, ZNF19, ZN404, ZN274, CBX3, ZN250, ZN570, ZN675, ZN695, ZN548, ZN132, ZN738, ZN420, ZN626, ZN559, ZN460, ZN268, ZN304, ZN605, ZN844, SUMO5, ZN101, ZN783, ZN417, ZN182, ZN823, ZN177, ZN197, ZN717, ZN669, ZN256, ZN251, CBX4, CDY2, CDYL2, ZN562, ZN461, Z324A, ZN766, ID2, ZN214, CBX7, ID1, CREM, SCX, ASCL1, ZN764, SCML2, TWST1, CREB1, TERF1, ID3, CBX8, GSX1, NKX22, ATF1, TWST2, ZNF17, TOX3, TOX4, ZMYM3, I2BP1, RHXF1, SSX2, I2BPL, ZN680, TRI68, HXA13, PHC3, TCF24, HXB13, HEY1, PHC2, ZNF81, FIGLA, SAM11, KMT2B, HEY2, JDP2, HXC13, ASCL4, HHEX, GSX2, ETV7, ASCL3, PHC1, OTP, I2BP2, VGLL2, HXA11, PDLI4, ASCL2, CDX4, ZN860, LMBL4, PDIP3, NKX25, CEBPB, ISL1, CDX2, PROP1, SIN3B, SMBT1, HXC11, HXC10, PRS6A, VSX1, NKX23, MTG16, HMX3, HMX1, KIF22, CSTF2, CEBPE, DLX2, PPARG, PRIC1, UNC4, BARX2, ALX3, TCF15, TERA, VSX2, HXD12, CDX1, TCF23, ALX1, HXA10, RX, CXXC5, SCML1, NFIL3, DLX6, MTG8, CEBPD, SEC13, FIP1, ALX4, LHX3, PRIC2, MAGI3, NELL1, PRRX1, MTG8R, RAX2, DLX3, DLX1, NKX26, NAB1, SAMD7, PITX3, WDR5, MEOX2, NAB2, DHX8, CBX6, EMX2, CPSF6, HXC12, KDM4B, LMBL3, PHX2A, EMX1, NC2B, DLX4, SRY, ZN777, ZN398, GATA3, BSH, SF3B4, TEAD1, TEAD3, RGAP1, PHF1, GATA2, FOXO3, ZN212, IRX4, ZBED6, LHX4, SIN3A, RBBP7, NKX61, R51A1, MB3L1, DLX5, NOTC1, TERF2, ZN282, RGS12, ZN840, SPI2B, PAX7, NKX62, ASXL2, FOXO1, GATA1, ZMYM5, LRP1, MIXL1, SGT1, LMCD1, CEBPA, SOX14, WTIP, PRP19, NKX11, RBBP4, DMRT2, SMCA2, 及其功能活性片段。

在一些实施方案中, 所述转录阻遏物结构域包含 SEQ ID NOs: 25-50 任一项所示的氨基酸序列。

在一些实施方案中, 所述复合物中: 1) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE -n×GCN4, 并且其中另一个融合物包含转录阻遏物结构域-scFv; 或 2) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE -scFv, 并且其中另一个融合物包含转录阻遏物结构域-GCN4; 或 3) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE -n×GFP11, 并且其中另一个融合物包含转录阻遏物结

构域-GFP1-10; 或 4) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE -GFP1-10, 并且其中另一个融合物包含转录阻遏物结构域-GFP11; 或 5) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE -n×GCN4, 并且其中另一个融合物包含 scFv-转录阻遏物结构域; 或 6) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE -scFv, 并且其中另一个融合物包含 GCN4-转录阻遏物结构域; 或 7) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE -n×GFP11, 并且其中另一个融合物包含 GFP1-10-转录阻遏物结构域; 或 8) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE -GFP1-10, 并且其中另一个融合物包含 GFP11-转录阻遏物结构域; 其中, -表示其两端的结构域按照从 N 端到 C 端的顺序直接或间接地连接; n×GCN4 或 n×GFP11 分别表示 n 个通过接头序列连接的 GCN4 拷贝或 n 个通过接头序列连接的 GFP11 拷贝, n 选自 1 至 20 的任一整数。

在一些实施方案中, 所述第一融合物和/或第二融合物包含 SEQ ID NO: 51-76、78-82、85-93、103-105、110-115、123、124、361 和 362 中任一项所述的氨基酸序列。

在一些实施方案中, 所述复合物包含 SEQ ID NO: 133-142、153、154、158-163、168、345 和 346 中任一项所示的氨基酸序列。

在一些实施方案中, 所述复合物中: 1) 所述第一融合物和所述第二融合物其中之一的融合物包含 n×GCN4-dCas9 或 dCas12 或 TALE -转录阻遏物结构域, 并且其中另一个融合物包含 DNA 甲基化结构域-scFv; 或 2) 所述第一融合物和所述第二融合物其中之一的融合物包含 scFv-dCas9 或 dCas12 或 TALE -转录阻遏物结构域, 并且其中另一个融合物包含 DNA 甲基化结构域-GCN4; 或 3) 所述第一融合物和所述第二融合物其中之一的融合物包含 n×GFP11-dCas9 或 dCas12 或 TALE -转录阻遏物结构域, 并且其中另一个融合物包含 DNA 甲基化结构域-GFP1-10; 或 4) 所述第一融合物和所述第二融合物其中之一的融合物包含 GFP1-10-dCas9 或 dCas12 或 TALE -转录阻遏物结构域, 并且其中另一个融合物包含 DNA 甲基化结构域-GFP11; 或 5) 所述第一融合物和所述第二融合物其中之一的融合物包含 n×GCN4-dCas9 或 dCas12 或 TALE -转录阻遏物结构域, 并且其中另一个融合物包含 scFv-DNA 甲基化结构域; 或 6) 所述第一融合物和所述第二融合物其中之一的融合物包含 scFv-dCas9 或 dCas12 或 TALE -转录阻遏物结构域, 并且其中另一个融合物包含 GCN4-DNA 甲基化结构域; 或 7) 所述第一融合物和所述第二融合物其中之一的融合物包含 n×GFP11-dCas9 或 dCas12

或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 GFP1-10-DNA 甲基化结构域；或 8) 所述第一融合物和所述第二融合物其中之一的融合物包含 GFP1-10-dCas9 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 GFP11-DNA 甲基化结构域；其中，- 表示其两端的结构域按照从 N 端到 C 端的顺序直接或间接地连接； $n \times \text{GCN4}$ 或 $n \times \text{GFP11}$ 分别表示 n 个通过接头序列连接的 GCN4 拷贝或 n 个通过接头序列连接的 GFP11 拷贝， n 选自 1 至 20 的任一整数。

在一些实施方案中，所述第一融合物和/或第二融合物包含 SEQ ID NO: 83、84、94-102、106-109、116-122、363 和 364 中任一项所示的氨基酸序列。

在一些实施方案中，所述复合物包含 SEQ ID NO: 143-152、155-157、164-167、347 和 348 中任一项所示的氨基酸序列。

在一些实施方案中，所述复合物还包含核定位信号和/或标记物结构域。

另一方面，本申请提供编码本申请所述复合物的核酸。

在一些实施方案中，所述核酸为重组载体。

在一些实施方案中，所述重组载体还包括非编码区。

在一些实施方案中，所述非编码区选自内含子、调节元件、启动子、增强子、终止序列以及 5'和 3'非翻译区。

在一些实施方案中，所述核酸包含编码所述第一融合物的第一核酸片段，以及编码所述第二融合物的第二核酸片段。

在一些实施方案中，其中所述第一核酸片段与第二核酸片段通过编码剪切肽的核酸片段连接。

在一些实施方案中，所述剪切肽为 2A 肽和/或 IRES。

在一些实施方案中，所述 2A 肽选自 P2A、T2A、E2A 和 F2A。

在一些实施方案中，所述核酸包含 SEQ ID NO: 169-335 和 349-360 中任一项所示的核酸序列。

另一方面，本申请提供一种递送载体，所述递送载体包含本申请所述的复合物和/或本申请所述的核酸，以及任选地包含脂质体和/或脂质纳米颗粒。

另一方面，本申请提供一种组合物，所述组合物包含本申请所述的复合物、本申请所述的核酸、和/或本申请所述的递送载体。

另一方面，本申请提供一种细胞，所述细胞包含本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、和/或本申请所述的组合物。

另一方面，本申请提供一种试剂盒，所述试剂盒包含本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、和/或本申请所述的细胞。

另一方面，本申请提供一种调控靶基因表达的方法，所述方法包括施用本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、本申请所述的细胞和/或本申请所述的试剂盒。

在一些实施方案中，所述方法包括将所述复合物、所述核酸、所述递送载体、所述组合物、所述细胞、和/或所述试剂盒引入含有所述靶基因的细胞。

在一些实施方案中，所述方法包括使所述复合物、所述核酸、所述递送载体、和/或所述组合物接触所述靶基因附近和/或所述靶基因的调控元件。

在一些实施方案中，所述调控元件包含核心启动子、近端启动子、远端增强子、沉默子、绝缘子元件、边界元件和/或基因座控制区。

另一方面，本申请提供一种治疗或缓解与靶基因表达异常和/或靶基因活性异常相关的疾病或其病症的方法，所述方法包括向有需要的受试者施用有效量的本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、本申请所述的细胞和/或本申请所述的试剂盒。

另一方面，本申请提供本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、本申请所述的细胞和/或本申请所述的试剂盒用于制备药物的用途，所述药物用于治疗或缓解与靶基因表达异常和/或靶基因活性异常相关的疾病或其病症。

另一方面，本申请提供本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、本申请所述的细胞或本申请所述的试剂盒，其用于治疗或缓解与靶基因表达异常和/或靶基因活性异常相关的疾病或其病症。

本领域技术人员能够从下文的详细描述中容易地洞察到本申请的其它方面和优势。下文的详细描述中仅显示和描述了本申请的示例性实施方式。如本领域技术人员将认识到的，本申请的内容使得本领域技术人员能够对所公开的具体实施方式进行修改而不脱离本申请所涉及发明的精神和范围。相应地，本申请的附图和说明书中的描述仅仅是示例性的，而非为限制性的。

附图说明

本申请所涉及的发明的具体特征如所附权利要求书所显示。通过参考下文中详细描述 of 示例性实施方式和附图能够更好地理解本申请所涉及发明的特点和优势。对附图简要说明如

下：

图 1 显示的是本申请复合物的结构示意图。

图 2 显示的是本申请复合物对 PTP1b 基因表达的抑制效果。

图 3 和图 4 显示的是本申请复合物对 PCSK9 基因表达的抑制效果。

具体实施方式

以下由特定的具体实施例说明本申请发明的实施方式，熟悉此技术的人士可由本说明书所公开的内容容易地了解本申请发明的其他优点及效果。

术语定义

在本申请中，术语“招募”通常是针对蛋白质分子间的招募作用，其具体是指蛋白质招募其他分子来执行特定的生物学功能。这种招募作用主要依赖分子间相互作用的亲和性，且通常其亲和性被认为与蛋白分子的空间结构相关，较为复杂。相互作用机制示例性地可包含但不限于氢键、离子相互作用、疏水相互作用、范德华力等非共价键作用，。例如，一些蛋白质可以招募酶来催化化学反应，或者招募其他蛋白质来形成复合物。这些招募作用对于许多细胞过程至关重要，如信号转导、DNA 复制和基因表达等。

在本申请中，术语“核酸结合结构域”通常是指能够与特定核酸结合的多肽或组合物的部分，其可以包含接触核酸的区域、核酸和/或蛋白质物质。核酸结合结构域的示例可包括但不限于：螺旋-转-螺旋结构域、锌指结构域、亮氨酸拉链(bZIP)结构域、有翼的螺旋结构域、有翼的螺旋转螺旋结构域、螺旋-环-螺旋结构域、HMG-box 结构域、Wor3 结构域、免疫球蛋白结构域、B3 结构域、TALE 结构域和/或 CRISPR/CasX 蛋白等的结构域。

在本申请中，术语“DNA 结合结构域”通常是指独立折叠的蛋白质结构域，其含有识别双链或单链 DNA 的至少一个基序。例如，所述 DNA 结合域可识别特异性 DNA 序列（识别或调节序列）或具有对 DNA 的一般亲和性。在某些情形下，DNA 结合域的其他结构域通常调节 DNA 结合域的活性；DNA 结合功能可以是结构性的或者包括转录调节，有时这两种作用是重叠的。在根据本申请所提供的方法和基因表达调节分子的某些实施方案中，DNA 结合域可包含(DNA)核酸酶，诸如能够以序列特异性方式靶向 DNA 或者能够被指导或指示以序列特异性方式靶向 DNA 的核酸酶，诸如 CRISPR-Cas 系统、锌指核酸酶(ZFN)、转录激活因子样效应因子核酸酶(TALEN)或大范围核酸酶。在一些实施方案中，DNA 结合域是源自 CRISPR-Cas 系统的 DNA 核酸酶。例如，该源自 CRISPR-Cas 系统的 DNA 核酸酶是 Cas 蛋白。

在本申请中，术语“TALE DNA 结合结构域”或“TALE”是包含一个或多个 TALE 重复结构域/单元的多肽。天然存在的 TALE 或“野生型 TALE”是由变形菌的众多物种分泌的核酸结合蛋白。TALE 多肽含有由高度保守的单体多肽的串联重复构成的核酸结合结构域，所述单体多肽长度主要为 33、34 或 35 个氨基酸，并且主要在氨基酸位置 12 和 13 中彼此不同。在优选的实施方案中，所述核酸是 DNA。如本文使用的，TALE 的多肽单体用于指在 TALE 核酸结合结构域内高度保守的重复多肽序列，并且术语“重复可变双残基”或“RVD”用于指在多肽单体的位置 12 和 13 处高度可变的氨基酸。包含在 DNA 结合结构域内的 TALE 单体的一般表示是 $X_{1-11}-(X_{12}X_{13})-X_{14-33 \text{ 或 } 34 \text{ 或 } 35}$ ，其中下标指示氨基酸位置，并且 X 表示任何氨基酸。 $X_{12}X_{13}$ 指示 RVD。在一些 TALE 多肽单体中，在位置 13 处的可变氨基酸缺失或不存在，并且在此类单体中，RVD 由单个氨基酸组成。在此类情况下，RVD 可以可替代地表示为 X^* ，其中 X 表示 X_{12} ，并且 (*) 指示 X_{13} 不存在。DNA 结合结构域包含 TALE 单体的几个重复，并且这可以表示为 $(X_{1-11}-(X_{12}X_{13})-X_{14-33 \text{ 或 } 34 \text{ 或 } 35})_z$ ，其中在优选的实施方案中，z 是至少 5-40。在进一步优选的实施方案中，z 是至少 10-26。

TALE 单体具有由在其 RVD 内的氨基酸类型决定的核苷酸结合亲和力。例如，具有 NI 的 RVD 的多肽单体优先与腺嘌呤 (A) 结合，具有 NG 的 RVD 的多肽单体优先与胸腺嘧啶 (T) 结合，具有 HD 的 RVD 的多肽单体优先与胞嘧啶 (C) 结合，并且具有 NN 的 RVD 的单体优先与腺嘌呤 (A) 和鸟嘌呤 (G) 结合。在另外一些实施方案中，具有 IG 的 RVD 的单体优先与 T 结合。因此，在 TALE 的核酸结合结构域中的多肽单体重复的数目和次序决定其核酸靶特异性。在本申请进一步的实施方案中，具有 NS 的 RVD 的单体识别所有四个碱基对，并且可以与 A、T、G 或 C 结合。TALE 的结构和功能例如在 Moscou 等人, *Science*326:1501(2009); Boch 等人, *Science*326:1509-1512(2009); 和 Zhang 等人, *Nature Biotechnology*29:149-153(2011) 中进一步描述，所述参考文献各自整体通过引用并入。TALE 的重复结构域参与 TALE 与其同源靶 DNA 序列的结合。这些重复单元(或称“重复序列”)展现与天然存在的 TALE 蛋白内的其它 TALE 重复序列的至少一些序列同源性。参见例如美国专利公布号 20110301073。本申请涉及的 TALE 结合结构域可以“工程改造”以结合于预定核苷酸序列，例如经由天然存在的 TALE 蛋白的识别螺旋区域的工程改造(改变一个或多个氨基酸)。因此，工程改造的 DNA 结合蛋白(TALE)是非天然存在的蛋白。用于工程改造 DNA 结合蛋白的方法的非限制性实例是设计和选择。所设计的 DNA 结合蛋白是非天然存在的蛋白，其设计和/或组成主要源于合理的设计标准。合理的设计标准包括应用替换规则和用于处理储存现有的 TALE 设计和结合数据的信息数据库中的信息的计算化算法。参见例如美国专利 6,140,081; 6,453,242; 和 6,534,261;

还参见 WO 98/53058; WO 98/53059; WO 98/53060; WO02/016536 和 WO 03/016496 以及美国公布号 20110301073。

在本申请中,“Cas 酶”可与“Cas 蛋白”、“CRISPR 蛋白”、“CRISPR 酶”、“CRISPR-Cas 蛋白”、“CRISPR-Cas 酶”、“Cas”、“CRISPR 效应子”或“Cas 效应子蛋白”互换地使用,其通常是指与 CRISPR 序列互补的一类酶,能够使用 CRISPR 序列作为指导(guide),从而识别和切割特定的 DNA 链。Cas 蛋白的非限制性实例包括: Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9 (也称为 Csn1 和 Csx12)、Cas10、Csy1、Csy2、Csy3、Cse1、Cse2、Cse1、Cse2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csl17、Csl14、Csl10、Csl16、CsaX、Csx3、Csl1、Csl15、Csf1、Csf2、Csf3、Csf4, 和/或他们的同系物、或其修饰形式。这些蛋白是已知的,例如,化脓链球菌 Cas9 蛋白的氨基酸序列可见于 SwissProt 数据库登录号 Q99ZW2 下。

在本申请中,术语“II 类 Cas 核酸酶”通常是指根据 CRISPR/Cas 基因座的更新分类方案 (Makarova 等人, (2015) Nat Rev Microbiol[自然微生物学综述], 13(11):722-36; Shmakov 等人, (2015) Mol Cell[分子细胞], 60:385-397) 所定义的,以单一蛋白形式发挥识别和/或切割功能的一类 Cas 蛋白。

在本申请中,术语“II 类 II 型 Cas 核酸酶和 II 类 V 型 Cas 核酸酶”通常是指 II 类 Cas 核酸酶中,是单蛋白的、RNA 指导的内切核酸酶。在这其中,II 型和 V 型中的 V-B 型 Cas 核酸酶需要 tracrRNA (反式激活 CRISPR RNA) 和 crRNA (CRISPR RNA) 共同作用才能正常发挥功能,且 crRNA 和 tracrRNA 可以人工组合成一个单一的向导 RNA (sgRNA); V 型中的 V-A 型 Cas 核酸酶则需要单独使用 crRNA 行使向导功能。II 类 II 型 Cas 核酸酶的非限制性示例包括 Cas9 及其家族相关核酸酶,II 类 V 型 Cas 核酸酶的非限制性示例包括 Cas12a (也称为 Cpf1)、Cas12b (也称为 C2c1)、Cas12c (也称为 C2c3)、Cas12d (CasY)、Cas12e(CasX)、Cas12g、Cas12h、Cas12i、C2c1、C2c4、C2c5、C2c8、C2c9、C2c10、Cas14a、Cas14b、Cas14c 核酸酶和/或 TnpB。

在本申请中,术语“dCas”可指 dCas 蛋白或其片段。例如,如本文中所用,“dCas9”可指 dCas9 蛋白或其片段。如本文中所用,术语“iCas”和“dCas”可互换使用,指无催化活性的 CRISPR 相关蛋白。在一个实施方案中,dCas 蛋白在 DNA 切割结构域中包含一个或多个突变。在一个实施方案中,dCas 蛋白在 RuvC 或结构域中包含一个或多个突变。在一个实施方案中,dCas 分子在 RuvC 和 HNH 结构域中都包含一个或多个突变。在一个实施方案中,dCas 蛋白是野生型 Cas 蛋白的片段。在一个实施方案中,dCas 蛋白包含来自野生型 Cas 蛋白

的功能结构域,其中该功能结构域选自 Reel 结构域、桥螺旋结构域或 PAM 相互作用结构域。在一个实施方案中,与相应的野生型 Cas 蛋白的核酸酶活性相比,dCas 的核酸酶活性降低了至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%或至少 99%。

在本申请中,术语“能够结合”可与“结合于”、“特异性地识别”、“靶向”等互换地使用,通常是指结合分子(例如,本申请的基因表达调节分子)能够与靶基因或靶位点上的核苷酸相互作用,或者该结合分子(例如,本申请的基因表达调节分子)对靶基因或靶位点具有足够的亲和力,这种相互作用可以通过缀合、偶联、附着、提供互补性、提供共价作用力或提供非共价作用力、提高结合稳定性等方式。

在本申请中,术语“引导 RNA”、“向导 DNA”和“gRNA”可互换地使用,其通常是指能够指导核酸酶(例如 Argonaute,或 Ago)结合和/或剪切靶标基因的 DNA 分子。在一些优选的实施方案中,向导 DNA 可以包括:为单链 DNA 分子(ssDNA)、为 5' 端磷酸化的单链 DNA 分子、为 5' 端羟基化的单链 DNA 分子、具有能够和靶标基因互补的碱基片段和/或具有 8-35nt 的长度。在本申请的一些实施方案中,术语“引导 RNA”是指包含以下的 RNA:

(1) 结合于向导 RNA 指导的核酸内切酶(例如 II 类 Cas 核酸酶,例如 II 型、V 型或 VI 型 Cas 核酸内切酶)且活化 RNA 指导的核酸内切酶的“活化”核苷酸序列;和(2) 包含与靶核酸杂交的核苷酸序列的“靶”核苷酸序列。“活化”核苷酸序列和“靶”核苷酸序列可以在分开的 RNA 分子(例如“双向导 RNA”)上;或可以在相同的 RNA 分子(“单向导 RNA”,也称为 sgRNA)上。

在本申请中,术语“DNA 甲基转移酶”通常是指催化甲基转移至 DNA 的酶。DNA 甲基转移酶的非限制性实例包括 DNMT1、DNMT 3A、DNMT 3B 和 DNMT 3L。例如,通过 DNA 甲基化,DNA 甲基转移酶可以在不更改 DNA 序列的情况下修饰 DNA 片段的活性(例如调控基因表达)。如本文所述,基因表达调节分子可以包括一个或多个(例如两个)DNA 甲基转移酶。当 DNA 甲基转移酶作为基因表达调节分子的一部分包括在内时,DNA 甲基转移酶可以被称为“DNA 甲基转移酶结构域”。在各方面中,DNA 甲基转移酶结构域包含与 DNMT 3A 具有至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%序列一致性的氨基酸序列的变异体或同源物。在各方面中,DNA 甲基转移酶结构域包含与 DNMT 3L 具有至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 91%、至少 92%、至少 93%、至少

94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%序列一致性的氨基酸序列的变异体或同源物。

在本申请中，术语“功能活性片段”通常是指具有全长蛋白质或核酸的部分区域，但保留或部分保留全长蛋白质或核酸的生物活性或功能的片段。例如，功能活性片段可以保留或部分保留全长蛋白质结合另一种分子的能力。例如，DNA 甲基转移酶的功能活性片段，可以保留或部分保留全长 DNA 甲基转移酶的催化甲基基团转移到 DNA 的生物活性功能。

在本申请中，术语“转录阻遏物”通常是指结合靶核酸序列且导致与靶核酸序列有关的基因产物的表达水平降低的物质和/或试剂，如蛋白质(例如转录因子或其片段)。例如，所述基因产物可以是从小基因转录的 RNA(例如 mRNA)或从自基因转录的 mRNA 翻译的多肽。通常 mRNA 水平中的增加或降低导致从其翻译的多肽水平的增加或降低。可以使用测量 mRNA 或蛋白的标准技术来测定表达水平。非限制性的转录阻遏物实例包括：mSin3 相互作用结构域(SID)蛋白、甲基-CpG-结合结构域 2(MBD2)、MBD3、DNA 甲基转移酶(DNMT)1(DNMT1)、DNMT2A、DNMT3A、DNMT3B、DNMT3L、视网膜母细胞瘤蛋白(Rb)、甲基 CpG 结合蛋白 2(Mecp2)、GATA-1 及其辅助因子 Fog1、MAT2 调节剂(ROM2)、拟南芥 HD2A 蛋白(AtHD2A)、赖氨酸特异性的脱甲基酶 1(LSD1)和/或 Krüppel-相关盒(KRAB)。

在本申请中，术语“KRAB”也称为“克鲁珀相关盒结构域”或“Krüppel 相关盒结构域”，其通常是指存在于人锌指蛋白的转录因子中的转录抑制结构域的约 45 至约 75 个氨基酸残基。在各方面中，KRAB 结构域可以包括与 ZIM3 KRAB 结构域或 KOX1 KRAB 结构域具有至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%序列一致性的氨基酸序列的变异体或同源物。

在本申请中，术语“分裂绿色荧光蛋白”通常是指能够分裂并在重新组合时立即形成活性绿色荧光蛋白的多肽。

在本申请中，术语“GCN4”是酿酒酵母 (*S. cerevisiae*) 中的一种转录因子，是酵母基因组中的“主调节因子” (“master regulator”)，调节接近十分之一的酵母基因组，它是一种高度保守蛋白，其在哺乳动物中的同源物是转录激活因子 (Activating Transcription factor)-4 (ATF4)。

在本申请中，术语“PDZ 蛋白”通常是指天然存在的含有 PDZ 结构域的蛋白。示例性的 PDZ 蛋白包括 CASK、MPPI、DLG1、DLG2、PSD95、NeDLG、TIP-33、SYN1a、TIP-43、LDP、LIM、LIMK1、LIMK2、MPP2、N0S1、AF6、PTN_4、prIL16、41.8kD、KIAA0559、RGS 12、KIAA0316、DVL1、TIP-40、TIAM1、MINT1、MAGI-1、MAGI-2、MAGI-3、KIAA0303、

CBP、MINT3、TIP-2、KIAA0561 和/或 TIP-I。

在本申请中，术语“单链抗体”或“scFv (Single Chain Antibody)”通常是指含有一个或多个抗原结合部位的单链多肽。另外，尽管 Fv 片段的 H 和 L 链是由不同的基因编码的，它们可直接或通过肽而连接在一起，例如，通过重组的方法，可用合成的衔接物(linker)将 H 和 L 链连接成单一蛋白链(称为单链抗体，sAb; Bird et al.1988 Science242: 423-426; and Huston et al.1988 PNAS 85: 5879-5883)。该单链抗体也被包括在术语“抗体”之中，可在设计和制造多特异性结合分子中用作结合决定簇，并且通过重组技术或完整抗体的酶促或化学切割可制备所述单链抗体。

在本申请中，术语“直接或间接结合”通常是指相对的“直接相连”或“间接相连”。“直接相连”通常是指直接连接。例如，所述直接相连可以为相连的物质(例如氨基酸序列区段)之间没有间隔成分(例如氨基酸残基或其衍生物)而直接相连接的情况；例如氨基酸序列区段 X 与另一氨基酸序列区段 Y 通过氨基酸序列区段 X 的 C 端氨基酸与氨基酸序列区段 Y 的 N 端氨基酸形成的酰胺键直接连接。“间接相连”通常是指相连的物质(例如氨基酸序列区段)之间有间隔成分(例如氨基酸残基或其衍生物)而间接相连接的情况。例如，本申请所用的间隔成分可以是一段氨基酸残基，其序列选自 SEQ ID NO: 125-132 (SEQ ID NO: 126 为 GSG) 中任一项所示的氨基酸序列。

在本申请中，“核定位序列”或“NLS”通常是指将蛋白导向至细胞核的肽。在某些实施方案中，NLS 包括五个碱性带正电的氨基酸。NLS 可以位于肽链上的任何位置。在某些实施方案中，NLS 为衍生自 SV40 的 NLS。在某些实施方案中，NLS 包括 SEQ ID NO:338-340 中任一项所示的序列。在一些实施方案中，NLS 具有与 SEQ ID NO:338-340 中任一项至少 50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或 100%序列一致性的氨基酸序列。

在本申请中，术语“标记物”指的是肽，可以将其引入表达载体中，其可以用来允许一个或多个载体插入片段的表达产物的删除和/或纯化。这样的标记物是本领域公知的并包括放射性标记的氨基酸或连接可以通过标记的抗生物素蛋白(例如，含有可以通过光学或比色方法检测的荧光标记或酶活性的抗生物素蛋白链菌素)检测的生物素部分的多肽。亲和性标记物如 FLAG，谷胱甘肽-S-转移酶，麦芽糖结合蛋白，纤维素结合结构域，硫氧还蛋白，NusA，mistin，几丁质结合结构域，角质酶，AGT，GFP 和广泛使用的其它标记物，如蛋白表达和纯化系统中所用的。用于多肽的更多非限制性实例包括，但不限于以下的：组氨酸标记物，放射性同位素或放射性核素(利如，³H，¹⁴C，³⁵S，⁹⁰Y，⁹⁹Tc，¹¹¹In，¹²⁵I，¹⁷⁷Lu，¹⁶⁶Ho 或 ¹⁵³Sm)；

荧光标记物(例如, FITC, 若丹明, 镧系磷), 酶标记物(例如, 辣根过氧化物酶, 荧光素酶, 碱性磷酸酶); 化学发光标记物; 生物素基团; 由第二报告子识别的悬垂多肽抗原决定部位(例如, 亮氨酸拉链对序列, 用于二抗的结合位点, 金属结合结构域, 抗原决定部位标记物); 和磁性试剂, 如钆螯合物。

在本申请中, 术语“核酸”与“多核苷酸”、“核苷酸”、“核苷酸序列”和“寡核苷酸”可互换地使用, 其通常是指核苷酸(例如, 脱氧核糖核苷酸或核糖核苷酸)和其呈单链、双链或多链形式的聚合物或其互补物。例如, 核苷酸可以为核糖核苷酸、脱氧核糖核苷酸或其修饰版本。例如, 核苷酸可以为单链和双链 DNA、单链和双链 RNA 以及具有单链和双链 DNA 和 RNA 的混合物的杂交分子。例如, 核苷酸可以包括但不限于任何类型的 RNA, 例如 mRNA、siRNA、miRNA、sgRNA 和引导 RNA, 以及任何类型的 DNA、基因组 DNA、质粒 DNA 和微环 DNA 以及其任何片段。所述术语还涵盖含有已知核苷酸类似物或经修饰的主链残基或键的核酸, 所述核酸为合成的、天然存在的和非天然存在的。

在本申请中, 术语“编码……的序列”或“编码……的核酸”通常是指包含编码蛋白质的核苷酸序列的核酸(RNA 或 DNA 分子)。编码序列还可包括与调控元件可操作地连接的起始和终止信号, 所述调控元件包含能够在对其施用了核酸的个体或哺乳动物的细胞中指导表达的启动子和多腺苷酸化信号。可对编码序列进行密码子优化。在本申请中, 术语“内含子”通常是指包括经过转录的, 但却从 RNA 转录本中通过将序列(外显子)两端的任一端拼接在一起而被去除的 DNA 片段。内含子被认为是基因的蛋白编码区内的干扰序列, 且通常不含有由该基因产生的蛋白所代表的信息。

在本申请中, 术语“重组载体”通常是指能够转运它以及与之连接的另一种核酸的核酸分子。一种类型的载体是“质粒”, 这指另外的 DNA 区段可以连接到其内的环状双链 DNA 环。可替代地, 载体可以是线性的。另一种类型的载体是病毒载体, 其中另外的 DNA 区段可以连接到病毒基因组内。特定载体能够在它们引入其内的宿主细胞内自主复制(例如, 具有细菌复制起点的细菌载体和附加型哺乳动物载体)。其他载体(例如, 非附加型哺乳动物载体)可以在引入宿主细胞内后整合到宿主细胞的基因组内, 并且从而连同宿主基因组一起复制。

在本申请中, 术语“调节元件”是指能够控制核酸序列表达的遗传元件。例如, 剪接信号、启动子序列、多腺苷酸化信号、转录终止序列、上游调节结构域、复制起点、内部核糖体进入位点(“IRES”)、增强子等, 它们共同提供了编码序列在受体细胞中的复制、转录和翻译。并非所有这些控制序列都需要存在。

在本申请中, 术语“启动子”通常是指控制或调节与启动子可操作地关联的核苷酸序列

(例如, 编码序列)的转录的核苷酸序列。由启动子控制或调节的编码序列可以编码多肽和/或功能性 RNA。通常,“启动子”是指包含 RNA 聚合酶 II 结合位点并指导转录起始的核苷酸序列。通常,启动子位于相对于相应编码序列的编码区起点的 5' 或上游。启动子可以包含其他作为基因表达调节因子的元件;例如,启动子区域。在一些实施方案中,启动子区域可包括至少一个内含子。启动子可以包括例如组成型、诱导型、时间调节的、发育调节的、化学调节的、组织优选的和/或组织特异性的启动子,用于制备重组核酸分子,例如“合成核酸构建体”或“蛋白质-RNA 复合物”。这些不同类型的启动子是本领域已知的。

在本申请中,术语“增强子”通常是指可以被蛋白质(激活蛋白)结合以刺激或增强一种基因或若干种基因的转录的调控 DNA 序列,例如 50-1500bp。这些激活蛋白(又称转录因子)与介体复合物相互作用并募集聚合酶 II 和一般转录因子,然后开始转录基因。增强子通常是顺式作用的,但可以位于基因或它们调控的基因的起始位点的上游或下游。此外,增强子可以是向前或向后方向,并且不需要位于转录起始位点附近以影响转录,因为已发现一些增强子位于起始位点上游或下游的数十万个碱基对。内含子中也可以发现增强子。

在本申请中,术语“剪切肽”是指能够实现剪切蛋白的功能的一类多肽。例如,所述剪切肽可以经核糖体跳跃而非蛋白酶水解来实现蛋白质剪切。例如,所述剪切肽可为剪切 2A 肽,其可包括 T2A、F2A、P2A 和/或 E2A。

在本申请中,术语“递送载体”通常是指能够将试剂(例如,核酸分子)递送至靶细胞的转移媒介物。递送载体可以将试剂递送到特定的细胞亚类。例如,借助递送载体的固有特征或者通过与载体相偶联的部分、包含在其内的部分(或者与载体结合的部分,从而使得该部分和该递送载体维持在一起,进而使得该部分足以靶向递送载体)使递送载体靶向某些类型的细胞。递送载体还可提高要递送的试剂的体内半衰期和/或要递送的试剂的生物利用度。递送载体可包括病毒载体、病毒样颗粒、聚阳离子载体、肽载体、脂质体和/或杂交载体。例如,如果靶细胞是肝细胞,所述递送载体的性质(例如,尺寸、电荷和/或 pH)可以有效地将所述递送载体和/或其中所载的分子递送至靶细胞、降低免疫清除和/或促进在该靶细胞中停留。

在本申请中,术语“脂质体”通常是指通过一个或多个双层的膜与外部介质隔离的具有内部空间的囊泡。在一些实施方案中,所述双层的膜可以通过两性分子形成,如包含空间隔离的亲水性和疏水性结构域的合成或天然来源的脂质;在另一些实施方案中,所述双层的膜可以通过两亲性聚合物和表面活性剂形成。在一些实施方案中,所述脂质体是球形囊泡结构,其由围绕内部水性区室的单层或多层脂质双分子层、和相对不可渗透的外部亲脂性磷脂双分子层组成。在一些实施方案中,脂质体是生物相容的、无毒的,可以递送亲水性和亲脂性药

物分子，保护它们的运载物不被血浆酶降解，并且将它们的负载运输穿过生物膜和血脑屏障(BBB)。脂质体可由几种不同类型的脂质例如磷脂制成。脂质体可包含天然磷脂和脂质(诸如1,2-二硬脂酰基-sn-甘油-3-磷脂酰胆碱(DSPC)、鞘磷脂、卵磷脂酰胆碱、单唾液酸神经节苷脂或其任意组合)。为了改变脂质体的结构和性质，可向脂质体中加入几种其它添加剂。例如，脂质体还可包含胆固醇、鞘磷脂和/或1,2-二油酰基-sn-甘油-3-磷酸乙醇胺(DOPE)，例如，以增加稳定性和/或防止脂质体内部运载物的泄漏。

术语“脂质纳米颗粒(LNP)”通常是指包含通过分子间力彼此物理结合(例如，共价或非共价)的多个(即多于一个)脂质分子的颗粒。LNP可以是例如微球(包括单层和多层囊泡，例如脂质体)、乳液中的分散相、胶团或悬浮液中的内相。LNP可将核酸包封在阳离子脂质颗粒(例如，脂质体)内，并且可被相对容易地递送至细胞。在一些实例中，脂质纳米颗粒不含任何病毒组分，这有助于最小化安全性和免疫原性问题。所述脂质颗粒可用于体外、离体和体内递送。所述脂质颗粒还可用于各种规模的细胞群。本申请的LNP可通过本领域已知的各种方法，例如通过混合有机相与水相来容易地制备。两相的混合可通过微流体装置和撞击流反应器来实现。有机相和水相混合越充分，获得的LNP的包埋率和粒径分布就越好。优选地，LNP的粒径可通过改变有机相与水相的混合速度来调节。混合速度越快，制备的LNP的粒径将越小。包埋效率可通过调节LNP系统的N/P(可电离脂质/核酸)比值来优化。在一些实例中，LNP可用于递送DNA分子和/或RNA分子(例如，Cas、sgRNA的mRNA)。在某些情况下，LNP可用于递送Cas/gRNA的RNP复合物。在一些实施方案中，LNP用于递送mRNA和gRNA。

在本申请中，术语“受试者”通常是指动物，通常是哺乳动物，诸如人、非人灵长类动物(猿、长臂猿、大猩猩、黑猩猩、猩猩、猕猴)、家畜(狗和猫)、农场动物(家禽如鸡和鸭、马、牛、山羊、绵羊、猪)和实验动物(小鼠、大鼠、兔、豚鼠)。人受试者包括胎儿、新生儿、婴儿、青少年和成人受试者。受试者包括动物疾病模型，例如小鼠和血液凝固疾病(诸如HemA)的其它动物模型，和本领域技术人员已知的其它动物模型。

在本申请中，术语“包含”通常是指包括明确指定的特征，但不排除其他要素。

在本申请中，术语“选自”通常是指包括选择的对象以及其所有组合。例如“选自(:)A、B和C”意指包括A、B和C的所有组合，例如，A、B、C、A+B、A+C、B+C或A+B+C。

在本申请中，术语“约”通常是指在指定数值以上或以下0.5%-10%的范围内变动，例如在指定数值以上或以下0.5%、1%、1.5%、2%、2.5%、3%、3.5%、4%、4.5%、5%、5.5%、6%、6.5%、7%、7.5%、8%、8.5%、9%、9.5%、或10%的范围内变动。

发明详述

一方面，本申请提供一种复合物，其包含第一融合物和第二融合物，所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域和至少一个招募结构域 A，并且其中另一个融合物包含转录阻遏物结构域和至少一个招募结构域 A'；其中，所述招募结构域 A 和招募结构域 A'能够相互作用，以使所述第一融合物和所述第二融合物其中之一的融合物或其部分能够被招募到另一个融合物附近。

另一方面，本申请提供编码本申请所述复合物的核酸。例如，所述核酸包含 DNA 和/或 mRNA。例如，所述核酸可用于治疗或缓解与靶基因表达异常和/或靶基因活性异常相关的疾病或其病症。在一些实施方案中，所述核酸为 mRNA；可使用一种或多种修饰技术用于产生更稳定的 mRNA。已知的 mRNA 修饰技术大致可分为三类：用人工合成的非天然核糖核酸代替天然核糖核酸合成 mRNA；添加 5' caps、3' poly (A) “尾”和 UTR（未翻译区）序列；采用特殊的新型配方技术，有效保护 mRNA。其中，优选的 mRNA 修饰技术可以通过人工合成非天然核糖核酸取代天然核糖核酸合成 mRNA。真核 mRNA 上的化学修饰大致可以分为三类：甲基化、伪尿苷(Ψ)和次黄嘌呤。例如，所述化学修饰可选自：假尿苷、N1-甲基假尿苷、N1-乙基假尿苷、2-硫代尿苷、4'-硫代尿苷、5-甲基胞嘧啶、2-硫代-1-甲基-1-脱氮-假尿苷、2-硫代-1-甲基-假尿苷、2-硫代-5-氮杂-尿苷、2-硫代-二氢假尿苷、2-硫代-二氢尿苷、2-硫代-假尿苷、4-甲氧基-2-硫代-假尿苷、4-甲氧基-假尿苷、4-硫代-1-甲基-假尿苷、4-硫代-假尿苷、5-氮杂-尿苷、二氢假尿苷、5-甲基尿苷、5-甲氧基尿苷和 2'-O-甲基尿苷。例如，所述核酸为重组载体，所述重组载体包含编码本申请所述复合物的核酸。例如，重组载体可以是指能够转运与其连接的另一种核酸的核酸分子。重组载体可以包括单链、双链或部分双链的核酸分子；包含一个或多个游离端，没有游离端（例如，环状）的核酸分子；包含 DNA、RNA 或两者的核酸分子；和本领域已知的其他种类的多核苷酸。例如，可以使用病毒载体。病毒载体可包含病毒衍生的 DNA 或 RNA 序列，用于包装成病毒(例如逆转录病毒、复制缺陷型逆转录病毒、腺病毒、复制缺陷型腺病毒、和腺相关病毒 AAV)。病毒和病毒载体可用于体外、离体和/或体内递送。

另一方面，本申请提供一种递送载体，所述递送载体包含本申请所述的复合物和/或本申请所述的核酸，以及任选地包含脂质体和/或脂质纳米颗粒。例如，可以通过物理递送方法将递送载体引入细胞。物理方法的例子包括显微注射、电穿孔和流体动力学递送。例如，LNPs 可以将核酸包裹在阳离子脂质颗粒(例如脂质体)中，并且可以相对容易地递送至细胞。在一些例子中，脂质纳米颗粒不含任何病毒成分，这有助于最大限度地减少安全性和免疫原性问题。

脂质颗粒可用于体外、离体和体内递送。LNP 的成分可包括阳离子脂质，可电离的脂质，聚乙二醇化脂质和/或支持脂质，以及任选的胆固醇组分。

另一方面，本申请提供一种组合物，所述组合物包含本申请所述的复合物、本申请所述的核酸、和/或本申请所述的递送载体。例如，组合物中的复合物、编码所述复合物的核酸（或重组载体）和递送载体可以同时包含在一个组合物中，或者分别包含在不同组合物中。例如，当使用组合物中的复合物、编码所述复合物的核酸（或重组载体）和/或递送载体时，可以同时使用，或者分开使用。

另一方面，本申请提供一种细胞，所述细胞包含本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、和/或本申请所述的组合物。

另一方面，本申请提供一种试剂盒，所述试剂盒包含本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、和/或本申请所述的细胞。例如，所述试剂盒还包含至少一个放置上述组分的容器。例如，所述试剂盒包含一种以上的上述组分，其还包含所述容器以外的第二个、第三个和/或其他容器，其内可以分开放置所述一种以上的上述组分。例如，所述试剂盒可以在容器中放置上述组分的各种组合形式。例如，所述试剂盒还进一步包括缓冲试剂、用于混合的装置、用于测量的装置、用于分选的装置和/或用于标记的装置。例如，所述试剂盒还包括用于容纳各种容器的包装。例如，所述试剂盒还包括关于使用试剂盒组分的说明书。例如，所述说明书包括纸质的实体形式和/或可机读的电子形式。

另一方面，本申请提供一种调控靶基因表达的方法，所述方法包括施用本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、本申请所述的细胞和/或本申请所述的试剂盒。例如，所述抑制靶基因表达的方法是将所述复合物、所述核酸、所述递送载体、所述组合物、所述细胞和/或所述试剂盒引入含有靶基因的细胞。例如，所述引入细胞可以是使用非病毒或基于病毒的转染方式引入细胞。例如，所述非病毒转染方法包括不使用病毒 DNA 或病毒颗粒作为递送系统引入细胞的任何适当方法，非限制性的非病毒转染方法示例包括编码复合物的核酸的纳米颗粒封装(例如脂质纳米颗粒、金纳米颗粒等)、磷酸钙转染、脂质体转染、核转染、声穿孔、通过热休克转染、磁转染和电穿孔。例如，基于病毒的转染方法包括任何适用于本申请所述方法的病毒载体，其非限制性示例包括但不限于反转录病毒、腺病毒、慢病毒和/或腺相关病毒载体。例如，所述抑制靶基因表达的方法还包括将所述复合物、所述核酸、所述递送载体、所述组合物、所述细胞和/或所述试剂盒从外部环境引入细胞中。再例如，所述抑制靶基因表达的方法包括使所述复合物、所述核酸、所述递送载体、和/或所述组合物接触靶基因附近和/或所述靶基因的转录调控元件。例如，所述接触

是指使本申请所述的第一融合物、第二融合物以及向导 RNA 与靶基因附近和/或所述靶基因的转录调控元件接触,并且向导 RNA 与包含 DNA 结合结构域的融合物形成复合物,该复合物特异识别靶基因中的特定区域并与之杂交,同时第一融合物和第二融合物通过其招募结构域 A 和招募结构域 A'直接或间接的相互作用而被募集到 DNA 结合结构域附近,从而对靶核酸的表达进行调控。例如,所述方法包括使本申请所述的第一融合物、第二融合物与向导 RNA 以复合物(例如,组装的核糖核蛋白复合物)的形式存在,并且使该复合物接触靶基因附近和/或所述靶基因的转录调控元件。

另一方面,本申请提供一种治疗或缓解与靶基因表达异常和/或靶基因活性异常相关的疾病或其病症的方法,所述方法包括向有需要的受试者施用有效量的本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、本申请所述的细胞和/或本申请所述的试剂盒。例如,所述治疗方法包含将所述复合物、所述核酸、所述递送载体、所述组合物、所述细胞和/或所述试剂盒与治疗剂混合,并对有需要的受试者进行全身递送,使其广泛暴露于身体的大部分,可以通过本领域已知的任何手段进行,包括但不限于向静脉内、动脉内、皮下、腔内和腹膜内递送。例如,所述治疗方法包含将所述复合物、所述核酸、所述递送载体、所述组合物、所述细胞和/或所述试剂盒与治疗剂混合,并对有需要的受试者进行局部递送,使其直接到达有机体内的靶部位,可以通过例如直接注射到疾病部位(如,肿瘤或炎症部位)或靶器官(如,肝脏、心脏、胰腺、肾脏等)中进行。例如,所述局部递送包括局部施用或局部注射技术,包括但不限于肌内、皮下或皮内注射。例如,所述局部递送不排除全身性的药理学作用。例如,所述疾病包括心血管疾病、非酒精性脂肪肝炎、AMD、老年黄斑病变、2 型糖尿病、肥胖症、肝衰竭、血脂异常(dislipidemia)、糖尿病性动脉粥样硬化和/或高血压。

另一方面,本申请提供本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、本申请所述的细胞和/或本申请所述的试剂盒用于制备药物的用途,所述药物用于治疗或缓解与靶基因表达异常和/或靶基因活性异常相关的疾病或其病症。

另一方面,本申请提供本申请所述的复合物、本申请所述的核酸、本申请所述的递送载体、本申请所述的组合物、本申请所述的细胞或本申请所述的试剂盒,其用于治疗或缓解与靶基因表达异常和/或靶基因活性异常相关的疾病或其病症。

第一融合物或第二融合物

在一些实施方案中,本申请复合物的第一融合物和第二融合物总体上可分为两种情形:

- (1) 两个融合物其中之一包含核酸结合结构域、DNA 甲基化结构域和招募结构域 A,且另

一个融合物包含转录阻遏物结构域和招募结构域 A'，或者 (2) 两个融合物其中之一包含核酸结合结构域、转录阻遏物和招募结构域 A，且另一个融合物包含 DNA 甲基化结构域和招募结构域 A'。

具体地，在上述 (1) 情形下的一些实施方案中，所述两个融合物其中之一从 N 端到 C 端依次可包含 DNA 甲基化结构域、核酸结合结构域和招募结构域 A。例如，在上述 (2) 情形下的一些实施方案中，所述两个融合物其中之一从 N 端到 C 端依次可包含招募结构域 A、核酸结合结构域和转录阻遏物结构域。例如，在上述 (1) 情形下的一些实施方案中，所述两个融合物中的另一个融合物从 N 端到 C 端依次可包含转录阻遏物结构域和招募结构域 A'，或者招募结构域 A' 和转录阻遏物结构域，即转录阻遏物结构域和招募结构域 A' 可互换顺序地连接。例如，在上述 (2) 情形下的一些实施方案中，所述两个融合物中的另一个融合物从 N 端到 C 端依次可包含 DNA 甲基化结构域和招募结构域 A'，或者招募结构域 A' 和 DNA 甲基化结构域，即 DNA 甲基化结构域和招募结构域 A' 可互换顺序地连接。

在一些更具体的实施方案中，所述核酸结合结构域为 DNA 结合结构域。例如，所述 DNA 结合结构域可选自：TALE 结构域、锌指结构域、tetR 结构域、大范围核酸酶、Cas 蛋白、Argonaute (Ago) 蛋白，以及其同系物、修饰形式或变体。例如，所述 DNA 结合结构域可以是 Cas 蛋白，且所述 Cas 蛋白为 II 类 Cas 核酸酶。进一步地，所述 Cas 蛋白可选自 II 类 II 型 Cas 核酸酶和 II 类 V 型 Cas 核酸酶；例如，所述 Cas 蛋白可以是 Cas9 或 Cas12 蛋白。在某些实施方案中，所述 Cas 蛋白可为失活 Cas9 (dCas9) 蛋白或失活 Cas12 (dCas12) 蛋白。例如，本申请所述 DNA 结合结构域可包含但不限于 SEQ ID NO: 1-9、343 和 344 中任一项所示的氨基酸序列。

在一些更具体的实施方案中，所述转录阻遏物选自下列所示结构域中的一种或多种：KRAB, ZIM3, ZNF680, ZNF554, ZNF264, ZNF582, ZNF324, ZNF669, ZNF354A, ZNF82, ZNF595, ZNF419, ZNF566, ZIM2, EHMT2, SUV39H1, ZFP1, TRIM28, EZH2, MXD1, SID, LSD1, HP1a, HDAC3, HDAC1, PRMT1, SETDB1, hSIRT1, ZNF436, ZNF257, ZNF675, ZNF490, ZNF320, ZNF331, ZNF816, ZNF41, ZNF189, ZNF528, ZNF543, ZNF140, ZNF610, ZNF350, ZNF8, ZNF30, ZNF98, ZNF677, ZNF596, ZNF214, ZNF37A, ZNF34, ZNF250, ZNF547, ZNF273, ZFP82, ZNF224, ZNF33A, ZNF45, ZNF175, ZNF184, ZFP28-1, ZFP28-2, ZNF18, ZNF213, ZNF394, ZFP1, ZFP14, ZNF416, ZNF557, ZNF729, ZNF254, ZNF764, ZNF785, ZNF10, CBX5, RYBP, YAF2, MGA, CBX1, SCMH1, MPP8, SUMO3, HERC2, BIN1, PCGF2, TOX, FOXA1, FOXA2, IRF2BP1, IRF2BP2, IRF2BPL IRF-2BP1_2 N-terminal

domain, HOXA13, HOXB13, HOXC13, HOXA11, HOXC11, HOXC10, HOXA10, HOXB9, HOXA9, ZFP28, ZN334, ZN568, ZN37A, ZN181, ZN510, ZN862, ZN140, ZN208, ZN248, ZN571, ZN699, ZN726, ZIK1, ZNF2, Z705F, ZNF14, ZN471, ZN624, ZNF84, ZNF7, ZN891, ZN337, Z705G, ZN529, ZN729, ZN419, Z705A, ZN302, ZN486, ZN621, ZN688, ZN33A, ZN554, ZN878, ZN772, ZN224, ZN184, ZN544, ZNF57, ZN283, ZN549, ZN211, ZN615, ZN253, ZN226, ZN730, Z585A, ZN732, ZN681, ZN667, ZN649, ZN470, ZN484, ZN431, ZN382, ZN254, ZN124, ZN607, ZN317, ZN620, ZN141, ZN584, ZN540, ZN75D, ZN555, ZN658, ZN684, RBAK, ZN829, ZN582, ZN112, ZN716, HKR1, ZN350, ZN480, ZN416, ZNF92, ZN100, ZN736, ZNF74, ZN443, ZN195, ZN530, ZN782, ZN791, ZN331, Z354C, ZN157, ZN727, ZN550, ZN793, ZN235, ZN724, ZN573, ZN577, ZN789, ZN718, ZN300, ZN383, ZN429, ZN677, ZN850, ZN454, ZN257, ZN264, ZN485, ZN737, ZNF44, ZN596, ZN565, ZN543, ZFP69, SUMO1, ZNF12, ZN169, ZN433, ZN175, ZN347, ZNF25, ZN519, Z585B, ZN517, ZN846, ZN230, ZNF66, ZN713, ZN816, ZN426, ZN674, ZN627, ZNF20, Z587B, ZN316, ZN233, ZN611, ZN556, ZN234, ZN560, ZNF77, ZN682, ZN614, ZN785, ZN445, ZFP30, ZN225, ZN551, ZN610, ZN528, ZN284, ZN418, ZN490, ZN805, Z780B, ZN763, ZN285, ZNF85, ZN223, ZNF90, ZN557, ZN425, ZN229, ZN606, ZN155, ZN222, ZN442, ZNF91, ZN135, ZN778, ZN534, ZN586, ZN567, ZN440, ZN583, ZN441, ZNF43, ZN589, ZN563, ZN561, ZN136, ZN630, ZN527, ZN333, Z324B, ZN786, ZN709, ZN792, ZN599, ZN613, ZF69B, ZN799, ZN569, ZN564, ZN546, ZFP92, ZN723, ZN439, ZFP57, ZNF19, ZN404, ZN274, CBX3, ZN250, ZN570, ZN675, ZN695, ZN548, ZN132, ZN738, ZN420, ZN626, ZN559, ZN460, ZN268, ZN304, ZN605, ZN844, SUMO5, ZN101, ZN783, ZN417, ZN182, ZN823, ZN177, ZN197, ZN717, ZN669, ZN256, ZN251, CBX4, CDY2, CDYL2, ZN562, ZN461, Z324A, ZN766, ID2, ZN214, CBX7, ID1, CREM, SCX, ASCL1, ZN764, SCML2, TWST1, CREB1, TERF1, ID3, CBX8, GSX1, NKX22, ATF1, TWST2, ZNF17, TOX3, TOX4, ZMYM3, I2BP1, RHXF1, SSX2, I2BPL, ZN680, TRI68, HXA13, PHC3, TCF24, HXB13, HEY1, PHC2, ZNF81, FIGLA, SAM11, KMT2B, HEY2, JDP2, HXC13, ASCL4, HHEX, GSX2, ETV7, ASCL3, PHC1, OTP, I2BP2, VGLL2, HXA11, PDLI4, ASCL2, CDX4, ZN860, LMBL4, PDIP3, NKX25, CEBPB, ISL1, CDX2, PROP1, SIN3B, SMBT1, HXC11, HXC10, PRS6A, VSX1, NKX23, MTG16, HMX3, HMX1, KIF22, CSTF2, CEBPE, DLX2, PPARG, PRIC1, UNC4, BARX2, ALX3, TCF15, TERA,

VSX2, HXD12, CDX1, TCF23, ALX1, HXA10, RX, CXXC5, SCML1, NFIL3, DLX6, MTG8, CEBPD, SEC13, FIP1, ALX4, LHX3, PRIC2, MAGI3, NELL1, PRRX1, MTG8R, RAX2, DLX3, DLX1, NKX26, NAB1, SAMD7, PITX3, WDR5, MEOX2, NAB2, DHX8, CBX6, EMX2, CPSF6, HXC12, KDM4B, LMBL3, PHX2A, EMX1, NC2B, DLX4, SRY, ZN777, ZN398, GATA3, BSH, SF3B4, TEAD1, TEAD3, RGAP1, PHF1, GATA2, FOXO3, ZN212, IRX4, ZBED6, LHX4, SIN3A, RBBP7, NKX61, R51A1, MB3L1, DLX5, NOTC1, TERF2, ZN282, RGS12, ZN840, SPI2B, PAX7, NKX62, ASXL2, FOXO1, GATA1, ZMYM5, LRP1, MIXL1, SGT1, LMCD1, CEBPA, SOX14, WTIP, PRP19, NKX11, RBBP4, DMRT2, SMCA2, 以及其功能活性片段。

在一些更具体的实施方案中,所述 DNA 甲基化结构域包含至少一种 DNA 甲基转移酶或其功能活性片段。例如,所述 DNA 甲基转移酶选自 DNMT3A、DNMT3B、DNMT3c、DNMT1、DNMT2 和 DNMT3L。例如,所述 DNA 甲基化结构域包含至少一个 DNMT3A 和至少一个 DNMT3L。例如,所述至少一个 DNMT3A 和所述至少一个 DNMT3L 可互换顺序地连接。例如,所述 DNA 甲基化结构域包含一个 DNMT3A 和一个 DNMT3L,且它们可互换顺序地连接。例如,所述 DNA 甲基转移酶包含 SEQ ID NO: 19-24 任一项所示的氨基酸序列。

本申请复合物的第一融合物和第二融合物是通过各自包含的招募结构域间的相互作用进而形成聚集的复合物。因此,本申请提供了非限制性的招募结构域 A 和招募结构域 A'的组合示例:(1)所述招募结构域 A 和所述招募结构域 A'其中之一的结构域为 GCN4,并且其中另一个结构域为 scFv;或(2)所述招募结构域 A 和所述招募结构域 A'其中之一的结构域为 GFP11 片段,并且其中另一个结构域为 GFP1-10;或(3)所述招募结构域 A 和所述招募结构域 A'其中之一的结构域为 GVKESLV,并且其中另一个结构域为 PDZ 蛋白结构域。同理于 GFP11 和 GFP1-10 分别源自分裂 GFP (SEQ ID NO: 15) 形成所述招募结构域 A 和所述招募结构域 A'的情形同样可适用于其他类别的荧光蛋白,例如 mCherry (SEQ ID NO: 16)、eYFP (SEQ ID NO: 18)、eCFP (SEQ ID NO: 17)等,即可通过分裂 mCherry、分裂 eYFP、或分裂 eCFP 分别获取不同组的招募结构域 A 和招募结构域 A'用于本申请提供的复合物中。在一些实施方案中,本申请复合物的所述第一融合物和所述第二融合物其中之一可包含两个或两个以上的招募结构域,且它们是通过接头序列连接的。示例性的招募结构域的氨基酸序列可包含如 SEQ ID NO: 10-14 中任一项所示。

综合以上情形,本申请可提供下列所示的第一融合物或第二融合物的氨基酸序列:

SEQ	结构域信	氨基酸序列 (dCas 或 TALE: 粗体, DNMT: <i>斜体斜体</i> , 转录阻遏物: <下划线粗体>; 招募结构域:
-----	------	--

ID	息	下划线斜体)
NO:		
51	DNMT3A (CD)- DNMT3L (CD)- dCas9- 10xGCN4	<i>MNHDQEFDPKVPVPAEKRPVLSLFDGIATGLLVKDLGIQVDRYASEVCEDSITVGMVR</i> <i>HQGKIMYVGDVRSVTQKHQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDA</i> <i>RPKEGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARYFWGNLPGMNRP</i> <i>LASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFG</i> <i>FPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSESSGLVPLSL</i> <i>RGSHMGPMIYKTVSAWKRPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRD</i> <i>VEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQ</i> <i>ETTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKV</i> <i>DLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPS</i> <i>EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKRKMVMDKKYSIGLAIGTNSVGWAVIT</i> <i>DEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE</i> <i>IFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD</i> <i>KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAK</i> <i>AILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDITYD</i> <i>DDLNLQAIGDQYADFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLL</i> <i>KALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED</i> <i>LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNS</i> <i>RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVY</i> <i>NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISG</i> <i>VEDRFNASLGTYHDLLKHKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDD</i> <i>KVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKE</i> <i>DIQKAQVSGQDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMAREN</i> <i>QTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE</i> <i>LDINRLSDYDVDAIVPQSFLKDDSIDNKVLTNSDKNRGKSDNVPSEEVVKKMKNYWRQLLN</i> <i>AKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI</i> <i>REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYG</i> <i>DYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV</i> <i>WDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG</i>

		FDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAP AAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDYPYDVPDYASLGSGSPK KKRKVEDPKKKRKVDGIGSGSNGSSG<u>EELLSKNYHLENEVARLKK</u>SGSGSGSGSGSGSGSGSGSGSG GGSGSGEELLSKNYHLENEVARLKKSGSGSGSGSGSGSGSGSGSGSGSGSGEELLSKNYHLENEVA <u>RLKK</u>SGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGEELLSKNYHLENEVARLKKSGSGSGSGSGSGSGSG GGSGSGSGSGSGEELLSKNYHLENEVARLKKSGEELLSKNYHL <u>ENEVARLKK</u>SGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGEELLSKNYHLENEVARLKKSGSGSGSGSGSGSGSG GSGSGSGSGSGSGSGEELLSKNYHLENEVARLKKSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGEELLSK <u>NYHLENEVARLKK</u>SGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGEELLSKNYHLENEVARLKKSGSGSGQRG SKRTADGSEFESPKKKRKV
52	scFv- KRAB	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LES GGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWQGTLVTVSSYPYDVPDYAGGGGSGGGGSGGGGSG</u> GGGGSSLDPGGGGSG-<<u>RTLVTFKDVFVDFTREEWKLLDTAQOIVYRNVMLENYKNLVSLGY</u> <u>QLTKPDVILRLEKGEEP</u>>GSGKRTADGSEFESPKKKRKV
53	scFv- Zim3	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LES GGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWQGTLVTVSSYPYDVPDYAGGGGSGGGGSGGGGSG</u> GGGGSSLDPGGGGSG-<<u>VT FEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLVSVGQGET</u> <u>TKPDVILRLEQGKEPWL</u>>GSGKRTADGSEFESPKKKRKV
54	scFv- Znf680	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LES GGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWQGTLVTVSSYPYDVPDYAGGGGSGGGGSGGGGSG</u> GGGGSSLDPGGGGSG-<<u>LTFRDVAIEFSLEEWQCLDTAQRNLYRKVMFENYRNLVFLGIAVSK</u> <u>PHLITCLEQGKEPWN</u>>GSGKRTADGSEFESPKKKRKV

55	scFv- Znf554	<u>MGPDIIVMTQSPSSLSASVGDRTTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>VTFEDVSMDFSQEEWELLEPAQKNLYREVMLENYRNVVSLEALKN</u> <u>QCTDVGIKEGPLSPAQT</u>>GSGKRTADGSEFESPKKKRKV
56	scFv- Znf264	<u>MGPDIIVMTQSPSSLSASVGDRTTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>VTFDDVAVTFTKEEWGQLDLAORTLYQEVMLENCGLLVSLGCPVP</u> <u>KAELICHLEHGQEPWT</u>>GSGKRTADGSEFESPKKKRKV
57	scFv- Znf582	<u>MGPDIIVMTQSPSSLSASVGDRTTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>ELFRDVAIVFSQEEWQWLAPAQRDLRYRDMLETYSNLVSLGLAVS</u> <u>KPDVISFLEQGKEPWM</u>>GSGKRTADGSEFESPKKKRKV
58	scFv- Znf324	<u>MGPDIIVMTQSPSSLSASVGDRTTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>MAFEDVAVYFSQEEWGLLDTAQRALYRRVMLDNFALVASLGLSTS</u> <u>RPRVVIQLERGEEPWV</u>>GSGKRTADGSEFESPKKKRKV
59	scFv- Znf669	<u>MGPDIIVMTQSPSSLSASVGDRTTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>VAFEDVAVNFTQEEWALLDSSQKNLYREVMQETCRNLASVGSQWK</u> <u>DQNIEDHFEKPGKDIRN</u>>GSGKRTADGSEFESPKKKRKV

60	scFv- Znf354a	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQTLITVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG-<<u>LT FEDVAVLFTRDEWRKLAPSQRNLYRDVMLENYRNLVSLGLPFT</u> <u>KPKVISLLQOGEDPWE</u>>GSGKRTADGSEFESPKKKRKV
61	scFv- Znf82	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQTLITVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG-<<u>VMFSDVSIDFSPEEWEYLDLEQKDLYRDVMLENYSNLVSLGCFISKP</u> <u>DVISSLEQGKEPWK</u>>GSGKRTADGSEFESPKKKRKV
62	scFv- Znf595	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQTLITVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG-<<u>VTFRDVAIEFSPEEWKCLDPAQQNLYRDVMLENYSNLVSLGFVISN</u> <u>PDLVTCLEQIKEPCN</u>>GSGKRTADGSEFESPKKKRKV
63	scFv- Znf419	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQTLITVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG-<<u>VT FEDVAVYFSQEEWRLLDDAQRLLYRNVMLENFTLLASLGLASSK</u> <u>THEITQLESWEEPFM</u>>GSGKRTADGSEFESPKKKRKV
64	scFv- Znf566	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQTLITVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG-<<u>VMFSDVSVDQSQEEWECLNDDQORDLYRDVMLENYSNLVSMGHSIS</u> <u>KPNVISYLEQGKEPWL</u>>GSGKRTADGSEFESPKKKRKV

65	scFv- Zim2	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLQPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGGSEVKL</u> <u>LES GGGLVQPGGSLKLSCAVSGFSLTDYGVNWVROAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> <u>GGGGSSLDPGGGGSG<VTFEDVLVDFSPEELSSLSAAQRNL YREVMLENYRNLVSLGHQFSKP</u> <u>DIISRLEEEESYAM></u> GSGKRTADGSEFESPKKKRKV
66	scFv- EHMT2(G9A)	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLQPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGGSEVKL</u> <u>LES GGGLVQPGGSLKLSCAVSGFSLTDYGVNWVROAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> <u>GGGGSSLDPGGGGSG<TGSAAIAEVLLNARCDLHAVNYHGDTP LHIAARES YHDCVLLFLSR</u> <u>GANPELRNKEGDTAWDLTPERSDVWFALQLNRKLRLGVGNRAIRTEKIIICRDVARGYENVP</u> <u>IPC VNGVDGEPCEPDYKYISENCETSTMNIDRNITHLOHCTCVDDCSSNCLCGQLSIRCWYD</u> <u>KDGRLLQEFNKIEPLIFE CNQACSWRNCKNRVYVQSGIKVRLQLYRTAKMGWGVRA LQT</u> <u>IPOGT FICEYVGELISDAEADVREDDSYLFDLDNKDGEVYCIDARYYGNISRFINHLCDPN IIP</u> <u>VRVFM LHQDLRFPRIAFFSSRDIRTGEELGFDYDGRFWDIKSKYFTCQC GSEKCKHSAEAIA</u> <u>LEQSRLARLDPHPELLPELGSLPPVNT></u> GSGKRTADGSEFESPKKKRKV
67	scFv- SUV39H1	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLQPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGGSEVKL</u> <u>LES GGGLVQPGGSLKLSCAVSGFSLTDYGVNWVROAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> <u>GGGGSSLDPGGGGSG<AENLKGSVCCKSSWNQLODL CRLAKLSCPALGISKRNL YDFEVEY</u> <u>LCDYKKIREQEYYLVKWRGYPDSESTWEPRQNLKCVRILKQFHKDLERELLRRHHRSKTP</u> <u>RHLDPSLANYLVQKAKQRRALRRWEQELNAKRSHLGRITVENEVDLDGPPRAFVYINEYRV</u> <u>GEGITL NQVAVGCECQDCLWAPTGGCCPGASLHKFAYNDQGQVRLRAGLPIYECNSRCRC</u> <u>GYDCPNRVVQKGIRYDLCIFRTDDGRGWGVRTLEKIRKNSFVMEYVGEIITSEEAEERRGOIY</u> <u>DRQGATYLFDLDYVEDVYTVDAAYYGNISHFVNHSCDPNLQVYNVFIDNLDERLPRIAFFAT</u> <u>RTIRAGEELTFDYNMQVDPVDMESTRMDSNFGLAGLPGSPKKRVRIECKCGTESCRKYLE></u> GSGKRTADGSEFESPKKKRKV
68	scFv-	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u>

	ZFPM1(F OG)	<u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG< <u>SRRKQSNPRQIKRSLGDMEAREEVQLVGASHMEQKATAPEAPSPPS</u> <u>A</u> >GSGKRTADGSEFESPKKKRKV
69	scFv- TRIM28(KAP1)	<u>MGPDIIVMTQSPSSLASVGDRTITCRSSTGAVTTSNYASWVQKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG< <u>AASAAAASAAAASASGSPGPEGSAGGEKRSTAPSAAAASASASAAA</u> <u>SSPAGGGAEALELLEHCGVCRERLRPEREPRLLPCLHSACSACLGPAAPAAANSSGDGGAA</u> <u>GDGTVVDCPVCKQQCFKSDIVENYFMRDSGSKAATDAQDANQCCTSCEDNAPATSYCVECS</u> <u>EPLCETCYEAHQRVKYTKDHTVRSTGPAKSRDGERTVYCNVHKHEPLVLFCECDTLTCRD</u> <u>COLNAHKDHQYQFLEDAVRNQRKLLASLVKRLGDKHATLQKSTKEVRSSIRQVSDVQKRV</u> <u>QVDVKMAILQIMKELNKRGRVLVNDQKVTEGQOERLERQHWTMTKIQKHQEHILRFAS</u> <u>WALESDNNTALLSKKLIYFQLHRALKMIVDPVEPHGEMKFQWDLNAWTKSAEAFGKIVA</u> <u>ERPGTNSTGPAPMAPPRAPGPLSKQSGSSQPMEOEGYGFSGDDPYSSAEPHVSQVGRSR</u> <u>SGEGEVSGLMRKVPRVSLERLDDLTADSQPPVFKVFPGSTTEDYNLIVIERGAAAAATGQP</u> <u>GTAPAGTPGAPPLAGMAIVKEEETEAAGAPPTATEGPETKPVLMALAEGPGAEGPRLASPS</u> <u>GSTSSGLEVVAPEGTSAPGGGPGTLDDSATICRVCKPGDLVMCNQCEFCFHLDCHLPALQ</u> <u>DVPGEWWSLCHVLPDLKEEDGSLSLDGADSTGVVAKLSPANQRKCERVLLALFCHEPCR</u> <u>PLHQLATDSTFSLDQPGGTLDLTLIRARLOEKLSPPYSSPQEFADVGRMEFKQFNKLTEDKA</u> <u>DVQSIIGLQRFETRMNEAFGDTKFSVLYVEPPMSLPGAGLSSQELSGGPGDGP</u> >GSGKRTA DGSEFESPKKKRKV
70	scFv- EZH2	<u>MGPDIIVMTQSPSSLASVGDRTITCRSSTGAVTTSNYASWVQKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG< <u>GQTGKKSEKGPVCWRKRVKSEYMRLRQLKRFRADEVKSMFSSN</u> <u>RQKILERTEILNQEWKQORRIOPVHILTSVSSLRGTRCSVTSDLDFPTQVIPLKTLNAVASVPI</u>

		<u>MYSWSPLQONFMVEDETVLHNIPYMGDEVLDQDGTFFIELIKNYDGKVHGDRECGFINDEI</u> <u>FVELYNALGQYNDDDDDDDDGDDPEEREKQKQDLEDHRDDKESRPPRKFPSDKIFEAISSMFP</u> <u>DKGTAHEELKEKYKELTEQQLPGALPPECTPNIDGPNKSVQREOSLHSFHTLFCRRCFKYD</u> <u>CFLHPFHATPNTYKRKNTETALDNKPCGPQCYQHLEGAKEFAAALTAERIKTPPKRPGGRR</u> <u>RGRLPNSSSRPSTPTINVLESKDTSDREAGTETGGENNDKEEEEKKDETSSSSEANSRCQTP</u> <u>IKMKPNIEPPENVEWSGAESMFRVLIGTYDNFCAIARLIGTKTCRQVYEFVKESSIIAPA</u> <u>PAEDVDTPPRKKKRKHRLWAAHCRKIQLKKGSSNHVYNYQPCDHPRQPCDSSCPCVIAQ</u> <u>NFCEKFCQCSSECONRFPGCRCKAQCNTKQCPCYLAVRECDPDLCLTCGAADHWDSKNVS</u> <u>CKNCISIQRGSKKHLLLAPSDVAGWGIFIKDPVQKNEFISEYCGEHSQDEADRRGKVYDKYM</u> <u>CSFLENLNNDFVVDATRKGNKIRFANHSVNPNCYAKVMMVNGDHRIGIFAKRAIQTGEELF</u> <u>FDYRYSQADALKYVGIEREMEIP</u>>GSGKRTADGSEFESPKKKRKV
71	scFv- MXD1(SI D FL)	<u>MGPDIIVMTQSPSSLASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>AAAVRMNIQMLLEAADYLERREREAEHGYASMLPYNNKDRDALK</u> <u>RRNKSCKNNSSSRSTHNEKNNRAHLRLCLEKLKGLVPLGPESRHTTSLSLTKAKLHIK</u> <u>KLEDCDRKAVHQIDQLQREQRHLKRQLEKLGIERIRMDSIGSTVSSERSDSREIDVDVESTD</u> <u>YLTGDLWSSSSVSDSDERGSMQSLGSDEGYSSTSIKRIKLQDSHKACLGL</u>>GSGKRTADGSE FESPKKKRKV
72	scFv-SID	<u>MGPDIIVMTQSPSSLASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>MNIQMLLEAADYLERREREAEHGYASMLP</u>>GSGKRTADGSEFESPK KRKV
73	scFv- LSD1(KD MA1)	<u>MGPDIIVMTQSPSSLASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGSGGGGSGGGGSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGQGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u>

		GGGGSSLDPGGGGSG<<u>LSGKKAIAAAAAAAAAATGTEAGPGTAGGSENGSEVAAQAGLSG</u> <u>PAEVGPGAVGERTPRKKEPPRASPPGGLAEPPGSAGPQAGPTTVVPGSATPMETGIAETPEGR</u> <u>RTSRRKRAKVEYREMDESLANLSEDEYYYSEEERNAKAEKEKKLPPPPQAPPEEENESEPEE</u> <u>PSGQAGGLQDDSSGGYGDOGASGVEGAAFQSRLPHDRMTSQEAACFPDIISGPQQTQKVFL</u> <u>FIRNRTLQWLDPNPKIQLTFEATLQOLEAPYNSDTVLVHRVHSYLERHGLINFGIYKRIKPLP</u> <u>TKKTGKVIIGSGVSGLAARQLQSFQMDVTLLLEARDRVGGRVATFRKGNVYADLGAMVV</u> <u>TGLGGNPMNAVSKQVNMELAKIKQKCPLYEANGQADTVKVPKEKDEMVEQEFNRLLEAT</u> <u>SYLSHQLDFNVLNKPVSLGOALEVVIQLOEKHKVDEQIEHWKKIVKTQEELKELLNKMV</u> <u>NLKEKIKELHQOYKEASEVKPPRDITAEFLYKSKHRDLTALCKEYDELAETQGKLEEKLOE</u> <u>LEANPPSDVYLSSRDROILDWHFANLEFANATPLSTLSLKHWDQDDDFEFTGSHLTVRNGYS</u> <u>CVPVALAEGLDIKLNTAVRQVRYTASGCEVIAVNTRSTSQTFFIYKCDAYLCTLPLGLVKQQP</u> <u>PAVQFVPPLPEWKTSVAVQRMGFGNLNKVVLCFDRVFWDPVNLFGHVGSTTASRGELFLF</u> <u>WNLYKAPILLALVAGEAAGIMENISDDVIVGRCLAAILKGIFGSSAVPOPKETVVSRRWADPW</u> <u>ARGSYSYVAAGSSGNDYDLMAQIPITPGPSIPGAPQPIPRLFFAGEHTIRNYPATVHGALLSGL</u> <u>REAGRIADQFLGAMYTLPROATPGVPAQQSPSM>GSGKRTADGSEFESPKKKRKV</u>
74	scFv- HP1a long	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGGSGGGGSGGGGSEVKL</u> <u>LES GGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGOGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>GKKTARTADSSSEDEEEYVVEKVLDRRVVKGOVEYLLKWKGFSE</u> <u>EHNTWEPEKNLDCPELISEFMKKYKKMKEGENNKPREKSES NKRSNFSNSADDIKSKKKR</u> <u>EQSNDIARGFERGLEPEKIIIGATDSCGDLMLMKWKDTDEADLVLAKEANVKCPQIVIAFYE</u> <u>ERLTWHAYPEDAENKEKETAKS>GSGKRTADGSEFESPKKKRKV</u>
75	scFv- HP1a short	<u>MGPDIIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRGGGGSGGGGSGGGGSEVKL</u> <u>LES GGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCVTGLFDYWGOGLTVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> GGGGSSLDPGGGGSG<<u>MKEGENNKPREKSES NKRSNFSNSADDIKSKKKREQSNDIARGFE</u> <u>RGLEPEKIIIGATDSCGDLMLMKWKDTDEADLVLAKEANVKCPQIVIAFYEEERLTWHAYPE</u> <u>DAENKEKETAKS>GSGKRTADGSEFESPKKKRKV</u>

76	scFv- HDAC3	<u>MGPDIVMTQSPSSLSASVGDRTITCRSSSTGAVTTSNYASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSG</u> <u>SLIGDKATLTISSLOPEDFATYFCALWYSNHWVFGQGTKVELKRRGGGSGGGGSGGGGSSGGGSEVKL</u> <u>LESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYNSALKDRFIISKDN</u> <u>GKNTVYLQMSKVRSDDTALYYCITGLFDYWGGQTLVTVSSYPYDVPDYAGGGGSGGGGSGGGGS</u> <u>GGGGSSLDPGGGGSG<AKTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMI</u> <u>VFKPYQASQHDMCRFHSEDIYDFLQRVSPTNMQGFTKSLNAFNVGDDCPVFPGLFEFCSR</u> <u>TGASLQATQLNNKICDIANWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDID</u> <u>IHHGDGVQEA FYLTDRVMTVSFHKYGN YFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQ</u> <u>SYKHLFQPVINQVVDFYOPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVL</u> <u>GGGGYTVRNVARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYL</u> <u>DOIRQTIFENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEEERGPEENYSRPEAPNEFYDGD</u> <u>HDNDKESDVEI</u>>GSGKRTADGSEFESPKKKRKV
77	3xGCN4- dCas9- 3xGCN4	MPAAKRVKLDGSE<u>EELLSKNYHLENEVARLKK</u>GSGSGSGSGSGSGSGSGSGSGSGSGSGGE<u>EELLSKNY</u> <u>HLENEVARLKK</u>GSGSGSGSGSGSGSGSGSGSGSGSGSGGE<u>EELLSKNYHLENEVARLKK</u>GSGSGSGSG GSGSGSGSGSGSGSGSGPKKKRKVMDDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNT DRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR LEESFLVEEDKKHERHPFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLLIYLALAHMI KFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLNL IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYA DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYFLKDNREKIEKILTRIPYYVGPLARGNSRFAMTRKSEETITP WNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTRNKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL LKHKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVP QSFLKDDSDKNVLTTRSDKNRGKSDNPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAE RGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDF

		RKDFQFYKREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSE QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKV EKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKR MLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPENEQQQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT STKEVLDATLIHQSITGLYETRIDLSQLGGDAYPYDVDPDYASLGSGSPKKRKVEDPKKKRKV DGSGETPGTSGSGSGSGSGSGSGSGEELLSKNYHLENEVARLKGSGSGSGSGSGSGSGSGS SGSGSGSGEELLSKNYHLENEVARLKGSGSGSGSGSGSGSGSGSGSGSGSGSGSGEELLSKNYHLENE VARLKGSGSGQRSGSKRTADGSEFESPKKKRKV*
78	scFv- KRAB	MPAAKRVKLDGSGPDIVMTQSPSSLSASVGDRVTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTN NRAPGVPSRFSGLIGDKATLTISSLOPEDFATYFCALWYSNHWWFGQGTKVELKRGGGGSGGGSGGG GGSGGGSGSEVKLLESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN SALKDRFIISKDNGKNTVYLQMSKVRSDDTALYYCVTGLFDYWGOGLTVTVSSYPYDVDPDYAGGGGGS GGGGSGGGSGGGGSLDPGGGGSG<RTLVTFKDVFVDFTREEWKLLDTAQOIVYRNVMLEN YKNLVSLGYQLTKPDVILRLEKGEEP>*
79	DNMMT 3A- DNMT3L -dCas9- 1xGCN4	MPAAKRVKLDGSGNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLSLGFLESSGSGGGGTLK YVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWI FMDNLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQA QVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSES ATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKVMDKKYSIGL AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP TIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQL FEENPINASGVDAKILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSGLTPNFKSNFDLA

		EDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILLS DILRVNTEITKAPLSASMI KRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTF RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGTYHDLLKHKDKDFLDNEENEDILEDIVLTLTLFEDREM IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGR HKPENIVIAMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLT RSDKNRGKSDNPSEEVV KKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS RMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALI KKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIAR KKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKPNIDFLEA KGYKEVKKDLIHKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDAYP YDVPDYASLGS GSPKKRKRKVEDPKKKR KVDGIGSGSNGSSG<u>EELLSKNYHLENEVARLKK</u>SGSGSG SGSGSGSGSGSGSGSGSGSGEELLSKNYHSGSGSGQRGSKRTADGSEFESPKKKRKV*
80	DNMT3A - DNMT3L -dCas9- 2xGCN4	MPAAKRVKLDGSNH<u>DQEFDP</u>PKVYPPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYASE VCEDSITVGMVRHQGKIMYVG D VRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRPVRVLSLFRNIDKVLKSLGFLES GSGSGGGTLK YVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRIQLYALPRQESQRPF FWI FMDNLLTEDDQETTTRFLQTEAVTLQDVGRDYNAMRVWSNIPGLKSKHAPLTPKEEEYLQA QVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSES ATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKVMDDKKYSIGL

		AIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR YTRRNKRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP TIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQL FEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLA EDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMI KRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTF RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREM IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGR HKPENIVIMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTNRSDKNRGKSDNVPSEEVV KKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDS RMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGITALI KKYPKLESEFVYGDYKVYDVRKMKIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR PLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIAR KKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEA KGYPEVKKDLIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDAYP YDVPDYASLGS GSPKKRKRVEDPKKKRKVDGIGSGSNGSSGE<u>ELLSKNYHLENEVARLKK</u>SGSGG SGSGSGSGSGSGSGSGSGGE<u>ELLSKNYHLENEVARLKK</u>SGSGSGQSGSKRTADGSEFESPKKKRK V*
81	DNMT3A - DNMT3L -dCas9- 4xGCN4	MPAAKRVKLDGSNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHQIEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSR

		SGSGQRGSKRTADGSEFESPKKKRKV*
82	DNMT3A	MPAAKRVKLDGNSHIDQEFDPKVPYPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASE
	-	VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEQWGPFDLVIGGSPCNDLSIVNPARKGLYEETGR
	DNMT3L	LFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARY
	-dCas9-	FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKGDQHFPVFMNEKEDI
	10xGCN4	LWCTEMERVFGFPVHYTDVSNMSRLARQRL LGRWSVPVIRHLFAPLKEYFACVSSGNSNANSR
		GPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRPVVRVLSLFRNIDKV LKSLGFLES GSGSGGGTLK
		YVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPF FWI
		FMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQA
		QVRSRSKLDAPKVDLLVKNCLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSES
		ATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEP KKKRKVMDDKKYSIGL
		AIGTNSVGWAVITDEYKVP SKKFVKLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
		YTRRKNRICYLQEIFS NEMAKVD D SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
		TIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQL
		FEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNLGFGNLIALSLGLTPNFKSNFDLA
		EDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMI
		KRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
		DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTF
		RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKV
		LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED
		YFKKIECFDSVEISGVEDRFNASLGT YHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREM
		IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
		FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA I KKGILQTVKVVDLVKVMGR
		HKPENIVIAMARENQTTQKGQKNSRERMKRIE EG I KELGSQILKEHPVENTQLQNEKLYLY
		YLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVV
		KKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS
		RMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA LI
		KKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
		PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIAR
		KKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSFEKNPIDFLEA

[illegible]

		KFRGHFLIEGDLNPDNSDVDFKFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENL IAQLPGEKKNGLFGNLIASLGLTPNFKSNDLAEDAKLQLSKDTYDDDLNLLAQIGDQYA DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYVVGPLARGNSRFAWMTRKSEETITP WNFEVVDKGGASQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVP QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDF RKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDKYKYDVRKMIKSE QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKV EKGKSKKLKSVKELLGITIMERSFFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKR MLASAGELQKGNELALPSKYVNFLLASHYEKLKGSPEDEQKQLFVEQHKHYLDEHEQIS EFSKRVLADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT STKEVLDTLIHQSTGLYETRIDLSQLGGDAYPYDVPDYASLGGSGPKKKRKVEDPKKKRKV DGGSGSETPGTSESATPES<<u>RTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG</u> <u>YQLTKPDVILRLEKGEEP</u>>*
85	GFP1-10- KRAB	MPAAKRVKLDGSS<u>SKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATIGKLTLCFICTTGKLPVP</u> <u>WPTLVTTLTYGVCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIE</u> <u>LKGTDFKEDGNILGHKLEYNFNSHNVIYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDGP</u> <u>VLLPDNHYLSTQTVLSKDPNEKLDPGGGGSG<RTLVTFKDVFVDFTREEWKLLDTAQQIVYRNV</u> <u>MLENYKNLVSLGYQLTKPDVILRLEKGEEP</u>>*
86	DNMT3A - DNMT3L	MPAAKRVKLDGNS<u>HDQEFDPKVPYPPVPAEKRRPIRVLSLFDGIATGLLVKDLGIQVDRYASE</u> <u>VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR</u> <u>LFFEFYRLLHDARPKEGDDRPFFWL FENYVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARY</u>

-dCas9-	<i>FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIQGKDQHFPVFMNEKEDI</i>
1xGFP11	<i>LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSPVIRHLFAPLKEYFACVSSGNSNANSR</i> <i>GPSFSSGLVPLSLRGSHMGPMIYKTVSAWKRPVRVLSLFRNIDKVLKSLGFLESGSGGGGTLK</i> <i>YVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWI</i> <i>FMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQA</i> <i>QVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSES</i> <i>ATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKRKRKVMDDKKYSIGL</i> <i>AIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR</i> <i>YTRRNKRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP</i> <i>TIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQ</i> <i>FEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLA</i> <i>EDAKLQLSKDQYDDDLNLLAQIGDQYADFLAAKNLSDAILSDILRVNTEITKAPLSASMI</i> <i>KRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM</i> <i>DGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTF</i> <i>RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASQSFIERMTNFDKNLPNEKV</i> <i>LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED</i> <i>YFKKIECFDSVEISGVEDRFNASLGTYHDLKKIHKDKDFLDNEENEDILEDIVLTLTLFEDREM</i> <i>IEERLKYAHLFDDKVMKQLKRRRYTGWRLSRKLINGIRDKQSGKTILDFLKSDFANRN</i> <i>FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDLVKVMGR</i> <i>HKPENIVIMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLY</i> <i>YLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVV</i> <i>KKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS</i> <i>RMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGITALI</i> <i>KKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR</i> <i>PLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKTEVQTGGFSKESILPKRNSDKLIAR</i> <i>KKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEA</i> <i>KGYKEVKKDLIHKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKL</i> <i>KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRRDKPIREQAE</i> <i>NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQISITGLYETRIDLSQLGGDAYP</i> <i>YDVPDYASLGSGSPKKKRKVEDPKKKRKVDGSGSETPGTSESATPESGSGSNGSSGSMRDHMLH</i>

		<i>EYVNAAGITGGSGGAAARSGSGQRGSKRTADGSEFESPKKKRKV*</i>
87	DNMT3A	<i>MPAAKRVKLDGSHDQEFDPKVPYPPVPAEKRPPIRVLSLFDGIATGLLVKDLGIQVDRYIASE</i>
	-	<i>VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEQWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR</i>
	DNMT3L	<i>LFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARY</i>
	-dCas9-	<i>FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIQGKDQHFPVFMNEKEDI</i>
	7xGFP11	<i>LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRWSVPVIRHLFAPLKEYFACVSSGNSNANSR</i> <i>GPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRPVVRVLSLFRNIDKVLSLGFLES GSGSGGGTLK</i> <i>YVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPF FWI</i> <i>FMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQA</i> <i>QVRSRSKLDAPKVDLLVKNCLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSES</i> <i>ATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEP KKKRKVMDDKKYSIGL</i> <i>AIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR</i> <i>YTRRKNRICYLQEIFS NEMAKVD D SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP</i> <i>TIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQL</i> <i>FEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALLSLGLTPNFKSNFDLA</i> <i>EDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMI</i> <i>KRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM</i> <i>DGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTF</i> <i>RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKV</i> <i>LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED</i> <i>YFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREM</i> <i>IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN</i> <i>FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDLVKVMGR</i> <i>HKPENIVIAMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLY</i> <i>YLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVV</i> <i>KKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS</i> <i>RMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA LI</i> <i>KKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR</i> <i>PLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIAR</i> <i>KKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERS SFEKNPIDFLEA</i>

		KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDAYP YDVPDYASLGSGSPKKKRKVEDPKKKRKVDGSGSETPGTSESATPESGSGSNGSSGSMRDHMLH EYVNAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYV NAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNA GITAAARSGSGQRGSKRTADGSEFESPKKKRKV*
88	DNMT3A - hDNMT3 L-dCas9- 1xGCN4	MPAAKRVKLDGSNH DQEFDP PKVYPPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKH IQEWGPF DLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISR FLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPF DLVYGATPPLGHTCDRPPSWYLFQFHRL LQYARPKPGSPRPFFWMFVDNL VLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRHWALVSEEELSLLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEP KKKRKVMDKKYSIGLAIGTNSVGWA VITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICY LQEIFS NEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRREN LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTRIPYYVGPLAR GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVL PKHSLLYEYF TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVE ISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLF DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNF MQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVD ELVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVD

		QELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFV YGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG GFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDL IHKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQK QLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDAYPYDVPDYASLGS GSPKKKRKVEDPKKKRKVDGIGSGSNGSSG<u>EELLSKNYHLENEVARLKK</u>SGSGSGSGSGSGSGSG SGSGSGSGSGEELLSKNYHSGSGSGRGSKRTADGSEFESPKKKRKV*
89	DNMT3A - hDNMT3 L-dCas9- 2xGCN4	MPAAKRVKLDGSNHDQEFDPKPYPPVPAEKRRPIRVLSLFDGIATGLLVKDLGIQVDRYASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHQIEWGPFDLVIGGSPCNDLSIVNPARKGLYEYGTGR LFFEFYRLLDARPKEGDDRPFFWL FENVVAMGVSDKRDISR FLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTV RKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWMFVDNL VLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRH WALVSEEELSLLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKVMDDKKYSIGLAIGTNSVGWA VITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICY LQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRIENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLAR GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYF TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVE

		ISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLF DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVD QELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTNRSDKNRGKSDNVPSEEVVKKMKNYWRQL LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFV YGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG GFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDL IIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQK QLFVEQHKHYLDEIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDAYPYDVPDYASLGS GSPKKKRKVEDPKKKRKVDGIGSGSNGSSG<u>EELLSKNYHLENEVARLKK</u>GSGSGSGSGSGSGSG GSGSGSGSGSG<u>EELLSKNYHLENEVARLKK</u>GSGSGQRGSKRTADGSEFESPKKKRKV*
90	DNMT3A - hDNMT3 L-dCas9- 10xGCN4	MPAAKRVKLDGSNHDQEFDPKVPYPPVPAEKRRPIRVLSLFDGIATGLLVKDLGIQVDRYIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPCNDLSIVNPARKGLYEYGTGR LFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKGDKQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKGPSRPFVWMFVDNL VLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRH WALVSEEELSLLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKVMDKKYSIGLAIGTNSVGWA VITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICY LQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL

[illegible]

	1xGCN4	GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFL ESGSDPGQLKHVVDVT DTV RKDVEEWGPF DLVYGATPPLGHTCDRPPSWYLFQFHRL LQYARPKPGSPRPFFWMFVDNL VLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRH WALVSEEELSLLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEP KKKRKVMSKLEKFTNCYLSKTLRF KAIPVGKTQENIDNKRLLEDEKRAEDYKGVKKLLDRYLSFINDVLHSIKLKNLNNYISLF RKKTRTEKENKELENLEINLRKEIAKAFKGN EGYKSLFKKDIIETILPEFLDDKDEIALVNSF NGFTTAFTGFFDNRENMFSEEAKSTSI AFR CINENLTRYISNMDIFEKVDAIFDKHEVQEIKEK ILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTKQKLP KFKPLYKQVLS DRESLSFYGEGYTSDEEVLEVFRNTLKNSEIFSSIKKLEKLFKNFDEYSSA GIFVKNGPAISTISKDIFGEWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSL EQLQEYADADLSVVEKLKEIIQKVDEIYKVYGSSEKLFDA DFVLEKSLKKND AVVAIMKDL LDSVKS FENYIKAFFGEGKETNRDES FYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKF KLYFQNPQFMGGWDKDKETDYRATILRYGSKY YLAIMDKKYAKCLQKIDKDDVNGNYEKI NYKLLPGPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLNDCHKLIDFFKD SISRYPKWSNAYDFNFSETEKYKDIA GFYREV EEQGYKVSFESASKKEVDKLV EEGKLYMFQ IYNKDFS D KSHGTPNLHTMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPI ANKNPDNPKKTTTLSYDVYKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYV IGIARGERNLLYIVVVDGKG NIVEQYSLNEIINN FN GIRIKTDYHSLDDKKEKERFEARQNW T SIENIKELKAGYISQVVHKICELVEKYDAVIALADLNSGFKNSRVKVEKQVYQKFEKMLIDK LNYMVDKKS NPCATGGALKGYQITNKFESFKSMSTQNGFIFYIPAWLT SKIDPSTGFVNLLK TKYTSIADSKKFISSFDRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPK KNNVFDWEEVCLTSAYKELFNKYGINYQQGDIRALLCEQSDKAFYSSFMALMSLMLQMRN SITGR TDVAFLISPVKNSDGIFYDSRNYEAQENAILPKNADANGAYNIARKVLWAIGQFKKAE DEKLDKVKIAISNKEWLEYAQTSVKHAYPYDVPDYASLGSGSPKKKRKVEDPKKKRKVDGIGS GSNGSSG<u>EELLSKNYHLENEVARLKK</u>SGSGSGSGSGSGSGSGSGSGSGSGSGSGSGEELLSKNYHGSGS GQRGSKRTADGSEFESPKKKRKV*
92	DNMT3A - hDNMT3	MPAAKRVKLDGSNHDQEFDPKVPYPPVPAEK RKPIRVLSLFDGIATGLLV LKDLGIQVDRYIASE VCEDSITVGMVRHQGKIMYVG D VRSVTQKH IQEWGPF DLVIGGSPCNDLSIVNPARKGLYEGTGR LF FEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFL ENPVMIDAKEVSAHRARY

[illegible]

[illegible]

		SGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNEFDLAEDAKLQL SKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHH QDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGP LARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD SVEISGVEDRFNASLGTYHDLKKIHKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDS LTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEM ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWR QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESE FVYG DYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPPKK YGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKK DLIHKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQ KQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNL GAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDAYPYDVPDYASLG SGSPKKKRKVEDPKKKRKVDGSGSETPGTSESATPES<<u>RTLVTFKDVFVDFTREEWKLLDTAQ</u> <u>QIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEE</u>>*
97	DNMT3A - hDNMT3 L-scFv	MPAAKRVKLDGSNH DQEFDPKPVYPPVPAEKRKPIRVLSLFDGIATGLLV LKDLGIQVDRYIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTV RKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRL LQYARPKPGSPRPFFWMFVDNL VLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRH WALVSEEELSLLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST

[illegible]

[illegible]

[illegible]

		AEDAKLQLSKDXYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASM IKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTF RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGTYHDLKHKDKDFLDNEENEDILEDIVLTLTLFEDREM IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGR HKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTNRSDKNRGKSDNPSEEVV KKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS RMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGITALI KKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR PLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIAR KKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKPNIDFLEA KGYKEVKKDLIHKLPKYSLEFENGKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDRDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQISITGLYETRIDLSQLGGDAYP YDVPDYASLGGSPKKKRVKVEDPKKKRVKVDGSGSETPGTSESATPES<<u>MNNSQGRVTFEDVTVN</u> <u>FTQGEWQRLNPEQRNLRYDVMLENYSNLVSYGQGETTKPDVILRL</u><u>EQGKEPWLEEEVVLG</u> <u>SGRAEKNGDIGGQIWPKDVKESL</u>>*
102	scFv- hDNMT3 L- DNMT3A	<u>MPAAKRVKLDGSGPDIVMTQSPSSLSASVGDRTTICRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTN</u> <u>NRAPGVPSRFSGLIGDKATLTISLQPEDFATYFCALWYSNHWFVFGQGTKVELKRGGGGSGGGGSGG</u> <u>GGSSGGGSEVKLESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN</u> <u>SALKDRFIISKDNGKNTVYLQMSKVRSDDTALYYCVTGLFDYWGQGTLVTVSSYPYDVPDYAGGGGGS</u> <u>GGGGSGGGGSGGGGSMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVTD</u> <u>TVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLQYARPKGPSRPFVWMFVDNLV</u> <u>LNKEDLDVASRFLMEPVVTIPDVHGGSLQNAVVRVWSNIPAIRSSRHVALVSEEELSLAQNKQSS</u> <u>KLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLSSGNSNANSRGPSFSSGLVPLSLRGSHMNHDQE</u> <u>FDPPKVYPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVRHQGKIM</u>

[illegible]

[illegible]

	KRAB	<u>GGSSGGGSEVKLLESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEWIGVIWGDGITDYN</u> <u>SALKDRFIISKDNGKNTVYLQMSKVRSDDTALYYCVTGLFDYWGGQTLVTVSSYPYDVPDYAGGGGGS</u> GGGGSGGGSGGGSGSGPKKKRKVMDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLG NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAH MIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQ YADFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGS IPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEET ITPWNFEEVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKHKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYT GWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDS LHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVAI VPQSFLKDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVS DFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPPKYGGFDSPTVAYSVLVAVK EKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIHKLPKYSLELENGRKR MLASAGELQKGNELALPSKYVNFLYLASHYEKLKGPEDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT STKEVL DATLIHQ SITGLYETRIDLSQLGGDAYPYDVPDYASLGS GSPKKRKVEDPKKKRKV DGSGSETPGTSESATPES<<u>RTLVTFKDVFVDFTREEWKLLDTAQOIVYRNVMLENYKNLVSLG</u> <u>YQLTKPDVILRLEKGEEP</u>>*
108	1xGCN4- DNMT3L -	MPAAKRVKLDGSGSGSGSGE<u>ELL SKNYHLENEV</u><u>ARLKK</u>GSGSGSGSGSGSGSGSGSGSGS GTPGTSEGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSESATPES<u>MFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKH</u>

	DNMT3A	VVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLQYARPKPGSPRPFFWM FVDNLVLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRHVALVSEELSLLA QNKQSSKLA AKWPTKL VKNCFLPLREYFKYFSTELTSSLSSGNSNANSRGPSFSSGLVPLSLRGSH MNHDQEFDPKVPYPPVPAEK RKPIRVLSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVR HQ GKIMYVGDVRSVTQKHQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLHDA RPKEGDDRPFFWL FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRP LASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGGKQHFVFMNEKEDILWCTEMERVFG FPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACV*
109	scFv- dCas9- HP1a	MPAAKRVKLDGSGPDIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGKLFKGLIGGTN NRAPGVPSRFSGLIGDKATLTISLQPEDFATYFCALWYSNHWWFGQGTKVELKRGSGGGSGGGSGGG GGSSGGGSEVKLLESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWWVROAPGRGLEWIGVWGDGITDYN SALKDRFIHSDNGKNTVYLQMSKVRSDDTALYYCVTGLFDYWGOGLTVTVSSYPYDVPDYAGGGGGS GGGGSGGGSGGGSGSGPKKKRKVMDDKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLG NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRNKRICYLQEIFSNEMAKVDDSSF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLNLLAQIGDQ YADFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGS IPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEET ITPWNFEVVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYT GWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDS LHEHIANLAGSPAIKKGILQTVKVVDLYVKVMGRHKPENIVIAMARENQTTQKGQKNSRER MKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVAI VPQSFLKDDSIDNKVLT RSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK AERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVS DFRKDFQFYKVREINNYHHHAHDAYLNAVVG TALIKKYPKLESEFVYG DYKVVDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL

		MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPPKKYGGFDSPTVAYSVLVAVK EKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKR MLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT STKEVL DATLIHQ SITGLYETRIDLSQLGGDAYPYDVPDYASLGSGSPKKKRKVEDPKKKRKV DGSGSETPGTSESATPESRTL<<u>GKKTKRTADSSSEDEEEYVVEKVLDRRVVKQVEYLLKWK</u> <u>GFSEEHNTWEPEKNLDCPELISEFMKKYKKMKEGENNPREKSESNNRKSNSNSADDIKS</u> <u>KKKREQSNDIARGFERGLEPEKIIGATDSCGDLMLMKWKDTDEADLVLAKEANVKCPOIV</u> <u>IAFYEERLTWHAYPEDAENKEKETAKSE>*</u>
110	DNMT3A - DNMT3L -dCas9- 7xGFP11	PAAKRVKLDGSNHDQEFDPKVPVPPVPAEKRRPIRVLSLFDGIATGLLVKDLGIQVDRYIASEVC EDSITVGMVRHQGKIMYVGDVRSVTQKHIEQEWGPFDLVIGGSPCNDLSIVNPARKGLYE GTGR LF FEFYRL LHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFW GNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKGDKQHFPVFMNEKEDILW CTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPS FSSGLVPLSLRGSHMGPMIYKTVSAWKRPVRLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVE DVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFWFWM DNLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYQLQAQV RSRSKLDAPKYDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPSGGSPAGSPTSTEEGTSESAT PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKVMDDKKYSIGLAI GTNSVGWAVITDEYKVP SKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRY TRRKNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFE ENPINASGVDAKAILSARLSKSRRL ENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTFR IPYVYGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVL P KHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF KKIECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF

		MLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRH KPENIVIAMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYY LQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSEEVVK KMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR MNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIK KYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRP LIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK KDWDPKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSSSF EKNPIDFLEAK GYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEK LK GSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENI IHLFTLTNLGAPAAFYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDAYPY DVPDYASLGSGSPKKRKVEDPKKKRKVDGSGSETPGTSESATPESGSGSNGSSGSMRDHMLHE <u>YVNAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNA</u> <u>AAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNAAGI</u> <u>TAAARSGSGQRGSKRTADGSEFESPKKKRKV*</u>
111	GFP1-10- ZIM3	MSKGEELFTGVVPILVELDGDYNGHKFSVRGEGEGDATIGKLT LKFICTTGKLPVPWPTLVTTLT YGYQ <u>CFSRYPDHMKRHDFEKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNIL</u> <u>GHKLEYNFNSHNVYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQT</u> <u>VLSKDPNEKLDPGGGGSG<MNNSOGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYS</u> <u>NLVSYGQGETTKPDVILRLLEQGKEPWLEEEVLGSGRAEKNGDIGGQIWKPKDVKESL>*</u>
112	DNMT3A - hDNMT3 L-dCas9- 1xGFP11	MPAAKRVKLDGNSHDQEFDPKVPYPPVPAEKRKPIRVLSLFDGIATGLLV LKDLGIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPV RVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKGSPRPFFWMFVDNL VLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRHWALVSEEELSLLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKVMDKKYSIGLAIGTNSVGWA

		VITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICY LQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRLLENIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLAR GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYF TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVE ISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLF DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVD QELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVG TALIKKYPKLESEFV YGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG GFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDL IHKLPKYSLFELENGKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQK QLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDAYPYDVPDYASLGS GSPKKKRKVEDPKKKRKVDGSGSETPGTSESATPESGSGSNGSSGSMRDHMLHEYVNAAGITAA ARSGSGQRGSKRTADGSEFESPKKKRKV*
113	GFP1-10- HDAC1	<u>MSKGEELFTGVVPIVVELDGDYNGHKFSYRGE GEGDATIGKLT LKFICTTGKLPVPWP TLT VTTLT YGYQ</u> <u>CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNIL</u> <u>GHKLEYNFNSHNVIYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQT</u> <u>VLSKDPNEKLDPGGGSG<MAQTQGTTRRKVCYYYDGDVGNYYYGQGHMPKPHRIRMTHNL</u> <u>LLNYGLYRKMEIYRPHKANAEEMTKYHSDDYIKFLRSIRPDNMSEYSKOMQRFNVGEDCPV</u> <u>FDGLFEFCQLSTGGSVASAVKLNKQQT DIAVNWAGGLHHAKKSEASGFCYVNDIVLAILEL</u>

		<u>LKYHQRVLYIDIDIHHGDGVEEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAV</u> <u>NYPLRDGIDDESIEAIFKPVMSKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCV</u> <u>EFVKSFNLPMLMLGGGGYTIRNVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISP</u> <u>SNMTNQNTNEYLEKIKQRLFENLRMLPHAPGVQMQAIPEDAIPESGDEDEDDEDPDKRISICSS</u> <u>DKRIACEEEFSDSEEEGEGGRKNSSNFKKAKRVKTEDEKEKDPEEKKEVTEEEKTKEEKPE</u> <u>AKGVKEEVKLA>*</u>
114	DNMT3A - hDNMT3 L- dCas12a- 1xGFP11	MPAAKRVKLDGNSHNDQEFDPKVPYPPVPAEKRKPIRVLSLFDGIATGLLVKLDLGIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHQIEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWLFEENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKGDKQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRWSVPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWMFVDNL VLNKEDLDVASRFLEMEPVVTIPDVHGGSLQNAVVRVWSNIPAIRSSRHVALVSEEELSLLAQNKQS SKLAAKWPTKLKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKVMKLEKFTNCYLSKTLRF KAIPVGKTQENIDNKRLLEDEKRAEDYKGVKKLLDRYVLSFINDVLHSIKLKNLNNYISLF RKKTRTEKENKELENLEINLRKEIAKAFKGNEGYKSLFKKDIETILPEFLDDKDEIALVNSF NGFTTAFTGFFDNRENMFSEEAKSTSIAPRCINENLTRYISNMDIFEKVDAIFDKHEVQEIKEK ILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQTKQKLP KFKPLYKQVLSDRRESLSFYGEGYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSSA GIFVKNGPAISTISKDIFGEWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSL EQLQEYADADLSVVEKLKEIIIQKVDEIYKVYGSSEKLFDAADFVLEKSLKKNDAVVAIMKDL LDSVKSFENYIKAFFGEGKETNRDESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKF KLYFQNPQFMGGWDKDKETDYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKI NYKLLPGPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLNDCHKLIDFFKD SISRPKWSNAYDFNFSETEKYKDIAGFYREVVEEQGYKVSFESASKKEVDKLVEEGKLYMFQ IYNKDFSDKSHGTPNLHTMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPI ANKNPDNPKKTTTSLSYDVYKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYV IGIARGERNLLYIVVVDGKGNIQYSLNEIINNFGIRIKTDYHSLDDKKEKERFEARQNW

		SIENIKELKAGYISQVVHKICELVEKYDAVIALADLNSGFKNSRVKVEKQVYQKFEKMLIDK LNYMVDKKSNPCATGGALKGYQITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLK TKYTSIADSKKFISSFDRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPK KNNVFDWEEVCLTSAYKELFNKYGINYQQGDIRALLCEQSDKAFYSSFMALMSLMLQMRN SITGRTDVAFLISPVKNSDGIFYDSRNYEAQENAILPKNADANGAYNIARKVLWAIGQFKKAE DEKLDKVKIAISNKEWLEYAQTSVKHAYPYDVPDYASLGSGSPKKRKRVEDPKKKRKYDVGSG SETPGTSESATPESGSGSNGSSGSMRDHMYLHEYVNAAGITAAARSGSGQRGSKRTADGSEFESPK KKRKV*
115	DNMT3A - hDNMT3 L- dCas12i- 1xGFP11	MPAAKRVKLDGNSHDQEFDPKVPYPPVPAEKRPPIRVLSLFDGIATGLLVKDLGIQVDRYIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEYRLLHDARPKEGDDRPFFWLFEENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLQYARPKPGSPRPFFWMFVDNL VLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRHWALVSEEELSLLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKMVGITMSD TTFRPYKSKLNP NERKMPMVVD TMNWC DNACSVFFDLFVGLYGGVKIGDVTNFLT EKQDTTTDADLLCAIN WFRAMPVSNPD AVVGPEQLVVRFRDYS GHNPSPMASMYLSAPVNPRYT WVD CRFLYQQM CQTLGANLDVDISTLVQTGLLPIQKDTQLNAYSFC SNVFGEGVKEDPSAKVTSYTA AANALS GATNPIITYEDYRQILSSAITNRPSRKGA KDILAADLDSTKTGVLP GNYVSQRIAMYQASAAS KARGVVLPNRLRVKAYLQNLVGLNYICTTWAAPMLRAVGD IKSNSNNYRFVGERLERQQ EIDNLLAVSNIRSAQAILQQFRANESNQFYIKKRHIGKGLPSLYDMWDL MDEKEGIDNYIQSL KGDKTERLPIRELLQHYPHRKSVSAELFLSAAELTDVEDRQNRCMVHPTVRGSLVINFGKSS TIFGAITPPDTLIKFRGEMVPAGQTGMIWFTMHLVDGGKWVKHHIPCHSSRYFEEIYAWKK DLPTLPFPRRSYGYAIGNKISDTRKIDLRRKKASKQYLRTLENMTHNVEFDPTTQFMVDQE LNV TITARVKKIPPVRDVPVGS RIMGLNKNQTIRD CYSIWERVEANTSGSFQHPVVQDVWLR LVADGHVTSTVNGPNGQIDQMSYSGPDVANC GNWMQERSNFITSVDDADQSIMKHFLWNN RVKLYAWNCVYAGTLMKAISKQMTTANRN LFRNEILSFVGGFRGCR LGSLAQVSLDLLRR

		MKSVISCYFSRNNAFTEERSAFDKELMELFESLDNKMNNKRKEKVNRTLSSIFQVAKQHNV VRLVVTKNLQTPDANVRSMNQRSVDWCARKIEEKIVDLGNVIGIYVDLVEALDLSHLDPIFI YIAKGDVGKECQFDEVESSLLPIHTSMFRDWHGLVNGRTTTTPIYHSALRSFASHYGLRFD DLARIKAAELAKMLPDQKVFIHRGGRFFMSTYPVTSTSRPVSYDGRQRFLNSGDMVAPVN TMLRGFRRARSAYPYDVPDYASLGSGSPKKKRKVEDPKKKRKVDGSGSETPGTSESATPESGSG SNGSSGSMRDHMLHEYYNAAGITAAARSGSGQQRGSKRTADGSEFESPKKKRKV*
116	DNMT3A - DNMT3L -GFP1-10	MPAAKRVKLDGNSHDQEFDPKVPYPPVPAEKRPPIRVLSLFDGIATGLLVKDLGIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEQWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLDARPKEGDDRPFFWLFEENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMGPMIYKTVSAWKRPVRLSLFRNIDKVLKSLGFLESGSGGGGTLK YVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWI FMDNLLTEDDQETTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLQA QVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSES ATMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATIGKLTLCFICTTGKLPVPWPTLVTTITYG VOCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDG NILGHKLEYNFNSHNVIYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLST QTVLSKDPNEK*
117	1xGFP11- dCas9- KRAB	MPAAKRVKLDGSGPDIVIGSGGRDHMLHEYYNAAGITAAARSGSGQQRGSKRTADGSEFESPKKK RKVGGSGSGPKKKRKVMDDKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKN LIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEE DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLIE GDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAI LRRQEDFYFPFLKDNREKIEKILTRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEVVDK GASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQK KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIKDKDFLD

		NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIIEGKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREI NNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQ TGGFSKESILPKRNSDKLIARKKDWDPPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKE LLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN ELALPSKYVNFYLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVLADANL DKVLSAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQ SITGLYETRIDLSQLGGDAYPYDVPDYASLGSGSPKKKKRKVEDPKKKRKVDGSGSETPGTSESA TPES<<u>RTLVTFKDVVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRL</u> <u>KGEEP</u>>*
118	7xGFP11- dCas9- ZIM3	MPAAKRVKLDGSMR<i>RDHMLHEYVNAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYV</i> <i>NAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNAAGITGGSGGRDHMLHEYVNA</i> <i>GITGGSGGRDHMLHEYVNAAGITAA</i>ARSGSGQRGSKRTADGSEFESPKKKRKVGGSGSGPKKKR KVMDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVSDTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKL FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLG LTPNFKSNFLAEDAKLQLSKDQYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVN TEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLK DNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT NFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKHKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILD FLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKV

		VDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT QLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRG KSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR QITKHVAQILDSRMNTKYDEENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDA YLNNAVGTALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTE ITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILP KRNSDKLIARKKDWDPPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERS FEKNPIDFLEAKGYKEVKKDLIKLPKYSLEFENGKRKMLASAGELQKGNELALPSKYVNF LYLASHYEKLKGPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETR IDLSQLGGDAYPYDVPDYASLGS GSPKKRKVEDPKKKRKVDGSGSETPGTSESATPES<<u>MNNSQG</u> <u>RVT FEDVT VNF TOGEWQRLNPEQRNLYRDVMLENYSNLVSVGOGETTKPDVILRLEQGKE</u> <u>PWLEEEVLGSGRAEKNGDIGGQIWKPKDVKESL</u>>*
119	DNMT3A - hDNMT3 L-GFP1- 10	MPAAKRVKLDGSNH DQEFDPK VYPPVPAEK RKPIRVLSLFDGIATGLLV LKDLGIQVDRYIASE VCEDSITVGMVRHQ GKIMYVGDVRSVTQKH IQEWGPF DLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPF DLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKGPSRPFFWMFVDNL VLNKEDLDVASRFLEMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRH WALVSEEELSLLAQNKQS SKLAAKWPTKL VKNCF LPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKVPSEM<<u>SKGEELFTGVVPILVELD</u> <u>GDVNGHKFSVRGEGEGDATIGKLT LKFICTTGKLPVPWPTLVTTLT YGVQCF SRYPDHMKRHDFFKSA</u> <u>MPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFN SHNYITADK</u> <u>QKNGIKANFTVRHNVEDG SVQLADHYQONTPIGDGPVLLPDNH YLSTQTVLSKDPNEK</u>*
120	DNMT3A - DNMT3L -1xGFP11	MPAAKRVKLDGSNH DQEFDPK VYPPVPAEK RKPIRVLSLFDGIATGLLV LKDLGIQVDRYIASE VCEDSITVGMVRHQ GKIMYVGDVRSVTQKH IQEWGPF DLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI

		<i>LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSPVIRHLFAPLKEYFACVSSGNSNANSR</i> <i>GPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLK</i> <i>YVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWT</i> <i>FMDNLLLTEDDQETTTRFLQTEAVTLQDVVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQA</i> <i>QVRSRSKLDAPKVDLLVKNCLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSES</i> <i>ATGGSGGRDHMVLEHYVNAAGITAAARSGSGQRGSKRTADGSEFESPKKKRKV*</i>
121	GFP1-10- dCas9- KRAB	<i>MPAAKRVKLDGSGPDIVMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATIGKLTCLKICTT</i> <i>GKLPVPWPTLVTTLTYGVCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDT</i> <i>LYNRIELKGTDFKEDGNILGHKLEYNFNHSHNYITADKQKNGIKANFTVRHNVEDGSGVQLADHYQONT</i> <i>PIGDGPVLLPDNHLYSTQTVLSKDPNEKLDPGGGGSGSGPKKKRKVMDDKYSIGLAIGTNSVGWA</i> <i>VITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRRKNRICY</i> <i>LQEIFSNEMAKVDDSFHRLSESLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV</i> <i>DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGV</i> <i>DAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD</i> <i>TYDDDLNLLAQIGDQYADFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDL</i> <i>TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIPILEKMDGTEELLVKLN</i> <i>REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLAR</i> <i>GNSRFAWMTRKSEETITPWNFEEVVDKGASQAQSFIERMTNFDKNLPNEKVLPKHSLLYEYF</i> <i>TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVE</i> <i>ISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLF</i> <i>DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTF</i> <i>KEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDLVKVMGRHKPENIVIMAR</i> <i>ENQTTQKGQKNSRERMKRIEKGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVD</i> <i>QELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTTRSDKNRGKSDNPSEEVVKKMKNYWRQL</i> <i>LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND</i> <i>KLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFV</i> <i>YGDYKVYDVVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI</i> <i>VWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG</i> <i>GFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDL</i> <i>IIKLPKYSLENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEQNEQK</i>

		QLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDAYPYDVPDYASLGS GSPKKKRKVEDPKKKRKVDGSGSETPGTSESATPES< <u>RTLVTFKDVFVDFTREEWKLLDTAQOI</u> <u>VYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP</u> >*
122	10xGFP1 1-dCas9- ZIM3	MPAAKRVKLDGSM <u>RDHMLHEYVNAAGIT</u> GGSGG <u>RDHMLHEYVNAAGIT</u> GGSGG <u>RDHMLHEYV</u> <u>NAAGIT</u> GGSGG <u>RDHMLHEYVNAAGIT</u> GGSGG <u>RDHMLHEYVNAAGIT</u> GGSGG <u>RDHMLHEYVNA</u> <u>GIT</u> GGSGG <u>RDHMLHEYVNAAGIT</u> AAAGSGG <u>RDHMLHEYVNAAGIT</u> GGSGG <u>RDHMLHEYVNA</u> <u>GIT</u> GGSGG <u>RDHMLHEYVNAAGIT</u> GGSGGSGSGQRGSKRTADGSEFESPKKKRKVGGSGSGPKK KRKVMDDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGET AEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRLEESFLVEEDKKHERHPIFG NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIY LALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL ENLIAQLPGEKKNGLFGNLIALS LGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADFLAAKNLS DAILLSDILR VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ EEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPF LKDNREKIEKILTRIPYVYGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIER MTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILED IVLTLTFLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTV KVVDLVKVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKN RGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVINNYHHAH DAYLNAVVG TALIKKYPKLESEFVYGDYKVYDV RKMIKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYN KH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYET

		RIDLSQLGGDAYPYDVPDYASLGSGSPKKRKRVEDPKKKRKRVDGSGSETPGTSESATPES<MNN SQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLVSVGOGETTKPDVILRLEQ GKEPWLEEEVLGSGRAEKNGDIGGQIWKPKDVKESL>*
123	scFv- ZIM3	MVPAAKRVKLDGSGPDIVMTQSPSSLSASVGDRTITCRSSTGAVTTSNYASWVQEKPGLFKGLIGGT NNRAPGVPSRFSGSLIGDKATLTISLQPEDFATYFCALWYSNHWWFGQGTKVELKRGGGSGGGGSG GGGSSGGGSEVKLLESGGGLVQPGGSLKLSCAVSGFSLTDYGVNWRQAPGRGLEWIGVIWGDGITY NSALKDRFIISKDNGKNTVYLOMSKVRSDDTALYYCVTGLFDYWGOGLVTVSSYPYDVPDYAGGGG GSGGGGSGGGGSGGGGSLDPGGGGSG<MNNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYR DVMLENYSNLVSVGOGETTKPDVILRLEQGKEPWLEEEVLGSGRAEKNGDIGGQIWKPK DVKESL>*
124	DNMT3A - hDNMT3 L-dCas9- 1xGCN4	MPAAKRVKLDGSGNHDQEFDPKVPYPPVPAEKRRPIRVLSLFDGIATGLLVKDLGIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHQIEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWLFEENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKGQDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWMFVDNL VLNKEDLDVASRFELEMPVTIPDVHGGSLQNAV RVWSNIPAIRSSRHVALVSEEELSLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKKRKRKVGSDKKYSIGLAIGTNSVGW AVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRIC YLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKL VDSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASG VDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK DTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKHKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAH

		LFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMA RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQ LLNAKLITQRKFDNLTKAERGGSELKAGFIKRLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKKPKLESEF VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG EIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKD LIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQK QLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGDAYPYDVPDYASLGS GSPKKKRKVEDPKKKRKVDGIGSGSNGSSGEELLSKNYHLENEVARLKKGSGSGSGSGSGSGSG GSGSGSGSGSGEELLSKNYHSGSGSGRGSKRTADGSEFESPKKKRKV*
361	DNMT3A - hDNMT3 L- XTEN80- TALE HPT8- 1xGCN4	MPAAKRVKLDGNSHNDQEFDPKVPYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHQIEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGR LFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLGRWSVPVIRHLFAPLKEYFACVSSGNSNANSR GPSFSSGLVPLSLRGSHMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLQYARPKGPSRPFVWMFVDNL VLNKEDLDVASRFLMEPV TIPDVHGGSLQNAV RVWSNIPAIRSSRHVALVSEEELSLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSGGPSSGAPPPSGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEPKKRKVMVDLRTLGYSSQQQKEIKIP KVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGK QWSGARALEALLTVAGELRGPPQLDGTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLT PDQVVAIASNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVL CQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETV QRLLPVLCQDHGLTPDQVVAIASNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGG GKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQV

[illegible]

		VVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQD HGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQR LLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQ ALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIA SHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLT PDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLC QDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALET VQRLLPVLCQDHGLTPDQVVAIASNIGGKQALESIVAQLSRPDPALAALTNDHLVALACLG GG RPAMDAVKKGLPHAPELIRRVRNRRIGERTSHRVAYPYDVPDYASLGSGSPKKKRKVEDPKKK RKVDGSGSETPGTSESATPES<<u>MNNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLEN</u> <u>YSNLYSVGQGETTKPDVILRLEQGKEPWLEEEVLGSGRAEKNGDIGGQIWKPKDYKESL</u>> *
--	--	---

不欲被任何理论所限，下文中的实施例仅仅是为了阐释本申请的复合物、制备方法和用途等，而并不用于限制本申请发明的范围。

实施例

实施例 1

设计和构建包含本申请的复合物的质粒

带有 HA 表位、P2A 的表观修饰招募系统和被招募元件（包括 Dnmt3a CD，Dnmt3l CD，dSpCas9 或 TALE，KRAB）的氨基酸序列是由 Genscript 公司优化成适于哺乳动物表达的核酸序列并合成，然后克隆到带有 CAG 启动子和 WPRE 的 pLV-CAG 载体，由 CAG 启动子表达被招募元件和自剪切的招募系统融合蛋白。

在优化不同功能元件时，不同功能元件是由 Genscript 公司优化成适于哺乳动物表达的核酸序列并合成。先通过 PCR 将需要替换的元件以外的载体扩增，再将需要替换的元件从公司合成的序列上扩增，同时引入同源臂序列，最后通过 NEBuilder 试剂将不同元件重组到载体中，构建成最终的表达质粒。

实施例 2

本申请的复合物对 PTP1b 基因表达的抑制作用

将包含本申请复合物（具有 SEQ ID NO: 51 所示氨基酸序列的第一融合物与具有 SEQ ID

NO: 52 所示氨基酸序列的第二融合物结合而成的复合物)的编辑工具和不同靶向到小鼠 PTP1b 基因的 gRNA (sg1、sg2, 或者 sg1 和 sg2 等量混合), gRNA 靶向序列的互补核苷酸序列如 SEQ ID NO: 336 和 337 所示, 以及对照 gRNA (NT gRNA, 其靶向序列的互补核苷酸序列如 SEQ IF NO: 365 所示) 共同转染到小鼠 N2a 细胞系 (700ng 编辑器+300ng gRNA, 24 孔板), 转染 72 小时后分选转染阳性的细胞, 用 Trizol 提取总 RNA, 通过 qPCR 定量 PTP1b 的相对表达量, 计算不同工具和不同 gRNA 的相对抑制效率 (图 2)。结果表明, 相比较于核 酸结合结构域 dCas9、甲基化因子 DNMT3A 和 DNMT3L、以及转录阻遏物 KRAB 直接融合 在一起的对照组 (图 2 右) 实施方案 (DNMT3A-DNMT3L-dCas9-KRAB), 本申请的复合物 对靶向 PTP1b 基因的 gRNA, 即 sg1 和/或 sg2, 均获得了更高或相当的抑制效率。

实施例 3

本申请的复合物对 PCSK9 基因表达的抑制作用

将不同版本的工具体外转录成 mRNA (待测工具信息参见下表总结), 然后和化学合成的 sgRNA 1:1 质量比混合 (第一融合物和第二融合物的 mRNA 与 sgRNA 的质量比为 0.5:0.5:1; sgRNA 靶向序列的互补核苷酸序列如 SEQ ID NO: 341 和/或 342 所示), 制备成 LNP (LNP 引用文献: Musunuru, K., Chadwick, A.C., Mizoguchi, T. *et al.* In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* **593**, 429–434 (2021).)。将制备好的 LNP 通过尾静脉注射到小鼠体内, 注射剂量为 4.5mg 每 kg 体重, 注射 4-10 天后, 从小鼠脸颊取 血, 通过 Elisa 测定血液中 PCSK9 蛋白的含量。PBS 组为注射等体积 PBS 的对照组。各工具 在注射后第 4 天的检测结果如图 3 所示, 与 PBS 对照组相比, 本申请的各组工具均对 PCSK9 基因表达显现出极为突出的抑制效果。

工具编号	结构域信息	氨基酸序列 (SEQ ID NO:)	mRNA 序列 (SEQ ID NO:)
V2+V3	scFv-KRAB、DNMT3A-DNMT3L-dCas9-1xGCN4	78、79	243、244
V2+V4	scFv-KRAB、DNMT3A-DNMT3L-dCas9-2xGCN4	78、80	243、245
V2+V5	scFv-KRAB、DNMT3A-DNMT3L-dCas9-4xGCN4	78、81	243、246
V2+V6	scFv-KRAB、DNMT3A-DNMT3L-	78、82	243、247

	dCas9-10xGCN4		
V7+V8	DNMT3A-DNMT3L-scFv、3xGCN4-dCas9-KRAB	83、84	248、249
V12	scFv-KRAB-2a-DNMT3A-DNMT3L-dCas9-1xGCN4	133	253
V13	scFv-KRAB-2a-DNMT3A-DNMT3L-dCas9-2xGCN4	134	254
V14	scFv-KRAB-2a-DNMT3A-DNMT3L-dCas9-10xGCN4	135	255
V15	scFv-ZIM3-2a-DNMT3A-hDNMT3L-dCas9-1xGCN4	136	256
V16	scFv-ZIM3-2a-DNMT3A-hDNMT3L-dCas9-2xGCN4	137	257
V17	scFv-ZIM3-2a-DNMT3A-hDNMT3L-dCas9-10xGCN4	138	258
V22	DNMT3A-DNMT3L-scFv-2a-1xGCN4-dCas9-KRAB	143	263
V24	DNMT3A-DNMT3L-scFv-2a-3xGCN4-dCas9-KRAB	145	265
V25	DNMT3A-DNMT3L-scFv-2a-5xGCN4-dCas9-KRAB	146	266
V26	DNMT3A-hDNMT3L-scFv-2a-3xGCN4-dCas9-ZIM3	147	267
V31	scFv-hDNMT3L-DNMT3A-2a-3xGCN4-dCas9-KRAB	152	272
V32	1xGCN4-ZIM3-2a-DNMT3A-hDNMT3L-dCas9-scFv	153	273
V37	GFP1-10-KRAB-2a-DNMT3A-DNMT3L-dCas9-7xGFP11	158	278
V9+V10	GFP1-10-KRAB、DNMT3A-	85、86	250、251

	DNMT3L-dCas9-1xGFP11		
V9+V11	GFP1-10-KRAB、DNMT3A- DNMT3L-dCas9-8xGFP11	85、87	250、252
V38	GFP1-10-KRAB-2a-DNMT3A- DNMT3L-dCas9-1xGFP11	159	279
V39	GFP1-10-ZIM3-2a-DNMT3A- hDNMT3L-dCas9-1xGFP11	160	280
V27	3xGCN4-dCas9-ZIM3-2a-DNMT3A- hDNMT3L-scFv	148	268
V23	DNMT3A-DNMT3L-scFv-2a-2xGCN4- dCas9-KRAB	144	264
V34	DNMT3A-DNMT3L-1xGCN4-2a-scFv- dCas9-KRAB	155	275
V45	DNMT3A-DNMT3L-1xGFP11-2a- GFP1-10-dCas9-KRAB	166	286
V43	DNMT3A-DNMT3L-GFP1-10-2a- 1xGFP11-dCas9-KRAB	164	284
V20	KRAB-scFv-2a-DNMT3A-DNMT3L- dCas9-2xGCN4	141	261
V40	GFP1-10-HDAC1-2a-DNMT3A- hDNMT3L-dCas9-1xGFP11	161	281
V33	KRAB-1xGCN4-2a-DNMT3A- DNMT3L-dCas9-scFv	154	274
V35	scFv-dCas9-KRAB-2a-1xGCN4- DNMT3L-DNMT3A	156	276
V36	scFv-dCas9-HP1a-2a-1xGCN4- DNMT3L-DNMT3A	157	277

实施例 4

本申请的复合物对 PCSK9 基因表达的抑制作用

根据实施例 3 的体外转录 mRNA（所用工具信息参见下表）、化学合成质量比为 1:1 的 sgRNA/mRNA 混合物（sgRNA 序列同实施例 3）、以及制备 LNP 的方法，将制备好的 LNP 加入到 Huh7 细胞中（1.25ug/mL 剂量），LNP 加入 4 天后，收取所有细胞，用 Trizol 提取总 RNA，通过 qPCR 定量 PCSK9 的相对表达量，计算不同工具的相对抑制效率。结果如图 4 所示，其中 NC 组为未加入 LNP 的对照，可以看出本申请提供的融合物对 PCSK9 基因的表达有着显著的抑制作用。

工具编号	结构域信息	氨基酸序列 (SEQ ID NO:)	mRNA 序列 (SEQ ID NO:)
V48	scFv-ZIM3-P2A-DNMT3A-hDNMT3L-XTEN80-TALE HPT8-1x GCN4	345	353
V49	scFv-ZIM3-P2A-DNMT3A-hDNMT3L-XTEN80-TALE HPT12-1x GCN4	346	354
V50	3xGCN4-TALE HPT8-ZIM3-P2A-DNMT3A-hDNMT3L-XTEN80-scFv	347	355
V51	3xGCN4-TALE HPT12-ZIM3-P2A-DNMT3A-hDNMT3L-XTEN80-scFv	348	356

权 利 要 求 书

1. 一种复合物，其包含第一融合物和第二融合物，所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域和至少一个招募结构域 A，并且其中另一个融合物包含转录阻遏物结构域和至少一个招募结构域 A'；其中，所述招募结构域 A 和招募结构域 A'能够相互作用，以使所述第一融合物和所述第二融合物其中之一的融合物或其部分能够被招募到另一个融合物附近。
2. 根据权利要求 1 所述的复合物，所述第一融合物或所述第二融合物包含核酸结合结构域。
3. 根据权利要求 1 或 2 所述的复合物，所述第一融合物包含 DNA 甲基化结构域、核酸结合结构域和至少一个招募结构域 A，所述第二融合物包含转录阻遏物结构域和至少一个招募结构域 A'。
4. 根据权利要求 1-3 中任一项所述的复合物，所述第一融合物中从 N 端到 C 端依次包含 DNA 甲基化结构域、核酸结合结构域和招募结构域 A。
5. 根据权利要求 1-4 中任一项所述的复合物，所述第二融合物中从 N 端到 C 端依次包含转录阻遏物结构域和招募结构域 A'，或者从 N 端到 C 端依次包含招募结构域 A'和转录阻遏物结构域。
6. 根据权利要求 1 或 2 所述的复合物，所述第一融合物包含转录阻遏物结构域、核酸结合结构域和至少一个招募结构域 A，所述第二融合物包含 DNA 甲基化结构域和至少一个招募结构域 A'。
7. 根据权利要求 1、2 和 6 中任一项所述的复合物，所述第一融合物中从 N 端到 C 端依次包含招募结构域 A、核酸结合结构域和转录阻遏物结构域。
8. 根据权利要求 1、2、6 和 7 中任一项所述的复合物，所述第二融合物中从 N 端到 C 端依次包含 DNA 甲基化结构域和招募结构域 A'，或者从 N 端到 C 端依次包含招募结构域 A'和 DNA 甲基化结构域。
9. 根据权利要求 1-8 中任一项所述的复合物，其特征在于：
 - 1) 所述第一融合物从 N 端到 C 端依次包含 DNA 甲基化结构域、核酸结合结构域和招募结构域 A，所述第二融合物从 N 端到 C 端依次包含转录阻遏物结构域和招募结构域 A'；或
 - 2) 所述第一融合物从 N 端到 C 端依次包含 DNA 甲基化结构域、核酸结合结构域和招募结构域 A，所述第二融合物从 N 端到 C 端依次包含招募结构域 A'和转录阻遏物结构域；或
 - 3) 第一融合物中从 N 端到 C 端依次包含招募结构域 A、核酸结合结构域和转录阻遏物结构域，所述第二融合物中从 N 端到 C 端依次包含 DNA 甲基化结构域和招募结构域 A'；

或

- 4) 第一融合物中从 N 端到 C 端依次包含招募结构域 A、核酸结合结构域和转录阻遏物结构域, 所述第二融合物从 N 端到 C 端依次包含招募结构域 A'和 DNA 甲基化结构域。
10. 根据权利要求 2-4、6、7 和 9 中任一项所述的复合物, 所述核酸结合结构域为 DNA 结合结构域。
11. 根据权利要求 10 所述的复合物, 所述 DNA 结合结构域选自: TALE 结构域、锌指结构域、tetR 结构域、大范围核酸酶、Cas 蛋白、Argonaute (Ago)蛋白, 以及其同系物、修饰形式或变体。
12. 根据权利要求 10 或 11 所述的复合物, 所述 DNA 结合结构域能够结合靶基因座的目标序列。
13. 根据权利要求 10-12 中任一项所述的复合物, 所述 DNA 结合结构域能够与引导 RNA 结合。
14. 根据权利要求 13 所述的复合物, 所述引导 RNA 能够特异性识别所述靶基因座的目标序列并与其杂交。
15. 根据权利要求 10-14 中任一项所述的复合物, 所述 DNA 结合结构域为 Cas 蛋白, 且所述 Cas 蛋白为 II 类 Cas 核酸酶。
16. 根据权利要求 15 所述的复合物, 所述 Cas 蛋白选自 II 类 II 型 Cas 核酸酶和 II 类 V 型 Cas 核酸酶。
17. 根据权利要求 15 或 16 所述的复合物, 所述 Cas 蛋白为 Cas9 或 Cas12 蛋白。
18. 根据权利要求 15-17 中任一项所述的复合物, 所述 Cas 蛋白为失活 Cas9 (dCas9)蛋白或失活 Cas12 (dCas12)蛋白。
19. 根据权利要求 10-14 中任一项所述的复合物, 所述 DNA 结合结构域包含 SEQ ID NO: 1-9、343 和 344 中任一项所示的氨基酸序列。
20. 根据权利要求 1-19 中任一项所述的复合物, 所述招募结构域 A 选自下列两组结构域其中一组中的任一个, 所述招募结构域 A'选自下列两组结构域中另一组中的任一个:
- 1) 通用控制非去阻遏蛋白 4 (GCN4)、来源于分裂绿色荧光蛋白(GFP)的 GFP11 片段或 GVKESLV 多肽; 和
- 2) 单链抗体(scFv)、来源于分裂绿色荧光蛋白(GFP)的 GFP1-10 片段或 PDZ 蛋白结构域。
21. 根据权利要求 20 所述的复合物, 其中:
- 1) 所述招募结构域 A 和所述招募结构域 A'其中之一的结构域为 GCN4, 并且其中另一个结构域为 scFv; 或

- 2) 所述招募结构域 A 和所述招募结构域 A' 其中之一的结构域为 GFP11 片段, 并且其中另一个结构域为 GFP1-10; 或
- 3) 所述招募结构域 A 和所述招募结构域 A' 其中之一的结构域为 GVKESLV, 并且其中另一个结构域为 PDZ 蛋白结构域。
22. 根据权利要求 1-21 中任一项所述的复合物, 所述 DNA 甲基化结构域包含至少一种 DNA 甲基转移酶或其功能活性片段。
23. 根据权利要求 22 所述的复合物, 所述 DNA 甲基转移酶选自 DNMT3A、DNMT3B、DNMT3c、DNMT1、DNMT2 和 DNMT3L。
24. 根据权利要求 1-23 中任一项所述的复合物, 所述 DNA 甲基化结构域包含至少一个 DNMT3A 和至少一个 DNMT3L。
25. 根据权利要求 22 或 23 所述的复合物, 所述 DNA 甲基转移酶包含 SEQ ID NO: 19-24 中任一项所示的氨基酸序列。
26. 根据权利要求 1-25 中任一项所述的复合物, 所述 DNA 甲基化结构域包含 DNMT3A-DNMT3L 结构域或 DNMT3L-DNMT3A 结构域; 其中, -表示其两端的结构域按照从 N 端到 C 端的顺序直接或间接地连接。
27. 根据权利要求 1-26 中任一项所述的复合物, 所述转录阻遏物选自下列所示结构域中的一种或多种: KRAB, ZIM3, ZNF680, ZNF554, ZNF264, ZNF582, ZNF324, ZNF669, ZNF354A, ZNF82, ZNF595, ZNF419, ZNF566, ZIM2, EHMT2, SUV39H1, ZFP1, TRIM28, EZH2, MXD1, SID, LSD1, HP1a, HDAC3, HDAC1, PRMT1, SETDB1, hSIRT1, ZNF436, ZNF257, ZNF675, ZNF490, ZNF320, ZNF331, ZNF816, ZNF41, ZNF189, ZNF528, ZNF543, ZNF140, ZNF610, ZNF350, ZNF8, ZNF30, ZNF98, ZNF677, ZNF596, ZNF214, ZNF37A, ZNF34, ZNF250, ZNF547, ZNF273, ZFP82, ZNF224, ZNF33A, ZNF45, ZNF175, ZNF184, ZFP28-1, ZFP28-2, ZNF18, ZNF213, ZNF394, ZFP1, ZFP14, ZNF416, ZNF557, ZNF729, ZNF254, ZNF764, ZNF785, ZNF10, CBX5, RYBP, YAF2, MGA, CBX1, SCM1, MPP8, SUMO3, HERC2, BIN1, PCGF2, TOX, FOXA1, FOXA2, IRF2BP1, IRF2BP2, IRF2BPL IRF-2BP1_2 N-terminal domain, HOXA13, HOXB13, HOXC13, HOXA11, HOXC11, HOXC10, HOXA10, HOXB9, HOXA9, ZFP28, ZN334, ZN568, ZN37A, ZN181, ZN510, ZN862, ZN140, ZN208, ZN248, ZN571, ZN699, ZN726, ZIK1, ZNF2, Z705F, ZNF14, ZN471, ZN624, ZNF84, ZNF7, ZN891, ZN337, Z705G, ZN529, ZN729, ZN419, Z705A, ZN302, ZN486, ZN621, ZN688, ZN33A, ZN554, ZN878, ZN772, ZN224, ZN184, ZN544, ZNF57, ZN283, ZN549,

ZN211, ZN615, ZN253, ZN226, ZN730, Z585A, ZN732, ZN681, ZN667, ZN649, ZN470, ZN484, ZN431, ZN382, ZN254, ZN124, ZN607, ZN317, ZN620, ZN141, ZN584, ZN540, ZN75D, ZN555, ZN658, ZN684, RBAK, ZN829, ZN582, ZN112, ZN716, HKR1, ZN350, ZN480, ZN416, ZNF92, ZN100, ZN736, ZNF74, ZN443, ZN195, ZN530, ZN782, ZN791, ZN331, Z354C, ZN157, ZN727, ZN550, ZN793, ZN235, ZN724, ZN573, ZN577, ZN789, ZN718, ZN300, ZN383, ZN429, ZN677, ZN850, ZN454, ZN257, ZN264, ZN485, ZN737, ZNF44, ZN596, ZN565, ZN543, ZFP69, SUMO1, ZNF12, ZN169, ZN433, ZN175, ZN347, ZNF25, ZN519, Z585B, ZN517, ZN846, ZN230, ZNF66, ZN713, ZN816, ZN426, ZN674, ZN627, ZNF20, Z587B, ZN316, ZN233, ZN611, ZN556, ZN234, ZN560, ZNF77, ZN682, ZN614, ZN785, ZN445, ZFP30, ZN225, ZN551, ZN610, ZN528, ZN284, ZN418, ZN490, ZN805, Z780B, ZN763, ZN285, ZNF85, ZN223, ZNF90, ZN557, ZN425, ZN229, ZN606, ZN155, ZN222, ZN442, ZNF91, ZN135, ZN778, ZN534, ZN586, ZN567, ZN440, ZN583, ZN441, ZNF43, ZN589, ZN563, ZN561, ZN136, ZN630, ZN527, ZN333, Z324B, ZN786, ZN709, ZN792, ZN599, ZN613, ZF69B, ZN799, ZN569, ZN564, ZN546, ZFP92, ZN723, ZN439, ZFP57, ZNF19, ZN404, ZN274, CBX3, ZN250, ZN570, ZN675, ZN695, ZN548, ZN132, ZN738, ZN420, ZN626, ZN559, ZN460, ZN268, ZN304, ZN605, ZN844, SUMO5, ZN101, ZN783, ZN417, ZN182, ZN823, ZN177, ZN197, ZN717, ZN669, ZN256, ZN251, CBX4, CDY2, CDYL2, ZN562, ZN461, Z324A, ZN766, ID2, ZN214, CBX7, ID1, CREM, SCX, ASCL1, ZN764, SCML2, TWST1, CREB1, TERF1, ID3, CBX8, GSX1, NKX22, ATF1, TWST2, ZNF17, TOX3, TOX4, ZMYM3, I2BP1, RHXF1, SSX2, I2BPL, ZN680, TRI68, HXA13, PHC3, TCF24, HXB13, HEY1, PHC2, ZNF81, FIGLA, SAM11, KMT2B, HEY2, JDP2, HXC13, ASCL4, HHEX, GSX2, ETV7, ASCL3, PHC1, OTP, I2BP2, VGLL2, HXA11, PDLI4, ASCL2, CDX4, ZN860, LMBL4, PDIP3, NKX25, CEBPB, ISL1, CDX2, PROP1, SIN3B, SMBT1, HXC11, HXC10, PRS6A, VSX1, NKX23, MTG16, HMX3, HMX1, KIF22, CSTF2, CEBPE, DLX2, PPARG, PRIC1, UNC4, BARX2, ALX3, TCF15, TERA, VSX2, HXD12, CDX1, TCF23, ALX1, HXA10, RX, CXXC5, SCML1, NFIL3, DLX6, MTG8, CEBPD, SEC13, FIP1, ALX4, LHX3, PRIC2, MAGI3, NELL1, PRRX1, MTG8R, RAX2, DLX3, DLX1, NKX26, NAB1, SAMD7, PITX3, WDR5, MEOX2, NAB2, DHX8, CBX6, EMX2, CPSF6, HXC12, KDM4B, LMBL3, PHX2A, EMX1,

NC2B, DLX4, SRY, ZN777, ZN398, GATA3, BSH, SF3B4, TEAD1, TEAD3, RGAP1, PHF1, GATA2, FOXO3, ZN212, IRX4, ZBED6, LHX4, SIN3A, RBBP7, NKX61, R51A1, MB3L1, DLX5, NOTC1, TERF2, ZN282, RGS12, ZN840, SPI2B, PAX7, NKX62, ASXL2, FOXO1, GATA1, ZMYM5, LRP1, MIXL1, SGT1, LMCD1, CEBPA, SOX14, WTIP, PRP19, NKX11, RBBP4, DMRT2, SMCA2, 以及其功能活性片段。

28. 根据权利要求 1-27 中任一项所述的复合物, 所述转录阻遏物结构域包含 SEQ ID NOs: 25-50 中任一项所示的氨基酸序列。
29. 根据权利要求 1-28 中任一项所述的复合物, 其中:
- 1) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE-n×GCN4, 并且其中另一个融合物包含转录阻遏物结构域-scFv; 或
 - 2) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE-scFv, 并且其中另一个融合物包含转录阻遏物结构域-GCN4; 或
 - 3) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE-n×GFP11, 并且其中另一个融合物包含转录阻遏物结构域-GFP1-10; 或
 - 4) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE-GFP1-10, 并且其中另一个融合物包含转录阻遏物结构域-GFP11; 或
 - 5) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE-n×GCN4, 并且其中另一个融合物包含 scFv-转录阻遏物结构域; 或
 - 6) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE-scFv, 并且其中另一个融合物包含 GCN4-转录阻遏物结构域; 或
 - 7) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE-n×GFP11, 并且其中另一个融合物包含 GFP1-10-转录阻遏物结构域; 或
 - 8) 所述第一融合物和所述第二融合物其中之一的融合物包含 DNA 甲基化结构域-dCas9 或 dCas12 或 TALE-GFP1-10, 并且其中另一个融合物包含 GFP11-转录阻遏物结构域; 其中, -表示其两端的结构域按照从 N 端到 C 端的顺序直接或间接地连接; n×GCN4 或 n×GFP11 分别表示 n 个通过接头序列连接的 GCN4 拷贝或 n 个通过接头序列连接的 GFP11 拷贝, n 选自 1 至 20 的任一整数。

30. 根据权利要求 1-29 中任一项所述的复合物，所述第一融合物和/或第二融合物包含 SEQ ID NO: 51-76、78-82、85-93、103-105、110-115、123、124、361 和 362 中任一项所述的氨基酸序列。
31. 根据权利要求 1-30 中任一项所述的复合物，其包含 SEQ ID NO: 133-142、153、154、158-163、168、345 和 346 中任一项所示的氨基酸序列。
32. 根据权利要求 1-28 中任一项所述的复合物，其中：
- 1) 所述第一融合物和所述第二融合物其中之一的融合物包含 $n \times \text{GCN4-dCas9}$ 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 DNA 甲基化结构域-scFv；或
 - 2) 所述第一融合物和所述第二融合物其中之一的融合物包含 scFv-dCas9 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 DNA 甲基化结构域-GCN4；或
 - 3) 所述第一融合物和所述第二融合物其中之一的融合物包含 $n \times \text{GFP11-dCas9}$ 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 DNA 甲基化结构域-GFP1-10；或
 - 4) 所述第一融合物和所述第二融合物其中之一的融合物包含 GFP1-10-dCas9 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 DNA 甲基化结构域-GFP11；或
 - 5) 所述第一融合物和所述第二融合物其中之一的融合物包含 $n \times \text{GCN4-dCas9}$ 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 scFv-DNA 甲基化结构域；或
 - 6) 所述第一融合物和所述第二融合物其中之一的融合物包含 scFv-dCas9 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 GCN4-DNA 甲基化结构域；或
 - 7) 所述第一融合物和所述第二融合物其中之一的融合物包含 $n \times \text{GFP11-dCas9}$ 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 GFP1-10-DNA 甲基化结构域；或
 - 8) 所述第一融合物和所述第二融合物其中之一的融合物包含 GFP1-10-dCas9 或 dCas12 或 TALE -转录阻遏物结构域，并且其中另一个融合物包含 GFP11-DNA 甲基化结构域；其中，-表示其两端的结构域按照从 N 端到 C 端的顺序直接或间接地连接； $n \times \text{GCN4}$ 或 $n \times \text{GFP11}$ 分别表示 n 个通过接头序列连接的 GCN4 拷贝或 n 个通过接头序列连接的 GFP11 拷贝， n 选自 1 至 20 的任一整数。
33. 根据权利要求 1-28 和 32 中任一项所述的复合物，所述第一融合物和/或第二融合物包含 SEQ ID NO: 83、84、94-102、106-109、116-122、363 和 364 中任一项所示的氨基酸序列。
34. 根据权利要求 1-28、32 和 33 中任一项所述的复合物，其包含 SEQ ID NO: 143-152、155-

157、164-167、347 和 348 中任一项所示的氨基酸序列。

35. 根据权利要求 1-34 中任一项所述的复合物，所述复合物还包含核定位信号和/或标记物结构域。
36. 编码权利要求 1-35 中任一项所述复合物的核酸。
37. 根据权利要求 36 所述的核酸，其为重组载体。
38. 根据权利要求 37 所述的重组载体，其还包括非编码区。
39. 根据权利要求 38 所述的非编码区，其选自内含子、调节元件、启动子、增强子、终止序列以及 5'和 3'非翻译区。
40. 根据权利要求 36-39 中任一项所述的核酸，其包含编码所述第一融合物的第一核酸片段，以及编码所述第二融合物的第二核酸片段。
41. 根据权利要求 40 所述的核酸，其中所述第一核酸片段与第二核酸片段通过编码剪切肽的核酸片段连接。
42. 根据权利要求 41 所述的核酸，所述剪切肽为 2A 肽和/或 IRES。
43. 根据权利要求 42 所述的核酸，所述 2A 肽选自 P2A、T2A、E2A 和 F2A。
44. 根据权利要求 36-43 中任一项所述的核酸，其包含 SEQ ID NO: 169-335 和 349-360 中任一项所示的核酸序列。
45. 一种递送载体，所述递送载体包含权利要求 1-35 中任一项所述的复合物和/或权利要求 36-44 中任一项所述的核酸，以及任选地包含脂质体和/或脂质纳米颗粒。
46. 一种组合物，所述组合物包含权利要求 1-35 中任一项所述的复合物、权利要求 36-44 中任一项所述的核酸、和/或权利要求 45 所述的递送载体。
47. 一种细胞，所述细胞包含权利要求 1-35 中任一项所述的复合物、权利要求 36-44 中任一项所述的核酸、权利要求 45 所述的递送载体、和/或权利要求 46 所述的组合物。
48. 一种试剂盒，所述试剂盒包含权利要求 1-35 中任一项所述的复合物、权利要求 36-44 中任一项所述的核酸、权利要求 45 所述的递送载体、权利要求 46 所述的组合物、和/或权利要求 47 所述的细胞。
49. 一种调控靶基因表达的方法，所述方法包括施用权利要求 1-35 中任一项所述的复合物、权利要求 36-44 中任一项所述的核酸、权利要求 45 所述的递送载体、权利要求 46 所述的组合物、权利要求 47 所述的细胞和/或权利要求 48 所述的试剂盒。
50. 根据权利要求 49 所述的方法，所述方法包括将所述复合物、所述核酸、所述递送载体、所述组合物、所述细胞、和/或所述试剂盒引入含有所述靶基因的细胞。
51. 根据权利要求 49 所述的方法，所述方法包括使所述复合物、所述核酸、所述递送载体、

和/或所述组合物接触所述靶基因附近和/或所述靶基因的调控元件。

52. 根据权利要求 51 所述的方法, 所述调控元件包含核心启动子、近端启动子、远端增强子、沉默子、绝缘子元件、边界元件和/或基因座控制区。
53. 权利要求 1-35 中任一项所述的复合物、权利要求 36-44 中任一项所述的核酸、权利要求 45 所述的递送载体、权利要求 46 所述的组合物、权利要求 47 所述的细胞和/或权利要求 48 所述的试剂盒用于制备药物的用途, 所述药物用于治疗或缓解与靶基因表达异常和/或靶基因活性异常相关的疾病或其病症。