Characterization, modelling and mitigation of gene expression burden in mammalian cells

Frei T[1]±, Cella F[2,3]±, Tedeschi F[2], Gutierrez J[1], Stan GB[4], Khammash M[1]*, Siciliano V[2]*.

1 Department of Biosystems Science and Engineering, ETH Zürich
2 Istituto Italiano di Tecnologia-IIT, Largo Barsanti e Matteucci, Naples ITA
3 University of Genoa, Genoa ITA
4 Department of Bioengineering and Centre for Synthetic Biology, Imperial College London, U.K.
± These authors contributed equally
*Co-corresponding authors: mustafa.khammash@bsse.ethz.ch and velia.siciliano@iit.it

Abstract

The ability to engineer synthetic circuits and devices in mammalian cells has enabled a multitude of exciting applications in industrial biotechnology and medical therapy. In spite of the recent availability of powerful genome engineering tools such as CRISPR-Cas9, the process of designing and implementing functioning genetic circuits remains painstakingly slow and fraught with inexplicable failures. The unexpected divergence between intended and actual function of synthetic circuits can be attributed to several factors, most notably the contextual background in which these circuits operate. In particular, the dependence of synthetic circuits on cellular resources which are limited leads to unintended dynamic couplings between the various exogenous components of the circuit, as well as with endogenous components of the host cell. The consumption of these resources by synthetic circuits thus exerts a burden on the host cell that reduces its capacity to support additional circuits, potentially resulting in counter-intuitive functional behaviors and detrimental effects on host physiology. In spite of its critical importance, gene expression burden in mammalian cells remains largely unstudied. Here, we comprehensively investigate the impact of host resource limitations on synthetic constructs in mammalian cells. We show that effects of both transcriptional and translational resource limitations can be observed and that each can lead to the coupling of independent, co-expressed synthetic genes, which in turn imposes trade-offs in their expression and diminishes performance. We next explore the role of post-transcriptional regulators, such as microRNAs (miRNAs) and RNA binding proteins (RBPs) and show that they can redistribute resources in a way that limits burden-induced coupling effects. To quantify and predict the influence of burden on gene expression in engineered cells, we describe a modelling framework that allows to incorporate the effect of limited resources into classical models of gene expression. Based on this framework, we identify network topologies that mitigate burden and then implement these topologies using endogenous and synthetic miRNA-based circuits that buffer the expression of genes of interest from fluctuations in cellular resources. Among other regulators, microRNAs can conveniently tailor synthetic device regulation in different cell lines and tissues, as well as during dynamic changes of cellular states and downstream information processing, or in pathological conditions. This study thus establishes a foundation for context-aware predictions of *in vivo* synthetic circuit performance and paves the way towards a more rational synthetic construct design in mammalian cells.

Introduction

Heterologous gene expression in mammalian cells facilitates the study of diverse biological processes including gene regulation[1], developmental patterns[2], mutations, and genomic rearrangements in cancer progression[3]. More recently, it has gained clinical relevance with gene therapy[4] and has become a fundamental tool for biotechnologists, genetic engineers, and synthetic biologists, with applications ranging from the engineering of recombinant protein-producing cells[5] to the creation of novel cell-based therapies for clinical use[6]. These critical applications require that gene expression, while being robust, does not impinge on normal host physiology. Engineering gene circuits with these features necessitates long design-build-and-test cycle iterations that are expensive and time consuming[7]. This is especially true for mammalian cells where the engineering process is slower compared to other organisms. At the core of the problem is poor predictability of gene expression[8] in engineered cells arising from the dependence of expression on the cellular context. In particular, the unmodeled dependence of exogenous genetic circuits on limited host resources shared with endogenous pathways frequently leads to unanticipated counter-intuitive circuit behaviors[9]. This remains one of the most important open problems in mammalian synthetic circuit design. In bacterial cells, substantial progress towards increasing the predictability of gene expression has been made by showing that the introduction of genetic material imposes a significant burden, resulting in decreased growth rate and degraded cellular performance[10]. This has been attributed to the fact that cells possess a finite pool of resources for the expression of genes[11,12], so that when new genes are delivered into cells, their transcription and translation require resources that become unavailable for the expression of other genes. These observations have prompted the development of models that consider gene expression in a resource-limited context[13–16], as well as to approaches for mitigating the impact of resource burden in bacteria[17], including the implementation of a system that directly senses gene expression burden and modulates the expression of output genes through a dCas9-based negative feedback[18]. Also in *S. Cerevisiae*, the contributions of transcription and translation were explored in relation to the cost of protein production, where it was demonstrated that both processes are limiting according to different growth conditions[19]. In mammalian cells, performance shortcomings of synthetic circuits due to limited resources have been observed[20], and control systems to enhance performance, proposed. However, a deeper examination of resource limitation in mammalian cells is still missing.

To evaluate the burden imposed by synthetic circuits on host cells we designed constructs to explore the effects of both transcriptional and translational resource limitations. We demonstrate in different cell lines that sharing of host cell resources is responsible for the effective coupling of otherwise independently co-expressed synthetic genes, the result of which is the presence of

inevitable trade-offs in their expression. Next, we investigate the role of post-transcriptional regulators, like RNA binding proteins (RBPs) and microRNAs (miRNAs) in mitigating the impact of burden-induced coupling. Interestingly, we observe that both are able to reallocate resources and alleviate burden.

In order to reduce the unpredictability of gene expression in engineered cells we describe a modelling framework that allows to incorporate the effect of limited resources into typical models of gene expression. This is achieved by introducing resource limited production rates that readily replace their non-resource limited counterparts. Similar modelling approaches have already been applied to bacterial systems[13,14,21,22] and conceptually date back to general models of multiple inhibitions in enzyme catalysis[23]. We demonstrate that this framework successfully recapitulates the non-trivial dose-responses of genetic circuits in the presence of resource constraints. Based on the model, we identify network topologies that mitigate burden based on miRNA regulation. MicroRNAs are important post-transcriptional regulators involved in several biological processes in mammalian and plant cells[24]. Their fast dynamics, cell specificity and tunability make them ideal biomarkers for interfacing with synthetic gene regulation. MicroRNA-based synthetic circuits have been constructed to buffer gene expression against noise[25] and fluctuations in external inducer concentration[26,27]. In the latter and other studies, miRNA-based Incoherent Feed-Forward Loops (IFFL), in which miRNAs and target mRNAs are transcribed simultaneously, have been shown to limit variability in gene expression in both endogenous and synthetic pathways[28,29]. We thus implemented circuits based on endogenous or synthetic miRNAs and demonstrated that they are able to preserve the expression levels of a gene of interest, in spite of changes in available cellular resources caused by the loading of other gene constructs.

To the best of our knowledge this is the first study assessing the individual contribution of transcriptional and translational processes of circuit-induced burden in mammalian cells, and the first to unveil mechanisms of resource reallocation as a consequence of post-transcriptional regulation. This new knowledge, accompanied by a refined model framework that describes biochemical reactions in a finite resources context was translated into the design of circuit topologies that enable tolerance to gene overload. This model-based approach to burden paves the way to more realistic output predictions and optimal synthetic construct design.

Results

Genetic circuits impose burden on mammalian cells and compete for shared cellular resources.

The outline of the paper is summarized in **Fig.1**. Since the expression of heterologous genes relies on finite amounts[30] of host transcriptional and translational resources, such as RNA polymerases, nucleotides, amino acids and ribosomes, we hypothesise that the availability of these

resources imposes a link between otherwise uncoupled genes (**Fig.1A**). Consequently, the increased expression of one gene comes at the expense of another, as resources employed in the first gene's expression are not available for the expression of the second (**Fig.1A, bottom**). Thus, resource limitation becomes a key parameter for the outcome of gene expression. We explicitly incorporate these effects in a resource-aware model, in which we replace the rates of resource-dependent reactions with adjusted effective rates (**Fig.1B**).

Our resource-aware model suggests that a gene (e.g. miRNA) that depends on a particular pool of resources (e.g. transcriptional resources) can be regarded as a sensor of the state of the pool and can thus be exploited to reduce the coupling of other co-expressed genes that depend on the same pool. In fact, such a topology mitigates the load of the regulated node on the available resources, which become more abundant for the production of other species. The model predicts that thanks to the mitigation, the expression of a tunable eXtra transgene (hereafter named *X-tra*) and a constitutively expressed reporter (hereafter named *capacity monitor*) become less interdependent (**Fig.1C**).

To investigate the effects of resource sharing in mammalian cells, we designed several regulatory circuits composed of *X-tra* and *capacity monitor*. In all the experimental setups *X-tra* is the extra, tunable load imposed on the host cells which is detected though the *capacity monitor's* response.

In general, we hypothesized that due to competition for finite resources, there is a threshold above which gene expression from synthetic circuits saturates. We therefore co-transfected HEK293T cells with two constitutively expressed fluorescent proteins, mCitrine and mRuby3 in molar ratios ranging from 1:4 to 4:1, for a total of 50 ng or 500 ng of encoding plasmid (**Fig. 2A**). A qualitative idea about the expected response in the presence of competition for limited resources is presented for both cases (**Fig. 2A, left**). The equations used to generate these plots were derived from the modelling framework introduced later on, in **Fig. 4A** (**Supplementary Note 1**). In accordance with our expectations, the total amount of 500ng of encoding plasmids results in a dramatic drop in global gene expression as compared to 50ng (**Fig. 2A, right**). Furthermore, in both experimental conditions mCitrine and mRuby3 fluorescence levels are negatively correlated; the higher the mCitrine, the lower the mRuby3 and vice versa (**Fig. 2A, right**). This correlation was also more severe for 500 ng of transfected plasmid than for 50 ng. The effect does not appear to be cell-type dependent, since we observe similar trends in H1299, U2OS and HeLa cells (**Supplementary Figures 1-3**). Analogously, the negative correlation of fluorescent proteins expression was observed by combining two constitutive promoters of different strengths. Here, we observed that the expression in one gene (e.g. mCitrine) changes depending on whether the second gene (e.g. mRuby3) is driven by a strong (hEF1a) or medium promoter (EFS) and vice versa (**Supplementary Figure 4**).

To confirm that there is a link between the expression of these allegedly uncoupled genes, we tuned *X-tra* activation with Doxycycline (Dox)-repressed promoter, while keeping mCitrine constitutively active (*capacity monitor*) (**Fig. 2B, left**), and evaluated protein levels in the absence or presence of Dox (from 0 to 1 µg/mL). Similar to previous observations, *X-tra* repression corresponds to increased *capacity monitor* levels (**Fig. 2B, right**). Furthermore, to exclude any experimental confounds as source of our observations, we investigated the effect of both cell seeding density and nutrient starvation on the outcome of a titration of a constitutively expressed gene, against a second constitutively expressed gene held at a constant copy number. In this experimental setting, cells were co-transfected with two plasmids. The first plasmid, which was provided at incremental levels, is composed of two transcriptional units (TU), one comprising strong promoter driving the expression of mCitrine (hEF1a), the other driving the expression of miRFP670 under a weak promoter (SV40). The second encodes for mRuby3 under a strong constitutive promoter (hEF1a). We observed that neither cell seeding nor nutrient supply had any apparent effect on the expression levels (**Supplementary Figure 5**).

These proof-of-concept experiments demonstrate that: *i)* gene expression in mammalian synthetic circuits is connected even in the absence of direct regulation, and *ii)* expression of exogenous genes is limited by cellular resource availability.

Since several different resource pools could be responsible for the observed effects, we set out to characterize the individual contributions of transcriptional and translational resource limitation to cellular burden.

Genes compete for limited transcriptional and translational resources

Following the line of reasoning presented so far, we speculate that any limitation in transcriptional resources will cause genes to compete for mRNA production. To investigate this hypothesis we increased *X-tra*/*capacity monitor* molar ratios from 1:1 to 2.5:1 in H1299 cells for a total of 500ng of plasmid DNA. We quantified mRNA levels for each condition, and show that as the plasmid encoding the *X-tra* increases so does the *X-tra* mRNA. However, this happens at the expense of the levels of *capacity monitor* mRNA, which decrease (**Fig. 2C, Supplementary Figure 1**). Thus, shared transcriptional resources are indeed a limiting factor in mammalian synthetic gene co-expression.

To confirm these observations we implemented a genetic circuit that can selectively overload the transcriptional resource pool without sequestering translational resources. The system is based on the self-cleaving hepatitis delta virus (HDV) ribozyme, which ensures that most of the transcribed mRNA is cleaved. Due to the cleavage and consequent transcript destabilization, the system allows to specifically observe limitations on transcriptional resources, while only minimally perturbing the

translational ones (**Fig. 2D, left**). We thus created a single plasmid with two TU. One TU consists of a tTA transcription factor co-expressed with the fluorescent protein mRuby3 (*capacity monitor*) via the P2A peptide, driven by a constitutive promoter. The second TU includes the TRE promoter (responsive to tTA) regulating the HDV-X-tra expression. In this setup, Dox can be used to modulate the amount of burden imposed, similar to what was already shown in **Fig. 2B**.

We compared this circuit to a catalytically inactive mutant of the HDV ribozyme. As expected, we observed that when the HDV ribozyme is inactive, *X-tra* protein levels increase with rising amounts of plasmid (**Supplementary Figure 6, top pale purple bar**), whereas those of the *capacity monitor* decrease (**Fig. 2D, bottom**). On the contrary, when the HDV ribozyme is active, *X-tra* expression is strongly reduced and only minorly increasing with lower Dox concentrations (**Supplementary Figure 6, top dark purple bar**). Here, the decrease of *capacity monitor* levels is still present, but to a lower extent than in the previous condition, supporting the observations in **Fig. 2C** that transcriptional resources are limited to a certain extent (**Fig. 2D, dark blue bar**). Interestingly, the expression levels of the *capacity monitor* with active HDV ribozyme are higher compared to the inactive mutant (**Supplementary Figure 6, bottom dark blue bar**). We suggest that, assuming that the *X-tra* with an active HDV ribozyme is rapidly degraded, it is likely to sequester less translational resources, which should result in higher expression of the *capacity monitor*. This suggests that the transcriptional resource pool is only partially responsible for the described trade-offs and that post-transcriptional resources may represent an additional bottleneck to the overall expression of synthetic genes.

Therefore, we aimed to characterize the impact of translational resources in cellular burden, by adding a synthetic intron[31] in the 5'UTR of the *X-tra* fluorescent protein (**Fig. 2E, top**). The synthetic intron enhances gene expression by augmenting mRNA export from the nucleus to the cytoplasm[31]. We observed higher expression of *X-tra* in H1299, HEK293T, HeLa, and U2OS cell lines in the presence of a synthetic intron, opposed by lower *capacity monitor* levels (**Fig.2E**, **Supplementary Figure 7**), confirming that resources employed for translational regulation are also limiting.

Our data collectively indicate that exogenous genes compete for resources both at the transcriptional and translational levels, overall imposing a gene expression burden on mammalian cells. Since one of the goals in synthetic biology is output predictability and reproducibility, gene expression burden becomes a key issue to address.

We reasoned that post-transcriptional and translational regulators such as miRNAs or RBPs, may free up cellular resources[32] by repressing target mRNA translation or inducing its degradation. If this hypothesis were true, they could be exploited in more robust circuit topologies to reduce gene expression load, resulting in improved performance and predictability of engineered circuits. Therefore, we tested two RBPs, L7Ae and Ms2-cNot7[33,34] (**Fig. 2F, Supplementary Figure 8**), as well as endogenous miRNAs, miR-31, miR-221 and miR-21 (**Fig. 2G**). Specifically, L7Ae downregulates protein

expression by inhibiting mRNA translation, Ms2-cNot7 induces mRNA degradation[33], and miRNAs operate by both mechanisms according to complete[35] or partial[36] complementarity to the mRNA target. For each system, a fluorescent protein encoding mRNA targeted by either RBPs or miRNAs (*X-tra*) was co-expressed with a second, constitutively expressed fluorescent readout (*capacity monitor*).

The archeal RBP L7Ae binds two kink turn motives in the 5'UTR of the *X-tra* mRNA, whereas Ms2 binds target sites in the 3'UTR of the *X-tra* transcript, allowing the deadenylase cNOT7 to cut the polyA tail of the target mRNA [33]. We tested the systems in U2OS, H1299, HEK293T and HeLa cell lines, consistently observing that *X-tra* downregulation by RBPs results in increased levels of the *capacity monitor* (**Fig. 2F, Supplementary Figure 8**).

To evaluate the effect of miRNA regulation on cellular resources re-allocation, we placed three target sites (TS) in the 3'UTR of the *X-tra*. Specifically, the TS respond to miR-31, miR-221, miR-21, highly expressed in H1299, U2OS, and HeLa cells respectively. The *capacity monitor* expression levels increased when the *X-tra* was downregulated by miRNAs, as compared to controls lacking microRNA target sites (**Fig. 2G**). These results confirm that post-transcriptional regulators can redistribute intracellular resources and, importantly, that this phenomenon is cell-context independent, albeit to different extents of negative correlation between *X-tra* and *capacity monitor* expression. The latter could be the consequence of several factors, such as relative abundance of post transcriptional/translational resources, as well as absolute miRNA levels in the cells.

A major advantage of miRNAs over RBPs is that they are endogenously expressed and cell-line specific. Thus, their expression does not impose additional burden, and any circuit's design can be easily tailored to the cell/tissue of interest. Based on the results presented here, we envision that genetic circuits that mitigate resource competition via miRNAs may be designed for any mammalian cell line with a very broad set of potential applications.

Effect of miRNAs on reshaping resource allocation

To outline the correlation between miRNA-mediated downregulation and resource redistribution, we designed a library of *miRNA sensors* for miR-31, which is highly expressed in H1299 lung cancer cells[37]. The *miRNA sensor* is composed of the fluorescent reporter mKate with or without miR-31 TS, encoded along with the *capacity monitor* (EGFP) on a single plasmid with a bidirectional promoter (**Fig. 3A, Supplementary Figure 9**). The library includes 0, 1, or 3 fully complementary miR-TS in the 3' or 5'UTR of mKate.

Similar to what was previously observed (**Fig. 2G**), when the *miRNA sensor's* levels decrease as a consequence of miR-31 regulation, expression of the *capacity monitor* increases. Stronger repression was appreciated with 3TS in the 5'UTR and was accompanied by higher capacity monitor

levels (**Fig. 3B**). Conversely, when miR-31 activity was impaired by a miR-31 inhibitor, mKate expression was rescued (**Fig. 3C, pink bars**) and *capacity monitor* levels decreased (**Fig. 3C, pale blue bars**), demonstrating that miRNA sensor and *capacity monitor* levels are linked. Interestingly, the effect of the miRNA inhibitor was more pronounced with TS placed in the 3'UTR. Synthetic miRNA inhibitors bind to endogenous miRNAs in an irreversible manner[38], but differences in their action when target sites are placed in the 3' or 5' UTR were never reported before, and mechanistic insights are missing.

To confirm that miRNA-mediated resource redistribution is independent of experimental setting and plasmid design, we encoded the *miRNA sensor* and *capacity monitor* on two separate plasmids. Similar to previous results, *miRNA sensor* and *capacity monitor* are negatively correlated (**Supplementary Figure 10A**), suggesting that cellular burden and miRNA-dependent resource reallocation are a common challenge and solution respectively. Downregulation of the *miRNA sensor* was also confirmed by qPCR (**Supplementary Figure 10B**) and by western blot (**Supplementary Figure 10C**). Finally, when *miR-31 sensor* was transfected in low miR-31 cell lines such as U2OS and HEK293T, neither the *miRNA sensor* nor the *capacity monitor* levels vary (**Supplementary Figure 11**), further confirming the correlation between mRNA downregulation by miRNA and mitigation of the cellular burden.

We showed in **Fig. 2G** that miRNA-dependent resource reallocation is observed across different cell lines, by expressing cell-specific *miRNA-sensors* which include 3TS in the 3'UTR. We then built a library of sensors with different number and location of TS for miRNAs -221 and -21 which are highly expressed in U2OS and HeLa cells respectively, and we confirmed that also in these cell lines *miRNA-sensor* and *capacity monitor* are inversely correlated, consistently with observations in H1299 cells (**Supplementary Figure 12-13**).

Overall these data show that miRNAs can be used to develop optimal resource-aware plasmid-designs with burden-mitigating circuit topologies.

A gene expression model extended to incorporate shared limited resources captures non-intuitive observations

In order to provide a better understanding of our results, we describe a general modelling framework which offers a simple and convenient recipe for extending existing models of biochemical reactions, allowing them to incorporate the effects of shared limited resources. This framework follows ideas originally used to capture the competitive interaction of multiple inhibitors with an enzyme[23,39] and has also been applied to the context of shared cellular resources in previous studies by other groups[13–16,40].

Fig. 4A illustrates an overview of the framework. The main idea is to replace the rates of reactions that involve a shared resource with an equivalent reaction rate that captures the reduced availability of that resource due to the presence of competing genes. To create a distinction between regular reactions and resource limited production, we introduce double headed reaction arrows as illustrated at the bottom of **Fig. 4A**. This arrow summarizes the set of intermediate interactions shown in more detail at the top left of **Fig. 4A**. Here, the substrate A_i binds resource R with rate k^+_i to form the complex C_i. This reaction is also assumed to be reversible with rate k^-_i. With rate k^{cat}_i the complex gives rise to the product B_i, while also freeing up both the substrate A_i and the resource R. It is assumed that the total amount of resource, R^{total} is conserved and remains constant. Considering all possible substrates of resource and assuming that C_i is in quasi-steady-state, the rate for resourced limited production can be expressed as k^{equiv}_i,shown in the top right corner of **Fig. 4A**, which will be a function of the total amount of resources and the current state of all competing substrates. This expression can be readily used to substitute all reaction rates that involve shared limited resources.

In Lillacci et al[41], new synthetic control systems have been introduced to reduce the effect of burden. The same study also reports a non-monotonic dose-response behavior in a simple inducible gene expression system. Such non-monotonic behavior is rather non-intuitive, especially considering that conventional methods of modelling gene expression via hill-type functions can only produce monotonically increasing (or decreasing) dose-response curves. Although Lillacci et al. proposed models for the different circuit topologies[41], the effect of competition for limited resources was not directly considered. To demonstrate the usefulness of our modelling framework, we extend these models to include limited resources and show that the refined models recapitulate the experimental observations.

The four topologies considered in[41] were split into two groups based on the presence of a negative feedback from the fluorescent protein DsRed to the transcriptional activator (tTA). The first group consists of the open loop (OLP) and incoherent feedforward (IFF) topologies. In both these circuits, the constitutively expressed transcriptional transactivator, fused to the fluorescent protein Cerulean (tTA-Cer), activates the expression of the fluorescent protein DsRed. Furthermore, the gene of DsRed intronically encodes the synthetic microRNA FF4 (miR-FF4). In the IFF topology, the matched target of this microRNA is present in the 3' untranslated region (UTR) of the DsRed gene. This target is replaced by a mismatched target for the microRNA FF5 in the OLP. These detailed interactions are depicted on the left side of **Fig. 4B**. To observe potential shifts in the allocation of resources, we generated dose-response curves by increasing the amount of transfected tTA-Cer plasmid. The other two plasmids, containing DsRed and the constitutively expressed fluorescent transfection reporter mCitrine, were held constant. As can be seen from the model fit, plotted as a

solid line in the data graph, the extended model reproduces the non-intuitive behavior of the dose-responses (**Fig. 4B, right**).

A second group of topologies considered by Lillacci et al consists of the feedback (FBK) and the FBK+IFF hybrid (HYB) topologies. In addition to all the interactions described for the OLP and IFF circuits, the FBK and the HYB circuits possess miR-FF4 targets in the 3' UTR of the tTA-Cer gene, which introduces negative feedback. Furthermore, the FBK and HYB differ from each other by the presence of a matched target for miR-FF4 in the HYB topology, which introduces incoherent feedforward and is replaced by a mismatched FF5 target in the FBK circuit. All the interactions are illustrated in detail on the left side of **Fig. 4C**. The dose-response curves for the two circuits were obtained as described above. Again, the fit of the extended model to the data is capable of capturing its non-intuitive behavior (**Fig. 4C, right**).

Here, we have described a simple framework that adapts existing models of gene expression to include pools of shared and limited resources. We show that it can be used to provide an explanation for non-intuitive dose-responses in tTA-based circuits. With this framework as a tool, we believe that issues due to burden can be addressed in the design phase of circuit building and will not have to be optimized in costly build-test cycles later on.

Mitigating cellular burden with incoherent feed-forward loop circuits

In the following section, we implement a strategy that exploits microRNA to reduce the indirect coupling between co-expressed genes. In particular, we take advantage of the fact that microRNA production also requires (pre-translational) cellular resources and therefore microRNA levels act as a sensor for resource availability. Because of this, it is possible to reduce the coupling between co-expressed genes to a common resource pool by introducing miRNA-mediated repression of those genes (as long as the miRNA itself is also affected by the same resource pool). Since both the miRNA and the miRNA-repressed gene are affected by the availability of resources, miRNA-mediated repression implements an incoherent feedforward loop similar to previously published circuits[20,27,28]. Furthermore, it represents a biological implementation of the microRNA circuit proposed by Zechner et al.[42]. In this setting, the microRNA can also be interpreted as an estimator of its cellular context (e.g. amount of free resources) and it acts to filter this context in the output of interest. Lastly, the incoherent feedforward motif was also presented at the 6th International Mammalian Synthetic Biology Workshop in Chicago by Jones et al. as a means to robustify against depletion of transcriptional resources.

We explore this strategy for both an endogenously expressed microRNA (**Fig. 5A and B**) and a synthetic microRNA expressed from a plasmid (**Fig. 5C and D**). More specifically, **Fig. 5A** displays an approach that exploits endogenous microRNA to reduce the coupling of a gene of interest (GOI) to

the expression level of other genes, introduced by the limitation in resources. To do so, it only requires adding target sites of an endogenous microRNA to the 5' UTR of the gene of interest (mKate). In our experimental setup, when the copy number of a second gene *(X-tra)* is increased, resources are drawn away from the expression of mKate and allocated to the expression of *X-tra*. The shift in resource allocation is expected to also affect the the microRNA miR-31, which acts as a *capacity monitor*. This leads to a reduction in the repression of mKate, effectively compensating for the imposed burden.

To demonstrate this mitigation approach experimentally, we co-transfected H1299 cells with increasing amounts of an X-tra reporter plasmid, containing EGFP, and an mKate plasmid (GOI) that either includes three miR-31 target sites (with mitigation) or none (without mitigation). As expected, the expression level of *X-tra* approaches saturation as the plasmid copy number increases, both for the targeted and non-targeted *GOI* variants (**Fig. 5B**). In agreement with previous results, the expression of the non-targeted *GOI* strongly decreases with increased expression of X-tra. Conversely, the decrease in expression of the targeted *GOI* is only about a third of that of the non-targeted variant, indicating improved robustness towards changes in resource availability. This observation is also captured well by a model of the system that explicitly considers resources, as described in the previous section. We also tested the same experimental model in U2OS (**Supplementary Figure 14**) and HEK293T (**Supplementary Figure 15**) cells using IFF circuits built to respond to miR21 and miR221 respectively. In both cases we demonstrate improved tolerance to burden imposed by increasing exogenous gene load.

Next, we show that a reduction of resource-mediated coupling can also be achieved without relying on endogenous factors, using a synthetic microRNA. We further extend the approach to more than one output, demonstrating that the general principle can extend to multiple outputs as well (**Fig. 5C**). Similar to the endogenous case, the genes-of-interest, mCitrine (GOI_1) and DsRed (GOI_2), encode target sites for the microRNA FF4 in their 3' UTRs. In contrast to endogenous microRNA expression, however, here the microRNA is expressed intronically from GOI_2. In this way, the circuit forms a self-contained unit that can be easily transferred between cell types.

To test this approach, we co-transfected mouse embryonic stem cells with a plasmid encoding constitutively expressed miRFP670 (X-tra) and a plasmid composed of two transcriptional units (TU), each expressed under the constitutive promoter hEF-1 α (**Fig. 5C**). The first TU leads to the expression of mCitrine, whereas the second one encodes the fluorescent protein DsRed. Furthermore, the 3' UTR of DsRed contains either three target sites for the synthetic miRNA-FF4 or three mismatched miR-FF5 target sites (negative control). The miRNA-FF4 is intronically encoded in the DsRed gene. Identically to the endogenous case, the amount of X-tra plasmid was increased while keeping the *GOIs'* plasmid constant. Again, expression of X-tra increases and approaches

saturation with increasing molar amounts. Conversely, the non-targeted variants of the *GOIs* (**TFF5 in Fig. 5D**) decrease as the expression of X-tra increases. Analogously to what was observed for endogenous miRNAs, the expression of the targeted variants (**TFF4 in Fig. 5D**) decreases less than for the non-targeted ones albeit with lower efficiency. Moreover, when fitted to data from transfections with plasmids where only one of the two *GOIs* is targeted by TFF4, our modelling framework is able to predict the experimental observations for two targeted *GOIs* (solid lines). Finally, to demonstrate that the observed effect is not cell-type specific, we also tested the approach in HEK293T cells (**Supplementary Figure 16**). In HEK293T cells, robustness towards shifts in resource availability was similar to the endogenous microRNA-based regulation (**Fig. 5C**). Thus, we show that also in entirely synthetic systems, robustness towards shifts in resources availability is achieved. To ensure that the observed mitigation is not caused by a higher tolerance to changes in availability at lower expression levels we show analytically using the described modelling framework that the normalized expression at lower levels is more sensitive to burden (Supplementary note 2).

Overall, these results suggest that our approach can be used to mitigate resource-mediated coupling of gene expression despite cell-to-cell variability, demonstrating the portability and broad applicability of our findings.

In summary, our results demonstrate that IFF circuits can mitigate burden from transgene expression in mammalian cells. Importantly, by using miRNA one can either opt for endogenous miRNAs to specifically tailor a circuit to a desired cell line, or create a portable circuit by using a synthetic miRNA such as miR-FF4.

Discussion

Here, we report on the first analysis of burden imposed by exogenous gene expression in mammalian cells. The cost of heterologous systems' expression in host cells had been previously demonstrated in bacteria[10,18] and yeast[19,43], where exogenous gene load results in slower or defective growth and poor predictability of gene expression. In bacteria, these studies have led to more advanced models of gene expression and to the implementation of feedback topologies that respond to burden, either by reducing the load to acceptable levels[18] or by increasing demand to compensate for reductions in available resources[17]. Our study demonstrates that finite cellular resources also represent a general bottleneck for the predictability and performance of synthetic circuits in mammalian cells, impairing efforts for optimal circuit design.

By characterizing the unexpected trade-offs in the expression of two allegedly independent, co-expressed genes in different cell lines, we demonstrated that burden due to heterologous gene expression is wide-spread in mammalian cells, imposed both at the transcriptional and translational level by even the most simple synthetic gene expression motifs, and that the expression of

independent genes is coupled through shared cellular resources. In addition, we show that endogenous and synthetic post-transcriptional and translational regulators such as RNA binding proteins and microRNAs can limit the burden by redistributing cellular resources.

Furthermore, we describe a modelling framework that captures the indirect interdependence of gene expression in a resource-limited context. Indeed, synthetic circuit designs usually assume that the expression of different components is independent. However, this might lead to unexpected and non-monotonous behaviors once the design has been implemented biologically. Specifically when considering transient transfections, where cells can contain up to several thousand copies of the transfected plasmid per nucleus[44,45], it is not unreasonable to assume that gene expression is limited by one of the aforementioned resources. A recent computational study has shown how host-circuit interactions in *E. coli* involving the use of shared resources can impair circuit function[46]. Furthermore, the family of controllers analyzed in **Fig. 4** exemplify the non-intuitive behavior that might arise due to burden. Here, limiting resources lead to a non-monotonous dose-response in a simple transcription factor and reporter pair. One way to address these issues is to reduce the amount of genetic material delivered to the cell. This concept is illustrated in **Fig 2A** by the fact that the intergene coupling is reduced when 50 ng of total plasmid DNA is transfected rather than 500 ng. To determine the scales under which burden can be neglected and the circuit components should function as expected, a framework for modelling gene expression with resource limitations, as described in this and other publications[13,16,40,47], could be applied. We believe that our simple framework for introducing resource-limited steps into pre-existing models of synthetic circuits provides a simple way of avoiding such downfalls, supporting the design of circuits that are apt to dealing with burden. However, the requirement of a small payload might make it hard to observe the circuit function, especially as the size of the circuit grows and the limited capacity will have to be distributed between an increased number of components.

Interestingly, the assumption of our model that the total amount of resources is constant seems to hold under the conditions and time-scales we study. This is in contrast to fast growing organisms such as bacteria, for which resource availability is strongly linked to growth rate and rapidly regulated[12]. In such cases, it is stringent that models take into account the adaptation mechanisms of the host[15]. Understanding the mechanisms and time-scales of adaptation to resource burden in mammalian cells would provide valuable insight into how this key aspect is regulated and would make it possible to design circuits that exhibit robust long-term behavior.

Based on the model and experimental findings, we propose an incoherent feed-forward loop (IFFL) motif, introduced in **Fig. 5**, to address the problem of coupling between components. The IFFL is a control architecture that is known to robustify gene expression against input perturbations[20,27,28,48].

We designed miRNA-dependent IFFL circuits which are capable of balancing the overall expression of selected output genes, regardless of their number. Similar architectures that use other components, such as endoribonucleases (Poster by Jones et al. at mSBW 6.0 in chicago), to implement IFFL motifs, also exhibit robustness to exogenous gene loading. In the present study, we opted for miRNA-based implementations because of their high versatility. For example, miRNAs enable the experimenter to choose between endogenous or synthetic miRNAs. Endogenous miRNAs offer cell type specificity and do not require expression of additional genes for regulation. On the other hand, synthetic miRNAs enable portability of circuits across different cell lines.

It is important to note that the use of the same miRNA to regulate multiple targets might lead to depletion of the miRNA, giving rise to an additional trade-off. This trade-off is observed for competing endogenous RNA (ceRNA), which are naturally occurring regulatory elements known to regulate other RNAs by competing for miRNA-binding. Furthermore, the principle of ceRNA has also been demonstrated synthetically on a gene circuit[49]. Additionally, one should note that this method of robustifying against resource availability comes with a tradeoff in the maximally achievable expression levels, due to the repression by the miRNA. Nevertheless, we believe the proposed strategy should enable the construction of evermore complex gene circuits capable of performing evermore transformative function.

Understanding the role of cellular composition and the availability of resources will have important implications for biological studies, as well as for improved mammalian cell engineering. Previous studies have reported on the impact of translationational capacity with respect to mammalian cell function. For instance, it has been shown that recombinant mRNAs in engineered CHO cells sequester up to 15% of the total ribosomes[32], and that blood disorders like Diamond-Blackfan anemia are due to mutations in ribosomal proteins that impair the availability of ribosomes[50]. Here, we provide evidence that, in addition to competition over translational resources, also transcriptional resources limit the full potential of synthetic circuits. We hope that our global assessment will help to unleash mammalian synthetic biology's potential for diverse applications.

Mammalian cells and especially human derived cell lines are essential for biopharmaceutics production[32,51,52], because they ensure post-translational modifications and correct folding of the expressed proteins. Our findings suggest that, when choosing a host cell line, one of the key factors to consider should be its transcriptional and translational capacity[19], not only in terms of productivity, but also in terms of the ability of the cells to maintain their fitness over the production phase. In regard to this issue, enhancing the insulation of synthetic genetic circuits from the cellular metabolism is fundamental. With our study, we implement a system that can efficiently sustain the production of up to three (and probably more) output proteins, reducing the unwanted coupling among their expression levels and with other cellular processes. However, it should be noted that

our system cannot reach high productivity levels. Its purpose is to allow the concurrent production of more outputs in a predictable manner.

Another issue in transgene expression is reliable output production. This is not only challenged by epigenetic silencing[53], but also by the coupling to the expression of other genes, an issue that has not been addressed so far in mammalian cells. Poor insulation of transgene expression can have dramatic consequences also in gene and cell therapy, where the success of the treatment could be compromised if cellular fitness is impaired. Finally, studies of biological functions that employ system perturbations by exogenous gene expression can suffer from highly variable results and lack of accuracy due to non-optimal genetic circuit design. This is critical when performing quantitative analysis, where predictability and reproducibility are highly desirable. For example, protein levels measured by flow cytometry are usually normalized to the expression levels of the transfection marker, which is also used as a measure of transfection efficiency. The same issue applies to other quantitative analyses, such as qPCR and Western Blot. Our study presents a portable design capable of enhancing insulation of transgene expression and will thus contribute to the development of applications in mammalian synthetic biology.

Future studies could explore the molecular mechanisms involved in resource redistribution and host adaptation, as well as investigate which resource pools are the most limiting and impactful on gene expression in overloaded systems. Understanding the interplay of different mechanisms underlying gene expression burden will allow more precise and optimal approaches to the minimization of gene expression coupling. To this end, a variety of technologies that uncover the processes involved in the passage from DNA to protein at the molecular level could help to refine the design of IFFL topologies.

By shedding light on the physiological limitations that hamper the exploitation of the full potential of genetic circuits, we believe that this work opens a new chapter of mammalian synthetic biology. Our proposed IFFL topology proposed an easy and portable example of how to overcome these limitations. Since transgene expression is rapidly becoming a popular tool in industry as well as in cell therapy, we believe that our work will be beneficial in many applied fields and will lead to new venues of investigation of resource competition and strategies for its mitigation in mammalian cells.

Methods

Cell culture. HEK293T, U2OS and HeLa cells (ATCC) used in this study were maintained in Dulbecco's modified Eagle medium (DMEM, Gibco); H1299 were maintained in Roswell Park Memorial Institute medium (RPMI, Gibco). Both media were supplemented with 10% FBS (Atlanta BIO), 1% penicillin/streptomycin/L-Glutamine (Sigma-Aldrich) and 1% non-essential amino acids (HyClone). HEK239T cells (ATCC, strain number CRL-3216) used for part of this study were maintained in DMEM (Sigma-Aldrich or Gibco) supplemented with 10% FBS (Sigma-Aldrich), 1X GlutaMAX (Gibco) and 1 mM Sodium Pyruvate (Gibco). E14 mouse embryonic stem (ES) cells (Figure 5C) were grown in DMEM (Gibco) supplemented with 15% FBS (PAN Biotech; specifically for ES cell culture), 1% penicillin/streptomycin (Sigma-Aldrich), 1% non-essential amino acids (Gibco), 2 mM L-Glutamine (GlutaMAX; Gibco), 0.1 mM beta-mercaptoethanol (Sigma-Aldrich) and 100 U/mL Leukemia inhibitory factor (LIF; Preprotech). At every passage the media was additionally supplemented with fresh CHIR99021 to 3 uM and PD0390125 to 1 uM to support naïve pluripotency (2i conditions[54]). All labware used was coated with 0.1% gelatin (prepared ourselves) prior to plating the ES cells. The cells were maintained at 37 °C and 5% CO2.

Transfection. Transfections were carried out in 24-well plate for flow cytometry analysis or in a 12-well plate format for flow cytometry and qPCR analysis run on the same biological replicate. H1299, HeLa and HEK293T cells were transfected with Lipofectamine® 3000 (ThermoFisher Scientific) according to m, manufacturer's instructions and 300 ng of total DNA in 24 well plates. U2OS were transfected with FuGENE® (Promega) according to manufacturer's instructions and 400 ng of total DNA for each sample in 24-well plates. HEK293T cells were plated approximately 24 hours before transfection at 62500 to 75000 cells per well in 24-well plates. The transfection solution was prepared using Polyethylenimine (PEI) "MAX" (Mw 40000; Polysciences, Inc.) in a 1:3 (ug DNA to ug PEI) ratio with a total of 500 ng of plasmid DNA per well. Both DNA and PEI were diluted in Opti-MEM I reduced serum media (Gibco) before being mixed and incubated for 25 minutes prior to addition to the cells. E14 mouse embryonic stem cells were transfected using Lipofectamine® 2000 (ThermoFisher Scientific) in a 1:3 (ug DNA to ug Lipofectamine® 2000) with 300 ng of plasmid DNA per well. The transfection was performed on cells in suspension immediately after plating at approximately 30000 cells per well. All wells were coated with 0.1% Gelatin before addition of the cells.

Flow cytometry and data analysis. HEK293T, U2OS and HeLa cells were analyzed with a BD Facsaria™ cell analyzer (BD Biosciences) using 488 and 561 lasers. For each sample >20000 singlet

events were collected and fluorescence data were acquired with the following cytometer settings: 488 nm laser and 530/30 nm bandpass filter for EGFP, 561 nm laser and 610/20 nm filter for mKate. Cells transfected in 12-well plate were washed with DPBS, detached with 100 µL of Trypsin-EDTA (0.25%) and resuspended in 600 µL of DPBS (Thermo Fisher). 200 µL of cell suspension were used for flow cytometry and 400 µL for RNA extraction. HEK293T cells were measured 48 hours after transfection on a BD LSRFortessa Special Order and Research Product (SORP) cell analyzer. mCitrine fluorescence was excited via a 488 nm laser and was detected through a 530/11 nm bandpass filter. mRuby3 was excited via 561 nm laser and measured through a 610/20 nm bandpass filter. miRFP670 was excited at 640 nm and measured through a 670/14 nm bandpass filter. The cells were collected for measurement by washing with with DBPS (Sigma-Aldrich or Gibco) and detaching in 100 to 180 µL of Accutase solution (Sigma-Aldrich). For each sample between 75000 to 200000 singlet events were collected. Fluorescence intensity in arbitrary units (AU) was used as a measure of protein expression. For each experiment a compensation matrix was created using unstained (wild type cells), and single-color controls (mKate only, EGFP only). Population of live cells and single cells were selected according to FCS/SSC parameters. Data analysis was performed with Cytoflow.

DNA cloning and plasmid construction. Plasmid vectors carrying gene cassettes were created using In-Fusion HD cloning kit (Clonetch), Gibson Assembly, via digestion and ligation or using the yeast toolkit (YTK)[55] with custom parts for mammalian cells. Gibson Assembly master mixes were created from Taq DNA Ligase (NEB), Phusion High-Fidelity DNA Polymerase (NEB) and T5 Exonuclease (Epicentre) in 5X isothermal buffer (Supplementary Table 6). Ligation reactions were performed in 1:2-5 molar ratios of plasmid backbone:gene insert starting with 50 to 100 ng of vector backbone digested with selected restriction enzymes. Assemblies using the YTK were performed according to the original publication[55]. Newly created constructs were transformed into XL10-Gold or TOP10 *E. coli* strains.

For plasmids with miRNA target sites, the target sequences were selected using miRBase database (http://www.mirbase.org/). All plasmids were confirmed by sequencing analysis and deposited to addgene.

mRNA extraction and reverse transcription. RNA extraction was performed with E.Z.N.A.® Total RNA Kit I (Omega Bio-tek). Protocol was followed according to manufacturer's instructions and RNA was eluted in 30 µL of RNAse free water. RNA samples were conserved at -80°C.

PrimeScript RT Reagent Kit with gDNA Eraser – Perfect Real Time (Takara) was used according to manufacturer's instructions. The protocol was performed on ice in a RNAse free environment to

avoid RNA degradation. A negative control without PrimeScript RT Enzyme Mix I was always prepared to investigate genomic DNA contamination.

qPCR. Fast SYBR Green Master Mix (ThermoFisher Scientific) was used to perform qPCR of cDNAs obtained from 500 ng of RNA and diluted 1:5. Samples were loaded in MicroAmp™ Fast Optical 96-Well Reaction Plate (0.1 mL) and the experiment was carried out with a CFX96 Touch Real-Time PCR Detection System (BioRad) machine. Each well contained 20 µL of final volume (7 µL SYBR Green Master Mix, 10 µL ddHO, 1 µL of each primer, 1 µL of template). Also, a control without template (blank) was set. Primers were designed to amplify a region of 60-200 pb (Supplementary table 5) and with a temperature of annealing of annealing between 50 and 65 °C 72. Data were analysed using the Comparative Ct Method according to Applied Biosystems Protocols.

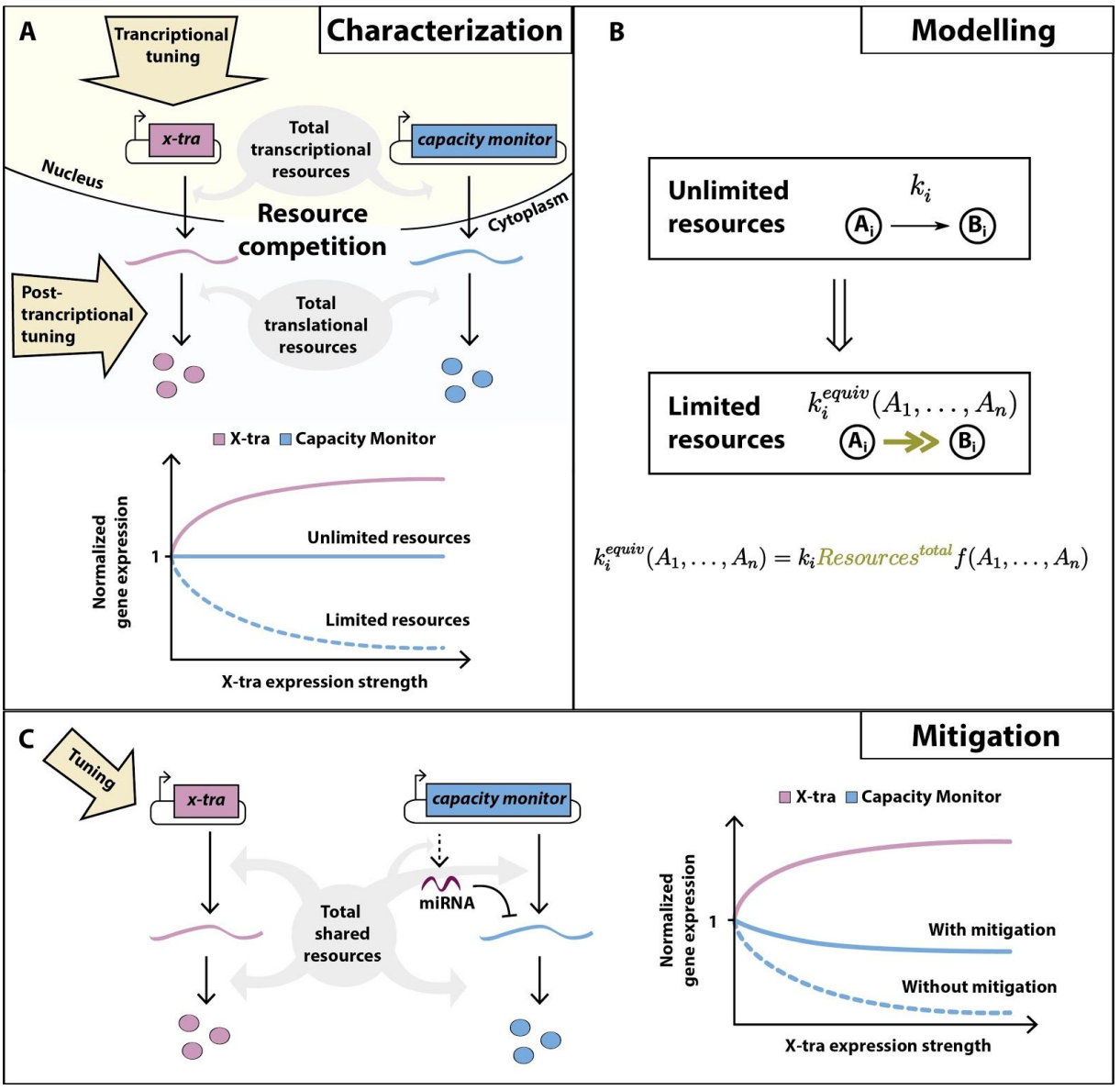

Fig 1. Graphical abstract. (A) Characterization of gene expression burden. Expression of independent exogenous genes impacts on host cellular resources. Thus, perturbations in one gene's expression (hereby named *X-tra*) affect the expression of a second gene (hereby named *capacity monitor*). **(B) Modelling of gene expression in a resource limited environment.** Modelling of gene expression is generally performed under the assumption of unlimited resources. A simple framework enables the straightforward transformation of such a model to a system that incorporates resources explicitly. The transformation involves a simple function that scales the original reaction rate. **(C) Mitigation of gene expression burden.** A simple microRNA-based circuit motif is capable of mitigating the burden-induced coupling of *X-tra* and the *capacity monitor*.

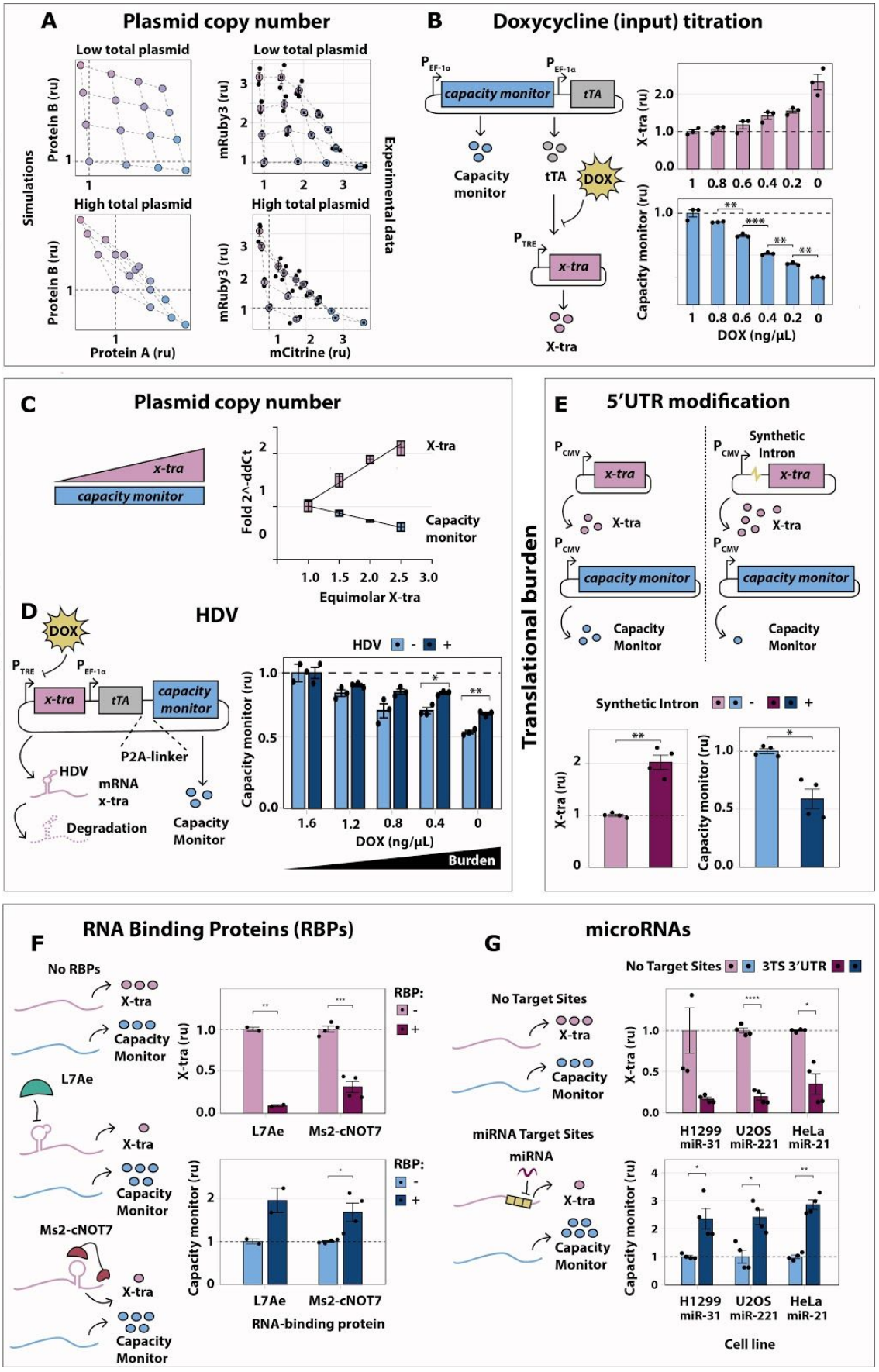
Generic burden
A
Plasmid copy number
Low total plasmid
High total plasmid
Simulations
Experimental data
Protein B (ru)
Protein A (ru)
mRuby3 (ru)
mCitrine (ru)
B
Doxycycline (input) titration
capacity monitor
tTA
Capacity monitor
DOX
x-tra
X-tra
X-tra (ru)
Capacity monitor (ru)
DOX (ng/µL)
Transcriptional burden
C
Plasmid copy number
x-tra
capacity monitor
Fold 2^-ddCt
Equimolar X-tra
D
HDV
P2A-linker
mRNA x-tra
Degradation
Capacity Monitor
DOX (ng/µL)
Burden
Translational burden
E
5'UTR modification
Synthetic Intron
Capacity Monitor
Resource redistribution
F
RNA Binding Proteins (RBPs)
No RBPs
L7Ae
Ms2-cNOT7
RBP:
RNA-binding protein
G
microRNAs
No Target Sites
3TS 3'UTR
miRNA Target Sites
miRNA
H1299 miR-31
U2OS miR-221
HeLa miR-21
Cell line

Fig 2. Burden imposed by genetic circuits in mammalian cells. (A) Left. A model accounting for the limited resources (**Supplementary Note 1**) shows the correlation of independent gene's expression. As the total plasmid amount increases, the total protein expression plateaus. Top right**.** Titration of two plasmids constitutively expressing the fluorescent proteins mCitrine and mRuby3 in ratios from 1:4 to 4:1, 50ng of total DNA. Bottom right. Titration of two plasmids constitutively expressing the fluorescent proteins mCitrine and mRuby3 in ratios from 1:4 to 4:1, 500ng of total DNA. The model effectively recapitulates the saturation of gene expression that we observed for a total amount of 50 or 500 ng of plasmids. **(B)** Inhibition of tTA by Dox treatment further unveil the hidden coupling of expression levels among two genes. Two plasmids were co-transfected, one constitutively expressing *capacity monitor* and tTA from a strong constitutive promoter and the other expressing *X-tra* from a tTA responsive promoter. Dox was added to the cells in decreasing concentrations, derepressing the expression of mRuby3. The data shows that *capacity monitor* levels counterbalance the increase in *X-tra* expression. The Data is normalized to the expression levels of each gene in absence of Dox. N = 3 biological replicates. **(C)** Transcriptional burden. qPCR analysis was performed on two plasmids encoding an *X-tra* and a *capacity monitor* respectively, expressed at different molar ratios. As the *X-tra* increases, the mRNA levels of the *capacity monitor* decreases. N=4 biological replicates **(D)** Increasing the expression of a translationally incompetent mRNA by Dox suggest transcriptional burden. Decreased expression of the mRuby3 *(capacity monitor)* with increasing expression of mCitrine *(X-tra)* with an inactive or active HDV ribozyme placed after its start codon suggest a sequestration of transcriptional resources. N=3 biological replicates. **(E)** Translational burden. We included a synthetic intron in the 5'UTR of a transcript to increase its translation and to selectively observe translational burden. As expected the synthetic intron shows higher translational load as compared to the control. Higher *X-tra* expression again leads to reduced *capacity monitor* levels. N=6 biological replicates. **(F)** Post-transcriptional regulators impact translational burden. The RBPs L7Ae and Ms2-cNOT7, downregulate *X-tra* expression by binding the 5' and 3'UTR respectively. Our data show that repressed *X-tra* expression leads to increased *capacity monitor* levels, indicating a shift in resources in the host cells. N = 4 biological replicates. **(G)** Similar to (F) also miRNAs can reallocate resources by targeting mRNA which include miRNA target sites (TS) in the 3' UTR. We engineered *X-tra* transcripts with 3TS that respond to either miRNA miR-31, miR-221 or miR-21, highly expressed respectively in H1299, U2-OS/HEK and HeLa cells. Co-transfection of the *X-tra* with the *capacity monitor* shows that *capacity monitor* levels are higher when the *X-tra* is downregulated by miRNAs, further demonstrating a negative correlation in protein expression. N = 4 biological replicates. Flow cytometry data were acquired 48 hours post-transfection and is plotted +/- SE. SE: standard error. ru: relative units. unpaired T-test.p-value: **** < 0.0001, *** < 0.0005, ** < 0.005, * < 0.05.

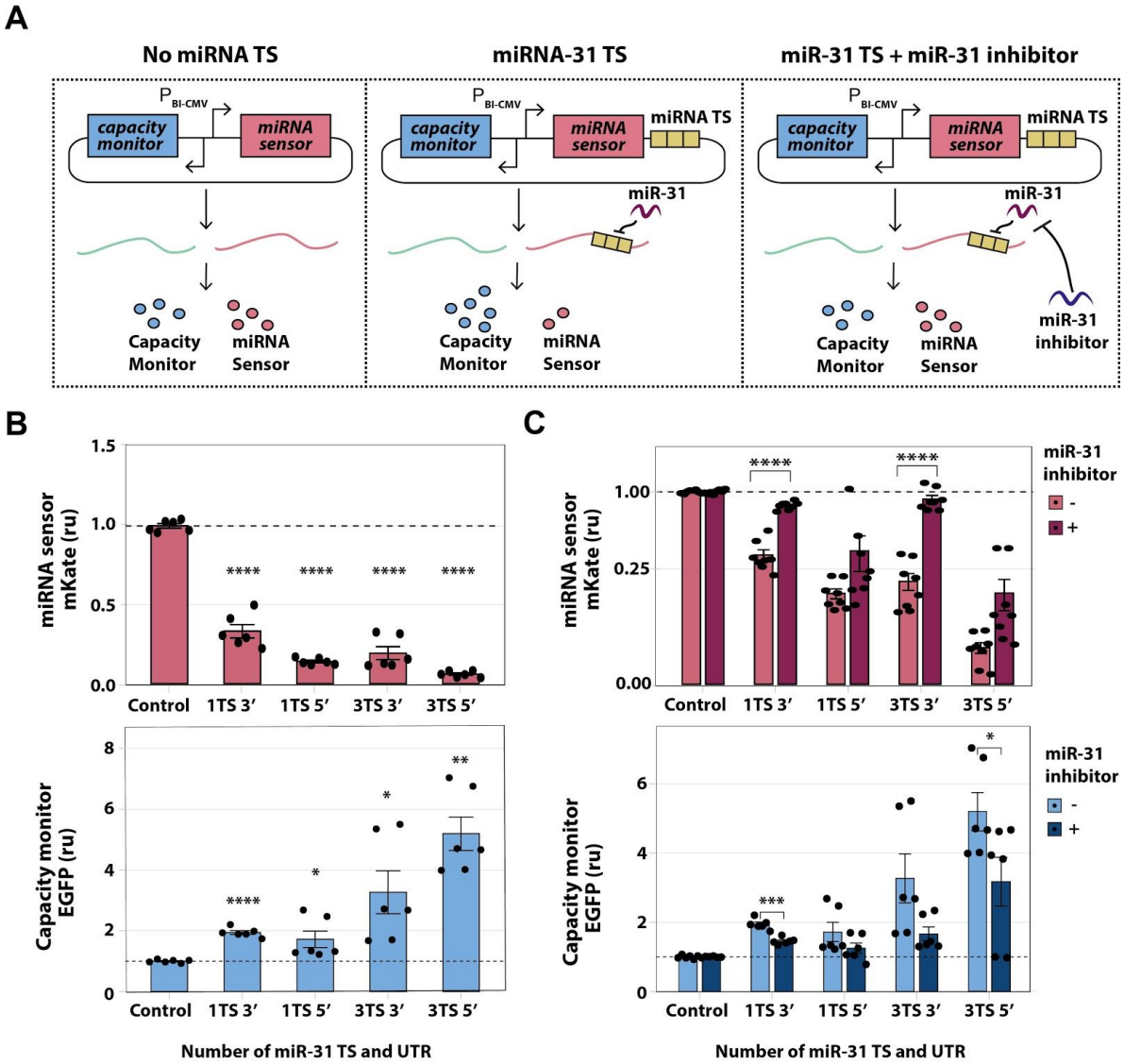

Fig 3. Impact of miRNA target sites number and location on burden. (A) Schematics of experimental design to infer miRNA-mediated cellular resources distribution. Two fluorescent proteins EGFP (*capacity monitor*) and mKate (*miRNA sensor*) are encoded on the same bi-directional CMV promoter plasmid. 1 or 3 TS for miR-31 (TS) are added either in the 3' or 5'UTR of mKate. Control: no miR31 TS. Hypothesis: in the absence of miR-31 regulation, *capacity monitor* and *miRNA sensor* are expressed to a certain level (left). In the presence of miR-31, lower *miRNA sensor* levels correlate with higher *capacity monitor* expression (middle). This condition is reversed by a miR-31 inhibitor (right). **(B)** Fold change of *miRNA sensor* (mKate) and *capacity monitor* (EGFP) compared to control (set to 1). EGFP levels increase up to 5 fold with the strongest downregulation of mKate (3TS 5'UTR). **(C)** When miR31 activity was impaired by a miR-31 inhibitor, the rescue of mKate expression corresponds to reduced EGFP levels, whereas both fluorescent proteins do not vary in the control. Data are expressed in logarithmic base 2 scale. Flow cytometry data was acquired 48 hours post-transfection and is plotted +/- SE. SE:standard error. Ru: relative units. N = 6 biological replicates, unpaired T-test. p-value: **** < 0.0001, *** < 0.0005, ** < 0.005, * < 0.05.

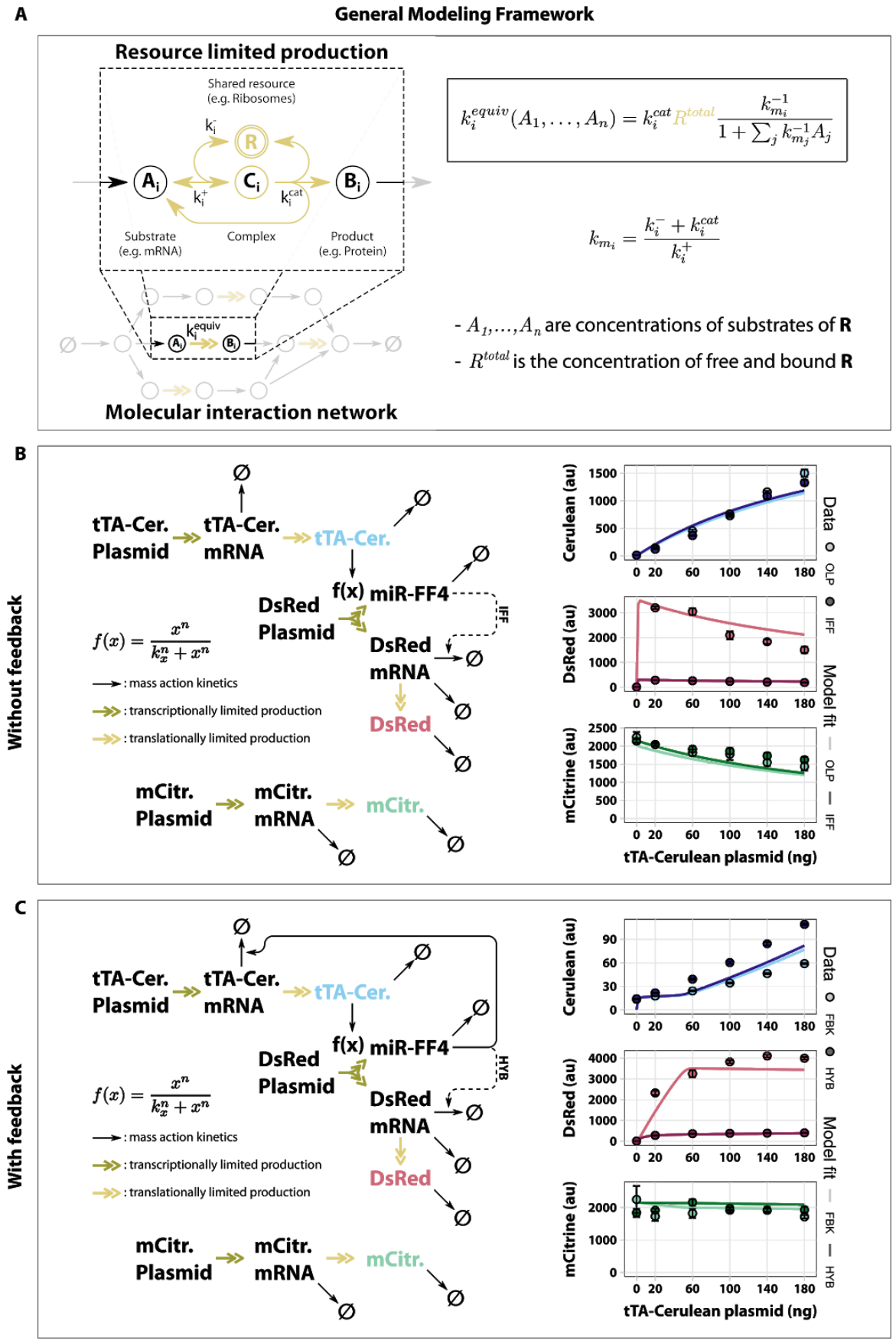
A
General Modeling Framework
Resource limited production
Shared resource
(e.g. Ribosomes)
Substrate
(e.g. mRNA)
Complex
Product
(e.g. Protein)
Molecular interaction network
- $A_1,\ldots,A_n$ are concentrations of substrates of R
- R^{total} is the concentration of free and bound R
B
Without feedback
tTA-Cer. Plasmid
tTA-Cer. mRNA
tTA-Cer.
DsRed Plasmid
f(x)
miR-FF4
IFF
DsRed mRNA
DsRed
: mass action kinetics
: transcriptionally limited production
: translationally limited production
mCitr. Plasmid
mCitr. mRNA
mCitr.
Cerulean (au)
DsRed (au)
mCitrine (au)
tTA-Cerulean plasmid (ng)
Data
Model fit
OLP
IFF
C
With feedback
HYB
FBK

Fig 4. A mathematical network model that factors in burden-induced coupling reproduces counter-intuitive experimental observations. (A) General framework for transforming molecular interaction network models. Existing models of molecular interaction networks can be transformed to include shared limiting resources by substituting k_i, the reaction rate of a resource limited production, with k^{equiv}_i. Shown above an exemplary resource limited production are the detailed interactions between the substrate and the shared resource. The substrate A_i binds and unbinds the shared resource R with rates k^+_i and k^-_i, respectively. Binding of the substrate A_i to the shared resource R forms a complex C_i which releases the substrate A_i, the shared resource R and the product B_i with rate k^{cat}_i. Under quasi steady-state assumptions the equivalent reaction rate k^{equiv}_i can be expressed as shown on the right. For simplification the parameters k^+_i, k^-_i and k^{cat}_i are lumped into k_{mi}. **(B) Limited shared translational resources reproduces non-monotonous dose-response in open-loop and incoherent feedforward circuit topologies.** On the left a graphical representation of a model for both the open-loop (OLP) and incoherent feedforward (IFF) topologies from Lillacci et al.[41] is shown. Transcriptional activation is modelled by a Hill-type function as shown. The solid arrows denote reactions assumed to follow the law of mass action. The model incorporates translational resources as introduced in panel A. These reactions are depicted as double headed arrows. The model was fit to data obtained by transiently transfecting HEK293T cells with increasing amounts of plasmid encoding tTA-Cerulean. The data and the fit are shown on the right. For a full description of the model see Supplementary Note 3 . **(C) Limited shared translational resources reproduces non-monotonous dose-response in feedback and hybrid circuit topologies.** The model shown on the left is the same as in panel B with an additional negative feedback from miR-FF4 to tTA-mRNA. These topologies correspond to the feedback (FBK) and hybrid (HYB) topologies from Lillacci et al.[41]. The activation of gene expression by tTA-Cerulean is modeled by a Hill-type function as shown in the center of the figure. Reactions with double headed arrows denote resource limited production reactions as introduced in panel A. Solid arrows are assumed to follow the law of mass action. The model was fit to experimental data obtained from transient transfections with increasing amounts of plasmid encoding tTA-Cerulean. A full description of the model equations can be found in Supplementary Note 4. Data were analysed 48 hours after transfection and is plotted +/- SE. SE: standard error. N = 3 biological replicates

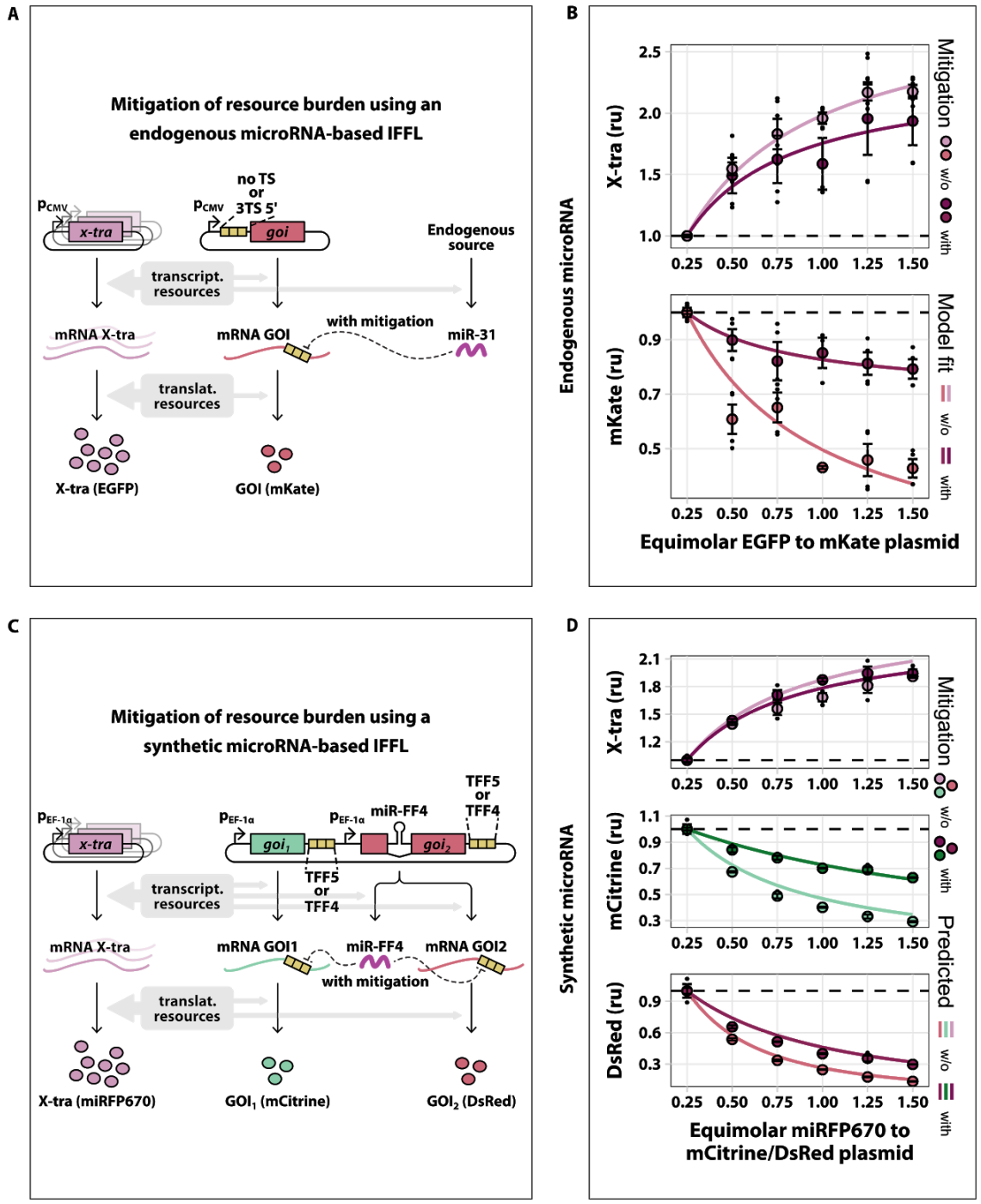

Fig 5. Mitigating the effects of resource limitation with microRNA. (A) Schematic representation of the mitigation system based on endogenous microRNA. At high copies of the *X-tra* resources are drawn away from the production of the *GOI* and the microRNA miR-31. By sensing the resource availability and repressing the *GOI* less when there are fewer resources, the microRNA approximately cancels out the effect of limited resources. **(B)** Two plasmids were co-transfected which respectively express the *X-tra* gene (EGFP, panel A) and mKate (panel A) from strong constitutive CMV promoters. To limit the amount of resources available to the expression of mKate, the copy number of the X-tra plasmid was increased relative to the copy number of mKate. A mKate variant, which encodes three targets for the microRNA miR-31 in it's 5' UTR, mitigates

effects due to resource sharing when compared to a variant lacking microRNA targets. N = 4 biological replicates. **(C)** Schematic of the mitigation system based on synthetic microRNA. In the presence of many copies of the *X-tra* gene, resources are drawn away from the production of both of the *GOIs* and the miR-FF4. Due to lower production of miR-FF4 the *GOIs* are less repressed. This compensates for the reduced availability of resources. **(D)** A plasmid encoding both the fluorescent protein mCitrine (*GOI*, panel C) and an intronically expressed microRNA (miR-FF4, panel C) under the control of strong constitutive promoters was co-transfected with increasing amounts of a plasmid expressing the *X-tra* gene (miRFP670, panel C) also under the control of a strong constitutive promoter. The impact of resource limitation on both *GOIs* was reduced when they contained three miR-FF4 targets in their 3' UTRs compared to when they contained three mismatched miR-FF5 targets. At least N = 2 biological replicates. Flow cytometry data were acquired 48 hours post-transfection and are plotted +/- SE. SE:standard error. ru: relative units.

References

1. Carrión, A. M., Link, W. A., Ledo, F., Mellström, B. & Naranjo, J. R. DREAM is a Ca2+-regulated transcriptional repressor. *Nature* **398**, 80–84 (1999).
2. Bamforth, S. D. *et al.* Cardiac malformations, adrenal agenesis, neural crest defects and exencephaly in mice lacking Cited2, a new Tfap2 co-activator. *Nat. Genet.* **29**, 469–474 (2001).
3. Liu, W. *et al.* Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating β-catenin/TCF signalling. *Nature Genetics* vol. 26 146–147 (2000).
4. Mendell, J. R. *et al.* Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N. Engl. J. Med.* **377**, 1713–1722 (2017).
5. Stuible, M. *et al.* Optimization of a high-cell-density polyethylenimine transfection method for rapid protein production in CHO-EBNA1 cells. *J. Biotechnol.* **281**, 39–47 (2018).
6. Xie, M. *et al.* β-cell-mimetic designer cells provide closed-loop glycemic control. *Science* **354**, 1296–1301 (2016).
7. Xiang, Y., Dalchau, N. & Wang, B. Scaling up genetic circuit design for cellular computing: advances and prospects. *Nat. Comput.* **17**, 833–853 (2018).
8. Bernardo, D. di, di Bernardo, D., Marucci, L., Menolascina, F. & Siciliano, V. Predicting Synthetic Gene Networks. *Methods in Molecular Biology* 57–81 (2012) doi:10.1007/978-1-61779-412-4_4.
9. Borkowski, O., Ceroni, F., Stan, G.-B. & Ellis, T. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Curr. Opin. Microbiol.* **33**, 123–130 (2016).

10. Ceroni, F., Algar, R., Stan, G.-B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415–418 (2015).

11. Li, G.-W., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624–635 (2014).

12. Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).

13. Carbonell-Ballestero, M., Garcia-Ramallo, E., Montañez, R., Rodriguez-Caso, C. & Macía, J. Dealing with the genetic load in bacterial synthetic biology circuits: convergences with the Ohm's law. *Nucleic Acids Research* vol. 44 496–507 (2016).

14. Qian, Y., Huang, H.-H., Jiménez, J. I. & Del Vecchio, D. Resource Competition Shapes the Response of Genetic Circuits. *ACS Synth. Biol.* **6**, 1263–1272 (2017).

15. Weiße, A. Y., Oyarzún, D. A., Danos, V. & Swain, P. S. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E1038–47 (2015).

16. Gyorgy, A. *et al.* Isocost Lines Describe the Cellular Economy of Genetic Circuits. *Biophys. J.* **109**, 639–646 (2015).

17. Huang, H.-H., Qian, Y. & Del Vecchio, D. A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nat. Commun.* **9**, 5415 (2018).

18. Ceroni, F. *et al.* Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387–393 (2018).

19. Kafri, M., Metzl-Raz, E., Jona, G. & Barkai, N. The Cost of Protein Production. *Cell Rep.* **14**, 22–31 (2016).

20. Lillacci, G., Benenson, Y. & Khammash, M. Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic Acids Res.* **46**, 9855–9863 (2018).

21. Gyorgy, A. & Del Vecchio, D. Limitations and trade-offs in gene expression due to competition for shared cellular resources. *53rd IEEE Conference on Decision and Control* (2014) doi:10.1109/cdc.2014.7040238.

22. Mather, W. H., Hasty, J., Tsimring, L. S. & Williams, R. J. Translational cross talk in gene networks. *Biophys. J.* **104**, 2564–2572 (2013).

23. Chou, T. C. & Talaly, P. A simple generalized equation for the analysis of multiple inhibitions of Michaelis-Menten kinetic systems. *J. Biol. Chem.* **252**, 6438–6442 (1977).

24. Jens, M. & Rajewsky, N. Competition between target sites of regulators shapes post-transcriptional gene regulation. *Nat. Rev. Genet.* **16**, 113–126 (2015).

25. Siciliano, V. *et al.* MiRNAs confer phenotypic robustness to gene networks by suppressing biological noise. *Nat. Commun.* **4**, 2364 (2013).

26. Carignano, A., Mukherjee, S., Singh, A. & Seelig, G. Extrinsic Noise Suppression in Micro RNA mediated Incoherent Feedforward Loops. *bioRxiv* 422394 (2018) doi:10.1101/422394.

27. Strovas, T. J., Rosenberg, A. B., Kuypers, B. E., Muscat, R. A. & Seelig, G. MicroRNA-based single-gene circuits buffer protein synthesis rates against perturbations. *ACS Synth. Biol.* **3**, 324–331 (2014).

28. Bleris, L. *et al.* Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Mol. Syst. Biol.* **7**, 519 (2011).

29. Li, X., Cassidy, J. J., Reinke, C. A., Fischboeck, S. & Carthew, R. W. A microRNA imparts robustness against environmental fluctuation during development. *Cell* **137**, 273–282 (2009).

30. McBride, C., Shah, R. & Del Vecchio, D. The Effect of Loads in Molecular Communications. *Proc. IEEE* **107**, 1369–1386 (2019).

31. Lu, J., Zhang, F. & Kay, M. A. A Mini-intronic Plasmid (MIP): A Novel Robust Transgene Expression Vector In Vivo and In Vitro. *Mol. Ther.* **21**, 954–963 (2013).

32. Kallehauge, T. B. *et al.* Ribosome profiling-guided depletion of an mRNA increases cell growth rate and protein secretion. *Sci. Rep.* **7**, 40388 (2017).

33. Wroblewska, L. *et al.* Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery. *Nat. Biotechnol.* **33**, 839–841 (2015).

34. Cella, F., Wroblewska, L., Weiss, R. & Siciliano, V. Engineering protein-protein devices for

multilayered regulation of mRNA translation using orthogonal proteases in mammalian cells. *Nature Communications* vol. 9 (2018).

35. Gam, J. J., Babb, J. & Weiss, R. A mixed antagonistic/synergistic miRNA repression model enables accurate predictions of multi-input miRNA sensor activity. *Nat. Commun.* **9**, 2430 (2018).
36. Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–488 (2013).
37. Thomson, D. W., Bracken, C. P. & Goodall, G. J. Experimental strategies for microRNA target identification. *Nucleic Acids Res.* **39**, 6845–6853 (2011).
38. Robertson, B. *et al.* Specificity and functionality of microRNA inhibitors. *Silence* **1**, 10 (2010).
39. Chou, T. C. & Talaly, P. A simple generalized equation for the analysis of multiple inhibitions of Michaelis-Menten kinetic systems. *J. Biol. Chem.* **252**, 6438–6442 (1977).
40. Rondelez, Y. Competition for catalytic resources alters biological network dynamics. *Phys. Rev. Lett.* **108**, 018102 (2012).
41. Lillacci, G., Benenson, Y. & Khammash, M. Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic Acids Res.* **46**, 9855–9863 (2018).
42. Zechner, C., Seelig, G., Rullan, M. & Khammash, M. Molecular circuits for dynamic noise filtering. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 4729–4734 (2016).
43. Huang, M., Bao, J., Hallström, B. M., Petranovic, D. & Nielsen, J. Efficient protein production by yeast requires global tuning of metabolism. *Nat. Commun.* **8**, 1131 (2017).
44. Cohen, R. N., van der Aa, M. A. E. M., Macaraeg, N., Lee, A. P. & Szoka, F. C., Jr. Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection. *J. Control. Release* **135**, 166–174 (2009).
45. Milo, R. & Phillips, R. *Cell Biology by the Numbers*. (Garland Science, 2015).
46. Nikolados, E.-M., Weiße, A. Y., Ceroni, F. & Oyarzún, D. A. Growth Defects and Loss-of-Function in Synthetic Gene Circuits. *ACS Synth. Biol.* **8**, 1231–1240 (2019).
47. Qian, Y. & Del Vecchio, D. Effective interaction graphs arising from resource limitations in gene

networks. *2015 American Control Conference (ACC)* (2015) doi:10.1109/acc.2015.7172024.

48. Segall-Shapiro, T. H., Sontag, E. D. & Voigt, C. A. Engineered promoters enable constant gene expression at any copy number in bacteria. *Nat. Biotechnol.* **36**, 352–358 (2018).
49. Yuan, Y. *et al.* Model-guided quantitative analysis of microRNA-mediated regulation on competing endogenous RNAs using a synthetic gene circuit. *Proceedings of the National Academy of Sciences* vol. 112 3158–3163 (2015).
50. Khajuria, R. K. *et al.* Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis. *Cell* **173**, 90–103.e19 (2018).
51. Hacker, D. L., Baldi, L., Adam, M. & Wurm, F. M. Transient Gene Expression in Mammalian Cells: Promises and Challenges for Medical Biotechnology. *Encyclopedia of Industrial Biotechnology* 1–8 (2009).
52. Walsh, G. Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.* **32**, 992–1000 (2014).
53. Rajeev Kumar, S. & Anunanthini, P. Epigenetic silencing in transgenic plants. *Front. Plant Sci.* (2015).
54. Ying, Q.-L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* vol. 453 519–523 (2008).
55. Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* **4**, 975–986 (2015).