CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation

Rubina Tuladhar, Yunku Yeu, John Tyler Piazza, Zhen Tan, Jean Rene Clemenceau, Xiaofeng Wu, Quinn Barrett, Jeremiah Herbert, David H. Mathews, James Kim, Tae Hyun Hwang & Lawrence Lum

The introduction of insertion-deletions (INDELs) by non-homologous end-joining (NHEJ) pathway underlies the mechanistic basis of CRISPR-Cas9-directed genome editing. Selective gene ablation using CRISPR-Cas9 is achieved by installation of a premature termination codon (PTC) from a frameshift-inducing INDEL that elicits nonsense-mediated decay (NMD) of the mutant mRNA. Here, by examining the mRNA and protein products of CRISPR targeted genes in a cell line panel with presumed gene knockouts, we detect the production of foreign mRNAs or proteins in ~50% of the cell lines. We demonstrate that these aberrant protein products stem from the introduction of INDELs that promote internal ribosomal entry, convert pseudo-mRNAs (alternatively spliced mRNAs with a PTC) into protein encoding molecules, or induce exon skipping by disruption of exon splicing enhancers (ESEs). Our results reveal challenges to manipulating gene expression outcomes using INDEL-based mutagenesis and strategies useful in mitigating their impact on intended genome-editing outcomes.
Technologies enabling the directed introduction of double-stranded DNA breaks such as CRISPR-Cas9 have transformed our ability to systematically identify DNA sequences important in biology\(^6\). The repair of these double-stranded breaks by non-homologous end-joining (NHEJ) results in insertion-deletions (INDELs) of unpredictable length that upon introduction into exonic sequences could alter the coding frame and install a premature termination codon (PTC). Ribosomes that encounter a PTC in nascent mRNAs, recognized by the assembly of a complex that includes proteins from the ribosome and a 3′ exon–splice junction complex, induces the destruction of the mutant mRNA\(^3,4\). On the other hand, INDELs that preserve the reading frame may yield proteins with altered sequences and thus shed light on determinants important for its function\(^5\).

Exonic sequences are laden with regulatory features that control many facets of the mRNA lifecycle including splicing and folding, two mRNA attributes that influence protein sequence composition and sites of initiation/termination, respectively\(^6\)–\(^8\). Yet, the frequency with which these elements are impacted by INDELs\(^3\) influence gene expression outcomes remains mostly unknown. Another potential obstacle to precision gene editing using INDEL-type mutagenesis is the presence of pseudo-mRNAs, mRNAs harboring a PTC that can nevertheless incorporate introduced INDELs thus altering their potential to produce proteins\(^9\).

To determine the extent to which these molecular events confound our ability to predict gene expression outcomes from CRISPR-Cas9 editing, we have taken inventory of the post-transcriptional and -translational effects of frameshift-inducing INDELs in a panel of CRISPR-edited cell lines. We observe changes in the array of transcripts or proteins expressed from CRISPR-targeted genes in ~50% of the cell lines studied. A mechanistic account of these phenomena is presented here.

**Results**

Unanticipated gene expression outcomes with CRISPR editing. To service several ongoing research programs, we had assembled a panel of commercially available HAP1 cell lines harboring frameshift-inducing INDELs that presumably eliminate effective protein production from the targeted gene by promoting nonsense-mediated decay (NMD) of the encoded mRNA (Fig. 1a; Supplementary Table 1). HAP1 cells harbor a single copy of each chromosome thus reducing the challenges frequently associated with achieving homozygosity in diploid cells for genetic studies\(^10\).

To confirm the effects of the INDEL on-target gene expression, we used two antibodies each recognizing a different epitope within the targeted protein (Fig. 1b; Supplementary Table 2). We observed in some cell lines the anticipated loss of protein presumably due to the introduced INDEL but in other instances the appearance of novel proteins detectable by western blot analysis using a single or both antibodies (4/13 cell lines or ~30%; Fig. 1b). For example, in the case of the TOP1, SIRT1, CTNNB1, and LRP6 knockout cell lines, we observed the substitution of the canonical protein for a faster migrating novel protein detected by western blot analysis.

Given our inability to account for the emergence of these novel proteins based on the annotated genetic alteration introduced by CRISPR-Cas9, we next examined the effects of the INDEL on mRNA splicing given that exonic sequences harbor splicing regulatory elements\(^8,11,12\) (Fig. 1c; Supplementary Table 3). In the case of the TOP1 knockout cell line where we had observed the appearance of a novel TOP1 protein, we also witnessed the emergence of a novel mRNA species. Sequencing a cDNA-derived amplicon from the novel splice variant revealed the absence of the INDEL-containing exon suggesting the mutant protein was generated by an INDEL-induced exon exclusion event (Supplementary Data 1). In addition to the use of two different antibodies to evaluate TOP1 protein in the CRISPR-edited cell line (Fig. 1b), we also observed enrichment of both the wt and truncated TOP1 protein in the nucleus where the protein is predominantly localized\(^13\) (Fig. 2a). The truncated TOP1 protein nevertheless retained catalytic activity as measured using an enzymatic assay for monitoring relaxation of supercoiled DNA (Fig. 2b). The retention of catalytic activity by the truncated TOP1 protein is consistent with the design of TOP1 as an essential gene in HAP1 cells from a gene trap mutagenesis screen that would preclude its elimination in viable cells\(^10,14\). In the case of the VPS35 and TLE3 cell lines, we observed changes in the splice variants harboring the CRISPR-targeted exons although no detectable novel proteins emerged (Fig. 1c).

In contrast to the TOP1 clones, the CTNNB1 and LRP6 cell lines exhibited no detectable change in mRNA splicing associated with the targeted exons suggesting the novel proteins are a consequence of alternative translation initiation (ATI) events presumably induced by the introduced INDELs (Fig. 1c). Consistent with this hypothesis, the mutant LRP6 protein is not glycosylated perhaps as a consequence of default expression in the cytoplasm in the absence of its N-terminal signal sequence (Supplementary Fig. 1A, C). Similarly, the novel β-catenin protein co-migrates on SDS-PAGE with an engineered β-catenin protein initiating from Met88 (Supplementary Fig. 1B). Similar events have previously been reported in transcripts with PTCs introduced proximal to the native initiation site in cancerous cells\(^15\). In summary, in ~50% of CRISPR-edited cell lines acquired from a commercial source, we observed unexpected changes in protein expression or mRNA splicing that challenge the notion that these reagents could be used to report the cellular effects of complete genetic ablation (Fig. 2c). Although not investigated here, conceivably the mutant proteins could also contribute to neomorphic cellular phenotypes.

ATI and pseudo-mRNAs confound CRISPR-based gene knockout. We had complemented our efforts to generate cells genetically null for various genes-of-interest with de novo CRISPR-Cas9-based gene targeting projects. As part of our focus on the tumor suppressor kinase LKB1, we observed the emergence of unexpected protein products—both smaller and larger proteins than the canonical protein—that were not readily explained by the presence of CRISPR-introduced INDELs (Fig. 3a–c). Given the INDELs created in LKB1 are localized to the first protein coding exon (Fig. 3d) and the antibody recognizing the C- but not the N-terminus epitope reported the shortened LKB1 protein on SDS-PAGE (Fig. 3b, c), we concluded that an ATI event induced by CRISPR-Cas9-introduced INDELs likely resulted in an LKB1 protein lacking a portion of its N-terminal sequence (ATI LKB1 protein).

We also noted in MIA, but not HAP1 cells, a slower migrating protein recognized by LKB1 antibodies emerged in CRISPR-Cas9-edited clones with frameshift-inducing INDELs (Fig. 3c; Super LKB1 protein). The appearance of Super LKB1 protein coincided with the appearance of a new mRNA splice variant that contained a 131 bp exon not included in the transcript that encodes the canonical LKB1 protein (Fig. 3e). Consistent with this exon belonging to an LKB1 pseudo-mRNA not previously annotated in MIA cells, the addition of cycloheximide (CHX) to disrupt NMD in parental MIA cells resulted in the emergence of an LKB1 splice variant that includes this exon (Supplementary Fig. 2A). Thus, the same INDELs that induced a frameshift in the canonical transcript now removed a PTC from an LKB1

---

**ARTICLE**

*Nature Communications* | https://doi.org/10.1038/s41467-019-12028-5

2 NATURE COMMUNICATIONS | (2019)10:4056 | https://doi.org/10.1038/s41467-019-12028-5 | www.nature.com/naturecommunications
Fig. 1 Unanticipated gene expression outcomes following on-target CRISPR editing. a The effect of CRISPR-introduced frameshift alterations on mRNA and protein expression was analyzed using a panel of CRISPR-Cas9-edited HAP1 cells that were commercially accessible. The targeted exon, anticipated PTC location following insertion/deletion mutation and the protein recognition sites of antibodies used in panel b are indicated. b Appearance of novel proteins in cells edited with CRISPR-Cas9. HAP1 cells were subjected to western blot analysis using two distinct antibodies. Asterisks (*) indicate novel proteins. c CRISPR-Cas9 gene editing induces expression of novel mRNA species. RT-PCR analysis of edited cells was performed using primers recognizing flanking exons and the amplicons generated were sequenced. Asterisks (*) indicate novel mRNA species. Source data are provided as a Source Data file.
We noted that a cDNA harboring the 1 bp insertion migrated with an engineered protein that initiates at methionine 51 using an algorithm for modeling conserved RNA structures (Supplementary Figs. 4–6). At least using this approach, we anticipate changes in RNA folding that may influence the location of ribosomal initiation. At the same time, we also evaluated the effects of these mutations on cDNAs that encode the predicted pseudo-mRNA sequence (with the 131 bp additional exon). As anticipated, we observed the emergence of proteins that co-migrated with Super LKB1 protein given that either CRISPR-introduced mutation in a transcript promotes INDEL in LKB1 (Supplementary Fig. 7E). More generally, our observations support our initial expectations that—at the presumed PTC introduced by the INDEL (Supplementary Fig. 7A).

We compared the effects on mRNA stability of an INDEL associated with ATI with an INDEL that yielded no detectable LKB1 polypeptides (Supplementary Fig. 7B C) in order to determine if ATI suppressed NMD as a potential mechanism for promoting C-terminally truncated proteins. Comparing the levels of the two LKB1 mRNAs, we observed greater loss of the mRNA in the CRISPR-edited clone lacking any detectable ATI events (Supplementary Fig. 7D). We observed little difference induced by CHX exposure in LKB1 mRNA abundance in the ATI-associated cells when compared to parental cells suggesting that NMD is not acting on the mRNA with an ATI-provoking mutation (Supplementary Fig. 7D). On the other hand, in the case of the CRISPR-edited cell line that expresses no LKB1 polypeptides, we observed a 10-fold change in LKB1 mRNA in the presence of CHX suggesting the mutant mRNA in this case is subject to robust NMD action (Supplementary Fig. 7D). In total, we observed the production of three polypeptides in lieu of the canonical LKB1 protein following the introduction of a frameshift-inducing INDEL: Super LKB1, ATI LKB1, and Short LKB1 (Supplementary Fig. 7E). More generally, our observations also suggest that introducing INDELS early in the transcript increases the potential for an ATI event that is able to clear off all of the splice junction complexes during the pioneer round thus enabling the synthesis of polypeptides with truncations in the C-terminal sequence.

ATI suppresses NMD. Despite the introduction of a frameshift-promoting INDEL in LKB1, we presumed that an ATI event, which restores codon usage to its native phase, would fail to elicit NMD during the pioneer round of translation. At the same time, having avoided destruction, the mutant mRNA is now able to support repeated rounds of translation including presumably short polypeptides initiating at the canonical start site and ending at the PTC. Given our initial western blot analysis of the LKB1 CRISPR-edited clones did not capture low molecular proteins (Fig. 3b, c), re-examination of LKB1 proteins in our CRISPR-edited clones indeed revealed the presence of a small LKB1 polypeptide. This protein (short LKB1) co-migrates with an engineered protein that initiates at the canonical start site but terminates at the presumed PTC introduced by the INDEL (Supplementary Fig. 7A).

Exon symmetry influences CRISPR outcomes. In the analysis of our assembled HAP1 cell line panel, we also observed ~30% of the clones exhibited exclusion of the targeted exon in the mRNA. Exons are replete with splicing regulatory motifs including exon splicing enhancers and suppressors (ESEs and ESSs, respectively). These degenerate hexameric sequences dictate the extent to which exons are included within a transcript. We suspected that exon exclusion was at least in part due to the disruption of ESEs by an INDEL event. As part of our efforts focused on studying the SUFU tumor suppressor protein, we had generated a collection of cells that presumably were null for SUFU based on western blot analysis (Fig. 4a, b). Yet, we noted that many of these clones exhibited exclusion of the targeted exon (Fig. 4c). The extent of…
Fig. 3 ATI and pseudo-mRNAs contribute to foreign protein production in CRISPR-edited cell line. a Genomic structure of the LKB1 gene and the exonic sequence targeted by the LKB1 exon 1 sgRNA. b Emergence of a small LKB1 protein (ATI LKB1) as a consequence of CRISPR-Cas9 gene editing. Lysates generated from CRISPR-edited HAP1 clones were subjected to western blot analysis using two distinct LKB1 antibodies recognizing either N- and C-terminus localized epitopes. c Western blot analysis of CRISPR-Cas9-edited MIA clones reveals the appearance of a large LKB1 protein (Super LKB1) in addition to the ATI LKB1 protein. d Genomic sequences of CRISPR-Cas9-edited HAP1 and MIA clones reveal on-target insertion/deletion mutations in the LKB1 gene. Predicted gene alteration for each clone is indicated. e CRISPR-Cas9-introduced INDELs are associated with the expression of an LKB1 pseudo-mRNA transcript. RT-PCR analysis was performed using primers mapping to 5′ UTR and exon 4 in LKB1 to generate amplicons from the cDNA of CRISPR-Cas9-edited clones. MIA clones M2 and M3, which express Super LKB1 protein, harbor an mRNA species that includes an additional exon. The 131 bp additional exon contains canonical splice acceptor and donor sequences. f A cDNA expression strategy for understanding allele-specific CRISPR-introduced INDELs on protein expression provides evidence for ATI. LKB1 and Super LKB1 cDNA expression constructs harboring genomic alterations found in LKB1 of MIA Clone M2 were introduced into HELA cells that lack endogenous LKB1 expression. The 1 bp insertion or 2 bp deletion in the Super LKB1 cDNA result in proteins that co-migrate with the Super LKB1 protein observed in MIA Clone M2. On the other hand, the same mutations in LKB1 cDNA give rise to proteins that co-migrate with the ATI LKB1 protein found in Clone M2, and with the protein that initiates at Met51. Source data are provided as a Source Data file.

exon exclusion notably differs suggesting other factors, perhaps RNA structure changes that contribute to exon splicing regulation, also may be compromised by the introduction of an INDEL at this position within the SUFU mRNA. We identified a cluster of potential ESEs in the targeted SUFU exon that was likely impacted by the INDEL in these clones (Fig. 4d). No ESSs were identified in this case. To determine how reliably we can induce exon exclusion by impacting a predicted ESE, we introduced INDELs at putative ESEs found in other SUFU exons and performed similar analysis of the protein and mRNA in RMS13 cell line (Fig. 4e–l). In every instance, we observed exon exclusion by targeted disruption of a putative ESE.

When all the clones presented so far from both commercial and de novo engineered were considered with respect to predicted impact on an ESE and exon exclusion, we observed a strong correlation between these two events (Fig. 4m; Supplementary Fig. 8). A subset of the clones exhibiting alternative splicing also expressed novel polypeptides (see TOP1 and SIRT1; Fig. 1b). We noted in both cases that the exons were symmetric—meaning the exon harbors a nucleotide number in multiples of three, and exclusion of this exon would result in a transcript that retains the original reading frame. In the case of the SUFU clones, the majority of exons skipped were asymmetric thus likely resulting in the lack of protein expression. However, we noted one targeted and skipped exon (exon 2) was symmetric yet the resulting transcript failed to generate a detectable protein perhaps due to misfolding of the mutant protein (Fig. 4e, f). Indeed, the skipped exon encodes part of an intrinsically disordered region of the protein that is essential for interaction with members of the pro-survival BCL2 family members18. From these SUFU clones, we expect that decreased SUFU mRNA seen in CRISPR-edited cells was due to NMD provoked by the introduction of a frameshift-inducing INDEL, or exclusion of the targeted asymmetric exon and the introduction of a PTC in an NMD-enabling position within the gene.

CRISPINator. Purposeful disruption of ESEs in asymmetric exons could improve gene knockout efficiency given that even INDELs that fail to alter the coding frame would have a second opportunity for introducing a PTC by skipping the exon altogether. In addition to the evidence provided here, the ability of mutations in ESEs to alter mRNA splicing have been documented

- **Fig. 6.** A subset of the clones exhibiting alternative splicing also expressed novel polypeptides (see TOP1 and SIRT1; Fig. 1b). We noted in both cases that the exons were symmetric—meaning the exon harbors a nucleotide number in multiples of three, and exclusion of this exon would result in a transcript that retains the original reading frame. In the case of the SUFU clones, the majority of exons skipped were asymmetric thus likely resulting in the lack of protein expression. However, we noted one targeted and skipped exon (exon 2) was symmetric yet the resulting transcript failed to generate a detectable protein perhaps due to misfolding of the mutant protein (Fig. 4e, f). Indeed, the skipped exon encodes part of an intrinsically disordered region of the protein that is essential for interaction with members of the pro-survival BCL2 family members18. From these SUFU clones, we expect that decreased SUFU mRNA seen in CRISPR-edited cells was due to NMD provoked by the introduction of a frameshift-inducing INDEL, or exclusion of the targeted asymmetric exon and the introduction of a PTC in an NMD-enabling position within the gene.

**CRISPINator.** Purposeful disruption of ESEs in asymmetric exons could improve gene knockout efficiency given that even INDELs that fail to alter the coding frame would have a second opportunity for introducing a PTC by skipping the exon altogether. In addition to the evidence provided here, the ability of mutations in ESEs to alter mRNA splicing have been documented
To systematize this strategy, we developed the CRISPinatoR, a website that identifies asymmetric exons found in a given gene and CRISPR-Cas9 guide sequences that help to deliver double-stranded breaks within proximity of a putative ESE (Fig. 5a, Supplementary Fig. 9). At the same time, the portal could be used to induce the skipping of an exon harboring a deleterious mutation in order to generate a novel protein that may retain function. We note that when analyzing genome-wide CRISPR libraries, that ratio of guides targeting symmetric and asymmetric exons was fairly consistent, suggesting that these algorithms do not factor in potential gene elimination efficiency based on exon symmetry (Supplementary Fig. 10A, B). Similarly, the CRISPinatoR could be used to re-evaluate previously reported phenotypes using CRISPR-Cas9 based on the potential for the sgRNA for inducing exon skipping.

Targeting RNA-regulatory elements for gene knockout agendas. We tested the ability of CRISPinatoR to design guides that induce exon skipping for either degradation of mRNA or production of novel protein-encoding mRNAs by targeting asymmetric or symmetric exons, respectively. Using the WNT receptor LRP5 as a case study, we asked the CRISPinatoR to identify sgRNAs that presumably would be able to induce exon skipping.
in each exon class (Fig. 5b). We identified clones that harbored INDELS at the anticipated LRP5 exon targeted by sequences of isolated genomic DNA (Fig. 5c). Using RT-PCR analysis coupled with targeted sequencing, we observed exon skipping in clones associated with both guides (Fig. 5d; Supplementary Fig. 11). We observed an absence of LRP5 protein in the clone exhibiting exclusion of an asymmetric exon (Fig. 5e). However, in the clone exhibiting exclusion of a symmetric exon, we observed the appearance of a faster migrating protein (Fig. 5e).

We confirmed that this new protein retains glycosyl motifs, suggesting that its signal sequence localized to the N-terminus is intact unlike in the case of the LRP6 edited HAP1 clone (Fig. 5f; Supplementary Fig. 1A, C). The presence of a secreted protein and evidence for skipping of the CRISPR-targeted exon suggest that the novel LRP5 protein formed would harbor a compromised β-propeller domain—one of two that contributes to WNT3A binding (Fig. 5g). Indeed, we observed response of a clone expressing the truncated LRP5 protein to exogenously supplied WNT conditioned medium using a WNT pathway reporter (Fig. 5h). The weakened response compared to WT HAP1 cells likely reflects reduced total LRP5 protein levels and/or reduced WNT-binding affinity with deletion of exon 16 sequence. On the other hand, the cell expressing the LRP5 mRNA excluding the CRISPR-edited asymmetric exon showed a loss of WNT pathway response consistent with the absence of LRP5 protein production from an mRNA lacking an asymmetric exon (Fig. 5h).

A number of considerations in guide design could be installed in our design workflow to increase the fidelity of DNA sequencing information for predicting protein translational outcomes. A map of RNA-regulatory motifs (such as ESEs) that might be impacted by a CRISPR-Cas9-delivered INDEL such as that generated by the CRISPinatoR for the human genome could help in improving gene elimination or protein engineering campaigns. We acknowledge that the impact of RNA structure and possibly other determinants that can influence the function of regulatory sequences involved in RNA splicing, for example, are not accounted for by our database. At the same time, an understanding of lineage-associated pseudo-transcripts that would be edited alongside the intended target transcripts would also help to anticipate the emergence of novel protein products such as Super LKB1 from conversion of a pseudo-mRNA to a protein-encoding mRNA.

Perhaps the most daunting challenge that we encountered from our analysis of CRISPR-edited cell lines is the emergence of IRESs likely due to INDEL-induced changes in RNA structure. We anticipate that the number of ATI events associated with INDELS will be higher than what is reported here given the shortage of antibodies useful for detecting native as well as potentially truncated proteins that emerge from ATI. In this regard, the use of translation inhibitors such as CHX combined with RT-PCR could be a simple method to flag mRNAs that harbor CRISPR-Cas9-introduced frameshift-inducing INDELS yet for reasons including ATI subversion are not substrates for NMD. Although we have attempted to account for the ATI events we observed in our LKB1 gene editing projects using an in silico RNA structure prediction strategy, admittedly other factors such as potential changes in RNA-binding protein interactions could contribute to alterations in translation initiation sites.

Our observations also have implications for the use of INDEL-based genome editing tools for gene rescue efforts where induced exon skipping can excise sequences that harbor a mutation thus producing a viable gene. These outcomes are currently achieved by using two CRISPR guides that flank a mutated exon or target an exon-specific splice junction using a single guide. However, the ability to use a single guide targeting an ESE to achieve a similar outcome should reduce the dangers of using two CRISPR guides and expand the number of single guide options with acceptable off-target risks. In this regard, guides identified by the CRISPinatoR targeting ESEs found in symmetric exons could be used to systematically identify such opportunities in genes involved in disease. Needless to say, mRNA splicing is a complex phenomenon and this approach should serve as a starting point.
Fig. 5 Targeting RNA-regulatory elements for gene knockout agendas. a CRISPinatoR: a web-based guide RNA design tool that utilizes targeted ESE disruption for achieving gene elimination. CRISPinatoR identifies sgRNA sequences that target ESEs in asymmetric exons and calculates off-targeting potential and the number of splice variants impacted by the sgRNAs. b A scoring system that integrates all three parameters is used to provide sgRNAs with high gene knockout potential and the number of splice variants impacted by the sgRNAs. c Genomic sequencing results of HAP1 clones edited using cRNA and scRNA sequences targeting the asymmetric exon 2 and the symmetric exon 16. d Targeted ESE disruption in asymmetric exon increases gene knockout potential. Western blot analysis of HAP1 clones edited with LRPS exon 2 and exon 16 sgRNAs. e Exclusion of an asymmetric or a symmetric exon with INDEL-induced changes to the putative ESE sequences. RT-PCR analysis and cDNA sequencing result of HAP1 cells edited with LRPS exon 2 sgRNA (Clone 21) and exon 16 sgRNA (Clone 3) was probed with two distinct antibodies indicated in ‘b’. f CRISPinatoR: a web-based guide RNA design tool that utilizes targeted ESE disruption for achieving gene elimination. g Western blot analysis of WT and LRP5ΔE16 HAP1 cells were incubated with the deglycosidase PNGase F then subjected to western blot analysis. h Trafficking of WNT and EGF-like domain protein is glycosylated. Lysates derived from WT or LRPS ΔE16 HAP1 cells were incubated with the deglycosidase PNGase F then subjected to western blot analysis. i Exclusion of a symmetric exon with INDEL-induced changes to the putative ESE sequences. RT-PCR analysis and cDNA sequencing result of HAP1 cells edited with LRPS exon 2 sgRNA (Clone 21) and exon 16 sgRNA (Clone 3) was probed with two distinct antibodies indicated in ‘b’. j Exclusion of an asymmetric or a symmetric exon with INDEL-induced changes to the putative ESE sequences. RT-PCR analysis and cDNA sequencing result of HAP1 cells edited with LRPS exon 2 sgRNA (Clone 21) and exon 16 sgRNA (Clone 3) was probed with two distinct antibodies indicated in ‘b’.

Materials and methods

Cell lines and reagents. WT and CRISPR-edited HAP1 knockout commercial cell lines were purchased from Horizon Discovery (Supplemental Table 1). HELa, MIA PaCa-2, and RMS13 cell lines were purchased from ATCC. Hela cells (listed in the database of commonly misidentified cell lines, ICLAC) lack endogenous LKB1 expression and therefore was used in an experiment to monitor the protein expression encoded LKB1 cDNA constructs harboring CRISPR-introduced INDELs. All the cell lines were tested for mycoplasma contamination. None of the cell lines were authenticated since all the cell lines were directly purchased from commercial suppliers.
Fig. 6 Cellular mechanisms for countering INDEL effects revealed by CRISPR failures. INDELS introduced by Cas9 and other enzymes used for gene editing elicit transcriptional and translational responses that may have evolved to buffer the transcriptome and proteome against common environmental insults.

Puromycin was purchased from Fisher Scientific (ICN0055225). Cytochalasin D was purchased from EMD Millipore (239765). NE-PER Nuclear and Cytoplasmic extraction reagent (78833) was purchased from ThermoFisher.

**Western blot analysis.** Cell lysates were generated with PBS/1% NP40 buffer supplemented with protease inhibitor cocktail (Sigma Cat. No. S8820). Protein sample loading buffer was added to cell lysates and proteins were separated on SDS-PAGE (BioRad Criterion TGX Precast Gel). The following primary antibodies were used for immunoblotting at the indicated dilutions: Cell Signaling Technology: CTNNB1 (610153; 1:1000), Genetex: VPS35 (GTC108058; 1:1000), CTNNB1 (C2206; 1:1000). Invitrogen: PTEN (44-1064; 1:1000). BD Biosciences: PPM1A (LS-C169090-100; 1:1000), PTPN11 (A302-589A and 302-590A; 1:1000), TOP1 (A302-243A-T; 1:1000), RICTOR (A300-459A; 1:1000), SIRT1 (A300-688A; 1:1000), SUFU (2520; 1:1000), LRP5 (5731; 1:1000). Bethyl Laboratories: BAP1 (A303-590A; 1:1000), LRP6 (2560; 1:1000), PTEN (9552; 1:1000), RICTOR (9476; 1:1000), SIRT1 (A300-688A; 1:1000), TOPI (A302-589A and 302-590A; 1:1000), VPS35 (A304-727A; 1:1000). Sigma-Aldrich: NUCLEAR and CYTOPSOMAL EXTRACT REAGENTS (ICN1005522). The target exon for each sgRNA and the symmetry of the target exon was identified using genomic annotation from Ensembi38. The number of ESEs within a given 20 bp sgRNA sequence was annotated. To score the off-target potential for each sgRNA candidate, we modified the bwa source code39. The total 23 bp sequence (20 bp sgRNA + 3 bp PAM) was aligned to the hg19 reference genome allowing up to three mismatches. The variable bp in the 5’ most position of the PAM sequence was not considered for mismatch scoring. An off-target score (0–100) for the sgRNA was calculated by a method used in Hsu et al.19 and a cutoff of 80 was considered acceptable.

**Data availability** The data supporting the findings of this study are available from the corresponding author upon request. The source data containing images of uncropped blots used in this paper are provided as a Source Data file.

Received: 6 May 2019 Accepted: 14 August 2019
Published online: 06 September 2019

**References**

1. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278 (2014).
2. Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering. *Science* 346, 1234315 (2014).
3. Garneau, N. L., Wilusz, J. & Wilusz, C. J. The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* 8, 113–126 (2007).
4. Lykke-Andersen, S. & Jensen, T. H. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat. Rev. Mol. Cell Biol.* 16, 665–677 (2015).
5. Maier, K. et al. Single amino acid deletion in Kindlin-1 results in partial protein degradation which can be rescued by chaperone treatment. *J. Invest. Dermatol.* 136, 920–929 (2016).
6. Dominguez, D. et al. Sequence, structure, and context preferences of human RNA binding proteins. *Mol. Cell* 70, 854–867.e9 (2018).
7. Julien, P., Minana, B., Baeza-Centurion, P., Vlachos, J. & Lehner, B. The complete local genotype-phenotype landscape for the alternative splicing of a human exon. *Nat. Commun.* 7, 11558 (2016).
8. Ray, D. et al. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499, 172–177 (2013).
9. Frith, M. C. et al. Pseudo-messenger RNA: phantoms of the transcriptome. *PloS Genet.* 2, e23 (2006).
10. Blomen, V. A. et al. Gene essentiality and synthetic lethality in haploid human cells. *Science* **350**, 1092–1096 (2015).

11. Ke, S. & Chasin, L. A. Context-dependent splicing regulation: exon definition, co-occurring motif pairs and tissue specificity. *RNA Biol.* **8**, 384–388 (2011).

12. Sterne-Weiler, T. & Sanford, J. R. Exon identity crisis: disease-causing mutations that disrupt the splicing code. *Genome Biol.* **15**, 201 (2014).

13. Wang, J. et al. DNA topoisomerases. *Annu. Rev. Biochem.* **75**, 359–365 (2006). 1996.

14. Wang, T. et al. Identification and characterization of essential genes in the human genome. *Science* **350**, 1096–1101 (2015).

15. Lindeboom, R. G., Supek, F. & Lehner, B. The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat. Genet.* **48**, 1112–1118 (2016).

16. Sriram A., Bohlen J. & Teleman A. A. Translation acrobatics: how cancer cells exploit alternate modes of translational initiation. *EMBO Rep.* **19**, e45947 (2018).

17. Cartegni, L., Chew, S. L. & Kainer, A. R. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.* **3**, 285–298 (2002).

18. Wu, X. et al. Extra-mitochondrial prosurvival BCL-2 proteins regulate gene transcription by inhibiting the SUFU tumour suppressor. *Nat. Cell Biol.* **19**, 1226–1236 (2017).

19. Baralle, M. et al. NF1 mRNA biogenesis: effect of the genomic milieu in splicing regulation of the NF1 exon 37 region. *FEBS Lett.* **580**, 4449–4456 (2006).

20. Zatkins, A. et al. Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. *Hum. Mutat.* **24**, 491–501 (2004).

21. Echeverri, C. J. et al. Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat. Methods* **3**, 777–779 (2006).

22. Wu, X. et al. Extra-mitochondrial prosurvival BCL-2 proteins regulate gene transcription by inhibiting the SUFU tumour suppressor. *Nat. Cell Biol.* **19**, 1226–1236 (2017).

23. Mou, H. et al. CRISPR/Cas9-mediated genome editing induces exon skipping by alternative splicing or exon deletion. *Genome Biol.* **18**, 108 (2017).

24. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).

25. Ousterout, D. G. et al. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat. Commun.* **6**, 6244 (2015).

26. Nelson, C. E. et al. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403–407 (2016).

27. Tabeobdar, M. et al. In vivo genome editing in dystrophic mouse muscle and muscle stem cells. *Science* **351**, 407–411 (2016).

28. Long, C. et al. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400–403 (2016).

29. Long, C. et al. Correction of diverse muscular dystrophy mutations in human engineered heart muscle by single-site genome editing. *Sci. Adv.* **4**, eaap9004 (2018).

30. Tan, Z., Fu, Y., Sharma, G. & Mathews, D. H. TurbFold II: RNA structural alignment and secondary structure prediction informed by multiple homologs. *Nucleic Acids Res.* **45**, 11570–11581 (2017).

31. Ke, S. et al. Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Res.* **21**, 1360–1374 (2011).

32. Fairbrother, W. G. et al. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res.* **32**, W187–W190 (2004).

33. Zerbino, D. R. et al. Ensembl 2018. *Nucleic Acids Res.* **46**, D754–D761 (2018).

34. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transforms. *Bioinformatics* **26**, 589–595 (2010).

35. Hsiu, P. D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).