






COMMENT



<https://doi.org/10.1038/s41467-021-22795-9>

OPEN

A call for caution in analysing mammalian co-transfection experiments and implications of resource competition in data misinterpretation

Roberto Di Blasi ^{1,2,4}, Masue M. Marbiah ^{1,2,4}, Velia Siciliano ³, Karen Polizzi ^{1,2} & Francesca Ceroni ^{1,2}✉

Transient transfections are routinely used in basic and synthetic biology studies to unravel pathway regulation and to probe and characterise circuit designs. As each experiment has a component of intrinsic variability, reporter gene expression is usually normalized with co-delivered genes that act as transfection controls. Recent reports in mammalian cells highlight how resource competition for gene expression leads to biases in data interpretation, with a direct impact on co-transfection experiments. Here we define the connection between resource competition and transient transfection experiments and discuss possible alternatives. Our aim is to raise awareness within the community and stimulate discussion to include such considerations in future experimental designs, for the development of better transfection controls.

Resource competition in engineered hosts

Expression of foreign DNA in microbes or mammalian cells for the expression of proteins of interest is common practice in molecular, cellular and synthetic biology for the study of gene function (foundational research) or for the production of new therapeutics or commercial biologics (translational research). Pioneering work in bacteria unveiled that cells possess a finite amount of intracellular resources available for gene expression¹ (e.g. RNA polymerases, ribosomes, proteases, energy sources) and competition for the utilisation of such resources arises between exogenous genes, as well as between exogenous and endogenous genes. Resource competition implies that if one transcriptional unit utilises more resources, this leads to a decrease in the resources available to operate cellular physiological processes or to express other synthetic transcriptional units, leading to an artificial decrease in their expression levels (Fig. 1). Thus, expression cassettes, otherwise independently regulated, become interlinked by

¹Department of Chemical Engineering, Imperial College London, South Kensington Campus, London, UK. ²Imperial College Centre for Synthetic Biology, South Kensington Campus, London, UK. ³Synthetic and Systems Biology lab for Biomedicine, Istituto Italiano di Tecnologia-IIT, Largo Barsanti e Matteucci, Naples (ITA), Italy. ⁴These authors contributed equally: Roberto Di Blasi, Masue M. Marbiah. ✉email: f.ceroni@imperial.ac.uk

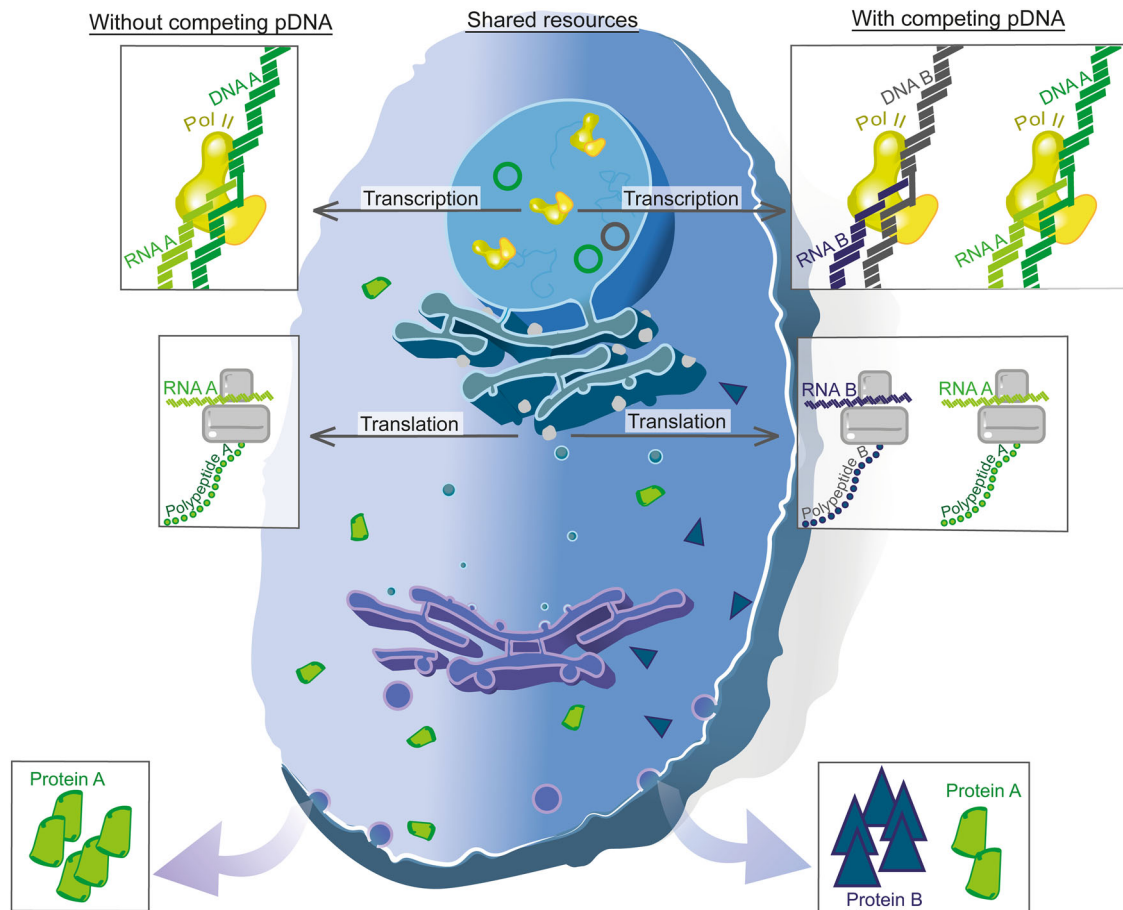


Fig. 1 A comparison of resource allocation in mammalian cells upon transfection of a genetic construct with and without a co-expressed control. Two scenarios are shown in parallel where a plasmid (pDNA) harbouring Protein A is transfected alone (left) or co-transfected with Protein B (right). In the nucleus, pDNA are processed by a finite number of RNA polymerase II (Pol II) molecules; therefore heterologous (A and B) and endogenous genes must compete for its availability. In a scenario with greater competition (right), less mRNA is transcribed from DNA A compared to a scenario with less competition (left). To compound this issue, cells also contain a finite number of ribosomes. As such, competition for translational machinery leads to unpredictable production of A and B proteins.

resource usage. Gyorgy et al. demonstrated that competition for gene expression resources between two independent transcriptional units causes a counterintuitive coupling in their expression in *E. coli*². The authors transformed *E. coli* with a plasmid bearing a constitutive GFP expression cassette and an AHL-inducible RFP expression cassette and observed how increasing inducer concentration led to increased RFP, but decreased GFP expression, despite the two proteins being functionally independent. Recently, Ceroni et al. quantified how the presence of synthetic constructs can cause decreased growth, global transcriptional rearrangements and emergence of escape mutants in engineered *E. coli* populations^{3,4}. In mammalian cells Huliak et al. observed interference between P_{CMV} -driven EGFP expression and other co-transfected luciferase-expressing promoters⁵. The authors reported a dramatic decrease in luminescence following co-transfection of the P_{CMV} -EGFP plasmid with increasing amounts of the latter leading to more drastic interference. Since different promoters displayed the same behaviour, resource competition was suggested to be the likely cause of the phenomenon.

More recently, resource competition in mammalian cells has been quantified and shown to undermine the performance of genetic circuits, prompting a deeper investigation of the issue⁶⁻⁸. Frei et al. demonstrated that, at very least, both transcription and translation are affected by intracellular resource availability⁷. In a

first pilot experiment, the authors showed that co-transfection of two functionally independent plasmids, one bearing a constitutively expressed reporter and the other a fluorescent protein under the control of doxycycline (Dox) via the TET-OFF promoter, led to increased constitutive reporter expression following Dox repression. The authors reasoned that this behaviour could be explained by the reliance of both modules on the same finite pool of cellular resources. Hence, when one module is repressed by Dox induction, more resources become available for the other one to be expressed. In a paper by Jones et al.⁸ resources limitation in the context of transient transfections was probed using a genetic model system in which the GAL4 DNA binding domain (DBD) was fused to activation domains of different strength. By looking at the expression of a co-transfected constitutive reporter driven by different promoters, the authors proved that the stronger the activation by the synthetic transcription factor, the more resource demanding the synthetic device, and the sharper the decrease in the reporter expression. Importantly, these works show that, even if to a different extent, resource limitation is independent of the cell type and the promoter used, pointing to the need for a solution that can be generally applied to several cellular contexts.

If overlooked and not included in experimental analysis, the effects of resource competition, both at the level of the host cell and the construct, can represent a source of variability⁹⁻¹².

In light of these observations, we are compelled to re-examine experimental co-transfection designs where resource competition may lead to data misinterpretation.

Transient transfection in mammalian cell experiments

Transient transfection is used for the expression of recombinant proteins or prototyping synthetic gene circuits to avoid the lengthy procedures involved in stable cell line generation. On a higher level and beyond the synthetic biology field, transient transfections provide a general means to induce gene perturbation in fundamental biology experiments and to study the overall response of the observed systems. In contrast to stably integrated heterologous genes, transiently transfected genes are episomal and are therefore lost as cells proliferate. Further, transfection efficiency has a narrow optimal range, relying on DNA complex formation, solution pH, cell membrane condition and cell type¹³ and differences in transfection efficiency and DNA copy number may confound experimental results. Finally, individual cells from the same batch may exhibit different ability to produce recombinant protein¹⁴, adding additional variability. To avoid experimental biases originating from these sources of variability, researchers often use a transfection control for normalisation. This longstanding and widespread practice is similar to that described by Huliak et al., where a constitutively expressed reporter gene is co-transfected with the gene of interest either on a separate plasmid or as a second cassette within the same plasmid. The transfection control can serve as a marker for plasmid uptake within the population. It also provides confidence in data interpretation by facilitating normalisation of the expression of the gene of interest against a constitutively expressed reporter. By using the same transfection control with the test conditions, variability in transfection efficiency can be accounted for through the above normalisation procedure, allowing the user to capture the behaviour of interest.

Drawbacks in the use of transfection controls

After transfection controls were adopted as a standard methodology, it became clear that there were issues with the practice. While some of the observed effects of co-transfection were linked to the specific constructs adopted in the experiments (e.g. interference between adopted promoter pairs, specific sequences that can respond to endogenous transcription factors, and presence of non-specific transcription events from the plasmid sequence imposing extra load and toxicity to the cells^{15–17}), the fluctuation in the transfection control expression due to resource loading arose as a systemic effect to be taken into account^{18–20}.

An example comes from the characterisation of promoter activity in mammalian cells. While determining the activity of putative DNA promoter sequences, Farr et al. noticed that the expression of the transfection control plasmid used to account for variability in transfection efficiency was heavily suppressed following co-transfection with the cognate positive control in HeLa cells¹⁸. The authors noted that this would result in overestimation of the promoter activity in the positive control compared to the test sequences. Similar problems can be envisioned when calculating the dose-response curve of inducible promoters, as displayed by Gyorgy et al. and Frei et al. Therefore, any normalisation against the transfection control will yield overestimated normalised expression values at high inducer concentrations.

The impact of misinterpreted data goes beyond the experiment it pertains to. Indeed, the current direction of synthetic biology and cell engineering toward standardisation²¹ and the ability to share genetic parts²² requires reliable characterisation of the behaviour of each component. Identifying methods for the in vivo

characterisation of genetic parts which are not confounded by resource competition is necessary to enable the development of predictable designs for cellular engineering.

Considering the mounting evidence for resource competition highlighted above, we believe it is now time for a deeper debate focussing on the development of novel and improved approaches for the normalisation of biological data that are robust to resource competition.

Routes toward more robust transfection controls

We discuss below recent experimental and theoretical approaches that we believe can help address the problem (Fig. 2). We highlight their advantages and disadvantages and describe the scenarios in which they can be adopted. However, we do note that this is very much an open area of research in which potential novel solutions are yet to be discovered and existing alternatives are still limited.

Orthogonal synthetic modules

One avenue to address this problem is the adoption of synthetic modules that are fully orthogonal to the cellular machinery, such as dedicated unnatural ribosomes that can only translate a specific set of mRNAs (Fig. 2A). This allows separation of resources for expression of orthogonal modules from expression of native genes and/or other co-transfected non-orthogonal constructs. Such designs have been demonstrated not only to provide stable and reliable expression, but also to reduce the burden imposed on the native machinery, thus representing a win-win approach that can prove useful to reduce experimental variability. Examples in bacteria from Cameron et al.²³, Aleksashin et al.²⁴ and Carlson et al.²⁵ show how orthogonal protease-based or ribosome-based decoupling of circuit function can lead to controlled expression. Similarly, we speculate that other orthogonal components may prove beneficial, such as unnatural nucleotides that go on to generate polypeptides containing unnatural amino acids²⁶, thus relieving the endogenous pool of natural precursors from feeding into the synthetic system. Although not fully implemented in mammalian cells yet, future developments in this direction may allow us to re-gain confidence in normalisation controls for transfection variability by providing modules that are more robust to resource loading than standard transfection controls. It must be noted, however, that using such orthogonal platforms reduces some, but not all, resource coupling, and further development in this direction is needed.

Mathematical transformation of data

Mathematical transformation of the data can prove useful as well. For example, Rudge et al.²⁷ introduced a ratiometric promoter characteristic (RPC) that is largely insensitive to extrinsic variability (i.e. expression variability due to differences in mRNA degradation, translation, fluorescent protein maturation and growth). They transformed *E. coli* with a plasmid bearing two constitutively expressed fluorescent reporters, one under the control of a test promoter (with a total of six promoters tested), the other under the control of a reference promoter. The authors observed a linear relationship between fluorescence intensity and biomass (absorbance) during the exponential phase for both promoters, showing how the slope of this relationship describes an intrinsic promoter characteristic (α_p) that is largely insensitive to growth rate but still impacted by other factors (e.g. growth media) (Fig. 2B). Further, they demonstrated that normalising the α_p of the test promoter to the α_p of the reference promoter gave an RPC that was very robust to variations of extrinsic factors. This and similar mathematical approaches can be useful to reduce variance in measures of promoter activities due to extrinsic

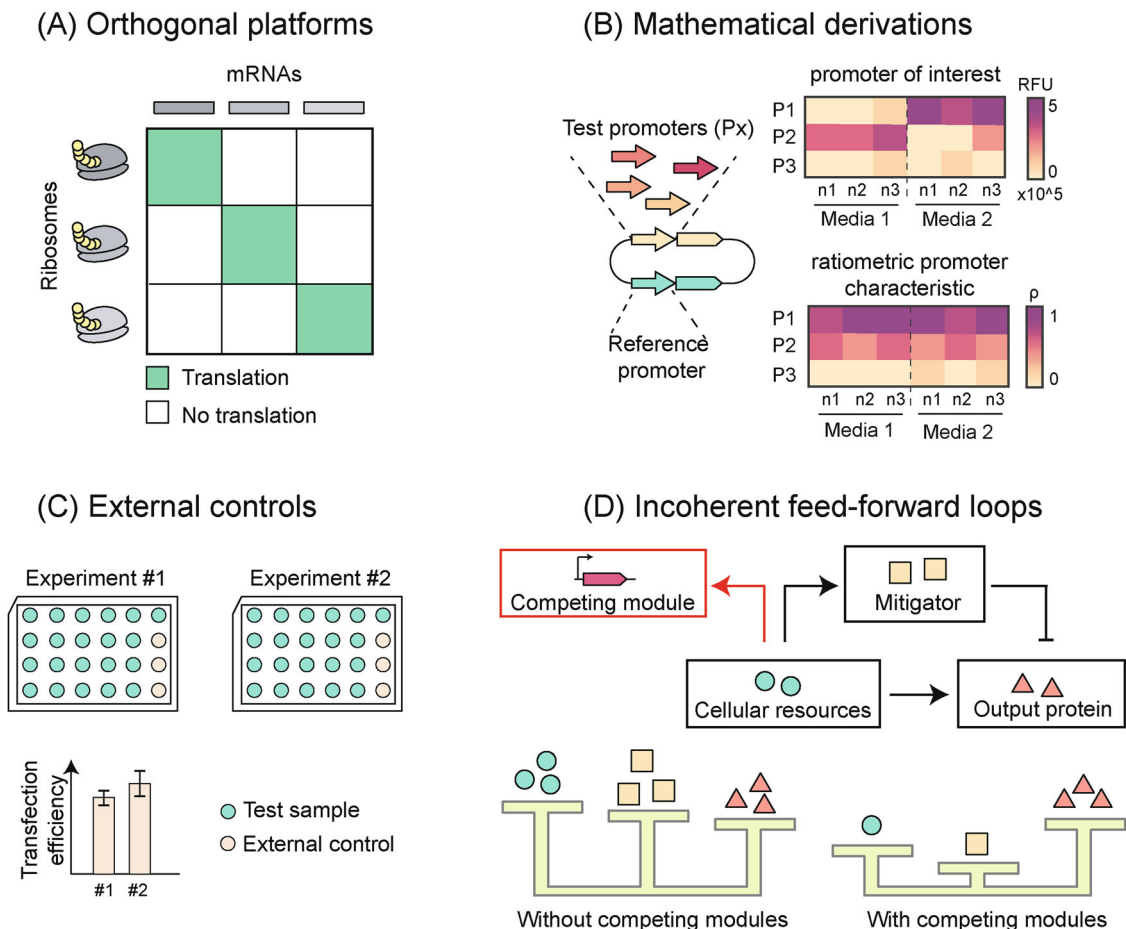


Fig. 2 Alternatives to traditional transfection controls. **A** Platforms functioning orthogonally to the cell machinery have been developed. These designs are uncoupled to the cell machinery with the double advantage of both stable expression and minimal burden imposed on the engineered cells (adapted from Carlson et al.²⁵). **B** Advanced mathematical derivations accounting for the global variation in part behaviour due to extrinsic factors allow precise definition of promoter characteristics (adapted from Rudge et al.²⁷). **C** Resource loading following co-transfection can be avoided by using external reporters to control for inter-experimental variability in transfection efficiency (adapted from Brown et al.³¹). **D** Incoherent feed-forward loops (iFFLs), where cellular resources act as the input regulating the expression of both a mitigator and a mitigator-repressed output gene, are an effective way to buffer the output gene expression to fluctuations in the pool of cellular resources (adapted from Jones et al.⁸ and Lillaci et al.⁶).

factors, thus representing a better alternative to the calculation of relative promoter units where promoter activity in different experimental conditions is normalised to a reference sample assessed in the same context^{28,29}. This can lead to deeper precision when annotating part behaviour and is applicable to co-expressed systems, thus greatly benefitting the standardisation agenda. However, the major caveat of this method is that it relies on normalising the reporter expression by the growth rate in the exponential phase, which is feasible for bacterial systems, but difficult to apply to situations such as transient transfection where plasmids number is diluted with each cell division³⁰. Advanced mathematical derivations better capturing variability in genetic part behaviour due to resource loading are awaited.

External controls and resource-decoupled modules

Approaches that eliminate the use of controls entirely can also be envisioned. For instance, the adoption of transfection controls could be replaced by the use of external reporters to keep track of transfection efficiency variability and confirm construct behaviour with no need for normalisation (Fig. 2C). This is supported by Brown et al. who, in a recent study on the design of synthetic promoters in CHO cells, mention their choice to omit transfection controls to avoid promoter interference, opting instead for the use of separate samples transfected with constitutively

expressed reporters with varying expression strength to assess the variability in transfection efficiency among experiments³¹. Although less elaborate, this approach is of immediate use, due to its simplicity. However, this approach relies on precise control over the transfection conditions to ensure that the experiments are not affected by variability in transfection efficiency as even small variations in cell state or plasmid DNA quality at the point of transfection can lead to large variations in the data.

Finally, the development of resource-decoupled modules to function as transfection controls can be envisioned. Previously, this strategy has been explored in microbial systems^{32–35}. In mammalian cells, a method recently proposed is the incoherent feedforward loop (iFFL) circuit topology, which represents a core motif in biology and has been often used in synthetic circuits for its adaptation properties^{6,8}. In iFFLs an input A regulates the expression of both an intermediate gene B (resource mitigator) and an output gene C, which is also negatively regulated by B (Fig. 2D). Jones et al. designed an endoribonuclease-based incoherent feed-forward loop (iFFL) where cellular resources act as the input A, regulating the Cas6 (B)-driven repression of the output gene C⁸. When cellular resource availability decreases (e.g., due the presence of competing modules), the expression of Cas6 diminishes, increasing the expression of the output gene. In this way, the decrease in the output expression due to resource

limitations is buffered by the partial release of Cas6-driven repression, and the expression of the transfection control can be maintained as constant. An alternative to the protein-based iFFL described by Jones et al., is the use of microRNA-based iFFLs as shown in Frei et al. Here, the use of endogenous or synthetic microRNAs as intermediate gene B in the iFFL designs was demonstrated to stabilize the expression of the gene of interest while perturbing the expression of a second gene, in either a two or three reporter system. Both systems show better adaptation to genetic load and DNA copy number variation as compared to the open loop version (no iFFL) of the same devices, mitigating the drawbacks of resource competition. Although conceptually more complex than standard transfection controls, iFFLs where a reporter gene acts as output C can guarantee a more stable expression under resource loading and can return higher-quality data when performing normalisation analyses.

Future directions

While the above-mentioned strategies are possible routes to undertake for mitigating resource competition, alternative and more optimal solutions to the resource-mediated coupling of expression systems in living cells need to be sought. To our knowledge, orthogonal modules have not been implemented in mammalian cells yet, and RPC mathematical derivations that are insensitive to resource loading are not available as of today. In contrast, iFFLs have been developed and characterised in mammalian cells and therefore may be the most immediate avenue to explore amongst those discussed above, even if they introduce complexity in the construct design. The exploitation of iFFLs also comes to a cost, namely the lower expression of the output gene C. While this should not be a problem in experiment where the downregulated gene is used as transfection control, this design may not be optimal in experimental settings seeking to maximize the iFFL-controlled output gene. Given that resource competition is just one of many mechanisms that cells use to couple and interlink their processes and pathways, a unique and easy solution to the problem may not be easily identified.

An overarching question is, how should the scientific community respond? We believe that a wider discussion within the molecular, cellular and synthetic biology communities is needed to evaluate the different approaches and propose possible alternatives. Further experimental and theoretical work aimed at identifying and quantifying the extent of interference between co-expressed systems and identifying novel strategies for resource-decoupled control of gene expression is also needed.

A possible direction for research might focus on the application of existing technologies to limit resource competition. For example, transcriptomic and proteomics approaches could be used to shed light on how resource allocation is regulated inside cells. This would not only provide a better understanding of the problem, but practically can also help us identify pathways that can be exploited for its mitigation. Similar approaches could be adopted for identification of more suitable transfection controls that overcome resource-coupled expression.

Another potential avenue may be to identify a set of genetic parts with lower footprint on the host resources. While *E. coli* construct design has benefitted from the development of automated tools taking into account the load imposed on the cellular machinery³⁶, such tools are still lacking for mammalian cells, and would hugely benefit the design of synthetic constructs with lower resource demand to be used for mammalian cell engineering or foundational experiments. While such tools are being developed and knowledge generated for the more complex mammalian systems, some may also opt for less time-consuming strategies to

mitigate resource competition. An example might be the use of fluorescent dyes to label intracellular plasmid DNA to characterise transfection efficiency and plasmid uptake and avoid the need for a reporter gene as the normalisation control^{37,38}. We believe there is no unique solution to the problem and a common effort, based on novel experimental and theoretical strategies as well as transparent data sharing and discussion, can pave the way for identification of possible solutions and mitigation approaches.

Received: 20 November 2020; Accepted: 29 March 2021;

Published online: 05 May 2021

References

1. Scott, M., Gunderson, C. W., Matesescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).
2. Gyorgy, A. et al. Isocost lines describe the cellular economy of genetic circuits. *Biophys. J.* **109**, 639–646 (2015).
3. Ceroni, F. et al. Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387 (2018).
4. Ceroni, F., Algar, R., Stan, G. B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415 (2015).
5. Huliak, I., Sike, A., Zencir, S. & Boros, I. M. The objectivity of reporters: interference between physically unlinked promoters affects reporter gene expression in transient transfection experiments. *DNA Cell Biol.* **31**, 1580–1584 (2012).
6. Lillacci, G., Benenson, Y. & Khammash, M. Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic Acids Res.* **46**, 9855–9863 (2018).
7. Frei, T. et al. Characterization and mitigation of gene expression burden in mammalian cells. *Nat. Commun.* **11**, 4641 (2020).
8. Jones, R. D. et al. An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nat. Commun.* **11**, 5690 (2020).
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. Tan, C., Marguet, P. & You, L. Emergent bistability by a growth-modulating positive feedback circuit. *Nat. Chem. Biol.* **5**, 842–848 (2009).
11. Klumpp, S., Zhang, Z. & Hwa, T. Growth rate-dependent global effects on gene expression in bacteria. *Cell* **139**, 1366–1375 (2009).
12. Liu, Q., Schumacher, J., Wan, X., Lou, C. & Wang, B. Orthogonality and burdens of heterologous AND gate gene circuits in *E. coli*. *ACS Synth. Biol.* **7**, 553–564 (2018).
13. Kim, T. K. & Eberwine, J. H. Mammalian cell transfection: the present and the future. *Anal. Bioanal. Chem.* **397**, 3173–3178 (2010).
14. Vitor, M. T., Sart, S., Barizien, A., Torre, L. G. & Baroud, C. N. Tracking the evolution of transiently transfected individual cells in a microfluidic platform. *Sci. Rep.* **8**, 1225 (2018).
15. Thirunavukkarasu, K., Miles, R. R., Halladay, D. L. & Onyia, J. E. Cryptic enhancer elements in luciferase reporter vectors respond to the osteoblast-specific transcription factor *Osf2/Cbfa1*. *Biotechniques* **28**, 506–510 (2000).
16. Nejepinska, J., Malik, R., Moravec, M. & Svoboda, P. Deep sequencing reveals complex spurious transcription from transiently transfected plasmids. *PLoS ONE* **7**, e43283 (2012).
17. Hu, Q., Suzuki, K., Hirschler-Laszkiwicz, I. & Rothblum, L. I. Paradoxical effect of eukaryotic expression vectors on reporters. *Biotechniques* **33**, 74, 76, 78 passim (2002).
18. Farr, A. & Roman, A. A pitfall of using a second plasmid to determine transfection efficiency. *Nucleic Acids Res.* **20**, 920 (1992).
19. Bergeron, D., Barbeau, B., Leger, C. & Rassart, E. Experimental bias in the evaluation of the cellular transient expression in DNA co-transfection experiments. *Cell Mol. Biol. Res.* **41**, 155–159 (1995).
20. Hofman, K., Swinnen, J. V., Claessens, F., Verhoeven, G. & Heyns, W. Apparent coactivation due to interference of expression constructs with nuclear receptor expression. *Mol. Cell Endocrinol.* **168**, 21–29 (2000).
21. Gardner, T. S. & Hawkins, K. Synthetic biology: evolution or revolution? A co-founder's perspective. *Curr. Opin. Chem. Biol.* **17**, 871–877 (2013).
22. Decoene, T. et al. Standardization in synthetic biology: an engineering discipline coming of age. *Crit. Rev. Biotechnol.* **38**, 647–656 (2018).

23. Cameron, D. E. & Collins, J. J. Tunable protein degradation in bacteria. *Nat. Biotechnol.* **32**, 1276–1281 (2014).
24. Aleksashin, N. A. et al. A fully orthogonal system for protein synthesis in bacterial cells. *Nat. Commun.* **11**, 1858 (2020).
25. Carlson, E. D. et al. Engineered ribosomes with tethered subunits for expanding biological function. *Nat. Commun.* **10**, 3920 (2019).
26. Fischer, E. C. et al. New codons for efficient production of unnatural proteins in a semisynthetic organism. *Nat. Chem. Biol.* **16**, 570–576 (2020).
27. Rudge, T. J. et al. Characterization of intrinsic properties of promoters. *ACS Synth. Biol.* **5**, 89–98 (2016).
28. Kelly, J. R. et al. Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* **3**, 4 (2009).
29. Keren, L. et al. Promoters maintain their relative activity levels under different growth conditions. *Mol. Syst. Biol.* **9**, 701 (2013).
30. Schaarschmidt, D., Baltin, J., Stehle, I. M., Lipps, H. J. & Knippers, R. An episomal mammalian replicon: sequence-independent binding of the origin recognition complex. *EMBO J.* **23**, 191–201 (2004).
31. Brown, A. J., Sweeney, B., Mainwaring, D. O. & James, D. C. Synthetic promoters for CHO cell engineering. *Biotechnol. Bioeng.* **111**, 1638–1647 (2014).
32. Mishra, D., Rivera, P. M., Lin, A., Del Vecchio, D. & Weiss, R. A load driver device for engineering modularity in biological networks. *Nat. Biotechnol.* **32**, 1268–1275 (2014).
33. 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).
34. Aoki, S. K. et al. A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature* **570**, 533–537 (2019).
35. Shopera, T., He, L., Oyetunde, T., Tang, Y. J. & Moon, T. S. Decoupling resource-coupled gene expression in living cells. *ACS Synth. Biol.* **6**, 1596–1604 (2017).
36. Nielsen, A. A. et al. Genetic circuit design automation. *Science* **352**, aac7341 (2016).
37. Srinivasan, C., Siddiqui, S., Silbart, L. K., Papadimitrakopoulos, F. & Burgess, D. J. Dual fluorescent labeling method to visualize plasmid DNA degradation. *Bioconjug Chem.* **20**, 163–169 (2009).
38. Rombouts, K., Braeckmans, K. & Remaut, K. Fluorescent labeling of plasmid DNA and mRNA: gains and losses of current labeling strategies. *Bioconjug Chem.* **27**, 280–297 (2016).

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council (grant BB/S006206/1 to K.P., M.M.M.) and the ERC starting grants (grant Synthetic T-rEX 852012 to V.S.).

Author contributions

R.D.B., M.M.M., K.P., and F.C. conceived the paper. R.D.B., M.M.M., V.S., K.P., and F.C. wrote and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to F.C.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021