Erik Jonsson School of Engineering and Computer Science

2014-05

Assembly and Validation of Versatile Transcription Activator-Like Effector Libraries

UTD AUTHOR(S): Yi Li, Kristina Ehrhardt, Michael Qiwei Zhang and Leonidas Bleris

Assembly and Validation of Versatile Transcription Activator-Like Effector Libraries

Yi Li[1,2], Kristina Ehrhardt[1,2], Michael Q. Zhang[2,3] & Leonidas Bleris[1,2,4]

[1]Bioengineering Department, The University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA, [2]Center for Systems Biology, The University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA, [3]Molecular and Cell Biology Department, The University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA, [4]Electrical Engineering Department, The University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA.

The ability to perturb individual genes in genome-wide experiments has been instrumental in unraveling cellular and disease properties. Here we introduce, describe the assembly, and demonstrate the use of comprehensive and versatile transcription activator-like effector (TALE) libraries. As a proof of principle, we built an 11-mer library that covers all possible combinations of the nucleotides that determine the TALE-DNA binding specificity. We demonstrate the versatility of the methodology by constructing a constraint library, customized to bind to a known p53 motif. To verify the functionality in assays, we applied the 11-mer library in yeast-one-hybrid screens to discover TALEs that activate human SCN9A and miR-34b respectively. Additionally, we performed a genome-wide screen using the complete 11-mer library to confirm known genes that confer cycloheximide resistance in yeast. Considering the highly modular nature of TALEs and the versatility and ease of constructing these libraries we envision broad implications for high-throughput genomic assays.

Transcription activator-like effectors (TALEs) are proteins secreted by *Xanthomonas* bacteria to target specific DNA sequences[1,2]. TALE family members contain a highly conserved and repetitive region (Fig. 1a) and their DNA binding capacities are dependent on the amino acid positions 12 and 13 (i.e. repeat variable diresidue or RVD) of each modular repeat[3,4]. Recent advances show that TALE proteins consist of a modular DNA-binding domain (DBD) that can be rapidly synthesized *de novo*[5]. The ability to complement the DBD with a functional domain permits a wide range of applications that include gene activation, repression, and nucleotide deletion and insertion, in a variety of model organisms and cell types[6–10].

In a parallel domain, high-throughput genome-wide screening experiments continue to provide critical insight into biology and medicine, shedding light on traditional cellular phenotypes, disease networks, and for reverse engineering endogenous pathways. Unfortunately, currently available methods are highly constrained, expensive, and often suffer from non-specific target effects[11–15]. Traditionally, these assays involve mutations or RNAi and explore the effect of loss of function coupled with selective pressure[15–17]. The complement, or gain-of-function, can be partially accomplished in yeast with introduction of cDNA[18]. In mammalian cells, a recently developed methodology provides a path for the use of zinc finger[19] proteins for genetic screens[20,21]. This protein-based technology can expand the repertoire of genetic screening in human cells by allowing the delivery of a variety of functional domains. In addition, RNAi-based screening is extensively used for genome-wide screens[11,22], and typically relies on shRNA delivery. Finally, the CRISPR-Cas9 system was recently used to perform genetic screens in human cells using genome-scale lentiviral single-guide RNA (sgRNA) libraries[23,24].

Here, we introduce novel transcription activator-like effector (TALE) libraries. Our assembly methodology is rapid, modular, and allows precise control of the resulting RVD content and overall DNA binding domain composition. As a proof of principle, we built and verified two libraries: a sequence-constrained library to target a known p53 motif and an 11-mer library that spans all possible 11-nucleotide DNA targets (i.e. resulting to a maximum of 4,194,304 combinations). In contrast, an alternative Golden-Gate assembly system that relies on 432 plasmids assembled using a multi-well plate format, produced a TALE Nuclease (TALEN) library that targets 18,740 exclusively protein-coding gene targets[25].
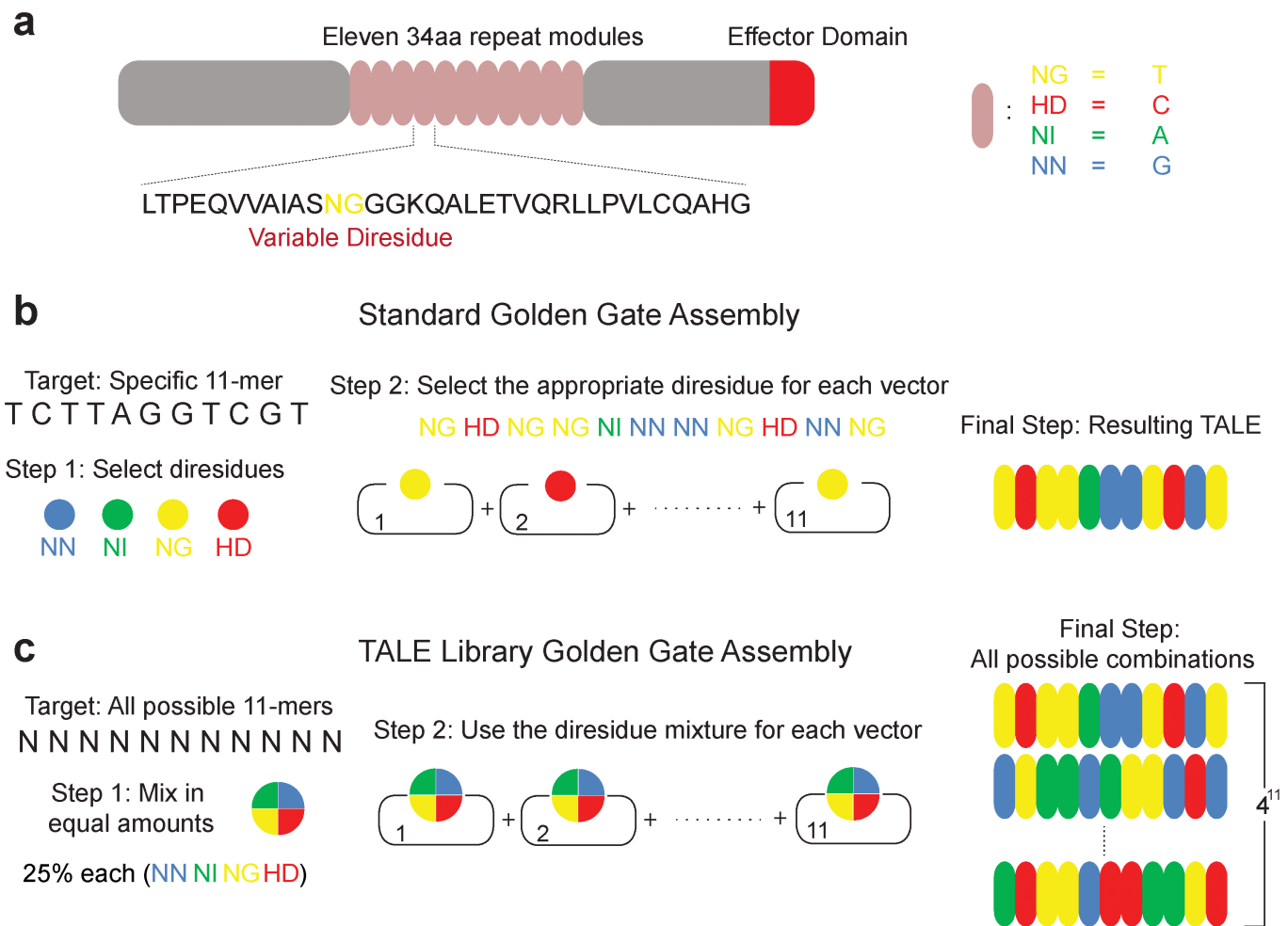
Figure 1 | Construction of an 11-mer TALE-VP16 library. (a) Schematic illustration of a TALE protein depicting the tandem repeat domain and the variable diresidues (RVDs). (b) Schematic illustration of a typical TALE assembling reaction. Corresponding RVDs were chosen for specific nucleotide targets (NI for A, HD for C, NG for T, and NN for G). (c) Schematic illustration of the TALE library assembling. For each position, equal amounts of all four building modules were used, which results in an 11-mer TALE library covering all possible 11-mer DNA targets.

Results

Assembly and verification of the libraries. To assemble our TALE libraries, we modified the Golden Gate protocol (Fig. 1b)[5,26]. The enabling modification is the use of a predetermined mixture of the four building modules during the first step of the assembly (Fig. 1c). The flanking sequences of the four modules are identical. Therefore, during the enzyme-based digestion and ligase-based ligation reactions these modules have an equal probability in appearing in the final TALE construct (Methods and Supplementary Material). To test the library fidelity we subjected our 11-mer TALE mixture to standard Sanger sequencing using primers flanking the TALE DNA binding domain (Supplementary Table 1, P23 and P24). The sequencing results show that there are 6-nucleotide long repeats, spaced by 102 nucleotides, which have variable signature (Supplementary Fig. 1a). This general profile perfectly matches the TALE tandem repeat structure which contains 102 nucleotides and has RVDs which are 6-nucleotides in length. Importantly the composition of different peaks within these 6-nucleotide noisy elements closely tracks the expected composition. In particular, when equal amounts of all four possible RVDs are used for all positions, the first nucleotide position should consist of 75% A and 25% C (Supplementary Fig. 1b) and indeed the results have the expected sequencing profile (Supplementary Fig. 1c). Similarly, at position 2 of the RVD sequence, we should have only nucleotide A which was confirmed by sequencing (Supplementary Fig. 1c). The sequencing results confirm all positions.

The versatility of the proposed methodology of TALE library construction, perhaps, can be best illustrated by the simplicity in generating specific variants, namely sequence or nucleotide-biased TALE libraries (Supplementary Fig. 2). A sequence-biased TALE library has constrained nucleotides, and is assembled to bind to a specific transcription factor motif, while a nucleotide-biased TALE library has a predetermined ratio of the building modules resulting in desired nucleotide distributions (e.g. GC-rich).

As an example, we assembled a sequence-biased library to target a p53 motif. The p53-responsive element has been identified[27] as two copies of the decamer motif 5′-RRRCWWGYYY-3′ (R=A or G, W=A or T, Y=C or T) separated by 0 to 13 random nucleotides. Accordingly, we prepared an 11-mer TALE library which targets 5′-RRRCWWGYYYN-3′ (Methods and Supplementary Material). The validity of this library was again verified by Sanger sequencing, and the results closely match the desired RVD patterns (Supplementary Fig. 3).

TALE library-based yeast one-hybrid assay. We then explored the functional integrity of our complete 11-mer TALE library using a yeast one-hybrid assay (**Methods, Yeast One-hybrid Assay**). A VP16 activation domain was added to the 11-mer library to enhance the transactivation activities. For the assay, we cloned part of the 5′untranslated region (UTR) and the open reading frame (ORF) of the human SCN9A gene (Supplementary Table 2) in front of an antibiotic resistance gene, Aureobasidin A (Aba) in yeast (bait),
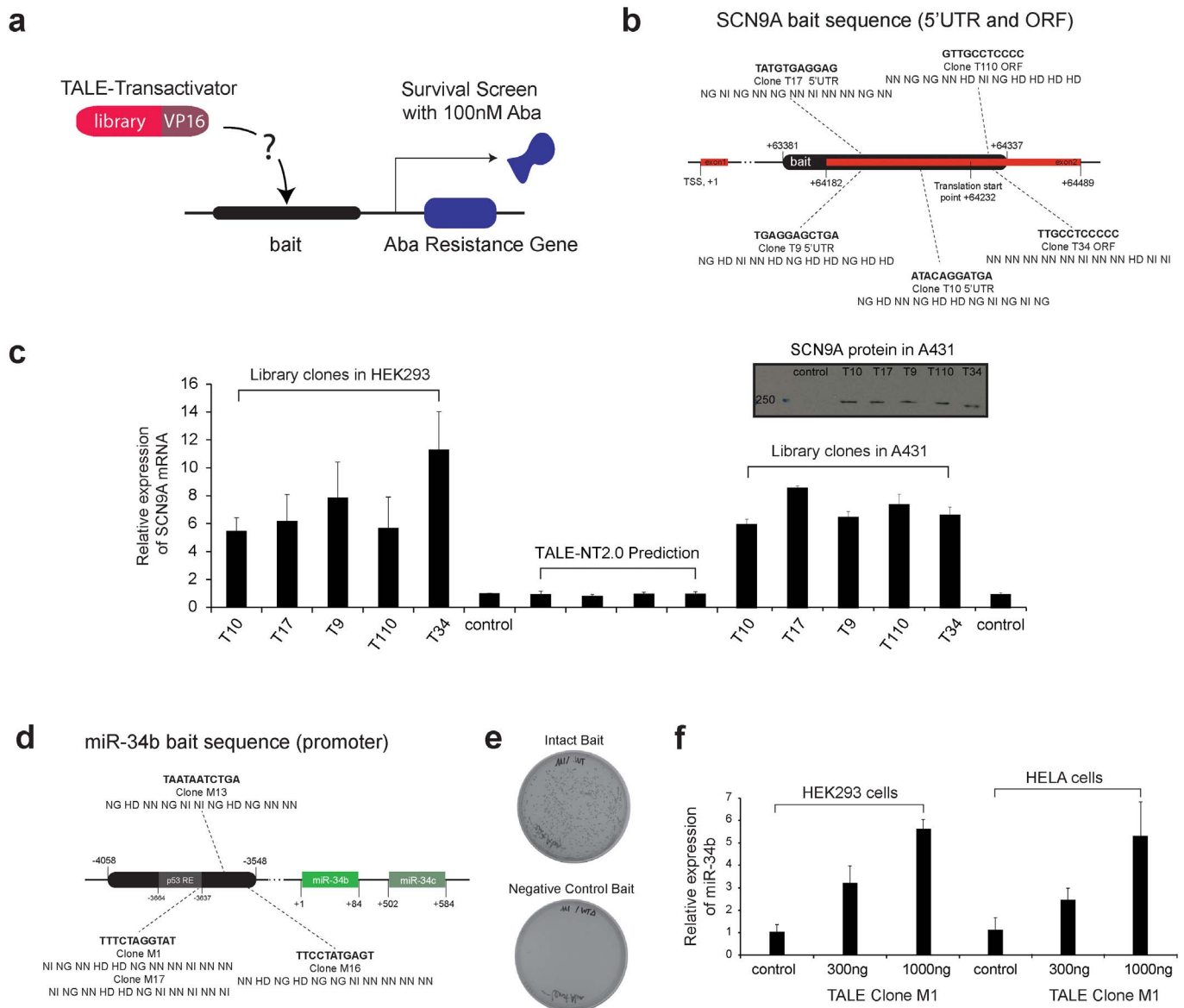
Figure 2 | Isolation of TALE-VP16 fusions targeting the human SCN9A gene using the 11-mer TALE-VP16 library and the yeast one-hybrid assay. (a) Schematic illustration of the yeast one-hybrid assay using the 11-mer TALE-VP16 library. A bait sequence was cloned in front of an antibiotic resistance gene (Aba resistance gene) in yeast. The 11-mer TALE-VP16 library was then transformed into this stable clone and a surviving assay was performed on -Leu plates containing 100 nM Aba. (b) The RVD sequences of isolated TALE-VP16 fusions and their targets within the human SCN9A bait sequence (scale not proportional). TALE-VP16 fusions were shown to bind to both the plus and the minus strands of the SCN9A bait sequence. (c) The isolated TALE-VP16 fusions induced overexpression of endogenous SCN9A in HEK293 cells and A431 cells. The mRNA levels of SCN9A were determined by quantitative RT-PCR. An empty vector (PEF-1) was used as the control. Columns 1-6: All TALE-VP16 fusions were able to effectively induce the overexpression of SCN9A in HEK293 cells (n=5). Columns 11-16: All TALE-VP16 fusions effectively induced the overexpression of SCN9A in A431 cells (n=3). Columns 7–10: All 4 TALEs designed according to TALE-NT 2.0 failed to induce the overexpression of SCN9A in HEK293 cells (n=3). Inlet: Western blot shows that all TALE-VP16 fusions induced the overexpression of SCN9A protein in A431 cells (representative data of two independent experiments). (d) The RVD sequences of isolated TALE-VP16 fusions and their targets within the human miR-34b/c bait sequence (scale not proportional). (e) Confirmation of the binding between isolated TALE-VP16 fusion M1 and its predicted gene target within the human miR-34b/c bait sequence. The isolated clone M1 was predicted to target 5′-TTTCTAGGTAT-3′ within the miR-34b/c bait sequence. The full-length bait sequence (pAbAi-miR-34b/c) or bait with the predicted target site deleted (pAbAi-miR-34b/c (ΔTTTCTAGGTAT)) was stably integrated into yeast cells. TALE-VP16 fusion clone M1 was then transformed into either cell line. Only cells which contained the intact bait sequence survived the 100 nM Aba selection. (f) The isolated TALE-VP16 fusion M1 effectively induced overexpression of miR-34b in a dose-dependent manner in both HEK293 and HeLa cells (n=3 for both cell lines).

and applied our library (prey) (Fig. 2a). We used 50 (15 cm diameter) plates, which resulted to approximately 1 million individual TALE clones (Supplementary Material, Yeast one-hybrid). For assays that require comprehensive library coverage the protocol can be scaled up by increasing the number of plates. The positive clones were confirmed by re-streaking on Aba-containing agar plates (Supple-

mentary Fig. 4). Thirteen of the TALE-VP16 expression plasmids were then rescued and sequenced to extract the RVD sequences (Supplementary Table 3).

Using TALE-NT 2.0 (TALE Effector Nucleotide Targeter 2.0)[28], we were able to confirm that these positive clones are predicted to bind to either the plus or minus strand of three specific locations

within the SCN9A bait sequence. Five of the TALE-VP16 clones are illustrated together with their corresponding position (Fig. 2b). To confirm these results, we first generated baits that exclude those predicted DNA target sites. For example, the positive clone T110 induced the expression of Aba resistance gene when the bait sequence is intact (Supplementary Fig. 5a) but it failed to produce surviving clones when the specific DNA target site (5′-GTTGC-CTCCCC-3′) was removed (Supplementary Fig. 5b). Subsequently, we cloned these TALE-VP16 fusions into a mammalian expression vector, transiently transfected the plasmids into human embryonic kidney cells (HEK293) and human squamous carcinoma cells (A431), and measured the expression level of SCN9A by quantitative RT-PCR (Methods, Quantitative reverse transcription-PCR). All fusions were able to effectively drive the overexpression of the SCN9A gene in both cell lines (up to 11-fold increase for HEK293 cells and 8.5-fold increase for A431 cells) (Fig. 2c).

In line with these results, the TALE-VP16 fusions also induced the overexpression of SCN9A protein in A431 cells (Fig. 2c, and Supplementary Fig. 6b). Interestingly, we were not able to detect SCN9A protein in HEK293 cell samples, including those transfected with the fusion constructs (Supplementary Fig. 6c). The discrepancy between the mRNA and protein measurements for HEK293 cells may be due to regulatory translational or post-translational mechanisms on the expression of SCN9A[29].

We also investigated whether the TALE fusions isolated from our screening were compatible with a virus-based delivery system for potential applications in primary cells. We prepared a TALE T34 recombinant adeno-associated viral vector using the AAV helper-free system (Agilent Technologies). This AAV viral stock was applied to HEK293 cells at a multiplicity of infection (MOI) of 400 and induced the overexpression of the SCN9A gene by 2 fold (Supplementary Fig. 7).

Furthermore, we tested the performance of our TALE-based yeast one-hybrid screening approach for microRNA gene targets. Specifically, we used part of the promoter sequence of the human miR-34b/c gene as the bait sequence (Supplementary Table 2) and isolated four positive clones, which were confirmed by re-streaking on Aba-containing plates (Supplementary Fig. 8). The RVD sequences of these four clones were extracted (Supplementary Table 3) and all are predicted to target the bait (Fig. 2d). Interestingly, two of the clones (M1, M17) target the same sequence within the miR-34b/c promoter (Fig. 2d). Again, we used two methods to confirm the results. We first generated a bait mutant which excludes the predicted target sequence for clone M1 (5′-TTTCTAGGTAT-3′) and prepared the according yeast stable cell line. As expected, in these yeast cells clone M1 failed to induce the expression of Aba resistance gene (Fig. 2e). Furthermore, the TALE_{M1}-VP16 cDNA was cloned into a mammalian expression vector and transiently transfected into HEK293 and HeLa cells. The TALE fusion effectively induced the overexpression of miR-34b in a dose-dependent manner (Fig. 2f).

Our results clearly demonstrate that the complete library can be used to isolate TALEs that bind to a desired DNA sequence. Importantly, we observed that TALE-VP16 can be designed to bind to the minus strand of a target sequence and in addition, can also be designed to target sequences within the ORF. We emphasize that one fundamental difference between our approach and all previous applications of TALEs is that our screening procedure does not rely on TALE-DNA binding prediction algorithms[28,30]. This is a considerable advantage as available algorithms may not yet fully capture the dynamics of TALE-DNA interactions. Indeed, we designed four 11-mer TALE-VP16 constructs for the same SCN9A bait sequence using the TALE-NT 2.0 program (Supplementary Table 3) and observed that all fusions, as well as combinations, failed to induce the overexpression of SCN9A mRNA levels in HEK293 (Fig. 2c, Supplementary Figs. 7 and 9). Additional insight can be gathered from recovered RVDs (Supplementary Table 3), as none of the iso-

lated positive TALE-VP16 clones contain RVD sequences exactly matching the canonical TALE-DNA binding code. As an example, for the target sequence 5′-GTTGCCTCCCC-3′, instead of the predicted canonical RVD sequence the four isolated clones contained at least one RVD difference. Therefore, the TALE-DNA binding information extracted from our library can be used in the future to refine the TALE-DNA binding algorithms.

TALE library-based benchmark genomic screen in yeast. We next applied our library to screen for TALE-VP16 fusion proteins which confer resistance to cycloheximide in yeast. The underlying mechanisms of cycloheximide resistance are well-studied. For example, in yeast *S. cerevisiae*, overexpression of ATP-binding cassette (ABC) transporters such as Pdr5p has been shown to contribute to cycloheximide resistance[31]. In addition, the expression of the PDR5 gene is known to be positively regulated by two homologous zinc finger-containing transcription regulators, Pdr1p and Pdr3p[32]. We first determined the lowest concentration cycloheximide (0.5 μg/ml) at which the wild-type yeast cells failed to grow (Supplementary Material). We applied the complete TALE-VP16 library and isolated eighteen positive clones which tolerate cycloheximide (e.g., selected clones illustrated in Supplementary Fig. 10). Subsequently, we isolated the TALE-VP16 fusion plasmids and re-transformed them back to the wild-type cells for confirmation, as in the original screen natural mutations of the yeast genome (both gain-of-function[33] and loss-of-function[34]) can artificially increase the resistance to cycloheximide. Five genuine positive clones were confirmed (Fig. 3a), isolated and sequenced to extract the TALE RVD sequences (Supplementary Table 3).

The sequencing results indicate that two clones (A8, A35) are predicted to target the promoter of the PDR3 gene and, in addition, A35 is also predicted to bind to the promoter of the PDR5 gene (Fig. 3a). Two methods were used to confirm the results. First, we cloned four copies of the predicted PDR3 and PDR5 promoter targets and their immediate adjacent sequences in front of a fluorescence reporter gene (mKate2) in yeast. We then transformed the yeast stable clones with plasmids carrying the TALEs. Indeed both clones induced the expression of mKate2 in cells containing the bait (Fig. 3b). Second, we prepared yeast cells transformed with TALE_{A8}-VP16, TALE_{A35}-VP16 or empty vector and measured the expression levels of PDR3 and PDR5 by quantitative RT-PCR. Again, both clones were able to effectively induce the expression of PDR3 and PDR5 (Fig. 3c).

Discussion

Here, we establish a methodology for the assembly of complete and biased TALE-based libraries. As proof-of-concept, we demonstrate that our methodology can be applied towards designing effective TALEs for any given DNA target (i.e. the yeast-one-hybrid experiments) as well as for genome-wide phenotype screens in yeast.

We show experimentally that the standard TALE design algorithm can fail to predict RVDs that bind to a desired target. Therefore, in the unforeseen cases where computational methods are not successful, the yeast-one-hybrid approach provides an algorithm-independent alternative towards designing TALEs.

A potential pitfall is that instead of the desired bait sequence a TALE may bind to the antibiotic resistance plasmid backbone sequence and induce the expression of the resistance gene. In our case, the negative control experiments (deletion of target) provide proof that the selected TALEs target exclusively the bait (Supplementary Fig. 5 and Fig. 2e). More generally, to avoid false-positives in the TALE-based yeast-one-hybrid assays, we recommend using bait sequences that are sufficiently long, ideally larger than 500 bp (Supplementary Material, Yeast-one-hybrid assay false-positives). We also presume that multiple copies of the targets can be used in cases when the target sequence is short.
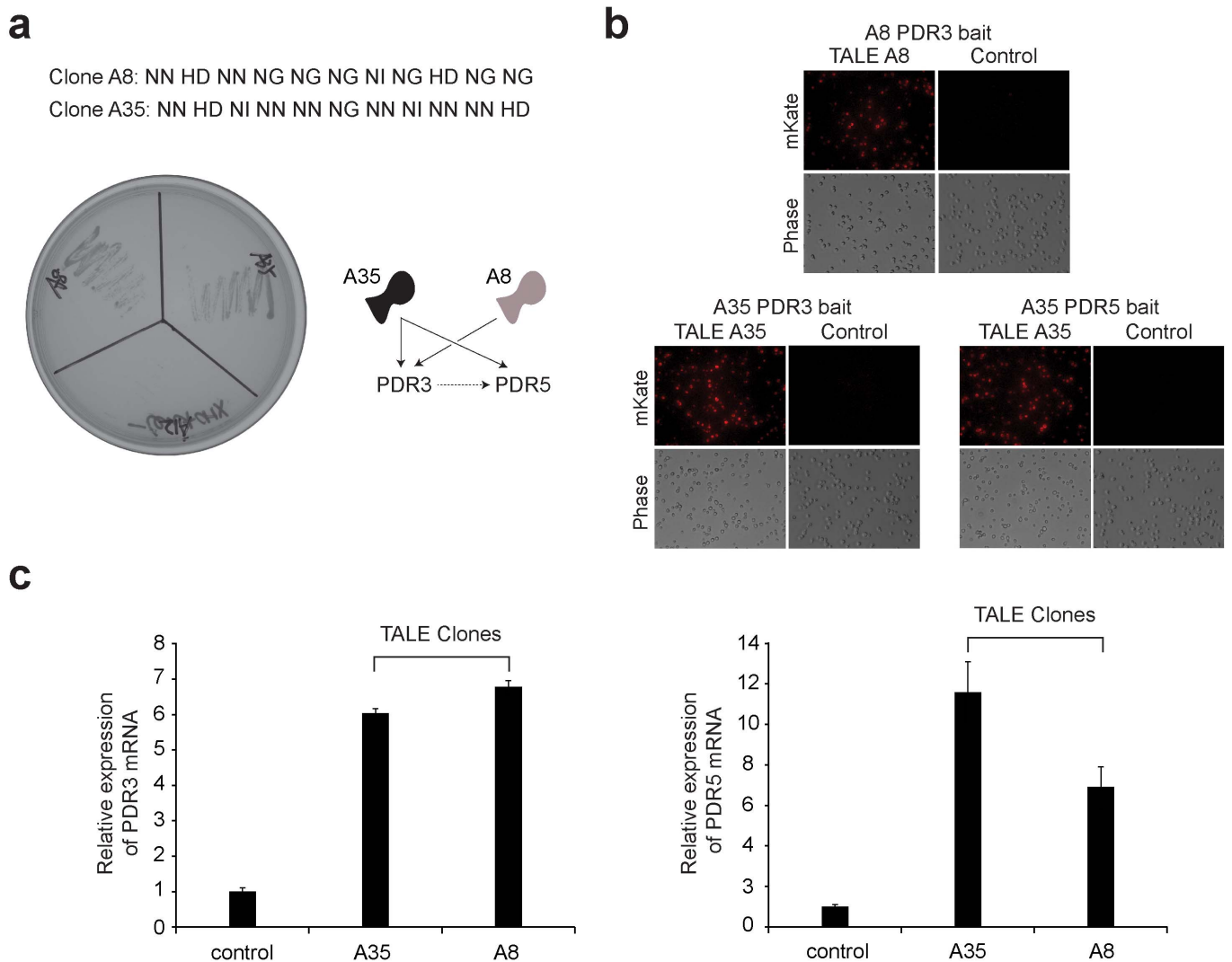
Figure 3 | Genetic screen for cycloheximide resistance in yeast using the TALE-VP16 plasmid library. (a) Confirmation of the genuine positive yeast clones conferring cycloheximide resistance. 18 positive clones were isolated from the cycloheximide resistance screening. Subsequently, these positive TALE-VP16 fusion plasmids were recovered and again re-transformed into the wild-type yeast cells. The transformed cells were then re-streaked onto –Leu plates containing 0.5 μg/ml of cycloheximide. After 3 days, cells transformed with genuine positive clones pGADT7-TALE-A8-VP16 and pGADT7-TALE-A35-VP16 were able to grow robustly. In contrast, cells transformed with the false positive clone pGADT7-TALE-A12-VP16 or the control pGADT7 failed to grow. (b) The isolated TALE-VP16 fusion clones A8 and A35 bind to the promoters of the PDR3 and PDR5 genes. TALE-VP16 fusion clones A8 and A35 were isolated from cycloheximide resistance screening. Both A8 and A35 were predicted to bind to the promoter of PDR3 gene. In addition, A35 was predicted to target the promoter of the PDR5 gene. Four copies of the predicted PDR3/PDR5 promoter targets and their immediate adjacent sequences were cloned in front of a fluorescence reporter gene (mKATE2) in yeast (bait). These yeast stable clones were then transformed with corresponding pGADT7-TALE-A8-VP16, pGADT7-TALE-A35-VP16 or pGADT7 (control). TALE-VP16 fusion clones A35 or A8 potently induced the expression of mKATE2 in yeast cells containing the corresponding baits, while pGADT7 failed to do so (right). (c) The isolated TALE-VP16 fusions A8 and A35 induced overexpression of endogenous PDR3 and PDR5 genes. Wild-type yeast cells were transformed with pGADT7-TALE-A8-VP16, pGADT7-TALE-A35-VP16 or pGADT7 (control). The expression levels of PDR3/PDR5 were measured by quantitative RT-PCR. Both clones were able to effectively induce the overexpression of PDR3 (n=3) and PDR5 (n=3).

An advantage of TALEs compared to other technologies (e.g., RNAi, CRISPR[35–37]) is that the size of the DNA target can be controlled. Increasing the DNA target size will result to superior specificity, albeit at the cost of increasing the library complexity and reducing the coverage in yeast screens. Therefore, a practical strategy towards eliminating off-target effects is to use a relatively short library (e.g., 11-mer) for screening purposes and a posteriori extend the DBDs of the screened TALEs to include flanking target sequences.

An alternative approach towards increasing the size of the targets of the library, without introducing additional experimental burden or increasing the complexity, involves utilizing the RVD properties.

For example, it is known that the RVD NN can target both A and G. Therefore, only three modules (i.e. NN, NG and HD) can be used for the assembly of the complete TALE libraries. As an example, a 14-mer TALE library with these three modules will result in a complexity of $3^{14}=4,782,969$. The assembly of larger TALE libraries can be experimentally cumbersome and in some cases practically prohibitive without the use of automated assembly systems.

One additional challenge lies in the fact that there is a certain degree of uncertainty in terms of deciphering the TALE-DNA binding given a RVD sequence, which may complicate the identification of potential target sequences in genome-wide assays. This issue can be resolved experimentally using ChIP-seq[38].

In summary, we have successfully established a methodology for the assembly of versatile TALE-based libraries and we demonstrated yeast-one-hybrid and genome-wide screens in yeast cells. We further emphasize that the TALE DBD can be fully customized to target any nucleotide content (e.g., GC-rich areas), nucleotide constraints (e.g., transcription factor motifs), or other modifications (e.g., methylations[39,40]). For the latter, during the construction of the TALE library, one additional RVD building block can be included in the assembly, namely N*. The star indicates that the amino acids at position 13 of TALE repeats are missing, which has been shown[39] to overcome the steric hindrance from the additional methyl moiety present at position 5 of thymine and thus allow efficient recognition of 5-methylated cytosine.

The modularity of the functional domain allows for a range of targeted manipulations including, but not limited to, transactivators, transrepressors, methylases, demethylases, recombinases, and nucleases[41]. More generally, with TALE libraries we have the ability to introduce multiple rounds of positive and negative screening, and even introduce a combination of positive and/or negative action to study synergistic effects. Finally, the TALE protein coding sequences can be easily isolated and sequenced, which can facilitate the rapid identification of their target genes. To conclude, considering the highly modular nature of TALEs, the ability to control their size and thus fine-tune their specificity, and the versatility and ease of constructing these libraries we foresee a broad range of applications in the future.

Methods

Library construction. A modified protocol based on the Golden Gate Assembly was used to construct the 11-mer TALE-VP16 library. First, a pFUS_A library containing all possible combinations of 10 tandem repeat (RVDs 1-10) and a pFUS_B library containing all combinations of 1 tandem repeat (RVD 11) were prepared as described in Cermak, T. et al.[5], except that for each position in a standard 30 µL reaction, 37.5 ng of each of the NN, NI, NG, and HD building blocks were mixed and included. Next, the pFUS_A and pFUS_B plasmid libraries were conjoined using the standard protocol. Similar procedures were used to prepare the p53-responsive element-targeting TALE-VP16 libraries. For each position of the Golden Gate Assembly reactions, the amounts of the four building blocks were adjusted according to the desired nucleotide target. All libraries were then subjected to standard Sanger sequencing using the forward primer P23 and the reverse primer P24 (Supplementary Table 1). The sequencing profiles were then analyzed using FinchTV (Geospiza). For full description, see Supplementary Material.

Yeast one-hybrid assay. The Matchmaker Gold Yeast One-Hybrid Screening System was used (Clontech, #630491). The TALE-VP16 library-based yeast one-hybrid assay was performed according to the protocol (Clontech, protocol #PT4087-1). The positive clones were screened by using SD/-Leu agar medium (Clontech, #630311) containing either 100 ng/ml Aureobasidin A (Clontech, #630466) or 0.5 µg/ml cycloheximide (Sigma-Aldrich, #C7698-1G). Subsequently, the positive clones were cultured in SD/-Leu broth (Clontech, #630310) and the pGADT7-TALE-VP16 plasmids were then prepared using the PrepEase Yeast Plasmid Isolation Kit (Affymetrix, #79220). The plasmids were then rescued into XL10-Gold Ultracompetent cells (Agilent, #200315) and prepared using the Qiagen Plasmid Mini Kit (Qiagen, #12123). For full description, see Supplementary Material.

Prediction of target sequences. All rescued pGADT7-TALE-VP16 plasmids were subjected to standard Sanger sequencing using the forward primer P23 and the reverse primer P24 (Supplementary Table 1), which are flanking sequences of the DNA binding domain of TALE proteins. After extracting the RVD sequences, the TALE potential target sequences were determined by the Target Finder function of TALE-NT 2.0[28] using the setting "upstream base T". Since TALE-NT 2.0 only supports sequences which are between 12 and 31 RVDs, for each TALE we conducted four searches by adding one additional RVD (HD or NN or NI or NG) to the end of our RVD sequences. The lowest score from one of these four searches was defined as "the lowest possible score". The targets were considered as potentially genuine only when the lowest possible score is less than 9, as lower scores indicate higher confidence of the bindings between TALE fusions and their predicted DNA target sequences.

Mammalian cell culture and transfection. All mammalian cells were maintained at 37°C, 100% humidity and 5% CO$_2$. HEK293, A431 and HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM, Invitrogen, #11965-1181) supplemented with 10% Fetal Bovine Serum (FBS, Atlanta Biologicals, #S11550), 0.1 mM MEM non-essential amino acids (Invitrogen, #11140-050), 0.045 units/mL of Penicillin and 0.045 units/mL of Streptomycin (Penicillin-Streptomycin liquid,

Invitrogen, #15140). For transient transfection, ~300 thousand cells in 1 mL of complete medium were plated into each well of 12-well culture treated plastic plates (Griener Bio-One, #665180) and grown for 16-20 hours. For jetPRIME (Polyplus Transfection, #114-15) transfection, up to 1 µg of the plasmid was added to 75 µL of jetPRIME buffer and 1.75 µL of jetPRIME. Transfection solutions were mixed and incubated at room temperature for 10 min. The transfection mixture was then applied to the cells and mixed with the medium by gently shaking.

Preparation of recombinant AAV vectors. The AAV helper-free system was purchased from Agilent Technologies (#240071). HEK293 cells were seeded on 10-cm cell culture treated petri dishes (Griener Bio-One, #664160). After reaching approximately 70–80% confluence, the cells were transfected with 3.3 µg of AAV-SCN9A-T34, 3.3 µg of pHELPER plasmid and 3.3 µg of pAAV-RC plasmid using 30 µL of jetPRIME. 72 hours post-transfection, the cells were scraped into the growth medium and pelleted by centrifugation at 1000 rpm for 5 min. The supernatant was then removed and the cell pellets were re-suspended in 1 mL of complete growth medium before being subjected to 4 rounds of freeze-thaw cycles between the dry ice-ethanol bath and the 37°C water bath. Subsequently, the cellular debris was collected by centrifugation at 10,000 × g for 10 min. The supernatant (primary recombinant AAV virus stock) was transferred into a fresh tube and stored at −80°C. The QuickTiter AAV Quantitation Kit (Cell Biolabs, #VPK-145-T) was used to determine the titers (genome copy/mL, GC/mL) of the recombinant AAV viral stocks. In brief, any free nucleic acids were first removed and the intact AAV viruses were captured using the QuickTiter AAV Capture Matrix. The captured viruses were then denatured to release the viral nucleic acids, which were subsequently labeled with fluorescence tags. The fluorescent intensities were measured using the FLUOstar Omega plate reader (BMG Labtech) with a 480/520 nm filter set.

Western blot assay. A431 or HEK293 cells were lysed in RIPA buffer (Sigma-Aldrich, #R0278) with protease inhibitors (Sigma-Aldrich, #P8340). Protein concentrations were measured using the Coomassie Plus Bradford assay kit (Thermo Scientific, #23236). For Western blot, 15 µg of whole cell lysates were mixed with loading buffer and boiled for 5 min, and then were run in 4–15% Mini-PROTEAN TGX polyacrylamide gel (Bio-Rad, #456-1083). After transferring, the nitrocellulose membranes were first blotted with 50 mM Tris, 150 mM NaCl and 0.1% Tween-20 (TBS-T) with 5% nonfat milk for 30 min at room temperature. The membranes were then incubated with either anti-SCN9A monoclonal antibody (1 : 500 dilution in TBS-T, Millipore, #MABN41) or anti-actin monoclonal antibody (1 : 5,000 dilution in TBS-T, Millipore, #MAB1501) at 4°C overnight. The membranes were then washed with TBS-T three times, each time for 10 min. Subsequently, the membranes were incubated with goat anti-mouse IgG (H+L) secondary antibody (1 : 10,000 dilution in TBS-T, Thermo Scientific, #31430) at room temperature for 45 min. The membranes were then washed with TBS-T four times, each time for 10 min. The membranes subsequently were incubated with ECL Western blotting substrates (Thermo Scientific, #32106) and then exposed to X-ray films (Sigma-Aldrich, #Z370398).

Quantitative reverse transcription-PCR (rtPCR). For measurement of mRNAs in mammalian cells, total RNA was extracted using the RNeasy Mini Kit (Qiagen, #74104) 48 hours post transfection. First strand synthesis was performed using the QuantiTect Reverse Transcription Kit (Qiagen, #205311). Quantitative PCR was performed using the KAPA SYBR FAST Universal qPCR Kit (KAPA Biosystems, #KK4601). GAPDH was used for normalization. For measurement of mRNAs in yeast cells, the same protocol was followed, except that total RNA was extracted using the MasterPure Yeast Purification Kit (Epicentre, #MPY80200), and yeast ACT1 was used for normalization. For measurement of miRNAs in mammalian cells, total RNA was extracted using the miRNeasy Micro Kit (Qiagen, #217084). First strand synthesis was performed using the miScript II RT Kit with HiSpec Buffer (Qiagen, #218161). Quantitative PCR was performed using miScript SYBR Green PCR Kit (Qiagen, #218073). RNU6 was used for normalization. The quantitative analysis was performed using the $2^{-\Delta\Delta Ct}$ method. The values of fold-change are reported as mean with standard deviation. For full description, see Supplementary Material.

Fluorescence microscopy. Bait yeast strains containing PDR3/PDR5 promoter target sequences were prepared as described in the yeast two-hybrid section. Subsequently, these yeast stable clones were transformed with corresponding pGADT7-TALE-VP16 plasmids and then grown on SD/-Leu agar medium. Individual positive clones (diameter 2–3 mm) were re-suspended in 1 mL of 0.9% (w/v) NaCl solution and transferred into 12-well culture treated plastic plates. Cells were allowed to fully settle and then imaged using an Olympus IX81 microscope in a Precision Control environmental chamber. The images were captured using a Hamamatsu ORCA-03 Cooled monochrome digital camera. The filter set (Chroma) for mKATE2 is ET560/40x (excitation) and ET630/75m (emission). Data collection and processing was performed using the software package Slidebook 5.0. All images were collected with the same magnification (40×) and exposure times and underwent identical processing.

1. Zhou, J. & Chai, J. Plant pathogenic bacterial type III effectors subdue host responses. *Curr Opin Microbiol* **11**, 179–185 (2008).
2. Boch, J. & Bonas, U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol* **48**, 419–436 (2010).

3. Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501–1501 (2009).

4. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).

5. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* **39**, e82 (2011).

6. Ding, Q. *et al.* A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* **12**, 238–51 (2013).

7. Cong, L., Zhou, R., Kuo, Y., Cunniff, M. & Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat Commun* **3**, 968 (2012).

8. Crocker, J. & Stern, D. L. TALE-mediated modulation of transcriptional enhancers in vivo. *Nat Methods* **10**, 762–7 (2013).

9. Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* **29**, 149–153 (2011).

10. Li, T., Liu, B., Spalding, M. H., Weeks, D. P. & Yang, B. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat Biotechnol* **30**, 390–392 (2012).

11. Kaelin, W. G. Use and Abuse of RNAi to Study Mammalian Gene Function. *Science* **337**, 421–422 (2012).

12. Zhang, X. D. Assessing the size of gene or RNAi effects in multifactor high-throughput experiments. *Pharmacogenomics* **11**, 199–213 (2010).

13. Cullen, B. Enhancing and confirming the specificity of RNAi experiments. *Nat Methods* **3**, 677–681 (2006).

14. Jackson, A. L. & Linsley, P. S. Noise amidst the silence: off-target effects of siRNAs? *Trends Genet* **20**, 521–524 (2004).

15. Esvelt, K. M. & Wang, H. H. Genome-scale engineering for systems and synthetic biology. *Mol Syst Biol* **9**, 641 (2013).

16. Mohr, S., Bakal, C. & Perrimon, N. Genomic Screening with RNAi: Results and Challenges. *Annu Rev Biochem* **79**, 37–64 (2010).

17. Root, D. E., Hacohen, N., Hahn, W. C., Lander, E. S. & Sabatini, D. M. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Methods* **3**, 715–719 (2006).

18. Jones, G. M. *et al.* A systematic library for comprehensive overexpression screens in Saccharomyces cerevisiae. *Nat Methods* **5**, 239–241 (2008).

19. Beerli, R. R., Dreier, B. & Barbas, C. F. Positive and negative regulation of endogenous genes by designed transcription factors. *Proc Natl Acad Sci USA* **97**, 1495–1500 (2000).

20. Gonzalez, B. *et al.* Modular system for the construction of zinc-finger libraries and proteins. *Nat Protoc* **5**, 791–810 (2010).

21. Blancafort, P., Magnenat, L. & Barbas, C. F. Scanning the human genome with combinatorial transcription factor libraries. *Nat Biotechnol* **21**, 269–274 (2003).

22. Ute, K. *et al.* Insights into Effective RNAi Gained from Large-Scale siRNA Validation Screening. *Oligonucleotides* **17**, 237–250 (2007).

23. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* **343**, 80–84 (2014).

24. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* **343**, 84–87 (2014).

25. Kim, Y. *et al.* A library of TAL effector nucleases spanning the human genome. *Nat Biotechnol* **31**, 251–258 (2013).

26. Li, Y., Moore, R., Guinn, M. & Bleris, L. Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. *Sci Rep* **2**, 897 (2012).

27. Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature* **408**, 307–310 (2000).

28. Doyle, E. L. *et al.* TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res* **40**, W117–W122 (2012).

29. Borghouts, C. *et al.* The Intracellular Delivery of a Recombinant Peptide Derived from the Acidic Domain of PIAS3 Inhibits STAT3 Transactivation and Induces Tumor Cell Death. *Mol Cancer Res* **8**, 539–553 (2010).

30. Grau, J. *et al.* Computational Predictions Provide Insights into the Biology of TAL Effector Target Sites. *PLoS Comput Biol* **9**, e1002962 (2013).

31. Balzi, E., Wang, M., Leterme, S., Van Dyck, L. & Goffeau, A. PDR5, a novel yeast multidrug resistance conferring transporter controlled by the transcription regulator PDR1. *J Biol Chem* **269**, 2206–2214 (1994).

32. Katzmann, D. J., Burnett, P. E., Golin, J., Mahe, Y. & Moye-Rowley, W. S. Transcriptional control of the yeast PDR5 gene by the PDR3 gene product. *Mol Cell Biol* **14**, 4653–4661 (1994).

33. Nourani, A., Papajova, D., Delahodde, A., Jacq, C. & Subik, J. Clustered amino acid substitutions in the yeast transcription regulator Pdr3p increase pleiotropic drug resistance and identify a new central regulatory domain. *Mol Gen Genet* **256**, 397–405 (1997).

34. Gerlinger, U., Gückel, R., Hoffmann, M., Wolf, D. H. & Hilt, W. Yeast cycloheximide-resistant crl mutants are proteasome mutants defective in protein degradation. *Mol Biol Cell* **8**, 2487–2499 (1997).

35. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822–826 (2013).

36. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).

37. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).

38. Furey, T. S. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat Rev Genet* **13**, 840–852 (2012).

39. Valton, J. *et al.* Overcoming Transcription Activator-like Effector (TALE) DNA Binding Domain Sensitivity to Cytosine Methylation. *J Biol Chem* **287**, 38427–38432 (2012).

40. Deng, D. *et al.* Recognition of methylated DNA by TAL effectors. *Cell Res* **22**, 1502–1504 (2012).

41. Yin, P. *et al.* Specific DNA-RNA Hybrid Recognition by TAL Effectors. *Cell Rep* **2**, 707–713 (2012).

Acknowledgments

Author contributions

Y.L. and L.B. conceived the project and designed the experiments. Y.L. performed the majority of the experiments. K.E. performed TALE-NT2.0 cloning and viral experiments. Y.L. and L.B. analyzed the data. Y.L., K.E., M.Z. and L.B. prepared the manuscript. L.B. supervised the project.

Additional information

Supplementary Information

Assembly and Validation of Versatile Transcription Activator-Like Effector Libraries

Yi Li[1,2], Kristina Ehrhardt[1,2], Michael Q. Zhang[2,3], and Leonidas Bleris[1,2,4]

[1]Bioengineering Department, The University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA

[2]Center for Systems Biology, The University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA

[3]Molecular and Cell Biology Department, The University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA

[4]Electrical Engineering Department, The University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA

[*]Corresponding author: Leonidas Bleris, The University of Texas at Dallas, NSERL 4.708, 800 West Campbell Road, Richardson TX 75080 USA

Email: bleris@utdallas.edu

Phone: 972-883-5785

Supplementary Methods

Library construction

A modified protocol based on the Golden Gate Assembly (Addgene, #1000000016) was used to construct the 11-mer TALE-VP16 library[1,2].

Part 1: Construction of a pFUS_A library containing all possible combinations of 10 tandem repeats (RVDs 1-10)

Step 1 (Day 1, 3 h)| The Golden Gate Assembling reactions were performed in strip PCR tubes (8 well, Applied Biosystems, #N801-0560). In a standard reaction, 37.5 ng of each of the 40 RVD building blocks (NN1 to NN10, NI1 to NI10, NG1 to NG10, and HD1 to HD10) were included. Then, 150 ng of pFUS_A plasmid, 1 µL of BsaI (New England Biolabs, #R0535S), 1 µL of BSA (3 mg/mL, New England Biolabs, #B9000S), 1 µL of T4 DNA Ligase (New England Biolabs, #M0202S), 3 µL of T4 DNA Ligase Reaction buffer (New England Biolabs, #B0202S) were added. Nuclease free water (Life Technologies, #10977) was then added to bring to 30 µL of total reaction volume.

Step 2 (Day 1, 10 min)| The mixture was mixed and briefly spun down.

Step 3 (Day 1-Day 2, 13 h)| The mixture was then incubated on a Thermal Cycler (Bio-Rad, #S1000). The thermal cycling conditions were 50 cycles of 5 min at 37°C and 10 min at 16°C, followed by 1 cycle of 5 min at 50°C, and 1 cycle of 5 min at 80°C. The number of digestion/ligation reaction cycles was increased to 50 to ensure the highest possible reaction efficiency.

Step 4 (Day 2, 1 hr)| Subsequently, the reaction mixture (30 µL) was treated with 1 µL of Plasmid-Safe ATP-Dependent DNase (Epicentre, #E3110K) and 1 µL of 10 mM ATP (New England Biolabs, #P0756S) at 37°C for 1 hour.

Step 5 (Day 2, 4 h)| 6 transformations were then performed. For each transformation, 5 µL of the reaction mixture was transformed into 100 µL of the XL10-Gold Ultracompetent cells (Agilent, #200314). The resulting 1 mL of cell suspension was then plated onto 2 10cm LB agar plates with Carbenicillin (100 µg/mL, Teknova, #C2113). In total, 12 plates were used and a lawn, which indicates the cells grew to full

confluency, was observed on all plates. This ensures that at least 1,048,576 independent clones for the pFUS_A library were collected.

Step 6 (Day 3, 6 h)| The cells were harvested after 16 hours and the plasmids were then prepared using the Qiagen Plasmid Midi Kit (Qiagen, #12143).

Part 2: Construction of a pFUS_B library containing all possible combinations of 1 tandem repeat (RVD 1)

Step 1 (Day 1, 3 h)| The Golden Gate Assembling reactions were performed in strip PCR tubes (8 well, Applied Biosystems, #N801-0560). In a standard reaction, 37.5 ng of each of the 4 RVD building blocks (NN1, NI1, NG1 and HD1) were included. Then, 150 ng of pFUS_B plasmid, 1 µL of BsaI (New England Biolabs, #R0535S), 1 µL of BSA (2 mg/mL, New England Biolabs, #B9000S), 1 µL of T4 DNA Ligase (New England Biolabs, #M0202S), 2 µL of T4 DNA Ligase Reaction buffer (New England Biolabs, #B0202S) were added. Nuclease free water (Life Technologies, #10977) was then added to bring to 20 µL of total reaction volume.

Step 2 (Day 1, 10 min)| The mixtures were mixed and briefly spun down.

Step 3 (Day 1-Day 2, 13 h)| The mixtures were then incubated on a Thermal Cycler (Bio-Rad, #S1000). The thermal cycling conditions were 50 cycles of 5 min at 37°C and 10 min at 16°C, followed by 1 cycle of 5 min at 50°C, and 1 cycle of 5 min at 80°C.

Step 4 (Day 2, 1 hr)| Subsequently, the reaction mixture (30 µL) was treated with 1 µL of Plasmid-Safe ATP-Dependent DNase (Epicentre, #E3110K) and 1 µL of 10 mM ATP (New England Biolabs, #P0756S) at 37°C for 1 hour.

Step 5 (Day 2, 4 h)| One transformation was then performed. 5 µL of the reaction mixture was transformed into 100 µL of the XL10-Gold Ultracompetent cells (Agilent, #200314). The resulting 1 mL of cell suspension was then plated onto 2 10cm LB agar plates with Carbenicillin (100 µg/mL, Teknova, #C2113). In total, 2 plates were used and a lawn, which indicates the cells grew to full confluency, was observed on both plates. This ensures that at least 4 independent clones for the pFUS_B library were collected.

Step 6 (Day 3, 6 h)| The cells were harvested after 16 hours and the plasmids were then prepared using the Qiagen Plasmid Midi Kit (Qiagen, #12143).

Part 3: Construction of a TALE-VP16 library containing all possible combinations of 11 tandem repeats (RVDs 1-11)

The pFUS_A (RVDs 1-10) and pFUS_B (RVD 1) plasmid libraries were conjoined.

Step 1 (Day 4, 3 h)| Specifically, in a standard 20 µL reaction, 150 ng of pFUS_A (RVDs 1-10) plasmid library, 150 of pFUS_B (RVD 1) plasmid library, 150 ng of pLR-HD vector (the last "half-repeat"), and 75 ng of PTAL1-VP16 were included. Then 1 µL of Esp3I (Thermo Scientific, #ER0452), 1 µL of T4 DNA Ligase (New England Biolabs, #M0202S), 2 µL of T4 DNA Ligase Reaction buffer (New England Biolabs, #B0202S) were added. Nuclease free water (Life Technologies, #10977) was then added up to 20 µL of total reaction volume. Four such reactions were set up.

Step 2 (Day 4, 10 min)| The mixtures were mixed and briefly spun down.

Step 3 (Day 4-Day 5, 13 h)| The mixtures were then incubated on a Thermal Cycler (Bio-Rad, #S1000). The thermal cycling conditions were 50 cycles of 5 min at 37°C and 10 min at 16°C, followed by 1 cycle of 15 min at 37°C, and 1 cycle of 5 min at 80°C. Similar to before, the number of digestion/ligation reaction cycles was increased to 50 to ensure the highest possible reaction efficiency.

Step 4 (Day 5, 1 hr)| The reaction mixtures were then treated with 1 µL of KpnI (New England Biolabs, #R0142S) at 37°C for 1 hour to linearize the unused PTAL1-VP16 plasmids.

Step 5 (Day 5, 4 h)| 20 transformations were then performed. For each transformation, 4 µL of the reaction mixture was transformed into 100 µL of the XL10-Gold Ultracompetent cells (Agilent, #200314). The resulting 1 mL of cell suspension was then plated onto 2 10cm LB agar plates with Carbenicillin (100 µg/mL, Teknova, #C2113). In total, 40 plates were used and a lawn, which indicates the cells grew to full confluency, was observed on all plates. This ensures that at least 4,194,304 independent clones for the 11-mer TALE-VP16 library were collected.

Step 6 (Day 6, 6 h)| The cells were harvested after 16 hours and the plasmids were then prepared using the Qiagen Plasmid Midi Kit (Qiagen, #12143).

Step 7 (Day 7)| The libraries were then subjected to standard Sanger sequencing using the forward primer (P23, 5'-CGTCGCTGTCACGTATCAGCACATAATC-3') and the reverse primer (P24, 5'-GCGTACGCTGGGTGGGAGTGGCAC-3'). The sequencing profiles were then analyzed using FinchTV (Geospiza).

Step 8 (Day 7, 2 h)| The pFUS_A (RVDs 1-10), pFUS_B (RVD 1), and the final 11-mer TALE-VP16 libraries were also subjected to 0.7% agarose gel electrophoresis along with the original pFUS_A, pFUS_B and PTAL1-VP16 mother vectors. The electrophoresis image showed that the libraries were free of noticeable contamination from their respective mother vectors (**Supplementary Fig. 11**, from left to right: 2-Log DNA ladder (New England Biolabs, #N3200S), pFUS_A (RVDs 1-10), pFUS_A, pFUS_B (RVD 1), pFUS_B, 2-Log DNA ladder, 11-mer TALE-VP16 library, and PTAL1-VP16).

Part 4: Construction of a biased TALE-VP16 library for the p53-responsive element

Similar procedures (Steps 1-3) were used to prepare for the p53-responsive element-targeting TALE-VP16 library, which targets 5'-RRRCWWGYYYN-3'. For each position of the Golden Gate Assembly reactions, the amounts of the 4 building blocks were adjusted according to the desired nucleotide target. Specifically, 37.5 ng of each of the four building blocks were mixed when the nucleotide target is N. 75 ng of HD and NG were mixed when the target is Y. 75 ng of NI and NG were mixed when the target is W. 75 ng of NI and NN were mixed when the target is R. For the final assembling reactions, at least 1,024 independent clones were harvested for preparation of the plasmid libraries. All libraries were then subjected to standard Sanger sequencing using the forward primer (P23, 5'-CGTCGCTGTCACGTATCAGCACATAATC-3') and the reverse primer (P24, 5'-GCGTACGCTGGGTGGGAGTGGCAC-3'). The sequencing profiles were then analyzed using FinchTV (Geospiza).

Yeast one-hybrid assay

The Matchmaker Gold Yeast One-Hybrid Screening System was purchased from Clontech (# 630491).

The TALE-VP16 library-based yeast one-hybrid assay was performed according to the manufacturer's

protocol (Clontech, #PT4087-1). To prepare for the bait yeast strain, the bait plasmids were linearized

using BstBI (New England Biolabs, #R0519S) and then transformed into Y1HGold yeast cells using

Yeastmaker Yeast Transformation System 2 (Clontech, #630439), closely following the manufacturer's

protocol (Clontech, #PT1172-1). The yeast bait stable clones were then selected by SD/-Ura agar medium

(Clontech, # 630315). The 11-mer TALE-VP16 cDNA library was PCR amplified using the 11-mer

TALE-VP16 plasmid library as the template and primers P27 and P28 (**Supplementary Table 1**). The

VP16 domain was fused with the TALE DBD (DNA binding domain) in addition to the GAL4 activation

domain within the pGADT7 destination plasmid to further enhance the transactivation activities. The 11-

mer TALE-VP16 cDNA library was transformed into the bait yeast cells using the Yeastmaker Yeast

Transformation System 2. In each experiment, ~1-1.2 x 10^6 individual clones were plated onto 50 150-

mm agar plates. Briefly, the library-scale transformation protocol was used, which resulted in 15 mL of

the transformation solution in 0.9% (w/v) NaCl solution. To calculate the number of total individual

clones, 10 µL of the transformation solution was diluted into 990 µL of the 0.9% (w/v) NaCl solution.

Subsequently, 100 µL of this diluent was plated onto one 10cm SD/-Leu agar plate. In a representative

experiment, 72 clones were recovered from this control plate, which indicates that the total individual

clone number was: 72 x (15 mL/1 µL) = 1.08 x 10^6. The positive clones were then selected by using SD/-

Leu agar medium (Clontech, #630311) containing either 100 ng/ml Aureobasidin A (Clontech, #630466)

or 0.5 µg/ml cycloheximide (Sigma-Aldrich, #C7698-1G). Subsequently, the positive clones were

cultured in SD/-Leu broth (Clontech, #630310) and the pGADT7-TALE-VP16 plasmids were then

prepared using the PrepEase Yeast Plasmid Isolation Kit (Affymetrix, #79220). The plasmids were then

rescued into XL10-Gold Ultracompetent cells (Agilent, #200314) and prepared using the Qiagen Plasmid

Mini Kit (Qiagen, #12123).

Yeast one-hybrid assay false-positives: To investigate the false-positive issue we used TALE-NT 2.0 to determine all possible target sites within the pAbAi backbone for the TALE fusions isolated from our experiments (13 for SCN9A, and 4 for miR-34b). Of the 68 possible TALE RVDs (the original 11-mer plus one additional RVD being NI, NN, NG or HD), only 9 of them are predicted to have one additional target within the pAbAi backbone (**Supplementary Table 4**). Importantly, all these targets are significantly far from the translation start site (ATG) of the Aba resistance gene (larger than 1 kb) and naturally have no/minimal effect on its expression activity (the average length of promoters in *s. cerevisiae* is ~455bp[3]).

Quantitative reverse transcription-PCR

In mammalian cells, 48 hours post transfection, total RNA was extracted using the RNeasy Mini Kit (Qiagen, #74104). First strand synthesis was performed using the QuantiTect Reverse Transcription Kit (Qiagen, #205311). Quantitative PCR was performed using the KAPA SYBR FAST Universal qPCR Kit (KAPA Biosystems, #KK4601, KAPA Biosystems). GAPDH was used for normalization. The forward primer for GAPDH was 5'-AATCCCATCACCATCTTCCA-3', and the reverse primer for GAPDH was 5'-TGGACTCCACGACGTACTCA-3'. The forward primer for SCN9A was 5'-CGTGGACAAACACTTGATGG-3', and the reverse primer for SCN9A was 5'-GTCACTGTGAGGCTGGGATT-3'. The forward primer for VP16 was 5'-CTCCACTTAGACGGCGAGGA-3', and the reverse primer for VP16 was 5'-GAAGTCGGCCGTATCCAGAG-3'.The thermal cycling conditions were 3 min at 95°C followed by 40 cycles of denaturation for 15 s at 95°C and annealing for 30 s at 60°C. The relative expression levels of SCN9A were determined by $2^{-\Delta\Delta Ct}$ analysis. The values of fold-change are reported as mean with standard deviation.

In yeast, the yeast transformants containing pGADT7-TALE-A8-VP16, pGADT7-TALE-A35-VP16 or pGADT7 were grown in SD-Leu broth (Clontech, #630310) until reaching stationary phase. Total RNA was extracted using the MasterPure Yeast Purification Kit (Epicentre, #MPY80200). First

strand synthesis was performed using the QuantiTect Reverse Transcription Kit (Qiagen, #205311). Quantitative PCR was performed using the KAPA SYBR FAST Universal qPCR Kit (KAPA Biosystems, #KK4601). Yeast ACT1 was used for normalization. The forward primer for ACT1 was 5'-TTGGCCGGTAGAGATTTGACTGAC-3', and the reverse primer for ACT1 was 5'-AGCGGTTTGCATTTCTTGTTCG-3'. The forward primer for PDR3 was 5'-ATACTGCCGAACGGAGAAGA-3', and the reverse primer for PDR3 was 5'-CTGAAATCCTTCGGCAAGAG-3'. The forward primer for PDR5 was 5'-TGACGCTTTTGCATCAGTTC-3', and the reverse primer for PDR5 was 5'-GAGAAAACCGCGACAATGTT-3'. The thermal cycling conditions were 3 min at 95°C followed by 40 cycles of denaturation for 15 s at 95°C and annealing for 30 s at 60°C. The relative expression levels of PDR3 and PDR5 were determined by $2^{-\Delta\Delta Ct}$ analysis. The values of fold-change are reported as mean with standard deviation.

In mammalian cells, 48 hours post transfection, total RNA was extracted using the miRNeasy Micro Kit (Qiagen, #217084). First strand synthesis was performed using the miScript II RT Kit with HiSpec Buffer (Qiagen, #218161). Quantitative PCR was performed using the miScript SYBR Green PCR Kit (Qiagen, #218073). RNU6 was used for normalization. The miScript Primer Assays for Hs_RNU6 (Qiagen, # MS00033740) and Hs_miR-34b (Qiagen, #MS00008190) were used. The thermal cycling conditions were 15 min at 95°C followed by 40 cycles of denaturation for 15 s at 94°C, annealing for 30 s at 55°C, and extension for 30 s at 70°C. The relative expression levels of miR-34b were determined by $2^{-\Delta\Delta Ct}$ analysis. The values of fold-change are reported as mean with standard deviation.

Molecular cloning

All restriction enzymes were purchased from New England Biolabs. All other reagents, unless specified, were purchased from Sigma-Aldrich.

pAbAi-SCN9A: Total DNA was isolated from HEK293 cells using the DNeasy Blood & Tissue Kit (Qiagen, #69504). Part of the 5'-UTR and ORF of the human SCN9A gene was amplified by using the HEK293 cells' total DNA as the template and primers P1 and P2, and then cloned into pAbAi (Clontech, #630491) using HindIII and XhoI.

pAbAi-SCN9A(ΔATACAGGATGA): A mutagenesis PCR reaction was performed using pAbAi-SCN9A as the template and primers P3 and P4. The PCR product was digested with SacII and then ligated to produce pAbAi-SCN9A(ΔATACAGGATGA).

pAbAi-SCN9A(ΔTATGTGAGGAGCTGA): A mutagenesis PCR reaction was performed using pAbAi-SCN9A as the template and primers P5 and P6. The PCR product was digested with SacII and then ligated to produce pAbAi-SCN9A(ΔTATGTGAGGAGCTGA).

pAbAi-SCN9A(ΔGTTGCCTCCCCC): A mutagenesis PCR reaction was performed using pAbAi-SCN9A as the template and primers P7 and P8. The PCR product was digested with SacII and then ligated to produce pAbAi-SCN9A(ΔGTTGCCTCCCCC).

PEF-1-SCN9A-T10/T17/T9/T110/T34-VP16: SCN9A-TALET10/T17/T9/T110/T34-VP16 were PCR amplified from pGADT7-SCN9A-T10/T17/T9/T110/T34-VP16 using primers P9 and P10, and then were cloned into PEF-1 using EcoRI.

PEF-1-SCN9A-T10/T17/T9/T110/T34-KRAB: The KRAB domain was extracted from PEF-1-TALETRE#3-KRAB using XbaI [14]. PEF-1-SCN9A-T10/T17/T9/T110/T34-VP16 were digested with XbaI and dephosphorylated with CIP (Calf Intestinal Alkaline Phosphatase, New England Biolabs, #M0290S), and then ligated with the KRAB domain to produce PEF-1-SCN9A-T10/T17/T9/T110/T34-KRAB.

pAbAi-miR-34b/c: Total DNA was isolated from HEK293 cells using the DNeasy Blood & Tissue Kit (Qiagen, #69504). Part of the promoter of human miR-34b/c gene was amplified by using the HEK293

cells' total DNA as the template and primers P11 and P12, and then cloned into pAbAi using HindIII and XhoI.

pAbAi-miR-34b/c(ΔTTTCTAGGTAT): A mutagenesis PCR reaction was performed using pAbAi-miR-34b/c as the template and primers P13 and P14. The PCR product was digested with SacII and then ligated to produce pAbAi-miR-34b/c(ΔTTTCTAGGTAT).

PEF-1-miR-34b/c-TALE-M1-VP16: miR-34b/c-TALE-M1-VP16 was PCR amplified from pGADT7-miR-34b/c-TALE-M1-VP16 using primers P15 and P16, and then was cloned into PEF-1 using EcoRI and KpnI.

pAbAi-mKATE2: MKATE2 was PCR amplified from PCMV-mKATE2 (Evrogen, #FP181) using primers P17 and P18. pAbAi was PCR amplified using primers P25 and P26. These two PCR products were digested with NheI and NotI and then ligated to produce pAbAi-mKATE2.

pAbAi-4XA35/PDR5-mKATE2: 4XA35/PDR5-mKATE2 was PCR amplified from pAbAi-mKATE2 using primers P19 and P20, and then cloned into pAbAi-mKATE2 using HindIII and NheI.

pAbAi-4XA35/PDR3-mKATE2: 4XA35/PDR3-mKATE2 was PCR amplified from pAbAi-mKATE2 using primers P21 and P20, and then cloned into pAbAi-mKATE2 using HindIII and NheI.

pAbAi-4XA8/PDR3-mKATE2: 4XA8/PDR3-mKATE2 was PCR amplified from pAbAi-mKATE2 using primers P22 and P20, and then cloned into pAbAi-mKATE2 using HindIII and NheI.

AAV-SCN9A-T34: The SCN9A-T34 TALE-VP16 cDNA transcript was PCR amplified from PEF-1-SCN9A-T34-VP16 using primers P31 and P32, and then cloned into AAV-MCS using ClaI and HindIII.
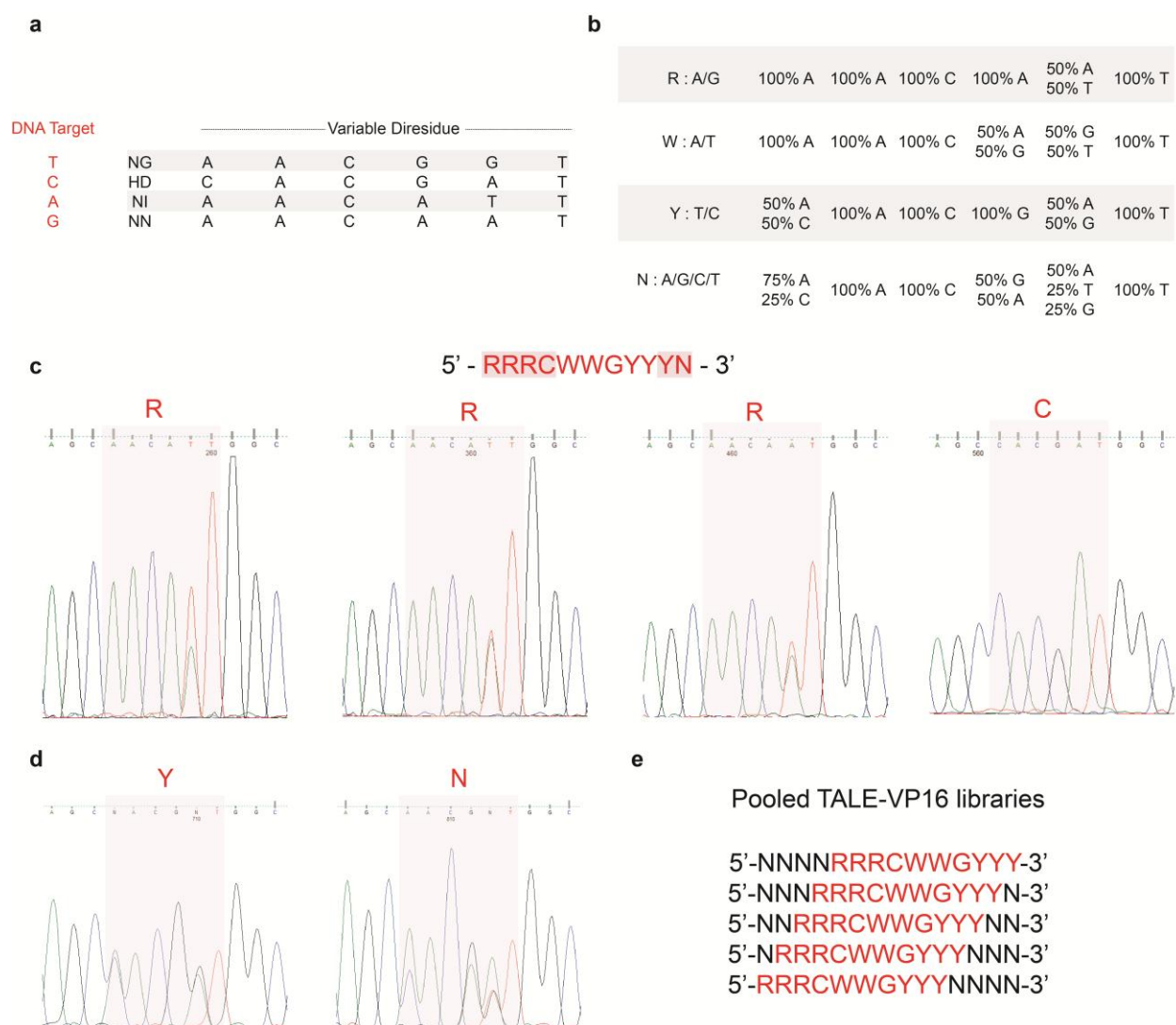
Supplementary Figures
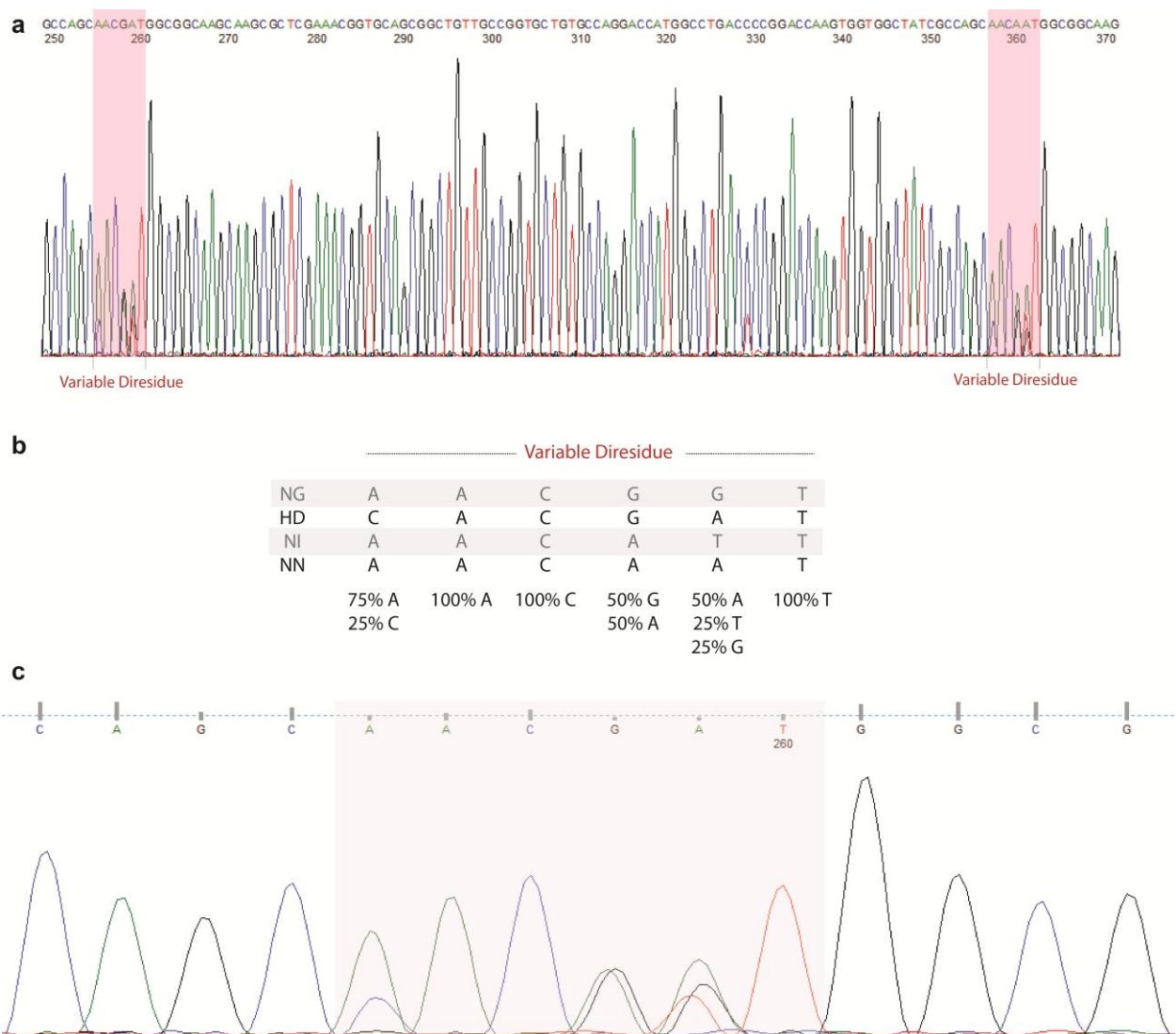
Supplementary Figure 1: Sanger sequencing profiles of the 11-mer TALE library. (a) The 11-mer TALE library was subjected to Sanger sequencing using the forward primer (P23, 5'-CGTCGCTGTCACGTATCAGCACATAATC-3') and the reverse primer (P24, 5'-GCGTACGCTGGGTGGGAGTGGCAC-3'). There are 6-nucleotide long repeats (RVDs), spaced by 102 nucleotides, showing "noisy" signals. (b) Expected nucleotide compositions of the RVD domain of our TALE library. (c) Observed nucleotide compositions of the RVD domain of our TALE library, which closely tracks the predicted composition.
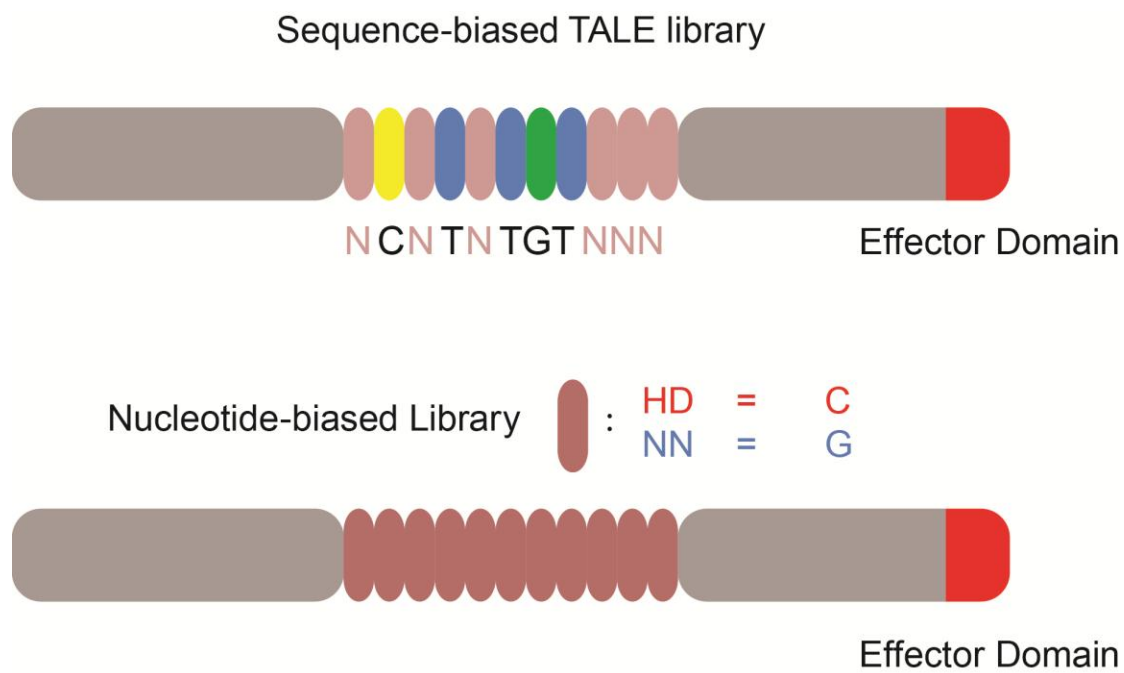
Supplementary Figure 2: Schematic illustration of sequence or nucleotide-biased TALE libraries. A sequence-biased TALE library has constrained nucleotides, while a nucleotide-biased library could be GC-rich.

Supplementary Figure 3: Sanger sequencing profiles of the p53-biased 11-mer TALE library. (a) The nucleotide compositions of the variable diresidues (RVDs) for four target nucleotides (T, C, A, G). (b) The expected nucleotide compositions of the RVD domains for various target nucleotides (R, W, Y, N). (c) Observed nucleotide compositions of the first four RVD domains of the p53-biased 11-mer TALE library using the forward primer (P23), which closely tracks the predicted composition. (d) Observed nucleotide compositions of the last two RVD domains of the p53-biased 11-mer TALE library using the reverse primer (P24), which closely tracks the predicted composition. (e) Since TALE binding sites are

preferentially preceded by a T, five 14-mer TALE-VP16 libraries which target 5'-NNNNRRRCWWGYYY-3', 5'-NNNRRRCWWGYYYN-3', 5'-NNRRRCWWGYYYNN-3', 5'-NRRRCWWGYYYNNN-3', and 5'-RRRCWWGYYYNNNN-3' can be prepared separately. Pooling these five libraries is predicted to cover at least $1-(0.75)^5 = 75\%$ of all possible 14-mer DNA target sequences which contain a p53-responsive element and are preceded by a T.

**Supplementary Figure 4: Secondary confirmation of the positive yeast clones for the SCN9A bait.**
The positive clones isolated from the yeast one-hybrid screening were re-streaked onto –Leu plates containing 100 nM Aureobasidin A. After 3 days, the TALE-VP16 positive clones (#109 and #110) showed robust growth, while the control (0, yeast cells which are transformed with empty prey vector pGADT7) failed to grow.
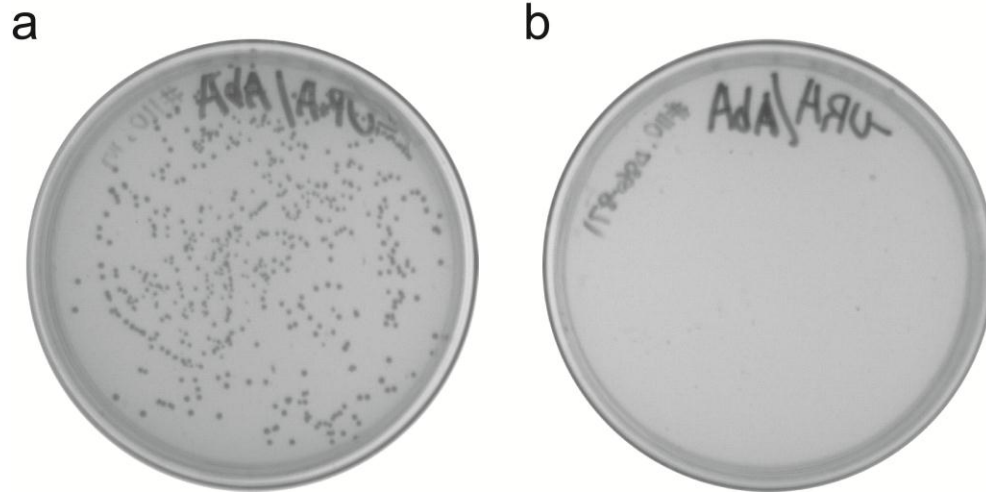
Supplementary Figure 5: Confirmation of the binding between the isolated TALE-VP16 fusion and its predicted target sequence within the SCN9A bait sequence. The isolated clone pGADT7-SCN9A-TALE#110-VP16 was predicted to target 5'-GTTGCCTCCCC-3' within the SCN9A bait sequence. The full-length bait sequence (pAbAi-SCN9A) or a bait with a deletion of the predicted target site (pAbAi-SCN9A(ΔGTTGCCTCCCCC)) was stably integrated into yeast cells. pGADT7-SCN9A-TALE#110-VP16 was then transformed into either cell line. Only cells which contained the intact bait sequence (panel a) survived the 100 nM Aureobasidin A selection.
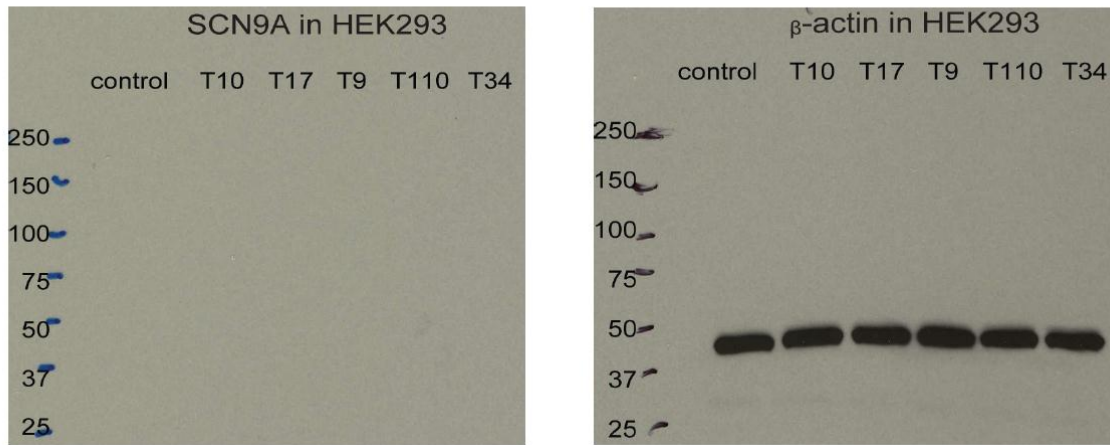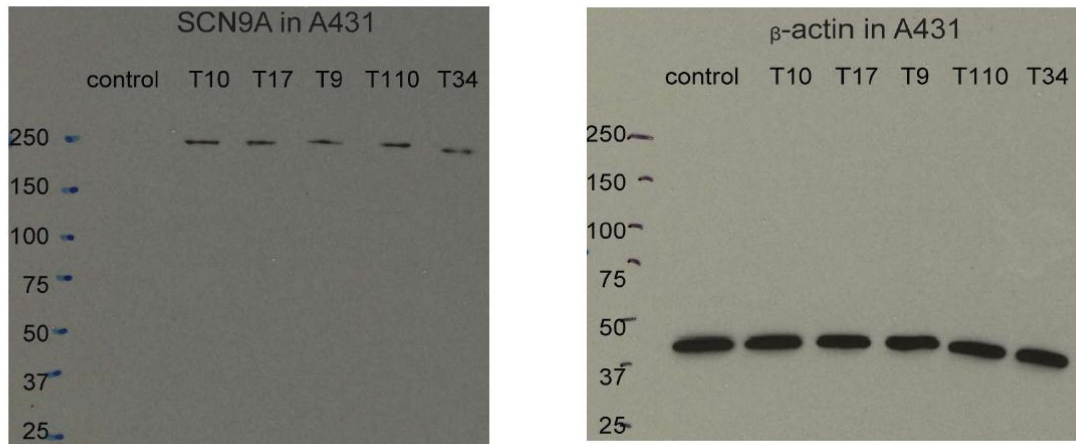
Supplementary Figure 6: The isolated TALE-VP16 fusions induced the overexpression of SCN9A protein in A431 cells, but not in HEK293 cells. 500,000 of A431 or HEK293 cells were seeded on each well of 6-well cell culture plates. After 16 hours, the cells were transiently transfected with 2 µg of each isolated TALE-VP16 constructs. The empty PEF-1 plasmid was used as the control. 72 hours later, the cells were lysed in RIPA buffer with protease inhibitors. For Western blot, 15 µg of whole cell lysates were used per lane. panel a: Western blot for HEK293 cells; panel b: Western blot for A431 cells. The TALE-VP16 fusions induced the overexpression of SCN9A protein in A431 cells, but not in HEK293 cells.
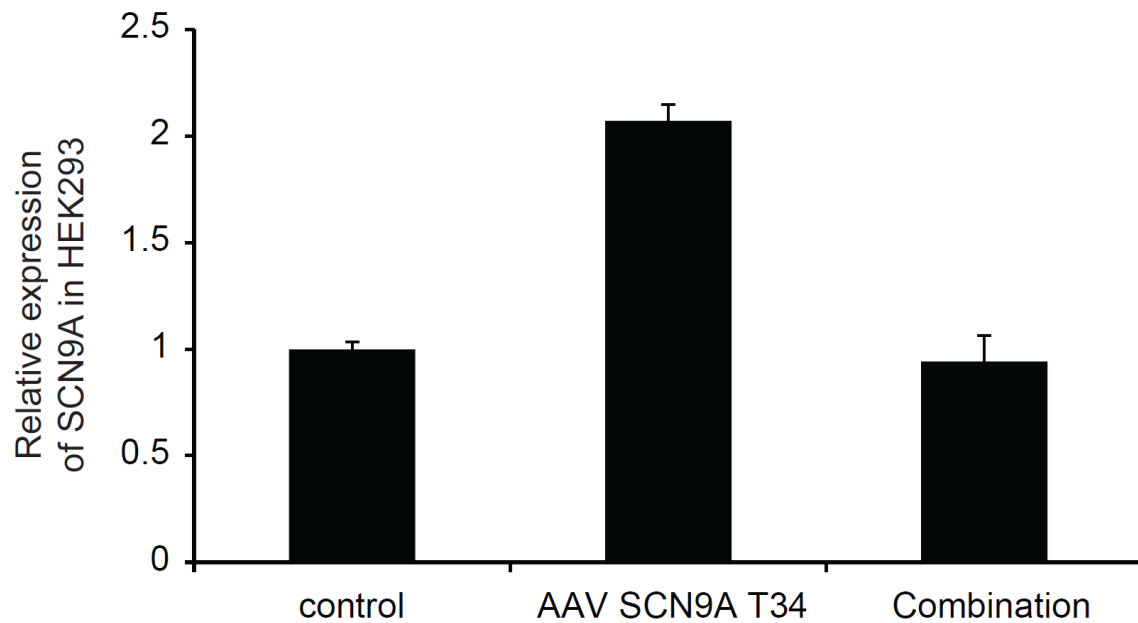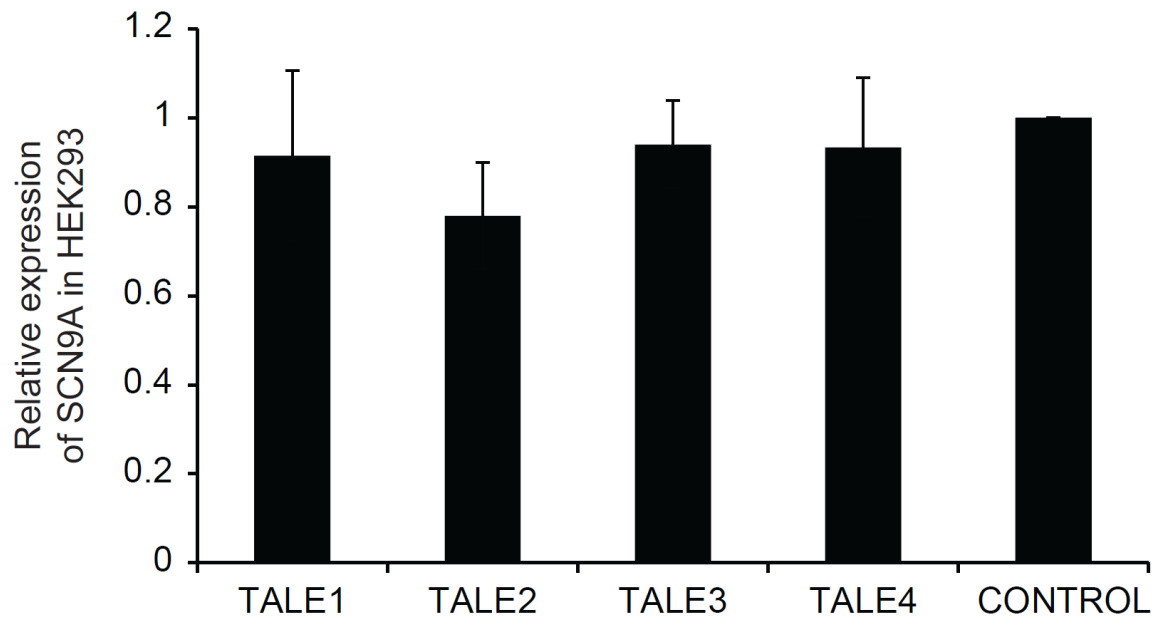
Supplementary Figure 7: The effects of SCN9A T34 AAV viral stock and the combinations of 4 TALE fusions designed by TALE-NT2.0 in HEK293 cells. Four TALE-VP16 fusions were designed according to TALE-NT2.0 and cloned into the PEF-1 vector. A combination of 250 ng of each TALE fusion construct was transiently transfected into HEK293 cells. An empty vector (PEF-1) was used as the control. In addition, the TALE-VP16 T34 cDNA was cloned into AAV-MCS vector and the corresponding AAV viral vectors were prepared. Subsequently, HEK293 cells were infected with this SCN9A T34 AAV viral stock at a MOI of 400. 48 hours later, the mRNA levels of SCN9A were determined by quantitative RT-PCR. The T34 AAV viral stock induced the overexpression of the SCN9A gene by 2 fold. In contrast, the combination of the four SCN9A bait-targeting TALE-VP16 fusions designed by TALE-NT2.0 failed to induce its overexpression.

Supplementary Figure 8: Secondary confirmation of the positive yeast clones for the human miRNA-34b/c bait. The positive clones isolated from the yeast one-hybrid screening were re-streaked onto –Leu plates containing 100 nM Aureobasidin A. After 4 days, the TALE-VP16 positive clones (M16 and M17) showed robust growth, while the control (Wild-type, yeast cells which are transformed with empty prey vector pGADT7) failed to grow.

Supplementary Figure 9: The SCN9A bait-targeting TALEs designed by TALE-NT2.0 failed to induce the overexpression of the SCN9A gene in HEK293 cells. Four TALE-VP16 fusions were designed according to TALE-NT2.0 and cloned into the PEF-1 vector. 1 μg of each TALE fusion construct was transiently transfected into HEK293 cells. 48 hours later, the mRNA levels of SCN9A were determined by quantitative RT-PCR. An empty vector (PEF-1) was used as the control. All 4 TALEs designed according to TALE-NT2.0 failed to induce the overexpression of SCN9A in HEK293 cells.

Supplementary Figure 10: Secondary confirmation of the positive yeast clones conferring cycloheximide resistance. 18 positive clones were isolated from the cycloheximide resistance screening. The positive clones were re-streaked onto –Leu plates containing 0.5 µg/ml of cycloheximide. After 3 days, the TALE-VP16 positive clones (A11 and A12) showed robust growth, while the control (Wild-type, yeast cells which are transformed with empty prey vector pGADT7) failed to grow.
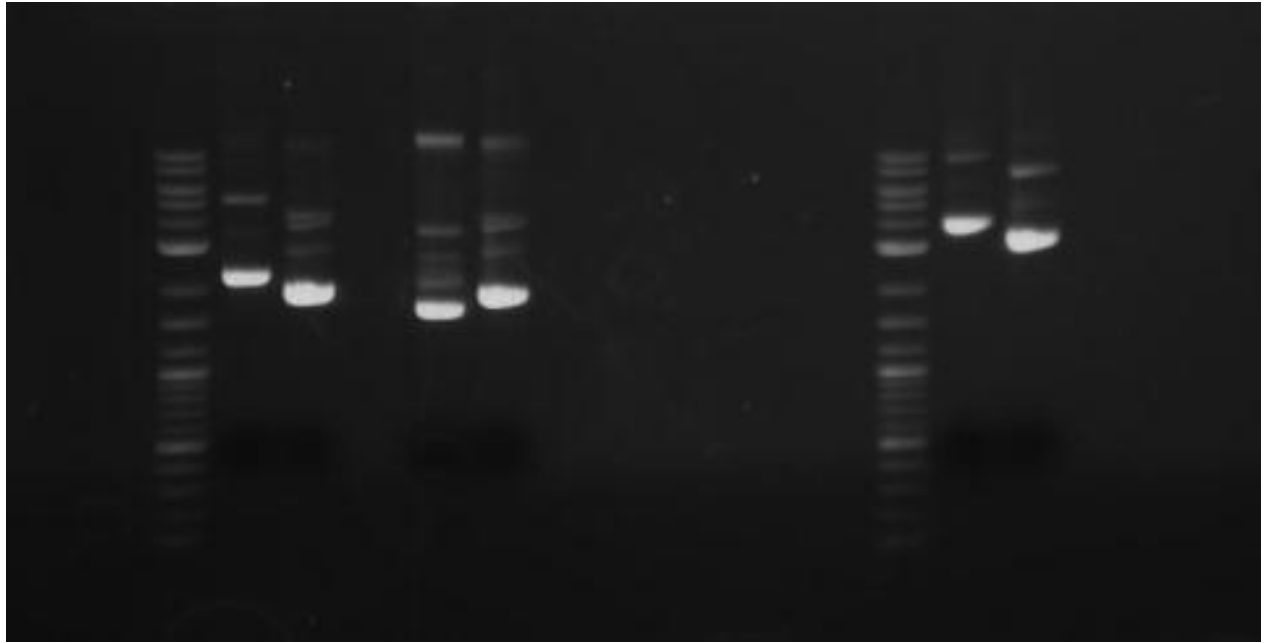
Supplementary Figure 11: Gel electrophoresis of 11-mer TALE-VP16 plasmid libraries in the PTAL1-VP16 plasmid. 600 ng of the pFUS_A (RVDs 1-10), pFUS_B (RVD 1), and the final 11-mer TALE-VP16 libraries (in the PTAL1-VP16 plasmid) were subjected to 0.7% agarose gel electrophoresis along with the original pFUS_A, pFUS_B and PTAL1-VP16 mother vectors. The electrophoresis image shows that the libraries were free of noticeable contamination from their respective mother vectors. From left to right: 2-Log DNA ladder (New England Biolabs, #N3200S), pFUS_A (RVDs 1-10), pFUS_A, pFUS_B (RVD 1), pFUS_B; 2-Log DNA ladder, 11-mer TALE-VP16 library, and PTAL1-VP16.

Supplementary Tables

Supplementary Table 1: Primers used in this study.

Primer ID	Primer sequence (5'->3')	Application
P1	CAGTACAAGCTTTCTGATGGGTCTGGGAAAAGTAAAT	forward primer for SCN9A
P2	CAGTACCTCGAGTGGGTTCCTTTGATTTTCTTTCAGCAA	reverse primer for SCN9A
P3	CAGTACCCGCGGAAAGATGGCAATGTTGCCTCCCCCA	forward primer for pAbAi-SCN9A(ΔATACAGGATGA)
P4	CAGTACCCGCGGATTTTAATTCCTCTTCAGCTCCTCAC	reverse primer for pAbAi-SCN9A(ΔATACAGGATGA)
P5	CAGTACCCGCGGATTAAAATATACAGGATGAAAAGATGG	forward primer for pAbAi-SCN9A(ΔTATGTGAGGAGCTGA)
P6	CAGTACCCGCGGAGAGGCCTGGATGGAAACAAAGAAAT	reverse primer for pAbAi-SCN9A(ΔTATGTGAGGAGCTGA)
P7	CAGTACCCGCGGAGGACCTCAGAGCTTTGTCCATTTC	forward primer for pAbAi-SCN9A(ΔGTTGCCTCCCCC)
P8	CAGTACCCGCGGATTGCCATCTTTTCATCCTGTATA	reverse primer for pAbAi-SCN9A(ΔGTTGCCTCCCCC)
P9	CAGTACGAATTCGTGCACCATGGATCCCATTCGTCCGCG	forward primer for PEF-1 cloning
P10	CAGTACGAATTCTCACTGAGGCAATAGCATATGTCTAGA	reverse primer for PEF-1 cloning
P11	CAGTACAAGCTTCATTGTCCATCTCACCAACTAGACC	forward primer for miR-34b/c
P12	CAGTACCTCGAGGCATATTTACTCATAGGAAATATAA	reverse primer for miR-34b/c
P13	CAGTACCCGCGGACTATCAAGAAGTTTAATAATCTGA	forward primer for pAbAi-miR-34b/c(ΔTTTCTAGGTAT)
P14	CAGTACCCGCGGCTAGTTTATATGTTGAAGCAAGACT	reverse primer for pAbAi-miR-34b/c(ΔTTTCTAGGTAT)
P15	CAGTACGAATTCCCGTGCACCATGGATCCCATTCGTCCGCG	forward primer for PEF-1 cloning
P16	CAGTACGGTACCCTGAGGCAATAGCATATGTCTAGA	reverse primer for PEF-1 cloning
P17	CAGTCCGCTAGCGCTACCGGTCGCCACCATGGTGAGC	forward primer for pAbAi-mKATE2
P18	CAGTCCGCGGCCGCTTAATCTGAGTCCGGAACCTCCTCCACC	reverse primer for pAbAi-mKATE2
P19	CAGTACAAGCTTCAATACAAACAAGGCCTCCAATACAAACAAGGCCTCCAATACAAACAAGGCCTCCAATACAAACAAGGCCTCGAATTCGAGCTCGGTACCCGGGGAT	forward primer for pAbAi-4XA35/PDR5-mKATE2
P20	CAGTACGCTAGCTATTAATTTAGTGTGTGTA	reverse primer for pAbAi-4XA35/PDR5-mKATE2
P21	CAGTACAAGCTTGCTTTCAGGTAAAGCCATGCTTTCAGGTAAAGCCATGCTTTCAGGTAAAGCCATGCTTTCAGGTAAAGCCATGAATTCGAGCTCGGTACCCGGGGAT	forward primer for pAbAi-4XA35/PDR3-mKATE2
P22	CAGTACAAGCTTAACTGCATTTCACTTTTCAACTGCATTTCACTTTTCAACTGCATTTCACTTTTCAACTGCATTTCACTTTTCGAATTCGAGCTCGGTACCCGGGGAT	forward primer for pAbAi-4XA8/PDR3-mKATE2
P23	CGTCGCTGTCACGTATCAGCACATAATC	forward primer for Sanger sequencing
P24	GCGTACGCTGGGTGGGAGTGGCAC	reverse primer for Sanger sequencing
P25	CAGTCCGCGGCCGCGGAAATCCATTATGTACTATTTAAAAA	forward primer for pAbAi-mKATE2
P26	CAGTCCGCTAGCTATTAATTTAGTGTGTGTATTTGTGTTTGC	reverse primer for pAbAi-mKATE2
P27	AAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCCGTGCACCATGGATCCCATTCGTCCGCG	forward primer for TALE-VP16 cDNA library
P28	TCTAGAGGCCGAGGCGGCCGACATGCTGAGGCAATAGCATATGTCTAGA	reverse primer for TALE-VP16 cDNA library
P29	CTCCACTTAGACGGCGAGGA	forward primer for VP-16
P30	GAAGTCGGCCGTATCCAGAG	reverse primer for VP-16
P31	TCGGCAATCGATCCACCATGGATCCCATTCGTCCGCGCAGGCC	forward primer for AAV-SCN9A-T34 cloning
P32	TCGGCAAAGCTTTCACTGAGGCAATAGCATATGTCTA	reverse primer for AAV-SCN9A-T34 cloning

Supplementary Table 2: The bait sequences for human SCN9A and miR-34b/c in the yeast one-hybrid assays.

Gene Target	Bait sequence (5'->3')
human SCN9A	TCTGATGGGTCTGGGAAAAGTAAATTGAAAATCTTCTGGAAAGGATTCACCATTCTAGATGTTATTAGAACATTAATGATTTGTGGGAGGAGGGCAATATACAAACATTAATAGGAAATTGGAATGAGTTGATTCCAACTCTCATGGATGACTATGAGAGGTTTAAGACTTCAGTGATGGAAAGAACTGCAGATGTGGTGGAAATAGCAAGAGAATTAGAATTAAGAGATGGAGCCTAAAATTGTGACTGAATTGCTGCCATCTCGCGGTAACTTGAATGGATGAGGAGTTGCTTCTGATTGATAAGCAAAGAAAGTGGTTTCGTGAGATGGAATCTTCTCCTGGTCAAGATGTTATGAACATTGTTGAAGTGACAGCAGAGGATTTAGAATATTACATAAGTTTAGTTGGTAAAACAGCATCAGGCTGGAGAGTGGTGCTACCAAAGGCATTATCATAAAGACTTTTAAAAGAGGAAAGTTGAGTTGAATTGCACCAAAGGTAAATATAGACGAAAAGAAGGGGAATTTAGGAAAAATTTACTCAGAAAATTGTCATTTTTTACTTTAAAGCTATGAAGCGTAAACTTCAAATTTCTGAGTATAAAGATGGTTAGAAGGCACGAAAGATAAGTCCCGCCCATTGCCTGACACATAGTAATCCCTTAACAAGCTGATAATTCTAGAACATGTTACCTTTTGTAGTTGAAACTAAAAATCTGTTTATGTTGTTATTATTAGTTTTTAATGGGCCTTTCTTGGCAGGCAAATAGTTAAGTCTTTATTTCTTTGTTTCCATCCAGGCCTCTTATGTGAGGAGCTGAAGAGGAATTAAAATATACAGGATGAAAAGATGGCAATGTTGCCTCCCCCAGGACCTCAGAGCTTTGTCCATTTCACAAAACAGTCTCTTGCCCTCATTGAACAACGCATTGCTGAAAGAAAATCAAAGGAACCCA
human miR-34b/c	CATTGTCCATCTCACCAACTAGACCACATAAATCCTTAGAGATTGTTCTCTAAAGATGGAAAATATGCAACCTGTTTGCTCCTACTGAAGCCATACCTAAGAGCTCCGGCAACTTTAATTATAGAGACTATTTTGTTAGACTGCTTATGATTTACCTGAAAGATAAATTAACATTTTATACTGAAAGCAACAGTATTGATTAATCAATAAATATAAACAAGAAGCTAATTAAGAATCTGGTGCACTCTTACTAATGAGAGTACAATGAAATGAAACTAGAAATATGGAATTATAAAAACCGCTGGCAGTTCATTTTAGCTCAAAAGAAAATCTTCTCCCATAAATAGATTCAATTATTTTTAGTTCTTCAAAGGCACTGAATTGACTTCAATTT***AGTCTTGCTTCAACATATAAACTAGTTT***CTAGGTATACTATCAAGAAGTTTAATAATCTGACATTTATTGATCTATAAACCAAGATATGGTTTTATATTTCCTATGAGTAAATATGC

*: p53 responsive element: p53RE

Supplementary Table 3: The RVD sequences for isolated TALE-VP16 clones.

human SCN9A clone ID	RVD sequence	gene target sequence (5'->3')	DNA strand	the lowest possible score
T42	NG HD NI NG HD HD NN NN NG NI NG	TCATCCTGTAT	minus	5.79 (addition of NI)
T10	NG HD NN NG HD HD NG NI NG NI NG	TCATCCTGTAT	minus	6.91 (addition of NI)
T72	NG HD NG NG HD HD NG NN NG NI NG	TCATCCTGTAT	minus	5.31 (addition of NI)
T54	NG HD NN NG HD HD NG NN NG NN NG	TCATCCTGTAT	minus	4.91 (addition of NI)
T17	NG NI NG NN NG NN NI NN NN NG NN	TATGTGAGGAG	plus	7.77 (addition of HD)
T49	NG NI NG NN NG NI NI NN NN NI NN	TATGTGAGGAG	plus	8.52 (addition of HD)
T33	NG HD NN NN HD NG HD HD NG HD HD	TCAGCTCCTCA	minus	6.09 (addition of HD)
T9	NG HD NI NN HD NG HD HD NG HD HD	TCAGCTCCTCA	minus	5.24 (addition of HD)
T110	NN NG NG NN HD NI NG HD HD HD HD	GTTGCCTCCCC	plus	5.63 (addition of HD)
T28	NN NG NG NN HD HD NG NN HD HD HD	GTTGCCTCCCC	plus	5.30 (addition of HD)
T66	NN NG NG NN HD NG NG HD HD HD NI	GTTGCCTCCCC	plus	7.76 (addition of HD)
T91	NN NG NG NN HD NN NG HD HD HD HD	GTTGCCTCCCC	plus	5.30 (addition of HD)
T34	NN NN NN NN NN NI NN NN HD NI NI	GGGGGAGGCAA	minus	6.65 (addition of HD)
human miR-34b/c clone ID	RVD sequence	gene target sequence (5'->3')	DNA strand	the lowest possible score
M1	NI NG NN HD HD NG NN NN NI NN NN	ATACCTAGAAA	minus	6.32 (addition of HD)
M17	NI NG NN HD HD NG NI NN NI NN NI	ATACCTAGAAA	minus	4.60 (addition of HD)
M13	NG HD NN NG NI NI NG HD NG NN NN	TAATAATCTGA	plus	7.14 (addition of HD)
M16	NN HD NG HD NG NG NI NN NN NN NN	ACTCATAGGAA	minus	8.25 (addition of NI)
cycloheximide resistance clone ID	RVD sequence	gene target sequence (5'->3')	DNA strand	the lowest possible score
A7	NG NI NG HD NG NG HD NG NG NG HD	n.d.	n.d.	n.d.
A8	NN HD NN NG NG NG NI NG HD NG NG	GCATTTCACTT (PDR3)	plus	8.33 (addition of NG)
A11	NN NG NG NN HD NG HD HD NG NN NN	n.d.	n.d.	n.d.
A17	NN HD NI HD HD HD HD NG NN NN NG	n.d.	n.d.	n.d.
A35	NN HD NI NN NN NG NN NI NN NN HD	TCAGGTAAAGC (PDR3)	plus	8.51 (addition of HD)
A35	NN HD NI NN NN NG NN NI NN NN HD	ACAAACAAGGC (PDR5)	plus	8.88 (addition of HD)
human SCN9A designed by TALE-NT2.0 clone ID	RVD sequence	gene target sequence (5'->3')	DNA strand	
TALE1	NI NG NI HD NI NI NI HD NI NG NG	ATACAAACATT	plus	
TALE2	NI HD HD NI NI NI NN NN HD NI NG	ACCAAAGGCAT	plus	
TALE3	NI HD NG HD NI NN NI NI NI NI NG	ACTCAGAAAAT	plus	
TALE4	HD HD NI NN NN HD HD NG HD NG NG	CCAGGCCTCTT	plus	
*n.d. : not determined				

References

1. Maxwell, K. L. *et al*. Protein folding: defining a "standard" set of experimental conditions and a preliminary kinetic data set of two-state proteins. *Protein Sci* 14, 602-616 (2005).

2. Rouillard, J., Zuker, M. & Gulari, E. OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucl. Acids Res.* 31, 3057-3062 (2003).

3. Kristiansson, E., Thorsen, M., Tamás, M. J. & Nerman, O. Evolutionary Forces Act on Promoter Length: Identification of Enriched Cis-Regulatory Elements. *Molecular Biology and Evolution* 26, 1299-1307 (2009).