Guidelines for minimal reporting requirements, design and interpretation of experiments involving the use of eukaryotic dual gene expression reporters (MINDR)

Gary Loughran, Dmitry E Andreev, Ilya M Terenin, Olivier Namy, Martin Mikl, Martina M Yordanova, C Joel Mcmanus, Andrew E Firth, John F Atkins, Christopher S Fraser, et al.

Guidelines for minimal reporting requirements, design and interpretation of experiments involving the use of eukaryotic dual gene expression reporters (MINDR).

Gary Loughran[1]*, Dmitry E. Andreev[2,3], Ilya M. Terenin[2], Olivier Namy[4], Martin Mikl[5], Martina M. Yordanova[1], C. Joel McManus[6], Andrew E. Firth[7], John F. Atkins[1], Christopher S Fraser[8], Zoya Ignatova[9], Shintaro Iwasaki[10,11], Joanna Kufel[12], Ola Larson[13], Sebastian A. Leidel[14], Alexander S. Mankin[15,16], Marco Mariotti[17], Marvin E. Tanenbaum[18,19], Ivan Topsirovic[20,21,22,23], Nora Vázquez-Laslop[15], Gabriela Viero[24], Neva Caliskan[25,26], Yiwen Chen[27], Patricia L. Clark[28], Jonathan D. Dinman[29], Philip J. Farabaugh[30], Wendy V. Gilbert[31], Pavel Ivanov[32,33], Jeffrey S. Kieft[34,35,36], Oliver Mühlemann[14], Matthew S. Sachs[37], Ivan N. Shatsky[2], Nahum Sonenberg[22], Anna-Lena Steckelberg[38], Anne E. Willis[39], Michael T. Woodside[40], Leos Shivaya Valasek[41]*, Sergey E. Dmitriev[2]*, Pavel V. Baranov[1]*.

[1]School of Biochemistry and Cell Biology, University College Cork, Ireland
[2]Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, 119234, Russia.
[3]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, RAS, Moscow, Russia.
[4]Institute for Integrative Biology of the Cell, Université Paris-Saclay, CEA, CNRS, 91198 Gif-sur-Yvette, France.
[5]Department of Human Biology, University of Haifa, Haifa, Israel
[6]Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, USA
[7]Division of Virology, Department of Pathology, Addenbrookes Hospital, University of Cambridge, Cambridge, UK.
[8]Department of Molecular and Cellular Biology, College of Biological Sciences, University of California, Davis, California, USA
[9]Institute of Biochemistry and Molecular Biology, University of Hamburg, Hamburg, Germany
[10]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan.
[11]RNA Systems Biochemistry Laboratory, RIKEN Cluster for Pioneering Research, Wako, Japan.
[12]Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, Pawinskiego 5a, Warsaw, 02-106, Poland.
[13]Department of Oncology-Pathology, Science for Life Laboratory, Karolinska Institute, 171 77, Stockholm, Sweden.
[14]Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, 3012 Bern, Switzerland.
[15]Department of Pharmaceutical Sciences, College of Pharmacy, University of Illinois Chicago, Chicago, Illinois 60612, United States.
[16]Center for Biomolecular Sciences, University of Illinois Chicago, Chicago, Illinois 60612, United States.
[17]Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Avinguda Diagonal 643, Barcelona, 08028, Catalonia, Spain.
[18]Oncode Institute, Hubrecht Institute-KNAW and University Medical Center Utrecht, Utrecht, the Netherlands.
[19]Department of Bionanoscience, Delft University of Technology, Delft, the Netherlands.
[20]Lady Davis Institute, McGill University, Montreal, Canada.

[21]Gerald Bronfman Department of Oncology, McGill University, Montréal, QC, Canada.
[22]Department of Biochemistry, McGill University, Montréal, QC, Canada.
[23]Division of Experimental Medicine, McGill University, Montréal, QC, Canada.
[24]Institute of Biophysics, CNR Unit at Trento, Povo, Italy.
[25]Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), Würzburg, Germany.
[26]Helmholtz Institute Faculty of Medicine, University of Würzburg, Würzburg Germany
[27]Department of Bioinformatics and Computational Biology, Division of Quantitative Sciences, The University of Texas MD Anderson Cancer Center, Houston, TX
[28]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana, USA.
[29]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD, 20742, USA.
[30]Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, Maryland, USA.
[31]Yale School of Medicine, Department of Molecular Biophysics & Biochemistry, New, Haven, CT, United States.
[32]Division of Rheumatology, Inflammation and Immunity, Brigham and Women's Hospital, Boston, MA 02115, USA.
[33]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA.
[34]Department of Biochemistry and Molecular Genetics, Aurora, CO 80045, USA.
[35]RNA BioScience Initiative, University of Colorado Denver School of Medicine, Aurora, CO 80045, USA.
[36]New York Structural Biology Center, New York, NY 10027, USA.
[37]Department of Biology, Texas A&M University, College Station, Texas 77843, USA.
[38]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA.
[39]MRC Toxicology Unit, University of Cambridge, Cambridge CB2 1QR, UK.
[40]Department of Physics, Li Ka Shing Institute of Virology, and Centre for Prions and Protein Folding Diseases, University of Alberta, Canada.
[41]Laboratory of Regulation of Gene Expression, Institute of Microbiology of the Czech Academy of Sciences, Videnska 1083, 142 20, Prague, the Czech Republic

* - corresponding authors: g.loughran@ucc.ie; valasekl@biomed.cas.cz; sergey.dmitriev@belozersky.msu.ru; p.baranov@ucc.ie

Abstract

Dual reporters, which encode two distinct proteins within the same mRNA, have played a crucial role in identifying and characterizing new instances of unconventional eukaryotic translation mechanisms. These mechanisms include initiation by internal ribosomal entry sites (IRESs), ribosomal frameshifting, stop codon readthrough, and reinitiation. This design allows one reporter's expression to be influenced by the specific mechanism under investigation, while the other reporter serves as an internal control. However, challenges arise when intervening sequences are placed between these two reporters. Such sequences can inadvertently impact the expression or function of either reporter, independently of translation-related changes. These effects may occur because of cryptic elements inducing or affecting transcription initiation, splicing, polyadenylation, and antisense transcription, as well as unexpected effects of the translated test sequences on the stability and activity of the reporters. Unfortunately, these unintended effects may lead to incorrect conclusions being published in the scientific literature. To address this issue and assist the scientific community in accurately interpreting dual reporter experiments, we have developed comprehensive guidelines. These guidelines cover experiment design, interpretation, and the minimal requirements for reporting results. They are designed to aid researchers conducting these experiments, as well as reviewers, editors, and other investigators who seek to evaluate published data.

Background

Fusions of genes expressing reporters have been used for characterizing mRNA translation mechanisms since the 1990s[1–5]. The popularity of this approach increased with the invention of the dual luciferase reporter in 1998[6]. Since then, dual reporters have been used to study ribosomal frameshifting, stop codon readthrough, reinitiation, and internal initiation[6–12]. The principle is based on encoding two reporters within the same mRNA so that expression of one reporter is dependent on the studied mechanism while the other reporter is used for normalization, i.e. as an internal control (Figure 1). As reporter proteins are synthesized from the same mRNA, it is assumed that any differences in the expression of the reporter used to characterize the mechanism under study is independent of confounding variables, e.g. transfection efficiency or RNA stability. Dual reporter systems have proven instrumental, in combination with site-directed or random mutagenesis, in characterizing mRNA features involved in specific translation mechanisms[11,13–23] and for the identification of underlying cellular factors impacting these mechanisms[24–32]. Dual reporters have also found application in the identification and study of drugs targeting specific mechanisms of translation[25,33–39] as well as the mechanisms of disease associated polymorphisms[40]. Increasingly, dual reporters are being used for discovery of novel translational mechanisms and processes[41,42].

However, differences in reporter expression can occur for several reasons other than differential translation, potentially leading to false interpretation of reporter readout (Figure 1). Firstly, DNA-encoded dual reporters may contain cryptic promoters, cryptic splice sites, or cryptic polyadenylation signals that could generate unexpected mRNA transcripts encoding only one of the two reporters[11,43–54]. Plasmid DNA can also produce unexpected antisense transcripts that may affect reporter expression[55]. Secondly, the protein extensions encoded by the test sequence may alter the stability or activity of one or both reporters if it is synthesized as a part of the same polypeptide chain[12,56]. Thirdly, in certain applications, the downstream reporter is placed under a known (control) or a putative (test) IRES. However, IRES activity may be influenced by surrounding sequences[20,21,57,58]; it has also been shown that the presence of an IRES may influence mRNA stability[59] and translation[60,61] of the upstream reporters in a sequence dependent manner.

Although various artefacts generated by experiments involving dual reporters have been discussed extensively[11,12,46–50,52,57,62–70], the misinterpretation of these assays continues to result in inaccurate conclusions[53,54,56].

To address this issue, we call for a community effort to create and maintain guidelines for the design, interpretation, and reporting of experiments involving dual reporters encoded within the same mRNA. This document consists of two parts, each with a specific purpose. The first part, Principles and Recommendations (PR), covers the design and interpretation of dual reporter experiments. We provide guidance on the selection of appropriate vectors to reduce risks of false positive findings and design of critical controls and additional experiments that may help to identify and avoid erroneous conclusions. These guidelines are tailored for a range of mechanisms and are organized into subsections accordingly. The second part, Minimal INformation on Dual expression Reporters (MINDR), briefly outlines our proposed reporting requirements, which establish the essential data that should accompany any publication describing dual reporter experiments. MINDR is designed to facilitate reproducibility and equip editors, reviewers, and readers with the information necessary for evaluating the reliability of the employed dual reporter strategies and assessment of the likelihood of false positive findings.

We expect these guidelines to evolve in line with the emergence of novel reporters and their applications during characterization of established and hitherto undiscovered mechanisms of regulation of mRNA translation.

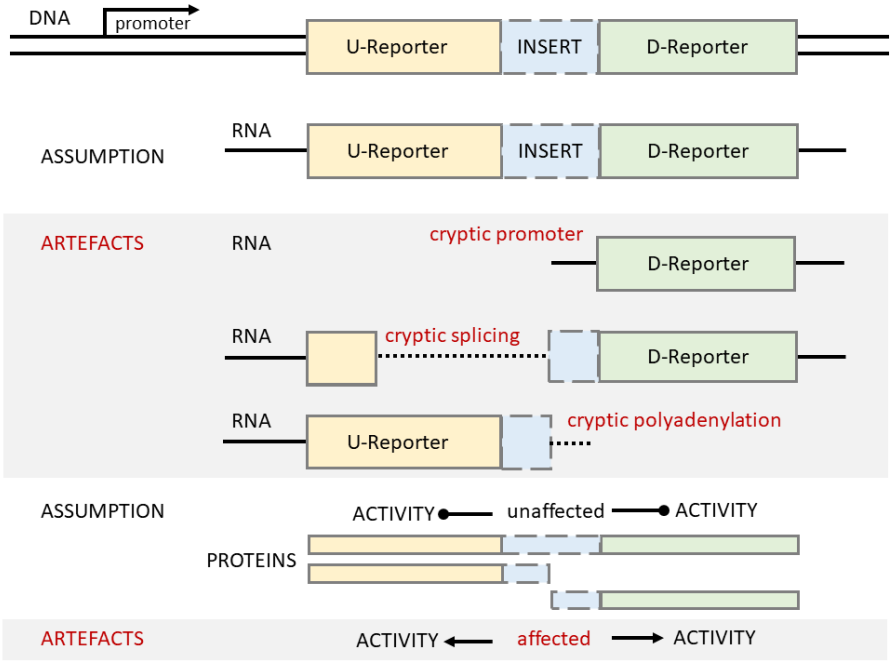

Figure 1. Principles of a dual reporter strategy and the most common artefacts (shaded background) arising from unexamined assumptions.

PR: Principles and Recommendations

In the following subsections we consider reporter design, possible controls, and data interpretation. These are intended to equip researchers with the means to investigate and avoid potential artefacts. We first describe general aspects regarding most dual reporter applications followed by considerations relevant to specific cases.

Cryptic promoters and splicing

A key assumption that does not always hold true is that all measured reporter activity is derived from translation of a single RNA species. As cryptic splicing and cryptic transcription may occur as part of the general nuclear events (Figure 1), the most straightforward way to control for these possibilities is to introduce *in vitro* transcribed RNA encoding the reporters instead of DNA-based reporters. Any discrepancies between DNA and RNA ratio readouts may indicate the existence of unexpected cryptic transcripts, although there may be other factors explaining these differences as discussed below. There may also be situations when RNA transfections are impractical as described in the RNA vs DNA Reporters section below.

For DNA transfections using reporters that are expected to express fusion proteins (e.g. by readthrough or frameshifting), western blotting is an effective and sensitive approach to ensure that downstream reporter expression is not derived from aberrant mRNA species producing shorter variants of fusion proteins.

Where dual reporters are not expected to be fused (e.g. in StopGo vectors and reinitiation or IRES studies), it may be desirable to identify possible aberrant mRNA species directly. Because a minor mRNA species could be highly translated, highly sensitive techniques for detecting mRNA are required. To this end, emerging long read sequencing technologies (such as nanopore or PacBio) or RT-qPCR/ddPCR may be sufficiently sensitive approaches to detect minor RNA species whereas

northern blotting is unlikely to be sensitive enough. RT-qPCR can be used for assessing the levels of different RNA segments[46] with specific primers as shown in Figure 2. Careful design and validation of RT-qPCR amplicons with *in vitro* transcript standards and melting curves is necessary to determine whether the amounts of each amplicon are equal[71]. As the same PCR amplicons can be derived from alternative transcripts and because RT-qPCR optimization is not always straightforward it is advisable that this method is not the sole approach used to analyze the occurrence of transcript isoforms[68]. An *in vitro* transcribed full-length control RNA can be used to assess the signal expected from a single RNA species containing both reporters (Figure 2).

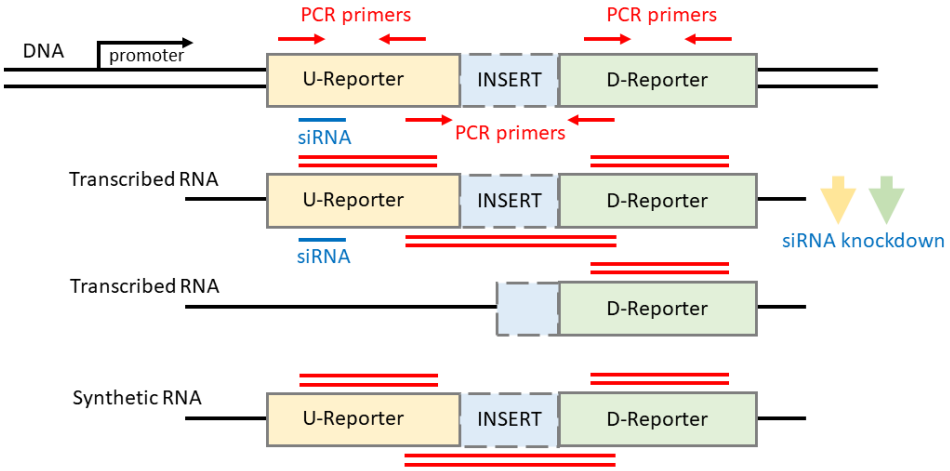

Figure 2. Assessment of dual reporter transcriptional integrity. Red arrows show the design of forward and reverse primers for the dual reporter construct and double red lines depict the resulting cDNA products. Using a synthetic RNA construct (bottom) it is possible to estimate the ratio between cDNA products for the transcript containing all three regions. Horizontal blue lines represent siRNAs for knockdown experiments expected to downregulate expression of both reporters if they are produced from the full-length RNA template as indicated with vertical yellow and green arrows.

Furthermore, an siRNA probe designed to knockdown expression of the upstream reporter is expected to have the same effect on the downstream reporter if they belong to the same RNA and no other transcripts are present (Figure 2). Transcript isoforms could also be explored by methods like 5'-CAGE[72] or with direct RNA sequencing with nanopore technologies, bearing in mind that minor alternative products can produce most of the reporter activity.

Potential existence of aberrant transcripts may also be explored with study-specific controls, the design of which is specific to study phenomena.

Reporters translated as fusion proteins may have altered activity or stability

A second key assumption is that reporter activities accurately reflect reporter translation. However, fusion of the test sequence can influence the stability or activity of one or both reporters (Figure 1). One solution to mitigate this is to use StopGo/2A sequences on both sides of the test sequence[12]. StopGo is a term used to describe the failure of the ribosome to form a peptide bond during the peptidyl transferase reaction without terminating protein synthesis[73]. A peptide motif enabling StopGo was initially discovered during translation of the 2A region of foot-and-mouth disease virus[74], and since then a number of 2A peptides with varying efficiencies have been described. The outcome of this phenomenon is the production of two peptide chains (from the same ORF) separated by the StopGo motif (Figure 3). While in theory StopGo use should produce reporters with identical amino acid sequences irrespective of the insert, this may not always be the case. StopGo is not 100% efficient (generally 80-90%[75]), therefore a certain amount of fusion between either of the reporters and the test sequence-encoded fragment will still be produced. The potential for ribosome drop-off

during StopGo may influence the ratio between two reporters although this can be tested with an appropriate in-frame control construct. When using a new StopGo-containing plasmid it is important to monitor the efficiency of StopGo by western blotting in the intended biological system. Although variability in StopGo activity has been reported[76], this is more likely to occur when a relatively short (<20 aa) StopGo motif is used making it vulnerable to the influence of proximal sequences, including test inserts. It is therefore recommended to use longer (>30 aa) StopGo motifs, which have higher activities and negligible ribosome drop-off[75]. Another important point is that although StopGo is active in all eukaryotes tested so far, different 2A peptide variants have different activities in various organisms, and none have been found to function in bacteria [77]. Furthermore, the kinetics of StopGo is still poorly understood and it is likely to cause some ribosome pausing, which may interfere with some of the studied phenomena.

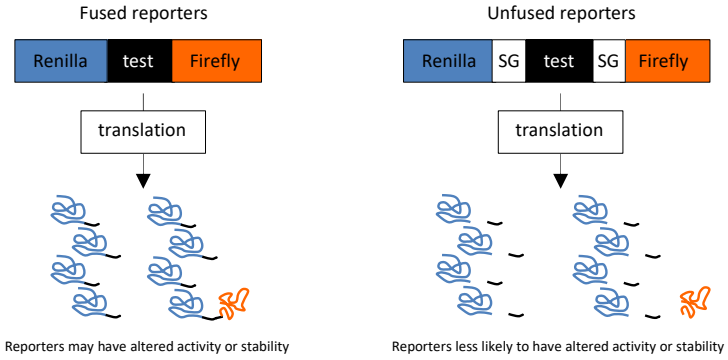

Figure 3 Fused versus unfused dual luciferase reporters. In this example, the test sequence promotes approximately 12.5% recoding. Both systems have their merits and drawbacks. The advantage of fused reporters is the ability to estimate recoding efficiencies by western blotting or *in vitro* translation reactions while controlling for unintended firefly translation by internal initiation or cryptic splicing. One potential downside of fused reporters is erroneous estimations of recoding efficiencies based solely on dual luciferase assays because of possible changes in firefly activity resulting from its fusion to the test and Renilla polypeptides. This can be mitigated with unfused reporter systems where the test peptide sequence is released from the reporters by flanking StopGo (SG) motifs. While western blotting or *in vitro* translations can still be used to estimate recoding efficiencies with unfused reporter systems, due to the similar sized firefly protein produced, spurious firefly expression by internal initiation or cryptic splicing may not be as obvious.

The RNA vs DNA reporters

The use of mRNA transfection is a powerful strategy to avoid the generation of unexpected RNA species by cryptic transcription and/or splicing events. However, its advantages are not limited to this. For several applications, RNA transfection may be preferred irrespective of transcriptional artefacts. Analysis of the immediate stress response is particularly difficult when using plasmid DNA reporters. There is little point in applying stress stimuli immediately after DNA transfection as substantial time is required for the accumulation of sufficient mRNA. In contrast, for RNA transfection, the stress stimuli can be applied immediately or shortly before/after the transfection (1-2 hours), here, newly synthesized protein products are responsible for most of the reporter activity. It is also difficult to synchronize the expression of DNA reporters in the entire cell population without synchronizing their cell cycle, as the plasmid only enters the nucleus during mitosis. Furthermore, transfection of non-dividing cells (i.e. matured neurons or cardiac myocytes) with plasmid DNA is highly inefficient.

Nonetheless RNA transfection is not a panacea and should be used appropriately. Firstly, activity values are generally much lower from mRNA transfections as compared to DNA transfections. For this reason, RNA transfections should use reporters with a high signal-to-background ratio such as luciferase reporters. Secondly, RNA is relatively unstable, so it is better to analyze the activity shortly after transfection, ideally in 2-4 hours post transfection. At later time-points than 8-10 hours post transfection, accumulated reporter products start to exceed newly synthesized ones - even more

so than in the case of DNA transfections. It is also important to use a transfection protocol that minimally stresses the cells and to avoid stress stimuli like cell re-plating, electroporation, or plate/dish cooling immediately before or during transfection[78]. Thirdly and most importantly, one should keep in mind that transfected artificial mRNA has not passed through the nucleus and may lack specific and potentially detrimental features including epitranscriptomic marks or associated mRNA-binding proteins of nuclear origin. Furthermore, in primary cells mRNA transfection often triggers innate immune responses, so additional efforts are needed to reduce its activation, like the use of transcripts with 5'-cap1 and modified nucleotides (e.g. m1Ψ in place of U)[79]. It is also important to note that lipocationic transfection impedes the analysis of reporter mRNA stability, as only a minor fraction of transfected mRNA-liposomal complexes is released into the cytosol[80]. Accordingly, total RNA extraction from the transfected cells yields mRNA predominantly not from the cytosolic fraction. Finally, the transfected RNA may still be processed by the cytoplasmic RNA cleaving enzymes, such as IRE1 or RNase L in mammals[81–83], therefore one cannot exclude the possibility that even in the case of mRNA transfection, aberrant RNA species may also be present. A more detailed comparison of DNA and RNA transfections can be found in a dedicated review[84].

Alterations in reporter ratios due to changes in reporters' absolute readouts

When transfecting the same reporter construct into two different cell lines, the readout of absolute reporter values may be quite different, sometimes spanning orders of magnitude. There may be several reasons for this, including differences in the reporter delivery between the cell lines. Reporter RNAs may trigger the innate immune response resulting in global suppression of translation, e.g. through activation of PKR or RIG-I-like receptors[79,85]. Reporter expression in particular cell lines can be compromised at either the transcriptional (e.g. weak promoter activity in DNA reporters) or translational (e.g. low global translation rates) level. Differential reporter activity can be explained by the presence of a cell-specific repressor/activator that acts on the reporter and affects its expression. Other factors to consider are reporter overexpression driven by strong promoters could further exacerbate cell-specific effects by potentially titrating essential translation components. The potential for test sequences to induce ribosome stalling and collisions, that may trigger ribosome quality control pathways, should also be considered[86–90].

When dual reporters are used to compare the efficiency of the studied mechanism across cell lines or under some conditions (including stress versus control, overexpression or depletion of specific factors, treatment with small molecule inhibitors, NMD activation) it is important to understand that differences in global translation may affect the ratio of the measured reporter activities without affecting their real relative activities. This is because any measured activity is a combination of genuine reporter activity with background levels due to biological noise and technical limitations. We can represent the ratio of measured activities as $(\boldsymbol{X}+\boldsymbol{Bx})/(\boldsymbol{Y}+\boldsymbol{By})$, where $\boldsymbol{X}$, $\boldsymbol{Y}$ are *bona fide* activities of the two reporters and $\boldsymbol{Bx}$, $\boldsymbol{By}$ are the corresponding background levels of their measurements. The relationship between background activities and reporter level activities is not linear, i.e. the level of background activity is disproportionately higher relative to genuine reporter activity when genuine activity is lower. Therefore, the measured ratio $(\boldsymbol{X}+\boldsymbol{Bx})/(\boldsymbol{Y}+\boldsymbol{By})$ may not stay constant when $\mathbf{X}/\mathbf{Y}$ stay the same but the absolute readouts of $\boldsymbol{X}+\boldsymbol{Bx}$ and $\boldsymbol{X}+\boldsymbol{By}$ change. Thus, it is critically important to report absolute readouts of reporter and background activities to reveal potential misinterpretations of observed changes in relative activities of the reporters.

The use of dual reporters for massively parallel assays

In recent years, dual reporter vectors have become popular in Massively Parallel Reporter Assays (MPRA) that simultaneously evaluate thousands of test sequences. This is a powerful approach that allows screening of a diverse pool of sequences for specific regulatory properties, e.g. driving internal initiation[41,42], ribosomal frameshifting[34] and other translation mechanisms. It can also be applied to screen a pool of all possible variants of a particular sequence to comprehensively

characterize *cis*-acting regulatory elements.

While many of the general principles that should guide the use of dual reporters outlined above also apply here, the high-throughput nature entails several specific considerations associated with measuring reporter activity in a pooled manner. Certain guidelines such as reporting absolute measured expression levels of the reporter genes is typically harder to achieve than in the case of single reporter measurements. On the other hand, dual reporter-MPRAs also allow for the inclusion of a much larger number of controls that can be measured in the same experiment.

Particularly suitable for MPRA are dual fluorescent reporter constructs, as they allow fluorescence-activated cell sorting based on relative reporter expression levels. This is then followed by DNA sequencing-based identification and quantification of the underlying sequences.

An important consideration for MPRA is to ensure equal vector copy number and similar expression levels among the cells. Utilization of systems for genome integration[91] enables the integration of a single copy of the test vector in the same genetic locus for all cells. However, this may not be strictly required. When this issue was carefully examined recently[92], a high degree of correlation was found between the expression of two cistrons across a polyclonal population of cells, regardless of integration site and number of integrated copies of the bicistronic construct.

A significant challenge in screening large pools of diverse sequences lies in discerning artefacts amidst the positive hits. A prevalent issue when it comes to interpreting data from MPRA is the reliance on validating only a subset of hits to argue that the remainder should be considered true positives. It is crucial to recognize that each individual sequence may possess specific properties, and that in some cases, even a single nucleotide change can induce effects, such as altering cryptic promoter activity.

MPRA-based screens typically yield numerous hits, and it is impractical to individually validate each one. Depending on the experimental effort necessary for individual validation experiments, estimating the rate of false positives can prove difficult[41,42]. In such cases, drawing general conclusions based on a few validated cases should be avoided. Conducting a meta-analysis of positive hit sequences could aid in identifying shared features among some positives, helping to pinpoint certain types of artefacts. When MPRAs are employed to dissect the regulatory code of a gene regulatory mechanism, rare artefacts will likely not affect the main conclusion. However, if an MPRA is used to identify specific “positive” events, conclusions should be restricted to candidates that were appropriately validated in order to minimize the potential for misinterpreting data obtained from MPRA screens.

Cell-free translation systems and other in vitro assays

Another potential source of false positives from dual reporter assays is inappropriate use of *in vitro* translation systems. For example, the commercially available and widely used nuclease-treated Rabbit Reticulocyte Lysate (ntRRL), which is prepared from specialized cells with a limited range of RNA-binding proteins, has been repeatedly shown to inaccurately reproduce conditions found in normal cells[62,69]. Firstly, exogenously added mRNAs translated in ntRRL exhibit a relatively weak reliance on the 5' cap[51,93], although optimizing the buffer conditions can substantially increase cap-dependency and start site recognition[94]. Moreover, as some eIF4G molecules are sequestered by the capped 5'-terminal mRNA fragments remaining in the ntRRL after a limited hydrolysis of endogenous reticulocyte transcripts, the addition of cap-dependent initiation inhibitors (such as m7GTP, 4E-BP, or proteases that cleave eIF4G) may release this factor and artificially stimulate translation of uncapped mRNAs. Then, this system does not recapitulate the cap/poly(A) synergy[95]. Taken together, all this is often misinterpreted as an indication of cap-independent translation of a particular studied mRNA. Secondly, ntRRL is prone to aberrant internal initiation at AUG codons

located within extended unstructured regions[96], causing artificial expression of the second cistron in bicistronic reporters even in the absence of *bona fide* IRESs. Finally, translation in RRL is highly sensitive to even moderately stable RNA secondary structures in the 5' leader[93].

It is likely that similar artefacts can be obtained in cell-free systems derived from budding yeast and wheat germ, at least under some conditions[67,97]. In contrast, such effects are not typically observed in cytosolic extracts of cultured mammalian cells[51,93]. However, in any *in vitro* system, results can greatly depend on the specific preparation conditions and component concentrations. For example, varying polyamine concentrations in the yeast extract can modulate stop-codon readthrough efficiency as much as 4-fold[98]. Therefore, caution is necessary when comparing findings from a specific cell-free system with cultured cells or *in vivo* observations.

Less commonly, bicistronic constructs are used in *in vitro* systems reconstituted from purified components[99,100]. Although such analysis is very informative, it also should be done with caution. Similar to RRL, in these systems, which are usually devoid of mRNA-binding proteins, the ribosomes are able to bind to internal AUG codons located within long unstructured regions[101,102]. This risks confusion of an authentic mechanism with an artificial one that should be excluded by validation in complete *in vitro* or *in vivo* systems.

Another potential source of artefacts specific for *in vitro* assays is related to partial hydrolysis of an *in vitro* synthesized reporter mRNA that can produce truncated versions of bicistronic constructs. This leads to inappropriate ribosome loading to the second cistron on transcripts that do not contain the first cistron. Thus, the extensive analysis of mRNA integrity should be carried out. In some cases (i.e. IRES studies), analysis of polysome-associated mRNA fractions can be helpful. In particular, total RNA isolated separately from monosome and polysome fractions and supplemented with equivalent spike-in RNA would be subject to RT-qPCR using primers covering either each cistron alone or both cistrons together. The resulting ddCq values would be normalized to the spike-in RNA and compared across all samples to determine the integrity of the full-length reporter mRNA.

Specific considerations in the assessment of ribosomal frameshifting

The assessment of ribosomal frameshifting using dual reporters involves a calculation of the relative reporter activities (downstream to upstream ratio). The relative reporter activity of the test construct is then compared to that of a positive control construct. The positive control has both reporters encoded within the same ORF (in-frame control). An in-frame control can be obtained with either insertion or deletion of a single nucleotide in the frameshifting site depending on the direction of frameshifting (-1 or +1). The reporters' ratio in such a construct is considered to correspond to 100% efficient frameshifting. Often a single in-frame control is compared to several test sequences. We caution against this practice; when the sequence of the test constructs is considerably different from the sequence of a single in-frame control, the test sequence may contain cryptic splice sites or promoters. The ideal in-frame control should have no amino acid differences from the expected frameshift product and minimal nucleotide differences. It is also advisable to introduce synonymous changes to disrupt the putative slippery sequence within the in-frame control so that it is a more accurate readout of 100% frameshifting. If the frameshifting site is not disrupted in an in-frame control and frameshifting occurs (say at 10% efficiency), only 90% of the ribosomes would synthesize the downstream reporter. Thus, ideally, each tested sequence should have its own positive in-frame control. Including a +1 or -1 frame termination codon (depending on the reading frame of the downstream reporter) 5' of the slip site is also worth considering to ensure that only frameshifting within the test sequence is reported.

When using fused reporters, orthogonal validation of frameshifting by western blotting is highly desirable when reporting novel instances. Most reporters can be detected by commercially available antibodies that can detect frameshifting efficiencies as low as 1%. As mentioned above, when using

fused reporters, western blotting can also control for cryptic splicing and cryptic promoter activity.

Another important way to validate ribosomal frameshifting is the use of a negative control. Ribosomal frameshifting normally occurs at specific frameshifting sites accompanied with stimulatory elements such as specific mRNA structures. While single point mutations in stimulatory signals rarely abolish frameshifting completely, disruptions of the frameshifting site are expected to eliminate frameshifting as tRNA re-pairing in the new frame is precluded. Constructs containing such disruptions can be used as a negative control. If the high activity of the downstream reporter persists, it would suggest that it may occur due to reasons other than frameshifting.

Although not always possible, it is desirable to avoid AUG codons within the test sequence that are in the same reading frame as the second reporter ORF, as they may serve as initiation codons on cryptic transcripts missing the upstream reporter, entirely or partly. However, a deliberate insertion of an AUG codon in-frame with the second reporter may be used as a control for the existence of cryptic transcripts. In the absence of such transcripts, introduction of an AUG codon in a good Kozak context should not substantially alter the activity of the second reporter unless there is reinitiation. Although translation initiation at near-cognate starts is less efficient, it is known to be highly productive and even comparable to that of AUG in certain contexts, therefore it is advisable to examine the sequences for the presence of such near-cognate start codons in a good Kozak context.

Specific considerations in the assessment of stop codon readthrough

When studying stop codon readthrough a number of considerations are similar to those of ribosomal frameshifting. Western blots are also a very useful orthogonal validation of stop codon readthrough. A positive control is required where a stop codon needs to be replaced with a sense codon, ideally one encoding the same amino acid that is expected to be inserted in place of the readthrough stop codon. However, the identity may not be known in advance and there may be one of several possible amino acids inserted[103,104]. The ratio of both reporters in the test construct is then compared to their ratio from the in-frame control to calculate readthrough efficiency. Like ribosomal frameshifting, each test construct should have its own in-frame control. Usually, a single nucleotide substitution is sufficient.

A negative control would ensure that no activity other than from stop codon readthrough exists. The negative control is usually a construct containing tandem in-frame stop codons (TAA_TAA) that in most organisms represent the most efficient terminators.

Similar to frameshifting, it is important to avoid AUG codons in the test sequence that are in- frame with the second reporter ORF. In addition to potential initiation at AUG codons occurring in the beginning of cryptic transcripts lacking first reporter there is also a potential for reinitiation of terminating ribosomes. Although reinitiation is extremely rare after translation of long ORFs, there are certain signals that could enable reinitiation even after translation of long ORFs[105–108].

Specific considerations in the assessment of reinitiation

Reinitiation is a process in which a ribosome initiates translation downstream of the stop codon at which it terminates (see[109,110] and references therein). Usually, this process is not very efficient in eukaryotes, unless the translated upstream ORF is short or mRNA-specific mechanisms are used[105–109,111]. However, reinitiation can be greatly facilitated under some physiological stress conditions or when ribosome recycling factors are artificially depleted[29,112].

The rate of translation reinitiation after long ORFs can be assessed with the dual reporter assay; other methods are more suitable to examine reinitiation after short uORFs (see[109,110] and references therein). However, due to the inefficiency of this process under normal conditions, absolute values

of reporter activity and appropriate background correction should be thoroughly considered. As in other cases, cryptic promoters in intercistronic spacers must be excluded and only appropriate cell-free systems should be used for *in vitro* studies to exclude false (or true) internal initiation. Moreover, stop-codon readthrough or frameshifting can be erroneously attributed to reinitiation, so these possibilities should also be excluded. Readthrough can be excluded with an additional stop codon between the translated upstream ORF and the putative reinitiation site, which would reduce readthrough, but not reinitiation. Similarly, the possibility of ribosomal frameshifting can be reduced by the insertion or deletion of a nucleotide to disrupt the reading frame and eliminate the possibility of frameshifting without affecting reinitiation (unless the nucleotide indel disrupts a cis-acting signal responsible for reinitiation). Alternatively, western blotting could be used to rule out the possibility of either readthrough or frameshifting.

Specific considerations in the assessment of internal initiation

When testing for IRES-dependent translation, the mRNA sequence that is suspected to promote internal initiation is inserted between two reporters to initiate translation of the downstream reporter and, unlike the cases of frameshifting or stop codon readthrough, no fusion product is expected. Therefore, any possibility of first cistron translation interfering with a putative IRES should be mitigated, e.g., by inclusion of extra stop codons to exclude any readthrough and placing the IRES at a sufficient distance from the stop codon so that terminating ribosomes do not disrupt the IRES structure. However, the risk that incorporation of the studied RNA fragment into an unnatural context may affect its IRES activity should be kept in mind.

Although distant elements affecting stop codon readthrough have been reported[113,114] and may also exist for frameshifting, often only short motifs or structures are tested. Since internal initiation occurs on more elaborate structures, testing longer sequences is essential. Accordingly, all the above issues regarding cryptic promoters and splice sites are of particular importance in the case of IRESs. It is important to keep in mind that almost any randomly chosen long fragment of a mammalian 5' UTR has potential cryptic promoter activity due to the abundance of transcription factor binding sites in these regions. Moreover, a large fraction of human genes possess alternative transcription start sites (TSS)[115] and TSS switching can occur during acute stress[116–118]. Long insertions also increase the risk of artefacts when using *in vitro* systems like ntRRL that are not strongly cap-dependent.

To reveal cryptic splicing (but not cryptic promoter) mediated events, the independence of first and second cistron expression can be verified by designing constructs containing a uORF in the 5' UTR or by the insertion of a stable hairpin at the very 5' end of the transcript to reduce translation of the first cistron. However, this approach requires analysis of the transcript level, as the second reporter may be affected if such modifications alter the stability of the whole mRNA. The insertion of one or more stop codons into the first cistron can also be used, although the effects can be complex and should be interpreted with caution.

In IRES research, the use of mRNA reporters is clearly preferable over DNA reporters[62]. Using *in vitro* transcribed reporters avoids artefacts arising from cryptic promoter activity, unintended splicing or premature transcription termination. When using DNA transfection to assess putative IRES activity, the promoterless and siRNA-mediated controls (see above) are necessary.

Most importantly, however, the bicistronic assay used to study IRES activity has a notable intrinsic limitation that was not initially apparent. Comparing the expression ratios of upstream and downstream cistrons between bicistronic reporters containing different IRESs is not very informative. This is because any two long non-specific arbitrary sequences placed between reporters are highly unlikely to give equal readouts, while different *bona fide* IRESs can also have significantly different activities (Figure 4A). As a result, it is not possible to confidently design

negative or positive controls for such assays.

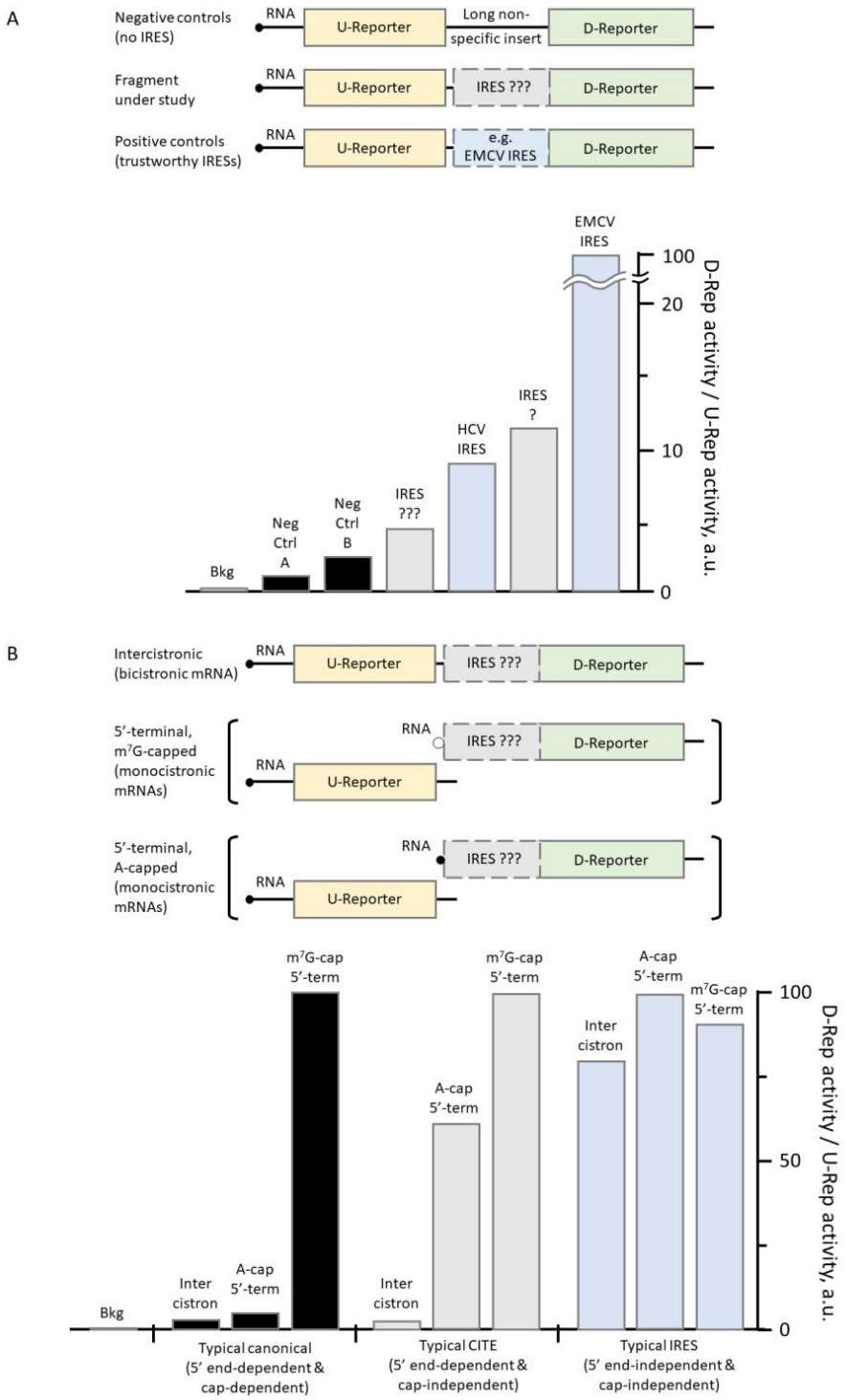

Figure 4. Peculiarities of dual reporter assay in the study of IRES activity. **A.** Simulated results of typical bicistronic assays and their potential interpretation. Unambiguous conclusions cannot be made on the basis of comparing the activities of different bicistronic mRNAs with each other, as control values (both negative and positive) may vary significantly. **B**. A mechanism of ribosome recruitment used by a particular mRNA fragment (for example, a 5' UTR) can be determined using three related dual reporter assays with mRNA constructs shown in the upper subpanel. Three different translation initiation mechanisms can be distinguished: canonical cap-dependent scanning, CITE-directed, or IRES-directed, as indicated below. The bottom panel represents typical simulated results. Bkg, background level of the reporter activity in mock-transfected cells.

Many putative IRES sequences are derived from natural 5' UTRs, which in the case of cellular and many viral mRNAs are naturally capped and therefore can be bound and scanned by the canonical initiation machinery. Therefore, even if the sequence under investigation suggests internal initiation in a bicistronic context, it is not trivial to assess the contribution of internal initiation to overall translation initiation in the natural context. These flaws can be tackled by comparing expression from bicistronic and monocistronic reporters, which allows the assessment of the relative contributions of the different mechanisms (Figure 4B).

Finally, it is important to note that the definition of "IRES" refers to internal ribosome entry and not to cap-independence. Although the latter is a consequence of the former, some mRNAs lacking an IRES and strictly requiring a free 5'-end can nevertheless be efficiently translated in an uncapped (or capped with the artificial non-functional A-cap analog) form. In this case, specific elements called cap-independent translation enhancers (CITE) promote cap-independent translation[62]. In contrast to IRESs, CITEs cannot direct translation of the second cistron in a bicistronic mRNA (Figure 4B). It is likely that many of so-called "cellular IRESs" are either artefacts or CITEs. Opinions on the wide representation of IRESs in 5' UTRs of cellular mRNAs is largely based on the premise of the low processivity of the translation initiation complex, which is, however, significantly underestimated[51,119,120]. This misconception helped form the view that any long 5' UTR must use a non-canonical translation initiation pathway, which is clearly not the case.

MINDR: Minimal INformation on Dual expression Reporters.

Based on the caveats described above, we propose the following three minimal reporting requirements that should mitigate many of these issues and should accompany any study with data obtained with dual reporters:

1. A list of positive and negative controls

As discussed above, the use of appropriate positive and negative controls is critical. However, it may be too impractical to design all the controls described in the previous sections. Nevertheless, the level of confidence in the reported results depends on the specific controls used in an experiment. Therefore, the authors should explicitly describe which positive and negative controls have been used for detecting potential artefacts, to help reviewers and readers assess the reliability of the study.

2. Full sequences of all vectors and inserts

Sequences responsible for potential artefacts, such as transcription enhancers or donor/acceptor splice sites, may be distant from the sequence encoding reporters[52,55]. Therefore, for the reproducibility and future interrogation of reported results, the exact sequence of all plasmids used should be provided.

3. Absolute readout values for each reporter and for the background

The use of dual reporters in general requires the analysis of their ratios rather than absolute values, which are subject to high variability due to technical reasons, such as transfection efficiencies.

Nonetheless, as discussed earlier, consistent differences in absolute values are often indicative of artefacts. Therefore, it is important that in addition to providing the ratios between the reporter activities, the absolute raw readouts should be made available for each replicate of each construct including background values from control experiments without reporter constructs.

ACKNOWLEDGMENTS

The authors wish to acknowledge many colleagues who informally supported and encouraged development of these guidelines. The following authors wish to acknowledge funding support.

MM by grants PID2020-115122GA-I00 from the Ministerio de Ciencia, Innovación y Universidades of Spain. MSS by the NIH (R01GM148702). PVB by the Wellcome Trust (210692/Z/18/]) and Science Foundation Ireland (20/FFP-A/8929). LSV by the Lead Agency (DFG & CSF) grant 23-08669L and the Praemium Academiae grant provided by the Czech Academy of Sciences and CZ.02.01.01/00/22_008/0004575 RNA for therapy by ERDF and MEYS. SED from the Russian Science Foundation (23-14-00218). CSF by the NIH (R01 GM092927). WVG. by the NIH (R01GM132358). YC by the NIH (R01GM130838).

CONFLICTS OF INTERESTS

G.L. and P.V.B. are cofounders and shareholders of EIRNABIO.

References

1. Pelletier, J. & Sonenberg, N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320–325 (1988).
2. Jang, S. K. *et al.* A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol* **62**, 2636–2643 (1988).
3. Reil, H., Höxter, M., Moosmayer, D., Pauli, G. & Hauser, H. CD4 Expressing Human 293 Cells as a Tool for Studies in HIV-1 Replication: The Efficiency of Translational Frameshifting Is Not Altered by HIV-1 Infection. *Virology* **205**, 371–375 (1994).
4. Stahl, G., Bidou, L., Rousset, J.-P. & Cassan, M. Versatile vectors to study recoding: conservation of rules between yeast and mammalian cells. *Nucl Acids Res* **23**, 1557–1560 (1995).
5. Bidou, L. *et al.* In vivo HIV-1 frameshifting efficiency is directly related to the stability of the stem-loop stimulatory signal. *RNA* **3**, 1153–1158 (1997).
6. Grentzmann, G., Ingram, J. A., Kelly, P. J., Gesteland, R. F. & Atkins, J. F. A dual-luciferase reporter system for studying recoding signals. *RNA* **4**, 479–486 (1998).
7. Chiba, S., Jamal, A. & Suzuki, N. First Evidence for Internal Ribosomal Entry Sites in Diverse Fungal Virus Genomes. *mBio* **9**, e02350-17 (2018).
8. Rakauskaite, R., Liao, P.-Y., Rhodin, M. H. J., Lee, K. & Dinman, J. D. A rapid, inexpensive yeast-based dual-fluorescence assay of programmed--1 ribosomal frameshifting for high-throughput screening. *Nucleic Acids Res* **39**, e97 (2011).
9. Cardno, T. S., Poole, E. S., Mathew, S. F., Graves, R. & Tate, W. P. A homogeneous cell-based bicistronic fluorescence assay for high-throughput identification of drugs that perturb viral gene recoding and read-through of nonsense stop codons. *RNA* **15**, 1614–1621 (2009).
10. Harger, J. W. & Dinman, J. D. An in vivo dual-luciferase assay system for studying translational recoding in the yeast Saccharomyces cerevisiae. *RNA* **9**, 1019–1024 (2003).
11. Bert, A. G., Grépin, R., Vadas, M. A. & Goodall, G. J. Assessing IRES activity in the HIF-1alpha and other cellular 5' UTRs. *RNA* **12**, 1074–1083 (2006).

12. Loughran, G., Howard, M. T., Firth, A. E. & Atkins, J. F. Avoidance of reporter assay distortions from fused dual reporters. *RNA* **23**, 1285–1289 (2017).

13. Jacobs, J. L., Belew, A. T., Rakauskaite, R. & Dinman, J. D. Identification of functional, endogenous programmed -1 ribosomal frameshift signals in the genome of Saccharomyces cerevisiae. *Nucleic Acids Res* **35**, 165–174 (2007).

14. Baranov, P. V. *et al.* Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology* **332**, 498–510 (2005).

15. Charbonneau, J., Gendron, K., Ferbeyre, G. & Brakier-Gingras, L. The 5' UTR of HIV-1 full-length mRNA and the Tat viral protein modulate the programmed -1 ribosomal frameshift that generates HIV-1 enzymes. *RNA* **18**, 519–529 (2012).

16. Gendron, K. *et al.* The presence of the TAR RNA structure alters the programmed -1 ribosomal frameshift efficiency of the human immunodeficiency virus type 1 (HIV-1) by modifying the rate of translation initiation. *Nucleic Acids Res* **36**, 30–40 (2008).

17. Howard, M. T. *et al.* Sequence specificity of aminoglycoside-induced stop condon readthrough: potential implications for treatment of Duchenne muscular dystrophy. *Ann Neurol* **48**, 164–169 (2000).

18. Harrell, L., Melcher, U. & Atkins, J. F. Predominance of six different hexanucleotide recoding signals 3' of read-through stop codons. *Nucleic Acids Res* **30**, 2011–2017 (2002).

19. Yordanova, M. M. *et al.* AMD1 mRNA employs ribosome stalling as a mechanism for molecular memory formation. *Nature* **553**, 356–360 (2018).

20. Hennecke, M. *et al.* Composition and arrangement of genes define the strength of IRES-driven translation in bicistronic mRNAs. *Nucleic Acids Res* **29**, 3327–3334 (2001).

21. Bochkov, Y. A. & Palmenberg, A. C. Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location. *Biotechniques* **41**, 283–284, 286, 288 passim (2006).

22. Vallejos, M. *et al.* The 5'-untranslated region of the mouse mammary tumor virus mRNA exhibits cap-independent translation initiation. *Nucleic Acids Res* **38**, 618–632 (2010).

23. Baranov, P. V. *et al.* Programmed ribosomal frameshifting in the expression of the regulator of intestinal stem cell proliferation, adenomatous polyposis coli (APC). *RNA Biol* **8**, 637–647 (2011).

24.Beznosková, P., Wagner, S., Jansen, M. E., von der Haar, T. & Valášek, L. S. Translation initiation factor eIF3 promotes programmed stop codon readthrough. *Nucleic Acids Res* **43**, 5099–5111 (2015).

25.Müller, C. *et al.* Broad-spectrum antiviral activity of the eIF4A inhibitor silvestrol against corona- and picornaviruses. *Antiviral Res* **150**, 123–129 (2018).

26.Zinshteyn, B., Sinha, N. K., Enam, S. U., Koleske, B. & Green, R. Translational repression of NMD targets by GIGYF2 and EIF4E2. *PLoS Genet* **17**, e1009813 (2021).

27.Kobayashi, Y., Zhuang, J., Peltz, S. & Dougherty, J. Identification of a cellular factor that modulates HIV-1 programmed ribosomal frameshifting. *J Biol Chem* **285**, 19776–19784 (2010).

28.Green, L., Houck-Loomis, B., Yueh, A. & Goff, S. P. Large ribosomal protein 4 increases efficiency of viral recoding sequences. *J Virol* **86**, 8949–8958 (2012).

29.Young, D. J. *et al.* Tma64/eIF2D, Tma20/MCT-1, and Tma22/DENR Recycle Post-termination 40S Subunits In Vivo. *Mol Cell* **71**, 761-774.e5 (2018).

30.Schult, P. *et al.* microRNA-122 amplifies hepatitis C virus translation by shaping the structure of the internal ribosomal entry site. *Nat Commun* **9**, 2613 (2018).

31.LaFontaine, E., Miller, C. M., Permaul, N., Martin, E. T. & Fuchs, G. Ribosomal protein RACK1 enhances translation of poliovirus and other viral IRESs. *Virology* **545**, 53–62 (2020).

32.Zhang, H., Song, L., Cong, H. & Tien, P. Nuclear Protein Sam68 Interacts with the Enterovirus 71 Internal Ribosome Entry Site and Positively Regulates Viral Protein Translation. *J Virol* **89**, 10031–10043 (2015).

33.Bidou, L. *et al.* Premature stop codons involved in muscular dystrophies show a broad spectrum of readthrough efficiencies in response to gentamicin treatment. *Gene Ther* **11**, 619–627 (2004).

34.Mikl, M., Pilpel, Y. & Segal, E. High-throughput interrogation of programmed ribosomal frameshifting in human cells. *Nat Commun* **11**, 3061 (2020).

35.Cencic, R., Robert, F. & Pelletier, J. Identifying small molecule inhibitors of eukaryotic translation initiation. *Methods Enzymol* **431**, 269–302 (2007).

36.Novac, O., Guenier, A.-S. & Pelletier, J. Inhibitors of protein synthesis identified by a high throughput multiplexed translation screen. *Nucleic Acids Res* **32**, 902–915 (2004).

37.Ahn, D.-G. *et al.* Interference of ribosomal frameshifting by antisense peptide nucleic acids suppresses

SARS coronavirus replication. *Antiviral Res* **91**, 1–10 (2011).

38.Sun, Y. *et al.* Restriction of SARS-CoV-2 replication by targeting programmed -1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **118**, e2023051118 (2021).

39.Feng, Y. *et al.* SBI-0640756 Attenuates the Growth of Clinically Unresponsive Melanomas by Disrupting the eIF4F Translation Initiation Complex. *Cancer Res* **75**, 5211–5218 (2015).

40.Hekman, K. E. *et al.* A conserved eEF2 coding variant in SCA26 leads to loss of translational fidelity and increased susceptibility to proteostatic insult. *Hum Mol Genet* **21**, 5472–5483 (2012).

41.Chen, C.-K. *et al.* Structured elements drive extensive circular RNA translation. *Mol Cell* **81**, 4300-4318.e13 (2021).

42.Weingarten-Gabbay, S. *et al.* Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science* **351**, aad4939 (2016).

43.Lidsky, P. V., Dmitriev, S. E. & Andino, R. *Robust expression of transgenes in* Drosophila melanogaster. http://biorxiv.org/lookup/doi/10.1101/2022.10.30.514414 (2022) doi:10.1101/2022.10.30.514414.

44.Mäkeläinen, K. J. & Mäkinen, K. Testing of internal translation initiation via dicistronic constructs in yeast is complicated by production of extraneous transcripts. *Gene* **391**, 275–284 (2007).

45.Han, B. & Zhang, J.-T. Regulation of gene expression by internal ribosome entry sites or cryptic promoters: the eIF4G story. *Mol Cell Biol* **22**, 7372–7384 (2002).

46.Van Eden, M. E., Byrd, M. P., Sherrill, K. W. & Lloyd, R. E. Demonstrating internal ribosome entry sites in eukaryotic mRNAs using stringent RNA test procedures. *RNA* **10**, 720–730 (2004).

47.Holcik, M. *et al.* Spurious splicing within the XIAP 5' UTR occurs in the Rluc/Fluc but not the betagal/CAT bicistronic reporter system. *RNA* **11**, 1605–1609 (2005).

48.Kozak, M. A second look at cellular mRNA sequences said to function as internal ribosome entry sites. *Nucleic Acids Res* **33**, 6593–6602 (2005).

49.Baranick, B. T. *et al.* Splicing mediates the activity of four putative cellular internal ribosome entry sites. *Proc Natl Acad Sci U S A* **105**, 4733–4738 (2008).

50.Young, R. M. *et al.* Hypoxia-mediated selective mRNA translation by an internal ribosome entry site-independent mechanism. *J Biol Chem* **283**, 16309–16319 (2008).

51. Andreev, D. E. *et al.* Differential contribution of the m7G-cap to the 5' end-dependent translation initiation of mammalian mRNAs. *Nucleic Acids Res* **37**, 6135–6147 (2009).

52. Lemp, N. A., Hiraoka, K., Kasahara, N. & Logg, C. R. Cryptic transcripts from a ubiquitous plasmid origin of replication confound tests for cis-regulatory function. *Nucleic Acids Res* **40**, 7280–7290 (2012).

53. Khan, Y. A. *et al.* Evaluating ribosomal frameshifting in CCR5 mRNA decoding. *Nature* **604**, E16–E23 (2022).

54. Akirtava, C., May, G. E. & McManus, C. J. False-positive IRESes from Hoxa9 and other genes resulting from errors in mammalian 5' UTR annotations. *Proc Natl Acad Sci U S A* **119**, e2122170119 (2022).

55. Nejepinska, J., Malik, R., Moravec, M. & Svoboda, P. Deep sequencing reveals complex spurious transcription from transiently transfected plasmids. *PLoS One* **7**, e43283 (2012).

56. Loughran, G., Fedorova, A. D., Khan, Y. A., Atkins, J. F. & Baranov, P. V. Lack of evidence for ribosomal frameshifting in ATP7B mRNA decoding. *Mol Cell* **82**, 3745-3749.e2 (2022).

57. Terenin, I. M., Andreev, D. E., Dmitriev, S. E. & Shatsky, I. N. A novel mechanism of eukaryotic translation initiation that is neither m7G-cap-, nor IRES-dependent. *Nucleic Acids Res* **41**, 1807–1816 (2013).

58. Payne, A. J., Gerdes, B. C., Kaja, S. & Koulen, P. Insert sequence length determines transfection efficiency and gene expression levels in bicistronic mammalian expression vectors. *Int J Biochem Mol Biol* **4**, 201–208 (2013).

59. Shikama, Y. *et al.* Transcripts expressed using a bicistronic vector pIREShyg2 are sensitized to nonsense-mediated mRNA decay. *BMC Mol Biol* **11**, 42 (2010).

60. Jünemann, C. *et al.* Picornavirus internal ribosome entry site elements can stimulate translation of upstream genes. *J Biol Chem* **282**, 132–141 (2007).

61. Yordanova, M. M., Loughran, G., Atkins, J. F. & Baranov, P. V. Stop codon readthrough contexts influence reporter expression differentially depending on the presence of an IRES. *Wellcome Open Res* **5**, 221 (2020).

62. Terenin, I. M., Smirnova, V. V., Andreev, D. E., Dmitriev, S. E. & Shatsky, I. N. A researcher's guide to the galaxy of IRESs. *Cell Mol Life Sci* **74**, 1431–1455 (2017).

63.Kozak, M. New ways of initiating translation in eukaryotes? *Mol Cell Biol* **21**, 1899–1907 (2001).

64.Schneider, R. *et al.* New ways of initiating translation in eukaryotes. *Mol Cell Biol* **21**, 8238–8246 (2001).

65.Kozak, M. Alternative ways to think about mRNA sequences and proteins that appear to promote internal initiation of translation. *Gene* **318**, 1–23 (2003).

66.Jacobs, J. L. & Dinman, J. D. Systematic analysis of bicistronic reporter assay data. *Nucleic Acids Res* **32**, e160 (2004).

67.Mardanova, E. S., Zamchuk, L. A. & Ravin, N. V. Contribution of internal initiation to translation of cellular mRNAs containing IRESs. *Biochem Soc Trans* **36**, 694–697 (2008).

68.Thompson, S. R. So you want to know if your message has an IRES? *Wiley Interdiscip Rev RNA* **3**, 697–705 (2012).

69.Jackson, R. J. The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harb Perspect Biol* **5**, a011569 (2013).

70.Shatsky, I. N., Dmitriev, S. E., Andreev, D. E. & Terenin, I. M. Transcriptome-wide studies uncover the diversity of modes of mRNA recruitment to eukaryotic ribosomes. *Crit Rev Biochem Mol Biol* **49**, 164–177 (2014).

71.Bustin, S. A. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**, 169–193 (2000).

72.Shiraki, T. *et al.* Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc Natl Acad Sci U S A* **100**, 15776–15781 (2003).

73.Atkins, J. F. *et al.* A case for 'StopGo': reprogramming translation to augment codon meaning of GGN by promoting unconventional termination (Stop) after addition of glycine and then allowing continued translation (Go). *RNA* **13**, 803–810 (2007).

74.Ryan, M. D. & Drew, J. Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *EMBO J* **13**, 928–933 (1994).

75.Minskaia, E., Nicholson, J. & Ryan, M. D. Optimisation of the foot-and-mouth disease virus 2A co-expression system for biomedical applications. *BMC Biotechnol* **13**, 67 (2013).

76.Kim, J. H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human

cell lines, zebrafish and mice. *PLoS One* **6**, e18556 (2011).

77.Donnelly, M. L. L. *et al.* The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. *J Gen Virol* **82**, 1027–1041 (2001).

78.Akulich, K. A. *et al.* Four translation initiation pathways employed by the leaderless mRNA in eukaryotes. *Sci Rep* **6**, 37905 (2016).

79.Sahin, U., Karikó, K. & Türeci, Ö. mRNA-based therapeutics--developing a new class of drugs. *Nat Rev Drug Discov* **13**, 759–780 (2014).

80.Barreau, C., Dutertre, S., Paillard, L. & Osborne, H. B. Liposome-mediated RNA transfection should be used with caution. *RNA* **12**, 1790–1793 (2006).

81.Karasik, A., Jones, G. D., DePass, A. V. & Guydosh, N. R. Activation of the antiviral factor RNase L triggers translation of non-coding mRNA sequences. *Nucleic Acids Res* **49**, 6007–6026 (2021).

82.Hollien, J. *et al.* Regulated Ire1-dependent decay of messenger RNAs in mammalian cells. *J Cell Biol* **186**, 323–331 (2009).

83.Han, D. *et al.* IRE1alpha kinase activation modes control alternate endoribonuclease outputs to determine divergent cell fates. *Cell* **138**, 562–575 (2009).

84.Andreev, D. E., Terenin, I. M., Dmitriev, S. E. & Shatsky, I. N. Pros and cons of pDNA and mRNA transfection to study mRNA translation in mammalian cells. *Gene* **578**, 1–6 (2016).

85.McAllister, C. S. & Samuel, C. E. The RNA-activated protein kinase enhances the induction of interferon-beta and apoptosis mediated by cytoplasmic RNA sensors. *J Biol Chem* **284**, 1644–1651 (2009).

86.Juszkiewicz, S. & Hegde, R. S. Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Molecular Cell* **65**, 743-750.e4 (2017).

87.Sundaramoorthy, E. *et al.* ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. *Molecular Cell* **65**, 751-760.e4 (2017).

88.Matsuo, Y. *et al.* Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat Commun* **8**, 159 (2017).

89.Ikeuchi, K. *et al.* Collided ribosomes form a unique structural interface to induce Hel2-driven quality

control pathways. *The EMBO Journal* **38**, e100276 (2019).

90.Kisly, I., Kattel, C., Remme, J. & Tamm, T. Luciferase-based reporter system for *in vitro* evaluation of elongation rate and processivity of ribosomes. *Nucleic Acids Research* **49**, e59–e59 (2021).

91.Matreyek, K. A., Stephany, J. J., Chiasson, M. A., Hasle, N. & Fowler, D. M. An improved platform for functional assessment of large protein libraries in mammalian cells. *Nucleic Acids Res* **48**, e1 (2020).

92.She, R., Luo, J. & Weissman, J. S. Translational fidelity screens in mammalian cells reveal eIF3 and eIF4G2 as regulators of start codon selectivity. *Nucleic Acids Res* **51**, 6355–6369 (2023).

93.Dmitriev, S. E., Andreev, D. E., Ad'ianova, Z. V., Terenin, I. M. & Shatskiĭ, I. N. [Efficient cap-dependent in vitro and in vivo translation of mammalian mRNAs with long and highly structured 5'-untranslated regions]. *Mol Biol (Mosk)* **43**, 119–125 (2009).

94.Kozak, M. Evaluation of the fidelity of initiation of translation in reticulocyte lysates from commercial sources. *Nucleic Acids Res* **18**, 2828 (1990).

95.Soto Rifo, R., Ricci, E. P., Décimo, D., Moncorgé, O. & Ohlmann, T. Back to basics: the untreated rabbit reticulocyte lysate as a competitive system to recapitulate cap/poly(A) synergy and the selective advantage of IRES-driven translation. *Nucleic Acids Res* **35**, e121 (2007).

96.Dmitriev, S. E., Bykova, N. V., Andreev, D. E. & Terenin, I. M. [Adequate system for investigation of translation initiation of the human retrotransposon L1 mRNA in vitro]. *Mol Biol (Mosk)* **40**, 25–30 (2006).

97.Iizuka, N., Najita, L., Franzusoff, A. & Sarnow, P. Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from Saccharomyces cerevisiae. *Mol Cell Biol* **14**, 7322–7330 (1994).

98.Tuite, M. F. & McLaughlin, C. S. Polyamines enhance the efficiency of tRNA-mediated readthrough of amber and UGA termination codons in a yeast cell-free system. *Curr Genet* **7**, 421–426 (1983).

99.Skabkin, M. A., Skabkina, O. V., Hellen, C. U. T. & Pestova, T. V. Reinitiation and other unconventional posttermination events during eukaryotic translation. *Mol Cell* **51**, 249–264 (2013).

100.Zinoviev, A., Hellen, C. U. T. & Pestova, T. V. Multiple mechanisms of reinitiation on bicistronic calicivirus mRNAs. *Mol Cell* **57**, 1059–1073 (2015).

101.Terenin, I. M. *et al.* A cross-kingdom internal ribosome entry site reveals a simplified mode of internal ribosome entry. *Mol Cell Biol* **25**, 7879–7888 (2005).

102.Abaeva, I. S., Pestova, T. V. & Hellen, C. U. T. Attachment of ribosomal complexes and retrograde scanning during initiation on the Halastavi árva virus IRES. *Nucleic Acids Res* **44**, 2362–2377 (2016).

103.Beznosková, P., Pavlíková, Z., Zeman, J., Echeverría Aitken, C. & Valášek, L. S. Yeast applied readthrough inducing system (YARIS): an invivo assay for the comprehensive study of translational readthrough. *Nucleic Acids Res* **47**, 6339–6350 (2019).

104.Beznosková, P., Gunišová, S. & Valášek, L. S. Rules of UGA-N decoding by near-cognate tRNAs and analysis of readthrough on short uORFs in yeast. *RNA* **22**, 456–466 (2016).

105.Powell, M. L. *et al.* Further characterisation of the translational termination-reinitiation signal of the influenza B virus segment 7 RNA. *PLoS One* **6**, e16822 (2011).

106.Luttermann, C. & Meyers, G. A bipartite sequence motif induces translation reinitiation in feline calicivirus RNA. *J Biol Chem* **282**, 7056–7065 (2007).

107.Pöyry, T. A. A., Kaminski, A., Connell, E. J., Fraser, C. S. & Jackson, R. J. The mechanism of an exceptional case of reinitiation after translation of a long ORF reveals why such events do not generally occur in mammalian mRNA translation. *Genes Dev* **21**, 3149–3162 (2007).

108.Gould, P. S., Dyer, N. P., Croft, W., Ott, S. & Easton, A. J. Cellular mRNAs access second ORFs using a novel amino acid sequence-dependent coupled translation termination-reinitiation mechanism. *RNA* **20**, 373–381 (2014).

109.Gunišová, S., Hronová, V., Mohammad, M. P., Hinnebusch, A. G. & Valášek, L. S. Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. *FEMS Microbiol Rev* **42**, 165–192 (2018).

110.Kozak, M. Constraints on reinitiation of translation in mammals. *Nucleic Acids Res* **29**, 5226–5232 (2001).

111.Sorokin, I. I. *et al.* Non-Canonical Translation Initiation Mechanisms Employed by Eukaryotic Viral mRNAs. *Biochemistry (Mosc)* **86**, 1060–1094 (2021).

112.Makeeva, D. S. *et al.* Relocalization of Translation Termination and Ribosome Recycling Factors to Stress Granules Coincides with Elevated Stop-Codon Readthrough and Reinitiation Rates upon Oxidative

Stress. *Cells* **12**, 259 (2023).

113.Brown, C. M., Dinesh-Kumar, S. P. & Miller, W. A. Local and distant sequences are required for efficient readthrough of the barley yellow dwarf virus PAV coat protein gene stop codon. *J Virol* **70**, 5884–5892 (1996).

114.Xu, Y. *et al.* A Stem-Loop Structure in Potato Leafroll Virus Open Reading Frame 5 (ORF5) Is Essential for Readthrough Translation of the Coat Protein ORF Stop Codon 700 Bases Upstream. *J Virol* **92**, e01544-17 (2018).

115.Kimura, K. *et al.* Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res* **16**, 55–65 (2006).

116.Makhnovskii, P. A. *et al.* Alternative transcription start sites contribute to acute-stress-induced transcriptome response in human skeletal muscle. *Hum Genomics* **16**, 24 (2022).

117.Huang, F. *et al.* Inhibiting the MNK1/2-eIF4E axis impairs melanoma phenotype switching and potentiates antitumor immune responses. *J Clin Invest* **131**, e140752, 140752 (2021).

118.Watt, K. *et al. Epigenetic coordination of transcriptional and translational programs in hypoxia*. http://biorxiv.org/lookup/doi/10.1101/2023.09.16.558067 (2023) doi:10.1101/2023.09.16.558067.

119.Wang, X.-Q. & Rothnagel, J. A. 5’-untranslated regions with multiple upstream AUG codons can support low-level translation via leaky scanning and reinitiation. *Nucleic Acids Res* **32**, 1382–1391 (2004).

120.Dmitriev, S. E. *et al.* Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5’ untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal ribosome entry site mediated. *Mol Cell Biol* **27**, 4685–4697 (2007).