Engineering of Metabolic Pathways in Mammalian Cells for the Biosynthesis and Utilization of Exogenous Small Molecules

Pascal Stücheli

DISS. ETH NO. 26721

DISS. ETH NO. 26721

ENGINEERING OF METABOLIC PATHWAYS IN MAMMALIAN CELLS FOR THE BIOSYNTHESIS AND UTILIZATION OF EXOGENOUS SMALL MOLECULES

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

Pascal STÜCHELI

MSc ETH in Biotechnology, ETH Zurich

born on 19.06.1988

citizen from

Lommis TG

accepted on the recommendation of

Prof. Dr. Martin Fussenegger

Prof. Dr. Sven Panke

Prof. Dr. Yaakov Benenson

2020

TABLE OF CONTENTS

SUMMARY

Synthetic biology has shown highly promising results in tackling real-world problems in medicine and bioprocessing. The ability to use biological tools from many different sources enabled engineers to develop intricate cellular systems. First smart cell-based and gene therapies are currently being applied in the clinic and the production of therapeutic antibodies using metabolically engineered cells is well established. To advance the development of engineered medical and biopharmaceutical systems, it is an essential task to establish the heterologous gene expression in mammalian cells and to develop conceptually novel tools for bioengineering. In this thesis we present the design of two new tools based on the metabolic engineering of a heterologous small molecule biosynthesis- and a utilization-pathway.

In chapter I we introduce a reporter system for mammalian cells, based on the expression of a combination of two human genes together with one gene from *Amanita muscaria*. Expressing all three genes resulted in the formation of a yellow, fluorescent dye derived from the metabolic amino acid L-tyrosine. The dye belongs to the betaxanthin class, it can be detected extra- and intracellularly and its production does not seem to impact cell viability. This offers multiple approaches for quantification, in particular plate readers, flow cytometry and microscopy. Additionally, we show how the system can be used as a reporter for gene expression in mammalian cells with comparable performance as standardized reporter systems.

In chapter II we show the approach to endow CHO-K1 cells with the ability to metabolize the disaccharide cellobiose using a cellobiose specific transporter and a β-glucosidase. This metabolic engineering approach enables cells to survive in cellobiose medium without glucose, whereas untransfected cells quickly died. This cellobiose utilization pathway was used to create the additive stable cell line selection method called CelloSelect. A detailed selection protocol was established to create long-term stable cells in a matter of days. We show that these stable cells can be grown in cellobiose- or glucose-culture medium and that the cells continue to express the cargo protein at the same expression level for the tested period of 30 days independent of medium composition. In a proof-of-concept bioprocessing experiment we conclude that the system can be used to produce a therapeutic protein.

ZUSAMMENFASSUNG

Synthetische Biologie hat schon zu vielversprechenden Resultaten in der Medizin und Bioproduktion geführt. Die Fähigkeit biologische Werkzeuge von verschiedensten Quellen zu verwenden, hat es Ingenieuren ermöglicht komplexe zelluläre Systeme zu entwickeln. Die ersten intelligenten Zell- und Gen-therapeutischen Medikamente werden in der Klinik verwendet und die Produktion von therapeutischen Antikörpern in metabolisch veränderten Zellen ist weit verbreitet. Es ist essentielle die heterologe Expression von Genen in Säugetierzellen weiter zu etablieren und konzeptionell neuartige Werkzeuge für das Bioengineering zu entwickeln. In dieser Doktorarbeit präsentieren wir zwei neuartige Werkzeuge, basierend auf der metabolischen Veränderung von Säugetierzellen für die heterologe Produktion und Verwendung von niedermolekularen Verbindungen.

In Kapitel I stellen wir ein Reportersystem für Säugetierzellen vor, das auf der Expression von zwei menschlichen Genen und einem Gen von *Amanita muscaria* basiert. Die Expression aller drei Gene resultiert in der Produktion eines gelben, fluoreszierenden Farbstoffes, der ausgehend von der metabolischen Aminosäure L-Tyrosin gebildet wird. Die Resultate zeigen, dass der Farbstoff zu der Klasse der Betaxanthine gehört, er intra- und extrazellulär nachgewiesen werden kann und die Produktion die Zellviabilität nicht beeinflusst. Der Farbstoff kann mittels mehrerer Methoden nachgewiesen werden, unter anderem mit Plattenlesern, Mikroskopie oder Durchflusszytometrie. Ausserdem zeigen wir, wie die Farbstoffproduktion als Reportersystem für die Genexpression verwendet werden mit vergleichbarer Performance wie bekannte Reportersysteme.

In Kapitel II zeigen wir eine Möglichkeit um CHO-K1 Zellen mit der Fähigkeit auszustatten das Disaccharid Cellobiose mittels eines spezifischen Transporters und einer β-Glucosidase zu metabolisieren. Diese metabolische Veränderung führt dazu, dass die Zellen in Kulturmedium überleben, welches Cellobiose anstatt Glucose enthält, während nicht-transfizierte Zellen sterben. Dieser Cellobiose Stoffwechsel wurde dann verwendet um ein CelloSelect genanntes Selektionssystem zu entwickeln um stabil-transfizierte Zelllinien zu selektionieren. Ein detailliertes Selektionsprotokoll wurde entwickelt, das verwendet werden kann um in wenigen Tagen langzeit-stabile Zelllinien zu generieren. Wir zeigen, dass die stabilen Zellen in Cellobiose- oder Glucose-Kulturmedium kultiviert werden können und dass sich die

Produktion des Cargo-Proteins während eines 30 tägigen Tests nicht verändert, unabhängig von der Zusammensetzung des Mediums. In einem konzeptionellen bioprozesstechnischen Experiment haben wir gezeigt, dass das System dazu verwendet werden kann ein therapeutisches Protein zu produzieren.

INTRODUCTION

Synthetic Biology

Synthetic biology is the application-driven modular engineering approach in biology, that functions under the elegant premise that cellular systems can be viewed as comparably complex machines. In mechanical or electrical engineering, components created by engineers are highly modular and can be adjusted quickly to a new design. Similarly, synthetic biologists are developing modules than can be used in a plug-and-play like manner to create novel functions in cells. Of course, in biological systems the separation of components, compartmentalization and productive interactions are not as trivial and well understood as in mechanical and electrical engineering. Where in mechanical systems the components are set in a fixed location and thus allows the interaction partners to be freely chosen, biological systems work in a chaotic diffusion driven environment where many things are constantly interacting. An additional difficulty arises when enzymes or whole enzyme cascades are heterologously used in different organisms. Enzymes are sensitive to the biochemical environment such as pH, temperature, salt concentration and required cofactors(1). One malfunctioning enzyme can lead to the collapse of a whole cascade. In many cases bioengineers have to modify enzymes that are to be used in a different organism, or screen different enzymes that could catalyze the same reaction.

The field of synthetic biology is a relatively new one that first appeared with the discovery of restriction enzymes in 1969(2) and gained some traction around the new millennium with the engineering of first synthetic gene networks(3,4). Since then have synthetic biologists designed and created an ever-growing number of building blocks that are akin to the basic building blocks of electrotechnology, such as molecular switches(5,6), pulse generators(7), time-delay circuits(8,9), logic gates(10,11), and oscillators(12,13). Using these modules, and in some cases combining them to more complex gene circuits or networks(14,15), has huge potential, primarily in medicine. Smart therapeutics in patients to precisely receive molecular information about a disease state, decide autonomously about actions to take, and then locally taking action by producing a therapeutic agent is an exciting development. Since the development of new methods, such as the Generalized Extracellular Molecule Sensor Platform(16) (GEMS) or especially the dawn of precise genome manipulation using the

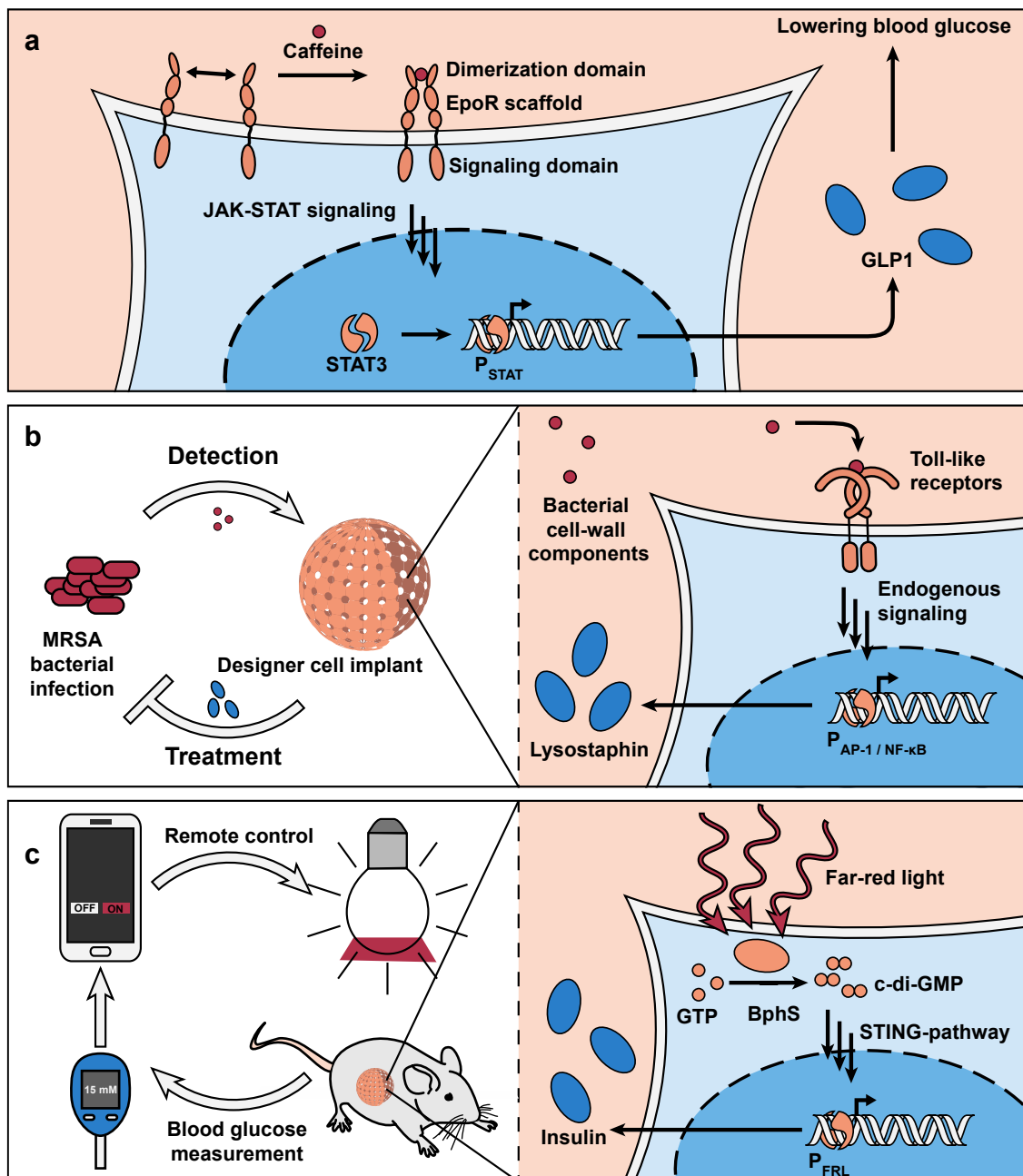CRISPR/Cas9 system(17), the development and use of smart designer drugs seems to be around the corner.

Figure 1: Examples of smart designer cells. (a) Cells were engineered to detect caffeine(18). This results in GLP1 production and secretion, which leads to decreased blood glucose levels. (b) Engineered cells in a cell implant capable of fighting multi drug resistant Staphylococcus aureus (MRSA) infections(19). Human Toll-like receptors detect the bacterial infection and react by producing the bacteriolytic enzyme lysostaphin. (c) Engineered cells in a cell implant capable of red light inducible insulin production(20). In the study the light was activated using a smartphone, thus enabling smartphone-controlled glucose levels. Figure based on the authors version of a manuscript that was published(21).

Although the use of viruses for the direct integration of a gene circuit into a patient's cells through gene therapy has been discussed(22), different limitations exist to date. On the other hand, the approach to engineer cells in a laboratory and subsequently injecting them into a patient is advancing rapidly with tests being done in animal models(23-25) (Fig. 1). The most famous system is likely the Chimeric Antigen Receptor T-Cell (CAR T) therapy developed and sold by Novartis. In this approach, patient-derived T-cells are engineered in the lab to recognize specific tumor antigens and reinjected, with highly promising results(26).
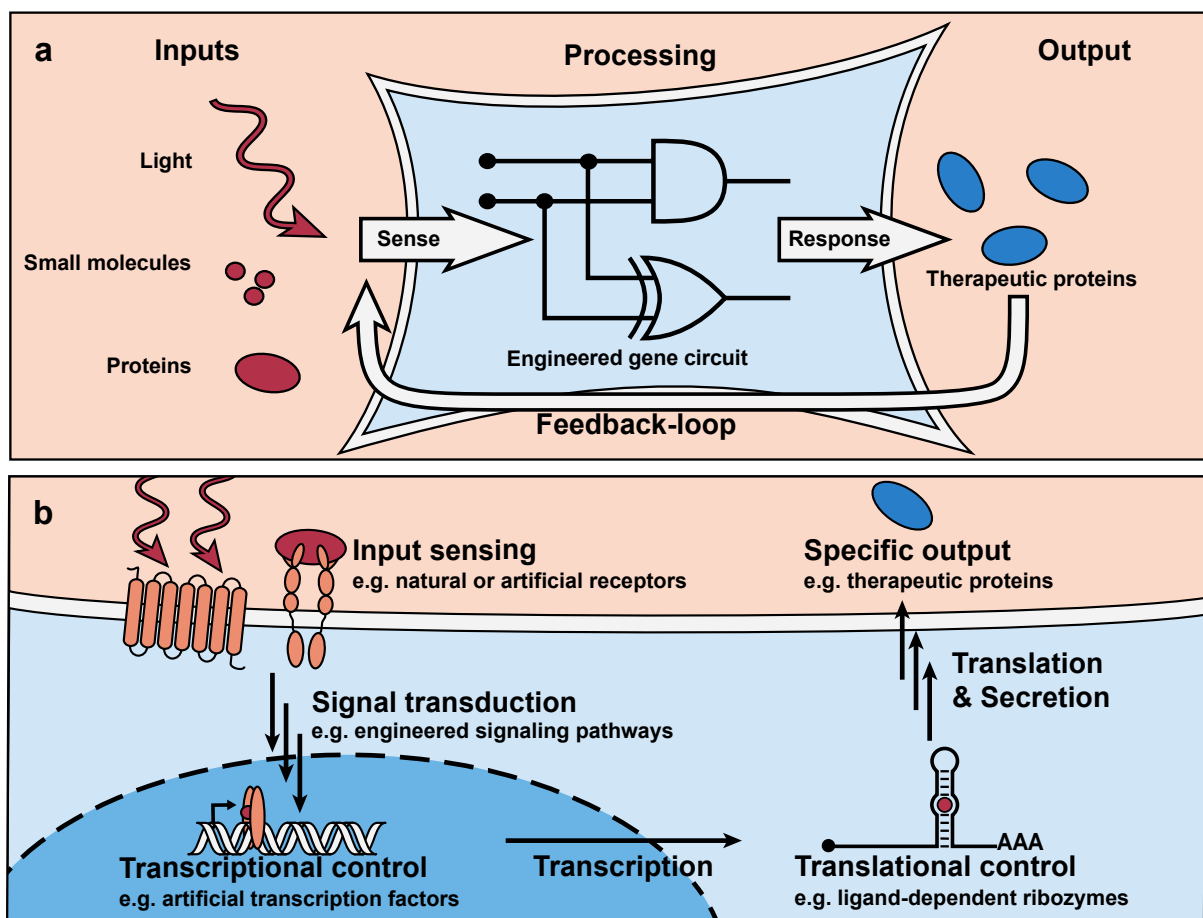
Figure 2: Schematic design of an engineered designer cell. (a) Generalized overview of the different abstract layers smart cells are built on. (b) Close-up of different simple levels on which the signal processing can appear in the cells. Figure based on the authors version of a manuscript that was published(21).

Therapeutic designer cells can be designed by combining an input detection module, processing module and effector output (Fig. 2). On a practical level, engineering such systems

generally requires three steps: 1. Creating genetic constructs that carry the desired genes, 2. Testing these modules individually and later combined for desired functionality, 3. Stably integrating the genetic modules into cells to create long-lasting cells with the desired functions.

Reporter Systems

Testing the genetic components of simple or complex synthetic systems for proper function generally requires a detectable output signal. Especially in complex interaction networks all systems have to be tested individually before being tested as a whole. Different conceptually diverse approaches exist to observe changes of gene expression inside cells, through which an engineer can assay the functionality of either an input signal or a signal processing part inside the cell. Depending on the molecular mechanism of the tested function, changes can be analyzed either on the mRNA level or on the protein level. mRNA quantification is usually done using reverse transcription quantitative PCR (RT-qPCR)(27) or more recently RNA sequencing(28) which yields information about a large number of gene expression levels. Highly powerful tools in basic research and synthetic biology alike, both methods are rather time consuming and expensive and therefore rarely used for routine tests if avoidable. Additionally, the mRNA levels do not necessarily correlate to the eventual protein production rate. As such, quantifying gene expression on the protein production level requires the production of a quantifiable protein. This is achieved either by using a protein that itself has optical properties, such as fluorescent proteins, or by using a protein that can be quantified through an external assay. The discovery of fluorescent proteins was seen as a huge milestone(29), especially in basic research where it enabled researchers to study sub-cellular interactions or the localization of other proteins(30). Is it also used in engineering for single-cell quantification with a high spatial resolution, using a microscope, or using flow cytometry to yield population distribution data.

Quantifying gene expression by producing a quantifiable protein is currently the go-to method in bioengineering. Although the palette of quantifiable proteins is vast due to universal technologies, such gel electrophoresis(31) or Enzyme-linked Immunosorbent Assays (ELISA)(32), bioengineers generally utilize enzymatic assays(33) due to their ease of use and them not being very time-consuming. Two well-known members of this assay class are the

secreted alkaline phosphatase(34) and luciferases(35). In both assays the analyzed genetic system produces the reporter enzyme that is harvested in the culture supernatant and subsequently quantified upon the addition of a substrate. In the case of the phosphatases a phosphate-containing dye is added that upon phosphate removal shows different spectral properties, in the case of the luciferases the catalyzed reaction leads to light being emitted. Using a laboratory plate reader, assays can be run on many samples in parallel and yield quantitative information about the expression of the reporter gene.

Industrial Biotechnology

Where currently the therapeutic use of synthetic biology by creating cell-based therapeutics or gene therapy is still in its infancy with only few clinically licensed examples(36) (e.g. Zolgensma gene therapy for the treatment of spinal muscular atrophy(37), or Zalmoxis for the restoration of the patient's immune system after partially incompatible hematopoietic stem cell transplantation(38)), its use in high throughput screening and biotechnological production is well established(39).

One of the most straight-forward approaches to use synthetic biology in industry is the bioproduction of drugs, or other valuable products in engineered microorganisms(40). Implementing a production cassette in bacteria or yeast is an efficient way to produce a product in a highly standardized large-scale reactor, that normally would be produced by in an organism that is not easily scalable or that produces it in a low concentration. The key advantage of producing small molecule drugs, or intermediates, in a bioreactor over chemical synthesis is likely the better control over chirality of the product. Examples of pharmaceuticals at the least partially produced by engineered microorganisms include the anti-malaria drug artemisinin produced by Sanofi(41) or the anti-cancer chemotherapeutic Taxol produced by BSM(42).

Another use for synthetic biology in an industrial setting can be found in drug-screenings in mammalian cell-based platforms. Pharmaceutical companies often have libraries of small molecules from many different sources (synthetic compounds, natural products or modified derivatives). The putative drugs in these libraries are each added individually to a cell population and different cellular changes observed in a format of up to 3456 tested cultures per culture plate. The most straight-forward cellular change is population growth and

viability. Assessing the effect of the tested substances on cell viability using a viability assay was used for example in the discovery of new anti-cancer drugs(43). When a specific cellular target is identified that leads to a certain disease type, specific model-cell lines can be used (wild-type or genetically engineered cells) to find compounds that interact in the desired fashion with the target protein. These cellular changes can be detected via different external methods (receptor activation-mediated changes in cAMP levels(44) or changes in calcium levels(45)) or through internal changes in gene expression. Synthetic biology comes into play to make the cell lines themselves report the changes made by the putative drug. Many cellular pathways lead to a direct or indirect change in gene expression which can be detected using an engineered cell line expressing a reporter gene that is expressed if the desired pathway is activated(46-48). An intriguing side-effect of this approach is the fact that, if a promoter can be used that is normally not active in the cell, reporter activation simultaneously indicates metabolic integrity of the tested cell. A toxic compound that damages the cell is likely to result in gene expression to decrease and thus lower levels of reporter proteins. In fact, reporter genes can be used to assess cell viability in general(49). More complex gene networks were developed to report information about specific and unspecific intracellular changes induced by the tested compounds(50).

Other opportunities for synthetic biology in industry can be found in mammalian bioproduction. Evermore of the current blockbuster drugs are proteins produced in a bioreactor instead of a chemical reactor(51). Therapeutic antibodies are quickly becoming the gold standard for many different diseases, such as cancer, autoimmune diseases, or even migraines(52). The general work-flow for the large-scale antibody production is as follows(53): a genetic construct is created containing the DNA sequence encoding the desired antibody. Suspension Chinese hamster ovary cells (CHO) are transfected and the genetic constructs are integrated into the genome by means of random integration. A polyclonal cell line is selected by having a gene encoding a selection marker on the genetic construct and using a selection agent to kill of all cells not expressing these markers. Often, cells from the polyclonal cell lines are then screened individually for highest production titer, resulting in a monoclonal cell line.

Metabolic Engineering

Increasing the production capacity of cell lines falls in the realm of metabolic engineering, a discipline closely related synthetic biology(54). Metabolic engineering aims primarily at three broad applications in bioprocessing: increasing the yield, reducing byproducts or extending the substrate range. In contrast to synthetic biology it uses rather simple approaches to influence the complex endogenous metabolic network to change fluxes of metabolites. It is frequently applied to microorganisms to create production cell lines for a large variety of natural products (such as amino acids, antibiotics, vitamins or biofuels(55)). Of course, there are many different approaches for each individual product and problem at hand but they tend to be based on similar methods. Creating knockout cell lines can remove a metabolic flux from a side branch which might then go into the desired production branch (e.g. (56)). Overexpressing key enzymes, as well as enzyme engineering, could remove bottlenecks and increase the flux to a desired product (e.g. (57)). And finally, heterologous gene expression offers the possibility to create completely new products in cells(58) or use different sources of nutrients(59). Today, many of these approaches are guided using mathematical models to improve the successful metabolic engineering(60).

In mammalian cell-based bioprocessing, metabolic engineering is used for multiple purposes. The main reason why mammalian cell lines are used as bioproduction system is the glycosylation of proteins, which other production systems based on yeast or bacteria would perform very differently(61). Antibody glycosylation is crucial regarding its binding affinity, antigenicity or half-life(62). As such, an effort is currently made to improve the glycosylation capabilities of CHO cells(63). Metabolic engineering helped to dramatically increase the protein production capacity in CHO-based systems from a few tenths of g/L to average large-scale production of more than 3 g/L (64). Important engineering approaches were developed on decreasing the glucose consumption of cells, decreasing the lactate accumulation, decreasing the growth rate and improving the protein secretion(65,66). An interesting discovery, in the context of this thesis, was that CHO cells are capable of utilizing maltose. Using this nutrient, instead of glucose, resulted in a remarkable decrease in lactate accumulation, one of the key waste products in mammalian cell culture.

Important work on metabolic engineering was done on knockout cell lines for stable integration selection systems. The two gold-standards for industrial cell line selection systems

are the (glutamine synthetase) GS(67) and the dihydrofolate reductase (DHFR)(68) systems. Using a chemical knockdown approach or a genetic knockout approach created cell lines that are dependent on glutamine or hypoxanthine and thymidine in the growth medium. As a selection marker the two genes are reintegrated into the genome allowing the cells to survive in regular growth medium and thus enable survival of only the desired cells.

REFERENCES

1.   Robinson, P.K. (2015) Enzymes: principles and biotechnological applications (vol 59, pg 1, 2015). *Understanding Biochemistry: Enzymes and Membranes*, **59**, 75-75.
2.   Arber, W. and Linn, S. (1969) DNA Modification and Restriction. *Annual Review of Biochemistry*, **38**, 467-+.
3.   Elowitz, M.B. and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**, 335-338.
4.   Gardner, T.S., Cantor, C.R. and Collins, J.J. (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature*, **403**, 339-342.
5.   Kramer, B.P., Viretta, A.U., Daoud-El-Baba, M., Aubel, D., Weber, W. and Fussenegger, M. (2004) An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.*, **22**, 867-870.
6.   Atkinson, M.R., Savageau, M.A., Myers, J.T. and Ninfa, A.J. (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. *Cell*, **113**, 597-607.
7.   Basu, S., Mehreja, R., Thiberge, S., Chen, M.T. and Weiss, R. (2004) Spatiotemporal control of gene expression with pulse-generating networks. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 6355-6360.
8.   Weber, W., Stelling, J., Rimann, M., Keller, B., Daoud-El Baba, M., Weber, C.C., Aubel, D. and Fussenegger, M. (2007) A synthetic time-delay circuit in mammalian cells and mice. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 2643-2648.
9.   Weber, W., Kramer, B.P. and Fussenegger, M. (2007) A genetic time-delay circuitry in mammalian cells. *Biotechnol Bioeng*, **98**, 894-902.
10.   Kramer, B.P., Fischer, C. and Fussenegger, M. (2004) BioLogic gates enable logical transcription control in mammalian cells. *Biotechnol Bioeng*, **87**, 478-484.
11.   Rinaudo, K., Bleris, L., Maddamsetti, R., Subramanian, S., Weiss, R. and Benenson, Y. (2007) A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat. Biotechnol.*, **25**, 795-801.
12.   Goh, K.I., Kahng, B. and Cho, K.H. (2008) Sustained oscillations in extended genetic oscillatory systems. *Biophys. J.*, **94**, 4270-4276.
13.   Stricker, J., Cookson, S., Bennett, M.R., Mather, W.H., Tsimring, L.S. and Hasty, J. (2008) A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**, 516-519.
14.   Auslander, D., Auslander, S., Pierrat, X., Hellmann, L., Rachid, L. and Fussenegger, M. (2018) Programmable full-adder computations in communicating three-dimensional cell cultures. *Nat. Methods*, **15**, 57-60.
15.   Auslander, S., Auslander, D., Muller, M., Wieland, M. and Fussenegger, M. (2012) Programmable single-cell mammalian biocomputers. *Nature*, **487**, 123-127.
16.   Scheller, L., Strittmatter, T., Fuchs, D., Bojar, D. and Fussenegger, M. (2018) Generalized extracellular molecule sensor platform for programming cellular behavior. *Nat. Chem. Biol.*, **14**, 723-729.
17.   Pickar-Oliver, A. and Gersbach, C.A. (2019) The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.*, **20**, 490-507.

18.   Bojar, D., Scheller, L., Hamri, G.C., Xie, M. and Fussenegger, M. (2018) Caffeine-inducible gene switches controlling experimental diabetes. *Nat. Commun.*, **9**, 2318.

19.   Liu, Y., Bai, P., Woischnig, A.K., Charpin-El Hamri, G., Ye, H., Folcher, M., Xie, M., Khanna, N. and Fussenegger, M. (2018) Immunomimetic Designer Cells Protect Mice from MRSA Infection. *Cell*, **174**, 259-270 e211.

20.   Shao, J., Xue, S., Yu, G., Yu, Y., Yang, X., Bai, Y., Zhu, S., Yang, L., Yin, J., Wang, Y. *et al.* (2017) Smartphone-controlled optogenetically engineered cells enable semiautomatic glucose homeostasis in diabetic mice. *Sci. Transl. Med.*, **9**.

21.   Tolle, F., Stucheli, P. and Fussenegger, M. (2019) Genetic circuitry for personalized human cell therapy. *Curr. Opin. Biotechnol.*, **59**, 31-38.

22.   Kitada, T., DiAndreth, B., Teague, B. and Weiss, R. (2018) Programming gene and engineered-cell therapies with synthetic biology. *Science*, **359**.

23.   Saxena, P., Charpin-El Hamri, G., Folcher, M., Zulewski, H. and Fussenegger, M. (2016) Synthetic gene network restoring endogenous pituitary-thyroid feedback control in experimental Graves' disease. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, 1244-1249.

24.   Kojima, R., Scheller, L. and Fussenegger, M. (2018) Nonimmune cells equipped with T-cell-receptor-like signaling for cancer cell ablation. *Nat. Chem. Biol.*, **14**, 42-49.

25.   Schukur, L., Geering, B., Charpin-El Hamri, G. and Fussenegger, M. (2015) Implantable synthetic cytokine converter cells with AND-gate logic treat experimental psoriasis. *Sci. Transl. Med.*, **7**, 318ra201.

26.   Miliotou, A.N. and Papadopoulou, L.C. (2018) CAR T-cell Therapy: A New Era in Cancer Immunotherapy. *Curr. Pharm. Biotechnol.*, **19**, 5-18.

27.   Taylor, S.C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M. and Fenrich, J. (2019) The Ultimate qPCR Experiment: Producing Publication Quality, Reproducible Data the First Time. *Trends Biotechnol.*, **37**, 761-774.

28.   Stark, R., Grzelak, M. and Hadfield, J. (2019) RNA sequencing: the teenage years. *Nat. Rev. Genet.*, **20**, 631-656.

29.   Tsien, R.Y. (1998) The green fluorescent protein. *Annual Review of Biochemistry*, **67**, 509-544.

30.   Chudakov, D.M., Matz, M.V., Lukyanov, S. and Lukyanov, K.A. (2010) Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiol. Rev.*, **90**, 1103-1163.

31.   Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic Transfer of Proteins from Polyacrylamide Gels to Nitrocellulose Sheets - Procedure and Some Applications. *Proc. Natl. Acad. Sci. U. S. A.*, **76**, 4350-4354.

32.   Engvall, E. and Perlmann, P. (1972) Enzyme-Linked Immunosorbent Assay, Elisa .3. Quantitation of Specific Antibodies by Enzyme-Labeled Anti-Immunoglobulin in Antigen-Coated Tubes. *J. Immunol.*, **109**, 129-+.

33.   Jiang, T., Xing, B. and Rao, J. (2008) Recent developments of biological reporter technology for detecting gene expression. *Biotechnol. Genet. Eng. Rev.*, **25**, 41-75.

34.   Cullen, B.R. and Malim, M.H. (1992) Secreted Placental Alkaline-Phosphatase as a Eukaryotic Reporter Gene. *Methods Enzymol.*, **216**, 362-368.

35.   Smale, S.T. (2010) Luciferase assay. *Cold Spring Harb Protoc*, **2010**, pdb prot5421.

36.   Shahryari, A., Saghaeian Jazi, M., Mohammadi, S., Razavi Nikoo, H., Nazari, Z., Hosseini, E.S., Burtscher, I., Mowla, S.J. and Lickert, H. (2019) Development and Clinical Translation of Approved Gene Therapy Products for Genetic Disorders. *Front Genet*, **10**, 868.

37.   Dangouloff, T. and Servais, L. (2019) Clinical Evidence Supporting Early Treatment Of Patients With Spinal Muscular Atrophy: Current Perspectives. *Ther. Clin. Risk Manag.*, **15**, 1153-1161.

38.   Vago, L., Oliveira, G., Bondanza, A., Noviello, M., Soldati, C., Ghio, D., Brigida, I., Greco, R., Lupo Stanghellini, M.T., Peccatori, J. *et al.* (2012) T-cell suicide gene therapy prompts thymic renewal in adults after hematopoietic stem cell transplantation. *Blood*, **120**, 1820-1830.

39.   Trosset, J.Y. and Carbonell, P. (2015) Synthetic biology for pharmaceutical drug discovery. *Drug Des. Devel. Ther.*, **9**, 6285-6302.

40.   Moses, T., Mehrshahi, P., Smith, A.G. and Goossens, A. (2017) Synthetic biology approaches for the production of plant metabolites in unicellular organisms. *J. Exp. Bot.*, **68**, 4057-4074.

41. Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J. *et al.* (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, **440**, 940-943.

42. Kundu, S., Jha, S. and Ghosh, B. (2017) Metabolic Engineering for Improving Production of Taxol. *Ref Ser Phytochem*, 463-484.

43. Lage, O.M., Ramos, M.C., Calisto, R., Almeida, E., Vasconcelos, V. and Vicente, F. (2018) Current Screening Methodologies in Drug Discovery for Selected Human Diseases. *Mar. Drugs*, **16**.

44. Fang, Y., Frutos, A.G. and Verkleeren, R. (2008) Label-free cell-based assays for GPCR screening. *Comb. Chem. High Throughput Screen.*, **11**, 357-369.

45. Arkin, M.R., Connor, P.R., Emkey, R., Garbison, K.E., Heinz, B.A., Wiernicki, T.R., Johnston, P.A., Kandasamy, R.A., Rankl, N.B. and Sittampalam, S. (2004) In Sittampalam, G. S., Grossman, A., Brimacombe, K., Arkin, M., Auld, D., Austin, C. P., Baell, J., Bejcek, B., Caaveiro, J. M. M., Chung, T. D. Y. *et al.* (eds.), *Assay Guidance Manual*, Bethesda (MD).

46. Xie, W., Silvers, R., Ouellette, M., Wu, Z., Lu, Q., Li, H., Gallagher, K., Johnson, K. and Montoute, M. (2016) A Luciferase Reporter Gene System for High-Throughput Screening of gamma-Globin Gene Activators. *Methods Mol. Biol.*, **1439**, 207-226.

47. Sedlmayer, F., Hell, D., Muller, M., Auslander, D. and Fussenegger, M. (2018) Designer cells programming quorum-sensing interference with microbes. *Nat. Commun.*, **9**.

48. Weber, W., Schoenmakers, R., Keller, B., Gitzinger, M., Grau, T., Baba, M.D.E., Sander, P. and Fussenegger, M. (2008) A synthetic mammalian gene circuit reveals antituberculosis compounds. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 9994-9998.

49. Lupold, S.E., Johnson, T., Chowdhury, W.H. and Rodriguez, R. (2012) A real time Metridia luciferase based non-invasive reporter assay of mammalian cell viability and cytotoxicity via the beta-actin promoter and enhancer. *PLoS One*, **7**, e36535.

50. Haefliger, B., Prochazka, L., Angelici, B. and Benenson, Y. (2016) Precision multidimensional assay for high-throughput microRNA drug discovery. *Nat. Commun.*, **7**.

51. Waegeman, H. and Soetaert, W. (2011) Increasing recombinant protein production in Escherichia coli through metabolic and genetic engineering. *J. Ind. Microbiol. Biotechnol.*, **38**, 1891-1910.

52. Kaplon, H., Muralidharan, M., Schneider, Z. and Reichert, J.M. (2020) Antibodies to watch in 2020. *MAbs*, **12**.

53. Shukla, A.A., Wolfe, L.S., Mostafa, S.S. and Norman, C. (2017) Evolving trends in mAb production processes. *Bioeng Transl Med*, **2**, 58-69.

54. Stephanopoulos, G. (2012) Synthetic Biology and Metabolic Engineering. *Acs Synth Biol*, **1**, 514-525.

55. Yang, S.T., Liu, X.G. and Zhang, Y.L. (2007) Metabolic Engineering - Applications, Methods, and Challenges. *Bioprocessing for Value-Added Products from Renewable Resources: New Technologies and Applications*, 73-118.

56. Kyla-Nikkila, K., Hujanen, M., Leisola, M. and Palva, A. (2000) Metabolic engineering of Lactobacillus helveticus CNRZ32 for production of pure L-(+)-lactic acid. *Applied and Environmental Microbiology*, **66**, 3835-3841.

57. Cremer, J., Eggeling, L. and Sahm, H. (1991) Control of the Lysine Biosynthesis Sequence in Corynebacterium-Glutamicum as Analyzed by Overexpression of the Individual Corresponding Genes. *Applied and Environmental Microbiology*, **57**, 1746-1752.

58. Liu, X., Li, L.L., Liu, J.C., Qiao, J.J. and Zhao, G.R. (2019) Metabolic engineering Escherichia coli for efficient production of icariside D2. *Biotechnol Biofuels*, **12**.

59. Kricka, W., Fitzpatrick, J. and Bond, U. (2014) Metabolic engineering of yeasts by heterologous enzyme production for degradation of cellulose and hemicellulose from biomass: a perspective. *Front. Microbiol.*, **5**.

60. Kerkhoven, E.J., Lahtvee, P.J. and Nielsen, J. (2015) Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Res.*, **15**, 1-13.

61. Clausen, H., Wandall, H.H., Steentoft, C., Stanley, P. and Schnaar, R.L. (2015) In rd, Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H. *et al.* (eds.), *Essentials of Glycobiology*, Cold Spring Harbor (NY), pp. 713-728.

62. Tripathi, N.K. and Shrivastava, A. (2019) Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. *Front Bioeng Biotechnol*, **7**, 420.

63. Tejwani, V., Andersen, M.R., Nam, J.H. and Sharfstein, S.T. (2018) Glycoengineering in CHO Cells: Advances in Systems Biology. *Biotechnol. J.*, **13**, e1700234.

64. Langer, E.S. and Rader, R.A. (2017) Top Trends in Biopharmaceutical Manufacturing, 2017 Innovation speeds discovery,drives down costs,and improves productivity. *BioPharm Int.*, **30**, 10-+.

65. Lim, Y., Wong, N.S., Lee, Y.Y., Ku, S.C., Wong, D.C. and Yap, M.G. (2010) Engineering mammalian cells in bioprocessing - current achievements and future perspectives. *Biotechnol Appl Biochem*, **55**, 175-189.

66. Dangi, A.K., Sinha, R., Dwivedi, S., Gupta, S.K. and Shukla, P. (2018) Cell Line Techniques and Gene Editing Tools for Antibody Production: A Review. *Front. Pharmacol.*, **9**, 630.

67. Bebbington, C.R., Renner, G., Thomson, S., King, D., Abrams, D. and Yarranton, G.T. (1992) High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. *Biotechnology. (N. Y.)*, **10**, 169-175.

68. Urlaub, G., Kas, E., Carothers, A.M. and Chasin, L.A. (1983) Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells. *Cell*, **33**, 405-412.

CONTRIBUTIONS OF THIS WORK

Chapter I: Genetically encoded betaxanthin-based small-molecular fluorescent reporter for mammalian cells

Reporter systems are an important tool in basic research and especially in bioengineering. The ability to visualize changes in cellular processes in a quick, simple and reliable fashion plays an essential part in shaping the modern field of bioengineering. Together with ever faster high-throughput DNA sequencing, engineers have the tools to develop highly modular approaches for engineering the biochemical functions in living cells. With the low cost of DNA synthesis, we are in a synthetic biology era where an idea can be genetically designed, the custom DNA sequence ordered, cells transfected and the setup tested with reporter systems in a short amount of time. Current available reporter systems are highly useful and all offer certain advantages, as well as disadvantages. It catches the eye that, with the exception of few niche reporters, all reporter systems are based on either enzymatic assays or proteins with optical properties, instead of the direct in cellula production of a small molecular dye. Life knows a huge variety of dyes and pigments and for many the exact metabolic pathways and genes are known. For example, plants and fungi offer a whole range of colors, but the transfer to mammalian systems is rather challenging. In this chapter we present a small molecule reporter system that uses a combination of human enzymes and enzymes from Amanita muscaria to produce yellow dyes from the betaxanthin class. We demonstrate how these enzymes can be used to create a novel reporter system that allows for direct quantification of gene expression in the culture supernatant.

Chapter II: CelloSelect – a synthetic cellobiose metabolic pathway for selection of stable transgenic CHO-K1 cell lines

Creating mammalian production cell lines is essential in the modern manufacturing of protein-based biopharmaceuticals. The process of creating these cell lines relies on the genomic integration of genes producing the therapeutic protein and subsequently applying and maintaining an artificial selection pressure to select only cells producing the cargo. Leaps forward have been made with respect to the stable integration strategies (such as transposon-based or CRISPR/Cas9-based), but the selection strategies used in industrial

settings seem to remain rather unchanged. Current selection methods are based on knockout cell lines or the use of toxic agents. In the knockout cell line approach, an essential metabolic gene is removed from the cells' genomes. Including this removed gene in the genetic construct as a selection marker only allows the stably transfected cells to survive. Akin to that, a detoxifying enzyme can be used to intracellularly neutralize a toxic agent. To date, the knockout approach is mostly used in industry but is, although effective, destructive in nature. Cell lines that are created for an industrial use are previously engineered with a focus not on production, but that the cells can be used for selection later. In this chapter we present an additive approach based on metabolic engineering that grants the cells the ability to consume a foreign nutrient source. We created a cellobiose utilization pathway that awards cells with the ability to metabolize cellobiose as sole source of glucose. In a proof-of-concept study, we developed the CelloSelect stable selection system that, in the absence of glucose, can be used to select CHO-K1 cell lines expressing the CelloSelect gene-cassette, containing a cargo gene. All, without the use of knockout cell lines or the use of antibiotics and toxic agents. We developed a step-by-step protocol for the use of the system and show how it can be used to create engineered cell lines, that stably produces the therapeutic protein erythropoietin (EPO). The cells can be grown in culture medium containing either glucose or cellobiose as major carbon and energy source.

CHAPTER I

GENETICALLY ENCODED BETAXANTHIN-BASED SMALL-MOLECULAR FLUORESCENT REPORTER FOR MAMMALIAN CELLS

This chapter is based on the authors version of a manuscript that has been submitted for publication.

Pascal Stücheli, Simon Sieber, David W. Fuchs, Leo Scheller, Tobias Strittmatter, Pratik Saxena, Karl Gademann, Martin Fussenegger

ABSTRACT

We designed and engineered a dye production cassette encoding a heterologous pathway, including human tyrosine hydroxylase and *Amanita muscaria* 4,5-DOPA dioxygenase, for the biosynthesis of the betaxanthin family of plant and fungal pigments in mammalian cells. The system does not impair cell viability, and can be used as a non-protein reporter system to directly visualize the dynamics of gene expression by profiling absorbance or fluorescence in the supernatant of cell cultures, as well as for fluorescence labeling of individual cells. Pigment profiling can also be multiplexed with reporter proteins such as mCherry or the human model glycoprotein SEAP (secreted alkaline phosphatase). Furthermore, absorbance measurement with a smartphone camera using standard application software enables inexpensive, low-tech reporter quantification.

Graphical abstract
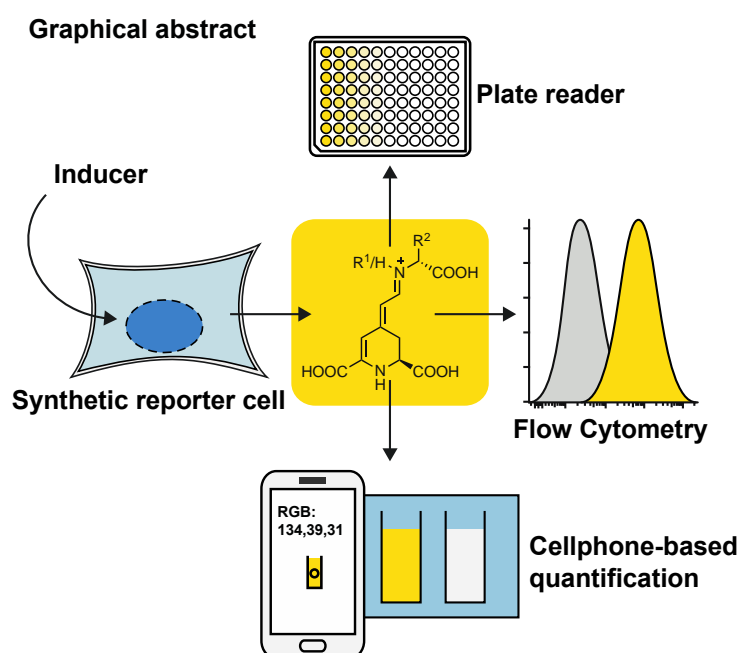
Synthetic mammalian reporter cells transfected with an engineered production cassette for heterologous betaxanthin production can be used to monitor gene expression. Upon induction with the designated inducer the cells produce a yellow fluorescent water-soluble dye that can be monitored with a variety of different methods.

INTRODUCTION

Colored or fluorescent proteins have been used extensively both as intracellular markers for microscopy and as reporter systems for gene expression(1) since the first isolation of green fluorescent protein from *Aequorea victoria* in 1998(2). In general, such reporters enable quantification of gene expression inside a single cell or across cell populations by producing a quantifiable protein(3). The most widely used reporter systems are based on fluorescent proteins, alkaline phosphatases(4) or luciferases(5). Fluorescent proteins are particularly well suited for single-cell analysis and for observing gene expression dynamics by continuous measurements(6). Another approach for continuous tracking of cellular behavior with enzymatic reporter systems is frequent sampling of the supernatant. However, this approach suffers from limited sampling frequency and labor-intensive sample preparation.

To date, the toolbox of fluorescent proteins for analyzing gene expression consists of more than a hundred members, with excitation and emission profiles ranging from near-UV to infrared, and numerous modifications are available for use in various experimental setups(7). Nevertheless, protein-based reporters can have disadvantages compared to small-molecular reporters. Small molecules are often able to passively penetrate cell membranes and can therefore diffuse into or out of cells, and enter most subcellular compartments. This behaviour enables measurements at the single cell or whole population level in the same setup, obviating the need for different reporter constructs. Additionally, secretion of protein reporters is not always trivial, as proteins may undergo glycosylation, form disulfide bonds, oligomerize while passing the endoplasmic reticulum, or require the addition of secretion signals, all of which can compromise cellular production capacity(8). In addition, small molecules tend to be resistant to denaturing conditions; this is particularly advantageous for sample preparations that require cell fixation, which often causes protein reporters to lose functionality. Lastly, small molecules with suitable optical properties can be directly quantified by absorbance or fluorescence measurements of the culture medium, without the need for laborious assays. Heterologous gene expression in mammalian cells is well established, but so far, only a few non-native small-molecular dyes or pigments, which are widespread in plants, have been successfully produced in mammalian cells(9,10). Differences in biochemical and biophysical properties (optimal temperature, salt concentration(11,12), as well as missing biochemical pathways(13), and even the absence of suitable reaction

compartments (organelles)(14) in mammalian cells make the task challenging. In addition, the actual biosynthetic pathways of dyes are often mediated by cascades of specialized enzymes that are all required to work in synchrony(15). Few reporter systems based on small molecules have been reported to date, and those that are available either employ an external substrate that is enzymatically converted(16) or are not water-soluble, so that supernatant sampling is not applicable(9,17).

Among the huge variety of plant dyes, the betalain class(18) appears to have suitable characteristics for heterologous production in mammalian cells, and indeed the use of betalains as reporters in plants has been proposed(19). The water-soluble betalains are L-DOPA-derived, yellow-orange to red-purple dyes produced by various plants and fungi(20), including the well-known *Amanita muscaria* (*A. muscaria*; fly agaric) Interestingly, the red dye betanin, found in red beet, is widely used in the food industry as natural food colorant(21). The biosynthesis of betalain family members follows the same core pathway from L-tyrosine(18) in a diverse set of organisms (Fig. 1a). Briefly, L-tyrosine is oxidized by tyrosinase (TYR) to L-DOPA, which is converted to betalamic acid by 4,5-DOPA dioxygenase (DODA)(22,23). Betalamic acid spontaneously(24) reacts with a variety of amine sources in the cytosol to yield the corresponding betaxanthins, which are a sub-group of the betalains.

Here, we describe the design and engineering of a betaxanthin production cassette consisting of a heterologous biosynthetic pathway, including human tyrosine hydroxylase and DODA from *A. muscaria*, for the biosynthesis of yellow-fluorescent indicaxanthin. We demonstrate the suitability of this system for continuous measurements of gene expression dynamics in mammalian cells at the population level, as well as for labeling individual cells. We present a simple, low-tech assay, using a smartphone-based set-up for the quantification of betaxanthins in cell-culture supernatants(25), that can easily be used by non-scientific personnel.

RESULTS AND DISCUSSION

Design and validation of betaxanthin production in mammalian cells

We initially examined whether DODA enzymatic activity could be achieved in mammalian cells by transfecting HEK293T cells with a plasmid encoding DODA from *A. muscaria*(29,30) (AmDODA, *DODA*, P87064). L-DOPA (1 mM) and ascorbic acid (0.05 mM) were added to the cell culture, and after 16 h we observed a yellow coloration of the supernatant, indicating successful biosynthesis of the dye, and confirming that it can cross the plasma membrane. Freshly prepared ascorbic acid solution was used in exp eriments with medium containing L-DOPA in order to prevent oxidation of L-DOPA. It was used at a concentration of up to 0.1 mM; this was confirmed to be non-toxic in HEK293T cells (Fig. S1). As betaxanthins were previously reported to be fluorescent(31), we recorded a 2D fluorescence scan of the supernatant (Fig. S2), in which we identified a hot spot at around 485 nm excitation and 507 nm emission wavelengths. Comparison of the supernatant fluorescence of cells transfected with the AmDODA-encoding plasmid with that of cells transfected with mock plasmid further supported the successful expression of functional DODA in mammalian cells for the first time (Fig. 1b).

We then focused on achieving autonomous production of the dye by mammalian cells. Intriguingly, the precursor L-DOPA is already produced by some mammalian cells *via* either the dopamine pathway (by tyrosine hydroxylase (TH)(32)) or the melanin pathway (by tyrosinase (TYR)(33)). However, human TYR targeted to the cytosol is unlikely to be functional, as it lacks endoplasmic reticulum-based glycosylation(34) and thus the combination with cytosolic DODA would probably be ineffective. In addition, human TYR is localized to endosomes in non-melanogenic cells, and these organelles would be a difficult engineering target for localization of DODA. Therefore, we decided to focus on the TH pathway for L-DOPA production in mammalian cells. For this purpose, we established ectopic L-DOPA production by transfecting HEK293T cells with plasmids coding for human TH(32) (hTH, *TH*, AAI04968) and human GTP-cyclohydrolase (hGCH, *GCH1*, NP_000152). We were able to observe the production of L-DOPA by UHPLC-MS (Fig. S3) and the concentration of the compound could be quantified. hGCH is involved in the biosynthesis of tetrahydrobiopterin (THB(35)), a cofactor of hTH, and its expression was necessary for high L-DOPA production in HEK293T cells.
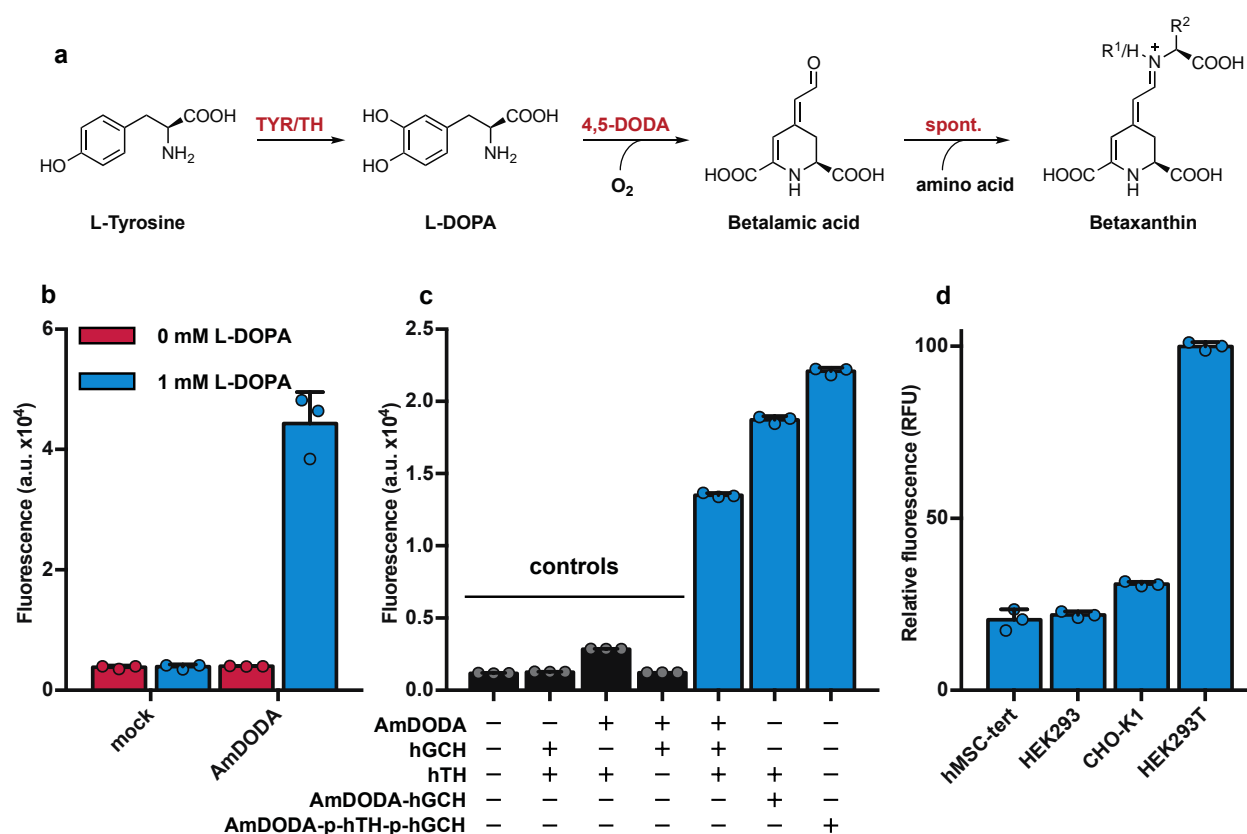
Figure 1: Overview and characterization of the heterologous betaxanthin production system. (a) Part of the biosynthetic pathway of the betaxanthin class of pigments. Tyrosine is oxidized by either tyrosine hydroxylase (TH) or tyrosinase (TYR) to form L-DOPA. A specific 4,5-DOPA dioxygenase (4,5-DODA) oxidizes L-DOPA to yield betalamic acid (BA), utilizing molecular oxygen. BA spontaneously reacts with various amino acids to yield betaxanthins. (b) Functionality test of DODA from *A. muscaria*. HEK293T cells were transfected with mock (pCOLADuet-1) or *DODA* (pPST320). After 48 h the medium was changed to a clear medium containing 1 mM L-DOPA and 0.05 mM ascorbic acid, and color development was measured 16 h later. ■, with L-DOPA; ■, with ascorbic acid only. (c) Functionality test of the complete betaxanthin production cascade. Cells were transfected according to the table, and color development was measured 48 h later. ■, negative controls; ■, complete production cascades. AmDODA (pPST320), hGCH (pPST321), hTH (pPST319), AmDODA-hGCH (pPST322), AmDODA-p-hTH-p-hGCH (pPST324). (d) Functionality test of the complete betaxanthin production cascade in different cell lines. The cells were transfected with pPST324 and color development was measured 48 h later. In the shown dataset the signal-to-noise ratios are between approximately 214 and 386 (calculated as average signal above background divided by the standard deviation of the background). Raw data for the lowest- and highest-producing cell lines can be found in Table S4. In experiments b and c, color development was measured in arbitrary fluorescence units, while in experiment d it was measured in relative fluorescence units normalized to the mock-transfection

samples (0) and HEK293T cells transfected with pPST324 (100). Graphs in b, c, and d show the mean ± s.d. of n = 3 independent samples and are representative of three independent experiments.

Next, having established both parts of the betaxanthin production pathway separately, we transfected HEK293T cells with plasmids encoding hTH, AmDODA and hGCH, and examined the functionality of the cells for fully autonomous betaxanthin production by means of fluorescence measurement (Fig. 1c). Indeed, the cells produced a fluorescent dye. This confirms the functional adaptation of these mammalian cells for heterologous secondary metabolite production through the introduction of a dye production cassette encoding a combination of human and fungal enzymes. Importantly, the system is functional in standard cell culture medium without the addition of ascorbic acid, L-DOPA or other special additives.

In order to increase the dye production and reduce the number of individual genetic components, we evaluated fusion versions of the introduced genes. We found that expression of an AmDODA-hGCH fusion protein in HEK293T cells resulted in increased dye production (Fig. 1c). The greatest increase in fluorescence was found in cells transfected with a plasmid encoding hTH, hGCH and AmDODA in a single mRNA fused together by means of a P2A sequence (plasmid pPST324, plasmid map in Fig. S9) (Fig. 1c). This construct was used for all further benchmarking experiments. P2A is a member of the *Herpes simplex* virus 2A ribosomal-skipping peptide family. During translation, there is no peptide bond formed at the P2A site, which results in two separate polypeptides.

To confirm the generality of this system, we next evaluated multiple cell types for DODA-based dye production by transfecting them with pPST324. The system was indeed functional in multiple cell lines. For the cell types shown, the measurements were significantly above background, with good signal-to-noise ratios of at least 200 (Fig. 1d). Naturally, the background in fluorescence measurements is highly dependent on experimental factors such as the specific culture medium, analysis device and settings used. Differences in reporter production levels in standard laboratory cell lines have been described before (e.g. (36)) and are likely due to specific intrinsic metabolic differences that result in different protein production capacities, or due to experimental differences (transfection efficiency or transgene dependency)(37)). As HEK293T cells produced the highest levels of fluorescence, we utilized this cell line for subsequent experiments.

Additionally, we investigated a secreted version of the dye production cassette to broaden the applicability of the system. For this purpose, we designed secreted variants of AmDODA (sAmDODA) and employed a secreted version of a tyrosinase(38) from *Celosia cristata* (sCcTYR, *CYP76AD4*, AGI78466), which was shown to be highly active in transgenic yeast and independent of specialized cofactors(39). Cotransfection of HEK293T cells with plasmids encoding sAmDODA and sCcTYR resulted in increased fluorescence (Fig. S4). However, the overall fluorescence of this secreted system was lower than that of the hTH-based system. Hence, we focused on the latter system to further characterize the dye and to explore possible applications.
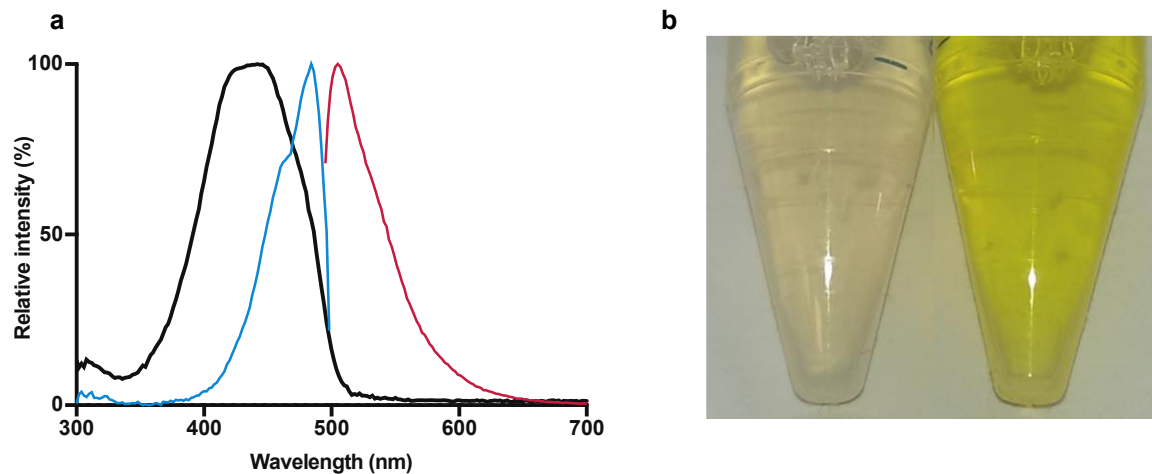

Figure 2: Characterization of the reaction product. (a) Fluorescence and spectroscopic analyses of the reaction product of AmDODA (pPST320) in mammalian cells. HEK293T cells were transfected with pPST320, and after 48 h the medium was changed to clear medium containing 1 mM L-DOPA and 0.05 mM ascorbic acid. Absorbance measurement and an excitation and emission scan were recorded 16 h later, and the background (supernatant from mock-transfected cells) was subtracted. ●, absorbance; ●, excitation; ●, emission. The data was normalized to the minimum and maximum intensities of each scan. The results of one measurement, which is representative of three independent experiments, are shown. (b) Photograph of betaxanthin-containing supernatant. Cells were transfected with a constitutive betaxanthin production plasmid (pPST324) or mock (pCOLADuet-1) and the image was recorded 72 h after transfection (right: supernatant containing betaxanthin).

Characterization of betaxanthins

To further characterize the reaction product in the present system we measured the absorbance and fluorescence spectra of the culture supernatant of dye-producing HEK293T cells (Fig. 2a; for a 2D fluorescence scan, see Fig. S2). An image of the supernatant of cells transfected with pPST324 after 72 h is shown in Fig. 2b. However, the absorbance maximum and the fluorescence excitation and emission maxima of 446 nm, 487 nm and 507 nm, respectively, did not correspond to the spectroscopic data of any single compound reported in the literature(24,31). We concluded that the sample contained a mixture of the precursor betalamic acid (405 nm λ_{max}) and betaxanthins (470 – 490 nm λ_{max}). To confirm the production of betaxanthins, we transfected HEK239T cells with pPST324 and analyzed the supernatant by UHPLC coupled to a DAD and MS/MS detector (Fig. S5). As various betaxanthins can be produced depending on the reaction partner of betalamic acid(40) (Fig. 1a), we added the amino acid L-proline (5 g/L) to the medium (containing 0.1 mM ascorbic acid) in order to push the reaction towards indicaxanthin (the reaction product of L-proline and betalamic acid) to facilitate its detection. The chromatogram extracted at the typical betaxanthin window (400-500 nm) exhibited a major peak at 1.24 min, which indeed showed the characteristic UV spectrum of indicaxanthin (absorption maximum at 478 nm; Fig. S5d). The identity of the compound was further confirmed by mass spectrometry in the selected reaction monitoring mode, detecting daughter ions at 263 Da (17 eV) and 217 Da (23 eV) (Fig. S5b,c)(41). Furthermore, we synthesized indicaxanthin according to a reported procedure(28), and confirmed that it was identical to the product in the supernatant of the transfected human cells. Different betaxanthins show very similar spectroscopic properties(24), so the exact composition of the mixture of pigments formed is not expected to significantly influence the absorbance characteristics.

To confirm the suitability of betaxanthin for use as a fluorescent reporter, we characterized its photostability in cell culture supernatant. Specifically, we analyzed the decrease in fluorescence of the supernatant of betaxanthin-producing cells in a plate reader with continuous measurement in comparison with that of the well-known small-molecular fluorophore fluorescein (Fig. S6). Fluorescein is less prone to photobleaching than betaxanthins, with its half-life being approximately four-fold longer.

Impact of betaxanthin production on cell viability and metabolism

An important consideration for a live-cell reporter system is its effect on cell viability and growth rate and its potential for interference with native signal transduction pathways. In particular, the toxicity of the system should be examined thoroughly, since the reaction intermediate L-DOPA is a biologically active compound that triggers apoptosis in some cell lines(42) and is converted to neurotoxic dopamine in some neuronal cell lines(43). Therefore we analyzed its effect on cell viability in three different assays. We assessed viability by using resazurin dye, which is converted to resorufin in viable cells only. Resorufin can then be quantified using spectroscopic methods(44). Clearly, the signal intensity is dependent on the metabolic activity of the cells, as well as the total cell number. As a first step, we looked at the viability of cells transfected with the betaxanthin production cassette compared to cells transfected with a plasmid encoding the fluorescent protein tGFP (Fig. S7a). Then, we analyzed the impact of medium containing betaxanthins on wild-type cells (Fig. S7b). Lastly, we measured growth curves and compared cells producing either betaxanthins or tGFP (Fig. S7c). Betaxanthin production did not impair the viability of HEK293T cells in any of these assays.

Moreover, in order to identify any impact on cellular metabolism and signaling pathways we quantified the changes in gene expression of 96 different endogenous genes belonging to 18 different signaling pathways and 13 housekeeping genes using an RT-qPCR array (Fig. S8). Of these 96 genes, 26 were below the detection limit and only 15 were significantly different in betaxanthin-producing cells compared to tGFP-producing control cells. Two of the 15 were upregulated in betaxanthin-producing cells; they were *FASN* (fatty acid synthase) and 18s rRNA (with $\Delta\Delta$ct of 3.1 ± 0.8; 2.4 ± 0.4). The two lowest-expressed genes were *EGR1* (early growth response protein 1) and *CDKN1B* (cyclin-dependent kinase inhibitor 1B) (with $\Delta\Delta$ct of -2.2 ± 0.3; -3.4 ± 0.8). The reasons for these differences in gene expression are unclear. One possible factor would be that fluorescent proteins increase the production of reactive oxygen species(45), and could thus lead to a relative increase of some markers in the GFP-producing control cells. Taking these results together with the previous findings, we do not consider these changes in gene expression particularly worrisome for the utility of the system, although use of this system in neuronal cell lines would need to be evaluated individually, as

they might react sensitively to L-DOPA production. Moreover, AmDODA alone could serve as an L-DOPA sensor in cells that already produce it.

Profiling mammalian gene expression with betaxanthin-based reporter systems

To further evaluate this reporter, we established a trigger-inducible system as illustrated in Fig. 3a (plasmid maps in Fig. S9). A fusion protein of AmDODA and hGCH was constitutively expressed, while hTH was put under the control of an inducible promoter. As transcription of all three genes is necessary for efficient dye production, only one gene needs to be inducible to fully regulate betaxanthin production. We chose hTH for this role since it is the first enzyme in the dye production cascade. Furthermore, it is likely that if hTH is expressed constitutively, L-DOPA would be produced in the cells and would slowly generate black melanin and potentially become toxic to the cells. Additionally, as cell-to-cell heterogeneity of AmDODA or hGCH would lead to variation in reporter production, we wanted to ensure that the two enzymes are highly overproduced and would not become a bottleneck. Heterogeneity in hTH levels is expected to lead to unpreventable variation in reporter production, as is the case for all transiently transfected reporter systems.

We created a model test setup by using a doxycycline-driven promoter for the inducible production based on a P2A-fusion protein of hTH in tandem with the benchmarking reporter phosphatase SEAP (human placental secreted alkaline phosphatase, *ALPP*, AAB64400.1)(4), to correlate the two reporter systems (plasmid map in Fig. S9). HEK293T cells were transfected with the constitutive AmDODA-hGCH and the inducible hTH-SEAP construct, then reseeded into medium containing different amounts of doxycycline, and the reporter activities were quantified 48 h later (Fig. 3b). Notably, the reporter activities were quantified from the same samples, thus demonstrating the multiplexing capabilities of the betaxanthin system. The induction curves of the two reporters are highly similar.

For real-time profiling of reporter production we transfected HEK293T cells with different amounts of pPST324 and followed the fluorescence generation at 30 min intervals for 72 h. A dose-dependent increase in fluorescence was observed (Fig. 3c). To get more insight into the temporal dynamics of reporter gene expression, we calculated the derivative of the curves in Fig. 3c. The maximum increase in fluorescence was observed after approximately 35 h (Fig. S10a). This curve presumably reflects the exponential growth of the cell population, the initial

lag before enzyme production following transfection, enzyme kinetics, dye diffusion, and nutrient or oxygen limitations. However, a comparison of the continuous color production with the growth curve of transfected cells (Fig. S10b) suggests that the peaking increase in fluorescence is not simply caused by the population growth dynamics.
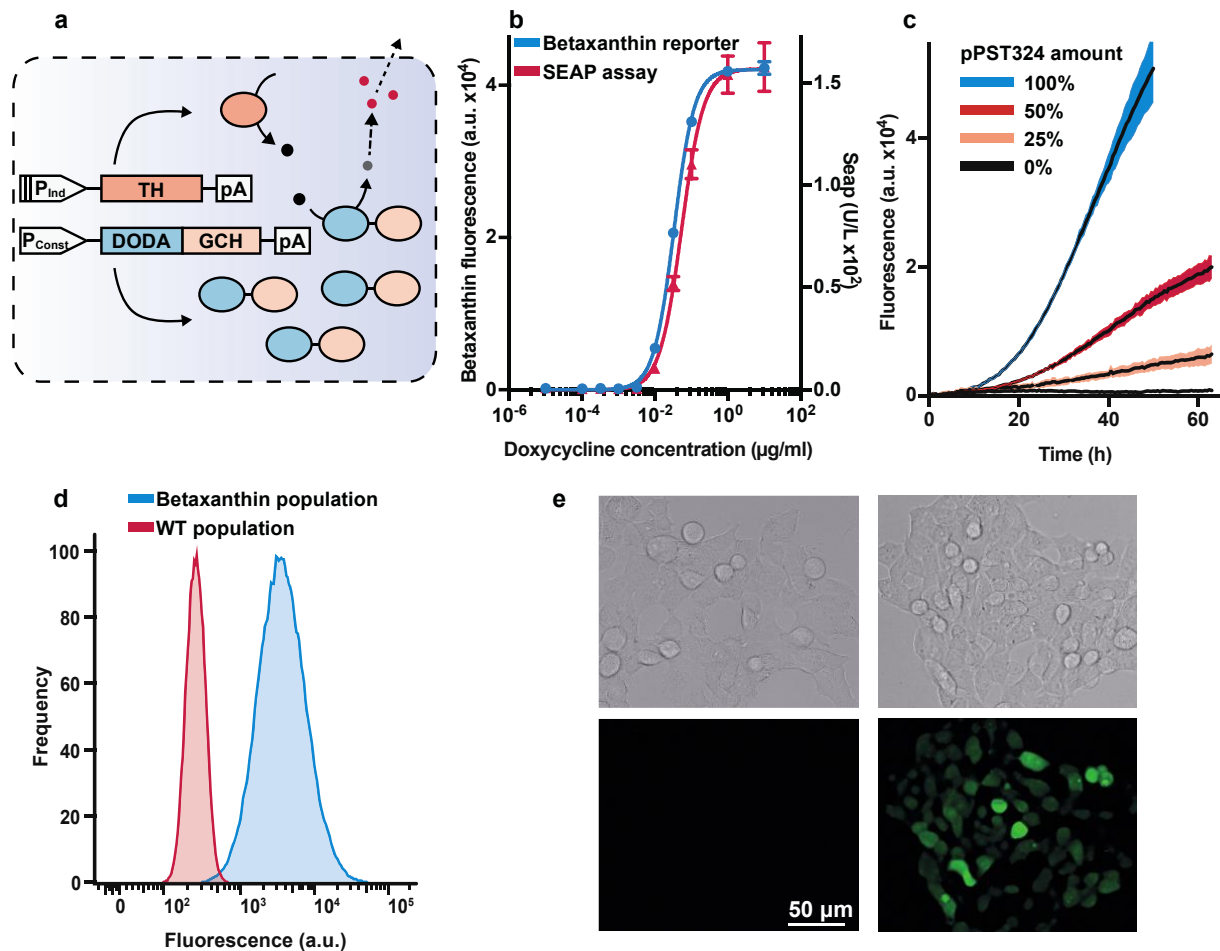

Figure 3: Application of the betaxanthin production system as a reporter. (a) Schematic representation of the betaxanthin reporter system. A direct fusion of DODA (4,5-DOPA dioxygenase) and GCH (GTP-cyclohydrolase) is constitutively expressed. Upon trigger-induced expression of TH (tyrosine hydroxylase) L-DOPA is produced, and in turn is converted to betalamic acid (BA), leading to the final reaction product (betaxanthins). P_{Ind}, inducible promoter; P_{Const}, constitutive promoter; pA, poly-A tail; ● , L-DOPA; ● , BA; ● , betaxanthin. (b) Comparison of betaxanthin as a reporter system with the widely used SEAP reporter system. Systems are expressed in tandem via a P2A under control of a doxycycline (dox)-inducible model promoter. pPST322 (AmDODA-hGCH1), pPST350 (SEAP-hTH) and pTS1105 (transactivator) were transfected. Each reporter activity was measured in the same supernatant first as betaxanthin fluorescence and second as SEAP activity, and a sigmoidal curve was fitted to the observations. The induction concentrations of doxycycline are shown below the graph. (c) Continuous betaxanthin

production assay. Cells were transfected with pPST324 and placed in a fluorescence reader. The fluorescence was measured every 30 min over the course of 72 h. The colors represent the percentage of active plasmid (pPST324), made up to 100% with mock plasmid (pCOLADuet-1). Fluorescence was normalized to the first data point in each measurement series to adjust for well-to-well differences. (d) Flow cytometry analysis of intracellular betaxanthin. Cells were transfected with pPST324 (●) or with empty vector pCOLADuet-1 (●) and analyzed 48 h later. Gating was done to exclude dead cells and doublets. A total of 100,000 raw data points were collected per sample. The population data shown is representative of three independent experiments. (e) Micrographs of cells producing betaxanthin. Left, mock-transfected cells; right, betaxanthin-producing cells; top, bright field; bottom, deconvoluted green fluorescence. Cells were transfected with either pColaDuet-1 or pPST324 and examined 72 h after transfection. The images are representative of three independent experiments. For graphs b and c dye development was measured in arbitrary fluorescence units (485 nm ex. | 507 nm em.). Graphs b and c show the mean ± s.d. of n = 3 biologically independent samples and are representative of three independent experiments.

The fluorescence of cells transfected with the betaxanthin production cassette was significantly above background within approximately 6 hours after transfection (Fig S10c). Various unspecific and system-specific limitations exist that limit the temporal resolution of this system. Unspecific delays, such as delays after transient transfection, initial transcription or translation are reporter system-independent. Thus, the reporter system would simply reflect target gene dynamics. Specific limitations are most likely due to the delay between protein translation and signal detection (enzyme secretion in the case of SEAP, fluorophore maturation in the case of fluorescent proteins(46), and enzyme and chemical reaction rates in the case of the betaxanthin system), as well as differences in the reporter sensitivity.

Dye localization and comparative analyses of fluorescent proteins with small-molecular betaxanthine-based reporter systems

As betalamic acid is produced in the cytosol, it is likely that intracellular betaxanthin can be measured dose-dependently and thereby could be used as a single-cell reporter. To test this, we transfected HEK293T cells with different amounts of pPST324. After 48 h, flow-cytometric analysis showed a dose-dependent shift in population fluorescence (Fig. 3d). To understand the distribution of the dye between cells and supernatant, we transfected cells with the betaxanthin production cassette and collected the supernatant and harvested PBS-

washed cells each day for 72 h (Fig. S11). Interestingly, the intra- and extracellular dye concentrations seemed to remain in an approximately constant ratio. We calculated the ratio of total fluorescence (the product of fluorescence and analysis volume) outside versus inside the cells to be 88 ± 8, meaning that at least 98% of the dye can be found extracellularly. To definitively visualize the intracellular dye accumulation we obtained a time-lapse microscopy recording (Supplementary Movie 1) of HEK293T cells transfected with pPST324. It can clearly be seen that, as is to be expected with transient transfection, different cells produce different amounts of dye, which subsequently appears in the background medium. Deconvoluted images of cells producing betaxanthin in fresh clear culture medium can be seen in Fig. 3e. Additionally, we used blue light to obtain bright-field micrographs (Fig. S12), which resulted in the dye absorbing the light and appearing dark in the images; this approach could potentially enable the detection of the dye in terms of absorbance instead of fluorescence.

To examine the multiplexing capabilities of this system at the single-cell level, we cotransfected HEK293T cells with pPST324 and a plasmid coding for mCherry fluorescent reporter protein. Flow cytometry revealed cells exhibiting both green and red fluorescence (Fig. S13).

Chemical fixation of cells prior to analysis with a flow cytometer or a microscope is required for many applications in cell biology. Therefore, we wanted to see whether betaxanthin-stained cells would offer an advantage over fluorescent proteins, as fluorescent proteins tend to lose their fluorescence during fixation. We fixed cells producing GFP or betaxanthin and analyzed them using a flow cytometer (Fig. S14). Although the GFP fluorescence intensity was stronger than the betaxanthin fluorescence prior to fixation, GFP fluorescence vanished completely following fixation, whereas betaxanthin fluorescence remained stable or even increased in intensity.

Simple, inexpensive betaxanthin assay using a smartphone

As the yellow dye is clearly visible to the naked eye, we next designed a smartphone-based set-up suitable for inexpensive sample profiling by non-scientific operators in remote field situations (see Fig. 4a for a schematic overview; Fig. S15 shows a photograph of the device and screenshots of the actual measurements). RGB values (red green blue standard additive color mode) were extracted from images of a sample-containing cuvette in front of a blue background, and the betaxanthin content of the sample was determined from the blue

portion of the RGB values and a calibration curve. We produced an indicaxanthin-specific calibration curve (0 - 2.24 mg/mL) to quantify the dye in cell-culture samples (Fig. 4b). With this methodology, we were able to measure the betaxanthin concentration of 1.04 mg/mL in the supernatant of HEK293T cells transfected with pPST324. Negative-control cells transfected with pColaDuet1 showed an insignificant dye concentration (0.06 mg/mL), confirming that the standard cell-culture medium does not interfere with assay performance (Fig. 4c).
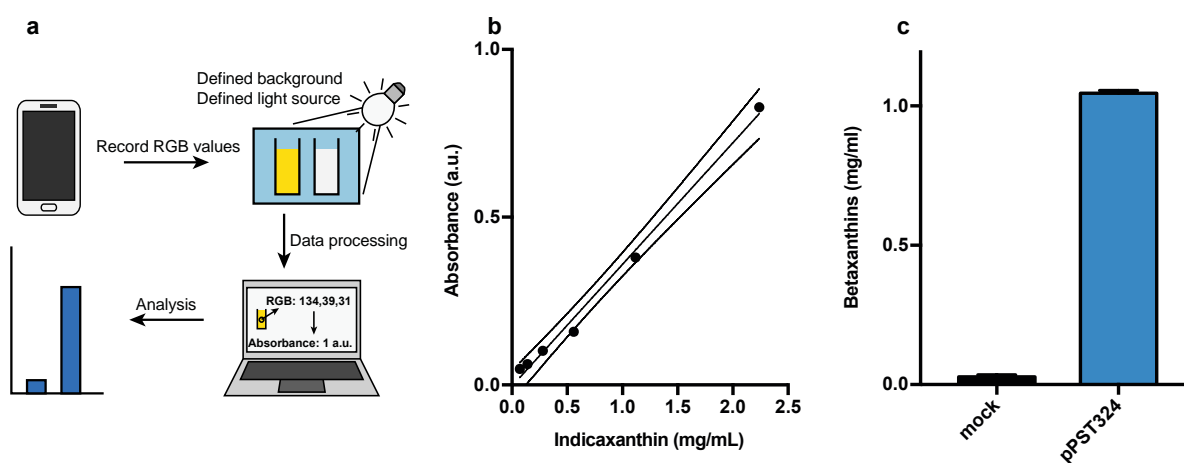

Figure 4: Smartphone-based quantification of betaxanthins in cell culture supernatant. (a) Schematic overview of the method. The color information is extracted in the form of RGB values from a sample in a cuvette in front of a blue background shielded from light interference, and processed to yield the absorbance or the betaxanthin concentration of the sample. RGB, red green blue standard additive color mode. (b) Standard curve of indicaxanthin absorbance created using the above method. The dashed line indicates the 95% CI (c) Quantification of betaxanthins in supernatant of cells transfected with either pPST324 or pColaDuet-1 after 72 h. Graphs b and c show the mean ± s.d. of n = 3 technical replicates and are representative of three independent experiments.

CONCLUDING REMARKS

We believe the toolbox presented here has considerable potential for further expansion in the future. For example, it might be feasible to produce red betanin in mammalian cells, since betanin has been produced as a mixture with betaxanthins in yeast(39); this could enhance the versatility of the betalain reporters. To make the betaxanthin system faster and more sensitive, the efficiency of the fungus-derived AmDODA in mammalian cells might be increased by mutagenesis. Furthermore, application of this system as an intracellular L-DOPA sensor, for example in dopaminergic neuronal cells, might provide a useful tool for basic research.

In summary, we present the first engineered device for heterologous biosynthesis of a plant- or fungus-derived dye in mammalian cells. We believe this genetically encoded fluorescent small-molecular reporter system will be especially useful for examining gene expression dynamics, but it should also complement or replace existing reporters in various applications, such as large-scale drug screening, where current reporter assay-based methods are time-consuming, often involving multiple pipetting and incubation steps, and the necessary reagents are relatively expensive. The dye shows bright yellow fluorescence, visible to the naked eye, which can be quantified using standard laboratory instruments and profiled with a smartphone by non-scientific personnel in field situations at remote locations. The dye can be directly quantified in regular cell culture medium, since there is no requirement for any additional reagent or dedicated sample processing. It does not impact substantially on cell viability, and it is equally functional in live and fixed cells. We believe this reporter system will be suitable for a broad range of studies, including whole-population measurements, single-cell quantification, continuous monitoring, and multiplexed reporter detection.

Funding

Author contributions

Conflict of interest

The authors declare no competing financial interests.

MATERIALS AND METHODS

DNA constructs

Construction of the plasmids is described in detail in Table S1 and S2. DNA sequences for AmDODA and CcTyr can be found in DNA sequences S1 and S2. *Escherichia coli* strain XL10-Gold (Agilent Technologies) was used for cloning.

Cell culture and transfection

HEK293T cells (DSMZ: ACC-635), HEK293 cells (DSMZ: ACC 305), CHO-K1 cells (ATCC: CCL-61) and human mesenchymal stem cells transgenic for the catalytic subunit of human telomerase (hMSC-TERT)(26) were cultivated in DMEM (Thermo Fisher, cat. No. 31053028) supplemented with 10% FCS (Sigma-Aldrich, cat. No. F7524) and 1x Glutamax (Thermo Fisher, cat. No. 35050061) at 37 °C in a humidified atmosphere containing 7.5% CO_2. For CHO-K1 cells, 0.15 μM L-proline (Fluka) was also added.

For serial passage of these cells, 0.05% trypsin–EDTA (Gibco) was used. The cells were generally passaged on 10 cm dishes at 80-90% confluency after 48 h. For transfection, 1.5 × 10^6 cells (counted with a CASY TTC Cell Counter) in 14.4 ml of medium were seeded on a 96-well cell culture plate (150 µl cell suspension per well) on the evening before transfection. For different plate formats, the amount of suspension per well was varied accordingly. The medium for seeding and transfecting the cells as well as for the cell-based assays was FluoroBrite™ DMEM (Thermo Fisher, cat. No. A1896701) supplemented with 10% FCS (Sigma-Aldrich), 1x Glutamax (Thermo Fisher) and 1% penicillin–streptomycin solution (Biowest) for all cell types, plus 0.15 μM L-proline (Fluka) for CHO-K1 cells. L-DOPA (Sigma-Aldrich, cat. No. D9628) as a 10 mM stock solution in DMEM and ascorbic acid (Sigma-Aldrich, cat. No. 11140) as a 1 M stock solution in dH_2O were added where necessary. For transfection in a 96-well plate format, a DNA–polyethyleneimine (PEI) mixture in DMEM without supplements (50 µL/well) was produced by incubating 0.75 µL PEI (40 kDa MW, Polysciences; stock solution 1 mg/mL in dH_2O) with 150 ng total plasmid DNA. The mixture was vortexed for 3 s and incubated at room temperature for 15 min. (When necessary for different plate formats, the transfection mix was scaled up accordingly.) The cells were incubated with the transfection mixture for 6.5-7.5 h, and then the medium was exchanged for 100 µL fresh, prewarmed

medium. For the photograph (Fig. 2b) only half of the medium (6 mL instead of 12 mL) was used in the medium change after the transfection in order to increase the color intensity. For long-term continuous measurement (Fig. 3c), the cells were placed in an incubator for 30 min to allow the medium and air in the plate to equilibrate to the desired pH, CO_2 concentration and humidity, then sealed and placed in a plate reader heated to 37 °C for cultivation. For the doxycycline-induction experiments, a 1.5 mg/ml stock solution of doxycycline hyclate (Sigma-Aldrich, D9891) in dH_2O was used and the cells were transfected in a 6-well plate format and reseeded after transfection into fresh clear medium containing doxycycline at a density of 3 $\times$ 10^6 cells per plate. Table S3 shows details of the transfection mixes.

L-DOPA quantification

The supernatant was filtered and analyzed by a UHPLC-MS/MS system composed of an Ultimate 3000 (Thermo Fisher) and a mass detector (TSQ Quantum Ultra, Thermo Fischer) equipped with a reverse-phase column (Kinetex® EVO C18; 50 x 2.1 mm, 1.7 μm, Phenomenex) in the SRM negative mode targeting the typical L-DOPA fragmentation of 196 Da to 135 Da with a collision energy of 19 eV. A calibration curve was built using standard solutions in the concentration range from 50 μg/mL to 0.5 μg/mL. Intracellular L-DOPA was extracted from the cells by suspending frozen cells in methanol, followed by centrifugation and filtration of the supernatant.

Fluorescence and absorbance analysis

For end-point measurements, 80 μL of the 100 μL supernatant was transferred to a clear 96-well plate. The fluorescence was measured in a Tecan infinite® M1000 pro plate reader unless otherwise noted, with 485 nm excitation and 507 nm emission (5 nm bandwidth) in a top-reading mode. The absorbance and fluorescence scans for Fig. 2a were created as above with a 2 nm step distance for absorbance and fluorescence. For the excitation scan, the emission was measured at 507 nm and for the emission scan, the excitation was kept at 485 nm. Continuous measurement and cultivation were performed in a Tecan infinite® M200 with excitation and emission wavelengths set to 482/521 nm (bandwidths of 9/20 nm, respectively) in a top-reading mode.

Microscopy

Samples for Fig. 3e were analyzed using a Nikon Eclipse Ti2 microscope with a Hamamatsu Orca Flash 4.0 camera set to 400 ms exposure, 40x objective, 100% LED strength. Betaxanthin detection settings included Lumencor Spectra X cyan excitation and 525 ± 25 nm emission. Image analysis and deconvolution were performed using Huygens Professional image processing software. Fig. S12 and Supplementary Movie 1 were recorded using a Nikon Eclipse Ti2 microscope with an Andor Sona 4.2B-11 using a 20x objective, 100 ms exposure and 440 ± 10 nm transmission bandpass filter for the bright-field recording. Betaxanthin detection was done with a Lumencor Spectra using 35% LED strength with a 488 ± 3 nm excitation filter, 495 nm dichroic mirror, 520 ± 17.5 nm emission filters and 800 ms exposure. Image analysis was performed using NIS-Elements software. Time-lapse recording was done as a green fluorescence/bright-field overlay with a frame rate of 1 frame every 2 h for 72 h.

Flow cytometry

Cell populations were analyzed with a Becton Dickinson LSRII Fortessa flow cytometer, equipped for EGFP detection (488 nm laser, 505 nm long-pass filter, 530 ± 15 nm emission filter) and mCherry detection (561 nm laser, 610 ± 10 nm emission filter), and set to exclude dead cells, debris and cell doublets. The live cell population was previously determined while adjusting the device for HEK-293T cells and could be identified in the front versus side scatter view (FSC-A/SSC-A). Cell doublets were identified by using the front scatter area versus height view (FSC-A/FSC-H) and assuming a linear relationship.

SEAP assay

Production of human placental secreted alkaline phosphatase was quantified in cell culture supernatant as described before(27).

Resazurin assay

The culture medium was replaced with fresh clear medium containing 8 mg/L resazurin sodium salt (Sigma Aldrich, R7017). The cells were placed back in the incubator at 37 ˚C for

one hour, and then the fluorescence (571 nm ex./ 585 nm em.) was measured (the background was subtracted). A stock solution of 0.8 g/L was prepared in dH_2O.

Cell viability assays

Initially, HEK293T cells were transfected with a betaxanthin production cassette, or empty plasmid, or tGFP control. At 48 h after transfection the viability of the cells was determined by means of resazurin assay. In a second approach, conditioned medium with or without betaxanthins was produced by allowing cells transfected with the betaxanthin production cassette or empty vector to grow for 72 h. Subsequently, wild-type HEK293T cells were grown in these conditioned media for 24 h and cell viability was determined by means of resazurin assay. In a third approach, a resazurin-independent method was chosen to determine the growth rate of cells transfected with plasmids encoding either tGFP or the betaxanthin production cassette. The cell number was measured every 24 h for 72 h using a flow cytometer.

Chemical fixation of cells

Cells were trypsinized (0.05% trypsin–EDTA (Gibco)), washed with PBS, and resuspended at the concentration of approximately 200,000 cells per mL in 2% formaldehyde solution. The formaldehyde solution was prepared by dilution of 35% formaldehyde solution (Sigma) with PBS. After one hour the cells were washed again with PBS and subsequently used for analysis. All steps after trypsinization were performed on ice.

Cell lysis

Cells in a 24-well plate were harvested by lysing them in 150 μL RIPA buffer for 15 min at 37˚C. RIPA buffer contained sodium chloride (140 mM), Tris-Cl (10 mM), EDTA (1 mM), EGTA (0.5 mM), Triton X-100 (1%), sodium deoxycholate (0.1%), and SDS (0.1%) in ddH_2O, and the pH was adjusted to 7.4.

Statistical analysis, curve fitting and determination of slope

All statistical analysis was performed with GraphPad Prism 7. The curve-fitting for Fig. 3b was performed using GraphPad Prism 7's [Agonist] vs. response -- Variable slope function. In the case of Fig. S7a, Student's T-test with Welch's correction for non-equal SD was applied to each independent experimental dataset, with n = 8 independent biological samples. The exponential growth fit for Fig. S7b and Fig. S10b was calculated using GraphPad Prism 7's exponential growth equation function. For Fig. S10a, differentiation was performed using Prism's integrated differentiation function with a smoothing factor of 20 neighbors without polynomial fit.

Synthesis of indicaxanthin

Indicaxanthin was synthesized following a reported procedure(28) with some modifications. To a cold (0°C) solution of H_2O (20 mL), degassed with argon for 5 min, betanin (500 mg, red beet extract diluted with dextrin, TCI) and an aqueous ammonium solution (0.3 mL, 25%) were added sequentially. The color of the solution changed from red to violet and the reaction was stirred at 0°C for 30 min. L-Proline (67 mg, 0.58 mmol) was added followed by glacial acetic acid until the pH of the solution reached 5 (around 0.8 mL). The resulting red solution was stirred for 1 h at 0°C, then directly loaded onto a reverse phase column (ZEOprep® 90, C18) and eluted with H_2O. The fractions containing the desired compound were combined and lyophilized to obtain indicaxanthin (8.2 mg) as an orange solid. The identity of the compound was confirmed by its characteristic MS/MS pattern.

Analysis of indicaxanthin

The sample, control and synthetic indicaxanthin were each dissolved in aqueous MeCN solution (60%). These solutions were filtered and injected (1 µL) into a UHPLC system (Ultimate 3000, Thermo Fisher) equipped with a reverse-phase column (Kinetex® EVO C18; 50 x 2.1 mm, 1.7 µm, Phenomenex), a DAD (Ultimate 3000) and a mass detector (TSQ Quantum Ultra, Thermo Fischer). The eluent was composed of MeCN (0.1% formic acid) and H_2O (0.1% formic acid), the flow rate was 0.4 µL/min, and the temperature of the column oven was 40°C. The compound was detected in positive SRM mode targeting the typical fragmentations of indicaxanthin of 309 Da to 217 Da (23 eV) and 309 Da to 263 Da (17 eV).

Smartphone-based betaxanthin quantification

The field set-up for betaxanthin quantification consists of a cardboard box to shield the sample from light and a blue paper (Artoz 10769614-427, Coop, Switzerland) as a defined background. For betaxanthin quantification, we used an iPhone running the Color Name AR Pro mobile application software. A calibration curve was obtained from ratios of the intensity (I) of standard solutions of the model compound indicaxanthin (2.24, 1.12, 0.56, 0.28, 0.14, 0.07 mg/mL) to that of water as a blank (I_0). The absorbance (A) was obtained according to Beer's law $A = -\log\left(\frac{I}{I_0}\right)$ with I equals the blue channel displayed on the smartphone application software. The slope (0.36 ± 0.01) and R^2 value (0.99) of the calibration curve were calculated using Prism software. The culture supernatants of cells producing betaxanthin (500 µL) and negative-control cells (500 µL) were diluted with H_2O (500 µL), the values of betaxanthin absorbance were recorded, and corresponding concentrations were obtained from the calibration curve.

RNA extraction and RT-qPCR

Cells were transfected in a 24-well plate format and harvested 48 h later using a Quick-RNA miniprep kit (Zymo Research, cat. No. R1054), according to the manufacturer's protocol. 400 ng of total RNA was used for cDNA synthesis with a High-Capacity cDNA Reverse Transcription Kit (Invitrogen, cat. No. 4368814), according to the manufacturer's instructions. Thereafter, the 20 µL cDNA reaction was diluted with 1100 µL water. Subsequently, 1080 µL of this cDNA mix was added to 1080 µL of KAPA 2x Taqman master mix (Sigma-Aldrich, cat. No. KK4703). Finally, 20 µL of cDNA-Master mix was added to each well of a 96-well Taqman array plate (Thermo Fisher, cat No. 4414130). The Eppendorf Realplex Mastercycler (Eppendorf GmbH) was used according to the Taqman array plate protocol. The relative threshold cycle (Ct) was normalized to GAPDH and ACTB genes and in a later step the normalized dataset from the active samples was normalized to the mock-transfected control (∆∆ct method). tGFP-producing cells were used as a benchmarking cell line in this setup, as wild-type cells would have resulted in a bias, due to the lower stress of being untransfected and not overproducing a foreign protein.

SUPPLEMENTARY INFORMATION

Table S1: Plasmids used in this work.

Plasmid	Description and Cloning Strategy (oligo sequences are listed in Table SII, sequences of AmDODA and CcTyr can be found below under DNA sequences 1-2; Key plasmid maps can be found in Fig. S9.)	Source
pBAD-mTagBFP2	Constitutive bacterial expression vector for mTagBFP2.	Addgene (34632) (47)
pCOLA Duet-1	Minimal vector lacking a mammalian promoter.	Merck (71406)
pDA171	Mammalian tGFP expression vector.	Ausländer et al.(48)
pDA251	Constitutive mammalian expression vector for GCH1 (P_{SV40}-GCH1-pA).	Ausländer et al.(49)
pDA326	Doxycycline inducible mammalian expression vector for Citrin-P2A-SEAP fusion (P_{teto7}-Citrin-P2A-SEAP-pA).	Ausländer et al.(49)
pDA701	Constitutive mammalian expression vector for Citrine-P2A-SEAP (P_{SV40}-Citrine-2A-SEAP-pA).	Ausländer et al.(49)
pDB114	Cloning template for P_{PGK} promoter. ODB166 and ODB167 were phosphorylated, annealed, digested (PsiI/SalI) and ligated into pSUPER retro puro (PsiI/SalI).	This work
pFS20	Constitutive mammalian expression vector for mCherry (P_{hCMV}-mCherry-pA).	Ausländer et al.(50)
pFOX13	Constitutive mammalian expression vector for tGFP (P_{hCMV}-tGFP-pA). tGFP was amplified from pDA171 with OFOX010 and OFOX011, digested (SpeI/BamHI) and ligated into pMM1 (SpeI/BamHI)	This work
pFOX41	Cloning vector for $P_{hCMVmin}$ promoter ($P_{hCMVmin}$-Citrin-pA). $P_{hCMVmin}$-Citrin was amplified from pMM545 with OFOX012 and OFOX033, digested (XhoI/BamHI) and ligated into pMM1 (XhoI/BamHI).	This work
pFOX76	Constitutive mammalian expression vector for SEAP (P_{hCMV}-SEAP-pA). SEAP was amplified from pDA701 with OFOX024 and OFOX028, digested (SpeI/BamHI) and ligated into (SpeI/BamHI).	This work
pMM1	Mammalian expression vector with a modified MCS (P_{hCMV}-MCS-pA; MCS, EcoRI-ATG-SpeI-NheI-BamHI-STOP-XbaI-HindIII-FseI-pA).	Müller et al.(51)

pMM328	Constitutive mammalian expression vector for SEAP (P_{PGK}-SEAP-pA). P_{PGK} was amplified from pDB114 with OMM284 and OMM285, digested (XhoI/EcoRI) and ligated into pFOX76 (XhoI/EcoRI). The whole plasmid was then amplified with OMM286 and OMM287 and religated.	This work
pMM545	Constitutive mammalian expression vector for Citrin (P_{hCMV}-Citrin-pA).	Müller et al.(51)
pMM585	Cloning template for SS (P_{hCMV}-SS-pA).	Scheller et al.(52)
pMM591	Constitutive mammalian expression vector for rtTA (P_{hCMV}-rtTA-pA). rtTA was amplified from pTetON-3G with OMM249 and OMM251, digested (EcoRI/XbaI) and ligated into pMM1 (EcoRI/XbaI).	This work
pPST319	Constitutive mammalian expression vector for human TH (P_{hCMV}-TH-pA). TH was amplified from HsCD00630964 (dnasu.org) with OPST521 and OPST522, digested (SpeI/BamHI) and ligated into pMM1 (SpeI/BamHI).	This work
pPST320	Constitutive mammalian expression vector for DODA from *A. muscaria* (P_{hCMV}-DODA-pA). Codon optimized DODA was amplified from a synthetic construct (Genscript, DNA Sequence S1) with OPST519 and OPST520 digested (SpeI/BamHI) and ligated into pMM1 (SpeI/BamHI).	This work
pPST321	Constitutive mammalian expression vector for human GCH1 (P_{hCMV}-GCH1-pA). GCH1 was amplified from pDA251 with OPST527 and OPST528, again amplified with OPST523 and OPST524, digested (SpeI/BamHI) and ligated into pMM1 (SpeI/BamHI).	This work
pPST322	Constitutive mammalian expression vector for DODA-GCH1 fusion protein (P_{hCMV}-DODA-GCH1-pA). GCH1 was excised from pPST321 (SpeI/BamHI) and ligated into pPST320 (NheI/BamHI).	This work
pPST323	Doxycycline inducible mammalian expression vector for human TH (P_{tetO7}-TH-pA). TH was excised from pPST319 (SpeI/BamHI) and ligated into pTS1017 (NheI/BamHI).	This work
pPST324	Constitutive mammalian expression vector for P2A based DODA-TH-GCH1 fusion protein (P_{hCMV}-DODA-P2A-TH-P2A-GCH1-pA). A DODA-P2A intermediate was created by excising P2A from pTS1018 (SpeI/BamHI) and ligating into pPST320 (NheI/BamHI). A TH-P2A intermediate was created by excising P2A from pTS1018 (SpeI/BamHI) and ligating into pPST319 (NheI/BamHI). A TH-P2A-GCH1 intermediate was created by excising GCH1 from pPST321 (SpeI/BamHI) and ligating into the TH-P2A fusion	This work

	(NheI/BamHI). TH-P2A-GCH1 was excised (SpeI/BamHI) and ligated into the DODA-P2A intermedia (NheI/BamHI).	
pPST325	Constitutive mammalian expression vector for secretion engineered, codon optimized TYR from *C. cristata* (P_{hCMV}-SS-TYR-pA). TYR was amplified from a synthetic construct (Genscript, DNA sequence S2) with OPST531 and OPST532, digested (SpeI/BamHI) and ligated into pMM585 (NheI/BamHI).	This work
pPST326	Constitutive mammalian expression vector for secretion engineered DODA from *A. muscaria* (P_{hCMV}-SS-DODA-pA). DODA was excised from pPST320 (SpeI/BamHI) and ligated into pMM585 (NheI/BamHI).	This work
pPST350	Doxycycline inducible mammalian expression vector for P2A based SEAP and TH fusion proteins (P_{tetO7}-SEAP-P2A-TH-pA). P2A was excised from pTS1018 (SpeI/BamHI) and ligated into pTS1015 (NheI/BamHI). The SEAP-P2A was excised from this (EcoRI/NheI) and ligated into pPST323 (EcoRI/SpeI).	This work
pSEAP2-Control	Constitutive mammalian SEAP expression vector (PSV40-SEAP-pA).	Clontech
pSUPER retro puro	Cloning vector.	Addgene (30519)
pTetON-3G	Constitutive mammalian expression vector for rtTA (P_{hCMV}-rtTA-pA).	Clontech
pTS1015	Mammalian minimal expression vector for SEAP ($P_{hCMVmin}$-SEAP-pA). pFOX76 was digested (SpeI/BamHI) and ligated into pFOX41 (SpeI/BamHI).	This work
pTS1017	Doxycycline inducible mammalian expression vector for SEAP (P_{tetO7}-SEAP-pA). pDA326 was amplified with OTS733 and OTS734, digested (MluI/XhoI) and ligated into pTS1015 (MulI/XhoI).	This work
pTS1018	Cloning template for P2A (P_{hCMV}-P2A-pA). OTS490 and OTS491 were phosphorylated, annealed and ligated into pMM1 (SpeI/BamHI).	This work
pTS1103	Constitutive mammalian expression vector for mTagBFP2 (P_{hCMV}-mTagBFP2-pA). mTagBFP2 was amplified from pBAD-mTagBFP2 with OTS381 and OTS382, digested (SpeI/BamHI) and ligated into pMM1 (SpeI/BamHI).	This work
pTS1105	Constitutive mammalian expression vector for rtTA (P_{PGK}-rtTA-pA). pMM591 was digested (SpeI/BamHI) and ligated into pMM328 (SpeI/BamHI).	This work

Abbreviations and additional information: DODA, 4,5-DOPA dioxygenase; GCH1, GTP cyclohydrolase 1; MCS, multiple cloning site; P2A, porcine teschovirus-1 self-cleaving 2A sequence; pA, poly A termination signal; P_{hCMV}, human cytomegalovirus immediate-early promoter; $P_{hCMVmin}$, minimal human cytomegalovirus immediate-early promoter requires transactivator binding for efficient transcription; P_{PGK}, phosphoglycerate kinase promoter; P_{SV40}, simian virus 40 promoter; P_{tetO2}, doxycycline-inducible promoter with two tetO operator sequences followed by a $P_{hCMVmin}$; SS, secretion signal peptide sequence; rtTA, reverse tetracycline-dependent transactivator (rTetR-VP16); tetR, *E. coli* transposon Tn10-derived tetracycline-dependent repressor; TH, tyrosine hydroxylase; TYR, tyrosinase; VP16, *Herpes simplex* virus-derived transactivation domain.

Table S2: Oligos used to construct plasmids used in this work.

ODB166	TCGttataaGTTCTGTATGAGACCACAGATCCCCaggcctctctctcCGTGTTCACAGCGGACCTTGATttaaatgtccatacaatTAAGGCAC
ODB167	TACgtcgacGGTATCGATAAGCTTAAGCTTTTCCAAAAAcagccccattcttGGCATTCACCGCGTGCCTTAattgtatggacatttaaATCAAG
OFOX012	CAACAACTCGAGGGTAGGCGTGTACGGTGG
OFOX024	GAAGCGGAATTCGCCACCATGACTAGTCTGCTGCTGCTGCTG
OFOX028	CGGTGGATCCGCTAGCGGTCTGCTCGAATCTGCC
OFOX033	TTTTTTTTTTTTTTTTTTTTctggcaactagaaggcacag
OMM249	AAGCTTTCTAGAcACCGGTGGATCCGCTAGCcccggggagcatgtcaag
OMM251	gcGGAATTCACCATGACTAGTGGATCAAGACTGGACAAGAG
OMM284	ggagatctccACGCGTGGTACCCTCGAGctaccgggtaggggaggcgc
OMM285	cgcgaattcggtctccctataccgagctcgggctggaggtcgaaaggcccgg
OMM286	ctcagtagtctcgtgcag
OMM287	acgtgctacttccatttg
OPST519	cctgtGAATTCACCATGACTAGTggtggttctggtGTACCTTCATTCGTCGTCTATAGTTC
OPST520	cctgtGGATCcgctagcAGCGTCCCGGTGCG
OPST521	cctgtGAATTCACCATGACTAGTggtggttctggtCCCACCCCCGACGC
OPST522	cctgtGGATCcgctagcGCCAATGGCACTCAGC
OPST523	tgtGAATTCACCATGACTAGTggtggttctggtGAGAAGGGCCCTGTGCGGGCACCGGCGGAGAAGCCGCGGGGGCGCCAGGTGCAGCAATGGGTTCCCCGAGCGCGATCCGCCGCGGCC
OPST524	cctgtGGATCcgctagcCAAGCTCCTAATGAGAGTCAGGAACTCTTCCCGAGTCTTTGGGTCCTCCCGGAACACACCCA
OPST527	GGTGCAGCAATGGGTTCCCCGAGCGCGATCCGCCGCGGCCC

OPST528	TCAGGAACTCTTCCCGAGTCTTTGGGTCCTCCCGGAACACACCCAAC
OPST531	cctgtGGATCcgctagcGTATCTTGAGACGGGGATTATCCTG
OPST532	cctgtGAATTCACCATGACTAGTggtggttctggtATGGATAACGCAACTCTGGCTATG
OTS381	ctgaACTAGTggtggttctggtGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATGCACATGAAG CTGT
OTS382	ctgaggatccgctagcATTAAGtTTGTGCCCCAGTTTGCTAGGGAG
OTS490	CTAGTggtggttctggtGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGG AGGAGAACCCTGGACCTgctagcg
OTS491	gatccgctagcAGGTCCAGGGTTCTCCTCCACGTCTCCAGCCTGCTTCAGCAGGCTGAAGTTAGTA GCTCCGCTTCCaccagaaccaccA
OTS733	CTGAACGCGTCCGTACACGCCTAAAGCATATACGTTC
OTS734	CCTCGACATACTCGAGTTTACTCCCTATC

Table S3: Amounts of plasmids transfected in this work.

All amounts are plasmid per well of a 96-well plate (or per well of the specified format).

Fig. 1b	150 ng pCOLADuet-1 or 150ng pPST320.
Fig. 1c	From left to right: 150 ng pCOLADuet-1; 75 ng pCOLADuet-1, 25 ng pPST321, 50 ng pPST319; 25 ng pCOLADuet-1, 75 ng pPST320, 50 ng pPST319; 50 ng pCOLADuet-1, 75 ng pPST320, 25 ng pPST321; 75 ng pPST320, 25 ng pPST321, 50 ng pPST319; 100 ng pPST322, 50 ng pPST319; 150 ng pPST324.
Fig. 1d	150 ng pPST324.
Fig. 2a	600 ng pPST320 in a 24-well plate.
Fig. 3b	80 ng pTS1105, 80 ng pPST350, 2240 ng pPST322 in a 6-well plate. Reseeded cells into a 96-well plate with $3*10^6$ cells per plate.
Fig. 3c	From top down: 150 ng pPST324; 75 ng pPST324, 75 ng pCOLADuet-1; 25 ng pPST324, 125 ng pCOLADuet-1; 150 ng pCOLADuet-1.
Fig. 3d	From left to right: 150 ng pCOLADuet-1; 50 ng pPST324, 100 ng pCOLADuet-1; 125 ng pPST324, 25 ng pCOLADuet-1.
Fig. 3e	For a: WT cells; For b: 150 ng pPST324.
Fig. 4c	14.4 µg pPST324; 14.4 µg pCOLADuet-1 for mock in a 10 cm dish.
Fig. S1	150 ng pSEAP2-Control
Fig. S2	600 ng pPST324; 600 ng pCOLADuet-1 for mock in a 24-well plate.

Fig. S3	From left to right: 600 ng pColaDuet-1; 600 ng pPST319; 600 ng pPST321; 300 ng pPST319, 300 ng pPST321 in a 24-well plate.
Fig. S4	From left to right: 150 ng pCOLADuet-1; 75 ng pCOLADuet-1, 75 ng pPST326; 75 ng pCOLADuet-1, 75 ng pPST325; 75 ng pPST325, 75 ng pPST326; 150 ng pPST324.
Fig. S5	14.4 µg pPST324; 14.4 µg pCOLADuet-1 for mock in a 10 cm dish.
Fig. S6	2400 ng pPST324; 2400 ng pCOLADuet-1 for mock in a 6-well plate.
Fig. S7a	For mock and negative control: 600 ng pCOLADuet-1; for tGFP control: 600 ng pFOX13; for pPST324: 600 ng pPST324 in a 24-well plate.
Fig. S7b	W/o betaxanthin: 2400 ng pCOLADuet-1; with betaxanthin 240 ng pPST324 in a 6-well plate. 2 wells transfected and medium pooled each.
Fig. S7c	For GFP producing: 600 ng pFOX13; for betaxanthin producing: 600 ng pPST324 in a 24-well plate.
Fig. S8	2400 ng pPST324; 2400 ng pFOX13 for tGFP control in a 6-well plate.
Fig. S11	600 ng pPST324; 600 ng pCOLADuet-1 for mock in a 24-well plate.
Fig. S12	600 ng pPST324; 600 ng pCOLADuet-1 for mock in a 24-well plate.
Fig. S13	From left to right: 120 ng pFS20, 480 ng pCOLADuet-1; 480 ng pPST324, 120 ng pCOLADuet-1; 480 ng pPST324, 120 ng pFS20 in a 24-well plate.
Fig. S14	From left to right: 2400 ng pCOLADuet-1; 2400 ng pPST324; 2400 ng pFOX13 in a 6-well dish.
Fig. S15	14.4 µg pPST324; 14.4 µg pCOLADuet-1 for mock in a 10 cm dish.

Table S4: Raw data for Fig 1d.

	Mock TF			pPST324		
hMSC-tert	7349	7438	7497	14386	15358	13296
HEK293T	7450	7612	7474	41287	41669	40861

DNA sequence S1: Codon-optimized DODA from *A. muscaria*

ATGGTACCTTCATTCGTCGTCTATAGTTCCTGGGTTAACGGACGACAACGCTATATACGCCAGGCGTTTGC
AAGTATCCTTTTCTACATCATACGCGATACGACGTTGAGTTTTCCTTCTCACACTACAATGTCAACAAAACCTGAGAC
GGATCTGCAAACTGTTTTGGACTCAGAAATCAAGGAGTGGCATTTCCATATTTATTTCCACCAAAATAACGCCGCAG
AACACCAAGCGGCTCTCGAACTTCGGGACGCGGTTCTCAGGCTCAGGCAGGACGGAGCCTTTGTGGCAGTACCGCT
TTTCCGAGTAAATATGGACCCGATGGGTCCTCATCCAGTGGGGTCATACGAGATATGGGTGCCCTCCGAAACGTTT
GCGAGCGTGTTTTCATATCTTTGCATGAACCGCGGTCGGCTCTCCATACTCGTCCACCCATTGACGCGCGAGGAACT
GCGCGACCATGAAATACGAAACGCTTGGATTGGGCCCTCCTTTCCATTGAATCTCGCCAACTTGCCAATCAAGTCAG
ATGAAATCCCCCTGCAGTATCCCTCACTGAAGCTGGGTTACTCAAGTACCGCTCATAAGATGAGCTTGGAGGAAAG
AAGGAAGCTCGGAGACGACATAGAGGCCGTGCTTAGGGGTGAGAAAGAGGCAGCAAGGGCACCGCACCGGGAC
GCTTGA

DNA sequence S2: Codon-optimized TYR from *C. cristata*

ATGGATAACGCAACTCTGGCTATGCTCCTTGCGATTTGGTTCATATCCTTTCATTTTATTAAAATGCTTTTCA
CGAACCAGTCTACCAAGCTCCTTCCGCCTGGCCCGAAACCTCTCCCAATAATCGGGAACATTCTCGAAGTCGGTAAA
AAGCCGCACCGAAGCTTTGCAAACCTTGCTAAAATACACGGACCTCTGATATCCCTCAAATTGGGATCTGTGACTAC
AATAGTTGTATCCTCAGCGGAAGTGGCAAAGGAAATGTTCCTGAAGAAAGATCAGCCGCTGAGTAACCGCACAGTT
CCTAACAGCGTCACGGCTGGCGACCATCACAAACTGACAATGAGCTGGCTCCCAGTATCCCCGAAATGGCGCAATT
TTCGGAAGATTACAGCAGTACACCTTCTCTCACCATTGCGCCTGGATGCGTGTCAGTCCCTCAGACATGCGAAAGTT
CAACAACTTTTTCAGTATGTCCAGGAGTGTGCGCAAAAGGGACAGGCGGTTGACATCGGTAAAGCGGCGTTTACCA
CCTCCTTGAATCTTCTTTCTAAACTTTTTTTCTCTAAGGAACTGGCCTCCCATAAGTCTAGAGAGTCTCAGGAATTCAA
ACAACTTATATGGAACATCATGGAGGACATAGGTAAGCCGAACTATGCTGATTATTTCCCAATCCTGGGGTGCGTC
GATCCTAGTGGGATAAGGCGACGGCTCGCGAGCAATTTCGACAAGCTCATTGAAGTCTTCCAGTGCATCATCAGAC
AAAGGCTCGAACGGAACCCCTCAACACCACCGACGAATGATGTCCTTGATGTTTTGTTGGAGTTGTATAAACAGAA
CGAACTGAGTATGGGAGAGATAAATCATCTGTTGGTTGATATTTTTGACGCAGGAACAGATACGACCAGTTCTACA
TTTGAATGGGTTATGGCTGAGCTCATCCGCAATCCAGAAATGATGGCCAAGGCACAGGACGAGATAGAACAAGTG
CTTGGAAAGGATCGGCAGATCCAGGAGTCAGATATCATAAAGTTGCCCTACTTGCAAGCGATTATCAAGGAAACAC
TTCGGCTCCACCCTCCGACTGTATTCTTGCTTCCCCGGAAAGCTGATACAGACGTCGAGCTTTACGGCTATATCGTAC
CCAAGGATGCACAAATCTTGGTCAATCTGTGGGCCATCGGTAGAGACTCTCAAGCGTGGGAAAACCCAAAGGTCTT
TAGTCCAGACCGGTTCTTGGGCTGCGAAATTGACGTAAAAGGTAGGGACTTTGGCCTTCTGCCCTTTGGTGCAGGT
AAGAGAATCTGCCCGGGAATGAACCTTGCCATCAGAATGCTTACCCTGATGCTGGCAACACTCCTCCAATTTTTTAA
CTGGAAGCTCCAAGACGGGATGAGTCTTGAAGACCTCGATATGGAAGAAAAATTTGGCATCGCACTCCAAAAAACA
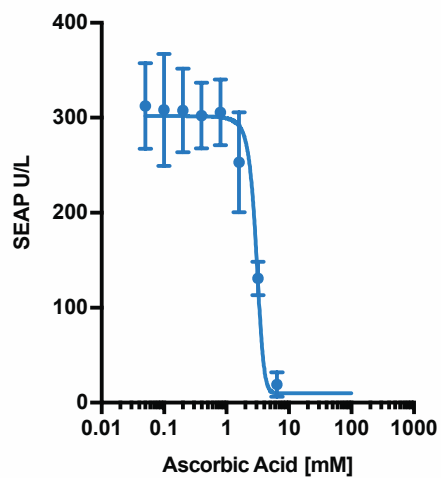AAGCCTCTCAGGATAATCCCCGTCTCAAGATACTAA

SUPPLEMENTARY FIGURES

Figure S1: Ascorbic acid toxicity in HEK293T cells. Cells were transfected with pSEAP2 control and 7.5 h later the medium was exchanged for fresh medium containing the indicated amounts of ascorbic acid. 48 h later SEAP reporter activity was evaluated as a measure of cell survival, and thereby ascorbic acid toxicity. Graph shows mean ± s.d. of n = 3 independent samples.
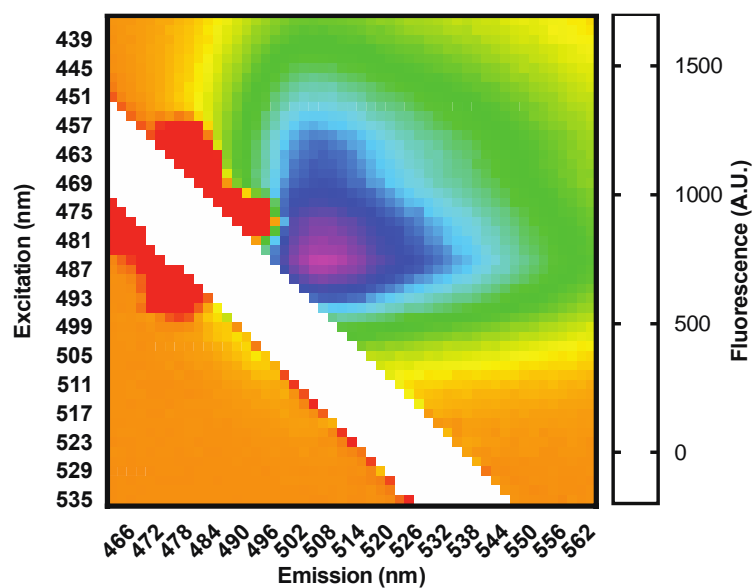
Figure S2: 2D fluorescence scan of cell supernatant containing betaxanthin. HEK293T cells were transfected with pPST320 (P$_{hCMV}$-AmDODA-pA). After 48 h the medium was changed to clear medium containing 1 mM L-DOPA and 0.05 mM ascorbic acid. After 16 h a 2D fluorescence scan was recorded and the background (supernatant from mock-transfected cells) was subtracted. The figure shows representative results from three independent experiments.
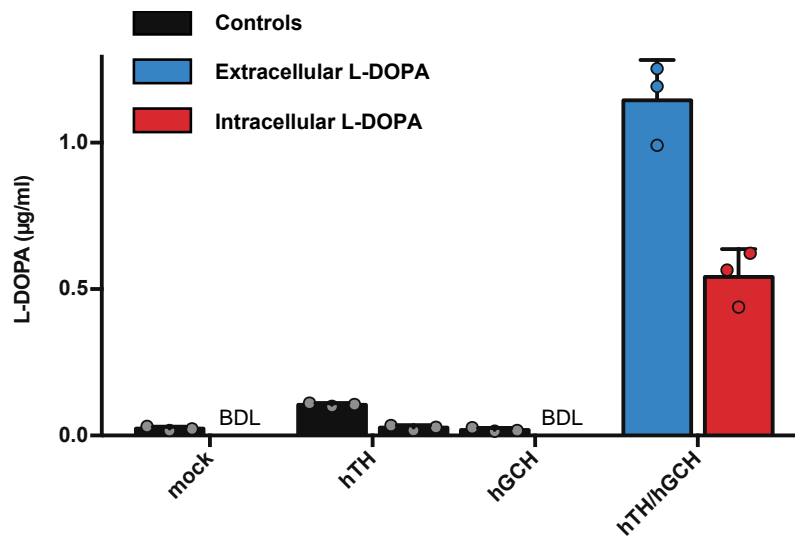
Figure S3: L-DOPA quantification of the human tyrosine hydroxylase system. HEK293T cells were transfected with plasmids encoding the indicated constructs. L-DOPA production was quantified by HPLC-MS with a standard curve between 0.1 and 50 µg/ml. Extracellular L-DOPA was quantified directly in the supernatant, while intracellular L-DOPA was extracted in a total volume of 2/5ths of the supernatant volume. From left to right: mock (pColaDuet-1); hTH (pPST319); hGCH (pPST321); hTH/hGCH (pPST319+pPST321). Results are the mean ± s.d. of n = 3 biologically independent samples and are representative of three independent experiments.
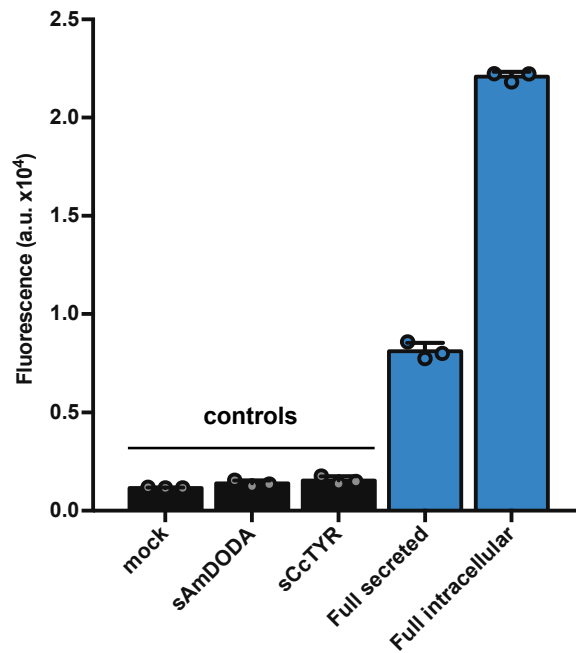
Figure S4: Functionality test of the secretion-engineered complete betaxanthin production system. HEK293T cells were transfected with plasmids encoding a secretion-engineered version of AmDODA and a secretion-engineered version of a tyrosinase from *C. cristata* (sCcTYR). Color production by the secreted system can be seen, albeit less than with the regular intracellular system. ■, negative controls; ■, complete production cascades. From left to right: mock (pCOLADuet-1); sAmDODA (pPST326); sCcTYR (pPST325); Full secreted (pPST325, pPST326); Full intracellular (pPST324). Color development was measured 48 h after transfection. Results are the mean ± s.d. of n = 3 biologically independent samples and are representative of three independent experiments.
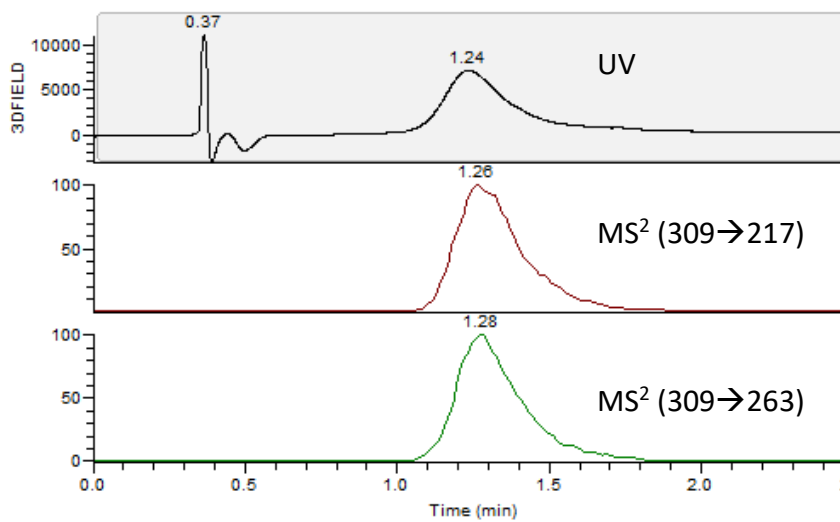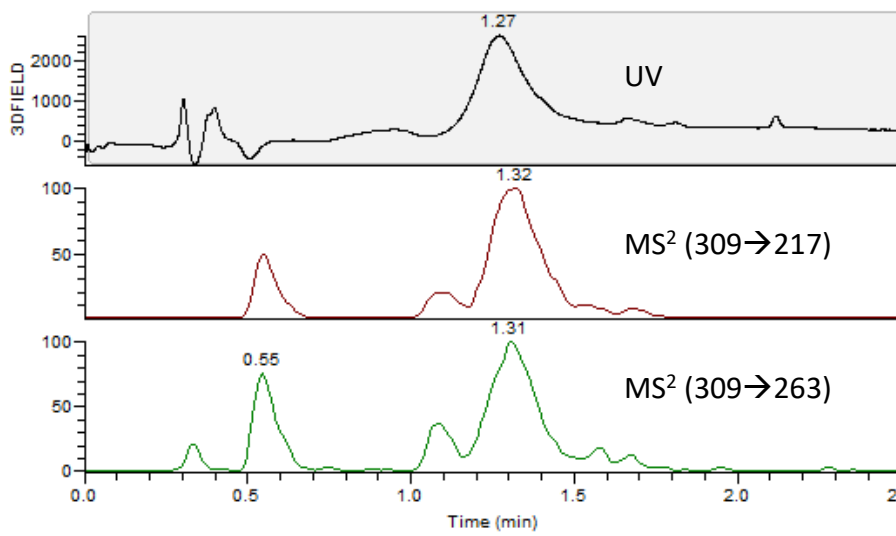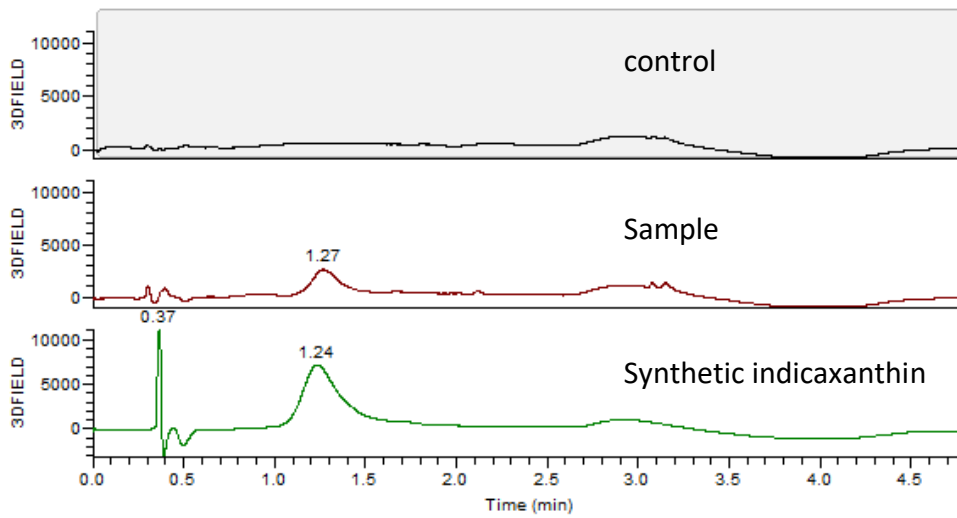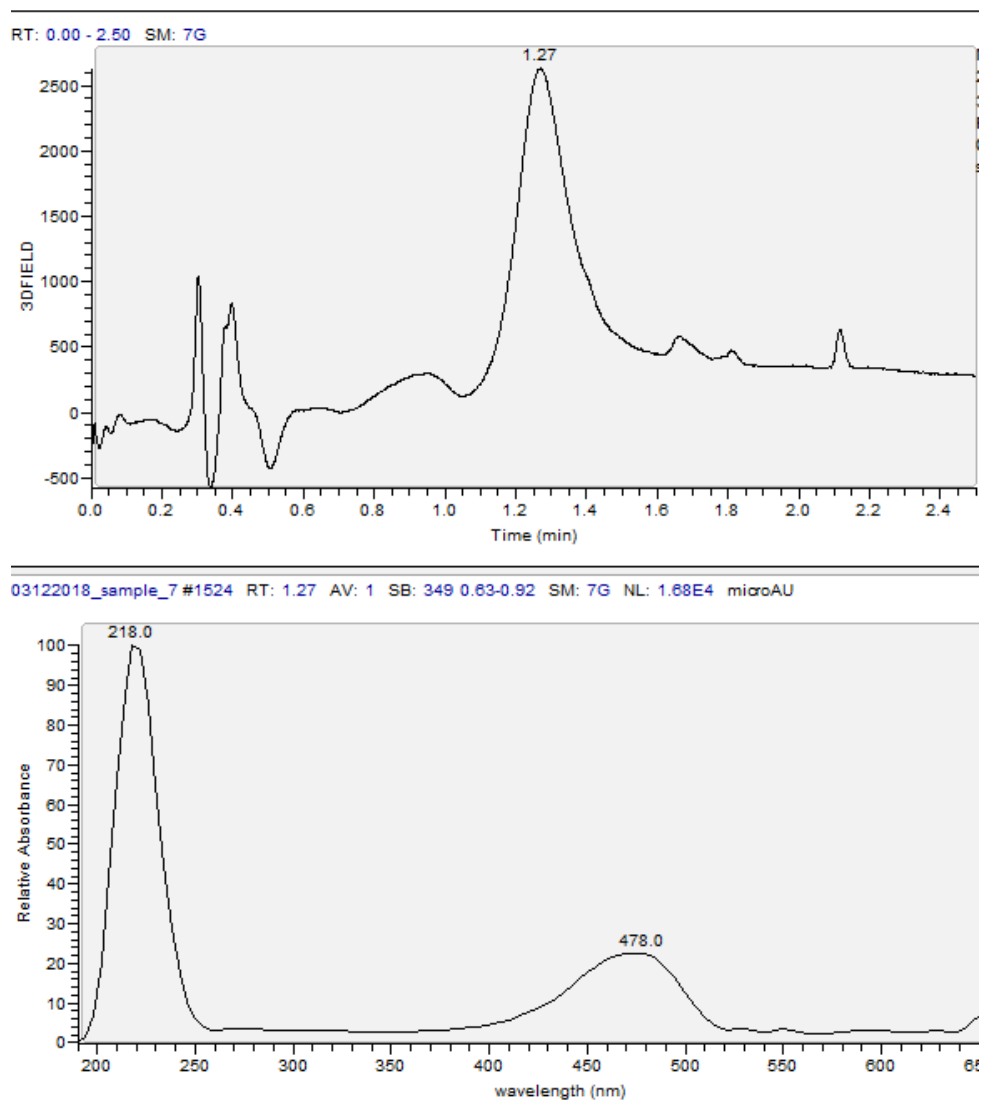
Figure S5: HPLC-UV/Vis-MS/MS analysis of the dye produced by HEK293T cells. Cells were transfected with either an empty mock plasmid (pColaDuet-1) or the full betaxanthin production cascade (pPST324). (a) Comparison of HPLC chromatograms extracted at the wavelength range from 400 to 500 nm. From top to bottom: Mock transfection control supernatant; full betaxanthin production system supernatant; synthetic indicaxanthin control. (b) HPLC and MS chromatograms of supernatant transfected with the full betaxanthin production system. From top to bottom: UV trace extracted at the wavelength range from 400 to 500 nm; MS chromatogram of the specific MS2 transition 309 Da → 217 Da; MS chromatogram of the specific MS2 transition 309 Da → 263 Da. (c) HPLC and MS chromatograms of synthetic indicaxanthin. From top to bottom, same as b (d) Top: HPLC chromatogram extracted at the wavelength range from 400 to 500 nm. Bottom: UV/Vis analysis of the compound at 1.27 min. HEK293T cells were transfected with pPST324 or pColaDuet-1, and the medium was changed 7 h later to medium containing 0.1 mM ascorbic acid and 5 g/L L-proline. The supernatant was harvested 72 h later, flushed with nitrogen, and stored below -20°C until analysis. The results are representative of three independent experiments.
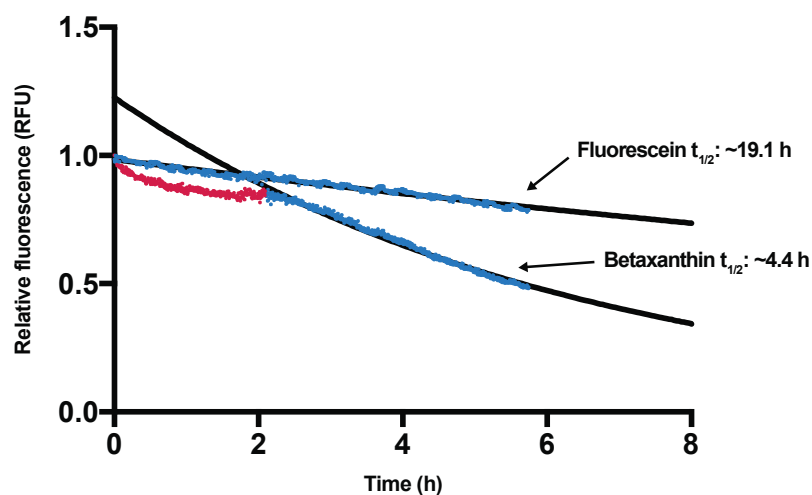
Figure S6: Photobleaching of supernatant containing betaxanthins compared to fluorescein. Cells were transfected with pColaDuet-1 or pPST324 and 7.5 h later, the medium exchanged for fresh clear medium. After 72 h the supernatant was harvested and the photobleaching experiment in a Tecan reader was started. Measurements were taken every 20 s for approximately 6 h, using a measuring cycle of 50 flashes each at 485 nm and 490 nm. The fluorescence was measured on a Tecan M1000 at 490/525 nm for the fluorescein curve and 485/507 nm for the betaxanthin curve with an optimized gain for each compound. The fluorescein sample was prepared by diluting a 4 g/L stock solution in the supernatant of cells transfected with pColaDuet-1 (which also counted as background) with a factor of 1/1000. The created data was background-subtracted, and normalized to the maximum fluorescence intensity, and a simple exponential decay function was fitted to part of the data, assuming first-order kinetics(53). The data used for fitting was marked with blue dots, and the data not used with red dots. The data used for the betaxanthin fit excluded the initial increase in fluorescence. The half-life was calculated from these fitted curves. Possible reasons for the unexpected initial increase in betaxanthin fluorescence are pH and oxygen equilibrations in the plate reader, shifting the fluorescence properties of the betaxanthins, or continued betaxanthin formation in the supernatant from precursors.
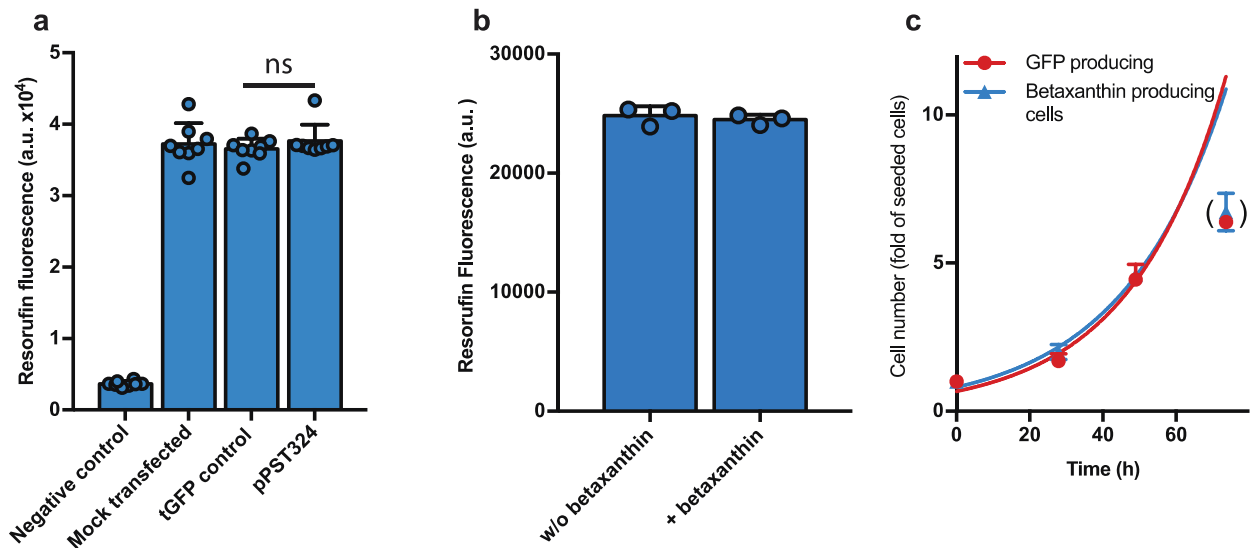
Figure S7: Cell viability assays of cells constitutively expressing the betaxanthin production cassette. (a) Resazurin cell viability assay of cells transfected with plasmids encoding either tGFP or no functional protein as controls or the constitutive betaxanthin production pathway. Negative controls were created by adding 10 μg/mL puromycin to mock-transfected cells after transfection. At 48 h after transfection, resazurin cell viability assay was performed. (b) Resazurin cell viability assay of wild-type HEK293T cells grown for 24 h in conditioned medium either containing betaxanthin or not. Conditioned medium was prepared by growing for 72 h cells transfected with pPST324 or pColaDuet-1, and diluted 2-fold with fresh culture medium. (c) Growth curve of cells transfected with plasmids encoding either tGFP or betaxanthin production pathway. Cell number was determined using a flow cytometer at the indicated times. The two curves show the exponential fit of the data with growth rates k = 0.038 ± 0.005 * h^{-1} for the GFP curve and k = 0.035 ± 0.002 * h^{-1} for the betaxanthin curve. The 72 h data points were not used for this calculation because growth was expected to have reached the stationary phase. Figure a shows the mean ± confidence interval of n = 8 independent samples, figures b and c show the mean ± confidence interval of n = 3 independent samples.
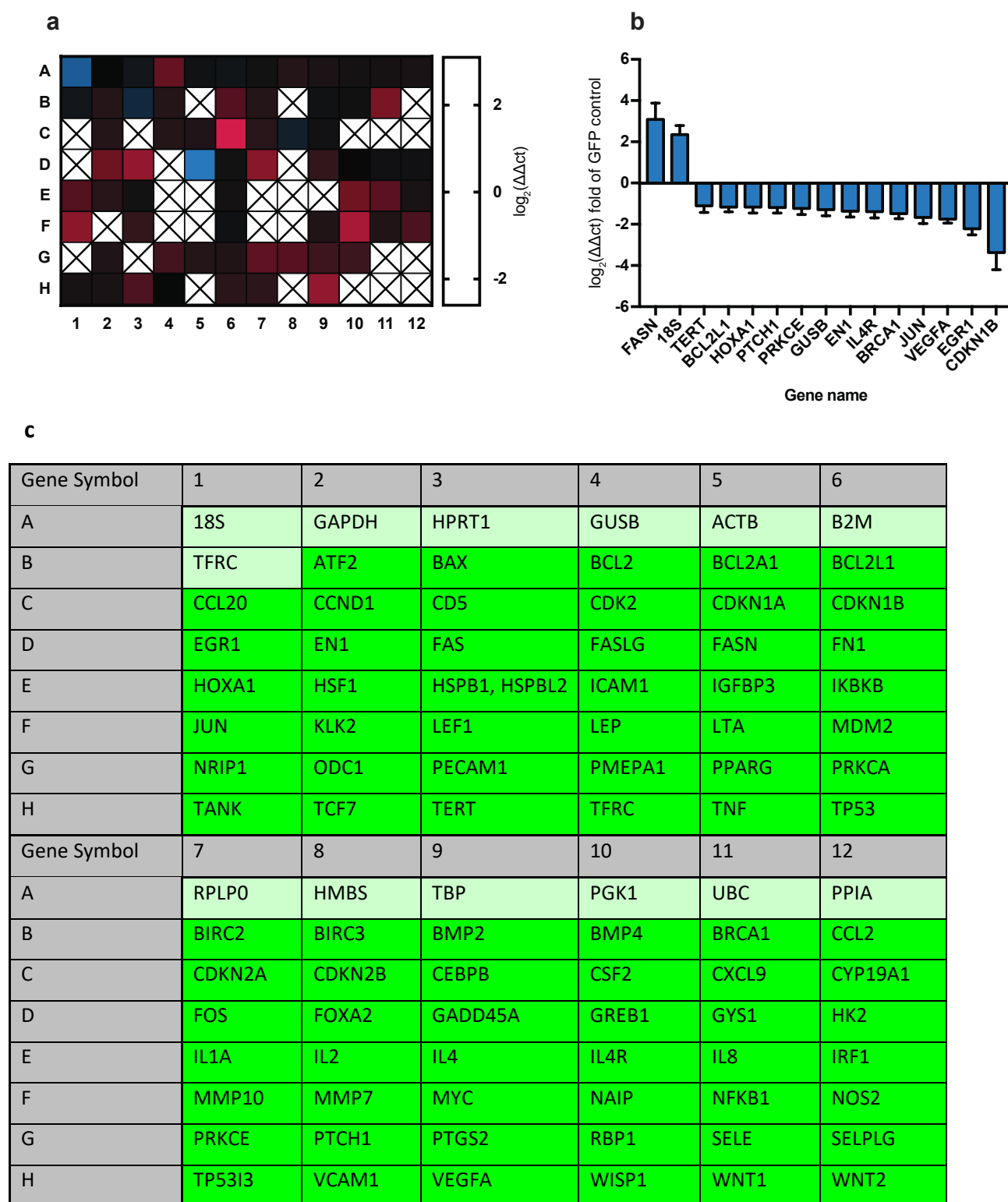
Figure S8: qPCR analysis of key metabolic and signaling pathways. (a) Heatmap showing $\log_2$ relative mRNA content of cells producing betaxanthin with respect to the tGFP-producing cells. Blue, higher gene expression in betaxanthin cells; red, lower gene expression; crossed boxes, gene expression below detection limit. (b) Graph showing genes that are significantly ($p<0.05$) over/under-expressed in betaxanthin-producing cells compared to GFP-producing cells as a control. Positive values mean the gene

is overexpressed in betaxanthin-producing cells. (c) Table showing the genes examined in this assay. Light green, housekeeping genes; dark green, genes important for signal transduction pathways. Graphs a and b show the mean (± s.d. for b) of n = 3 independent samples normalized to the two housekeeping genes *GAPDH* and *beta-actin*.
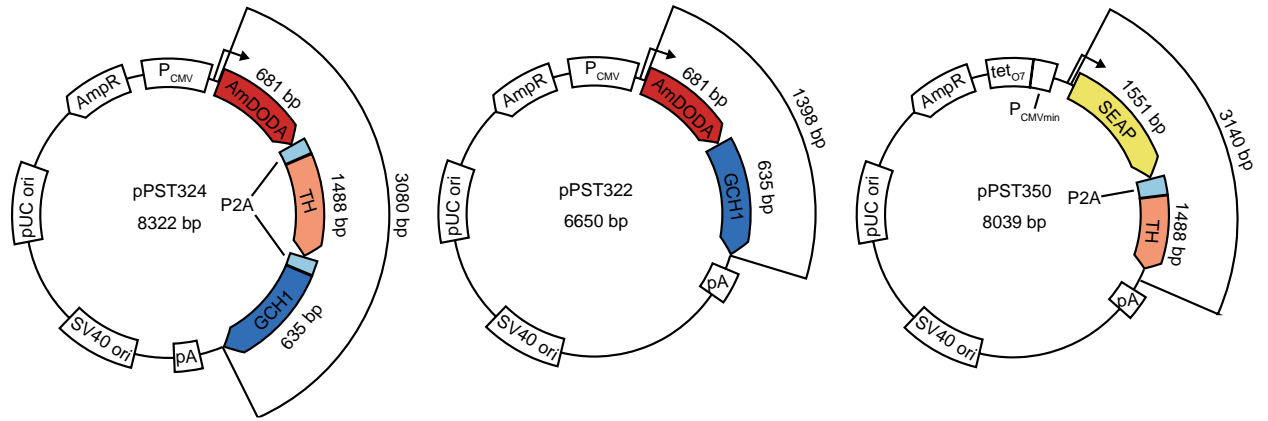
Figure S9: Plasmid maps of key constructs. For abbreviations see Table S1.
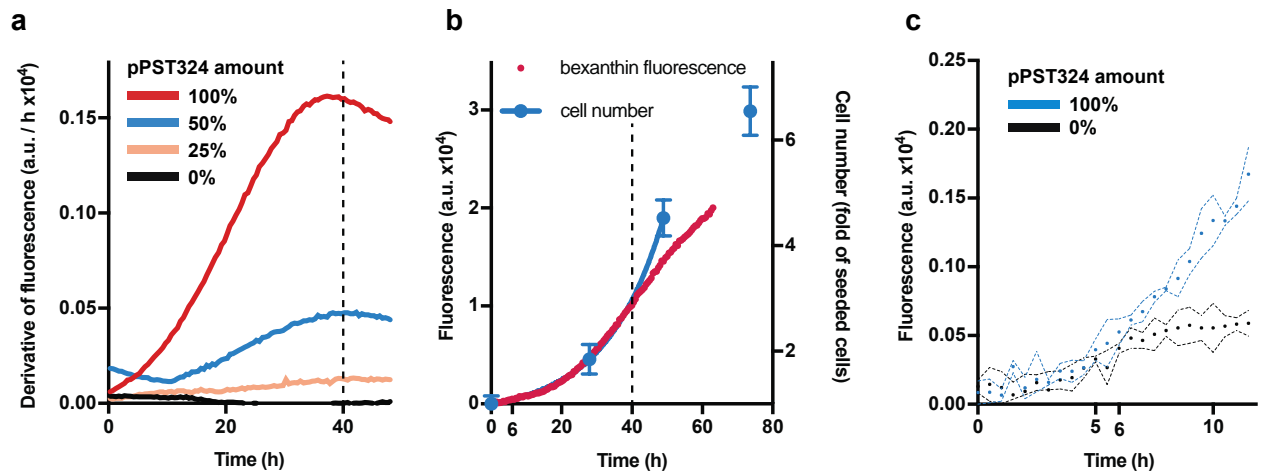
Figure S10: Derivative of the continuous measurement (Fig. 3c), comparison with growth curve and early-period changes in the continuous measurement (Fig. 3c). (a) Calculation of the increase in fluorescence instead of total fluorescence can be useful to analyze the kinetic features of the system. In the example at hand it can be easily seen that the maximum fluorescence production is reached at around 35 h, a feature that is hidden in Fig. 3c. The derivative values were calculated by GraphPad Prism 7 software. (b) Comparison of betaxanthin fluorescence (Fig. 3c) with the growth curve (Fig. S7c). The axes were chosen to allow visual comparison of the two curves. (c) Enlargement of Fig. 3c to show the initial increase in fluorescence. Graphs b and c show the mean ± s.d. of n = 3 independent samples.
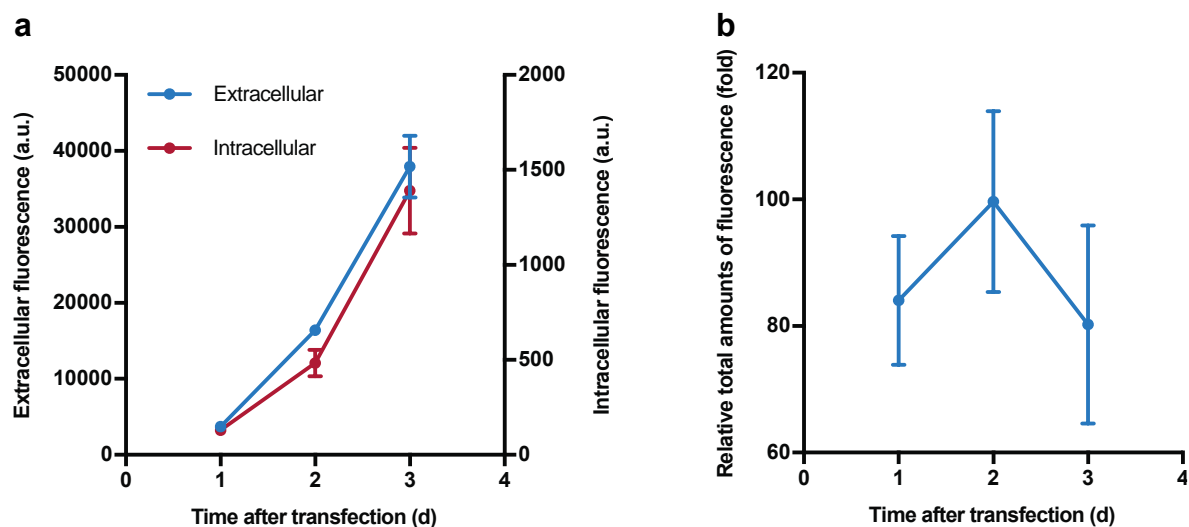
Figure S11: Analysis of the extra-/intracellular accumulation of fluorescent dye. (a) HEK293T cells were transfected with pPST324 and sampled every 24 h for intra- and extracellular dye fluorescence. Sampling was done by harvesting color-containing supernatant, washing the cells with PBS to remove residual extracellular dye, and lysing the cells with RIPA cell lysis buffer for 15 min at 37˚. Samples were then frozen until analysis with a plate reader. (b) Ratio of the total amount of extra- and intracellular fluorescence. The ratio was calculated by using the following equation: $(f_e*v_e)/(f_i*v_i)$, where f_e and f_i are the measured extra-/intracellular fluorescence values; v_e and v_i are the volumes used (culture medium for the extracellular measurement, amount of lysis buffer for intracellular measurement). Graphs show the mean ± s.d. of n = 3 independent samples.
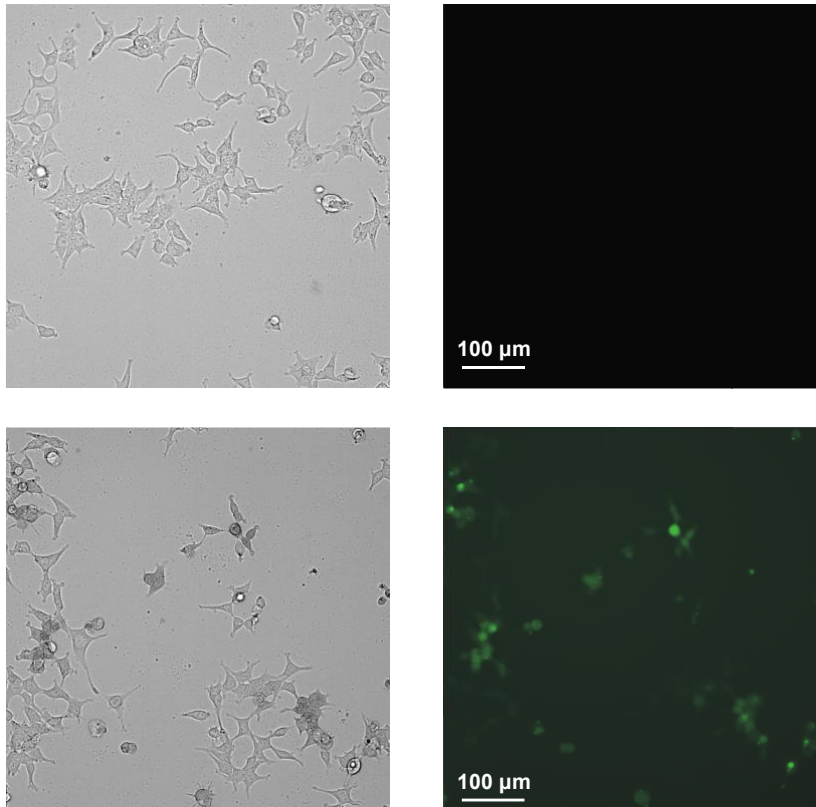
Figure S12: Micrographs of cells producing betaxanthin. Cells were transfected with pColaDuet-1 (top) or pPST324 (bottom), then 7.5 h later cells were reseeded to a density of $1.5*10^6$ cells per plate and measured 48 h later. Left shows bright-field images recorded with a blue light 440/20 nm transmission bandpass filter to allow for absorbance-based color detection. Right shows green fluorescence images recorded with a 488/6 nm excitation filter, 495 nm dichroic mirror and 520/35 nm emission filter.
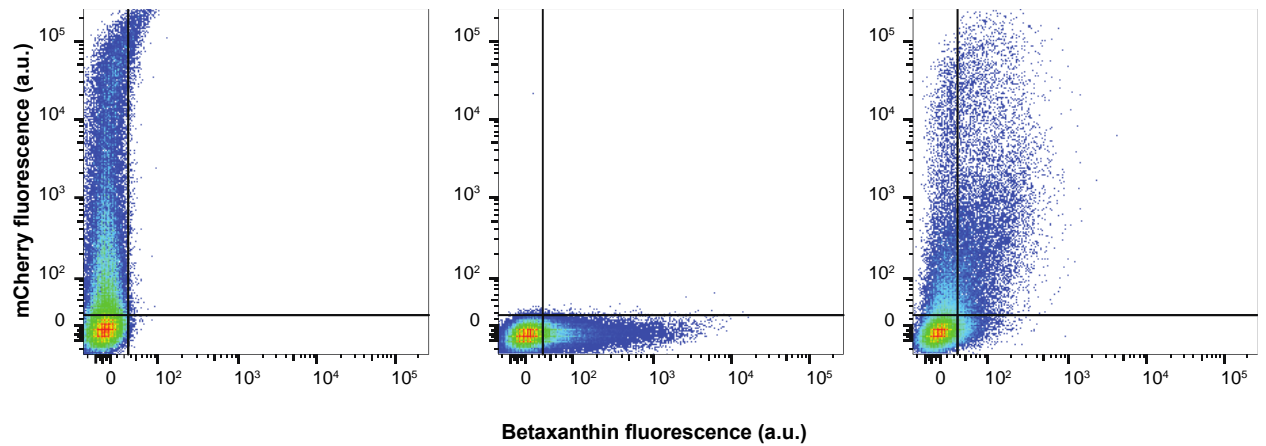
Figure S13: Dual reporter capabilities of the betaxanthin production system with mCherry fluorescent protein reporter. Cells producing red fluorescent mCherry and betaxanthin in the same cell can be analyzed using a flow cytometer. Left, cells transfected with mCherry (pFS20); middle, cells transfected with betaxanthin production cassette (pPST324); right side, cells transfected with both. Cell populations were analyzed with a Becton Dickinson LSRII Fortessa flow cytometer, equipped for EGFP detection (488-nm laser, 505-nm long-pass filter, 530 ± 15 nm emission filter) and mCherry detection (561nm laser, 610 ± 10 nm emission filter), and set to exclude dead cells, debris and cell doublets. Approximately 30000 cells were measured 48 h after transfection. The figure shows data representative of three independent experiments.
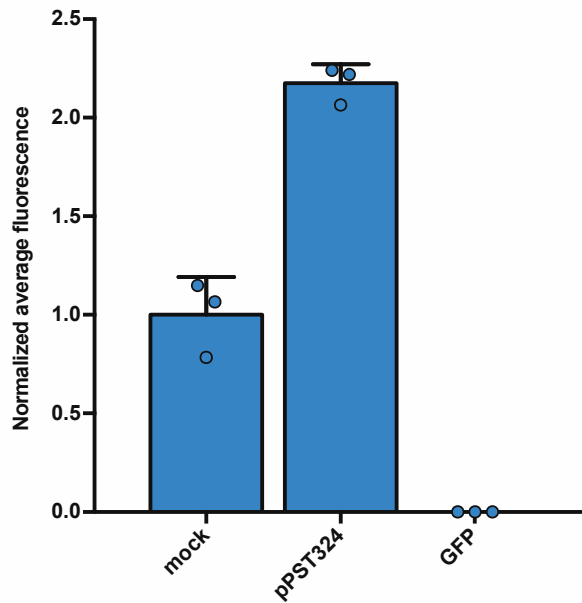
Figure S14: Fluorescence loss in cells fixed with formaldehyde. In experiments that require the fixation of cells, the fluorescence of fluorescent proteins is generally lost or drastically diminished. We tested whether this also applies to the intracellular fluorescence of betaxanthin-producing cells. Cells producing tGFP fixed with formaldehyde completely lost fluorescence (>100x reduction in fluorescence), whereas cells producing betaxanthin actually exhibited an increase in fluorescence (x2.2). Cells were transfected with mock (pCOLADuet-1), pPST324 and tGFP (pFOX13), fixed 48 h later with 2% formaldehyde for 1 h, and then analyzed using a flow cytometer with the same settings as for Fig. S13. The figure shows the average population fluorescence ± s.d. normalized to the average population fluorescence prior to fixation of n = 3 independent samples.
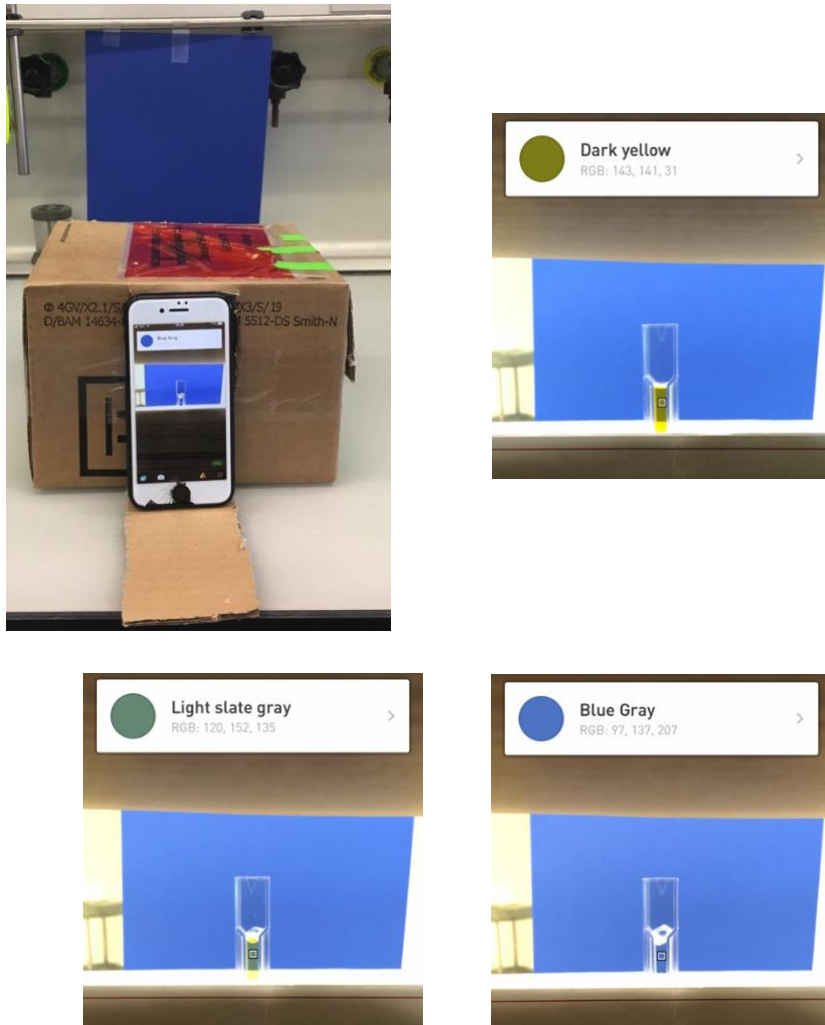
Figure S15: Smartphone-based betaxanthin quantification set-up and screenshots of data generation. Top left: Experimental set-up consisting of a cardboard box, a blue paper as background and a smartphone using the Color Name AR Pro application software. Top right: highest concentration of indicaxanthin used for the standard curve. Bottom left: Supernatant of cells transfected with pPST324, harvested and stored frozen after 72 h, bottom right: supernatant of cells transfected with pCola-Duet1, harvested and stored frozen after 72 h.

REFERENCES

1. Kremers, G.J., Gilbert, S.G., Cranfill, P.J., Davidson, M.W. and Piston, D.W. (2011) Fluorescent proteins at a glance. J. Cell Sci., 124, 157-160.
2. Tsien, R.Y. (1998) The green fluorescent protein. Annu. Rev. Biochem., 67, 509-544.
3. Jiang, T., Xing, B. and Rao, J. (2008) Recent developments of biological reporter technology for detecting gene expression. Biotechnol. Genet. Eng. Rev., 25, 41-75.
4. Berger, J., Hauber, J., Hauber, R., Geiger, R. and Cullen, B.R. (1988) Secreted placental alkaline phosphatase: a powerful new quantitative indicator of gene expression in eukaryotic cells. Gene, 66, 1-10.
5. Gould, S.J. and Subramani, S. (1988) Firefly luciferase as a tool in molecular and cell biology. Anal. Biochem., 175, 5-13.
6. Muzzey, D. and van Oudenaarden, A. (2009) Quantitative time-lapse fluorescence microscopy in single cells. Annu. Rev. Cell Dev. Biol., 25, 301-327.
7. Rodriguez, E.A., Campbell, R.E., Lin, J.Y., Lin, M.Z., Miyawaki, A., Palmer, A.E., Shu, X., Zhang, J. and Tsien, R.Y. (2017) The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. Trends Biochem. Sci., 42, 111-129.
8. Costantini, L.M., Baloban, M., Markwardt, M.L., Rizzo, M., Guo, F., Verkhusha, V.V. and Snapp, E.L. (2015) A palette of fluorescent proteins optimized for diverse cellular environments. Nat. Commun., 6, 7670.
9. Muller, M., Auslander, S., Auslander, D., Kemmer, C. and Fussenegger, M. (2012) A novel reporter system for bacterial and mammalian cells based on the non-ribosomal peptide indigoidine. Metab. Eng., 14, 325-335.
10. Muller, K., Engesser, R., Timmer, J., Nagy, F., Zurbriggen, M.D. and Weber, W. (2013) Synthesis of phycocyanobilin in mammalian cells. Chem. Commun., 49, 8970-8972.
11. Crichton, R.R. (2008) Sodium and Potassium-Channels and Pumps. Biological Inorganic Chemistry: An Introduction, 151-164.
12. Halperin, S.J. and Lynch, J.P. (2003) Effects of salinity on cytosolic Na+ and K+ in root hairs of Arabidopsis thaliana: in vivo measurements using the fluorescent dyes SBFI and PBFI. J. Exp. Bot., 54, 2035-2043.
13. Yonekura-Sakakibara, K., Higashi, Y. and Nakabayashi, R. (2019) The Origin and Evolution of Plant Flavonoid Metabolism. Front. Plant Sci., 10, 943.
14. Rolland, N., Bouchnak, I., Moyet, L., Salvi, D. and Kuntz, M. (2018) The Main Functions of Plastids. Plastids: Methods and Protocols, 1829, 73-85.
15. Tanaka, Y., Sasaki, N. and Ohmiya, A. (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant J., 54, 733-749.
16. Jerome, V., Freitag, R., Schuler, D. and Mickoleit, F. (2019) SEAP activity measurement in reporter cell-based assays using BCIP / NBT as substrate. Anal. Biochem., 585, 113402.
17. Tastanova, A., Folcher, M., Muller, M., Camenisch, G., Ponti, A., Horn, T., Tikhomirova, M.S. and Fussenegger, M. (2018) Synthetic biology-based cellular biomedical tattoo for detection of hypercalcemia associated with cancer. Sci. Transl. Med., 10.
18. Gandia-Herrero, F. and Garcia-Carmona, F. (2013) Biosynthesis of betalains: yellow and violet plant pigments. Trends Plant Sci., 18, 334-343.
19. Zrÿd, J.P. and Christinet, L. (2003). Betalain Pigments. Université de Lausanne.
20. Khan, M.I. and Giridhar, P. (2015) Plant betalains: Chemistry and biochemistry. Phytochemistry, 117, 267-295.
21. Khan, M.I. (2016) Plant Betalains: Safety, Antioxidant Activity, Clinical Efficacy, and Bioavailability. Compr. Rev. Food Sci. Food Saf., 15, 316-330.
22. Christinet, L., Burdet, F.X., Zaiko, M., Hinz, U. and Zryd, J.P. (2004) Characterization and functional identification of a novel plant 4,5-extradiol dioxygenase involved in betalain pigment biosynthesis in Portulaca grandiflora. Plant Physiol., 134, 265-274.

23. Gandia-Herrero, F. and Garcia-Carmona, F. (2012) Characterization of recombinant Beta vulgaris 4,5-DOPA-extradiol-dioxygenase active in the biosynthesis of betalains. Planta, 236, 91-100.

24. Schliemann, W., Kobayashi, N. and Strack, D. (1999) The decisive step in betaxanthin biosynthesis is a spontaneous reaction. Plant Physiol., 119, 1217-1232.

25. Kuntzleman, T.S. and Jacobson, E.C. (2016) Teaching Beer's Law and Absorption Spectrophotometry with a Smart Phone: A Substantially Simplified Protocol. J. Chem. Educ., 93, 1249-1252.

26. Simonsen, J.L., Rosada, C., Serakinci, N., Justesen, J., Stenderup, K., Rattan, S.I., Jensen, T.G. and Kassem, M. (2002) Telomerase expression extends the proliferative life-span and maintains the osteogenic potential of human bone marrow stromal cells. Nat. Biotechnol., 20, 592-596.

27. Schlatter, S., Rimann, M., Kelm, J. and Fussenegger, M. (2002) SAMY, a novel mammalian reporter gene derived from Bacillus stearothermophilus alpha-amylase. Gene, 282, 19-31.

28. Bartoloni, F.H., Goncalves, L.C.P., Rodrigues, A.C.B., Dorr, F.A., Pinto, E. and Bastos, E.L. (2013) Photophysics and hydrolytic stability of betalains in aqueous trifluoroethanol. Monatsh. Chem., 144, 567-571.

29. Mueller, L.A., Hinz, U. and Zryd, J.P. (1997) The formation of betalamic acid and muscaflavin by recombinant dopa-dioxygenase from Amanita. Phytochemistry, 44, 567-569.

30. Hinz, U.G., Fivaz, J., Girod, P.A. and Zyrd, J.P. (1997) The gene coding for the DOPA dioxygenase involved in betalain biosynthesis in Amanita muscaria and its regulation. Mol. Gen. Genet., 256, 1-6.

31. Gandia-Herrero, F., Escribano, J. and Garcia-Carmona, F. (2005) Betaxanthins as pigments responsible for visible fluorescence in flowers. Planta, 222, 586-593.

32. Nagatsu, T. (1995) Tyrosine hydroxylase: human isoforms, structure and regulation in physiology and pathology. Essays Biochem., 30, 15-35.

33. Lai, X., Wichers, H.J., Soler-Lopez, M. and Dijkstra, B.W. (2018) Structure and Function of Human Tyrosinase and Tyrosinase-Related Proteins. Chemistry, 24, 47-55.

34. Mishima, Y. and Imokawa, G. (1983) Selective aberration and pigment loss in melanosomes of malignant melanoma cells in vitro by glycosylation inhibitors: premelanosomes as glycoprotein. J. Invest. Dermatol., 81, 106-114.

35. Thony, B., Auerbach, G. and Blau, N. (2000) Tetrahydrobiopterin biosynthesis, regeneration and functions. Biochem. J., 347 Pt 1, 1-16.

36. Saxena, P., Charpin-El Hamri, G., Folcher, M., Zulewski, H. and Fussenegger, M. (2016) Synthetic gene network restoring endogenous pituitary-thyroid feedback control in experimental Graves' disease. Proc. Natl. Acad. Sci. U. S. A., 113, 1244-1249.

37. Delafosse, L., Xu, P. and Durocher, Y. (2016) Comparative study of polyethylenimines for transient gene expression in mammalian HEK293 and CHO cells. J. Biotechnol., 227, 103-111.

38. Strack, D. and Schliemann, W. (2001) Bifunctional Polyphenol Oxidases: Novel Functions in Plant Pigment Biosynthesis. Angew. Chem., Int. Ed. Engl., 40, 3791-3794.

39. DeLoache, W.C., Russ, Z.N., Narcross, L., Gonzales, A.M., Martin, V.J. and Dueber, J.E. (2015) An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. Nat. Chem. Biol., 11, 465-471.

40. Gandia-Herrero, F., Garcia-Carmona, F. and Escribano, J. (2006) Development of a protocol for the semi-synthesis and purification of betaxanthins. Phytochem. Anal., 17, 262-269.

41. Mata, A., Ferreira, J.P., Semedo, C., Serra, T., Duarte, C.M.M. and Bronze, M.R. (2016) Contribution to the characterization of Opuntia spp. juices by LC-DAD-ESI-MS/MS. Food Chem., 210, 558-565.

42. Liedhegner, E.A., Steller, K.M. and Mieyal, J.J. (2011) Levodopa activates apoptosis signaling kinase 1 (ASK1) and promotes apoptosis in a neuronal model: implications for the treatment of Parkinson's disease. Chem. Res. Toxicol., 24, 1644-1652.

43. Stansley, B.J. and Yamamoto, B.K. (2013) L-dopa-induced dopamine synthesis and oxidative stress in serotonergic cells. Neuropharmacology, 67, 243-251.

44. Czekanska, E.M. (2011) Assessment of cell proliferation with resazurin-based fluorescent dye. Methods Mol. Biol., 740, 27-32.

45. Ganini, D., Leinisch, F., Kumar, A., Jiang, J.J., Tokar, E.J., Malone, C.C., Petrovich, R.M. and Mason, R.P. (2017) Fluorescent proteins such as eGFP lead to catalytic oxidative stress in cells. Redox Biology, 12, 462-468.

46. Balleza, E., Kim, J.M. and Cluzel, P. (2018) Systematic characterization of maturation time of fluorescent proteins in living cells. Nat. Methods, 15, 47-51.

47. Subach, O.M., Cranfill, P.J., Davidson, M.W. and Verkhusha, V.V. (2011) An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. PLoS One, 6, e28674.

48. Auslander, D., Auslander, S., Charpin-El Hamri, G., Sedlmayer, F., Muller, M., Frey, O., Hierlemann, A., Stelling, J. and Fussenegger, M. (2014) A synthetic multifunctional mammalian pH sensor and CO2 transgene-control device. Mol. Cell, 55, 397-408.

49. Auslander, D., Auslander, S., Pierrat, X., Hellmann, L., Rachid, L. and Fussenegger, M. (2018) Programmable full-adder computations in communicating three-dimensional cell cultures. Nat. Methods, 15, 57-60.

50. Auslander, S., Stucheli, P., Rehm, C., Auslander, D., Hartig, J.S. and Fussenegger, M. (2014) A general design strategy for protein-responsive riboswitches in mammalian cells. Nat. Methods, 11, 1154-1160.

51. Muller, M., Auslander, S., Spinnler, A., Auslander, D., Sikorski, J., Folcher, M. and Fussenegger, M. (2017) Designed cell consortia as fragrance-programmable analog-to-digital converters. Nat. Chem. Biol., 13, 309-316.

52. Scheller, L., Strittmatter, T., Fuchs, D., Bojar, D. and Fussenegger, M. (2018) Generalized extracellular molecule sensor platform for programming cellular behavior. Nat. Chem. Biol., 14, 723-729.

53. Song, L., Hennink, E.J., Young, I.T. and Tanke, H.J. (1995) Photobleaching kinetics of fluorescein in quantitative fluorescence microscopy. Biophys. J., 68, 2588-2600.

CHAPTER II

CELLOSELECT-A SYNTHETIC CELLOBIOSE METABOLIC PATHWAY FOR SELECTION OF STABLE TRANSGENIC CHO-K1 CELL LINES

This chapter describes a work in progress and has not yet been submitted for publication.

CONTRIBUTIONS

Pascal Stücheli, Simon Ausländer and Martin Fussenegger developed the project idea and designed the project. Pascal Stücheli, Simon Ausländer and Martin Fussenegger wrote the text. Pascal Stücheli conducted most of the experiments and analyzed the results. Pascal Schönenberger and Samuel Hürlemann conducted initial experiments, Simon Ausländer and David Ausländer analyzed initial results that are not present in the current manuscript. Simon Ausländer created figures 1c and parts of 1b and 3a.

ABSTRACT

Current protocols for generating stable transgenic cell lines mostly rely on antibiotic selection or the use of specialized cell lines lacking an essential part of their metabolic machinery, but these approaches require working with either toxic chemicals or knockout cell lines, which can reduce productivity. Since most mammalian cells cannot utilize cellobiose, a disaccharide consisting of two β-1,4-linked glucose molecules, we designed an antibiotic-free selection system, CelloSelect, which consists of a selection cassette encoding Neurospora crassa cellodextrin transporter CDT1 and β-glucosidase GH1-1. When cultivated in glucose-free culture medium containing cellobiose, CelloSelect-transfected cells proliferate by metabolizing cellobiose as a primary energy source, and are protected from glucose starvation. We show that the combination of CelloSelect with a PiggyBac transposase-based integration strategy provides a platform for the swift and efficient generation of stable transgenic cell lines. Growth rate analysis of metabolically engineered cells in cellobiose medium confirmed the expansion of cells stably expressing high levels of a cargo fluorescent marker protein. We further validated this strategy by applying the CelloSelect system for stable integration of a sequence encoding the biopharmaceutical protein erythropoietin, and confirmed that the protein is efficiently produced in either cellobiose- or glucose-containing medium in a model bioproduction system. We believe coupling heterologous metabolic pathways additively to the endogenous metabolism of mammalian cells has the potential to complement or to replace current cell-line selection systems.

INTRODUCTION

Genomic integration and stable propagation of custom transgenes within mammalian cells is of great importance in basic research(1-3), biotechnology(4) and biomedicine(5,6). Chinese hamster ovary (CHO) cell lines stably producing medically relevant proteins, e.g. monoclonal antibodies (mAb), serve as the primary workhorse in biopharmaceutical production(7-9). The creation of stable producer cell lines requires integration of an expression cassette into the genome and subsequent selection of stably integrated cells (**Fig. 1a**). Genomic integration is achieved by one of three methods: random unaided integration, semi-targeted transposon(10) or virus-based integration(11), and targeted integration by CRISPR/Cas9-mediated homologous recombination(12). These methods differ in terms of integration

frequency, location specificity and available genomic payload complexity and can be selected as appropriate for the task at hand. Random integration(13) is still the current industry standard because of its relatively long history and ease of use. However, transposon-based integration is likely to replace random integration in the near future, by virtue of affording a vastly higher number of cells with genomic integration of the transgenic cargo and an enhanced copy number of the transgenic cargo per cell (integration frequency)(10). In a laboratory setting, the major strategies to select cells with genomic integration of expression cassettes make use of culture systems containing antibiotics with toxic effects on mammalian cells, coupled with the simultaneous genomic co-integration and expression of heterologous proteins that detoxify these antibiotics only in transgenic cells. Examples include antibiotics that function by ribosome inhibition (geneticin-G418) or DNA cleavage (zeocin), and resistance based on enzymatic detoxification (with *neoR* for G418(14)) or toxin binding (with *ble* for zeocin(15)). Cultivating the transgenic cells for multiple passages in medium containing the antibiotic allows only stably transgenic cells to survive. However, the applied antibiotics are toxic to humans, negatively affect our environment if not correctly disposed of, and are often expensive. Therefore, the industrial gold standard(16) in selecting stable cell lines is not antibiotics-based, but rather uses selection systems that involve eliminating endogenous enzymes catalyzing the synthesis of essential metabolites, such as dihydrofolate reductase (DHFR)(17) or glutamine synthetase (GS)(18) and including an expression cassette for these enzymes as a selection marker of transgenic cells. However, endogenous enzyme elimination relies on knockout cell lines and toxic (methotrexate/DHFR(19) and methionine sulfoximine/GS(20)) small-molecule-mediated inhibition and transgene amplification cycles(21). This approach has the disadvantage that knockout cell lines are difficult and time-consuming to create, and small-molecule-mediated inhibition suffers from the same limitations as antibiotics-based systems. Additionally, these cell lines (e.g. CHO-DG44(22), CHO-DXB11(23)) can have inferior production capacity(24).

In this work we present a novel method, which we call CelloSelect, to metabolically engineer cells for cultivation on cellobiose, making it possible to select stable cell lines by culture in a medium containing cellobiose as the sole source of glucose. CelloSelect is independent of antibiotics and does not require elimination of endogenous enzymes. In contrast to cellulolytic microorganisms and fungi, most mammalian cells cannot easily metabolize cellobiose(25). Degradation of ingested disaccharides already in the mammalian

gut is generally mediated by specific brush border disaccharidases,(26) and these enzymes are not expressed in non-specialized cells(25). Therefore, we explored the possibility of expanding the metabolism of mammalian cells by introducing the heterologous *Neurospora crassa*-derived(27) cellobiose utilization pathway, thereby endowing mammalian cells with the capacity to survive and grow on cellobiose as the sole carbon source (**Fig. 1b**). Our results show that this approach provides a new and effective strategy for stable cell line selection based on expansion of the primary cellular metabolism.

RESULTS

Design of the cellobiose utilization pathway.

To enable CHO-K1 cells to utilize cellobiose as an energy source, we created three genetic constructs (**Fig. 1c**) encoding for P_{CAG}-driven expression of either the *N. crassa*-derived cellodextrin transporter (CDT1), or β-glucosidase (GH1-1), or a combination of the two separated by a T2A cleavage peptide,(28) flanked by PiggyBac transposase(10)-based inverted terminal repeats (ITRs). To assess the functionality of the constructs, we transfected CHO-K1 cells separately with each of the above three plasmids or with both the CDT1 and GH1-1 plasmids together. After 48 h recovery, the transfected cells were cultured in glucose-rich (positive control), cellobiose-rich and glucose-depleted (selection conditions), or glucose and cellobiose-depleted (negative control) media. Cell viability was measured with a resazurin-based assay after 48 hours. Cells transfected with the empty vector or only the cellodextrin transporter died when cultivated in cellobiose-containing/glucose-depleted medium or sugar-free medium, but showed high viability when cultivated in glucose-containing culture medium (**Fig. 1d**). Expression of β-glucosidase alone or both β-glucosidase and the cellodextrin transporter rescued CHO-K1 cells from glucose starvation. To confirm the functionality of the cellobiose utilization pathway in CHO-K1 cells, we cultivated transfected cells in medium containing resorufin-β-D-cellobioside, and observed release of the fluorescent product (resorufin) only in cells that expressed β-glucosidase alone or that expressed both β-glucosidase and the cellodextrin transporter (whether in separate constructs or combined) (**Fig. 1e**).
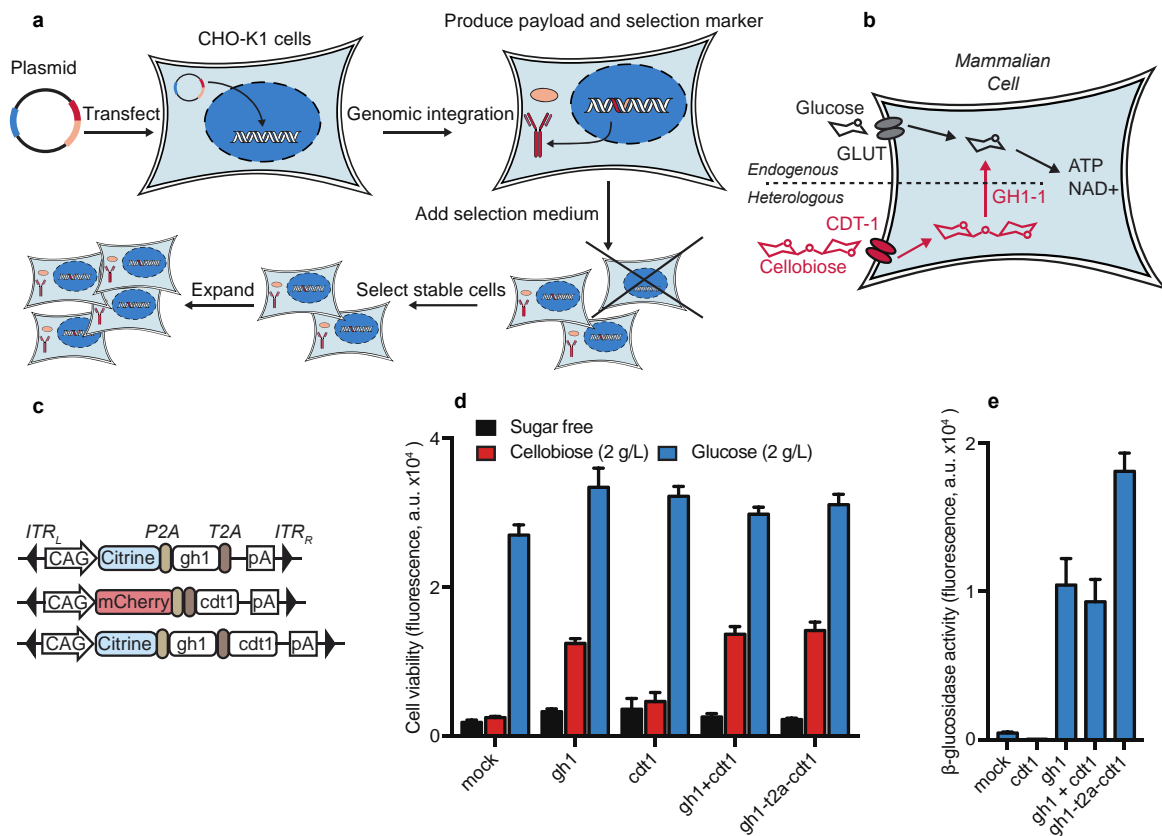
Figure 1: Characterization of the cellobiose utilization pathway in mammalian cells. a) Schematic overview of stable cell line selection in mammalian cells. b) Schematic overview of CelloSelect. Expression of the cellodextrin transporter (CDT1) and β-glucosidase (GH1-1) enables cells to metabolize cellobiose. c) Schematic illustration of PiggyBac-based transposon donor plasmids encoding for P_{CAG}-driven gene expression cassettes. Constructs from top to bottom correspond to the plasmids pSA845, pSA846, pSA847. ITR, Inverted terminal repeat sequences; P2A/T2A, self-cleaving peptides; Citrine/mCherry, fluorescent proteins. d) Resazurin-based cell viability assay of CHO-K1 cells transfected with plasmids encoding β-glucosidase or cellodextrin transporter, or both, or the complete CelloSelect system and cultivated in different sugar-containing cell culture media. e) Ability of CHO-K1 cells expressing components of the CelloSelect system or the full system to degrade resorufin-β-D-cellobioside. Panels d and e show mean ± s.d. of n = 3 independent samples and are representative of three independent experiments.

Characterization of CelloSelect medium

We further investigated the influence of the number of seeded cells, the FCS concentration, and the cellobiose concentration on cell survival. CHO-K1 cells were co-transfected with plasmids encoding for a hyperactive PiggyBac transposase (hyPBase) and the full CelloSelect system with the fluorescent protein Citrine as a model payload. At 48 h after

transfection, different numbers of cells were seeded in cell culture media with varying FCS and cellobiose concentrations and analyzed three days later by flow cytometry to measure the percentage of fluorescence-positive cells (**Fig. 2a**) and the cell number (**Fig. 2b**). We found that lower concentrations of cellobiose and FCS increased the percentage of fluorescent cells and higher numbers of seeded cells improved cell survival. The number of surviving cells was inversely correlated with the percentage of fluorescence-positive cells. Therefore, for better selection efficiency, we chose to prioritize fluorescent cell enrichment over cell survival. We define the selection conditions as: 3×10^6 seeded cells/well or dish, 5% FCS, 0.125 g/L cellobiose (DMEM$_{select}$); in the experiment shown in **Fig. 2a** these conditions yielded 73% fluorescent cells.
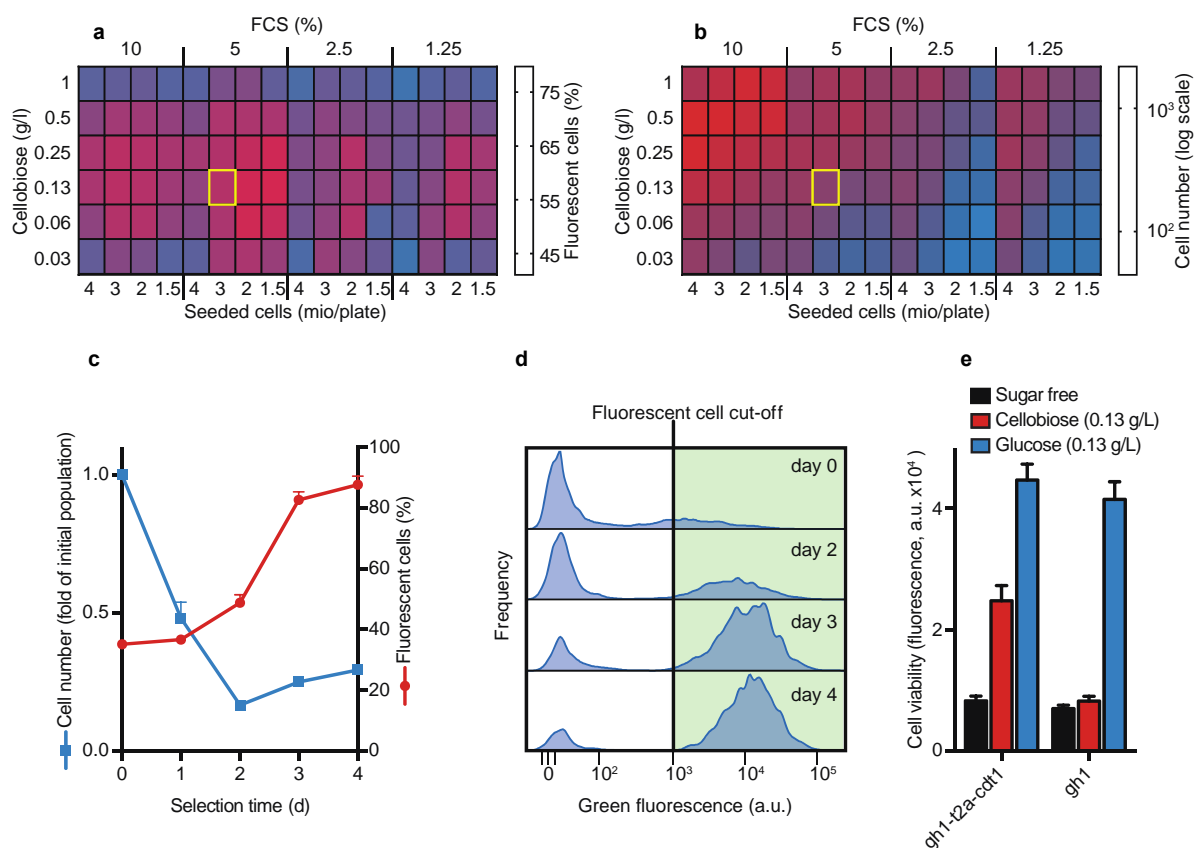

Figure 2: Development of CelloSelect, a cellobiose-based selection method. a) Screening of medium composition for stable cell-line selection, evaluated in terms of fluorescent cell enrichment. CHO K1 cells transfected with the CelloSelect system were cultured under the indicated conditions. Blue color represents the proportion of fluorescent cells at seeding. In panels a and b the yellow square shows the conditions chosen for further experiments. Fluorescent cell percentage and cell number (fold increase) were recorded 72 h after seeding into selection medium. b) Screening of medium composition for stable

cell-line selection, evaluated in terms of cell survival. c) Changes of fluorescent cell enrichment and cell number in a selection cycle of cells transfected with the CelloSelect system. Sugar-free and mock-transfected controls are shown in Fig. S1. d) Flow cytometry data of populations at various time points during a selection cycle. This panel shows four populations from panel c with >5000 cells per histogram. The vertical line shows the fluorescent cell cut-off, and the green area indicates fluorescent cells. e) Resazurin-based cell viability assay of CHO-K1 cells transfected with plasmids encoding the complete CelloSelect system or only β-glucosidase, and cultivated in media containing a low level of glucose or cellobiose, or no sugar. Panels a and b show the mean values of n = 3 independent populations. Panels c and e show the mean ± s.d. of n = 3 independent samples and are representative of three independent experiments.

Establishing an optimized CelloSelect protocol

We repeated CelloSelect with the above selection conditions in a 24-well format and analyzed the cell number and percentage of fluorescent cells daily for five days (**Figs. 2c-d**). In transfected cultures in DMEM$_{select}$, fluorescent cells accounted for more than 85% of the population after four days (**Fig. 2c**). The untransfected and sugar-free control populations were almost completely depleted after three days (**Fig. S1**). The selected population was drastically diminished, but the number of cells started to recover after two days. The viability of cells transfected with either the full CelloSelect system or with GH1-1 only was measured after 4 d selection with the optimized conditions, and the results confirmed that the cellodextrin transporter CDT1 is required for cell survival under selection conditions (**Fig. 2e**).

Stable population selection and long-term expansion.

We transfected cells with the full CelloSelect selection cassette, seeded them 48 h later in DMEM$_{select}$ at 3×10^6 cells/plate, and analyzed them periodically with a flow cytometer. Every 3-4 days, cells were passaged and analyzed using flow cytometry. Up to 100% fluorescent cells were observed during 4-11 days in three independent experimental runs (**Fig. 3a, b**). After the initial selection cycle, cells multiplied rapidly, and seeding 1×10^6 cells/plate was sufficient for cycles 3 and above. Since cellobiose simultaneously serves as a selection agent and an energy source, modified DMEM$_{select}$ with different sugar and FCS concentrations can likely be used for continuous maintenance of selection pressure during cell cultivation. To identify an appropriate medium composition for long-term culture, we seeded 1×10^6 stable cellobiose-

selected Citrine-expressing cells per plate in DMEM$_{select}$ with 10% FCS and different amounts of cellobiose or glucose. After 48 h, the cell number was determined using flow cytometry. The amount of glucose or cellobiose had a negligible impact on the 48 h growth rate, with growth in cellobiose medium being only slightly (approx. 23%) slower than growth in glucose medium (**Fig. 3c**). Monitoring the long-term stability of transgenic populations in 12 cultures with 4 different media compositions for 3 weeks showed no loss of fluorescent cells, independent of the amount of glucose or cellobiose in the medium (**Fig. 3d, Fig. S2**).
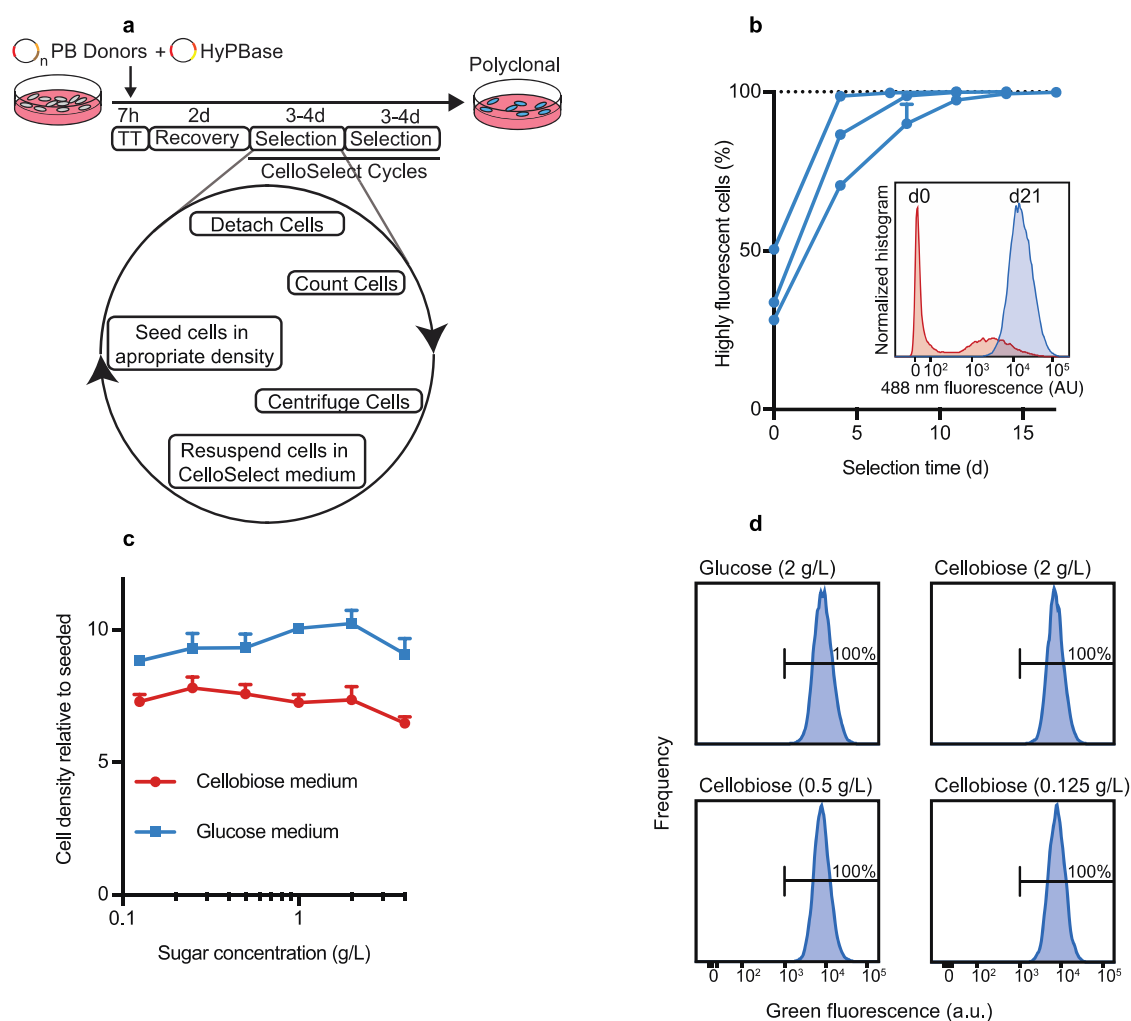

Figure 3: Application of CelloSelect to generate stable CHO-K1 cell lines. a) Schematic illustration of the procedure for generating stable cell lines with the CelloSelect system. b) Time-course of three complete CelloSelect selection runs for generating stable Citrine-expressing cell lines. CHO-K1 cells were transfected with the CelloSelect system and a constitutive transposase expression plasmid. The selection was conducted as shown in panel a, and after each cycle samples of the cell populations were analyzed by flow cytometry. The inset shows a representative flow-cytometry histogram in which the red area

corresponds to the population at first seeding into DMEM$_{select}$ and the blue area corresponds to the population after selection. c) Influence of medium composition on the growth rate of CHO-K1 cells stably expressing the CelloSelect selection cassette. d) Fluorescence of CHO-K1 cells stably expressing the CelloSelect selection cassette 21 days after selection in different culture media. The green area indicates fluorescent cells. Data for all populations can be found in Fig. S2. Panel b shows the mean ± s.d. of 3 independent experiments with n = 3 populations in each experiment, panel c shows the mean ± s.d. of n = 3 independent samples and is representative of three independent experiments, and panel d shows a representative example of three independent populations for each condition.

Proof of concept: application for stable biopharmaceutical protein production.

To exemplify the use of the CelloSelect system for the stable integration of a sequence encoding a biopharmaceutical protein we created a slight variation (pPST347, **Fig. 4a,b**) of the construct used for benchmarking the selection system (pSA847, **Fig. 1c**). As a model biopharmaceutical cargo we chose the therapeutic protein erythropoietin (EPO, *epo*, GenBank: AAA52400). To create an exactly full-length protein without additional C-terminal amino acids we used a furin-T2A fusion to connect the EPO gene to the selection cassette, following a previously reported approach.(29) Briefly, T2A stuttering occurs during translation, restarting translation for the rest of the fusion protein. As the EPO protein is a secretory protein it is transported into the Golgi, where furin cleavage occurs, followed by carboxypeptidase-mediated removal of C-terminal lysine and arginine residues.(30)
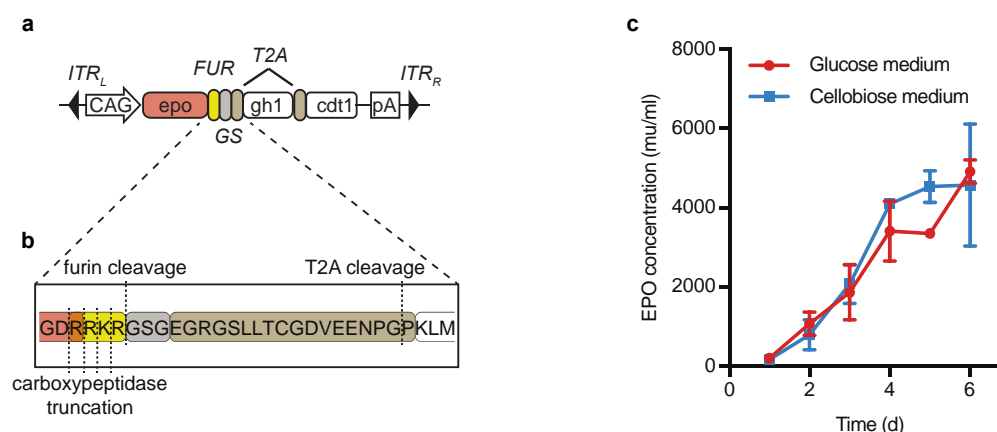

Figure 4: Using the CelloSelect system to create cell lines stably producing a biopharmaceutical protein. a) Schematic illustration of the genetic construct containing an erythropoietin-encoding cargo gene. FUR, furin cleavage site; GS, small GS linker peptide; ITR, inverted terminal repeat sequences; T2A,

self-cleaving peptides. b) Schematic illustration of the different cleavage- and processing-modules used to separate the cargo from the selection cassette. c) Model EPO production run for six days in DMEM containing either 4 g/L glucose or cellobiose. Panel c shows the mean ± s.d. of n = 2 independent samples.

We transfected CHO-K1 cells in parallel with the CelloSelect selection cassette encoding either Citrine or EPO as cargo and performed a full selection run according to the standardized method. The progress of selection was observed by analyzing the Citrine expressing cells using flow cytometry (**Fig. S3**). When the fluorescent cell count reached 100%, one more selection cycle was added, and then both populations were expanded in DMEMselect. The complete selection and expansion process required 13 days. To profile EPO production, EPO-producing cells were seeded into culture medium containing either glucose or cellobiose. The supernatant from individual samples was harvested and frozen daily for six days, and finally the EPO concentration was determined by means of ELISA (**Fig. 4c**). The EPO concentration increased at approximately the same rate in the glucose- and cellobiose-grown populations throughout the course of the experiment.

DISCUSSION

Metabolic engineering in mammalian cells mainly focuses on improving and engineering central metabolic pathways to improve productivity(31,32), to reduce the generation of by-products(33) or to increase cell viability(34,35). In this work, we set out to transfer a heterologous metabolic pathway from *N. crassa* to mammalian cells. We demonstrated that the cellobiose utilization pathway enabled engineered cells to survive and proliferate on cellobiose as the sole energy source. Since mammalian cells do normally not metabolize cellobiose(25), this system can facilitate the specific and efficient selection of CHO-K1 cells that stably express the cellobiose utilization pathway. The proof-of-concept bioprocessing experiment establishes the ability of our system to create stable EPO-producing cells that can be grown for an extended period of time in regular culture medium, as well as in cellobiose-containing medium. Compared to existing selection methods, CelloSelect has the advantage of employing a cheap and non-toxic selection agent, permitting continuous maintenance of selection pressure and potentially reducing the necessary amount of downstream processing. Using this added-function approach instead of a destructive(24) knockout approach should enable cell-line engineers to focus more on production capacity than on selection-enabling

mutations. Future studies directed at amplifying transgene integration could employ similar strategies to those used in some current GS or DHFR systems; for example, adding a β-glucosidase inhibitor(36) to only allow cells with higher cargo production activity to survive or using a less active β-glucosidase to achieve more stringent selection.

This work also serves as a proof of concept regarding the introduction of heterologous metabolic pathways into mammalian cells to feed into or augment endogenous metabolism. For example, other heterologous carbohydrate degradation pathways, such as the cellulose degradation pathway, could be implemented in mammalian cells to engineer cells capable of metabolizing lignocellulose(37,38) or other cheap carbon sources. It may also be possible to utilize carbon sources that afford lower levels of toxic metabolic by-products or that enable higher cell densities, as exemplified by maltose(39). Building on the design principles of CelloSelect, it should be possible to generate specialized selection systems based on orthogonal sugars feeding into essential metabolism of a variety of mammalian or insect cell lines, offering the potential for higher production titers, which are needed to match the ever-growing demand for biotherapeutics.

Competing financial interests.

The authors declare no competing financial interests.

MATERIALS AND METHODS

DNA constructs

Comprehensive design and construction details for all expression vectors are provided in **Tables S1 and S2**.

Mammalian cell culture and transfection.

Chinese hamster ovary (CHO-K1, ATCC: CCL-61) cells were cultured in full DMEM (Dulbecco's modified Eagle's medium (DMEM, ThermoFisher, cat. no.: 31053028) supplemented with 10% (v/v) fetal calf serum (FCS, Sigma-Aldrich, cat. no.: F7524) and 0.15 mM L-proline (Fluka, cat. no.: 81710)) at 37°C in a humidified atmosphere of 7.5% CO_2 in air. For routine tasks, cell viability and cell number were quantified using an electric field multichannel cell-counting device (Casy Cell Counter and Analyzer Model TT, Roche Diagnostics GmBH). For experiments, 1.5×10^6 cells were seeded into each well of a multiwell plate or into a 10 cm dish one day prior to transfection in full DMEM. Plasmid DNA was mixed with polyethylene imine (PEI, MW 40 000, stock solution: 1 mg/mL in ddH_2O, Polysciences, cat. no. 24765-2) in a 1:5 ratio (w/w) and incubated in FCS-free DMEM for 15 min at room temperature before being added to the cells. After 6 h, the transfection medium was replaced with fresh full DMEM. For detailed information regarding transfections see **Table S3**. Cell viability assays were done using the resazurin assay as described below, and analytical cell number was determined using flow cytometry.

CelloSelect

At 48 h after transient transfection, the cells were trypsinized and centrifuged. 3×10^6 cells/well were reseeded in a multiwell plate in CelloSelect selection medium DMEM$_{select}$ consisting of glucose- and pyruvate-free DMEM (ThermoFisher, cat. no.: A1443001) supplemented with 1x Glutamax (ThermoFisher, cat. no.: A1443001, cat. no.: 35050061), 5% dialyzed FCS (ThermoFisher, cat. no.: A3382001) and 0.125 g/L cellobiose (Roth, cat. no.: 5840.2). Part of the cells was used for flow cytometric analysis. Selection cycles consisted of 3-4 days cultivation in selection medium. After each selection cycle, cells were trypsinized, and examined by flow cytometry. Then 3×10^6 cells/well were reseeded in a multiwell plate

(if cultures were confluent in less than 3 days the seeded cell number was reduced to 1×10^6 cells/well). If indicated, polyclonal cell populations were further cultivated in glucose-containing full DMEM or in DMEM$_{select}$ with 10% instead of 5% dialyzed FCS and varying amounts of cellobiose.

Flow Cytometry

Cell populations were analyzed with a LSRII Fortessa flow cytometer (Becton Dickinson) equipped for Citrine (488 nm laser, 505 nm long-pass filter and 530/30 emission filter) detection and set to exclude cell debris and cell doublets. At least 5,000 single cells were recorded per sample and were analyzed with FlowJo software (version no. 10; FlowJo LLC, Oregon, USA).

Resazurin assay

The culture medium was replaced with fresh, clear medium containing 25 mg/L resazurin sodium salt (Sigma Aldrich, R7017). The cells were placed back in the incubator at 37 °C for 3 h, and then the fluorescence (571 nm ex./ 585 nm em.) was measured and the background (culture medium with resazurin) was subtracted. A stock solution of 2.5 g/L was prepared in dH$_2$O.

Beta-glucosidase activity assay

Cells were incubated with 0.5 mM resorufin-β-D-cellobioside (5 mM stock solution in full DMEM, Marker Gene Technologies Inc., cat. no.: M1238) for 6 h in cell culture multiwell plates until fluorescence measurement (571 nm ex./ 585 nm em.).

EPO production assay

Stably transfected cells producing EPO were seeded at a density of 1x106 cells per plate into a 24-well plate. Standard DMEM, containing either 4 g/L glucose or cellobiose was used. Every day the supernatant was harvested and frozen until used for EPO concentration measurement. The concentration was determined with a human EPO ELISA kit (Sigma Aldrich, RAB0654) according to the manufacturer's protocol. The dilution factor was determined in

two steps to allow measurement in the linear range. In the first step the maximum dilution factor was determined from the day six sample using serial dilutions. In the second step, one sample from each day was analyzed using this maximum dilution factor to adjust the dilution factor for less concentrated samples. Then, the actual measurements were performed.

SUPPLEMENTARY INFORMATION

Table S1: Plasmids used in this work

Plasmid Name	Detailed Cloning Description	Source
pcDNA3.1(+)	Cloning vector for constitutive expression of target genes (P_{hCMV}-MCS-pA).	ThermoFisher
pCMV-hyPBase	Vector encoding for P_{CMV}-driven hyPBase expression cassette (P_{CMV}-hyPBase-pA).	Yusa et al. *Proc Natl Acad Sci U S A* (2011).
PB.CAG-Venusnucmem.rbpA	PiggyBac transposon vector encoding for P_{CAG}-driven Venusnucmem expression cassette (P_{CAG}-Venusnucmem-pA).	Plasmid was a gift from Timm Schröder.
pFS29	Vector encoding for a P_{SV40}-driven mCherry expression cassette (P_{SV40}-mCherry-pA).	Ausländer et al. *Nat. Methods* (2014).
pDF145	Plasmid without mammalian promoter as empty transfection vector.	Scheller et al. *Nat Chem Bio* (2018).
pCK222	Vector encoding for a constitutive cellobiose transporter expression cassette (P_{CMV}-cdt1-pA). A human codon-optimized version of cdt1 was synthesized by gene synthesis, digested with EcoRI/XbaI and ligated into pcDNA3.1(+)(EcoRI/XbaI).	This work
pCK223	Vector encoding for a constitutive beta-glucosidase expression cassette (P_{CMV}-gh1-pA). A human codon-optimized version of cdt1 was synthesized by gene synthesis, digested with EcoRI/XbaI and ligated into pcDNA3.1(+)(EcoRI/XbaI).	This work
pSA835	Vector encoding for a mutated constitutive cellobiose transporter expression cassette (P_{CMV}-cdt1(D433G)-pA). The mutation D433G was introduced into cdt1 by whole-plasmid PCR-directed site-directed mutagenesis from pCK222 with oligonucleotides oSA832 and oSA831 and the plasmid was religated.	This work
pSA836	Vector encoding for a mutated constitutive cellobiose transporter expression cassette (P_{CMV}-cdt1(C82S-D433G)-pA). The mutation C82S was encoded on a gBlock(cdt1(C82S),	This work

	digested with EcoRI/BsrGI and ligated into pSA835 (EcoRI/BsrGI).	
pSA837	Vector encoding for a mutated constitutive beta-glucosidase expression cassette (P$_{CMV}$-gh1(H23L)-pA). The mutation H23L was introduced into gh1 by whole-plasmid PCR-directed site-directed mutagenesis from pCK223 with oligonucleotides oSA826 and oSA825 and the plasmid was religated.	This work
pSA838	Vector encoding for a mutated constitutive beta-glucosidase expression cassette (P$_{CMV}$-gh1(H23L-L137H)-pA). The mutation L137H was introduced into gh1(H23L) by whole-plasmid PCR-directed site-directed mutagenesis from pSA837 with oligonucleotides oSA827 and oSA828 and the plasmid was religated.	This work
pSA839	Transposon-based vector encoding for a P$_{CAG}$-driven empty expression cassette (P$_{CAG}$-P2A-T2A-rbpA). P2A-T2A fragment was generated with oligonucleotide annealing using oligonucleotides oSA833 and oSA834 (insert 1) and oligonucleotides oSA835 and oSA836 (insert 2). Inserts 1 and 2 were digested with HindIII, and ligated. The P2A-T2A product was digested with EcoRI/NotI and ligated into PB.CAG-Venusnucmem.rbpA (EcoRI/NotI).	This work
pSA840	Transposon-based vector encoding for a P$_{CAG}$-driven Citrine expression cassette (P$_{CAG}$-Citrine-P2A-T2A-pA). The Citrine fragment was PCR-amplified from pDA701 with oligonucleotides oSA837 and oSA838, digested with EcoRI/AgeI and ligated into pSA839 (EcoRI/AgeI).	This work
pSA841	Transposon-based vector encoding for a P$_{CAG}$-driven mCherry expression cassette (P$_{CAG}$-mCherry-P2A-T2A-pA). The mCherry fragment was PCR-amplified from pFS29 with oligonucleotides oSA837 and oSA838, digested with EcoRI/AgeI and ligated into pSA839 (EcoRI/AgeI).	This work
pSA845	Transposon-based vector encoding for a P$_{CAG}$-driven Citrine-P2A-gh1(H23L-L137H)-T2A expression cassette (P$_{CAG}$-Citrine-P2A-gh1(H23L-L137H)-T2A-pA). The gh1(H23L-L137H) fragment was PCR-amplified from pSA838 with	This work

	oligonucleotides oSA839 and oSA840, digested with HindIII/XhoI and ligated into pSA840 (HindIII/XhoI).	
pSA846	Transposon-based vector encoding for a P$_{CAG}$-driven mCherry-P2A-T2A-cdt1(C82S-D433G) expression cassette (P$_{CAG}$-mCherry-P2A-T2A-cdt1(C82S-D433G)-pA). The cdt1(C82S-D433G) fragment was PCR-amplified from pSA836 with oligonucleotides oSA841 and oSA842, digested with BamHI/NotI and ligated into pSA841 (BamHI/NotI).	This work
pSA847	Transposon-based vector encoding for a P$_{CAG}$-driven Citrine-P2A-gh1(H23L-L137H)-T2A-cdt1(C82S-D433G) expression cassette (P$_{CAG}$-Citrine-P2A-gh1(H23L-L137H)-T2A-cdt1(C82S-D433G)-pA). The cdt1(C82S-D433G) fragment was PCR-amplified from pSA836 with oligonucleotides oSA841 and oSA842, digested with BamHI/NotI and ligated into pSA845 (BamHI/NotI).	This work
pPST347	Transposon-based vector encoding for a P$_{CAG}$-driven EPO-FUR-GS-T2A-gh1(H23L-L137H)-T2A-cdt1(C82S-D433G) expression cassette. The synthetic EPO gene fragment (Twist Bioscience) was PCR-amplified with oligonucleotides oPST616 and oSA617, digested with EcoRI/HindIII and ligated into pSA847 (EcoRI/HindIII).	This work

Abbreviations and additional information: Citrine, improved version of YFP derived from *Aequorea victoria*; **FUR**, furin cleavage site; **GS**, GSG short peptide linker sequence; **mCherry**, red-fluorescent protein; **MCS**, multiple cloning site; **EPO**, erythropoietin; **P2A**, P2A cleavage site; **pA**, polyadenylation signal; **P$_{CAG}$** strong synthetic mammalian promoter; **P$_{hCMV}$**, human cytomegalovirus immediate-early promoter; **P$_{SV40}$**, simian virus 40 promoter; **SEAP**, human placental secreted alkaline phosphatase; **T2A**, T2A cleavage site; **Venusnucmem**, Improved version of a yellow fluorescent protein with a nuclear membrane anchor.

Table S2: Oligonucleotide sequences used in this work

Oligo Name	Oligonucleotide Sequence (5'-3')
oSA825	AATAGCTCCCTCAATCTGGTAAG
oSA826	CTGGCTGATGGAAGAGGACCAAGCATCTG
oSA827	GATACTAGAACACCAAGGCTCG

oSA828	CAC GGATACAATAGCGGGTATTTCGCTCC
oSA831	CTGCTCATTCATGCCCTGATTC
oSA832	GGC GCAAAGGATAACGCCTACCTG
oSA835	CGAGCAAGCTTGAGCCGCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTCTTCT AACATGCGGTGACGTGGAGGAGAATCCCGGCCCTGGATCCGAGTAAGCGGCCGCGCTGA
oSA836	TCAGCGCGGCCGCTTACTCGGATCCAGGGCCGGGATTCTCCTCCACGTCACCGC ATGTTAGAAGACTTCCTCTGCCCTCTCCGCTTCCCTCGAGCGGCTCAAGCTTGCTCG
oSA837	GCGACGAATTCCCACCATGGTGAGCAAGGGCGAGGAG
oSA838	GCAGGTACCGGTCTTGTACAGCTCGTCCATGCC
oSA839	CGAGCAAGCTTATGTCTCTGCCCAAGGACTTC
oSA840	GAGCCGCTCGAGATCTTTCTTGATCAGGGAGTCG
oSA841	CGGCCCTGGATCCATGTCCTCCCACGGCTCCCACGAC
oSA842	GAGTAAGCGGCCGCTTATGCCACGATTGCTTCTGAG
oPST616	GTGGCAAAGCTTAGGGCCGGGATTCTCCTCCACGTCACCGCATGTTAG AAGACTTCCTCTGCCCTCTCCGCTTCCGCGTTTCCGGCGATCACCAGTTCTACATGCTTC
oPST617	TTTGGCAAAGAATTCCCACCATGGGCGTACATGAGTGTCC

Table S3: Transfection sheet

| Fig. 1d	In 24-well format, per well: mock, 600 ng pDF145; gh1, 600 ng pSA845; cdt1, 600 ng pSA846; gh1+cdt1, 300 ng pSA845, 300 ng 846; gh1-T2A-cdt1, 600 ng pSA847.
Fig. 1e	In 24-well format, per well: vector, 600 ng pDF145; cdt1 600 ng pSA846: gh1+vector, 300 ng pDF145, 300 ng pSA846; gh1+cdt1, 300 ng pSA845, 300 ng 846; gh1-T2A-cdt1, 600 ng pSA847.
Fig. 2a-d	In 10 cm dish format: positive, 12000 ng pSA847, 3000 ng pCMV-HypBase; negative, 15000 ng pDF145.
Fig. 2e	In 24-well format, per well: gh1, 600 ng pSA846; gh1-T2A-cdt1, 600 ng pSA847.
Fig. 3b	In 10 cm dish format: positive, 12000 ng pSA847, 3000 ng pHypBase.
Fig. 4c	In a 6-well format per well, 2000 ng pPST347, 500 ng pHypBase.
Fig. S1	See negative for Fig. 2a-d
Fig. S3	In a 6-well format per well, 2000 ng pSA847, 500 ng pHypBase.

SUPPLEMENTARY FIGURES
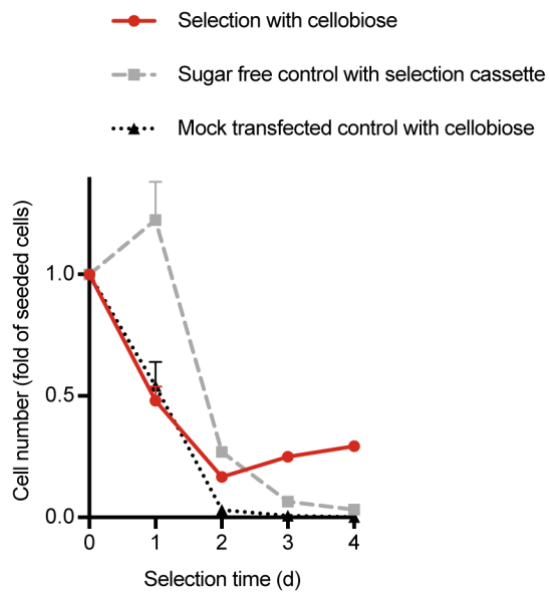
Figure S1: Decrease in cell number in sugar-free and mock-transfected controls (Fig. 2c). Sugar-free control cells were transfected with the CelloSelect selection cassette; mock-transfected cells were grown in cellobiose-containing medium. The figure shows the mean ± s.d. of n = 3 independent samples and is representative of three independent experiments.
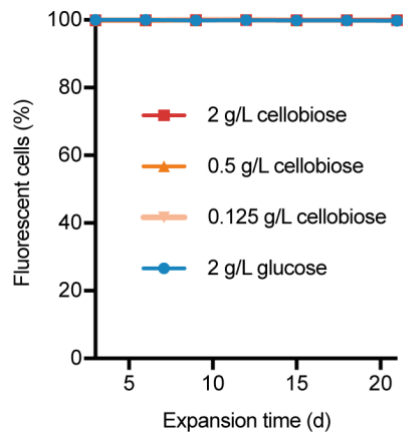
Figure S2: Loss of fluorescence of cells stably expressing Citrine-P2A-GH1-T2A-CDT1 in different media. Cells were reseeded every 3 days and fluorescence was analyzed using flow cytometry. The figure shows the mean ± s.d. of n = 3 independent populations for each medium composition.
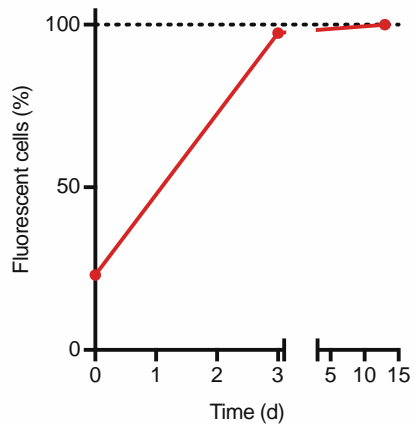
Figure S3: Co-selection of fluorescent cells and cells producing the biopharmaceutical model protein EPO. Cells were transfected with pHypBase and either pSA847 or pPST347, and stable cells were selected using the standardized CelloSelect method. Fluorescent cells were analyzed using a flow cytometer at the indicated time points. The end-point was measured after expanding the population in DMEMselect. These results only illustrate the completion of the selection, and do not serve an analytical purpose. The figure shows the percentage population fluorescence of approximately 10,000 cells.

REFERENCES

1. Counter, C.M., Avilion, A.A., LeFeuvre, C.E., Stewart, N.G., Greider, C.W., Harley, C.B. and Bacchetti, S. (1992) Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.*, 11, 1921-1929.
2. Alattia, J.R., Matasci, M., Dimitrov, M., Aeschbach, L., Balasubramanian, S., Hacker, D.L., Wurm, F.M. and Fraering, P.C. (2013) Highly efficient production of the Alzheimer's gamma-secretase integral membrane protease complex by a multi-gene stable integration approach. *Biotechnol Bioeng*, 110, 1995-2005.
3. Bussow, K. (2015) Stable mammalian producer cell lines for structural biology. *Curr. Opin. Struct. Biol.*, 32, 81-90.
4. Zhang, J.Y. (2010) Mammalian Cell Culture for Biopharmaceutical Production. *Manual of Industrial Microbiology and Biotechnology, Third Edition*, 157-178.
5. Kojima, R., Scheller, L. and Fussenegger, M. (2018) Nonimmune cells equipped with T-cell-receptor-like signaling for cancer cell ablation. *Nat. Chem. Biol.*, 14, 42-49.
6. Scheller, L. and Fussenegger, M. (2019) From synthetic biology to human therapy: engineered mammalian cells. *Curr. Opin. Biotechnol.*, 58, 108-116.
7. Kim, J.Y., Kim, Y.G. and Lee, G.M. (2012) CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl Microbiol Biotechnol*, 93, 917-930.
8. Walsh, G. (2018) Biopharmaceutical benchmarks 2018. *Nat. Biotechnol.*, 36, 1136-1145.
9. Birch, J.R. and Racher, A.J. (2006) Antibody production. *Adv Drug Deliv Rev*, 58, 671-685.
10. Matasci, M., Baldi, L., Hacker, D.L. and Wurm, F.M. (2011) The PiggyBac transposon enhances the frequency of CHO stable cell line generation and yields recombinant lines with superior productivity and stability. *Biotechnol Bioeng*, 108, 2141-2150.
11. Elegheert, J., Behiels, E., Bishop, B., Scott, S., Woolley, R.E., Griffiths, S.C., Byrne, E.F.X., Chang, V.T., Stuart, D.I., Jones, E.Y. *et al.* (2018) Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins. *Nat. Protoc.*, 13, 2991-3017.
12. Dangi, A.K., Sinha, R., Dwivedi, S., Gupta, S.K. and Shukla, P. (2018) Cell Line Techniques and Gene Editing Tools for Antibody Production: A Review. *Front. Pharmacol.*, 9, 630.
13. Wurm, F.M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.*, 22, 1393-1398.
14. Southern, P.J. and Berg, P. (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *Journal of molecular and applied genetics*, 1, 327-341.
15. Mulsant, P., Gatignol, A., Dalens, M. and Tiraby, G. (1988) Phleomycin resistance as a dominant selectable marker in CHO cells. *Somatic cell and molecular genetics*, 14, 243-252.
16. Lai, T., Yang, Y. and Ng, S.K. (2013) Advances in Mammalian cell line development technologies for recombinant protein production. *Pharmaceuticals (Basel)*, 6, 579-603.
17. Simonsen, C.C. and Levinson, A.D. (1983) Isolation and expression of an altered mouse dihydrofolate reductase cDNA. *Proceedings of the National Academy of Sciences of the United States of America*, 80, 2495-2499.
18. Cockett, M.I., Bebbington, C.R. and Yarranton, G.T. (1990) High level expression of tissue inhibitor of metalloproteinases in Chinese hamster ovary cells using glutamine synthetase gene amplification. *Bio/technology*, 8, 662-667.
19. Kaufman, R.J. and Sharp, P.A. (1982) Amplification and expression of sequences cotransfected with a modular dihydrofolate reductase complementary dna gene. *J. Mol. Biol.*, 159, 601-621.
20. Bebbington, C.R., Renner, G., Thomson, S., King, D., Abrams, D. and Yarranton, G.T. (1992) High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. *Biotechnology. (N. Y.)*, 10, 169-175.
21. Noh, S.M., Shin, S. and Lee, G.M. (2018) Comprehensive characterization of glutamine synthetase-mediated selection for the establishment of recombinant CHO cells producing monoclonal antibodies. *Sci. Rep.*, 8, 5361.

22. Urlaub, G., Kas, E., Carothers, A.M. and Chasin, L.A. (1983) Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells. *Cell*, 33, 405-412.

23. Urlaub, G. and Chasin, L.A. (1980) Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc. Natl. Acad. Sci. U. S. A.*, 77, 4216-4220.

24. Reinhart, D., Damjanovic, L., Kaisermayer, C., Sommeregger, W., Gili, A., Gasselhuber, B., Castan, A., Mayrhofer, P., Grunwald-Gruber, C. and Kunert, R. (2019) Bioprocessing of Recombinant CHO-K1, CHO-DG44, and CHO-S: CHO Expression Hosts Favor Either mAb Production or Biomass Synthesis. *Biotechnol. J.*, 14.

25. Eagle, H., Barban, S., Levy, M. and Schulze, H.O. (1958) The utilization of carbohydrates by human cell cultures. *J. Biol. Chem.*, 233, 551-558.

26. Van Beers, E.H., Buller, H.A., Grand, R.J., Einerhand, A.W. and Dekker, J. (1995) Intestinal brush border glycohydrolases: structure, function, and development. *Crit Rev Biochem Mol Biol*, 30, 197-262.

27. Tian, C., Beeson, W.T., Iavarone, A.T., Sun, J., Marletta, M.A., Cate, J.H. and Glass, N.L. (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus Neurospora crassa. *Proc Natl Acad Sci U S A*, 106, 22157-22162.

28. Liu, Z., Chen, O., Wall, J.B.J., Zheng, M., Zhou, Y., Wang, L., Ruth Vaseghi, H., Qian, L. and Liu, J. (2017) Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci. Rep.*, 7, 2193.

29. Lin, J., Neo, S.H., Ho, S.C.L., Yeo, J.H.M., Wang, T., Zhang, W., Bi, X., Chao, S.H. and Yang, Y. (2017) Impact of Signal Peptides on Furin-2A Mediated Monoclonal Antibody Secretion in CHO Cells. *Biotechnol. J.*, 12.

30. Fang, J., Yi, S., Simmons, A., Tu, G.H., Nguyen, M., Harding, T.C., VanRoey, M. and Jooss, K. (2007) An antibody delivery system for regulated expression of therapeutic levels of monoclonal antibodies in vivo. *Mol. Ther.*, 15, 1153-1159.

31. Tastanova, A., Schulz, A., Folcher, M., Tolstrup, A., Puklowski, A., Kaufmann, H. and Fussenegger, M. (2016) Overexpression of YY1 increases the protein production in mammalian cells. *J. Biotechnol.*, 219, 72-85.

32. Tigges, M. and Fussenegger, M. (2006) Xbp1-based engineering of secretory capacity enhances the productivity of Chinese hamster ovary cells. *Metab. Eng.*, 8, 264-272.

33. Bojar, D., Fuhrer, T. and Fussenegger, M. (2019) Purity by design: Reducing impurities in bioproduction by stimulus-controlled global translational downregulation of non-product proteins. *Metab. Eng.*, 52, 110-123.

34. Majors, B.S., Betenbaugh, M.J., Pederson, N.E. and Chiang, G.G. (2009) Mcl-1 overexpression leads to higher viabilities and increased production of humanized monoclonal antibody in Chinese hamster ovary cells. *Biotechnol. Prog.*, 25, 1161-1168.

35. Dreesen, I.A. and Fussenegger, M. (2011) Ectopic expression of human mTOR increases viability, robustness, cell size, proliferation, and antibody production of chinese hamster ovary cells. *Biotechnol Bioeng*, 108, 853-866.

36. Tiwari, P., Misra, B.N. and Sangwan, N.S. (2013) beta -Glucosidases from the fungus trichoderma: an efficient cellulase machinery in biotechnological applications. *Biomed Res Int*, 2013, 203735.

37. Galazka, J.M., Tian, C., Beeson, W.T., Martinez, B., Glass, N.L. and Cate, J.H. (2010) Cellodextrin transport in yeast for improved biofuel production. *Science*, 330, 84-86.

38. Parisutham, V., Chandran, S.P., Mukhopadhyay, A., Lee, S.K. and Keasling, J.D. (2017) Intracellular cellobiose metabolism and its applications in lignocellulose-based biorefineries. *Bioresour. Technol.*, 239, 496-506.

39. Leong, D.S.Z., Tan, J.G.L., Chin, C.L., Mak, S.Y., Ho, Y.S. and Ng, S.K. (2017) Evaluation and use of disaccharides as energy source in protein-free mammalian cell cultures. *Sci. Rep.*, 7.

CONCLUSION

It is in the nature of engineering as a discipline, but also of engineers themselves to focus on real life applications, rather than abstract theoretical constructs. Standing on the solid foundation of basic research and mathematics and coupled with the intrinsic human urge to tinker, create and overcome challenges, engineers of all generations managed to devise inventions that seemed completely unthinkable in the past. The modern accomplishments in mechanical-, electronic- or biological-engineering keep changing people's lives to the better. Bioengineering has come a long way from the first manipulations of DNA(1) to where it is today. The possibility to use and change the genome of living cells shows a fascinating opportunity to, for example, fight diseases. Especially synthetic biology is a discipline that is promoting great progress to changing the pharmaceutical industry and medicine in particular(2). With conceptually novel approaches, such as gene therapy or cell-based therapy, we envision that synthetic biology in the near future can solve some of the medical problems that to date remain unsolved(3). As synthetic biologists design evermore intricate circuits in mammalian cells to be used as therapeutic designer cells, the potential is unveiling(4). Creating such complex systems currently is still relying on a trial and error-based approach, as many biochemical interactions are difficult to precisely predict(5). This difficulty is made worse when non-native components are used, as they might function differently in another organism or not at all. Nevertheless, heterologous gene expression can likely be seen as one of the core pillars of synthetic biology. Enabling engineered cells to display a functionality that was previously only known in different organisms shows a huge potential for industrial and therapeutic applications. Although remarkable advances(6,7) have been made in the adaptation of enzymes from one organism to another, often the fastest and most successful approach is to test enzymes from different sources that catalyze the same reaction. In this thesis we show that enzymes from two different fungi are functioning in mammalian cells, as well as one enzyme from the plant *Celosia cristata.* Fungi, especially *Amanaita muscaria* and *Neurospora crassa*, should be considered viable sources for heterologous enzymes for the use in mammalian cells. It should be noted, that DOPA-dioxygenases from *Beta vulgaris*(8) and *E. coli*(9) did not show any signs of functional activity in our tested cell lines.
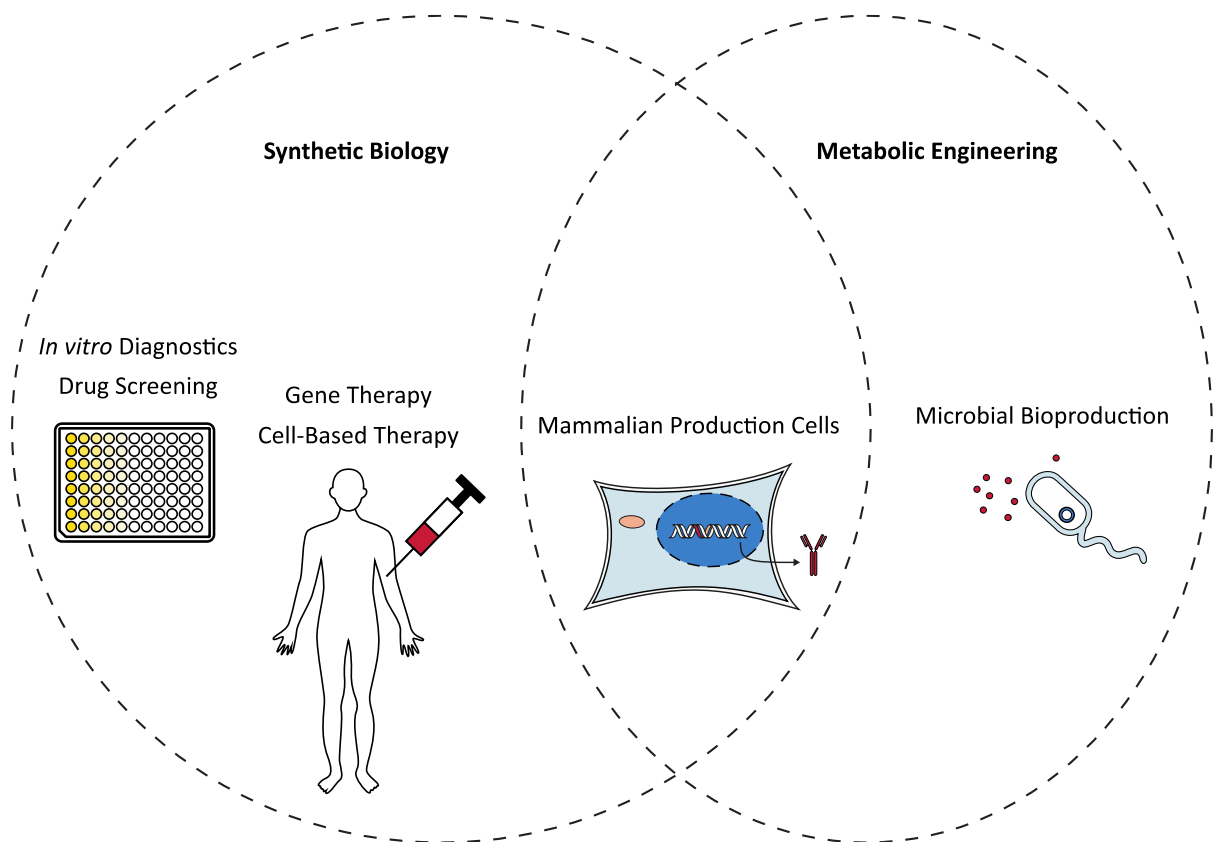
Figure 1: Examples of industrial applications of synthetic biology and metabolic engineering. Synthetic and metabolic engineering are widely used in the biopharmaceutical industry. With synthetic biological methods cell-based screening platforms (10) and diagnostic tools(11) were developed, as well as the recent approaches in creating smart cell- or gene therapy-based therapeutics. Additionally, synthetic biology was applied to create smart mammalian protein production cell lines that, for example allow the trigger-inducible cargo production(12). Metabolic engineering is used to enhance the production capacity or decrease the waste-product formation in microbial- or mammalian systems.

It is crucial for most approaches in engineering biological systems to acquire information about what exactly is happening inside the cell. One essential tool for this is, as discussed before, the technology of reporter systems for gene expression. Currently, in a research laboratory setting multiple different reporter systems are generally used, each offering a certain analysis method for the task at hand. For example, if a quick cursory observation about a tested system is required, an enzymatic reporter may be used in the supernatant or if more precise population dynamic is relevant, a fluorescent protein combined with flow cytometry

is a good candidate. A more universal reporter that offers simultaneously a quick whole-population analysis and a more precise dynamic analysis would be advantageous.

In chapter I we introduced a novel type of reporter system based on the direct, *in cellula* production of a dye. The detection and quantification of gene expression with this system is intrinsically simple. No need for time-consuming assays to assess the level of gene expression in a tested setup. Simply by sampling the supernatant, placing it in a standard fluorescent reader and measuring the fluorescence, gene expression can be quantified. We additionally displayed the ease of use by measuring reporter production using a cell phone camera. In particular, this illustrates how well the dye can be detected, which allows first cursory assessment of the functionality of a tested system by the naked eye. Another advantage of our system is that it offers multiple measurement modes simultaneously. It can either be directly measured in the culture supernatant, yielding whole-population data or with a flow cytometer or microscope, resulting in single cell data. This might decrease the number of genetic constructs that need to be created in bioengineering. Whether this system will be used for actual development of synthetic biological circuits remains to be seen, but the ease of measurement might also find its niche in industrial drug screenings. The high-throughput screening of functional small molecules in cell culture often relies on simple methods to assess changes in gene expression(13). In our system, gene expression can be assayed directly in the culture plate continuously without any additional reagents, potentially making these approaches faster and less expensive.

Of course, our betaxanthin production system still has room for optimization and, at the very least, could inspire the development of other direct small molecule reporter systems. The main point in improving the system likely has to do with the origin of the heterologous enzyme DOPA-dioxygenase from *Amanita muscaria*. It might be possible that the enzyme only functions sub-optimally in mammalian cells. Or that the enzyme in general has a low turn-over rate for the formation of betalamic acid, as this does not need to be a very fast process in the native fly agaric. Two different strategies come to mind to improve the reaction rate of the enzyme in mammalian cells, rational design and random screening. With a rational approach, the enzyme stability, solubility or localization could be adjusted. Using a random screening strategy might yield point mutations that potentially increase the enzymatic activity. The second optimization that comes to mind for the system as a whole is a possible broadening of the range of colors the system is capable of producing. In nature, plants

producing betalains can display a whole range of colors between red and yellow. The red dye belongs to the closely related betacyanin class and might be produced by an additional pathway starting from L-DOPA. The issue is, that the spontaneous color production is not highly specific. Betalamic acid would react with amino acids in the cytosol and betacyanin precursor cyclo-DOPA alike, thus making the color production process concentration-dependent. This issue might be solved by compartmentalization of the reaction in specific organelles.

Similarly to how synthetic biology is impacting current medicine, albeit in a less revolutionizing fashion, are advances in metabolic engineering impacting the biotechnological industry(14). Using proteins produced in cell culture, in particular antibodies, is well established and cancer or autoimmune disease patients world-wide rely on these products. Increasing the production capacity and decreasing the production cost should hopefully make these products more affordable, which would foster their use in less developed places. As previously mentioned, many therapeutic proteins are primarily produced by mammalian cells stably producing the cargo of choice, due to the importance of appropriate glycosylation patterns. With the development of CRISPR/Cas9- or transposase-based integration methods a high genomic cargo integration frequency can be assured. Selecting the highest-producing cells afterwards is still performed using a knockout approach for essential metabolic pathways. The used knockout cell lines were shown to be sub-optimally suited for high production titers(15) and a different approach would be required that is not based on the knockout of a gene but on the additive gain of the ability to utilize a different nutrient source. As no selection-permitting knockout cell lines have to be used, such a system could be applied to cells that were optimized for protein production and, as the selection agent is simultaneously an essential nutrient, would allow the selection pressure to be applied for the complete duration of pharmaceutical production. In this case the selection agent is not a toxic chemical, which might reduce downstream processing requirements. Also, as modern bioreactors are up to 600 m^3(16), every additional reduction in waste treatment requirements is beneficial. From an environmental point of view in the age of wide spread antibiotic resistance, the large-scale use of antibiotics seems rather irresponsible if other options exist, which could be part of the reason why antibiotic selection systems are rarely used in industry.

Conclusion

In chapter II we introduced a novel, additive approach for the generation of stably transfected cell lines called CelloSelect. The cellobiose-utilization pathway uses two heterologous proteins to endow cells with the ability to use cellobiose as glucose source and, if glucose is not present in the culture medium, to survive while non-producing cells starve. This selection system does not require toxic chemicals or knockout cell lines and we showed that selected cells could simply be fed with cellobiose. In a model bioproduction experiment we managed to produce the therapeutic protein erythropoietin (EPO) and we showed that the produced amount is comparable regardless of whether the cells were fed with glucose or cellobiose. This clearly indicates that the selection pressure could be applied during a complete production run. In this proof-of-concept study, we clearly showed the potential for industrial applications. Industrial cell lines and selection strategies have come a long way and antibody production capacity keeps increasing. It is difficult to compare the industrially optimized production systems to proof-of-concept model laboratory setups. Thus, we hope that our project inspires similar work in the development of engineered metabolic pathways into mammalian cells, in particular for stable cell line generation, and that the system can be adapted to an industrial setting.

To potentially apply the presented system in an industrial setting, different strategies come to mind. Large-scale bioprocessing in a bioreactor is normally performed using an adapted suspension cell line (usually a CHO-derived cell line). Our system was developed for a standard adherent CHO-K1 adherent cell line. Application of this system to a suspension cell line requires the formulation of specialized suspension media without glucose. One advantage that our system offers is, that it gives the engineer the option to use a suspension cell line that is already optimized for production capacity. In comparison, industrial cell lines are generally based on GS or DHFR knockout cell lines that might not be optimal for production of a protein of choice. A second strategy to improve the production capacity would be to further engineer the general features of the genetic constructs. Of course, our system is not limited to the current setup using a tandem expression vector having all three genes expressed at the same rate. Lowering the expression strength of one or both of the cellobiose utilization genes (e.g. using a weak internal ribosomal entry site (IRES)(17) instead of the P2A sequence) might yield an even higher integration frequency or lower the metabolic burden the two proteins put on the cell. A third strategy that may increase production of cellobiose selected cells is to apply a specific inhibitor for the cellobiose-splitting β-glucosidase(18) to

potentially amplify the CelloSelect cassette copy number, including the cargo gene. The same effect could likely be achieved by continually decreasing the cellobiose concentration during the selection process or using enzymes with a lower turn-over rate. Lastly, it is a goal to lower the production of toxic waste products (e.g. lactate) of production cell lines by limiting the glucose uptake or by using a sugar different than glucose for feeding. Waste product-formation in mammalian bioreactor production processes is a major obstacle. For example, a high lactate concentration was shown to quickly damage cells and slow growth[19]. It might be worth analyzing the waste product formation in CelloSelect cells fed with cellobiose, which might be superior to normal production cells fed with glucose[20]. Establishing novel selection systems based on different disaccharides could open up a completely new branch of optimized biopharmaceutical production. An additional factor to consider might be the use of agriculture waste products for bioreactors. Cellobiose can be enzymatically created from cellulose, a polysaccharide that cannot be digested by mammalian cells. Using cellulose for bioprocessing was extensively studied for biofuel production, where researchers managed to create cellulose-consuming yeast[21]. Combining such an approach with cellobiose-fed cells might lead to mammalian bioreactors fed with only marginally processed cellulose, possibly reducing cost or making culture medium manufacturing compete less with food production.

In summary, this PhD thesis contributes to the field of mammalian synthetic biology with the development of two novel tools, based on the engineering of new metabolic pathways. We successfully developed the systems, tested and characterized them under model conditions and offer strategies for improvement, as well as for the application in advanced systems. We hope to see the systems getting attention and leading to new discoveries, biological tools, or platforms to advance synthetic biology, metabolic engineering and bioprocessing.

REFERENCES

1. Lederberg, J. (1994) The Transformation of Genetics by DNA - an Anniversary Celebration of Avery, Macleod and Mccarty (1944). *Genetics*, 136, 423-426.
2. Bueso, Y.F. and Tangney, M. (2017) Synthetic Biology in the Driving Seat of the Bioeconomy. *Trends Biotechnol.*, 35, 373-378.
3. Caliendo, F., Dukhinova, M. and Siciliano, V. (2019) Engineered Cell-Based Therapeutics: Synthetic Biology Meets Immunology. *Front Bioeng Biotech*, 7.
4. Ye, H. and Fussenegger, M. (2014) Synthetic therapeutic gene circuits in mammalian cells. *FEBS Lett.*, 588, 2537-2544.
5. Purcell, O., Jain, B., Karr, J.R., Covert, M.W. and Lu, T.K. (2013) Towards a whole-cell modeling approach for synthetic biology. *Chaos*, 23.
6. Hunter, M., Yuan, P., Vavilala, D. and Fox, M. (2019) Optimization of Protein Expression in Mammalian Cells. *Curr Protoc Protein Sci*, 95, e77.
7. Rosano, G.L. and Ceccarelli, E.A. (2014) Recombinant protein expression in Escherichia coli: advances and challenges. *Front. Microbiol.*, 5.
8. Gandia-Herrero, F. and Garcia-Carmona, F. (2012) Characterization of recombinant Beta vulgaris 4,5-DOPA-extradiol-dioxygenase active in the biosynthesis of betalains. *Planta*, 236, 91-100.
9. Gandia-Herrero, F. and Garcia-Carmona, F. (2014) Escherichia coli protein YgiD produces the structural unit of plant pigments betalains: characterization of a prokaryotic enzyme with DOPA-extradiol-dioxygenase activity. *Appl Microbiol Biotechnol*, 98, 1165-1174.
10. Sedlmayer, F., Hell, D., Muller, M., Auslander, D. and Fussenegger, M. (2018) Designer cells programming quorum-sensing interference with microbes. *Nat. Commun.*, 9.
11. Auslander, D., Eggerschwiler, B., Kemmer, C., Geering, B., Auslander, S. and Fussenegger, M. (2014) A designer cell-based histamine-specific human allergy profiler. *Nat. Commun.*, 5.
12. Auslander, D., Auslander, S., Charpin-El Hamri, G., Sedlmayer, F., Muller, M., Frey, O., Hierlemann, A., Stelling, J. and Fussenegger, M. (2014) A synthetic multifunctional mammalian pH sensor and CO2 transgene-control device. *Mol. Cell*, 55, 397-408.
13. Zhang, Z.Y., Guan, N., Li, T., Mais, D.E. and Wang, M.W. (2012) Quality control of cell-based high-throughput drug screening. *Acta Pharm Sin B*, 2, 429-438.
14. Richelle, A. and Lewis, N.E. (2017) Improvements in protein production in mammalian cells from targeted metabolic engineering. *Curr Opin Syst Biol*, 6, 1-6.
15. Reinhart, D., Damjanovic, L., Kaisermayer, C., Sommeregger, W., Gili, A., Gasselhuber, B., Castan, A., Mayrhofer, P., Grunwald-Gruber, C. and Kunert, R. (2019) Bioprocessing of Recombinant CHO-K1, CHO-DG44, and CHO-S: CHO Expression Hosts Favor Either mAb Production or Biomass Synthesis. *Biotechnol. J.*, 14.
16. Noorman, H. (2011) An industrial perspective on bioreactor scale-down: What we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnol. J.*, 6, 934-943.
17. Koh, E.Y.C., Ho, S.C.L., Mariati, Song, Z.W., Bi, X.Z., Bardor, M. and Yang, Y.S. (2013) An Internal Ribosome Entry Site (IRES) Mutant Library for Tuning Expression Level of Multiple Genes in Mammalian Cells. *PLoS One*, 8.
18. Tiwari, P., Misra, B.N. and Sangwan, N.S. (2013) beta -Glucosidases from the fungus trichoderma: an efficient cellulase machinery in biotechnological applications. *Biomed Res Int*, 2013, 203735.
19. Buchsteiner, M., Quek, L.E., Gray, P. and Nielsen, L.K. (2018) Improving culture performance and antibody production in CHO cell culture processes by reducing the Warburg effect. *Biotechnology and Bioengineering*, 115, 2315-2327.
20. Leong, D.S.Z., Tan, J.G.L., Chin, C.L., Mak, S.Y., Ho, Y.S. and Ng, S.K. (2017) Evaluation and use of disaccharides as energy source in protein-free mammalian cell cultures. *Sci. Rep.*, 7.

21. Tsai, S.L., Oh, J., Singh, S., Chen, R.Z. and Chen, W. (2009) Functional Assembly of Minicellulosomes on the Saccharomyces cerevisiae Cell Surface for Cellulose Hydrolysis and Ethanol Production. *Applied and Environmental Microbiology*, 75, 6087-6093.

ACKNOWLEDGEMENTS

I would like to thank…

…Prof. Dr. Martin Fussenegger, for giving me the opportunity to work in his group and for all the helpful advice and support I received during my PhD

…my referees Prof. Dr. Sven Panke & Prof. Dr. Kobi Benenson, as well as Prof. Dr. Niko Beerenwinkel as chair, for their time and all the great questions and comments

… my co-authors, for the fruitful collaborations

… Simon, Marius, Ferdi & David, for showing me how to science

…all the current and former lab members, in particular Krzysztof, Tolle, Bozho, Pratik, Ana, Nik, HJ, Richard, Elsa, Niela & Laila, for all the help with the projects, the fun times we had, and putting up with all the fighting and the discussions

… Tobi, Fuchs, Leo, Viktor & Adi, for the many amusing hours outside the lab and all the help inside

…all my friends from Weinfelden, Riehen, Basel and beyond, amongst many others Su, Joe, Matthias, Sämmi, Steffi, Yannick, Jordi, Tobi, Sarah, Cheran, Dave, Silas, Päsci, Ramon & my HK friends, for listening to my whining and still spending time with me over the years

…Belinda, for loving me

…my family, especially my parents Beatrice and Markus, my sister Tanja, my grandma Irma & the Heizes, for believing in me and being there for me.

Thank you!

CURRICULUM VITAE

Pascal Stücheli

Education

2014 – 03.2020 | **Doctor of Sciences in Biotechnology**

Swiss Federal Institute of Technology (ETH)

Thesis entitled: Engineering of Metabolic Pathways in Mammalian Cells for the

Biosynthesis and Utilization of Exogenous Small Molecules

2012-2013 | **Master of Science ETH in Biotechnology**

Swiss Federal Institute of Technology (ETH)

Research Focus: Synthetic biology

2008-2012 | **Bachelor of Science ETH in Biotechnology**

Swiss Federal Institute of Technology (ETH)

Main Subjects: Chemistry and biotechnology

Work Experience

2014 | **Research Assistant**

Swiss Federal Institute of Technology (ETH)

Skills

Language Skills: German (native), English (excellent, TOEFL iBT 111/120), French (basic)

Computer Skills: Microsoft Office Suit, Graphpad Prism, Adobe Illustrator & Photoshop, FlowJo

Programming Skills: Matlab (proficient), Python (advanced), C++/C#/Pascal (basic)

Publications

Ausländer S., **Stücheli P.**, Rehm C., Ausländer D., Hartig J.S., Fussenegger M. A general design strategy for protein-responsive mammalian riboswitches., *Nat. Methods*, 2014.

Tolle F., **Stücheli P.**, Fussenegger M., Genetic circuitry for personalized human cell therapy., *Curr. Opin. Biotechnol.,* 2019

Stücheli P., Sieber S., Fuchs D. W., Scheller L., Strittmatter T., Saxena P., Gademann K., Fussenegger M., Genetically encoded betaxanthin-based small-molecular fluorescent reporter for mammalian cells., *Nucleic Acids Res.*, 2020

Stücheli P., Ausländer S., Ausländer D., Schönenberger P., Hürlemann S., Fussenegger M., CelloSelect – a synthetic cellobiose metabolic pathway for selection of stable transgenic CHO-K1 cell lines., in preparation, 2020