

METHOD AND COMPOSITION FOR INHIBITING TELOMERASE ACTIVITY

[0001] This application claims the benefit of priority of Singapore application No. 10202112459Q, filed 9 November 2021, the contents of it being hereby incorporated by
5 reference in its entirety for all purposes.

[0002] The invention relates generally to the field of biotechnology. In particular, the invention relates to methods and compositions for inhibiting telomerase activity in cells.

[0003] Human telomerase reverse transcriptase (hTERT), the enzymatic component of telomerase holoenzyme, is expressed in stem cells and germ cells to maintain telomere length
10 and chromosome integrity. However, *hTERT* is transcriptionally repressed in terminally differentiated somatic cells. Accumulation of driver mutations during oncogenesis and ensuing replicative stress in preneoplastic cells lead to re-activation of *hTERT* expression for the maintenance of critically short telomeres.

[0004] Mutations in human telomerase reverse transcriptase (hTERT) promoter is seen in
15 19% of human cancers, irrespective of the cancer type. The mutations in mutant *hTERT* (Mut-*hTERT*) promoter create binding sites for ETS family of transcription factors, which enhance *hTERT* gene expression and hence telomerase activity. Therefore, there are distinctions between the transcriptional regulation of wild-type (WT) hTERT promoter and of mutant (Mut) hTERT promoter. While conventional screens have identified regulators of wild-type hTERT
20 promoter, no genome-wide screens have been carried out to identify specific regulators of Mut-hTERT promoters.

[0005] There is thus a need to identify specific regulators of Mut-hTERT promoter that can be modulated to inhibit telomerase activity in cancer cells. Furthermore, other desirable features and characteristics will become apparent from the subsequent detailed description
25 and the appended claims, taken in conjunction with the accompanying drawings and this background of the disclosure.

[0006] In one aspect, the present invention provides a use of a moiety that inhibits MED12 in the manufacture of a medicament for treating cancer.

[0007] By "moiety", it is meant to include any substance, compound or agent.

30 [0008] In another aspect of the present invention, there is provided a moiety that inhibits MED12 for use in treating cancer.

[0009] In another aspect, the present invention provides a method for treating cancer comprising administering to a subject a composition comprising a moiety that inhibits MED12.

[0010] The moiety that inhibits MED12 can be administered by any of the following routes:
35 oral, intravenous, intramuscular, or subcutaneous. The treatment regimens can be dynamic

treatment regime (DTR), adaptive intervention, or adaptive treatment strategy. The inhibitor (drug) can be prescribed as a single agent or in combination with chemotherapeutic/ targeted therapeutic agents as a daily, monthly, quarterly, semi-annual, or yearly course.

[0011] In various embodiments, the cancer to be treated has a MED12 mutation, a hTERT promoter mutation, or both. The hTERT promoter mutation may be a C250T mutation or a C228T mutation. The MED12 mutation may be a gain-of-function mutation that causes activation of the MED12 protein. The activation of MED12 protein can in turn have synergistic effect with the hTERT promoter mutation. Therefore, the moiety that inhibits MED12 can inhibit cancer growth in these cancers.

[0012] In various embodiments, the cancer as described herein is selected from the group consisting of T-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia, colorectal cancer, gastric cancer, breast cancer, non-anaplastic thyroid cancer, central nervous system cancer, and skin cancer.

[0013] In various embodiments, the central nervous system cancer is glioblastoma and the skin cancer is melanoma.

[0014] As used herein, the term "inhibit" or "inhibiting" is intended to mean a substantially slowing, interference, suppression, prevention, delay and/or arrest of a chemical action. For example, the phrase "inhibiting telomerase activity" refers to a partial decrease or a complete loss in the activity of the telomerase when compared to the original activity of the telomerase.

The term "inhibiting" may also refer to an ability of an inhibitor to block, partially block, interfere with, reduce or inactivate target molecules (e.g. gene, mRNA, protein). The term "inhibit" in the context of transcriptional inhibition may refer to the epigenetic repression of a gene or reducing or blocking transcription of the gene. The term "inhibit" in the context of translational inhibition may include the modulation of splicing of a pre-mRNA, or facilitating the nonsense mediated decay of an mRNA, or facilitating cleavage of an mRNA. The term "inhibit" in the context of post-translational inhibition may include modifications to a protein to modulate its activity, function or interactions with other moieties, or facilitating the degradation of a protein.

[0015] The term "inhibitor" is meant to include to a substance (such a compound, drug or agent) that interferes with (i.e. reduces or eliminates or suppresses) a resulting target molecule or target compound or target process. An inhibitor could be a transcriptional inhibitor, a translational inhibitor, or a post-translational inhibitor. A transcriptional inhibitor may inhibit a target gene through epigenetic repression, by blocking transcription initiation, or by blocking transcription elongation, thereby causing a reduction in the level of expression of the target gene. A translational inhibitor may inhibit a target mRNA by blocking translation initiation, by blocking translation elongation, or by modulating the splicing of the target mRNA. A

translational inhibitor may also trigger/facilitate cleavage or degradation of the target mRNA. A post-translational inhibitor may inhibit a target protein by blocking the interactions of the target protein with other moieties, reduce its activity by post-translational modification of the target protein, or trigger/facilitate degradation of the target protein.

5 [0016] An inhibitor may decrease or silence the expression of a gene. An inhibitor may also decrease or silence the activity of a protein. The phrase "MED12 inhibitor" or "inhibitor of MED12" may include an inhibitor of *MED12* gene, an inhibitor of MED12 mRNA or an inhibitor of MED12 protein. Examples of inhibitors include but are not limited to an oligonucleotide, a small molecule or a compound. The oligonucleotide may be an interfering RNA (iRNA),
10 including but not limited to small interfering RNA (siRNA) or short hairpin RNA (shRNA). A small molecule will generally be understood in the art as a compound that has a low molecular weight.

[0017] Small molecule inhibitors of MED12 may include Pracinostat (SB939), PCI-24781 (Abexinostat), CUDC-101, UNC0379, Belinostat (PXD101), M344 and (+)-JQ1. Preferably,
15 the small molecule inhibitor is an epigenetic inhibitor that causes transcriptional inhibition of MED12.

[0018] As used herein, the term "promoter" refers to a region of DNA capable of binding an RNA polymerase in a cell (directly or through other promoter-bound proteins or substances) that initiates transcription of a gene. Within the promoter sequence may be found a
20 transcription initiation site as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. A promoter may be a major promoter, a minor promoter or an alternative promoter. A major promoter is a promoter that is the most frequently used for the transcription of a gene. A promoter may be a constitutive promoter or an inducible promoter. As used herein, the term "operably linked" refers to the association of nucleic acid
25 sequences on a single nucleic acid fragment such that the function of one is affected by the other. For example, a polynucleotide sequence is said to be "operably linked" to a promoter if the two sequences are situated such that the promoter affects the expression of the polynucleotide sequence (i.e., the polynucleotide sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense
30 or antisense orientation.

[0019] The term "nucleic acid" as used herein refers to a deoxyribonucleotide or ribonucleotide polymer in either single or double stranded form, and unless otherwise limited, encompassing known analogues of natural nucleotides that hybridize to nucleic acids in a manner similar to naturally occurring nucleotides. "Nucleotide" includes, but is not limited to,
35 a monomer that includes a base linked to a sugar, such as a pyrimidine, purine or synthetic

analogous thereof, or a base linked to an amino acid, as in a peptide nucleic acid (PNA). A nucleotide is one monomer in a polynucleotide. A nucleotide sequence refers to the sequence of bases in a polynucleotide.

[0020] By "gene", it is meant to include a polynucleotide sequence that expresses a specific protein. The polynucleotide sequence may include regions preceding and following the coding region involved in the transcription/translation of the gene product and the regulation of the transcription/translation, as well as intervening sequences (introns) between individual coding segments (exons).

[0021] In another aspect, the present invention provides a pharmaceutical composition comprising (a) a therapeutic amount of a moiety that inhibits MED12 and (b) one or more pharmaceutically acceptable carriers and/or diluents.

[0022] The term "pharmaceutically acceptable carrier" refers to a substance or composition that, when combined with an active ingredient, is compatible with the active ingredient and does not cause a toxic or otherwise unnecessary reaction when administered to a subject, particularly a mammal. Examples of pharmaceutically acceptable carriers include solvents, surfactants, suspending agents, buffers, lubricants, emulsifiers, absorbents, dispersion media, coatings and stabilizers.

[0023] The pharmaceutical compositions of the disclosure may be administered in a number of ways depending upon whether local or systemic treatment is desired and upon the area to be treated. Administration may be topical (e.g. cream based, ophthalmic or to mucous membranes including rectal delivery), pulmonary, e.g., by inhalation of powders or aerosols, (including by nebulizer, intratracheal, intranasal, epidermal and transdermal), oral or parenteral. Parenteral administration includes intravenous, intra-arterial, subcutaneous, intraperitoneal or intramuscular injection or infusion; or intracranial, e.g., intrathecal or intraventricular, administration.

[0024] The pharmaceutical formulations of the disclosure, which may conveniently be presented in unit dosage form, may be prepared according to conventional techniques well known in the pharmaceutical industry. Such techniques include the step of bringing into association the active ingredients with the pharmaceutical carrier(s) or excipient(s). In general the formulations are prepared by uniformly bringing into association the active ingredients with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the product.

[0025] Combination therapy with an additional therapeutic agent may also be contemplated by the present disclosure. The additional therapeutic agent may be a TGF- β R Inhibitor, a tyrosine kinase inhibitor (TKI), or a bromodomain and extra terminal motif (BET) inhibitor.

[0026] In one embodiment, the moiety that inhibits MED12 is an siRNA. By “siRNA”, it is meant to include siRNAs formed from the processing of Dicer-substrate small interfering RNAs (dsiRNAs) by the enzyme DICER. DsiRNAs refer to duplex RNAs that are synthesized as single-stranded RNA (ssRNA).

5 [0027] **Table 1.** DsiRNAs that inhibit expression of MED12 mRNA.

SEQ ID NO	Sequence of dsiRNA (shown as single-stranded RNAs used to form dsiRNA duplex)	dsiRNA catalogue ID (IDT)
1	5'- rArCrC rUrGrA rGrUrU rCrArU rGrArU rCrGrA rCrArC rUrCT A	hs.Ri.MED12.13.9
2	5'- rUrArG rArGrU rGrUrC rGrArU rCrArU rGrArA rCrUrC rArGrG rUrGrC	
3	5'- rArArG rGrArU rArUrC rUrUrG rArArG rGrUrU rCrUrG rArAC C	hs.Ri.MED12.13.1
4	5'- rGrGrU rUrCrA rGrArA rCrCrU rUrCrA rArGrA rUrArU rCrCrU rUrGrG	

[0028] In one embodiment, the siRNA as described herein is formed from a Dicer-substrate small interfering RNA (dsiRNA), wherein the dsiRNA comprises the sequence of: (i) SEQ ID NOs 1 and 2; or (ii) SEQ ID NOs 3 and 4. “r” is indicative of ribonucleotides, U: uracil, T: thymine, A: adenine, G: guanine, C: cytosine. The dsiRNA sequences may comprise both ribonucleotides and deoxyribonucleotides to increase the stability of the dsiRNA within the cell.

[0029] In one aspect of the present invention, there is provided a method for inhibiting telomerase activity in a cell comprising a mutant hTERT promoter, the method comprising contacting the cell with an inhibitor of MED12.

[0030] In one embodiment, the mutant hTERT promoter comprises a C250T mutation or a C228T mutation.

[0031] The term “mutation” or “mutant” refers to, in the context of a nucleic acid sequence, an alteration or difference in the nucleic acid sequence (nucleotide substitution, addition or deletion) compared to a reference. Mutations can occur within the protein-coding region of a gene as well as in regions other than the protein-coding sequence, such as regulatory or promoter sequences. Mutations may include point mutations, frame-shift mutations, nonsense mutations, insertions or deletions. The term “mutation”, in the context of a protein or polypeptide, refers to the deletion, addition or substitution of amino acid residues in the amino

acid sequence of a protein or polypeptide compared to the amino acid sequence of a reference protein or polypeptide.

[0032] In one embodiment, the inhibitor of MED12 inhibits telomerase activity via transcriptional inhibition of MED12, translational inhibition of MED12 or post-translational inhibition of MED12.

[0033] For example, an epigenetic inhibitor could cause gene repression of MED12 thus blocking its transcriptional activity. An inhibitor could also influence the splicing of MED12 pre-mRNA, creating vulnerabilities for nonsense mediated decay of MED12 transcript. MED12 protein has been shown to be methylated at arginine 1862 (R1862) and arginine 1912 (R1912) by coactivator-associated arginine methyltransferase 1 (CARM1), and these methylations are crucial for the cancer-specific functions of MED12. Thus, an inhibitor targeting the methylation could affect the function of MED12. Kinase inhibitors and proteolysis-targeting chimera (PROTAC) may promote the degradation of MED12 protein, facilitating its clearance from cells. An inhibitor could also block the interaction of MED12 protein with its activating non-coding RNA, thus disrupting the function of MED12 protein.

[0034] In one embodiment, the inhibitor of MED12 is a small interfering RNA (siRNA) that inhibits expression of MED12 mRNA.

[0035] In one embodiment, the siRNA as described herein is formed from a Dicer-substrate small interfering RNA (dsiRNA), wherein the dsiRNA comprises the sequence of: (i) SEQ ID NOs 1 and 2; or (ii) SEQ ID NOs 3 and 4.

[0036] In one aspect of the invention, there is provided a host cell comprising a nucleic acid construct, wherein the nucleic acid construct comprises a MED12 promoter; and an open reading frame, wherein the open reading frame comprises a translation start site, a polynucleotide sequence encoding a reporter element, a polynucleotide sequence encoding a self-cleaving peptide and a polynucleotide sequence encoding MED12. In the nucleic acid construct as described herein, the translation start site located after the MED12 promoter may be the only translation start site in the nucleic acid construct. In other words, the polynucleotide sequence encoding the reporter element, the polynucleotide sequence encoding the self-cleaving peptide and the polynucleotide sequence encoding MED12 do not comprise a translation start site. The "translation start site" may also be known as "start codon".

[0037] Advantageously, the position of the polynucleotide sequence encoding the reporter element after the MED12 promoter allows quantification of MED12 promoter activity by measuring the intensity of the signal produced by the reporter element. A high activity of the MED12 promoter would result in a high intensity of the signal produced by the reporter element. The reporter protein and the MED12 protein will be expressed as a fusion protein.

Subsequent cleavage by T2A will separate the reporter protein and MED12 protein into independent proteins.

[0038] In one embodiment, the open reading frame comprises a translation start site, a polynucleotide sequence encoding a reporter element, a polynucleotide sequence encoding a self-cleaving peptide and a polynucleotide sequence encoding MED12, in that order.

[0039] In another embodiment, the open reading frame comprises a translation start site, a polynucleotide sequence encoding MED12, a polynucleotide sequence encoding a self-cleaving peptide and a polynucleotide sequence encoding a reporter element, in that order.

[0040] Advantageously, positioning the polynucleotide sequence encoding the reporter element upstream of the polynucleotide sequences encoding the self-cleaving peptide and MED12 allows a stronger signal from the reporter element to be captured.

[0041] In one embodiment, the reporter element is a fluorescent protein or a luminescent protein. Examples of fluorescent proteins include red fluorescent protein (RFP), green fluorescent protein (GFP) and cyan fluorescent protein (CFP). An example of a luminescent protein is NanoLuc. The expression of the reporter element may be quantified using fluorescence-activated cell sorting (FACS), flow cytometry, microscopy or high-throughput fluorescence scanners.

[0042] In one embodiment, the reporter element is red fluorescent protein (RFP).

[0043] In one embodiment, the self-cleaving peptide is T2A.

[0044] In one embodiment, the nucleic acid construct is integrated into a genome of the host cell. In this embodiment, the MED12 promoter is the endogenous MED12 promoter, and the polynucleotide sequences encoding the reporter element, T2A and MED12 would be inserted after the ATG translation start side downstream of the MED12 promoter. Advantageously, integration of the nucleic acid construct into the genome reduces the variability caused by copy number variation and allows expression to be driven solely by the endogenous MED12 promoter.

[0045] In one aspect, the invention provides a method for identifying an inhibitor of MED12, the method comprising: (i) contacting the host cell as described herein with a composition comprising a CRISPR-Cas system for knocking out a target gene, wherein the CRISPR-Cas system comprises: (a) a guide RNA (gRNA) capable of binding to the target gene; and (b) a polynucleotide sequence encoding a CRISPR enzyme; (ii) quantifying expression of the reporter element in a host cell having the CRISPR-Cas system for knocking out the target gene as compared to expression of the reporter element in a control host cell without knockout of the target gene, wherein expression of the reporter element is indicative of MED12 expression, and wherein a higher expression of the reporter element in the host cell having

the CRISPR-Cas system for knocking out the target gene indicates that the target gene encodes an inhibitor of MED12.

[0046] For example, if a target gene encodes an inhibitor of MED12 gene expression, the knocking out of this target gene with the CRISPR-Cas system would lift the inhibition on MED12 gene expression, thereby increasing the expression of the reporter element. The method as described above may be used in a screening for transcriptional inhibitors and translational inhibitors of MED12.

[0047] In one embodiment, the composition comprises a viral vector and the viral vector comprises the CRISPR-Cas system.

10 [0048] In one embodiment, the viral vector is a lentiviral vector.

[0049] In one embodiment, the CRISPR enzyme is Cas9.

[0050] The "control host cell" of the method for identifying an inhibitor of MED12 as described herein refers to a comparator host cell for determining the effect of knocking out the target gene in the host cell comprising the CRISPR-Cas system for knocking out the target gene.

15 [0051] In one embodiment, the control host cell without knockout of the target gene in step (ii) of the method for identifying an inhibitor of MED12 as described herein refers to a cell that expresses Cas 9 but does not express a gRNA.

[0052] The terms "vector" and "expression vector" refer to a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular polynucleotide sequence in a host cell. An expression vector may be part of a plasmid, viral genome, or nucleic acid fragment. Typically, an expression vector includes a polynucleotide to be transcribed, operably linked to a promoter. Vectors may be viral or non-viral. A non-viral vector may comprise a plasmid, liposomes, nanoparticles, microbubble plus ultrasound, dendrimers, cationic magnetic nanoparticles, lipoplexes (lipid-based), inorganic molecules, etc. A viral vector may be of any kind, but in specific embodiments the viral vector is an adenoviral vector, an adeno-associated viral vector, a retroviral vector, or a lentiviral vector.

20 [0053] The term "viral vector" can refer to either a virus or a viral particle capable of transferring nucleic acid into a cell, or the transferred nucleic acid itself. Viral vectors and transfer plasmids contain structural and/or functional genetic elements that are primarily derived from a virus. The term "retroviral vector" refers to a viral vector or plasmid containing structural and functional genetic elements, or parts thereof, which are mainly derived from a retrovirus. The term "lentiviral vector" refers to a retroviral vector or plasmid containing

structural and functional genetic elements, or portions thereof, including LTRs, which are primarily derived from a lentivirus.

[0054] The term "CRISPR-Cas system" refers to an enzyme system including a guide RNA sequence that contains a nucleotide sequence complementary or substantially complementary to a region of a target polynucleotide/gene, and a protein with nuclease activity. CRISPR-Cas systems include Type I CRISPR-Cas system, Type II CRISPR-Cas system, Type III CRISPR-Cas system, and derivatives thereof. CRISPR-Cas systems include engineered and/or programmed nuclease systems derived from naturally accruing CRISPR-Cas systems. CRISPR-Cas systems may contain engineered and/or mutated Cas proteins. CRISPR-Cas systems may contain engineered and/or programmed guide RNA. The CRISPR-Cas system may be used to inhibit gene expression by modifying the genome. Modification of the genome includes but is not limited to deletion, insertion or substitution of nucleotides.

[0055] As used herein, the term "guide RNA" refers to an RNA containing a sequence that is complementary or substantially complementary to a region of a target DNA sequence. A guide RNA may contain nucleotide sequences other than the region complementary or substantially complementary to a region of a target DNA sequence. A guide RNA may be a crRNA or a derivative thereof.

[0056] In another aspect, the present invention provides a host cell comprising a nucleic acid construct, wherein the nucleic acid construct comprises an inducible promoter operably linked to a polynucleotide sequence encoding an RNA that inhibits gene expression of MED12 via RNA interference (RNAi).

[0057] In various embodiments, the RNA as described herein is selected from the group consisting of a short hairpin RNA (shRNA), small interfering RNA (siRNA), Dicer-substrate small interfering RNA (dsiRNA), peptide nucleic acid (PNA), antisense oligonucleotide (ASO) and microRNA (miRNA).

[0058] Examples of the inducible promoter include but are not limited to tetracycline-inducible promoters, doxycycline-inducible promoters and rapamycin-inducible promoters.

[0059] Long-term MED12 inhibition prevents cancer cells from proliferating and may not provide sufficient time to perform high-throughput screens. Advantageously, generating inducible shRNAs specific to MED12 gene (i.e. shMED12) and only activating the inducible shRNA during the screening process allows reduction or elimination of cellular cytotoxicity. The host cell as described herein may be used in a screen to identify MED12-dependent hTERT inhibitors by perturbing MED12 expression levels in the cell. For example, if a target protein reduces hTERT expression in the presence of MED12 (e.g. in a hTERT reporter cell line comprising shControl) and but does not reduce hTERT expression when MED12 is

knocked down (e.g. in a hTERT reporter cell line comprising shMED12), the target protein is a potential MED12-dependent hTERT inhibitor.

[0060] In one embodiment, the nucleic acid construct is integrated into a genome of the host cell. For example, the nucleic acid construct comprising the inducible promoter operably linked to the polynucleotide sequence encoding shRNA may or may not be integrated into the genome of the host cell. The nucleic acid construct comprising the inducible promoter operably linked to the polynucleotide sequence encoding siRNA, dsRNA, PNA, ASO and miRNA may not be integrated into the genome of the host cell.

[0061] In another aspect of the present invention, there is provided a method for identifying a MED12-dependent inhibitor of hTERT, the method comprising: (i) providing a host cell comprising an expression construct for inducible inhibition of MED12 expression; (ii) contacting the host cell with a composition comprising a CRISPR-Cas system for knocking out a target gene, wherein the CRISPR-Cas system comprises: (a) a guide RNA (gRNA) capable of binding to the target gene; and (b) a polynucleotide sequence encoding a CRISPR enzyme; (iii) quantifying hTERT expression in the host cell in which inhibition of MED12 expression has been induced as compared to hTERT expression in a control host cell without induced inhibition of MED12 expression, wherein a higher hTERT expression in the control host cell without induced inhibition of MED12 expression indicates that the target gene encodes a MED12-dependent inhibitor of hTERT.

[0062] For example, if a target gene is an inhibitor of hTERT but is dependent on MED12 for its action, it will cause a reduction in hTERT expression only in the presence of MED12 but not when MED12 is knocked out by induced inhibition of MED12 expression. When this target gene is knocked out by the CRISPR-Cas system, a higher hTERT expression in the presence of MED12 indicates that the target gene encodes a MED12-dependent inhibitor of hTERT.

[0063] In one embodiment, the host cell of step (i) of the method for identifying a MED12-dependent inhibitor of hTERT is the host cell as described herein and the expression construct is the nucleic acid construct comprising an inducible promoter operably linked to a polynucleotide sequence encoding an RNA that inhibits gene expression of MED12 via RNA interference (RNAi).

[0064] For example, the RNA that inhibits gene expression of MED12 via RNAi may be an shRNA. In this embodiment, the control host cell in the method for identifying a MED12-dependent inhibitor of hTERT as described herein may further include an inducible shControl for off-target control. "Inducible shControl" refers to an expression construct comprising an inducible promoter operably linked to a polynucleotide sequence encoding an "shControl" RNA, wherein this "shControl" RNA is a random sequence that does not target any specific

RNA in the cell. The inducible shControl is an experimental control and serves the purpose of normalizing cell viability after transfection.

[0065] Alternative ways of inducing inhibition of MED12 or achieving transient inhibition of MED12 include using siRNAs, peptide nucleic acid (PNA) oligonucleotides and miRNAs. In such methods, the control host cell may also include an inducible control RNA (equivalent to the inducible shControl described in the paragraph above), where the control RNAs are random sequences that do not have a specific target in the cell. Another way is to insert a recombinase sequence at the MED12 open reading frame and create a MED12 knockout with a stimulant. Another way is to perform CRISPR interference by targeting the MED12 gene or its promoter by sequence-specific gRNAs. Creating mutations or deletions at the MED12 promoter may reduce or increase the MED12 expression levels.

[0066] The method for identifying a MED12-dependent inhibitor of hTERT may be used to identify inhibitors that affect MED12 protein-protein interaction or MED12 protein-DNA interactions. The inhibitor may also be an inhibitor of MED12 levels or an inhibitor of MED12 interactions and functions.

[0067] In one embodiment, the composition comprises a viral vector and the viral vector comprises the CRISPR-Cas system.

[0068] In one embodiment, the viral vector is a lentiviral vector.

[0069] In one embodiment, the CRISPR enzyme is Cas9. The CRISPR enzyme can also be a derivative of Cas9, such as dCas9 (enzymatically inactive) fused with transcriptional regulatory domains (inhibitory and activatory).

[0070] In order that the present invention may be fully understood and readily put into practical effect, there shall now be described by way of non-limitative examples only preferred embodiments of the present invention, the description being with reference to the accompanying illustrative figures.

[0071] In the Figures:

[0072] **Fig. 1** shows the study design. **A)** Schematic of the reporter cell lines created by CRISPR genome editing. Either GFP or NanoLuc gene is inserted under the control of endogenous *hTERT* promoter in exon 1 with in-frame fusion. GFP/NanoLuc insertion and promoter mutation corrections were done using a single homologous recombination (HR) template designed to span homology arms in *hTERT* promoter and downstream of exon 1. **B)** Workflow of the genome-wide screens are shown. The reporter lines and the algorithms that are used to identify hits from the CRISPR and siRNA screens are indicated. Hits from siRNA and CRISPR knockout screen were overlapped to identify commonality factors and were then chosen for the validation study.

[0073] **Fig. 2** shows the establishment and validation of reporter cell lines. **A)** Representative plots of GFP sorting in U251-Mut and HCT116-Mut reporter cells. The x-axis shows increased signal (log) of GFP in clones after CRISPR mediated insertion of GFP into the endogenous *hTERT* locus. GFP positive reporter lines are enriched using FACS and named as U251-Mut-GFP and HCT116-Mut-GFP. **B)** qPCR data showing increased expression of the luciferase gene in clones after CRISPR mediated insertion of NanoLuc into the endogenous *hTERT* locus. **C)** U251-Mut-GFP (left) and HCT116-Mut-GFP cells were transfected with siRNA against *GABPA*. Cells were isolated for protein extraction 48, 72 and 96h post-transfection. Protein expression analysis was performed for the indicated molecules by western blot. HSP90 was used as a loading control. **D)** HCT116-Mut-NLuc, DLD1-Mut-NLuc and LN382-Mut-NLuc reporter lines were transfected with siControl (siCTRL) and siMYC. After 48 hours, luminescence signal of NanoLuc reporter was measured using Luciferase assay. The error bar indicates standard deviation from at least 3 biological replicates (p-value < 0.00005). **E)** HCT116-Mut-NLuc, DLD1-Mut-NLuc and LN382-Mut-NLuc reporter lines were transfected with siCTRL and siMYC. Gene expression analysis was performed for MYC and Luciferase by RT-qPCR. Actin was used as a control. **F)** HCT116-Mut-GFP, U251-Mut-GFP reporter lines were transfected with siControl (siCTRL) and siGABPA. qPCR results indicating the expression level of *GABPA* and GFP after transient knockdown of *GABPA* by siRNA. The error bars indicate standard deviation (SD) from at least 3 biological replicates.

[0074] **Fig. 3** shows siRNA screen in NanoLuc reporter cell lines. **A)** Workflow of siRNA screen illustrating the performance and selection methods used to identify the hits. **B)** Plots showing the inhibition and toxicity thresholds used in different lines to select hits from each siRNA screen. **C)** Table listing the number of hits identified from each siRNA screen utilizing 2 different algorithms using threshold level from panel B. **D)** Venn diagram showing the overlap of hits obtained from 2 algorithms in different cell lines (DLD-1-Mut-NLuc –top, LN382-Mut-NLuc –middle, HCT116-Mut-NLuc -bottom). **E)** Bar plots showing the enrichment of pathway terms (top 15, adj. p-value <0.05) for the overlapping hits identified from panel D. The black bars indicate common cell survival-related pathways seen among all the lines; dark grey bars indicate the pathway terms related to *hTERT* regulation seen in specific cell lines **F)** Overlap of hits identified from 2 different algorithms (Cutoff and SSMD) in DLD-1-Mut-NLuc and LN382-Mut-NLuc cell lines. The methods of Cutoff and SSMD are described in the method section. **G)** Graph shows the pathway terms of the common genes between DLD-1-Mut-NLuc and LN382-Mut-NLuc reporter cell lines. The pathway terms were obtained from the gProfiler (g:GOst) using its inbuilt default functions.

[0075] **Fig. 4** shows CRISPR knockout screen in GFP reporter cell lines. **A)** Workflow for CRISPR knockout screen. The GFP reporter lines were transduced with lentiviral encoded sgRNA library and selected with puromycin. GFP sorting was done on day 14 to segregate GFP high and low populations for calculating sgRNA enrichment score. **B-C)** Enriched KEGG pathways for depleted (negative selected) sgRNA from reporter cells at day 14 compared to day1. sgRNA targeting important cell survival pathways (dark grey bars) were depleted in both HCT116-Mut-GFP (B) and U251-Mut-GFP (C) cells at day14. **D)** Representative FACS histograms of GFP sorted reporter cells (HCT116 WT-GFP-left: HCT116-Mut-GFP - middle: U251-Mut-GFP -right) before and after Tv3 sgRNA library expression. The population with high GFP fluorescence were sorted as GFP high while the remaining populations were sorted into GFP low for subsequent genomic DNA isolation to look for sgRNA enrichment in the populations. **E)** Schematic showing the comparison scheme for sorted cells to find activators of *hTERT*. For activators in HCT116 isogenic reporter lines HCT116-WT-GFP and HCT116-Mut-GFP the comparison was made between same population of either GFP high or GFP low to look for depleted or enriched sgRNA targeting genes, respectively. For non-isogenic line (U251-Mut-GFP), sgRNA depleted in GFP high Vs. GFP low was derived to look for negatively selected sgRNA in GFP high population. **F)** Plot showing the p-value distribution of sgRNA targeted genes and top hits of activator genes (GNB2L1, SMG9, TP53, C7orf26, CYP2A13, POP5, TRUB2, PLK1, PPA1 and DUSP4) and known regulators of *hTERT* among all hits (NFKB2, CPSF4, STAT3, HNRNPK and SUPT5H) identified from the HCT116-Mut-GFP cell line. **G)** Plot showing the p-value distribution of sgRNA targeted genes and top hits of activator genes (UBE3D, HIRA, NXT1, GTF2H5, PIN1, UBE2A, NHSL2, BEX5, API5 and RBMXL3) and known regulators of *hTERT* among all hits (HIF1A, KLF2, HNRNPD, SIRT1 and GLI2) identified from U251-Mut-GFP cells line.

[0076] **Fig. 5** shows a comparison of screen hits and validation. **A)** Venn diagram showing the overlap of all combined hits from various cell lines done either by CRISPR sgRNA library or siRNA library. The MED complex genes were identified as common among many different lines. **B)** Telomerase activity assay was measured in non-reporter T98G and U251 cell lines after transient knockdown (with 2 siRNAs) of 3 selected hits from the overlap of screen data. **C)** qPCR data showing the expression of *hTERT* in non-reporter T98G and U251 cell lines after transient knockdown (with 2 siRNAs) of DOCK11, POLR2L and MED12. Actin was used as a control. **D)** Non-reporter T98G and U251 cell lines were transfected with siControl (siCTRL) and *siMED12* (#1 and #2). Western blot image shows the protein levels of GABPA and ETS1-2 in T98G and U251 cell lines after transient knockdown of MED12 (with 2 siRNAs). HSP90 was used as a loading control. **E)** qPCR data showing the expression of *hTERT* and

MED12 in different cell lines (A549, MCF-7 and Hep3B) with WT *hTERT* promoter after transient knockdown (with 2 siRNAs) of MED12. **F)** Telomerase activity assay was measured in WT *hTERT* promoter lines after transient knockdown (with 2 siRNAs) of MED12. Error bars indicate mean $\pm$ SD of three independent experiments. P values were calculated by Student's

5 t-test method (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

[0077] **Fig. 6** shows MED12 regulates Mut-*hTERT* promoter-driven gene expression. **A)** Graph shows the relative telomerase activity of T98G-WT and T98G-Mut cells with (*siMED12* #1 and *siMED12* #2) and without (*siCTRL*) MED12 depletion by RT-qPCR. Relative telomerase activity was quantified as compared to *siControl* sample. **B-C)** Gene expression

10 analysis was performed for *hTERT* and MED12 genes by qPCR in T98G-WT and T98G-Mut cells transfected with *siControl* or *siMED12*. Ct values were normalized to actin gene. **D)** Graph shows relative telomerase activity in BLM-WT and BLM-Mut cells transfected with *siControl* or *siMED12*. **E-F)** Gene expression analysis was performed for *hTERT* and MED12

15 genes by qPCR in BLM-WT and BLM-Mut cells transfected with *siControl* or *siMED12*. **G-H)** ChIP-qPCR was performed for MED12, GABPA and IgG in T98G-Mut and BLM-Mut cells transfected with *siControl* and *siMED12*. Enrichment in the proximal *hTERT* promoter was calculated by using % input method. **I)** Chromatin interaction frequency between Mut-*hTERT* promoter and T-INT1 region was analysed by 3C-qPCR in T98G-Mut and BLM-Mut cells transfected with *siControl* and *siMED12*. Error bars indicate mean $\pm$ SD of three independent

20 experiments. P values were calculated by Student's t-test method (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). **J)** Picture illustrates the *hTERT* promoter with its distal regulatory elements. In the absence of *hTERT* promoter mutation (left panel), *hTERT* promoter and T-INT1 interaction does not form as GABPA and MED12 is not bound to the WT-*hTERT* promoter. Upon acquiring cancer-specific *hTERT* promoter mutation, GABPA and MED12 binds to the Mut-

25 *hTERT* promoter and mediate long-range chromatin interaction between Mut-*hTERT* promoter and T-INT1 region which recruits Pol II to the promoter to transcribe *hTERT* expression.

[0078] **Fig. 7** shows **A)** Schematics of CRISPR genome editing in *hTERT* gene to insert GFP or NanoLuc gene. Either GFP or NanoLuc gene is inserted under the control of endogenous *hTERT* promoter in exon 1 with in-frame fusion. **B)** Gene expression of *hTERT*

30 and **C)** telomerase activity was measured in all our reporter cell lines after successfully inserting promoter mutation (C228T/C250T) and reporter gene (GFP/NLuc). Actin was used as a control gene to normalise *hTERT* expression. **D)** Chromatogram images showing the mutated residues in the promoter of *hTERT* gene in GFP reporter lines (top) and NanoLuc reporter lines (bottom). **E)** Chromatogram images showing in-frame fusion of GFP in reporter

lines HCT116-Mut-GFP and U251-Mut-GFP (top); NanoLuc gene in reporter lines LN382-Mut-NLuc, DLD-1-Mut-NLuc and HCT116-Mut-NLuc (bottom).

[0079] **Fig. 8** shows **A)** Scatter plot showing signals from PrestoBlue in 3 different replicates of DLD-1-Mut-NLuc cells treated with siRNA library. **B)** Scatter plot showing luminescence-based NanoLuc readings (obtained using Nano-Glo luciferase assay system) from different replicates from DLD-1-Mut-NLuc cells treated with siRNA library. **C-D)** Scatter plot showing FLuc and NLuc activity from LN382-Mut-NLuc cells after siRNA treatment. FLuc activity (left) is a measure of cell viability while NanoLuc activity (right) is a measure of the *hTERT* promoter activation. **E-F)** Scatter plot showing FLuc and NLuc activity from HCT116-Mut-NLuc cells after siRNA treatment. FLuc activity (left) is a measure of cell viability while NanoLuc activity (right) is a measure of the *hTERT* promoter activation.

[0080] **Fig. 9** shows **A)** Box plot showing the sgRNA count (in log2 scale) across various conditions and harvesting days from HCT116-WT-GFP cells (left) and HCT116-Mut-GFP cells (right). **B)** Box plot showing the sgRNA count (in log2 scale) across various conditions and harvesting days from U251-Mut-GFP cells (right). The sgRNA representation is equal across various days and conditions used in both cell lines. **C)** Venn diagram showing the overlap of hits obtained from cell lines with the same *hTERT* mutation status (C228T). The common overlapped genes are listed in a box. The known regulators are GABPA, GABPB1 and Myc, while the novel hits are DOCK11, MED6, MED12 and POLR2I. **D)** Venn diagram showing the overlap of hits obtained from cell lines with different *hTERT* mutation statuses (C250T and C228T). The common overlap genes are listed in a box. The known regulators are GABPA, GABPB1 and Myc while the novel hits are DOCK11, MED6, MED7 and MED27. **E)** Cell cycle analysis was performed in T98G isogenic cell lines after transient knockdown (with 2 siRNAs) of MED12. X-axis represents PI staining and y-axis represents cell count. For Mut-*hTERT* promoter cell line (T98G-C250T), the analysis was performed at early (p1) and late passage (p10) with continuous knockdown (with 2 siRNAs) of MED12. **F)** Telomere length assay was performed in T98G isogenic cell lines after continuous transient knockdown (with 2 siRNAs) of MED12 for two months. The mean telomere length of each sample is indicated below (TRF: Telomere restriction fragment).

[0081] **Fig. 10** shows **A)** qPCR data showing the expression of *GABPA* in non-reporter T98G and U251 cell lines with Mut *hTERT* promoter after transient knockdown (with 2 siRNAs) of MED12. **B)** Gene expression analysis was performed for the MED12 gene by qPCR in T98G-WT and T98G-Mut after stable MED12 knockdown (with 2 shRNAs). Ct values were normalized to actin gene. **C)** Western blot image shows the protein levels of MED12 in T98G-Mut lines after stable MED12 knockdown (with 2 shRNAs) and rescued with MED12

overexpression. HSP90 was used as a loading control. **D)** Telomerase activity assay was measured in T98G-Mut lines after stable MED12 knockdown (with 2 shRNAs) and rescued with MED12 overexpression. Error bars indicate the mean $\pm$ SD of three independent experiments. P values were calculated by Student's t-test method (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). **E)** Image showing the colony formation by crystal violet staining of T98G-WT and T98G-Mut cells after MED12 knockdown (with 2 shRNAs) and rescued with MED12 overexpression.

[0082] **Fig. 11** shows **A-B)** ChIP-qPCR was performed for MED12 and GABPA in T98G-WT **(A)** and BLM-WT **(B)** cells transfected with siControl and *siMED12*. Enrichment in the proximal *hTERT* promoter was calculated by using % input method. **C)** CoIP was performed in isogenic T98G lines against endogenous MED12 and IgG. The fractions were separated on SDS-PAGE and probed with anti-MED12 and anti-GABPA. **D-E)** ChIP-qPCR was performed for MED12, GABPA and IgG in T98G-Mut **(D)** and BLM-Mut **(E)** cells transfected with siControl and *siGABPA*. Enrichment in the proximal *hTERT* promoter was calculated by using % input method. **F)** Chromatin interaction frequency between Mut-*hTERT* promoter and T-INT1 region was analysed by 3C-qPCR in WT *hTERT* promoter lines transfected with siControl and *siMED12*. Error bars indicate the mean $\pm$ SD of three independent experiments. P values were calculated by Student's t-test method (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

[0083] **Fig. 12** shows the Editing strategy for developing reporter assay for quantification of MED12 expression.

[0084] **Fig. 13** shows the agarose gel electrophoresis result of pUC19 and MED12 repair construct.

[0085] **Fig. 14** shows the Sanger sequencing result of MED12 repair template cloned into pUC19 plasmid.

[0086] **Fig. 15** shows Removing puromycin and replacing hygromycin in the pLKO-TET-On cassette. Agarose gel electrophoresis image shows successful incorporation of hygromycin to the pLKO-TET-On plasmid (Lane1). Negative control and undigested plasmids are shown in lanes 2 and 3, respectively.

[0087] **Fig. 16** shows amplification of CRISPR library for CRISPR screening. **A-B)** Representative bacterial plates are shown with various dilutions. **C)** Number of colonies is indicated in the table obtained from bacterial plates.

[0088] The invention is described in greater detail below.

[0089] The invention illustratively described herein may suitably be practiced in the absence of any element or elements, limitation or limitations, not specifically disclosed herein.

Thus, for example, the terms "comprising", "including", "containing", etc. shall be read expansively and without limitation. Additionally, the terms and expressions employed herein have been used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the inventions embodied therein herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention.

[0090] **METHODS AND MATERIALS**

[0091] Telomerase Activity Assay (TRAP)

[0092] TRAP assay was performed from cells 36h of post-transfection as described in Akincilar SC, *et al.* (2015) Quantitative assessment of telomerase components in cancer cell lines. *FEBS Lett* 589(9):974-984, which is hereby incorporated by reference in its entirety.

[0093] Statistical Analysis

[0094] The two-tailed Student's t-test was used to analyse groups for the qPCR, 3C-qPCR, ChIP-qPCR and TRAP assay. The mean $\pm$ SD of each gene expression, ChIP, chromatin interaction and TRAP assays were obtained from at least three independent experiments, as indicated in the figure legends.

[0095] Generation of firefly luciferase reporter cell lines

[0096] pLenti CMV Puro LUC (w168-1) was a gift from Eric Campeau & Paul Kaufman (Addgene plasmid # 17477, RRID: Addgene_17477). The firefly luciferase expression constructs and ViraSafe™ lentiviral packaging system (Cell Biolabs) were co-transfected into HEK293T cells using Lipofectamine 3000. After 24 hrs, the medium was replenished with the full growth medium and incubated for another 24 hrs to harvest the viral particles. The viral particles were collected and target cells were infected in the presence of 8 μ g/mL of polybrene (Santa Cruz Biotechnology). Firefly luciferase expressing nano-luciferase (NLuc) reporter cells were selected with Puromycin (Gibco).

[0097] Whole genome siRNA screening

[0098] Silencer™ Select Human whole-genome siRNA library (Invitrogen) was used for whole-genome siRNA screening against the reporter cell lines. This library targets >21,000 human genes arrayed in microplates with each well containing a pool of 3 siRNAs targeting a particular gene. The siRNAs were first dispensed into white, 384-well plates (Corning) at a

concentration of 200 nM and 5 μ L per well. Silencer™ Select Negative Control No. 1 and No. 2 siRNAs (Invitrogen) were included as negative controls. Silencer™ KIF11 siRNA (Invitrogen), Silencer™ Firefly Luciferase siRNA (Invitrogen) and a customized Silencer™ Select nano-luciferase siRNA (Invitrogen) were also included as positive controls. Each gene was tested in triplicate.

[0099] Reverse transfection was carried out by first dispensing 5 μ L of Opti-MEM™ I Reduced Serum Medium (Gibco) containing 0.1 μ L of Lipofectamine™ RNAiMAX Transfection Reagent (Invitrogen) into each well containing siRNA. The transfection reagent and siRNAs were allowed to complex for 20 min at room temperature before addition of 50 μ L per well of cells in DMEM (Gibco) containing 12% FBS (Gibco) to yield final transfection mixtures containing about 17 nM siRNA in DMEM and 10% FBS. The cell numbers used for each cell line were 2,000 cells per well for DLD-1-Mut-NLuc and HCT116-Mut-NLuc, and 4,000 cells per well for LN382-Mut-NLuc. Cell viability and nano-luciferase reporter activity were assayed 72 hr after reverse transfection. For DLD-1-Mut-NLuc cells, cell viability was measured with PrestoBlue cell viability reagent (Invitrogen) and nano-luciferase reporter activity was measured with Nano-Glo Luciferase assay system (Promega) and following the manufacturer's protocols. For HCT116-Mut-NLuc and LN382-Mut-NLuc cells, cell viability and nano-luciferase reporter activity were measured with Nano-Glo Dual-Luciferase assay system (Promega), and following the manufacturer's protocols. Signals were detected using Tecan Infinite M1000 or Perkin Elmer EnSpire multimode plate readers.

[00100] siRNA screening data analysis

[00101] Readouts from the whole genome siRNA screens were uploaded onto the ChIP-GIS database portal (<https://www.chip-phenomics.org/portal.pl>). Percentage toxicity and percentage inhibition was calculated using the following formula:

$$[00102] \quad \% \text{Toxicity or \%Inhibition} = \left(1 - \frac{(\text{RFU or RLU})_x}{(\text{RFU or RLU})_y} \right) \times 100\%$$

where x represents the signal from each well containing siRNA and y represents the median signal of the negative control siRNAs in the same plate. A cumulative plot of all the %Inhibition scores for all the siRNAs investigated was then obtained for each cell line and a %Inhibition cutoff was determined by identifying the inflection point. Genes for which the siRNAs resulted in %Inhibition score greater than or equal to the cutoff were plotted in a second cumulative plot based on their %Toxicity scores and a cutoff for %Toxicity score was determined similarly as for the %Inhibition cutoff.

[00103] Simultaneously, the readouts from the whole genome siRNA screens were analysed using GUITars, a siRNA screen analysis tool based on the SSMD method. Wells containing negative control siRNAs were selected as negative controls while wells containing

KIF11 siRNAs were selected as positive controls for both cell viability and reporter inhibition. For HCT116-Mut-NLuc and LN382-Mut-NLuc cells, wells containing firefly luciferase siRNAs were also selected as a second positive control. Hits were identified based on a Toxicity cutoff of SSMD ≥ -3 and an Inhibition cutoff of SSMD < -3 . Overlapping hits identified from both the analysis methods were then selected as the final hits for each cell line.

[00104] siRNA secondary screening

[00105] The siRNA secondary screening was carried out in a similar fashion to the primary whole-genome screen. The difference between the secondary screen and the primary screen was that in the secondary screen, siRNAs against hit genes identified from the primary screen were cherry-picked and arrayed in white, 384-well plates. Cell lines used for the secondary screen include DLD-1 WT-NLuc (2,000 cells per well), and DLD-1 Mut-NLuc (2,000 cells per well). Data analysis for the secondary screen was carried out by calculating the percentage toxicity and percentage inhibition scores as per the primary screen.

[00106] Lentivirus production of sgRNA library

[00107] The Toronto KnockOut Library v3 (TKOv3) (Addgene plasmid #90294, a gift from Jason Moffat) was purchased. The library of sgRNAs were amplified followed by lentiviral preparation of the sgRNA library in 293FT cells as described in Mair B, *et al.* (2019) Essential Gene Profiles for Human Pluripotent Stem Cells Identify Uncharacterized Genes and Substrate Dependencies. *Cell Rep* 27(2):599-615 e512, which is hereby incorporated by reference in its entirety. Briefly, 293FT cells were seeded at a density of 3×10^6 cells/10cm² plate. The next day, cells were transfected with a mixture of pRSV-Rev (2.5 μ g), pMDG.2 (2.5 μ g), pMDLg/pRRE (7.5 μ g), TKOv3 plasmid library (10 μ g), and Lipofectamine LTX (42 μ l), following the manufacturer's protocol. At 8 hr after transfection, the medium was changed to the growth medium (DMEM, 10% FBS, 1% penicillin/streptomycin). Virus-containing medium was harvested 24 hr and 48 hr after transfection and centrifuged at 1200 rpm for 5 min. Viral titers in HCT116 cells were determined by infecting cells with a titration of TKOv3 lentiviral library in the presence of polybrene (8 μ g/ml). 24 hr post infection, puromycin (2 μ g/ml) containing medium was added onto cells to select the transduced cells. After 72 hours, the multiplicity of infection (MOI) of the titrated virus was determined by comparing the percent survival of infected cells to non-infected control cells.

[00108] Generation of CRISPR cell lines and screening

[00109] Target cell lines (HCT116-WT-GFP, HCT116-Mut-GFP and U251-Mut-GFP) were grown in 15cm plates and a total of 40×10^6 cells were infected with TKOv3 lentiviral library (71,090 gRNAs) at an MOI of ~ 0.5 to achieve ~ 200 -fold coverage of the library after selection. Polybrene at 8 μ g/ml was added to increase transduction efficiency. All screenings were

performed in duplicates. After 24 hr, the virus-containing medium was removed and replaced with growth medium (DMEM, 10% FBS, 1% penicillin/streptomycin) supplemented with puromycin (1 µg/ mL) and cultured for 5 days. The cells were maintained at 20×10^6 cells at each passage every 3–4 d to have ~200-fold coverage. Post selection, a total of 20×10^6 cells were collected for genomic DNA extraction at day 0 and day 14. 20×10^6 cells were fluorescence sorted by FACS (BD FACSAria II) on day 14 to GFP low and GFP high. The cells were collected and grown until they reach a 20×10^6 cells and were harvested for genomic DNA isolation. 20 million cells post-CRISPR screen were used to isolate genomic DNA using QIAGEN DNA Maxi kit (Qiagen), according to the manufacturer's instruction. Concentration of the isolated DNA was quantified using Qubit 2.0 fluorometer (Thermo Fisher Scientific).

[00110] Library preparation and sequencing of CRISPR knockout libraries

[00111] To generate Illumina libraries for deep sequencing, 2 PCR reactions were performed. (1) enrich guide-RNA regions in the genome and (2) amplify guide-RNA with i5 and i7 indices. Both PCR reactions were performed using Q5 Hot Start High-Fidelity 2 × Master Mix and forward/reverse primers pair (v2.1-F1 and v2.1-R1, Table 2) in a 50 µl reaction. For PCR1, 40 independent PCR reactions were set up for each sample using 2.5 µg of the genomic DNA. Using this concentration, the library is maintained at 200-fold coverage. The PCR conditions were as follows: 98°C for 30 s, 98°C for 10 s, 66°C for 30 s and 72°C for 15 s (X25 cycles), and the final extension, 72°C for 2 min. The PCR products (600bp) were resolved on 1% agarose gel and individual reactions were pooled. About 5 µl of pooled PCR products from all 40 reactions and purified using QIAquick PCR purification kit (QIAGEN). The concentration of purified PCR products were quantified using the Qubit 2.0 fluorometer (Thermo Fisher Scientific). 5 µl of the PCR pooled purified product was used in a second PCR performed with 2x NEBNext Ultra II Q5 Master Mix. 2.5 µl of different i5 primer and 2.5 µl of different i7 primers (Table 2) were used. The PCR conditions were as follows: 98°C for 30 s, 98°C for 10 s, 55°C for 30 s and 65°C for 15 s (10–12 cycles), and the final extension, 65°C for 5 min. The PCR product (200bp) was excised after running on 2% agarose gel. Finally, QIAquick Gel Extraction Kits (Qiagen) were used to purify the PCR products. Purified libraries were resuspended in 50 µl nuclease-free water and analysed using Agilent High Sensitivity DNA kit (Agilent Technologies) on Agilent 2100 Bioanalyzer (Agilent Technologies).

[00112] Purified libraries of the duplicate screens tagged with different i5 and i7 tags were pooled and sequenced at about 300x coverage on Illumina HiSeq 2500. About 30–40 million reads were obtained for each replicate.

[00113] **Table 2.** List of primers used to amplify the sgRNA from genomic DNA and primers used to create the sgRNA library for Illumina sequencing.

SEQ ID NO	Primer name	Sequence
5	v2.1-F1	GAGGGCCTATTTCCCATGATTC
6	v2.1-R1	GTTGCGAAAAAGAACGTTACGCG
7	D501	AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCTTTCCCTA CACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
8	D502	AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCTTTCCCTA CACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
9	D503	AATGATACGGCGACCACCGAGATCTACACCCTATCCTACACTCTTTCCCTA CACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
10	D504	AATGATACGGCGACCACCGAGATCTACACGGCTCTGAACACTCTTTCCCTA CACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
11	D505	AATGATACGGCGACCACCGAGATCTACACAGGCGAAGACACTCTTTCCCT ACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
12	D506	AATGATACGGCGACCACCGAGATCTACACTAATCTTAACACTCTTTCCCTA CACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
13	D701	CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTACTTGCTATTTCTAGCTCTAAAC
14	D702	CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTACTTGCTATTTCTAGCTCTAAAC
15	D703	CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTACTTGCTATTTCTAGCTCTAAAC
16	D704	CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTACTTGCTATTTCTAGCTCTAAAC
17	D705	CAAGCAGAAGACGGCATACGAGATTCTGAATGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTACTTGCTATTTCTAGCTCTAAAC
18	D706	CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTACTTGCTATTTCTAGCTCTAAAC
19	D707	CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTACTTGCTATTTCTAGCTCTAAAC

[00114] CRISPR data processing

[00115] Quality control step was performed on sequencing reads by using Trimmomatic v0.32 to filter low-quality bases and adapter sequences. Reads that passed quality control were used as input in downstream analysis using MAGeCK v0.5.9 tool. Read counts corresponding to single-guided RNAs (sgRNAs) were first quantified and normalized using median normalization. Subsequently, the relationship of replicates was learned by mean-variance modeling and the statistical significance of each sgRNA was calculated using the learned mean-variance model. Essential genes were then positively and negatively selected (p<0.01) using robust rank aggregation (RRA) method. Essential genes identified in different cell lines and screens (CRISPR and RNAi) were compared to identify common genes. Finally, by reference to KEGG (Kyoto Encyclopedia of Genes and Genomes) database, enriched pathways (p<0.05) were identified by applying the gene set enrichment analysis (GSEA) algorithm to the ranked list of genes.

15 [00116] RNA isolation and gene expression analysis

[00117] Total RNA was isolated from cells after 36h of transfection by Trizol method followed by column purification. cDNA synthesis was performed from 1µg RNA using iScript cDNA synthesis kit, BioRad. RT-qPCR analysis was performed using diluted cDNAs (5ng/reaction) by SSO-Sybr Greener qPCR master mix (BioRAD). Expression values were

5 normalized to actin and relative expression was calculated by delta-delta Ct method.

[00118] Chromatin Immunoprecipitation (ChIP) Assays

[00119] Cells were harvested 36h post-transfection and cross-linked for 10 minutes with 1% Formaldehyde. After isolation of nuclei fraction, sonication was performed by 30 sec on 30 sec off intervals to obtain DNA smear between 150bp -750 bp. Soluble chromatin fraction was

10 pre-cleared with Protein-A/G beads and incubated over-night with IgG (Sant Cruz), anti-MED12 (Bethyl A300-774A) and anti-Gabpa (Santa Cruz Biotechnology; sc-22810) antibodies. ChIP-qPCR was performed with hTERT promoter-specific primers (forward: GCGGCGCGAGTTTCAG, reverse: AGCACCTCGCGGTAGTGG)

[00120] TRF Assay

15 [00121] TRF assay was performed in isogenic T98G cell line using TeloTAGGG telomere length assay kit (Sigma, Cat No: 12209136001) following manufacturer instruction. Briefly, the cells were harvested and counted. 2×10^6 cell were used to lyse and extract pure genomic DNA using The Wizard® Genomic DNA Purification Kit (Promega, Cat No: A1120). The extracted DNA was quantified using Nanodrop. 4µg of total DNA was used in the assay. The

20 telomere specific probes were used to bind to the digested DNA after they are transferred to the nylon membrane and UV-crosslinked. The DEG labelled probes are visualised using HRP substrate and the signal intensity was quantified using image J software.

[00122] Cell cycle analysis

[00123] Cell cycle analysis was performed by DNA content analysis using propidium iodide (PI)/RNase staining. Briefly, the cell were transfected with siRNA using RNAiMAX reagent. After specific time point, the cells were harvested and fixed with absolute ethanol over night. Next day, the cell were pelleted and washed 3 times. PI was used at 1:5000 dilution in PBS and incubated with cells for 30-60 mins. After the incubation, the cells were analysed using

25 FACS with specific gates using 532-nm excitation with a 585/42-nm bandpass filter.

30 [00124] 3C- chromatin interaction assay

[00125] Cells were fixed with 2% formaldehyde for 10 mins and lysed with cold lysis buffer (50mM Tris-HCL pH7.5, 150mM NaCl, 5mM EDTA, 0.5% NP-40, 1% Triton X-100) on ice for 20 mins. Isolated nuclei were resuspended in DNase-free water and chromatin was digested using HindIII enzyme for 36 hours. Fragmented chromatin was ligated overnight at 16C and

35 after phenol-chloroform extraction, 3C template DNAs were quantified using Qubit dsDNA

High Sensitivity Kit (Thermo Fischer). 3C-qPCR was performed using primers specific to *hTERT* promoter and the *T-INT1* region as described in Akincilar SC, *et al.* (2016) Long-Range Chromatin Interactions Drive Mutant TERT Promoter Activation. *Cancer Discov* 6(11):1276-1291, which is hereby incorporated by reference in its entirety.

5 [00126] Constructing Med12 Repair Template

[00127] MED12 repair construct (MED12 promoter-RFP-T2A-MED12 gene) was ordered as a gene block from IDT with the EcoR1 and Xba1 restriction sites at the 5 and 3 primes. 200ng of gene block was digested using EcoR1 and Xba1 enzymes at 37°C for 1 hour. In parallel pUC19 (1µg) plasmid was also digested to create flanking sites. After purification of gene block and pUC19 from agarose gel, insert and plasmid were ligated overnight at 16°C and transformed into competent cells. Plasmid isolation was performed using a plasmid-mini kit (Qiagen) according to the manufacturer's instructions. Cloning was validated by agarose gel and sanger sequencing.

[00128] Constructing inducible shRNA plasmid

15 [00129] The puromycin gene in the pLKO-TET-ON plasmid was disrupted by digesting the plasmid with BamHI and CsiI restriction enzyme. The hygromycin gene was amplified from another plasmid, pLV-EF1a-IRES-Hygro, and then digested with the same enzymes, BamHI and CsiI. The digested hygromycin was then ligated with the digested pLKO-TET-On plasmid and transformed into DH5α competent cells. The colonies were picked and inoculated for plasmid isolation. The isolated plasmid was confirmed for the hygromycin integration using sanger sequencing.

20 [00130] Amplification of genome-wide CRISPR libraries

[00131] Toronto V3 and Gecko genome-wide libraries were amplified according to the manufacturer's instructions. Briefly, the gRNA plasmid pool was electroporated into electrocompetent bacteria (Lucigen) cells. Bacterial cultures were grown overnight at 32°C and plasmid purification was performed using a Plasmid-midi isolation kit (Qiagen). DNA concentrations were measured using Nanodrop.

[00132] **EXAMPLES**

30 [00133] Non-limiting examples of the invention and comparative examples will be further described in greater detail by reference to specific Examples, which should not be construed as in any way limiting the scope of the invention.

[00134] Shanmugam, R. et al (2022). Genome-wide screens identify specific drivers of mutant *hTERT* promoters. *Proceedings of the National Academy of Sciences*, 119(3) is hereby incorporated by reference in its entirety.

[00135] Example 1: Generation of *hTERT* reporter cell lines

[00136] To comprehensively identify physiological regulators of wild-type (WT) and mutant (Mut) *hTERT* promoters, using CRISPR/Cas9 mediated genome editing, several isogenic reporter cells were generated that carry either WT, or C228T/C250T Mut promoters in the endogenous *hTERT* locus (Fig. 1A). In these isogenic cells with 3 different versions of *hTERT* promoters (WT/C250T/C228T), either GFP or luciferase reporter genes were knocked-in to utilize them in genome-wide loss of function screens (Fig. 1B). Both reporter genes (GFP or luciferase) were inserted downstream of the *hTERT* transcription start site (TSS) (Fig. 7A). To identify *bona-fide* regulators of WT and Mut-*hTERT* promoters, irrespective of cancer type, several colon cancer and glioblastoma lines were generated as representatives of cancers that are at the opposite spectrum of *hTERT* promoter mutation frequencies. Viable engineered cells from 4 different cell lines representing these two cancer types were obtained (Fig. 1B). It is important to note that GFP and luciferase reporter genes were introduced only in one of the alleles as it has been reported that 5' tagging of *hTERT* gene results in decreased *hTERT* expression. It was evident that the remaining unedited copy was sufficient to maintain *hTERT* expression (Fig. 8B-C). Appropriate editing of the promoter mutations in the engineered lines was confirmed by Sanger sequencing (Fig. 7D). In addition to confirming the precise editing of the sites, the in-frame insertion of GFP and Nano luciferase genes in *hTERT* loci downstream of the promoter was sequence-verified using Sanger sequencing (Fig. 7E). Next, GFP and luciferase expression were checked and significant and detectable levels of GFP (Fig. 2A) or Luciferase (Fig. 2B) were found in the reporter lines indicating that in-frame inserted reporter genes are faithfully driven by the endogenous *hTERT* promoter. Here onwards the engineered cell lines are represented based on the following nomenclature *i.e.* cell line (original name)- WT/Mut (*hTERT* promoter)-GFP/NLuc (reporter-gene). To evaluate if the reporter genes are functionally driven by the endogenous *hTERT* promoter and its relevant chromatin context, these reporter lines were transfected with siRNAs targeting known regulators of *hTERT* such as Myc and GA Binding Protein Transcription Factor Alpha (GABPA). Myc binds to the two E-box sequences in *hTERT* proximal promoter and positively regulates *hTERT*. Cancer-specific mutations in the proximal *hTERT* promoter create binding sites for GABPA, binding of which enhances *hTERT* transcription. Depletion of GABPA in the GFP reporter lines led to a significant reduction in GFP expression (Fig. 2C). Similarly, transient knockdown of Myc in luciferase reporter cell lines, led to a significant reduction of luciferase signal (Fig. 2D). Furthermore, a significant reduction in the mRNA levels of Myc and luciferase genes were observed upon Myc knockdown (Fig. 2E). Similarly, the functionality of

GFP reporter cell lines was evaluated by transient knockdown of GABPA. Both reporter cells with siRNAs targeting GABPA showed reduced levels of GABPA and GFP (Fig. 2F). Collectively, these observations indicate the functionality and usability of our engineered reporter lines with promoter mutations for genome-wide loss/gain of function screens. These
5 unique resources are the first such resources generated and could be used for many chemical and genomic screens in the future.

[00137] Example 2: Whole genome siRNA screen to identify regulators of *hTERT* promoter

[00138] For the comprehensive and unbiased discovery of regulators of *hTERT*, whole-
10 genome siRNA screens were performed in three cell lines which survived the editing process and could be expanded for large-scale screening. These lines, DLD-1-Mut-NLuc, LN382-Mut-NLuc, and HCT116-Mut-NLuc (Fig. 1B), are representative of different cancer types (i.e. DLD-1 vs LN382) and different *hTERT* promoter mutation status (i.e. DLD-1 harbours the C228T while HCT116 harbours the C250T mutations) and it was reasoned that common hits from
15 these lines would likely help to narrow down the genuine regulators of Mut-*hTERT* promoter. The workflow for each siRNA screen is illustrated in Figure 3A. Briefly, on Day 0, cells were seeded and reverse-transfected with arrayed siRNAs in 384-well plate format for 3 days. Signals obtained using PrestoBlue reagent had a poor correlation between replicates (Fig. 8A). In contrast, luminescence-based NanoLuc readings (obtained using Nano-Glo luciferase
20 assay system) to measure *hTERT* promoter activity displayed an excellent correlation between replicates (Fig. 8B). Consequently, for the LN382-Mut-NLuc and HCT116-Mut-NLuc siRNA screens, the inventors switched to using cytomegalovirus (CMV) promoter-driven firefly luciferase (FLuc) expression to quantify and normalize the cell viability. This is also a luminescence-based readout similar to the NanoLuc reporter, except that the substrates used
25 are different. Results from the subsequent screens using LN382-Mut-NLuc and HCT116-Mut-NLuc cells showed an excellent correlation between replicates with FLuc activity as a measure of cell viability (Fig. 8C-F). To identify hits, a combination of percentage activity scores cutoff (cutoff method) and strictly standardized means difference (SSMD) method were used and these results were overlapped for common hit identification. The main consideration in the
30 selection of hits was that the siRNAs should have potent inhibition of the *hTERT* promoter-driven reporter activity (i.e. NanoLuc) while having minimal effects on the viability of the cells (i.e. low toxicity). This is to enable us to identify specific and direct regulators of *hTERT* promoter activity. In the first approach, the percentage toxicity (%Toxicity) and percentage *hTERT* promoter inhibition (%Inhibition) scores were calculated for each gene in the siRNA
35 screen by normalizing their signals against the respective signals obtained from negative

control siRNA (siNeg)-treated cells. A cumulative plot of all the %Inhibition scores for all the siRNAs investigated was then obtained for each cell line (Fig. 3B). Based on these cumulative plots, cell line-specific differences were observed in the effects of siRNA on *hTERT* promoter activity. This was expected as each cell line screened has inherently different mutational and gene expression profiles due to their tissue of origin. Based on the % inhibition cumulative plots, a cutoff was decided for each cell line based on the point of inflection as this point marks the boundary between specific and non-specific effects (Fig. 3B). Genes for which the siRNAs resulted in % inhibition scores greater than or equal to the cutoff (35% for DLD-1-Mut-NLuc, 20% for LN382-Mut-NLuc, and 85% for HCT116-Mut-NLuc) were plotted a second cumulative plot based on their % toxicity scores (Fig. 3B). From the second cumulative plot, a cutoff for % toxicity score was determined similarly as for the % inhibition cutoff. Genes for which the siRNAs resulted in % toxicity scores less than or equal to the cutoff (75% for DLD-1-Mut-NLuc, 50% for LN382-Mut-NLuc, and 95% for HCT116-Mut-NLuc) were selected. Consequently, based on this cutoff method, 593, 370, and 894 hits were identified for DLD-1-Mut-NLuc, LN382-Mut-NLuc, and HCT116-Mut-NLuc cells respectively (Fig. 3C).

[00139] Simultaneously, the siRNA screen results were analysed using GUItars, a siRNA screen analysis tool based on the SSMD method. For each cell line screened, 2 SSMD scores were obtained for each gene investigated; one SSMD score for toxicity (toxicity SSMD score), and one SSMD score for *hTERT* promoter activity inhibition (inhibition SSMD score). The use of SSMD as the criteria for hit selection has been reported as a means to control for false positive and false negative rates. The SSMD score obtained is also a measure for the strength of the effect with SSMD scores > 3 representing very strong activation or upregulation effects and SSMD scores < -3 representing very strong inhibition or downregulation effects. As mentioned above, the criteria used for hit selection was that the siRNAs should have potent *hTERT* promoter activity inhibition with minimal toxicity. Since a toxicity SSMD score < -3 indicated strong toxic effects, we set a cutoff of ≥ -3 for toxicity SSMD score. Conversely, for *hTERT* promoter activity inhibition, siRNAs with inhibition SSMD score of < -3 were selected (Fig. 3A). Based on these criteria, 1666, 1878, and 2410 hits were identified for DLD-1-Mut-NLuc, LN382-Mut-NLuc, and HCT116-Mut-NLuc cells respectively (Fig. 3D). Combining the hits identified from the cutoff method and the SSMD method, 556, 236, and 111 common hits were found for DLD-1-Mut-NLuc, LN382-Mut-NLuc, and HCT116-Mut-NLuc cells respectively (Fig. 3D). Pathway enrichment analysis of the common hits for each cell line revealed similarities across the cell lines investigated. Amongst the top 15 enriched pathways for each cell line, pathways involved in transcription and development biology (Fig. 3E, black bars) are common across all 3 cell lines. Interestingly, in the LN382-Mut-NLuc cell line, chromatin

organization and modification pathways were identified (Fig. 3E; middle panel, dark grey bars). Pathway enrichment of the hits identified in the HCT116-Mut-NLuc cell line also identified hematopoietic stem cell (HSC) gene regulation by GABPA/B (Fig. 3E; bottom panel, dark grey bar) as one of the significantly enriched pathways. This was particularly reassuring as this

5 corroborates previous findings on the importance of GABPA in the regulation of *hTERT* expression specifically by Mut-*hTERT* promoters. The hits identified from DLD-1-Mut-NLuc and LN382-Mut-NLuc siRNA screens were compared and 64 common hits were found (Fig. 3F). Pathway enrichment of the 64 common genes also identified HSC gene regulation by GABPA/B as one of the enriched pathways, as well as the interactome of the polycomb

10 repressive complex 2 (Fig. 3G, dark grey bars).

[00140] **Table 3.** Gene symbol and IDs of top hits of inhibitors identified from secondary siRNA screen performed in DLD-1 cell line with WT or C228T mutant hTERT promoter.

	Gene Symbol	Gene ID	WT Toxicity	WT Inhibition	C228T Toxicity	C228T Inhibition	Mutant specific inhibition score
1	C7orf26	79034	-34.43	-259.95	-44.88	-52.58	-207.36
2	MED12	9968	86.98	-32.70	78.53	57.02	89.71
3	CHD4	1108	31.35	-35.91	31.09	46.79	82.70
4	TNRC18P1	644962	40.19	-13.51	50.87	47.58	61.09
5	SNRPD2	6633	-4.68	-3.61	-34.28	51.72	55.33
6	TAF7	6879	-35.29	-52.09	-8.28	-1.12	50.96
7	EIF3C	8663	75.56	4.95	70.98	54.48	49.53
8	GNB2L1	10399	28.93	-22.97	47.22	19.45	42.42
9	RG9MTD1	54931	-33.14	-45.36	-5.80	-7.80	37.57
10	DPP9	91039	36.92	-19.52	35.01	17.86	37.38
11	RPL35	11224	78.22	18.38	81.01	54.56	36.17
12	PSMD12	5718	65.63	9.45	60.18	42.68	33.23
13	Luc	Luc	98.62	-35.60	97.38	-3.64	31.96
14	LOC73027	730271	60.12	-5.36	37.69	25.12	30.48
15	SNRPD3	6634	28.23	25.87	-18.81	51.24	25.36
16	RAPGEF4	11069	68.29	24.99	73.93	48.80	23.81
17	PMVK	10654	-65.37	2.62	-32.55	25.08	22.46
18	MORN3	283385	16.19	27.04	16.76	49.42	22.38
19	Neg1	Neg1	17.70	-13.55	-3.75	8.04	21.59
20	KIAA2013	90231	33.66	24.43	32.01	45.24	20.81
21	FAM19A3	284467	66.49	31.86	56.92	51.96	20.11
22	GABPA	2551	21.04	28.36	-4.95	48.42	20.07
23	CYP4F3	4051	-1.36	-7.75	-2.73	12.12	19.87
24	EDDM3B	64184	10.67	18.38	12.29	37.81	19.43

After top hits were identified from the initial siRNA screen, a secondary siRNA screen was performed in isogenic cell line (DLD-1) with WT or C228T promoter mutation in *hTERT* promoter. Table 3 indicates the top hits identified from this secondary screen where MED12 is the topmost identified hit with known function in cells. GABPA, a known activator of Mut *hTERT* promoter, also came up in the hit, thus serving to validate the screen performance.

[00141] Example 3: Whole genome CRISPR knockout screen to identify regulators of *hTERT* expression

[00142] In addition to siRNA screens which have their own limitations such as off-target effects and toxicity, Toronto v3 CRISPR knockout library containing a pool of 70948 single guide RNA (sgRNA) targeting 18053 genes in the human genome was utilized. The advantage of this library is it is an all-in-one system wherein the sgRNA and Cas9 are in the same plasmid, allowing single antibiotic selection after transduction and it is hence superior over other CRISPR knockout libraries. As outlined in Figure 4A, the screening process involves transducing the lentiviral pool of sgRNA+Cas9 in the target cells followed by puromycin selection to eliminate non-transduced cells. Sufficient cells were collected on day 1 and day 14 to compare the sgRNA representation of essential genes and pathways. On day 14, the cells were sorted based on GFP expression into GFP high and GFP low as compared to non-transduced cells. sgRNA enrichment scores from each sample were plotted to identify genes essential for *hTERT* regulation (Fig. 9A-B). Consistent sgRNA scores were observed across various samples (Fig. 9A-B). Besides, it was also observed, in both HCT116-Mut-GFP and U251-Mut-GFP cells, from the pathway terms comparing day 14 with day 1 that sgRNA targeting essential pathways like cell cycle, spliceosome, proteasome and ribosomes were significantly depleted at day 14 (Fig. 4B, C). These results uphold the functionality and usability of our CRISPR knockout screens for the identification of physiological regulators of Mut-*hTERT* promoter. The sgRNAs among GFP-high populations of HCT116-WT-GFP vs. HCT116-Mut-GFP cells were further analysed and the hits were overlapped with GFP-low population of HCT116-WT-GFP vs. HCT116-Mut-GFP cells (Fig. 4D). The sgRNAs depleted in GFP-high population most likely represent specific activators of Mut-*hTERT* promoter (Fig. 4E). CRISPR-cas9 engineered U251 line harbouring WT-*hTERT* promoter was not viable due to dramatic loss of telomerase activity. Therefore, sgRNAs that are enriched in the GFP-high were compared to GFP-low population as activators (Fig. 4E). All the activators that are specific to each cell line are reported in Table 4. The gene rank was plotted as indicative of the Mut-*hTERT* inhibition against the p-value for each of the two cells lines (Fig. 4F, G). In HCT116-GFP cells, Mut-*hTERT* promoter-specific novel hits were identified including PLK1,

DUSP4, TP53, TRUB2 and PPA1 (Fig. 4F). In addition, previously reported regulators of *hTERT* expression namely Myc, GABPA, CPSF4 and hnRNPK were found among significant top hits ($p < 0.01$) (Fig. 4F). Hits from U251-Mut-GFP cells identified genes like HIRA, GTF2H5 and PIN1 among top 10 hits while known regulators of *hTERT* expression like HIF1a, GLI2, KLF2, SIRT1 and hnRNPD came along as significant hits ($p < 0.01$) (Fig. 4G). The hits from each screen were overlapped and varying degrees of overlap was found depending on the cell lines and promoter mutation status (Fig. 9C-D).

[00143] **Table 4.** Gene IDs of all the hits identified from siRNA and CRISPR screen performed in various reporter cell lines.

siRNA screen			CRISPR knockout screen	
DLD-1-Mut-Nluc	LN382-Mut-Nluc	HCT116-Mut-Nluc	U251-Mut-GFP	HCT116-Mut-GFP
ALDOA	ABCC10	BCL9	ABCB10	ABHD5
AMD1	ABHD9	CACNA1S	ACOT9	ACBD5
APEX1	ANKAR	S100G	ADAM19	ACCSL
CCND1	APOL4	CFL1	AGRP	ACOX1
BCL9	ARHGAP22	DPYD	ALDH5A1	ACSBG2
BGLAP	ASB14	EIF2S3	ALG2	ACSF3
BOK	ASTN2	EPHB2	API5	ACTN2
ZFP36L1	ATG3	ETV2	ARL6IP1	ACVR2B
MPPED2	ATG7	GABPA	ARMCX1	AIM1L
C11orf10	B3GAT2	GABPB1	ARMCX2	AKAP6
S100G	BCL9	GZMK	ARMCX3	AKT1
CAPN3	BGLAP	HRG	ARMCX6	AKTIP
CAPN6	C19orf18	IL17A	ASCL4	APOBEC1
CAPS	C19orf39	PDX1	ASGR2	ARL13B
RUNX1T1	C19orf44	LHX1	ATG4A	ATP13A4
CCND2	C19orf47	SMAD2	ATP1B4	B9D1
CDH8	C1orf70	MCM2	ATP6V1B2	BCL7A
CDK4	C3orf62	MYC	AWAT2	BCL7C
CDK7	C5orf22	RPL10A	BEX4	BIRC2
CDR2	C9orf135	NTF4	BEX5	C17orf82
CDX1	C9orf9	ODC1	C11orf83	C1orf137
CFL1	CACNA1S	OR1F1	C1GALT1C1	C21orf59
CHD4	CACUL1	PDK2	C3orf55	C7orf55
CNN1	CALML4	PFDN4	C9orf50	CALCRL
COL4A2	CARD8	PLN	CABIN1	CAMTA2
COPA	CASTOR1	PPP2R3A	CAPN6	CBLN1
COX11	CCDC181	CCL4	CAPRIN1	CCDC146
CSNK1A1	CCDC67	SGCD	CAST	CCDC6
CSNK1G2	CCDC88A	SLC15A2	CCDC115	CCND3
CTNNB1	CCDC92	SNRPD2	CCDC22	CD180

DAB1	CDR2L	ZNF19	CCDC83	CD4
DAZL	CHCHD5	ZNF79	CDC42BPG	CD96
DBI	CHL1	CCDC6	CDC45	CDC37L1
DIO1	CHST5	PRPF18	CDCA8	CDCA4
DPYD	CHTF18	SART1	CDX4	CDH6
DSC3	CLEC19A	NCR2	CHIC1	CHRNA6
DUSP8	COL6A6	MED27	CHRD1	CLSTN2
E4F1	COPG1	MED7	CINP	CMIP
EEF1A1	CPLX3	SH3BP5	CITED1	CPXM1
EIF1AX	CPSF6	ZNF432	CLEC5A	CREG1
EIF2S1	CTRC	SDC3	COPA	CSTF3
EIF2S3	CXorf22	SART3	COPS5	CYP4F3
EIF5	CYBASC3	IQSEC1	CPXCR1	DAK
ELK3	DDX19B	MED6	CRB3	DAPK2
ENDOG	DDX21	AVIL	CT47B1	DEAF1
STOM	DDX43	DDX19B	CTAG2	DEPDC4
EPHB2	DEFB112	PUF60	CTDSPL2	DPYSL5
ETF1	DENND1B	COTL1	CTHRC1	EPS15
EZH1	DIO1	BACE2	CUL4B	FAM111A
FAAH	DLX4	SOSTDC1	CXCR3	FAM122A
FAH	DNAJC7	VSX1	CXorf65	FAM221A
FBN1	DOCK11	ATP6V0A4	CXorf66	FAM69A
FGB	DOK4	RNF138	DACH2	FCER1G
FGFR2	DPP9	RIN2	DCAF12L2	FLAD1
FOXF1	DPYD	CCSER2	DCX	FNBP1L
FLII	DPYSL2	RRN3	DDX39B	FNDC1
FLNA	EHMT1	ZNF586	DDX55	FNIP2
GABPA	EIF4G3	TRPM7	DGAT2L6	FOXL1
GABPB1	EP400	SVOP	DOCK11	FRAS1
GBA	EPC1	ZNF83	DPPA3	FSD1L
GDF2	EPHB2	SLC22A11	DRP2	GAB1
GPC1	ESPNL	UBQLN4	EDA	GABBR1
HELLS	EZH1	PRDM8	EDDM3B	GDPGP1
HLA-DQB2	F11	ACN9	EFNB1	GFI1B
HNRNPH1	FAAH	C16orf62	EIF3B	GLRA4
PRMT1	FABP2	NAT14	EP300	GORASP1
HSPA5	FAM120AOS	ZNF250	ERCC3	GPR108
HSPA9	FAM98B	HERPUD2	EXOSC5	GPR135
IL5	FBXO11	MRPL36	FAM199X	GVQW1
IL18	FGF9	ACD	FAM19A2	H2AFV
IMPDH1	FKBP9	PPDPF	FAM19A3	HHEX
ITGB6	FOLR4	FCRL2	FAM26F	HN1L
JARID2	GABPA	TMCO7	FAM46D	HOXB5
KCNJ15	GABPB1	SEL1L2	FBXL14	HOXD12

KRAS	GFOD2	OR2B2	FDFT1	HSPB6
KTN1	GLYATL1	SF3B5	FITM2	HYI
LLGL1	GPC1	HS6ST2	FOXK2	IKBKG
SMAD2	GPR112	ERI1	FRK	IPO4
MCM2	GPX8	TMEM88	FRMD7	JPH1
MPST	GTF3C4	RG9MTD2	GABRR2	JUP
MYC	H1FOO	PLCD3	GGCX	KATNAL2
NCK1	HAND1	LRRC15	GGPS1	KCNA7
RPL10A	HCFC1	DOCK11	GK	KCNAB1
NONO	HCN1	MDP1	GLRX3	KCTD14
NTN3	HEY1	CCDC17	GLUD2	KIAA1211L
NUP98	HEYL	FAAH2	GPC3	KIF3C
OPRK1	HIST2H3A	ZNF836	GPR174	KLHL7
P2RX3	HTT	DENND1B	GREM2	KLRG2
FURIN	IKBKG	C1orf177	GRIA3	KRIT1
PDGFRB	INPP5F	C1QL2	GRWD1	KRTAP10-4
PKD2	ISM2	OR5AS1	GTF2E2	LAMB2
PGF	ISOC1	RNF169	GTF2H5	LCA5
PML	ITGB6	WDR51B	H2BFM	LRFN4
POLR2C	ITSN2	NKAIN3	H2BFWT	LRRC46
POLR2E	JARID2	VN1R5	HIF1A	LRRK1
POLR2G	KCNA3	TCEAL5	HINFP	LZIC
POLR2I	KCNH2	PALM3	HIRA	MAFB
PPP2R3A	KIAA2013	OTOP3	HNRNPH2	MAPKAPK5
PRIM1	KLC4	TRIM77	HTR2C	MCOLN2
PRKAR2A	L1TD1	OR5AU1	HTR3E	MESDC2
PRSS1	LAMA3	RPL18P4	IL13RA1	MEX3B
PSMA1	LCE1F	TEX19	IL13RA2	MORN3
PSMA5	LGALS9B	C14orf53	IRS4	MPND
PSMA6	LHX9	LOC646570	ITGB1BP2	MSX1
PSMA7	LIME1	LOC648672	ITM2A	MTR
PSMB3	LINC01501	RPL7P9	KAT6A	MTRR
PSMB6	LOC401442	BCL9	KCNK10	NCAN
PSMB7	LOC646570	CACNA1S	KHSRP	NDST4
PSMC1	LOC649975	S100G	KIAA2022	NFKBIB
PSMC2	LOC727894	CFL1	KLHL22	NOL7
PSMC3	LOC728685	DPYD	KLHL4	OAF
PSMC5	LOC729296	EIF2S3	KRT13	ODF3
PSMC6	LOC729421	EPHB2	LAMP2	OMP
PSMD2	LOC729823	ETV2	LDLR	OR51A7
PSMD3	LOC729827	GABPA	LMAN1	OR7D2
PSMD7	LOC730271	GABPB1	LPCAT1	OTOA
PSMD11	LOC732028	GZMK	LRRTM4	PAQR9
PSMD12	LOXL4	HRG	LSMEM2	PDS5A

PTPRB	LYZL4	IL17A	LTV1	PIM1
RFC2	MCM2	PDX1	LYPD6	PJA1
BRD2	MDS1	LHX1	MAPKAPK3	PLCE1
RPLP1	MED6	SMAD2	MATR3	PLEKHF2
RPLP2	MPND	MCM2	MAVS	PLOD2
RPS3	MYF5	MYC	MCEE	POU2F3
RPS3A	NAA10	RPL10A	MED12	PPFIA3
RPS4X	NAT14	NTF4	MED19	PPM1D
RPS5	NCOA2	ODC1	MED27	PRICKLE3
RPS6	NDRG3	OR1F1	METTL23	PROM2
RPS7	NDUFS7	PDK2	MFSD12	PRPH
RPS8	NGEF	PFDN4	MPL	PSD4
RPS9	NKAIN3	PLN	MRPL41	PTGS1
RPS10	NPIPB7	PPP2R3A	MSN	PTPN22
RPS11	NRSN1	CCL4	MTMR8	PTPRZ1
RPS13	OR13G1	SGCD	MX1	RAB39B
RPS14	OR51T1	SLC15A2	NABP2	RBP3
RPS15	OR52E6	SNRPD2	NAGK	REPS2
RPS27	OR8B2	ZNF19	NDP	RGL4
RPS27A	OTOP3	ZNF79	NHSL2	RGS12
SDC4	PABPC4L	CCDC6	NSMCE2	RGS18
ST3GAL1	PAK4	PRPF18	NUTF2	RHNO1
SLC1A3	PALM3	SART1	NXF5	RND3
SLC4A2	PALMD	NCR2	NXT1	RPL12
SLC8A3	PCNT	MED27	OPHN1	S100Z
SLC15A2	PCSK2	MED7	OTOS	SAT2
SLC16A1	PHF16	SH3BP5	OTUD6A	SEMA4C
SMARCB1	PHF21A	ZNF432	P2RY10	SGPP2
SNRPD1	PHKB	SDC3	P2RY4	SKIL
SNRPD2	PIWIL2	SART3	PATL2	SLC25A43
SNRPD3	PMVK	IQSEC1	PCDH19	SLC5A5
SNRPF	PPWD1	MED6	PCDHGB1	SNAPC4
SON	PRC1	AVIL	PCGF6	SPATA24
SUPT5H	PRDM8	DDX19B	PIAS1	SPDYA
TAT	PRKAR2A	PUF60	PIGO	SPNS2
TBX1	PRPF18	COTL1	PIGV	SRSF2
TKT	PRR12	BACE2	PIN1	ST8SIA4
TPR	PRR17	SOSTDC1	PIN4	STRIP2
UBTF	PRR4	VSX1	PLA2G12A	SVOPL
UPK1B	PRRX2	ATP6V0A4	PLEKHM1	TBL2
UQCRC2	PRSS1	RNF138	PLP1	TCHHL1
WEE1	PSG1	RIN2	PNMA5	THBS3
ZNF19	PSMA7	CCSER2	PNPLA4	TICRR
ZNF79	PTPRB	RRN3	POF1B	TLR1

ZNF84	RAD50	ZNF586	POLR2I	TMEM220
ZNF193	RAPGEF4	TRPM7	POU3F2	TMEM235
MLL2	RASGRP1	SVOP	PRPS1	TMPPE
AP3B2	RBPMS2	ZNF83	PRR7	TNNT2
SMC1A	RILP	SLC22A11	PRRG3	TSKU
SLC10A3	RNF138	UBQLN4	PSMD10	TTLL10
HIST3H3	RPLP2	PRDM8	RAB40A	UCK2
HIST1H3J	RXRG	ACN9	RAB9B	VWA3A
MAD1L1	S100G	C16orf62	RBM41	WDR41
NIPSNAP1	S100PBP	NAT14	RBMXL3	WDR86
IFITM1	SAMD13	ZNF250	RIPPLY1	XYLT1
CAMK1	SART3	HERPUD2	RNF128	YOD1
BARX2	SEC22B	MRPL36	RNPS1	ZBTB10
CGGBP1	SH3TC2-DT	ACD	RPL30	ZBTB47
CDC14B	SIGLEC12	PPDPF	RPS19	ZFP92
PRPF18	SLC43A1	FCRL2	RRN3	ZNF197
PDLIM4	SMAD2	TMCO7	RTP3	ZNF268
RUVBL1	SON	SEL1L2	S1PR2	ZNF384
NCOA1	SPINT2	OR2B2	SAPCD2	ZNF575
DDX3Y	SRMS	SF3B5	SATL1	ABCA2
EIF3A	SSTR1	HS6ST2	SEC11A	ABHD2
EIF3B	STOM	ERI1	SEC63	ACAN
EIF3C	SUCLG1	TMEM88	SEMA5B	AHNAK
EIF3G	SYNGR1	RG9MTD2	SH3BGRL	AHR
EIF3I	TCEAL6	PLCD3	SLC15A4	ALDH6A1
GALNT4	TFAP2E	LRRC15	SLC16A2	ALG14
SUCLG1	TFPI2	DOCK11	SLC25A53	ALOXE3
CES2	TMCC2	MDP1	SLC29A2	ANAPC13
TOP3B	TMEM98	CCDC17	SLC33A1	ANTXR1
HIST1H2BJ	TMEM9B	FAAH2	SLC38A10	AP1G2
SYNGR1	TMSB15A	ZNF836	SLC7A3	AP2S1
DDX21	TMTC3	DENND1B	SLK	APOO
MED14	TOMM40	C1orf177	SNRPB2	ARMC1
NDST3	TRAK1	C1QL2	SNRPD2	ASB6
MED21	TRAP1	OR5AS1	SPATA6	ATF7IP
DDX23	TRIM50	RNF169	SPINK2	ATP5A1
HAND1	TRIM68	WDR51B	SPTLC1	ATP5SL
MED17	TRPM7	NKAIN3	SRPX2	ATP6AP2
MED7	TTC22	VN1R5	STX17	BACH2
STX8	TTPAL	TCEAL5	SYMPK	BCAP29
CTR9	TUBB4B	PALM3	SYTL4	BRD2
MDC1	TUBGCP4	OTOP3	TAC1	C11orf91
ZNF432	TXLNB	TRIM77	TADA1	C16orf92
NUP93	UBXN11	OR5AU1	TAF5L	C18orf21

RAPGEF2	UGT2B15	RPL18P4	TAF9B	C19orf70
SART3	UPF1	TEX19	TBC1D12	C2orf68
CKAP5	VWF	C14orf53	TBL1X	C7orf26
NUPL1	WDR51B	LOC646570	TBPL1	C8orf74
TMCC2	WFIKK1	LOC648672	TBX22	CABP1
MED12	WIPF1	RPL7P9	TCEAL1	CAD
MED6	WNK2		TCEAL7	CAP1
CHAF1A	YWHAB		TERT	CCDC106
PDIA6	ZNF187		TEX11	CCDC39
PSMD14	ZNF250		THOC6	CCT7
PEMT	ZNF418		TIMM8A	CD6
TOMM40	ZNF432		TMED2	CDC37
HAX1	ZNF552		TMEM230	CDC5L
NXF1	ZNF778		TMEM259	CDIPT
IPO8	ZNF799		TMEM31	CDT1
TNFSF13B	ZNF836		TMEM35	CEP57
AVIL			TMEM47	CHMP1A
TOB2			TNMD	CHRM1
RUVBL2			TOX4	CHST8
GPR83			TREX2	CIRH1A
RAB10			TRMT10C	CLGN
ERP29			TRMT2B	CLVS2
CPSF6			TRPC5	CMPK1
OGFR			TSPAN6	CPZ
RAPGEF4			TSR3	CUL4A
TOPBP1			UACA	CYP2A13
CLASRP			UBE2A	DDA1
RPL35			UBE2H	DHX37
DDX19B			UBE3D	DHX9
COPZ1			USHBP1	DPY19L1
CARD8			USP15	DPYS
TRAK1			UTP23	DUSP4
SEPHS1			VASN	EEF1G
CDK19			VSIG1	EIF2B5
WDR43			VSIG4	EIF3G
NUP205			WDR78	ENC1
SF3B3			WDTC1	ENDOD1
SF3B1			WFDC8	ENDOG
TMEM50A			XPOT	ENKUR
NUP62			YIF1A	ERCC5
PLD3			YIPF6	ETNK1
IL17RA			YJEFN3	EYA1
BRD1			YPEL5	FAM186B
MTCH2			ZC3H14	FAM20B

FBXO7			ZC4H2	FAM47E
SPDEF			ZCCHC12	FAM57B
TXN2			ZCCHC16	FBXO11
DAK			ZNF280C	FERD3L
ACOT11			ZNF414	GABRB1
ATRNL1			ZRANB2	GABRB3
SLITRK5				GDAP1L1
GIGYF2				GJC2
CCDC9				GNB2L1
GTPBP5				GNPDA1
OR2F1				GRB10
HEYL				H2AFX
MTBP				HBM
CECR6				HBZ
TRIB2				HDGF
CCDC113				HIRIP3
NOSIP				HOXA9
CUTC				HSPB9
FCF1				IFNB1
TUBD1				IGLL1
THEM6				IL17D
POMP				IL27
AIG1				INO80
ANAPC5				IPO11
RNF138				ISCA1
LARS				ISCU
C14orf100				ITPRIP
AZIN1				KARS
PCF11				KCNQ4
ZNF44				KDM5C
LUC7L3				KLF17
SSH1				L3MBTL2
STAG3L1				LARP4B
CCSER2				LAT2
SMCR7L				LCE3C
APBB1IP				LDB3
HES2				LECT1
GATAD2A				MADCAM1
TRPM7				MAML1
C1orf26				MAP2K1
ASPN				MCCD1
WDR55				MDM2
COMMD8				MED10
MED9				MERTK

AKIRIN2				METTL14
RBM23				MKI67IP
CEP55				MKS1
TYW1				MMP26
TMEM100				MORF4L1
DDX43				MRPL21
SUPT20H				MRPL34
RAB20				MRPL51
DENND4C				MRPS23
FRMD4A				NCL
DOK4				NDUFAB1
MUDENG				NEK11
SLC22A11				NID2
PCDHGA1				NOL9
BEX4				NUP133
CLDND1				OR5AC2
TMEM9B				PAFAH1B3
UBQLN4				PALB2
MCCC1				PARD6B
DUSP22				PCDHA3
XAB2				PDCL
OTUD7B				PGD
KIF15				PHYH
C16orf62				PIGL
ADAMTSL3				PLIN1
NHSL1				PLK1
SMAGP				PLXNB2
NDRG3				PMVK
GATAD2B				PNMAL2
KIDINS220				POLD2
STARD9				POLD3
TMEM181				POLR1E
WDR19				POP5
ZNF250				POP7
ARHGAP22				PPA1
GNB4				PPP1R11
PBOV1				PROM2
AASDHPPT				PRPS2
ELAC2				PSMD12
RIC8A				RAB11B
ZFAND3				RAN
MRPL17				RASGRP1
SH2D4A				RASSF8
CARD9				RBM23

EDDM3B				RHOT2
LPIN3				RIOK3
UBE2Z				RNGTT
CYP4F12				RPL11
LYNX1				RPL23
GID4				RPL35
PDCL3				RPS2
PPDPF				RSU1
IRX1				RTCA
SAP130				RUSC2
TMEM53				SAE1
ZNF552				SCO1
ARHGAP28				SERPINB2
PIP4K2C				SLC25A47
RMI1				SLC38A2
CCDC92				SLC50A1
ABTB1				SLX4
KCNIP4				SMG9
NIPA2				SMPD2
OR2B2				SMUG1
TTC25				SNRPD3
PARP9				SNX12
WDR87				SNX22
ING5				SPC25
C14orf142				SRPRB
C6orf168				SRXN1
CYorf15B				STK40
MYO18B				SV2B
CBX2				TAF7
MPV17L2				TECPR1
PARP10				TFAM
LRP11				TFB1M
MAEL				TLN1
C12orf62				TMC5
STON2				TNNT2
WNT3A				TNPO3
LMLN				TP53
NAV2				TPI1
C9orf140				TPT1
KIAA2013				TRAIP
DCAF15				TRIB1
ZNF799				TRMT10C
ANKRD40				TRPM1
RFT1				TRUB2

IMP4				TSFM
ZBTB47				TTC30A
C13orf27				TWIST1
RG9MTD2				USP34
SERHL				WARS2
NACC1				WDR3
PANX3				WDR33
OR51A7				WDR66
AK7				WRAP53
C19orf47				XPO1
TPRG1L				YRDC
C1orf216				YWHAE
EDARADD				ZNF280C
IQGAP3				
C1orf88				
CST9LP1				
CIB4				
SCLT1				
C9orf135				
UPRT				
GPR119				
DOCK11				
CACUL1				
RPL14L				
C13orf31				
HAPLN3				
C17orf77				
AMAC1				
C19orf84				
COMMD7				
CAMSAP1				
TTLL11				
KIAA2026				
C9orf84				
PATE1				
ZNF600				
FAM47B				
CLEC4C				
GPHA2				
DAND5				
DHRS7C				
TMEM192				
CDHR3				
NEGR1				

NALCN				
TAS2R50				
MDGA1				
WDR51B				
C11orf72				
MORN3				
LGALS9B				
BOD1P				
C19orf54				
TPRX1				
OR2Z1				
ZNF844				
FAM19A3				
ERICH2				
RPS26P40				
PRSS38				
LINC01139				
ESPNL				
KLHL38				
POTEA				
STAC2				
ZNF404				
BARHL2				
RSHL3				
ZNF322B				
TMEM212				
GRXCR1				
USP27X				
OR5B12				
LOC390335				
THA1P				
PRAMEF8				
RPS12P4				
RPS15AP17				
LOC391722				
LINC01061				
RPS26P10				
LOC401677				
RPL15P11				
OR4A47				
LOC440055				
RPS26P35				
TSPAN11				
RPS16P1				

RPS15AP9				
PRR9				
POTEM				
FAM230B				
RPS26P28				
LOC643950				
FNDC10				
BCRP3				
LOC644828				
RPS26P50				
LOC644982				
RPL17P39				
C14orf34				
RPS26P31				
LOC646051				
RPS8P4				
RPS28P7				
LOC646561				
LOC646570				
LOC646736				
LOC646831				
LOC647089				
UQCRFS1P2				
LOC647315				
LOC648672				
LOC648894				
LOC648934				
LINC00273				
LOC651732				
LOC652728				
LOC652788				
RPS23P8				
LOC727786				
LOC727804				
LOC727894				
C21orf122				
DUX4L16				
LOC728200				
GTF2F2P2				
LOC728721				
LOC729036				
LOC729398				
LOC729592				
LOC729622				

LOC729720				
LOC729807				
LOC729833				
RPSAP19				
LOC730271				
LOC731517				
HGC6.3				
FAM196B				
TIMM23				
THEGL				

[00144] Example 4: Hit comparison, validation of MED12 as mutant specific *hTERT* regulator

[00145] After creating a list of activators for each cell line used in our screens utilizing either

5 CRISPR knockout library or siRNA library, the hits were overlapped to identify if there are any common regulators. This analysis revealed 30 genes that surfaced as common hits between the two different screens performed in various isogenic cell lines (Fig. 5A). Among the common activators, the focus was on genes related to transcription factors or DNA binding proteins, including mediator complex subunit (MED) genes, DOCK11 and POLR2I. Mediator complex

10 subunits are important for gene transcription and have been shown to play role in controlling expression of cell identity genes like Sox9, Nanog. In recent years, MED complex has been linked to the pathogenesis of cancer and is frequently mutated in different cancers. Particularly, *MED12* has a high frequency of mutation in phylloid cancer and *MED12* gene has been reported to play a role in the control of cell identity genes and super-enhancer associated

15 gene expression. However, the mechanism by which MED12 functions in cancers in general and cancers with Mut-*hTERT* promoter remains elusive. With this backdrop on the identified common hits, 3 candidates namely DOCK11, POLR2I, and MED12 were further analysed for their ability to controlling Mut-*hTERT* promoters. Transient knockdown of these candidate genes in parental (non-reporter) T98G and U251 cell lines containing inherent C250T and

20 C228T mutations, respectively showed a significant reduction of telomerase activity upon MED12 knockdown (Fig. 5B). *POLR2I* knockdown significantly reduced telomerase activity only in T98G line (Fig. 5B). A similar trend was observed in expression of endogenous *hTERT* mRNA indicating that MED12 exerts control of *hTERT* at gene expression level supporting its reported role in the control of gene expression (Fig. 5C). As a proof of concept that our screens

25 can identify *bona-fide* regulators of Mut-*hTERT* promoter, the results prompted further validation of MED12 as a candidate. To evaluate the mechanism of action of MED12, the influence of MED12 on the expression of well-known activators of Mut-*hTERT* promoters such

as ETS1/2 and GABPA was first analysed. Knocking down MED12 in both T98G and U251 lines did not show a reduction of either of these positive regulators of Mut-*hTERT* promoters (Fig. 5D and Fig. 10A). Instead, increased expression of GABPA in U251 lines was observed with MED12 knockdown. While MED12 has been shown to regulate Mut-*hTERT* promoters in two different cell lines, its influence on WT-*hTERT* promoter was investigated. Depletion of MED12 by siRNAs in cell lines A549 (lung cancer), MCF-7 (breast cancer) and Hep3B (liver cancer) which harbour WT-*hTERT* promoter, showed no significant effect on *hTERT* expression and telomerase activity (Fig. 5E-F). These observations further reiterate that MED12 is a specific positive regulator of Mut-*hTERT* promoters.

[00146] Example 5: MED12 specifically regulates Mut-*hTERT* promoter via long-range chromatin interaction

[00147] Having identified MED12 as a positive regulator that specifically activates Mut-*hTERT* promoters in parental lines with inherent C228T and C250T promoter mutations, the precise mechanism by which MED12 regulates Mut-*hTERT* promoter was further evaluated using non-reporter isogenic T98G (glioma) and BLM (melanoma) lines with WT or Mut-*hTERT* promoters. These isogenic lines had been instrumental in demonstrating the importance of *T-INT1* region in regulation of Mut-*hTERT* expression. In these lines, depletion of MED12 using siRNAs inhibited telomerase activity specifically in Mut-*hTERT* promoter isogenic variants of both T98G and BLM cells (Fig. 6A and 6D). A significant decrease of *hTERT* expression was observed in these isogenic lines only with the Mut-*hTERT* promoter (Fig. 6B and 6E). It is important to note that this specific effect of MED12 knockdown on the isogenic mutant line is not simply due to differences in knockdown efficiency of MED12 in the isogenic lines (Fig. 6C and 6F). It is also not related to the immediate effect of MED12 knockdown on cell cycle (Fig. 9E). Interestingly, this effect is attributed to reduced telomere length specifically in Mut-*hTERT* promoter cells (Fig. 9F). Colony formation assays also showed that upon MED12 knockdown, significant reduction in colony formation is observed particularly in Mut-*hTERT* promoter cells and this was rescued by restoration of *hTERT* activity by ectopic expression of MED12 (Fig. 10C-E). Next, the role of MED12 in epigenetic control of mutant *hTERT* promoter was analysed by chromatin occupancy and chromatin interaction analyses. GABPA has been shown to bind to the *de novo* C250T and C228T sites to drive *hTERT* expression. GABPA dimers on the proximal *hTERT* promoter have been shown to interact with GABPA dimers located in a distal region (~260kb upstream of *hTERT* promoter) and form a long-range chromatin interaction to make a stable transcriptional hub. Therefore, it was first tested if MED12 influences GABPA occupancy on the proximal *hTERT* promoter in isogenic T98G and

BLM cells. ChIP-qPCR results demonstrated a significant reduction of GABPA occupancy on the proximal *hTERT* promoter upon MED12 depletion in both T98G-Mut and BLM-Mut cells (Fig. 6G-6H and Fig. 11A-B). Given that MED12 regulates GABPA occupancy, biochemical interaction between them was checked by coimmunoprecipitation analysis. Using isogenic

5 T98G cells, GABPA and MED12 was shown to interact irrespective of *hTERT* promoter status suggesting that this complex might regulate other genes in the genome (Fig. 11C). As these proteins interact, MED12 occupancy after siGABPA was tested in T98G-Mut and BLM-Mut lines. MED12 occupancy is reduced in these lines after siGABPA treatment (Fig. 11D-E). Interestingly, MED12 was identified to bind to the proximal *hTERT* promoter and its occupancy

10 was reduced upon MED12 depletion (Fig. 6G-6H and Fig. 11A-B). Further, depletion of MED12 was observed to reduce interaction between Mut-*hTERT* proximal promoter and T-INT1 chromatin interaction region as reduction of MED12 causes dissociation of GABPA subunits from the proximal *hTERT* promoter (Fig. 6I and Fig. 10F). In absence of GABPA recruitment to proximal Mut-*hTERT* promoter, the long-range interaction with T-INT1 region is

15 disrupted and hence the productive 3D chromatin architecture along with the necessary chromatin modifiers required for activation of *hTERT* does not occur. Together, results from this study identify MED12, as a novel and specific regulator of Mut-*hTERT* promoter (Fig. 6J). These results also suggest that the isogenic reporter lines described here could serve as

20 wonderful resources for discovering context and stimuli-specific genetic and chemical regulators of *hTERT* in cancers.

[00148] Example 6: Targeting of telomerase activity in cancer using MED12 inhibition

[00149] Mutations in *hTERT* promoter is seen in over 19% of human cancers, irrespective of the cancer type. As described above, the inventors devised the first genome-wide functional

25 screen and identified 30 specific regulators of Mut-*hTERT* promoters. MED12 is one among them which was functionally characterised in the inventors' study. Interestingly genetic depletion of MED12 led to decrease in telomerase activity in cells with Mut-*hTERT* promoters. However, MED12 is a protein in the kinase sub-unit of MED complex that regulate generic POL II driven transcription. Due to this limitation, there were no attempt of targeting of MED12

30 by small molecule. However, recent findings show that MED12 regulates a subset of gene preferentially. Moreover, MED12 is mutated in different cancers and its over expression is seen in therapeutic resistance. MED12 mutations are mostly gain of function that can enhance telomerase activity in cancers when it co-occurs with mutant *hTERT* promoter. Hence, targeting MED12 by small molecule might proves beneficial for arresting cancer growth. A

small molecule library screen has been devised to identify a candidate for clinical development.

[00150] Example 7: Developing MED12 reporter assay

5 [00151] One direct approach for targeting the MED12 function is to halt its expression transcriptionally. To identify the molecules and drug candidates, a MED12 reporter assay was generated that quantifies MED12 promoter activity. In particular, a DNA construct was generated in which red fluorescent protein (RFP) was inserted after the translation start site of MED12 (Fig. 12). The RFP gene will quantify the MED12 promoter activity; hence the
10 molecular inhibitors that regulate MED12 expression can be identified. The repair construct is to be inserted into the genome with Crispr editing technique.

[00152] The repair construct has been successfully cloned into a carrier plasmid puc19 (Fig. 13). Successful insertion was also validated by Sanger sequencing using primers that cover the gene-plasmid junction (Fig. 14).

15

[00153] Example 8: Developing MED12-dependent *hTERT* reporter assay

[00154] A second approach to identify MED12-specific *hTERT* inhibitors involves the perturbation of MED12 levels in the cell. The inventors' preliminary result showed that long-term MED12 inhibition prevents cancer cells from proliferating and doesn't provide sufficient
20 time to perform high-throughput screens. One alternative approach is to generate inducible shRNAs specific to the MED12 gene. Inducible shRNAs are only activated upon external stimuli that allow sufficient inhibition to perform screening experiments and eliminate cytotoxicity as knockdown is not constitutive. However, commercial shRNA plasmids are not compatible with the Crispr libraries as they harbour the same antibiotic resistance gene.
25 Therefore, the puromycin resistance gene was replaced with hygromycin in the inducible shRNA plasmid and subsequently cloned MED12 targeting shRNAs (Fig. 15).

[00155] Example 9: Amplification of genome-wide Crispr screen libraries

[00156] Using the aforementioned reporter systems, genome-wide screens were performed
30 to identify MED12 inhibitors. Two genome-wide libraries were used identify regulators of the MED12 gene. Crispr libraries were amplified to obtain a sufficient amount of gRNA clones (Fig. 16). These libraries were utilized to make lentivirus which are used to infect cells with the various reporter systems mentioned above. Hit candidates identified from these screens will be further validated by molecular biology experiments using non-reporter lines.

[00157] As stem cells harbour WT hTERT promoter, small molecule hTERT inhibitors have been unsuccessful in clinical trials as these wipe out stem cells before killing cancer cells. Mut-hTERT promoters drive 1 in 5 human cancers. Advantageously, selectively targeting Mut-hTERT promoter allows the inhibition of telomerase activity specifically in cancer cells
5 harbouring the Mut-hTERT promoter, while leaving stem cells intact. As such, the inhibition of MED12, which is a specific positive regulator of Mut-hTERT promoter, can treat cancer by inhibiting the telomerase activity specifically in cancer cells.

[00158] While embodiments of the invention have been particularly shown and described with reference to specific embodiments, it should be understood by those skilled in the art that
10 various changes in form and detail may be made therein without departing from the scope of the invention as defined by the appended claims. The scope of the invention is thus indicated by the appended claims and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced.

CLAIMS

1. Use of a moiety that inhibits MED12 in the manufacture of a medicament for treating cancer.
2. The use of claim 1, wherein the cancer has a MED12 mutation, a hTERT promoter mutation, or both.
3. The use of claims 1 or 2, wherein the cancer is selected from the group consisting of T-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia, colorectal cancer, gastric cancer, breast cancer, non-anaplastic thyroid cancer, central nervous system cancer, and skin cancer.
4. The use of claim 3, wherein the central nervous system cancer is glioblastoma and wherein the skin cancer is melanoma.
5. The use of any one of claims 1 to 4, wherein the moiety that inhibits MED12 is an siRNA.
6. The use of claim 5, wherein the siRNA is formed from a Dicer-substrate small interfering RNA (dsiRNA), wherein the dsiRNA comprises the sequence of:
 - (i) SEQ ID NOs 1 and 2; or
 - (ii) SEQ ID NOs 3 and 4.
7. A moiety that inhibits MED12 for use in treating cancer.
8. The moiety of claim 7, wherein the cancer has a MED12 mutation, a hTERT promoter mutation, or both.
9. The moiety of claims 7 or 8, wherein the cancer is selected from the group consisting of T-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia, colorectal cancer, gastric cancer, breast cancer, non-anaplastic thyroid cancer, central nervous system cancer, and skin cancer.
10. The moiety of claim 9, wherein the central nervous system cancer is glioblastoma and wherein the skin cancer is melanoma.

11. The moiety of any one of claims 7 to 10, wherein the moiety that inhibits MED12 is an siRNA.
12. The moiety of claim 11, wherein the siRNA is formed from a Dicer-substrate small
5 interfering RNA (dsiRNA), wherein the dsiRNA comprises the sequence of:
- (i) SEQ ID NOs 1 and 2; or
 - (ii) SEQ ID NOs 3 and 4.
13. A method for treating cancer comprising administering to a subject a therapeutically
10 effective amount of a composition comprising a moiety that inhibits MED12.
14. The method of claim 13, wherein the cancer has a MED12 mutation, a hTERT
promoter mutation, or both.
15. The method of claims 13 or 14, wherein the cancer is selected from the group
consisting of T-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia,
colorectal cancer, gastric cancer, breast cancer, non-anaplastic thyroid cancer, central
nervous system cancer, and skin cancer.
16. The method of claim 15, wherein the central nervous system cancer is glioblastoma
20 and wherein the skin cancer is melanoma.
17. The method of any one of claims 13 to 16, wherein the moiety that inhibits MED12 is
25 an siRNA.
18. The method of claim 17, wherein the siRNA is formed from a Dicer-substrate small
interfering RNA (dsiRNA), wherein the dsiRNA comprises the sequence of:
- (i) SEQ ID NOs 1 and 2; or
 - (ii) SEQ ID NOs 3 and 4.
19. A pharmaceutical composition comprising (a) a therapeutic amount of a moiety that
inhibits MED12 and (b) one or more pharmaceutically acceptable carriers and/or
35 diluents.
20. The pharmaceutical composition of claim 19, wherein the moiety that inhibits MED12
is an siRNA.

21. The pharmaceutical composition of claim 20, wherein the siRNA is formed from a Dicer-substrate small interfering RNA (dsiRNA), wherein the dsiRNA comprises the sequence of:
- 5 (i) SEQ ID NOs 1 and 2; or
 (ii) SEQ ID NOs 3 and 4.
22. A method for inhibiting telomerase activity in a cell comprising a mutant hTERT promoter, the method comprising contacting the cell with an inhibitor of MED12.
- 10 23. The method of claim 22, wherein the mutant hTERT comprises a C250T mutation or a C228T mutation.
24. The method of claims 22 or 23, wherein the inhibitor of MED12 inhibits telomerase activity via transcriptional inhibition of MED12, translational inhibition of MED12 or
- 15 post-translational inhibition of MED12.
25. The method of any one of claims 22 to 24, wherein the inhibitor of MED12 is a small interfering RNA (siRNA) that inhibits expression of MED12 mRNA.
- 20 26. The method of claim 25, wherein the siRNA is formed from a Dicer-substrate small interfering RNA (dsiRNA), wherein the dsiRNA comprises the sequence of:
- (i) SEQ ID NOs 1 and 2; or
 (ii) SEQ ID NOs 3 and 4.
- 25 27. A host cell comprising a nucleic acid construct, wherein the nucleic acid construct comprises:
- a MED12 promoter; and
- an open reading frame, wherein the open reading frame comprises a translation start site, a polynucleotide sequence encoding a reporter element, a
- 30 polynucleotide sequence encoding a self-cleaving peptide and a polynucleotide sequence encoding MED12.
28. The host cell of claim 27, wherein the open reading frame comprises a translation start site, a polynucleotide sequence encoding a reporter element, a polynucleotide
- 35 sequence encoding a self-cleaving peptide and a polynucleotide sequence encoding MED12, in that order.

29. The host cell of claim 27, wherein the open reading frame comprises a translation start site, a polynucleotide sequence encoding MED12, a polynucleotide sequence encoding a self-cleaving peptide and a polynucleotide sequence encoding a reporter element, in that order.

5

30. The host cell of any one of claims 27 to 29, wherein the reporter element is a fluorescent protein or a luminescent protein.

31. The host cell of claim 30, wherein the reporter element is red fluorescent protein (RFP).

10

32. The host cell of any one of claim 27 to 31, wherein the self-cleaving peptide is T2A.

33. The host cell of any one of claims 27 to 32, wherein the nucleic acid construct is integrated into a genome of the host cell.

15

34. A method for identifying an inhibitor of MED12, the method comprising:

(i) contacting the host cell of any one of claims 27 to 33 with a composition comprising a CRISPR-Cas system for knocking out a target gene, wherein the CRISPR-Cas system comprises:

20

(a) a guide RNA (gRNA) capable of binding to the target gene; and

(b) a polynucleotide sequence encoding a CRISPR enzyme;

(ii) quantifying expression of the reporter element in a host cell having the CRISPR-Cas system for knocking out the target gene as compared to expression of the reporter element in a control host cell without knockout of the target gene, wherein expression of the reporter element is indicative of MED12 expression, and wherein a higher expression of the reporter element in the host cell having the CRISPR-Cas system for knocking out the target gene indicates that the target gene encodes an inhibitor of MED12.

25

30

35. The method of claim 34, wherein the composition comprises a viral vector and wherein the viral vector comprises the CRISPR-Cas system.

36. The method of claim 35, wherein the viral vector is a lentiviral vector.

35

37. The method of any one of claims 34 to 36, wherein the CRISPR enzyme is Cas 9.
38. A host cell comprising a nucleic acid construct, wherein the nucleic acid construct comprises an inducible promoter operably linked to a polynucleotide sequence encoding an RNA that inhibits gene expression of MED12 via RNA interference (RNAi).
39. The host cell of claim 38, wherein the RNA is selected from the group consisting of a short hairpin RNA (shRNA), small interfering RNA (siRNA), Dicer-substrate small interfering RNA (dsiRNA), peptide nucleic acid (PNA), antisense oligonucleotide (ASO) and microRNA (miRNA).
40. The host cell of claims 38 or 39, wherein the nucleic acid construct is integrated into a genome of the host cell.
41. A method for identifying a MED12-dependent inhibitor of hTERT, the method comprising:
- (i) providing a host cell comprising an expression construct for inducible inhibition of MED12 expression;
 - (ii) contacting the host cell with a composition comprising a CRISPR-Cas system for knocking out a target gene, wherein the CRISPR-Cas system comprises:
 - (a) a guide RNA (gRNA) capable of binding to the target gene; and
 - (b) a polynucleotide sequence encoding a CRISPR enzyme;
 - (iii) quantifying hTERT expression in the host cell in which inhibition of MED12 expression has been induced as compared to hTERT expression in a control host cell without induced inhibition of MED12 expression, wherein a higher hTERT expression in the control host cell without induced inhibition of MED12 expression indicates that the target gene encodes an MED12-dependent inhibitor of hTERT.
42. The method of claim 41, wherein the host cell of step (i) is the host cell of any one of claims 38 to 40 and wherein the expression construct is the nucleic acid construct comprising an inducible promoter operably linked to a polynucleotide sequence

encoding an RNA that inhibits gene expression of MED12 via RNA interference (RNAi).

5 43. The method of claims 41 or 42, wherein the composition comprises a viral vector and wherein the viral vector comprises the CRISPR-Cas system.

44. The method of claim 43, wherein the viral vector is a lentiviral vector.

10 45. The method of any one of claims 41 to 44, wherein the CRISPR enzyme is Cas 9.