MODULATION OF GENE EXPRESSION AND SCREENING FOR DEREGULATED PROTEIN EXPRESSION

CROSS-REFERENCE

[0001] This application claims the benefit of UK Patent Application No: 1517937.7, filed on October 9, 2015, and UK Patent Application No: 1614744.9, filed on August 31, 2016, each of which are incorporated herein by reference in their entireties.

BACKGROUND OF THE INVENTION

[0002] The ATM protein belongs to the PI3/PI4-kinase family and is involved in the developments of the nervous system and the immune system. The ATM protein kinase is activated upon DNA damage and subsequently coordinates the DNA repair mechanism.

SUMMARY OF THE INVENTION

[0003] In certain embodiments, described herein include methods of screening a subject susceptible to functional-ATM protein deficiency and associated conditions, methods for selecting subjects for treatment, methods for treatment or prevention of conditions associated with functional-ATM protein deficiency, methods of modifying a cells susceptibility to DNA damaging radio- and chemotherapy, methods for treatment of cancer, and associated compositions and kits.

[0004] Disclosed herein, in certain embodiments, is a method of screening a subject for susceptibility to functional-ATM protein deficiency, wherein the screening comprises determining the presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome, wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency. In some embodiments, the NSE comprises the sequence tctacaggttggctgcatagaagaaaaag. In some embodiments, the NSE repressor agent is arranged to bind to NSE within the sequence agTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or tcttagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or tctcagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag. In some embodiments, the NSE repressor agent is arranged to bind to the NSE or its 5' or 3' splice site in ATM intron 28.

[0005] Disclosed herein, in certain embodiments, is a method of selecting a subject for treatment, wherein the subject is susceptible to functional-ATM protein deficiency, the method comprising determining the presence of a non-thymine variant residue rs609261 located at position -3 relative to

the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome, wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and selecting such subject for treatment with an agent arranged to increase functional-ATM levels in the subject. In some embodiments, the method further comprises administering the agent for treatment of the selected subject. In some embodiments, the agent comprises a NSE repressor agent. In some embodiments, the NSE comprises the sequence tctacaggttggctgcatagaagaaaaag. In some embodiments, the NSE repressor agent is arranged to bind to NSE within the sequence agTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or tcttagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or tctcagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag. In some embodiments, the NSE repressor agent is arranged to bind to the NSE or its 5' or 3' splice site in ATM intron 28.

[0006] Disclosed herein, in certain embodiments, is a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome, wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to increase functional-ATM levels. In some embodiments, the NSE comprises the sequence tctacaggttggctgcatagaagaaaaag. In some embodiments, the NSE repressor agent is arranged to bind to NSE within the sequence agTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or tcttagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or tctcagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag. In some embodiments, the NSE repressor agent is arranged to bind to the NSE or its 5' or 3' splice site in ATM intron 28.

[0007] Disclosed herein, in certain embodiments, is a method of treatment or prevention of a condition associated with a functional-ATM protein deficiency, comprising the administration of a NSE repressor agent arranged to increase levels of functional ATM protein, wherein the agent is arranged to bind to a NSE in ATM intron 28 of the pre-mRNA transcript to decrease inclusion of the NSE in the mature RNA transcript. In some embodiments, decreasing inclusion of the NSE in the mature RNA transcript provides an increase in functional ATM protein expression. In some embodiments, the method of treatment or prevention of functional-ATM protein deficiency in a subject or an at-risk population of subjects is for treatment or prevention of a condition or symptoms associated with functional-ATM protein deficiency. In some embodiments, the condition is ataxia-telangiectasia; cancer; immune deficiency; cellular radiosensitivity; or chromosomal instability. In

some embodiments, the NSE comprises the sequence tctacaggttggctgcatagaagaaaaag. In some embodiments, the NSE repressor agent is arranged to bind to NSE within the sequence agTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or tcttagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or tctcagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag. In some embodiments, the NSE repressor agent is arranged to bind to the NSE or its 5' or 3' splice site in ATM intron 28.

[0008] Disclosed herein, in certain embodiments, is a method of treatment or prevention of a condition associated with deregulation of ATM expression in a subject comprising the administration of a NSE-activator agent, wherein the NSE-activator agent is arranged to increase NSE inclusion in ATM mature RNA transcript by binding to regulatory motifs in ATM intron 28, optionally wherein the regulatory motifs in ATM intron 28 compete with NSE for spliceosomal components, and further optionally wherein such motifs comprise a 24 nucleotide pseudoexon (PE) located 3' of NSE in ATM intron 28 of the pre-mRNA transcript or binding to a U2AF65 binding site upstream of the pseudoexon.

[0009] Disclosed herein, in certain embodiments, is a method of treatment or prevention of cancer in a subject comprising the administration of a NSE-activator agent arranged to increase a cancer cell's susceptibility to cytotoxic therapy with DNA damaging agents such as radiotherapy, wherein the NSE-activator agent is arranged to increase NSE inclusion in ATM mature RNA transcript by binding to regulatory motifs in ATM intron 28, optionally wherein the regulatory motifs in ATM intron 28 compete with NSE for spliceosomal components, and further optionally wherein such motifs comprise a 24 nucleotide pseudoexon (PE) located 3' of NSE in ATM intron 28 of the pre-mRNA transcript or binding to a U2AF65 binding site upstream of the pseudoexon; and treating the subject with the cytotoxic therapy, such as radiotherapy or chemotherapy. In some embodiments, increasing inclusion of the NSE in the mature RNA transcript provides a decrease in functional ATM protein expression. In some embodiments, the pseudoexon comprises the sequence tcatcgaatacttttggaaataag.

[0010] Disclosed herein, in certain embodiments, is a method of increasing a cell's susceptibility to cytotoxic therapy with DNA damaging agents such as radiotherapy or chemotherapy comprising the reduction of ATM protein expression by administration of a NSE-activator agent arranged to increase NSE inclusion in ATM mature RNA transcript by binding to motifs in ATM intron 28, optionally wherein the regulatory motifs in ATM intron 28 compete with NSE for spliceosomal components, and further optionally wherein such motifs comprise a 24 nucleotide pseudoexon (PE)

located 3' of NSE in ATM intron 28 of the pre-mRNA transcript or binding to a U2AF65 binding site upstream of the pseudoexon.

[0011] Disclosed herein, in certain embodiments, is a method of tailoring functional ATM expression in a subject, cell or tissue, comprising the administration of a NSE-activator agent and/or a NSE-repressor agent described herein. In some embodiments, the NSE repressor agent and/or NSE activator agent comprise a polynucleic acid polymer. In some embodiments, the NSE repressor agent and/or NSE activator agent is an SSO (Splice Switching Oligonucleotide). In some embodiments, the NSE repressor agent and/or NSE activator agent is associated with a delivery vehicle suitable for delivering the NSE repressor agent and/or NSE activator agent to cells. In some embodiments, the NSE repressor agent comprises: an SSO of the sequence cuucuaugcagccaaccuguagacu (SSO -NSE3), or a nucleic acid analogue thereof; or an SSO of the sequence accuuuuucuucuaugcagccaac (SSO -NSE5), or a nucleic acid analogue thereof; and/or the NSE repressor agent comprises or consists of any one SSO selected from the group comprising: aacauuucuauuuaguuaaaagc (SSO A11); uuaguauuccuugacuuua (SSO A17); gacugguaaauaauaaacauaauuc (SSO B2); auauauuagagauacaucagcc (SSO B4); and uuagagaaucauuuuaaauaagac (SSO AN3), or combinations thereof; or the method according to any of claims 13 to 22 wherein the NSE activator agent comprises the SSO PEkr/PEdel; and/or the NSE activator agent comprises or consists of any one SSO selected from the group comprising: aacuuaaagguuauaucuc (SSO A2); uauaaauacgaauaaaucga (SSO A4); caacacgacauaaccaaa (SSO A9); gguaugagaacuauagga (SSO A23); gguaauaagugucacaaa (SSO A25);guaucauacauuagaagg (SSO A26); and ugugggugaccacagcuu (SSO B11), or combinations thereof.

[0012] Disclosed herein, in certain embodiments, is use of rs609261 and/or rs4988000 genotyping to predict a subject's response to therapy for conditions associated with ATM deregulation.

[0013] Disclosed herein, in certain embodiments, is a composition comprising the NSE repressor agent and/or the NSE activator agent described herein. In some embodiments, the composition is a pharmaceutically acceptable formulation.

[0014] Disclosed herein, in certain embodiments, is a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome, wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to replace the non-thymine variant residue rs609261 with a thymine residue.

[0015] Disclosed herein, in certain embodiments, is a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising replacing a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome with a thymine residue. In some embodiments, replacing the non-thymine variant residue rs609261 comprises administration of an agent to the subject, which is arranged to replace the non-thymine variant residue rs609261 with a thymine residue. In some embodiments, the agent for replacement of the non-thymine residue is a genomic editing molecule. In some embodiments, the agent for replacement of the non-thymine residue is CRISPR-Cas9, or a functional equivalent thereof, together with an appropriate RNA molecule arranged to target rs609261.

[0016] Disclosed herein, in certain embodiments, is a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to replace the guanine variant residue at rs4988000 with adenine.

[0017] Disclosed herein, in certain embodiments, is a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising replacing a guanine variant residue at rs4988000 of the human genome with an adenine residue; or blocking the guanine residue by the binding of an SSO. In some embodiments, replacing the guanine variant residue at rs4988000 comprises administration of an agent to the subject, which is arranged to replace the guanine variant residue at rs4988000 with an adenine residue. In some embodiments, the agent for replacement of the guanine residue is a genomic editing molecule. In some embodiments, the agent for replacement of the guanine residue is CRISPR-Cas9, or a functional equivalent thereof, together with an appropriate RNA molecule arranged to target rs4988000.

[0018] Disclosed herein, in certain embodiments, is a method of screening a subject or a population of subjects for susceptibility to functional-ATM protein deficiency, wherein the screening comprises determining the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject (or group of subjects) has, or is susceptible to, functional-ATM protein deficiency.

[0019] Disclosed herein, in certain embodiments, is a method of selecting a subject or a population of subjects for treatment or prophylaxis, wherein the subject is susceptible to functional-ATM protein deficiency, the method comprising determining the presence of a guanine variant residue at

rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and selecting such subject for treatment with an agent arranged to increase functional-ATM levels in the subject.

[0020] Disclosed herein, in certain embodiments, is a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to increase functional-ATM levels. In some embodiments, the method is in combination to modify a CG haplotype to TA. In some embodiments, the method is in combination to identify a CG haplotype in a subject, and optionally treat or select the patient for treatment according to any preceding claim.

[0021] Disclosed herein, in certain embodiments, is a method of modifying regulation of NSE inclusion in a mature RNA transcript, the method comprising the insertion or deletion of one or more splicing regulatory motifs upstream or downstream of the NSEs that compete with the NSE for spliceosomal components, said regulatory motifs comprising cryptic splice sites or pseudo-exons.

[0022] Disclosed herein, in certain embodiments, is a method of modifying regulation of a functional protein expression, wherein the functional protein expression is regulated by NSE inclusion in a mature RNA transcript of the gene encoding protein, the method comprising the insertion or deletion of one or more splicing regulatory motifs upstream or downstream of the NSE that compete with the NSE for spliceosomal components, said regulatory motifs comprising cryptic splice sites or pseudo-exons. In some embodiments, the insertion or deletion of one or more splicing regulatory motifs is in genomic DNA of ATM intron 28. In some embodiments, insertion of one or more splicing regulatory motifs causes a reduction in NSE inclusion in the mature RNA transcript. In some embodiments, the deletion of one or more splicing regulatory motifs causes an increase in NSE inclusion in the mature RNA transcript. In some embodiments, the insertion or deletion of one or more splicing regulatory motifs comprises the use of genome editing technology, such as CRISPR-Cas9.

[0023] Disclosed herein, in certain embodiments, is a kit comprising one or more oligonucleotide probes for identifying rs609261 and/or rs4988000 variants. In some embodiments, the oligonucleotide probes are primers for use to PCR amplify regions of nucleic acid comprising rs609261 and/or rs4988000.

[0024] Disclosed herein, in certain embodiments, is a vector comprising nucleic acid encoding the NSE activating agent and/or NSE repressor agent.

[0025] Disclosed herein, in certain embodiments, is a method of screening for an agent capable of modifying regulation of a gene's expression comprising: identifying a nonsense-mediated RNA decay switch exon (NSE) that limits functional gene expression; identifying one or more splicing regulatory motifs upstream or downstream of the NSE that compete with the NSE for spliceosomal components, said regulatory motifs comprising cryptic splice sites or pseudoexons; targeting the one or more splicing regulatory motifs with antisense polynucleic acid arranged to hybridize to the splicing regulatory motifs through Watson-Crick base pairing; and determining if there is an increased or decreased inclusion of the NSE in a mature RNA transcript of the gene.

[0026] Disclosed herein, in certain embodiments, is a method of modulating gene's expression comprising providing an agent arranged to bind to splicing regulatory motifs, such as cryptic splice sites or pseudoexons, that compete with a nonsense-mediated RNA decay switch exon (NSE) for spliceosomal components.

[0027] Disclosed herein, in certain embodiments, is an agent arranged to bind to a gene splicing regulatory motif, such as a cryptic splice site or a pseudoexon, that competes with a nonsense-mediated RNA decay switch exon (NSE) for spliceosomal components, wherein the splicing regulatory motif controls inclusion of the NSE into a mature RNA transcript of the gene.

[0028] Disclosed herein, in certain embodiments, is a method, use, composition, vector, or agent substantially described herein, optionally with reference to the accompanying figures.

INCORPORATION BY REFERENCE

[0029] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0030] Fig. 1A-Fig. 1C illustrate an identification of a U2AF-repressed cryptic exon in ATM intron 28. Fig. 1A shows a schematics of the cryptic exon (termed here NSE for NMD-switch exon) activation. NSE sequence (upper panel) is boxed, asterisk denotes rs609261, and black rectangles show the indicated antisense oligonucleotides. Genome browser views of RNA-Seq data from RNAi- or SSO-mediated depletions of both U2AF35 isoforms (ab-), U2AF35a (a-), U2AF35b (b-) and controls (c) are shown in the lower panel. SSOs targeting 3' ss of U2AF1 exons Ab and 3 and U2AF35 siRNA were as previously described. Y axis, read densities. NSE inclusion/exclusion is schematically shown by dotted lines at the top. ATM exons (gray boxes) are numbered. The 29-nt

NS E introduced a stop codon in the ATM mRNA. Fig. 1B shows validation of the NSE activation by RT-PCR (upper panel) in independent depletions (lower panel). RT-PCR primers (ATM-F, ATM-R, Fig. 20) are denoted by arrows in panel A. Spliced products are shown to the right, the percentage of transcripts with NSE is at the top. Error bars denote SDs of two transfections experiments (***, $p<0.0001$, **, $p<0.001$). Fig. 1C shows NSE inclusion in mature transcripts inversely correlates with residual U2AF (r = Pearson correlation). Estimates of heterodimer levels were determined.

[0031] Fig. 2A-Fig. 2I show NSE activation and ATM expression modified by rs609261. Allelic frequencies at rs609261 are shown in the indicated populations (Fig. 2A). Fig. 2B shows exemplary minigene schematics. An XhoI/XbaI segment of ATM containing NSE and exon 29 was cloned between U2AF1 exons 2 and 4 (black boxes). RT-PCR primers to amplify exogenous transcripts (PL3 and ATM-R, Fig. 20) are denoted by arrows. Fig. 2C shows the rs609261-dependent NSE activation in exogenous pre-mRNAs. HEK293 cells depleted of U2AF35 or U2AF65 were transiently transfected with T (black) and C (grey) minigenes. Final concentration of the U2AF35 and U2AF65 siRNAs was 30 and 60 nM, respectively. Fig. 2D illustrates the identification of cell lines homozygous at rs609261 (asterisk). NSE is boxed. Fig. 2E and Fig. 2F show allele-specific activation of NSE in endogenous transcripts limits ATM expression in a dose-dependent manner. The source of endogenous transcripts is at the bottom, antibodies are to the right. Concentration of siRNAs in cultures was 3, 10 and 30 nM. C1, C2, control siRNAs. Transfection efficiency was monitored by a GFP-plasmid and fluorescent microscopy. Fig. 2G shows UPF1 depletion increased NSE activation (upper panel) and upregulated isoform U2AF1c (lower panel). The U2AF1c isoform contains both exons Ab and 3 and is repressed by NMD. Final concentration of the UPF1 siRNA was 7, 20 and 60 nM (SC = a scrambled control). Error bars are SDs of independent transfections. Fig. 2H shows NSE inclusion levels in cells depleted of U2AF-related proteins and a subset of heterogeneous nuclear RNPs. Error bars denote SDs of two transfections. Immunoblots are shown to the right. Final concentration of the U2AF35 siRNA was 25 nM; the remaining siRNAs were at 60 nM (C = controls). Fig. 2I shows overexpression of PUF60 induced NSE skipping. Immunoblots are shown below, antibodies to the right.

[0032] Fig. 3A-Fig. 3D illustrate rescue of U2AF-repressed ATM expression by SSOs targeting NSE. Fig. 3A and Fig. 3B show efficient SSO-mediated NSE inhibition in exogenous (Fig. 3A) and endogenous (Fig. 3B) ATM transcripts. Mean NSE inclusion levels of two transfection experiments are shown in the right panels. Fig. 3C shows restoration of ATM protein levels by SSOs that blocks access to NSE. Cells lacking U2AF35 and control cells were transfected with the SSO targeting the

NSE 3'ss and a control SSOs (Fig. 1A and Fig. 20). After 48 hrs, the cells were exposed to ionizing radiation (IR, 10 Gy) and harvested 1 hr later. Cell lysates were separated using a gradient SDS-PAGE. Western blotting was with antibodies shown to the right. Fig. 3D shows dose-dependent reconstitution of ATM expression SSO-NSE3 in depleted cells.

[0033] Fig. 4A-Fig. 4H show identification of intronic cis-elements and SSOs that modulate NSE activation. Fig. 4A shows schematics of two pseudoexons in ATM intron 28. Canonical exons (numbered) are shown as grey boxes, NSE as a white box, and PE as a checkered box. Asterisk indicates location of the IVS28-159A>G substitution, causing A-T. In this A-T case, both NSE and PE were included in the ATM mRNA together with the intervening sequence because NSE is separated from PE by less than the minimal size of human intron. Canonical and aberrant transcripts are denoted by dotted lines above and below the pre-mRNA, respectively. Middle panel shows RNA-Seq read densities for NSE in cells depleted of both U2AF35 isoforms (ab-) together with U2AF65 tags/high-confidence binding sites (horizontal lines/rectangles) identified by crosslinking and immunoprecipitation. The 100 basewise vertebrate conservation by Phylop (100 VC) is shown at the bottom. Lower panel shows mutations (in red and underlined) introduced in the C-minigene. Fig. 4B shows splicing pattern of wildtype and mutated C minigenes. Mutations are shown in panel A; RNA products are shown schematically to the right. The largest product produced by clone PE delPPT/AG contains the shortened pseudointron (42 nt). Fig. 4C shows splicing pattern of C minigenes mutated in NSE (lanes 2, 3, 7 and 8) or PE (lanes 4, 5, 9 and 10) in (mock) depleted HEK293 cells. Mutations are at the bottom and minigene sequences in Fig. 21. Spliced products are schematically shown to the right; a hairpin symbol above PE denotes the MIR stem-loop insertion. Fig. 4D and Fig. 4E illustrate SSO-induced pseudoexon switching. Transfected minigenes are shown at the top, spliced products to the right and SSOs at the bottom. SSO sequences are in Fig. 20. Final concentration of SSOs shown in panels D-G was 3, 10 and 30 nM. Fig. 4F shows SSOs targeting PE induced NSE skipping. Fig. 4G shows SSOs targeting a sequence activating NSE upon deletion (PEdelPPT/AG; panel A and B) inhibit PE. Fig. 4H shows NSE activation is haplotype-dependent. Minigene haplotypes at the indicated variants are shown at the bottom. Columns represent mean NSE inclusion, error bars are SDs, and asterisks denote statistically significant differences as in Fig. 1B.

[0034] Fig. 5A-Fig. 5G show exon-centric regulation of ATM signaling. Fig. 5A shows U2AF-regulated gene- and exon-level expression changes in MRN-ATM-CHEK2-CDC25-cdc2/cyclin B pathway (left panel). Log2fold- and q-values are shown in parentheses. Exon usage of CHEK2 and CDC25A genes is shown by RNA-Seq browser shots; PCR validation gels are in the right panels.

CHEK2 exon 9 is a NMD switch exon; exon 11 encodes a portion of the kinase domain. Full spectrum of U2AF-mediated expression changes in the ATM signaling pathway is shown in Fig. 9; examples of the U2AF-mediated splicing regulation are in Fig. S3-S6. Fig. 5B shows impaired ATM signaling in U2AF35 depleted cells following IR. HEK293 cells were (mock) depleted of U2AF35 and subjected to IR (10 Gy) 48 hrs later. Expression was examined by immunoblotting at the indicated time points. Antibodies are shown to the right. CHEK2 exon 9 skipping levels are at the bottom; their measurements in control (U2AF35+) and depleted cells (U2AF35-) are in panel Fig. 5C. Fig. 5D shows CHEK2 exon 9 inclusion in UPF1 depleted cells. Final concentration of the UPF1 siRNA (Fig. 20) was 12.5, 25, 50, and 100 nM. Fig. 5E shows repression of CHEK2 exon 9 by SSO reduced CHEK2 levels and promoted NSE inclusion. Final concentration of SSO targeting CHEK2 exon 9 was 3, 10 and 30 nM. Fig. 5F shows CHEK2 exon 9 inclusion upon transfection of HEK293 cells with the indicated SSOs. Fig. 5G shows a lack of SF3B1 induced CHEK2 exon 9 skipping but did not alter NSE activation. Final concentration of each siRNA targeting SF3B1 was 20 nM.

[0035] Fig. 6 shows rescue of NSE repression by cancer-associated mutations in U2AF35. Rescue of U2AF35-dependent NSE splicing of the C minigene by zinc finger 1 and 2 substitutions in U2AF35 (upper panel). All substitutions were made in the U2AF1a construct (35a). Cancer-associated mutations (bottom) are boxed; splice products are to the right. Immunoblot with U2AF35 and GFP antibodies is shown in the lower panel (ex = exogenous; en = endogenous U2AF35).

[0036] Fig. 7 shows SSO-based modulation of gene expression by pseudoexon targeting. Canonical exons are shown as grey boxes, a nonsense-mediated RNA decay (NMD) switch exon as a black box, pseudoexons as white boxes. Canonical splicing is shown by dotted lines. Pseudosplice sites competing with the NMD exon are shown below the RNA precursor. SSO activators/repressors are denoted by horizontal black/grey bars, respectively. Splicing regulatory motifs or secondary structures that compete with NMD switch exons for spliceosome components such as U2AF, heterogeneous nuclear ribonucleoproteins, or serine/arginine-rich proteins, for inclusion to mature transcripts are not shown for simplicity. They can be predicted by computational methods described in details previously (for example, Kralovicova, J. and Vorechovsky, I. (2007) Global control of aberrant splice site activation by auxiliary splicing sequences: evidence for a gradient in exon and intron definition . *Nucleic Acids Res.*, 35, 6399-6413, and references therein) or determined by experimental methods , including RNA crosslinking and immunoprecipitation, mutagenesis of splicing substrates and RNA folding studies.

[0037] Fig. 8A-Fig. 8C show SSO-mediated NSE repression enhances ATM expression. Fig. 8A shows SSO-NSE3 increased expression of total and activated ATM. HEK293 cells were (mock)-depleted of U2A F35, cotransfected with X press-tagged CHEK2 and SSO NSE3/control (SSO-C), exposed to ionizing radiation (IR) and harvested 30 minutes later. Cell lysates were immunoblotted with the indicated antibodies. Final concentration of siRNA and SSOs was 30 nM. The amount of plasmids expressing CHEK2 was 30, 90 and 270 ng; DNA from the empty vector was added to a final concentration of 270 ng/mL. Ex/enCHEK2, signal from exogenous and endogenous CHEK2, as detected by the D9C6 antibody. Fig. 8B and Fig. 8C show increased expression of exogenous CHEK2 by an SSO targeting NMD switch exon 9 (SSO CHEK2). Constant amounts of SSO CHEK2 were cotransfected with increasing amounts of Xpress-CHEK2 and constant amounts of GFP plasmids as transfection and loading control (B) and vice versa (C). Antibodies are to the right.

[0038] Fig. 9 illustrates an exemplary map of U2AF-regulated functional ATM interactions. U2AF-regulated ATM signaling network is highlighted by red arrows/pink background. Genes up-/down-regulated in cells depleted of U2AF35 are shown in red/dark green, respectively. Genes exhibiting significantly altered exon usage are shown in yellow. The ATM signaling map shows ATM-interacting proteins (purple)/protein complexes (light green). Arrows correspond to activation, T-shaped edges to inhibition and open circles denote unknown regulations. Containment links are shown as green edges.

[0039] Fig. 10A-Fig. 10C show exon usage in CDC25B and CDC25C in cells depleted of U2AF35. Genomic browser views of RNA-Seq data in control (ctr) and depleted (ab-) cells (left panels in Fig. 10A and Fig. 10B). PCR primers are shown by arrows, differentially used exons are denoted by black rectangles. RefSeq exon annotation is shown at the bottom. Validation of RNA-Seq data using RT-PCR with RNA extracted from cells depleted of each U2AF subunits and U2AF-related proteins (right panels in Fig. 10A and Fig. 10B).

[0040] Fig. 11A-Fig. 11C shows U2AF-regulated exon usage in TTK, PIN1and CDK1. Genomic browser views of RNA-Seq data in control (ctr) and depleted (ab-) cells (in Fig. 11A, left panel of Fig. 11B, and Fig. 11C). PCR primers are shown by arrows, differentially used exons are denoted by black rectangles. RefSeq exon annotation is shown at the bottom. Validation of RNA-Seq data using RT-PCR with RNA extracted from cells depleted of each U2AF subunits and U2AF-related proteins (right panel in Fig. 11B).

[0041] Fig. 12A-Fig. 12D show RNA processing of RAD50 and EZH2 in depleted cells. Genomic browser views of RNA-Seq data in control (ctr) and depleted (ab-) cells (left panels in Fig. 12A and Fig. 12B and in Fig. 12C and Fig. 12D). PCR primers are shown by arrows, differentially used exons

are denoted by black rectangles. RefSeq exon annotation is shown at the bottom. Validation of RNA-Seq data using RT-PCR with RNA extracted from cells depleted of each U2AF subunits and U2AF-related proteins (right panels in Fig. 12A and Fig. 12B).

[0042] Fig. 13A-Fig. 13B show U2AF35-controlled exon usage of the peptidyl-prolyl isomerase PIN2 and components of the shelterin complex.

[0043] Fig. 14A-Fig. 14D show U2AF control of RARA fusion partners.

[0044] Fig. 15A-Fig. 15E show NSE activation in normal tissue and leukemic cells. NSE inclusion levels were measured in 19 human tissues (Fig. 15A) and 17 AML/CMML bone marrow samples (Fig. 15B) using primers ATM-F and ATM-R (Fig. 1, Fig. 20). Exon inclusion was quantified. Means were compared with an unpaired t-test (Fig. 15C). Fig. 15D and Fig. 15E show inclusion levels of U2AF-repressed (Fig. 15D) and –activated (Fig. 15E) exons in lymphoblastoid cell lines (top). Cells were exposed to cold and heat shock at the indicated temperatures. ES, exon skipping; EI, exon inclusion.

[0045] Fig. 16A-Fig. 16B show identification of transposed elements in ATM intron 28 that influence NSE activation. Fig. 16A shows the location of transposed elements in intron 28 and schematics of NSE activation. Canonical exons are shown as grey boxes, the NSE as a white box, introns flanking the NSE as lines and their splicing by dotted lines. Transposed elements are shown as horizontal white rectangles below the primary transcript; UC, a unique sequence lacking recognizable transposons. Their deletions are numbered 1-6, which corresponds to lane numbers in panel B. RT PCR primers are denoted by black arrows. A scale is at the top. The NSE sequence is boxed in the lower panel. Constructs lacking the sense Alu (Alu+) repeatedly failed to ligate/propagate and were not examined. Fig. 16B shows deletion of antisense Alu and MER51 elements alters NSE activation. Wild-type (WT) and mutated constructs (designated 1-6) were transiently transfected into HEK293 cells (mock)depleted of U2AF35. NSE+/-, RNA products with/without NSE. Columns represent mean NSE inclusion (%), error bars SDs of 2 transfection experiments. Asterisks denote two-tailed P values <0.01 (t-test).

[0046] Fig. 17A-Fig. 17C show identification of intronic SSOs that activate or repress NSE. Fig. 17A shows the location of tested SSOs in intron 28 relative to transposed elements. For legend, see Fig. 16A. Fig. 17B shows the identification of intron 28 SSOs that alter NSE activation in exogenous transcripts. Illustrative SSOs are listed in Table 2. The “x” symbol denotes multiple negative controls, dotted line the average NSE inclusion, error bars SDs of two transfections experiments. Columns represent mean inclusion levels, asterisks show significant P values. Fig. 17C shows SSOs

targeting single-stranded regions tended to repress endogenous NSE. r, Pearson correlation coefficient. The P value is in parentheses.

[0047] Fig. 18A-Fig. 18B show TMC-SA-assisted delivery of SSO-NSE3 to human cell lines leads to NSE repression. Fig. 18A shows NSE inclusion in HEK293 cells is inhibited upon exposure of SSO-NSE3/TMC-SA nanocomplexes. N/P ratio was 20, 40 and 80 (Sc = a scrambled control with the same modification , M = size marker). Error bars denote SDs of two transfections experiments. P values are shown at the top for the indicated comparisons. Fig. 18B shows NSE repression in VAVY cells exposed to SSO-NSE3/TMC-SA complexes.

[0048] Fig. 19 shows inverted repeats in the MER51 consensus sequence with ATM intron 28 (v, transversions; i, transitions). Most stable inverted repeats in the ATM MER51A are underlined and highlighted; purine-rich single-stranded regions are in red; the long terminal repeat homology originally described for the MER51 family is in italics. The aligned segment corresponds to deletion 4 shown in Fig. 16a. The MER51A consensus sequence is in the antisense orientation.

[0049] Fig. 20 illustrates exemplary synthetic DNA and RNA sequences.

[0050] Fig. 21 shows exemplary sequences of splicing reporter constructs mutated in NSE and PE.

[0051] Fig. 22 shows auxiliary splicing elements in NSE and PE.

[0052] Fig. 23 shows a summary of U2AF35-regulated transcripts involved in NMD.

DETAILED DESCRIPTION OF THE INVENTION

[0053] Intervening sequences or introns are removed by a large and highly dynamic RNA-protein complex termed the spliceosome, which orchestrates complex interactions between primary transcripts, small nuclear RNAs (snRNAs) and a large number of proteins. Spliceosomes assemble *ad hoc* on each intron in an ordered manner, starting with recognition of the 5‘ splice site (5’ss) by U1 snRNA or the 3’ss by the U2 pathway, which involves binding of the U2 auxiliary factor (U2AF) to the 3’ss region to facilitate U2 binding to the branch point sequence (BPS). U2AF is a stable heterodimer composed of a U2AF2-encoded 65-kD subunit (U2AF65), which binds the polypyrimidine tract (PPT), and a U2AF1-encoded 35-kD subunit (U2AF35), which interacts with highly conserved AG dinucleotides at 3‘ss and stabilizes U2AF65 binding. In addition to the BPS/PPT unit and 3’ss/5’ss, accurate splicing requires auxiliary sequences or structures that activate or repress splice site recognition, known as intronic or exonic splicing enhancers or silencers. These elements allow genuine splice sites to be recognized among a vast excess of cryptic or pseudo-sites in the genome of higher eukaryotes, which have the same sequences but outnumber authentic sites

by an order of magnitude. Although they often have a regulatory function, the exact mechanisms of their activation or repression are poorly understood.

[0054] Exome sequencing studies have revealed a highly restricted pattern of somatic mutations in U2AF1/U2AF2 and other genes involved in 3'ss recognition (SF3B1, ZRSR2, SF1, SF3A1, PRPF40B, and SRSF2) in cancer cells, most prominently myelodysplastic syndromes. These genes encode products that often interact during spliceosome assembly, suggesting the existence of shared pathways in oncogenesis, which is further supported by a high degree of mutual exclusivity of cancer-associated mutations. Genome-wide transcriptome profiling in leukemic samples carrying these mutations detected numerous alterations in splicing of mRNA precursors, however, key links between specific RNA processing defects and cancer initiation or progression have remained obscure, despite the great promise of these targets for therapeutic modulation. The interconnections between these RNA-binding proteins and DNA damage response (DDR) pathways remain to be fully characterized.

[0055] Mutations in traditional (BPS/PPT/3'ss/5'ss) and auxiliary splicing motifs often cause aberrant splicing, such as exon skipping or cryptic exon or splice-site activation, and contribute significantly to human morbidity and mortality. Both aberrant and alternative splicing patterns can be influenced by natural DNA variants in exons and introns, which play an important role in heritability of both Mendelian and complex traits. However, the molecular mechanisms that translate the allele- or haplotype-specific RNA expression to phenotypic variability as well as interactions between intronic and exonic variant alleles and *trans*-acting factors are largely obscure.

[0056] Antisense technology has now reached important clinical applications. For example, antisense splice-switching oligonucleotides (SSOs) targeting the ATM gene have been used to repair splicing mutations in ataxia-telangiectasia (A-T) and were successful in normalizing ATM protein levels (Du *et al.*, 2011; Du *et al.*, 2007).

[0057] A large fraction of both leukemias and solid tumors show deregulation of ATM expression (for example, Stankovic *et al.*, 1999; Starczynski *et al.*, 2003). Chemical inhibitors of ATM (wortmannin, CP-466722, KU-55933, and KU60019) have not reached clinical trials, largely because of nonspecific effects and/or high toxicity, although KU-559403 has shown good bioavailability and reliably conferred radiosensitivity.

[0058] In some instances, the ability to up or down regulate gene expression in a sequence-specific manner is desirable.

[0059] In certain embodiments, provided herein is a method of screening a subject or a population of subjects for susceptibility to functional-ATM protein deficiency, wherein the screening comprises

determining the presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic or nonsense-mediated RNA decay switch exon in ATM intron 28) of the human genome, wherein the presence of a non-thymine variant residue rs609261 indicates that the subject (or group of subjects) has, or is susceptible to, functional-ATM protein deficiency.

[0060] The term "functional ATM-protein deficiency" means the reduction in the presence/expression of ATM protein that is functional in a subject, cell or tissue. Functional ATM-deficiency is the result of a functional variant rs609261 in ATM intron 28 that alters RNA processing of ATM precursor messenger RNA (pre-mRNA). Cytosine allele at rs609261 results in a higher inclusion of a nonsense-mediated RNA decay switch exon (termed here NSE) in ATM mRNA than a thymine allele at this position, limiting the expression of ATM protein more efficiently than the thymine allele. This limitation can be removed or modulated by novel SSOs that block access to NSE or to NSE-regulatory sequences in the same intron, leading to derepression or inhibition of ATM protein, respectively.

[0061] In some embodiments, provided herein is a method of selecting a subject or a population of subjects for treatment or prophylaxis, wherein the subject is susceptible to functional-ATM protein deficiency, the method comprising determining the presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome, wherein the presence of a non-thymine variant residue rs609261indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and selecting such subject for treatment with an agent arranged to increase functional-ATM levels in the subject.

[0062] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE of the human genome, wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to increase functional-ATM levels.

[0063] According to another aspect of the invention, there is provided a method of treatment or prevention of a condition associated with a functional-ATM protein deficiency, comprising the administration of a NSE repressor agent arranged to increase levels of functional ATM protein, wherein the agent is arranged to bind to a NSE in ATM intron 28 of the pre-mRNA transcript or to NSE-activating regulatory sequences in the same intron to decrease inclusion of the NSE in the mature transcript.

[0064] According to another aspect of the invention, there is provided a method of treatment or prevention of a condition associated with deregulation of ATM expression in a subject comprising the administration of a NSE-activator agent, wherein the NSE-activator agent is arranged to increase NSE inclusion in the ATM mature RNA transcript by binding to NSE-inhibiting regulatory motifs in ATM intron 28.

[0065] NSE-inhibiting regulatory motifs in ATM intron 28 may comprise sequences that compete with NSE for spliceosomal components, such as a 24 nucleotide pseudoexon (PE) located 3' of NSE in ATM intron 28 of the pre-mRNA transcript or U2AF65 binding site upstream of the pseudoexon.

[0066] According to another aspect of the invention, there is provided a method of treatment or prevention of cancer in a subject comprising the administration of a NSE-activator agent arranged to increase a cancer cell's susceptibility to DNA damaging agents that induce double strand DNA breaks, such as radiotherapy, wherein the NSE-activator agent is arranged to increase NSE inclusion in the ATM mature RNA by binding NSE regulatory motifs in ATM intron 28; and treating the subject with DNA damaging agents that cause double strand breaks, such as radiotherapy or chemotherapy.

[0067] According to another aspect of the invention, there is provided a method of increasing a cell's susceptibility to cytotoxic therapy, such as radiotherapy treatment, comprising the reduction of ATM protein expression by administration of a NSE-activator agent arranged to increase NSE inclusion in ATM mature RNA transcript by binding to regulatory motifs in ATM intron 28.

[0068] The regulatory motifs in ATM intron 28 may compete with NSE for spliceosomal components, wherein such motifs may comprise a 24 nucleotide pseudoexon (PE) located 3' of NSE in ATM intron 28 of the pre-mRNA transcript or U2AF65 binding site upstream of the pseudoexon.

[0069] According to another aspect of the invention, there is provided a method of tailoring functional ATM expression in a subject, cell or tissue, comprising the administration of a NSE-activator agent and/or a NSE-repressor agent described herein.

[0070] According to another aspect of the invention, there is provided use of rs609261 genotyping to predict a subject response to therapy for conditions associated with ATM deregulation.

[0071] According to another aspect of the invention, there is provided a composition comprising the NSE repressor agent of the invention herein.

[0072] According to another aspect of the invention, there is provided a composition comprising the NSE activator agent of the invention herein.

[0073] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the

presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome, wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to replace the non-thymine variant residue rs609261 with a thymine residue.

[0074] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising replacing a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome with a thymine residue.

[0075] According to another aspect of the invention, there is provided a vector comprising the polynucleic acid polymer of the invention.

[0076] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to replace the guanine variant residue at rs4988000 with adenine.

[0077] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising replacing a guanine variant residue at rs4988000 of the human genome with an adenine residue.

[0078] According to a first aspect of the invention, there is provided a method of screening a subject or a population of subjects for susceptibility to functional-ATM protein deficiency, wherein the screening comprises determining the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject(or group of subjects) has, or is susceptible to, functional-ATM protein deficiency.

[0079] According to another aspect of the invention, there is provided a method of selecting a subject or a population of subjects for treatment or prophylaxis, wherein the subject is susceptible to functional-ATM protein deficiency, the method comprising determining the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and selecting such subject for treatment with an agent arranged to increase functional-ATM levels in the subject.

[0080] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to increase functional-ATM levels.

[0081] According to another aspect of the invention, there is provided a method of screening for an agent or a combination of agents capable of modifying regulation of a gene's expression (Fig. 7) comprising identifying a nonsense-mediated RNA decay switch exon (NSE) that limits functional gene expression; identifying one or more splicing regulatory motifs upstream or downstream of the NSE that compete with the NSE for spliceosomal components, said regulatory motifs comprising cryptic splice sites or pseudo-exons; targeting the one or more splicing regulatory motifs with antisense polynucleic acid that are arranged to hybridize to the splicing regulatory motifs through Watson-Crick base pairing; and determining if there is an increased or decreased inclusion of the NSE in a mature RNA transcript of the gene.

[0082] According to another aspect of the invention, there is provided a method of modulating gene's expression comprising providing an agent arranged to bind to NSE splicing regulatory motifs.

[0083] According to another aspect of the invention, there is provided an agent arranged to bind to a gene splicing regulatory motif of NSE, wherein the splicing regulatory motif controls inclusion of the NSE into a mature RNA transcript of the gene.

[0084] According to another aspect of the invention, provided herein is a method of a treatment or prevention of a disease pathology caused by an NSE inclusion in an mRNA gene transcript comprising providing an agent arranged to bind to a gene NSE splicing regulatory motif that controls inclusion of the NSE into a mature RNA transcript of the gene.

[0085] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

[0086] The determination may use any suitable assay or genetic analysis available to the skilled person. In some instances, detection is done at a nucleic acid level with nucleic acid-based

techniques such as in situ hybridization and RT-PCR. Sequencing technologies can include next-generation sequencing technologies such as Helicos True Single Molecule Sequencing (tSMS) (Harris T.D. *et al.*, (2008) Science 320:106-109); 454 sequencing (Roche) (Margulies, M. *et al.*, 2005, Nature, 437, 376-380); SOLiD technology (Applied Biosystems); SOLEXA sequencing (Illumina); single molecule, real-time (SMRT™) technology of Pacific Biosciences; nanopore sequencing (Soni GV and Meller A. (2007) Clin Chem 53: 1996-2001); semiconductor sequencing (Ion Torrent; Personal Genome Machine); DNA nanoball sequencing; sequencing using technology from Dover Systems (Polonator), and technologies that do not require amplification or otherwise transform native DNA prior to sequencing (*e.g.*, Pacific Biosciences and Helicos), such as nanopore-based strategies (*e.g.*, Oxford Nanopore, Genia Technologies, and Nabsys). Sequencing technologies can also include Sanger sequencing, Maxam-Gilbert sequencing, Shotgun sequencing, bridge PCR, mass spectrometry based sequencing, microfluidic based Sanger sequencing, microscopy-based sequencing, RNAP sequencing, or hybridization based sequencing.

[0087] Sequencing of a gene transcript of interest may also include an amplification step. Exemplary amplification methodologies include, but are not limited to, polymerase chain reaction (PCR), nucleic acid sequence based amplification (NASBA), self-sustained sequence replication (3SR), loop mediated isothermal amplification (LAMP), strand displacement amplification (SDA), whole genome amplification, multiple displacement amplification, strand displacement amplification, helicase dependent amplification, nicking enzyme amplification reaction, recombinant polymerase amplification, reverse transcription PCR, ligation mediated PCR, or methylation specific PCR.

[0088] Additional methods that can be used to obtain a nucleic acid sequence include, *e.g.*, whole-genome RNA expression array, enzyme-linked immunosorbent assay (ELISA), genome sequencing, de novo sequencing, Pacific Biosciences SMRT sequencing, immunohistochemistry (IHC), immunocytochemistry (ICC), mass spectrometry, tandem mass spectrometry, matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS), in-situ hybridization, fluorescent in-situ hybridization (FISH), chromogenic in-situ hybridization (CISH), silver in situ hybridization (SISH), digital PCR (dPCR), reverse transcription PCR, quantitative PCR (Q-PCR), single marker qPCR, real-time PCR, nCounter Analysis (Nanostring technology), Western blotting, Southern blotting, SDS-PAGE, gel electrophoresis, and Northern blotting.

[0089] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome, wherein the presence of a non-

thymine variant residue rs609261 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to increase functional-ATM levels.

[0090] According to another aspect of the invention, there is provided a method of treatment or prevention of a condition associated with a functional-ATM protein deficiency, comprising the administration of a NSE repressor agent arranged to increase levels of functional ATM protein, wherein the agent is arranged to bind to a NSE in ATM intron 28 of the pre-mRNA transcript to decrease inclusion of the NSE in the mature RNA transcript.

[0091] Decreasing inclusion of the NSE in the mature RNA transcript may provide an increase in functional ATM protein expression.

[0092] The method of treatment or prevention of functional-ATM protein deficiency in a subject or an at-risk population of subjects may be a method of treatment or prevention of a condition associated with functional-ATM protein deficiency. The condition may be any symptom of ataxia-telangiectasia; cerebellar ataxia; oculocutaneous angiectasia; cancer; immune deficiency; cellular radiosensitivity; or chromosomal instability. The cancer may comprise lymphoblastoid leukemias, or lymphomas. In one embodiment, the condition is ataxia-telangiectasia. In another embodiment, the condition is cancer. The cancer may comprise a non-Hodgkin or Hodgkin lymphoma.

[0093] In one embodiment, the NSE comprises the sequence tctacaggttggctgcatagaagaaaaag (SEQ ID NO: 57). The NSE repressor agent may be arranged to bind to NSE within the sequence agTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag (SEQ ID NO: 58) (respective 3' and 5' splice site dinucleotides of flanking intervening sequences are underlined). The NSE repressor agent may be arranged to bind to the 5' or 3' splice site of the NSE in ATM intron 28. In another embodiment, the NSE repressor agent is arranged to bind to the 3' splice site of the NSE in ATM intron 28. In another embodiment, the NSE repressor agent may be arranged to bind to NSE within the sequence tcttagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag (SEQ ID NO: 59) (respective 3' and 5' splice site dinucleotides of flanking intervening sequences are underlined). In another embodiment, the NSE repressor agent may be arranged to bind to NSE within the sequence tctcagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag (SEQ ID NO: 60) (respective 3' and 5' splice site dinucleotides of flanking intervening sequences are underlined).

[0094] According to another aspect of the invention, there is provided a method of treatment or prevention of a condition associated with deregulation of ATM expression in a subject comprising the administration of a NSE-activator agent, wherein the NSE-activator agent is arranged to increase

NSE inclusion in ATM mature RNA transcript by binding to splicing regulatory motifs in ATM intron 28.

[0095] Increasing inclusion of the NSE in the mature RNA transcript may provide a decrease in functional ATM protein expression.

[0096] According to another aspect of the invention, there is provided a method of treatment or prevention of cancer in a subject comprising the administration of a NSE-activator agent arranged to increase a cancer cell's susceptibility to cytotoxic therapy with DNA damaging agents such as radiotherapy, wherein the NSE-activator agent is arranged to increase NSE inclusion in ATM mature RNA transcript by binding to splicing regulatory motifs in ATM intron 28; and treating the subject with the cytotoxic therapy, such as radiotherapy or chemotherapy.

[0097] Chemotherapy may comprise a therapeutic that induces double strand DNA breaks. The skilled person will understand that there are several chemotherapy/therapeutic agents that are capable of inducing double strand DNA breaks. In one embodiment, the chemotherapy agents may comprise bleomycin.

[0098] Increasing inclusion of the NSE in the mature RNA transcript may provide a decrease in functional ATM protein expression.

[0099] The radiotherapy or chemotherapy may be following the administration of the agent. The radiotherapy or chemotherapy may one or more days following the administration of the agent. The radiotherapy or chemotherapy may be one or more weeks following the administration of the agent.

[0100] In one embodiment the pseudoexon comprises the sequence tcatcgaatacttttggaaataag.

[0101] According to another aspect of the invention, there is provided a method of increasing a cell's susceptibility to cytotoxic therapy with DNA damaging agents such as radiotherapy comprising the reduction of ATM protein expression by administration of a NSE-activator agent arranged to increase NSE inclusion in ATM mature RNA transcript by binding to NSE regulatory motifs in ATM intron 28.

[0102] In one embodiment the cell is a cancerous cell. In another embodiment the cell is a pre-cancerous cell.

[0103] According to another aspect of the invention, there is provided a method of tailoring functional ATM expression in a subject, cell or tissue, comprising the administration of a NSE-activator agent and/or a NSE-repressor agent described herein.

Nonsense-mediated mRNA Decay

[0104] Nonsense-mediated mRNA decay (NMD) is a surveillance pathway that exists in all eukaryotes. Its main function is to reduce errors in gene expression by eliminating mRNA transcripts that contain premature stop codons. NMD targets transcripts with premature stop codons but also a broad array of mRNA isoforms expressed from many endogenous genes, suggesting that NMD is a master regulator that drives both fine and coarse adjustments in steady-state RNA levels in the cell.

[0105] A nonsense-mediated RNA decay switch exon (NSE) is an exon or a pseudoexon that activates the NMD pathway if included in a mature RNA transcript. A NSE inclusion in mature transcripts downregulates gene expression.

[0106] Cryptic (or pseudo- splice sites) have the same splicing recognition sequences as genuine splice sites but are not used in the splicing reactions. They outnumber genuine splice sites in the human genome by an order of a magnitude and are normally repressed by thus far poorly understood molecular mechanisms. Cryptic 5' splice sites have the consensus NNN/GUNNNN or NNN/GCNNNN where N is any nucleotide and / is the exon-intron boundary. Cryptic 3' splice sites have the consensus NAG/N. Their activation is positively influenced by surrounding nucleotides that make them more similar to the optimal consensus of authentic splice sites, namely MAG/GURAGU and YAG/G, respectively, where M is C or A, R is G or A, and Y is C or U.

[0107] Cryptic (or pseudo-) exons have the same splicing recognition sequences as genuine exons but are not used in the splicing reactions. They outnumber genuine exons by an order of a magnitude and are normally repressed by thus far poorly understood molecular mechanisms.

[0108] Splice sites and their regulatory sequences can be readily identified by a skilled person using suitable algorithms publicly available, listed for example in Kralovicova, J. and Vorechovsky, I. (2007) Global control of aberrant splice site activation by auxiliary splicing sequences: evidence for a gradient in exon and intron definition. *Nucleic Acids Res.*, 35, 6399-6413,(http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2095810/pdf/gkm680.pdf).

[0109] The cryptic splice sites or splicing regulatory sequences may compete for RNA-binding proteins such as U2AF with a splice site of the NSE. In one embodiment, the agent may bind to the cryptic splice site or splicing regulatory sequences to prevent the binding of RNA-binding proteins and thereby favoring utilization of the NSE splice sites.

[0110] In one embodiment, the cryptic splice site may not comprise the 5' or 3' splice site of the NSE. The cryptic splice site may be at least 10 nucleotides upstream of the NSE 5' splice site. The cryptic splice site may be at least 20 nucleotides upstream of the NSE 5' splice site. The cryptic splice site may be at least 50 nucleotides upstream of the NSE 5' splice site. The cryptic splice site

may be at least 100 nucleotides upstream of the NSE 5' splice site. The cryptic splice site may be at least 200 nucleotides upstream of the NSE 5' splice site.

[0111] The cryptic splice site may be at least 10 nucleotides downstream of the NSE 3' splice site. The cryptic splice site may be at least 20 nucleotides downstream of the NSE 3' splice site. The cryptic splice site may be at least 50 nucleotides downstream of the NSE 3' splice site. The cryptic splice site may be at least 100 nucleotides downstream of the NSE 3' splice site. The cryptic splice site may be at least 200 nucleotides downstream of the NSE 3' splice site.

The NSE repressor agent and NSE activator agent

[0112] The NSE repressor agent and/or NSE activator agent may comprise a polynucleic acid polymer. In one embodiment, the NSE repressor agent and/or NSE activator agent is an SSO (Splice Switching Oligonucleotide).

[0113] In an embodiment wherein the NSE repressor agent and/or NSE activator agent comprises a polynucleic acid polymer the following statements may apply equally to both the NSE repressor agent and the NSE activator agent unless otherwise indicated. The polynucleic acid polymer may be about 50 nucleotides in length. The polynucleic acid polymer may be about 45 nucleotides in length. The polynucleic acid polymer may be about 40 nucleotides in length. The polynucleic acid polymer may be about 35 nucleotides in length. The polynucleic acid polymer may be about 30 nucleotides in length. The polynucleic acid polymer may be about 24 nucleotides in length. The polynucleic acid polymer may be about 25 nucleotides in length. The polynucleic acid polymer may be about 20 nucleotides in length. The polynucleic acid polymer may be about 19 nucleotides in length. The polynucleic acid polymer may be about 18 nucleotides in length. The polynucleic acid polymer may be about 17 nucleotides in length. The polynucleic acid polymer may be about 16 nucleotides in length. The polynucleic acid polymer may be about 15 nucleotides in length. The polynucleic acid polymer may be about 14 nucleotides in length. The polynucleic acid polymer may be about 13 nucleotides in length. The polynucleic acid polymer may be about 12 nucleotides in length. The polynucleic acid polymer may be about 11 nucleotides in length. The polynucleic acid polymer may be about 10 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 50 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 45 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 40 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 35 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 30 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 25

nucleotides in length. The polynucleic acid polymer may be between about 10 and about 20 nucleotides in length. The polynucleic acid polymer may be between about 15 and about 25 nucleotides in length. The polynucleic acid polymer may be between about 15 and about 30 nucleotides in length. The polynucleic acid polymer may be between about 12 and about 30 nucleotides in length.

[0114] The sequence of the polynucleic acid polymer may be at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% complementary to a target sequence of the partially processed mRNA transcript. The sequence of the polynucleic acid polymer may be 100% complementary to a target sequence of the pre-mRNA transcript.

[0115] The sequence of the polynucleic acid polymer may have 4 or less mismatches to a target sequence of the pre-mRNA transcript. The sequence of the polynucleic acid polymer may have 3 or less mismatches to a target sequence of the pre-mRNA transcript. The sequence of the polynucleic acid polymer may have 2 or less mismatches to a target sequence of the pre-mRNA transcript. The sequence of the polynucleic acid polymer may have 1 or less mismatches to a target sequence of the pre-mRNA transcript.

[0116] The polynucleic acid polymer may specifically hybridize to a target sequence of the pre-mRNA transcript. The specificity may be at least a 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or 100% sequence complementarity of the polynucleic acid polymer to a target sequence of the pre-mRNA transcript. The hybridization may be under high stringent hybridization conditions.

[0117] The polynucleic acid polymer may have a sequence with at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity to a sequence illustrated in Table 2 or Fig. 20. The polynucleic acid polymer may have a sequence with 100% sequence identity to a sequence illustrated in Table 2 or Fig. 20. In some instances, the polynucleic acid polymer may have a sequence with at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity to a sequence illustrated in Table 2. In some cases, the polynucleic acid polymer may have a sequence with 100% sequence identity to a sequence illustrated in Table 2.

[0118] In some instances, the polynucleic acid polymer has a sequence with at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 50% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least

60% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 70% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 80% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 85% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 90% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 91% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 92% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 93% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 94% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 95% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 96% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 97% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 98% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 99% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with at least 99.5% sequence identity to a sequence selected from SEQ ID NOs: 18-52. In some cases, the polynucleic acid polymer has a sequence with 100% sequence identity to a sequence selected from SEQ ID NOs: 18-52.

[0119] In some embodiments, a polynucleic acid polymer hybridizes to a motif within a transposed element, upstream of a transposed element, or downstream of a transposed element. In some instances, the transposed element is Alu, MER51, UC or L4C. In some instances, the transposed element is Alu (e.g., Alu- or Alu+) or MER51. In some cases, the transposed element is Alu (e.g., Alu- or Alu+). In other cases, the transposed element is MER51. In some instances, the polynucleic acid polymer hybridizes to a target motif within Alu (e.g., Alu- or Alu+). In other instances, the polynucleic acid polymer hybridizes to a target motif downstream of MER51. In some instances, the polynucleic acid polymer has a sequence with at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%,

90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity to a sequence selected from SEQ ID NOs: 18-52.

[0120] In some embodiments, the polynucleic acid polymer hybridizes to a target motif that is either upstream or downstream of Alu (e.g., Alu- or Alu+). In some instances, the polynucleic acid polymer hybridizes to a target motif that is upstream of Alu. In some cases, the target motif is about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, or more bases upstream of Alu. In some cases, the target motif is about 5 or more bases upstream of Alu. In some cases, the target motif is about 10 or more bases upstream of Alu. In some cases, the target motif is about 20 or more bases upstream of Alu. In some cases, the target motif is about 30 or more bases upstream of Alu. In some cases, the target motif is about 40 or more bases upstream of Alu. In some cases, the target motif is about 50 or more bases upstream of Alu. In some cases, the target motif is about 80 or more bases upstream of Alu. In some cases, the target motif is about 100 or more bases upstream of Alu. In some cases, the target motif is about 150 or more bases upstream of Alu. In some cases, the target motif is about 200 or more bases upstream of Alu. In some cases, the target motif is about 300 or more bases upstream of Alu. In some cases, the target motif is about 500 or more bases upstream of Alu. In some cases, the target motif is about 800 or more bases upstream of Alu. In some instances, the polynucleic acid polymer has a sequence with at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity to a sequence selected from SEQ ID NOs: 18-52.

[0121] In some instances, the polynucleic acid polymer hybridizes to a target motif that is downstream of Alu (e.g., Alu- or Alu+). In some cases, the target motif is about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, or more bases downstream of Alu. In some cases, the target motif is about 5 or more bases downstream of Alu. In some cases, the target motif is about 10 or more bases downstream of Alu. In some cases, the target motif is about 20 or more bases downstream of Alu. In some cases, the target motif is about 30 or more bases downstream of Alu. In some cases, the target motif is about 40 or more bases downstream of Alu. In some cases, the target motif is about 50 or more bases downstream of Alu. In some cases, the target motif is about 80 or more bases downstream of Alu. In some cases, the target motif is about 100 or more bases downstream of Alu. In some cases, the target motif is about 150 or more bases downstream of Alu. In some cases, the target motif is about 200 or more bases downstream of Alu. In some cases, the target motif is about 300 or more bases downstream of Alu. In some cases, the target motif is about 500 or more bases

downstream of Alu. In some cases, the target motif is about 800 or more bases downstream of Alu. In some instances, the polynucleic acid polymer has a sequence with at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity to a sequence selected from SEQ ID NOs: 18-52.

[0122] In some embodiments, the polynucleic acid polymer hybridizes to a target motif that is either upstream or downstream of MER51. In some instances, the polynucleic acid polymer hybridizes to a target motif that is upstream of MER51. In some cases, the target motif is about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, or more bases upstream of MER51. In some cases, the target motif is about 5 or more bases upstream of MER51. In some cases, the target motif is about 10 or more bases upstream of MER51. In some cases, the target motif is about 20 or more bases upstream of MER51. In some cases, the target motif is about 30 or more bases upstream of MER51. In some cases, the target motif is about 40 or more bases upstream of MER51. In some cases, the target motif is about 50 or more bases upstream of MER51. In some cases, the target motif is about 80 or more bases upstream of MER51. In some cases, the target motif is about 100 or more bases upstream of MER51. In some cases, the target motif is about 150 or more bases upstream of MER51. In some cases, the target motif is about 200 or more bases upstream of MER51. In some cases, the target motif is about 300 or more bases upstream of MER51. In some cases, the target motif is about 500 or more bases upstream of MER51. In some cases, the target motif is about 800 or more bases upstream of MER51. In some instances, the polynucleic acid polymer has a sequence with at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity to a sequence selected from SEQ ID NOs: 18-52.

In some instances, the polynucleic acid polymer hybridizes to a target motif that is downstream of MER51. In some cases, the target motif is about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, or more bases downstream of MER51. In some cases, the target motif is about 5 or more bases downstream of MER51. In some cases, the target motif is about 10 or more bases downstream of MER51. In some cases, the target motif is about 20 or more bases downstream of MER51. In some cases, the target motif is about 30 or more bases downstream of MER51. In some cases, the target motif is about 40 or more bases downstream of MER51. In some cases, the target motif is about 50 or more bases downstream of MER51. In some cases, the target motif is about 80 or more bases downstream of MER51. In some cases, the target motif is about 100 or more bases downstream of MER51. In some cases, the target motif is about 150 or more bases downstream of MER51. In some cases, the

target motif is about 200 or more bases downstream of MER51. In some cases, the target motif is about 300 or more bases downstream of MER51. In some cases, the target motif is about 500 or more bases downstream of MER51. In some cases, the target motif is about 800 or more bases downstream of MER51. In some instances, the polynucleic acid polymer has a sequence with at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity to a sequence selected from SEQ ID NOs: 18-52.

[0123] Where reference is made to a polynucleic acid polymer sequence, the skilled person will understand that one or more substitutions may be tolerated, optionally two substitutions may be tolerated in the sequence, such that it maintains the ability to hybridize to the target sequence, or where the substitution is in a target sequence, the ability to be recognized as the target sequence. References to sequence identity may be determined by BLAST sequence alignment (www.ncbi.nlm.nih.gov/BLAST/) using standard/default parameters. For example, the sequence may have 99% identity and still function according to the invention. In other embodiments, the sequence may have 98% identity and still function according to the invention. In another embodiment, the sequence may have 95% identity and still function according to the invention.

[0124] A polynucleic acid polymer, such as the SSOs, may comprise RNA or DNA. The polynucleic acid polymer, such as the SSOs, may comprise RNA. The polynucleic acid polymer, such as the SSOs, may comprise natural or synthetic or artificial nucleotide analogues or bases, having equivalent complementation as DNA or RNA. The polynucleic acid polymer, such as the SSOs, may comprise combinations of DNA, RNA and/or nucleotide analogues. Nucleotide analogues may comprise PNA or LNA. In another embodiment, the nucleic acid, such as the SSOs, may comprise or consist of PMO.

[0125] In some instances, the synthetic or artificial nucleotide analogues or bases can comprise modifications at one or more of ribose moiety, phosphate moiety, nucleoside moiety, or a combination thereof. For example, a nucleotide base may be any naturally occurring, unmodified nucleotide base such as adenine, guanine, cytosine, thymine and uracil, or any synthetic or modified base that is sufficiently similar to an unmodified nucleotide base such that it is capable of hydrogen bonding with a base present on a target pre-mRNA. Examples of modified nucleotide bases include, without limitation, hypoxanthine, xanthine, 7-methylguanine, 5,6-dihydrouracil, 5-methylcytosine, and 5-hydroxymethoylcytosine.

[0126] Sometimes, the polynucleic acid polymers described herein also comprise a backbone structure that connects the components of an oligomer. The term “backbone structure” and “oligomer linkages” may be used interchangeably and refer to the connection between monomers of

the polynucleic acid polymer. In naturally occurring oligonucleotides, the backbone comprises a 3'-5' phosphodiester linkage connecting sugar moieties of the oligomer. The backbone structure or oligomer linkages of the polynucleic acid polymers described herein may include (but are not limited to) phosphorothioate, phosphorodithioate, phosphoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phosphoraniladate, phosphoramidate, and the like. See *e.g.*, LaPlanche *et al.*, Nucleic Acids Res. 14:9081 (1986); Stec *et al.* , J. Am. Chem. Soc. 106:6077 (1984), Stein *et al.* , Nucleic Acids Res. 16:3209 (1988), Zon *et al.* , Anti Cancer Drug Design 6:539 (1991); Zon *et al.* , Oligonucleotides and Analogues: A Practical Approach, pp. 87-108 (F. Eckstein, Ed., Oxford University Press, Oxford England (1991)); Stec *et al.* , U.S. Pat. No. 5,151,510; Uhlmann and Peyman, Chemical Reviews 90:543 (1990).

[0127] In embodiments, the stereochemistry at each of the phosphorus internucleotide linkages of the polynucleic acid polymer backbone is random. In embodiments, the stereochemistry at each of the phosphorus internucleotide linkages of the polynucleic acid polymer backbone is controlled and is not random. For example, U.S. Pat. App. Pub. No. 2014/0194610, "Methods for the Synthesis of Functionalized Nucleic Acids," incorporated herein by reference, describes methods for independently selecting the handedness of chirality at each phosphorous atom in a nucleic acid oligomer. In embodiments, a polynucleic acid polymer described herein comprises a polynucleic acid polymer having phosphorus internucleotide linkages that are not random. In embodiments, a composition used in the methods of the invention comprises a pure diastereomeric polynucleic acid polymer. In embodiments, a composition used in the methods of the invention comprises a polynucleic acid polymer that has diastereomeric purity of at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, about 100%, about 90% to about 100%, about 91% to about 100%, about 92% to about 100%, about 93% to about 100%, about 94% to about 100%, about 95% to about 100%, about 96% to about 100%, about 97% to about 100%, about 98% to about 100%, or about 99% to about 100%.

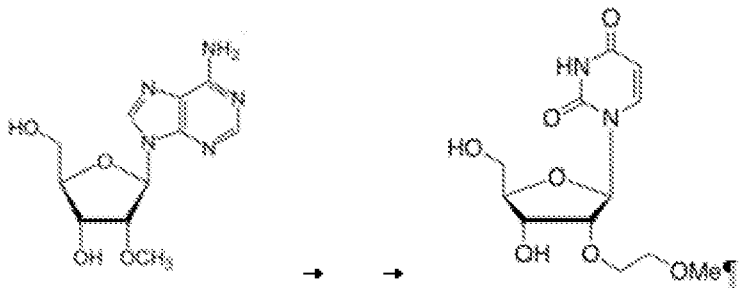
[0128] In embodiments, the polynucleic acid polymer has a nonrandom mixture of Rp and Sp configurations at its phosphorus internucleotide linkages. For example, a mix of Rp and Sp may be required in antisense oligonucleotides to achieve a balance between good activity and nuclease stability (Wan, *et al.* , 2014, "Synthesis, biophysical properties and biological activity of second generation antisense oligonucleotides containing chiral phosphorothioate linkages," *Nucleic Acids Research* 42(22): 13456-13468, incorporated herein by reference). In embodiments, a polynucleic acid polymer described herein comprises about 5-100% Rp, at least about 5% Rp, at least about 10%

Rp, at least about 15% Rp, at least about 20% Rp, at least about 25% Rp, at least about 30% Rp, at least about 35% Rp, at least about 40% Rp, at least about 45% Rp, at least about 50% Rp, at least about 55% Rp, at least about 60% Rp, at least about 65% Rp, at least about 70% Rp, at least about 75% Rp, at least about 80% Rp, at least about 85% Rp, at least about 90% Rp, or at least about 95% Rp, with the remainder Sp, or about 100% Rp. In embodiments, a polynucleic acid polymer described herein comprises about 10% to about 100% Rp, about 15% to about 100% Rp, about 20% to about 100% Rp, about 25% to about 100% Rp, about 30% to about 100% Rp, about 35% to about 100% Rp, about 40% to about 100% Rp, about 45% to about 100% Rp, about 50% to about 100% Rp, about 55% to about 100% Rp, about 60% to about 100% Rp, about 65% to about 100% Rp, about 70% to about 100% Rp, about 75% to about 100% Rp, about 80% to about 100% Rp, about 85% to about 100% Rp, about 90% to about 100% Rp, or about 95% to about 100% Rp, about 20% to about 80% Rp, about 25% to about 75% Rp, about 30% to about 70% Rp, about 40% to about 60% Rp, or about 45% to about 55% Rp, with the remainder Sp.

[0129] In embodiments, a polynucleic acid polymer described herein comprises about 5-100% Sp, at least about 5% Sp, at least about 10% Sp, at least about 15% Sp, at least about 20% Sp, at least about 25% Sp, at least about 30% Sp, at least about 35% Sp, at least about 40% Sp, at least about 45% Sp, at least about 50% Sp, at least about 55% Sp, at least about 60% Sp, at least about 65% Sp, at least about 70% Sp, at least about 75% Sp, at least about 80% Sp, at least about 85% Sp, at least about 90% Sp, or at least about 95% Sp, with the remainder Rp, or about 100% Sp. In embodiments, a polynucleic acid polymer described herein comprises about 10% to about 100% Sp, about 15% to about 100% Sp, about 20% to about 100% Sp, about 25% to about 100% Sp, about 30% to about 100% Sp, about 35% to about 100% Sp, about 40% to about 100% Sp, about 45% to about 100% Sp, about 50% to about 100% Sp, about 55% to about 100% Sp, about 60% to about 100% Sp, about 65% to about 100% Sp, about 70% to about 100% Sp, about 75% to about 100% Sp, about 80% to about 100% Sp, about 85% to about 100% Sp, about 90% to about 100% Sp, or about 95% to about 100% Sp, about 20% to about 80% Sp, about 25% to about 75% Sp, about 30% to about 70% Sp, about 40% to about 60% Sp, or about 45% to about 55% Sp, with the remainder Rp.

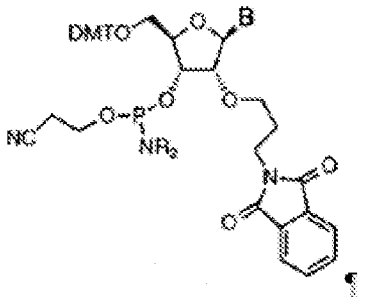
[0130] Nucleotide analogues or artificial nucleotide base may comprise a nucleic acid with a modification at a 2' hydroxyl group of the ribose moiety. The modification can be a 2'-O-methyl modification or a 2'-O-methoxyethyl (2'-O-MOE) modification. The 2'-O-methyl modification can add a methyl group to the 2' hydroxyl group of the ribose moiety whereas the 2'O-methoxyethyl modification can add a methoxyethyl group to the 2' hydroxyl group of the ribose moiety.

Exemplary chemical structures of a 2'-O-methyl modification of an adenosine molecule and 2'O-methoxyethyl modification of an uridine are illustrated below.



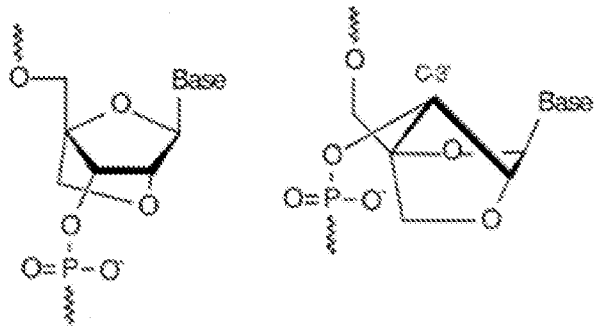
2'-O-methyl-adenosine → → 2'-O-methoxyethyl-uridine¶

[0131] An additional modification at the 2' hydroxyl group can include a 2'-O-aminopropyl sugar conformation which can involve an extended amine group comprising a propyl linker that binds the amine group to the 2' oxygen. This modification can neutralize the phosphate derived overall negative charge of the oligonucleotide molecule by introducing one positive charge from the amine group per sugar and can thereby improve cellular uptake properties due to its zwitterionic properties. An exemplary chemical structure of a 2'-O-aminopropyl nucleoside phosphoramidite is illustrated below.



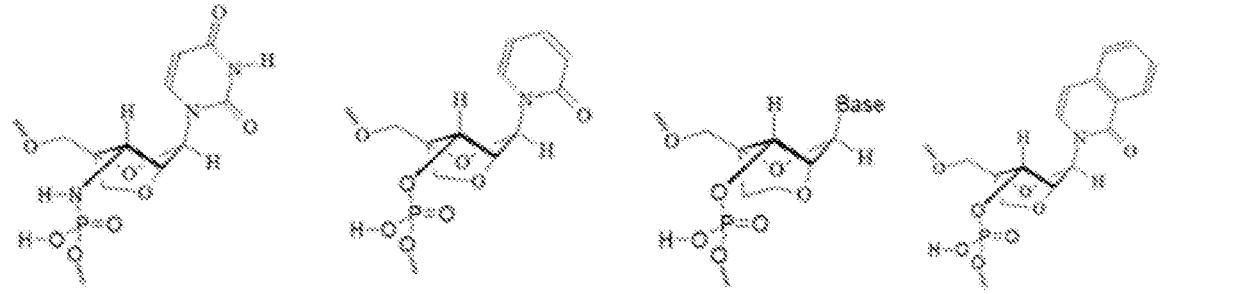
2'-O-aminopropyl nucleoside phosphoramidite¶

[0132] Another modification at the 2' hydroxyl group can include a locked or bridged ribose conformation (*e.g.*, locked nucleic acid or LNA) where the 4' ribose position can also be involved. In this modification, the oxygen molecule bound at the 2' carbon can be linked to the 4' carbon by a methylene group, thus forming a 2′-C,4′-C-oxy-methylene-linked bicyclic ribonucleotide monomer. Exemplary representations of the chemical structure of LNA are illustrated below. The representation shown to the left highlights the chemical connections of an LNA monomer. The representation shown to the right highlights the locked 3′-endo (3E) conformation of the furanose ring of an LNA monomer.



LNA (Locked Nucleic Acids)¶

[0133] A further modification at the 2' hydroxyl group may comprise ethylene nucleic acids (ENA) such as for example 2'-4'-ethylene-bridged nucleic acid, which locks the sugar conformation into a C3'-endo sugar puckering conformation. ENA are part of the bridged nucleic acids class of modified nucleic acids that also comprises LNA. Exemplary chemical structures of the ENA and bridged nucleic acids are illustrated below.

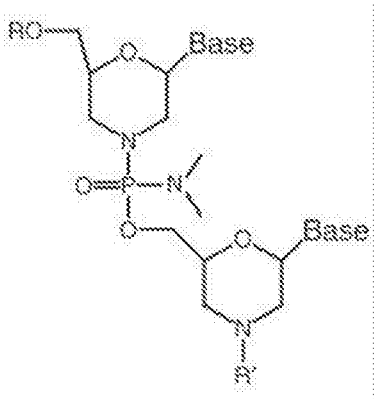


3'-amino-2',4'-BNA 2',4'-BNA-2-pyridone 2',4'-ENA 2',4'-BNA-1-isoquinolone

[0134] Still other modifications at the 2' hydroxyl group can include 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA).

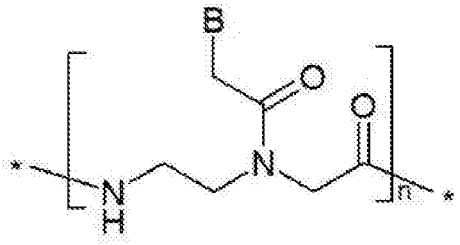
[0135] Nucleotide analogues may further comprise Morpholinos, peptide nucleic acids (PNAs), methylphosphonate nucleotides, thiolphosphonate nucleotides, 2'-fluoro N3-P5'-phosphoramidites, 1', 5'- anhydrohexitol nucleic acids (HNAs), or a combination thereof. Morpholino or phosphorodiamidate morpholino oligo (PMO) comprises synthetic molecules whose structure mimics natural nucleic acid structure by deviates from the normal sugar and phosphate structures. Instead, the five member ribose ring can be substituted with a six member morpholino ring containing four carbons, one nitrogen and one oxygen. The ribose monomers can be linked by a phosphordiamidate group instead of a phosphate group. These backbone alterations can remove all positive and negative charges making morpholinos neutral molecules that can cross cellular

membranes without the aid of cellular delivery agents such as those used by charged oligonucleotides.



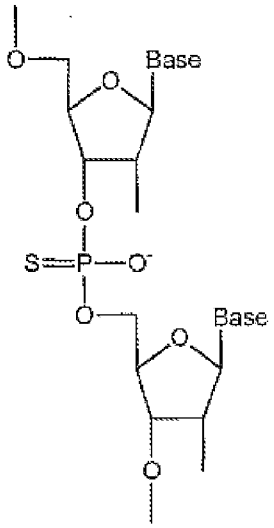
Morpholino

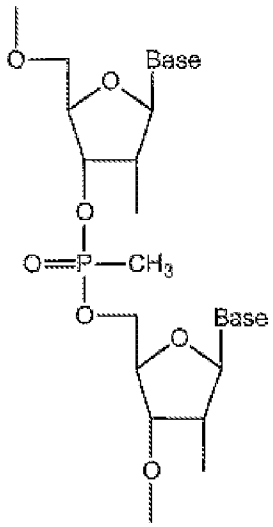
[0136] Peptide nucleic acid (PNA) does not contain sugar ring or phosphate linkage. Instead, the bases can be attached and appropriately spaced by oligoglycine-like molecules, therefore, eliminating a backbone charge.



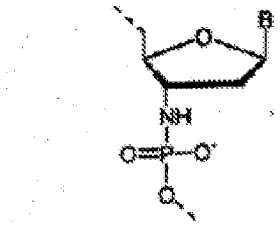
PNA

[0137] Modification of the phosphate backbone may also comprise methyl or thiol modifications such as methylphosphonate nucleotide and. Exemplary thiolphosphonate nucleotide (left) and methylphosphonate nucleotide (right) are illustrated below.



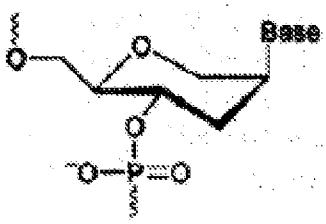


[0138] Furthermore, exemplary 2'-fluoro N3-P5'-phosphoramidites is illustrated as:



N3'-P5' Phosphoroamidate

[0139] And exemplary hexitol nucleic acid (or 1', 5'- anhydrohexitol nucleic acids (HNA)) is illustrated as:



Hexitol Nucleic Acid

[0140] In addition to modification of the ribose moiety, phosphate backbone and the nucleoside, the nucleotide analogues can also be modified by for example at the 3' or the 5' terminus. For example, the 3' terminus can include a 3' cationic group, or by inverting the nucleoside at the 3'-terminus with a 3'-3' linkage. In another alternative, the 3'-terminus can be blocked with an aminoalkyl group, *e.g.*, a 3' C5-aminoalkyl dT. The 5' -terminus can be blocked with an aminoalkyl group, *e.g.*, a 5'-O-

alkylamino substituent. Other 5' conjugates can inhibit 5'-3' exonucleolytic cleavage. Other 3' conjugates can inhibit 3'-5' exonucleolytic cleavage.

[0141] Unless specified otherwise, the left-hand end of single-stranded nucleic acid (*e.g.*, pre-mRNA transcript, oligonucleotide, SSO, etc.) sequences is the 5′ end and the left-hand direction of single or double-stranded nucleic acid sequences is referred to as the 5′ direction. Similarly, the right-hand end or direction of a nucleic acid sequence (single or double stranded) is the 3′ end or direction. Generally, a region or sequence that is 5′ to a reference point in a nucleic acid is referred to as "upstream," and a region or sequence that is 3′ to a reference point in a nucleic acid is referred to as "downstream." Generally, the 5′ direction or end of an mRNA is where the initiation or start codon is located, while the 3′ end or direction is where the termination codon is located. In some aspects, nucleotides that are upstream of a reference point in a nucleic acid may be designated by a negative number, while nucleotides that are downstream of a reference point may be designated by a positive number. For example, a reference point (*e.g.*, an exon-exon junction in mRNA) may be designated as the "zero" site, and a nucleotide that is directly adjacent and upstream of the reference point is designated "minus one," *e.g.*, "-1," while a nucleotide that is directly adjacent and downstream of the reference point is designated "plus one," *e.g.*, "+1."

[0142] In some cases, one or more of the artificial nucleotide analogues described herein are resistant toward nucleases such as for example ribonuclease such as RNase H, deoxyribonuclease such as DNase, or exonuclease such as 5'-3' exonuclease and 3'-5' exonuclease when compared to natural polynucleic acid polymers. In some instances, artificial nucleotide analogues comprising 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) modified, LNA, ENA, PNA, HNA, morpholino, methylphosphonate nucleotides, thiolphosphonate nucleotides, 2'-fluoro N3-P5'-phosphoramidites, or combinations thereof are resistant toward nucleases such as for example ribonuclease such as RNase H, deoxyribonuclease such as DNase, or exonuclease such as 5'-3' exonuclease and 3'-5' exonuclease. 2'-O-methyl modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'O-methoxyethyl (2'-O-MOE) modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-aminopropyl modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-deoxy modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3'

exonuclease or 3'-5' exonuclease resistance). T-deoxy-2'-fluoro modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-aminopropyl (2'-O-AP) modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-dimethylaminoethyl (2'-O-DMAOE) modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-dimethylaminopropyl (2'-O-DMAP) modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). T-O-dimethylaminoethyloxyethyl (2'-O-DMAEOE) modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-N-methylacetamido (2'-O-NMA) modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). LNA modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). ENA modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). HNA modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). Morpholinos may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). PNA can be resistant to nucleases (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). Methylphosphonate nucleotides modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). Thiolphosphonate nucleotides modified polynucleic acid polymer may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). Polynucleic acid polymer comprising 2'-fluoro N3-P5'-phosphoramidites may be nuclease resistance (*e.g.*, RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance).

[0143] In some instances, one or more of the artificial nucleotide analogues described herein have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. The one or more of the artificial nucleotide analogues comprising 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-0-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O-dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) modified, LNA, ENA, PNA, HNA, morpholino, methylphosphonate nucleotides, thiolphosphonate nucleotides, or 2'-fluoro N3-P5'-phosphoramidites can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-methyl modified

polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-methoxyethyl (2'-O-MOE) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-aminopropyl modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-deoxy modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. T-deoxy-2'-fluoro modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-aminopropyl (2'-O-AP) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-dimethylaminoethyl (2'-O-DMAOE) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-dimethylaminopropyl (2'-O-DMAP) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. T-O-dimethylaminoethyloxyethyl (2'-O-DMAEOE) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-N-methylacetamido (2'-O-NMA) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. LNA modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. ENA modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. PNA modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. HNA modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. Morpholino modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. Methylphosphonate nucleotides modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. Thiolphosphonate nucleotides modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. Polynucleic acid polymer comprising 2'-fluoro N3-P5'-phosphoramidites can have increased binding affinity toward their mRNA target relative to an

equivalent natural polynucleic acid polymer. The increased affinity can be illustrated with a lower Kd, a higher melt temperature (Tm), or a combination thereof.

[0144] In additional instances, a polynucleic acid polymer described herein may be modified to increase its stability. In an embodiment where the polynucleic acid polymer is RNA, the polynucleic acid polymer may be modified to increase its stability. The polynucleic acid polymer may be modified by one or more of the modifications described above to increase its stability. The polynucleic acid polymer may be modified at the 2' hydroxyl position, such as by 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-0-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) modification or by a locked or bridged ribose conformation (*e.g.*, LNA or ENA). The polynucleic acid polymer may be modified by 2'-O-methyl and/or 2'-O-methoxyethyl ribose. The polynucleic acid polymer may also include morpholinos, PNAs, HNA, methylphosphonate nucleotides, thiolphosphonate nucleotides, or 2'-fluoro N3-P5'-phosphoramidites to increase its stability. Suitable modifications to the RNA to increase stability for delivery will be apparent to the skilled person.

[0145] A polynucleic acid polymer described herein can be constructed using chemical synthesis and/or enzymatic ligation reactions using procedures known in the art. For example, a polynucleic acid polymer can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the polynucleic acid polymer and target nucleic acids. Exemplary methods can include those described in: US5,142,047; US5,185,444; WO2009099942; or EP1579015. Additional exemplary methods can include those described in: Griffey *et al.* , "2'-O-aminopropyl ribonucleotides: a zwitterionic modification that enhances the exonuclease resistance and biological activity of antisense oligonucleotides," *J. Med. Chem.* 39(26):5100-5109 (1997)); Obika, *et al.*, "Synthesis of 2'-O,4'-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -endo sugar puckering". Tetrahedron Letters38 (50): 8735(1997); Koizumi, M. "ENA oligonucleotides as therapeutics". Current opinion in molecular therapeutics8 (2): 144–149 (2006); and Abramova *et al.* , "Novel oligonucleotide analogues based on morpholino nucleoside subunits-antisense technologies: new chemical possibilities," *Indian Journal of Chemistry* 48B:1721-1726 (2009). Alternatively, the polynucleic acid polymer can be produced biologically using an expression vector into which a polynucleic acid polymer has been

subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted polynucleic acid polymer will be of an antisense orientation to a target polynucleic acid polymer of interest).

[0146] A polynucleic acid polymer may be bound to any nucleic acid molecule, such as another antisense molecule, a peptide, or other chemicals to facilitate delivery of the polynucleic acid polymer and/or target the nucleic acid to a specific tissue, cell type, or cell developmental stage. The polynucleic acid polymer may be bound to a protein or RNA. The protein tethered to the polynucleic acid polymer may comprise a splicing factor to enhance, inhibit or modulate splicing and intron removal. RNA tethered to the polynucleic acid polymer may comprise an aptamer or any structure that enhance, inhibit or modulate splicing and intron removal. The polynucleic acid polymer may be isolated nucleic acid.

[0147] A polynucleic acid polymer may be conjugated to, or bound by, a delivery vehicle suitable for delivering the polynucleic acid polymer to cells. The cells may be a specific cell type, or specific developmental stage. The delivery vehicle may be capable of site specific, tissue specific, cell specific or developmental stage-specific delivery. For example, the delivery vehicle may be a cell specific viral particle, or component thereof, alternatively, the delivery vehicle may be a cell specific antibody particle, or component thereof. The polynucleic acid polymer may be targeted for delivery to beta cells in the pancreas. The polynucleic acid polymer may be targeted for delivery to thymic cells. The polynucleic acid polymer may be targeted for delivery to malignant cells. The polynucleic acid polymer may be targeted for delivery to pre-malignant cells (that are known to develop into overt malignant phenotypes within a foreseeable future, such as pre-leukemias and myelodysplastic syndromes or histopathologically defined precancerous lesions or conditions.

[0148] A polynucleic acid polymer may be conjugated to, or bound by, a nanoparticle-based delivery vehicle. A nanoparticle may be a metal nanoparticle, *e.g.*, a nanoparticle of scandium, titanium, vanadium, chromium, manganese, iron, cobalt, nickel, copper, zinc, yttrium, zirconium, niobium, molybdenum, ruthenium, rhodium, palladium, silver, cadmium, hafnium, tantalum, tungsten, rhenium, osmium, iridium, platinum, gold, gadolinium, aluminum, gallium, indium, tin, thallium, lead, bismuth, magnesium, calcium, strontium, barium, lithium, sodium, potassium, boron, silicon, phosphorus, germanium, arsenic, antimony, and combinations, alloys or oxides thereof. Sometimes a nanoparticle may be prepared from polymeric materials. Illustrative polymeric materials include, but are not limited to, poly(ethylenimine) (PEI), poly(alkylcyanoacrylates), poly(amidoamine) dendrimers (PAMAM), poly(ε-caprolactone) (PCL), poly(lactic-co-glycolic acid) (PLGA), or polyesters (poly(lactic acid) (PLA). Sometimes a nanoparticle may be further coated with molecules for attachment of functional elements. In some cases, a coating comprises

chondroitin sulfate, dextran sulfate, carboxymethyl dextran, alginic acid, pectin, carragheenan, fucoidan, agaropectin, porphyran, karaya gum, gellan gum, xanthan gum, hyaluronic acids, glucosamine, galactosamine, chitin (or chitosan), polyglutamic acid, polyaspartic acid, lysozyme, cytochrome C, ribonuclease, trypsinogen, chymotrypsinogen, α-chymotrypsin, polylysine, polyarginine, histone, protamine, graphene, ovalbumin or dextrin or cyclodextrin. A nanoparticle may include a core or a core and a shell, as in a core-shell nanoparticle. Sometimes, a nanoparticle may have at least one dimension of less than about 500nm, 400nm, 300nm, 200nm, or 100nm.

[0149] In some embodiments, a polynucleic acid polymer may be formulated with a nanoparticle-based delivery vehicle for delivery to a site of interest (*e.g.*, a malignant tissue site or a cell with deregulated protein expression). In some cases, a polynucleic acid polymer may be formulated with a nanoparticle-based delivery vehicle to facilitate and/or enable transport across the blood-brain barrier (BBB).

[0150] Sometimes, a polynucleic acid polymer is coupled to a substance, known in the art to promote penetration or transport across the blood-brain barrier, *e.g.*, an antibody to the transferrin receptor. In some embodiments, the polynucleic acid polymer is linked with a viral vector, *e.g.*, to render the compound more effective or increase transport across the blood-brain barrier. In some embodiments, osmotic blood brain barrier disruption is assisted by infusion of sugars, *e.g.*, meso erythritol, xylitol, D(+) galactose, D(+) lactose, D(+) xylose, dulcitol, myo-inositol, L(-) fructose, D(-) mannitol, D(+) glucose, D(+) arabinose, D(-) arabinose, cellobiose, D(+) maltose, D(+) raffinose, L(+) rhamnose, D(+) melibiose, D(-) ribose, adonitol, D(+) arabitol, L(-) arabitol, D(+) fucose, L(-) fucose, D(-) lyxose, L(+) lyxose, and L(-) lyxose, or amino acids, *e.g.*, glutamine, lysine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glycine, histidine, leucine, methionine, phenylalanine, proline, serine, threonine, tyrosine, valine, and taurine. Methods and materials for enhancing blood brain barrier penetration are described, *e.g.*, in U.S. Pat. No. 4,866,042, U.S. Pat. No. 6,294,520 and U.S. Pat. No. 6,936,589, each incorporated herein by reference.

[0151] In one embodiment the polynucleic acid polymer may be bound to a chemical molecule (*e.g.*, non-peptide or nucleic acid based molecule), such as a drug. The drug may be a small molecule (*e.g.*, having a MW of less than 900Da).

[0152] In one embodiment of the invention, the delivery vehicle may comprise a cell penetrating peptide (CPP). For example, the polynucleic acid polymer may be bound or complexed with a CPP. The skilled person will understand that any suitable CPP may be conjugated with the polynucleic acid polymer to aid delivery of the polynucleic acid polymer to and/or into cells. Such CPPs may be any suitable CPP technology described by Boisguérin *et al.*, Advanced Drug Delivery Reviews

(2015), which is herein incorporated by reference. Suitable delivery vehicles for conjugation to the polynucleic acid polymer are also described in Lochmann *et al.*, ((*European Journal of Pharmaceutics and Biopharmaceutics* 58 (2004) 237–251), which is herein incorporated by reference).

[0153] The CPP may be an arginine and/or lysine rich peptide, for example, wherein the majority of residues in the peptide are either lysine or arginine. The CPP may comprise a poly-L-lysine (PLL). Alternatively, the CPP may comprise a poly-arginine. Suitable CPPs may be selected from the group comprising Penetratin; R6-Penetratin; Transportan; oligo-arginines; F-3; B-peptide; B-MSP; Pip peptides, such as Pip1, Pip2a, Pip2b, Pip5e, Pip5f, Pip5h, Pip5j; Pip5k, Pip5l, Pip5m, Pip5n, Pip5o, Pip6a, Pip6b, Pip6c, Pip6d, Pip6e, Pip6f, Pip6g, or Pip6h; peptide of sequence PKKKRKV; Penatratin; Lys4; SPACE; Tat; Tat-DRBD (dsRNA-binding domain); (RXR)4; (RFF)3RXB; (KFF)3K; RgF2; T-cell derived CPP; Pep-3; PEGpep-3; MPG-8; MPG-8-Chol; PepFect6; P5RHH; R15; and Chol-R9; or functional variants thereof (*e.g.*, see Boisguérin *et al.*, Advanced Drug Delivery Reviews (2015)).

[0154] In one embodiment, the CPP comprises or consists of a Pip peptide. The Pip peptide may be selected from the group comprising Pip1, Pip2a, Pip2b, Pip5e, Pip5f, Pip5h, Pip5j; Pip5k, Pip5l, Pip5m, Pip5n, Pip5o, Pip6a, Pip6b, Pip6c, Pip6d, Pip6e, Pip6f, Pip6g, and Pip6h.

[0155] In one embodiment of the invention, the delivery vehicle may comprise a peptide-based nanoparticle (PBN), wherein a plurality of CPPs (for example one or more suitable CPPs discussed herein) form a complex with the polynucleic acid polymer through charge interactions. Such nanoparticles may be between about 50nm and 250nm in size. In one embodiment the nanoparticles may be about 70-200nm in size. In another embodiment the nanoparticles may be about 70-100nm in size or 125-200nm in size.

[0156] In one embodiment, the polynucleic acid polymer may be complexed with a delivery vehicle, for example by ionic bonding. Alternatively, the polynucleic acid polymer may be covalently bound to the delivery vehicle. Conjugation/binding methods are described in Lochmann *et al.*, ((*European Journal of Pharmaceutics and Biopharmaceutics* 58 (2004) 237–251), which is herein incorporated by reference). For example, a conjugation method may comprise introducing a suitable tether containing a reactive group (*e.g.*, $-NH_2$ or $-SH_2$) to the polynucleic acid polymer and to add the delivery vehicle, such as a peptide, post-synthetically as an active intermediate, followed by carrying out the coupling reaction in aqueous medium. An alternative method may comprise carrying out the conjugation in a linear mode on a single solid-phase support.

[0157] The delivery vehicle and polynucleic acid polymer may be thiol and/or maleimide linked, such as thiol-maleimide linked. The conjugation of the polynucleic acid polymer and the delivery vehicle may be by click-chemistry, such as reaction of azido or 2'-O-propyargyl functional groups and alkyne groups on the respective molecules to be conjugated. In one embodiment, the delivery vehicle and polynucleic acid polymer may be linked by a thioether bridge. In another embodiment, the delivery vehicle and polynucleic acid polymer may be linked by a disulphide bridge. The skilled person will readily identify suitable linking groups or reactions for conjugation of polynucleic acid polymer and the delivery vehicle, such as a peptide.

[0158] In one embodiment the NSE repressor agent may comprise an SSO of the sequence cuucuaugcagccaaccuguagacu (SSO –NSE3) (SEQ ID NO: 53), or a nucleic acid analogue thereof. In one embodiment the NSE repressor agent may comprise an SSO of the sequence accuuuuucuucuaugcagccaac (SSO –NSE5) (SEQ ID NO: 54), or a nucleic acid analogue thereof. The skilled person will note that NSE3 (cuucuaugcagccaaccuguagacu) (SEQ ID NO: 53) and NSE5 (accuuuuucuucuaugcagccaac) (SEQ ID NO: 54) overlap in sequence. In one embodiment, the NSE repressor agent may comprise an SSO having a sequence of, or within, this overlapping sequence (*i.e.* accuuuuucuucuaugcagccaaccuguagacu) (SEQ ID NO: 55).

[0159] In one embodiment, the NSE repressor or activator agent comprises or consists of any one SSO selected from the group comprising:

aacuuaaagguuauaucuc (SSO A2) (SEQ ID NO: 18);

uauaaauacgaauaaaucga (SSO A4) (SEQ ID NO: 19);

caacacgacauaaccaaa (SSO A9) (SEQ ID NO: 21);

aacauuucuauuuaguuaaaagc (SSO A11) (SEQ ID NO: 23);

uuaguauuccuugacuuua (SSO A17) (SEQ ID NO: 26);

gguaugagaacuauagga (SSO A23) (SEQ ID NO: 32);

gguaauaagugucacaaa (SSO A25) (SEQ ID NO: 34);

guaucauacauuagaagg (SSO A26);

gacugguaaauaauaaacauaauuc (SSO B2);

auauauuagagauacaucagcc (SSO B4);

ugugggugaccacagcuu (SSO B11);

uuagagaaucauuuuaaauaagac (SSO AN3); and

cuguaaaagaaaauaga (PEkr),

or combinations thereof.

[0160] In another embodiment, the NSE activator agent comprises or consists of any one SSO selected from the group comprising:

aacuuaaagguuauaucuc (SSO A2) (SEQ ID NO: 18);

uauaaauacgaauaaaucga (SSO A4) (SEQ ID NO: 19);

caacacgacauaaccaaa (SSO A9) (SEQ ID NO: 21);

gguaugagaacuauagga (SSO A23) (SEQ ID NO: 32);

gguaauaagugucacaaa (SSO A25) (SEQ ID NO: 34);

guaucauacauuagaagg (SSO A26) (SEQ ID NO: 35);

ugugggugaccacagcuu (SSO B11) (SEQ ID NO: 45); and

cuguaaaagaaaauaga (PEkr) (SEQ ID NO: 56),

or combinations thereof.

[0161] The NSE activator agent may comprise or consist of an SSO of the sequence aacuuaaagguuauaucuc (SSO A2) (SEQ ID NO: 18). The NSE activator agent may comprise or consist of an SSO of the sequence uauaaauacgaauaaaucga (SSO A4) (SEQ ID NO: 19). The NSE activator agent may comprise or consist of an SSO of the sequence caacacgacauaaccaaa (SSO A9) (SEQ ID NO: 21). The NSE activator agent may comprise or consist of an SSO of the sequence gguaugagaacuauagga (SSO A23) (SEQ ID NO: 32). The NSE activator agent may comprise or consist of an SSO of the sequence gguaauaagugucacaaa (SSO A25) (SEQ ID NO: 34). The NSE activator agent may comprise or consist of an SSO of the sequence guaucauacauuagaagg (SSO A26) (SEQ ID NO: 35). The NSE activator agent may comprise or consist of an SSO of the sequence ugugggugaccacagcuu (SSO B11) (SEQ ID NO: 45).

[0162] In one embodiment the NSE-activator agent may comprise the SSO PEkr herein described. In one embodiment the NSE-activator agent may comprise an SSO of the sequence CUGUAAAAGAAAAUAGA (PEkr) (SEQ ID NO: 56). PEkr may also be referred to as PEdel and it is understood that these terms are interchangeable.

[0163] In one embodiment, the NSE repressor agent comprises or consists of any one SSO selected from the group comprising:

cuucuaugcagccaaccuguagacu (SSO –NSE3) (SEQ ID NO: 53);

accuuuuucuucuaugcagccaac (SSO –NSE5) (SEQ ID NO: 54);

aacauuucuauuuaguuaaaagc (SSO A11) (SEQ ID NO: 23);

uuaguauuccuugacuuua (SSO A17) (SEQ ID NO: 26);

gacugguaaauauauaaacauaauuc (SSO B2) (SEQ ID NO: 37);

auauauuagagauacaucagcc (SSO B4) (SEQ ID NO: 39); and

uuagagaaucauuuuaaauaagac (SSO AN3) (SEQ ID NO: 51), or combinations thereof.

[0164] The NSE repressor agent may comprise or consist of an SSO of the sequence cuucuaugcagccaaccuguagacu (SSO –NSE3) (SEQ ID NO: 53). The NSE repressor agent may comprise or consist of an SSO of the sequence accuuuuucuucuaugcagccaac (SSO –NSE5) (SEQ ID NO: 54). The NSE repressor agent may comprise or consist of an SSO of the sequence aacauuucuauuuaguuaaaagc (SSO A11) (SEQ ID NO: 23). The NSE repressor agent may comprise or consist of an SSO of the sequence uuaguauuccuugacuuua (SSO A17) (SEQ ID NO: 26). The NSE repressor agent may comprise or consist of an SSO of the sequence gacugguaaauauaaaacauaauuc (SSO B2) (SEQ ID NO: 37). The NSE repressor agent may comprise or consist of an SSO of the sequence auauauuagagauacaucagcc (SSO B4) (SEQ ID NO: 39). The NSE repressor agent may comprise or consist of an SSO of the sequence uuagagaaucauuuuaaauaagac (SSO AN3) (SEQ ID NO: 51).

[0165] In one embodiment the NSE repressor agent, such as an SSO, may be arranged to bind to guanine variant residue at rs4988000.

[0166] The skilled person will understand that combinations of two or more SSOs described herein may be provided and/or used for treatment. For example, combinations of two, three, four, five or more NSE repressor agents may be provided or combinations of two, three, four, five or more NSE activating agents may be provided.

[0167] Where reference is made to reducing NSE inclusion in the mature RNA, the reduction may be complete, *e.g.*, 100%, or may be partial. The reduction may be clinically significant. The reduction/correction may be relative to the level of NSE inclusion in the subject without treatment, or relative to the amount of NSE inclusion in a population of similar subjects. The reduction/correction may be at least 10% less NSE inclusion relative to the average subject, or the subject prior to treatment. The reduction may be at least 20% less NSE inclusion relative to an average subject, or the subject prior to treatment. The reduction may be at least 40% less NSE inclusion relative to an average subject, or the subject prior to treatment. The reduction may be at least 50% less NSE inclusion relative to an average subject, or the subject prior to treatment. The reduction may be at least 60% less NSE inclusion relative to an average subject, or the subject prior to treatment. The reduction may be at least 80% less NSE inclusion relative to an average subject, or the subject prior to treatment. The reduction may be at least 90% less NSE inclusion relative to an average subject, or the subject prior to treatment.

[0168] Where reference is made to increasing active-ATM protein levels, the increase may be clinically significant. The increase may be relative to the level of active-ATM protein in the subject

without treatment, or relative to the amount of active-ATM protein in a population of similar subjects. The increase may be at least 10% more active-ATM protein relative to the average subject, or the subject prior to treatment. The increase may be at least 20% more active-ATM protein relative to the average subject, or the subject prior to treatment. The increase may be at least 40% more active-ATM protein relative to the average subject, or the subject prior to treatment. The increase may be at least 50% more active-ATM protein relative to the average subject, or the subject prior to treatment. The increase may be at least 80% more active-ATM protein relative to the average subject, or the subject prior to treatment. The increase may be at least 100% more active-ATM protein relative to the average subject, or the subject prior to treatment. The increase may be at least 200% more active-ATM protein relative to the average subject, or the subject prior to treatment. The increase may be at least 500% more active-ATM protein relative to the average subject, or the subject prior to treatment.

[0169] The terms active-ATM and functional-ATM may be used interchangeably herein.

[0170] According to another aspect of the invention, there is provided use of rs609261 genotyping to predict a subject response to therapy for conditions associated with ATM deregulation.

[0171] The conditions associated with ATM deregulation may comprise A-T or cancer.

[0172] In one embodiment, the presence of an rs609261 cytosine residue is associated with a higher NSE activation, less efficient response of ATM to DNA double-strand break signaling, a higher cancer risk and lower survival relative to non-cytosine residue at the same position.

[0173] According to another aspect of the invention, there is provided a composition comprising the NSE repressor agent of the invention herein.

[0174] According to another aspect of the invention, there is provided a composition comprising the NSE activator agent of the invention herein.

[0175] In one embodiment, the composition is a pharmaceutically acceptable formulation.

[0176] The composition may comprise at least one other biologically active molecule in addition to the polynucleic acid polymer. The biologically active molecule may be drug or a pro-drug. The biologically active molecule may comprise nucleic acid or amino acid. The biologically active molecule may comprise a small molecule (*e.g.*, a molecule of <900 Daltons).

[0177] In some embodiments, pharmaceutical formulations described herein are administered to a subject by an enteral administration route, by a parenteral administration route, or by a topical administration route. In some cases, pharmaceutical formulations described herein are administered to a subject by an enteral administration route. In other cases, pharmaceutical formulations described herein are administered to a subject by a parenteral administration route. In additional cases,

pharmaceutical formulations described herein are administered to a subject by a topical administration route.

[0178] Illustrative administration routes include, but are not limited to, parenteral (*e.g.*, intravenous, subcutaneous, intramuscular, intra-arterial, intracranial, intracerebral, intracerebroventricular, intrathecal, or intravitreal), oral, intranasal, buccal, topical, rectal, transmucosal, or transdermal administration routes. In some instances, the pharmaceutical composition describe herein is formulated for parenteral (*e.g.*, intravenous, subcutaneous, intramuscular, intra-arterial, intracranial, intracerebral, intracerebroventricular, intrathecal, or intravitreal) administration. In other instances, the pharmaceutical composition describe herein is formulated for oral administration. In still other instances, the pharmaceutical composition describe herein is formulated for intranasal administration.

[0179] Pharmaceutical formulations described herein may include, but are not limited to, aqueous liquid dispersions, self-emulsifying dispersions, solid solutions, liposomal dispersions, aerosols, solid dosage forms, powders, immediate release formulations, controlled release formulations, fast melt formulations, tablets, capsules, pills, delayed release formulations, extended release formulations, pulsatile release formulations, multiparticulate formulations, and mixed immediate and controlled release formulations.

[0180] Pharmaceutical formulations may include a carrier or carrier materials which may include any commonly used excipients in pharmaceutics and should be selected on the basis of compatibility with the composition disclosed herein, and the release profile properties of the desired dosage form. Exemplary carrier materials include, *e.g.*, binders, suspending agents, disintegration agents, filling agents, surfactants, solubilizers, stabilizers, lubricants, wetting agents, diluents, and the like. Pharmaceutically compatible carrier materials may include, but are not limited to, acacia, gelatin, colloidal silicon dioxide, calcium glycerophosphate, calcium lactate, maltodextrin, glycerin, magnesium silicate, polyvinylpyrrollidone (PVP), cholesterol, cholesterol esters, sodium caseinate, soy lecithin, taurocholic acid, phosphotidylcholine, sodium chloride, tricalcium phosphate, dipotassium phosphate, cellulose and cellulose conjugates, sugars sodium stearoyl lactylate, carrageenan, monoglyceride, diglyceride, pregelatinized starch, and the like. Liposomes can include sterically stabilized liposomes, *e.g.*, liposomes comprising one or more specialized lipids. These specialized lipids can result in liposomes with enhanced circulation lifetimes. Sometimes, a sterically stabilized liposome can comprise one or more glycolipids or is derivatized with one or more hydrophilic polymers, such as a polyethylene glycol (PEG) moiety. See, *e.g.*, *Remington: The Science and Practice of Pharmacy*, Nineteenth Ed (Easton, Pa.: Mack Publishing Company, 1995);

Hoover, John E., Remington's *Pharmaceutical Sciences*, Mack Publishing Co., Easton, Pennsylvania 1975; Liberman, H.A. and Lachman, L., Eds., *Pharmaceutical Dosage Forms*, Marcel Decker, New York, N.Y., 1980; and *Pharmaceutical Dosage Forms and Drug Delivery Systems*, Seventh Ed. (Lippincott Williams & Wilkins 1999).

[0181] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome, wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to replace the non-thymine variant residue rs609261 with a thymine residue.

[0182] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising replacing a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of NSE (cryptic exon in ATM intron 28) of the human genome with a thymine residue.

[0183] In one embodiment, replacing the non-thymine variant residue rs609261 may comprise administration of an agent to the subject, which is arranged to replace the non-thymine variant residue rs609261 with a thymine residue.

[0184] The agent for replacement of the non-thymine residue may be a genomic editing molecule, such as CRISPR-Cas9, or a functional equivalent thereof, together with an appropriate RNA molecule arranged to target rs609261.

[0185] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to replace the guanine variant residue at rs4988000 with adenine.

[0186] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising replacing a guanine variant residue at rs4988000 of the human genome with an adenine residue.

[0187] In one embodiment, replacing the guanine variant residue at rs4988000 may comprise administration of an agent to the subject, which is arranged to replace the guanine variant residue at rs4988000 with an adenine residue.

[0188] The agent for replacement of the guanine residue may be a genomic editing molecule, such as CRISPR-Cas9, or a functional equivalent thereof, together with an appropriate RNA molecule arranged to target rs4988000.

[0189] According to a first aspect of the invention, there is provided a method of screening a subject or a population of subjects for susceptibility to functional-ATM protein deficiency, wherein the screening comprises determining the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject (or group of subjects) has, or is susceptible to, functional-ATM protein deficiency.

[0190] According to another aspect of the invention, there is provided a method of selecting a subject or a population of subjects for treatment or prophylaxis, wherein the subject is susceptible to functional-ATM protein deficiency, the method comprising determining the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and selecting such subject for treatment with an agent arranged to increase functional-ATM levels in the subject.

[0191] According to another aspect of the invention, there is provided a method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying the presence of a guanine variant residue at rs4988000 of the human genome, wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, functional-ATM protein deficiency, and administration of an agent to the subject, which is arranged to increase functional-ATM levels.

[0192] The methods of the invention herein may comprise blocking a guanine variant residue at rs4988000, for example using an SSO.

[0193] PE contains a natural DNA variant rs4988000 (G/A), which also influences NSE recognition (Fig. 4H). Transfections of C and T minigenes systematically mutated at rs4988000 revealed that the rare A allele decreased NSE inclusion on each pre-mRNA, both in U2AF35- and mock-depleted cells. Therefore, replacement of the guanine residue with adenine will decrease NSE inclusion, and increase the level of functional ATM-protein.

[0194] The highest NSE inclusion is produced by the haplotype that is most frequent in Caucasians (CG), followed by haplotypes CA>TG>TA (referring tors609261 and rs4988000 respectively). Therefore, the methods and compositions of the invention may be used in combination (concurrently or sequentially) to modify a CG haplotype to CA, TG, or TA. In one embodiment, the methods and compositions of the invention may be used to modify a CG haplotype to TA. In one embodiment, the

methods and compositions of the invention may be used to modify a CA haplotype to TG or TA. In one embodiment, the methods and compositions of the invention may be used to modify a CA haplotype to TA. In one embodiment, the methods and compositions of the invention may be used to modify a TG haplotype to TA.

[0195] The methods and compositions of the invention may also be used in combination (concurrently or sequentially) to identify a CG haplotype in a subject, and optionally treat or select the patient according to the invention. The methods and compositions of the invention may also be used in combination (concurrently or sequentially) to identify a CA haplotype in a subject, and optionally treat or select the patient according to the invention. The methods and compositions of the invention may also be used in combination (concurrently or sequentially) to identify a TG haplotype in a subject, and optionally treat or select the patient according to the invention.

[0196] According to another aspect of the invention, there is provided a method of modifying regulation of NSE inclusion in a mature RNA transcript, the method comprising the insertion or deletion of one or more splicing regulatory motifs upstream or downstream of the NSEs that compete with the NSE for spliceosomal components, said regulatory motifs comprising cryptic splice sites or pseudo-exons.

[0197] According to another aspect of the invention, there is provided a method of modifying regulation of a functional protein expression, wherein the functional protein expression is regulated by NSE inclusion in a mature RNA transcript of the gene encoding protein, the method comprising the insertion or deletion of one or more splicing regulatory motifs upstream or downstream of the NSE that compete with the NSE for spliceosomal components, said regulatory motifs comprising cryptic splice sites or pseudo-exons.

[0198] In one embodiment, the insertion or deletion of one or more splicing regulatory motifs is in genomic DNA of ATM intron 28.

[0199] The insertion of one or more splicing regulatory motifs may cause a reduction in NSE inclusion in the mature RNA transcript. The deletion of one or more splicing regulatory motifs may cause an increase in NSE inclusion in the mature RNA transcript.

[0200] The insertion or deletion of one or more splicing regulatory motifs may comprise the use of genome editing technology, such as CRISPR-Cas9. CRISPR-Cas9 may be provided with an appropriate targeting RNA molecule.

[0201] The subject or cells that are treated or screened according to the invention may be mammalian. In one embodiment, the subject is a human. In one embodiment, the cells are human.

[0202] Kits and articles of manufacture are provided herein for use with one or more methods described herein. The kits can contain one or more of the polynucleic acid polymers described herein.

[0203] According to another aspect of the invention, there is provided a kit comprising one or more oligonucleotide probes for identifying rs609261 and/or rs4988000 variants.

[0204] The skilled person will be familiar with techniques for probing the presence or absence of genetic sequence features. For example, the oligonucleotide probes may comprise primers for use in PCR amplifying a region of a nucleic acid comprising rs609261 and/or rs4988000. In another embodiment the oligonucleotide probes may directly bind rs609261 or rs4988000, wherein the binding may be detectable. The binding of the probe may be detectable for example using SERS or SERRS technology.

[0205] The kits can also include a carrier, package, or container that is compartmentalized to receive one or more containers such as vials, tubes, and the like, each of the container(s) comprising one of the separate elements, such as the polynucleic acid polymers and reagents, to be used in a method described herein. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers can be formed from a variety of materials such as glass or plastic.

[0206] The articles of manufacture provided herein contain packaging materials. Examples of pharmaceutical packaging materials include, but are not limited to, bottles, tubes, bags, containers, bottles, and any packaging material suitable for a selected formulation and intended mode of administration and treatment.

[0207] A kit typically includes labels listing contents and/or instructions for use, and package inserts with instructions for use. A set of instructions will also typically be included.

[0208] According to another aspect of the invention, there is provided a vector comprising the polynucleic acid polymer of the invention.

[0209] The vector may comprise a viral vector. The viral vector may comprise adeno-associated viral vector. The vector may comprise any virus that targets the polynucleic acid polymer to malignant cells or specific cell type.

Indications

[0210] In some instances, compositions and methods described herein are used to treat a genetic disorder or condition such as a hereditary disease. Compositions and methods described herein can be used to treat a genetic disorder or condition such as a hereditary disease that is characterized by an impaired production of a protein. Compositions and methods described herein can be used to treat

a genetic disorder or condition such as a hereditary disease that is characterized by a defective splicing.

[0211] Compositions and methods described herein can also be used to treat a genetic disorder or condition such as an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder. Compositions and methods described herein can be used to treat an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder, in which the disorder or condition is characterized by an impaired production of a protein. Compositions and methods described herein can also be used to treat an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder, in which the disorder or condition is characterized by a defective splicing.

[0212] The condition associated with deregulated ATM expression may comprise cancer. Compositions and methods described herein can be used to treat cancer. In one embodiment the cancer comprises breast cancer. Cancer can be a solid tumor or a hematologic malignancy. A solid tumor can be a sarcoma or a carcinoma. Sarcoma can be a cancer of bone, cartilage, fat muscle, vascular or hematopoietic tissues. Exemplary sarcoma can include alveolar rhabdomyosarcoma, alveolar soft part sarcoma, ameloblastoma, angiosarcoma, chondrosarcoma, chordoma, clear cell sarcoma of soft tissue, dedifferentiated liposarcoma, desmoid, desmoplastic small round cell tumor, embryonal rhabdomyosarcoma, epithelioid fibrosarcoma, epithelioid hemangioendothelioma, epithelioid sarcoma, esthesioneuroblastoma, Ewing sarcoma, extrarenal rhabdoid tumor, extraskeletal myxoid chondrosarcoma, extraskeletal osteosarcoma, fibrosarcoma, giant cell tumor, hemangiopericytoma, infantile fibrosarcoma, inflammatory myofibroblastic tumor, Kaposi sarcoma, leiomyosarcoma of bone, liposarcoma, liposarcoma of bone, malignant fibrous histiocytoma (MFH), malignant fibrous histiocytoma (MFH) of bone, malignant mesenchymoma, malignant peripheral nerve sheath tumor, mesenchymal chondrosarcoma, myxofibrosarcoma, myxoid liposarcoma, myxoinflammatory fibroblastic sarcoma, neoplasms with perivascular epitheioid cell differentiation, osteosarcoma, parosteal osteosarcoma, neoplasm with perivascular epitheioid cell differentiation, periosteal osteosarcoma, pleomorphic liposarcoma, pleomorphic rhabdomyosarcoma, PNET/extraskeletal Ewing tumor, rhabdomyosarcoma, round cell liposarcoma, small cell osteosarcoma, solitary fibrous tumor, synovial sarcoma, telangiectatic osteosarcoma.

[0213] Carcinoma can be a cancer developed from epithelial cells. Exemplary carcinoma can include adenocarcinoma, squamous cell carcinoma, adenosquamous carcinoma, anaplastic carcinoma, large cell carcinoma, small cell carcinoma, anal cancer, appendix cancer, bile duct cancer (i.e., cholangiocarcinoma), bladder cancer, brain tumor, breast cancer, cervical cancer, colon cancer, cancer of Unknown Primary (CUP), esophageal cancer, eye cancer, fallopian tube cancer, gastroenterological cancer, kidney cancer, liver cancer, lung cancer, medulloblastoma, melanoma, oral cancer, ovarian cancer, pancreatic cancer, parathyroid disease, penile cancer, pituitary tumor, prostate cancer, rectal cancer, skin cancer, stomach cancer, testicular cancer, throat cancer, thyroid cancer, uterine cancer, vaginal cancer, or vulvar cancer. Hematologic malignancy is a malignancy of the blood system and can include T-cell based and B-cell based malignancies. Exemplary hematologic malignancy can include myeloid leukemia, myeloproliferative neoplasias, peripheral T-cell lymphoma not otherwise specified (PTCL-NOS), anaplastic large cell lymphoma, angioimmunoblastic lymphoma, cutaneous T-cell lymphoma, adult T-cell leukemia/lymphoma (ATLL), blastic NK-cell lymphoma, enteropathy-type T-cell lymphoma, hematosplenic gamma-delta T-cell lymphoma, lymphoblastic lymphoma, nasal NK/T-cell lymphomas, treatment-related T-cell lymphomas, chronic lymphocytic leukemia (CLL), small lymphocytic lymphoma (SLL), high risk CLL, non-CLL/SLL lymphoma, prolymphocytic leukemia (PLL), follicular lymphoma (FL), diffuse large B-cell lymphoma (DLBCL), mantle cell lymphoma (MCL), Waldenstrom's macroglobulinemia, multiple myeloma, extranodal marginal zone B cell lymphoma, nodal marginal zone B cell lymphoma, Burkitt's lymphoma, non-Burkitt high grade B cell lymphoma, primary mediastinal B-cell lymphoma (PMBL), immunoblastic large cell lymphoma, precursor B-lymphoblastic lymphoma, B cell prolymphocytic leukemia, lymphoplasmacytic lymphoma, splenic marginal zone lymphoma, plasma cell myeloma, plasmacytoma, mediastinal (thymic) large B cell lymphoma, intravascular large B cell lymphoma, primary effusion lymphoma, or lymphomatoid granulomatosis.

[0214] According to another aspect of the invention, there is provided a method of a treatment or prevention of a disease pathology caused by an NSE inclusion in a pre-mRNA gene transcript comprising providing an agent arranged to bind to a cryptic splice site of a pseudoexon present on the pre-mRNA gene transcript, wherein the cryptic splice site is capable of regulating inclusion of a nonsense-mediated RNA decay switch exon (NSE)into a mature RNA transcript of the gene.

[0215] Wherein the binding of the agent to the cryptic splice site of the pseudoexon present on the pre-mRNA gene transcript reduces the NSE inclusion.

[0216] The method may comprise a step of determining if a disease pathology is caused by an NSE inclusion in a gene transcript prior to treatment.

[0217] The skilled person will understand that optional features of one embodiment or aspect of the invention may be applicable, where appropriate, to other embodiments or aspects of the invention.

EXAMPLES

[0218] Embodiments of the invention will now be described in more detail, by way of example only, with reference to the accompanying figures. These examples are provided for illustrative purposes only and not to limit the scope of the claims provided herein.

[0219] ABBREVIATIONS

NSE	nonsense-mediated RNA decay switch exon in *ATM* intron 28
PE	a 24-nt pseudoexon located 3' of NSE in *ATM* intron 28
NMD	nonsense-mediated RNA decay
A-T	ataxia-telangiectasia
ATM	gene deficient in ataxia-telangiectasia
SSO	splice-switching oligonucleotide
DSB	double-strand DNA break
DDR	DNA damage response
MIR	mammalian-wide interspersed repeat
BPS	branch point sequence
PPT	polypyrimidine tract
IR	ionizing radiation
U2AF	auxiliary factor of U2 small nuclear ribonucleoprotein
U2AF35	a 35-kD subunit of U2AF encoded by *U2AF1*
U2AF65	a 65-kD subunit of U2AF encoded by *U2AF2*
snRNA	small nuclear RNAs

Example 1

SUMMARY

[0220] Phenotypic diversity and susceptibility to genetic disease is influenced by natural intronic variants, but their interactions with RNA-binding proteins are largely unknown. Here a single-nucleotide polymorphism in a detained *ATM* intron was shown to gain functionality in cells lacking the auxiliary factor of U2 small nuclear ribonucloprotein (U2AF). Each U2AF subunit was required for repression of a nonsense-mediated RNA decay switch exon (NSE) in *ATM* intron 28. NSE was

activated to a greater degree in the presence of cytosine than thymine at rs609261 located at position -3 relative to the NSE 3' splice site. The cytosine allele, which is predominant in Caucasians, resulted in a more efficient NSE-mediated inhibition of ATM expression than thymine, the principal allele in Asian populations. NSE activation was deregulated in leukemic cells and was influenced by the amino acid identity at U2AF35 residue 34. Exploiting competition between NSE and a downstream pseudoexon, splice-switching oligonucleotides (SSOs) that repress or activate NSE to modulate ATM expression were identified. Using RNA-Seq, U2AF-regulated exon usage in the ATM signaling pathway was shown to be centered on the MRN/ATM-CHEK2-CDC25-cdc2/cyclin B axis and that U2AF preferentially controls RNA processing of transcripts involved in cancer-associated fusions and chromosomal translocations. These results reveal important links between 3' splice-site control and ATM-dependent response to double strand DNA breaks, illustrate functional plasticity of intronic variants in response to RNA-binding factors, demonstrate versatility of SSOs to modify gene expression by targeting pseudo-splice sites in introns and may explain ethnic differences in cancer risk and survival.

INTRODUCTION

[0221] Here, U2AF was shown to repress a nonsense-mediated decay (NMD) switch exon (NSE) in the *ATM* gene (ataxia-telangiectasia, A-T, mutated) and other proteins involved in 3'ss recognition that regulate NSE inclusion in mature transcripts were identified. The extent to which this event limits ATM expression depends on a common C/T variant rs609261 located in the NSE 3'ss consensus deep in intron 28. Also identified are intronic *cis*-elements that control NSE inclusion in mature transcripts and splice-switching oligonucleotides (SSOs) that modulate NSE activation by targeting a competing pseudoexon in the same intron. Using RNA-Seq, it was next shown that the U2AF-mediatedregulation of DNA damage response (DDR) pathway is centered on the ATM-CHEK2-CDC25-cdc2/cyclin B axis, suggesting that it has coevolved with cellular responses to double-strand DNA breaks (DSBs). Finally, a preferential involvement of U2AF-regulated transcripts is demonstrated in cancer-associated gene fusions and chromosome translocations.

RESULTS

Identification of a U2AF-repressed cryptic exon in *ATM*

[0222] It has been recently shown that depletion of each U2AF subunit resulted in down- and up-regulation of a large number of exons that were predominantly alternatively spliced. When inspecting global RNA processing changes in cells depleted of U2AF35, an unexpectedly strong activation of a cryptic, 29-nt *ATM* exon that was not annotated by RefSeq (termed NSE, Fig. 1A) was found. The NSE activation was observed also in cells individually depleted of each U2AF35

isoform with isoform-specific small interfering RNAs (siRNAs) and with SSOs targeting 3'ss of alternatively spliced *U2AF1* exons Ab and 3, which encode isoform U2AF35b and U2AF35a, respectively (Fig. 1A). Validation of RNA-Seq data using RT-PCR showed that NSE was present in ~10-20% of polyadenylated transcripts in untreated HEK293 cells, similar to levels observed in lymphoblastoid cell lines. The NSE inclusion levels increased to ~75% in cultures depleted of ~90% U2AF35 and to ~50% in cells depleted of ~75% U2AF65 (Fig. 1B), were siRNA dose-dependent and inversely correlated with the amount of available U2AF heterodimers (Fig. 1C), consistent with the requirement of each U2AF subunit for NSE repression. Inspection of RNA-Seq data revealed retention of intronic sequences surrounding NSE (Fig. 1A), suggesting that intron 28 is 'detained' and could be spliced post-transcriptionally. Retention levels of intron 28 were affected neither by SSO- nor siRNA-mediated depletion of U2AF35 (Fig. 1A) and no other cryptic exon in this gene was activated to the same extent as NSE. Thus, NSE plays a key role in the exon-centric regulation of *ATM* expression by U2AF.

NSE activation and ATM expression is modified by rs609261

[0223] Examination of genomic sequences surrounding NSE revealed that position -3 relative to the NSE 3'ss is polymorphic (rs609261, Fig. 2A) in which thymine (T) is predominant in African and Asian populations and cytosine (C) in Caucasians (Fig. 2A). The base identity at this position is important for universal exon recognition, with a CAG>TAG>AAG>GAG hierarchy of exon inclusion levels both at authentic and U2AF35-dependent 3'ss. To confirm that the NSE usage is allele-specific, splicing of two reporter constructs that contained C or T at this position was examined following transient transfections into human embryonic kidney (HEK) 293 cells (Fig. 2B). The T construct yielded lower NSE inclusion than the C reporter, both in untreated cells and cells individually depleted of each U2AF subunit (Fig. 2C).

[0224] To test whether the allele-specific NSE usage results in differential protein expression in cells lacking U2AF35, DNA was first sequenced from available cell lines across rs609261 to obtain transfectable cells homozygous for each allele. HEK293 cells were found to be homozygous for the C allele and HeLa cells were homozygous for the T allele (Fig. 2D). Immunoblots from the U2AF35-depleted cells and untreated controls confirmed efficient depletion in each cell line and a greater U2AF-mediated decline of ATM expression in the presence of the C allele than the T allele (Fig. 2E,F). Depletion of UPF1, a key component of the NMD pathway, revealed a dose-dependent increase of NSE inclusion in *ATM* mature RNAs (Fig. 2G). No signal from a putative truncated ATM was detected on immunoblots from depleted cells.

[0225] Because U2AF-repressed and -activated exons show preferential responses to U2AF-related proteins, HEK293 cells were depleted of PUF60 and CAPERa, and several heterogeneous nuclear RNPs, including hnRNP A1. PUF60 interacts with uridine-rich motifs at 3'ss and hnRNP A1 forms a ternary complex with the U2AF heterodimer on AG-containing U-rich RNAs. Depletion of either PUF60 or hnRNP A1 increased NSE inclusion (Fig. 2H) while PUF60 overexpression led to NSE skipping (Fig. 2I). Thus, the rs609261- and population-dependent NSE activation deep in *ATM* intron 28 is regulated by U2AF, PUF60 and hnRNP A1, demonstrating how functionality of a common intronic polymorphism varies with cellular levels of RNA-binding proteins that facilitate 3'ss recognition.

NSE inhibition by SSOs promotes ATM expression

[0226] To test if NSE activation in cells lacking U2AF can be repressed to restore ATM expression, the C-allele reporter construct was individually cotransfected with SSOs targeting each NSE splice site (Fig. 1A). SSOs were modified at each phosphorothioate linkage and 2'-*O*-methyl ribose and were designed to avoid the PPT of NSE, stable Mfold-predicted stems and rs609261. Each SSO diminished NSE inclusion in a dose-dependent manner both in exogenous (Fig. 3A) and endogenous (Fig. 3B) transcripts and the SSO targeting the NSE 3'ss was more efficient than the SSO bridging its 5'ss at the same concentrations.

[0227] Whether the NSE 3'ss SSO can increase ATM protein expression and activation in cells exposed to ionizing radiation (IR) was next examined. The low ATM expression in cells lacking U2AF35 was partially rescued by this SSO, both in unexposed and IR-exposed cells (lanes 1 vs 2 and 5 vs 6, Fig. 3C, lanes 5-8 vs 9-12, Fig. 8A) and the increase was dose-dependent (Fig. 4D). Following IR, activated ATM autophosphorylated at S1981 showed reduced signal in depleted cells as compared to untreated cells (lane 6 vs 8, Fig. 3C, and lanes 1-4 vs 5-8, Fig. 8A). Exposure to the NSE 3'ss SSO slightly increased also activated ATM (lanes 5-8 vs 9-12, Fig. 8A, lane 5 vs 6, Fig. 3C). To begin to explore putative effects of SSO-mediated NSE repression on ATM signaling, wild type CHEK2 was also overexpressed in (mock)irradiated cells (mock)depleted of U2AF (Fig. 8A). CHEK2 is a serine/threonine kinase phosphorylated by ATM at T68 in response to DNA double-strand breaks (DSBs). However, cells lacking U2AF had markedly lower levels of endogenous CHEK2 compared to controls, which did not appear to be altered by the NSE 3'ss SSO (lanes 1-4 vs 5-8 vs 9-12) whereas exogenous CHEK2 was increased in depleted cells both in IR-exposed and -unexposed cells (lanes 1-4 vs 5-8, see also Fig. 5 and Fig. 8B,C further below).

[0228] Taken together, NSE activation was efficiently inhibited by SSOs that block access to NSE splice sites and do not support RNase H cleavage. The more efficient SSO partially rescued the NSE-mediated inhibition of ATM.

Activation of a NMD switch exon is influenced by a downstream pseudoexon

[0229] To identify intronic regulatory *cis*-elements that control NSE inclusion in mature transcripts, a previously reported A-T mutation IVS28-159A>G was utilized. This mutation was observed to activate the NSE 3'ss while repressing its 5'ss and promoting a downstream 5'ss instead, introducing a 112-nt cryptic exon in the mRNA. There is a strong 3'ss consensus preceded by optimal BPS/PPT motifs observed within this exon, which may bind U2AF and activate a smaller, 24-nt pseudoexon (termed PE; Fig. 4A). Examination of published RNA crosslinking/immunoprecipitation data in *ATM* showed U2AF65 binding upstream and downstream of NSE and upstream of PE, suggesting that NSE activation may be controlled by competition between partially productive spliceosomes assembled at the PE 3'ss and the NSE 3'ss. The two 3'ss are conserved in mammals but are separated by a distance smaller than the minimal size of human introns, sterically preventing simultaneous recognition of NSE and PE (Fig. 4A). In agreement with this hypothesis, deletion of the PE PPT/3'ss introduced in the C minigene, which should alleviate NSE repression through diminished U2AF binding to PE, increased NSE inclusion (Fig. 4B). This deletion also brought about retention of the intron that separates NSE and PE, mimicking the splicing pattern of the A-T mutation IVS28-159A>G. Increasing the intron length from 59 to 79 nt, thereby overcoming a steric hindrance imposed by the insufficient distance between the two pseudo-3'ss, also improved NSE inclusion and diminished the intron retention (Fig. 4B).

[0230] To test if NSE inactivation can influence PE inclusion in mRNA, the NSE 3'ss was first eliminated. This mutation activated a cryptic 3'ss 7-nt downstream of the authentic NSE 3'ss (lanes 1, 2 and 6, 7, Fig. 4C, Fig. 21). This cryptic 3'ss showed a diminished requirement for U2AF. Because extending the intron length between NSE and PE on this background failed to activate PE (Fig. 4C, lanes 3 and 8) and PE lacks exonic splicing enhancers and has a suboptimal BPS (Fig. 22), a 24-nt stem loop derived from a mammalian-wide interspersed repeat (MIR) was inserted in the middle of PE. This MIR hairpin acts as a nearly universal exon definition module through an exposed splicing enhancer in a terminal RNA triloop. The enlarged PE was strongly activated in mock-depleted cells, but was outcompeted by NSE in cells lacking U2AF35 (lanes 4 and 9), indicating that NSE inclusion is more dependent on U2AF35 than PE. The construct containing both the MIR insertion in PE and the extended intron finally generated mRNAs containing both NSE and PE (lanes 5 and 10).

Intronic SSOs targeting competing pseudoexons to modulate gene expression

[0231] Next, the MIR reporter was employed to test the impact of NSE and PE SSOs on exon usage and ATM expression. Fig. 4D shows that the NSE 3'ss SSO repressed transcripts containing NSE and upregulated those with PE whereas the opposite effect was found for SSOs targeting the MIR enhancer loop in PE. The same pattern was observed for the reporter in which NSE and PE were separated by a distance insufficient for their simultaneous inclusion in mRNA (Fig. 4E). These results suggested that SSOs targeting PE and/or U2AF65 binding sites upstream of PE may potentially promote NSE inclusion and reduce ATM expression while the NSE SSOs should have the opposite effect. This approach would provide a broad strategy to modulate gene expression in either direction by antisense-based targeting of competing pseudoexons, one of which is critical for gene regulation. To test this concept, SSOs targeting PE 3'ss and 5'ss were examined. Although each PE SSO induced NSE skipping, both on exogenous and endogenous transcripts (Fig. 4F), SSOs targeting U2AF65 binding sites just upstream of PE (Fig. 4A), i.e. the NSE-repressing sequence (construct delPPT/AG, Fig. 4B), reduced PE inclusion and slightly increased NSE in the MIR reporter (Fig. 4G). In contrast, a longer oligo extended in the 5' direction (SSO-PEBP, Fig. 20) did not show any effect.

[0232] PE contains a natural DNA variant rs4988000 (G/A), which may also influence NSE recognition (Fig. 4H). Transfections of C and T minigenes systematically mutated at rs4988000 revealed that the rare A allele decreased NSE inclusion on each pre-mRNA, both in U2AF35- and mock-depleted cells. Thus, the highest NSE inclusion was produced by the haplotype that is most frequent in Caucasians (CG), followed by haplotypes CA>TG>TA.

[0233] Taken together, the haplotype-dependent activation of the U2AF-repressed NSE can be modified by SSOs that target U2AF65 intronic binding sites upstream of competing pseudo-3'ss, potentially providing a general method to manipulate exon-centric gene expression by antisense-based targeting of NMD switch exons and their regulatory motifs in introns.

Regulation of ATM signaling by U2AF: DSBs at the focal point

[0234] Because ATM is a key apical kinase in the DDR pathway and NMD switch exons often regulate genes encoding protein interaction partners, U2AF35-induced RNA processing changes of currently known ATM substrates and other constituents of the ATM signaling network were systematically characterized. Interestingly, although genes involved in the DDR and cell cycle control that contained U2AF35-dependent exons were only marginally enriched (FDR=0.08), each component in the ATM-CHEK2-CDC25-CDC2/cyclin B axis showed RNA processing alterations (Fig. 5A, Fig. 9). This pathway is critical for ATM signaling of DSBs.

[0235] First, reduced ATM expression in cells lacking U2AF (Fig. 8) was associated with decreased CHEK2 mRNA, increased retention of *CHEK2* intron 10, and skipping of exons 9 and 11 (Fig. 5A). RNA processing alterations of known *CHEK2* substrates were limited to genes regulating the cell cycle (*CDC25A*, *CDC25B*, *CDC25C* and *TTK*; Fig. 5A, S3A-B, 11A) and were not apparent in genes involved in DNA repair (*BRCA1*/2, *XRCC1*, *FOXM1*, *TRIM28*) or p53 signaling (*TP53*, *MDM4*, *CABIN1*, *STRAP*, *AATF*). *CHEK2* exon 9 skipping, which would be predicted to activate NMD, was only marginally increased 24 hrs after IR and did not contribute to the decline of total CHEK2 observed as early as 30 min after IR (Figs. 5B and 5C). As *CHEK2* exon 9 inclusion was increased only for the highest concentration of *UPF1* siRNAs (Fig. 5D), HEK293 cells were transcfected with an SSO targeting its 3'ss. This treatment induced exon 9 skipping and reduced expression of the CHEK2 protein, however, it also increased NSE activation (Fig. 5E). In contrast, SSOs targeting NSE or PE did not have any effect on *CHEK2* exon 9 inclusion (Fig. 5F). Exon 9 skipping, but not NSE, was also dramatically increased in cells lacking SF3B1 (Fig. 5G). To address why exogenous expression of CHEK2 was increased in cells lacking U2AF35 as compared to controls (Fig. 8A), HEK293 cells were cotransfected with the CHEK2-repressing SSO and a CHEK2-expressing plasmid (Figs. 8B, and 8C). Reduced endogenous CHEK2 was associated with a significant increase of exogenous CHEK2 also in U2AF-proficient cells, pointing to a tight homeostatic regulation of the total CHEK2 protein in the cell.

[0236] Second, U2AF was required for full activation of *CDC25A* exon 6 (Fig. 5A), which encodes a residue (S178) that is phosphorylated by CHEK2 and CHEK1, facilitating binding of 14-3-3. U2AF35 was also required for inclusion of exon 3 of *CDC25B* and *CDC25C* (Figs. 10A and 10B), confirming previous microarray data. *CDC25B* exon 3 encodes multiple phosphorylated residues, including a B-domain residue S169, phosphorylated by MAPKAP kinase 2 and pEg3. This isoform localizes to the centrosomes and accumulates during mitosis. *CDC25C* exon 3 encodes T67 phosphorylated by cdc2/cyclin B as a part of the auto amplification loop. Phosphorylated T67 inCDC25C creates a binding site recognized by the WW domain of PIN1, which sustained activation of a U2AF-repressed NMD switch exon (Fig. 11B), possibly modifying catalytic activity of this abundant peptidyl-prolyl isomerase. Finally, cyclin B1 and B2 mRNAs were upregulated in cells lacking U2AF35 as well as cyclin B1-interacting protein (CCNB1IP1, also known as HEI10), although their RNA processing pattern did not appear to be altered (Fig. 5A). Cyclin B upregulation was associated with a detained *CDK1* intron (Fig. 11C), which may be spliced post-transcriptionally.

[0237] ATM recruitment to DSB is facilitated by the MRN complex, consisting of MRE11, RAD50 and NBN. *NBN* showed no obvious RNA processing changes in cells lacking U2AF35, but *RAD50*

mRNA was downregulated, possibly through activation of a NMD switch exon and/or additional splicing alterations (Fig. 12A-C and Fig. 9). The last *MRE11A* exon was upregulated as a result of a promotion of distal alternative polyadenylation site in depleted cells, which is present in most cell types, but not in B cells. DEXSeq analysis did not detect significant RNA processing changes in transcripts encoding other members of the phosphatidylinositol 3 kinase-like family of serine/threonine protein kinases (*ATR* and *PRKDC*), nor in *BRCA1/2*, *RNF168* and the ATM interactor *ATMIN*. Additional ATM interacting partners with altered exon or gene expression included *RPS6*, *SRSF1* and other SR proteins, *EP300*, *RPA2*, *BLM*, *FANCD2* and *FANCI*, *PPP2R5C* and *PPP2R5D*, and *SMC3*, a central component of the cohesin complex (Fig. 9).

[0238] Depletion of U2AF35 was associated with preferential alterations of genes/exons involved in chromatin modification, which have numerous functional links to ATM signaling (Fig. 9). For example, the INO80 chromatin remodeling complex is phosphorylated by ATM and is functionally linked to checkpoint regulators, including CHEK2. U2AF inhibited INO80C isoforms containing 54-nt exons that encode peptides that are absent in the yeast *Ies6* homolog, which is critical to INO80 function in vivo and is likely to alter heterodimer formation with ACTR5 and nucleosome binding. Expression of multiple components of the INO80 complex was altered in depleted cells, including ACTR5, ACTL6A and RUVL2B. Many INO80 subunits localize preferentially in telomeres and their mutations result in telomere elongation. U2AF is required for full inclusion of *TERF1* exon 7 in mRNA (Fig. 13A), regulating the abundance of TRF1 (exon 7+)/PIN2 (exon 7-) isoforms, important components of the shelterin complex. Exon 7 encodes multiple phosphorylated serine residues and both isoforms can heterodimerize through the dimerization domain. TRF1 binding to telomeres is promoted by ATM inhibition whereas ATM-mediated phosphorylation impairs TRF1 interaction with telomeric DNA. TRF1 association with telomeres is also negatively regulated by RAD50. TRF1-interacting TIF2 is another shelterin protein localized in nuclear matrix and encoded by *TINF2*. TIF2 exists in at least two isoforms produced by alternative splicing, termed TIN2S and TIN2*L*. TIN2*L* contains an extra NM binding domain and associates more strongly with the nuclear matrix than TIF2*S*, which is encoded by a transcript with retained 3' introns that form a long 3' untranslated region. This mRNA isoform was repressed by U2AF (Fig. 13B).

[0239] Collectively, these results show that the MRN/ATM-CHEK2-CDC25-cdc2/cyclin B axis is at the center of the U2AF35-mediated control of DDR, although the U2AF regulation extends into additional ATM substrates involved in chromatin modification and telomere length control.

U2AF preferentially controls RNA processing of transcripts involved in leukemia-associated fusions

[0240] CHEK2 phosphorylates PML (Promyelocytic Leukemia) and appears to require PML for subsequent autophosphorylation. Depletion of U2AF35 promoted the use of proximal alternative polyadenylation site of PML, leading to the upregulation of the shortest *PML* isoform, which lacks the last exon coding for the nuclear export signal (Fig. 14A). The long and short PML isoforms have distinct functions; for example, nuclear PML isoforms, but not the cytoplasmic isoform, are positive regulators of IFNγ signaling. The C-terminus of the longest PML isoform specifically interacts with AML1 to enhance AML1-mediated transcription, suggesting that U2AF deficiency could impair PML-AML1 interactions. PML also binds PIN1 and this interaction promotes PML degradation in a phosphorylation-dependent manner. U2AF depletion increased a *PIN1* NMD exon (Fig. 11 B), potentially limiting expression of this highly abundant peptidyl-prolyl isomerase, which interacts with many phosphoproteins to regulate mitosis, including phosphorylated CDC25.

[0241] Apart from PML, U2AF35 depletion upregulated other RARA partners, including *NPM1* (Fig. 14B). This event was associated with promotion of a proximal polyadenylation site, thus increasing the abundance of shorter, presumably more stable transcripts. An alternatively spliced exon of *BCOR*, a BCL6 corepressor that forms BCOR-RARA fusions and interacts with several histone deacetylases to increase BCL6 transcriptional repression, was also downregulated (Fig. 14C).

[0242] Interestingly, the overlap between U2AF35-sensitive genes/exons and 1,187 genes involved in cancer-associated gene fusions and 300 genes involved in recurrent chromosome translocations was greater than expected, with more significant P values observed for genes with differentially used exons than those implicated by Cufflinks at the transcript level (Table 1). Similarly, sharing of genes frequently mutated in the myelodysplastic syndrome and genes differentially expressed upon U2AF35 depletion was significantly higher than expected ($P<0.01$, hypergeometric test). Thus, RNA processing of transcripts involved in cancer-associated gene fusions and chromosome translocations is preferentially regulated by U2AF.

[0243] To test the function of cancer-associated *U2AF1* mutations in NSE splicing, reconstitution experiments were performed with wild-type and mutated U2AF35 constructs that were cotransfected with the C minigene into cells (mock)-depleted of U2AF35 (Fig. 6). NSE activation was repressed by either U2AF35 isoform to a similar extent, as well as U2AF35a containing substitutions of S34 in the zinc finger 1 domain, the most frequently mutated U2AF35 residue in cancer. In contrast, only a partial rescue was achieved by substitutions of Q157 in the second zinc finger domain where these mutations are less frequent. Other S34 mutations failed to fully reconstitute the defect, including S34T and substitutions with small amino acids, although a large residue at this position (S34R) was

efficient. Thus, the identity of the residue at position 34 of U2AF35 is important for NSE recognition.

[0244] Finally, a low degree of NSE activation was detected in diverse human tissues, both in hexamer-primed samples and polyadenylated transcripts (Fig. 15A). The proportion of NSE-containing RNAs was on average higher in leukemic cells than in normal cells, with some samples exhibiting very high levels not observed in normal tissues (Figs. 15B and 15C), potentially contributing to reduced ATM expression previously observed both in leukemias and solid tumors. NSE inclusion was also examined in polyadenylated RNAs extracted from a panel of lymphoblastoid cell lines exposed to cold and heat shock at the indicated temperatures prior to lysis (Figs. 15D and 15E). Interestingly, NSE appeared to be activated to a minor extent by exposing cells to 42 °C but not at subphysiological temperatures (Fig. 15D), suggesting that markedly higher NSE inclusion levels in malignant cells are unlikely to be explained by a cold shock encountered during storage of patients' samples. Since proteomic profiling of Jurkat cells exposed to a heat stress at 43 °C revealed diminished expression of several proteins including U2AF35a, these results further support U2AF35 as a specific NSE repressor.

DISCUSSION

[0245] The work described herein significantly expands currently known links between RNA processing and DDR pathways (Figs. 5 and 9). An alternative splicing-coupled NMD switch exon critical for ATM expression was identified(Figs. 1 and 3) and its importance in cancer risk was examined (Fig. 2, Figs. 6 and 15). How intronic haplotypes influence inclusion of this exon in mature transcripts and their functional dependence on cellular levels of RNA-binding proteins involved in 3'ss selection was also shown(Figs. 2 and 4H). Finally, SSOs were identified that modulate activation of this exon by targeting its regulatory sequences and propose a novel antisense strategy to modify gene expression.

[0246] U2AF is an important 3'ss recognition complex and a critical regulator of alternative splicing. In addition to expanding protein-protein interactions, alternative splicing has evolved to fine-tune quantitative gene expression through NMD, in agreement with alterations of many NMD exons in cells lacking this factor (Figs. 1, 5 and 13). Peptides encoded by alternatively spliced exons are enriched in disordered regions and post-translation modification (PTM) sites, which are required for dynamic and reversible switching between two or more isoforms. Conversely, PTMs regulate numerous splicing factors, including proteins involved in NSE regulation. This complexity represents a clear challenge ahead and can be exemplified by the observed NSE activation upon targeting of *CHEK2* exon 9 (Fig. 5E). Reduced CHEK2 expression may alter interactions with other

kinases such as CDK11, which is constitutively phosphorylated at S737 in a CHEK2-dependent manner and interacts with U2AF65 and PUF60 , creating a regulatory loop that controls NSE levels (Fig. 2H,I).

[0247] These results suggest that U2AF is an integral part of the DDR control, contributing to fine-tuning of its PTM network and subject to PTMs itself. U2AF35 was found among proteins that showed increased phosphorylation at S59 upon DNA damage. This serine residue is present only in U2AF35a and is replaced by alanine in U2AF35b. Exogenous expression of U2AF35b was higher than U2AF35a and the relative abundance of U2AF35b increased upon depletion of U2AF65, suggesting that the two U2AF35 isoforms may differentially interact with U2AF65 and may not have equivalent roles in DDR. However, U2AF35- and U2AF65-regulated exons vastly overlap and most, but not all, RNA processing changes found in U2AF35 depleted cells are attributable to the lack of the U2AF heterodimer, including the NSE activation (Fig. 1C).

[0248] U2AF-repressed exons have a distinct 3'ss organization and response to U2AF-related proteins as compared to U2AF -activated exons, suggesting that the exon repression involves direct RNA binding. This is supported by the observed NSE activation on exogenous transcripts that do not undergo NMD and by the SSO-induced NSE blockage (Figs. 2 and 4). NSE lacks AG dinucleotides between the predicted BPS and 3'ss, its AG exclusion zone is longer than the average and has an unusual stretch of 5 conserved guanines upstream of the BPS, which may contribute to stable secondary structures across 3'ss that might be required for the repression. The adenine-rich 3' portions of both NSE and PE are more conserved in evolution than their 5' parts (Fig. 4A), potentially providing important ligand interactions, given the propensity of adenine to occupy unpaired positions in structured RNAs. Interestingly, primate NSEs have uridine at position -3 and longer PPT than lower mammals, which have cytosine at this position. Although direct RNA binding appears to be the simplest explanation for exon repression by U2AF, U2AF35 depletion led to downregulation of several proteins involved in NMD (Table S4), which may contribute to NSE activation on endogenous transcripts. In addition, physical interactions between U2AF65 and the C-terminus of TRF1 or other components of the ATM signaling network may also participate in NSE regulation.

[0249] Apart from *U2AF1/U2AF2*, additional genes involved in 3'ss selection have been found mutated in cancer. Interestingly, chronic lymphocytic leukemias with *SF3B1* mutation were associated with a cryptic 3'ss activation of *ATM* exon 46, leading to ATM truncation. Recently, splicing of an *EZH2* exon as a result of cancer-associated *SRSF2* mutation was implicated in impaired hematopoietic differentiation and the same NMD exon was upregulated also upon U2AF35

depletion (Fig. 12D). Whether these exons are targets of a common 3'ss recognition pathway underlying leukemogenesis remains to be established. In contrast, NSE inclusion did not appear altered in cells depleted of SF3B1, which produced almost complete skipping of *CHEK2* exon 9 (Fig. 5G).

[0250] Because NSE activation may restrict ATM expression both in normal and cancer cells (Figs. 1, 2, and 15) and ATM is a limiting factor in the DDR pathway, cytosine at rs609261 may confer a relative ATM deficiency not only in (pre-)malignant cells but also in the germline. ATM kinase activity appears to be a good predictor of A-T severity, however, the diversity of A-T alleles does not fully account for the spectrum of clinical symptoms. Genes involved in NSE activation (Fig. 1, 2) might contribute to clinical heterogeneity of A-T patients, particularly those with 'leaky' mutations. Natural variants modifying NSE inclusion (Figs. 2C-F and 4H) may also contribute to the phenotypic complexity of A-T or even A-T heterozygotes that lack overt clinical features but may display increased radiosensitivity and cancer risk, consistent with the central focus of U2AF-regulated exon usage within the ATM signaling network (Fig. 9).

[0251] These results predict that NSE activation is on average more efficient in Caucasians than in Asian populations as a result of a higher frequency of the C allele at rs609261 in the former (Fig. 2A). Asian Americans have lower mortality rates for common malignancies than Caucasians that persist over a long-period of time. The risk of hematopoietic malignancies also varies greatly by ethnic group and their incidence is the highest in white populations, including non-Hodgkin and Hodgkin lymphomas, which are associated with A-T. This trend also persists in migrants and continues in subsequent generations. Although lymphoblastoid leukemias, lymphomas and other cancer types show distinct incidence rates across Asian and Caucasian populations, no significant ethnic differences in the age-standardized incidence rates were found for myeloid leukemias, which does not appear to be more prevalent in A-T, unlike lymphoid malignancies or other cancers. Asian cancer patients respond more favorably than Caucasian patients to cytotoxic therapy and have on average a longer median survival. Asian cancer patients were also reported to have a lower prevalence of some gene fusions than Caucasians, potentially reflecting their capacity to respond to DSBs. rs609261 showed the lowest p-value of *ATM* variants in Cochrane-Armitage tests of association with glioma. rs2235006 (ATM allele F582L), which is located only ~35 kb upstream of rs609261 in a region of minimal recombination, was associated with a high risk (OR 11.2) of chronic lymphocytic leukemia. This study genotyped 1467 coding nonsynonymous SNPs in 865 candidate genes and implicated variants in genes encoding the ATM-BRCA2-CHEK2 DDR axis as the most significant risk pathway. Allelic association studies of nonagenarians/centenarians and younger

controls also suggested association between *ATM* and longevity. Finally, ethnic differences were noted also for mutation rates in genes frequently altered in hematological malignancies; for example, *SF3B1* mutations in chronic lymphocytic leukemias were less frequent in Chinese than in European populations.

[0252] Although these considerations collectively support the importance of rs609261-dependent NSE activation in cancer risk and survival, the U2AF- and hnRNP A1-dependency of NSE inclusion (Fig. 2H, S8B) demonstrates that it is by no means fixed. Variable expression patterns of these proteins from one malignancy to another would imply a 'capricious functionality' of this variant. Many more polymorphic sites with this attribute are likely to be established in future, contributing not only to the inter-individual variability of gene expression through restrictive capacity of 'poison' cryptic exons, but potentially also to the 'missing heritability' of complex traits and failures of genome-wide association studies, particularly in cancer.

[0253] Although RNA-Seq is a powerful tool to examine global transcriptome in response to DNA damage, rigorous standards that correctly estimate biological and statistical significance of the observed alterations in RNA processing are yet to be implemented. Given a high stringency of the DEXSeq algorithm, the existence of additional biologically important RNA processing events responsive to U2AF cannot be excluded. For example, upregulation of a proximal polyadenylation site in *CHEK1*, which was coupled with upregulation of 24-nt and 27-nt exons in *CLASP1*, would implicate the ATM apoptotic pathway. These events were not detected by DEXSeq but were see genomic browsers and require confirmation. The apoptotic pathways are of particular interest in the myelodysplastic syndrome which shows susceptibility of myeloid progenitors to the programmed cell death and where deregulation of genes involved in ATM signaling was found in more advanced but not initial clinical stages. Interestingly, *U2AF1* mutations were also found to be more frequent in advanced stages and were associated with shorter survival. This study also highlights current limitations of incomplete transcript annotation and the importance of examining cryptic exons in RNA-Seq data. Future RNA-Seq studies should therefore attempt at global detection of NMD events associated with alternative splicing, which has been hindered by the instability of stop codon-containing transcripts.

[0254] Finally, this study demonstrates efficient repression of a key NMD switch exon in *ATM* by SSOs that also increased ATM protein levels (Fig. 3A-D, Fig. 8). It also reveals competing regulatory motifs of NSE in the same intron (Fig. 4A-C, H) that could be exploited as a target for SSO-mediated modulation of gene expression (Fig. 4D-G). This approach can be combined with genome-editing such as CRISPR-Cas9 to delete or introduce splicing regulatory motifs or protein

binding sites implicated by minigene studies (Fig. 4C) and should also help us to find efficient intronic SSOs with desired outcomes for RNA processing. The search for such SSOs is more challenging than for those targeting human exons. For example, most SSOs systematically covering *SMN2* exon 7 stimulated exon skipping, an event exploited for treatment of spinal muscular atrophy, however, ~20% induced exon inclusion. By analogy, the desired stimulation of intron splicing was found only for 10% of SSOs targeting *INS* intron 1 while the majority failed to show this effect. The proposed strategy takes advantage of a much higher information content of human auxiliary splicing sequences as compared to lower organisms and should be greatly facilitated by future global pre-mRNA folding studies. For example, unlike the SSO that efficiently blocked the NSE 3'ss (SSO-NSE3, Fig. 3A,B), a partially overlapping morpholino extending only 7-nt into NSE failed to repress the same 3'ss to rescue splicing of mutation IVS28-159A>G, despite targeting U2AF binding sites (Fig. 4A). This suggests that the morpholino oligo may have blocked access to structures or motifs that are not responsible for exon activation, but exon repression, in agreement with these finding (Fig. 1A-C). Administration of antisense-based RNA processing activators or inhibitors that target or avoid binding sites of splicing factors in introns could be exploited therapeutically to shape beneficial or detrimental consequences of NMD in cancer cells. This approach is supported by a broad recognition that NMD serves primarily a regulatory function across a wide range of transcripts and may also promote translation of NMD substrates that produce truncated polypeptides, which may stimulate anti-tumor immunity.

MATERIAL AND METHODS

Plasmid constructs

[0255] *ATM* minigenes were prepared by cloning ~0.9-kb amplicons into *XhoI*/*XbaI* sites of the *U2AF1* construct. Cloning primers are shown in Table S1. Full inserts were sequenced to confirm the identity of intended changes and exclude undesired mutations. PUF60 expression vectors were also used. The hnRNP A1 construct was a generous gift of Gideon Dreyfuss (University of Pennsylvania).

Cell cultures and transfections

[0256] Cell cultures were maintained in standard conditions in DMEM supplemented with 10% (v/v) bovine calf serum (Life Technologies). Depletion of U2AF subunits and U2AF35 isoforms with small interfering RNAs (siRNAs) and splice-switching oligonucleotides (SSOs), were carried out following a time course experiment that established depletion levels of each isoform. Oligo(ribo)nucleotides and siRNAs are listed in Table S1. Transfections were carried out in 6- or 12-well plates using jetPRIME (Polyplus) according to manufacturer's recommendations. The cells

were harvested 48 hrs after the second hit, except for those exposed to IR, which received a single hit. For SF3B1 depletion, HEK293 cells were exposed to a siRNAs mixture (S23850, S23852, and S223598 (LifeTechnologies)) and were harvested 48 hrs later.

RNA-Seq

[0257] Analysis of differential exon usage was performed using DEXSeq (v. 1.12.1), based on q-values less than 0.05. Differential gene and isoform expression between sample sets was analyzed with Cufflinks (v. 2.1.1), which normalizes the reads using a fragments per kilobase of exon model per million reads measure. Selection of significantly differentially expressed genes was made on the basis of FDR-adjusted P-values ($q<0.05$).

NSE expression in human tissues and cell lines

[0258] The FirstChoice human total RNA survey panel containing total RNA samples from 19 different tissues was purchased from LifeTechnologies. Each tissue sample contained a pool of RNAs from different donors. Lymphoblastoid cell lines were exposed to cold and heat shock. Total RNA samples were reverse transcribed with the Moloney murine leukemia virus reverse transcriptase (Promega) and random hexamer or oligo-d(T) primers. cDNA samples were amplified using primers shown in Fig. 20. Total RNA extracted from leukocytes from bone marrow samples of randomly selected patients with acute myeloid leukemia or chronic myelomonocytic leukemia was reverse transcribed with random hexamer primers. The study was approved by the National Research Ethics Service (UK) Committee South West.

Splice-switching oligonucleotides

[0259] SSOs were designed to maximize interactions with single-stranded regions and avoid secondary structures predicted by Mfold. All SSOs were purchased from Eurofins, diluted in water and their aliquots were stored at -80 °C. All transfections were carried out with jetPRIME (Polyplus) according to manufacturer's recommendations.

Exposure of cell cultures to ionizing irradiation

[0260] (Mock)-depleted HEK293 cells were exposed to IR 48 hours after the first hit using a Gulmay Medical (X-Strahl) D3225 Orthovoltage X-ray system at a dose-rate of 0.63 Gy/min at room temperature. The actual dose rate was monitored by a constancy meter. Cells were harvested as indicated in figure legends.

Immunoblotting

[0261] Antibodies against ATM (D2E2), ATM-pS1981 (D6H9), CHEK2 (D9C6) and CHEK2pThr68 (C13C1) were purchased from the Cell Signaling Technology, Inc. RBM39 antibodies were purchased from Thermo Fisher Scientific (PA5-31103). Antibodies detecting X-

press tag, U2AF35, U2AF65, and tubulin were used. SF3B1 immunoblotting was performed with mouse monoclonal anti-SAP155 antibody (D138-3, MBL). Preparation of cell lysates and immunoblotting was carried out.

TABLE 1 U2AF35-dependent transcripts are more common than expected among genes involved in cancer-associated gene fusions and recurrent chromosomal translocations

Database	Source	Number of Genes	Overlap with U2AF35-sensitive exons[2]	P-value / representation factor[3]	Overlap with U2AF35-sensitive transcripts[2]	P-value / representation factor[3]
ChimerDB 2.0	[69]	1187	66	P<0.00004/1.7	204	P<0.02/1.1
Genes involved in recurrent structural abnormalities in cancer	[70]	300	19	P<0.006/1.9	56	P<0.05/1.2

[1]Gene list downloaded on 2 April 2014. [2]Exon- and gene-level analysis of RNA-Seq data was carried out for 23,263 genes using DEX-Seq and Cufflinks, respectively. [3]Number of overlapping genes divided by the expected number of overlapping genes drawn from two independent groups. A representation factor >1 indicates a greater overlap than expected of two independent groups, the value <1 indicates less overlap than expected. P-values were derived by hypergeometric tests.

Example 2 - Antisense macrowalk targeting a regulated nonsense-mediated RNA decay switch exon in the ATM gene

Summary

[0262] ATM is an important cancer susceptibility gene that encodes a critical kinase of the DNA damage response (DDR) pathway. ATM deficiency results in ataxia-telangiectasia (A-T), a rare genetic syndrome exhibiting a high susceptibility to lymphoid malignancies. ATM expression is limited by a nonsense-mediated RNA decay (NMD) switch exon (termed NSE) located in intron 28, which is tightly controlled by the spliceosome. NSE inclusion in mature transcripts can be modulated by splice-switching oligonucleotides (SSOs), but their optimal targets in the intron are unknown and their delivery to lymphoid cells has not been tested. Here a systematic search for efficient SSOs targeting intron 28 to identify NSE activators and inhibitors was employed. Discovery of these antisense compounds was assisted by a segmental deletion analysis of intronic transposed elements, revealing NSE repression upon removal of a distant antisense Alu and NSE

activation upon elimination of a long terminal repeat transposon MER51A. Efficient NSE repression upon SSO delivery with chitosan-based nanoparticles to embryonic and lymphoblastoid cells was also demonstrated, opening a possibility for NSE-mediated modulation of ATM expression in cancer and A-T. Taken together, these results highlight an important role of transposed elements in regulating NMD switch exons and the power of intronic SSOs to modify gene expression.

Introduction

[0263] Eukaryotic genes contain intervening sequences or introns that need to be removed by a large and highly dynamic RNA protein complex termed the spliceosome to ensure accurate protein synthesis. The cell requires excessive energy and time to complete transcription of intron containing precursor messenger RNAs (pre-mRNAs) from at least a quarter of the human genome and also needs to synthesize non-coding RNAs and >200 different spliceosomal proteins to achieve this task. Although once regarded a 'selfish' or 'junk' DNA, introns are now recognized as critical functional components of eukaryotic genes that enhance gene expression, regulate alternative RNA processing, mRNA export and RNA surveillance. They are also an important source of new gene-coding and -regulatory sequences and noncoding RNAs, including microRNAs and circular RNAs. Their removal process is tightly coupled with transcription, mRNA export and translation, with most human introns eliminated from pre-mRNA co-transcriptionally, however, their potential as targets for nucleic acid therapy is only beginning to be unleashed.

[0264] Spliceosomes assemble ad hoc on each intron in an ordered manner, starting with recognition of the 5' splice site (5'ss) by U1 small nuclear RNP or the 3'ss by the U2 pathway. In addition to traditional splice site recognition sequences (5'ss, branch point, polypyrimidine tracts and 3'ss), accurate splicing requires auxiliary sequences or structures that activate or repress splice sites, known as intronic or exonic splicing enhancers or silencers. These elements allow genuine splice sites to be recognized among a vast excess of cryptic or pseudo-sites in eukaryotic genomes that have similar sequences but outnumber authentic sites by an order of magnitude. Activation of cryptic splice sites can introduce premature termination codons (PTCs) in translational reading frames that may lead to genetic disease. Such transcripts are usually recognized by a NMD pathway and downregulated. However, cryptic exons and NMD have also an important role in controlling the expression of naturally occurring transcripts and for differentiation stage-specific splicing switches, as exemplified by terminal stages of hematopoiesis. In addition, cryptic splice sites may permit unproductive or partial spliceosome assemblies that may compete with natural splice sites, facilitating their accurate selection at a single-nucleotide resolution. Cryptic splice sites activating such 'pseudo-exons' (also known as 'poison' or 'NMD switch' exons) that limit gene expression and

regulate the pool of mRNA isoforms could thus provide interesting targets for nucleic acid therapeutics, however, exploitation of such approaches is in its infancy.

[0265] Splice-switching oligonucleotides (SSOs) are antisense reagents that modulate intron splicing by binding splice-site recognition or regulatory sequences and competing with *cis*- and *trans*-acting factors for their targets. They have been shown to restore aberrant RNA processing, modify the relative abundance of existing mRNA isoforms or produce novel splice variants that are not normally expressed by the cell. Most SSOs employed in pre-clinical and clinical development have targeted exonic sequences. Functional intronic SSOs are more difficult to identify, unless SSOs block access to intronic cryptic splice sites activated by a disease-causing mutation. First, a large fraction of intronic sequences may not affect RNA processing, despite the wealth of intronic auxiliary splicing motifs in the human genome. In addition, their identification is costly and inefficient in long introns. Most exonic SSOs designed to induce exon skipping have usually a desired effect. For example, most SSOs systematically covering SMN2 exon 7 stimulated exon skipping, a prerequisite for antisense therapy of spinal muscular atrophy, however, ~20% increased exon inclusion. By contrast, stimulation of intron splicing was found only for ~10% of SSOs targeting INS intron 1 while the majority failed to show this effect. Identification of effective SSOs may be facilitated by global pre-mRNA folding and ultraviolet crosslinking and immunoprecipitation studies that identify binding sites for components of the spliceosome or the exon junction complex. However, these binding sites may not reflect optimal antisense targets and their resolution may not be sufficient. Thus, a search for intronic SSOs with desired effects on RNA processing remains challenging.

[0266] The RNA-Seq studies have recently revealed activation of a NMD switch exon (termed NSE) deep in ATM intron 28 in cells depleted of each subunit of the auxiliary factor of U2 small nuclear RNP (U2AF). U2AF binds to polypyrimidine tracts coupled with highly conserved 3'ss AG dinucleotides at intron ends and this binding promotes U2 recruitment to the branch site and formation of lariat introns. However, the recent identification of a large number of exons that were activated in cells depleted of each U2AF subunit (U2AF35 and U2AF65) and exhibited a distinct 3'ss organization suggested that a subset of both canonical and NMD switch exons is repressed by U2AF, similar to exon-repressing and -activating activities found for a growing number of RNA binding proteins. The NSE levels were responsive to knockdown of additional splicing factors involved in 3'ss recognition and were influenced by two natural DNA variants (rs4988000 and rs609261) located in the NSE itself and its 3'ss, respectively. SSOs that modulate NSE inclusion levels in the ATM mRNA by targeting NSE and its competing pseudoexon in the same intron have

also been identified. The ATM NSE provides an interesting and promising target for anticancer therapy for several reasons: (i) the ATM kinase is activated in response to double-strand breakage, mobilizing an extensive signaling network with a broad range of targets, influencing cellular sensitivity to DNA-damaging agents; (ii) the U2AF-regulated exon usage in the ATM signaling pathway was centered on the MRN/ATM-CHEK2-CDC25 axis and preferentially involved transcripts implicated in cancer-associated gene fusions and chromosomal translocations; and (iii) the ATM NSE activation limits ATM expression in cells lacking each U2AF subunit. However, optimal NSE SSOs are unknown and their delivery to lymphoid cells has not been tested.

[0267] In the present study, SSOs covering the entire intron 28 were systematically screened and additional SSOs that activate or repress NSE and could be exploited as putative NSE-based ATM inhibitors and activators in therapeutic strategies were identified. Distant transposed elements in the same intron that influence NSE inclusion were also identified. Finally, efficient NSE repression upon SSO delivery to embryonic and lymphoblastoid cell lines using chitosan-based nanoparticles was also shown.

Materials and Methods

Plasmid constructs

[0268] Reporter constructs containing full ATM intron 28 and flanking exons were cloned in the HindIII/XbaI site of pCR3.1 (Invitrogen) using amplification primers ATM26 and ATM27 (Table 2). Deletion constructs (Fig. 16) were obtained by overlap extension PCR with mutagenic primers (Table 2). Hybrid ATM minigenes were prepared by cloning ~0.9-kb amplicons containing NSE and exon 29 into XhoI/XbaI sites of the U2AF1 construct. Plasmids were propagated in *E. coli* (DH5α) and plasmid DNA was extracted with the Gene JET Plasmid Miniprep kit (ThermoScientific). Full inserts were sequenced to confirm the identity of intended changes and exclude undesired mutations.

Splice-switching oligonucleotides (SSOs)

[0269] To test SSOs with both endogenous and exogenous pre-mRNAs, SSOs were designed to avoid transposed elements in intron 28. Transposons were confirmed in sequences of the constructs using RepeatMasker. The SSO GC content was at least 24% (mean 31%) and their average length was ~20 nt. The SSOs comprehensively covered three unique regions in ATM intron 28 (termed A, B and AN, Fig. 17), avoiding only homopolymeric tracts. SSOs (Eurofins) were modified at each ribose by 2'-O-methyl and by a phosphorothioate at each end linkage to ensure adequate stability for the ex vivo screening. SSOs were diluted in double distilled water and quantified using Nanodrop (ThermoScientific). Their normalized aliquots were stored at -80 °C.

Determination of PU values

[0270] The PU (probability of unpaired) values estimate RNA single-strandedness using the equilibrium partition function by considering all possible RNA structures of short sequences, permitting their comparison at each nucleotide position. Higher PU values indicate a higher single-strandedness of an RNA motif. The PU values were computed as described using the three intronic regions and their 30-nt flanks as an input. PU values for each position of an SSO target were averaged and correlated with SSO-induced NSE inclusion levels.

Preparation of stearylated trimethyl chitosan

[0271] Trimethyl chitosan, originally derived from ultrapure chitosan obtained from *Agaricus bisporus*, was provided by KitoZyme (Belgium).

[0272] Purified products had the number average molecular weight (Mn) of 43.3±5.5 kDa and the polydispersity index (Mw/Mn) of 2.4±0.3, as determined by gel permeation chromatography in a 0.33 M $NaCH_3COOH$/0.28 M CH_3COOH eluent at a flow rate of 1 mL/min. The degrees of acetylation and quaternization, determined by the Fourier-transform infrared spectroscopy and 1H-nuclear magnetic resonance spectroscopy (^1H NMR), respectively, were 11.1±0.9% and 30.1±4.6%. Trimethyl chitosan was functionalized with N-succinimidyl stearate (Santa Cruz BioTechnologies), achieving a final degree of substitution of 2.1±0.6% (mol %), as determined by 1H NMR.

Formation of nanocomplexes

[0273] The nanocomplexes were prepared by mixing equal volumes (30 μL) of SSO and polymer solutions. Briefly, SSOs were diluted in buffer A (20 mM HEPES, pH 7.3, 5% (w/v) glucose) and supplemented with 1 M Na_2SO_4 to a final concentration of 50 mM. Both the polymer and SSO solutions were heated at 60 °C for 5 min before mixing with vortex at 1,000 rpm for 15 s. The tested complexes were prepared with molar ratios of quaternized amines (N) to phosphate groups (P) of 20, 40 and 80, as previously optimized, and had a hydrodynamic diameter between 110-130 nm for N/P ratios between 20-80. The complexes were allowed to stabilize for 30 min at room temperature before adding to a 240 μL of the culture medium (DMEM) without serum and antibiotics. Final concentration of SSOs in chitosan-containing cultures was 300 nM. Twenty four hours after transfections, 300 μL of the culture medium with serum/antibiotics was added. The cells were harvested 24 hrs later.

[0274] Cell cultures and transfections. HEK293 and lymphoblastoid VAVY cells were maintained in standard culture conditions in DMEM supplemented with 10% (v/v) bovine calf serum. Cells were seeded at 70% confluency 24 hrs prior to transfections. Transfections of wild-type and deletion constructs were carried out in 12- or 24-well plates using jetPRIME (Polyplus) according to manufacturer's recommendations. The cells were harvested 24 hrs later for total RNA extraction.

Each SSO was transfected with or without the full-length ATM construct at 50 nM and cells were harvested 48 hours later for RNA extraction.

[0275] Analysis of spliced products. RNA samples were isolated using TRI-reagent (Ambion). Total RNA samples from chitosan experiments were extracted with the RNeasy kit (Qiagen). RNA was quantified and 1 μg of total RNA was reverse transcribed with the Moloney murine leukemia virus reverse transcriptase (Promega) and random hexamer or oligo-d(T) primers. Exogenous cDNA samples were amplified using primers PL4 and ATM-F and endogenous products were amplified with primers ATM-F and ATM-R (Table 2). Spliced products were separated on agarose and polyacrylamide gels and their signal intensities were measured. Statistical analysis was carried out with Stat200 (BioSoft, UK).

Table 2 Oligonucleotide primers

Primer	**5'-3' sequence**	**SEQ ID NO:**	**Primer**	**5'-3' sequence**	**SEQ ID NO:**
Cloning primers			A16	caaccaguuugcauucgu	25
ATM26	ataaagcttcttgttataaggttttgattcc	1	A17	uuaguauuccuugacuuua	26
ATM27	atatctagatgtacataccctgaaaagtcac	2	A18	uucuguacacuguuuaguauucc	27
RT-PCR primers			A19	gaagagggagugaagguu	28
PL4	agtcgaggctgatcagcgg	3	A20	aaagcuuggugagauuga	29
ATM-F	gagggtaccagagacagtgggatggc	4	A21	uuucuugaaaaguggaaagcuug	30
ATM-R	ggctcatgtaacgtcatcaat	5	A22	uggaaugagggacgguuguuuuuc	31
Mutagenic primers			A23	gguaugagaacuauagga	32
del-1F	atacaatttaccataatttacttttgaattatgtt	6	A24	aaacaaacagcaggguau	33
del-1R	aagtaaattatggtaaattgtatcatacattag	7	A25	gguaauaagugucacaaa	34
del-2F	ccttgccagaccagtttcctagttatctatattgaac	8	A26	guaucauacauuagaagg	35
del-2R	taactaggaaactggtctggcaaggtggctta	9	B1	ucaaaaguaaauuauggucu	36
del-3F	cttcaagggaccttggccgggtgcggtggct	10	B2	gacugguaaauaauaaacauaauuc	37
del-3R	gcacccggccaaggtcccttgaagtttatctaa	11	B3	aaauguauacuggagaagacu	38
del-4F	acacaaacaaagcttaggtttctttcttgtcaccttcta	12	B4	auauauuagagauacaucagcc	39
del-4R	agaaagaaacctaagctttgtttgtgtgttttatacaa	13	B5	gacaaacauuuaaugaauacucaa	40
del-5F	tgcctcatttacgtcatacaacttaatgatagacct	14	B6	uugacuccuucuuuugacaaacau	41
del-5R	ttaagttgtatgacgtaaatgaggcagggcaa	15	B7	uuuaaauccuuccuuacuu	42
del-6F	tgatacaatttacctcatacaacttaatgatagacct	16	B8	gauuauaaaacaaacgaagc	43
del-6R	attaagttgtatgaggtaaattgtatcatacattag	17	B10	uguuuuaauauaaguugcuucaa	44
2'-O-methyl/ PTO SSOs[a]			B11	ugugggugaccacagcuu	45
A2	aacuuaaagguuauaucuc	18	B12	ucccuuacuuauauccaa	46
A4	uauaaauacgaauaaaucga	19	B13	ccaaguuugguuacuuauc	47
A8	cauggguuggcuaugcuag	20	B14	gaaguuuaucuaauauugacc	48
A9	caacacgacauaaccaaa	21	AN1	ggucuaucauuaaguuguauga	49

Primer	5'-3' sequence	SEQ ID NO:	Primer	5'-3' sequence	SEQ ID NO:
A10	aagccaaucagagggagaca	22	AN2	uuaaauaagacuucaggucua	50
A11	aacauuucuauuuaguuaaaagc	23	AN3	uuagagaaucauuuuaaauaagac	51
A15	ucguguauuacaacaguuaa	24	AN4	cuuaauccaauucuucaauuuuag	52

[a] PTO, phosphorothioate

Results

[0276] SSOs targeting either 3' or 5'ss of the NSE efficiently repress this exon in a haplotype dependent manner. To facilitate identification of optimal intronic SSOs that activate NSE, splicing reporter constructs with the entire ATM intron 28 (Fig. 16A) were first prepared. The construct was obtained by PCR using the HEK293 DNA as a template. The reference sequence (hg19) of intron 28 is ~3,100 nt long, which is similar to the average human intron. Transposed elements occupy ~64% of intron 28, filling completely its middle part, except for a ~350 nt region in the 5' half of the intron and exonic flanks (Fig. 16A). Plasmid DNA sequencing revealed the same organization of transposed elements without any additional transposon copies. It also showed the C and G allele at rs4988000 and rs609261, respectively, indicating that the construct contains the haplotype most permissive for NSE inclusion in the ATM mRNA. After transfections into HEK293 cells, total RNA was extracted and reverse transcribed prior to amplification with a vector primer PL4 (Table 2) and an exon primer (Fig. 16A). Examination of spliced products showed that most transcripts entirely lacked intronic sequences (NSE-) whereas ~36% mRNA contained NSE (Fig. 16B, lane 1), a fraction slightly higher than for a hybrid reporter reported previously.

[0277] To determine the importance of transposed elements for NSE inclusion, each transposon from intron 28 was individually deleted using overlap-extension PCR (deletions 1-5, Fig. 16A). A large middle part of the intron was also deleted along with all transposons, leaving the NSE and its upstream sequences intact (~75% of the intron, deletion 6). Transfection of validated mutated constructs, which all had identical genotypes to the wildtype construct at rs4988000 and rs609261, revealed that the large deletion promoted NSE-containing transcripts (deletion 6, Fig. 16B). Deletion of the MER51 element increased NSE inclusion to a lesser extent. In contrast, deletion of the antisense Alu inhibited NSE while deletion of long interspersed repeats (deletions 3 and 5) or a unique intronic segment (deletion 2) had no effect on NSE activation. The variability of NSE inclusion levels was much higher following a two-hit knockdown of U2AF35, with a significant increase of NSE levels maintained only for deletion 6 (Fig. 16B), consistent with a major stress component of NSE responses. A series of SSOs were then designed targeting three intronic regions

that have unique sequences in the genome (termed A, B and AN) while avoiding a predicted branch site upstream of NSE (Fig. 17A, Table 2). Each SSO was modified with 2'-O-methyl at each ribose and phosphorothioate at each end linkage to ensure their RNase H resistance and sufficient stability in transient transfections. As positive and negative controls, SSO-NSE3 was used, which was highly efficient in blocking the NSE 3'ss, and a series of scrambled SSOs and SSOs targeting other genes, including INS and BTK which were not expressed in HEK293 cells, as confirmed by RNA-Seq. Each SSO was individually transfected with or without the wild-type ATM construct.

[0278] Measurements of spliced products revealed that SSO-NSE3 yielded the most efficient NSE repression (Fig. 17B). About a half of tested SSOs significantly altered NSE inclusion levels as compared to controls, with similar numbers of repressor and activator SSOs. The Pearson correlation coefficient between replicate transfections was highly significant, reaching 0.88 ($P<10^{-8}$); however, the overall correlation between exogenous and endogenous NSE levels was only 0.35 ($P<0.01$).

[0279] Experiments in Fig. 16 showed that the NSE inclusion is controlled by distant splicing regulatory sequences within and outside transposons. Experimentally determined splicing enhancer and silencer motifs in their natural pre-mRNA context occur preferentially in single-stranded regions, suggesting that they are more accessible to RNA binding proteins or other ligands that control exon selection. Preferential targeting of SSOs to unpaired regions could thus improve a search for intronic SSOs. To test this assumption, NSE inclusion levels induced by each SSO were correlated with their average PU values (Fig. 17C). These values estimate single-strandedness of their RNA targets using an equilibrium partition function, with higher values signaling a higher probability of single-stranded conformation. Interestingly, SSO targets with higher average PU values tended to induce exon skipping, suggesting that efficient blocking of unpaired interactions as far as 2 kb from the exon can impair its activation.

[0280] The experiments described above identified a small set of intronic SSOs that activated NSE inclusion in mature exogenous and endogenous transcripts. Since NSE can limit ATM expression through NMD, activator and repressor SSOs could serve as tunable gene-specific inhibitors. Transient ATM repression by NSE-activating SSOs could be advantageous for cancer treatment by inhibiting the double-strand break signaling pathway and radiosensitization.

[0281] To test if ATM SSOs can be delivered to cells that have much lower transfection efficiency than HEK293 cells, a stearylated trimethylated chitosan (TMC-SA) was employed. Chitosan is a natural copolymer of D-glucosamine and N-acetyl-D-glucosamine known for biocompatibility, biodegradability and low toxicity and immunogenicity. When trimethylated, chitosan acquires a permanent positive charge that improves its solubility at neutral pH. Stearylation was found

necessary for formation of stable nanocomplexes with SSOs and their transfection activity in a HeLa/pLuc705 system, which makes use of a luciferase gene interrupted by a mutated HBB1 intron.

[0282] Wheher TMC-SA can facilitate delivery of SSO-NSE3 into HEK293 cells was first tested. Fig. 18A shows reduction of NSE levels following exposure to SSO-NSE3-TMC nanoparticles as compared to a scrambled SSO. This decline was significant for the TMC-SA/SSO-NSE3 (N/P) ratios of 20 and 40. The NSE decline was also apparent when comparing NSE inclusion in cells exposed to uncomplexed SSO-NSE3, consistent with their significant uptake by this highly transfectable cell line. However, the reduction of NSE levels was less efficient for TMC-SA/SSO-NSE3 than for the same oligo transfected with jetPrime to the same cell line at a lower final concentration. A significant NSE repression upon exposure to TMC-SA/SSO-NSE3 nanocomplexes was observed also for a lymphoblastoid cell line where uncomplexed SSO-NSE3 failed to reduce NSE (Fig. 18B). Collectively, these results provide the first proof-of-principle that a chitosan-based delivery system of intronic SSOs can repress NMD switch exons in human cells.

Discussion

[0283] This work shows the first example of transposed elements that promote and repress activation of a NMD switch exon (Fig. 16). Alu sequences themselves have a propensity to exonize through 3'ss or 5'ss activation or auxiliary splicing motifs, which contributes significantly to human morbidity. They can also be exonized by outlying deletions and cause genetic disease, suggesting that they can promote inclusion of distant intronic sequences in mature transcripts. This is further supported by a higher fraction of Alus that flank alternatively spliced exons than those spliced constitutively. Although the exact mechanism of these distant effects is not understood, secondary structure of these GC-rich transcripts is likely to play a major role.

[0284] Mutation-induced exonizations have been shown for all other classes of transposed elements, including more ancient short interspersed elements termed mammalian interspersed repeats. In the present study, an intronic transposed element with the highest similarity to MER51A (Medium Reiterated frequency repeat, family 51) repressed NSE, acting as a buffer to counteract the Alu-mediated NSE activation (Figs. 16A and 16B). The ATM MER51 is relatively GC-rich (~44%), which may facilitate intramolecular interactions with GC-rich Alus during co-transcriptional folding. The element contains several inverted repeats, possibly forming stable hairpins containing exposed purine-rich loops that may control NSE inclusion (Fig. 19). About 250,000 copies of recognizable MER sequences were estimated to exist in the human genome and many were found in mature transcripts of protein-coding genes, contributing to the diversity of protein interactions. A mutation-induced MER exonization event was also shown to cause Gitelman syndrome. The 3' part of

MER51 is similar to a long terminal repeat of retroviruses (Fig. 19), which account for ~15% of disease-causing exonizations. The origin of most MERs was placed after the decline of mammalian interspersed repeats before the spread of Alus, coinciding with expansion of mammals and suggesting that MERs may offer insight into early mammalian radiation. However, the molecular mechanisms underlying MER-mediated exon activation are not understood and will require further studies. Taken together, these results suggest that the interplay of transposed elements in long introns could influence inclusion levels of many NMD switch exons, fine-tuning gene expression.

[0285] In this work, candidate sequence-specific ATM inhibitors that act by promoting a regulated NMD switch exon critical for ATM expression were also identified (Fig. 17). ATM inhibitors sensitize cancer cells to cytotoxic therapy that induces double-strand breaks, including local radiotherapy, which is an integral part of treatment regimens of many cancer types. Although chemical ATM inhibitors showed a great promise for cancer radiotherapy, their undesired pharmacokinetic profiles, high toxicity or poor efficacy have hampered their progression into the clinic. In contrast, newly identified SSOs target unique sequences in the human genome, their mechanism of action is well-defined and they can be delivered to cells using natural biodegradable compounds (Fig. 18). In addition, the availability of NSE-activating and - repressing SSOs provides an opportunity to titrate gene expression more accurately than chemical inhibitors. The approach described herein makes use of SSO-mediated modulation of cryptic exons that activate NMD. These exons are usually present in natural transcripts at very low levels but their inclusion levels can be efficiently upregulated in response to various stimuli. Recently, a gene-specific antisense inhibition of NMD employed SSOs targeting exon junction complex deposition sites, thus permitting NMD repression without relying on skipping of a PTC-containing exon. The two approaches, the former relying on intronic sequence and the latter one on exonic targets, might complement each other in the future to expand the repertoire of antisense strategies that inhibit NMD.

[0286] The average length of SSOs employed in the screening was close to the minimum for unique targets (Table 2). The short SSOs may induce more off-target effects than longer SSOs, which could contribute to the low correlation between inclusion levels of endogenous and exogenous NSE transcripts. Apart from the possible suboptimal target specificity, intron 28 splicing and NSE inclusion can be influenced by adjacent introns that were absent in exogenous transcripts. In addition, intron 28 splicing may not be entirely co-transcriptional and folding and folding kinetics of RNAs transcribed from different promoters are likely to be distinct, contributing to the low correlation. Nevertheless, this study clearly demonstrates a wealth of candidate intronic target sites for SSOs in the human genome, consistent with a higher information content of intronic auxiliary

splicing sequences as compared to lower organisms, which have smaller introns with a lower regulatory potential for alternative splicing. Although SSO-NSE3 and other SSOs can repress endogenous NSE-containing mRNAs (Figs. 17B and 17C) and NMD transcripts with the relative abundance as low as ~1% can contribute to the mRNA consumption, it remains to be tested if their reduction can lead to a sustained increase of ATM protein levels in normal cells. This approach may have a potential to alleviate phenotypic consequences of leaky A-T alleles in a mutation-independent manner, especially in homozygous A-T patients carrying the C allele at rs609261, which facilitates 3'ss recognition of the NSE. Finally, chitosan-based nanoparticles have been shown to penetrate the blood-brain barrier and accumulate in cerebellum without affecting histomorphology, opening a possibility to deliver NSE repressors and putative ATM activators to neural cells to ameliorate cerebellar symptoms of AT.

CLAIMS

WHAT IS CLAIMED IS:

1. A method of screening a subject for susceptibility to a functional-ATM protein deficiency, wherein the screening comprises determining a presence of a non-thymine variant residue rs609261 located at position -3 relative to a 3' splice site of a NSE (cryptic exon in ATM intron 28) of a human subject's genome,
wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, a functional-ATM protein deficiency.
2. A method of selecting a subject for treatment, wherein the subject is susceptible to a functional-ATM protein deficiency, the method comprising determining a presence of a non-thymine variant residue rs609261 located at position -3 relative to a 3' splice site of a NSE (cryptic exon in ATM intron 28) of a human subject's genome,
wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, a functional-ATM protein deficiency, and selecting the subject for treatment with an agent thereby increasing a functional-ATM level in the subject.
3. The method according to claim 2, wherein the method further comprises administering the agent to the selected subject, thereby treating the subject.
4. The method according to claim 2 or claim 3, wherein the agent comprises a NSE repressor agent.
5. A method of treatment or prevention of a functional-ATM protein deficiency in a subject, the method comprising identifying a presence of a non-thymine variant residue rs609261 located at position -3 relative to a 3' splice site of NSE (cryptic exon in ATM intron 28) of a human subject's genome,
wherein the presence of a non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, a functional-ATM protein deficiency, and administering an agent to the subject, wherein the agent increases a functional-ATM level.
6. A method of treatment or prevention of a condition associated with a functional-ATM protein deficiency, comprising administering a NSE repressor agent to a subject, thereby increasing a functional ATM protein level, wherein the agent binds to a NSE in ATM intron 28 of a pre-mRNA transcript, thereby decreasing inclusion of the NSE in a mature RNA transcript of the pre-mRNA transcript.
7. The method according to claim 6, wherein the decreasing inclusion of the NSE in the mature RNA transcript increases functional ATM protein expression.

8. The method according to any one of claims 5 to 7, wherein the method is for treatment or prevention of a condition or symptoms associated with a functional-ATM protein deficiency.
9. The method according to claim 8, wherein the condition is ataxia-telangiectasia, cancer, immune deficiency, cellular radiosensitivity or chromosomal instability.
10. The method according to any preceding claim, wherein the NSE comprises a sequence comprising tctacaggttggctgcatagaagaaaaag.
11. The method according to any one of clams 4 to 10, wherein the NSE repressor agent binds to the NSE within a sequence comprising
agTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or
tcttagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag; or
tctcagTCTACAGGTTGGCTGCATAGAAGAAAAAGgtagag.
12. The method according to any one of claims 4 to 10, wherein the NSE repressor agent binds to the NSE or s 5' or 3' splice site in ATM intron 28 of the NSE.
13. A method of treatment or prevention of a condition associated with deregulation of ATM expression in a subject comprising administering a NSE-activator agent to the subject, wherein the NSE-activator agent increases inclusion of a NSE in an ATM mature RNA transcript by binding to a regulatory motif in ATM intron 28, or by binding to a U2AF65 binding site upstream of a pseudoexon located 3' of a NSE in ATM intron 28 of an ATM pre-mRNA transcript.
14. A method of treatment or prevention of cancer in a subject comprising administering a NSE-activator agent to the subject, wherein the NSE-activator agent increases a cancer cell's susceptibility to cytotoxic therapy with DNA damaging agents such as radiotherapy, wherein the NSE-activator agent increases inclusion of a NSE in an ATM mature RNA transcript by binding to a regulatory motif in ATM intron 28, or by binding to a U2AF65 binding site upstream of a pseudoexon located 3' of a NSE in ATM intron 28 of an ATM pre-mRNA transcript, and
treating the subject with the cytotoxic therapy, such as radiotherapy or chemotherapy.
15. A method of increasing a cell's susceptibility to cytotoxic therapy with DNA damaging agents such as radiotherapy or chemotherapy comprising reducing ATM protein expression by administering a NSE-activator agent, wherein the NSE-activator agent increases inclusion of a NSE in an ATM mature RNA transcript by binding to motifs in ATM intron 28, or by binding to a U2AF65 binding site upstream of a pseudoexon located 3' of a NSE in ATM intron 28 of an ATM pre-mRNA transcript.

16. The method according to any one of claims 13 to 15, wherein increasing inclusion of the NSE in the mature RNA transcript provides a decrease in functional ATM protein expression.
17. The method according to any one of claims 13 to 16, wherein the pseudoexon comprises the sequence tcatcgaatacttttggaaataag.
18. The method according to any one of claims 13 to 17, wherein the regulatory motif in ATM intron 28 competes with the NSE for spliceosomal components.
19. The method method according to any one of claims 13 to 18, wherein the regulatory motif in ATM intron 28 comprises a 24 nucleotide pseudoexon (PE) located 3' of the NSE in ATM intron 28 of the pre-mRNA transcript.
20. A method of tailoring functional ATM expression in a subject, cell or tissue, comprising adminstering a NSE-activator agent and/or a NSE-repressor agent described herein.
21. The method according to any preceding claim, wherein the NSE repressor agent and/or NSE activator agent comprises a polynucleic acid polymer.
22. The method according to any preceding claim, wherein the NSE repressor agent and/or NSE activator agent is an SSO (Splice Switching Oligonucleotide).
23. The method according to any preceding claim, wherein the NSE repressor agent and/or NSE activator agent is associated with a delivery vehicle suitable for delivering the NSE repressor agent and/or NSE activator agent to cells.
24. The method according to any one of claims 4 to 12 or 20 to 23, wherein the NSE repressor agent comprises:

an SSO of the sequence cuucuaugcagccaaccuguagacu (SSO -NSE3), or a nucleic acid analogue thereof; or

an SSO of the sequence accuuuuucuucuaugcagccaac (SSO -NSE5), or a nucleic acid analogue thereof; and/or

the NSE repressor agent comprises or consists of any one SSO selected from the group comprising: aacauuucuauuuaguuaaaagc (SSO A11); uuaguauuccuugacuuua (SSO A17); gacugguaaauaauaaacauaauuc (SSO B2); auauauuagagauacaucagcc (SSO B4); and uuagagaaucauuuuaaauaagac (SSO AN3), or combinations thereof.
25. The method according to any one of claims 13 to 19 wherein the NSE activator agent comprises the SSO PEkr/PEdel; and/or

the NSE activator agent comprises or consists of any one SSO selected from the group comprising: aacuuaaagguuauaucuc (SSO A2); uauaaauacgaauaaaucga (SSO A4);

caacacgacauaaccaaa (SSO A9); gguaugagaacuauagga (SSO A23); gguaauaagugucacaaa (SSO A25);guaucauacauuagaagg (SSO A26); and ugugggugaccacagcuu (SSO B11), or combinations thereof.

26. Use of rs609261 and/or rs4988000 genotyping for predicting a subject's response to therapy for conditions associated with ATM deregulation.

27. A composition comprising a NSE repressor agent and/or a NSE activator agent described herein.

28. The composition according to claim 27, wherein the composition is a pharmaceutically acceptable formulation.

29. A method of treatment or prevention of a functional-ATM protein deficiency in a subject, the method comprising identifying a presence of a non-thymine variant residue rs609261 located at position -3 relative to a 3' splice site of a NSE (cryptic exon in ATM intron 28) of a human subject's genome,

wherein the presence of the non-thymine variant residue rs609261 indicates that the subject has, or is susceptible to, the functional-ATM protein deficiency, and administering an agent to the subject, wherein the agent replaces the non-thymine variant residue rs609261 with a thymine residue.

30. A method of treatment or prevention of a functional-ATM protein deficiency in a subject, the method comprising replacing a non-thymine variant residue rs609261 located at position -3 relative to the 3' splice site of a NSE (cryptic exon in ATM intron 28) of a human subject's genome with a thymine residue.

31. The method according to claim 29 or claim 30, wherein replacing the non-thymine variant residue rs609261 comprises administration of an agent to the subject, which is arranged to replace the non-thymine variant residue rs609261 with a thymine residue.

32. The method according to any one of claims 29 to 31, wherein the agent for replacement of the non-thymine residue is a genomic editing molecule.

33. The method according to any of claims 29 to 32, wherein the agent for replacement of the non-thymine residue is CRISPR-Cas9, or a functional equivalent thereof, together with an RNA molecule that targets rs609261.

34. A method of treatment or prevention of a functional-ATM protein deficiency in a subject, the method comprising identifying a presence of a guanine variant residue at rs4988000 of a human subject's genome,

wherein the presence of the guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, the functional-ATM protein deficiency, and administering an agent to the subject, wherein the agent replaces the guanine variant residue at rs4988000 with an adenine.

35. A method of treatment or prevention of a functional-ATM protein deficiency in a subject, the method comprising replacing a guanine variant residue at rs4988000 of a human subject's genome with an adenine residue, or blocking the guanine variant residue at rs4988000 by the binding an SSO to the guanine variant residue at rs4988000.

36. The method according to claim 34 or claim 35, wherein replacing the guanine variant residue at rs4988000 comprises administration of an agent to the subject, which is arranged to replace the guanine variant residue at rs4988000 with an adenine residue.

37. The method according to any one of claims 34 to 36, wherein the agent for replacement of the guanine residue is a genomic editing molecule.

38. The method according to any one of claims 31 to 34, wherein the agent for replacement of the guanine residue is CRISPR-Cas9, or a functional equivalent thereof, together with an RNA molecule that targets rs4988000.

39. A method of screening a subject or a population of subjects for susceptibility to a functional-ATM protein deficiency, wherein the screening comprises determining a presence of a guanine variant residue at rs4988000 of a human subject's genome,
wherein the presence of a guanine variant residue at rs4988000 indicates that the subject (or group of subjects) has, or is susceptible to, functional-ATM protein deficiency.

40. A method of selecting a subject or a population of subjects for treatment or prophylaxis, wherein the subject is susceptible to a functional-ATM protein deficiency, the method comprising determining a presence of a guanine variant residue at rs4988000 of a human subject's genome,
wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, the functional-ATM protein deficiency, and selecting the subject for treatment with an agent that increases functional-ATM levels in the subject.

41. A method of treatment or prevention of functional-ATM protein deficiency in a subject, the method comprising identifying a presence of a guanine variant residue at rs4988000 of a human subject's genome,

wherein the presence of a guanine variant residue at rs4988000 indicates that the subject has, or is susceptible to, the functional-ATM protein deficiency, and administering an agent to the subject, wherein the agent increases functional-ATM levels.

42. The method of any one of claims 29-33 and any one of claims 34-38 in combination to modify a CG haplotype to TA.

43. The method of claim 1 and claim 36 in combination to identify a CG haplotype in a subject, and optionally to treat or to select a subject for treatment according to any preceding claim.

44. A method of modifying regulation of inclusion of a NSE in a mature RNA transcript, the method comprising inserting or deleting one or more splicing regulatory motifs upstream or downstream of the NSE that compete with the NSE for spliceosomal components, said one or more splicing regulatory motifs comprising a cryptic splice site or a pseudo-exon.

45. A method of modifying regulation of expression of a functional protein, wherein the expression of a functional protein is regulated by inclusion of a NSE in a mature RNA transcript of a gene encoding the functional protein, the method comprising inserting or deleting one or more splicing regulatory motifs upstream or downstream of the NSE that compete with the NSE for spliceosomal components, said one or more splicing regulatory motifs comprising cryptic splice sites or pseudo-exons.

46. The method according to claim 44 or 45, wherein the insertion or the deletion of the one or more splicing regulatory motifs is in genomic DNA of ATM intron 28.

47. The method according to any of claims 44 to 46, wherein insertion of the one or more splicing regulatory motifs causes a reduction in the inclusion of the NSE in the mature RNA transcript.

48. The method according to any of claims 44 to 46, wherein the deletion of the one or more splicing regulatory motifs causes an increase in the inclusion of the NSE in the mature RNA transcript.

49. The method according to any one of claims 44 to 48, wherein the insertion or the deletion of the one or more splicing regulatory motifs comprises the use of genome editing technology, such as CRISPR-Cas9.

50. A kit comprising one or more oligonucleotide probes for identifying rs609261 and/or rs4988000 variants.

51. The kit according to claim 50, wherein the one or more oligonucleotide probes are primers for use in PCR amplifying a region of anucleic acid comprising the rs609261 and/or the rs4988000 variants.

52. A vector comprising a nucleic acid encoding a NSE activating agent and/or a NSE repressor agent.

53. A method of screening for an agent capable of modifying regulation of expression of a gene comprising:

identifying a nonsense-mediated RNA decay switch exon (NSE) that limits functional gene expression;

identifying one or more splicing regulatory motifs upstream or downstream of the NSE that compete with the NSE for spliceosomal components, said one or more splicing regulatory motifs comprising cryptic splice sites or pseudoexons;

targeting the one or more splicing regulatory motifs with an antisense polynucleic acid comprising a sequence that hybridizes to a splicing regulatory motif of the one or more splicing regulatory motifs through Watson-Crick base pairing; and

determining if there is an increased or decreased inclusion of the NSE in a mature RNA transcript of the gene.

54. A method of modulating expression of a gene comprising providing an agent that binds to a splicing regulatory motif, such as a cryptic splice site or a pseudoexon, that competes with a nonsense-mediated RNA decay switch exon (NSE) for spliceosomal components.

55. An agent that binds to a gene splicing regulatory motif, such as a cryptic splice site or a pseudoexon, that competes with a nonsense-mediated RNA decay switch exon (NSE) for spliceosomal components, wherein the gene splicing regulatory motif controls inclusion of the NSE into a mature RNA transcript of the gene.

56. A method, use, composition, vector, or agent substantially described herein, optionally with reference to the accompanying figures.

57. A method of modulating protein expression comprising:

a) contacting an isolated polynucleic acid polymer to a target cell of a subject;

b) hybridizing the contacted polynucleic acid polymer to a target motif on a pre-processed mRNA transcript, wherein a hybridization of the contacted polynucleic acid polymer to the target motif either promotes or represses activation of a non-sense mediated RNA decay switch exon (NSE);

c) processing a mRNA transcript of the pre-processed mRNA transcript, wherein the NSE is either present or absent in the mRNA transcript; and

d) translating the processed mRNA transcript of step c), wherein the presence or absence of the NSE modulates protein expression.

58. The method of claim 57, wherein the protein is expressed from the processed mRNA transcript.
59. The method of claim 57, wherein the presence of the NSE downregulates protein expression.
60. The method of claim 57, wherein the absence of the NSE upregulates protein expression.
61. The method of claim 57, wherein the polynucleic acid polymer hybridizes to a motif within ATM intron 28.
62. The method of claim 61, wherein the motif is a splicing regulatory motif that competes with the NSE for a spliceosomal component.
63. The method of claim 62, wherein the splicing regulatory motif comprises a cryptic splice site or a pseudoexon.
64. The method of claim 63, wherein the pseudoexon is a 24 nucleotide pseudoexon located at 3' of a NSE in ATM intron 28 of the pre-mRNA transcript.
65. The method of claim 62, wherein the motif is a U2AF65 binding site.
66. The method of claim 62, wherein the motif is a motif within a transposed element, upstream of a transposed element, or downstream of a transposed element.
67. The method of claim 66, wherein the transposed element is Alu or MER51.
68. The method of claim 67, wherein the isolated polynucleic acid polymer hybridizes to a target motif within Alu.
69. The method of claim 67, wherein the isolated polynucleic acid polymer hybridizes to a target motif that is either upstream or downstream of Alu.
70. The method of claim 67, wherein the isolated polynucleic acid polymer hybridizes to a target motif downstream of MER51.
71. The method of any one of claims 57-70, wherein the polynucleic acid polymer is from about 10 to about 50 nucleotides in length.
72. The method of any one of claims 57-71, wherein the isolated polynucleic acid polymer comprises a sequence with at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity to a sequence selected from the group consisting of SEQ ID NOs: 18-52.
73. A method of modulating protein expression comprising:
    a) contacting an isolated polynucleic acid polymer to a target cell of a subject;
    b) hybridizing the contacted polynucleic acid polymer to a target motif within a transposed element, wherein a hybridization of the contacted polynucleic acid polymer to the target motif either promotes or represses activation of a non-sense mediated RNA decay switch exon (NSE);

c) processing a mRNA transcript of the pre-processed mRNA transcript, wherein the NSE is either present or absent in the mRNA transcript; and

d) translating the processed mRNA transcript of step c), wherein the presence or absence of the NSE modulates protein expression.

74. A method of modulating protein expression comprising:

a) contacting an isolated polynucleic acid polymer to a target cell of a subject;

b) hybridizing the contacted polynucleic acid polymer to a target motif either upstream or downstream of a transposed element, wherein a hybridization of the contacted polynucleic acid polymer to the target motif promotes or represses activation of a non-sense mediated RNA decay switch exon (NSE);

c) processing a mRNA transcript of the pre-processed mRNA transcript, whereinthe NSE is either present or absent in the mRNA transcript; and

d) translating the processed mRNA transcript of step c), wherein the presence or absence of the NSE modulates protein expression.

75. The method of claim 73 or 74, wherein the protein is expressed from the processed mRNA transcript.

76. The method of claim 73 or 74, wherein the presence of the NSE downregulates protein expression.

77. The method of claim 73 or 74, wherein the absence of the NSE upregulates protein expression.

78. The method of claim 73 or 74, wherein the transposed element is Alu or MER51.

79. The method of claim 73 or 74, wherein the isolated polynucleic acid polymer comprises a sequence with at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity to a sequence selected from the group consisting of SEQ ID NOs: 18-52.

80. The method of any one of the claims 73, 74 or 79, wherein the polynucleic acid polymer is from about 10 to about 50 nucleotides in length.

81. The method of claim 73, wherein the isolated polynucleic acid polymer hybridizes to a target motif within Alu.

82. The method of claim 74, wherein the isolated polynucleic acid polymer hybridizes to a target motif that is either upstream or downstream of Alu.

83. The method of claim 74, wherein the isolated polynucleic acid polymer hybridizes to a target motif downstream of MER51.

84. The method of claim 73 or 74, wherein activation of the NSE further induces exon skipping.

85. The method of any one of the claims 73, 74 or 84, wherein the NSE is located in intron 28.

86. The method of any one of the claims 73, 74, 84 or 85, wherein the NSE modulates ATM protein expression.

87. A method of treating or preventing a disease or condition associated with deregulation of ATM expression in a subject in need thereof, the method comprising:

administering to the subject a pharmaceutical composition comprising:

(i) a non-sense mediated RNA decay switch exon (NSE)-activator agent that interacts with a pre-processed mRNA transcript to promote inclusion of a NSE into a processed mRNA transcript; and

(ii) a pharmaceutically acceptable excipient and/or a delivery vehicle;

wherein the disease or condition associated with deregulation of ATM expression is treated or prevented in the subject by the administration of the NSE-activator agent.

88. A method of treating or preventing a disease or condition associated with a functional-ATM protein deficiency in a subject in need thereof, the method comprising:

administering to the subject a pharmaceutical composition comprising:

(i) a non-sense mediated RNA decay switch exon (NSE)-repressor agent that interacts with a pre-processed mRNA transcript to promote exclusion of a NSE into a processed mRNA transcript; and

(ii) a pharmaceutically acceptable excipient and/or a delivery vehicle;

wherein the disease or condition associated with a functional-ATM protein deficiency is treated or prevented in the subject by the administration of the NSE-repressor agent.

89. The method of claim 87, wherein the NSE-activator agent is an isolated polynucleic acid polymer.

90. The method of claim 88, wherein the NSE-repressor agent is an isolated polynucleic acid polymer.

91. The method of claim 89 or 90, wherein the polynucleic acid polymer hybridizes to a motif within ATM intron 28.

92. The method of claim 89, wherein the polynucleic acid polymer hybridizes to a splicing regulatory motif that competes with the NSE for spliceosomal components.

93. The method of claim 92, wherein the splicing regulatory motif comprises a cryptic splice site or a pseudoexon.

94. The method of claim 93, wherein the pseudoexon is a 24 nucleotide pseudoexon located at 3' of a NSE in ATM intron 28 of the pre-mRNA transcript.

95. The method of claim 89, wherein the polynucleic acid polymer hybridizes to a U2AF65 binding site.

96. The method of claim 89 or 90, wherein the polynucleic acid polymer hybridizes to a motif within a transposed element, upstream of a transposed element, or downstream of a transposed element.

97. The method of claim 96, wherein the transposed element is Alu or MER51.

98. The method of claim 97, wherein the isolated polynucleic acid polymer hybridizes to a target motif within Alu.

99. The method of claim 97, wherein the isolated polynucleic acid polymer hybridizes to a target motif that is either upstream or downstream of Alu.

100. The method of claim 97, wherein the isolated polynucleic acid polymer hybridizes to a target motif downstream of MER51.

101. The method of any one of the claims 89-100, wherein the polynucleic acid polymer is from about 10 to about 50 nucleotides in length.

102. The method of any one of the claims 89-101, wherein the isolated polynucleic acid polymer comprises a sequence with at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity to a sequence selected from the group consisting of SEQ ID NOs: 18-52.

103. The method of claim 87, wherein the disease or condition is cancer.

104. A method of treating or preventing a disease or condition in a subject in need thereof, the method comprising:

administering to the subject a pharmaceutical composition comprising:

(i) a non-sense mediated RNA decay switch exon (NSE)-activator agent that interacts with a pre-processed mRNA transcript to promote inclusion of NSE into a processed mRNA transcript; and

(ii) a pharmaceutically acceptable excipient and/or a delivery vehicle;

wherein the disease or condition is treated or prevented in the subject by the administration of the NSE-activator agent.

105. A method of treating or preventing a disease or condition in a subject in need thereof, the method comprising:

administering to the subject a pharmaceutical composition comprising:

(i) a non-sense mediated RNA decay switch exon (NSE)-repressor agent that interacts with a pre-processed mRNA transcript to promote exclusion of an NSE into a processed mRNA transcript; and

(ii) a pharmaceutically acceptable excipient and/or a delivery vehicle;

wherein the disease or condition is treated or prevented in the subject by the administration of the NSE-repressor agent.

106. The method of claim 104, wherein the NSE-activator agent is an isolated polynucleic acid polymer.

107. The method of claim 105, wherein the NSE-repressor agent is an isolated polynucleic acid polymer.

108. The method of claim 106 or 107, wherein the polynucleic acid polymer hybridizes to a motif within ATM intron 28.

109. The method of claim 106, wherein the polynucleic acid polymer hybridizes to a splicing regulatory motif that competes with the NSE for spliceosomal components.

110. The method of claim 109, wherein the splicing regulatory motif comprises a cryptic splice site or a pseudoexon.

111. The method of claim 110, wherein the pseudoexon is a 24 nucleotide pseudoexon located at 3' of a NSE in ATM intron 28 of the pre-mRNA transcript.

112. The method of claim 108, wherein the polynucleic acid polymer hybridizes to a U2AF65 binding site.

113. The method of claim 106 or 107, wherein the polynucleic acid polymer hybridizes to a motif within a transposed element, upstream of a transposed element, or downstream of a transposed element.

114. The method of claim 113, wherein the transposed element is Alu or MER51.

115. The method of claim 114, wherein the isolated polynucleic acid polymer hybridizes to a target motif within Alu.

116. The method of claim 114, wherein the isolated polynucleic acid polymer hybridizes to a target motif that is either upstream or downstream of Alu.

117. The method of claim 114, wherein the isolated polynucleic acid polymer hybridizes to a target motif downstream of MER51.

118. The method of any one of the claims 104-117, wherein the polynucleic acid polymer is from about 10 to about 50 nucleotides in length.

119. The method of any one of the claims 104-118, wherein the isolated polynucleic acid polymer comprises a sequence with at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity to a sequence selected from the group consisting of SEQ ID NOs: 18-52.

120. The method of claim 104, wherein the disease or condition is cancer.

121. The method of claim 104, wherein the disease or condition is a disease or condition associated with deregulation of ATM expression.

122. The method of claim 105, wherein the disease or condition is a disease or condition associated with a functional-ATM protein deficiency.

123. The method of any one of the claims 57-122, wherein the polynucleic acid polymer is modified at a nucleoside moiety, at a phosphate moiety, at a 5' terminus, at a 3' terminus, or a combination thereof.

124. The method of any one of the claims 57-123, wherein the polynucleic acid polymer comprises an artificial nucleotide.

125. The method of claim 124, wherein the artificial nucleotide is selected from the group consisting of 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), 2'-O-N-methylacetamido (2'-O-NMA), a locked nucleic acid (LNA), an ethylene nucleic acid (ENA), a peptide nucleic acid (PNA), a 1',5'- anhydrohexitol nucleic acid (HNA), a morpholino, a methylphosphonate nucleotide, a thiolphosphonate nucleotide, and a 2'-fluoro N3-P5'-phosphoramidite.

126. The method of claim 87, 88, 104 or 105, wherein the delivery vehicle comprises a nanoparticle-based delivery vehicle.

127. A method of modulating protein expression comprising:
    a) contacting an isolated polynucleic acid polymer to a target cell of a subject;
    b) hybridizing the contacted polynucleic acid polymer to a target motif within a transposed element, wherein a hybridization of the contacted polynucleic acid polymer to the target motif either promotes or represses activation of an alternative splice site;
    c) processing a mRNA transcript of the pre-processed mRNA transcript, whereinthe alternative splice site is either present or absent in the mRNA transcript; and
    d) translating the processed mRNA transcript of step c), wherein the presence or absence of the alternative splice site modulates protein expression.

128. A method of modulating protein expression comprising:
    a) contacting an isolated polynucleic acid polymer to a target cell of a subject;
    b) hybridizing the contacted polynucleic acid polymer to a target motif either upstream or downstream of a transposed element, wherein a hybridization of the contacted

polynucleic acid polymer to the target motif promotes or represses activation of an alternative splice site;

c) processing a mRNA transcript of the pre-processed mRNA transcript, whereinthe alternative splice site is either present or absent in the mRNA transcript; and

d) translating the processed mRNA transcript of step c), wherein the presence or absence of the alternative splice site modulates protein expression.

129. The method of claim 127 or 128, wherein the transposon element is on the pre-processed mRNA transcript.

130. A method of modulating protein expression comprising:

a) contacting an isolated polynucleic acid polymer to a target cell of a subject;

b) hybridizing the contacted polynucleic acid polymer to a target motif on a pre-processed mRNA transcript, wherein hybridization of the contacted polynucleic acid polymer to the target motif either promotes or represses activation of an alternative splice site;

c) processing a mRNA transcript of the pre-processed mRNA transcript, whereinthe alternative splice site is either present or absent in the mRNA transcript; and

d) translating the processed mRNA transcript of step c), wherein the presence or absence of the alternative splice site modulates protein expression.

131. The method of claim 130, wherein the protein is expressed from the processed mRNA transcript.

132. The method of claim 130, wherein the presence of NSE downregulates protein expression.

133. The method of claim 130, wherein the absence of NSE upregulates protein expression.

134. The method of claim 130, wherein the polynucleic acid polymer hybridizes to a motif within ATM intron 28.

135. The method of claim 134, wherein the motif is a splicing regulatory motif that competes with NSE for a spliceosomal component.

136. The method of claim 135, wherein the splicing regulatory motif comprises a cryptic splice site or a pseudoexon.

137. The method of claim 136, wherein the pseudoexon is a 24 nucleotide pseudoexon located at 3' of a NSE in ATM intron 28 of the pre-mRNA transcript.

138. The method of claim 134, wherein the motif is a U2AF65 binding site.

139. The method of claim 134, wherein the motif is a motif within a transposed element, upstream of a transposed element, or downstream of a transposed element.

140. The method of claim 139, wherein the transposed element is Alu or MER51.

141. The method of claim 140, wherein the isolated polynucleic acid polymer hybridizes to a target motif within Alu.

142. The method of claim 140, wherein the isolated polynucleic acid polymer hybridizes to a target motif that is either upstream or downstream of Alu.

143. The method of claim 140, wherein the isolated polynucleic acid polymer hybridizes to a target motif downstream of MER51.

144. The method of any one of the claims 130-143, wherein the polynucleic acid polymer is from about 10 to about 50 nucleotides in length.

145. The method of any one of the claims 130-144, wherein the isolated polynucleic acid polymer comprises a sequence with at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity to a sequence selected from the group consisting of SEQ ID NOs: 18-52.

146. A pharmaceutical composition comprising:

a) a non-sense mediated RNA decay switch exon (NSE)-activator agent that interacts with a pre-processed mRNA transcript to promote inclusion of NSE into a processed mRNA transcript, or a non-sense mediated RNA decay switch exon (NSE)-repressor agent that interacts with a pre-processed mRNA transcript to promote exclusion of an NSE into a processed mRNA transcript; and

b) a pharmaceutically acceptable excipient and/or a delivery vehicle.

147. The pharmaceutical composition of claim 146, wherein the NSE-activator agent is an isolated polynucleic acid polymer.

148. The pharmaceutical composition of claim 147, wherein the NSE-repressor agent is an isolated polynucleic acid polymer.

149. The pharmaceutical composition of claim 147 or 148, wherein the polynucleic acid polymer hybridizes to a motif within ATM intron 28.

150. The pharmaceutical composition of claim 147, wherein the polynucleic acid polymer hybridizes to a splicing regulatory motif that competes with the NSE for a spliceosomal component.

151. The pharmaceutical composition of claim 150, wherein the splicing regulatory motif comprises a cryptic splice site or a pseudoexon.

152. The pharmaceutical composition of claim 151, wherein the pseudoexon is a 24 nucleotide pseudoexon located at 3' of NSE in ATM intron 28 of the pre-mRNA transcript.

153. The pharmaceutical composition of claim 150, wherein the polynucleic acid polymer hybridizes to a U2AF65 binding site.

154. The pharmaceutical composition of claim 147 or 148, wherein the polynucleic acid polymer hybridizes to a motif within a transposed element, upstream of a transposed element, or downstream of a transposed element.

155. The pharmaceutical composition of claim 154, wherein the transposed element is Alu or MER51.

156. The pharmaceutical composition of claim 154, wherein the isolated polynucleic acid polymer hybridizes to a target motif within Alu.

157. The pharmaceutical composition of claim 154, wherein the isolated polynucleic acid polymer hybridizes to a target motif that is either upstream or downstream of Alu.

158. The pharmaceutical composition of claim 154, wherein the isolated polynucleic acid polymer hybridizes to a target motif downstream of MER51.

159. The pharmaceutical composition of any one of the claims 146-158, wherein the polynucleic acid polymer is from about 10 to about 50 nucleotides in length.

160. The pharmaceutical composition of any one of the claims 146-159, wherein the isolated polynucleic acid polymer comprises a sequence with at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity to a sequence selected from the group consisting of SEQ ID NOs: 18-52.

161. A cell comprising the pharmaceutical composition of any one of claims 146-160.