Refactoring yeast signalling pathways for tuneable extracellular biosensing

William Shaw

Submitted for the Degree of Doctor of Philosophy

Supervised by Dr Tom Ellis

Imperial College London

Department of Bioengineering

The Imperial College Centre for Synthetic Biology

October 1st, 2018

Abstract

G protein-coupled receptors (GPCRs) present themselves as an attractive class of membrane protein for use within eukaryotic whole-cell biosensors due to their responsiveness to an extensive and diverse range of ligands. Despite being a comparatively simple organism, *Saccharomyces cerevisiae* has the complex machinery necessary for coupling heterologous GPCRs to a cellular output, thus enabling this highly amenable chassis with the sensing abilities of higher eukaryotes. Although examples are beginning to emerge within the field of synthetic biology, GPCR signalling in yeast remains an underutilised foundation for the creation of new biosensors. Current designs are often limited due to a mismatch between the input concentrations to which these biosensors respond and the application requirements.

Here we present a new platform for rationally tuning GPCR-based biosensors in *Saccharomyces cerevisiae* to address these limitations. While previous efforts to manipulate GPCR signalling in yeast have involved a top-down approach or overlaying complexity, we sought to identify the minimal requirements to achieve fully tuneable behaviour from the bottom-up. Using genome engineering, we constructed an insulated, modular GPCR signal transduction system to study how the response to stimuli can be predictably tuned using synthetic tools. By systematically refactoring the system, we delineated the contributions of a minimal set of components, identifying robust and straightforward design rules for tuning the sensitivity, leakiness, and signal output. Using these principles, we then established novel community-based approaches for tuning the final dose-response property – the Hill slope. This work enables the development of diverse yeast biosensors that are well-suited to their applications demands, while also providing a framework to guide the reprogramming of GPCR-based signalling in more complex systems.

Declaration

I hereby certify that everything included in this thesis is my own work, and in the instances where work has been used from other sources, it is acknowledged appropriately.

Acknowledgements

First and foremost, I would like to thank Tom and Mark for being amazing supervisors. I understand I can be quite hard to keep grounded, and I truly appreciate your unyielding commitment to keeping me on track and for helping me develop as a researcher. I hope to go forward in my career following the examples you have set and the lessons you have taught me along the way.

I would like to thank all the wonderful collaborators I have had the privilege in working with. Thank you to Graham for making so many aspects of this project possible and offering countless wise advice. Thanks to Hitoshi and Jack for your incredible efforts realising the computational side of this work. And thank you to David and Niklas for the many stimulating discussions.

To all the members of the Centre for Synthetic Biology who have given their time to helping me - Rochelle, David, Ciaran, and so many others - thank you for all your support. And thank you to the rest of the Centre for providing such a friendly environment to be part of. I would also like to thank the honorary member of our lab, Kealan, for always being around for a chat or a break when needed. And thank you to Edina and Song for holding the labs together.

I especially want to thank all the awesome members of the Ellis lab, both past and present, that have made my time at Imperial such a great experience in and outside the lab. The last few years have genuinely been some of the best, and this would not have happened without you all. Thanks to Ali for mentoring me during the early stages of my PhD and for showing me the ropes. Thank you to Ben for your continued guidance. Thanks to Olivier for not taking the pranks too far. Thanks to Robert Chen for showing me the ways of the YTK. And thank you to Charlie for being such a good friend throughout.

Finally, I would like to thank my friends and family who have supported me these last four years. I could not have done it without you all.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Summary

In a sustainable bioeconomy, microbial cell factories will convert inexpensive renewable feedstocks into many commodities and high-value chemicals. Advances in synthetic biology are making it possible to generate vast phenotypic libraries to identify the optimal cell factories for producing these products. However, our ability to design and build far exceeds our capacity to test for variants with optimal performance, slowing our progress towards this sustainable future. Yeast GPCR-based biosensors represent a possible solution to this current bottleneck in metabolic engineering projects. Indeed, a number of recent studies have demonstrated their potential for diverse biosensing applications, albeit with some fundamental limitations that will need addressing to meet the demands of real-world applications. In this section, we discuss the field of genetically encoded biosensors and the emergence of yeast GPCR-based biosensors, their limitations, and possible solutions to these limitations.

1.2 Background

1.2.1 Towards a sustainable bioeconomy

Some of the biggest challenges facing the world today include climate change, environmental degradation, food shortages, water crises, and the spread of infectious disease, caused by our dependence on fossil fuels, waste production, and over population[1]. With a global economic model that is intrinsically linked to energy consumption and a population that is only set to increase, it is clear we are trapped in a system that is unsustainable[2]. While we urgently need to address the way we live as a society, we also need to be realistic in coming up with new technologies and strategies that will provide us with securities of today, tomorrow[3].

Synthetic biology, with its aims of reprogramming biology using principles borrowed from engineering, is a field expected to address such real-world issues[4–6]. By engineering biological systems, we can harness cellular metabolism for the industrial production of many commodities and high-value chemicals from renewable feedstocks[7,8]. This shift from an oil-based economy to a "bioeconomy" will not only reduce our environmental burden but also lead to sustainable value chains[9,10]. Already, the bioeconomy is shaping up to be a profitable sector, with revenues estimated to have grown on average of >10% each year over the past decade[11] (**Figure 1**). These advances are due to our increasing knowledge of engineering biological systems, aided by foundational innovations in the "design-build-test" development cycle that underpins much of synthetic biology[12].

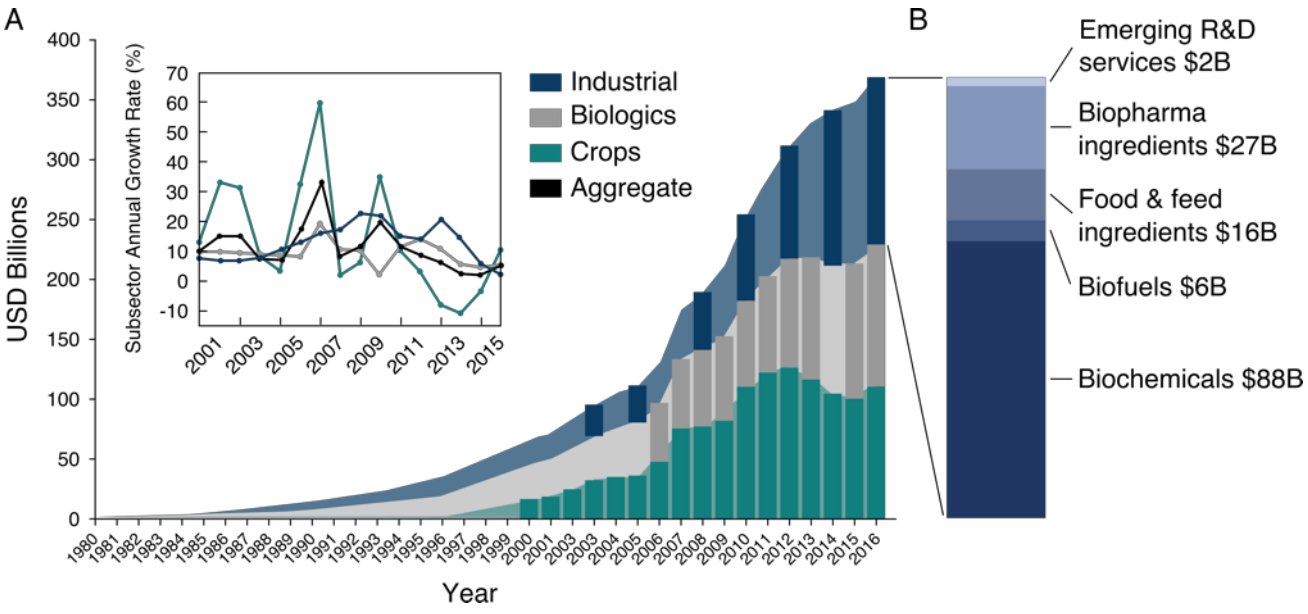


Figure 1. Biotech revenues in the US. (*A*) Biotech revenues from industry, biologics, and GM crops have grown on average >10% each year since the turn of the century, with the most stable increases from the industrial sector (*B*). Figure adapted from Carlson[11] and Bioeconomy Capital under the Creative Commons image licence.

A significant amount of progress has been made within the first two phases of the development cycle. The “design” step has benefited from substantial efforts to standardise, abstract, and characterise biological parts, with innovative community efforts in academia, such as the International Genetically Engineered Machine (iGEM) competition[13], and non-profit organisations, such as Addgene[14], encouraging the open sharing of DNA constructs. Powerful bioinformatics tools are enabling us to mine our newfound data-richness for new parts[15,16]; a direct consequence of the exponential decrease in the cost of sequencing[17–19]. Suites of computational tools have then been developed to seamlessly bridge the gap to the next phase in the development cycle[20,21].

The “build” step has also seen an explosion of advances in the past decade. The impact of CRISPR/Cas9 and multiplexed genome editing tools are enabling biologists to engineer organisms with unprecedented precision and ease[22–24]. The progressive reduction in the cost of oligonucleotide synthesis and innovations in DNA assembly has allowed us to expand natures diversity and explore larger biological hypotheses[25–27]. Designer genomes can now be built from scratch[28,29] or recoded to suit out needs[30]. Innovations in cloning techniques, such as the Golden Gate assembly method[31], have inspired a number of modular plasmid toolkits for the assembly of complex genetic constructs from reusable parts (e.g. MoClo[32], YTK[33], EcoFlex[34], CIDAR[35], and Plant Tool Kit[36]), making it easier and easier for biologists to engineer a wide diversity of organisms[37].

Finally, central facilities for genetic engineering, or “biofoundaries”, are consolidating the foundational technologies in the “design” and “build” steps, integrating the genetic tools with software, automation, and manufacturing processes to streamline the engineering of biological systems[12,38,39]. Although the automation process is highlighting gaps in the first two phases of the development cycle, the throughput to design and build within a single project is incredibly high and is enabling researchers to explore a more significant proportion of the biological solution space to any particular problem[38].

The ability to design and build vast libraries of genetic diversity is of particular importance to metabolic engineering. Strategies based on rational design remain difficult as important pathway information is often unavailable and these approaches often fail to take a holistic view of cellular metabiolism[40,41]. Biosynthetic pathways also tend to be generated from parts sourced from diverse organisms which may have unknown context dependency within the host chassis[42]. Furthermore, cells have evolved robust metabolic networks hard-wired to resist diverting resources and often require extensive genome engineering to bypass[42]. Because of the many complex variables and unknowns, engineering biosynthetic pathways is incredibly challenging, and so metabolic engineers have begun to shift towards an engineering paradigm based on the principles of Darwinian evolution[42–44]. Using combinatorial approaches, billions of genetic variants can be used to scan the genetic landscape, rather than rationally work towards an optimum from the bottom-up[42–44].

This top-down approach is achieved by defining the outcome of the complete system, such as the production yield of a particular metabolite, without specifying the parameters of the individual components. The optimal combination of genetic parts can then be determined without a complete understanding of the individual contribution of each component[42–44]. However, this strategy imposes a need for efficient screening methods to identify individuals with the desired phenotype[45], which is beginning to highlight our woeful inadequacy to test chemical entities in any meaningful volume[38].

Identification of most microbially produced chemicals currently relies on conventional chromatography-based methods, such as gas or liquid chromatography coupled with mass spectrometry (MS). Although advances are being made to the throughput of MS-based methods[46,47], they are still limited to the processing of up to 10^3 samples per day, creating a severe bottleneck in the development cycle (**Figure 2**)[38,43]. Frequently, we are selecting projects based on easy-to-read colourimetric end products, such as carotenoid or violacein production[48]. The field of metabolic engineering has, therefore, often been criticised as a collection of demonstrations, rather than a systematic practice of engineering towards a chosen compound of value[49]. Flag-ship projects which do break the mould, such as the microbial production of the anti-malarial drug precursor artemisinic acid, often represent decades of man-hours to produce economically viable titres[50,51].

New and promising strategies exist to reduce the size of these vast libraries, such as statistical model-based multivariate modular metabolic engineering (MMME; also known more broadly as design of experiments (DoE)), that employ algorithms to efficiently sample multidimensional gene expression

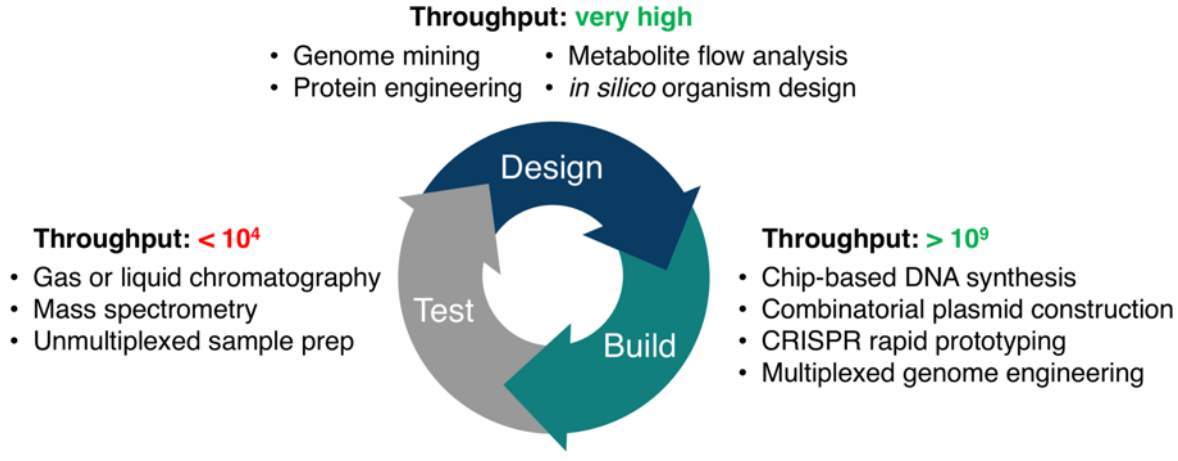


Figure 2. Throughput in the metabolic engineering "design-build-test" development cycle. Progress in metabolic engineering proceeds through an iterative "design-build-test" development cycle. Our current capacities for designing and building far exceed our ability to screen the microbial production of chemical entities, creating a bottleneck for top-down approaches to engineering biosynthetic pathways. The gold standard of metabolite measurement, mass spectrometry, is limited to numbers approaching 10^4 using the best equipment and an optimised sample prep workflow[46]. More commonly fewer than 10^3 measurements can be made per day. (Figure adapted from Rogers *et al.*[38]).

space and identify the desired overproduction phenotype[40,52,53]. MMME strategies reduce the number of variants to a fraction of the possible combinatorial space, making it practical to screen with our current throughput[41]. However, although this strategy does not require an in-depth knowledge of the metabolism, it does require a working pathway producing measurable levels of the end product[41]. Therefore, while this may be a powerful approach for improving the yields from a previously characterised pathway, it remains an inappropriate approach for the development of novel pathways which may be using many variants of previously uncharacterised genes. For such pathways, a genuinely combinatorial approach using billions of genetic variants may still be necessary.

It is clear that we lack a standard, universal principle for the engineering of biosynthetic pathways[41]. As these pathways are predicted to make up such a large segment of the rapidly expanding "bioeconomy"[11], we are in desperate need of innovations and foundational technologies to increase the throughput during the test phase of the development cycle to accelerate all approaches to metabolic engineering.

1.2.2 Biosensors and their application in metabolic engineering

Genetically encoded biosensors (referred to simply as "biosensors" hereafter) represent a revolutionary technology that has shown a potential to address this shortcoming[43–45,54]. Biosensors provide *in vivo* monitoring of cellular metabolism by responding to a molecule of interest, often with high specificity and sensitivity. When this activity is coupled to a reporter gene such as GFP or a selection marker (e.g. antibiotic resistance), biosensors offer the capability for mass library screening via high-throughput techniques such as fluorescence-activated cell sorting (FACS) or selection via cell survival (**Figure 3A**)[43,45]. These two methods can then be used to enrich cell populations for productive cells at a capacity approaching the limits of the first two steps in the development cycle. Enriched populations can then be characterised by deep sequencing of the entire population to uncover genetic trends, or single cells can be individually characterised at a lower throughput more amenable to techniques such as LC-MS[44].

Alternatively, a two-cell screening system can be employed, where the biosensing element has been separated from the producer strain (**Figure 3B**). Although the throughput of this technique is limited compared to the single-cell system, decoupling production from sensing has benefits[55]: (i) preparation of samples can ensure metabolite concentrations are always within the linear range of the biosensor, allowing it to report on production as titres from the producer strain improve, (ii) random, genome-wide mutagenesis strategies can be applied to the producer strain without affecting biosensor function

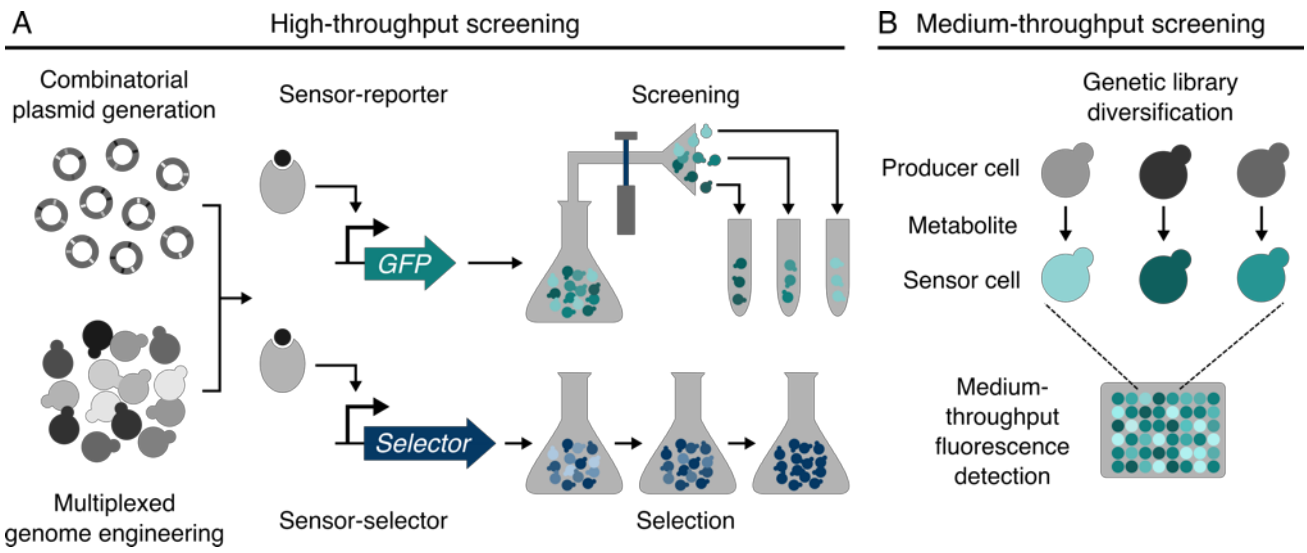


Figure 3. Application of biosensors to high- and medium-throughput screening. (*A*) High-throughput screening and selection in a producer-sensor strain. Large libraries of genetic diversity can be created from the combinatorial assembly of genetic constructs and multiplexed genome engineering. Sensor-reporter or sensor-selector combinations can then be used for high-throughput screening via FACS or by applying a selection to enrich cell populations with the desired phenotype. (*B*) Medium-throughput screening using a two-cell system. Samples prepared from producer cells can be applied to the sensor cell which are then measured using a medium-throughput detection apparatus such as a plate reader or flow cytometer.

or creating cheater cells, (iii) a suite of biosensing tools can be generated for a single organism and used to screen diverse industrially relevant strains which may not have the same tools at their disposal, and (iv) extracellular production can be separated from intracellular production if the secreted yield of a metabolite is of interest. There is also the future potential for high-throughput screening of producer microbes using FACS by encapsulating the producer and sensor cells[56]. Such advances would boost the throughput of this approach towards the single-cell system, while maintaining the benefits of the two-cell system, and will likely be a powerful strategy in the future[55].

1.2.3 Other applications of biosensors

Although biosensors are poised to revolutionise metabolic engineering, their application is not limited to development of biosynthetic pathways. Many of the attributes that make biosensors useful for screening and selection of metabolites also makes them useful for a diverse range of applications[54]. Once created, biosensors offer an inexpensive, rapidly deployable diagnostic that can be coupled to an easy-to-measure output. The untethering of analysis from specialised laboratory equipment makes them particularly attractive as analytics in the healthcare and environmental setting, where a point-of-use strategy can save significant amounts of time and money[57]. Furthermore, their coupling to biological processes makes them suitable as *in vivo* medical diagnostics, which could provide real-

time measurement of multiple pathological conditions and respond with an appropriate therapeutic programme according to the individual patient's disease state[58].

Real-time monitoring by integrating biosensors into closed-loop circuits will also play a significant role in the design of microbes more robust to unpredictable environments. By sensing the intra- and extracellular environment, cells can be programmed to autonomously correct themselves by appropriately up- or downregulating a particular process within a synthetic network[59–61]. Being able to sense a greater diversity of states will allow us to build designer microbes able to thrive in more diverse environments with greater predictability[60,61].

Finally, as synthetic microbial communities gain popularity as a method for distributing biological tasks, biosensors will play a significant role in the coordination of population stability and dynamics[62,63]. Communication systems, such as the extensively studied lux system from *Vibrio fischeri*, are being extensively characterised to expand our capacity for engineering consortia[64,65]. Although still highly proof-of-principle, complex microbial communities with defined social interactions are beginning to emerge[66].

1.2.4 Biosensor architecture

Biosensor design and construction, like many areas in synthetic biology, is a multidisciplinary endeavour, often requiring experience from diverse fields, such as affinity chemistry, protein engineering, molecular biology, nucleic acid molecular dynamics, and material science[54]. At their most basic, biosensors are composed of two parts: the sensitive element, which interfaces with the target ligand and sets the detection thresholds, and the transducer module, which processes the input from the sensitive element and generates a cellular response through gene regulation[54,67] (**Figure 4A**).

The input and output within this process have an enormous diversity of possible configurations. Target inputs can range from stimuli such as photons[68] and single atoms, such as calcium[69], through the biological spectrum to whole proteins, such as tumour necrosis factor[70]. The output can vary from fluorescence, luminescence, enzymatic activity, colourimetric, generation of an electrical current, or integration with synthetic gene networks[54,71]. To achieve the desired response between the input and output, the two basic sensing modules must be finely balanced[67]. This is achieved by engineering specificity and modularity into the biosensor at the transcriptional, translational, and post-translational levels[67].

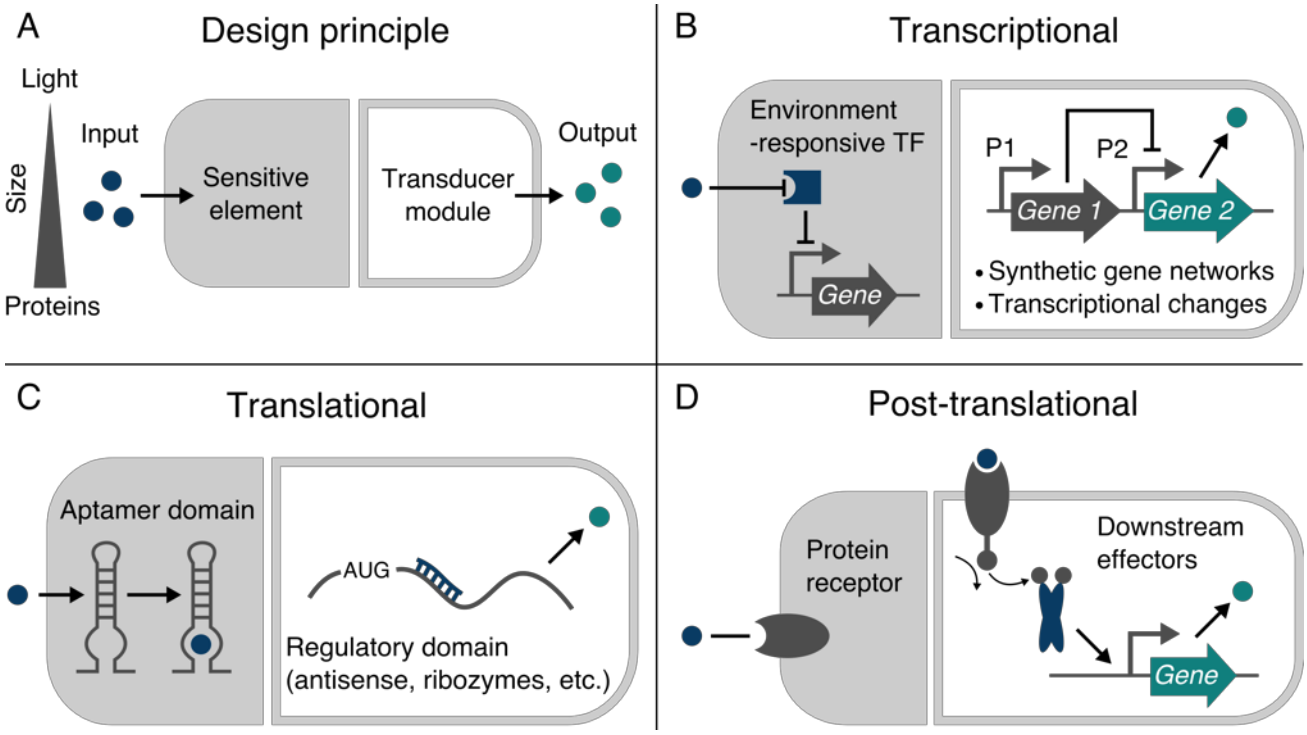


Figure 4. Biosensor architecture. At their most basic, biosensors are composed of two sensing modules; the sensitive element and the transducer domain. The processing can act through transcriptional, translation or post-translational events, all ultimately resulting in the up- or downregulation of a desired gene. Figure adapted from Khalil and Collins[67].

1.2.5 Transcriptional biosensing

As the first stage in gene expression, transcription offers a direct platform for generating a cellular response from an environmental perturbation. Consequently, promoters, RNA polymerases, transcription factors, and any other component of the transcription machinery all serve as an engineerable module for creating transcriptional biosensors[67]. Although biosensors based on RNA polymerase have been explored[72], the vast majority of transcriptional biosensors have focused on environmentally-responsive promoters or transcription factors (TFs) (**Figure 4B**), such as the highly-studied and extensively used promoters of the bacterial *lac*, *tet* and *ara* operons[67]. Indeed, the first early design strategies for inducible gene expression were based on these environmental-responsive promoter sequences[73]. To tune promoter-based biosensors, activator or repressor sites can be added, subtracted, or modified, and libraries of variants with a wide range of sensitivities have been characterised[67]. Furthermore, because these designs are transcriptionally modular, additional control over the inducibility of transcriptional biosensing can also be achieved by incorporating the environmental-responsive promoters into synthetic gene networks[67].

Synthetic eukaryotic transactivation systems represent a generic version of this strategy by fusing the environmentally-responsive TF to a eukaryotic transactivation domain for ligand-programmable gene expression[74,75]. The modularity of eukaryotic promoters lends this approach to a high degree of tuneability which is extremely predicable[76]. Initially, antibiotics such as tetracycline or doxycycline were used to control gene expression through the use of the bacterial transcriptional repression domain of TetR[77], and the activatable TetR variant, TetA[78]. Depending on the application, the two domains offered the flexibility to act as an OFF or ON switch for gene regulation[77,78]. Due to concerns over antibiotic resistance and side effects in the therapeutic setting, next-generation transactivation systems were developed to have safe and orthogonal inducers, such as the apple tree leaf metabolite phloretin[79], vanillic acid[80], caffeine[81], electrical currents[82], and light[83]. Mammalian transactivation systems now bolster an impressive diversity of stimuli[81].

TF-based biosensors have been used for a wide variety of metabolite detection systems[84], including fatty acids[85], zylose[86], malonyl-CoA[87], amino acids such as L-lysine[88] and L-valine[89], and muconic acid[90]. New pairs can readily be identified via transcriptional analysis coupled with random mutagenesis while perturbing cells with the ligand of interest[54]. Once a TF and promoter/operator pair have been identified, only simple genetic manipulation is required to link the desired input to a transcriptional output. Furthermore, because of their direct link to gene expression, transcriptional biosensors are highly-amenable to rapid directed evolution strategies, such as phage-assisted continuous evolution (PACE)[91,92]. Because of their availability, simple implementation, and evolvability, TF-based biosensors have become a popular option in synthetic biology and metabolic engineering[45,54].

1.2.6 Translational biosensing

As the next stage in gene expression, translation offers another platform for linking an environmental perturbation to the control of gene expression via interaction with the mRNA transcript[54,67]. Natural RNA molecules play an incredibly diverse and important set of cellular functions due to their versatile interaction with proteins, small molecules, and other nucleic acids[93,94]. Indeed, many RNA molecules are natural environmental sensors[95,96]. Because of the versatile nature of non-coding RNA and a structure which is defined by its sequence, there has been a growing interest in designing RNA counterparts for engineering biosensors[97,98] (**Figure 4C**). For example, riboswitches incorporated into the 5' UTR of their own mRNA can bind specific small-molecules through aptamer domains, causing a conformational change that acts in cis to control gene expression[99]. Aptamer domains inspired by riboswitches are widely used as sensitive elements for RNA-based biosensing[67].

Creating an RNA-based biosensor requires typically linking the aptamer domain with a translational regulatory domain, such as the ribosome binding site, on the mRNA structure to sterically hinder the translational machinery. To achieve this, antisense RNA sequences are often used[67,100]. Conformational changes in the aptamer caused by ligand interaction can be designed to either sequester the antisense domain in a stable stem-loop structure, freeing the mRNA to create an ON switch, or release it to inhibit translation by binding the mRNA, creating an OFF switch[101]. Alternatively, translational regulation using RNA-based biosensors can be achieved through RNA self-cleavage[102]. This is realised using a naturally occurring strategy where some ribozymes possess aptameric domains responsive to metabolites. In the absence of the cognate ligand, constitutive cis autocleavage of the reporter mRNA attenuates all signal. This is then reversed when the ligand binds the aptamer domain, inhibiting ribozyme acitivty[103]. The sensitivity of any of these RNA-based biosensor strategies can then be tuned by altering the RNA sequence of the aptamer, and therefore the thermodynamic properties of the structure[101].

The most limiting factor in RNA-based biosensor development is the identification of RNA sequences that bind specifically to a new ligand[54]. Fortunately, because riboswitches do not rely on protein-metabolite interactions, RNA-based biosensors are more amenable to systematic evolution. Pre-existing aptamer libraries can be evolved towards new ligands using both *in vitro*[104] and *in vivo* aptameric selection strategies[105]. Combining a high-throughput screen with next-generation sequencing allows thousands of ligand-responsive RNA elements to be characterised for consensus sequences *in silico*, which can then allow for the rational design of new biosensors using the underlying sequence-structure-function relationships[105]. Accordingly, synthetic aptamer libraries now exist for a wide range of ligands, such as the amino acids L-lysine and L-tryprophan[106], the small molecule therapeutics theophylline[107] and folinic acid[104], and the vitamin thiamine[108].

1.2.7 Post-translational biosensing

Signal transduction pathways offer the third and most diverse platform for controlling gene expression from an environmental perturbation. The overall signalling dynamics of signalling pathways are dependent upon many properties, including the number of interconnected proteins in the signalling cascade, the nature of these molecular interactions, and the use of spatial organisation[67,109]. Furthermore, complex feedback mechanisms are often overlaid on signal transduction pathways to improve their robustness to environmental noise[109,110]. Despite the complexity, the hierarchical organisation of these pathways follows a common format; a sensitive element followed by a sequence of downstream transducers modules, ultimately resulting in in the regulation of one or more

genes[67,109]. The number of components and mechanisms together with the modularity of operating at the protein level often bestows signal transduction pathways with many points of intervention, allowing for a larger degree of control and tuning[109]. Additionally, because signal transduction pathways operate at the protein level, signalling dynamics, including feedback and deactivation mechanisms, can be rapidly achieved[109].

The sensitive element of a signal transduction pathway is often a membrane-bound receptor, for example, the two-component system[67] (**Figure 4D**). Two-component systems (TCSs) are widely found in prokaryotes and represent a powerful platform for coupling an environmental stimulus to an appropriate cellular response[111]. A prototypical TCS comprises a histidine kinase (HK) sensitive element and response regulator (RR) effector module[112]. The HK detects stimuli from outside of the cell and autophosphorylates conserved histidine residues inside the kinase, which in turn regulates the RR by phosphorylating aspartate residues within the effector module. The phosphorylated RR is then able to create a transcriptional output by binding to one or more promoters and activating or repressing gene expression[112]. TCS biosensors have been developed for diverse stimuli, such as heavy metals[113], metabolites[114,115], and light[116]. Furthermore, promiscuous TCSs can be multiplexed to distinguish against environmental factors, such as osmolarity, temperature, and pH[117].

Due to the requirement for multiple components, not all organisms are amenable to the full diversity of signal transduction pathways that exist in nature. For example, the complex G protein-coupled receptor (GPCR) signal transduction pathways of eukaryotes require a large number of components not present in bacterial systems[118]. Although receptors similar to GPCRs exist in some bacteria, such as the light sensing microbial rhodopsins[119], the mechanisms of G protein signalling are absent[118]. Ambitions to port the entire system into prokaryotes would represent serious technical effort, notwithstanding the fact that organelles, such as the nucleus, often play an essential role in the spatial organisation in the native system[120]. However, once a signal transduction pathway has been established in any chassis organism, the modularity of the components often lends the receptor to be swapped out for a new one with a different ligand specify in a “plug-and-play” manner[109]. Creating new biosensors may, therefore, simply be a case of porting new receptors and coupling them to the pre-existing signalling machinery.

Furthermore, the conserved nature of protein-based sensor domains lends them to rational engineering[67]. While many TF- and RNA-based biosensors are identified from nature and then evolved towards a new ligand, protein receptors can be designed *de-novo* using well established structural information[121]. When structural information is absent, large-scale synthetic site saturation mutagenesis techniques can be employed to approximate the ligand binding pocket to alter receptor signalling[122].

While signal transduction systems have been successfully engineered in both prokaryotes and eukaryotes, because of their small size, prokaryotic systems frequently rely on the diffusion of a small number of components. Larger eukaryotic systems, conversely, tend to have longer signalling cascades to spatially organise signal transduction, exemplified by mitogen-activated protein kinase (MAPK) pathways[109]. The additional steps involved in eukaryotic signalling pathways present more opportunities for rewiring and tuning the system. Furthermore, the larger average size and modularity of eukaryotic proteins[123] provides greater flexibility to replace or add functional domains to change the binding partners and localisations[124]. Although this adds a higher level of complexity, this also makes eukaryotic signal transduction pathways a more powerful system to engineer.

1.2.8 GPCRs as a sensitive element for eukaryotic biosensors

As the largest and most diverse superfamily of cell surface receptor in eukaryotes, comprising over 800 members in humans[125,126], G protein-coupled receptors (GPCRs) represent the ideal sensory module in eukaryotic post-translational biosensors[127]. They are responsive to a plethora of ligands and stimuli, often operate with high specificity, and have one of the key principles of synthetic biology written into their signalling architecture, modularity[4,127–129]. Furthermore, downstream effectors of GPCR signalling are diverse and represent a rich library of systems for engineering new biological responses (**Figure 5**)[127,130]. Indeed, GPCRs have been the focus a large number of mammalian synthetic biology applications, particularly within the area of theranostics; engineered cells that have been microencapsulated and implanted into the body to sense specific stimuli and respond with an appropriate therapeutic output tailored to the individual[131].

In theranostic systems, GPCRs sensitive to orthogonal dietary supplements such as oleanolic acid[132], vanillic acid[133], or a physical stimulus such as light[134], can be used to reprogram cells to respond to exogenously supplied signals in a dose-dependent manner. Linking this response to a therapeutic output means hard to administer drugs can be supplied to the body *in situ* without concerns over delivery or pharmacodynamics[131]. Furthermore, multiple drugs could be produced by the same implant using the same stimulus, reducing the complex drug regimens that often lead to failure due to nonadherence[135].

Perhaps a more exciting angle on this approach is the design of closed-loop synthetic gene networks, which detect a disease state as it presents itself and reacts with an appropriate therapeutic response. By both monitoring and treating the disease as it presents itself, such engineered devices would work as a preventer of a disease state, rather than the post-diagnostic treatment[131]. An example of this

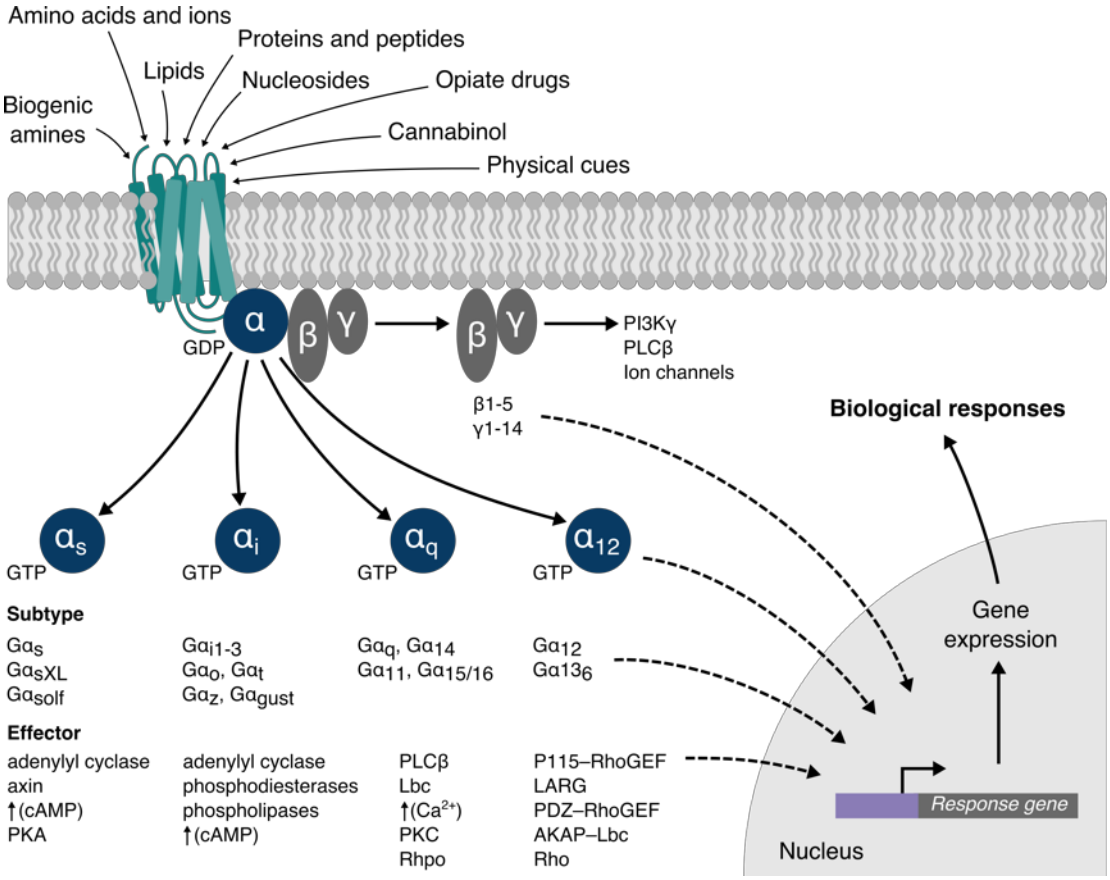


Figure 5. Diversity of G protein-coupled receptor signalling. A wide diversity of ligands and stimuli are able to stimulate a cellular response through interaction with GPCRs. This interaction induces a conformation change in the receptor which enables them to catalyse the exchange of GDP for GTP in the Gα subunit and the subsequent dissociation from the Gβγ. The Gα subunits of G proteins are divided into 4 subfamilies; Gαs, Gαi, Gαq, and Gα12. A single GPCR can couple one or more of the Gα family members. Depending upon the type of Gα and Gβγ within he heterotrimeric G protein complex, either the Gα or Gβγ activates the downstream effector. Gα subunits activate a wide range of downstream effectors which vary within each family, whereas Gβγ dimer signalling typically activate a phospholipase, ion channel or lipid kinase. Besides the regulation of these classical second-messenger generating systems, Gα and Gβγ can activate signal transduction pathways, such as the mitogen-activated protein kinase (MAPK) family of serine-threonine kinases. Ultimately, the biological response is determined by the combinations of interacting receptor/G protein complexes and downstream effectors present within the cell. Figure adapted from Dorsam and Gutkind[130].

being the closed-loop signalling network utilising the G-protein-coupled bile acid receptor, TGR5, for the protection of liver disease in mice[136]. After implantation of the genetically engineered human cells into mice, the encapsulated liver-protection device detected pathologic concentrations of serum bile acid, responding by producing therapeutic levels of hepatocyte growth factors that successfully protected the animals from acute drug-induced liver failure[136].

Although GPCR-based theranostics are a long way from becoming a viable medical treatment, their importance as pharmacological targets and responsiveness to a diversity of metabolites relevant to humans also makes them a desirable system for creating biosensors for metabolic engineering (**Table 1**). Of particular interest is the olfactory class of GPCRs, which through combinatorial mechanisms

allow animals to detect the enormous number of structurally diverse chemical cues in their environment with extreme sensitivity and specificity[137–139]. Although we are only beginning to work out how the structure and topographic distribution of olfactory receptors (ORs) in sensory neurons can differentiate between the seemingly infinite number of odorous substances[138], there has been a high interest over the past two decades in developing OR-based biosensors due to their promising potential in various applications[139,140]. However, many challenges still surround OR-based biosensors, and studies in which OR-mediated signal transduction pathways have been used to detect odorant molecules are very few[141].

Table 1. Examples of GPCR ligand/stimuli diversity.

Inductive ligands/stimuli		**GPCR(s)**	**Reference**
Biogenic amines	Histamine	Histamine receptor (HRH1-4)	142
	Serotonin	Serotonin receptor (HTR1-7)	143
	Acetylcholine	Muscarinic acetylcholine receptor (CHRM1-5)	144
Amino acids and ions	Ca^{2+}	Calcium-sensing receptor (CASR)	145
	Glutamate	Metabotropic glutamate receptor (GRM1-8)	146
Proteins and peptides	Thrombin	Protease activated receptor (FR2, F2RL2+3)	147
	Glucagon	Glucagon receptor (GCGR)	148
	Follicle-stimulating hormone (FSH)	FSH receptor (FSHR)	149
Lipids	Lysophosphatidic acid (LPA)	LPA receptor (LPA1-6)	150
	Sphingosine-1-phosphate	Sphingosine-1-phosphate receptor (S1PR1-5)	151
Nucleosides	Adenosine	Adenosine receptor (ADORA1-3)	152
Physical cues	Light	Rhodopsin (RHO)	153
	Temperature	Drosophila rhodopsin receptor (ninaE)	154
Odorants		Olfactory receptors	138
Cannabinol		Cannabinoid receptors (CB1+2)	155
Opiate drugs		Opioid receptors (OPRD1, K1, M1, L1)	156

Notably, G protein-coupled receptors also exist for the end products of one of the most ambitious metabolic engineering projects to date, the complete biosynthesis of opioids in yeast[156,157]. This 23 enzyme pathway, borrowing genes from plants, mammals, bacteria, and yeast itself, was able to produce detectable levels of hydrocodone (< 1 μg/L), the selective agonist of the μ-opioid receptor[157,158]. At this titre, thousands of litres of fermentation broth would be required to produce a single dose of hydrocodone, as used in Vicodin (5 mg). Such improvements would not be economically viable through fermentation scale-up alone and would require many strain and pathway improvements[157]. Currently limited to screening via LC-MS[157], exploring the entire genetic landscape of a pathway a fraction of this size would be challenging. The development of an opioid receptor-

based biosensor would be revolutionary and allow researchers to rapidly prototype strain improvements and fermentation conditions, possibly increasing the yields to economically viable levels.

For decades, researchers have been depositing high-quality information to databases explicitly curated for GPCRs, including structural, functional, and ligand binding data, providing an enormous library of potential receptor/ligand pairs for generating novel biosensors from[159,160]. Furthermore, GPCRs by their very nature are incredibly evolvable[128,129]. Their highly-conserved, yet incredibly plastic 7-TM structure has allowed GPCRs to detect things as simple as light[161], all the way through the biological spectrum to entire proteins[162]. Platforms to diversify GPCRs are well established, enabling researchers to evolve GPCRs towards new ligand specificities for orthogonal control of GPCR-mediated biological responses[163]. Designer receptors exclusively activated by designer drugs (DREADDs), permit spatial and temporal control of GPCR signalling *in vivo* and are a popular chemogenetic tool for the non-invasive study of GPCRs in a large number of therapeutically relevant contexts[164,165]. Developing novel GPCR-based biosensors in the future may simply involve choosing a receptor with a ligand specificity similar to the chemical of interest and then rapidly evolving it towards the new target[166]. Such efforts would expand the already extensive list of chemicals GPCRs can detect, enabling biosensors for diverse value chemicals.

Finally, as cell surface receptors that detect the extracellular environment, GPCRs are particularly well suited to the two-cell screening system previously mentioned in section **1.2.2**. Although currently limited to medium-throughput screening, two-cell biosensing systems are poised to become a powerful technique that can be used to analyse levels of secreted products[55]; a highly-desired property that can significantly reduce the cost of microbial chemical production[167,168]. Furthermore, if a universal GPCR-based biosensing platform were developed, this would facilitate the screening of diverse production hosts, by either screening prepared samples or by growing the producer and sensor cells in co-culture.

However, many of the current GPCR-based sensing systems are limited to mammalian cell lines[169,170]. The applications of these systems to microbial metabolic engineering are limited due to their complex and expensive growth requirements[171], difficulties scaling to higher-throughput plates formats[172], and the need for specialised experience[173]. Furthermore, the modular and flexible nature that makes GPCRs so attractive as a biosensing platform in mammalian cells also makes them an exceptionally complicated system to understand, let alone engineer[174,175]. Within a single cell line, many G protein signalling pathways are seen, and the downstream pathway effectors are complex and intertwined with each other[174,175]. As with many other biological systems that are complex, we often turn to a simpler model organism for the answer[176–178].

1.2.9 Yeast as the modern workhorse for metabolic engineering

Despite being a comparatively simple eukaryotic organism, mechanisms of GPCR signalling in the yeast, *Saccharomyces cerevisiae*, are similar to those in mammalian systems[179]. This can often allow for the functional exchange of components, enabling this simple chassis with the sensory function of higher eukaryotes[180]. Furthermore, yeast only has two GPCR signalling pathways, the pheromone response and glucose sensing pathways[181], between which there is minimal crosstalk[180]. Both of these pathways are well understood, and either of them can be altered as they are non-essential to cell survival[181]. Yeast, therefore, represents one of the simplest organisms in which we can engineer GPCR-based signal transduction pathways[180].

As well as being a simplistic chassis for engineering GPCR-based biosensors, yeast is also one of the most relevant organisms to metabolic engineering[157,182]. Indeed, some of the most ambitious metabolic engineering projects to date have been achieved in yeast, including the aforementioned production of opioids[157] and the flagship synthetic biology project; the production of the antimalarial drug precursor artemisinic acid[50,51]. Yeast has also been engineered to make a number of other chemicals at industrial scales, including the production of resveratrol, succinic acid, lactic acid, and advanced biofuels[183].

Yeast has been widely adopted within the industrial setting because of a number of intrinsic biological properties: (i) it is a generally regarded as safe (GRAS) organism, making it suitable for large-scale operation[184], (ii) genetic manipulation is greatly facilitated due to a preference for homologous recombination (HR) over non-homologous end joining (NHEJ)[185], (iii) unlike prokaryotes, multiple organelles provide distinctive environments and compartments for biosynthesis[186], and (iv) a tolerance for harsh industrial conditions[184]. Because of its popularity, a dedicated repertoire of techniques and tools have been developed for yeast to extend these intrinsic biological properties, further extending its relevance to metabolic engineering[187].

As the most basic element in a biosynthetic pathway, characterising genetic parts has been a significant focus of yeast synthetic biology[187]. We now understand a considerable amount of detail regarding the engineering of biological parts at the DNA level (i.e. yeast shuttle vectors[188]), RNA level (i.e. promoter[76] and terminator engineering[189,190]), protein level (i.e. degradation[191], localisation[192], and secretion[193]), and metabolite level (i.e. co-factor engineering[194]). Technologies have also been developed for the continuous evolution of targeted parts, a strategy that used to exist only in bacterial systems, allowing us to expand the rich library of parts yeast synthetic biologists already have at their disposal[195,196]. To facilitate the quick and easy assembly of these parts into genetic constructs for use within yeast, powerful modular cloning toolkits have been created, such as the Yeast MoClo Toolkit[33].

To extend the capabilities of yeast beyond the introduction of heterologous biosynthetic pathways, CRISPR-based tools have been developed to combinatorially rewire host metabolism[197] (**Figure 6**). This approach allows metabolic engineers to optimise the strain by overexpressing, knocking-down, or knocking out multiple gene targets simultaneously, enabling the perturbation of metabolic and regulatory networks in a modular, parallel, and high-throughput manner. Such approaches can be used to create large combinatorial strain libraries which, together with libraries of genetic constructs, can be used to scan a more comprehensive genetic landscape to find the optimal phenotype[197].

Finally, yeast is poised to become the first eukaryotic organism with a custom-designed genome created entirely from synthetic DNA as part of the international Sc2.0 project[29]. Included within the synthetic genome design was the addition of loxPsym sites downstream of all non-essential genes that provide the basis for the inducible evolution system, known as SCRaMbLE (synthetic chromosome rearrangement and modification by loxP-mediated evolution)[198]. The SCRaMbLE

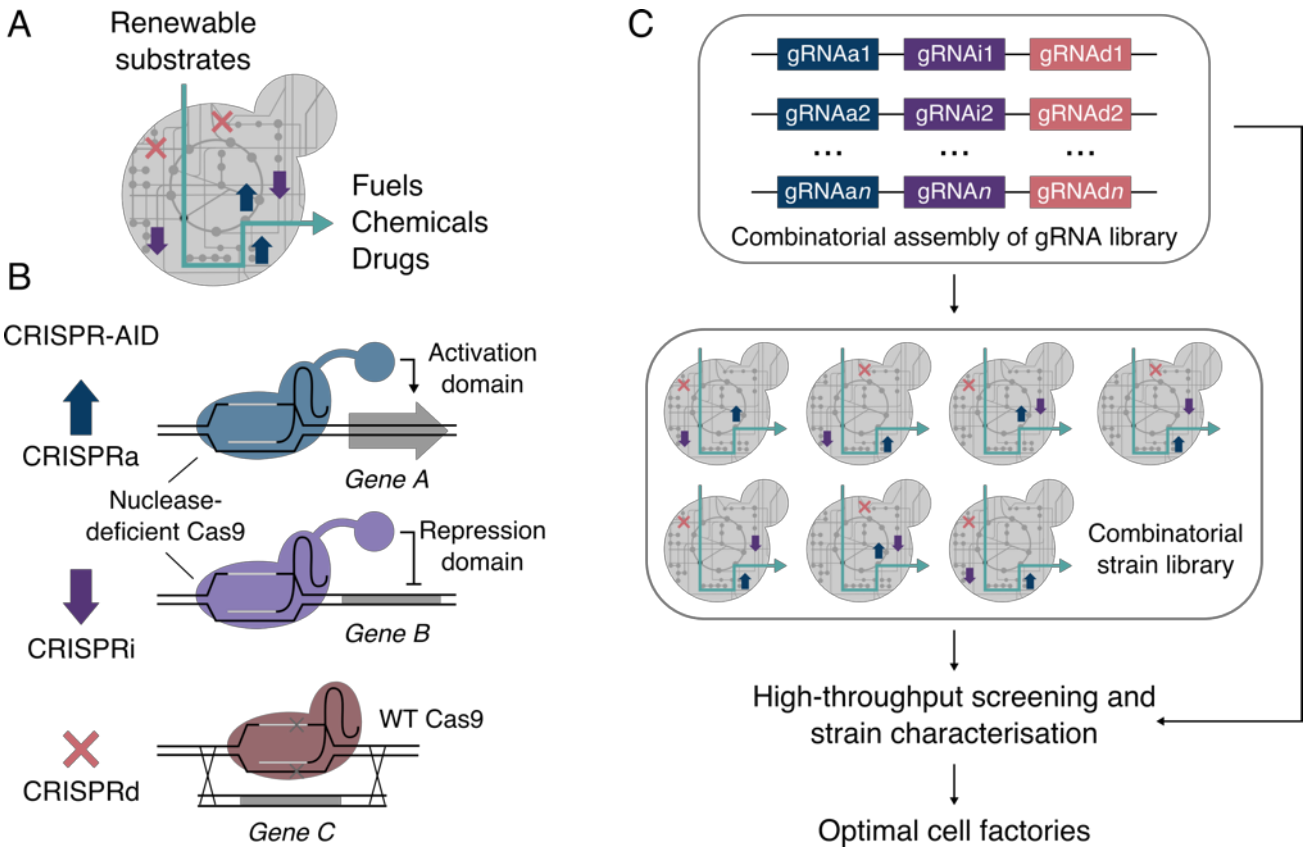


Figure 6. The CRISPR-AID platform for combinatorial metabolic engineering in *S. cerevisiae*. (*A*) Optimising cellular metabolism for producing commodity and high value chemicals from renewable resources often requires the upregulation, downregulation and deletion of many genes. (*B*) The CRISPR-AID platform, consisting of 3 orthogonal CRISPR/Cas proteins; one nuclease-deficient CRISPR protein fused to an activation domain for CRISPRa, another nuclease-deficient CRISPR protein fused to a repression domain for CRISPRi, and a catalytically active CRISPR protein for CRISPRd. (*C*) CRISPR-AID can be used to combinatorially explore metabolic rewiring by programming gRNA libraries which can then be screened for the optimal cell factories. Figure adapted from Lian *et al.*[197].

system was designed to allow on-the-fly genome rearrangements, producing a population of combinatorially diverse yeast cells with correspondingly diverse phenotypes (**Figure 7**). The SCRaMbLEing of single synthetic chromosomes has already been demonstrated as a powerful method for rapid host strain improvement, creating genetic backgrounds favourable to various biosynthetic pathways[199]. The final construction of the complete Sc2.0 strain will expand the capabilities of the SCRaMbLE system, providing more possible genetic combinations, and therefore greater phenotypic diversity from which to select.

The intrinsic properties of yeast supported by this versatile range of tools guarantees yeast a prosperous future at the forefront of metabolic engineering for many years to come[182]. Developing yeast as a platform for engineering GPCR-based biosensors would, therefore, enable one of the most industrially relevant organisms with one of the most versatile sensing systems in biology. This would allow for the integrated screening of diverse yeast libraries, complementing our ability to build, and increase the throughput of yeast metabolic engineering projects significantly[71]. Furthermore, many of the properties that make yeast industrially relevant, such as genetic amenability, tolerance to harsh conditions, lost cost, as well as an ability to store as 'active dry', and compatibility with high-throughput technologies, make yeast GPCR-based biosensors attractive more generally, and holds great promise from a variety of applications[71,180,200,201].

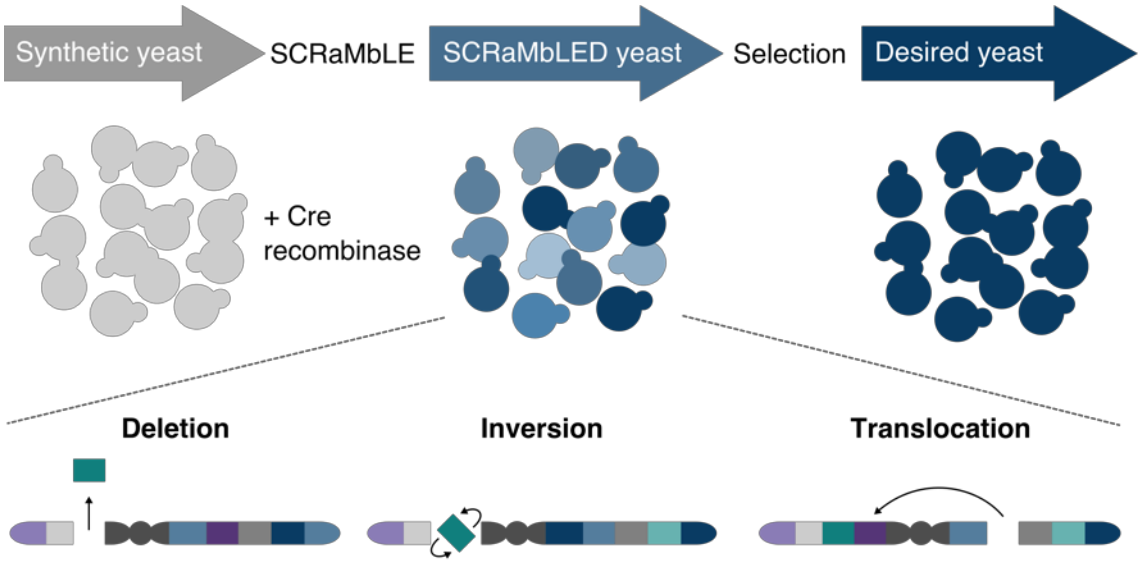


Figure 7. SCRaMbLEing the synthetic yeast genome. Induction of the SCRaMbLE system in yeast strains containing synthetic chromosomes (grey) results the deletion, inversion, and translocation of any combination of the recoded genomic regions, leading to vast phenotypic diversity in the yeast population (blue shades). A selection or screen can then be applied to the SCRaMbLEd yeast to identify improved host strains (dark blue).

1.2.10 The pheromone response pathway

The pheromone response pathway, or mating pathway for short, is one of two native GPCR signalling pathways in *S. cerevisiae*[181], and has long been the go-to choice for coupling heterologous GPCRs for functional study or for evolving towards a desired target[163,180]. The intracellular signal transduction pathway from agonist activation to gene expression is one of the most extensively studied in eukaryotes[179]. Although this system has been a pathfinding experimental paradigm for GPCR-mediated signalling and its regulation for the last 30 years, we continue to learn how the signal is transmitted from the cell-surface GPCR, via a MAP kinase cascade, to the nucleus[202]. However, the list of crucial parts is essentially complete, and the order in which they function with regards to the transmission of the signal from the environment into the nucleus is well understood[179]

Despite its deceptively simple lifestyle as a unicellular eukaryotic microorganism, yeast exists as three distinct cell types. There are two haploid mating types, termed **a**- and α-cells, and a third type, the **a**/α diploid. The pheromone response pathway exists only in the haploid cells and is directly responsible for initiating mating between cells of the opposite type, controlled by the exchange of peptide pheromones; **a**-cells express the α-factor receptor (Ste2) and secrete the **a**-factor peptide, whereas α-cells express the **a**-factor receptor (Ste3) and secrete the α-factor peptide. Both of these pathways couple to the same heterotrimeric G protein complex and have the same signal transduction pathway[179]. As the response in **a**-cells has been subject to far more study and is typically the cell type for coupling heterologous receptors, for simplicity I will specifically refer only to the response in **a**-cells (**Figure 8**). Several of the components of the mating pathway are also involved in other distinctive pathways that regulate the response to certain stresses and filamentous growth. Again, for simplicity, these will not be discussed, but recent reviews can be found here - Cullen and Sprague[203], and Hohmann[204].

Binding of the α-factor peptide to the Ste2 GPCR on the cell surface, causes the α subunit of the heterotrimeric G protein complex, Gpa1, to dissociate from the Gβγ dimer, Ste4/Ste18. This event results in the recruitment of the Ste5 scaffold protein to the plasma membrane, which in turn leads to the assembly the MAP kinase cascade, MAPKKK Ste11, MAPKK Ste7, and MAPK Fus3. A key result of Gβγ binding Ste5 and recruiting the cascade to the membrane is that Ste11 and Ste20 are brought close together. This leads to the phosphorylation of Ste11 by Ste20, enhanced by Ste50, initiating a phosphorylation cascade down the MAP kinase pathway, resulting in the release of Fus3 from the Ste5 scaffold. In the cytoplasm, Fus3 phosphorylates a number of targets, including Far1 which mediates the response for cell cycle arrest and the downregulation of a large number of genes. In the nucleus, Fus3 phosphorylates the pheromone-responsive TF, Ste12, and the negative regulators of Ste12, Dig1 and Dig2. The subsequent release of the Ste12 from negative repression and the

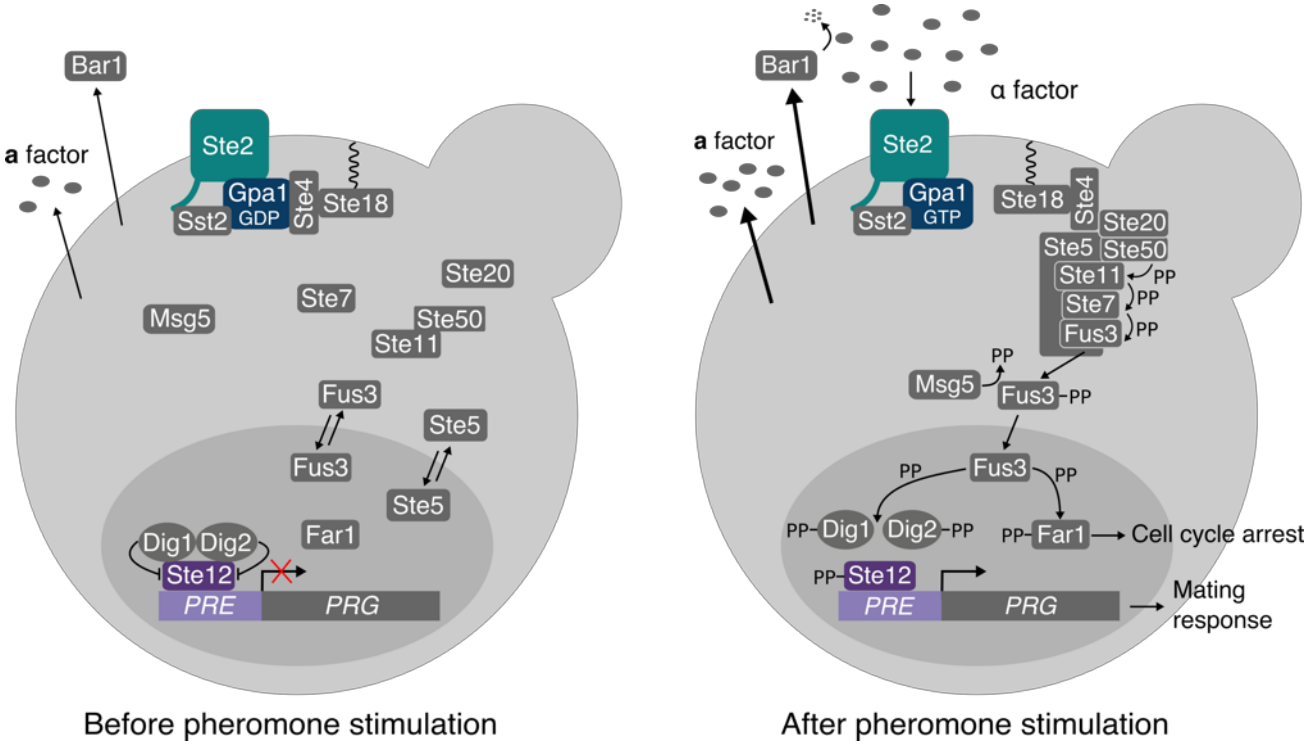


Figure 8. The yeast pheromone response pathway in MATa cells. Binding of the peptide pheromone, α-factor, to the Ste2 receptor causes the Gα subunit, Gpa1, to dissociate from the Gβγ dimer, Ste4/Ste18, leading to the membrane localisation and activation of the MAPK cascade and the detachment of the MAPK, Fus3, from the scaffold. Fus3 then phosphorylates a number of targets in the cell, including Far1, leading to cell cycle arrest, and the Dig1/Dig2/Ste12 complex, leading to the upregulation of approximately 100 genes involved in the mating response, identified by a pheromone response element (PRE).

activation of its transcriptional activity leads to the upregulation of over 100 genes involved in the mating response, including positively-acting components (Ste2 and Fus3) and negative feedback regulators (Gpa1, Sst2, Msg5) of the mating pathway[179]. Ste12 also acts in a positive autoregulatory manner, by upregulating its own expression[205].

Once the pheromone response pathway has been activated, multiple layers of regulatory proteins act to return the pathway to the OFF-state to promote desensitisation/adaptation and recovery. This ensures the resource-costly response is only successful if a viable mating partner is present[179]. This modulation of signal intensity is also crucial for accurate gradient sensing and faithful signalling through the pathway[206]. (i) An α-factor specific protease, Bar1, is secreted into the periplasmic space to cleave and inactivate α-factor in the environment, allowing cells to resist activation from low levels of extracellular peptide pheromone. (ii) The pheromone-bound receptor is desensitised by phosphorylation sites on the cytoplasmic C-terminal tail, which also acts as a tag for endocytosis and degradation. (iii) The regulator of G protein signalling (RGS), Sst2, catalyses the conversion of GTP to GDP in the Gα subunit, returning it to the deactivated state and sequestering the Gβγ dimer, preventing further signalling. (iv) Fus3 controls a negative feedback loop which limits the duration of

its own activity. (v) A number of phosphatases, including Msg5, dephosphorylate the MAPK cascade, particularly Fus3, to damped signalling through the pathway[179].

The result of successful pathway activation above the threshold for a response is a series of physiological changes in preparation for mating. These include cell-cycle arrest during the G1 phase, significant changes in the expression of around 200 genes, asymmetrical growth in the direction of the mating partner, and ultimately, the fusion of the plasma membrane between two mating partners, shortly followed thereafter by the fusion of their nuclei. Cells are not irreversibly committed to this process, and the failure to successfully mate eventually results in cells re-entering the cell-cycle as haploids[179]. As soon as the two mating partners fuse, the pheromone response is shut down permanently when the transcription of many pathway components are repressed by the a1/α2 diploid-specific heterodimer. Considering the number of steps involved in the mating pathway, the signal transduction is surprisingly fast. 30 seconds after pheromone stimulation, the G protein can be seen dissociating. Several minutes after that, activation of the MAPK cascade can be detected, and changes to gene expression have already begun by 10 minutes. The entire process, from the activation of the mating response to the formation of a diploid cell, takes around 4 hours[179].

Although quite distinctive from the majority of GPCR pathways in mammalian cells, most similar perhaps to ERK1/2 activation[207] (some of which is mediated by Gβγ in mammalian systems), nevertheless, the pheromone response pathway has played a predominant role in establishing many paradigms in GPCR signalling. Contributions include the demonstration that the heterotrimeric G protein complex is responsible for transmitting the signal from the receptor to the downstream effectors, the strategy for ordering gene function in signalling pathways using a combined use of loss- and gain-of-function mutants, and the discovery of the first MAPK scaffold and RGS[179]. Yeast has also been instrumental in determining the function of certain mammalian GPCRs in isolation by heterologously expressing and coupling them to the G protein complex[180].

1.2.11 The heterologous expression of GPCRs in yeast

The first heterologous GPCR to be functionally coupled to the mating pathway was the human β_2-adrenergic receptor[208]. Since then, many other receptors, such as the adenosine[209], GLP-1[210], melatonin[211], and muscarinic[212] receptors, have been successfully coupled for functional study. This has been aided by modifications to yeast to improve coupling and for more convenient outputs[180] (**Figure 9**).

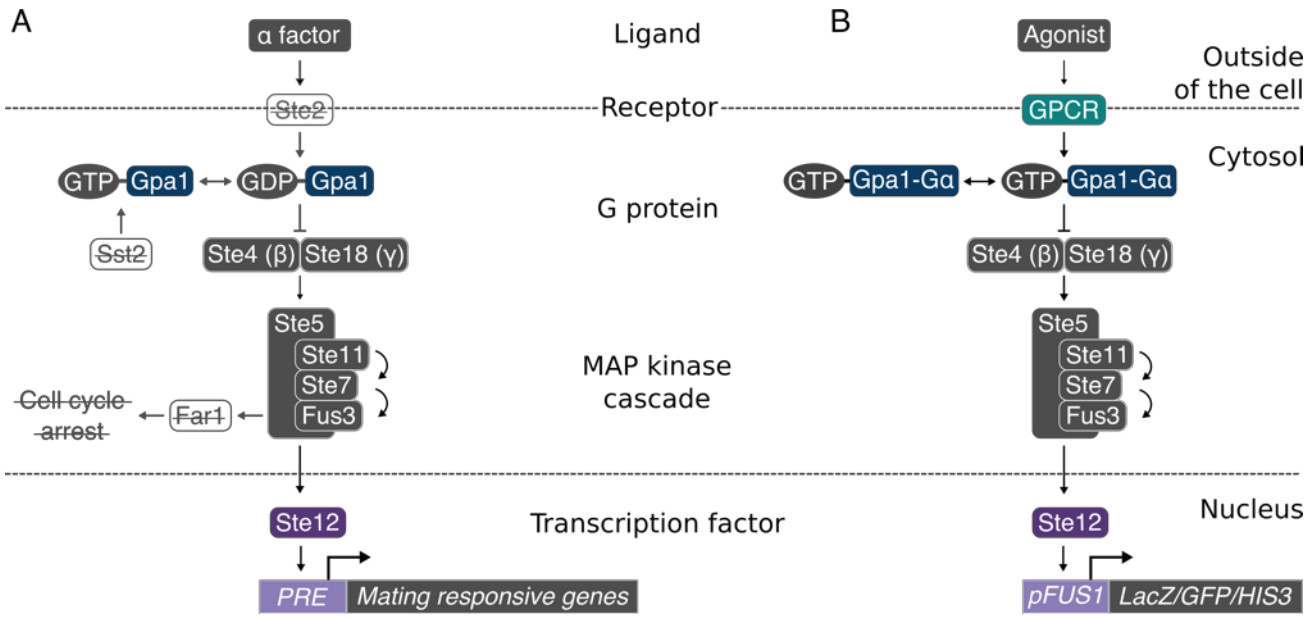


Figure 9. Engineering of *S. cerevisiae* for analysing heterologous GPCRs. (*A*) Modifications introduced to the yeast pheromone response pathway to facilitate the analysis of heterologous GPCRs. Gene disruption is used to remove the endogenous α-factor receptor Ste2 to prevent sequestering of the heterotrimeric G protein complex into inactive complexes, the RGS Sst2 to reduce desensitisation and enhance signalling, and Far1 to prevent cell cycle arrest in response to signalling. (*B*) Heterologous GPCRs can then be introduced to the system and coupled to the pathway via a Gpa1-Gα chimera, where the c-terminal residues of Gpa1 have been substituted for the mammalian equivalent to improve the interaction. Readouts of colour (*LacZ*), fluorescence (*GFP*), or growth (*HIS3*) are achieved by coupling the expression of these reporter genes to the pheromone responsive *FUS1* promoter. Figure adapted from Ladds *et al.*[180]

Introducing new heterologous receptors into yeast has been met with varying success. Some receptors can be expressed in yeast unmodified, whereas others get retained within the intercellular compartments or get degraded[180]. Sometimes issues with expression can be solved by merely tuning the method of expression[213], other times this can require additional engineering. Adding leader sequences or substituting the N-terminal sequence with that of a highly-expressing receptor can occasionally improve this situation[214]. The stability or activity of some receptors can also be improved by modifying or removing regions within the receptor, for example, deletion of the central portion of the third intracellular loop of the muscarinic receptor increased the expression considerably[212]. Other receptors can require the co-expression of accessory proteins to aid their progress through the secretory pathway to the cell surface or for their correct activation[215]. As these proteins are not likely to be present in yeast, their properties can be studied in isolation to determine their effects on the receptor[216].

Coupling of the heterologous receptor to the endogenous Gα, Gpa1, is often successful but can be inefficient. A number of approaches to improve this coupling have been attempted, including the outright substitution of Gpa1 with a mammalian Gα, although this has had limited success owing to their low affinities for the Gβγ[180]. To date, the best strategy for improving coupling to the mating

pathway is to use Gα chimeras, which consist of the Gpa1 protein in which the last five residues have been substituted for the various mammalian equivalents, often called Gpa1/Gα transplants[217]. The ability of Gα chimeras to couple different receptors reflects the plasticity of the mammalian system[180].

Finally, to provide a measure of receptor activity, most yeast GPCR assays take advantage of the strongly inducible Ste12-mediated *FUS1* promoter to drive reporter expression[218,219]. Several reporter genes have typically been used for these studies, including *HIS3*[220], lacZ (β-galactosidase)[221], and GFP[222]. Although *HIS3* and lacZ have been used very successfully over the years, the rapid signal and easy integration with sensitive measurement apparatus and high-throughput sorting technologies, such as FACS, has meant GFP has become the recent reporter of choice.

1.2.12 Yeast GPCR-based biosensors

Due to a rich history of coupling heterologous GPCRs for functional study, it is unsurprising researchers have turned their attention to repurposing the yeast mating pathway for biosensing. Indeed, a recent series of publications have highlighted the potential for reprogramming this system for a diverse number of applications (**Figure 10**). Although a number of publications, stretching back to the early 2000s, have highlighted the use of olfactory receptors in yeast for screening volatile chemicals[223–227], a consistent lack of convincing data, appropriate controls, or the use of suitable reporter assays makes it hard to trust these reports, aspects of which have not been repeatable[228]. Instead, we will focus on the more recent contributions to the field of yeast GPCR-based biosensors in the context of synthetic biology.

The first paper to lay down a serious framework for creating yeast GPCR-based biosensors for the detection of industrially relevant chemicals was Mukherjee *et al.* in 2015[228] (**Figure 10A**). In this paper, the authors defined the pheromone response pathway as three separate units; the sensing unit (a heterologous receptor), the processing unit (the G protein complex and the MAPK signalling cascade), and the response unit (the TF and reporter). Using the human OR1G1 and GPR40 receptors as the sensing units, they were able to show a sensitivity to decanoic acid, linking this to a response unit utilising a GFP output. By varying the response unit, using either the native Ste12 TF with the pheromone-responsive *FUS1* and *FIG1* promoters or synthetic TFs targeting synthetic promoters, they were able to modulate the pathway response. Although the results were variable using these particular receptors, the framework description of the pathway as three distinct units, the contribution of synthetic TF/promoter pairs, and the use of flow cytometry to measure the corresponding output laid important foundations for the future of GPCR-based biosensing in yeast.

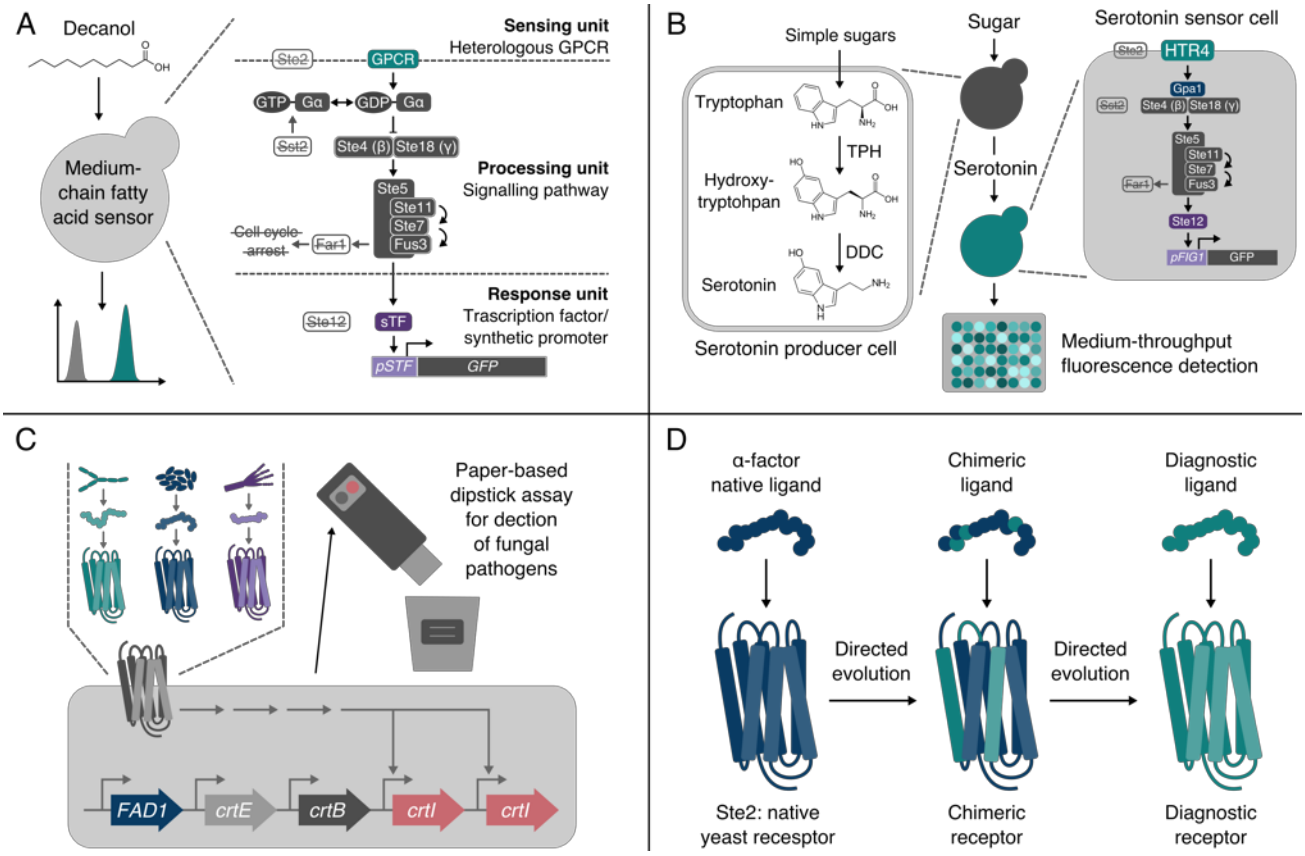


Figure 10. Examples of yeast GPCR-based sensors. (*A*) The biosensing of medium-chain fatty acids using olfactory receptors[228]. (*B*) A medium-throughput screen of microbially produced serotonin using the HTR4 receptor[55]. (*C*) A modular yeast biosensor for low-cost point-of-care pathogen detection using fungal pheromone peptide receptors[229]. (*D*) The directed evolution of the native Ste2 receptor towards peptide biomarkers to support clinical diagnostics[166].

The next publication from this team extended their work to the sensing of microbially produced serotonin, focusing on the development of a medium-throughput assay to screen producer strains[55] (**Figure 10B**). Using the human HTR4 receptor, they developed and optimised a two-cell 96-well screening platform for the detection of serotonin from the spent media of yeast producer strains. Although the dynamic range and sensitivity of the sensor strains were low, by diluting the spent media or supplementing with exogenous serotonin they were able to provide statistically significant separation of producer strains, demonstrating, for the first time, the utility of GPCR-based biosensors for increasing the throughput of strain characterisation for optimising a biosynthetic pathway.

Around the same time as the serotonin sensor, Ostrov *et al.* published a modular yeast biosensor for low-cost point-of-care pathogen detection[229] (**Figure 10C**). By coupling various peptide pheromone receptors from nine major human, plant, and food fungal pathogens to the mating pathway, they were able to detect the presence of these microbial contaminants with sensitivities and specificity levels comparable to antibody and nucleic acid assays. Furthermore, exchanging the GFP reporter for the production of lycopene allowed them to convert the output of the system to a simple, reagent-free

colourimetric readout. They then developed a cheap, one-step dipstick-based device that could be used for the diagnosis of fungal pathogens in complex samples, such as blood cultures. This system could easily be expanded to other fungi by mining new GPCR-pheromone pairs, offering a cheap, easy-to-read diagnostic which is not reliant on cold-chain storage.

Finally, we have the recent publication by Adeniran *et al.*[166]. In this study, the researchers sequentially evolved the native Ste2 receptor towards a new target specificity through a series of chimeric intermediates with increasing similarity to cystatin C, a new and promising peptide biomarker for detecting early declines in kidney function[230] (**Figure 10D**). They then demonstrated the clinical utility of the yeast-based biosensor by showing specific activity for C-terminally amidated cystatin C peptide in human urine. Although the system was not sensitive enough for clinical application, this substrate walking approach demonstrates the evolvability of GPCRs and how we can expand our ligand sensing capabilities through the evolution of our current receptor library.

1.2.13 Limitations of yeast GPCR-based biosensors

The recent examples of yeast GPCR-based biosensors demonstrate an emerging tool in synthetic biology for medical diagnostics, pathogen detection, and metabolite measurement. However, natural biosensing systems respond to their cognate inputs over finite concentration ranges that are often mismatched with artificial application demands[231]. Although the choice of strain, specificity of the receptor, and coupling to downstream responses are all vitally important considerations when designing a biosensor, ultimately, the dose-response characteristics will determine whether the biosensor is suited for the application (**Figure 11A**).

When describing the properties of a biosensor response curve, there is currently no standardised nomenclature and many alternative terms are used[43,231–233]. Before beginning discussions on the characteristics of dose-response curves, we here define a set of terms that are used consistently throughout this work and have been carefully chosen from across these many studies to allow each property to be examined independently from the others.

The range of concentrations over which a biosensor responds is the operational range. This range is determined by measuring the concentrations that give a graded, concentration-dependent change in the response. The lower and upper bounds of the operational range relate to the minimum and maximum sensitivities of the biosensor, respectively. A biosensor is unable to detect differences below the minimum sensitivity or above the maximum sensitivity, as without the graded response differences in the output are indistinguishable. The operational range is intrinsically linked to the slope

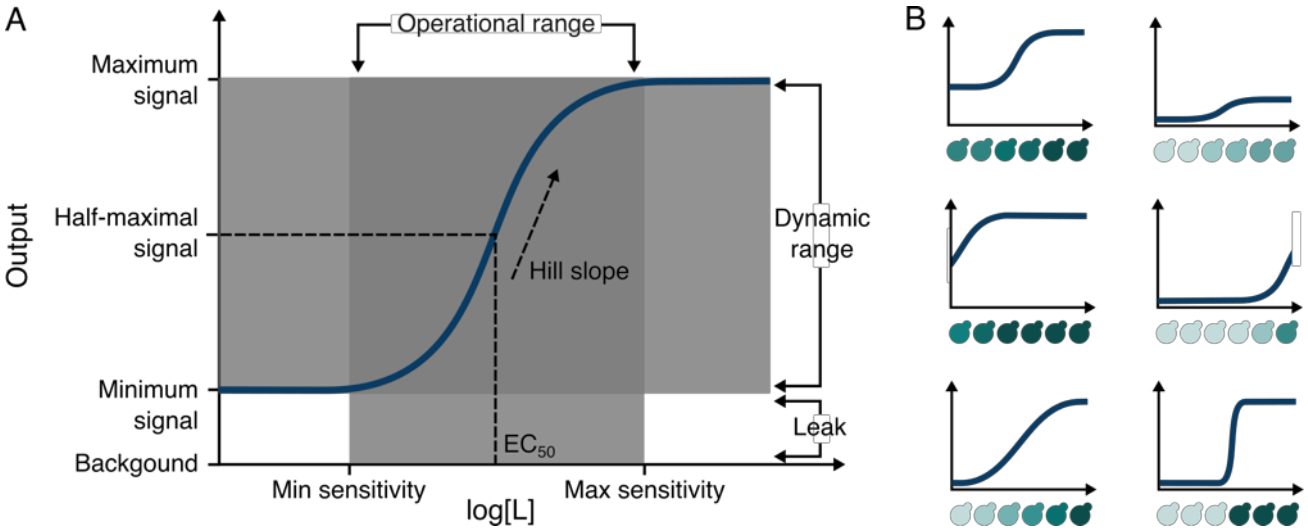


Figure 11. Properties of a dose-response curve. (*A*) The relationship between the biosensor output and ligand concentration is defined by the dose-response curve. The range of concentrations over which a biosensor operates is the operational range, the ratio of the maximum and minimum signal is the dynamic range, and the basal activity of the biosensor in the absence of ligand is the leak. (*B*) Example biosensor responses over a concentration range for a desired application. The top four designs are undesirable due to properties such as leakiness, low dynamic range, and inappropriate operational ranges. The bottom two responses are more appropriate for the range of concentrations, but depending upon the specific application, either the linear curve on the left or the digital curve on the right may be better suited to the demands of the biosensor.

factor, or Hill slope. Typically, lower Hill slope values relate to linear systems with large operational ranges, whereas higher Hill slope values relate to more digital-like systems with narrow operational ranges[234].

However, the Hill slope is also affected by the signal-to-noise of the biosensor, also referred to as the dynamic range of the system. This can be quantified as the ratio of the maximum signal output to the minimum signal output of the biosensor. High dynamic ranges are essential for producing reliable signals that can be discerned from noise, especially when the exact quantification of a ligand is required[43]. Halfway between the minimum and maximum signal of a dose-response curve is the potency of the inducer, or half-maximal effective concentration (EC_{50}). This is often a good measure of the system's sensitivity, although it can be unreliable for systems where the operational range varies.

Finally, we have the leakiness of the system. This is defined as the ratio between the minimal signal when the biosensor is in the OFF-state (also referred to as the basal level of activity) and the signal of biosensor strain lacking the reporter output (also referred to as the background). The leakiness of the system is an important consideration when creating a biosensor, as it significantly affects the overall dynamic range. Furthermore, when embedded within a complex synthetic gene network, any form of leaky activity may erroneously activate downstream processes, leading to a system

malfunction or a higher rate of false positives in selection processes[43]. For this reason, much effort is usually focused on reducing the leakiness of biosensing systems.

Depending upon the combination of these properties, a large number of possible dose-response curves can be created (**Figure 11B**). The ability to tune individual components to control properties in artificial systems, such as the dose-response of a biosensor, is central to the process of system design in all fields of engineering, and synthetic biology should be no exception[231]. Although we are beginning to turn our heads towards addressing the need for tunability in other biosensing systems[232,233], GPCRs have yet to receive the same attention, limiting the initial promise of their widespread application. In the four previously given examples of yeast GPCR-based biosensors, each system had a specific property which was not well suited for the application.

The biosensors designed for measuring medium-chain fatty acids and serotonin both had a very narrow operational range[55,228]. Applying the serotonin sensor strain to the detection of microbial production required significant assay optimisation, such as the supplementation of exogenous ligand to shift the samples into the linear range of the biosensor for quantifying concentrations[55]. This additional sample preparation created another laborious step to an approach specifically focused on removing a bottleneck in the development cycle. Conversely, the fungal pathogen detector had a large operational range, spanning four orders of magnitude[229]. As a device designed for visual readouts of pathogen levels above a certain threshold, a narrow operational range with a digital-like response would have been more appropriate, allowing users to more easily distinguish whether there is fungal contamination or not[229]. Finally, although the cystatin C biosensor was highly selective for the peptide biomarker and had demonstrated its utility in complex biological samples, it was not sensitive enough for clinical application[166].

The ability to tune GPCR-based biosensors would allow for more appropriate detection windows and thresholds, simplify sample preparation, and provide a more suitable output for the individual needs of each application. Furthermore, tuneability of GPCR-sensing modules would support complex artificial systems by matching the output/input levels across connected networks[235,236]. As yeast GPCR-based biosensors are poised to become a powerful tool for synthetic biology, there is a need to address the tuneability of these systems to unlock their full potential.

1.2.14 Tuning the pheromone response pathway

As a complex signal transduction pathway with many steps and layers of regulation, the mating pathway has a multitude of potential points of intervention for altering the signalling dynamics. Indeed,

the Ste5 scaffold gained much attention in the mid to late 2000s by revealing itself as more than a passive assembly platform, but a hub for precisely tuning the pathway input-output properties[237,238]. Inspired by the manner in which Ste5 recruits components into an organised complex and modulates the pathway response, in a landmark study in rewiring signal transductions pathways, Bashor *et al.* artificially altered the signalling dynamics of the mating pathway by overlaying synthetic feedback loops on the MAPK cascade[239] (**Figure 12**).

By recruiting the positive pathway modulator Ste50 (which promotes the interaction of Ste20 with Ste11) and the negative pathway modulator Msg5 (which dephosphorylates and deactivates Fus3) to the Ste5 scaffold via a leucine zipper docking interaction, they were able to tune the output from the yeast mating pathway. They then built synthetic feedback loops by dynamically recruiting the modulators to the Ste5 scaffold by expressing them from the mating responsive *FIG1* promoter. Stimulation of the mating pathway induced the expression of the modulators, which were then recruited to the Ste5 scaffold resulting in altered signalling dynamics and dose-response curves. They then demonstrated the tuneability of the system by either altering the recruitment affinity or inducible expression levels of the pathway modulators. These design principles were extended to various response behaviours such as a pulse generation, acceleration, delay, and ultrasensitivity by using different combinations of the pathway modulators and decoy units that compete for the binding site on Ste5.

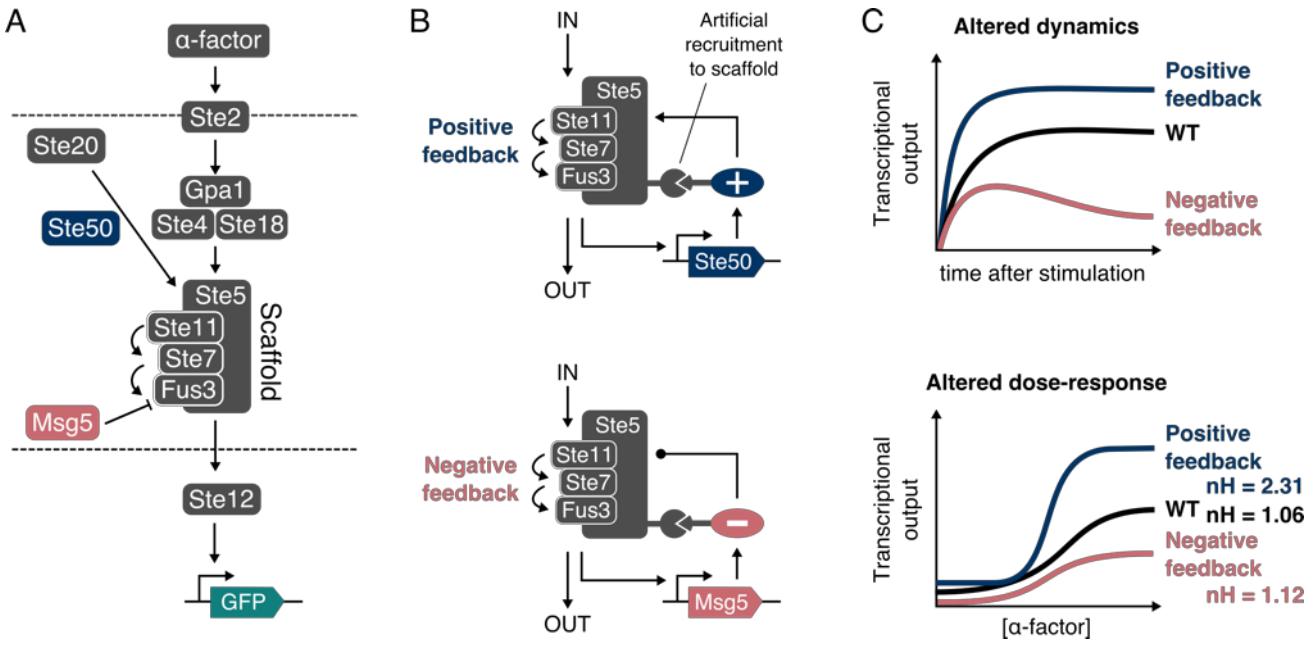


Figure 12. Tuning the output from the yeast MAPK pathway by artificially recruiting pathway modulators. (*A*) Ste50 (blue) positively modulates the MAPK cascade by promoting the interaction of Ste20 with Ste11. Msg5 (red) negatively modules the MAPK cascade by deactivating Fus3. (*B*) The MAPK modulators are expressed as an output of the pathway and artificially recruited to the Ste5 scaffold via a via a leucine zipper interaction, thus altering signalling dynamics and the dose-response characteristics independent of the receptor properties (*C*). Figure adapted from Bashor *et al.*[239].

Using pathway modulators without artificial recruitment, another two studies took a similar approach to manipulate the response of the mating pathway, programming bistability[240] and cell fate[241] in response to specified environmental signals. The successful and predictive manipulation of the MAPK cascade demonstrates the high level of plasticity inherent to complex signal transduction pathways. However, the presence of feedback mechanisms controlling one characteristic, such as the Hill slope, often makes it difficult to independently modify other signalling properties, such as the sensitivity[109]. Furthermore, these systems represent complexity overlaid onto complexity. As the mating pathway contains many of its own inherent feedback mechanisms to ensure faithful signalling, adding additional feedback would further obscure a system we have yet to achieve a full working model of.

An alternative strategy for fine-tuning GPCR activity is through the co-expression of receptor-interacting proteins[242]. In mammalian systems, the signalling and trafficking properties of GPCRs is often highly diverse depending upon the cellular context. This modification of GPCR function can, in part, be attributed to receptor-interacting proteins that are differentially expressed across different cell types. In some cases, these proteins act mainly as scaffolds to assist trafficking and interactions with other proteins, whereas in other cases, these proteins directly interact with the receptor to modulate signalling[242]. Some of these interacting proteins have been successfully ported into yeast to modulate the receptor pharmacology, such as the receptor activity-modifying protein-2 (RAMP2) which was able to significantly enhance the sensitivity and signal output of the glucagon receptor[216].

Other receptor modifying proteins, such as G protein-coupled receptor kinases (GRKs) and β-arrestins which have a role in the desensitisation of GPCRs[243], could also be used to alter the dynamics of GPCR signalling in yeast. This wide diversity of receptor interacting and activity modifying proteins offers a potential strategy for fine-tuning heterologous GPCR signalling yeast. However, as receptors have different specificities for different interacting proteins, this approach would require a significant effort to characterise all GPCR/interacting protein pairings. The toolkit would provide a method for tuning heterologous GPCR-based biosensing systems on an individual basis, rather than a generalisable set of rules for tuning the response. As these proteins also tend to affect the dose-response characteristics of a receptor in one direction, for example, increasing ligand sensitivity, a combinatorial approach may also be required to gain full tuneability in all directions.

Taken together, MAPK engineering and the use of modifier proteins to tune signalling properties both increase the complexity of the system and require the expression of additional components. Alongside the several components already required to generate a heterologous GPCR-based biosensor in yeast, including the receptor, G protein, and reporter, the addition of supplementary components would add further complication to a system that can already require significant effort to develop.

1.2.15 Refactoring natural biological systems

To gain control of a natural biological system, synthetic biologists occasionally take a systematic approach to completely defining its genetics by reconstructing the system from the bottom-up using well-characterised synthetic parts[244,245]. This strategy is known as “refactoring”, a term borrowed from computer science whereby existing computer code is restructured to improve an attribute (e.g. stability or readability) without changing functionality[246]. Refactoring can be used to reduce the complexity of the biological system by removing integrated host regulation and non-essential genes and allow manipulation by placing the components under synthetic control. These engineered surrogate systems can be used to probe fundamental biological questions or facilitate further engineering for a particular application.

The first study to take on this approach was the refactoring of the T7 phage[244]. The original design goals were to physically separate and enable unique manipulation of all primary genetic elements in the bacteriophage T7 genome. Although they never achieved full refactoring, replacing only a third of wild-type genome, the chimera encoded a viable bacteriophage that maintained key features of the original strain, while being more straightforward to model and easier to manipulate. Understanding what they were able to build and what they were not was a clear demonstration of the areas that need more research.

The next landmark study to take on this approach was the refactoring of the nitrogen fixation gene cluster from *Klebsiella oxytoca*[245] (**Figure 13**). This native gene cluster which converts atmospheric nitrogen into ammonia consists of twenty genes organised into seven operons, encoded over 23.5 kb of DNA. Their goal was to take the entire gene cluster and to systematically eliminate native regulation and replace it with synthetic genetic parts and circuits to produce a rewritten version of the gene cluster encoding the same function. First, they removed all non-coding DNA and TFs, and then recoded the sequence of all essential pathway genes so that they were as distant as possible from the WT sequence to remove internal regulation. The recoded sequences were then organised into artificial operons controlled by synthetic ribosome-binding sites and placed downstream of synthetic control circuits to control the dynamics of the system. The complete refactored gene cluster consisting of 89 genetic parts was synthesised and shuttled into *K. oxytoca*, successfully fixing nitrogen in strains deficient of the WT cluster. This refactored system now serves as a dynamic platform for addressing fundamental questions in basic biology.

Although these studies remain as two of the few true examples to fully embrace the approach of refactoring, the idea of simplifying and modularising natural biological systems to gain control remains a powerful one. As our understanding of basic biology improves and we expand our synthetic toolkit,

it is likely we will see this approach to the study and repurposing of natural systems used more widely in synthetic biology[245,247]. As of yet, no one has demonstrated an attempt to refactor an endogenous signalling pathway. This is surprising considering the principles that make the approach attractive for systems such as metabolic pathways (i.e. the removal of feedback and precise control of component levels) would also be extremely valuable for the study and engineering of signal transduction pathways[109].

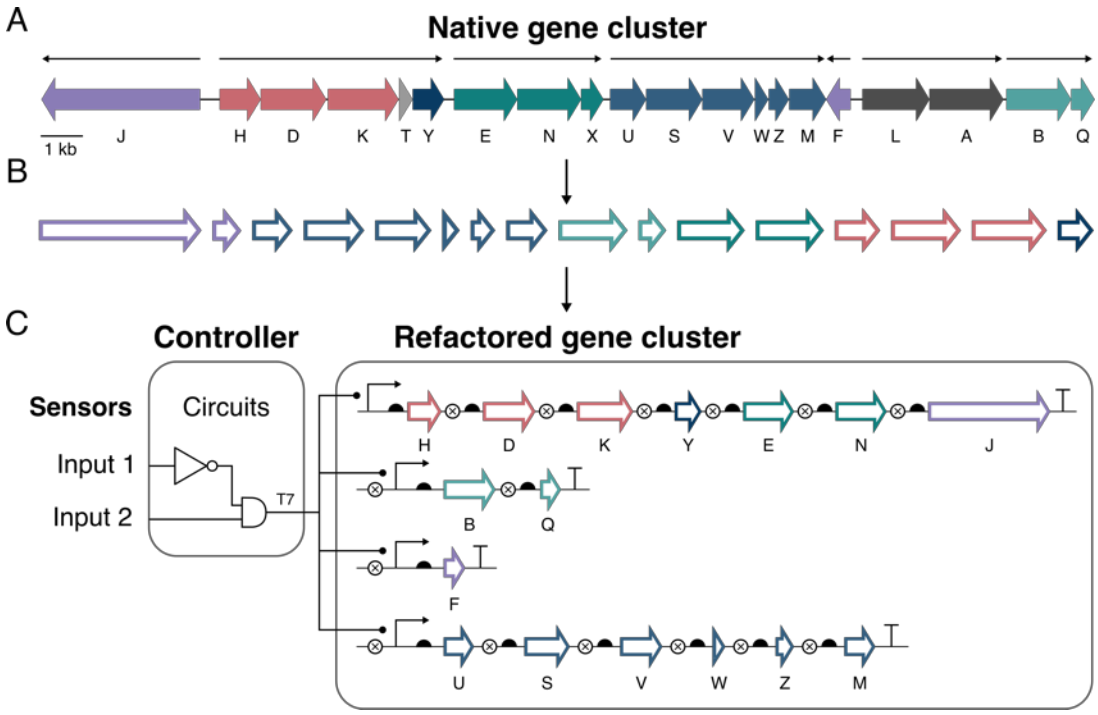


Figure 13. The process of refactoring the *K. oxytoca* nitrogen fixation gene cluster. (*A*) the native *K. oxytoca* nitrogen fixation gene cluster, coloured by gene function. (*B*) First, all non-coding DNA, non-essential genes, and transcription factors are removed, and the codons are randomised. (*C*) The recoded genes are then organised into operons and synthetic regulation is added to complete the refactored gene cluster. Figure adapted from Temme *et al.*[245].

1.3 Refactoring the pheromone response pathway

Refactoring the pheromone response pathway from the bottom-up to gain control of the system would represent a stark contrast from all previous efforts to manipulate this pathway, with past approaches being top-down (e.g. knockouts/mutations or overlaying complexity)[202]. Although there are far simpler endogenous signal transduction pathways that may be better suited for trialling the refactoring approach, the yeast mating pathway has benefits.

As a model eukaryotic signalling pathway, all the events required for the transduction of a signal from agonist to gene activation have been mapped out in great detail[179,202], allowing us to make educated decisions on the exact specification of the refactored pathway without requiring guesswork. We also have models of different aspects of the mating pathway and G protein signalling which can be employed at various stages of the refactoring to predict the system behaviour and guide the next step[248–250]. The genome engineering tools available to yeast are arguably more powerful than in any other organism[251], facilitating the large-scale modifications that would be required to strip out the endogenous pathway. Finally, the synthetic tools that have been developed for yeast are extensive and would enable the predictable reintroduction of the signalling components[33].

These attributes support the basis for refactoring the yeast mating pathway as an exercise in understanding fundamental biology. However, our goal is to gain control of the system for generating tuneable GPCR-based biosensors. Before we consider this approach, we need to identify where the sources of tuning could be achieved.

Firstly, it has been shown that in the absence of the RGS, Sst2, receptor number is proportional to the signal[252]. By varying the number of receptors in the system using variable strengths of promoter, this should provide the sensitivity dial. Secondly, by overexpressing the Gα, it is possible to reduce the basal expression of constitutive receptors[253,254]. Fine tuning the expression of the Gα should, therefore, provide the dial for leakiness. Finally, the levels of TF in a system often dictates the level of gene expression[255]. Titrating the TF copy number should provide the tuning dial for tuning the maximum output. Alternatively, the DNA binding domain (DBD) of the pheromone-responsive TF Ste12 could be substituted for an orthogonal DBD and redirected to a synthetic promoter[228,256]. This would allow the modulation of the maximum output by varying the number of TF binding sites and core promoter identity[257,258]. The use of orthogonal DBDs would also have the benefit of decoupling the signalling pathway from the 100+ genes normally upregulated in the mating response[259]. Gaining control of the entire pathway should, therefore, allow the tuning of several key dose-response properties using these three components.

Refactoring a genuinely minimal mating pathway to give insulated, unidirectional signalling that is free from feedback regulation and decoupled from the mating response would involve the deletion of 27 genes, 12 of which would then be reintroduced (**Table 2, Minimal**). This refactored pathway, comprising only the receptor, heterotrimeric G protein complex, MAPK signalling cascade, and TF complex, would not only represent an extensive amount of genome engineering but a tour de force in genetic design, as the pathway components likely require precision expression to achieve the desired response. To reasonably achieve our goals, this approach needs to be orchestrated in steps, where the first intermediate towards the fully refactored pathway has direct application to tuneable GPCR-based biosensing. This intermediate can then be studied to understand a reduced number of components before attempting to refactor the entire pathway, while also exploring strategies for tuning the dose-response for biosensing applications.

Table 2. Refactoring the minimal and minimised mating pathways. Descriptions from *Saccharomyces* Genome Database.

Gene	Description	Minimal	Minimised
STE2	α-factor GPCR receptor	**Refactored**	**Refactored**
SST2	Negative regulator of Gpa1	**Deleted**	**Deleted**
GPA1	G protein α subunit – interacts with Ste2/ Ste3	**Refactored**	**Refactored**
STE4	G protein β subunit – activates pathway in dimer with Ste18	**Refactored**	**Fixed**
STE18	G protein γ subunit – activates pathway in dimer with Ste4	**Refactored**	**Fixed**
STE20	Cdc42p-activated signal transducing kinase – activates Ste11	**Refactored**	**Fixed**
STE50	Promotes interaction between Ste20 and Ste11	**Deleted**	**Fixed**
STE5	Scaffold protein for the MAPK cascade	**Refactored**	**Fixed**
STE11	Signal transducing MAPKKK	**Refactored**	**Fixed**
STE7	Signal transducing MAPKK	**Refactored**	**Fixed**
FUS3	MAPK – activates various targets in the pathway	**Refactored**	**Fixed**
KSS1	MAPK – less active role than Fus3	**Deleted**	**Deleted**
FAR1	Induces cell cycle arrest in response to pheromone	**Deleted**	**Deleted**
DIG1	Negative regulator of Ste12 with Dig2	**Refactored**	**Fixed**
DIG2	Negative regulator of Ste12 with Dig1	**Refactored**	**Fixed**
STE12	Pheromone responsive transcription factor	**Refactored**	**Refactored**
MSG5	Protein phosphatase – minor role in deactivating Fus3	**Deleted**	**Fixed**
PTP2	Protein phosphatase - minor role in deactivating Fus3	**Deleted**	**Fixed**
PTP3	Protein phosphatase - major role in deactivating Fus3	**Deleted**	**Fixed**
BAR1	α-factor protease	**Deleted**	**Deleted**
STE3	**a**-factor GPCR receptor	**Deleted**	**Deleted**
MF(ALPHA)1	Mating pheromone α-factor	**Deleted**	**Deleted**
MF(ALPHA)2	Mating pheromone α-factor	**Deleted**	**Deleted**
MFA1	Mating pheromone **a**-factor	**Deleted**	**Deleted**
MFA2	Mating pheromone **a**-factor	**Deleted**	**Deleted**
GPR1	Glucose GPCR receptor	**Deleted**	**Deleted**
GPA2	G protein α subunit – interacts with Gpr1	**Deleted**	**Deleted**

We thus propose a partially refactored, minimised pathway as an intermediate towards the fully minimal design (**Table 2**, **Minimised**). This pathway has all of the extensive engineering to generate the insulated, unidirectional response that is free from feedback regulation and decoupled from the mating response but keeps a number of the core signalling components fixed (wild-type). By refactoring only the receptor, Gα, and TF, the number of components required in the initial refactoring is reduced to a more manageable number, while allowing us to delineate their contribution to the overall response as potential tuning knobs for developing GPCR-based biosensors. As these are also the components that are required for coupling a heterologous GPCR to a measurable output, this approach has the added benefit of providing an interchangeable platform for creating new biosensors in a "plug-and-play" manner (**Figure 14**).

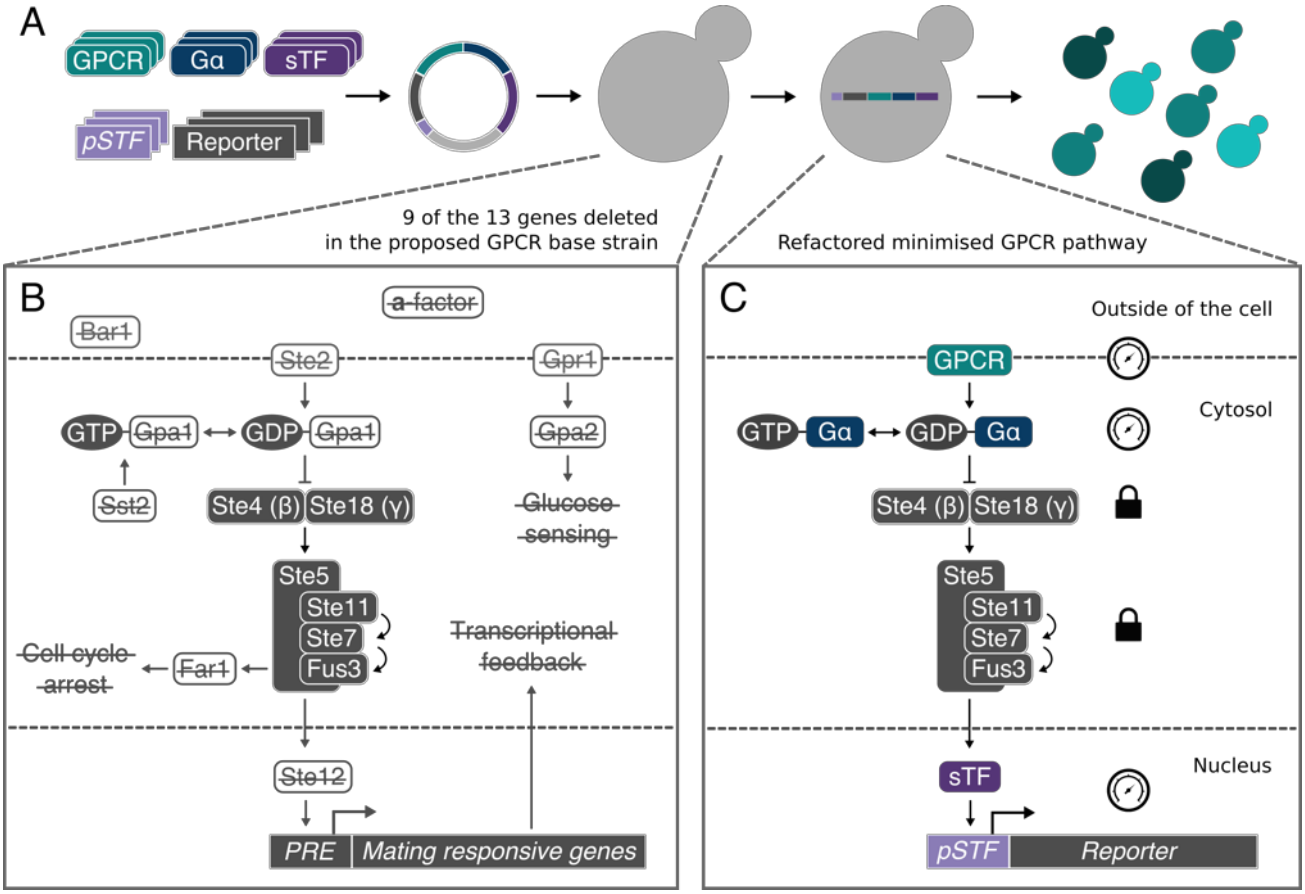


Figure 14. Overview of the proposed platform for developing GPCR-based biosensors in yeast. (*A*) Workflow for generating a GPCR-based biosensor. New biosensor designs are generated by assembling parts from a library of signalling components into a single multigene cassette. The newly assembled constructs are then linearised, transformed and integrated into the genome of the base strain to reconstitute a minimised GPCR signalling pathway. (*B*) The proposed modifications to the pheromone response and glucose sensing pathways in the GPCR base strain, leaving only the Gβγ and core signalling elements of the MAPK cascade, as described in **Table 2, Minimised**. (*C*) A reconstituted signalling pathway, incorporating a heterologous GPCR coupled to the pathway via a chimeric Gpa1-Gα subunit (Gα), and the output of the pathway redirected through a synthetic transcription factor (sTF) to an orthogonal promoter.

The proposed “base” strain for refactoring the minimised GPCR signalling pathways consists of 14 gene disruptions, serving 7 key purposes: (i) to remove negative feedback within the signalling pathway (Sst2), (ii) to avoid unwanted cell-cycle arrest (Far1), (iii) to prevent α-factor signal degradation (Bar1), (iv) to route all signalling through a single MAPK (Kss1), (v) to remove all mechanisms for pheromone-mediated cell-cell communication (α-factor, **a**-factor, and Ste3), (vi) to insulate heterologous GPCR receptors from alternative signalling pathways (Gpr1 and Gpa2), and (vii) to be refactored using synthetic tools (Ste2, Gpa1, and Ste12).

The protein phosphatases, Msg5, Ptp2, and Ptp3, are fixed in the proposed strain as they act downstream in the pathway, playing a minimal role in changing signalling dynamics, and will reduce unwanted basal pathway activity in the absence of a stimulus[179]. Ste50 is also not essential for pathway activity, but enhances signal transduction via Ste20 and so is also fixed[179]. The remaining components that have been left wild type comprise the essential elements of the minimal signalling pathway: the Gβγ (Ste4 and Ste18), MAPK signalling cascade (Ste20, Ste5, Ste11, Ste7, and Fus3), and the negative repressors of Ste12 (Dig1 and Dig2).

Finally, once the proposed GPCR base strain has been created, a genetic system is required for reintroducing the receptor, Gα, and TF, alongside a reporter. This system should provide several key features that will be essential for effective pathway refactoring: (i) chromosomal integration of the genetic constructs to ensure stable and robust expression[260,261], (ii) well-characterised regulatory elements for predictably fine-tuning the expression of the refactored components[245], and (iii) a modular and standardised assembly method to create complex genetic constructs from parts[37]. Although several yeast genetic toolkits meet these criteria, including the yGG[262] and YeastFab[263] systems, the Yeast MoClo Toolkit (YTK) system by Lee *et al.*[33] stands out as the toolkit of choice due to the extensive list of parts and the thoroughly characterised library of promoters and terminators.

In this proposed platform, new biosensor designs are created by refactoring the receptor, Gα, and TF alongside a reporter on a single multigene plasmid that is then integrated into the genome of a single, highly-engineered GPCR base strain, giving full control over the components chosen and their level of expression.

1.4 Aims and objectives

Yeast GPCR-based biosensors present themselves as a promising tool for a wide variety of applications in synthetic biology, such as an analytical tool for supporting metabolic engineering projects. However, we have yet achieved a generic set of rules for tuning the dose-response properties, limiting their wide-spread application. While previous work to manipulate the pheromone response pathway has involved a top-down approach, such as overlaying complex feedback loops, no one has yet attempted a bottom-up approach of refactoring to gain control of the pathway properties. The aim of this PhD project was, therefore, to refactor a minimised signalling pathway to create a platform for rationally tuning GPCR-based biosensors in yeast.

Firstly, in Chapter 2, we set out to create the extensively engineered base strain for refactoring the minimised GPCR signalling pathway. To achieve the large number of genomic modifications necessary for developing this strain, we explore novel CRISPR-based strategies to iterate through successive edits rapidly. Additional genetic features were also pursued to enable the downstream engineering of the base strain and integration with the YTK system to facilitate rapid and predictable refactoring of the minimised pathway.

In Chapter 3, we systematically refactor the endogenous receptor, Gα, and TF using synthetic tools to understand how their expression levels can affect the dose-response of the signalling pathway. By treating the cell as an "*in vivo* model", we delineate the contribution of each component by individually varying their expression using constitutive promoter libraries.

In Chapter 4, we explore the use of synthetic transcription factors for redirecting the pathway response to orthogonal promoters. By exchanging the DNA binding domain of the pheromone-responsive transcription factor Ste12 for synthetic alternatives, we demonstrate how the pathway output can be manipulated using modular promoters and secondary inducers. We then use these tools for optimising the response of the minimised mating pathway to its native ligand.

Finally, In Chapter 5, we combine the lessons learned in the previous chapters to the design and implementation of heterologous GPCR-based biosensors, using receptors sensitive to peptides, metabolites, and hormones relevant to human health. We then address the limitations of single cells by moving towards community-based systems for tuning the remaining properties of the dose-response curve.

1.5 Note to the reader

Before beginning the results chapters, there are a couple of items to clarify in order to understand the structure and other aspects of this thesis.

Firstly, a significant amount of work has been appended to the supplementary information, including a small number of tables and figures, as well as two additional sections. Although this work was essential to this thesis, it has been left out of the core chapters so that it does not distract from the main story. The first section covers updates we have made to the YTK system over the course of this project to improve the general usability and integration with a CRISPR toolkit which we describe in the second section. These sections are not crucial to the understanding of the following chapters but are included as a reference that will be necessary for reproducing or extending this work.

Secondly, all of the computational modelling presented in this work was developed and performed by Hitoshi Yamauchi, Jack Mead, and Graham Ladds from the Department of Pharmacology at the University of Cambridge as part of a collaboration, details of which can be found in Shaw *et al.*[264]. These models support the experimental findings within this thesis and are referenced appropriately wherever used.

We hope that by presenting the work in this manner, the thesis follows a single coherent story that was outlined in the aims and objectives.

Publications resulting from this work

Shaw, W. M., Yamauchi, H., Mead, J., Gowers, G-F., Öling, D., Larsson, N., Wigglesworth, M., Ladds, G., and Ellis, T. Engineering a model cell for rational tuning of GPCR signaling. *BioRxiv* 1–59 (2018).

2 A highly-engineered base strain for GPCR-based biosensing

2.1 Introduction

In section **1.3**, we proposed a highly-engineered yeast strain to enable the refactoring of a minimised signalling pathway for creating tuneable GPCR-based biosensors. The base strain comprises 14 gene disruptions to the yeast mating and glucose sensing pathways to remove feedback regulation, prevent cross-talk, reduce non-essential components, and to provide a null background for refactoring the receptor, Gα, and TF. Previous efforts to manipulate this system have relied on traditional recombination-based methods, such as the *Dellito Perfetto* approach[265], to generate the several gene disruptions necessary for effective coupling of a heterologous GPCR to the endogenous yeast machinery (as described in section **1.2.11**)[55,166,228,229]. However, while these genome engineering methods have been a major driving force behind yeast molecular biology, they rely on the integration and eventual removal of markers to iterate through successive edits[185]. As a single edit can take between 1-2 weeks, the application of these technologies to the development of the proposed GPCR base strain would be far too laborious and time-consuming.

CRIPSR-mediated genome engineering offers a promising alternative for site-directed mutagenesis in yeast, with a large number of groups reporting the highly-efficient, simultaneous editing of multiple genomic targets, without the need for integrating a marker[266]. By maintaining the marker on a non-integrative plasmid, yeast cells can easily be cured, thereby creating a clean strain containing the multiple edits, ready for application. Using this approach, a single round of editing can be achieved in a timeframe similar to the *Dellito Perfetto* approach, but with a throughput several times greater. However, while these CRISPR-based systems have focused heavily on the multiplexing of genomic edits, they are often geared toward the immediate use of the modified strains, with no concerns over further engineering[266]. The requirement for plasmid curing prior to additional editing and the limited number of markers available in these systems precludes the rapid iteration required for more ambitious genome engineering projects.

In this chapter, we investigate novel strategies to accelerate the iteration of CRIPSR-mediated editing to support the development of the proposed GPCR base strain, while formalising an approach for future yeast engineering projects. We also take special consideration into the downstream engineering of the GPCR base strain by exploring genetic features to allow the multiplexed (re)introduction of genes and improve the compatibility with the YTK system. Finally, we characterise the suitability of the base strain for GPCR-based biosensing and create a reference strain for benchmarking the performance of future pathway designs.

2.2 Results and discussion

2.2.1 The design of the GPCR base strain

Before exploring strategies for iteratively editing of the yeast genome, we needed to define the specifications of our GPCR base strain. We chose the S288C derivative BY4741[267,268] (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) as our parental strain for several reasons: (i) MATa genotype to keep the base strain consistent with previous systems[180], (ii) compatibility with the integration vectors in the YTK toolkit[33], (iii) as a derivative of the S288C reference strain, high quality sequence data is available[269], and (iv) Kss1 is non-functional[270], requiring one fewer deletion to achieve the final base strain.

To improve the downstream engineering of the final strain, we designed the gene KOs as a precise substitution of the open reading frame (ORF) for an addressable 24 bp sequence comprising a unique CRISPR/Cas9 target (**Figure 15A**). If required, the deleted genes could then be reintroduced in any combination by targeting these "landing pads" with CRISPR/Cas9 to rapidly access a desired genotype. Additionally, the landing pads could also be used for the introduction of gene variants to study mutants in their native context or introduce heterologous genes to study their expression from the endogenous regulatory elements (**Figure 15B**). Landing pads were also designed to sit between the two regions of homology used by the chromosomal integration vectors in the YTK system to improve the efficiency of routine plasmid integration using CRISPR/Cas9.

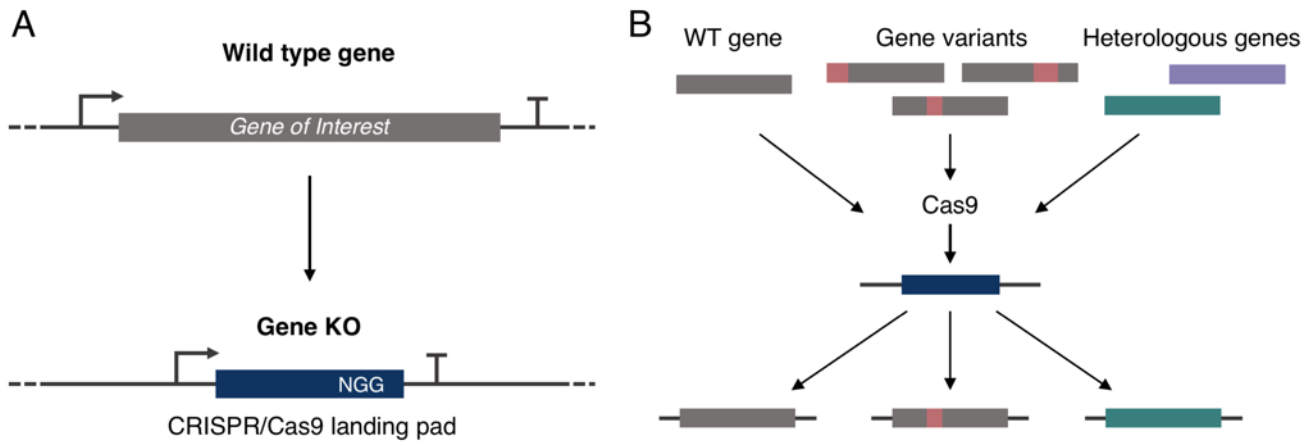


Figure 15. CRISPR landing pads for the (re)introduction of genes. (*A*) To facilitate downstream editing, genes are deleted by precisely substituting their ORF with an addressable 24 bp sequence containing a unique CRISPR/Cas9 target (landing pad). (*B*) In the final base strain, these landing pads can be targeted for the efficient reintroduction of genes to restore the WT genotype, introduction of gene variants to combinatorially assess mutants in the context of their WT regulation or introduce a heterologous gene between the native regulatory elements of the deleted gene. Landing pad sequences and predicted on-target scores can be found in **Table 3**. Cas9 was chosen as the CRISPR endonuclease for the landing pad system as the on- and off-target score calculators existed only for Cas9 at the time of design.

The landing pad sequence of each edit was designed in Benchling using the "gRNA Design Tool" to optimise the on-target score (for efficient CRISPR/Cas9 targetting[271]) and the off-target score (to reduce sequence similarity with the genome) of biologically neutral 20 bp sequences[272]. On-target scores were optimised to achieve a minimum score of 50 (higher the better), and off-target scores were all optimised to be 100 (no predicted off-target activity). Including the 13 gene KOs and the additional installation of 3 landing pads to support the CRISPR-aided integration of YTK plasmids, 16 edits were required to generate the final base strain design (**Table 3**).

Table 3. List of edits and landing pad sequences in the final GPCR base strain design.

Edit	Location	CRISPR LP sequence	On-target score	Off-target score
***SST2* KO**	chrXII	AATGCAATCGTAGTCCACCTCGG	71	100
***FAR1* KO**	chrX	GATCGTACTTAGAAATGAGGCGG	65	100
***BAR1* KO**	chrIX	AATGGGGTTAGCAAGTCGCACGG	68	100
***STE2* KO**	chrVI	CTAGCTTTCGTGTTAGTACGCGG	60	100
***GPA1* KO**	chrVIII	TAGCATGGTGACACAAGCAGCGG	76	100
***STE12* KO**	chrVIII	CATCGCTTCCTACTTCCGCTCGG	59	100
***STE3* KO**	chrXI	AATGTTTCTTGTCCAAGCGGCGG	60	100
***MF(ALPHA)1* KO**	chrXVI	ACACGAGTTCCCAAAACCAGCGG	73	100
***MF(ALPHA)2* KO**	chrVII	GTTCCGATAGGCCAGCATATCGG	55	100
***MFA1* KO**	chrIV	GCAGTAACGCTCATCAGCTACGG	53	100
***MFA2* KO**	chrXIV	CTTCTCCTGGAGATCAAGGACGG	59	100
***GPR1* KO**	chrIV	TCTAACCGTCGACTTTGGCGCGG	65	100
***GPA2* KO**	chrV	GCTGTTATCCTGCATCGGAACGG	65	100
***URA3* LP**	chrV	ATATTATTGTACACCTACCGCGG	73	100
***LEU2* LP**	chrIII	GCATCAGGTGGACTAGCATGCGG	71	100
***HO* LP**	chrIV	ATGGACGAAATGCTTCACCACGG	70	100

2.2.2 Multiplexed CRISPR-mediated genome engineering

During the early stages of this work, we developed a CRISPR toolkit for the efficient and flexible multiplexed editing of yeast. The CRISPR toolkit consists of a modular system for editing using the well-known Cas9 endonuclease from *Streptococcus pyogenes* and the increasingly popular alternative Cpf1 endonuclease from *Lachnospiraceae bacterium*. We will not cover the details of the system here as these are not important to the main story and would serve as a distraction. Instead, we direct the reader to section **8.4** for a detailed explanation of the CRISPR toolkit.

To determine the multiplexed editing efficiencies of the CRISPR tools for precisely substituting the ORF of selected genes for a new sequence, we developed an easy-to-measure efficiency assay. Two guides were targeted within the ORF of the first two genes in **Table 3**, *FAR1* and *SST2*, as close as possible to the 5' and 3' ends. Donor DNA containing the in-frame coding region of sfGFP and mRuby2, flanked by 500 bp of direct homology to the up- and downstream regions of the two genes, was then used as a template for repair (**Figure 16A**). Successful deletion of the ORFs of both genes would, therefore, result in the expression of the two fluorescent proteins, providing a measurable indicator of successful gene deletion. 24 transformants were randomly picked and run on a flow cytometer to determine the fraction of cells that had successfully substituted the ORF of the two genes for the coding region of the two fluorescent proteins, as measured by a positive green and red fluorescence (**Figure 16B**).

At the time of performing cell line development, we had only created the tools for Cas9 editing using a single method of plasmid delivery (**Figure 16C**). However, as these tools were updated to include Cpf1 and an alternative gRNA delivery method, experiments were performed *ex post facto* to

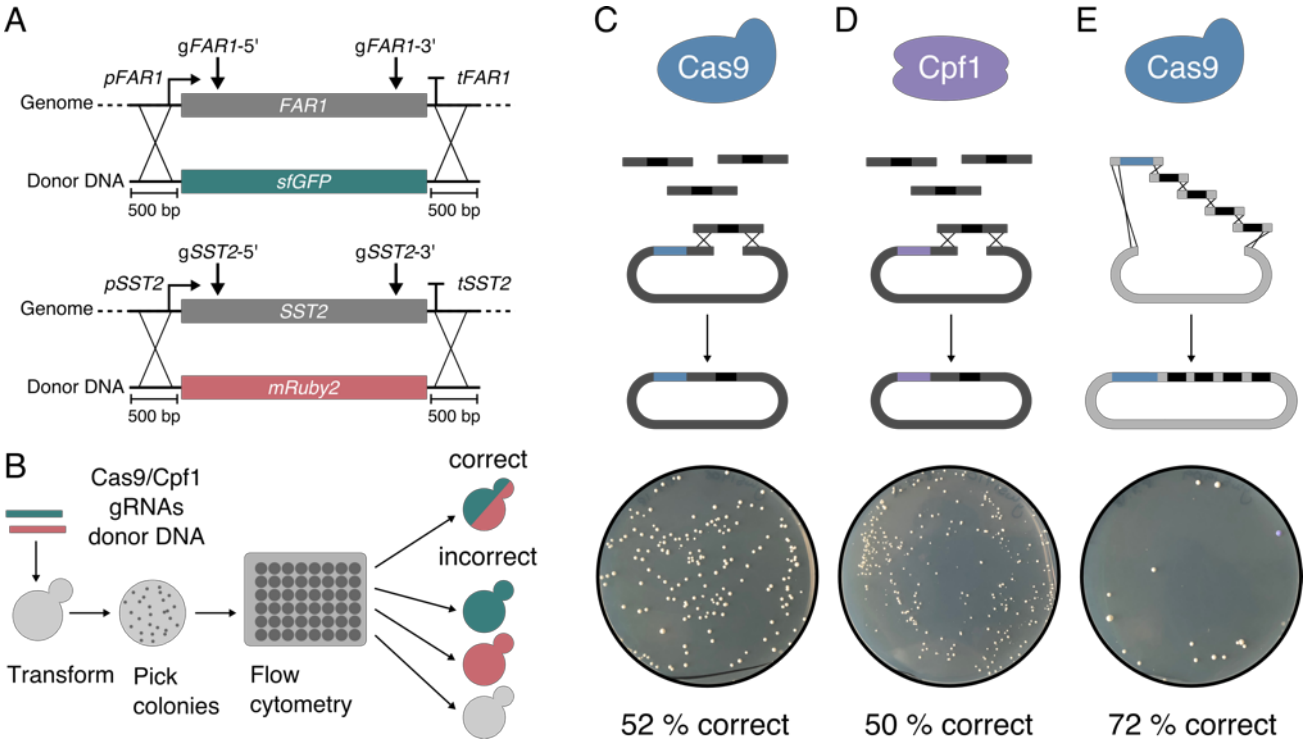


Figure 16. Strategies for multiplexed CRISPR-mediated editing. (*A+B*) Overview of the assay for determining multiplexed CRISPR editing efficiency. (*A*) CRISPR-mediated substitution of the *FAR1* and *SST2* ORFs for the fluorescent proteins sfGFP and mRuby2, respectively. (*B*) Successful deletion of the *FAR1* and *SST2* ORFs in clonal populations were determined by measuring the green and red fluorescence of individual colonies using flow cytometry. (*C-E*) Three distinct CRISPR strategies for multiplexing edits using either Cas9 (blue) or Cpf1 (purple) with the standard (dark grey) or alternative (light grey) plasmid delivery methods in the CRISPR toolkit (see section **8.4** for a detailed description of the toolkit). 24 random colonies from each strategy were measured to determine their efficiency for double gene deletion. Blue colonies represent incorrectly assembled CRISPR plasmid and were avoided.

determine the editing efficiencies of all CRISPR approaches and provide a reference for future engineering projects (**Figure 16D+E**). Over 50% of the cells edited using the first CRISPR method were correct for the double ORF substitution, indicating this approach would be appropriate for our purposes. The updates to the toolkit demonstrated a similar editing efficiency when using the alternative CRISPR endonuclease, Cpf1 (**Figure 16D**), and improved editing efficiencies when using the alternative plasmid delivery method (**Figure 16E**).

We decided the method of editing in a pairwise manner and achieving efficiencies around 50% would be preferable over pursuing marginal performance gains or stretching the system to its limits. This would allow us to screen a small number of colonies during each round, as well as providing a buffer for fluctuations in the editing efficiencies. Instead, we explored strategies for reducing the time required between successive rounds of editing to accelerate genome engineering projects.

2.2.3 CRISPR marker cycling for rapid iterative editing

Using the standard approach of editing, validating, and then curing the cells of the CRISPR plasmids, it would take around 1-2 weeks per round of editing. With edits performed in a pairwise manner, iterating through the 16 modifications in the final base strain design would take up to 16 weeks to complete. We hypothesised that given a set of CRISPR plasmids with different selectable markers we could skip the plasmid curing step, thus halving the time taken to complete each round of editing. By transforming the cells with a different marker and changing the selection, the edits would be performed by the CRISPR machinery, maintained by the new selection pressure, and the previous plasmids would eventually be lost due to an absence of a selective pressure. The markers could then be cycled at a periodicity that would ensure the previous markers had been lost before they were used again. If successful, this marker cycling strategy would enable rapid and unlimited iterative editing of yeast.

To determine whether this approach was viable, we adapted a simple GFP to BFP conversion assay that would allow the iterative editing of a measurable phenotype[273] (**Figure 17**). In this assay, a chromosomally integrated GFP reporter can be converted to BFP by the substitution of two amino acids (T65S and Y66H). By targeting CRISPR/Cas9 to this site and supplying donor DNA to switch the protein identity, the editing efficiency can be determined by measuring the number of cells which have converted from green to blue fluorescence using flow cytometry. This process can then be repeated by targeting the new BFP coding sequence and converting it back to GFP. Assuming faithful editing, this process can be iterated indefinitely.

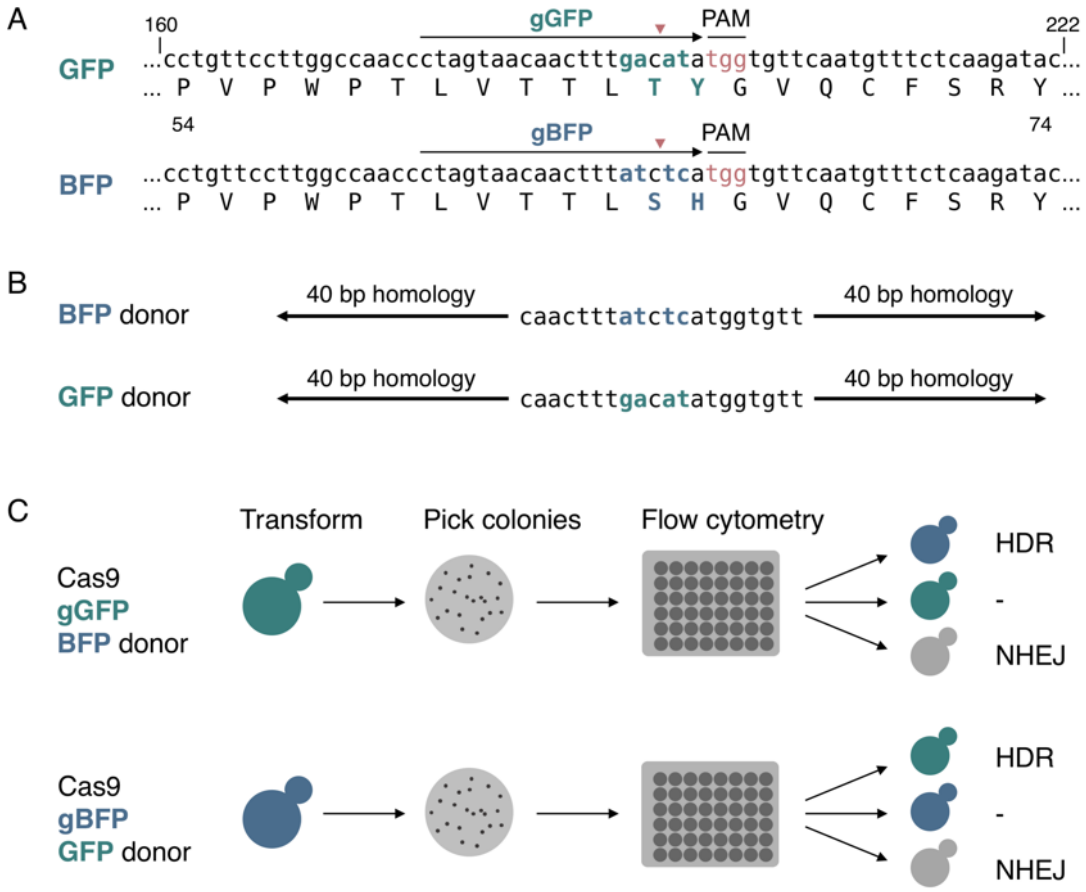


Figure 17. A reversible GFP to BFP conversion assay for iterative CRISPR/Cas9-mediated editing. (*A*) Sequence alignment between the chromophore regions of GFP and BFP. A double T65S and Y66H amino acid substitution in GFP corresponds to a significant shift in the excitation and emission spectra of the protein, converting it to BFP, and vice versa. Cas9 is targeted to the region to generate a double stranded break (DSB) at the site of difference (red indicators) using unique and orthogonal guides specific for either GFP or BFP. (*B*) Double-stranded donor DNA created by overlap extension PCR to repair the DSB in GFP and BFP, generating BFP and GFP, respectively. (*C*) Scheme for converting GFP to BFP and vice versa. Random colonies transformed with Cas9, the gRNA, and the appropriate donor DNA are picked and measured for their green and blue fluorescence by flow cytometry. A switch in fluorescence represents a successful edit, no change fluorescence represents an unsuccessful edit, and no fluorescence represents a mutation caused by either non-homologous end joining (NHEJ) or erroneous donor DNA.

Using three auxotrophic markers (*URA3*, *LEU2*, and *HIS3*), we performed the iterative GFP to BFP conversion assay for a total of six rounds, comprising two full cycles of the marker set (**Figure 18A+B**). The cycle was structured to finish on the *URA3* marker so that once all six rounds of editing had been completed, the final CRISPR plasmids could be cured from the cells using 5-FoA counter selection. We plated the transformed yeast onto the selective media of all three markers during each round of editing to determine the retention of the previous plasmids in the cycle (**Figure 18C**). Typically, the previous marker in the cycle would yield approximately 10% of the colonies of the current marker, demonstrating a two-marker set would not be sufficient for iterative editing. However, no colonies ever resulted from the marker used two cycles before the current marker, thus demonstrating a set of three markers is sufficient for iterative editing.

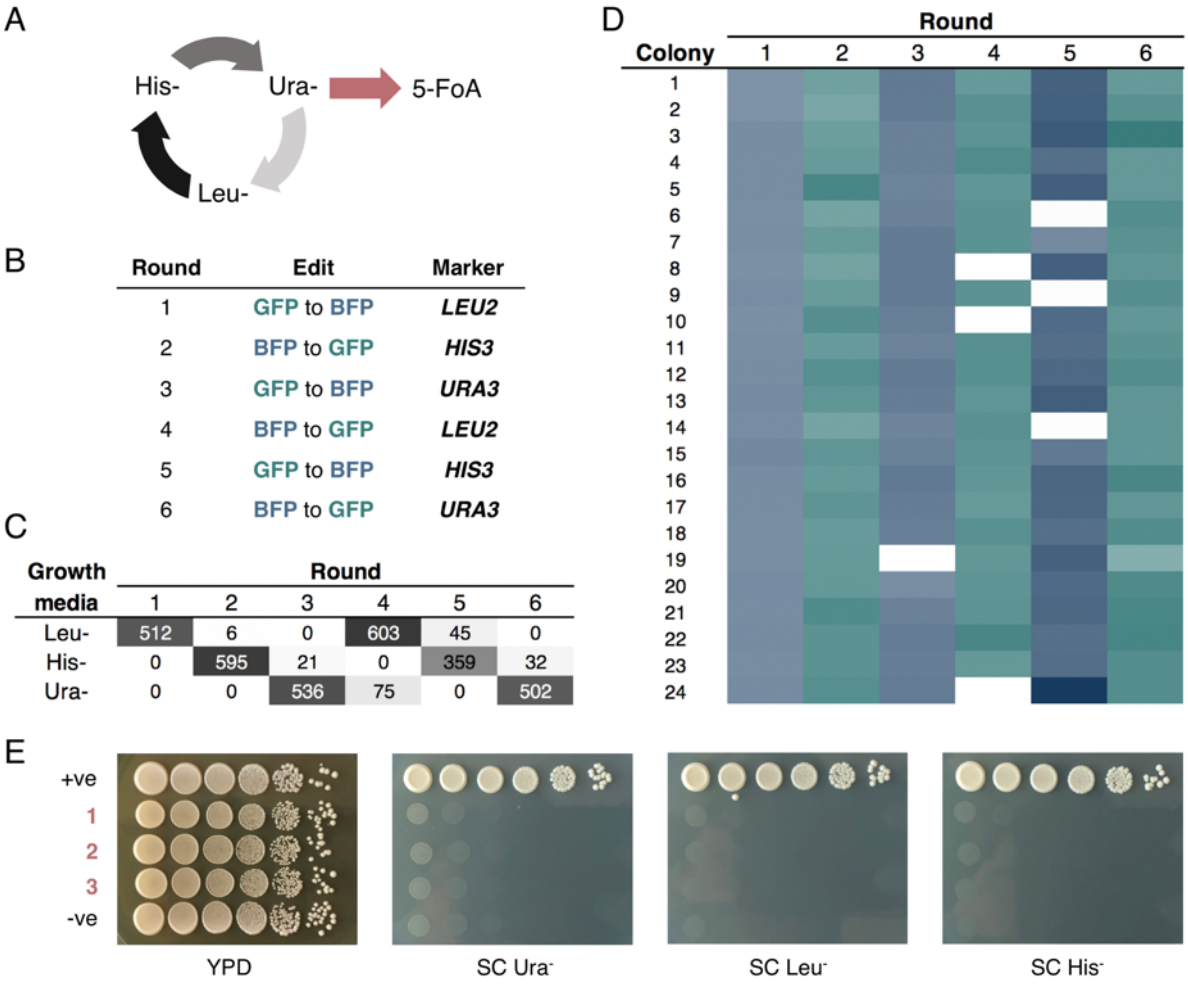


Figure 18. Iterative CRISPR/Cas9 editing using marker cycling. (*A*) Media conditions for the iterative CRISPR/Cas9 marker cycling scheme using a set sequence of three auxotrophic markers (*URA3*, *LEU2*, and *HIS3*). Ending the cycle on the *URA3* marker facilitates the removal of the CRISPR plasmids from cells using 5-FoA counter selection. (*B*) Marker cycling scheme using GFP to BFP conversion to assay editing efficiencies and the loss of marker plasmids during 6 rounds of editing (two full cycles). (*C*) Colony counts on selective media after transformation during each round of editing. (*D*) GFP and BFP fluorescence of 24 randomly picked colonies from each round of editing. Results are displayed as a colour change ranging from green (GFP fluorescence) through white (no fluorescence) to blue (BFP fluorescence). (*E*) Spot test of 3 round 6 strains on rich (YPD) and selective media (SC dropout) after counterselection on 5-FoA media to confirm the loss of all CRISPR plasmids. Negative cells are WT BY4741 (*his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*). Positive cells are BY4741 with the *URA3*, *LEU2* and *HIS3* genes restored. Experimental measurements are sfGFP and mTagBFP2 levels per cell determined by flow cytometry.

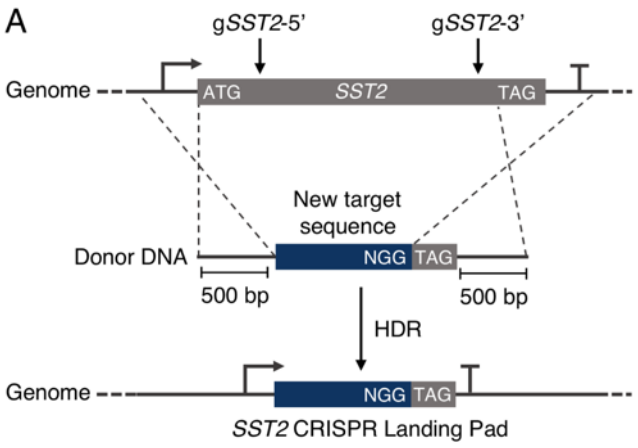
To determine the editing efficiency at each round, we picked 24 random colonies from the current marker plate and ran them on the flow cytometer to determine their fluorescence (**Figure 18D**). 100% of the colonies picked demonstrated the fluorescent protein had been edited. However, some of these resulted in the loss of fluorescence (white), suggesting either non-homologous end joining or faulty donor DNA. This indicated the CRISPR/Cas9 generated double-stranded break (DSB) was efficient during the six rounds, but extra attention would be required when creating and supplying donor DNA in the future. After validating a switch in fluorescence, the first correct colony was then back diluted in preparation for the next transformation.

After completing the six rounds of editing, the first three colonies in the last round were back diluted, cultured overnight in non-selective media, and plated onto 5-FoA media to counter select the CRISPR plasmids using the *URA3* marker. A single colony from each plate was then grown up in rich media (YPD) and spotted onto YPD and the selective media of all three markers to confirm the loss of all CRISPR plasmids (**Figure 18E**). No growth was detected on the selective media for any of the replicates, demonstrating the back culturing and then growth on 5-FoA was sufficient for losing the previous markers in the cycle and counter selecting the *URA3* plasmid.

This marker cycling strategy demonstrated it was possible to efficiently iterate CRISPR-mediated edits using a three-marker plasmid set and then cure the cells of all plasmids to generate an extensively engineered, clean strain. Furthermore, this approach allowed us to perform a new edit every four days. With eight rounds of editing required to create the base strain, this would represent five weeks of work. With this new protocol, we progressed on to the development of the base strain.

2.2.4 Engineering the GPCR base strain

Before performing the first edit in the cell line development of the GPCR base strain, we designed and created all of the necessary donor DNAs and gRNAs required for the 16 different edits. The donor DNA was designed using the unique CRISPR targeting sequence (**Table 3**) flanked by 500 bp of up- and downstream homology to the target gene or YTK integration locus (**Figure 19A**). gRNAs were



B

Round	Edit	Marker
1	*SST2* + *FAR1* KO	***HIS3***
2	*BAR1* + STE12 KO	***URA3***
3	*STE2* + *GPA1* KO	***LEU2***
4	*GPR1* + *GPA2* KO	***HIS3***
5	*MF(ALPHA)1* + *MF(ALPHA)2* KO	***URA3***
6	*MFA1* + *MFA2* KO	***LEU2***
7	*STE3* KO + *URA3* LP	***HIS3***
8	*LEU2* + *HO* LP	***URA3***

5 weeks

Figure 19. Iterative editing of BY4741 to generate the GPCR base strain, yWS677. (*A*) Genes were deleted by precisely substituting their ORF with an addressable 24 bp sequence containing a new and unique CRISPR/Cas9 targeting sequence (landing pad). Landing pads for YTK plasmid integration were generated by introducing the target sequence between the two arms of integration homology. Landing pad sequences and predicted on-target scores can be found in **Table 3**. (*B*) Order of the pair-wise edits to generate the GPCR base strain over a 6-week period, finishing on the *URA3* marker for counter selection of the CRISPR plasmids.

then designed to cut the genome between the two arms of homology to facilitate homology-directed repair with the donor DNA and installation of the CRISPR landing pads (See **8.5.2** for a list of gRNAs used in this study).

The edits were organised in a pairwise manner, eventually taking five weeks to generate the final GPCR base strain, yWS677 (BY4741 *sst2Δ0 far1Δ0 bar1Δ0 ste2Δ0 ste12Δ0 gpa1Δ0 ste3Δ0 mf(alpha)1Δ0 mf(alpha)2Δ0 mfa1Δ0 mfa2Δ0 gpr1Δ0 gpa2Δ0)* (**Figure 19B**). 12 colonies were screened during each round using colony PCR as an initial confirmation of a successful edit, yielding at least one correct colony at all stages. The locus of each confirmed edit was then sent for verification of identity by direct Sanger sequencing, while the next round of editing took place (See section **8.5.5** for a list of primers used in this study). All 16 edits were confirmed by exact alignment with the expected sequence, as derived from the S288C reference (see **8.5.4** for the Sanger sequencing results of 16 modifications in yWS677). Following the validation of the final two edits, yWS677 was cured of the CRISPR plasmids using 5-FoA counter selection of the final *URA3* marker, as previously described. We confirmed for the loss of all plasmids by absence the of growth on selective media and colony PCR.

Finally, we used long-read nanopore sequencing to confirm the identity of the entire genome and determine whether any large-scale mutations, such as inversions, deletions or insertions, had occurred outside the boundaries of the locus sequencing (**Figure 20**). The sequencing reads from a single Nanopore run were sufficient to assemble contigs de-novo, representing the full set of 16

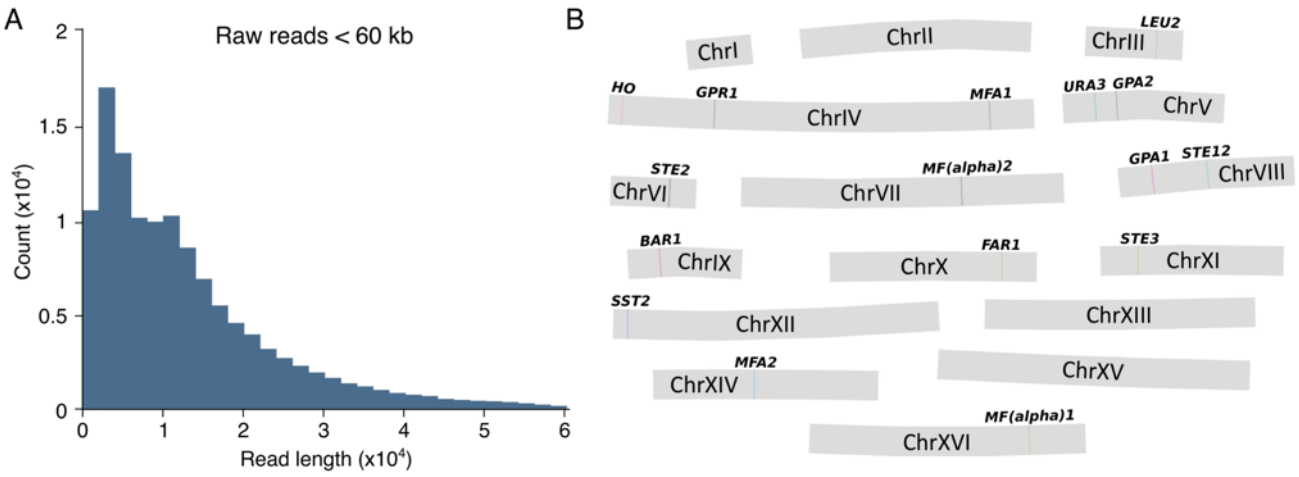


Figure 20. De-novo assembly of the yWS677 genome from Nanopore sequencing. (*A*) Read length histogram for the Nanopore sequencing run, displaying all reads below 60 kb. (B) De-novo contigs assembled using SMARTdenovo from reads corrected by Canu, representing the full set of 16 chromosomes from *Saccharomyces cerevisiae*, confirmed by exact alignment to S288C reference genome using a minimum alignment of 100 bp. All discrepancies with the reference genome are highlighted and correspond to the 16 edits described in **Table 3**. No other discrepancies were detected, suggesting precise and clean CRISPR/Cas9 editing during the 8 rounds.

chromosomes from *Saccharomyces cerevisiae*, confirmed by alignment to the S288C reference genome. We then probed all discrepancies between the yWS677 genome and the S288C reference using a minimum alignment length of 100 bp, identifying 16 areas that did not align, exactly corresponding to the expected changes we had made (**Table 4**). The absence of any further discrepancies confirmed no large-scale genome mutations had occurred, indicating precise editing during the five weeks of iterative CRISPR-mediated genome engineering.

Table 4. Expected changes and confirmation of their positioning in the yWS677 genome.
Note, all alignments are approximately 1000 bp, as this was the size of the donor DNA transformed, except for *STE3*. Due to cloning issues with the *STE3* KO donor DNA, a smaller fragment generated by overlapping oligo PCR was used instead.

CRISPR donor DNA	de novo contig ID	Query identity	Alignment length	Query coverage
SST2	chrXI	99.02%	1,024	100.00%
FAR1	chrX	98.93%	1,026	100.00%
BAR1	chrIX	98.73%	1,024	100.00%
STE2	chrVI	99.30%	993	100.00%
GPA1	chrVIII	98.72%	1,018	100.00%
STE12	chrVIII	98.68%	987	100.00%
STE3	chrXI	98.06%	103	100.00%
MF(ALPHA)1	chrXVI	99.22%	1,024	100.00%
MF(ALPHA)2	chrVII	98.73%	1,024	100.00%
MFA1	chrIV	98.05%	1,024	100.00%
MFA2	chrXIV	98.34%	1,025	100.00%
GPR1	chrIV	97.66%	1,025	100.00%
GPA2	chrV	91.64%	1,040	100.00%
URA3	chrV	98.15%	1,024	100.00%
LEU2	chrIII	98.82%	1,014	100.00%
HO	chrIV	99.61%	1,024	100.00%

2.2.5 Characterising the yWS677 GPCR base strain

After confirming the sequence identity of yWS677, we compared its growth to the BY4741 parental strain in rich (YPD) and synthetic complete (SC) media (**Figure 21**). yWS677 exhibited a very slight reduction in the maximum growth rate in both conditions, likely due to the constitutive activation of the MAPK signalling cascade in the absence of the Gα, Gpa1[274]. Although the Ste12 TF had also been removed, preventing a cellular response to pathway activation, the signalling through the MAPK alone would spend energy[179]. However, this slight defect was unnoticeable when handling the strain compared to BY4741. Additionally, refactoring of the base strain to produce a GPCR-based biosensor would reintroduce the Gα, preventing constitutive MAPK signalling in the final biosensor designs.

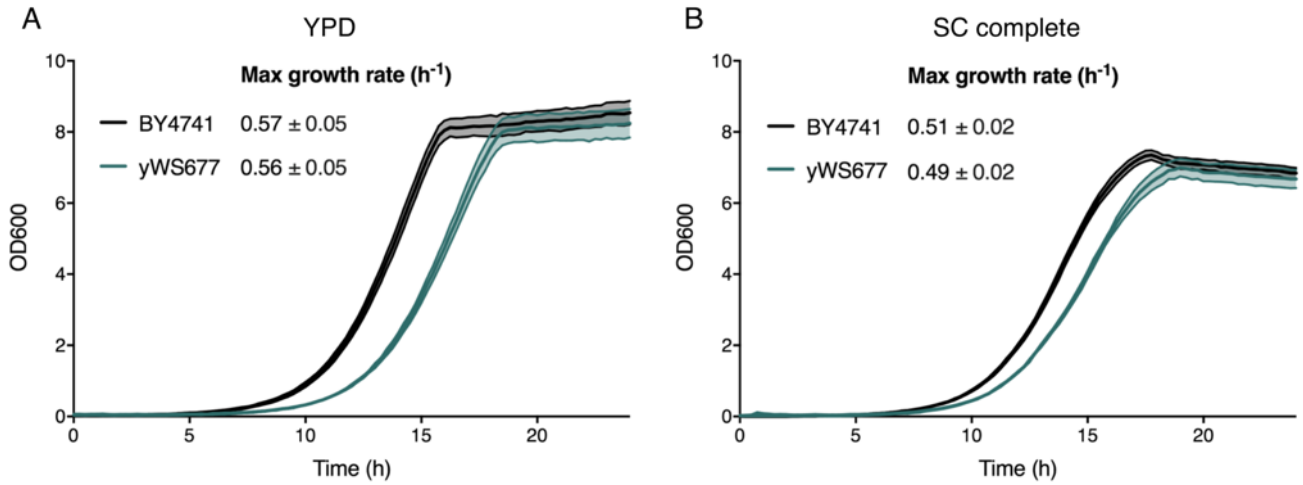


Figure 21. Growth rates of the yWS677 base strain compared to WT BY4741. (*A*) Growth of yWS677 (green) and BY4741 (black) in rich media (YPD). (*B*) Growth of yWS677 (green) and BY4741 (black) in synthetic complete media. Results are OD600 measurements from a plate reader and shown as the mean ± standard deviation from triplicate isolates.

Next, we characterised the landing pads at the *URA3*, *LEU2*, and *HO* loci for improving the efficiency of YTK plasmid integration using transient CRISPR/Cas9 and gRNA expression. To assay this efficiency, we assembled transcriptional units for sfGFP, mRuby2, and mTagBFP2 into the integration vectors of these three loci, so that their correct integration and expression could be measured using flow cytometry (**Figure 22A**). The plasmids were mixed with transient expression vectors for Cas9 and the gRNAs for targeting the respective landing pads. This one pot reaction was then digested with BpiI to linearise the plasmids and transformed directly into yeast (see **8.3** for a description of the plasmid tools and methods used in this study).

The transformation efficiency of a single (*URA3*), double (*URA2* and *LEU2*), and triple (*URA3*, *LEU2*, and *HO*) integration was compared with and without CRISPR/Cas9 and the gRNAs for targeting the three loci (**Figure 22B**). Multiplexed plasmid integration was greatly improved using the transient expression of CRISPR/Cas9 and the gRNAs, making it possible to achieve the efficient integration of three plasmids simultaneously. Without the aid of CRISPR/Cas9, the double integration resulted in only two colonies and no viable colonies were seen for the triple integration.

To determine the fraction of cells which had correctly integrated the three plasmids, we picked 96 random colonies from the CRIPSR-aided triple integration and measured them on the flow cytometer to determine their relative levels of fluorescent protein expression (**Figure 22C**). 90/96 colonies were monogenic, demonstrating a correct CRISPR-aided triple integration rate of 94%. The remaining six colonies contained a mixture of multiple integrations or missing fluorescence proteins.

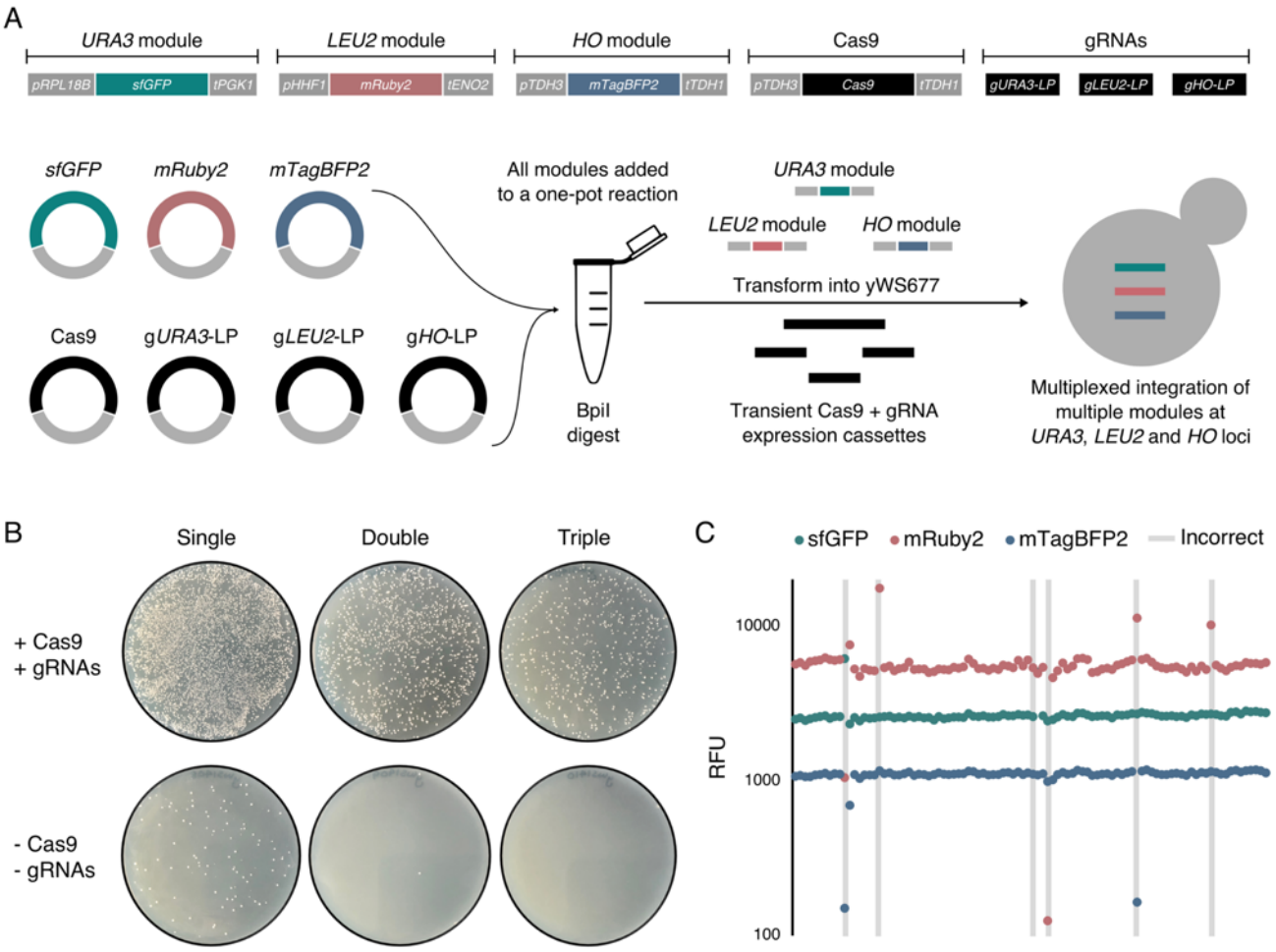


Figure 22. CRISPR/Cas9-aided multiplexed integration of marker plasmids. (*A*) Workflow for integrating multiple plasmids into yWS677. To assay the efficiency of multiplexed plasmid integration, sfGFP, mRuby2, and mTagBFP2 were assembled into the *URA3*, *LEU2*, and *HO* integration plasmids, respectively, and digested with BpiI in a one-pot reaction with transient Cas9 and gRNA expression cassettes, targeting the landing pads of the three genomic loci. Digestions were then directly transformed into yWS677 and plated on Ura$^-$/ Leu$^-$/ His$^-$ media. (*B*) Single, double, and triple integration of the *URA3*, *LEU2* and *HIS3* marker cassettes with and without Cas9 and the gRNAs required to generate DSBs at their respective landing pads. (*C*) Green, red, and blue fluorescence of 96 random colonies from the CRIPSR/Cas9-aided triple integration, 90 of which were correct for triple integration. The remaining six colonies contained a mixture of multiple integrations or missing fluorescence proteins (grey lines). Experimental measurements are sfGFP, mRuby2, and mTagBFP2 levels per cell determined by flow cytometry.

Following the success of the landing pads for improving multiplexed plasmid integration, we explored their use for restoring the deleted genes in yWS677 using markerless CRISPR-mediated editing. We designed gRNAs for targeting the landing pads at the *STE2*, *GPA1* and *STE12* KOs and amplified the wild type ORFs of these three genes, flanked by 500 bp, from the WT BY4741 genome to serve as donor DNA (**Figure 23A**). We then transformed the gRNAs and donor DNA alongside a CRISPR/Cas9 plasmid using the *URA3* marker (**Figure 23B**). Ten colonies were randomly picked, and colony PCR was used to validate the reintroduction of all three genes. 8/10 colonies screened were WT at the three loci, demonstrating an incredibly high success rate for the multiplexed reintroduction of genes using the landing pads (**Figure 23C**).

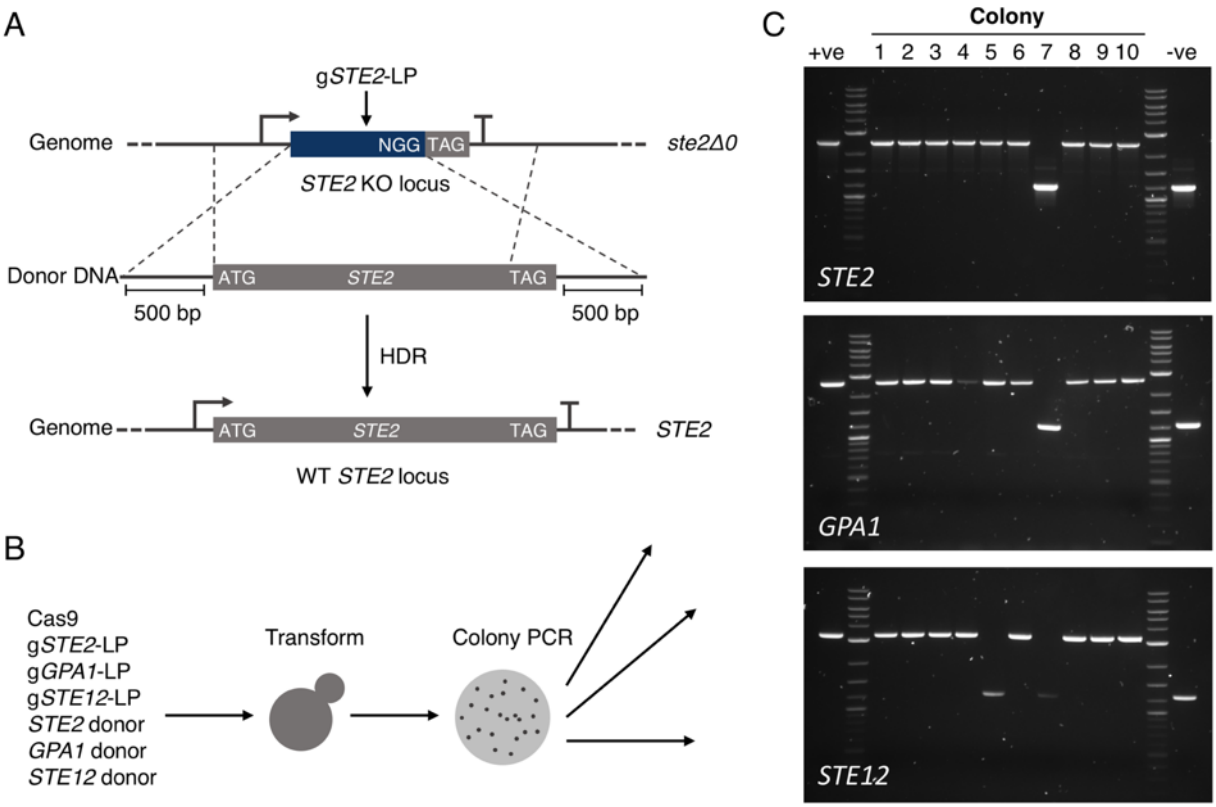


Figure 23. Multiplexed restoration of deleted genes using markerless CRISPR editing. (*A*) Restoration of genomic changes in the yWS677 base strain by targeting CRISPR/Cas9 to the landing pads of deleted genes and repairing with donor DNA PCR amplified from the WT genome. (*B*) Transformation of DNA for simultaneously restoring *STE2*, *GPA1,* and *STE12* to generate the "Quasi-WT" strain. (*C*) Colony PCR of 10 random transformants to identify correctly restored genotypes. The positive and negative controls are WT BY4741 and yWS677, respectively. 8/10 random colonies screened were correct for the restoration of all three genes.

As well as demonstrating the efficient multiplexed reintroduction of wild type genes in the GPCR base strain, the restoration of *STE2*, *GPA1,* and *STE12* produced a strain with useful properties. This strain contains the minimised signalling pathway composed entirely of wild type genes, and so can be used as a useful benchmark for comparing future refactored pathway designs. As this strain is wild type for all the genes we are interested in, but still highly modified (BY4741 *sst2Δ0 far1Δ0 bar1Δ0 ste3Δ0 mf(alpha)1Δ0 mf(alpha)2Δ0 mfa1Δ0 mfa2Δ0 gpr1Δ0 gpa2Δ0)*, we designated it "Quasi-WT". To confirm the reintroduced genes were in fact wild type identity, we sequenced the *STE2*, *GPA1* and *STE12* loci of the first confirmed strain (colony 1). We then cured this strain of the CRIPSR plasmids using 5-FoA counter selection, confirming plasmid loss by the absence of growth on uracil-deficient media and colony PCR.

To determine the utility of the Quasi-WT strain as a benchmark of pathway performance and the impact of minimising the pheromone response pathway, we compared the α-factor dose-response to the parental BY4741 yeast using the pheromone-responsive *FUS1* promoter driving sfGFP expression to report pathway activity[218,225] (**Figure 24A**). As expected, a substantial shift in the signal

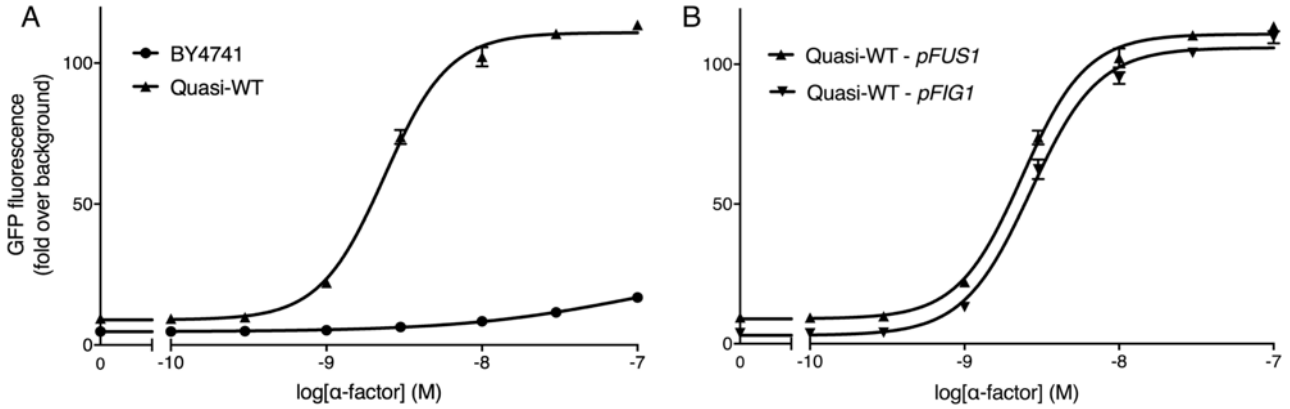


Figure 24. Benchmarking the dose-response characteristics of the Quasi-WT strain. (*A*) α-factor dose-response curves of the Quasi-WT and WT BY4741 strains, using the pheromone inducible *FUS1* promoter driving the expression of sfGFP to measure pathway activity. (*B*) α-factor dose response curves of the Quasi-WT strain using the *FUS1* and *FIG1* promoters to drive sfGFP expression. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

output and the sensitivity of the pathway was observed, due to the loss of negative feedback regulation, such as the RGS, Sst2[179,180]. However, a consequence of improving the sensitivity and the signal output of the system was an increase in the basal activity of the pathway (leakiness).

Finally, we compared the *FUS1* promoter to the alternative and increasingly popular *FIG1* promoter for reporting pathway activity to assess whether this would be a better fit for our refactoring purposes[228] (**Figure 24B**). The *FIG1* promoter exhibited more favourable biosensing characteristics, such as low leakiness and a high dynamic range, while maintaining an equal potency to α-factor. However, we decided to stick with the *FUS1* promoter as it displayed more intrinsic basal activity, and so discrete changes in pathway designs would be more measurable (i.e. we could work out why the system was leaky and address it directly, rather than mask it with a naturally tight promoter).

2.2.6 The GPCR-based biosensor platform

Demonstrating the minimal pathway was viable and able to produce a significant and highly sensitive output in the Quasi-WT strain suggested we could recreate a similar response by refactoring the receptor, Gα, and TF in the yWS677 base strain, using synthetic tools. To do this, we standardised a format for refactoring these three components, alongside a reporter, that would allow us to vary the promoter and terminator of each component, thus tuning their levels of expression (**Figure 25**), with all parts and cloning steps conforming to the YTK standard and hierarchical assembly strategy[33]. For a description of the updates we have made to this toolkit and parts we have added outside of the basic YTK starter set, see sections **8.3** and **8.5.1**, respectively.

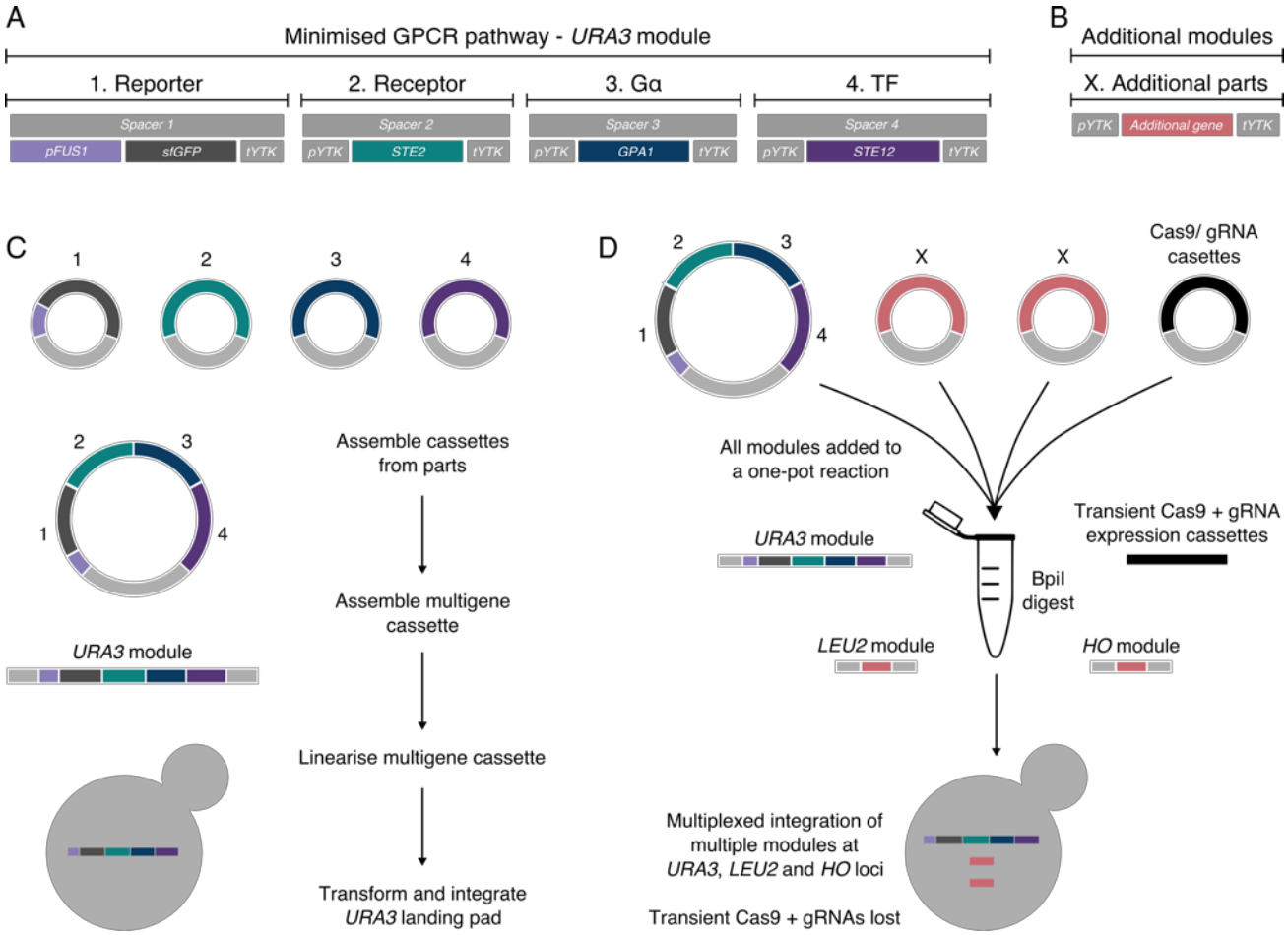


Figure 25. Overview of the GPCR-based biosensor toolkit format and workflow. (*A*) Format of the *URA3* module, used for refactoring the minimised GPCR signalling pathway, demonstrating the positions of the reporter, receptor, Gα, and TF in the final multigene plasmid. Spacer sequences are provided to exclude the instance of any combination of the four components. (*B*) Additional modules for integrating constructs at the *LEU2* and *HO* loci. (*C*) Assembling and integrating the *URA3* module for generating a minimised GPCR-based sensor in the yWS677 base strain, following the YTK hierarchical assembly strategy. (*D*) CRISPR/Cas9-aided integration of multiple modules. Here, a module refers to either a cassette or multigene cassette that integrates into the yeast genome at one of the three sites provided in the YTK toolkit starter set (*URA3*, *LEU2* or *HO* loci). All parts and cloning steps conform to the YTK standard.

In this format, the components are organised into four transcriptional units which are then assembled into a multigene plasmid for integration at the *URA3* locus (*URA3* module). In situations where any one of the four components are not required, spacer sequences have been provided to allow the construction of the complete module without needing to redefine the structure of pre-existing transcriptional units. An additional two modules exist for integration at the *LEU2* and *HO* loci. As the integration of all three modules can be multiplexed (see **Figure 22**), these modules can be used for integrating additional genetic constructs alongside the *URA3* module. Furthermore, components from the *URA3* module can be transferred to the additional modules for implementing combinatorial pathway designs.

2.3 Conclusions

In this chapter, we established a novel CRISPR-based method for rapidly iterating multiplexed genomic edits in yeast by cycling through a set of three markers. This allowed us to engineer our extensively modified GPCR base strain, yWS677, comprising 13 gene knockouts and a further three edits to improve the compatibility with the YTK system. We demonstrated how the use of landing pads could be used to improve the efficiency of multiplexed plasmid integration, as well as the multiplexed reintroduction of WT genes. Although these precise edits put significant constraints on the design of the gRNAs and donor DNA, their downstream application has already proved useful by providing a means to restore the receptor, Gα, and TF, thereby creating a benchmark strain for future refactoring of the minimised pathway (Quasi-WT strain). This benchmark strain displayed desirable biosensing characteristics compared to the WT response, signifying the benefits of the minimising the signalling pathway by removing feedback regulation. Finally, we defined the format and workflow of our platform for creating GPCR-based biosensors.

3 Refactoring a minimised pathway with native components

3.1 Introduction

Restoring the receptor, Gα, and TF in the GPCR base strain demonstrated the benefits of minimising the mating pathway for producing a highly sensitive response with a significant output. However, the loss of negative feedback regulation also introduced unwanted leakiness into the system, limiting the dynamic range. Refactoring the three components using a constitutive promoter library would, therefore, give us an opportunity to recreate the desired characteristics of the Quasi-WT response while looking for ways to reduce the unwanted properties. Furthermore, the use of non-PRE promoters would allow us to decouple the three components from positive transcriptional feedback by Ste12, which upregulates its own pathway components in response to pheromone[205].

When trying to understand a complex biological system, *in silico* approaches typically model only the key components, while removing other nonessential interactions from consideration[275,276]. These models can then be used to probe the individual contribution of each component within the reduced system by varying important parameters, such as their concentration. With the extensive genome engineering we have performed in our base strain to strip the cell of complexity, and the synthetic tools available to finely-tune the expression of the several key components, it is now possible to take an equivalent approach *in vivo*. This approach would allow us to delineate the contribution of the receptor, Gα, and TF, and determine whether their levels of expression could be used to tune the overall dose-response, as discussed in section **1.3**.

In this chapter, we explore our GPCR platform as an "*in vivo* model" of the minimised mating pathway. By individually refactoring the receptor, Gα, and TF, while keeping all other components fixed, we examine their contribution to the overall system. In parallel to this work, an *in silico* model of the minimised mating pathway was developed in collaboration with Graham Ladds at The University of Cambridge, using an equivalent computational approach. By performing the same analysis both experimentally and computationally, we attempt to strengthen our understanding of the biology, while also formalising a set of rules for tuning the pathway for biosensing applications.

3.2 Results and discussion

3.2.1 Initial refactoring of the minimised mating pathway

Before we began our refactoring efforts, we needed to identify the suitability of the promoter library in the YTK toolkit for stably expressing our reintroduced components. The wild type mating response results in massive changes to the transcriptome, significantly affecting over 200 genes, most of which have no involvement in the operation of the pathway[259]. Indeed, some well-known constitutive promoters have been used as inverted reporters of the pheromone response pathway, due to their decrease in expression after pheromone induction[277]. Although the repression of constitutive promoters should be eliminated in the absence of Far1, there is still the means of upregulation by Ste12[179].

To determine the effect of pathway activation on the stability of the YTK promoter library, we integrated a series of constructs into the Quasi-WT strain, each containing one of the 19 constitutive promoters from the YTK toolkit driving the expression of sfGFP (**Figure 26A**). We then measured the fluorescence of this panel of promoter characterisation strains in the presence and absence of

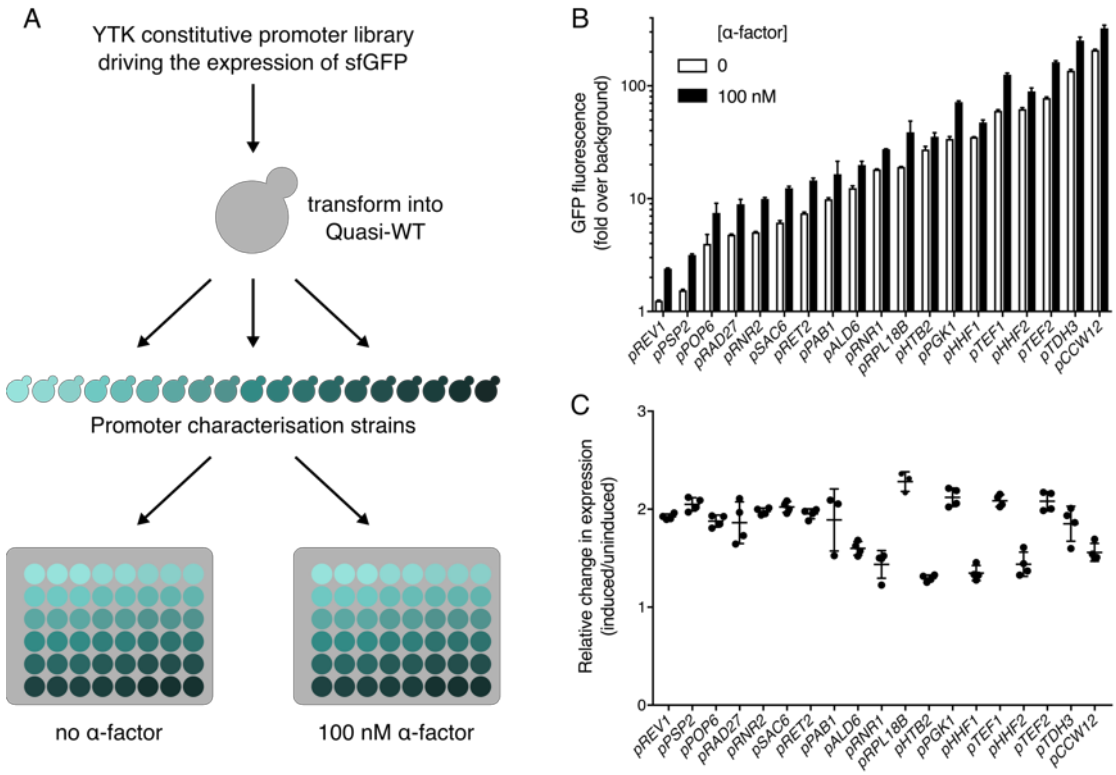


Figure 26. Characterising the YTK promoter library after pheromone stimulation. (A) Integration of the YTK promoter library driving the expression of sfGFP into the Quasi-WT strain to create a panel of promoter characterisation strains. (*B*) GFP fluorescence measurements of the promoter characterisation strains in response to saturating levels of α-factor. (C) Relative changes to the expression of sfGFP in the promoter characterisation strains after pheromone stimulation. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates.

saturating levels of α-factor (**Figure 26B+C**). An increase in the expression of sfGFP after pheromone induction was measured for all promoters. However, an increase was seen across the panel, resulting in between a 1.2- and 2.2-fold increase in relative fluorescence, suggesting a common mechanism was responsible for the increase in sfGFP protein. This was likely caused by morphological changes as part of the pheromone response leading to larger cells and therefore an increase in total protein per cell[278]. As all of the promoters followed a similar trend, we decided they were all suitable for refactoring the minimised pathway.

Next, we needed to define the promoters to drive the expression of the receptor, Gα, and TF in our initial pathway design. With 19 promoters and six terminators in the YTK toolkit and three different components, there would be over a million ways we could put this pathway back together. As the expression levels of each component are key for producing faithful signalling, the vast majority of these combinations would likely be non-functional, with the optimum existing in a very small parameter space[109]. While combinatorial approaches to creating these variants would be trivial, measuring the response of each variant to determine the optimal solution would not. Assuming an eight-point dose-response curve to assess all pathway characteristics (as shown in **Figure 24**), one person could process a maximum of 100 variants a day. This would take a lifetime to scan the entire genetic space. Reduced libraries based on statistical models would also be challenging as the dose-response is a complex and non-linear behaviour with many interlinked properties. Therefore, we needed a rational starting point to begin our refactoring efforts.

We decided to determine the native expression levels of the receptor, Gα, and TF so that we could match these levels using the YTK promoters. Using the landing pads in yWS677, we integrated the in-frame ORF of sfGFP between the promoter and terminator of the *STE2*, *GPA1*, and *STE12* genes so that fluorescence could be used a proxy for expression (**Figure 27A**). We then compared the GFP-ORF substitution strains to the previously established promoter characterisation strains using flow cytometry, all of which demonstrated relatively low levels of expression (**Figure 27B+C**).

We then chose four YTK promoters that had a similar level of expression to the three genes (grey ellipse; *pSAC6*, *pPOP6*, *pRNR2*, and *pRAD27*) and used these to create a new set of characterisation strains by combining these four promoters with the six terminators in the YTK toolkit. This produced an expanded promoter/terminator library with 24 discrete expression profiles similar to the three genes (**Figure 27D**). We then used this new library to select promoter/terminator combinations for the initial pathway design (coloured ellipses: green, receptor; blue, Gα; purple, TF). These combinations were chosen to be as similar to the native expression as possible, while also being unique to each other to avoid recombination issues downstream (*pSAC6-STE2-tSSA1*, *pPOP6-GPA1-tENO2*, and *pRAD27-STE12-tENO1*).

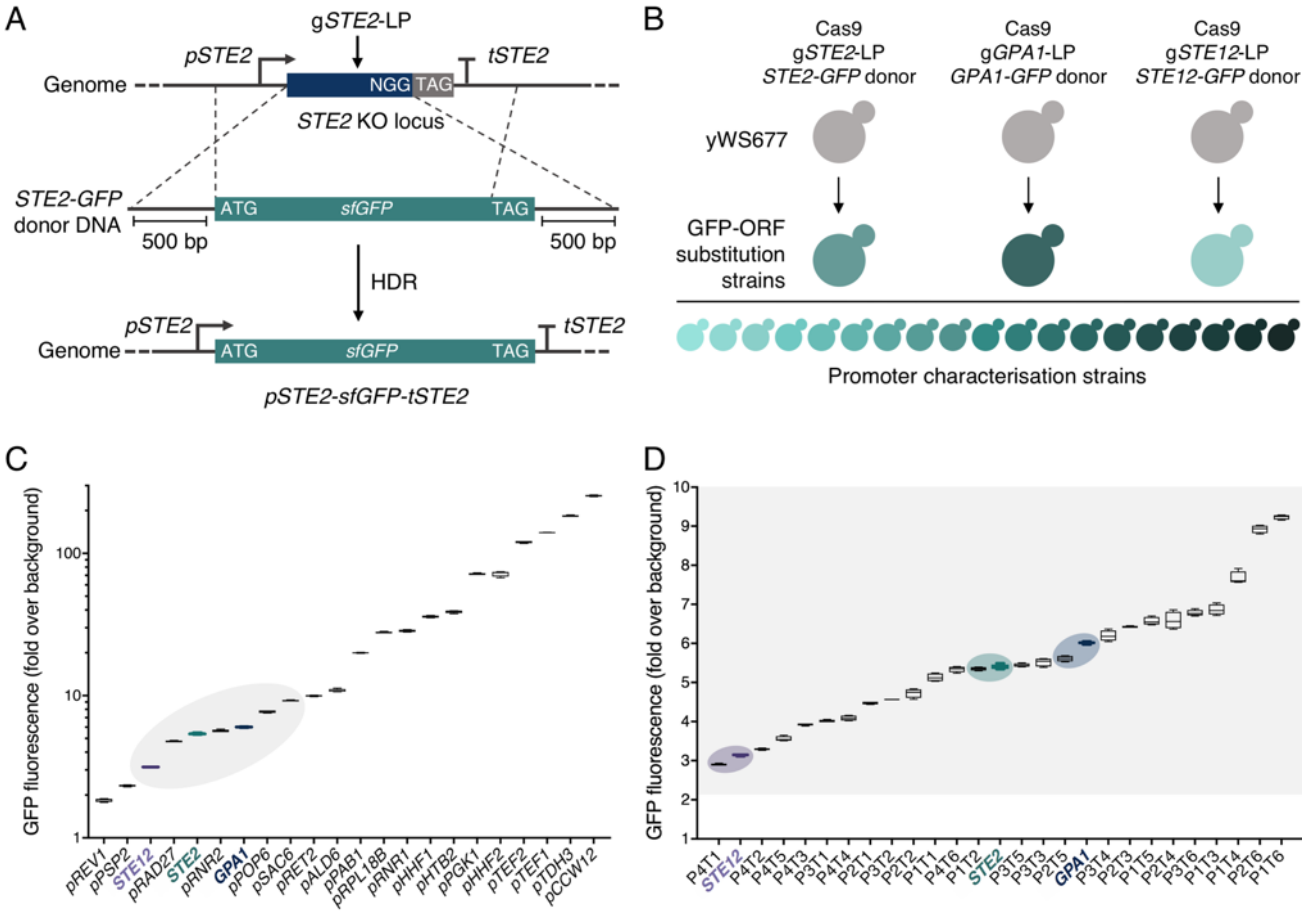


Figure 27. Characterising the native regulation of *STE2*, *GPA1*, and *STE12*. (*A*) Integrating the sfGFP ORF in-frame between the promoter and terminator of *STE2* in yWS677. (*B*) Generating GFP-ORF substitution strains for *STE2*, *GPA1*, and *STE12* in yWS677. (*C*) GFP fluorescence of the promoter characterisation and GFP-ORF substitution strains. The grey ellipse represents the GFP-ORF substitution strains and four of the promoter characterisation strains with fluorescence levels similar to the GFP-ORF substitution strains. (*D*) GFP fluorescence of the expanded promoter characterisation and GFP-ORF substitution strains. The grey area represents the boundaries of the grey ellipse in (*C*). The coloured ellipses represent promoter and terminator combinations both unique to each other and similar to native expression of *STE2* (green), *GPA1* (blue), and *STE12* (purple) chosen for the initial refactoring of the minimised pheromone response pathway (identity of the promoter terminator pairs listed in **Figure 28C**). Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates.

Using these promoter/terminator combinations we refactored the minimised mating pathway following the format outlined in **Figure 25**, with the *FUS1* promoter driving sfGFP expression to report pathway activity (*pFUS1-sfGFP-tTDH1*) (**Figure 28A-C**). We then compared the α-factor dose-response of the refactored pathway (Design 1) to the Quasi-WT strain (**Figure 28D**). The Design 1 response exhibited a substantial loss in sensitivity and maximum signal output while maintaining the same level of leak as the Quasi-WT response. This highly unfavourable biosensing response had essentially lost all of the attractive biosensing properties of the Quasi-WT strain while keeping all the undesirable properties. However, while the Design 1 strain had not produced the output we were expecting, it did provide a measurable dose-response, providing a starting point and a platform for future engineering.

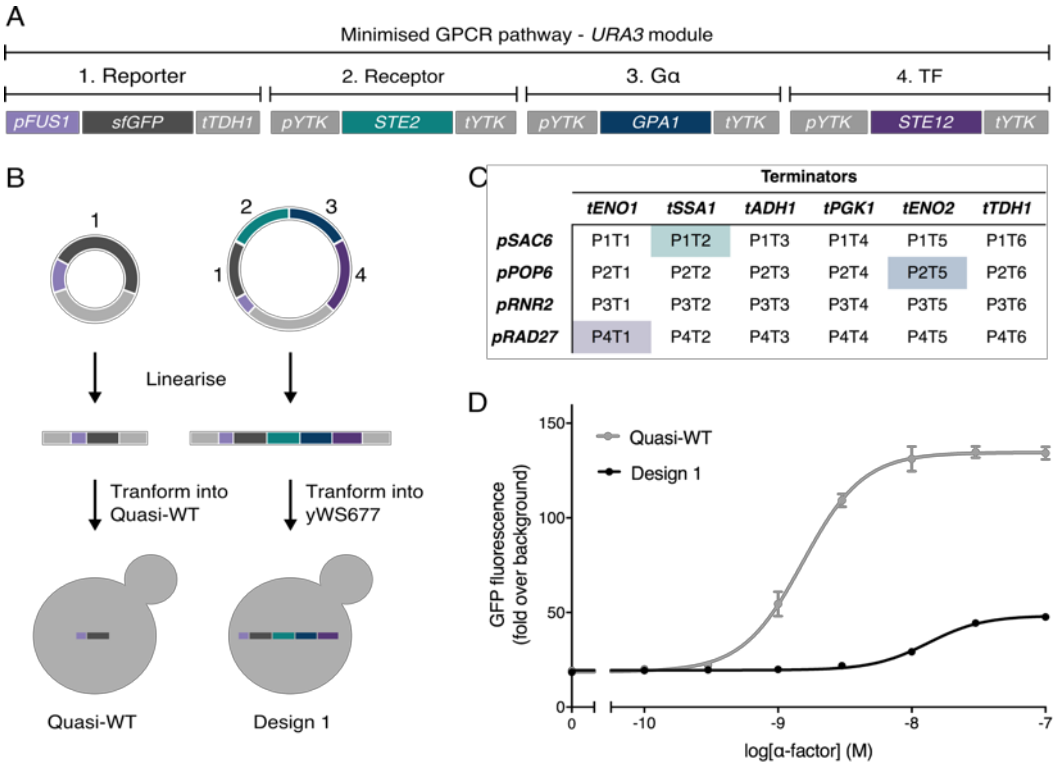


Figure 28. Initial refactoring of the minimised pheromone response pathway. (*A*) Format of the *URA3* module for refactoring the minimised pheromone response pathway. (*B*) Workflow for generating the Quasi-WT and Design 1 strains. (*C*) Promoter and terminator combinations from **Figure 27D**. The promoter and terminator combinations chosen to refactor the minimised pathway in Design 1 are highlighted in green (Ste2), blue (Gpa1), and purple (Ste12). (*D*) α-factor dose-response of the Quasi-WT and Design 1 strains. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

3.2.2 Transcriptional feedback is responsible for the Quasi-WT response

As we were confident we were expressing the three components at the approximate levels they are seen at in the native system, we decided to investigate why the Quasi-WT and Design 1 responses were so diverse. Essentially, the only difference between the two systems was at the promoter level. Therefore, we hypothesised the differences between the two responses was due to the loss of transcriptional feedback, as these three genes normally contain PRE elements in their promoters (**Supplementary Figure S1**) and are upregulated by Ste12 in response to pheromone[179] (**Figure 29A**).

To quantify the effect of this positive feedback, we integrated the *URA3* module from Design 1 into the three GFP-ORF substitution strains so that we could stimulate Ste12 transcriptional activity. The cells were then induced with saturating concentrations of α-factor, and their fluorescence was measured over time (**Figure 29B**). This revealed a significant upregulation of the receptor, Gα, and to a lesser extent, the TF, demonstrating that in the native context the expression of these components is not fixed during the mating response.

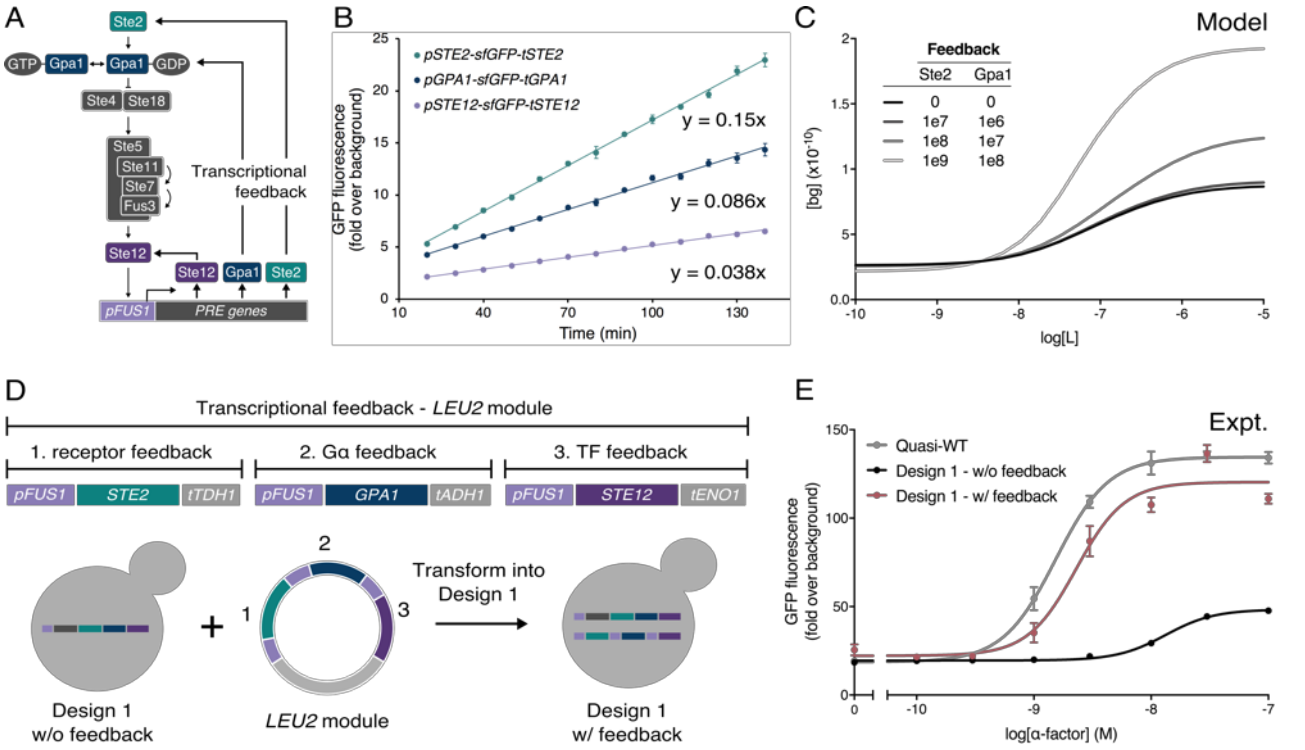


Figure 29. Wild-type and engineered transcriptional feedback. (*A*) Native transcriptional feedback of the receptor, Gα, and TF in the mating pathway after stimulation. (*B*) Upregulation of sfGFP in the GFP-ORF substitution strains containing the Design 1 *URA3* module after pathway stimulation with 1 µM α-factor. (*C*) Model of the pheromone response pathway incorporating reaction rates to enable the concentration of Ste2 and Gpa1 to increase as a product of pathway output. (*D*) The format of the secondary *LEU2* module for introducing transcriptional feedback of the receptor, Gα, and TF in response to pathway stimulation, and integration of the module into the previous established Design 1 strain. (*E*) α-factor dose-response of the Quasi-WT strain and the Design 1 strains with and without the secondary transcriptional feedback module. All modelling developed and performed by Hitoshi Yamauchi, Jack Mead, and Graham Ladds[264]. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

To investigate whether the transcriptional feedback of these components alone could lead to a significantly altered dose-response, we collaborated with Graham Ladds at Cambridge University to develop a model of the minimised pheromone response pathway to probe this hypothesis *in silico*. This model captured the receptor/G protein signalling, incorporating reaction rates to enable the concentration of the receptor and Gα to be increased as a function of the pathway output, thus simulating positive transcriptional feedback. Positive feedback of the TF was omitted as the model extended only as far as the Gβγ and the experimental data suggested minimal upregulation of this component in the native system. Performing the model over various levels of receptor and Gα feedback demonstrated an increase in sensitivity and maximal signalling while maintaining a consistent level of basal pathway activity (**Figure 29C**).

The model, therefore, supported the hypothesis that the differences between the Quasi-WT and Design 1 response were due to the loss of transcriptional feedback of the refactored components. To

confirm this experimentally, we constructed a second module (*LEU2* module) to artificially emulate transcriptional feedback of the receptor, Gα, and TF, by placing each these components downstream of the pheromone-inducible *FUS1* promoter, using various strengths of terminator to reflect the differences in the natural feedback levels (**Figure 29D**). We then transformed this module into the Design 1 strain and measured the α-factor dose-response of the system with the engineered feedback (**Figure 29E**). The addition of the feedback module significantly enhanced the sensitivity and maximum output of the system while maintaining a similar basal activity, as predicted by the model and almost matching the Quasi-WT response.

3.2.3 An *in vivo* model of receptor/G protein signalling

We have demonstrated the use of transcriptional feedback to alter the dose-response characteristics of the minimised mating pathway by introducing a second module for expressing the refactored components as a product of the pathway output. However, while we could explore the use of positive feedback for tuning minimised mating pathway, this would be a direct contradiction of our initial aims to reduce the complexity of the system. Furthermore, as demonstrated in **Figure 29**, the use of positive feedback can only tune the dose-response curve in one direction. As we would like independent control of all dose-response properties for creating fully-tuneable GPCR-based biosensors, this approach would not be suitable.

Instead of pursuing the use of feedback, we next investigated whether we could tune the dose-response of the minimised pathway by individually altering the expression levels of the refactored components. We first focused on the receptor and Gα as it was possible to simulate these components using an adaptation of a model of heterotrimeric G protein signalling by Bridge *et al.*[249] (**Figure 30B**). This model captures all of the interactions and possible states of the receptor and G protein complex, and so was suitable for probing changes to the levels of the receptor and Gα in the minimised mating pathway. In this model, the output of the system is measured as the free Gβγ. As this dimer is responsible for activating the MAPK signalling cascade[179], which itself displays a graded, linear response with respect to agonist concentration[248,279], the rest of the pathway could be blackboxed, and the Gβγ used as a proxy of pathway output.

Using this model of receptor/G protein signalling, we individually varied the initial concentrations of the receptor and Gα in the system, while keeping all other components fixed, and simulated the response of the pathway to agonist stimulation. This demonstrated a clear relationship between the receptor number, the sensitivity, and the maximum signal, as previously shown by Bush *et al.*[252]

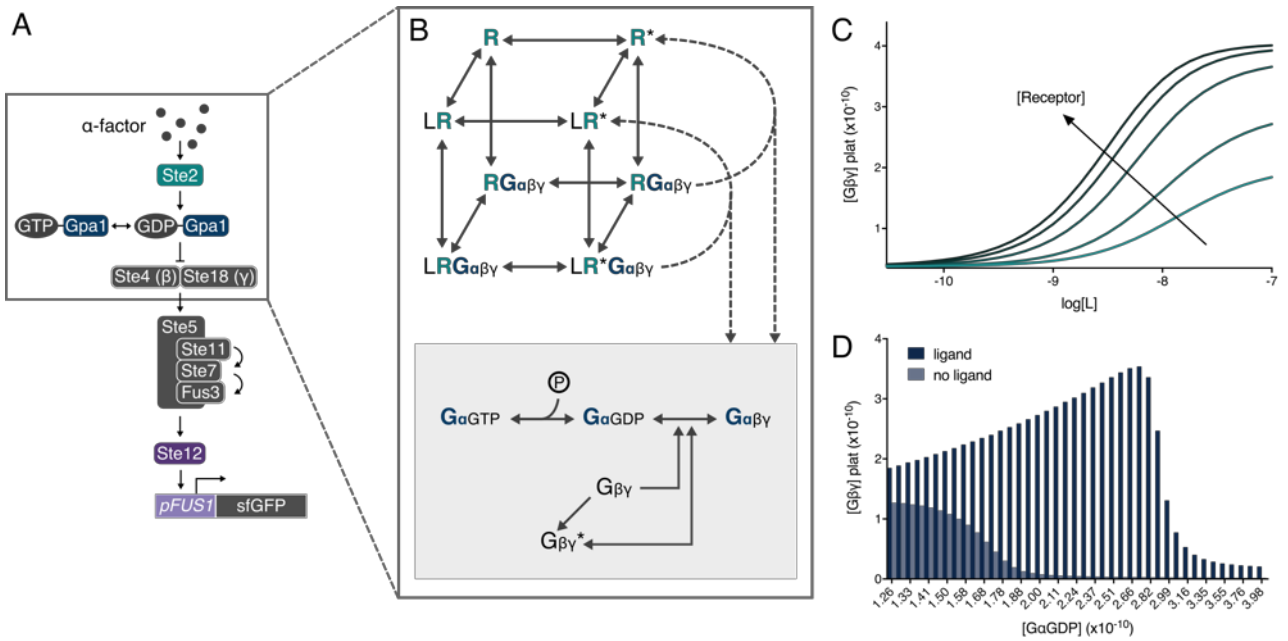


Figure 30. Model of receptor/G protein signalling in the minimised pheromone response pathway. (*A*) The minimised pheromone response pathway. The box represents the components incorporated into the model of receptor/G protein signalling. (*B*) Single cubic ternary complex model of receptor/G protein signalling in the minimised pheromone response pathway. (*C*) Model of the pathway dose-response to α-factor over a range of initial receptor concentrations. (*D*) Endpoint response of the model in the presence and absence of saturating agonist over a range of initial GαGDP concentrations. All modelling developed and performed by Hitoshi Yamauchi, Jack Mead, and Graham Ladds[264].

(**Figure 30C**), suggesting that we would indeed be able to tune the sensitivity of our minimised pathway by varying the expression of the receptor.

The relationship between the Gα number and the pathway response was more complex (**Figure 30D**). At lower concentrations of Gα, constitutive expression of the pathway was observed due to increased free Gβγ. This, in combination with reduced receptor-mediated signalling caused by lower receptor/Gαβγ concentrations, results in a lower maximum fold change in pathway activation. At higher Gα concentrations free Gβγ is rapidly sequestered, also leading to a decrease in the pathway activity by acting as a "sponge" to signalling. This model, therefore, predicts a "sweet spot" of Gα expression between the two extremes where all three members of the heterotrimeric G protein appear to be in balance, leading to a high fold change in the signal after activation. Altering the expression of the Gα in the system would, therefore, provide a means to tune the leakiness of the system and was likely the source of the high basal activity in our Design 1 strain.

Next, we sought to experimentally validate the findings of the receptor/G protein model by applying the same analysis to our refactored system. To first address the leakiness in the system, a property that may be exacerbated by increasing the expression of the receptor[219], we explored the changes to the Gα by driving the expression of Gpa1 using a promoter library and measuring the ON/OFF response to saturating levels of α-factor (**Figure 31**). The experimental findings were in remarkable

agreement with the simulations, demonstrating similar profiles for the ON/OFF response and the maximum fold change of the system, displaying the predicted sweet spot of Gα expression using the *PGK1* promoter.

Finally, we experimentally investigated the receptor behaviour in the leak-free system by fixing the expression of Gpa1 using the *PGK1* promoter and varying the expression of Ste2 using five different promoters, spanning the entire range of promoter strengths (**Figure 32**). Increasing the expression of the receptor significantly enhanced the sensitivity and the maximum signal output in a very predictable manner, as expected by the model. Using receptor tuning, the sensitivity of the minimised pathway could be tuned over 1.5 orders of magnitude, demonstrating the flexibility of the system.

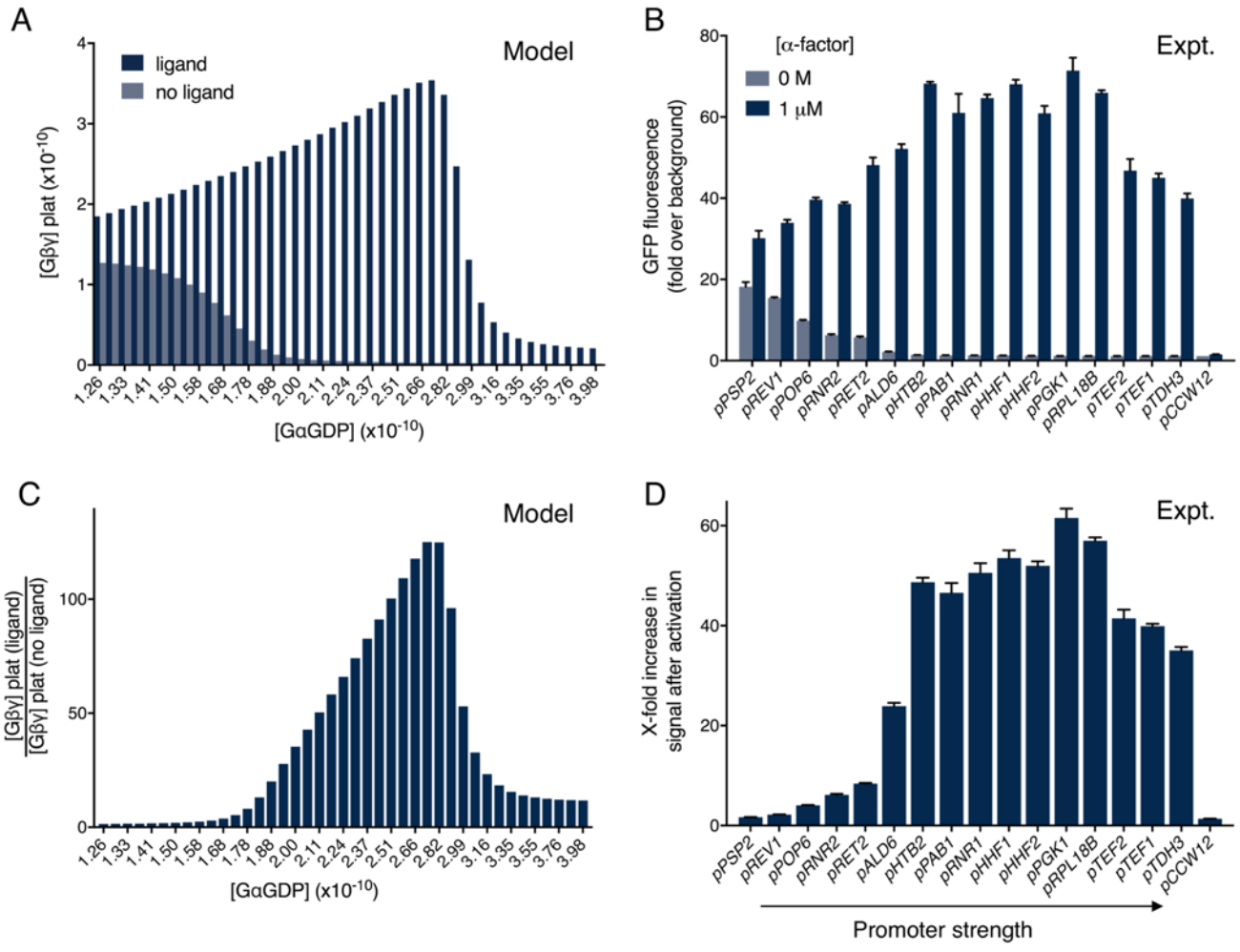


Figure 31. Experimental validation of the Gpa1 concentration model. (*A*) Endpoint response of the minimised pheromone response pathway model in the presence and absence of saturating agonist over a range of initial GαGDP concentrations. (*B*) Experimental ON/OFF response of 17 minimised pathway designs where the intracellular levels of Gα are varied using a promoter library driving the expression of Gpa1. (*C*) Model of maximum pathway activation over a range of initial GαGDP concentrations. (D) Experimental maximum x-fold change in signal of the 17 pathway designs. All modelling developed and performed by Hitoshi Yamauchi, Jack Mead, and Graham Ladds[264]. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates.

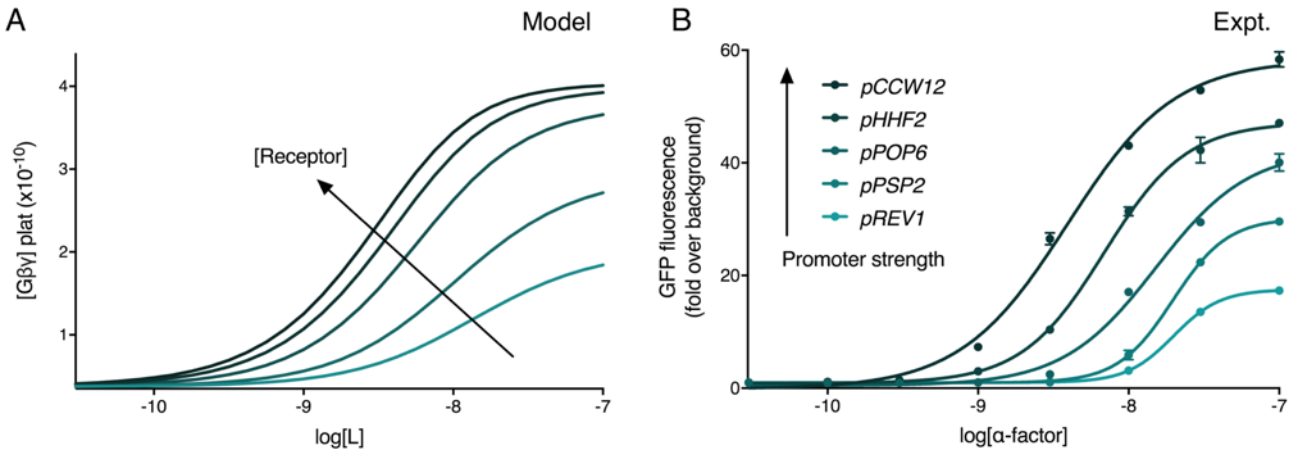


Figure 32. Experimental validation of the Ste2 concentration model. (*A*) Model of the pathway dose-response to α-factor over a range of initial receptor concentrations. (*B*) Experimental α-factor dose-response curves for 6 minimised pathway designs, varying the expression of the receptor, Ste2, with a promoter library. All modelling developed and performed by Hitoshi Yamauchi, Jack Mead, and Graham Ladds[264]. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

3.2.4 Increasing the expression of the Ste12 transcription factor is toxic

Although our model did not extend as far as the Ste12 TF, we decided to continue with experimental refactoring nonetheless. Keeping the Gα levels fixed with the *PGK1* promoter and tuning a sensitive response to α-factor by driving receptor expression with the strong *CCW12* promoter, we varied the expression of the TF. Expressing Ste12 on anything greater than a low-mid strength promoter resulted in toxicity issues (**Figure 33A**), a phenomenon that has previously been reported but never adequately explained[280]. We hypothesised this toxicity might be due to an imbalance between Ste12 and the negative repressors of Ste12, Dig1, and Dig2.

While the pathway is inactive, Dig1 and Dig2 sit in a fine balance with Ste12 and the presence of these two negative regulators stabilise an inactive transcriptional complex, allowing the system to rapidly respond once a signal is recieved[250]. This complex is highly regulated and prevents transcription in the absence of pathway activity[179]. However, as the levels of Ste12 increase, the available pool of Dig1 and Dig2 is spread thin, resulting in partially or fully unregulated TFs. These would then constitutively activate any of the 100+ genes usually upregulated in the mating response, leading to cellular burden and toxicity. In this hypothesis, the greater the number of TFs in the system, the more mating-responsive genes are constitutively switched on, and the larger the degree of cellular burden (**Figure 33B**).

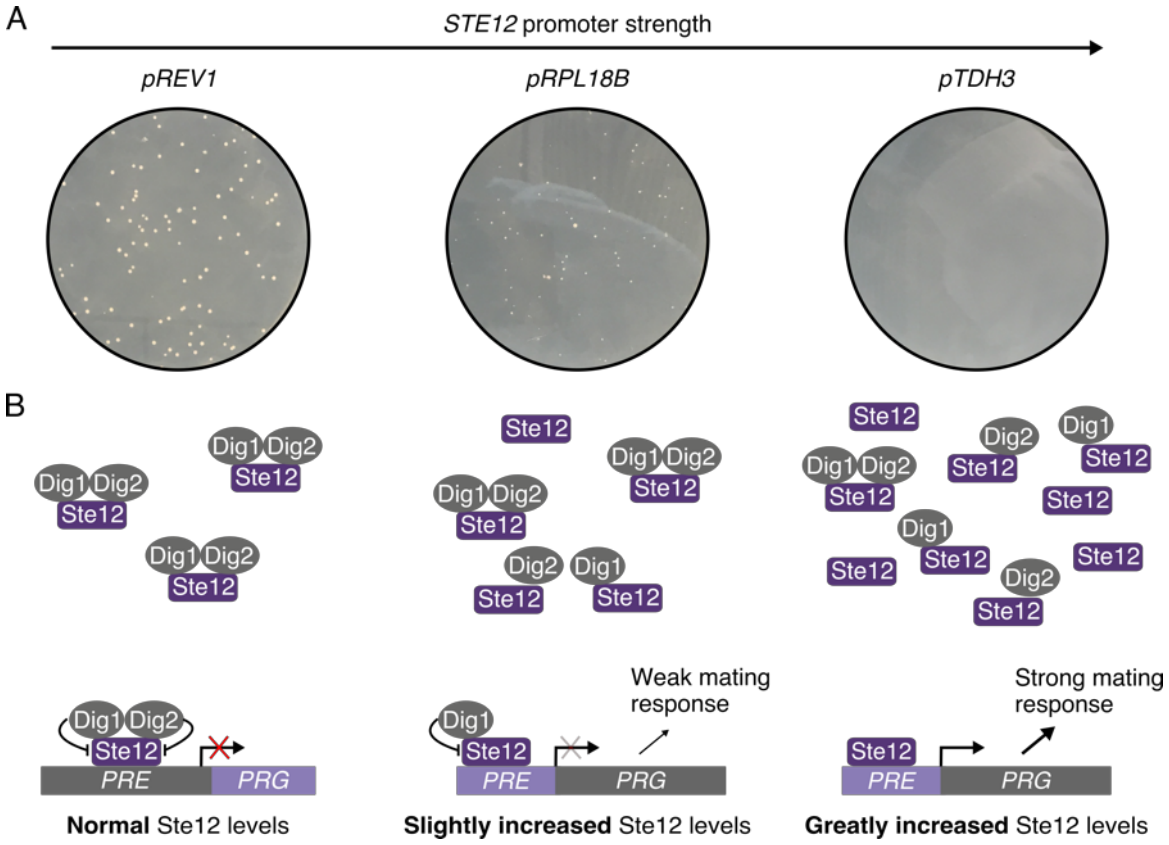


Figure 33. Experimental changes to the expression of Ste12. (*A*) Transformants of 3 pathway designs with Ste12 under the control of a weak (*pREV1*), medium (*pRPL18B*), and strong (*pTDH3*) constitutive promoter. Driving the expression of Ste12 using the weak and medium strength *REV1* and *RPL18B* promoters resulted in a similar number of colonies, although the *pRPL18B* variants were a fraction of the size of the *pREV1* variants. Driving the expression of Ste12 using the strong *TDH3* promoter resulted in no visible colonies. (*B*) The proposed mechanism of cellular burden from increased Ste12 levels. Normal Ste12 levels result in a fine balance with the negative regulators of Ste12, Dig1 and Dig2, resulting in the tight repression of *PRE*-genes when the pathway is in the OFF-state. Slightly increased levels of Ste12 spread the available pool of Dig1 and Dig2 thin, resulting in some half- or fully unrepressed Ste12 molecules leading to the leaky expression of some *PRE*-genes and a weak mating response in the OFF-state. Greatly increased Ste12 levels significantly outnumber the available pool of Dig1 and Dig2 molecules, leading to many unrepressed Ste12 proteins and a strong mating response in the OFF-state. The constitutive activation of the 100+ genes involved in the mating response likely creates a high level of cellular burden, preventing growth. Pictures of plates taken 3 days post transformation.

To investigate whether or not the limited availability of Dig1 or Dig2 could be causing the toxicity, we co-transformed the Ste12 promoter library with an additional module for overexpressing Dig1, Dig2, or DIG1 and Dig2 under the control of strong constitutive promoters (**Figure 34**). The overexpression of Dig1 or Dig2 was able to partially recover the Ste12 library, with burden still present for the highest levels of TF expression. However, the overexpression of both transcriptional regulators was able to recover the full library, suggesting that both repressors are needed for the absolute regulation of Ste12. As it seems there is a fine balance between the TF and the two negative regulators, manipulating the pathway via the expression of Ste12 would require the concerted tuning of both Dig1 and Dig2. Due to the combinatorial complexity of this problem, we decided to keep the expression of all pheromone-responsive TFs fixed at low levels using the *RAD27* promoter in all future designs.

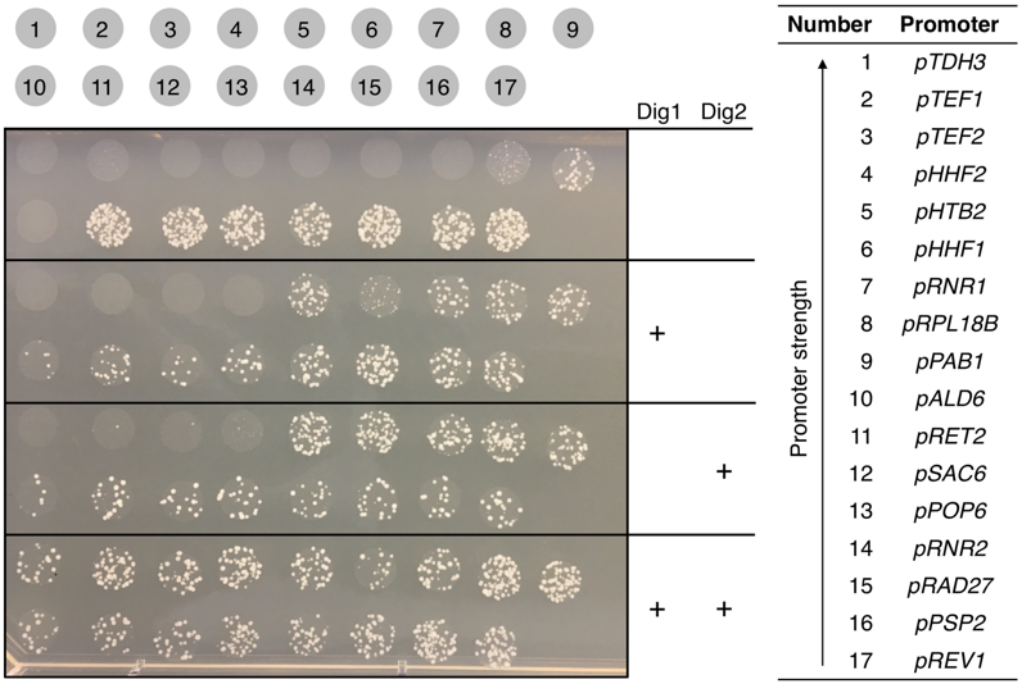


Figure 34. Recovery of the Ste12 promoter library with the overexpression of Dig1 and Dig2. A library of 17 constitutive promoters, from the YTK library, driving the expression of Ste12 with and without the overexpression of Dig1 and Dig2 in all combinations. The spotted yeast are direct transformants of two plasmids: the first containing the refactored pathway, with Ste12 under varying strengths of promoter, and the second containing either a blank spacer sequence, Dig1, Dig2, or Dig1 and Dig2 under the control of strong constitutive promoters. Identical handling was performed, and equal amounts of DNA was used for each condition. Picture taken 2 days post transformation.

3.2.5 Tuning the leakiness and sensitivity of the refactored pathway

Experimentally varying the expression of the receptor and Gα had demonstrated two individual tuning knobs for predictably programming the sensitivity and the leakiness of the system. To demonstrate their ability to tune the minimised mating pathway, we compared the α-factor dose-response of two designs presented earlier in the chapter to the Quasi-WT response (**Figure 35A**). These designs built on the Design 1 architecture by first reducing the leakiness from the system by increasing the expression of Gα, using the *PGK1* promoter (Design 2), and then increasing the expression of the receptor, using the strong *CCW12* promoter, to improve the sensitivity (Design 3).

At each stage of the redesign, the properties of the dose-response curve were enhanced (**Figure 35B**). The Design 2 strain demonstrated significant improvements to the tightness of the system (the opposite of leakiness), resulting in a tight OFF-state with no measurable levels of basal activity over background. Consequently, this improved the dynamic range of the system compared to the Quasi-WT response. The Design 3 strain built on the improvements of the previous design by enhancing the sensitivity to α-factor, while maintaining the tightness. This resulted in a highly-sensitive response

with a large operational range, comparable to the Quasi-WT strain. Although the maximum output of the system was lower in the Design 3 than the Quasi-WT strain, improvements to dynamic range coupled with no measurable leak made this response more suitable for biosensing applications.

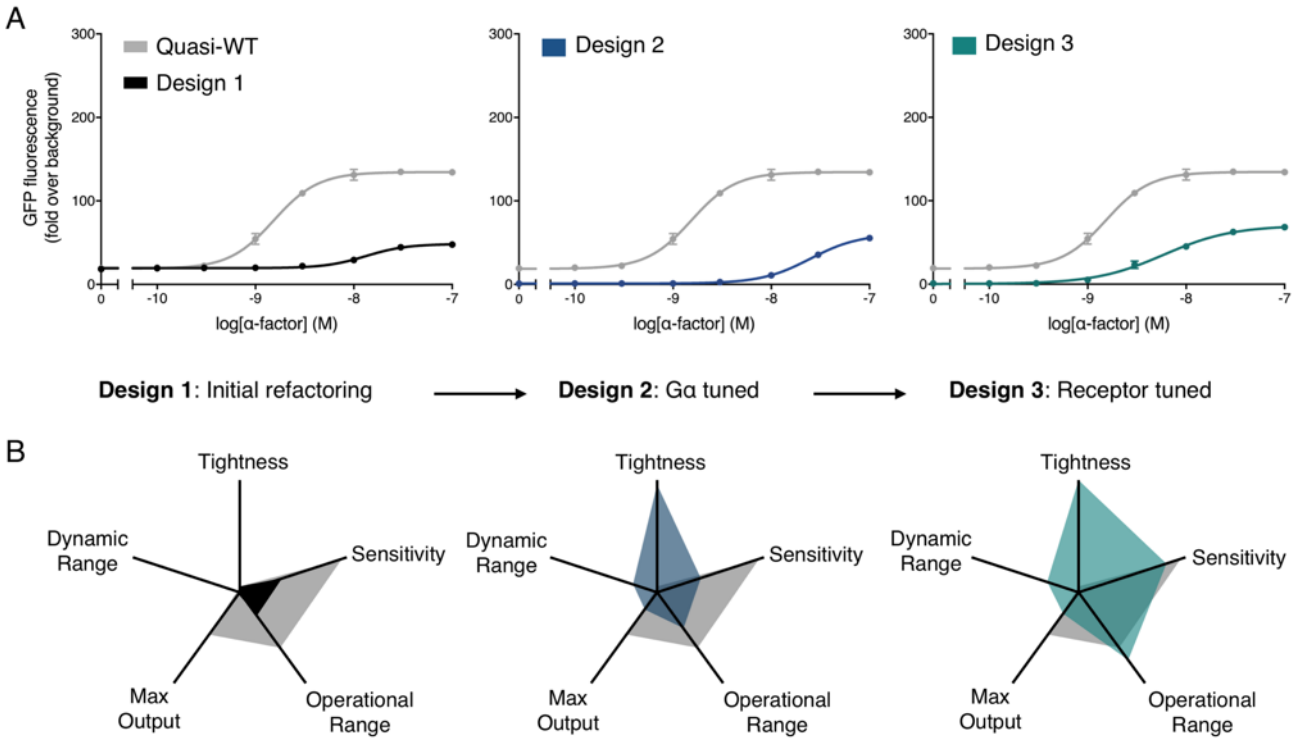


Figure 35. Tuning the minimised pheromone response pathway through iterative refactoring (Design 1-3). (*A*) α-factor dose-response curves of 3 sequentially refactored pathway designs compared to the Quasi-WT strain, starting with the initial refactoring from **Figure 28** (Design 1), then tuning the expression of Gpa1 using the *PGK1* promoter to reduce the leakiness (Design 2), and finally, tuning the expression of Ste2 using the *CCW12* promoter to improve the sensitivity and operational range (Design 3). Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit. (*B*) Dose-response characteristics of the three refactored pathway designs. Tightness is defined as the reciprocal of basal activity and the dynamic range is defined as (maximum output/basal activity). Sensitivity and operational range were determined from the fitted curve, defining sensitivity as the lowest concentration for which a >2-fold change in GFP expression is seen, and operational range as the concentration span between the sensitivity and the lowest concentration that gives a GFP expression within 2-fold of the maximum. All values were then normalized to the minimum measurable value and the maximum calculated value in the dataset.

3.3 Conclusions

In this chapter, we began our refactoring efforts by approaching the problem in a rational manner. We characterised the approximate expression of the receptor, Gα, and TF in the native system and created an initial pathway design to emulate these levels using synthetic regulation. Although we were able to recreate a measurable pathway response, the initial design had lost considerable sensitivity and pathway activity compared to the Quasi-WT system, revealing transcriptional feedback as the mechanism responsible for these differences. To explore points of tuning in the refactored pathway and improve these properties without using feedback, we developed an *in vivo* model of receptor/G protein signalling. By computationally and experimentally varying the amount of receptor and Gα in the system, we discovered predictable tuning knobs for individually altering the sensitivity and leakiness. However, this approach was not possible with the TF, as a fine balance with the negative regulators, Dig1 and Dig2, was required to attenuate aberrant signalling in the absence of an input. Finally, by iteratively refactoring the Gα and receptor, we were able to create a refactored pathway design with improved biosensing characteristics compared to the Quasi-WT response.

4 Redirecting the mating pathway using sTFs

4.1 Introduction

So far, we have identified rules for predictably tuning the sensitivity and leakiness of the system by varying the expression of the receptor and Gα. However, we have not yet identified a rule for independently modulating the maximum signal output of the system that was independent of the other dose-response properties. As increasing the levels of TF caused toxicity issues, exploring whether varying the expression of this component would lead to changes in the dose-response was not possible. Consequently, we needed an alternative strategy for modulating the signal output of the system. As outlined in section **1.3**, synthetic transcription factors (sTFs) may hold the key to this problem. By substituting the DNA binding domain (DBD) of Ste12 for an orthogonal substitute, the transcriptional response of the mating pathway can be redirected to a synthetic promoter[228,256]. We could then take advantage of the modular architecture of eukaryotic promoters for tuning the pathway output[257].

Typically, eukaryotic promoters consist of two distinct regions: (i) the core promoter, which can be identified as the minimal region responsible for initiating transcription, and (i) the upstream enhancer, which determines the number and organisation of the TFs as a means of imparting regulation on the core promoter[257]. Both of these elements can be modified to tune the expression from the promoter. If the upstream enhancer imparts positive regulation of the core promoter, it is commonly referred to as an upstream activating sequence (UAS)[257]. Altering the number of transcription factor binding sites (TFBSs) within the UAS can recruit different numbers of TFs to regulate the core promoter in a predictable manner[281]. Similarly, exchanging the core promoter region for other endogenous or synthetic core promoter regions with different transcription initiation rates can further fine-tune the promoter output[281,282]. Accordingly, a number of toolkits have embraced this modularity to allow for the tuneable expression of genes in yeast[258,283].

In this chapter, we use sTFs for redirecting the pheromone response pathway to synthetic promoters. We then explore how the promoter identity can be used as a method for tuning the maximum output of the system. We also consider sTFs with unique properties for adding further control or programmability to the system through the action of the DBD, rather than the promoter. Finally, we look at the effect of decoupling the pheromone response from transcriptional feedback and use the tools developed in this chapter to revisit the tuning of the minimised mating pathway.

4.2 Results and discussion

4.2.1 Pheromone responsive synthetic transcription factors

The first demonstration of pheromone-responsive sTFs was published two decades ago as a way to investigate the domains within the Ste12 TF[256]. In this study, the authors used a series of hybrid proteins of Ste12 with the DBD of the endogenous transcriptional activator Gal4 to define several key domains in Ste12: the DNA binding domain (D, 1-215), the induction region (I, 216-383), and the transcription activation sequences (A, 384-688) (**Figure 36B**). The minimal pheromone induction region (I) was shown to be dependent on the MAPK pathway for induction activity and interacted with the Dig1 and Dig2 transcriptional repressors. This domain controlled transcriptional activation by the relief of repression and synergistic activation with the transcriptional activation domain (A)[256].

To exploit this modular protein architecture for biosensing applications, Mukherjee *et al.*[228] transplanted the induction region of Ste12 between the DBD and activation domain (AD) of endogenous and heterologous substitutes to generate two synthetic transcription factors, STF1 ($Gal4_{DBD}$-I-$Gal4_{AD}$) and STF2 ($LexA_{DBD}$-I-$B42_{AD}$) (**Figure 36C**). These sTFs could then be targeted to synthetic promoters containing TFBSs appropriate for the respective DBDs, upstream of a core promoter driving GFP expression. In the absence of pathway activity, the transcriptional repressor proteins would bind the induction region between the DBD and AD, preventing transcriptional activity. Pathway activation would then release the transcriptional repressors from the induction region, allowing transcriptional activity, driven by either the yeast Gal4 or bacterial B42 AD.

To assess the performance of these sTFs in our system, we ported STF1 and STF2 using the conditions identified in Design 3 to create a refactored pathway, using the synthetic promoters reported in Mukherjee *et al.*[228] to report pathway activity. We then measured the α-factor dose-response of these two new systems compared to the Design 3 strain utilising the wild type Ste12 TF (**Figure 36D**). The STF1-mediated pathway produced a response similar to the Design 3 strain. Although the maximum signal was slightly improved, a small increase in the leakiness resulted in an almost identical maximum fold change in signal output after activation (**Figure 36E**). STF2, on the other hand, had a very low activation, barely producing an output above background, and in its current form would not be suitable for this work.

STF1 had, therefore, established a means to redirect the mating pathway to a synthetic promoter while maintaining a response in alignment with the native Ste12 TF. However, as the DBD of STF1 was from an endogenous transactivation protein (Gal4) it was not truly orthogonal in the host chassis.

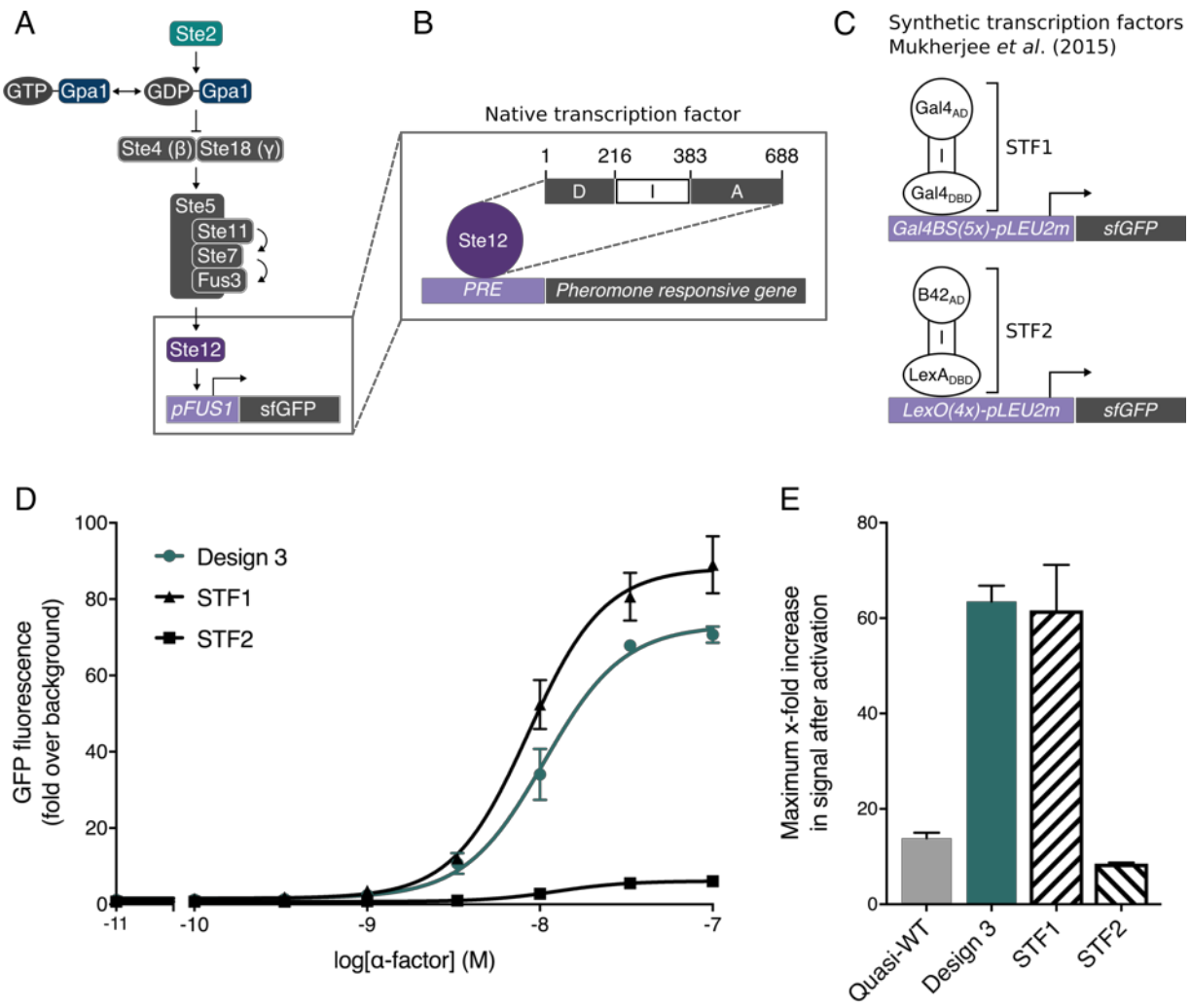


Figure 36. Redirecting the minimised mating pathway to synthetic promoters using STFs from Mukherjee *et al.*[228]. (*A*) Overview of the minimised mating pathway using native components. (*B*) The native pheromone responsive TF, Ste12, targeting a PRE-gene. The Ste12 protein is composed of 3 distinct regions: the DNA binding domain (D, 1-215), the induction region (I, 216-383), and the transcription activation sequences (A, 384-688). (*C*) The architecture of the two pheromone responsive synthetic transcription factors from Mukherjee *et al.*[228]. STF1 is composed of the DNA binding domain and activation domain of Gal4 from *S. cerevisiae*, connected by the induction domain of Ste12, and is targeted to a synthetic promoter consisting of five repeats of the Gal4 binding sequence (Gal4BS) followed by the *LEU2* core promoter driving the expression of sfGFP. STF2 is composed of the orthogonal DNA binding domain from the bacterial repressor protein LexA followed by the induction domain of Ste12 and the bacterial B42 activation domain and is targeted to a synthetic promoter composed of 4 repeats of the LexA operator (LexO) sequence upstream of the *LEU2* core promoter driving the expression of sfGFP. (*D*) α-factor dose-response curves of the three pathway variants composed of the receptor and Gα conditions of Design 3, varying at the TF/reporter, utilising either the native Ste12 TF (Design 3) or the synthetic transcription factors (STF1 and STF2). (*E*) Maximum x-fold change in signal of the TF/reporter variants compared to the Quasi-WT strain. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

Before progressing with STF1 and exploring modular synthetic promoters, we decided to briefly detour by focusing on improving STF2, as this was a genuinely orthogonal transcription factor using the bacterial LexA DBD.

Our first consideration for improving STF2 was the LexA DBD. In the native bacterial system, LexA proteins form dimers to stabilise the DNA bound complex[284]. Dimerisation requires the full-length LexA protein, consisting of the DBD and a carboxy-terminal dimerisation and latent protease domain[285]. However, STF2 only consists of the LexA DBD, preventing dimerisation. Interestingly, a similar organisational constraint has been demonstrated for Ste12, revealing multimerisation of the C-terminal domain is important for transcriptional activation[286]. Therefore, we reasoned that the full-length LexA protein would form more stable complexes on the synthetic promoter, while facilitating the interaction of neighbouring Ste12 C-terminal domains, potentially enhancing the transcriptional activation. As the bacterial proteins required for activating the protease in LexA are absent in yeast, this domain should remain latent[287]. However, for this to work, the entire Ste12 C-terminal domain would need to be considered a single unit, and so we redefined the I and A regions as one; the pheromone-responsive domain (PRD, 216-688) (**Figure 37A**).

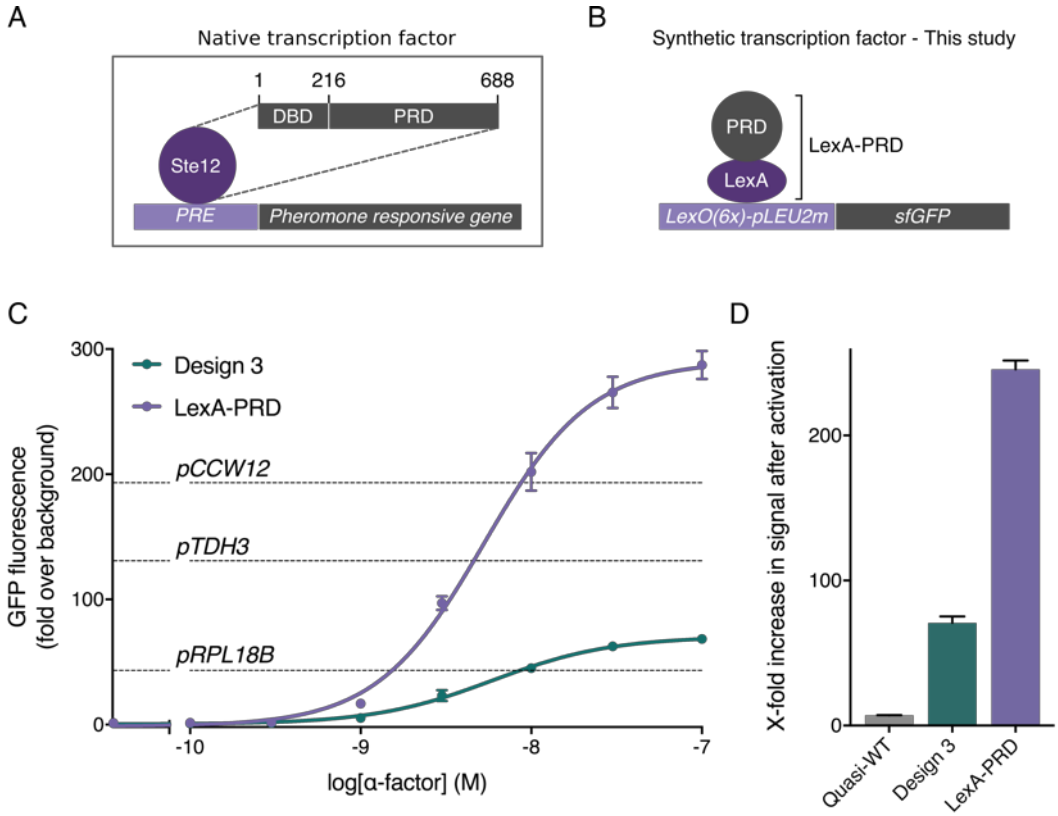


Figure 37. Redesigning the sTF architecture for a more responsive output. (*A*) Redefining the Ste12 domains. Here the induction and activation domains of Ste12 are considered as one unit, the pheromone responsive domain (PRD, 216-688). (*B*) Design of sTFs created in this study. An example sTF fusing the full length LexA bacterial repressor protein to the Ste12 PRD via the SV40 nuclear localisation signal (NLS). The LexA-PRD sTF is targeting a synthetic promoter consisting of six LexO sequences upstream of the *LEU2* core promoter driving the expression of sfGFP. (*C*) α-factor dose-response of the minimised pheromone response pathway using the native Ste12 TF (Design 3) and the LexA-PRD sTF (LexA-PRD). Mean fluorescence levels for the medium *RPL18B*, and strong *TDH3*, and *CCW12* constitutive promoters driving the expression of sfGFP have been included as a reference. (*D*) Maximum x-fold change in signal after activation of the two pathway variants compared to the Quasi-WT strain. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

Following these new design considerations, we constructed a new sTF from the direct fusion of the full-length LexA bacterial repressor and the Ste12 PRD, separated by the SV40 nuclear localisation sequence (NLS)[288] (LexA-PRD) (**Figure 37B**). We then targeted the LexA-PRD sTF to a synthetic promoter consisting of six repeats of the LexA operator sequence (LexO) upstream of the *LEU2* core promoter driving sfGFP expression, using the same conditions as Design 3 for refactoring the pathway components. The α-factor dose-response of the LexA-PRD-mediated pathway exhibited a significantly improved signal output compared to the Ste12 equivalent (Design 3) while maintaining a tight OFF-state (**Figure 37C**). The maximum output using LexA-PRD was also significantly greater than the strongest constitutive promoter in the YTK toolkit (*pCCW12*), actually causing some cellular burden from the sheer amount of sfGFP protein being produced over the 4h assay (discussed in section **5.2.3**). This resulted in a 245-fold change in expression after activation, representing a 3.5- and 35-fold improvement over the Design 3 and Quasi-WT responses, respectively (**Figure 37D**).

4.2.2 Tuning the maximum pathway output using modular promoters

The newly designed LexA-PRD had demonstrated all of the properties we were looking for from an sTF: a tight OFF response in the absence of a signal, a large maximum signal output after activation, and an orthogonal DBD to decouple the mating pathway from endogenous PRE-genes. With the success of the new sTF format, we decided to drop STF2 and focus our efforts solely on LexA-PRD. To tune the output of the LexA-PRD-mediated pathway, we created a library of UASs, containing different numbers of the LexO sequence, and a library of core promoters, comprising the minimal promoter region from a number of endogenous yeast promoters.

The repetitive UAS sequences were manually optimised to reduce possible recombination in yeast by varying the regions between the LexO consensus, using the Evolutionary Failure Mode (EFM) calculator from the Barrick Lab to assess the overall stability[289]. The core promoters were designed as the sequence downstream of the consensus TATA box of constitutive YTK promoters and other popular yeast promoters[281], using a neutral AT-rich spacer to provide optimal spacing with the UAS[282]. These split promoter sequences were then formatted as type 2a and type 2b parts to allow the modular exchange of UAS and core promoter sequences during assembly (for a description of this new formatting see section **8.3.2**). We also addressed the modularity of the sTF by designing the LexA protein followed by the SV40 NLS as a type 3a part, and the Ste12 PRD as a type 3b part. The resulting parts, therefore, allow for the modular assembly of the sTF and promoter, allowing us to easily alter the synthetic promoter identity or create a new sTF (**Figure 38A**).

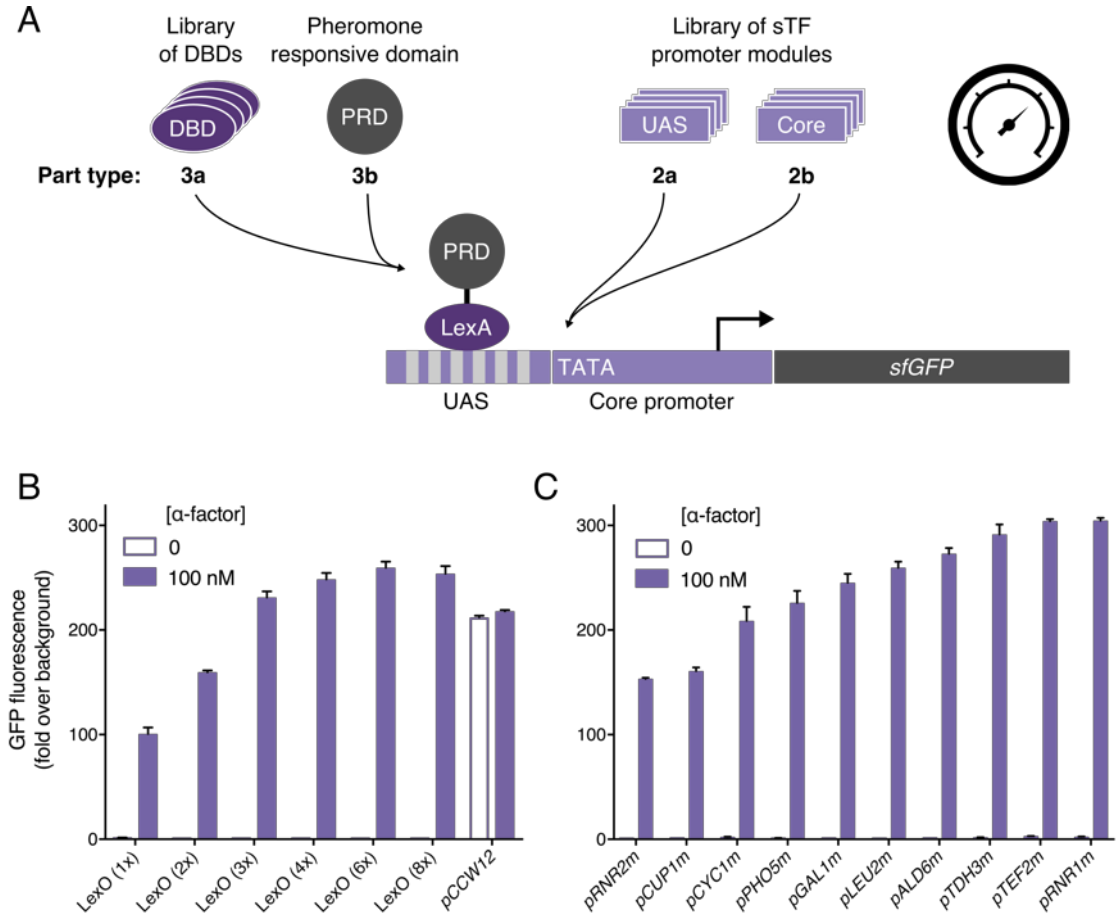


Figure 38. Tuning the pathway output using modular promoters. (*A*) sTFs created from the fusion of the full length LexA protein and the Ste12 PRD can be targeted to synthetic promoters with an interchangeable UAS and core regions. Changing the number of sTF binding sites in the UAS and the core promoter identity allows the maximum output of the system to be modulated. (*B+C*) Maximum α-factor-activated pathway expression mediated by the LexA-PRD sTF driving the expression from a synthetic promoter, using variants of the UAS and core promoter module, respectively. For maximum fold change in expression of all UAS and core promoter conditions see **Supplementary Figure S2**. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates.

Using these new tools, we investigated changes to the synthetic promoter identify as a way of tuning the maximum pathway output. Fixing the core region as the minimal *LEU2* promoter (*pLEU2m*) and varying the number of LexO sites allowed us to modulate the maximum output of the pathway over an almost 3-fold range (**Figure 38B**). Similarly, fixing the UAS as 6x repeats of the LexO sequence and changing the core promoter identity allowed us to vary the maximum output of the system over a 2-fold range (**Figure 38C**). The graded distribution in both libraries offers 60 discrete levels of maximum signal output using all combinations of the promoter modules. Importantly, these changes to the maximum pathway output were achieved without changing the concentration of the TF.

4.2.3 Ligand tuneable synthetic transcription factors

We have demonstrated a new way to tune the output of the mating pathway by using the LexA bacterial repressor to redirect the response to modular promoters. In this context, the LexA protein functions as an immutable DNA binding domain that sits on available LexO sequences, with the PRD awaiting activation to initiate transcription. The tuning of the output is decided entirely by the synthetic promoter identity. However, as described in section **1.2.5**, TFs exist which themselves can be induced to bind or release DNA in response to a ligand. Using this mechanism, the copy number of the transcription factor would remain fixed, but the DNA binding or nuclear localisation could be altered as an alternative means of tuning the maximum pathway output.

To explore whether ligand-inducible DBDs was another viable strategy for tuning the maximum output of the system, we created two new type 3a parts: the bacterial TetR repressor protein from bacteria[77] and the synthetic Z3E domain (a fusion of the Zif268 DNA binding domain and the ligand binding domain of the human estrogen receptor)[290]. First, we targeted the TetR domain fused to the Ste12 PRD (TetR-PRD) to a synthetic promoter comprising six repeats of the TetR operator (TetO) upstream of the *LEU2* core promoter driving sfGFP expression (**Figure 39A**). We then probed the maximum response of the system over a range of anhydrotetracycline (aTc) concentrations to determine the ligand tunability (**Figure 39B**). The TetR-PRD displayed a binary behaviour to the concentration of aTc, resulting in either full or no pathway activity. Although we were unable to obtain a graded tunability of the maximum signal output using this sTF, it may offer a useful way to prevent the cells from reporting the pathway activity by acting as an effective OFF-switch, demonstrating no detectable leak in the presence of saturating levels of aTc (**Figure 39C**).

Next, we targeted the Z3E domain fused to the Ste12 PRD (Z3E-PRD) to the *pZ3* promoter (a modified *GAL1* promoter containing six Zif268 binding sequences[290]) driving the expression of sfGFP (**Figure 39D**). As before, we examined the maximum pathway output over a range of β-estradiol concentrations. However, unlike TetR-PRD, Z3E-PRD demonstrated a highly-graded variation in the maximum output, albeit with a much lower maximum output at saturating levels of inducer (**Figure 39E**). This sTF would, therefore, be very suitable for tuning the maximum signal output from the refactored pathway as any output over a 60-fold range should be achievable using fine-tuning of β-estradiol concentrations. Furthermore, this sTF also displayed very clean ON-switch properties, providing a way to report the pathway activity only in the presence of β-estradiol (**Figure 39F**).

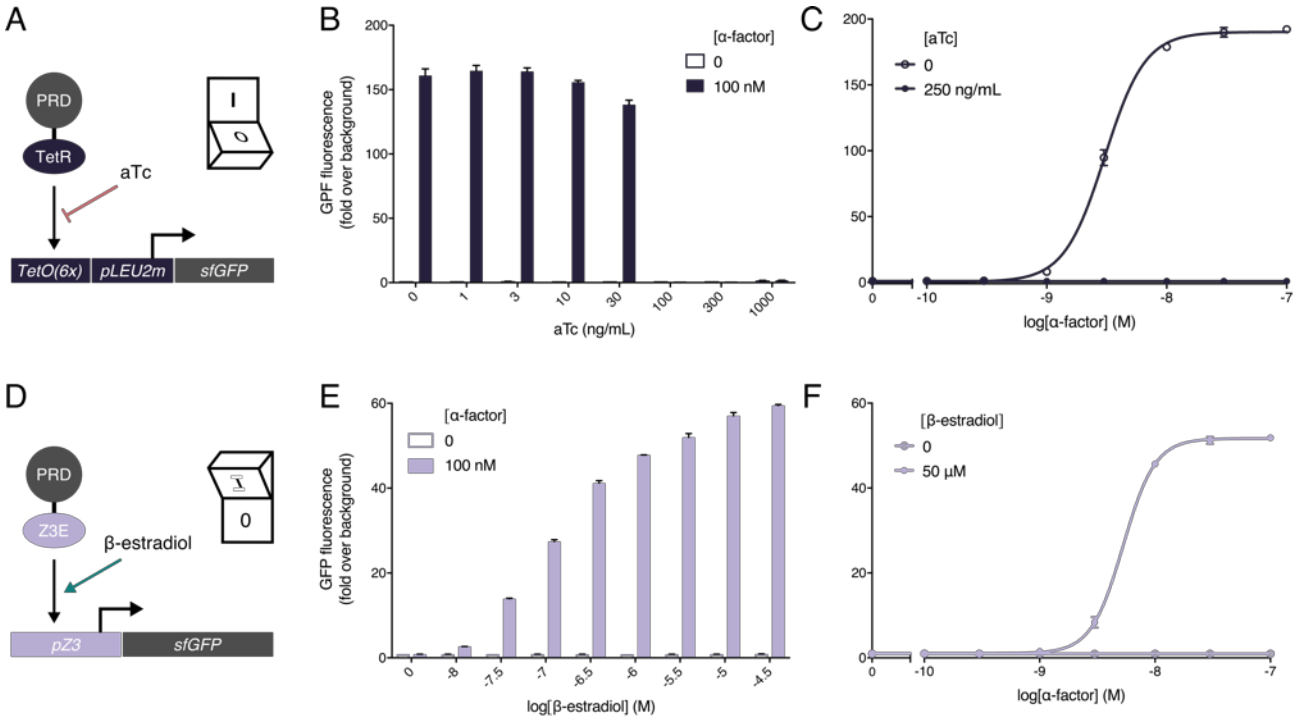


Figure 39. Ligand tuneable synthetic transcription factors. (*A*) A fusion of the DNA binding domain of the bacterial transcriptional repressor TetR with the Ste12 PRD, targeting a synthetic promoter comprising 6x repeats of the TetO binding sequence followed by the *LEU2* core promoter driving the expression of sfGFP. (*B*) Inducing maximum α-factor-induced expression of the TetR-PRD-mediated signalling pathway over a range of aTc concentrations. (*C*) α-factor dose-response curve of the TetR-PRD-mediated pathway with and without saturating levels of aTc. (*D*) A fusion of the Z3E (itself a fusion of the Zif268 DNA binding domain and the ligand binding domain of the human estrogen receptor) with the Ste12 PRD, targeting the *pZ3* promoter (a modified *GAL1* promoter containing six Zif268 binding sequences) driving the expression of sfGFP. (*E*) Inducing maximum α-factor-induced expression of the Z3E-PRD-mediated signalling pathway over a range of β-estradiol concentrations. (*F*) α-factor dose-response curve of the Z3E-PRD-mediated pathway with and without saturating levels of β-estradiol. aTc and β-estradiol concentrations were added at 0 h in the assay. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

4.2.4 An RNA programmable synthetic transcription factor

With the demonstration of the flexibility and modularity of the Ste12 PRD for targeting new promoters by substituting the DBD for synthetic alternatives, we decided to explore whether we could use the nuclease deficient Cas9 (dCas9) protein for creating an RNA programmable sTF[291]. Although this would not aid our work in tuning the dose-response characteristics, CRISPR logic is starting to become a powerful platform for programming complex transcriptional gene circuits in yeast[292]. This system may, therefore, be useful for integrating extracellular signals directly into these circuits. Furthermore, this would provide a new means for probing native yeast genes; a tool which has been recently developed for mammalian systems[293].

To assess whether a fusion between dCas9 and the Ste12 PRD (dCas9-PRD) would work, we designed a synthetic promoter which consisted of a random 20 bp sequence followed by a

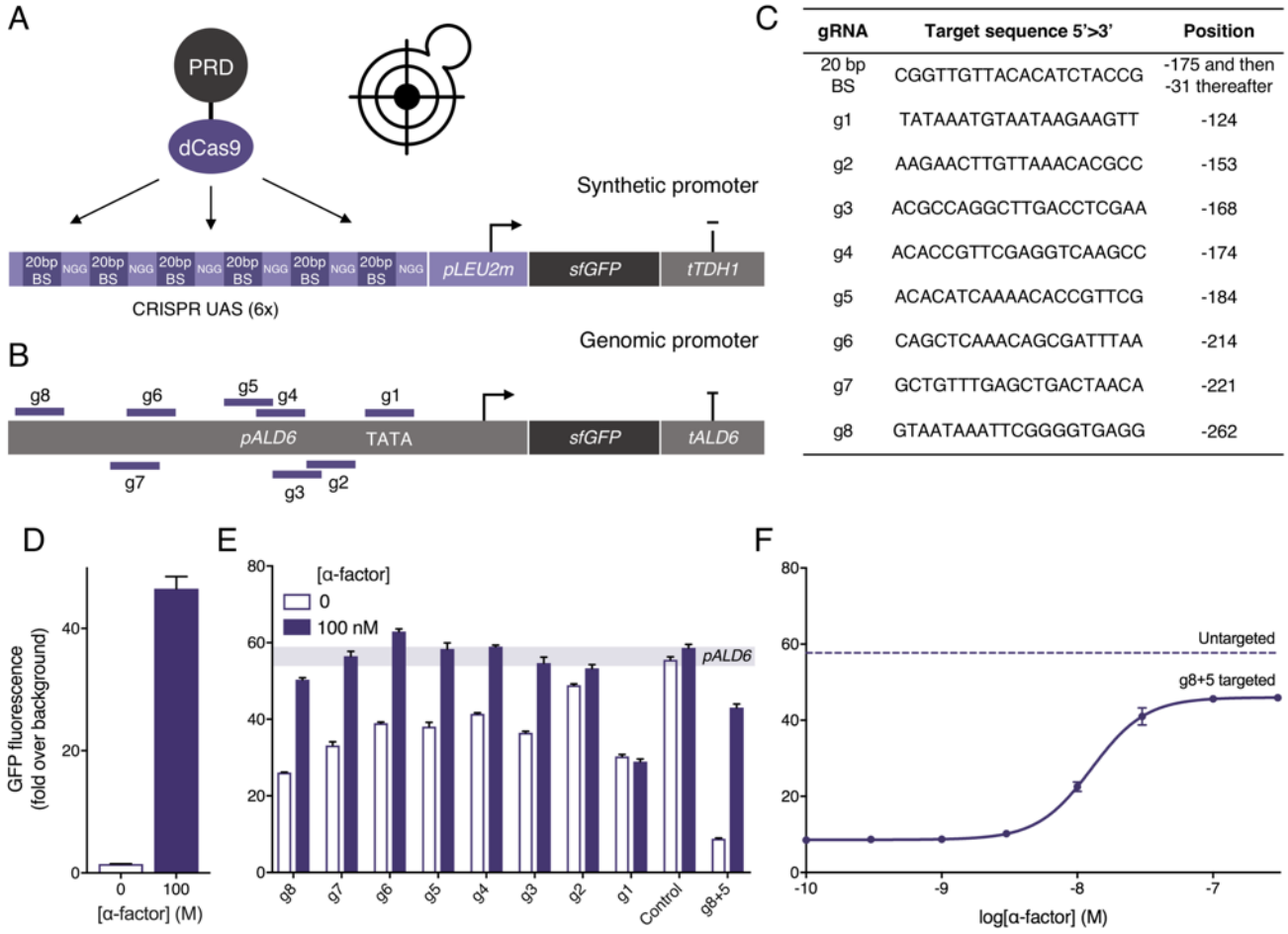


Figure 40. A programmable synthetic transcription factor. (*A*) Fusion of the Ste12 PRD to catalytically inactive Cas9 (dCas9) targeting a synthetic promoter consisting of six identical 20 bp CRISPR targeting sequences followed by a protospacer adjacent motif (PAM) sequence upstream of the *LEU2* core promoter driving the expression of sfGFP. (*B*) 8 locations on promoter of the *ALD6* gene where the ORF has been substituted for sfGFP. (*C*) Guide sequences for the repeated 20 bp targeting sequence in the synthetic promoter and 8 targets on the genomic *ALD6* promoter (*D*) On/off response of dCas9-PRD targeting the synthetic promoter. (*E*) On/off response of the minimised mating pathway targeting the genomic *pALD6-sfGFP-tALD6*. Simultaneous targeting of g5 and g8 is also included as it produced a more significant on/off response (see **Supplementary Figure S3** for details on the multiplexed targeting of the ALD6 promoter). (*F*) α-factor dose-response of the dCas9-PRD-mediated pathway targeting *pALD6-sfGFP-tALD6* at g5 and g8. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

protospacer adjacent motif (PAM) sequence repeated six times, upstream of the *LEU2* core promoter driving sfGFP expression (**Figure 40A**). We then targeted dCas9-PRD to the synthetic promoter by co-expressing a gRNA coding the 20 bp sequence and measured the ON/OFF response of the system (**Figure 40C**). Although the response was reasonable, it was not as great as we had seen with our other sTF variants. This was possibly due to the organisation of the PRD, as the size and spacing of dCas9 may have limited multimeric interactions. Alternatively, the DNA binding efficiency may not have been optimal due to the sequence of the guide[294]. Nonetheless, activating the synthetic promoter was a clear demonstration that the dCas9-PRD worked as an RNA programmable sTF.

Next, we investigated the targeting of dCas9-PRD to endogenous yeast promoters. We substituted the ORF of the *ALD6* gene for the in-frame ORF of sfGFP in the yWS677 base strain and targeted dCas9-PRD to eight locations on the *ALD6* promoter on, and upstream of the TATA consensus (**Figure 40B**). We then measured the ON/OFF response of the system, which exposed a novel behaviour (**Figure 40D**). In the absence of a stimulus, the expression of sfGFP was reduced compared to the non-targeted control. The output from the *ALD6* promoter could then be returned to wild type levels by stimulating the pathway. However, targeting at the TATA box resulted in a decrease in sfGFP expression with no measurable inducibility, highlighting the importance of this sequence for transcriptional initiation in Pol II promoters[295]. Multiplexed targeting of dCas9-PRD using several guides was used to increase the dynamic range of this system, reducing the OFF-state expression to 10%. This could then be restored to 90% of the wild type with pathway activation (**Figure 40E**).

To our knowledge, this is the first instance of CRISPR transcriptional regulation being used to repress the expression of a gene which can then be relieved in a dose-dependent manner. This could be a useful tool for probing the function of any particular gene by effectively knocking down gene expression and then titrating it back at tuneable levels. Although it should be mentioned that this system would not be immediately transferrable to all yeast strains, as this would require a number of the gene disruptions required in our yWS677 base strain.

4.2.5 Tuning maximum signal output of the refactored pathway

With the newly developed tuning knob for modulating the maximum signal output, we decided to readdress the iterative tuning of the minimised mating pathway in section **3.2.5**. Following on from the Design 3 pathway, we swapped the Ste12 TF for the LexA-PRD sTF and used the best performing synthetic promoter (*LexO(6x)-pLEU2m*) to drive sfGFP expression and report on pathway activity. The dose-response of the final refactored pathway (Design 4) demonstrated significantly improved biosensing characteristics across the board compared to the Quasi-WT response (**Figure 41**). The substantial improvement to the maximum output of the system enhanced both the dynamic and operational range over the previous design, albeit with a very slight increase in the basal activity (1.2-fold over background (Design 4) compared to 1.0-fold over background (Design 3)). This system was now marginally more sensitive than the Quasi-WT strain, outperforming it in every other class. Importantly, we had achieved this using the constitutive expression of the refactored pathway components without using feedback.

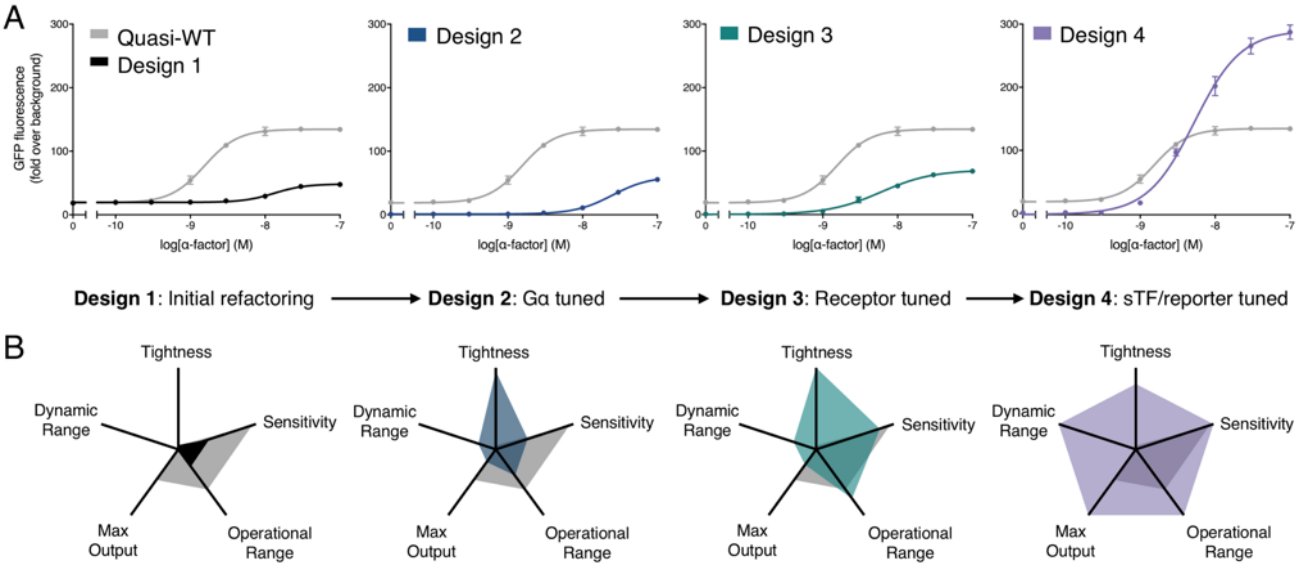


Figure 41. Tuning the minimised pheromone response pathway through iterative refactoring (Design 1-4). (*A*) α-factor dose-response curves of four sequentially refactored pathway designs compared to Quasi-WT strain. Following on from Design 3 (**Figure 35**), the native Ste12 TF was substituted for the LexA-PRD sTF and targeted to the best performing synthetic promoter (*LexO(6x)-pLEU2m*) (Design 4). Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit. (*B*) Dose-response characteristics of the four refactored pathway designs. The final pathway design (Design 4) demonstrated significantly improved dose-response characteristics compared to all previous refactored pathway designs and the Quasi-WT strain. Tightness is defined as the reciprocal of basal activity and the dynamic range is defined as (maximum output/basal activity). Sensitivity and operational range were determined from the fitted curve, defining sensitivity as the lowest concentration for which a >2-fold change in GFP expression is seen, and operational range as the concentration span between the sensitivity and the lowest concentration that gives a GFP expression within 2-fold of the maximum. All values were then normalized to the minimum measurable value and the maximum calculated value in the dataset. For a list of the obtained values ± standard deviation see **Supplementary Table S1**.

We have now successfully demonstrated the rational tuning of the sensitivity, leakiness, and maximum signal output of the minimised mating pathway. Remarkably, all of these points of tuning were performed by changes to promoter identity of the receptor, Gα, and reporter. This is significant for biosensing applications. As the changes in the system are limited to the non-coding regions, this should represent a generalisable set of rules for tuning the response of any heterologous GPCR. It should now be possible to port new receptors into the system without requiring extensive protein engineering to achieve the desired biosensing response by simply adjusting the expression of the components. We could now explore the rational tuning of heterologous GPCR-based biosensors in yeast using this platform. However, before beginning this work, there were a couple of system properties we needed to validate. First, we needed to ensure we were expressing the Gα at the appropriate level for heterologous GPCRs, and secondly, we needed to validate the sTFs had actually decoupled the mating pathway from the mating response.

4.2.6 A novel constitutive coupling assay validates appropriate Gα expression levels

Constitutive receptor activity is a common property of many wild type GPCRs[296,297]. This ligand-independent mode of action is finely-tuned in mammalian cell types and is important for correct physiological function[297]. As a result of constitutive receptor activity, basal G protein activity also increases, leading to increased levels of pathway signalling in the absence of a stimulus (i.e. a leaky system). In fact it was the study of the constitutively active histamine and muscarinic receptors that suggested overexpressing the Gα was a viable approach for tuning the leakiness in our refactored pathway[253,254]. If the Ste2 receptor was also a constitutively active receptor, there was a chance we were overexpressing the Gα in a way that could mask the response of non-leaky heterologous receptors when ported into our system. Considering the context of the native Ste2 receptor to respond only in the presence of a viable mating partner, it would be surprising if this receptor had constitutive activity. However, while mutant variants of Ste2 have demonstrated an increased constitutive activity when compared to the wild type receptor[298], the actual basal activity of WT Ste2 has never been reported. To ensure we were expressing the Gα at a level appropriate for the full activity of all GPCRs, we decided to study the constitutive activity of WT Ste2.

To measure the constitutive activity of Ste2, we needed a simple assay that would allow us to fix the expression of the receptor while somehow probing its constitutive activity in the absence of a stimulus. We decided to revisit our receptor/G protein model from section **3.2.3**. In this model, we demonstrated the leakiness of the system was dependent on the concentration of the Gα. By setting the receptor to 2% constitutive activity and simulating this model with and without the receptor, we demonstrated a shift in the basal activity of the system (**Figure 42A+B**). The difference between these two conditions revealed a characteristic bell-shaped distribution would be produced from a constitutively active receptor (**Figure 42C**). If this model was accurate, we could then use the shape of the distribution over the Gα promoter library to determine whether the Ste2 receptor was constitutively active or not.

To validate this behaviour experimentally, we chose the human adenosine-sensitive A2BR receptor, which has recently been shown to have constitutive activity[299] and is also a well-studied receptor in yeast[209,300]. Varying the expression of Gpa1 using a promoter library and measuring the basal activity of the system with and without the co-expression of the A2BR receptor revealed a similar trend to the model (**Figure 42D**). Moreover, the relative change in basal expression with and without the receptor demonstrated the predicted bell-shaped distribution, indicative of a constitutive receptor (**Figure 42E**). We then reperformed this experiment using a C-terminally truncated version of the A2BR receptor (A2BR-CΔ-tail[301]) to confirm the receptor/G protein interaction was the mechanism of action, as the C-terminal tail is essential for G protein signalling[301,302] (**Figure 43B**). The loss of the bell-shaped

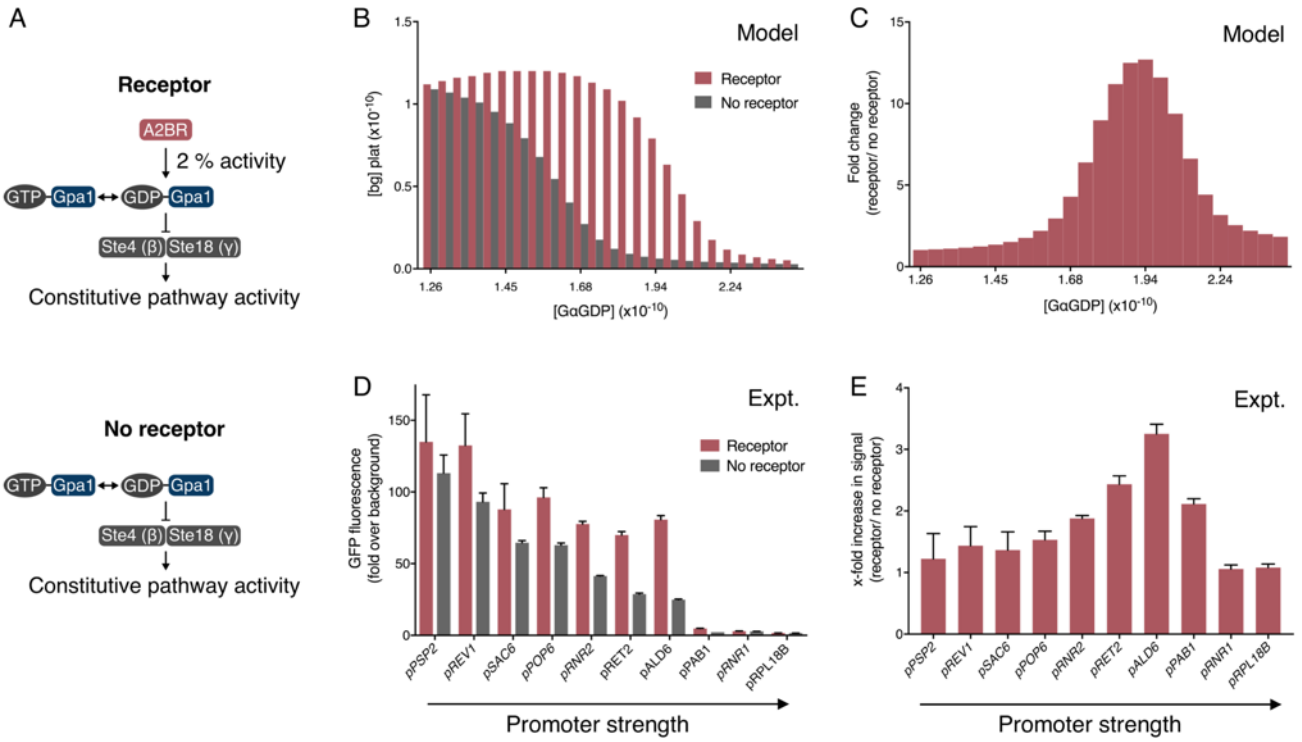


Figure 42. Measuring the constitutive activity of GPCR receptors. (*A*) Pathway activity in the presence and absence of the constitutively active adenosine A2BR receptor coupled to the pathway via the native Gα, Gpa1. (*B*) Model of the minimised pheromone response pathway over a range of initial GαGDP concentrations in the presence and absence of a constitutively active receptor, set to 2% activity in the absence of a stimulus. (*C*) Fold change difference between the receptor and no receptor model demonstrating a bell curve distribution. (*D*) Experimentally coupling the A2BR receptor to a library of the pathway variants with varied levels of Gpa1 expression compared to a no receptor control. (*E*) Fold change difference between the A2BR pathway coupling and no receptor control. All modelling developed and performed by Hitoshi Yamauchi, Jack Mead, and Graham Ladds[264]. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates.

distribution when expressing the A2BR-CΔ-tail mutant demonstrated the behaviour was indeed G protein-dependent and this assay would be suitable for measuring the constitutive activity of the WT Ste2 receptor.

Finally, we performed the constitutive coupling assay with the WT Ste2 receptor (**Figure 43C**). The flat distribution across the Gα library confirmed Ste2 as a tight receptor, with no significant levels of constitutive activity. This suggested that we were expressing the Gα at precisely the right levels in the Design 2-4 strains for reducing intrinsic pathway activity without masking ligand-induced receptor activity. The Design 4 conditions should, therefore, be appropriate for porting heterologous GPCRs, with the option of increasing Gα expression for receptors displaying higher levels of constitutive activity.

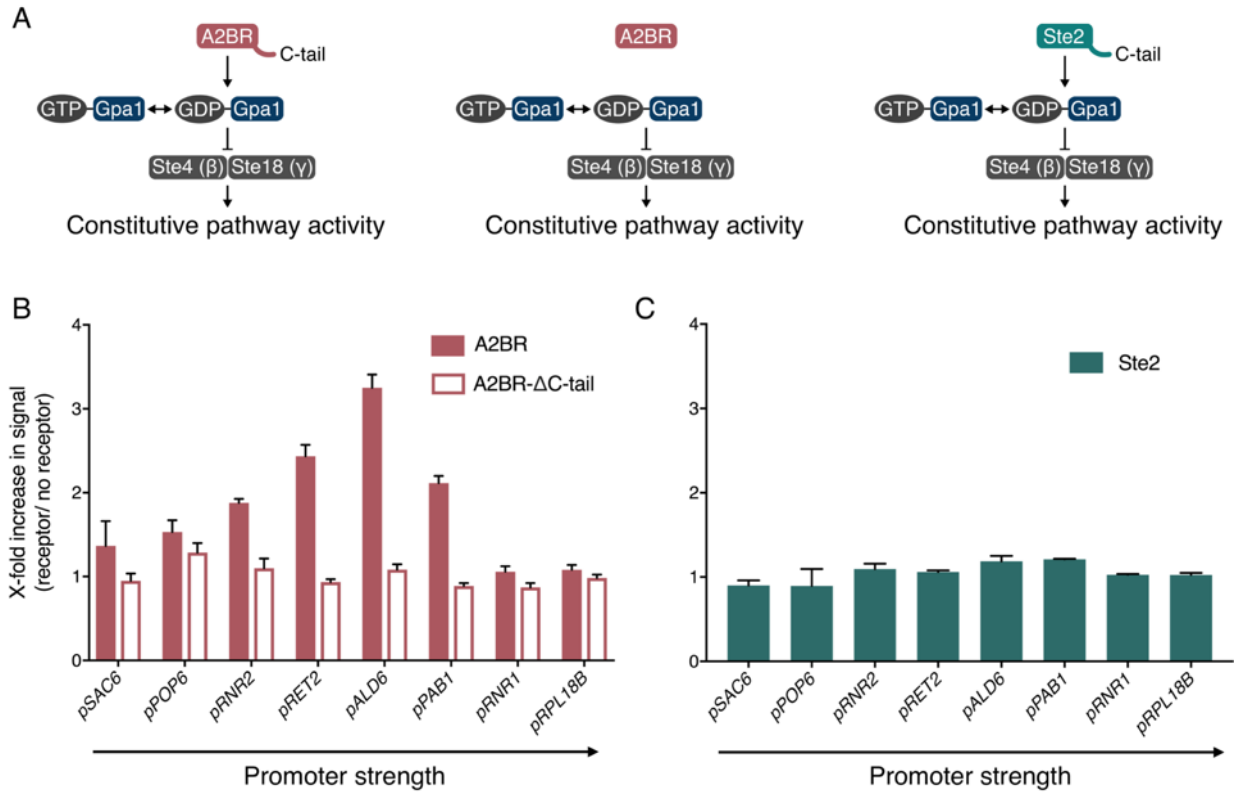


Figure 43. Demonstrating the tightness of the Ste2 receptor using the constitutive coupling assay. (*A*) Expression of the A2BR receptor, a C-terminally truncated A2BR receptor (A2BR-ΔC-tail, 1-293), and the Ste2 receptor using the strong *TDH3* promoter. (*B*) Constitutive pathway activity in the presence of the full length and C-terminally truncated A2BR receptors. A flat distribution when expressing the C-terminally truncated A2BR receptor provides evidence for receptor/Gα interactions as the sole mechanism for constitutive pathway activity from a leaky receptor. (*C*) Constitutive pathway activity in the presence of the Ste2 receptor. A flat distribution suggests the Ste2 receptor is very tight and has minimal constitutive activity in the off-state. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates.

4.2.7 Confirming the decoupling of the mating response

Our final validation to perform before exploring heterologous GPCRs was the decoupling of the minimised signalling pathway from the mating response. The substitution of the Ste12 DBD for LexA, or any of the other orthogonal domains, should have prevented the TF from targeting the PRE-genes normally upregulated in the pheromone response. To confirm whether this was indeed true, we performed RT-qPCR to measure the relative change in expression of the refactored pathway components and several of the most highly activated genes in the mating response[259,303] in the Quasi-WT and Design 4 strains after stimulation with α-factor (**Figure 44A**). We used *HTB2* as the reference gene, as it showed minimal variation during the mating response (as shown in **Figure 26**), and GFP as the positive control (12.4-fold change in Quasi-WT and 664.3-fold change in Design 4). All of the genes in the Quasi-WT strain demonstrated a change in expression, with the most significant changes in the *FIG1*, *PRM2*, *CIK1* genes, as expected. Conversely, none of the genes in the Design 4 strain

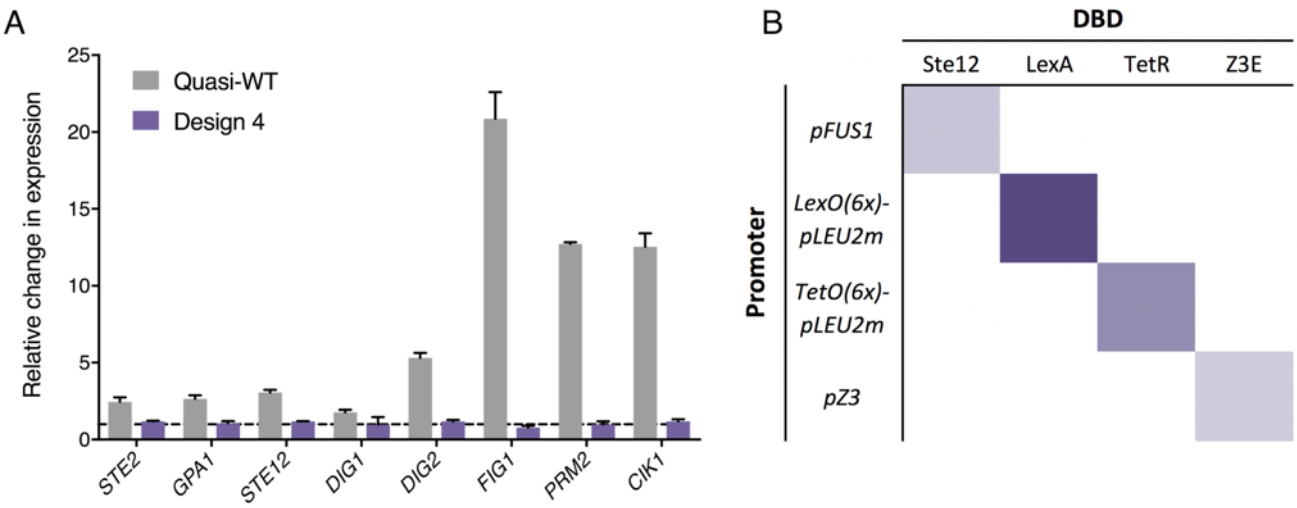


Figure 44. Decoupling the minimised pathway from the mating response using orthogonal sTFs. (*A*) Relative changes to the expression of the key pathway components and highly induced *PRE*-genes after pheromone stimulation in the Quasi-WT and Design 4 strains as determined by RT-qPCR. No change in the expression of *PRE*-genes in the Design 4 strain demonstrates the use of sTFs completely decouples the pathway from the mating response. Results are means ± standard deviation from triplicate technical repeats. (*B*) Orthogonality between the sTF/promoter systems. Maximum fold-change in expression of sfGFP using the Ste12, LexA-PRD, TetR-PRD, and Z3E-PRD sTFs in all combinations with the *pFUS1*, *LexO(6x)-pLEU2m*, *TetO*(6x)-LEU2m, and *pZ3* promoters after induction with 100 nM of α-factor. Purple colour represents the mean fold-change in expression after activation from triplicate isolates.

were upregulated, confirming the use of the sTF, LexA-PRD, had decoupled the pathway from the mating response, removing the final source of feedback within the minimised pathway.

We also measured the relative change in the expression of the negative repressors of Ste12, Dig1 and Dig2, after pheromone stimulation. In the Quasi-WT strain, both Dig1 and Dig2 are both seen increasing. This may explain why the maximum signal in the refactored designs using Ste12 was limited compared to the Quasi-WT response, as the negative repressors would be increasing while the TF remained fixed. Switching to sTFs removed this positive transcriptional feedback, so the levels of the sTF, Dig1, and Dig2 would remain balanced at their initial concentrations, enabling a high level of transcriptional output throughout the entire response.

Next, we looked at the orthogonality of the DBDs between each other to investigate whether they could be used within the same system. By combining the WT Ste12, LexA-PRD, TetR-PRD, and Z3E-PRD TFs in all combinations with their cognate promoters, *pFUS1*, *LexO(6x)-pLEU2m*, *TetO(6x)-pLEU2m*, and *pZ3*, we were able to measure their on- and off-target activity (**Figure 44B**). The TFs were highly specific for their cognate promoters, showing no measurable activity when targeting any of the other promoters. TF orthogonality ensures any additional gene circuits introduced into the system, specifically using these DBDs, will remain insulated from the GPCR signalling pathway (discussed in section **6.2.3**).

4.2.8 A heterologous GPCR-based biosensor for *S. pombe* P-factor pheromone

As an initial demonstration of our platform for developing heterologous GPCR-based biosensors, we decided to domesticate the Mam2 receptor from the equivalent mating pathway in the yeast *Schizosaccharomyces pombe*, which is sensitive to a 23 amino acid peptide pheromone called P-factor[304]. Following the optimised conditions, we identified for the Ste2-mediated α-factor response, we constructed a Mam2 sensor strain, using a Gα chimera where the final five residues of Gpa1 had been substituted for the *S. pombe* equivalent. (**Figure 45A**). We then measured the response of the Mam2 sensor to P-factor and compared it to previously characterised data of the receptor in its native context, from Croft *et al*.[305] (**Figure 45B**). The Mam2 sensor strain behaved almost exactly as in *S. pombe*, achieving an almost identical potency (pEC50) to P-factor. Furthermore, the Mam2 sensor displayed no detectable leakiness and exhibited a 180-fold change in the signal after activation.

The successful domestication of the *S. pombe* Mam2 receptor into our GPCR-based biosensor platform suggests the optimisation we had performed for the Ste2 α-factor response would also be suitable for other GPCRs (**Table 5**). This would allow us to bring in new heterologous GPCRs and expect a reasonable response, without requiring the several weeks of iterative refactoring that were required to optimise the α-factor response. Further tuning to create the desired response could then be performed from this starting point, allowing the user to focus on the higher-level aspects of the biosensor design.

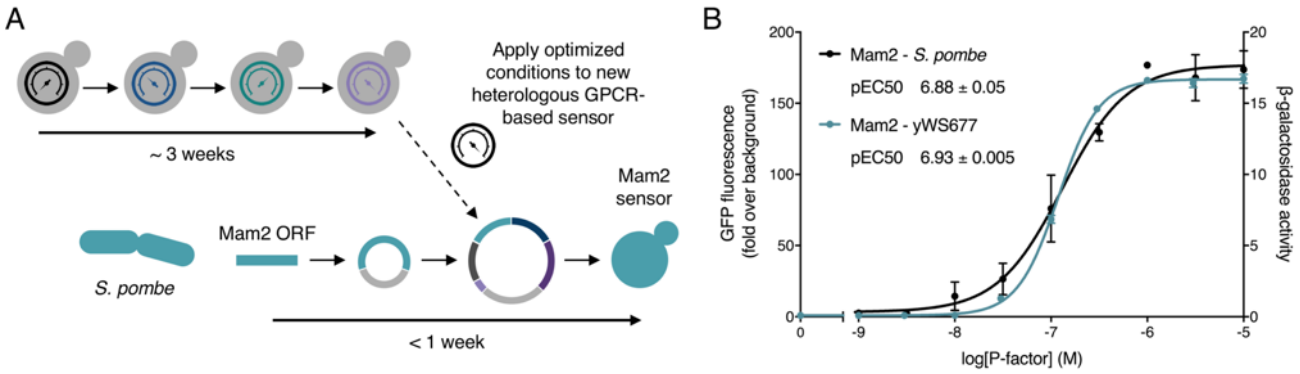


Figure 45. Domesticating the *S. pombe* Mam2 receptor. (*A*) The conditions identified during the 3-week optimisation of the α-factor response in **Figure 41** were directly applied to the design of a P-factor sensitive sensor strain using the *S. pombe* Mam2 receptor, which was built in less than a week. To improve coupling of the Mam2 receptor to the *S. cerevisiae* Gα, the final 5 residues of Gpa1 were substituted with the *S. pombe* Gα equivalent. (*B*) P-factor dose-response curves of the Mam2 sensor (turquoise) compared to the WT Mam2 response in its native *S. pombe* background (black) using previously characterised data from Croft *et al*.[305]. Slight differences in the shape of the curves are likely due to differences in the assay length and choice of reporter. The Mam2 sensor strain displayed a tight response with a high dynamic range and an almost identical potency (pEC50) to P-factor compared to the *S. pombe* system. This suggests the optimised conditions identified during the iterative refactoring of Ste2-mediated α-factor response are transferable to heterologous GPCRs. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. *S. pombe* Mam2 dose-response data was kindly provided by Graham Ladds and represents P-factor-dependent transcription of β-galactosidase using the sxa2 promoter, taking measurements 16 hours after stimulation. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

Table 5. Conditions for domesticating heterologous GPCRs.

Module position	Part Type				
	2a	2b	3a	3b	4
1	*LexO(6x)*	*pLEU2m*	*sfGFP*		*tTDH1*
2	*pCCW12*		*Receptor*		*tSSA1*
3	*pPGK1*		*Gα*		*tENO2*
4	*pRAD27*		*LexA*	*PRD*	*tENO1*

4.3 Conclusions

In this chapter, we explored the use of sTFs for achieving new behaviours in our minimised signalling pathway. We redefined the induction and activation regions in the Ste12 TF as a single pheromone-responsive domain. This allowed us to create a range of modular and highly-responsive sTFs that we could redirect to synthetic promoters or endogenous genes, later demonstrating how this completely decouples the minimised pathway from the mating response. We then used modular promoters and ligand-inducible DNA binding domains as a new approach for tuning the maximum output of the system. Using this new tuning knob, we further optimised the refactored signalling pathway, achieving a response with improved biosensing characteristics compared to the Quasi-WT strain, with the final sources of feedback removed from the system. A novel constitutive coupling assay was then developed to determine the constitutive activity of the Ste2 receptor to ensure the optimisation conditions for α-factor biosensing were appropriate for other receptors. Finally, we demonstrated our platform for porting heterologous GPCR receptors by creating a biosensor for the *S. pombe* peptide pheromone, P-factor.

5 Community-based approaches for tuning the Hill slope

5.1 Introduction

With the platform we have developed over the last three chapters, we are now able to rationally tune the sensitivity, leakiness, and response output of GPCR-based biosensors in yeast. However, we had yet to identify a way to tune one final and important characteristic of a biosensor – the operational range (the Hill slope of its response curve). This is a far more difficult task, as this property is largely determined by the ligand-binding properties of the receptor[306]. Some receptors will confer a narrow range switch-like behaviour, only requiring a small increase in signal to trigger maximum output (i.e. a digital response), while others will produce a wide operational range where the relationship between the signal and output is proportional (i.e. a linear response)[231]. For quantitative biosensor applications, a linear response is typically required, whereas a digital response is more desirable for a gene circuit[234].

As discussed in section **1.2.14**, previous efforts to tune the Hill slope of the pheromone response pathway have overlaid synthetic feedback loops onto the MAPK cascade[239–241]. By expressing positive and negative modulators of the pathway as an output of the response, the Hill slope can be modulated to either increase or decrease the operational range. This approach, independent of the receptor, could, therefore, provide a modular and generalisable way to address the final point of tuning in our GPCR-based biosensor platform. Although this strategy would reintroduce complexity back into our system, we had thus far stripped back the pathway back to its most basic structure. Without another round of major refactoring efforts, involving the remaining nine genes (**Table 2, Minimal**), building regulation on top of our current minimised pathway represents the only way to create this behaviour within our cells.

In this chapter, we explore strategies for tuning the Hill slope of GPCR-based biosensors. Starting with two heterologous receptors with comparatively digital and linear dose-response curves, we first investigate the use of intracellular feedback loops for modulating the operational range. We then move towards multicellular consortia as an alternative engineering approach, so that the Hill slope for a digital-like biosensor can be reduced to extend the operational range, while the Hill slope for a linear-like biosensor can be increased to narrow the operational range. Throughout the chapter we apply the lessons learned during the previous chapters, attempting to formalise a generalisable set of rules for tuning the remaining properties of the dose-response curve.

5.2 Results and discussion

5.2.1 Domesticating the human adenosine and melatonin responsive GPCRs

To design a new strategy for the linearisation and digitisation of the biosensor dose-response curve, we needed two receptors with contrasting Hill slopes. One response would be comparatively digital-like, and the other would be comparatively linear. We could then explore methods to reverse these behaviours, thus demonstrating the ability to tune the Hill slope. For this, we chose the human adenosine-responsive A2BR receptor, (as used in section **4.2.6**) previously shown to give a digital-like response in yeast[209], and the human melatonin-responsive MTNR1A receptor, previously shown to give a linear-like response in yeast[211].

We ported these two receptors into yWS677 using the conditions defined in **Table 5**, using a library of Gα chimeras comprising C-terminal transplants from all mammalian Gα variants (**Figure 46A-C**).

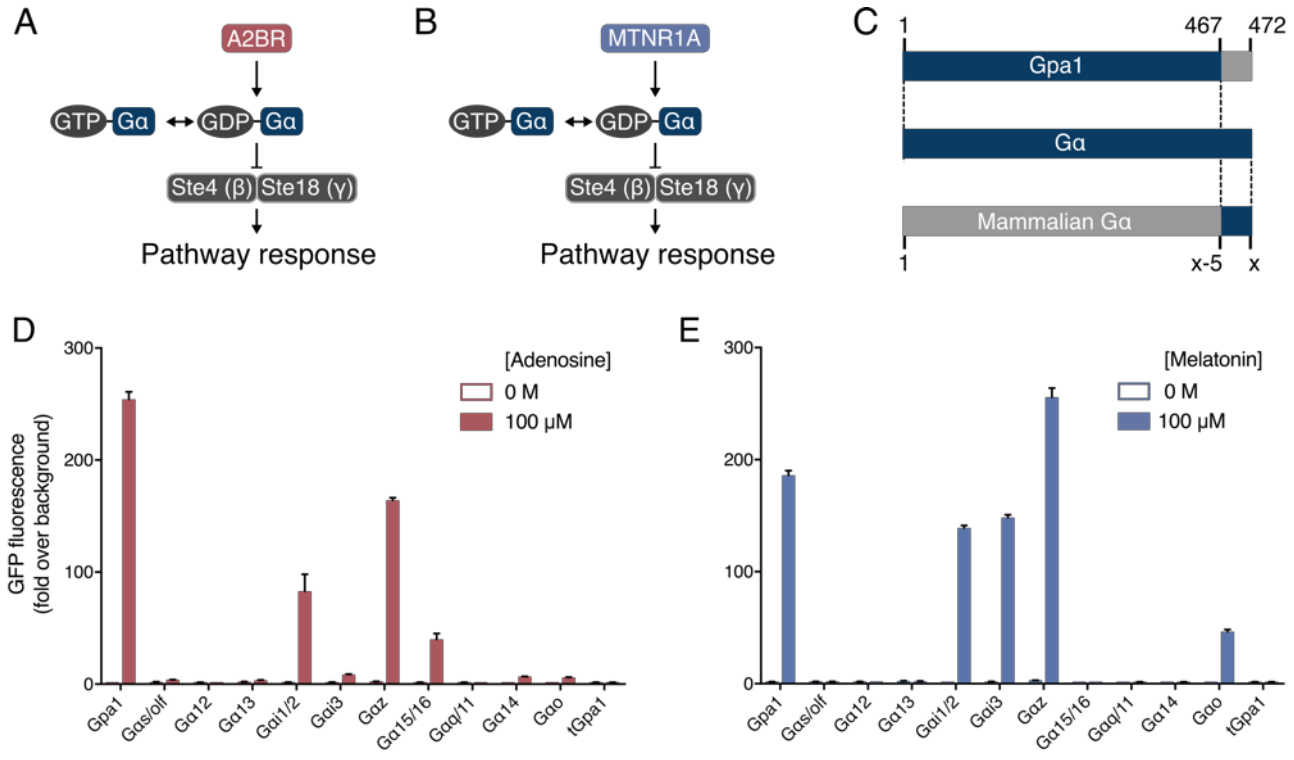


Figure 46. Domesticating the human A2BR and MTNR1A receptors. (*A+B*) Screening the adenosine-responsive A2BR and melatonin-responsive MTNR1A receptors against a chimeric Gpa1-Gα library (simply designated Gα here) to identify optimal coupling to the mating pathway. (*C*) Gpa1-Gα chimeras are composed of the Gpa1 protein where the final five residues have been substituted for the final five residues of the mammalian Gα subunits. The mammalian Gα C-terminal transplants are displayed on the x-axis of (*D+E*). (*D*) Coupling of the A2BR receptor to the Gpa1-Gα library in the presence and absence of saturating concentrations of adenosine (100 μM). (*E*) Coupling of the MTNR1A receptor to the chimeric Gpa1-Gα library in the presence and absence of saturating concentrations of melatonin (100 μM). As WT Gpa1 demonstrated highly responsive coupling to both A2BR and MTNR1A it was used for all future experiments using these two receptors. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates.

We then measured the ON/OFF response of the two receptors across the Gα library with saturating levels of their respective agonists to determine the optimal receptor/Gα coupling (A2BR, 100 µM adenosine; MTNR1A, 100 µM melatonin). The A2BR receptor coupled most efficiently with WT Gpa1 and showed a reasonable affinity for the Gαz chimera (**Figure 46D**). The MTNR1A receptor, on the other hand, displayed the most efficient coupling with the Gαz chimera, good coupling with WT Gpa1, and a reasonable affinity for the Gα1/2 and Gαi3 chimeras (**Figure 46E**). As both receptors exhibited good or optimal coupling to WT Gpa1, we decided to use this Gα for all future experiments using these receptors to reduce the amount of cloning that we would otherwise require.

Next, we reduced the expression of the receptor using the *HHF2* promoter, so that the response would sit in the middle of the sensitivity spectrum and allow us to manipulate the response in either direction, and then measured the dose-response of the two biosensor strains to their respective agonists (**Figure 47**). The A2BR sensor strain gave a relatively high Hill slope value of 1.32, whereas the MTNR1A sensor strain gave a comparatively low Hill slope value of 0.8. With the Hill slope values of the two sensor strains sitting either side of the standard value of 1.0, they represent the ideal starting points for engineering linearisation and digitisation.

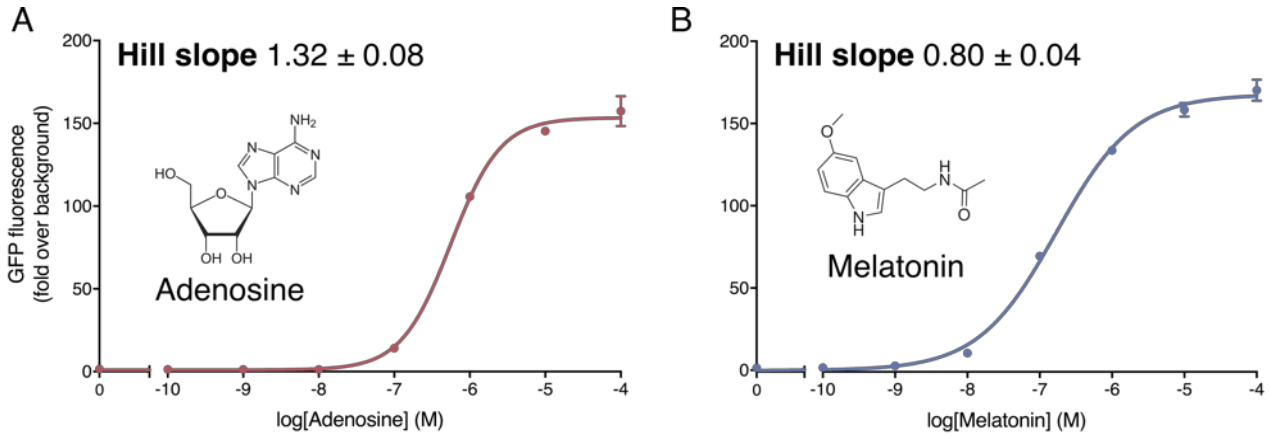


Figure 47. The comparatively digital and linear response of the A2BR and MTNR1A sensor strains. (*A*) Adenosine dose-response curve of the A2BR sensor strain. (*B*) Melatonin dose-response curve of the MTNR1A sensor strain. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

5.2.2 Overlaying intracellular feedback loops for tuning the Hill slope

Now that we had established a digital and linear sensor, we could explore the use of feedback loops for tuning the Hill slopes. To linearise the dose-response we would need to express negative modulators of the pathway as an output of the system[239]. We identified four negative modulators that

targeted different locations in the minimised pathway: (i) the negative regulator of G protein signalling, Sst2, (ii) the Gα, Gpa1, (iii) the negative regulator of Fus3 activity, Msg5, and (iv) the negative regulator of Ste12-mediated transcription, Dig1[179]. We then introduced these components into the A2BR sensor strain on a second module, as an output of pathway activity (**Figure 48A**) and measured the adenosine dose-response of the new system (**Figure 48B**). A significant decrease in the maximum signal output of the system was seen with the feedback of all negative components, demonstrating the negative feedback was working. To make the curves more comparable to each other, we normalised the data (**Figure 48C**). Although the curves had been significantly altered, the changes to the Hill slope were minimal (**Figure 48D**). Sst2 and Gpa1 demonstrated the most

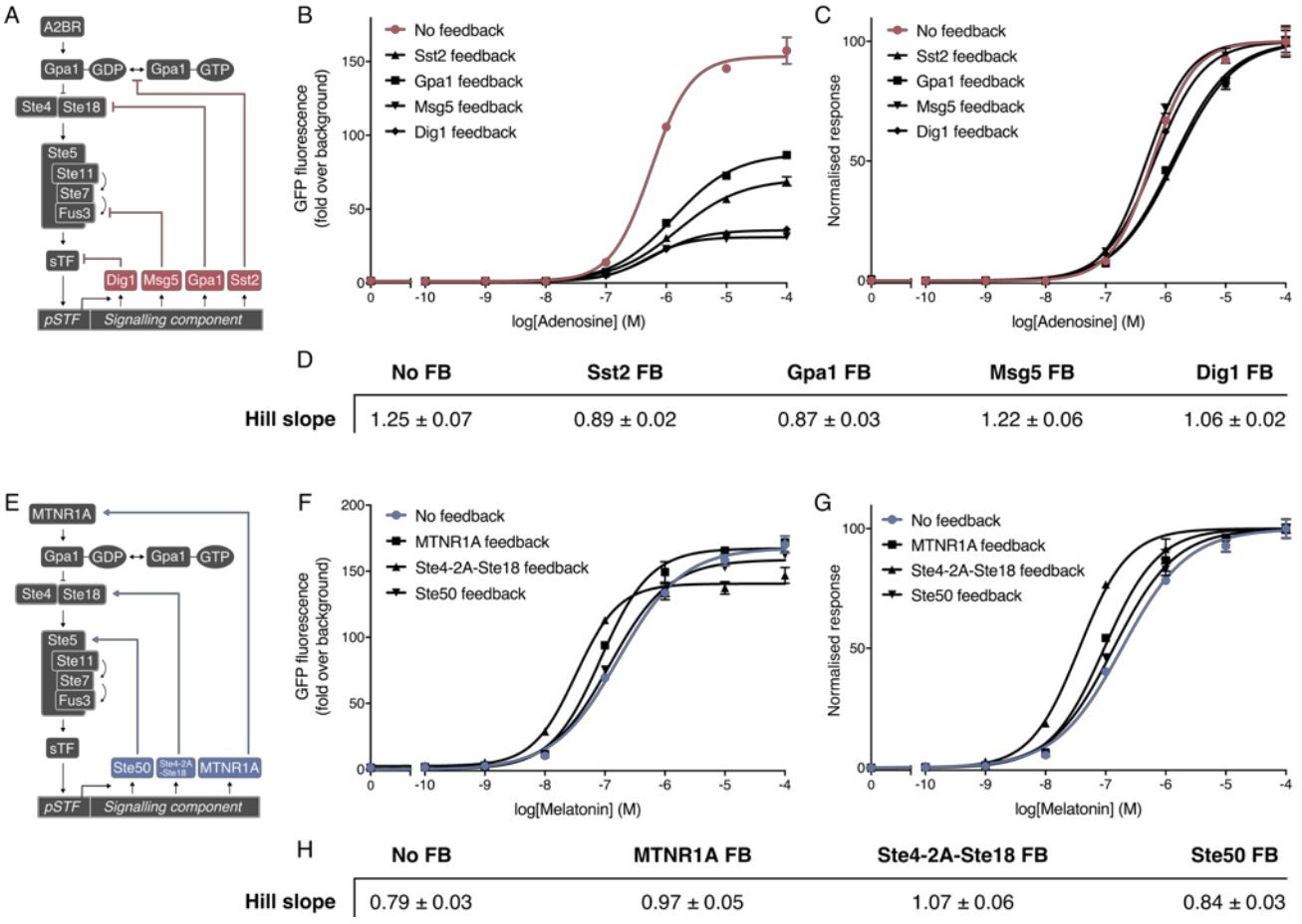


Figure 48. Overlaying feedback loops on the signalling pathway to modulate the Hill slope. (*A*) Transcriptional feedback of the negative pathway modulators Sst2, Gpa1, Msg5, and Dig1 to reduce the Hill slope of the A2BR sensor response. (*B*) Adenosine dose-response curves of the A2BR sensor (red) compared to the A2BR sensor with the negative feedback modules (black). (*C*) Normalisation of the A2BR negative feedback dose-response data demonstrating minimal changes to the slope. (*D*) Hill slope values from the normalised A2BR negative feedback dose-response curves. (*E*) Transcriptional feedback of the positive pathway modulators MTNR1A, Ste4-2A-Ste18, and Ste50 to increase the Hill slope of the MTNR1A sensor response. (*F*) Melatonin dose-response curves of the MTNR1A sensor (blue) compared to the MTNR1A sensor with the positive feedback modules (black). (G) Normalisation of the MTNR1A positive feedback dose-response data, demonstrating minimal changes to the slope. (*H*) Hill slope values from the normalised MTNR1A positive feedback dose-response curves. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

significant linearisation of the dose-response, as well as the smallest decrease in the maximum signal output and would, therefore, be the best choices out of the negative modulators.

Next, we applied the opposing strategy to the digitisation of the MTNR1A response. To digitise the dose-response we would need to express positive modulators of the pathway as an output of the system[239]. We identified three positive modulators that targeted different locations in the minimised pathway: (i) the MTNR1A receptor, (ii) a bicistronic Ste4-2A-Ste18 construct, which forms a Gβγ dimer (**Supplementary Figure S4**), and (iii) the positive modulator of Ste20 and Ste11, Ste50[179]. As before, we introduced these positive modulators to the MTNR1A sensor strain on a second module as an output of pathway activity (**Figure 48E**) and measured the melatonin dose-response of the new systems (**Figure 48F**). Unlike the A2BR sensors, the maximum output of the system did not change much, suggesting the system without feedback was already saturating signalling through the pathway. Normalising the data to compare between curves once again demonstrated that changes to the Hill slope were quite small, with the Ste4-2A-Ste18 construct exhibiting the most substantial increase (**Figure 48G+H**).

Although we had identified negative and positive modulators that were able to alter the Hill slope of both systems, none of them produced as large of a shift as we had anticipated. This may be due to a couple of reasons. As we had transitioned to orthogonal sTFs, we would not have the Ste12 autoregulatory feedback that usually amplifies transcriptional feedback mechanisms similar to our engineered system[205]. Additionally, we did not recruit the pathway modulators to the Ste5 scaffold, as realised by Bashor *et al.*[239], although this has previously been achieved without this mechanism[240,241]. Before committing to the adaptation of our system to the Bashor *et al.* strategy, we decided to search for novel alternatives in which we could leverage our established tuning principles for altering the operational range. As the effects within the intracellular system were limited, we decided to explore multicellular community-based approaches as an alternative.

5.2.3 Linearisation of the A2BR response using mixed populations

First, we tackled the problem of linearising the A2BR sensor response. To create a population that linearises the steep response of our adenosine-sensing cells, we took inspiration from a strategy employed by previous artificial biosensor systems, where receptors with different sensitivities are combined, and their average response determines the output[307,308]. To achieve this in our system, we first needed to create a panel of A2BR sensor strains with different sensitivities to adenosine. Using the weak *RPL18B*, medium *HHF2*, and strong *CCW12* promoters, we created three strains with a low, medium, and high sensitivity to adenosine, respectively (**Figure 49A**). We could then use

promoter output tuning to normalise the curves experimentally, so they would have an equal output when combined in a mixed population of cells.

However, before we began tuning the output levels of the sensor strains, we noticed the growth rate of the high sensitivity, and to a lesser extent, the mid sensitivity strains was slower than the low sensitivity strain at the higher concentrations of adenosine (**Figure 49B**). This was likely being caused by the higher levels of GFP expression, compounded by higher levels of receptor expression, leading to an increased cellular burden. As flow cytometry data represents the individual fluorescence from cells, without considering the growth rate of the population, it would not be an appropriate comparator between cells which would then be grown together, as the total fluorescence difference between the low and high sensitivity strains would be smaller the flow cytometry data would suggest. To take the fluorescence contribution of the entire population into account, we decided to perform all future measurements with a plate reader and use raw fluorescence to quantify GFP expression from the entire population.

As the low sensitivity strain was expressing sfGFP from a strong synthetic promoter combination (*LexO(6x)-pLEU2m*), we would keep the conditions of this strain fixed while tuning the output of the mid and high sensitivity strains to match it. These cells could then be mixed together to create a population of cells which respond over an extended operational range (**Figure 50A**). First, we remeasured the dose-response curves of the three A2BR sensor strains using a plate reader (**Figure 50B**). Considering the contribution of sfGFP production from the entire population, using raw

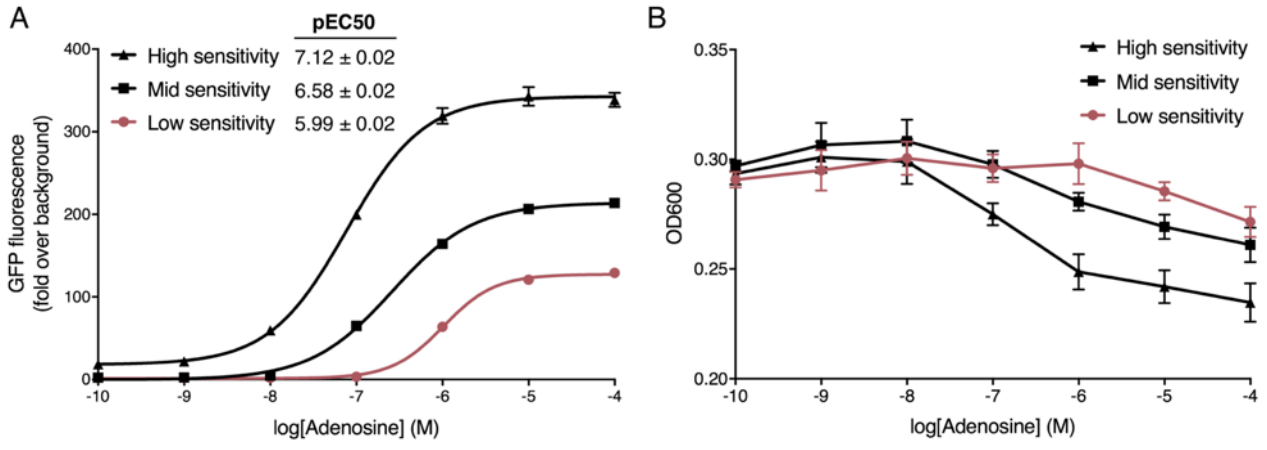


Figure 49. Varying the A2BR sensor sensitivity to adenosine. (*A*) Adenosine dose-response of several adenosine biosensors expressing A2BR using the weak *RPL18B* (low sensitivity), medium *HHF2* (mid sensitivity), and strong *CCW12* (high sensitivity) promoters. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit. (*B*) Adenosine dose-dependent OD measurements of the three different sensitivity A2BR sensor strains after the standard 6 h assay time. All strains set up at the starting OD of 0.175 at time 0 h. Results are means ± standard deviation from triplicate isolates.

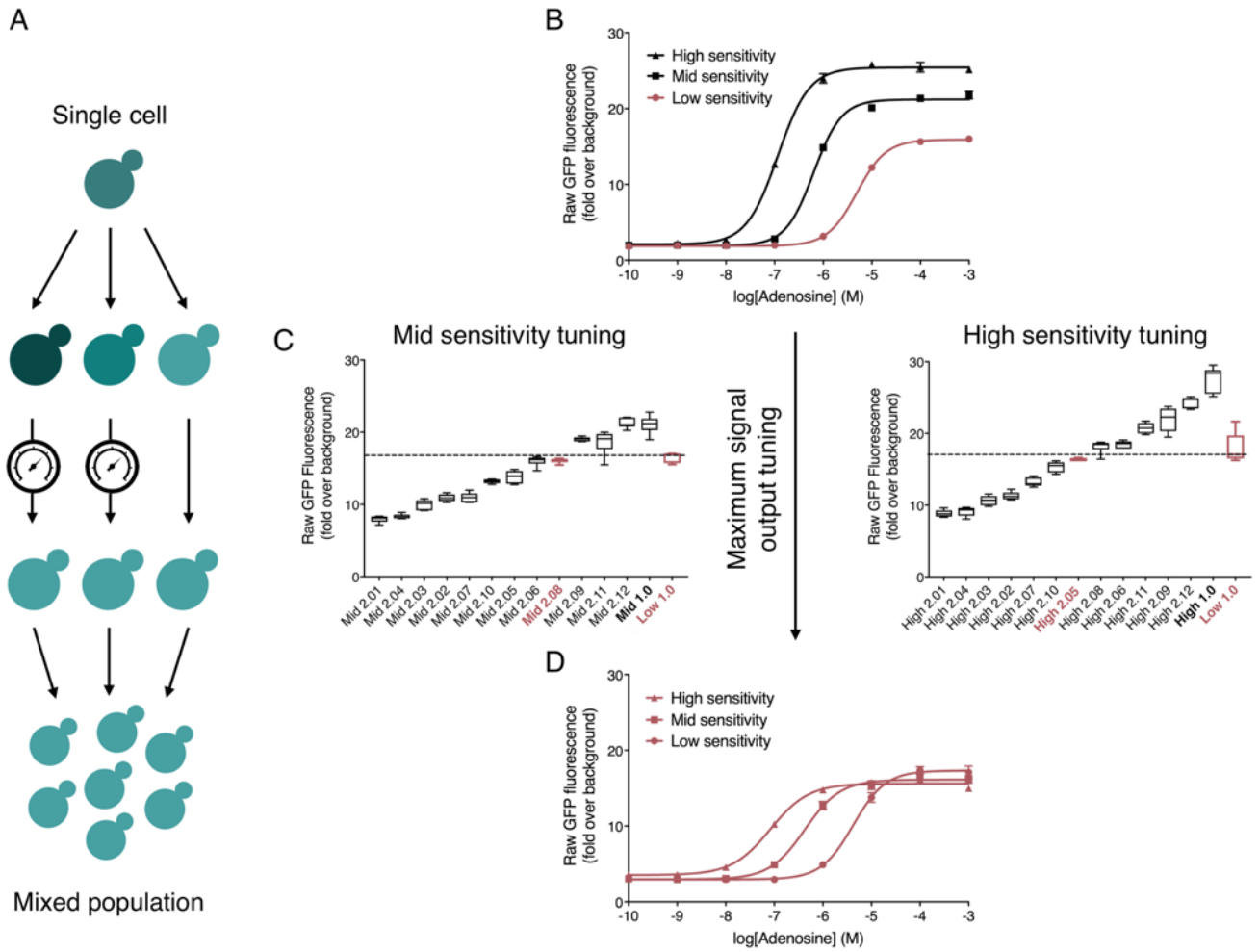


Figure 50. Experimentally normalising a panel of A2BR sensors with different sensitivities. (*A*) Engineered cells combined to produce a linear response from a comparatively digital sensor. First, a range of cells are produced with different sensitivities to a particular ligand by expressing the receptor at varying strengths. Next, the dose-response curves are balanced to produce similar maximum outputs. Finally, the cells are combined in equal parts to create a mixed population of cells with an extended operational range. (*B*) Adenosine dose-response of the Low, Mid, and High sensitivity A2BR sensor strains, utilizing the *RPL18B*, *HHF2*, and *CCW12* promoters to drive the expression of A2BR, respectively. (*C*) Maximum signal output tuning of the Mid and High sensitivity A2BR sensors, to match the maximum signal output of the Low sensitivity A2BR sensor, using different combinations of the LexO UAS and core promoter driving the expression of sfGFP. Chosen promoters for the Mid and High sensitivity A2BR sensors are highlighted in red. (*D*) Adenosine dose-response of the original Low and tuned Mid and High sensitivity A2BR sensors. For details on the LexO UAS and core promoters used here see **Supplementary Table S2**. Unnormalized, raw fluorescence readings were taken using a plate reader to account for growth during the 6h assay. Experimental measurements are sfGFP levels determined using a plate reader and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

fluorescence, demonstrated dose-response curves with a smaller difference between the maximum outputs and a greater separation of sensitivities. Essentially, the low and high sensitivity curves had complete separation of their operational ranges.

Next, we tuned the outputs of the Mid and High sensitivity sensor strains so that they would match the Low sensitivity strain. We did this, by first creating a small library of synthetic output promoters using all combinations of the 1x, 2x, 3x, and 4x LexO UASs with the *RNR2*, *PHO5*, and *LEU2* core

promoters (for a list of promoter combinations see **Supplementary Table S2**). We then used these to vary the maximum output of the Mid and High sensitivity sensor strains and measured their maximum pathway output compared to the original Low sensitivity sensor strain (**Figure 50C**). We then selected strains with a similar maximum output (red) and re-performed the adenosine dose-response experiments (**Figure 50D**). The tuned curves maintained their original sensitivity to adenosine but now shared similar maximum outputs. Finally, we cultured the three tuned sensing strains in a 1:1:1 ratio to create a consortium whose average response integrates the signal from all cells (**Figure 51**). This almost halved the Hill slope of the response while maintaining a similar potency (pEC50), yielding an operational range 50-fold greater than the initial biosensor response.

This simple approach to extending the operational range of the adenosine biosensor is highly-modular and should be easily transferable to other heterologous GPCRs. Furthermore, as the collective response of individual strains determines the dose-response properties of the population, new

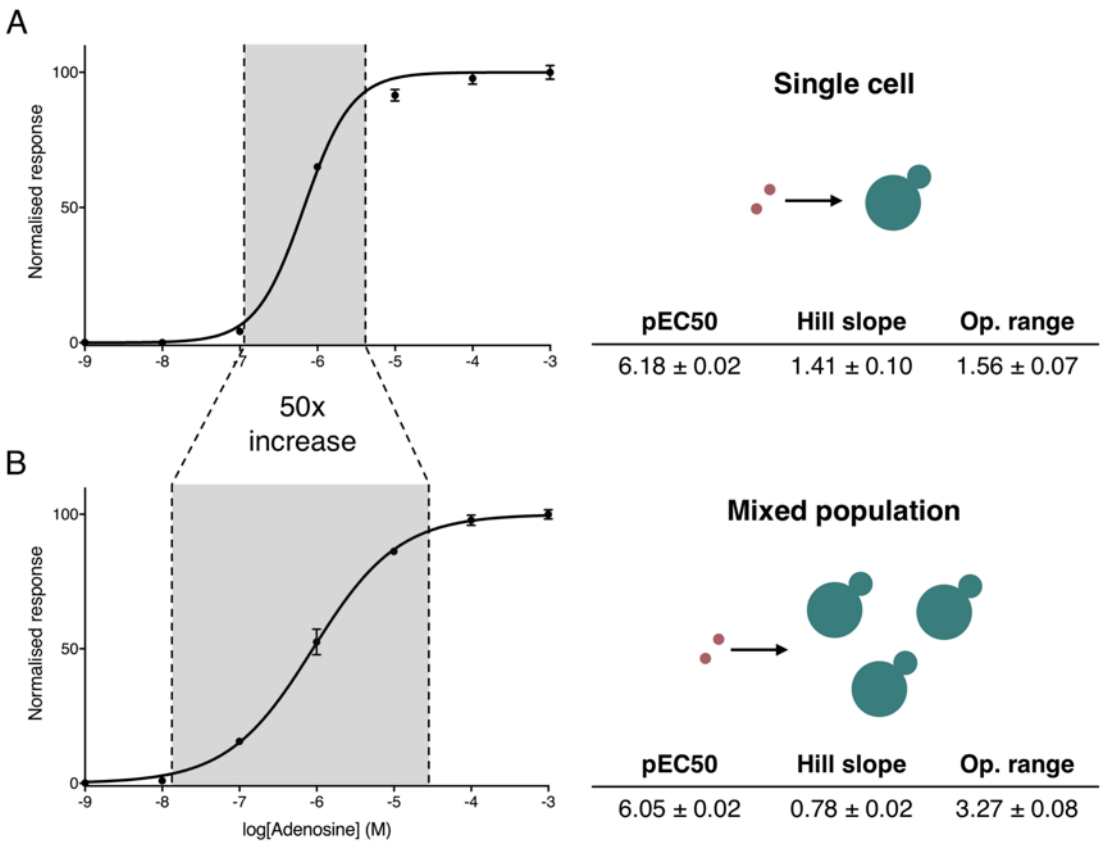


Figure 51. Linearising adenosine sensing. (*A*) The limited dose-response of the human A2BR receptor, operational over 1.6 orders of magnitude. (*B*) The extended dose-response of the engineered population of cells in **Figure 50D** to adenosine, operational over 3.3 orders of magnitude – a 50-fold increase compared to the single cell system, whilst maintaining a similar potency (pEC50). The Low, Mid, and High sensitivity strains in the mixed population were inoculated in a 1:1:1 ratio at the beginning of the assay and the data was normalised from raw fluorescence measurements from a plate reader. Here, the operational range is defined as the concentration difference between 5% and 95% of the activated response. Experimental measurements are sfGFP levels determined using a plate reader and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

operational ranges should be easily accessible by combining strains in different combinations. Using only the points of tuning identified in the previous chapters, this is a good demonstration of the power of these tools when used in combination for generating a desired biosensing for a particular application. We were also able to combine the adenosine dose-response curves of computationally normalised A2BR sensors strains, separated from a mixed population using flow cytometry gating, demonstrating an alternative approach with improved signal-to-noise (**Supplementary Figure S6**).

5.2.4 Digitisation of the MTNR1A response using cell-cell communication

Next, we focused on narrowing the operational range of the melatonin biosensor. This would require more complex engineering, as the Hill slope of a response can only be increased via mechanisms such as cooperativity[234], sequestration[309], or positive feedback[310]. However, these mechanisms are only achievable in interconnected systems, such as synthetic gene circuits. As we intended on using multicellular consortia, we would need a way for the cells to communicate with each other. To achieve engineered cell-cell communication in yeast, previous efforts have repurposed the peptide signalling from the pheromone response pathway[236,311]. As these were components we had stripped out of our yWS677 base strain, this would, therefore, allow us to create a noise-free background in which to reintroduce them for precise cell-cell communication.

With these considerations, we, therefore, designed a two-cell system, where the first cells sensing via MTNR1A would respond to stimulation with melatonin with the production of α-factor. This would then be detected by the second population of cells, which would respond with the production of sfGFP. This would create a chemical “wire” between the cells, which would act as a linkage in a distributed synthetic gene network (**Figure 52A**). To achieve an increase in the Hill slope, firstly, an amplification step would occur in the first cell with the production of α-factor. If a single molecule of melatonin stimulated the pathway past the threshold of activation, this would lead to the transcription of the α-factor prepropeptide (α-factor is processed from a larger precursor polypeptide in the secretory pathway to create multiple copies of the peptide[312]). Hundreds of mature peptides could potentially be translated from this transcript and secreted into the environment, leading to a 10^2 to 10^3-fold amplification. Secondly, we would express the α-factor Bar1 protease in the second cell to create a thresholding effect. The presence of Bar1 would degrade low levels of α-factor, preventing reporter strain activation until levels of α-factor were high enough to saturate the capacity of Bar1-mediated degradation. An amplification step followed by a thresholding effect would, therefore, create an increase in the Hill slope at the population level.

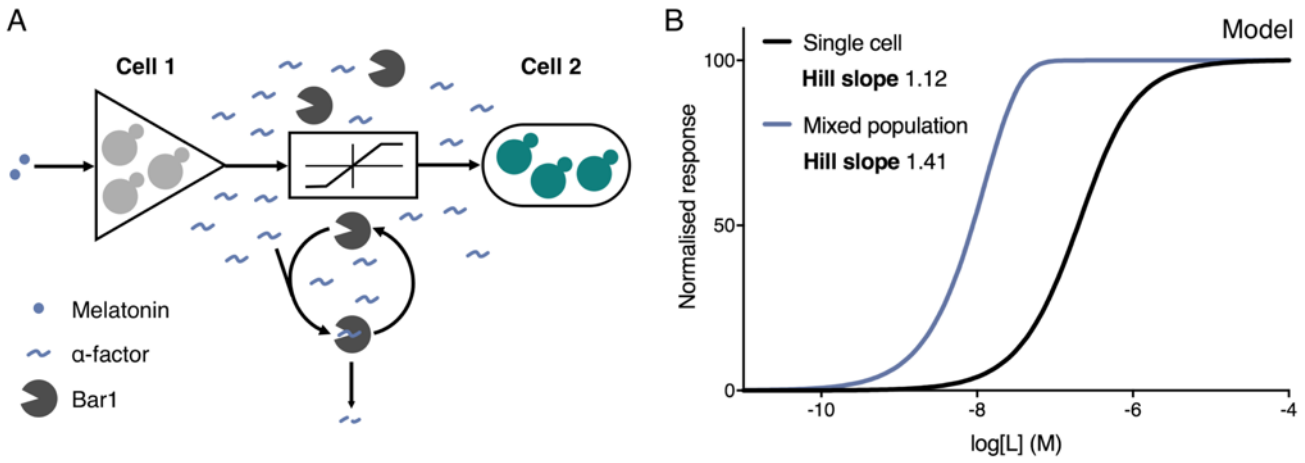
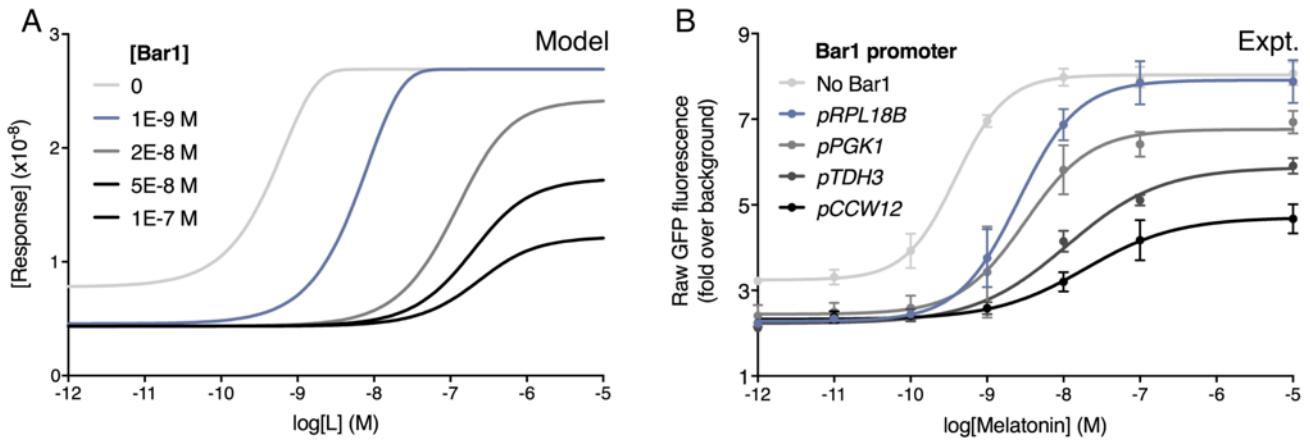


Figure 52. Two-cell amplifier and reporter system to increase the Hill slope of sensor strains. (*A*) A mixed population of amplifier (cell 1) and reporter (cell 2) strains designed to create a digital response from an otherwise linear sensor. In response to ligand, amplifier cells release α-factor which is detected by reporter cells constitutively secreting the α-factor degrading protease, Bar1. The presence of Bar1 degrades low levels of α-factor preventing reporter strain activation until levels of α-factor are high enough to saturate the capacity of Bar1-mediated degradation, creating an amplification followed by a thresholding effect. (*B*) Model of the two-cell amplifier-reporter system with Bar1-mediated threshold response compared to the single-cell response. The combination of amplification and thresholding is able to create a system with a Hill slope two times greater than the single cell sensor. All modelling developed and performed by Hitoshi Yamauchi, Jack Mead, and Graham Ladds[264].

To investigate whether or not this approach would work, we modelled the cell-cell system *in silico* (**Figure 52B**). This revealed an increase in both the sensitivity of the system, as a result of the amplification, and an increase in the Hill slope, as a combined result of the amplification and thresholding. As we were unsure what the effects of Bar1 expression would be, we simulated the two-cell model over different initial concentrations of Bar1 (**Figure 53A**). As the conversion of the input in the first cell into α-factor would amplify leaky expression, a high level of basal activity was predicted without the presence of Bar1. However, this leaky system was predicted to have a high Hill slope. The expression of Bar1 would reduce the leakiness, at the expense of losing sensitivity, and if expressed too highly, the loss of maximum signal and reduction in the Hill slope. The high Hill slope, therefore, looked like a product of the amplification step, whereas Bar1 controlled the leak.

The two-cell model predicted the precise expression of Bar1 was essential for creating a tight and sensitive system with a high Hill slope and maximum signal output. To identify the level of Bar1 expression that would produce this experimentally, we developed the two-cell system. The first cell expressed MTNR1A at medium levels using the *HHF2* promoter and produced α-factor in response to melatonin sensing. The second cell expressed the Ste2 receptor at high levels using the *CCW12* promoter and produced sfGFP as an output. We then varied the expression of Bar1 in the second cell using four promoters of varying strengths and a no Bar1 control. We cultured the two strains in a 1:1 ratio and measured the melatonin dose-response of the mixed population (**Figure 53B**). The



C

	[Bar1] / promoter	1e-7 / *pCCW12*	5e-8 / *pTDH3*	2e-8 / *pPGK1*	1e-9 / *pRPL18B*	0 / No Bar1
Hill slope	Model	1.10	1.11	1.14	1.44	1.49
	Expt.	0.74 ± 0.14	0.71 ± 0.08	1.02 ± 0.21	1.10 ± 0.14	1.32 ± 0.13

Figure 53. Tuning the expression of Bar1 in the two-cell amplifier-reporter system. (*A*) Varying the concentration of Bar1 in the two-cell amplifier-reporter model. (*B*) Experimentally varying the expression of Bar1 in the two-cell amplifier-reporter system using a select promoter library. (*C*) Hill slope values from computationally and experimentally varying the levels of Bar1 in the amplifier-reporter system. All modelling developed and performed by Hitoshi Yamauchi, Jack Mead, and Graham Ladds[264]. Experimental measurements are sfGFP levels determined using a plate reader and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

experimental results confirmed the predictions of the model, revealing the *RPL18B* promoter as the ideal choice for driving the expression of Bar1, as it created minimal system leak and maintained a high Hill slope and maximum signal output, without losing too much sensitivity.

We had now established the optimal conditions for the two-cell amplifier-reporter system. Next, we set out to digitise the melatonin biosensor, while maintaining a similar potency as the original one-cell system. We did this in two steps. First, we digitised the MTNR1A response using the two-cell amplifier-reporter system using the optimised conditions identified from the Bar1 tuning experiment (**Figure 54B**). This produced a more sensitive response with a higher Hill slope compared to the initial biosensor. Secondly, we tuned the expression of the MTNR1A receptor expression in the first cell, using the weaker *ALD6* promotor to match the potency of the initial biosensor (**Figure 54C**). The final digitised melatonin biosensor produced a 2.3-fold increase in the Hill slope, representing a 200-fold decrease in the operational range of the initial sensor strain while maintaining a similar potency (**Figure 55**).

We have, therefore, successfully developed a new strategy for reducing the operational range of GPCR-based biosensors, creating a near digital response from a previously linear curve. While this method was more complicated than the approach used to linearise a digital-like biosensing response,

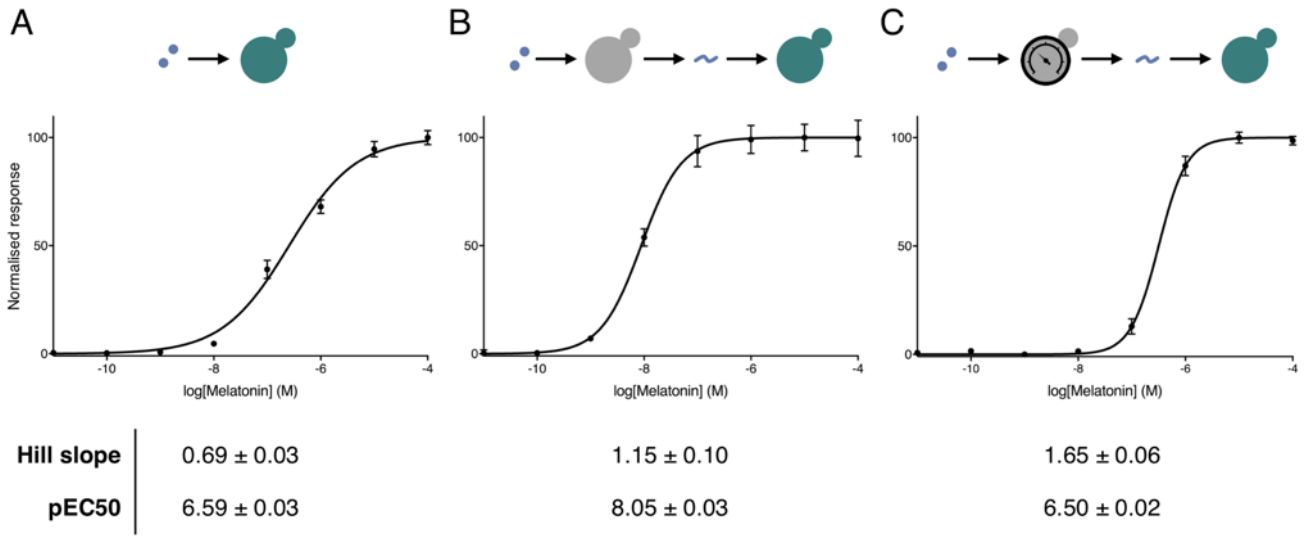


Figure 54. Digitising and fine tuning the MTNR1A sensor response. (*A*) Melatonin dose-response of the MTNR1A sensor strain in a monogenic population of cells. (*B*) Digitisation of the MTNR1A response via α-factor mediated cell-cell communication. In response to melatonin, the first cell produces large quantities of α-factor peptide that is then detected by the second cell, which responds by producing sfGFP. (*C*) Fine tuning the MTRN1A digital response by reducing the receptor expression in the first cell, so that the pEC50 matches the response of the single cell system. By lowering the expression of the MTNR1A receptor in the first cell using the *ALD6* promoter, we were able to shift the potency (pEC50) of the melatonin dose-response right by 1.5 orders of magnitude, to match the potency of the first, single cell system, whilst maintaining a high Hill slope. Data normalised to the minimum and maximum values within each dataset. Experimental measurements are sfGFP levels determined using a plate reader and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

it is still reliant on the rules developed during the previous chapters, while reintroducing a couple of components from the mating response we had previously stripped out. As with the linearisation strategy, the digitisation approach tunes properties of the system without requiring protein engineering of any of the components, thus ensuring the platform will work for any GPCR we are able to port into yeast. Furthermore, as digitisation is performed using two strains, engineering can be distributed into more manageable units. Indeed, the second cell should not require further engineering and thus be universal, as changes to the receptor will occur only in the first cell.

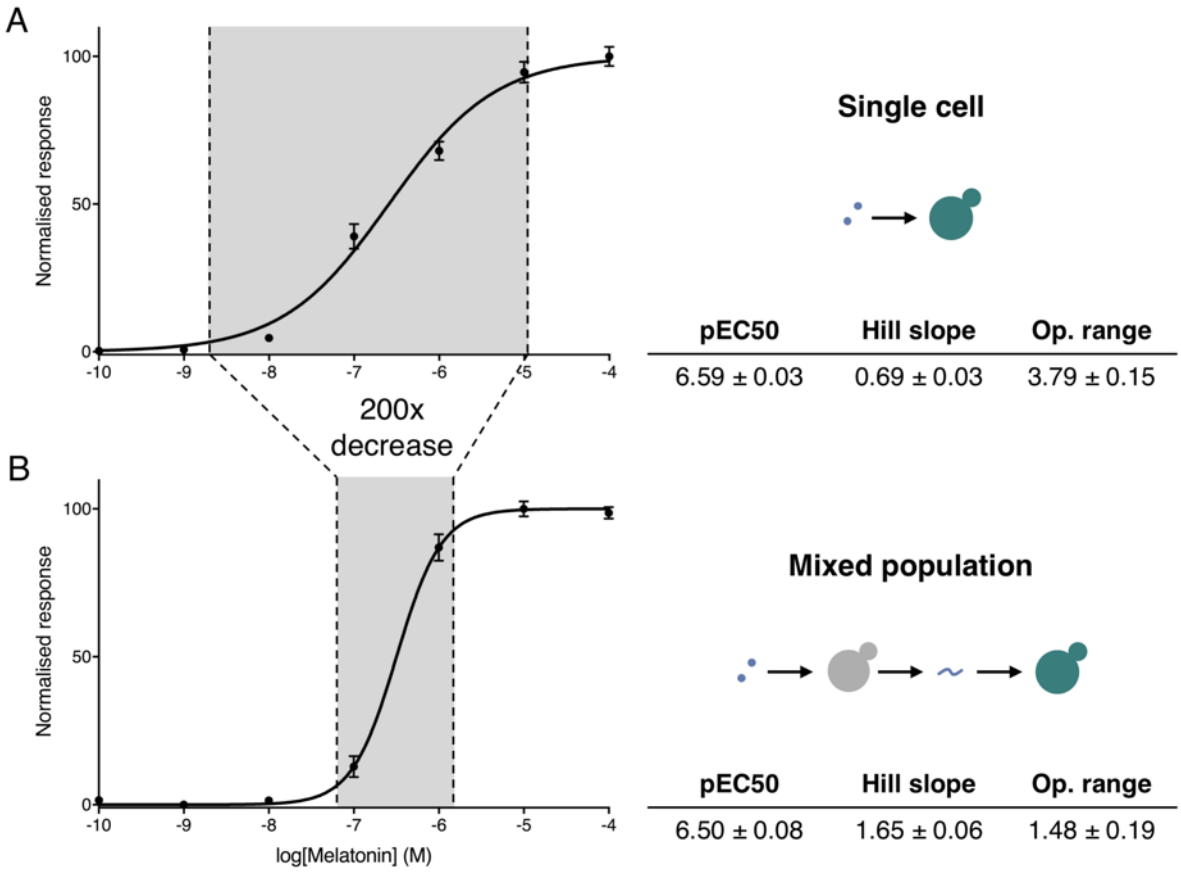


Figure 55. Digitising melatonin sensing. (*A*) The broad dose-response of the human MTNR1A receptor to melatonin in a single cell line, operational over 3.8 orders of magnitude. (*B*) Digitisation of melatonin sensing in the mixed population, reducing the operational range of the response to 1.5 orders of magnitude, 200-fold lower than a single population. Values for the potency (pEC50), Hill slope (nH) and operational range (O.R.) for the single and mixed population A2BR sensor strains are included. Here, the operational range is defined as the concentration difference between 5% and 95% of the activated response. Experimental measurements are sfGFP levels determined using a plate reader and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit.

5.3 Conclusions

In this chapter, we explored strategies for tuning the operational range of GPCR-based biosensors. To provide a starting point for this work, we domesticated two human receptors with naturally contrasting Hills slopes: the comparatively digital adenosine-sensitive A2BR and linear melatonin-sensitive MTNR1A receptors. We then overlaid feedback loops on the minimised pathway of these sensor strains, using negative and positive modulators to linearise and digitise the A2BR and MTNR1A sensors, respectively. As intracellular regulation had an insubstantial effect on the Hill slope in both of these systems, we instead looked into community-based approaches as an alternative engineering strategy. First, we linearised the A2BR sensor by creating a panel of cells with different sensitivities which were then combined to respond over an extended operational range. Finally, we created a two-cell amplifier-receiver system to convert the linear MTNR1A curve into a near digital response. Both of these approaches required no change to the protein identity of the components, representing a generalisable strategy for tuning the Hill slope of GPCR-based biosensors, thus completing the toolkit for tuning all aspects of the dose-response curve.

6 Discussion

6.1 Summary

In this thesis, we have used genome engineering and synthetic biology tools to refactor a minimised yeast signalling pathway so that it could be used for rationally tuning GPCR-based biosensors. We uncovered simple mechanisms for tuning the sensitivity, leakiness, and signal amplitude of the pathway dose-response curve, which we then combined in community-based approaches for tuning the final biosensing property - the Hill slope. In this chapter, we will discuss this new platform, including how it compares to previous tools, the limitations of the yeast platform, and further applications for yeast GPCR-based biosensors, using a few additional pieces of work to illustrate our points. We then discuss the concept of refactoring signal transduction pathways, what we have learned from the minimised pathway, what has been left unanswered, and our expectations of this approach for synthetic biology in the future.

6.2 A platform for rationally tuning yeast GPCR-based biosensors

6.2.1 A new paradigm in yeast GPCR research and application

Since the first report of a heterologous GPCR being successfully coupled to the pheromone response pathway almost 30 years ago, we have had the necessary tools for creating yeast GPCR-based biosensors. In this landmark study, King *et al.*[208] demonstrated we could enable the comparatively simple eukaryotic organism, *Saccharomyces cerevisiae*, with the adrenaline sensing capabilities of higher eukaryotes, opening the door for fundamental GPCR research and eventually manifesting in the development of purpose-built biosensors for detecting diverse stimuli. However, since this initial study, the sophistication of these yeast GPCR-based systems has not changed that much.

As discussed in section **1.2.11**, adaptations have been made over the years for improving receptor coupling, increasing the signal, and providing an easy-to-measure output[180]. While these adaptations have provided incremental improvements, yeast GPCR-based biosensors are still created as they were back then. Heterologous receptors are expressed in standard yeast vectors, in combination with various yeast strains, which are engineered to different extents and contain pre-installed components, such as reporters and Gα chimeras[55,166,228,229]. This out-dated paradigm is limiting the wide-spread application of GPCR-based biosensors, as new designs are constrained by the immutable dose-response characteristics they are naturally bestowed.

In this work, we set out to completely transform the state of the yeast GPCR landscape, rather than provide an incremental improvement. In our new platform, biosensor designs are created by assembling modular parts into a single multigene plasmid under the control of synthetic regulation, which is then integrated into the genome of a single highly-engineered base strain. Transitioning to a modular cloning toolkit allows for the precise control over the receptor, Gα, transcription factor, and reporter while reducing the number of strains required to generate any desired biosensor to one. By providing researchers with a standardised and modular platform for creating GPCR-based biosensors, new parts and designs can be shared, hopefully accelerating the collective advancement of the field. Furthermore, the abstraction of these collective tools into a single platform reduces the complexity of creating GPCR-based biosensors, making it more accessible to labs without in-depth knowledge of yeast or the mating pathway.

Using this new platform, we have developed a general set of rules for rationally tuning all dose-response properties, which to our knowledge is the first instance in any GPCR-based system. Importantly, these points of tuning do not require protein engineering of the receptor. This modularity will allow for the tuning of any heterologous receptor which can be ported into the system, facilitating

the creation of GPCR-based biosensors tailored for their specific application requirements and overcoming their current limitations in yeast. As a brief demonstration of this platform for creating biosensors with an industrial application, we applied the MTNR1A sensor to the detection of microbially produced melatonin.

In 2016, Germann *et al.* reported the glucose-based microbial production of melatonin in *Saccharomyces cerevisiae*[313]. This highly-engineered biosynthetic pathway converts *L*-tryptophan into melatonin through three non-native intermediates (**Figure 56A**). Currently, the only means to measure the production of this metabolite is by LC-MS, limiting the authors to a small number of

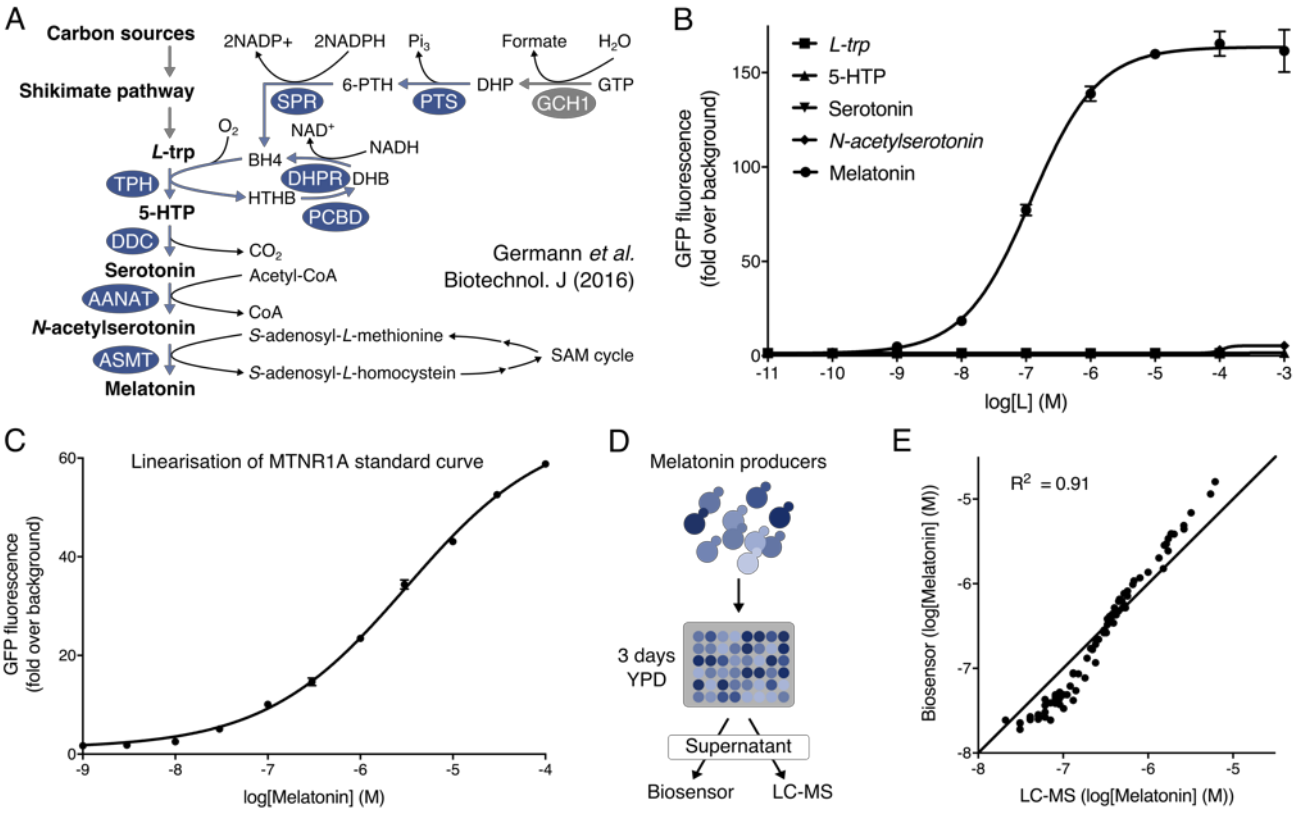


Figure 56. A medium-throughput screen for microbially produced melatonin. (*A*) Overview of the melatonin biosynthesis pathway in yeast from Germann *et al.*[313]. Melatonin is produced from *L*-tryptophan via four enzymatic steps and with the BH4 biosynthesis and regeneration pathways to supply BH4 cofactor. Blue ellipses and arrows indicate the heterologous enzymes and biosynthetic steps required to generate melatonin from *L*-tryptophan. Grey ellipses and arrows represent enzymes and biosynthetic steps endogenous to yeast. (*B*) Dose-response curves of the MTNR1A sensor to melatonin and the intermediates in the melatonin biosynthesis pathway, demonstrating exquisite specificity for melatonin. (*C*) A linearised MTNR1A sensor population consisting of two cells with different sensitives to extend the operational range of melatonin biosensing. Experimental measurements are sfGFP levels per cell determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates. Curves were fitted using GraphPad Prism variable slope (four parameter) nonlinear regression fit. (*D*) Measuring the production of melatonin from the spent media of 88 random yeast producer strains using the extended MTNR1A sensor and LC-MS. (*E*) The measured production of melatonin from the 88 producer strains as determined from single measurements using the extended MTNR1A sensor and LC-MS. A linear y=x curve was fitted to the dataset. Melatonin producer strains were kindly provided by Susanne Germann, Simo Baallal Jacobsen, and Jochen Foster (Technical University of Denmark).

pathway designs that were optimised by rational engineering[313]. Although they achieved a reasonable maximum yield of 14.5 mg/L, if a higher throughput measurement was available to them at the time it is possible they could have used a more top-down approach to scan a wider genetic landscape, perhaps identifying a more optimal biological solution to this complex biosynthetic pathway.

To determine whether or not our MTNR1A sensor strain was suitable for reporting the yields of melatonin, we first measured the response to all the intermediates and the end product of the pathway (**Figure 56B**). The MTNR1A sensor demonstrated exquisite specificity for melatonin, with only a slight activation at very high levels of the immediate precursor, *N*-acetylserotonin. As the dose-response of the MTNR1A sensor was flat at the concentration of the highest reported yield of melatonin (6.2e-5 M), we linearised the curve using the method outlined in section **5.2.3** (**Figure 56C**). The resulting curve was linear over four orders of magnitude, but more importantly was linear over a range that was appropriate to the reported yields of melatonin production, providing a means of quantification. We then prepared the supernatant from 88 random melatonin producer strains, kindly given to us by Germann *et al.*, and measured their production using our extended biosensor and LC-MS (**Figure 56D**). The calculated yields of melatonin, as determined by the melatonin biosensor, agreed strongly with the LC-MS results, with the biosensor arguably more accurate over the large span of concentrations, due to the limitations of the linear range in LC-MS[314] (**Figure 56E**).

This demonstration highlights several key features of this platform and GPCRs in general. Firstly, the ability to tune the operational range allowed us to create a biosensing population which we could use to directly measure the supernatant from the producer strains without any sample preparation. All 88 samples, spanning three orders of magnitude difference in concentration, fell into the linear range of the biosensor, allowing us to quantify melatonin production accurately. Secondly, this method is economical, untethered from specialist equipment, and operates at a medium throughput. The use of fluorescence detection using flow cytometry allowed us to measure all 88 samples in less than an hour, rather than the entire day for LC-MS (plus repeated sample measurements due to the broad span of concentrations). If we switched over to a plate reader, this would allow for the measurement of 96 samples every five minutes. Medium throughput screening of small libraries would now be possible for optimising this pathway. Finally, the biosensor was able to distinguish between a single methyl group on the melatonin end product and the *N*-acetylserotonin intermediate, demonstrating the high level of specificity GPCRs can achieve.

6.2.2 Limitations of the yeast platform

Although we have addressed a number of the key limitations of yeast GPCR-based biosensors in this work, issues remain. Following the completion of our GPCR platform, we attempted to domesticate a number of additional receptors for sensing ligands relevant to metabolic engineering. First, we attempted to port the human serotonin-sensitive HTR4 receptor, as reported by Ehrenworth *et al*.[55]. Similar to their experience, the HTR4 sensor strains exhibited poor activation and sensitivity to serotonin (**Figure 57A**). This turned out to be a problem with receptor expression. As suggested by the authors, reducing the pre-incubation temperature to 15 °C improved the expression of the receptor, leading to a greater dynamic range, although a significant fraction of the cells still did not produce a measurable output. However, the dynamic range was over 2x better than the published biosensor with minimal optimisation, demonstrating the benefits of our system over previous tools.

Next, we tried to port a small library of mammalian olfactory receptors responsive to ligands that can be produced microbially, such as vanillin[315] and raspberry ketone[316], as identified by OlfactionDB (a database of olfactory receptors and their ligands)[317]. We tried coupling these receptors with the WT Gpa1 and a Gpa1-Gα(olf) chimera at 30 °C and 15 °C pre-assay incubation temperatures. None of these conditions produced a biosensor able to detect any of the ligands above the control (**Figure 57B**). It is clear that while yeast may have the complex machinery required for GPCR signalling, it does not have the capacity for functionally expressing all heterologous receptors.

To make yeast a more flexible chassis for GPCR-based biosensing, we next need to address this issue with the expression of heterologous receptors. As discussed in section **1.2.11**, several approaches to resolving this problem exist, ranging from engineering receptors on the individual basis, through more broadly-applicable strategies, such as the co-expression of accessory proteins, to universal strategies, such as humanising the cell membrane (**Figure 57C**). Alternatively, we could take our existing GPCR library and evolve these receptors towards new ligands. By setting the threshold of agonist detection, we could control the directed evolution of receptors to ensure sensitivity and specificity. As nature has demonstrated through the vast diversity of GPCR ligand specificities, this could prove to be a powerful way to create biosensors for any application.

Whatever strategy chosen to address this problem, our platform should provide the ideal means for carrying it out: (i) receptor variant libraries can be introduced into the system at higher transformation efficiencies using CRISPR-aided integration, (ii) accessory proteins can be combinatorially co-transformed using the additional integration modules, and (iii) improvements to the cell biology can be performed on a single strain, which would then succeed the yWS677 base strain.

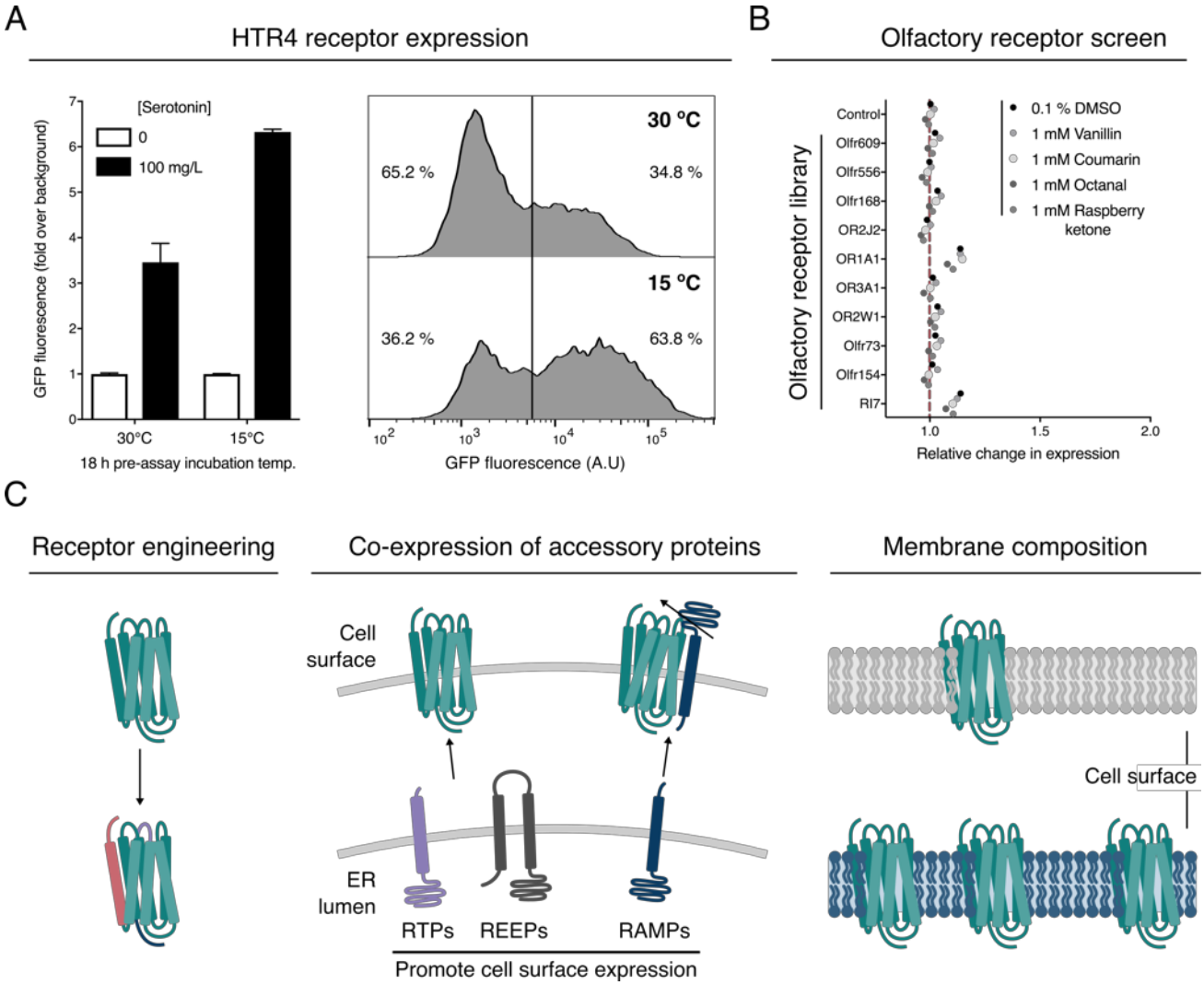


Figure 57. Limitations of heterologous GPCRs in yeast. (*A*) The low activation of the serotonin responsive HTR4 receptor in yeast, demonstrating the mean GFP fluorescence (left) and fluorescence histogram (right) of yeast cells after activation with saturating serotonin levels when incubated for 18 h at the standard (30 °C) and reduced (15 °C) pre-assay temperatures. In both of these conditions a large number of cells are unresponsive, resulting in a low dynamic range at the population level. Results are means ± standard deviation from triplicate isolates. (*B*) Relative activation of a panel of olfactory receptors known to be responsive to one or more of the odorous compounds, vanillin, coumarin, octanal, and raspberry ketone, added at saturating concentrations. This experiment was performed with Gpa1 and a Gpa1-Gα(olf) chimera at both 30°C and 15°C pre-assay incubation temperatures. None of the screened receptors produced a response greater than the DMSO control in any of the conditions. Results are means from triplicate isolates. (*C*) Possible strategies to improve the expression and functional coupling of heterologous GPCRs in the yWS677 base strain, ranging from engineering receptors on the individual basis to extensive strain engineering to humanise the cell membrane. Experimental measurements are sfGFP levels per cell determined by flow cytometry.

6.2.3 Further applications of yeast GPCR-based biosensing

While we have mainly focused on the application of yeast GPCRs as a tool for measuring a ligand of interest, the yeast pheromone response pathway is also a versatile system for cell-cell communication. As demonstrated in section **5.2.4**, pheromone communication can be used to create

new system properties at the population level. Effectively, this is the yeast equivalent of the popular bacterial quorum sensing systems which are advancing prokaryotic synthetic biology in many new areas[318]. Already, pheromone systems from other yeast species have been shown to function well in *S. cerevisiae* and have been used to program diverse behaviours[235,236,311,319,320]. As we expand our toolkit of peptide/receptor pairings in yeast, this will become a powerful communication language for coordinating community behaviour.

The yeast pheromone communication system offers some unique qualities: (i) as peptides are a direct translational output, the response is incredibly fast (**Supplementary Figure S7**), can be controlled at the level of transcription, and does not require a precursor molecule, (ii) pheromone receptors are highly-specific for their cognate peptides[229], facilitating the simultaneous use of multiple receptor/peptide pairs within a complex consortia, and (iii) many pheromones have a specific protease which can be used to inactivate the peptide in the extracellular matrix, shifting precise regulation of these distributed synthetic gene networks outside of the cell[319]. However, one limitation of this system is that receptors share a common pathway. Individual cells would therefore not be able to distinguish between two inputs if expressing two receptors. We would need to find an intelligent solution this problem, perhaps similar to how olfactory neurons differentiate the vast diversity of volatile compounds by multiplexing receptors that also share a common pathway[138].

The GPCR platform we have developed here will undoubtedly aid in the progress of yeast community-based applications. As yeast GPCR receptors naturally exhibit different potencies for their cognate peptide[229], being able to tune the input-output of these systems would balance such discrepancies while setting digital thresholds of activation will improve the flow of information[234]. The tools we have developed to redirect the response to synthetic promoters will also allow the extracellular cell-cell communication to be coupled with intracellular synthetic gene networks. Seamless integration between these two modes of regulation would enable synthetic biologists to distribute gene networks between multiple cells, thus expanding the capabilities of current designs[321]. Finally, our GPCR base strain has all the pheromone communication components stripped out, providing a clean chassis for building these systems without the interference of endogenous cell-cell signalling.

To demonstrate how these systems can be developed and layered using our GPCR platform, we created a distributed biological circuit with half adder logic (**Figure 58**). In this half adder design, the complex genetic circuitry is distributed between 5 different cells, which themselves comprise of nothing more complicated than a single AND NOT gate. In contrast to previous designs of this biological circuit, for example, Regot *et al.*[235], the inputs are all within the first layer of cells, which are chemically ‘wired’ to the next layer using peptide communication. Regulation of these chemical wires

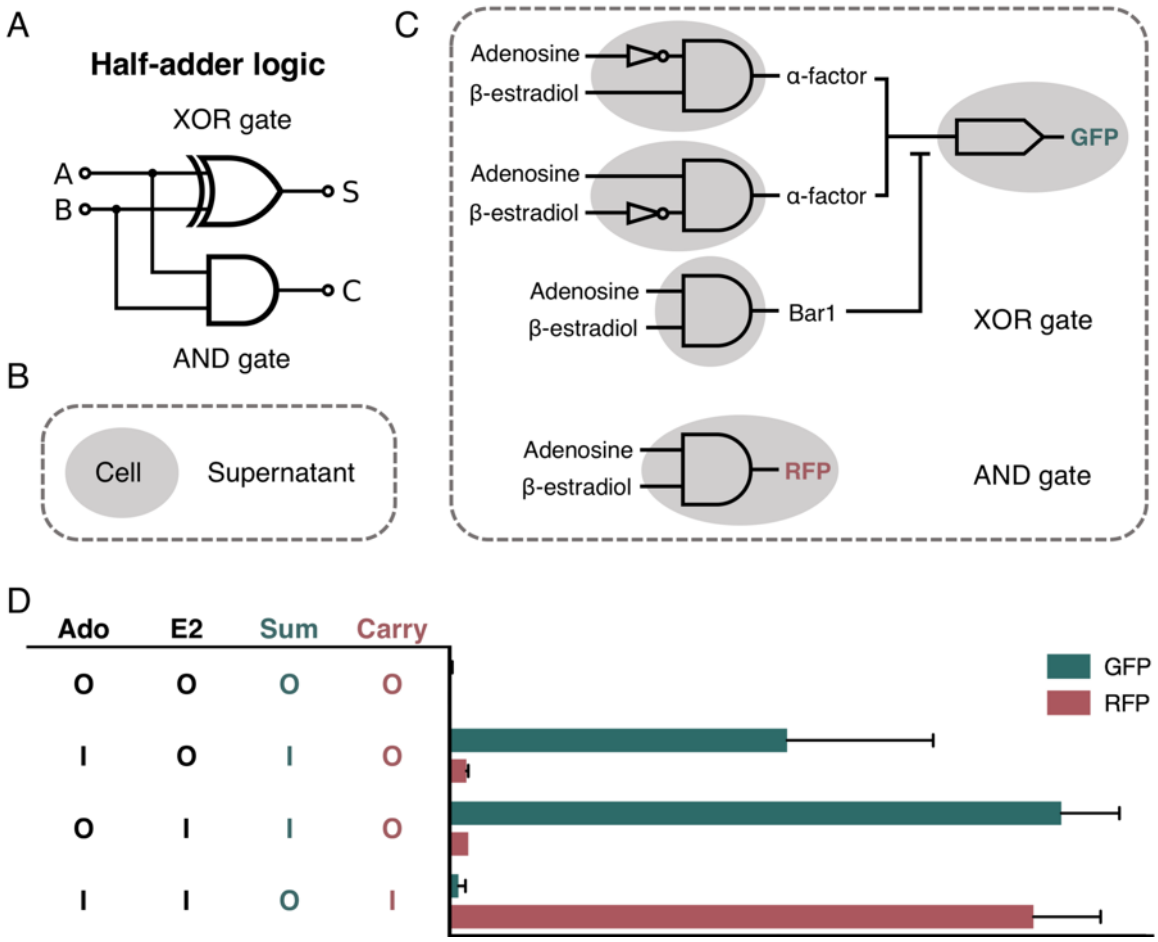


Figure 58. Design an *in vivo* implementation of a distributed biological circuit with half adder logic. (*A*) Half adder logic diagram. The half adder adds two single binary inputs, A and B and produces two outputs, sum (S) and carry (C). In the simplest form, a half adder can be constructed from a XOR gate for the sum and an AND gate for the carry. (*B*) Representation of yeast cells and the extracellular environment in (*C*). (*C*) Schematic representation of the community-based half adder. Four cells that together create an XOR gate, responding to the presence of either adenosine or β-estradiol, but not both, to produce a GFP output (sum). A fifth cell acts an AND gate, responding only in the presence of both adenosine and β-estradiol to produce an RFP output (carry). (*D*) Truth table and measured responses from the community-based half adder. See **Supplementary Figure S8** and **Supplementary Figure S9** for the design of the AND and AND NOT gates used to create the half adder. Experimental measurements are the percentage of the population of sfGFP and mRuby2 positive cells determined by flow cytometry and shown as the mean ± standard deviation from triplicate isolates.

can then be achieved with the secretion of the corresponding peptide protease. As each cell represents a single logic gate, it is possible to imagine a library of cells with different inputs/outputs and logic. Cells could then be mixed in different combinations to rapidly prototype new circuit designs. Such modularity would not be possible using synthetic gene networks within single cells.

6.3 Refactoring the mating pathway

6.3.1 The highly engineered GPCR base strain

In section **1.3**, we outlined our grand vision for refactoring a minimal pheromone response pathway to gain complete control of the dose-response properties, representing the first time this bottom-up approach would be applied to an endogenous signal transduction pathway. To make this a more achievable goal within the scope of this thesis, we proposed an intermediate design on the road towards full pathway refactoring that would also be useful for biosensing applications. This minimised pathway consisted of 14 gene disruptions to remove feedback regulation, insulate the pathway, and allow the reintroduction of the receptor, Gα, and TF. To achieve this, we developed a new CRIPSR-based strategy that allowed us to rapidly iterate the large number of modifications to the yeast genome, resulting in the creation of a single highly-engineered GPCR base strain.

During the development of the CRISPR methodology, we established a well-defined and robust protocol for future engineering projects and explored genetic features that would aid the downstream engineering of the final base strain, in the form of landing pads. The resulting strategy comprised the pairwise deletion of chosen genes by precisely substituting the ORFs for a new CRISPR targeting sequence. By implementing a novel marker cycling strategy, we were able to achieve 16 genome modifications in only five weeks; an achievement that would have taken more than twice as long with pre-existing approaches. The final base strain was easy to reverse engineer and was more compatible with the YTK system. Although engineering as extensive as this will not be appropriate for all projects, we believe this work is a good template when designing strains that will have many downstream applications, such as ours.

6.3.2 An *In vivo* model of cell signalling

Once the GPCR base strain was complete, we turned towards the refactoring of the receptor, Gα, and TF. By individually varying the expression of these key components while keeping all others fixed, we were able to treat the cell in a manner highly analogous to an *in silico* model. We developed the experimental and computational models in parallel, with each approach validating the findings of the other. Using this methodology, we were able to uncover fundamental properties of the pheromone response pathway. First, we revealed the sensitivity of the pheromone response pathway is achieved through positive transcriptional feedback of the receptor, Gα, and transcription factor. Although it is well known these components are upregulated in response to pheromone[179], this is the first time the

collective behaviour of these components has been demonstrated. We then go on to show how fine-tuning of the pathway can be used to recreate the sensitivity of the wild type system without using feedback.

Our final refactored design achieved significantly improved biosensing characteristics with a fraction of the wild type components. While this may seem like a strange evolutionary adaptation to include more components than necessary to accomplish a similar response, it is worth noting that our system is probably far more expensive to the cell. In the native system, the components seem to be expressed at low levels in the absence of a stimulus. When the pathway is then activated, these become significantly upregulated to achieve the desired response. This means that while there may be more components in wild type cells, their collective expression is likely lower than in our system while the mating pathway is latent (i.e. during normal growth). Moreover, our cells have lost the ability to desensitise to the agonist. While this is what we would like for an end-point biosensor, continual activation to low levels of pheromone peptide in the environment or a failure to recover from an unsuccessful mating event would be disastrous for the cell. Selectively reintroducing feedback into the minimised pathway and exposing cells to different perturbations could prove to be a useful way to study how the native response is robust to changes in environmental conditions.

Through the parallel development of the *in silico* and *in vivo* models, we have arguably created the most accurate computational model of receptor/G protein signalling in the mating pathway to date. Although the model extends only as far as the Gβγ, the agreement between the computational and experimental observations is remarkable, demonstrating the power of the refactoring approach to understanding complex biological systems. Using this model, we were able to predict the system behaviour when introducing a constitutively active receptor, allowing us to design a new assay for detecting the basal activity of Ste2. Being able to reliably predict a new system behaviour before it has been observed experimentally is a good demonstration of the model accuracy.

Varying the expression of the receptor, Gα, and TF, both *in silico* and *in vivo*, revealed new insights into the importance of protein stoichiometries in GPCR signalling pathways. Most notably, the promoter identity of signalling components can significantly affect the dose-response in predictable ways that held true for the four receptors we tested, indicating we have uncovered generic principles for tuning the dose-response of the system. This also demonstrates each GPCR-mediated response is not defined by the intrinsic properties of the receptor, such as ligand binding affinity. Instead, it seems signal transduction is a function of the properties of all components in the signalling pathway and particularly their relative levels.

These findings have important consequences for GPCR signalling more broadly in biology. It explains how changes to the levels of components can have a significant effect on signal transduction, whether due to environmental perturbation or altered expression and protein turnover in different tissues. Cells can, therefore, have different sensitivities and activation thresholds from the same agonist using identical receptors. This highlights the importance of examining non-coding variation alongside protein polymorphisms when assessing how genetic variation links to health and the efficacy of treatments for disease-relevant pathways[322].

Receptor variation has recently been recognised as a major cause of GPCR-targeted drugs being ineffective in individuals[323]. It is possible non-coding mutations leading to differences in protein stoichiometries could help to explain this further. The importance of protein stoichiometries may also explain past failures in GPCR-targeted drug discover projects which used dramatically overexpressed receptor numbers, as these models would have borne little relevance to the endogenous systems they aimed to represent. Indeed, overexpression of a receptor without considering the Gα levels will typically lead to a leaky system which may not reflect the natural biological state[219].

6.3.3 Refactoring the minimal mating pathway

To achieve the full refactoring of the minimal pathway, a large amount of work remains. First, we would need to achieve a further 12 gene knockouts to completely strip the current base strain of the remaining mating pathway components. With the tools we developed early in this project, this should be trivial. However, we would then be required to reintroduce a total of 13 pathway components, rather than the three we achieved in his work (**Figure 59A**). Although we now have a robust methodology for refactoring the signalling components, this would still be a monumental amount of work. To break this down into more manageable sized pieces, we propose deleting and refactoring the signalling components in three defined stages (**Figure 59B**).

The first unit we would tackle would be the receptor/G protein signalling, comprising the receptor, heterotrimeric G protein complex, and Ste20 (*URA3* module). As we already have a working model for these components (excluding Ste20), this should be relatively straightforward. The second unit we would then address would be the transcriptional complex, composed of the TF, the Dig1 and Dig2 transcriptional repressors, and the reporter (*HO* module). As a model of this complex already exists, we could adapt this to fit our needs[250]. Finally, we would refactor the MAPK signalling cascade (*LEU2* module). This section may be challenging as this part of the pathway is often blackboxed, and the signalling assumed to be linear[248]. However, it is well documented that the interactions with the Ste5

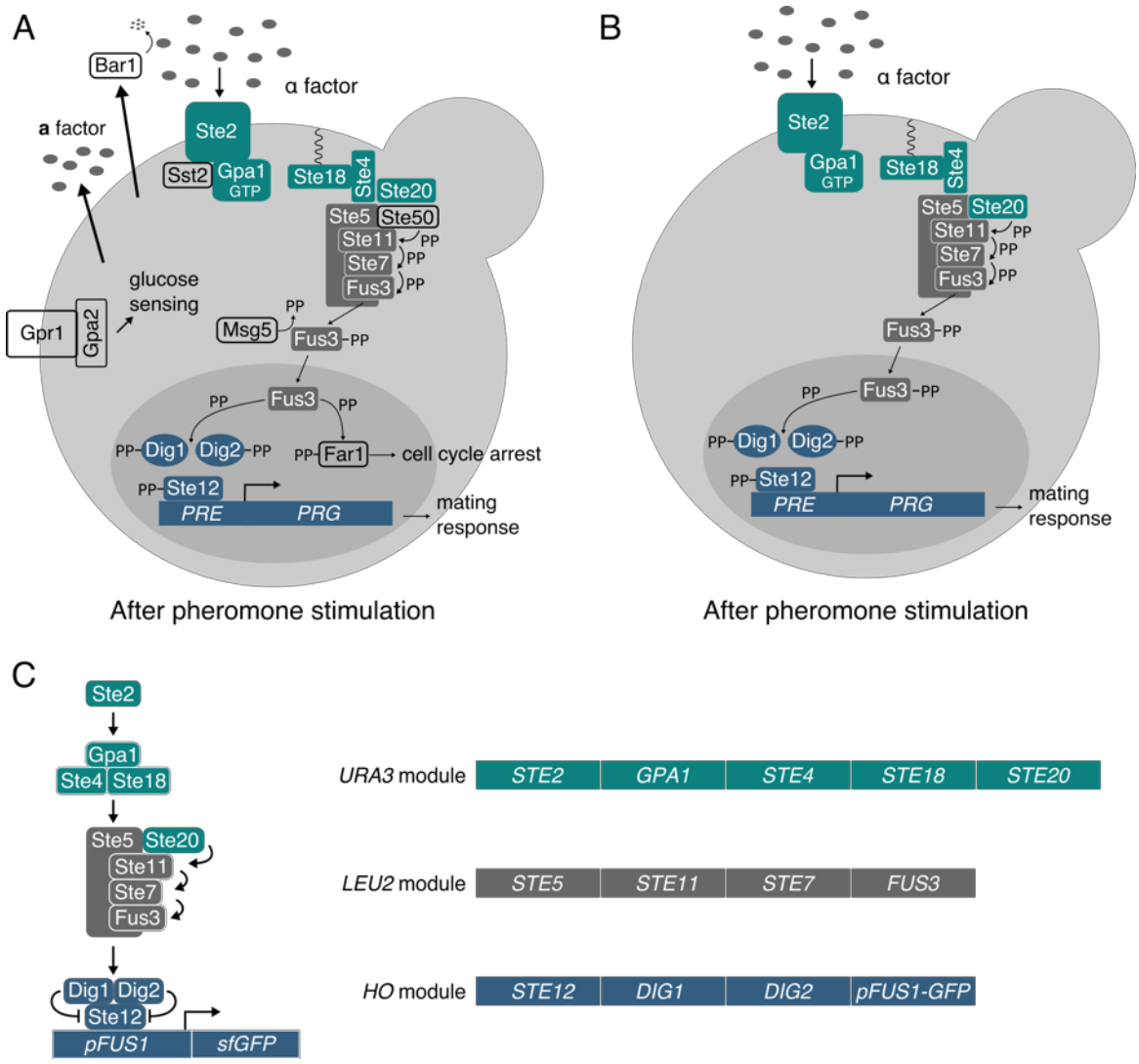


Figure 59. Complete refactoring of a minimal mating pathway. (*A*) A detailed overview of the WT pheromone response pathway in MAT**a** cells after stimulation with α-factor. (*B*) A minimal pheromone response pathway in MAT**a** cells after stimulation with α-factor. The pathway has been reduced to the minimal components for a unidirectional response from agonist activation to a transcriptional output. (*C*) A simplified schematic of the minimal pheromone response pathway linked to a GFP reporter output using the *FUS1* promoter. The colours correspond to functional units that can be separated into 3 modules for full refactoring of the minimised mating pathway; the receptor/G protein unit (green, *URA3* module), the MAPK cascade (grey, *LEU2* module) and the transcriptional response unit (blue, *HO* module).

scaffold modulate the activity of the recruited MAPK proteins[237,238,324]. Changing the concentrations of these components may not be as simple as if this were a simple docking interaction.

Once this has been achieved, feedback regulation can then be overlaid onto the pathway to restore the behaviour of the wild type system. This approach would build up our understanding of the complete pathway from the bottom-up and may prove to be a promising strategy for reigniting the ambitions of the of the Alpha Project – a multidisciplinary effort to model all of the interactions in the yeast pheromone response pathway[325]. This minimal pathway would also provide a biological "breadboard" for the rapid exploration of how genetically-encoded pathway components interact to

produce signal transduction behaviour. As the pathway is composed of many themes common throughout biology, it could provide fundamental insights into more than just the yeast mating pathway.

6.3.4 Refactoring mammalian GPCR signalling pathways

We anticipate the tuning principles uncovered here in yeast will also be relevant for GPCR signalling in mammalian systems. However, it is first worth re-stressing the diversity of components and pathway architecture of GPCR signalling in different cell types and organisms. To apply the same approach to mammalian systems by first stripping out and then refactoring the signalling components would involve the removal of hundreds of genes, all of which would be intricately cross-linked with each other. This approach may be outside of our current engineering capabilities, although this may change as the construction and testing of large genomes become commonplace as part of the Genome Project (GP)-Write Consortium[326]. In the future, we may see custom genomes designed with all of the GPCR signalling pathways removed. These could then be selectively reintroduced in a similar manner as performed in this work to create fully tuneable mammalian GPCR-based systems.

For now, a simple strategy of removing individual components and refactoring them with synthetic promoter libraries and measuring their contribution to the pathway behaviour would be the immediate next step. However, at the very minimum this would require the tuning of not one G protein, but the many G proteins that exist, all of which could interact to influence the activity seen via the others[327]. Whatever path chosen to refactor mammalian GPCR-signalling pathways, it would be a beneficial endeavour, as any way in which we could tune mammalian GPCR-signalling would be beneficial to applications in pharmacology that utilise GPCRs, such as cell-based theranostics, as discussed in section **1.2.8**.

Refactoring signalling pathways may also provide a method for probing mechanisms of disease, for example, cancer, where malignant cells often hijack the physiological roles of GPCRs to survive[130]. As we have shown, changes to the expression of pathway components is an easy route to altering the pathway dynamics. We have demonstrated how a pathway can be turned into a constitutively active system or completely switched off by changing the expression of a single component. This may turn out to be a common theme in the progression of cancer that is receiving less attention than protein polymorphisms due to the complexity of interpreting non-coding mutations[322]. Refactoring signalling pathways may, therefore, provide a new means to recreate these disease states artificially, and thus narrow down the search area for more in-depth pharmacogenomic analysis.

6.4 Conclusions

In section **1.3**, we outlined our proposal for refactoring a minimised mating pathway to allow for the rational tuning of yeast GPCR-based biosensors. Using extensive genome engineering and synthetic tools combined with novel community-based approaches, we were able to achieve these goals. Although limitations remain with the expression of heterologous receptors, we believe the platform we have developed in this study lays the foundations for significantly advancing the field of yeast GPCR-based biosensors. Being able to tune the dose-response properties rationally will allow researchers to design biosensors appropriate for their application demands, bringing us one step closer towards standardising genetically encoded biosensors as a tool for metabolic engineering, as well as other diverse applications.

In addition to the biosensor platform, we have also achieved a significant milestone on the road towards full refactoring of the pheromone response pathway. So far, we revealed the importance of protein stoichiometry, developed a robust model to predict the receptor/G protein signalling, and explained fundamental pathway properties, such as agonist sensitivity through transcriptional feedback. Our current GPCR base strain now offers a starting point for refactoring the remaining pathway components and extending the *in silico* model. As we get closer to the fully refactored pathway and begin to rebuild complex behaviours, such as feedback regulation, we expect to learn much more about this model pathway and signal transduction in general. However, we believe the overall strategy we have taken in this work for simplifying and refactoring a natural pathway will be applicable to other biological systems in and beyond signal transduction. With the accelerating capabilities of genome engineering and synthetic biology in all organisms, it is likely that we will see the creation of equivalent *in vivo* model strains to rationally explore and exploit other important and diverse biological systems.

7 Materials and methods

7.1 Cells manipulation

7.1.1 Bacterial and yeast strains

7.1.1.1 E. coli

NEB Turbo Competent *E. coli* (F' *proA*$^{+}$*B*$^{+}$ *lacIq ΔlacZM15 / fhuA2 Δ(lac-proAB) glnV galK16 galE15 R(zgb-210::Tn10)Tet*^{S} *endA1 thi-1 Δ(hsdS-mcrB)5*) were used for all cloning experiments. An initial glycerol stock of the commercial cell line was created at the beginning of this study, from which in-house competent cells were produced, following the *E. coli* cell competency protocol below.

7.1.1.2 Yeast

Table 6 shows the yeast strains used in this study along with their genotype and parental lineage. For a description of the altered genotypes of these cells used for experiments, see section **8.5.6**.

Table 6. Yeast strains used in this study.

Name	Genotype	Parental Strain
BY4741	*MAT**a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	S288C
GFP	*pTDH3-sfGFP-tTDH1 at URA3 (markerless)*	BY4741
BFP	*pTDH3-sfGFP(T65S, Y66H)-tTDH1 at URA3 (markerless)*	GFP
yWS677	*sst2Δ0 far1Δ0 bar1Δ0 ste2Δ0 ste12Δ0 gpa1Δ0 ste3Δ0 mf(alpha)1Δ0 mf(alpha)2Δ0 mfa1Δ0 mfa2Δ0 gpr1Δ0 gpa2Δ0*	BY4741
Quasi-WT	*STE2 GPA1 STE12*	yWS677
STE2-GFP	*pSTE2-sfGFP-tSTE2*	yWS677
GPA1-GFP	*pGPA1-sfGFP-tGPA1*	yWS677
STE12-GFP	*pSTE12-sfGFP-tSTE12*	yWS677
ALD6-GFP	*pALD6-sfGFP-tALD6*	yWS677

7.1.2 Media and culturing

7.1.2.1 *E. coli*

Selection and growth of *E. coli* was performed in Lysogeny Broth (LB) medium at 37 °C with aeration. With the exception of generating competent cells, the LB medium was supplemented with appropriate antibiotics (ampicillin 100 µg/mL, chloramphenicol 34 µg/mL, or kanamycin 50 µg/mL).

7.1.2.2 Yeast

Yeast extract peptone dextrose (YPD) was used for culturing cells in preparation for transformation: 1% (w/v) Bacto Yeast Extract (Merck), 2% (w/v) Bacto Peptone (Merck), 2% glucose (VWR). Cells were cultured at 30 °C shaking at 250 rpm.

Selection of yeast transformants was performed on synthetic complete (SC) dropout agar medium: 2% (w/v) glucose (VWR), 0.67% (w/v) Yeast Nitrogen Base without amino acids (Sigma), 0.14% (w/v) Yeast Synthetic Drop-out Medium Supplements without histidine, leucine, tryptophan, and uracil (Sigma) supplemented with 20 mg/L tryptophan (Sigma), and 20 g/L bacteriological agar (VWR). Depending on the required selection, SC dropout media was supplemented with 20 mg/L uracil (Sigma), 100 mg/L leucine (Sigma), and 20 mg/L histidine (Sigma). Cells were grown at 30 °C static.

All liquid experiments were performed in synthetic complete (SC) medium with 2% (w/v) glucose (VWR), 0.67% (w/v) Yeast Nitrogen Base without amino acids (Sigma), 0.14% (w/v) Yeast Synthetic Drop-out Medium Supplements without histidine, leucine, tryptophan, and uracil (Sigma), 20 mg/L uracil (Sigma), 100 mg/L leucine (Sigma), 20 mg/L histidine (Sigma), and 20 mg/mL tryptophan (Sigma). Unless otherwise stated in the text, all yeast strains were cultured in 500 µL of synthetic complete media and grown in 2.2 mL 96 deep-well plates at 30 °C in an Infors HT Multitron, shaking at 700 rpm.

7.1.3 Cell competency and transformation

7.1.3.1 E. coli

Chemically competent cells were created following the TSS protocol for KCM transformations[328]. A colony of *E. coli* was grown to saturation overnight in 10 mL of LB and then split into two 2 L baffled flasks with 500 mL of LB. The culture was grown for 2-3 h to an OD600 ~ 1.0, chilled on ice to stop growth, split between 50 mL conical tubes, and centrifuged at 4000 rpm at 4 °C for 10 minutes. The supernatant was then discarded, and the cell pellets resuspended by aspiration in ice-cold TSS (85 mL LB, 10 g PEG-3350, 5 mL DMSO, and 2 mL 1 M $MgCl_2$). 200 µL of the cell suspension was then aliquoted into PCR reaction tubes, flash frozen on dry ice, and put into a -80 °C freezer for long term storage. To transform the DNA, 50 µL of 5 x KCM (500 mM KCl, 150 mM $CaCl_2$, 250 mM $MgCl_2$) was added to 200 µL of the competent cell prep after 10 minutes for thawing on ice. 50 µL of the competent cell-KCM cocktail was then added to DNA and transferred to a thermocycler with the following protocol: 4 °C for 10 minutes, 42 °C for 1 minute, 4 °C for 1 minutes, and then 30-60 minutes recovery at 37 °C. Cells were then plated on solid LB media supplemented with the appropriate antibiotics.

7.1.3.2 Yeast – low-throughput

Chemically competent yeast cells for low-throughput, tube-based transformations were created following the lithium acetate protocol[329]. Yeast colonies were grown to saturation overnight in YPD. The following morning the cells were diluted 1:100 in 15 mL of fresh YPD in a 50 mL conical tube and grown for 4-6 h to OD600 0.8-1.0. Cells were pelleted and washed once with 10 mL 0.1 M lithium acetate (LiOAc) (Sigma). Cells were then resuspended in 0.1 M LiOAc to a total volume of 100 µL/transformation. 100 µL of cell suspension was then distributed into 1.5 mL reaction tubes and pelleted. Cells were resuspended in 64 µL of DNA/salmon sperm DNA mixture (10 µL of boiled salmon sperm DNA (Invitrogen) + DNA + ddH_2O), and then mixed with 294 µL of PEG/LiOAc mixture (260 µL 50% (w/v) PEG-3350 (Sigma) + 36 µL 1 M LiOAc). The yeast transformation mixture was then heat-shocked at 42 °C for 40 mins, pelleted, resuspended in 200 µL 5 mM CaCl2 and plated onto the appropriate solid agar media.

7.1.3.3 Yeast – high-throughput

Chemically competent yeast cells for high-throughput, plate-based transformations were created following a variation on the lithium acetate protocol[329]. Yeast colonies were grown to saturation overnight in YPD. The following morning, the cells were diluted 1:100 in 100 mL of fresh YPD in a 500 mL baffled flask and grown for 4-6 h to OD600 0.8-1.0. Cells were split into two 50 mL conical tubes, pelleted and washed once with 25 mL 0.1 M lithium acetate (LiOAc) (Sigma). Cells were then resuspended to a total volume of 3.38 mL in H_2O and mixed with 12.7 mL 50% PEG-3350 (w/v), 1.9 mL 1 M LiOAc and 528 µL salmon sperm DNA (10 mg/mL). 175 µL of the yeast transformation mixture was then added to each well of a 96-well reaction plate (STARLAB) containing 10 µL of the transformation DNA and mixed by aspirating gently several times. The transformation plate was then transferred to a 96-well thermocycler block, set at 42 °C, for 40 minutes. Following the heat-shock, the yeast suspension was transferred to a 96-well V-bottom plate and pelleted at 2000 rpm for 10 minutes. The supernatant was then removed and 200 µL of 5 mM CaCl2 was added to each well and left to recover for 10 minutes shaking at 700 rpm, room temperature. 10 µL of the cell suspension was spotted onto the appropriate solid agar media in an OmniTray (Thermo Scientific) and left to dry next to an open flame. This was repeated another 2 times on separate plates with 1:10 serial dilutions of the cell suspension to ensure varied dilution of the transformed yeast to account for transformation efficiencies.

7.2 Yeast MoClo Toolkit (YTK) system

All plasmids used in this study were created using the Yeast MoClo Toolkit (YTK) system[33]. The YTK system uses Golden Gate assembly to combine user generated parts and a library of highly-characterised parts supplied with the toolkit into single- and multigene plasmids for expression in yeast. Most of the parts used here were generated within this study. However, several parts were kindly provided by colleagues in the Ellis Lab, and are acknowledged in the parts list. A list of all parts and assembled expression cassettes used in this study can be found in sections **8.5.1** and **8.5.6**, respectively. Aspects of this system were updated during this work and are explained in section **8.3**. This section assumes a working knowledge of the YTK system. See Lee *et al.*[33] main text and supporting information for a description of the YTK system and section **8.3** of this work for amendments to the system, which should be read together.

7.2.1 General DNA manipulation

All parts generated in this study were constructed using the cloning techniques, or combinations of the cloning techniques, described below. Unless otherwise stated, restriction endonucleases, Phusion-HF DNA polymerase, T4 polynucleotide kinase, and T7 DNA ligase were obtained from NEB. Oligonucleotides were obtained from IDT. Benchling (www.benchling.com) was used for the design all of oligonucleotides and DNA constructs used in this study. Standard Phusion-HF DNA polymerase PCR conditions and protocols were used for the amplification of DNA in all instances. Zymoclean Gel DNA Recovery kit (Zymo Research) was used for cleaning up DNA after gel electrophoresis. Plasmids were prepared from bacterial culture and PCR reactions purified following the Qiagen plasmid miniprep and PCR purification protocol, respectively, using homemade Qiagen buffers (https://openwetware.org/wiki/Qiagen_Buffers) and purification columns from an alternative supplier (NBS Biologicals). DNA was quantified using a NanoDrop 1000 Spectrophotometer following the manufacturers guidelines.

7.2.2 Yeast genomic DNA isolation for PCR-based applications

Genomic DNA from yeast was isolated using a lithium acetate/SDS extraction protocol[330]. Colonies were picked from a plate or spun down from 100-200 µL of saturated liquid yeast culture, resuspended in 200 µL of 200 mM LiOAc, 1% SDS solution, and incubated at 70 °C for 5 minutes. 300 µL of 100% ethanol was then added to the suspension which was vortexed and spun down at 15,000 g for 3 minutes. The supernatant was removed, and the pellet resuspended in 100 µL H_2O. The suspension was spun down for a final 20 seconds at 15,000 g, and 80 µL of the resulting supernatant was moved to a fresh tube for immediate usage or transferred to the -20 °C freezer for long term storage.

7.2.3 Part construction

To generate new parts conforming to the YTK standard, native yeast sequences were domesticated by removing all instances of the BsmBI, BsaI, BpiI and NotI recognition sequences using Multiplexed site-directed mutagenesis (see section **7.2.3.3**). Non-yeast and de novo sequences were synthesised by GeneArt or IDT using their DNA fragment and gene synthesis services, or in-house using combinations of the assembly methods below.

7.2.3.1 Golden Gate assembly

Golden Gate assembly was the assembly method of choice when the type IIs restriction enzymes, BsmBI, BsaI, or BsaI, were freely available to use within the reaction without interfering with internal recognition sites within the final sequence.

All DNA for Golden Gate reactions was set to equimolar concentrations of 50 fmol/µL (50 nM) prior to experiments. Golden Gate reactions were prepared as follows: 0.1 µL of entry vector (backbone), 0.5 µL of each DNA fragment or plasmid, 1 µL T4 DNA ligase buffer (Promega), 0.5 µL T7 DNA Ligase (NEB), 0.5 µL restriction enzyme (BsaI, BsmBI or BbsI-HF) (NEB), and H_2O to bring the final volume to 10 µL. Reaction mixtures were then incubated in a thermocycler using the following program: (42 °C for 2 min, 16 °C for 5 min) x 25 cycles, followed by a final digestion step of 60 °C for 10 min, and then heat inactivation at 80 °C for 10 min.

7.2.3.2 Gibson assembly

Gibson assembly was the assembly method of choice when the type IIs restriction enzymes, BsmBI, BsaI, or BbsI, were not freely available to use within the reaction without interfering with internal recognition sites within the final sequence.

A Gibson assembly master mix was created in two stages by combining 160 µL of an isothermal mix (500 mM Tris-HCl (pH 7.5), 50 mM $MgCl_2$, 1 mM Deoxynucleotide (dNTP) Solution Mix (NEB), 5 mM DTT, 25% PEG-8000, 5 mM NAD) with 0.32 µL T5 exonuclease (NEB), 10 µL Phusion-HF DNA polymerase (2 U/ µL), 80 µL Taq DNA ligase (NEB), and 5.75 mL H_2O. The resulting master mix was distributed into PCR reaction tubes in 15 µL aliquots, and stored at -80 °C. To perform a Gibson assembly reaction, a vial of the Gibson master mix was thawed on ice, added to 5 µL of DNA and then incubated at 50 °C for 1 hour. Completed reactions were kept on ice until transforming half the reaction volume into *E. coli*.

7.2.3.3 Multiplexed site-directed mutagenesis

To domesticate DNA sequences containing internal BsaI, BsmBI, BbsI, or NotI recognition sites into the YTK format, a parallel assembly method for multiplexed site-directed mutagenesis of DNA developed by Yan *et al.*[331] was employed. Based on Golden Gate assembly, this method uses type IIs restriction enzymes and T7 DNA ligase to assemble multiple PCR-generated fragments into a single plasmid according to a user defined order. The Benchling Golden Gate assembly wizard was used to design the primers for generating the fragments and mutating the desired sites to remove the recognition sequences. The standard Golden Gate assembly protocol was then used to assemble the fragments.

7.2.3.4 Small fragment assembly

Assembly of the gRNAs and the C-terminal G protein transplants into their respective entry vectors was achieved by annealing two phosphorylated oligonucleotides with sticky ends and performing a Golden Gate reaction with the small fragments and the entry vectors. When ordering overlapping oligos they are designed such that when annealed, they have the desired sequence and sticky ends (at least 15 bases shared between the oligos). For example, inserting a new 20 bp target sequence into the SpCas9 sgRNA Dropout (pWS2061).

Forward: 5’ agatNNNNNNNNNNNNNNNNNNNN 3’

Reverse: 5’ aaacNNNNNNNNNNNNNNNNNNNN 3’

When annealed these primers will form the following small DNA fragment:

5’ agatNNNNNNNNNNNNNNNNNNNN 3’
3’ NNNNNNNNNNNNNNNNNNNNcaaa 5’

The resulting small fragment represents an artificially “pre-cut” DNA fragment, analogous to a type IIs restriction enzyme digested DNA intermediate found within a Golden Gate reaction, with the sticky ends determining the position within the assembly. This method was also used to create new DNA sequences up to 200-300 bp, by ligating multiple fragments together in an order defined by their sticky ends. To decrease the likelihood of a misassembly between small fragments, the number of base pairs within a sticky end can be varied between fragments to increase the specificity of the ligation partners.

To create the small fragment, oligonucleotides are first resuspended at 100 µM concentration in H_2O. Each oligonucleotide is then treated separately with T4 polynucleotide kinase (PNK) in the following reaction: 1 µL oligonucleotide (100 µM), 1 µL 10 x T4 DNA ligase buffer (Promega), 0.5 µL T4 PNK (NEB), and 7.5 µL H_2O. The mixture is then incubated at 37 °C for 1 hour. The 10 µL reactions for both oligonucleotides in the fragment pair are then added together and brought to a total volume of 200 µL in H_2O (10 µL oligo (sense) + 10 µL oligo (antisense) + 180 µL H_2O). The oligonucleotides are then annealed under slowly decreasing temperatures using the following programme: 96 °C for 6 minutes followed by 0.1 °C/s ramp down to 20 °C, and then hold at 20 °C. 1 µL of the resulting reaction is used to ligate into the desired vector using standard a Golden Gate assembly reaction.

7.2.3.5 Synthesis design

Non-yeast and novel sequences were first designed in Benchling and then ordered either through GeneArt or IDT as gene fragments or sequenced gene inserts. All open reading frames were codon optimised for *Saccharomyces cerevisiae* using the codon optimisation service of the synthesis provider. Synthetic DNA sequences were cloned into pYTK001.

7.2.3.6 Verification of new parts

All new part constructs were verified by restriction enzyme digest followed by direct Sanger sequencing of the full insert (Source Bioscience). Once confirmed, parts were retransformed into *E. coli*, glycerol stocked at -80 °C and a working stock set to 50 nM stored in the -20 °C.

7.2.4 Cassette construction

Cassettes and multi-gene cassettes were constructed following the YTK hierarchical workflow using pre-assembled cassettes (see section **8.3.6** for a list of pre-assembled cassettes) using the standard Golden Gate assembly reaction. See **Figure 25** for details on the position of the cassettes used for creating GPCR-based biosensors.

7.2.4.1 Verification of cassettes

Plasmids were validated at the single- and multigene cassette levels using NotI-HF restriction digest analysis. No sequencing of cassettes was performed at the cassette level, as the mutational rate of *E. coli* plasmid amplification is low enough to assume no errors would be introduced during the brief period of cloning.

7.2.4.2 Transformation and validation of plasmid integration

Single integration of marker plasmids was performed with 300 ng of plasmid. Double, and triple integration of marker plasmids were performed with 50, 100, and 200 ng of plasmid, respectively, with 100 ng of Cas9 and 200 ng of each gRNA expression cassette. All integration plasmids were first linearized by digestion before transformation using BpiI (Thermo).

Due to a high likelihood of an incorrect integration of any particular cassette (6% for a triple integration – see **Figure 22**), all biosensing strains were screened for correct integration before collecting data for analysis. Screening was achieved by performing a pre-experiment on 4 colonies from a single transformation and determining their ON/OFF response. A single strain, consistent with the other strains, was then re-streaked onto the appropriate selection media and then incubated for an additional 2 days before storing at 4 °C for a minimum of 1 day. The addition of the 1 day 4 °C storage was essential for consistency; especially with biosensor strains using heterologous GPCR receptors.

7.2.4.3 *in vivo* gap repair assembly

The YTK toolkit was updated in this study to allow cassettes to be assembled into multigene plasmids in yeast via homology directed repair using terminal homology between pre-digested cassettes, following a method similar to the Versatile Genetic Assembly System (VEGAS) from the Boeke Lab[332]. See section **8.3.3** for a description of the new gap repair method.

DNA for gap repair assembly was prepared as follows: 50 ng acceptor cassette, 500 ng of each insert cassette, 2 µL of 10 x Buffer G (Thermo), 1 µL BpiI (BbsI) (Thermo), and H_2O to bring the final volume to 20 µL. Reaction mixtures were then incubated overnight at 37 °C followed by 20 minutes heat inactivation at 65 °C the next day. The entire reaction was then transformed directly into yeast.

7.3 CRISPR-mediated genome editing

7.3.1 Markerless genome editing using CRISPR/Cas

Reagents for CRIPSR markerless genome editing was prepared as follows: 50 ng of CRISPR/Cas plasmid, 500 ng of each sgRNA expression vector, 2 µL of 10 x Buffer G (Thermo), 1 µL of BpiI, made up to 20 µL total with H_2O. The digestion mixture was then incubated > 4 h at 37 °C, followed by heat inactivation at 65 °C for 20 minutes. The donor DNA was then added to the mixture to a total volume of 54 µL in H_2O and transformed directly into yeast using the low throughput yeast transformation protocol.

7.3.1.1 Marker cycling

Using the above protocol, yeast cells were transformed with the CRISPR plasmids and donor DNA and plated on the appropriate selective media. Once colonies had appeared, overnight cultures were inoculated into YPD. The following day, the yeast was validated for the edits and correct strains were back diluted 1:100 in YPD and prepared for transformation. The yeast was then transformed with the CRISPR plasmids and donor DNA for the next edit using a different selectable marker. This procedure can be repeated until all required edits have been performed, requiring at least three markers within the cycle to prevent background from the previous markers (see section **2.2.3** for more information).

7.3.1.2 Donor DNA

Donor DNA was created in one of two ways depending on the application. For single edits, with lower efficiency requirements, donor DNA was produced using PCR amplification of a template or overlap PCR extension from oligos, to add ~ 40 bp of flanking homology to an insert. For multiplexed edits, with much greater efficiency requirements, donor DNA was produced by first cloning the sequence into the pYTK001 CamR-ColE1 backbone, which would consist of ~ 500 bp of homology flanking an insert. The plasmid was then sequence verified, and primers were designed to amplify the donor DNA from the plasmid.

7.3.2 Validation of markerless edits

7.3.2.1 Colony PCR

Genomic DNA for colony PCR was prepared following the protocol in section **7.2.2**. GoTaq Green Master Mix (Promega) was used for PCR amplification or the genomic DNA for gel analysis. Using 1 µL of genomic DNA prep, 2 µL of the forward and reverse primer (10 µM), and 5 µL of the GoTaq

Green Master Mix, PCR reactions were performed following the manufacturers thermocycling protocol.

7.3.2.2 Direct sequencing

Genomic DNA for direct sequencing was amplified from ~ 100 bp upstream and downstream of the edited region, including the homology arms of the donor DNA, using Phusion DNA polymerase, according to the manufacturer's specifications. The DNA fragment was then Sanger sequenced (Source Bioscience) using the amplification primers, plus additional primers if required, to obtain the full sequence of the edited region. Sanger reads were aligned to the expected sequence in Benchling.

7.3.2.3 Nanopore sequencing of yWS677

DNA was isolated from yWS677 for Nanopore sequencing using the 100/G Genomic-Tip kit (QIAGEN), sheared to 20 kb using a g-TUBE (Covaris) and prepared for sequencing using a Ligation Sequencing Kit $1D^2$ R9.5 (Oxford Nanopore Technologies). The genomic DNA was then run on an R9.5 flow cell using a MinION Mk 1B (Oxford Nanopore Technologies). A standard 48h sequencing run was performed using the MinKnow 1.5.5 software using local basecalling. Reads were exported directly to fastq using MinKNOW. Canu (v1.5) was used to correct raw reads (www.canu.readthedocs.io) and smartdenovo (www.github.com/ruanjue/smartdenovo) was used to de novo assemble the reads into contiguous sequences (contigs) using default flags. The resulting contigs were compared to the WT reference genome (s288c, SGD) using lastdb/lastal (www.last.cbrc.jp) and viewed on integrative genome viewer (IGV) (www.software.broadinstitute.org) to inspect genomic changes.

7.4 Data collection and analysis

7.4.1 Biosensing protocol

7.4.1.1 Assay format

All biosensing strains were picked into 500 µL of synthetic complete media and grown in 2.2 mL 96 deep-well plates at 30 °C in an Infors HT Multitron, shaking at 700 rpm overnight. The next day, saturated strains were then diluted 1:100 into fresh media (for mixed populations the volumes are stacked). After 2 h of incubation the strains were induced with their respective ligands and incubated for a further 4 h, for a total assay time of 6 h. All ligands were dissolved in DMSO, and the final concentration in all cultures was 1%. For strains using the Z_3E-PRD and TetR-PRD transcription factor, aTc and β-estradiol was added during the back dilution at time 0 h. To perform flow cytometry and plate reader measurements, 200 µL from each well was directly transferred to a 96-well clear, flat-bottom microplate (Corning).

7.4.1.2 Flow cytometry

Unless otherwise stated, flow cytometry was used for all monoclonal biosensing experiments. Cell fluorescence was measured using an Attune NxT Flow Cytometer (Thermo Scientific) with the following settings: FSC 300 V, SSC 350 V, BL1 500 V, YL2 450 V, or by a LSRFortessa (BD Biosciences) with the following settings: FSC 330 V, SSC 250 V, 405 nm line (450/50) 378 V, 488 nm line (530/30) 450 V, 640 nm line (685/35) 590 V. Fluorescence data was collected from 10,000 cells for each experiment and analysed using FlowJo software. The data presented is the geometric mean of the height of measurements from the respective channels. Compensation was applied to data sets using more than one fluorescent protein.

7.4.1.3 Plate reader

Unless otherwise stated, a plate reader was used for all polyclonal biosensing experiments. Cell fluorescence was measured with a Synergy HT Microplate Reader (BioTek) with the following settings for measuring sfGFP: excitation 485/20, emission 528/20, gain 80. The machine was set to 30 °C for time course experiments with shaking between measurements.

7.4.1.4 Analysis

All presented dose-response fittings were generated in GraphPad Prism 7, which was used to determine the pEC50 and Hill slope wherever shown. To determine all remaining properties of the dose-response curve, curve fitting was performed using Python (SciPy and Matplotlib) using the 4PL model, where x is the concentration, A is the minimum asymptote, B is the steepness, C is the inflection point and D is the maximum asymptote:

$$f(x) = \frac{A - D}{1 + \left(\frac{x}{C}\right)^B} + D$$

7.4.2 RT-qPCR

For RNA purification, RNA was isolated from yeast culture grown to an OD600 of 1.0 ± 0.1 using a YeaStar RNA Kit (Zymo Research) according to the manufacturer's instructions. RNA was quantified by a NanoDrop spectrophotometer (Thermo Fisher) and cDNA was generated from each RNA prep using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). All quantitative PCR (qPCR) reactions were performed in a MasterCycler RealPlex 4 (Eppendorf) using SYBR FAST Universal qPCR Master Mix (Kapa Biosystems) according to the manufacturer's instructions. Each qPCR reaction contained 20 ng of cDNA. qPCR results were normalized to the housekeeping gene *HTB2*. All qPCR primers were designed manually using Benchling, selecting an amplicon length between 100-120 bp and a Tm of 60-61 °C. Primer pairs were validated with an initial qPCR run to check for secondary products.

7.4.3 Measurement of melatonin in media by LC-MS and MTNR1A biosensor

Samples for mass spectrometry and the MTNR1A biosensor were prepared by centrifuging yeast cultures at 4000 rpm in a large desktop centrifuge for 10 minutes at 4 °C and extracting the supernatant. Supernatant samples were kept on ice before running on the LC-MS or transferred to the -20 °C for later use. No further sample preparation was performed on the supernatant sample before running on the LC-MS. Standards were kept in 100% DMSO before being diluted in spent media. Spent media was prepared from BY4741 in the same manner as the measured yeast.

An LC-MS method was developed for the measurement of melatonin in media, using an Agilent 1290 UPLC and 6550 quadrupole – time-of-flight (Q-ToF) mass spectrometer with electrospray ionization (Santa Clara, CA). The UPLC column was an Agilent Zorbax Eclipse Plus C-18, 2.1 x 50mm and 1.8um particle size. The UPLC buffers were 0.1% formic acid in water and 0.1% formic acid in acetonitrile (v/v). The gradient elution method is detailed in **Table 7**.

Table 7. The LC gradient elution method for the measurement of melatonin in media.

Time (minutes)	% Solvent A	% Solvent B	Flow rate (mL/min)
0	100	0	0.5
0.5	100	0	0.5
1.5	70	30	0.5
2	5	95	0.5
2.5	5	95	0.5
2.6	100	0	0.5
3.6	100	0	0.5

Quantitation was based on the LC retention time from melatonin standard solutions and the area of accurately measured diagnostic ions from the molecule, namely the protonated molecule, $[M+H]^+$, along with an in-source fragment (**Table 8**). The solutions of a melatonin standard in media were used to generate calibration curves.

Table 8. The MS ions used for the measurement of melatonin.

	Quantifier ion [M+H]+	Qualifier ion [M+H-C_2H_5NO]
Melatonin ($C_{13}H_{16}N_2O_2$)	233.1285	174.0913

To measure the melatonin from producer strains using the MTNR1A sensor strains, 50 µL of supernatant was added to an adjusted 450 µL volume of biosensing cells according to the protocol in section **7.4.1.1** and run on the flow cytometer according to the protocol in section **7.4.1.2**. Melatonin concentrations were then calculated from a standard curve.

8 Supplementary information

8.1 List of Supplementary Tables and Figures

8.2 Supplementary data

Supplementary Table S1. Properties of the Design 1-4 α-factor dose-response curves in Figure 35 and Figure 41.

Property	Quasi-WT	Design 1	Design 2	Design 3
Tightness	0.052 ± 0.001	0.054 ± 0.0003	0.997 ± 0.007	1.039 ± 0.014
Sensitivity (log(M))	-9.79 ± 0.05	-8.45 ± 0.1	-8.48 ± 0.04	-9.48 ± 0.21
Operational range (log(M))	1.89 ± 0.02	1.09 ± 0.15	1.4 ± 0.05	2.19 ± 0.27
Max output (GFP fluorescence (fold over background))	134 ± 3.28	47.65 ± 0.31	55.56 ± 1.03	68.36 ± 3.72
Dynamic range	6.97 ± 0.12	2.58 ± 0.02	55.40 ± 0.68	71.02 ± 4.51

Supplementary Table S2. UAS and core promoter combinations in Figure 50.

Low/Mid/High	UAS	Core promoter	Final sensor
1.0	***LexO (6x)***	***pLEU2m***	**Low sensitivity**
2.01	*LexO (1x)*	*pRNR2m*	
2.02	*LexO (1x)*	*pPHO5m*	
2.03	*LexO (1x)*	*pLEU2m*	
2.04	*LexO (2x)*	*pRNR2m*	
2.05	***LexO (2x)***	***pPHO5m***	**High sensitivity**
2.06	*LexO (2x)*	*pLEU2m*	
2.07	*LexO (3x)*	*pRNR2m*	
2.08	***LexO (3x)***	***pPHO5m***	**Mid sensitivity**
2.09	*LexO (3x)*	*pLEU2m*	
2.10	*LexO (4x)*	*pRNR2m*	
2.11	*LexO (4x)*	*pPHO5m*	
2.12	*LexO (4x)*	*pLEU2m*	

STE2 (-700...-1)

CTGATTAGTATGAAATAATGACCTTAATGGCATCGTATTTTGTTAATTGGCTTGTTCTATGTCATTTGCTAAATTTTCTCT
TTATTTCTTTACTTTGCCAATTCAAGATAATCTTCCCTTCCCAGAGAGAAAAAAGGAAAAATTTAGCTATGAAACCTCAA
TAAGCTTTTTAATACACCAAAGATTCAAGATAAGAGCATAGAACGAACTGTAGAATAGTCCGGATATGTTATCCAATGC
CTGCCAAAATGCATTGTCACACGCTGTAGTGCTCGAATAGGTGTTGCAATCCGTCAATATACGTCTTGCTCTGTGGGT
AAATGTCTCGTGCATTAAGACAGGCTAGTATAAACGAGAAGAAGTATCCTGCTTTGCAATGAAACAATAGTATCCGCTA
AGAATTTAAGCAGGCCAACGTCCATACTGCTTAGGACCTGTGCCTGGCAAGTCGCAGATTGAAGTTTTTTCAACCATG
TAAATTTCCTAATTGGGTAAGTACATGATGAAACACATATGAAGAAAAAAGCTTTCCTACATATTCAAGATTTTTTTCTG
TGGGTGGAATACTATTTAAGGAGTGCTATTAGTATCTTATTTGACTTCAAAGCAATACGATACCTTTTCTTTTCACCTGC
TCTGGCTATAATTATAATTGGTTACTTAAAAATGCACCGTTAAGAACCATATCCAAGAATCAAAA

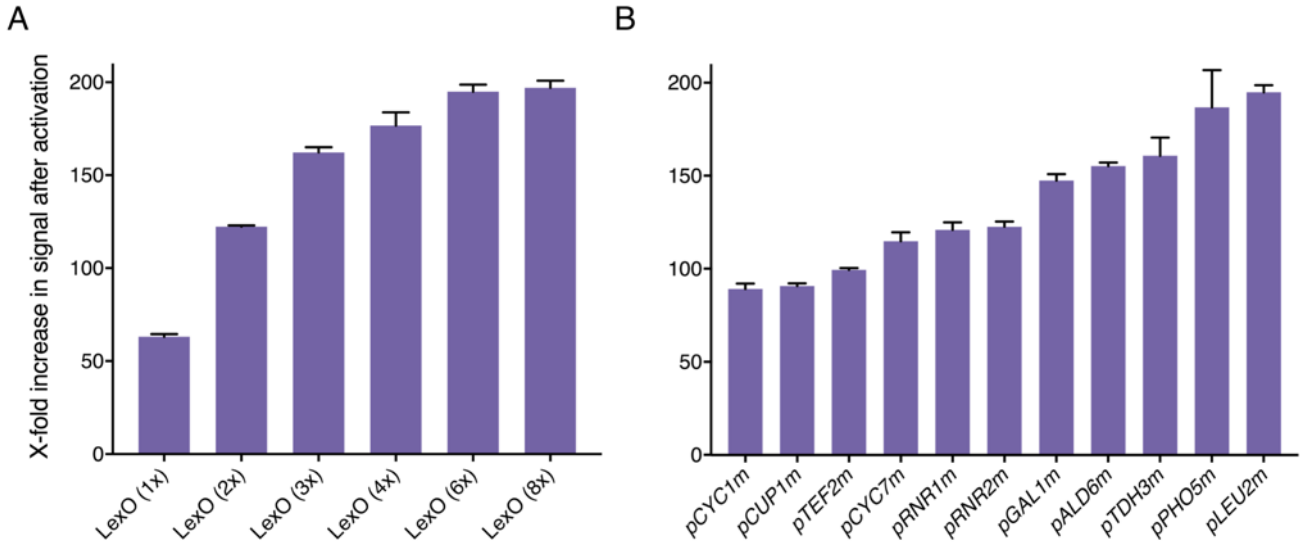
GPA1 (-700...-1)

ATGTGCATTAAAGCAGTAATGATAAGACGAAAATAAGAAAAGAAATCCATAAGCTGTTTTACTCGACTCAACGTTATAA
TTAGTATTATTGATTCATATCCTGTATATACAAGTAACATTATACTCTTTTCTGTACTTCATCTTTACCTTTTTATCTTACA
TGGCACATGTTGTTTGAAAACAAGATCATAGGTGGATAAAGCAAGCCGAATCTAAAAAAAAAAAAAATGTCTCTATTGGA
AAACTGAATGCATAACGATATTCCCTTTTCATGCAGATAACACTGACTAGTTTCAATTTGAAACGCATCTTCGTGTTATT
TCACCGAAACGCACTCGGCTCAGCATGTTAAAAAGCACATCAATTTAGGGCTCTGCGCGTCCTTCTGCGTATTCTTCC
TTGTAGAAATGCAATTAAATGGAGAGCAGAAATTTTTTTGTTACATATTGTTTTCCTTAAAGGGAAATATTAAAAAATAG
TCTAAAATGAAGAGGATAGTAGAATTCCACCAATTTCTTTACGTTTTATATTATTCGTAATCTTTTGATCTGTTATTCATT
TTTTCTTGTCACTCCGTTTCTAACATTTTTGACCATTTCTAAGACCAAACTGAGTAGAAGCTATTCATACTGTAAATTGG
TATTTTAGCATCACATCAATAATCCAGAGGTGTATAAATTGATATATTAAGGTAGGAAATA

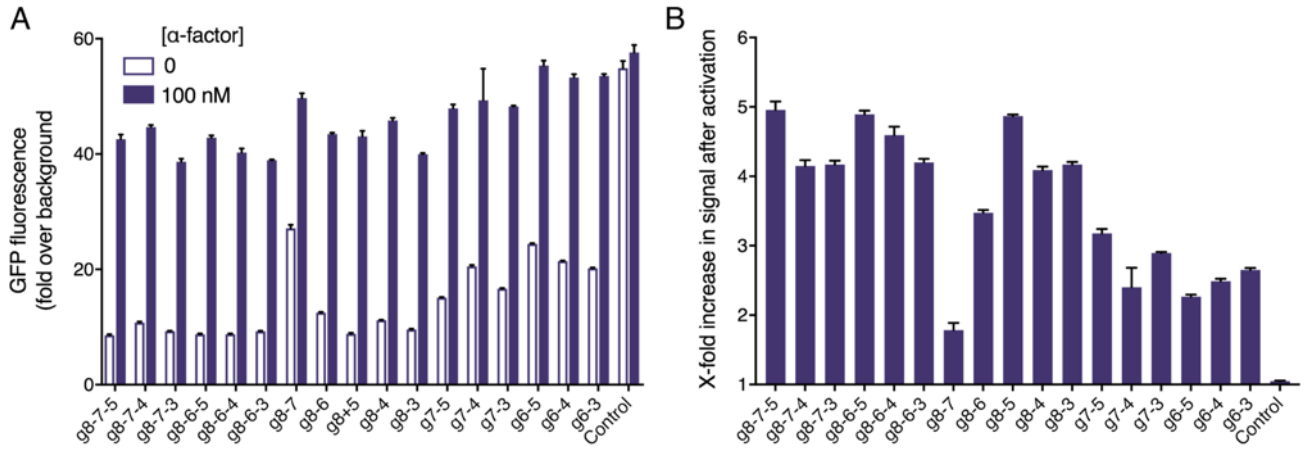
STE12 (-700...-1)

CTTACTTACATCTGAAAATTGCAAGTTACATTCTTTGTATAACGAACGTTAAGGAACCCATAAAGCTAAAGACATTTGTT
GAAAACGAATGTAAAGAATTGGTCCAGTTTGCACAAGACACCCTGAAGAACTTCGTTCAGTAATCACTTTCAAGCTGTA
GTATGTAAACGATATAGATGAAGTTTTCGTGTGTATAAATATATGAACTCTAGAGTGTTGCATAATTTGAAACACAGCAT
TTCTTTTCGGAGAGCTCGTTTCAAAATGAAACAAACGCCGTTGTCCGTTTTCGTCTCAATAGAAAAAGTGAAACAGATA
AAAATTGTTTTAAAAGAAACGAATTTGCAACATCTTAAGATATATCAAAACTAATACAAAACAGCCTAAAAAAGATTGAA
CAACTCTTCGCGGTCAGGTCTCGACACCATAAATCGAAGTACTCGTACGCTAGTTTTCTCGCACATAGTACCACTACG
TTCCTTTTACAATTAGATTACTTCTTTTTAGTTGACTTTTTTGAGACGTTCGTGCCATTCATAAAATAGGAAAAGATAACA
GGTAAGCACTGAAGACTTGTTTTATAAGTGTCCCAAGCGAGACCTAGAGTGGATATTGATATTTCTCAAACAAGACTC
GTCGAAGAAAACACACTTTTATAGCGGAACCGCTTTCTTTATTTGAATTGTCTTGTTCACCAAGG

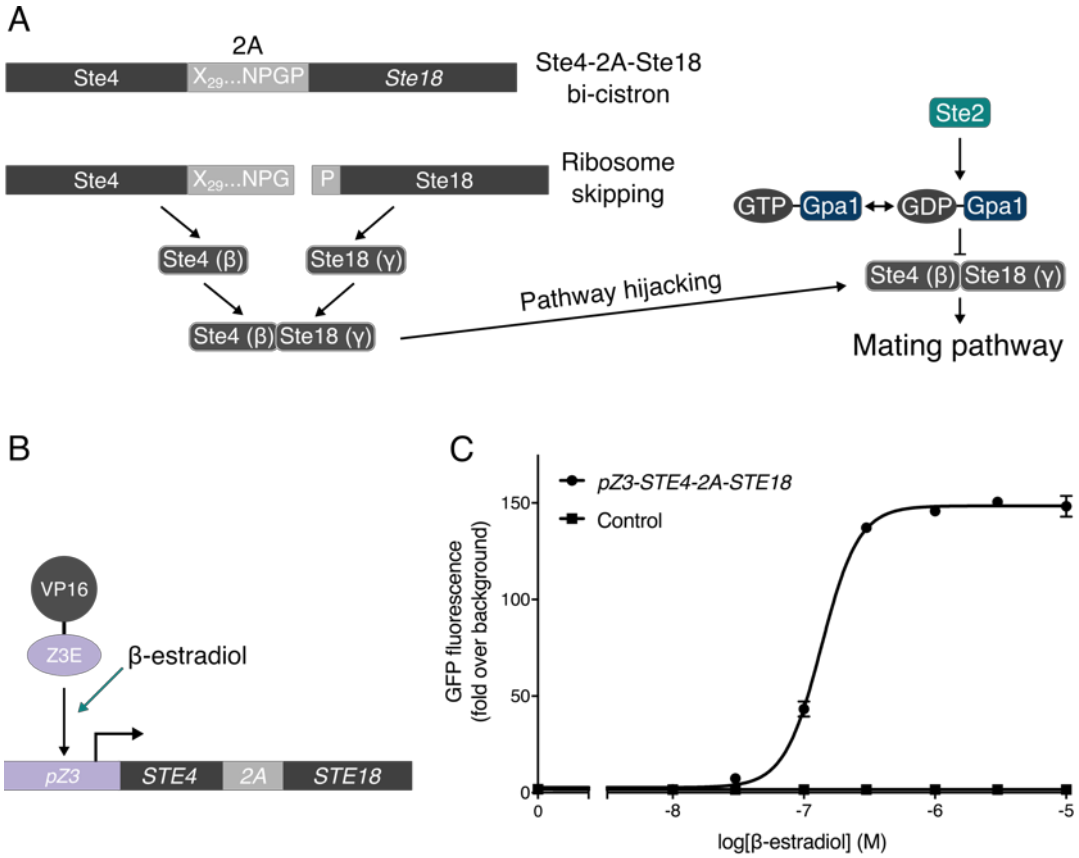
Supplementary Figure S1. PRE consensus sequences in the promoter regions of *STE2*, *GPA1*, and *STE12*.
Promoter regions of the STE2, GPA1, and STE12 including 700 bp upstream of the coding sequence. Consensus PRE sequences highlighted yellow.



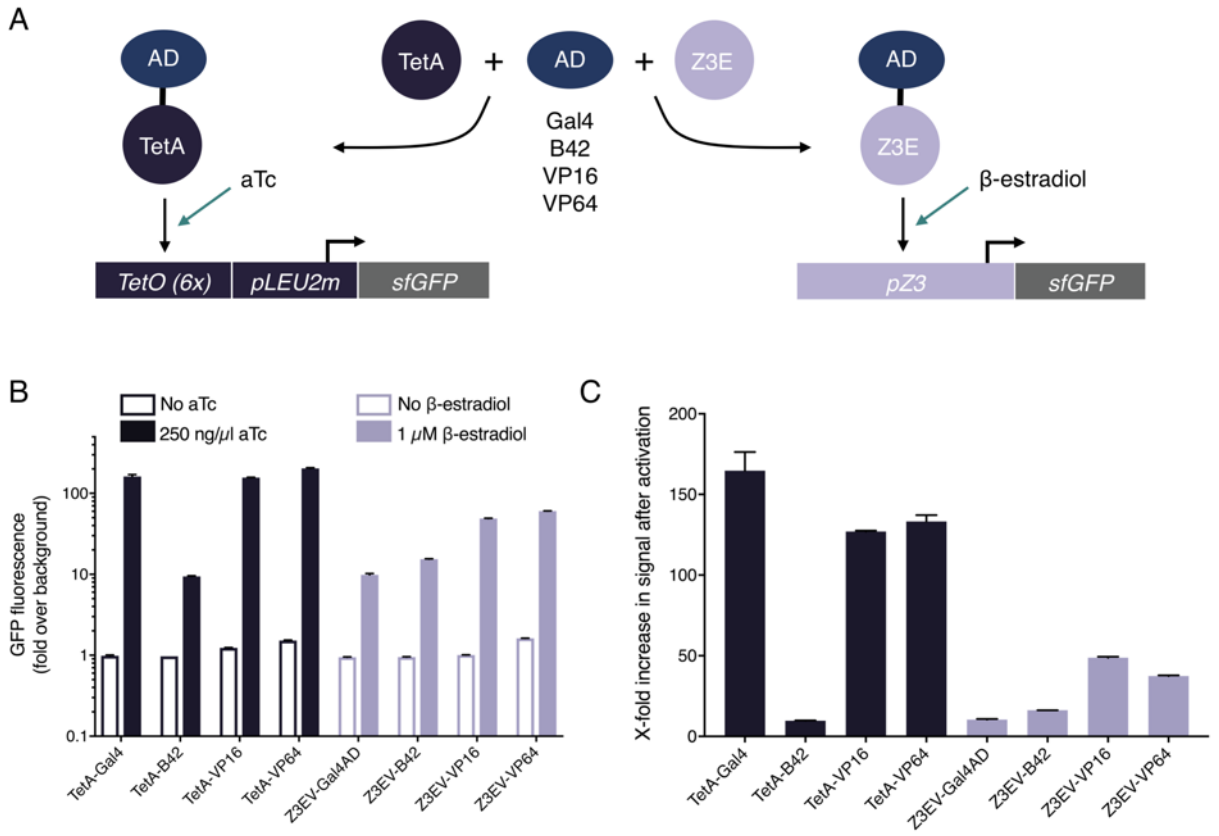
Supplementary Figure S2. Maximum fold-change in expression of the synthetic promoter libraries in Figure 38. (*A*) The LexA-PRD sTF targeted to synthetic promoter variants with different numbers of LexO sites upstream of the *LEU2* core promoter driving GFP expression. (*B*) The LexA-PRD sTF targeted to synthetic promoter variants with 6 repeats of the LexO upstream of different core promoter regions driving sfGFP expression. Results are means ± standard deviation from triplicate isolates.



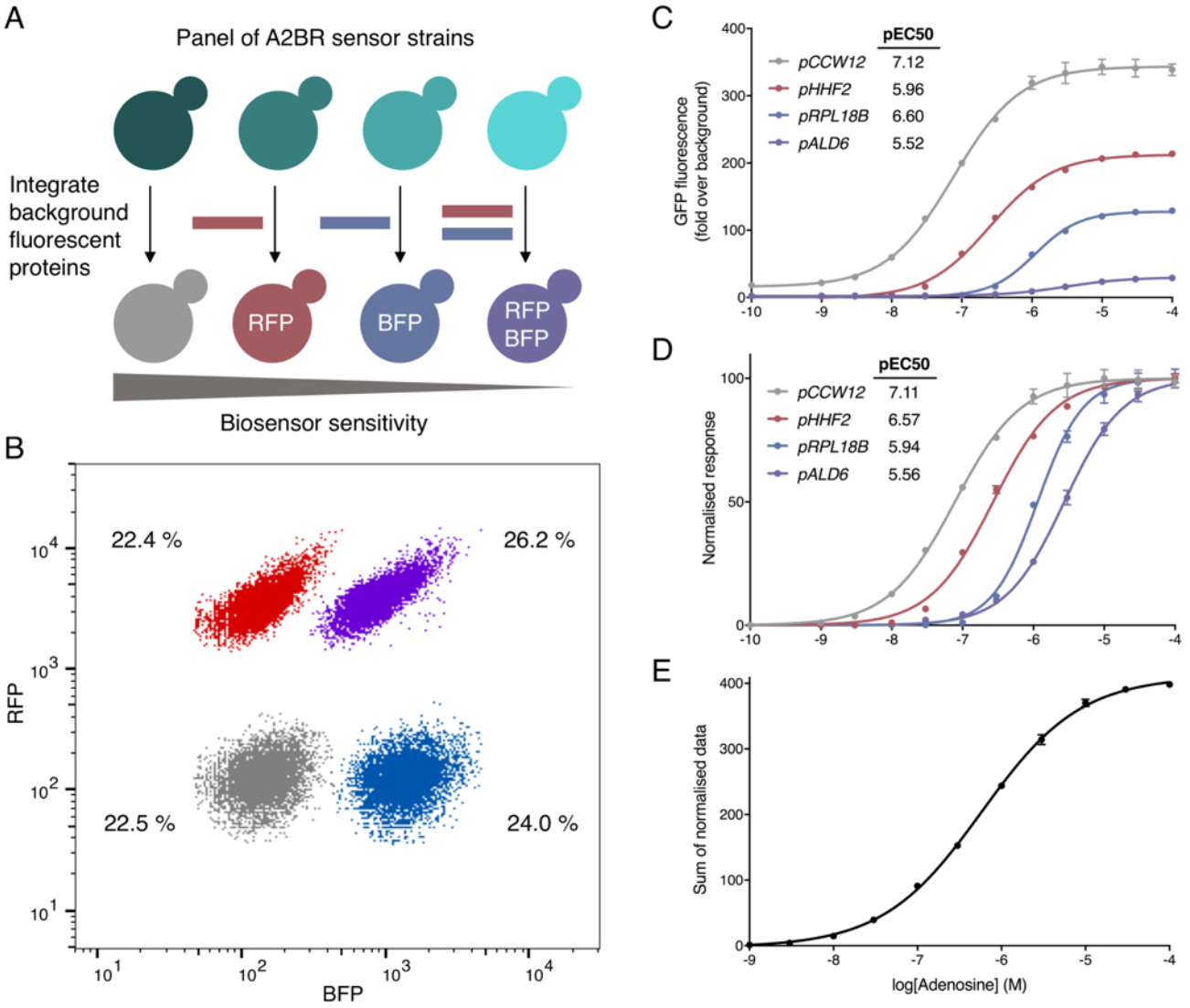
Supplementary Figure S3. Multiplexed targeting of dCas9-PRD to the genomic *ALD6* promoter. All combinations of g1-8 which did not overlap were sampled in all combinations up to the simultaneous expression of 3 gRNAs (*A*) ON/OFF response of the mating pathway targeted to multiple locations on the genomic *ALD6* promoter via the dCas9 sTF. (*B*) Maximum fold change in sfGFP expression from multiplexed targeting of the genomic *ALD6* promoter after stimulation with saturating α-factor concentrations. Results are means ± standard deviation from triplicate isolates.



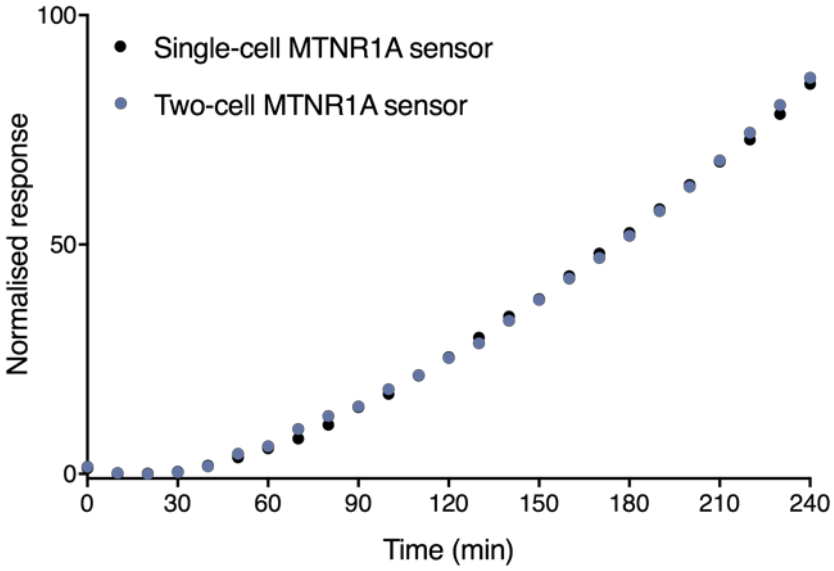
Supplementary Figure S4. Activating the mating pathway using a Ste4-2A-Ste18 bicistron.
To create the Gβγ dimer as a single transcript, we employed the 2A "self-cleaving" peptide sequence between the CDS of Ste4 and Ste18 to create a bicistron that would be "cleaved" during translation to create equimolar concentrations of the two subunits[333]. (*A*) Overview of the Ste4-2A-Ste18 bicistron demonstrating ribosome skipping at the 2A peptide to generate individual Ste4 and Ste18 subunits in equimolar ratios. After the Ste4 and Ste18 subunits have expressed and dimerised, excess Gβγ in the system initiates pathway activity, bypassing the receptor. (*B*) The Ste4-2A-Ste18 bicistron under the control of the β-estradiol inducible Z3EV system. (*C*) β-estradiol dose-response curve of the pathway output in the absence of a receptor agonist demonstrating the Ste4-2A-Ste18 bicistron bypasses receptor activity. For information of the Z3E transactivation system see **Supplementary Figure S5**. Results are means ± standard deviation from triplicate isolates.



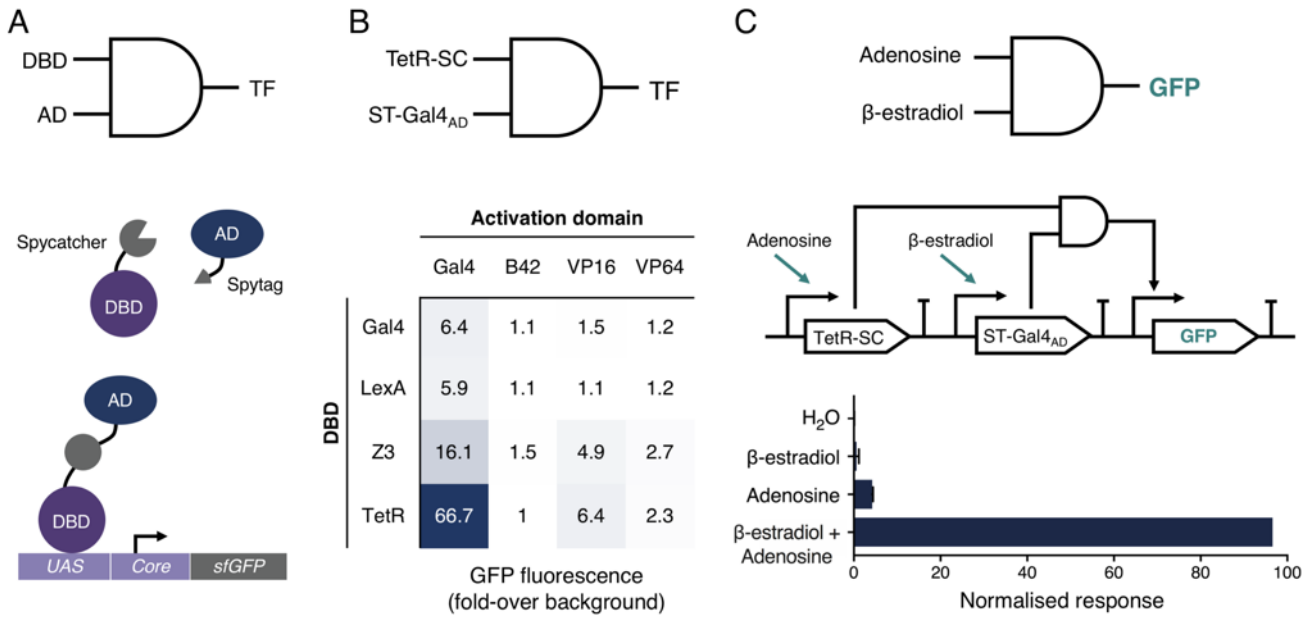
Supplementary Figure S5. Characterising the inducible TetA- and Z3E-based transcription factors.
(*A*) Fusing a library of transcriptional activation domains to the aTc and β-estradiol inducible DNA binding domains, TetA[78] and Z3E[290], for inducible gene regulation. (*B*) Characterising the ON/OFF response of the two transactivation systems with the activation domains, $Gal4_{AD}$, $B42_{AD}$, $VP16_{AD}$, and $VP64_{AD}$. (*C*) Maximum x-fold change in signal after activation for the TetA and Z3E systems in combination with the different activation domains. The combination of TetA with the $Gal4_{AD}$ and Z3E with $VP16_{AD}$ demonstrated the greatest dynamic ranges and so were used for future uses involving these two systems. Results are means ± standard deviation from triplicate isolates.



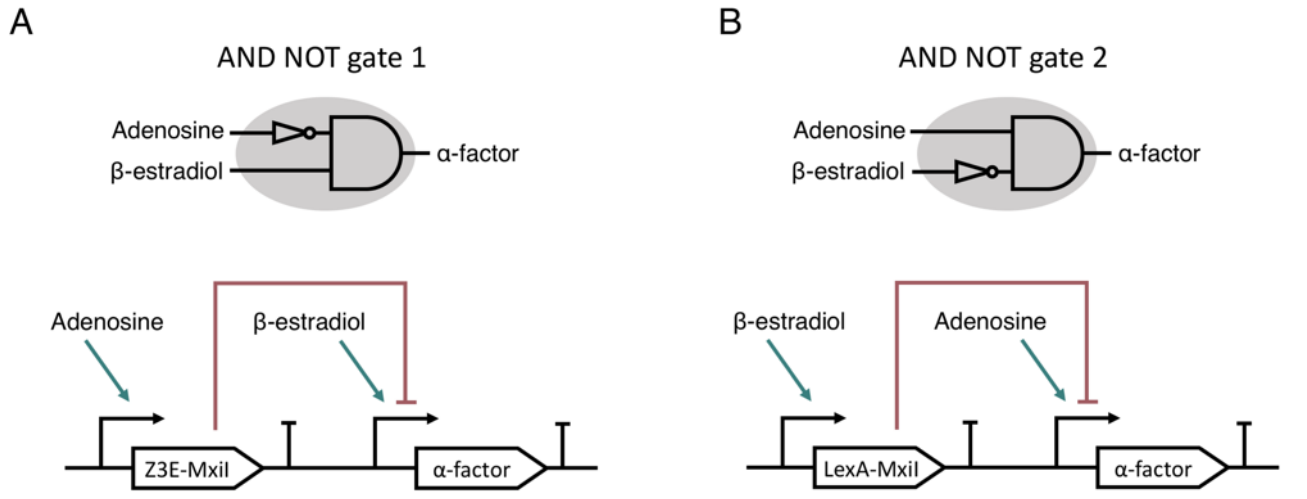
Supplementary Figure S6. Computational addition of flow cytometry data to linearise the adenosine sensing.
As an alternative strategy to experimentally normalise the dose-response outputs of the A2BR sensors strains, we integrated four A2BR sensors with different combinations of the mRuby2 (RFP) and mTagBFP2 (BFP) fluorescent proteins so they could be separated from a mixed population using flow cytometry. The curves could then be computationally normalised and added together. This strategy allows for the use of multiple strains in the same system without needing to normalise the output. Furthermore, the difference between the strains can be greater as there in no limitation from the reporter tuning. This allowed us to use an additional strain with a very low sensitivity and maximum output and thus extend the operational range of the mixed population even further. (*A*) Strategy for separating four individual A2BR sensor strains with different sensitivities to adenosine using combinations of two background fluorescent proteins, mRuby2 and mTagBFP2. (*B*) Gating the four A2BR strains by their unique combinations of red and blue fluorescence in 1:1:1:1 ratio of the mixed cells. (*C*) Individual adenosine dose-response curves of the 4 A2BR sensor strains from a mixed population of cells. (*D*) Computationally normalising the adenosine dose-response curves. (*E*) Addition of the four normalised curves to produce a single adenosine dose-response curve with an extended operational range. Fluorescence data was collected from 40,000 viable cells. Results are means ± standard deviation from triplicate isolates.



Supplementary Figure S7. Time course measurements from the single-cell and two-cell MTNR1A biosensor.
To demonstrate the dynamics of cell-cell communication, we induced the single-cell and two-cell MTNR1A sensors with saturating concentrations of melatonin and measured the GFP fluorescence over time. This releveled no difference between the final output of both systems, demonstrating peptide communication is incredibly fast (i.e. quicker than sfGFP protein maturation). Within 30 minutes after inducing the system, the second cell is seen to be responding. (Data) Normalised GFP output from single-cell (**Figure 55A**) and two-cell (**Figure 55B**) MTNR1A biosensor populations after stimulation with saturating levels of melatonin (100 µM). Results are means from triplicate isolates.



Supplementary Figure S8. AND gate development to support the half adder in Figure 58.
(*A*) AND gate logic using split domain transcription factors utilising the SpyTag/SpyCatcher system. Expression of a DBD-SpyCatcher domain and a SpyTag-AD results in the formation of a complete transcription factor which can be targeted to synthetic promoter driving the expression of sfGFP. (*B*) Constitutive expression of four DBD-SpyCatcher and four SpyTag-AD domains in all combinations targeted to synthetic transcription factors appropriate for each individual DBD driving the expression of GFP. Results are means from triplicate isolate. The TetR-SpyCatcher in combination with the SpyTag-Gal4$_{AD}$ demonstrated the greatest signal and was therefore used for all future AND gate logic. (*C*) Inducible expression of the TetR-SpyCatcher and the SpyTag-Gal4$_{AD}$ by adenosine and β-estradiol, respectively, resulting in the expressing of GFP only in the presence of both inducers. Results are means ± standard deviation from triplicate isolates.



Supplementary Figure S9. Design of the AND NOT gate in the half adder in Figure 58.
(*A*) AND NOT gate 1 responds only in the presence of β-estradiol, producing a secondary transcription factor in the presence of adenosine that competes for the inducible promoter of α-factor, repressing its expression using the MxiI repressor domain[291]. (*B*) AND NOT gate 2 responds only in the presence of adenosine, producing a secondary transcription factor in the presence of β-estradiol that competes for the inducible promoter of α-factor, repressing its expression using the MxiI repressor domain.

8.3 Updates to the YTK system

8.3.1 Introduction

The Yeast MoClo Toolkit (YTK) system is a versatile engineering platform for yeast, consisting of a rapid, modular assembly method and a basic set of highly-characterised parts[33]. The platform provides a framework for creating new genetic designs for reprogramming yeast, including data on promoters, terminators, and other genetic parts, that help to inform such designs. Furthermore, basic CRISPR/Cas9 tools were included within the toolkit to facilitate efficient yeast genome engineering. Indeed, the popularity of this toolkit is reflected by its position as the most read publication on ACS synthetic biology (as of September 2018) and its adaptation by other research groups for engineering organisms as similar as *Pichia pastoris*[334], and as diverse as the bacteria from the gut microbiome of bees[335]. For these reasons, the YTK system has been chosen for all creating all genetic constructs throughout this entire project.

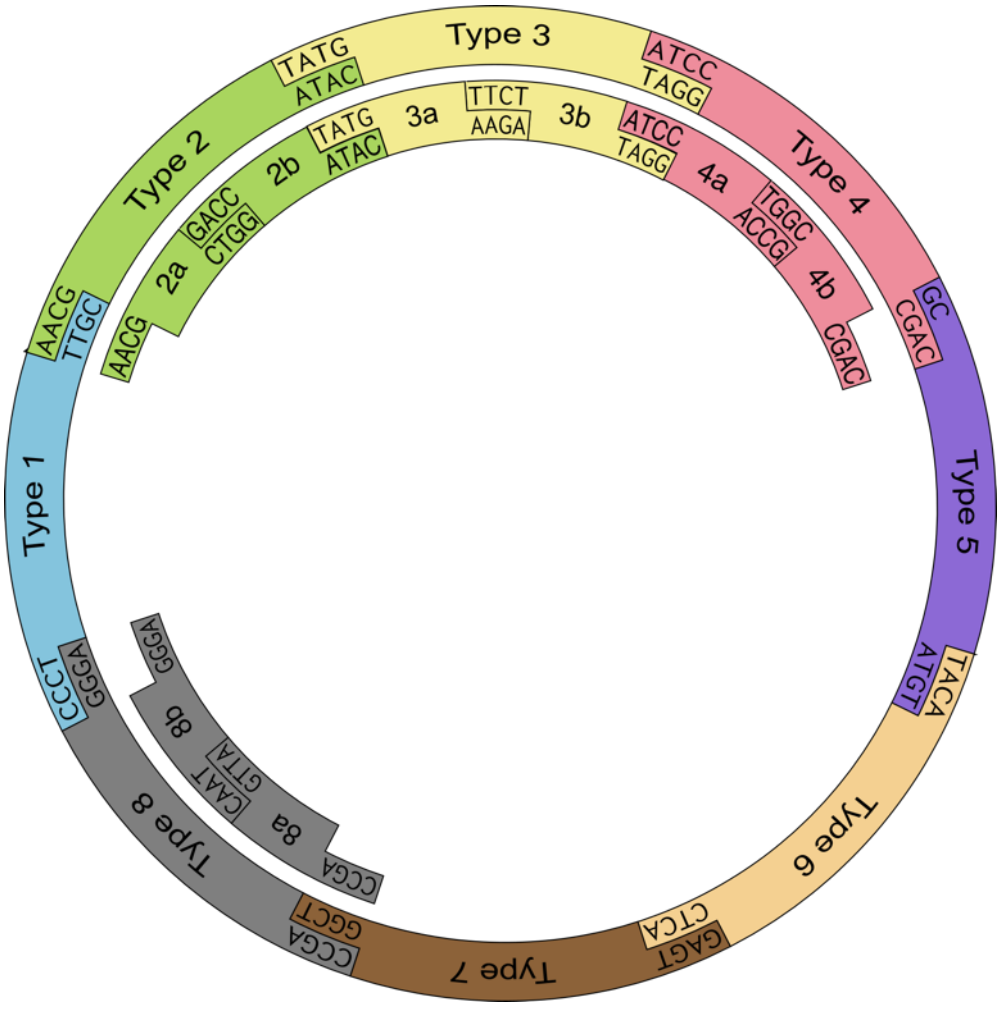
Although the YTK system is an extensive and incredibly flexible platform, there are limitations we have encountered over the course of this work or general aspects we believe could be improved. Accordingly, we have implemented design changes within the YTK system to address these, including the introduction of a new part subtype for split promoter usage, a quick and easy gap repair method for assembling multigene plasmids *in vivo***,** and the integration with a powerful CRISPR toolkit presented following this section. The implemented changes do not affect the formatting of the original system or the hierarchical assembly workflow but simply add additional features.

All possible combinations of the assembly cassettes and yeast vectors have also been pre-assembled to reduce the number of parts assembled at all stages of the hierarchical workflow, which we have found to drastically improve the efficiency and time taken to set up individual Golden Gate reactions. Included in the library of assembly cassettes is the addition of spacer cassettes, which act as a blank position within a multigene plasmid to provide flexibility when creating new multigene plasmids from pre-existing cassettes.

The following section is intended as an amendment to the original publication, and so should be read in parallel. To simplify this section for the user, the style used in the following figures has been adapted from those in the original publication. For a full description of the YTK system see Lee *et al.*[33]. Plasmids presented here and in section **8.4** were under development during the entire course of this work and represent the final versions of a series of iterative improvements. Validation of plasmids presented in both sections is their use in the results chapters, and wherever appropriate will reference the main text.

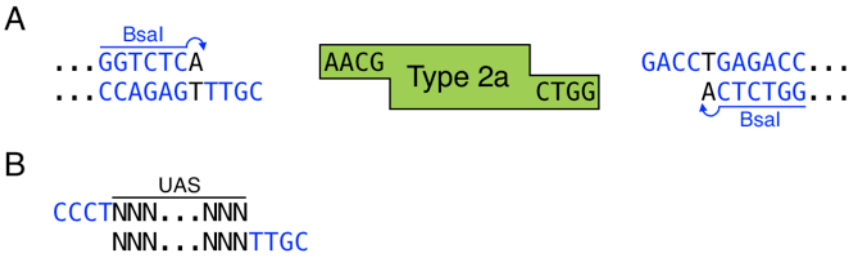
8.3.2 Update to definition of part types

To allow for the modular use of promoters with a different upstream activating sequence (UAS) and core promoter identity (as used in Chapter 4), the type 2 part in the YTK standard has been split into two subtypes to create an additional type 2a and 2b part (**Supplementary Figure S10**). The overhang chosen between the type 2a and 2b parts, 'GACC', was checked for orthogonality and compatibility to all other overhangs in the YTK assembly standard, and through extensive usage has been found to be very robust.

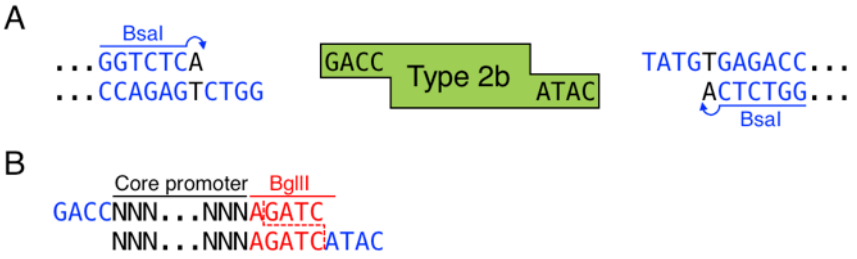


Supplementary Figure S10. Definition of YTK part types.
There are eight primary part types in the YTK assembly standard and four of those have options to split into subtypes. Part types are technically defined only by their flanking overhangs, and the contents need not necessarily match the biologically defined functions described by Lee *et al.*[33]. Here the type 2 part has been divided into two subtypes to enable the use of split domain promoters. Figure and legend adapted from Lee *et al.*[33].

8.3.2.1 Split domain promoters



Supplementary Figure S11. Type 2a: Upstream activating sequence (UAS).
Type 2 parts can be split into 2a and 2b for greater flexibility when creating promoters, allowing the user to combine different upstream activating sequences and core promoter identities. Type 2a parts are flanked by AACG and GACC and typically comprises the UAS region, containing different numbers of DNA binding domains for various DNA binding proteins. Figure and legend adapted from Lee *et al.*[33].



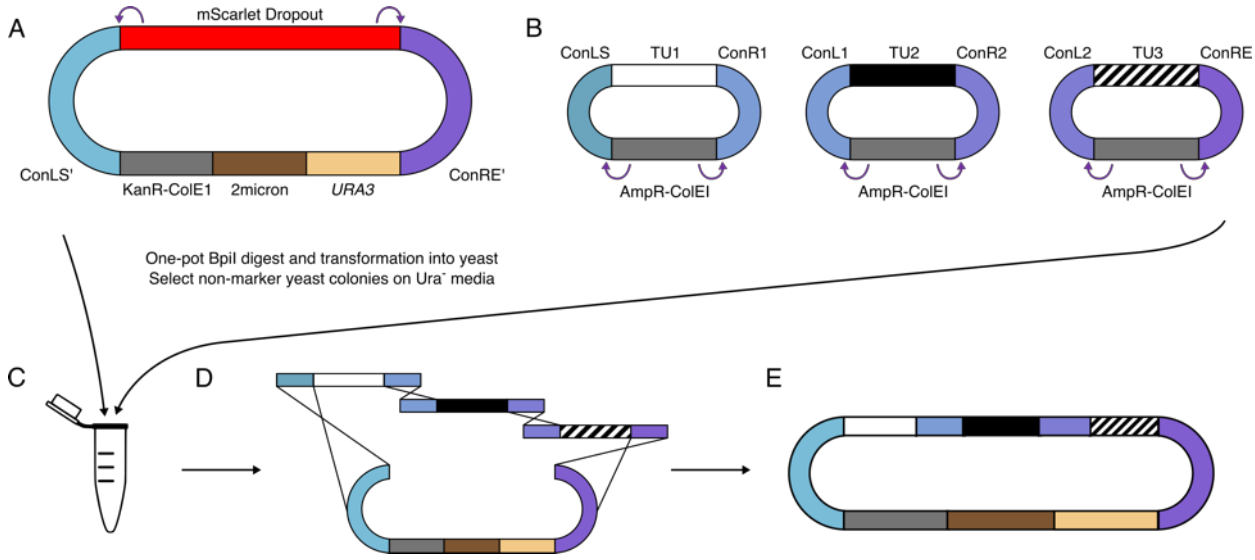
Supplementary Figure S12. Type 2b: Core promoter.
Type 2b parts are flanked by GACC and ATAC and typically comprise the core promoter region. The downstream overhang doubles as the start codon for the subsequent type 3 or 3a coding sequence. Additionally, all the promoters in the YTK system have a BglII site immediately preceding the start codon (overlapping the downstream overhang) for BglBrick compatibility. Figure and legend adapted from Lee *et al.*[33].

8.3.3 Yeast gap repair assembly of multigene plasmids

Although Golden Gate is the preferred method of assembly in the YTK system, sometimes it is more desirable to use *in vivo* yeast gap repair assembly. To facilitate gap repair assembly of multigene plasmids in the YTK system, 143 bp sequences were included in the assembly connectors, where ConLX and ConRX sequences with the same value of X are identical. Amplification of pre-assembled yeast cassettes outwards from the inside of the Con sequences and amplification inwards from the outside of the Con sequences generates linear fragments with terminal homology. These fragments can then be assembled *in vivo* using the highly efficient homology-directed repair mechanisms of *S. cerevisiae* to create a fully-assembled multigene plasmid, bypassing the additional Golden Gate reaction that would otherwise be necessary.

While convenient, the use of PCR to generate linear DNA fragments for assembly in yeast adds an unnecessary step that is also likely to introduce mutations from replication error. To avoid this, we introduced recognition sites for the type IIs restriction enzyme, BpiI, at the terminal ends of the Con sequences to facilitate the linearization of the fragments by digest. BpiI was selected due to its absence from all YTK toolkit plasmids and the recommendation in the publication to avoid its use when creating new parts.

In this way, the yeast vector, containing a BpiI dropout with a visible yeast marker (such as mScarlet), can be combined with the cassettes, digested in a one-pot reaction with BpiI, and transformed directly into yeast, where it will then be assembled by gap repair, analogous to the PCR-based method (**Supplementary Figure S13**). Yeast colonies can then be screened by the loss of the visible marker.



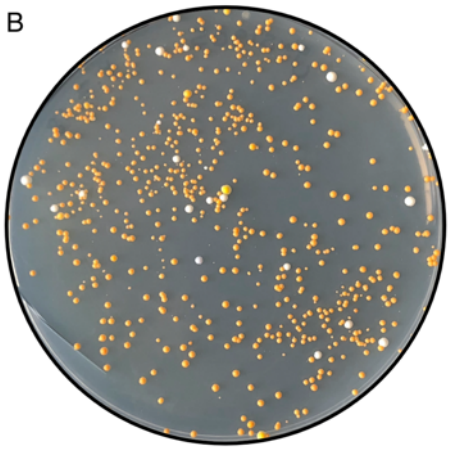
Supplementary Figure S13. Gap repair assembly of multigene plasmids in the updated YTK system.
(*A*) A pre-assembled yeast low- or high-copy vector with BpiI restriction sites flanking a mScarlet dropout. (*B*) Cassette plasmids with BpiI restriction sites flanking the ConL and ConR sequences. (*C*) One-pot digestion of the pre-assembled vector and cassette plasmids and direct transformation into yeast. (*D*) Yeast gap repair assembly of the BpiI-generated fragments, with their order determined by flanking homology of the ConL and ConR sequences, where ConLX and ConRX sequences with the same value of X are identical. (*F*) Fully assembled multigene plasmid, identical in sequence to a gap repair assembled multigene plasmid using the original YTK system. Note the final sequence of the multigene plasmid will be different to a BsmBI Golden Gate assembled multigene plasmid.

To demonstrate the performance of this new gap repair method, the β-carotene biosynthesis pathway, encoded by four genes, was assembled into transcriptional units (TUs) using the new BpiI assembly cassettes. The TUs were then combined with a *URA3* high-copy vector containing an AmilCP BpiI dropout marker, digested with BpiI, and transformed directly into yeast (**Supplementary Figure S14**). The majority of the resulting colonies exhibited a homogenous orange colour, indicating > 95% rate of correct assembly.

A

Plasmid Name	Cassette Type	Cassette Position 2	3	4	E. coli Marker	Yeast Marker	Yeast Vector
pWS1742	S/1	*pTEF2*	*crtE*	*tADH1*	AmpR		
pWS1743	1/2	*pTEF1*	*crtI*	*tPGK1*	AmpR		
pWS1744	2/3	*pTDH3*	*crtYB*	*tENO2*	AmpR		
pWS1745	3/E	*pPGK1*	*tHMG1*	*tTDH1*	AmpR		
pWS1746	O	AmilCP BpiI Dropout Marker			KanR	URA3	High-copy

B

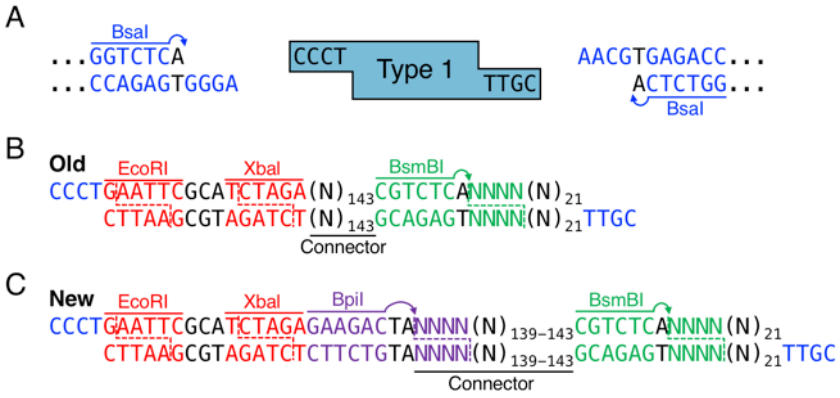


Supplementary Figure S14. Gap repair assembly of the β-carotene biosynthesis pathway.
(*A*) Transcriptional units encoding the β-carotene pathway genes (*CrtE*, *CrtI*, *CrtYB*, and *tHMG1*) and a *URA3* high-copy plasmid containing the AmilCP BpiI dropout marker. (*B*) *S. cerevisiae* colonies expressing the assembled pathways to produce an orange colour on medium lacking uracil. White colonies indicate either the misassembly of the multigene plasmid or random insertion into the genome. Blue colonies represent undigested yeast vector expressing the AmilCP dropout marker.

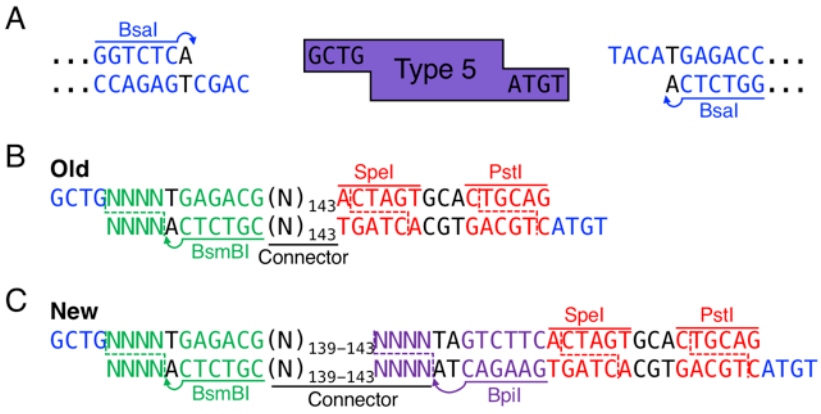
8.3.3.1 BpiI compatible sequences

8.3.3.1.1 Assembly connectors

To allow TUs to be excised from their backbone for gap repair assembly, BpiI sequences were added to terminally flank the 143 bp connector sequences of the 5' and 3' assembly connectors.



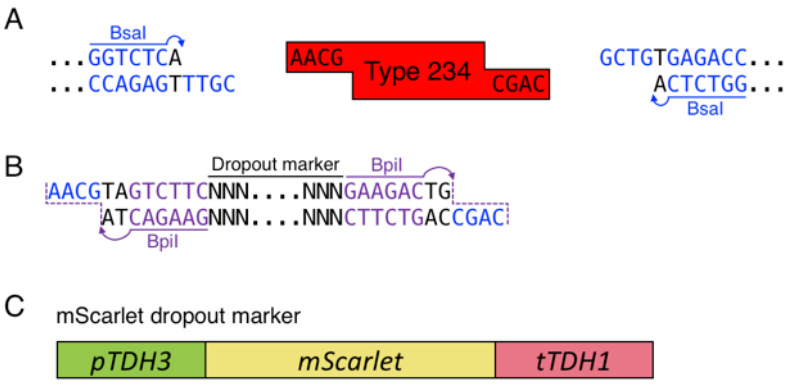
Supplementary Figure S15. Redesign of the Type 1: 5' assembly connector.
(*A*) Type 1 Parts are flanked by CCCT and AACG and contain non-coding, non-regulatory sequences that are used to direct assembly of multigene plasmids. (*B*) The sequence of the old 5' assembly connectors, containing a 143 bp ConL sequence, a BsmBI recognition site and unique overhang, and a 21 bp barcode scar. Type 1 parts also include an EcoRI and XbaI site for BioBrick compatibility of the assembled cassettes and multigene plasmids. (*C*) The sequence of the updated 5' assembly connectors, similar to the old sequences, but with the addition of a BpiI recognition site 5' of the ConL sequence, orientated to cut into the connector without leaving residual sequence. Figure and legend adapted from Lee *et al.*[33].



Supplementary Figure S16. Redesign of the Type 5: 3' assembly connector.
(*A*) Type 5 parts are flanked by GCTG and ATGT and contain non-coding, non-regulatory sequences that are used to direct the assembly of multigene plasmids. (*B*) The sequence of the old 3' assembly connectors, containing a 143 bp ConR sequence and a BsmBI recognition site and unique overhang. Type 5 parts also include a SpiI and PstI site for BioBrick compatibility of the assembled cassettes and multigene plasmids. (*C*) The sequence of the updated 3' assembly connectors, identical in all ways to the old sequence, except with the addition of a BpiI recognition site 3' of the ConR sequence, orientated to cut into the connector without leaving residual sequence. Figure and legend adapted from Lee *et al.*[33].

8.3.3.1.2 Yeast dropout marker

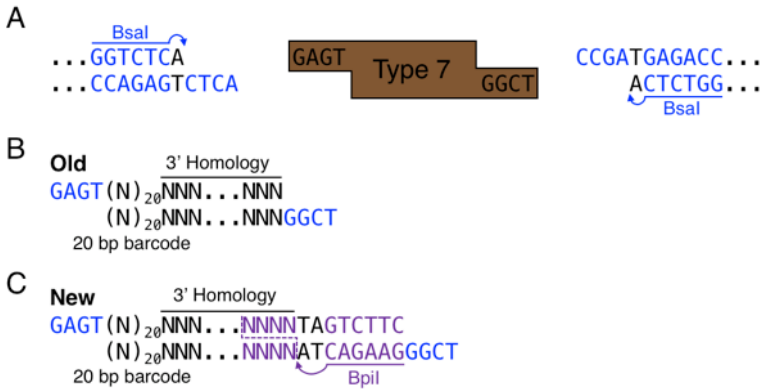
To select multigene plasmids that have been correctly assembled by yeast gap repair, a new part has been created to indicate whether or not the backbone vector was digested properly – the yeast dropout marker.



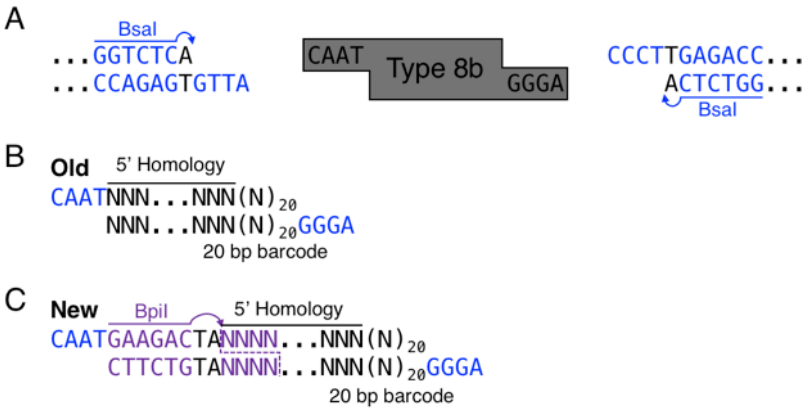
Supplementary Figure S17. Type 234: Yeast dropout marker.
(*A*) Type 234 Parts are flanked by AACG and CGAC and, in this instance, contain a yeast expression cassette containing a visible marker. (*B*) The sequence of the yeast dropout marker, containing a marker expression cassette flanked by BpiI recognition sites directed outwards. The BpiI recognition sites are designed so that upon assembly into a yeast vector, the old BsaI overhangs become the new BpiI overhangs, and so allow the marker expression cassette to be re-excised from the vector backbone in a BpiI digest. (*C*) An example of a dropout marker, comprising the yeast *TDH3* promoter, *mScarlet* ORF, and *TDH1* terminator.

8.3.3.1.3 Yeast chromosomal integration

For integration of plasmids into the yeast chromosome, the YTK system includes type 7 and 8b parts containing homology to a specific locus in the yeast genome. The type 8a part then contains NotI sites, so that when fully assembled, plasmids can be linearised to create a DNA fragment flanked by these homology arms, directing integration in the chromosome by homologous recombination. To convert the system from NotI to BpiI (for compatibility with the CRISPR toolkit and create consistency with yeast pre-transformation prep) BpiI sites were included on the 3' of the 3' homology arms and 5' of the 5' homology arms.



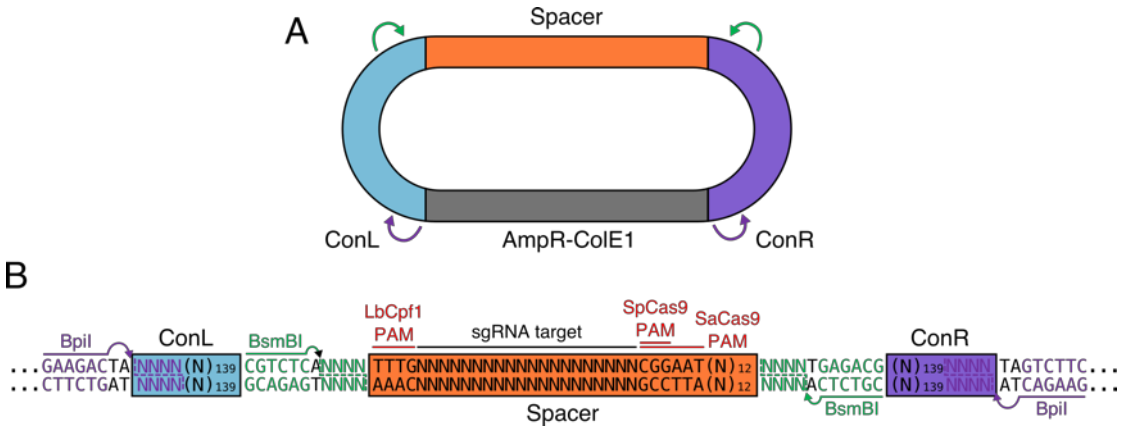
Supplementary Figure S18. Redesign of Type 7: 3' homology.
Type 7 parts, flanked by GAGT and CCGA, contain sequences of homology that is downstream (3') of the target locus. (*B*) Sequence of the old 3' homology containing a 20 bp barcode sequence for validation correct integration of the plasmid and the 3' homology, typically 500 bp in length. (*C*) Sequence of the redesigned 3' homology arm with the addition of a BpiI recognition site, orientated to cut into the homology region. Figure and legend adapted from Lee *et al.*[33].



Supplementary Figure S19. Redesign Type 8b: 5' homology.
Type 8b parts, flanked by CAAT and GGGA, contain sequences of homology that is upstream (5') of the target locus. (*B*) Sequence of the old 3' homology containing a 20 bp barcode sequence for validation correct integration of the plasmid and the 5' homology, typically 500 bp in length. (*C*) Sequence of the redesigned 5' homology with the additional a BpiI recognition sequence, orientated to cut into the homology region. Figure and legend adapted from Lee *et al.*[33].

8.3.4 Spacer cassettes

To increase the flexibility of multigene plasmid assemblies from pre-made cassettes, ten spacer cassettes have been created to substitute for any particular TU position within a multigene plasmid. (**Supplementary Figure S20** and **Supplementary Table S3**). The spacer sequences themselves consist of 42 bp containing a unique sgRNA target for each TU position. The sgRNA targets have been designed to have a high on-target score (SpCas9) and perfect off-target score, using Benchling's CRISPR wizard, to facilitate further editing of a multigene cassette in vivo if required.



Supplementary Figure S20. Spacer cassettes.
(*A*) A spacer cassette consisting of a short spacer sequence flanked by ConL and ConR with BsmBI and BpiI recognition sites on the inside and outside of the Con sequences, respectively, both pairs facing inwards. (*B*) Sequence of the spacer cassette. The 42 bp spacer sequence (orange box) consists of a 20 bp sgRNA target flanked by the LbCpf1 PAM at the 5' and the SpCas9 and SaCas9 PAM at the 3'.

Supplementary Table S3. Spacer sequence sgRNA targets.

Spacer	sgRNA target	On-target score	Off-target score
Spacer 1	CCGATAATTGCAGACGAACG	64	100
Spacer 2	ATGTCAACACAGCTACAACG	77	100
Spacer 3	ACTGGCTTAAGATGACAACG	87	100
Spacer 4	GGGCACAGACAACCTAAACG	88	100
Spacer 5	CCCATGAGTCACAATGAACG	75	100
Spacer 6	AAGCTCCACACAGTCGAACG	61	100

These spacer cassettes, flanked by the ConL and ConR, can be assembled into a multigene plasmid via the traditional BsmBI Golden Gate or by the new *in vivo* gap repair method demonstrated in this work. They are recommended whenever a new multigene plasmid design is made using pre-existing cassettes which are not able to fill all required positions. Instead of remaking the cassettes from scratch, it may be possible to complete the multigene plasmid using these spacer cassettes to substitute a TU position.

8.3.5 Assembly and spacer cassette definitions

Supplementary Table S4. Assembly cassette definitions.

	Plasmid Name	Cassette Type	Cassette position 1	2	3	4	5	6
Assembly Cassettes	pWS1431	S/1	TU1					
	pWS1432	1/E		TU2				
	pWS1433	1/2		TU2				
	pWS1434	2/E			TU3			
	pWS1435	2/3			TU3			
	pWS1436	3/E				TU4		
	pWS1437	3/4				TU4		
	pWS1438	4/E					TU5	
	pWS1439	4/5					TU5	
	pWS1440	5/E						TU6
Spacer Cassettes	pWS2011	S/1	Spacer 1					
	pWS2012	1/E		Spacer 2				
	pWS2013	1/2		Spacer 2				
	pWS2014	2/E			Spacer 3			
	pWS2015	2/3			Spacer 3			
	pWS2016	3/E				Spacer 4		
	pWS2017	3/4				Spacer 4		
	pWS2018	4/E					Spacer 5	
	pWS2019	4/5					Spacer 5	
	pWS2020	5/E						Spacer 6

Key

- Assembly cassette
- Spacer cassette
- Open position
- End of multigene plasmid

8.3.6 New YTK toolkit plasmids

Supplementary Table S5. List of updated toolkit plasmids.

Plasmid Name	Part Type	Part Description	*E. coli* Marker
pWS1415	1	ConLS (BpiI)	CamR
pWS1416	1	ConL1 (BpiI)	CamR
pWS1417	1	ConL2 (BpiI)	CamR
pWS1418	1	ConL3 (BpiI)	CamR
pWS1419	1	ConL4 (BpiI)	CamR
pWS1420	1	ConL5 (BpiI)	CamR
pWS1546	234	sfGFP BpiI Dropout Marker	CamR
pWS1547	234	mScarlet BpiI Dropout Marker	CamR
pWS1548	234	AmilCP BpiI Dropout Marker	CamR
pWS1421	5	ConR1 (BpiI)	CamR
pWS1422	5	ConR2 (BpiI)	CamR
pWS1423	5	ConR3 (BpiI)	CamR
pWS1424	5	ConR4 (BpiI)	CamR
pWS1425	5	ConR5 (BpiI)	CamR
pWS1426	5	ConRE (BpiI)	CamR
pWS1208	7	URA3 3’ Homology (BpiI)	CamR
pWS1209	7	LEU2 3’ Homology (BpiI)	CamR
pWS1451	7	HO 3’ Homology (BpiI)	CamR
pWS1212	8b	URA3 5’ Homology (BpiI)	CamR
pWS1213	8b	LEU2 5’ Homology (BpiI)	CamR
pWS1452	8b	HO 5’ Homology (BpiI)	CamR

Supplementary Table S6. Pre-assembled yeast vectors and assembly cassettes created in this study.

Plasmid Name	Cassette Type	Part Description	*E. coli* Marker
Yeast vectors			
pWS1287	-	**URA3** Integration Vector	**KanR**
pWS035	o	**URA3** Low-copy Vector	**KanR**
pWS038	O	**URA3** High-copy Vector	**KanR**
pWS1288	-	**LEU2** Integration Vector	**KanR**
pWS036	o	**LEU2** Low-copy Vector	**KanR**
pWS039	O	**LEU2** High-copy Vector	**KanR**
pWS1624	-	**HIS3** Integration at HO Vector	**KanR**
pWS037	o	**HIS3** Low-copy Vector	**KanR**
pWS040	O	**HIS3** High-copy Vector	**KanR**
pWS2084	o	**KanR** Low-copy vector	**KanR**
pWS2088	O	**KanR** High-copy vector	**KanR**
pWS2085	o	**NatR** Low-copy vector	**KanR**
pWS2089	O	**NatR** High-copy vector	**KanR**
pWS2086	o	**HygR** Low-copy vector	**KanR**
pWS2090	O	**HygR** High-copy vector	**KanR**
pWS2087	o	**ZeoR** Low-copy vector	**KanR**
pWS2091	O	**ZeoR** High-copy vector	**KanR**
Yeast control integration vectors			
pWS1295	-	**URA3** Empty Control Integration Vector	**KanR**
pWS1296	-	**LEU2** Empty Control Integration Vector	**KanR**
pWS1698	-	**HIS3** Empty Control Integration at HO Vector	**KanR**
Assembly cassettes			
pWS1431	S/1	**S/1** Assembly Cassette	**AmpR**
pWS1432	1/E	**1/E** Assembly Cassette	**AmpR**
pWS1433	1/2	**1/2** Assembly Cassette	**AmpR**
pWS1434	2/E	**2/E** Assembly Cassette	**AmpR**
pWS1435	2/3	**2/3** Assembly Cassette	**AmpR**
pWS1436	3/E	**3/E** Assembly Cassette	**AmpR**
pWS1437	3/4	**3/4** Assembly Cassette	**AmpR**
pWS1438	4/E	**4/E** Assembly Cassette	**AmpR**
pWS1439	4/5	**4/5** Assembly Cassette	**AmpR**
pWS1440	5/E	**5/E** Assembly Cassette	**AmpR**
Spacer cassettes			
pWS2011	S/1	**S/1** Spacer Cassette	**AmpR**
pWS2012	1/E	**1/E** Spacer Cassette	**AmpR**
pWS2013	1/2	**1/2** Spacer Cassette	**AmpR**
pWS2014	2/E	**2/E** Spacer Cassette	**AmpR**
pWS2015	2/3	**2/3** Spacer Cassette	**AmpR**
pWS2016	3/E	**3/E** Spacer Cassette	**AmpR**
pWS2017	3/4	**3/4** Spacer Cassette	**AmpR**
pWS2018	4/E	**4/E** Spacer Cassette	**AmpR**
pWS2019	4/5	**4/5** Spacer Cassette	**AmpR**
pWS2020	5/E	**5/E** Spacer Cassette	**AmpR**

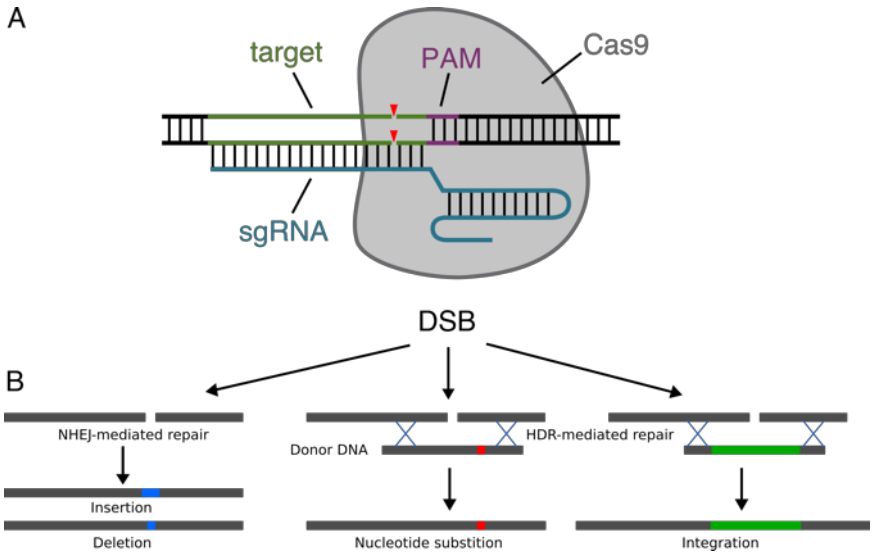
Cassette type definitions: Yeast vector cassette types, integration (-), low-copy (o), high-copy (O). Assembly and spacer cassettes types, (ConLX/ConRX).

8.4 CRISPR toolkit

8.4.1 Introduction

Saccharomyces cerevisiae has long been the most tractable organism for eukaryotic cell biology, owing to its genetic malleability, greatly facilitated by a preference for homologous recombination (HR) over non-homologous end joining (NHEJ) for double-stranded break (DSB) repair. Over the years, biologists have taken advantage of this preference, allowing for the site-specific installation of genetic material and genomic edits with base-pair precision[185].

However, the demands on yeast engineering have significantly increased with the design of more complex systems and extensive metabolic pathways. Genetic techniques that have historically relied on the recycling of markers are proving too cumbersome to keep up with the ambitions of synthetic biologists. Seemingly, the advent of CRISPR/Cas9 in yeast, as it was for many other organisms, was incredibly timely. The introduction of targeted DSBs, by RNA programmable Cas9, boosted the rate of homology-directed repair (HDR) with linear DNA at the DSB locus by several orders of magnitude[336,337] (**Supplementary Figure S21**). This new capability made it practically feasible for yeast synthetic biologists to rapidly edit the yeast genome without the need for a marker, saving significant amounts of time and labour.



Supplementary Figure S21. Overview of CRISPR/Cas-mediated genome editing.
(*A*) The sgRNA/Cas9 in complex with target DNA, creating a blunt end DSB 3 bp upstream of the NGG PAM sequence. (*B*) Possible DNA repair mechanisms following the DSB. Non-homologous end joining (NHEJ) is the method of repair in the absence of a template. This usually results in the insertion or deletion of base pairs from the DNA sequence (indel). Homology-directed repair (HDR) is the favoured repair mechanism in yeast when a template is present. By introducing donor DNA into the cells alongside the CRISPR/Cas9 machinery, it is possible to create small edits, integrate new DNA, or delete large sections of DNA. It is this highly-efficient pathway in yeast that makes CRISPR/Cas9 such a powerful engineering tool.

From the beginning, yeast CRISPR/Cas9-mediated genome engineering was highly-efficient at targeting a single locus, with near 100% donor DNA recombination efficiency being achieved in the first reported system by DiCarlo *et al.*[337]. Quickly, the focus turned towards multiplexing edits. If it takes up to a week to achieve a single round of editing, being able to multiplex edits would save an entire week for each additional edit performed within a single experiment. It was not long before the first multiplexed edits were achieved in yeast, although with a drastic drop in efficiency with each additional edit[260]. From this point forward, a plethora of new CRISPR/Cas9 platforms were developed to improve the efficiency of performing multiple edits simultaneously.

There is far too much yeast CRISPR literature to possibly cover in this short introduction, so we refer the reader to Jakočiūnas *et al.*[266] for a detailed review on numerous and diverse CRIPSR/Cas9 platforms that emerged between 2013 and 2016, and Alexander[251] for a brief history and the future prospects of genome engineering in *S. cerevisiae*. Instead of an in-depth literature review, in the next section we highlight the contributions to the CRISPR field we believe have had the most significant impact on yeast CRISPR/Cas9-mediated genome engineering. We then go on to present our new CRISPR platform that builds upon these advancements and integrates them with the YTK system to produce a highly flexible and efficient toolkit for yeast engineering.

8.4.2 Advances in CRISPR-mediated engineering in *Saccharomyces cerevisiae*

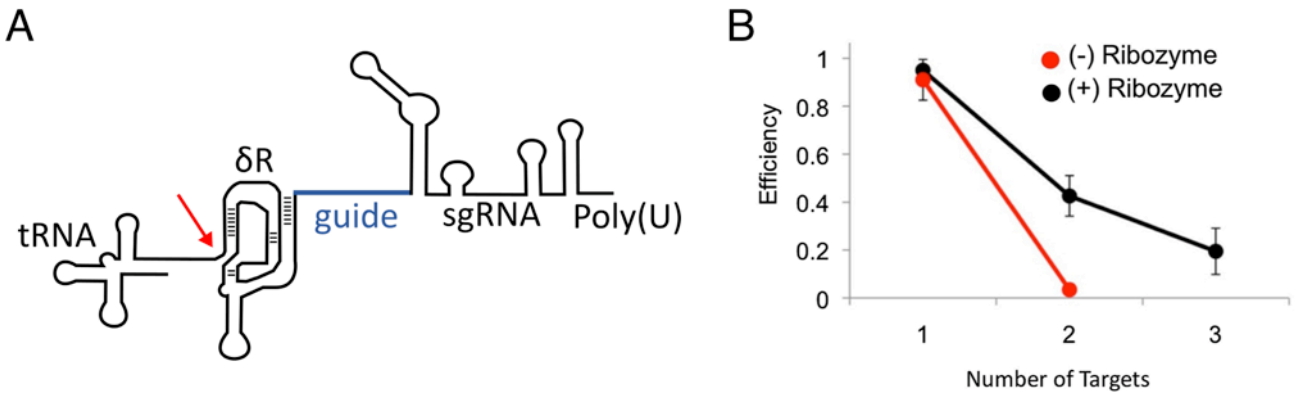
8.4.2.1 sgRNA expression using ribozymes

One of the biggest challenges to multiplexing in any organism is the expression of the Cas9 protein and mature sgRNA transcripts. Low levels of either of these components often fails to elicit the desired change, with cellular levels of the sgRNAs correlating with the efficiency of CRISPR/Cas9-mediated engineering, particularly when the on-target effectiveness of the sgRNA is also low[338,339]. The expression of Cas9 was an easy problem to solve in yeast due to the tools available to molecular biologists. The initial DiCarlo paper cloned a mammalian codon optimised Cas9 ORF, C-terminally tagged with the SV40 nuclear localisation signal, downstream of a strong/medium strength constitutive promoter[337]. This has been one of the most prevalent strategies[340–344], with a few other studies using either the native *S. pyogenes* codon bias or yeast codon optimised Cas9[260,340,345,346].

The expression of mature sgRNA transcripts was less trivial as the optimal transcription and correct folding of this component is of paramount importance to ensure correct interaction with Cas9[347]. sgRNA expression was initially achieved using the RNA pol III *SNR52* promoter and the 3' flanking sequence of the *SUP4* gene as a terminator[337]. Although many other studies have used this sgRNA

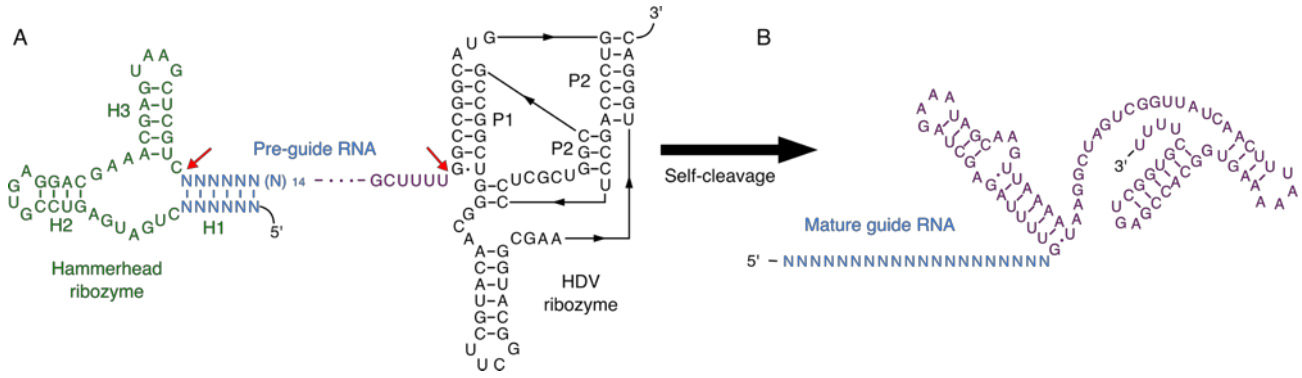
strategy[341–343,345,346], being restricted to one particular promoter was an obvious limitation, as low abundances of intracellular sgRNAs were reported[260] with no alternatives available.

Ryan *et al.*[260] showed the advantages of fusing the sgRNA to the 3' end of the self-cleaving hepatitis delta virus (HDV) ribozyme. This structured domain cleaves RNA immediately 5' of the ribozyme, removing extraneous sequences to produce a mature sgRNA, providing flexibility over the choice of promoter. This allowed the fusion of the sgRNA to the 3' of a tRNA and, therefore, take advantage of a strong RNA polymerase III tRNA promoter (**Supplementary Figure S22A**). This strategy boosted the efficiency of CRISPR/Cas9-mediated genome engineering, allowing the authors to achieve multiplexed edits (**Supplementary Figure S22B**).



Supplementary Figure S22. The ribozyme strategy from Ryan *et al.*[260].
(*A*) The mature sgRNA contains a 5′ hepatitis delta virus (HDV) ribozyme (δR), 20mer target sequence (guide), sgRNA and RNA polymerase III terminator. The RNA polymerase III promoter tRNA is catalytically removed by the HDV ribozyme (red arrow). B) Efficiency of multiplex insertion of DNA in diploid yeast cells with the 5′ HDV ribozyme (black) and without the 5′ HDV ribozyme (red). Triplex targeting without the 5′ HDV ribozyme was not tested. The $tRNA^{Tyr}$ promoter was used in these experiments. (Figure and legend taken from Ryan *et al.*[260] under the terms of the Creative Commons Attribution License).

An alternative strategy by Gao and Zhao[348] demonstrated a more elegant solution to the problem, by flanking the sgRNA with the Hammerhead (HH) ribozyme at the 5' side and the HDV ribozyme at the 3' side (HH-sgRNA-HDV) (**Supplementary Figure S23**). By programming the first six bp of the HH sequence with the reverse complement of the first 6 bp of the target sequence, it is possible to generate a fully mature sgRNA transcript with no extraneous sequence after self-cleavage. The flanking ribozymes also have the benefit of separating the sgRNA sequence from both 5' and 3' regulatory sequences, providing full flexibility over the choice of promoter and terminator.



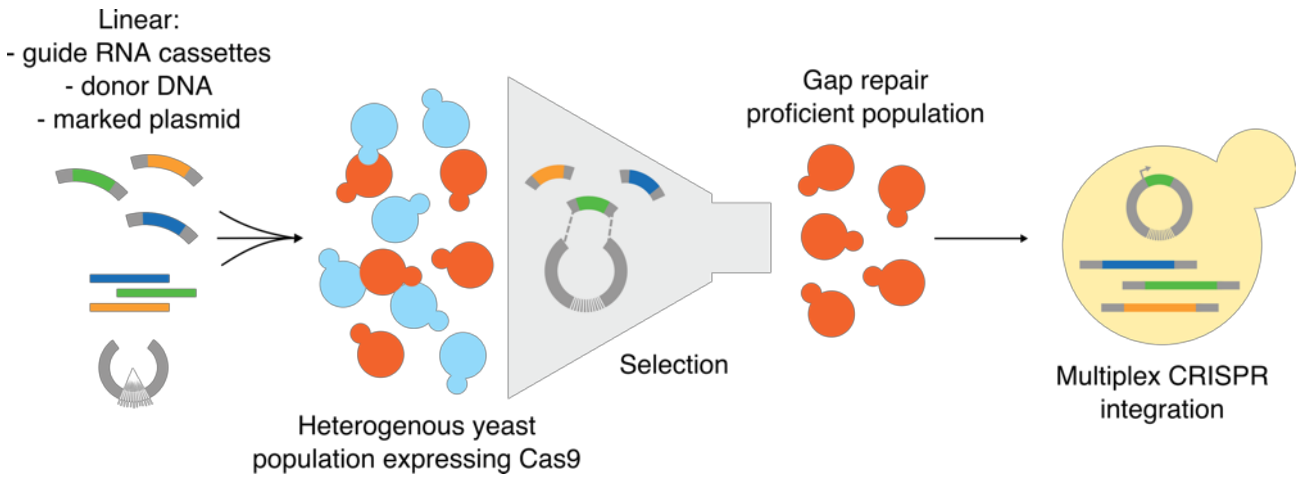
Supplementary Figure S23. The HH-sgRNA-HDV ribozyme strategy from Gao and Zhao[348].
(*A*) Structure of the transcribed HH-sgRNA-HDV sequence. The CRISPR/Cas9 pre-guide RNA consisting of the target sequence (blue) followed by the Cas9 sgRNA scaffold (purple) is directly flanked by the Hammerhead ribozyme (green) at the 5' side and the HDV ribozyme (black) at the 3' side. Note, the first 6 bp of the Hammerhead ribozyme is the reverse compliment of the target sequence. (*B*) Self-cleavage of the HH-sgRNA-HDV transcript to produce a fully-mature guide RNA with no extraneous sequence. Figure adapted from Gao and Zhao[348].

8.4.2.2 Selection of yeast cells proficient in gap repair

The next important development in yeast CRISPR/Cas9-mediated genome engineering was a novel method to select for yeast cells proficient in gap repair, by Horwitz *et al.*[346]. This was achieved by delivering the sgRNAs into yeast cells as linear fragments of DNA that have to be assembled by the yeast HDR machinery to reconstitute the sgRNA expression vector (**Supplementary Figure S24**). The authors proposed that this necessity for the yeast cells to reconstitute a plasmid via gap repair, and so be able to grow on selective media, identifies a special subpopulation of transformed cells more competent in HDR and, therefore, more likely to perform any particular CRISPR edit, via the same mechanisms.

As the competency for homologous recombination is known to vary throughout the cell cycle, when transforming cells there will be a mixture of log-phase cells in all states of the cell cycle. If by selecting for cells which have gap repaired the sgRNA vector, and thus will be the cells most proficient in gap repair, these cells will also be best equipped to accomplish multiplexed integrations. This capability may, therefore, facilitate the selection of cells that have performed extensive markerless, multiplex edits. Typically, engineering rates targeting three loci using this protocol exceed 50%[346,349], allowing the recovery of desired clones with minimal screening.

This method also has the benefit of streamlining the DNA preparation, as sgRNAs are introduced on separate fragments that can be delivered in a combinatorial manner. Furthermore, because this method is independent of the format of sgRNA and endonuclease expression, there is an opportunity to combine this strategy with improvements made elsewhere.

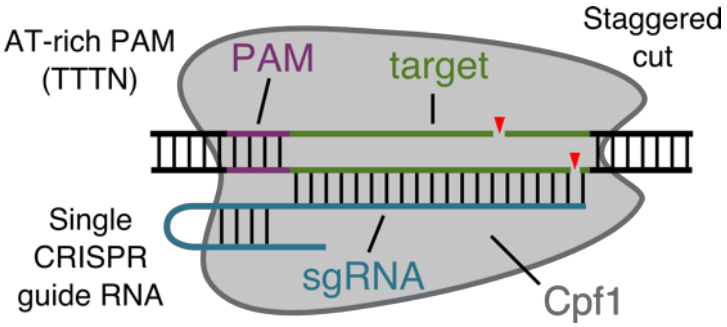


Supplementary Figure S24. Gap repair of the sgRNA expression vector from Horwitz *et al.*[346].
The sgRNA expression vector is introduced into yeast cells constitutively expressing Cas9 in two linearised fragments with homology to each other. Once transformed into yeast, the HDR machinery will gap repair the plasmid, reconstituting the sgRNA expression vector. As the plasmid also contains a marker, only cells that have repaired the plasmid may survive in selective media. This selects for a subpopulation of cells that are most proficient in gap repair, increasing the likelihood that colonies will have integrated the donor DNA that was introduced alongside the linearised sgRNA vector components. Therefore, high efficiency gene editing can be achieved. (Figure adapted from Horwitz *et al.*[346] and Walter *et al.*[349]).

8.4.2.3 Expanding the RNA-guided endonuclease toolkit with Cpf1

The final advance in CRISPR-mediated genome engineering is the introduction of the alternative RNA-guided endonuclease, Cpf1, from *Lachnospiraceae bacterium*[350] (**Supplementary Figure S25**). The addition of Cpf1 to the yeast CRISPR toolbox adds further flexibility, providing an alternative CRISPR/Cas system with similar editing efficiencies to Cas9[351]. Rather than the GC-rich PAM sequence of Cas9, Cpf1 uses the AT-rich TTTN sequence, making Cpf1 more suited to editing AT-rich regions, such as gene regulatory elements[352]. Secondly, Cpf1 cleaves the target at the distal end of the protospacer, rather than near the seed region, as is the case with Cas9. Cpf1 generated indels will, therefore, be located at a distance from the target, possibly preserving the sequence for subsequent rounds of Cpf1 cleavage, whereas indels resulting from Cas9 cleavage are more likely to disrupt the target sequence. This may have the benefit of improving HDR. If the first round of cleavage results in an indel, a subsequent round of targeting could still be repaired by HDR[350].

Finally, in this work, we use dCas9 as a programmable DNA binding domain to target transcriptional regulation via the Ste12 pheromone responsive domain. As this requires the constitutive expression of Cas9 specific guides, the prospect of further engineering of cells using this system is limited due to cross-talk between guides introduced for CRISPR/Cas9 editing and the pre-existing guides for programming dCas9. Cpf1, however, is fully orthogonal to Cas9 in yeast[197], therefore, providing a means for editing yeast cell lines with the latter system already in use.



Supplementary Figure S25. CRISPR/Cpf1 overview.
The sgRNA/Cpf1 in complex with target DNA, creating a 4 bp staggered DSB at the distal end of target sequence from the TTTN PAM. The DNA is repaired via the mechanisms demonstrated in **Supplementary Figure S21B**.

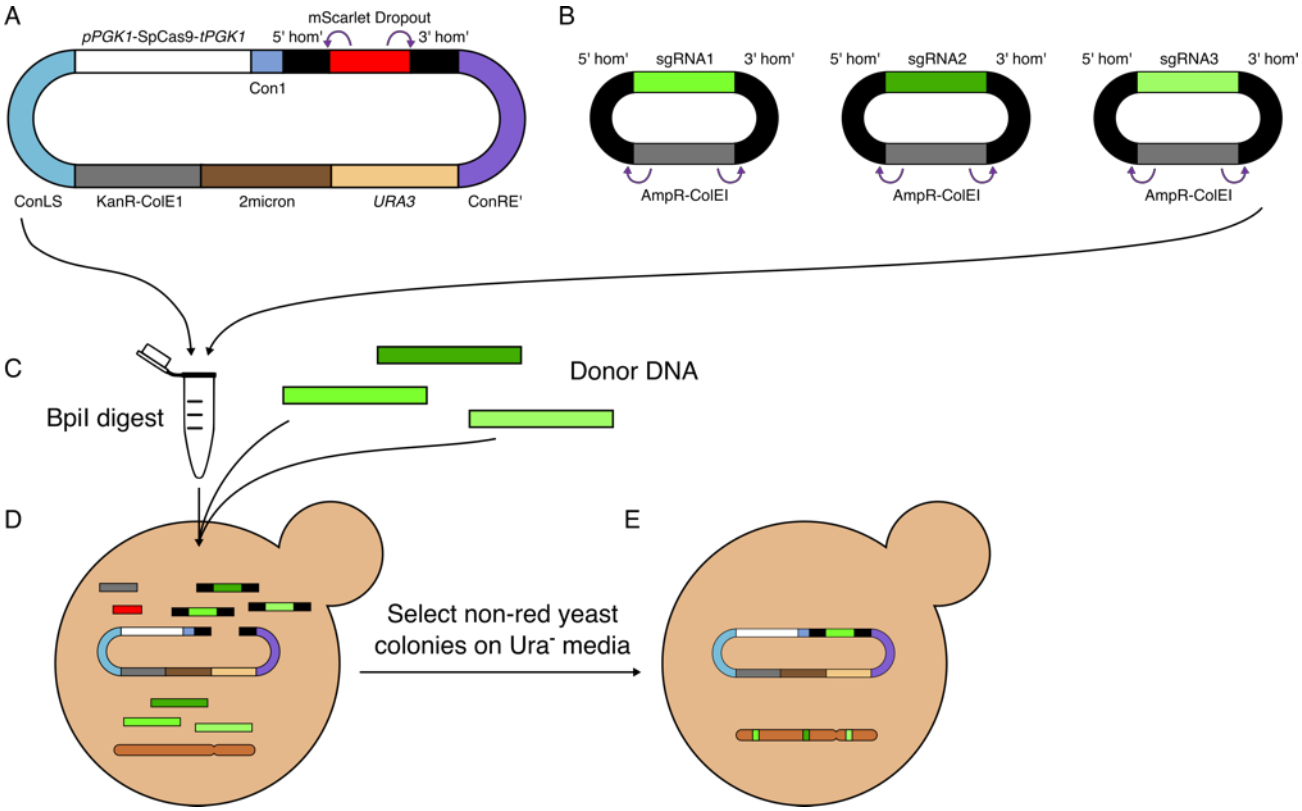
8.4.2.4 Next steps for CRISPR-mediated engineering in *Saccharomyces cerevisiae*

We have presented three of the key advances to yeast CRISPR engineering we believe have been the most significant over the last five years. Individually, these three advances have improved the efficiency of multiplexed editing, streamlined the preparation of DNA materials, and increased the flexibility and versatility of the CRISPR-mediated engineering in yeast. We have reached a point where the efficiency of multiplexing is good enough for routine engineering. The limiting factor is now the preparation of DNA and the complexity of the current toolkits. The next challenge, therefore, is not to improve the efficiency of CRISPR multiplexing, but to streamline the process by abstracting the tools and making it easier for the yeast synthetic biologist to design, set up, and perform CRISPR experiments.

8.4.3 Overview of CRISPR toolkit

Here we present a new CRISPR toolkit, consisting of a minimal set of parts that in combination with the YTK system can be built up to generate an extensive and flexible set for tools for highly-efficient multiplexed editing in *Saccharomyces cerevisiae*. The toolkit brings together the three advances covered in the previous section, along with a number of other properties, including modularity, simplicity, minimal cloning requirements, and integration with the YTK system.

At the highest level, the toolkit consists of two plasmids. The first plasmid (CRISPR/Cas plasmid), comprises the endonuclease, yeast replicon, yeast selection marker, and a yeast dropout marker, flanked by two 500 bp sequences with BpiI recognition sites on the inside facing outwards (**Supplementary Figure S26A**). The second plasmid (sgRNA plasmid) contains a single sgRNA expression unit, flanked by two 500 bp sequences, with exact complementarity to the first plasmid, with BpiI recognition sites at their terminus, facing inwards (**Supplementary Figure S26B**).



Supplementary Figure S26. Overview of CRISPR toolkit.
(*A*) An example CRISPR/Cas plasmid containing Cas9 under medium strength expression, *URA3* selectable marker, 2micron replication element, and mScarlet dropout marker. (*B*) Cas9 sgRNA plasmids containing different target sequences. (*C*) The CRISPR/Cas and sgRNA plasmids are combined in a one-pot reaction and digested with BpiI. (*D*) Donor DNA is added to the linearised plasmids and transformed directly in yeast. (*E*) The CRISPR/Cas plasmid has been gap-repaired with one of the sgRNA fragments to reconstitute a selectable plasmid and the three edits have been performed during the transient expression of the remaining guides. The yeast dropout marker and remaining sgRNA fragments have been lost.

The CRISPR/Cas plasmid can be built up to include any combination of the endonuclease, yeast selection marker, yeast replicon, or yeast dropout marker that exists in the CRISPR toolkit or YTK system. The sgRNA plasmid either contains a Cas9 or Cpf1 sgRNA expression unit. Once created the CRISPR/Cas plasmid can be used for all future engineering experiments using those particular conditions (**Supplementary Figure S29D**). New guides will have to be designed for each new target using a simple one-step assembly protocol from a pre-assembled sgRNA entry vector (**Supplementary Figure S29F**).

The CRISPR/Cas plasmid can be combined with any number of sgRNA plasmids (reasonably up to six) and then digested by BpiI in a one-pot reaction (**Supplementary Figure S29C**). Following digestion and heat kill, donor DNA is added, and the resulting mixture is directly transformed into yeast (**Supplementary Figure S26D**). The CRISPR machinery then performs the edits, targeted by the sgRNAs and repaired by the donor DNA, and the linearised CRISPR/Cas plasmid is repaired with

one of the linear sgRNA fragments to reconstitute a selectable vector (**Supplementary Figure S26E**). Yeast colonies lacking the dropout marker can then be selected for genotyping.

8.4.3.1 Integration with the YTK system

The CRISPR toolkit has been designed for full integration with the updated YTK system. Having BpiI as the restriction enzyme that prepares both the CRISPR plasmids and integration vectors for yeast transformation allows the two systems to be used together in a one-pot pre-transformation digest.

This can be used to markerlessly integrate plasmids (a yeast marker spacer was designed to replace the marker on an integration vector) or more commonly, CRISPR can be used to target the genome between the YTK integration homology arms to increase the efficiency of marker plasmid integration, allowing multiple plasmids to be integrated simultaneously. See **Figure 22** for a demonstration of CRISPR-aided integration of marker plasmids and section **8.4.6** for a detailed description.

8.4.3.2 SpCas9 and LbCpf1 expression

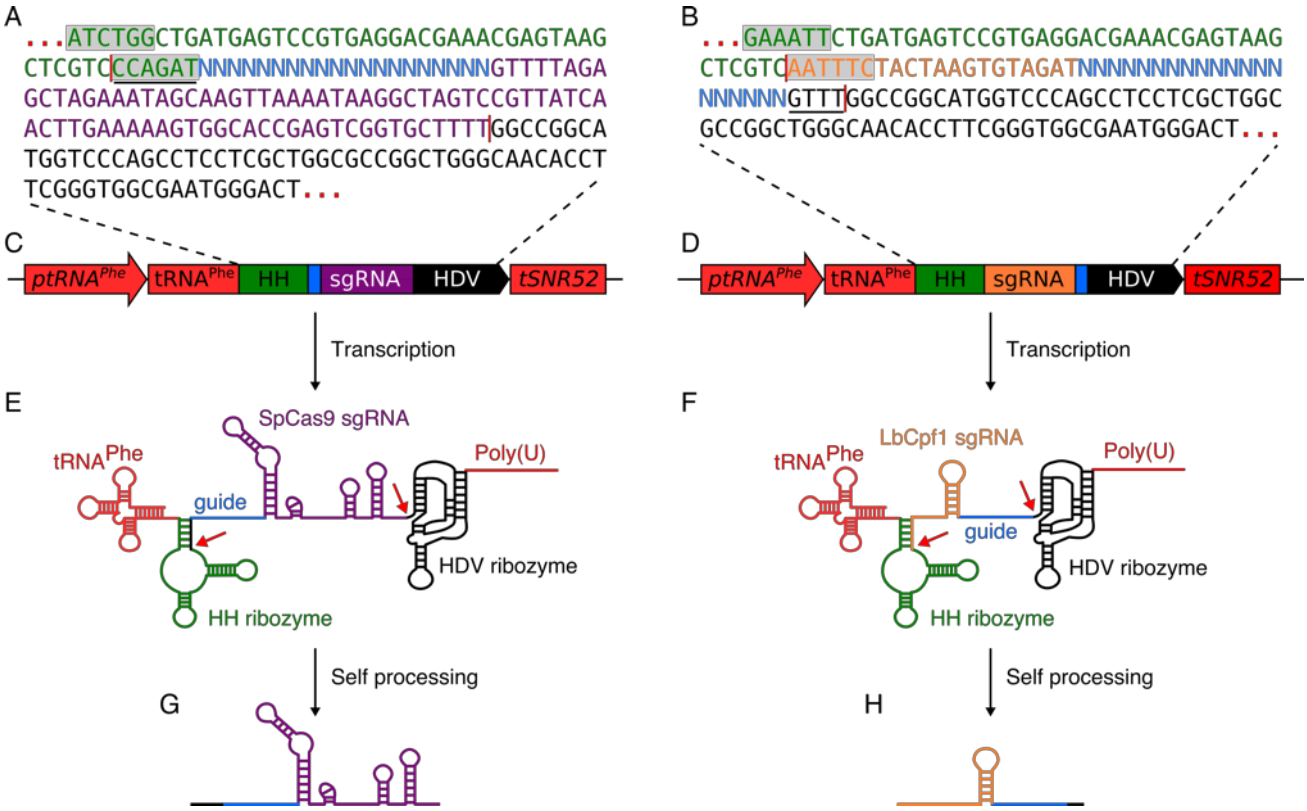
The YTK system already provided SpCas9 within the starting set of parts. For Cpf1, the variant from *Lachnospiraceae bacterium* has repeatedly been shown to be the better performing endonuclease[197,351], and so we synthesised a yeast codon optimised version with an SV40 nuclear localisation signal at both termini, as suggested by Lian *et al.*[197].

Overexpressing CRISPR/Cas endonucleases has shown to be toxic to the cell[260]. To avoid toxicity issues, moderate SpCas9 and LbCpf1 expression using the *PGK1* promoter and terminator was chosen for all CRIPSR experiments, as suggested by Ryan *et al.*[260].

8.4.3.3 SpCas9 and LbCpf1 sgRNA expression

8.4.3.3.1 SpCas9 and LbCpf1 sgRNAs structure

To generate an efficient and modular sgRNA structure that would work for both SpCas9 and LbCpf1 sgRNAs, we combined the Ryan *et al.* 5' tRNA structure with the Gao and Zhao HH-sgRNA-HDV format (**Supplementary Figure S27**). This allows for a fully mature sgRNA while still using a highly active RNA pol III promoter, enabling continuous high level production of the transcript[353].

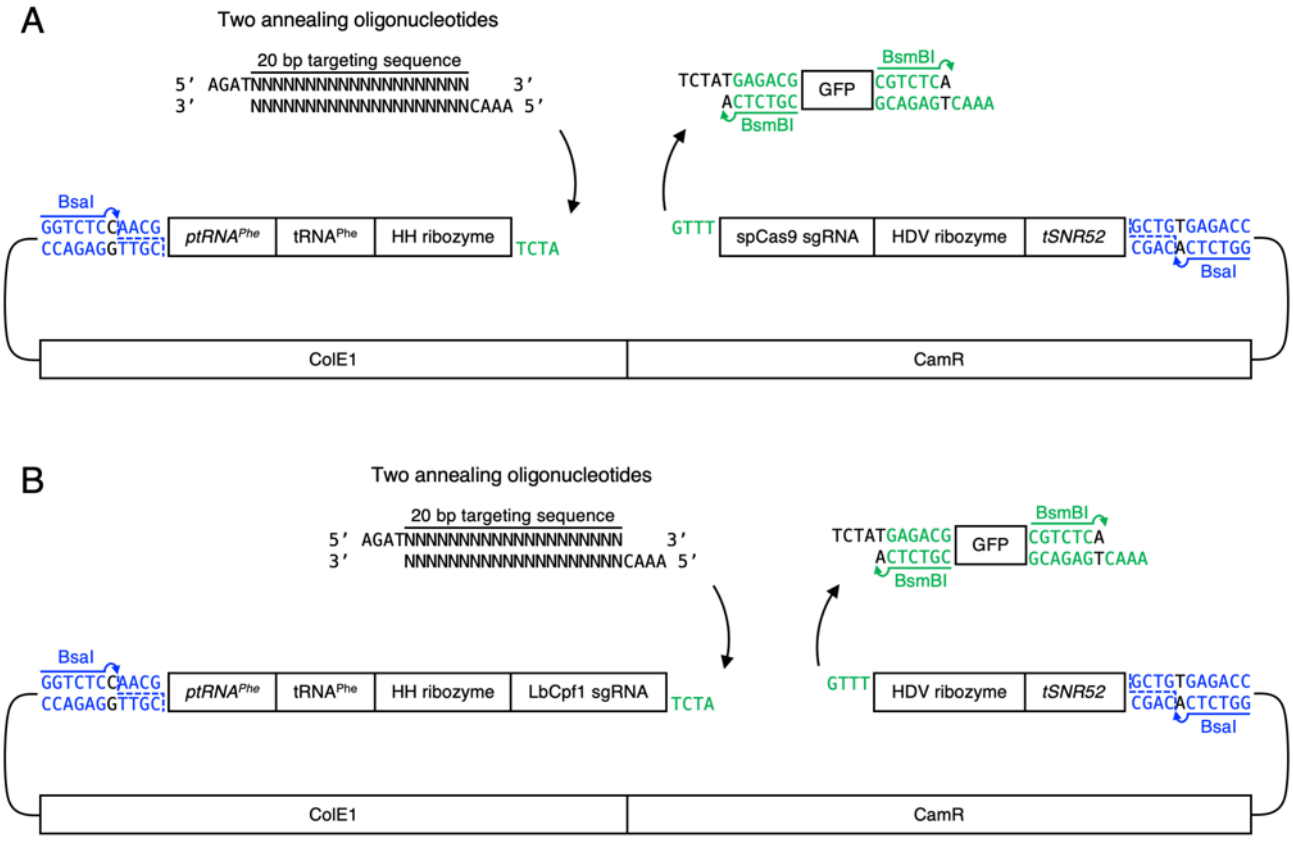


Supplementary Figure S27. SpCas9 and LbCpf1 sgRNA structure.
(*A+B*) Sequence of the CRISPR/SpCas9 and CRISPR/LbCpf1 sgRNA, respectively, between the *tRNA*^{Phe} and *SNR52* terminator, showing the Hammerhead ribozyme (HH) (green), 20 bp target sequence (blue), SpCas9 sgRNA scaffold (purple), LbCpf1 sgRNA scaffold (orange), and HDV ribozyme (HDV) (black). Grey highlights indicate the first and last 6 bp of the Hammerhead ribozyme with reverse complementarity, the vertical red lines indicate the cleavage sites of the two ribozymes, and the horizontal black lines indicate additional sequences that remain on the mature sgRNAs. (*C+D*) Production of the tRNA-HH-sgRNA-HDV constructs from the *tRNA*^{Phe} promoter. (*E+F*)

The resulting sgRNA architecture consists of the *tRNA*^{Phe} promoter driving the expression of tRNA^{Phe} (taken from the sgRNA dropout in the YTK system), followed by a Hammerhead ribozyme, sgRNA, HDV ribozyme, and then the *SNR52* terminator (also from the YTK sgRNA dropout). The last 4 bp of the LbCpf1 sgRNA scaffold was used within the HH hairpin of the SpCas9 sgRNA, and the last 4 bp of the SpCas9 sgRNA scaffold was used between the LbCpf1 sgRNA and HDV to allow consistent assembly overhangs between the two sgRNA entry vectors and prevent continual redesign of the 5' HH sequence. This results in a short but unobtrusive additional sequence on the 5' and 3' of the mature SpCas9 and LbCpf1 sgRNA, respectively.

8.4.3.3.2 Construction of SaCas9 and LbCpf1 sgRNAs

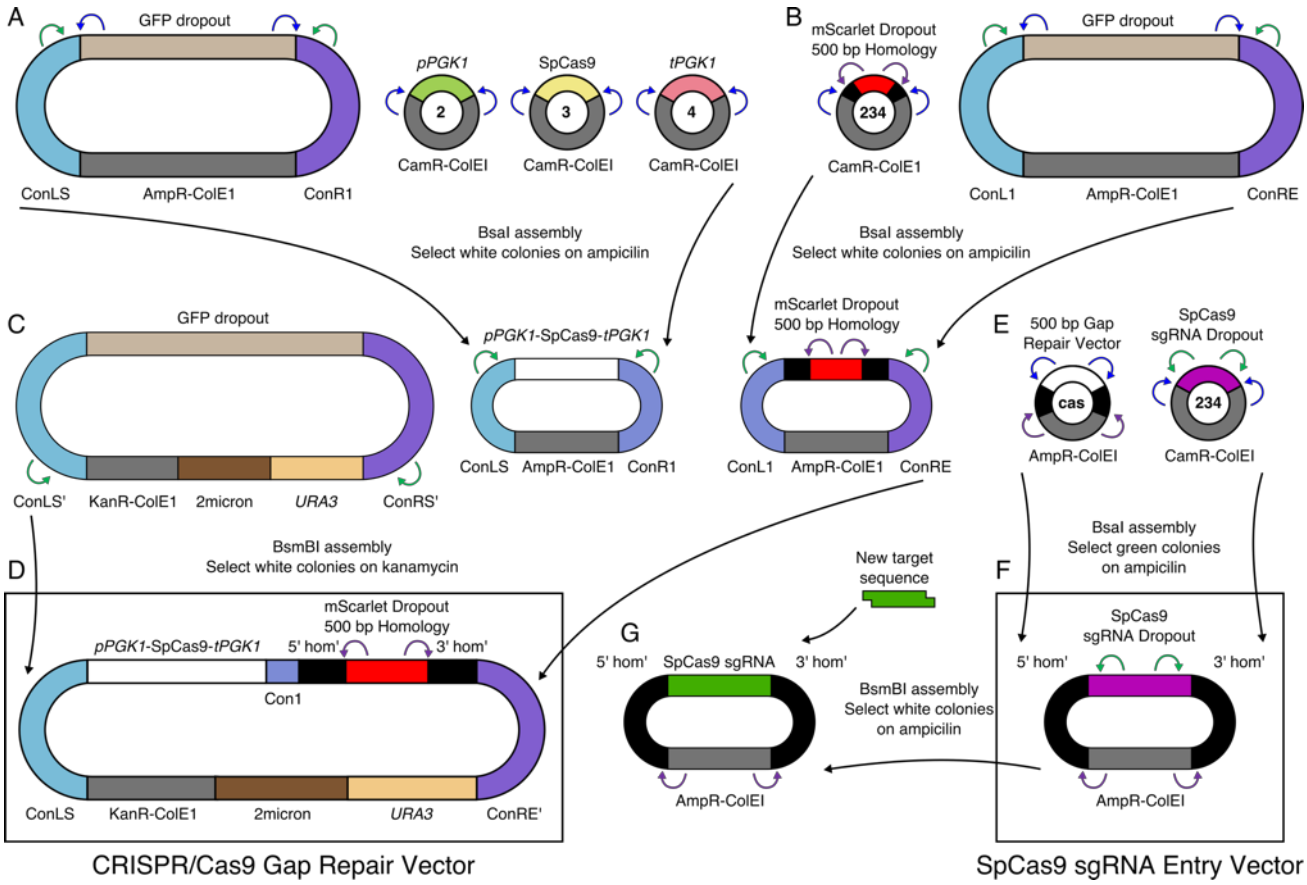
To construct sgRNAs to guide SpCas9 or LbCpf1 to the genome, the toolkit includes two type 234 sgRNA dropouts, which act as an entry vector for new targets (**Supplementary Figure S28**). These vectors contain the entire sgRNA structure with the targeting sequence substituted for a BsmBI-flanked GFP expression cassette. The GFP dropout cassette can then be replaced by annealing oligonucleotides using the overhang TCTA at the 5' end and GTTT at the 3' to complete the full sgRNA sequence, using the protocol in **7.2.3.4**. Once assembled, the vector can then be transferred to any assembly cassette for use in the YTK hierarchical workflow, or into the 500 bp Gap Repair Vector for use in the CRISPR toolkit system. Alternatively, the Type 234 sgRNA Entry Vectors can be pre-assembled into the 500 bp gap repair vector to create the sgRNA entry vector, in which oligos can then be assembled for immediate use with the CRISPR system (**Supplementary Figure S29F**).



Supplementary Figure S28. Construction of SpCas9 and LbCpf1 sgRNAs.
These vectors are based on the gRNA architecture described in **Supplementary Figure S27**. For the dropout, the targeting sequence has been replaced by a BsmBI-flanked GFP expression cassette. To keep the system simple, the overhangs "TCTA" and "GTTT" are used at the 5' and 3' of both SpCas9 and LbCpf1 sgRNA dropouts. (*A*) Type 234 SpCas9 sgRNA Dropout. (*B*) Type 234 LbCpf1 sgRNA dropout.

8.4.4 CRISPR toolkit assembly

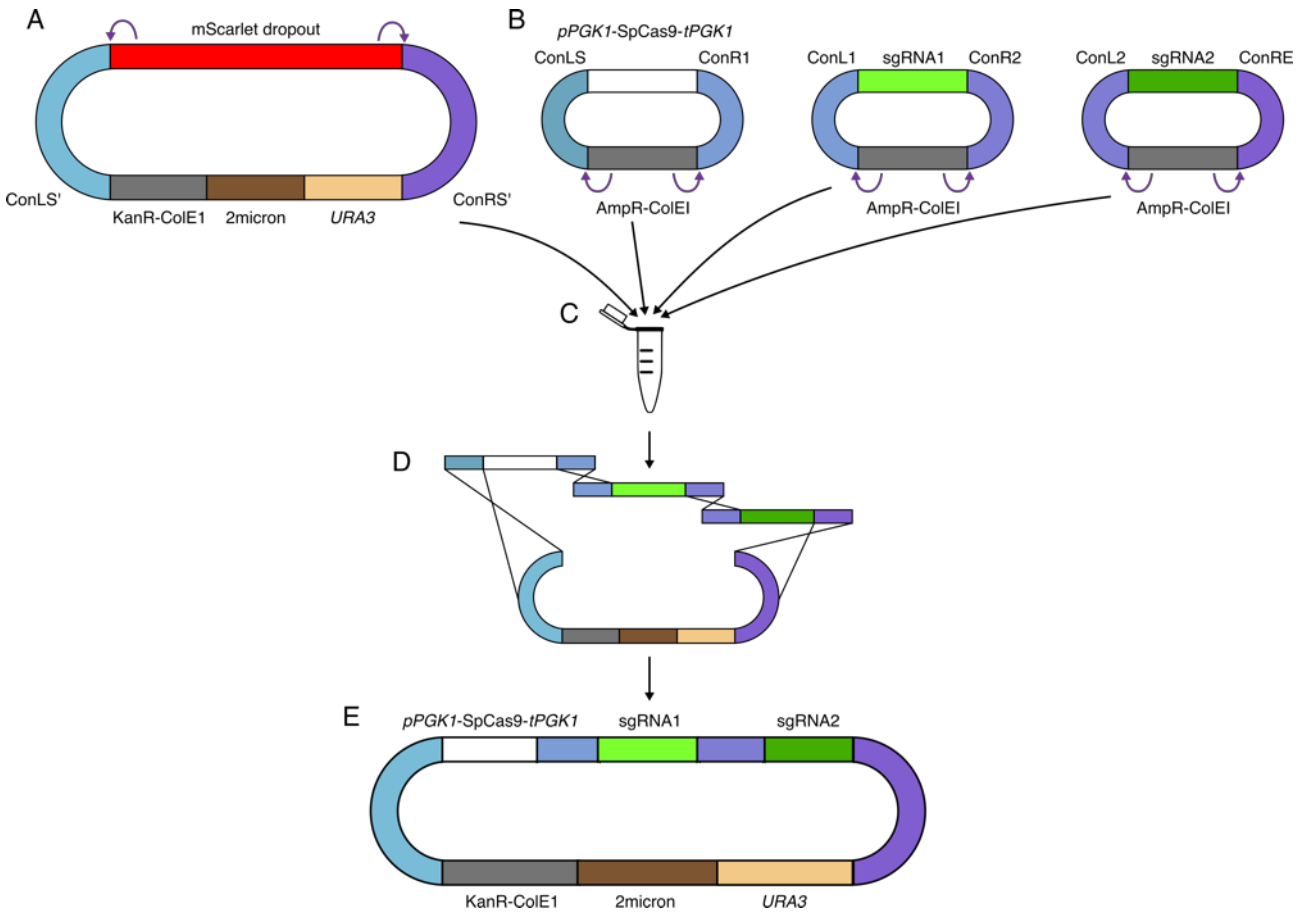
To allow for a large number of CRISPR/Cas plasmid variants for experimental flexibility, the toolkit has been designed as a small set of parts, that in combination with the YTK system can be quickly assembled to create a new plasmid with the desired endonuclease, yeast selection marker, and yeast dropout marker (**Supplementary Figure S29**). With two endonucleases, seven yeast selection markers, three yeast dropout markers, and keeping the plasmid high-copy with 2micron, 42 different CRISPR/Cas Gap Repair Vectors can be assembled to suit the experimental requirements. This can be readily expanded by adding additional parts to the YTK system.



Supplementary Figure S29. CRISPR toolkit assembly.
(*A-D*) Assembly of an example CRISPR/Cas gap repair vector containing Cas9, mScarlet dropout, *URA3* selection marker, and 2micron sequence (recommended). (*A*) Assembly of the TU1 for the CRISPR gap repair vector containing a Cas9 expression cassette. (*B*) Assembly of the TU2 for the CRISPR gap repair vector containing the mScarlet dropout with 500 bp homology. (*C*) Assembly of the CRISPR/Cas gap repair vector from TU1 and TU2 into a pre-assembled *URA3* high-copy vector. (D) CRISPR/Cas gap repair vector. (*E*) Assembly of the SpCas9 sgRNA entry vector from the 500 bp gap repair vector and the SpCas9 sgRNA dropout. (*F*) Assembled SpCas9 sgRNA entry vector. (*G*) Assembly of SpCas9 sgRNA expression vector from the SpCas9 sgRNA entry vector and two annealed oligonucleotides containing a new target sequence.

8.4.5 Alternative CRISPR protocol

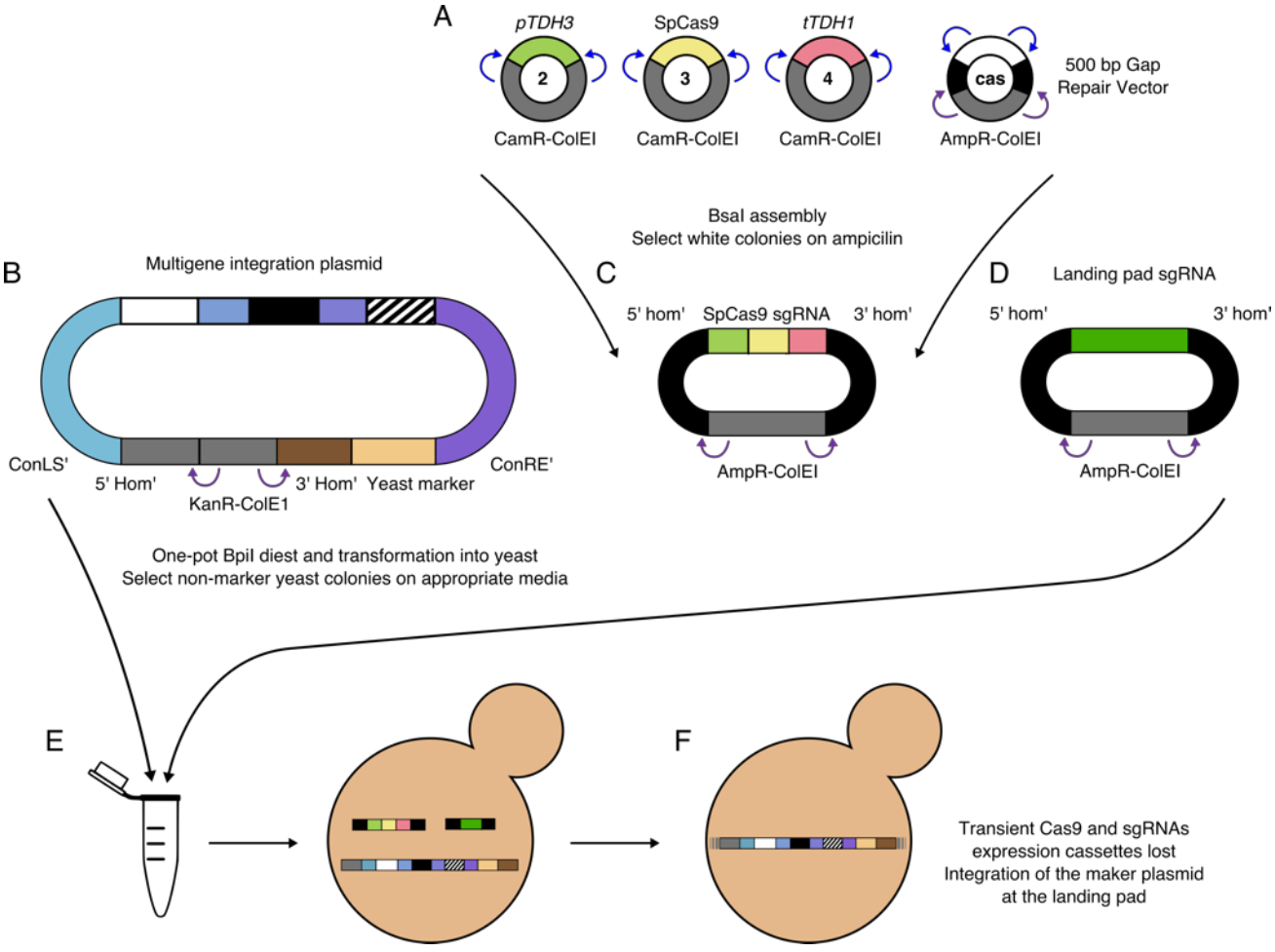
An alternative approach to the main CRISPR protocol is to use the gap repair method demonstrated in **Supplementary Figure S13**. By setting TU1 as the expression cassette for the CRISPR/Cas endonuclease and the remaining TU positions as sgRNAs, the plasmids can be combined with a yeast high-copy vector containing a BpiI dropout cassette, digested in a one-pot reaction and then assembled *in vivo* (**Supplementary Figure S30**). In this way, the cells are still selected for those highly-efficient in gap repair, however, the final multigene plasmid maintains all sgRNAs. This can lead to higher editing efficiencies, although with the trade-off of being less flexible due to the positioning requirements of the sgRNAs. For applications where a slight improvement in efficiency is worth the decrease in flexibility, this approach may be useful. For a comparison of the two methods, see **Figure 16**.



Supplementary Figure S30. Alternative CRISPR method using multigene gap repair.
(*A*) Yeast *URA3* high-copy vector using the mScarlet dropout. (*B*) Assembly vectors containing SpCas9 at TU1 and two sgRNAs at TU2 and TU3. (*C*) All four vectors are mixed together, digested with BpiI, and transformed directly into yeast. (*D*) The linearized DNA fragments are assembled *in vivo* by gap repair to create a multigene plasmid containing the SpCas9 TU1 and all sgRNAs (*E*).

8.4.6 CRISPR-aided integration of marker plasmids

In section **2.2.5**, we demonstrate the use of the landing pads for improving the efficiency of multiplexed marker plasmid integration. To achieve this, transient Cas9 and sgRNA expression cassettes are required. The transient sgRNA expression cassettes used here are the standard sgRNA plasmids, as shown in **Supplementary Figure S29G**. The transient Cas9 cassette is built by assembling a Cas9 TU into the 500 bp gap repair vector. The resulting cassettes have no selectable markers or yeast replicons and cannot be maintaining in yeast. These are eventually lost due to growth/degradation. These cassettes can then be mixed with a multigene integration plasmid, digested with BpiI, and directly transformed into yeast (**Supplementary Figure S31**).



Supplementary Figure S31. CRISPR-aided integration of marker plasmids.
(*A*) Assembly of the transient Cas9 expression cassette using the strong *TDH3* promoter and *TDH1* terminator. (*B*) Multigene integration plasmid using the updated integration arms for digestion with BpiI. (*C*) Transient Cas9 expression cassette flanked by BpiI recognition sites. (*D*) sgRNA for targeting any one of the landing pads, flanked by BpiI recognition sites. (*E*) The integration plasmid and transient Cas9 and sgRNA cassettes are added together, digested with BpiI, and transformed directly into yeast. (*F*) The transient expression of Cas9 and the sgRNAs create DSBs in the appropriate landing pads, facilitating the efficient integration of the maker plasmid. As the Cas9 and sgRNA expression cassettes contain no markers, replicons, or flanking homology with the yeast genome they are eventually lost.

8.4.7 CRISPR toolkit plasmids

Supplementary Table S7. CRISPR toolkit parts.

Plasmid Name	Part Type	Part Description	E. coli Marker
pWS2064	3	LbCpf1	CamR
pWS2005	234	sfGFP Dropout 500bp Homology	CamR
pWS2006	234	mScarlet Dropout 500bp Homology	CamR
pWS2007	234	AmilCP Dropout 500bp Homology	CamR
pWS2061	234	SpCas9 sgRNA Dropout	CamR
pWS2063	234	LbCpf1 sgRNA Dropout	CamR
pRC163	6	Yeast Marker Spacer	CamR
pWS2024	cas	500 bp Gap Repair Vector	AmpR

Supplementary Table S8. Pre-assembled CRISPR toolkit plasmids used in this study.

Plasmid Name	Part Type	Part Description	E. coli Marker
pWS2069	EV	SpCas9 sgRNA Entry Vector	AmpR
pWS2071	EV	LbCpf1 sgRNA Entry Vector	AmpR
pWS2081	O	SpCas9 URA3 Gap Repair Vector (mScarlet dropout)	KanR
pWS2082	O	SpCas9 LEU2 Gap Repair Vector (mScarlet dropout)	KanR
pWS2083	O	SpCas9 HIS3 Gap Repair Vector (mScarlet dropout)	KanR
pWS2100	O	SpCas9 URA3 Gap Repair Vector (AmilCP dropout)	KanR
pWS2102	O	LbCpf1 URA3 Gap Repair Vector (AmilCP dropout)	KanR

8.5 DNA and oligonucleotides

8.5.1 Parts

All part sequences are in YTK format (Lee *et al.*[33]) with the part type in square brackets. Only the insert is shown here, with the backbone listed following the part ype.

Blue highlight – BsaI recognition site. Green highlight – BsmBI recognition site. Purple highlight – BpiI recognition site. Light grey – NotI recognition site. **Bold** – BsaI/BsmBI/BpiI generated overhang. Light blue highlight – GFP dropout. Pink highlight – RFP dropout. Yellow highlight – start codon. Red highlight – stop codon. Underlined – open reading frame.

pWS1415 – ConLS (BpiI) [1] CamR-ColE1
GGTCTCA**ACCCT**GAATTCGCATCTAGAGAAGACTA**TGGT**AGAGCCACAAACAGCCGGTACAAGCAACGATCTCCAGGACCATCTGAAT
CATGCGCGGATGACACGAACTCACGACGGCGATCACAGACATTAACCCACAGTACAGACACTGCGACAACGTGGCAATTCGTCGCA
ATACAACGCGTCTCA**CTGA**ACTGGCCGATAATTGCAGACG**AACG**TGAGACC

pWS1416 – ConL1 (BpiI) [1] CamR-ColE1
GGTCTCA**ACCCT**GAATTCGCATCTAGAGAAGACTA**ACGG**GGTCATCACGGCTCATCATGCGCCAAACAAATGTGTGCAATACACGCTC
GGATGACTGCATGATGACCGCACTGACTGGGGACAGCAGATCCACCTAAGCCTGTGAGAGAAGCAGACACCCGACAGATCAAGGCA
GTTACGTCTCA**CCAA**ACCAGATGTCAACACAGCTAC**AACG**TGAGACC

pWS1417 – ConL2 (BpiI) [1] CamR-ColE1
GGTCTCA**ACCCT**GAATTCGCATCTAGAGAAGACTA**ATAA**ATCGCAGCCAAGTGAGTGAATAGATGACGCACCACGGTCAGACACGGC
CACATCGTATCTCACAGGAGCAAGCGCGATAGGAGCACTCACACATAGTACGGTGATCCGCTGACTCCTTTGCCCAAATAAGACGTG
AGCCCGTCTCA**GATG**CACACACTGGCTTAAGATGAC**AACG**TGAGACC

pWS1418 – ConL3 (BpiI) [1] CamR-ColE1
GGTCTCA**ACCCT**GAATTCGCATCTAGAGAAGACTA**GCAG**ACCTAGACCACATGAGGCTGATGTAGGACAGCCACCAGTGGCAGCTAAT
CAACACCGCAAGATGCCGATGCACGCTCATATCATCGTCGTCAGCCTGGTAGCCATTCGACATACGGATCAGGGAACTCGAACCCA
GTACCGTCTCA**GTTC**TGTATGGGCACAGACAACCTA**AACG**TGAGACC

pWS1419 – ConL4 (BpiI) [1] CamR-ColE1
GGTCTCA**ACCCT**GAATTCGCATCTAGAGAAGACTA**GCCT**GCACTAGACGAACTAGGCAAGATGCGTCCAATCCGTCTAAACATGGTGA
CAACGCTGGACAGATGACGTAACACCGAGCCACATCCTGAAATCGAGGCAGGCTAACCGAAACCGTGACAATGCAAAGAGACAGCC
TGACCGTCTCA**GGTA**GACTACCCATGAGTCACAATG**AACG**TGAGACC

pWS1420 – ConL5 (BpiI) [1] CamR-ColE1
GGTCTCA**ACCCT**GAATTCGCATCTAGAGAAGACTA**ACGC**ACAAGGTCAGGGCACTCATGCGACAATCAACTCGATGCATGATCCGCAC
CATTGTCGAGGGGCCAGCGTCAATAGTGCCGATGACCACAGACCCGGTTAAGACATAGCCGAATGGAGCCGCGCCGACCACAGAAT
GATACGTCTCA**AAGT**GAATAAAGCTCCACACAGTCG**AACG**TGAGACC

pWS060 - pFUS1 [2] CamR-ColE1
GGTCTCA**AACG**CCATATTTACCATGTGGACCCTTTCAAAACAGAGTTGTATCTCTGCAGGATGCCCTTTTTGACGTATTGAATGGCAT
AATTGCACTGTCACTTTTCGCGCTGTCTCATTTTGGTGCGATGATGAAACAAACATGAAACGTCTGTAATTTGAAACAAATAACGTAAT
TCTCGGGATTGGTTTTATTTAAATGACAATGTAAGAGTGGCTTTGTAAGGTATGTGTTGCTCTTAAAATATTTGGATACGACATCCTTTA
TCTTTTTTCCTTTAAGAGCAGGATATAAGCCATCAAGTTTCTGAAAATCAAAGATCT**ATG**TGAGACC

pWS053 – pFIG1 [2] CamR-ColE1
GGTCTCA**AACG**GAACTGGTTGATATTATTACTGGTGCTTCCTCTTTGGGATGATAAAAAAATCACCCTGCATTGCCTCTTTATTTGACG
TTGTTTTTGTAGAACATGAAACGAATTTTGACTTGATGACACGAAGTATATATCCAAAGAATACCTTAAATAGAAAAGGAAAGATAATAA
ATACTAAACACTACTATATATTCAGGTAAAATACAAAAATTATAACATTTTTAAAACTTTTTTTTTTTGAAAGTCCTTCTCGCTTTAGGATT
TTTTCCCATTAAGATTATGATGGTTTCATGTATGTGTCAGTTAAAAAAAATATGGCTAAGTAGCAATGAAACGAACCAAGAAAATGAAAA
CAATATATAGTGCTGTTGAAATAACAAAGACATTGGTATATATTTGTAAAATGTCTGTTAAATGTTTTTTATCTCAGGTTCTTGCTTGTCT
TTGGTAGAAGAAATTATAGTAAACAAACAAACAAACAAACAAAAAAAAAAAAAAAAAAAGATCT**ATG**TGAGACC

pWS1078 - pZ3 [2] CamR-ColE1
GGTCTCA**AACG**CCCCATTATCTTAGCCTAAAAAAACCTTCTCTTTGGAACTTTCAGTAATACGCTTAACTGCTCATTGCTATATTGAAG
TACGGCCGCGTGGGCGTGCGTGGGCGGGCGTGGGCGTGCGTGGGCGGGCGTGGGCGTGCGTGGGCGTCTAGACCGTGCGTCCT
CGTCATCACCGGTCGCGTTCCTGAAACGCAGATGTGCCTCGCGCCGCACTGCTCCGAACAATAAAGATTCTACAATACTAGCTTTTAT
GGTTATGAAGAGGAAAAATTGGCAGTAACCTGGCCCCACAAACCTTCAAATTAACGAATCAAATTAACAACCATAGGATGATAATGCG
ATTAGTTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTTGATCTATTAACAGATATATAAATGGAAAAGCTGCAT
AACCACTTTAACTAATACTTTCAACATTTTCAGTTTGTATTACTTCTTATTCAAATGTCATAAAAGTATCAACAAAAAATTGTTAATATACC
TCTATACTTTAACGTCAAGGAGAAAAAACTATAAGATC**TATG**TGAGACC

pWS334 – Gal4BS(5x)-pLEU2m [2] CamR-ColE1
GGTCTCA**AACG**CCGAGCTCTTACGCGGGTCGAAGCGGAGTACTGTCCTCCGAGTGGAGTACTGTCCTCCGAGCGGAGTACTGTCCT
CCGAGTCGAGGGTCGAAGCGGAGTACTGTCCTCCGAGTGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGTCGACTCTAGA
GGGTATACAATATTATTTAAGGACCTATTGTTTTTTCCAATAGGTGGTTAGCAATCGTCTTACTTTCTAACTTTTCTTACCTTTTACATTT
CAGCAATATATATATATATTTCAAGGATATACCATTCTAAGATC**TATG**TGAGACC

pWS334 – Gal4BS(5x)-pLEU2m [2] CamR-ColE1
GGTCTCA**AACG**CCGAGCTCTTACGCGGGTCGAAGTGCTGTATATACTCACAGCAAGTGGAGTACTGTCCTCCGAGAACTGTATATAC
ACCCAGGGAGTCGAGGGTCGAAGTACTGTATGAGCATACAGTAAGTGGAGTACTGTCCTCCGAGAACTGTATATAAATACAGTTAGT
CGACTCTAGAGGGTATACAATATTATTTAAGGACCTATTGTTTTTTCCAATAGGTGGTTAGCAATCGTCTTACTTTCTAACTTTTCTTAC
CTTTTACATTTCAGCAATATATATATATATATTTCAAGGATATACCATTCTAAGATC**TATG**TGAGACC

pWS480 – LexO (8x) [2a] CamR-ColE1
GGTCTCA**AACG**GGTAGTCCATCGTTGTAGGATACTGTATATACATACAGTAGTCACGTCGTTTACTGTATATACTCACAGTACGCTATT
GCAACTACTGTATATACACCCAGTAGAGTAGGTGACTACTGTATGAGCATACAGTAGTGACAACCACTTACTGTATATAAATACAGTAG
TGGTCATCGATACTGTATATAAAACCAGTAGGCAATCGTTTGTACTGTATGTACATACAGTATACCTCGCAACTACTGTATATAAACAC
AGTAATTAACTTGTAATATTCTAATCAA**TTGA**TGAGACC

pWS481 – LexO (6x) [2a] CamR-ColE1
GGTCTCA**AACG**GGTAGTCCATCGTTGTAGGATACTGTATATACACCCAGTAGAGTAGGTGACTACTGTATGAGCATACAGTAGTGAC
AACCACTTACTGTATATAAATACAGTAGTGGTCATCGATACTGTATATAAAACCAGTAGGCAATCGTTTGTACTGTATGTACATACAGT
ATACCTCGCAACTACTGTATATAAACACAGTAATTAACTTGTAATATTCTAATCAA**TTGA**TGAGACC

pWS482 – LexO (4x) [2a] CamR-ColE1
GGTCTCA**AACG**GGTAGTCCATCGTTGTAGGATACTGTATATAAATACAGTAGTGGTCATCGATACTGTATATAAAACCAGTAGGCAAT
CGTTTGTACTGTATGTACATACAGTATACCTCGCAACTACTGTATATAAACACAGTAATTAACTTGTAATATTCTAATCAA**TTGA**TGAGA
CC

pWS483 – LexO (3x) [2a] CamR-ColE1
GGTCTCA**AACG**GGTAGTCCATCGTTGTAGGATACTGTATATAAAACCAGTAGGCAATCGTTTGTACTGTATGTACATACAGTATACCT
CGCAACTACTGTATATAAACACAGTAATTAACTTGTAATATTCTAATCAA**TTGA**TGAGACC

pWS484 – LexO (2x) [2a] CamR-ColE1
GGTCTCA**AACG**GGTAGTCCATCGTTGTAGGATACTGTATGTACATACAGTATACCTCGCAACTACTGTATATAAACACAGTAATTAACT
TGTAATATTCTAATCAA**TTGA**TGAGACC

pWS485 – LexO (1x) [2a] CamR-ColE1
GGTCTCA**AACG**GGTAGTCCATCGTTGTAGGATACTGTATATAAACACAGTAATTAACTTGTAATATTCTAATCAA**TTGA**TGAGACC

pWS938 – TetO (6x) [2a] CamR-ColE1
GGTCTCA**AACG**GGTAGTCCATCGTTGTAGGATACTCCCTATCAGTGATAGAGAAATCTATGCGGCATCCCTATCAGTGATAGAGATC
GGTAACGGAGTCCCTATCAGTGATAGAGAAGTGGTGTTCAGTTCCCTATCAGTGATAGAGAAGTTTGACACTATCCCTATCAGTGATA
GAGAGAACAGCAATGACTCCCTATCAGTGATAGAGATCAAATGGCGTATCCCTATCAGTGATAGAGATAATTAACTTGTAATATTCTAA
TCAA**TTGA**TGAGACC

pWS890 – Gal4BS(6x) [2a] CamR-ColE1
GGTCTCA**AACG**GGTAGTCCATCGTTGTAGGATACGGATTAGAAGCCGCCGAGTGTCTATTATCCCGGGCGACAGCCCTCCGTCAAA
TATCTCAATACGGATGACTCTCCTCCGCAATAGGTATAGATCGGATTAGAAGCCGCCGAATCGTACTTTCAAACGGGCGACAGCCCT
CCGTAATTGAATTCTTTCGGATGACTCTCCTCCGTAATTAACTTGTAATATTCTAATCAA**TTGA**TGAGACC

pWS628 – CRISPR UAS (6x) [2a] CamR-ColE1
GGTCTCA**AACG**AACAGGTTATCGTAGTCTATCGGTTGTTACACATCTACCGAGGCTATTATGCGGTTGTTACACATCTACCGTGGCAC
TATGATCGGTTGTTACACATCTACCGTGGTATTTTAACGGTTGTTACACATCTACCGCGGAATTCTAATCGGTTGTTACACATCTACCG
CGGTTTAGAAACGGTTGTTACACATCTACCGGGGTAATTAACTTGTAATATTCTAATCAA**TTGA**TGAGACC

pWS494 – pRNR1m [2b] CamR-ColE1
GGTCTCA**TTGA**TATATAAAGGAGCTAATATTTCATTGTTGGAAAATTACTCTACCATAATTGAAGCATATCTCATCCTTTTCATCCTTTTC
AACGCAAGAGAGACACCAACGAACAACACTTTATTTGTTGATATATTAACATCAGATC**TATG**TGAGACC

pWS491 – pTEF2m [2b] CamR-ColE1
GGTCTCA**TTGA**TATATAAATTCTCTTGCATTTTCTATTTTTCTCTCTATCTATTCTACTTGTTTATTCCCTTCAAGGTTTTTTTTTAAGGAG
TACTTGTTTTTAGAATATACGGTCAACGAACTATAATTAACTAAACAGATC**TATG**TGAGACC

pWS487 – pTDH3m [2b] CamR-ColE1
GGTCTCA**TTGA**TATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTC
TTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAACAAAAGATC**TATG**TGAGACC

pWS493 – pALD6m [2b] CamR-ColE1
GGTCTCA**TTGA**TATATAAATGTAATAAGAAGTTTGGTAATATTCAATTCGAAGTGTTCAGTCTTTTACTTCTCTTGTTTTATAGAAGAAAA
AACATCAAGAAACATCTTTAACATACACAAACACATACTATCAGAATACAAGATC**TATG**TGAGACC

pRC246 – pLEU2m [2b] CamR-ColE1 (Kindly provided by Robert Chen)
GGTCTCA**TTGA**CAATATTATTTAAGGACCTATTGTTTTTTCCAATAGGTGGTTAGCAATCGTCTTACTTTCTAACTTTTCTTACCTTTTAC
ATTTCAGCAATATATATATATATATTTCAAGGATATACCATTCTAAGATCT**ATG**TGAGACC

pWS486 – pGAL1m [2b] CamR-ColE1
GGTCTCA**TTGA**TATATAAATGGAAAAGCTGCATAACCACTTTAACTAATACTTTCAACATTTTCAGTTTGTATTACTTCTTATTCAAATGT
CATAAAAGTATCAACAAAAAATTGTTAATATACCTCTATACTTTAACGTCAAGGAGAAAAAACTATAAGATCT**TATG**TGAGACC

pWS498 – pPHO5m [2b] CamR-ColE1
GGTCTCA**TTGA**TATATAAGCGCTGATGTTTTGCTAAGTCGAGGTTAGTATGGCTTCATCTCTCATGAGAATAAGAACAACAACAAATAG
AGCAAGCAAATTCGAGATTACCAAGATCT**ATG**TGAGACC

pWS496m – pCYC1m [2b] CamR-ColE1
GGTCTCA**TTGA**TATATAAAACTCTTGTTTTCTTCTTTTCTCTAAATATTCTTTCCTTATACATTAGGACCTTTGCAGCATAAATTACTATA
CTTCTATAGACACACAAACACAAATACACACACTAAATTAATAAGATCT**ATG**TGAGACC

pWS497 – pCUP1m [2b] CamR-ColE1
GGTCTCA**TTGA**TATAAAAGAGAAGCAAATAACTCCTTGTCTTGTATCAATTGCATTATAATATCTTCTTGTTAGTGCAATATCATATAGA
AGTCATCGAAATAGATATTAAGAAAAACAAACTGTACAATCAATCAATCAATCATCACATAAAAGATC**TATG**TGAGACC

pWS495 – pRNR2m [2b] CamR-ColE1
GGTCTCA**TTGA**TATATATAGCGGTAGTGTTTGCGCGTTACCATCATCTTCTGGATCTATCTATTGTTCTTTTCCTCATCACTTTCCCCTT
TTTCGCTCTTCTTCTTGTCTTTTATTTCTTTCTTTTTTTTAATTGTTCCCTCGATTGGCTATCTACCAAAGAATCCAAACTTAATACACGT
ATTTATTTGTCCAATTACCAGATC**TATG**TGAGACC

pWS067 – STE2 [3] CamR-ColE1
GGTCTCA**TATG**TCTGATGCGGCTCCTTCATTGAGCAATCTATTTTATGATCCAACGTATAATCCTGGTCAAAGCACCATTAACTACACT
TCCATATATGGGAATGGATCTACCATCACTTTCGATGAGTTGCAAGGTTTAGTTAACAGTACTGTTACTCAGGCCATTATGTTTGGTGT
CAGATGTGGTGCAGCTGCTTTGACTTTGATTGTCATGTGGATGACATCGAGAAGCAGAAAAACGCCGATTTTCATTATCAACCAAGTT
TCATTGTTTTTAATCATTTTGCATTCTGCACTCTATTTTAAATATTTACTGTCTAATTACTCTTCAGTGACTTACGCTCTCACCGGATTTC
CTCAGTTCATCAGTAGAGGTGACGTTCATGTTTATGGTGCTACAAATATAATTCAAGTCCTTCTTGTGGCTTCTATTGAGACTTCACTG
GTGTTTCAGATAAAAGTTATTTTCACAGGCGACAACTTCAAAAGGATAGGTTTGATGCTGACGTCGATATCTTTCACTTTAGGGATTGC
TACAGTTACCATGTATTTTGTAAGCGCTGTTAAAGGTATGATTGTGACTTATAATGATGTTAGTGCCACCCAAGATAAATACTTCAATG
CATCCACAATTTTACTTGCATCCTCAATAAACTTTATGTCATTTGTCCTGGTAGTTAAATTGATTTTAGCTATTAGATCAAGAAGATTCCT
TGGTCTTAAGCAGTTCGATAGTTTCCATATTTTACTCATAATGTCATGTCAATCTTTGTTGGTTCCATCGATAATATTCATCCTCGCATA
CAGTTTGAAACCAAACCAGGGAACAGATGTCTTGACTACTGTTGCAACATTACTTGCTGTATTGTCTTTACCATTATCATCAATGTGGG
CCACGGCTGCTAATAATGCATCCAAAACAAACACAATTACTTCAGACTTTACAACATCCACAGATAGGTTTTATCCAGGCACGCTGTCT
AGCTTTCAAACTGATAGTATCAACAACGATGCTAAAAGCAGTCTCAGAAGTAGATTATATGACCTATATCCTAGAAGGAAGGAAACAAC
ATCGGATAAACATTCGGAAAGAACTTTTGTTTCTGAGACTGCAGATGATATAGAGAAAAATCAGTTTTATCAGTTGCCCACACCTACGA
GTTCAAAAAATACTAGGATAGGACCGTTTGCTGATGCAAGTTACAAAGAGGGAGAAGTTGAACCCGTCGACATGTACACTCCCGATA
CGGCAGCTGATGAGGAAGCCAGAAAGTTCTGGACTGAAGATAATAATAATTTAGG**ATCC**TGAGACC

pWS939 – STE12 [3] CamR-ColE1
GGTCTCATATG**TCTGATGCGGCTCCTTCATTGAGCAATCTATTTTATGATCCAACGTATAATCCTGGTCAAAGCACCATTAACTACACT**
TCCATATATGGGAATGGATCTACCATCACTTTCGATGAGTTGCAAGGTTTAGTTAACAGTACTGTTACTCAGGCCATTATGTTTGGTGT
CAGATGTGGTGCAGCTGCTTTGACTTTGATTGTCATGTGGATGACATCGAGAAGCAGAAAAACGCCGATTTTCATTATCAACCAAGTT
TCATTGTTTTTAATCATTTTGCATTCTGCACTCTATTTTAAATATTTACTGTCTAATTACTCTTCAGTGACTTACGCTCTCACCGGATTTC
CTCAGTTCATCAGTAGAGGTGACGTTCATGTTTATGGTGCTACAAATATAATTCAAGTCCTTCTTGTGGCTTCTATTGAGACTTCACTG
GTGTTTCAGATAAAAGTTATTTTCACAGGCGACAACTTCAAAAGGATAGGTTTGATGCTGACGTCGATATCTTTCACTTTAGGGATTGC
TACAGTTACCATGTATTTTGTAAGCGCTGTTAAAGGTATGATTGTGACTTATAATGATGTTAGTGCCACCCAAGATAAATACTTCAATG
CATCCACAATTTTACTTGCATCCTCAATAAACTTTATGTCATTTGTCCTGGTAGTTAAATTGATTTTAGCTATTAGATCAAGAAGATTCCT
TGGTCTTAAGCAGTTCGATAGTTTCCATATTTTACTCATAATGTCATGTCAATCTTTGTTGGTTCCATCGATAATATTCATCCTCGCATA
CAGTTTGAAACCAAACCAGGGAACAGATGTCTTGACTACTGTTGCAACATTACTTGCTGTATTGTCTTTACCATTATCATCAATGTGGG
CCACGGCTGCTAATAATGCATCCAAAACAAACACAATTACTTCAGACTTTACAACATCCACAGATAGGTTTTATCCAGGCACGCTGTCT
AGCTTTCAAACTGATAGTATCAACAACGATGCTAAAAGCAGTCTCAGAAGTAGATTATATGACCTATATCCTAGAAGGAAGGAAACAAC
ATCGGATAAACATTCGGAAAGAACTTTTGTTTCTGAGACTGCAGATGATATAGAGAAAAATCAGTTTTATCAGTTGCCCACACCTACGA
GTTCAAAAAATACTAGGATAGGACCGTTTGCTGATGCAAGTTACAAAGAGGGAGAAGTTGAACCCGTCGACATGTACACTCCCGATA
CGGCAGCTGATGAGGAAGCCAGAAAGTTCTGGACTGAAGATAATAATAATTTAGG**ATCC**TGAGACC

pWS341 – STF1 [3] CamR-ColE1

GGTCTCATATGAAGTTGTTGTCTAGTATAGAACAGGCGTGTGATATTTGCAGATTAAAGAAATTGAAATGCTCTAAAGAAAAACCGAAG
TGTGCAAAATGTCTAAAAAATAACTGGGAATGTAGATATTCCCCGAAAACAAAAAGATCACCACTAACACGTGCTCATCTAACAGAAGT
CGAGAGTAGATTGGAAAGATTAGAACAGTTATTTCTTCTAATATTCCCACGTGAAGATCTTGATATGATATTGAAAATGGATAGTCTAC
AGGATATTAAAGCATTGTTAACTGGTCTATTTGTTCAAGACAATGTAAACAAAGATGCTGTCACTGATAGACTAGCTTCTGTTGAAACC
GATATGCCTCTAACCTTGCGTCAACATAGAATCAGTGCTACATCCTCTTCTGAAGAATCTAGTAACAAGGGTCAAAGACAACTAACAG
TCTCAAGACCATCTTCTACAACTAAATCTGACAATTCACCCCCCAAGCTGGAGTCCGAAAATTTCAAGGATAATGAGTTGGTAACGGT
TACTAATCAACCATTACTAGGCGTTGGCTTAATGGACGATGACGCTCCAGAAAGTCCGTCACAAATAAATGACTTTATCCCACAAAAAT
TAATTATCGAACCAAATACACTTGAACTAAACGGCCTTACTGAAGAGACTCCTCATGATCTTCCAAAAAATACAGCCAAAGGTAGAGAT
GAGGAAGATTTTCCATTGGATTATTTCCCAGTGAGCGTAGAATATCCAACTGAAGAAAATGCATTCGATCCATTTCCTCCCCAAGCCTT
TACCCCAGCAGCGCCTTCAATGCCAATTAGTTATGATAATGTCAACGAGAGAGACAGCATGCCTGTTAATTCCTTACTAAATCGTTAC
CCGTATCAATTGTCTGTTGCCCCCACCTTTCCCGTGCCTCCTTCTTCTAGTCGTCAACATTTTATGTACCCATATGATGTCCCTGACTA
CGCCATTGACTCCGCCGCCCATCATGATAATTCAACGATTCCACTAGACTTCATGCCTAGAGATGCTTTGCACGGCTTCGACTGGAG
CGAAGAAGATGATATGTCCGACGGACTTCCTTTTTTGAAAACTGATCCTAATAATAATGGGTTCGG**ATCC**TGAGACC

pWS342 – STF2 [3] CamR-ColE1

GGTCTCATATGGGTGCCCCGCCAAAGAAGAAGAGAAAGGTCGCTAAAGCTCTAACAGCGAGGCAGCAGGAAGTCTTTGACCTTATTA
GAGATCACATCTCCCAAACAGGTATGCCACCGACTAGAGCAGAGATTGCTCAAAGACTTGGTTTCAGGAGCCCTAATGCCGCAGAGG
AGCACCTAAAAGCTCTAGCCAGGAAGGGAGTCATAGAGATAGTCTCTGGTGCATCAAGAGGTATTAGATTGCTTCAAGAAGAAGAAG
AAGGGCTGCCATTGGTTGGTAGGGTTGGTAGGCCATCTTCAACAACAAAAAGTGATAATTCTCCTCCCAAATTGGAATCAGAAAACTT
CAAAGACAATGAATTAGTGACGGTTACTAATCAACCATTGTTAGGCGTGGGACTAATGGATGACGATGCGCCAGAGTCACCTTCCCA
GATAAACGATTTTATCCCGCAAAAGCTTATTATCGAACCTAATACATTAGAACTTAACGGATTAACTGAAGAAACACCGCACGACTTGC
CAAAAAACACCGCCAAAGGTCGTGATGAAGAAGATTTCCCATTGGATTACTTCCCTGTTAGCGTGGAATATCCTACCGAAGAGAATGC
TTTTGATCCATTTCCCCCACAAGCCTTTACACCAGCTGCCCCAAGTATGCCCATATCATACGATAACGTGAACGAAAGAGATTCCATG
CCAGTGAACTCTTTGTTGAACAGGTATCCGTACCAACTGTCAGTCGCTCCGACATTCCCCGTCCCACCAAGTTCCTCCAGACAACACT
TTATGGGTATTAACAAGGACATTGAAGAGTGCAACGCTATAATCGAACAGTTTATTGACTATTTACGTACTGGTCAAGAAATGCCTATG
GAAATGGCAGATCAGGCTATCAACGTCGTCCCCGGCATGACACCAAAAACTATCCTGCATGCTGGTCCTCCTATACAGCCAGATTGG
TTAAAATCAAATGGCTTCCATGAGATAGAGGCTGACGTGAATGACACTTCTTTATTGTTGAGCGGCGATGCATCAGG**ATCC**TGAGACC

pWS033 – sfGFP [3] CamR-ColE1

GGTCTCATATGTCCAAGGGTGAAGAGCTATTTACTGGGGTTGTACCCATTTTGGTAGAACTGGACGGAGATGTAAACGGACATAAATT
CTCTGTTAGAGGTGAGGGCGAAGGCGATGCCACCAATGGTAAATTGACTCTGAAGTTTATATGCACTACGGGTAAATTACCTGTTCCT
TGGCCAACCCTAGTAACAACTTTGACATATGGTGTTCAATGTTTCTCAAGATACCCAGACCATATGAAAAGGCATGATTTCTTTAAAAG
TGCTATGCCAGAAGGCTACGTGCAAGAGAGAACTATCTCCTTTAAGGATGACGGTACGTATAAAACACGAGCAGAAGTGAAATTCGA
AGGGGATACACTAGTTAATCGCATCGAATTAAAGGGTATAGACTTTAAGGAAGATGGTAATATTCTCGGCCATAAACTTGAGTATAATT
TCAACTCGCATAATGTGTACATTACAGCTGACAAACAAAAGAACGGAATTAAAGCGAATTTTAAAATCAGGCACAACGTCGAAGATGG
GTCTGTTCAACTTGCCGATCATTATCAGCAAAACACCCCTATTGGTGATGGTCCAGTCTTGTTACCCGATAATCACTACTTAAGCACAC
AGTCTAGATTGTCAAAAGATCCGAATGAAAAGCGTGATCACATGGTTTTATTGGAATTTGTCACCGCTGCAGGAATAACTCACGGAAT
GGACGAGCTTTATAAGGG**ATCC**TGAGACC

pRC063 – mTagBFP2 [3] CamR-ColE1 **(Kindly provided by Robert Chen)**

GGTCTCATATGAGTGAGTTGATTAAGGAAAACATGCACATGAAGTTGTACATGGAAGGAACAGTAGATAATCATCATTTCAAATGTAC
GTCAGAAGGTGAGGGCAAACCTTATGAAGGAACTCAAACTATGAGAATTAAAGTCGTTGAGGGTGGCCCACTACCATTTGCATTCGA
TATCCTCGCAACATCTTTTCTGTATGGTTCTAAAACCTTTATCAACCATACACAAGGCATACCAGATTTCTTTAAGCAATCATTTCCTGA
AGGTTTCACATGGGAAAGAGTAACTACCTATGAAGATGGTGGTGTCCTTACAGCTACTCAAGACACGTCTTTGCAGGATGGTTGTTTA
ATCTACAATGTTAAGATAAGAGGAGTGAACTTTACATCCAATGGACCAGTCATGCAAAAGAAAACTCTAGGTTGGGAGGCATTCACAG
AAACTTTGTACCCAGCCGATGGGGGCTTAGAAGGGAGAAATGACATGGCTTTAAAACTGGTTGGCGGTAGCCATTTGATTGCGAACG
CTAAAACAACCTACAGAAGCAAAAAGCCAGCTAAAAATCTTAAGATGCCTGTATACTACGTGGATTACAGATTAGAAAGGAAGGAAGC
TAATAATGAAACTTACGTTGAACAGCACGAGGTGGCCGTTGCCCGTTATTGCGACCTTCCTTCAAAATTGGGACACAAACTAAACGG**A
TCC**TGAGACC

pWS803 – mScarlet [3] CamR-ColE1

GGTCTCATATGGTTTCTAAAGGTGAAGCAGTTATTAAGGAATTCATGAGATTCAAGGTACACATGGAAGGTAGTATGAACGGTCACGA
ATTTGAAATTGAAGGTGAAGGTGAAGGTAGACCATATGAAGGTACTCAAACTGCTAAGTTGAAGGTTACTAAAGGTGGTCCATTGCCA
TTTTCTTGGGATATCTTGTCTCCACAATTCATGTACGGTTCTAGAGCTTTTACAAAACATCCAGCAGATATCCCAGATTACTACAAGCA
ATCATTCCCAGAAGGTTTTAAATGGGAAAGAGTTATGAACTTCGAAGATGGTGGTGCAGTTACTGTTACACAAGATACTTCTTTGGAA
GATGGTACATTGATCTATAAGGTTAAGTTGAGAGGTACTAATTTTCCACCAGATGGTCCAGTTATGCAAAAGAAAACTATGGGTTGGG
AAGCTTCAACAGAAAGATTGTACCCAGAAGATGGTGTTTTGAAGGGTGACATCAAGATGGCATTGAGATTGAAGGATGGTGGTAGAT
ATTTGGCTGATTTCAAGACTACATACAAGGCTAAGAAACCAGTTCAAATGCCAGGTGCTTACAACGTTGATAGAAAGTTGGATATCACT
TCTCATAATGAAGATTACACAGTTGTTGAACAATATGAAAGAAGTGAAGGTAGACACTCAACTGGTGGTATGGACGAATTATACAAGG
G**ATCC**TGAGACC

pWS625 – A2BR [3] CamR-ColE1

GGTCTCATATGTTGTTGGAAACACAAGATGCGTTGTATGTTGCGTTGGAATTGGTAATTGCTGCGCTATCGGTTGCGGGAAATGTTTT
GGTTTGTGCTGCGGTTGGAACGGCGAATACCTTGCAAACGCCTACTAATTATTTTTGGTTTCATTGGCAGCGGCTGATGTTGCTGTT
GGACTATTTGCTATTCCTTTCGCTATTACTATTTCTTTGGGATTTTGTACCGATTTTTATGGATGTCTATTTCTAGCTTGTTTTGTTTTGG
TTCTAACGCAATCTTCAATTTTTTCTCTATTGGCTGTTGCCGTAGATAGGTATTTGGCTATTTGCGTACCGCTAAGGTACAAGTCCCTT
GTAACGGGAACTCGTGCCAGGGGAGTAATAGCAGTACTATGGGTACTAGCTTTCGGAATTGGACTTACCCCTTTTTTGGGATGGAAT
TCCAAGGATTCCGCTACTAATAATTGTACAGAGCCTTGGGACGGAACTACGAACGAGTCTTGTTGTCTAGTTAAATGCCTATTCGAAA
ACGTTGTACCTATGTCTTATATGGTGTACTTTAACTTTTTCGGATGCGTGTTGCCTCCTTTGCTAATCATGTTGGTTATTTATATAAAAA
TTTTTTTGGTTGCTTGTAGGCAACTACAACGTACCGAATTGATGGATCATTCGAGGACTACTCTACAAAGGGAAATTCACGCCGCTAA
ATCCTTGGCTATGATAGTTGGAATATTCGCTTTGTGTTGGCTCCCTGTTCACGCAGTGAATTGCGTAACCCTATTTCAACCTGCACAA
GGCAAGAACAAACCTAAATGGGCCATGAACATGGCTATACTATTGTCCCACGCTAACTCCGTGGTAAACCCTATAGTATACGCATATA
GGAATCGTGATTTTCGTTATACCTTCCATAAGATAATTTCAAGGTACCTACTATGTCAGGCCGACGTAAAATCCGGAAACGGCCAAGC
AGGAGTGCAACCTGCACTAGGCGTTGGACTCGG**ATCC**TGAGACC

pWS626 – Mam2 [3] CamR-ColE1

GGTCTCATATGAGGCAACCTTGGTGGAAAGACTTTACTATTCCCGATGCAAGCGCAATTATTCACCAAAATATTACCATTGTGTCTATT
GTCGGAGAGATTGAAGTGCCTGTTTCAACAATTGATGCATATGAAAGGGATAGGTTATTAACTGGAATGACTTTATCTGCCCAATTAG
CTTTAGGAGTGTTAACCATTTTAATGGTTTGTCTGTTATCATCAAGCGAAAAACGTAAACACCCTGTTTTTGTTTTTAATTCGGCAAGTA
TTGTTGCAATGTGTTTACGCGCCATTTTAAATATCGTGACCATCTGCTCGAATTCGTACAGTATCCTCGTTAATTACGGGTTTATCTTAA
ACATGGTTCATATGTATGTGCATGTGTTTAATATTTTAATTTTATTATTAGCACCGGTGATCATTTTTACTGCTGAGATGTCGATGATGA
TTCAAGTTAGGATCATTTGTGCACATGATAGGAAAACACAACGTATCATGACTGTTATTAGTGCCTGCTTAACTGTTTTAGTTCTCGCA
TTTTGGATTACTAACATGTGTCAACAGATTCAGTATCTCTTATGGTTAACTCCTTTATCGTCGAAAACCATTGTTGGATACTCTTGGCCC
TACTTTATTGCTAAAATCTTATTTGCTTTTTCGATTATTTTTCACAGTGGTGTTTTTTCATACAAACTCTTTAGGGCCATCTTAATCCGCA
AAAAAATTGGGCAATTTCCTTTTGGTCCGATGCAGTGTATTTTAGTTATTTCGTGCCAATGTTTAATTGTTCCTGCTACCTTTACTATCA
TCGATAGTTTTATCCATACGTATGATGGCTTTTCGTCTATGACTCAATGTCTCCTGATCATTTCTTTACCATTATCGAGTTTATGGGCGT
CTAGTACAGCTCTCAAATTACAATCGATGAAAACTTCATCTGCGCAAGGAGAAACCACCGAGGTTTCGATTAGGGTTGATAGGACGTT
TGATATCAAACATACTCCCAGTGACGATTATTCGATTTCTGATGAATCTGAAACTAAAAAGTGGACGGG**ATCC**TGAGACC

pWS1806 – MTNR1A [3] CamR-ColE1

GGTCTCATATGCAAGGTAATGGTTCTGCTTTGCCAAATGCTTCTCAACCAGTTTTGAGAGGTGATGGTGCTAGACCTTCTTGGTTGGC
TTCTGCTTTAGCTTGTGTTTTGATTTTCACCATCGTTGTCGATATCTTGGGTAACTTGTTGGTTATCTTGTCCGTCTACCGTAACAAGAA
ATTGAGAAACGCTGGTAACATCTTCGTTGTTTCTTTGGCTGTTGCTGATTTGGTTGTTGCTATCTATCCATATCCACTGGTCTTGATGT
CCATTTTTAACAACGGTTGGAACTTGGGTTACTTGCATTGTCAAGTTTCTGGTTTCTTGATGGGTTTGTCCGTTATTGGTTCCATTTTCA
ACATTACCGGTATCGCCATTAACAGGTACTGTTACATTTGCCACTCACTGAAGTACGACAAGTTGTACTCTTCTAAGAACTCCTTGTGC
TACGTTTTGTTGATCTGGTTGTTAACTTTGGCTGCTGTTTTGCCTAATTTGAGAGCTGGTACATTGCAATACGATCCAAGAATCTACTC
TTGTACCTTCGCTCAATCTGTTTCTTCTGCTTACACTATTGCCGTTGTCGTTTTCCATTTTTTGGTCCCAATGATTATCGTCATCTTCTG
CTACTTGAGAATCTGGATTTTGGTCTTGCAAGTCAGACAAAGAGTTAAGCCAGATAGAAAGCCAAAATTGAAGCCACAAGACTTCAGA
AACTTCGTTACCATGTTTGTGGTTTTCGTTTTGTTCGCTATTTGTTGGGCTCCATTGAACTTTATTGGTTTGGCAGTTGCTTCTGATCCA
GCTTCTATGGTTCCAAGAATTCCAGAATGGTTGTTCGTTGCTTCTTACTACATGGCTTACTTCAACTCTTGTTTGAACGCCATTATCTAC
GGCTTGTTGAATCAGAACTTTAGGAAAGAGTACAGGCGTATCATCGTTTCTTTGTGTACTGCTAGAGTTTTCTTCGTCGATTCCTCTAA
TGATGTTGCCGATAGAGTTAAGTGGAAACCATCTCCATTGATGACCAACAACAATGTTGTCAAGGTTGACTCCGTTGG**ATCC**TGAGAC
C

pWS788 – BAR1 [3] CamR-ColE1

GGTCTCATATGTCTGCAATTAATCATCTTTGTTTGAAACTTATTTTGGCGAGTTTCGCGATTATTAACACCATTACTGCTTTAACAAACG
ATGGCACTGGTCACTTAGAATTCCTTTTACAACACGAAGAGGAGATGTATTACGCAACAACCTTAGATATAGGTACACCGTCCCAAAG
TCTGACAGTGTTGTTTGATACCGGATCTGCCGATTTTTGGGTTATGGATTCTAGCAATCCCTTCTGCTTACCAAATTCAAATACGTCAT
CCTATTCAAACGCAACTTATAATGGCGAAGAAGTTAAGCCTTCAATTGATTGCAGGTCTATGAGTACTTATAATGAGCATAGATCTTCC
ACCTACCAATATCTGGAAAATGGTAGGTTTTACATCACATATGCTGACGGAACATTTGCTGACGGTAGTTGGGGGACGGAAACTGTAT
CAATTAATGGAATTGACATCCCCAATATCCAGTTCGGAGTTGCCAAGTATGCTACGACACCCGTTAGTGGTGTTCTTGGAATTGGGTT
TCCTAGAAGAGAGTCCGTTAAGGGCTATGAAGGTGCTCCTAATGAATATTATCCTAATTTTCCTCAGATTTTAAAAAGTGAAAAAATAA
TCGATGTGGTCGCGTATTCGCTGTTCTTAAACTCACCTGATTCAGGTACTGGTTCGATTGTTTTTGGTGCCATTGATGAATCAAAGTTT
TCTGGTGATTTGTTCACTTTCCCTATGGTAAATGAATATCCCACAATAGTCGACGCTCCTGCAACTTTAGCAATGACTATACAAGGATT
AGGTGCCCAAAACAAAAGTAGTTGTGAACATGAAACGTTTACGACGACCAAGTATCCAGTTTTGTTGGACTCAGGAACCTCGCTATTG
AATGCGCCCAAGGTCATAGCAGATAAAATGGCTTCTTTTGTAAATGCGTCCTATAGTGAAGAGGAAGGTATATATATATTAGACTGTCC
AGTATCTGTAGGTGACGTGGAATACAATTTTGATTTCGGCGATTTGCAAATAAGTGTTCCACTGTCTAGTTTGATTTTAAGTCCCGAGA
CAGAAGGCAGCTATTGTGGGTTTGCGGTCCAGCCAACAAACGATTCGATGGTTCTGGGTGATGTGTTCCTGTCCTCTGCATACGTCG
TATTCGATCTCGATAATTATAAGATATCTTTAGCACAGGCAAATTGGAACGCAAGCGAAGTTTCGAAAAAGCTAGTAAATATTCAAACA
GATGGGTCTATTTCAGGTGCCAAAATTGCTACAGCTGAACCCTGGTCCACCAATGAACCATTTACAGTCACCTCTGACATTTATTCATC
TACAGGCTGCAAGAGTAGGCCTTTTCTTCAATCATCGACAGCCTCTTCGCTTATTGCAGAAACCAACGTACAAAGTCGCAACTGCTCT
ACGAAGATGCCAGGCACTAGATCAACTACTGTCTTAAGTAAGCCTACTCAAAATAGTGCTATGCATCAAAGTACAGGCGCTGTCACAC
AAACCTCAAATGAAACTAAATTAGAATTATCCTCGACTATGGCAAATTCGGGCAGTGTCTCGCTTCCCACTTCGAATTCAATAGACAAA
GAGTTCGAACATTCGAAATCTCAAACTACCAGCGATCCAAGTGTAGCAGAGCATTCTACGTTTAACCAAACGTTTGTACATGAAACTAA
ATATCGGCCTACTCATAAGACAGTCATAACAGAAACTGTCACGAAGTATTCTACAGTCTTAATAAATGTCTGTAAACCAACATATGG**AT**
CCTGAGACC

pWS081 – SST2 [3] CamR-ColE1

GGTCTCATATGGTTGACAAGAACAGGACATTGCATGAATTGTCATCTAAGAATTTCAGTAGGACGCCAAATGGTCTTATCTTTACAAAT
GATCTAAAAACCGTATATTCAATATTTTTGATTTGCCTTGACCTAAAGGAAAAGAAACACAGTTCAGATACTAAATCCTTTCTACTAACA
GCTTTCACCAAACATTTCCATTTTACATTCACGTATCAGGAGGCAATCAAGGCCATGGGTCAACTGGAACTAAAAGTGGATATGAATA
CTACGTGCATCAATGTGAGTTATAATATTAAGCCGAGCTTAGCAAGACATCTACTAACCCTATTCATGAGTTCAAAGTTATTACATACG
CCTCAAGATCGTACAAGAGGTGAGCCCAAGGAGAAGGTATTATTCCAACCTACGCCCAAAGGTGTCGCGGTTCTGCAGAAATACGTC
AGAGACATTGGTCTAAAGACCATGCCGGACATATTGCTGTCTAGTTTCAACTCTATGAAGCTGTTTACATTCGAGCGTTCTTCCGTTAC
TGACAGCATTATTCATTCAGACTATCTGATACATATCCTATTCATTAAAATGATGGGTGCTAAGCCCAACGTTTGGAGCCCAACCAATG
CCGACGACCCTCTGCCATGCCTGAGTAGTTTATTAGAGTATACGAATAACGACGACACATTTACCTTCGAAAAGTCTAAACCGGAGCA
GGGATGGCAAGCTCAAATCGGAAACATAGACATCAATGATCTAGAGAGAGTCTCCCCGCTTGCCCACAGGTTTTTTACGAACCCTGA
TAGTGAATCCCACACTCAATATTATGTAAGTAATGCTGGTATAAGACTTTTCGAAAATAAGACGTTCGGCACGTCAAAGAAGATAGTGA
TTAAATACACGTTTACGACAAAAGCAATTTGGCAATGGATCATGGACTGCACAGATATAATGCATGTAAAAGAGGCAGTGTCCCTAGC
CGCACTTTTCTTGAAAACAGGATTGATCGTACCCGTACTATTGCAACCAAGCCGTACGGATAAGAAAAAATTTCAAATCAGTAGGTCTA
GCTTCTTCACCCTTAGCAAACGTGGCTGGGATCTAGTATCCTGGACCGGGTGCAAGTCTAACAATATTCGTGCACCAAACGGCTCCA
CGATAGACCTGGACTTTACCCTGAGAGGCCATATGACAGTTAGGGACGAGAAAAAGACTCTAGATGATTCTGAAGGCTTTAGCCAAG
ATATGTTGATCTCAAGCAGCAATCTTAACAAATTAGACTACGTGCTAACTGATCCTGGCATGCGTTACTTGTTCCGTAGACATCTGGAG
AAAGAGCTTTGCGTGGAGAATTTAGATGTTTTCATTGAGATAAAACGTTTCCTAAAGAAGATGACCATTTTAAAAAAACTGATCGATTC
AAAGCACTGCGATAAAAAGTCCAATACGAGCACAAGTAAGAATAACATCGTGAAAACCATCGATAGTGCCTTAATGAAGCAGGCCAAC
GAGTGTTTGGAAATGGCCTATCACATCTATTCCTCTTATATAATGATTGGAAGTCCATATCAATTGAACATTCATCACAACTTGAGACA
GAACATTTCTGACATTATGTTACACCCACATAGTCCTTTGTCAGAACACTTTCCAACGAATCTGTACGATCCATCCCCCGCCAGCGCC
GAGTCCGCTGCATCTTCAATTTCTAGTACAGAGGCGGACACCCTTGGCGAGCCACCAGAAGTGAGCCTGAAACCAAGTAAGAACCTG
AGCAATGAAAATTGCTCATTCAAGAAGCAGGGTTTCAAACATCAGTTGAAAGAATATAAGCCTGCCCCCTTAACGCTGGCGGAGACTC
ATTCCCCTAACGCTAGTGTGGAAAACAGCCATACGATCGTGAGATATGGTATGGACAATACGCAGAACGATACGAAGTCAGTGGAAT
CATTTCCTGCTACCCTTAAAGTATTAAGGAAGCTTTACCCCTTATTTGAAATAGTGTCAAATGAGATGTATAGGTTGATGAACAACGAC
TCCTTCCAAAAATTTACCCAGAGTGATGTTTACAAGGACGCTTCTGCCTTGATTGAAATACAGGAAAAATGCGGATCCTGAGACC

pWS509 – MSG5 [3] CamR-ColE1

GGTCTCCTATGCAATTTCACTCAGATAAGCAGCATTTGGACAGTAAAACCGACATCGATTTCAAGCCAAACTCACCGCGTTCCTTACA
AAATAGGAATACCAAAAATTTATCTTTAGATATAGCAGCACTCCATCCATTAATGGAATTCTCATCGCCAAGCCAAGATGTGCCAGGTT
CAGTAAAATTCCCATCGCCGACACCTTTGAATCTATTTATGAAGCCCAAACCTATTGTGTTGGAGAAATGTCCACCAAAAGTAAGTCCA
AGGCCAACGCCACCATCGCTGTCGATGAGGCGAAGCGAGGCCTCTATATACACACTACCAACATCTTTGAAGAACCGAACTGTTTCT
CCAAGCGTGTATACAAAATCATCCACAGTATCGTCTATCAGTAAGCTGTCATCATCATCACCGTTATCGTCATTTTCAGAAAAACCTCA
TCTGAATAGAGTCCATTCATTATCCGTGAAAACTAAAGACTTGAAGTTGAAGGGAATTAGAGGACGTTCTCAAACGATCTCAGGGTTA
GAGACATCCACGCCAATTTCTAGTACTCGTGAAGGTACTTTAGATAGTACGGATGTCAACAGATTTTCTAACCAAAAGAATATGCAAAC
AACATTGATTTTCCCGGAGGAGGACTCGGATCTGAATATTGATATGGTGCATGCAGAGATTTATCAACGAACGGTTTATTTAGATGGA
CCATTGCTGGTACTGCCGCCTAATTTGTATCTATATTCAGAACCCAAACTAGAAGATATATTATCGTTTGATTTAGTCATTAATGTTGCC
AAAGAAATACCGAACCTGGAGTTTTTAATACCGCCGGAAATGGCACATAAAATAAAATATTACCATATTGAATGGACACACACATCCAA
GATCGTCAAGGACTTATCCCGATTGACACGCATTATACATACCGCTCATTCGCAAGGCAAGAAAATACTCGTACACTGTCAGTGTGGA
GTATCAAGATCGGCGTCATTGATTGTGGCGTATATCATGCGATATTATGGCTTGAGTTTAAATGATGCATACAATAAGCTGAAAGGTGT
TGCTAAGGATATAAGTCCAAACATGGGTCTGATCTTCCAACTTATGGAATGGGGAACCATGTTGTCCAAGAACTCACCGGGCGAAGA
AGGAGAGACTGTTCACATGCCTGAGGAAGATGACATCGGAAACAACGAAGTTTCCTCGACTACGAAGTCCTACTCTTCTGCGTCCTTT
AGAAGTTTTCCCATGGTAACGAATCTATCATCGTCGCCGAATGACAGTTCTGTCAATTCTTCGGAAGTAACGCCAAGAACTCCTGCTA
CGTTGACTGGAGCAAGGACCGCACTGGCCACAGAACGCGGGGAAGATGATGAGCACTGTAAAAGTTTGTCTCAACCCGCAGATTCA
CTGGAAGCTTCTGTGGACAACGAATCAATATCTACTGCCCCGGAACAGATGATGTTTCTTCCTGG**ATCC**TGAGACC

pWS510 – DIG1 [3] CamR-ColE1

GGTCTCCTATGGCCGTATCAGCCCGTTTGAGAACGACTGCCGAGGATACATCCATTGCTAAATCAACACAAGATCCAATTGGTGATA
CAGAAATCAGTGTAGCAAATGCTAAGGGCAGCAGCGATAGCAACATTAAGAATTCGCCAGGCGGAAACAGCGTTGGTCAGGAGTCG
GAGCTAGAGCATGTCCCTGAGGAGGATGACTCTGGTGACAAGGAAGCAGATCATGAGGATTCTGAGACAGCCACTGCGAAGAAGAG
GAAAGCCCAACCATTGAAGAATCCGAAGAAATCATTGAAGAGGGGCAGAGTCCCGGCGCCTTTGAATTTATCGGATAGCAACACTAA
TACACACGGTGGTAATATTAAGGACGGCAACTTGGCTTCGTCTAACTCTGCACATTTTCCTCCTGTTGCTAATCAAAACGTGAAAAGC
GCGCCCGCACAAGTTACTCAGCATTCCAAGTTCCAGCCCCGTGTCCAATACTTGGGAAAGGCCAGTTCTAGACAATCTATACAAGTG
AATAATAGCAGCAATAGTTATGGGAAACCACACATGCCCTCGGCGGGCATCATGAGCGCCATGAACCCTTACATGCCCATGAATCGC
TACATAATGTCACCATATTATAATCCGTACGGTATCCCTCCACCTCACATGCTGAACAAGCCCATAATGACACCCTACGTGTCGTATCC
ATATCCAATGGGACCGCGGACCTCCATTCCTTATGCTATGCAAGGTGGCAACGCTAGGCCCTACGAAGAAAATGAGTATAGTGCTAG
CAATTACAGAAACAAGAGGGTTAACGACTCATATGATTCGCCTTTGAGTGGCACCGCTTCCACTGGGAAAACCAGACGATCCGAGGA
AGGCTCAAGAAATTCTAGCGTAGGATCAAGTGCGAATGCCGGTCCTACGCAACAGCGCGCGGATCTACGCCCAGCAGATATGATAC
CTGCTGAAGAATACCACTTTGAACGGGATGCATTACTCTCGGCCAACACGAAAGCCAGAAGCGCAAGCACAAGCACAAGTACAAGCA
CAAGCACAAACCGGGACAGGAGTTCATGGCATGAGGCAGAACCCAATAAAGACGAAGAAGAGGGTACGGACTTGGCCATCGAGGAC
GGAGCGGTCCCCACTCCCACTTTTACTACGTTCCAGCGGACCTCGCAGCCGCAGCAGCAATCACCTAGTCTTCTTCAAGGCGAAATC
CGACTCTCATCGCATATTTTTGCCTTCGAGTTCCCCCTGAGCTCTAGCAATGTAGACAAGAAAATGTTTATGAGCATATGTAATAAAGT
ATGGAATGAATCAAAGGAGCTGACAAAAAAATCATCATCACATCACAGAACCGGAAAAGG**ATCC**TGAGACC

pWS511 – DIG2 [3] (Note: Internal BpiI sites – use NotI for integration) CamR-ColE1

GGTCTCCT**ATG**AATAAAGAAGAGCAAGAAGACCCACAGCAAGAGCAAATATCAACTGTTCAGGAAAATGATCCCAGGAATTTGCAGC
AACTGGGAATGCTGTTAGTATCTCCAGGGCTTGATGAAGACAGATTGAGTGAGAAAATGATTTCGAAGATCAAAAAATCCAGGGATAT
CGAAAAAAATCAAAAATTGCTCATATCTAGGCTGTCGCAAAAAGAGGAAGATCATAGTGGTAAGCCTCCTACCATTACGACTTCTCCA
GCAGAGAAAACCGTACCCTTTAAGTCGCTGAATCATTCTTTAAAGAGGAAAAGGGTACCTCCAGCGCTAAATTTTTCCGATATACAAG
CATCTTCTCATTTGCATGGATCTAAAAGTGCTCCACCAAACATAACAAGATTTCCGCAGCACAAAAATAGCCTTAGGGTCAGATATATG
GGTAGGATGGCTCCTACGAATCAAGATTATCATCCTTCAGTGGCCAATTCATATATGACAGCAACCTACCCTTATCCATATACTGGACT
GCCACCAGTACCATGCTACCCATATTCTTCAACTCCAACACAAACGCACGCATACGAAGGCTATTATTCCCCGATGTATCCCGGCCCC
TTGTATAATAATGGTATAATACCAGCCGACTACCATGCAAAGAGGAAGAAGTTGGCTGGTAGATCACCACACTTGGAAGATTTGACAT
CGAGAAAAAGAACCTTTGTCTCCAAACACCACAACGGAGATCCAATCATAAGTAAGACTGATGAAGACATTGAATGCTCTGTCACGAA
AAATTCATTAAGTGAGGGCGCTTCACTTAACGACGATGCCGATGATGACAACGACAAAGAAAGGATCATTATTGGAGAAATCTCTCTG
TATGATGATGTTTTCAAATTTGAAGTTCGCGACGACAAAAATGACTATATGAAAGCATGTGAAACAATCTGGACTGAATGGCATAACTT
GAAGAAAGG**ATCC**TCAGACC

pWS1859 – STE50 [3] (Note: Internal BpiI sites – use NotI for integration) CamR-ColE1

GGTCTCAT**ATG**GAGGACGGTAAACAGGCCATCAATGAGGGATCAAACGATGCTTCGCCGGATCTGGACGTGAATGGCACAATATTGA
TGAATAATGAAGACTTTTCCCAGTGGTCGGTTGATGATGTGATAACTTGGTGTATATCCACGCTGGAGGTGGAAGAAACCGATCCATT
ATGTCAGAGACTGCGAGAAAATGATATTGTAGGAGATCTTTTGCCGGAATTGTGCTTGCAAGATTGCCAGGACTTGTGTGACGGTGAT
TTGAATAAGGCCATAAAATTCAAGATACTGATCAATAAGATGAGAGACAGCAAGTTGGAGTGGAAGGACGACAAGACTCAAGAGGAC
ATGATAACGGTACTGAAAAACTTGTACACTACTACATCTGCGAAATTGCAAGAATTTCAATCGCAGTACACAAGGCTGAGGATGGATG
TCTTGGACGTAATGAAGACCAGCTCAAGCTCTTCTCCGATTAACACACATGGAGTGTCCACTACGGTACCTTCTTCAAACAACACAAT
TATACCCAGTAGTGACGGTGTGTCTCTTTCACAAACAGACTATTTCGACACAGTTCATAACCGACAATCACCGTCAAGGAGAGAATCC
CCGGTAACGGTATTTAGGCAACCCAGTCTTTCCCACTCAAAATCTTTGCACAAGGATAGCAAAAACAAAGTACCCCAAATATCTACAA
ACCAATCTCACCCATCTGCCGTTTCAACAGCGAACACACCGGGGCCATCACCTAACGAGGCGTTAAAACAGTTGCGTGCATCTAAAG
AAGACTCCTGCGAACGGATCTTGAAAAACGCAATGAAAAGACATAACTTAGCAGATCAGGATTGGAGACAATATGTCTTGGTCATTTG
CTATGGGGATCAAGAGAGGCTGTTAGAATTGAACGAAAAGCCTGTGATCATATTCAAGAACTTAAAGCAACAGGGTTTGCACCCCGC
CATTATGTTAAGAAGAAGAGGTGATTTCGAAGAAGTAGCAATGATGAACGGAAGTGACAATGTCACCCCCGGTGGAAGATTGGG**ATC**
CTCGAGACC

pWS1715 – HTR4 [3] CamR-ColE1

GGTCTCAT**ATG**GATAAGTTGGATGCTAACGTTTCTTCCGAAGAAGGTTTCGGTTCTGTTGAAAAGGTTGTCTTGCTAACATTCTTGTCT
ACTGTCATCTTGATGGCTATCTTGGGTAACTTGTTGGTTATGGTTGCTGTTTGTTGGGATAGACAATTGAGAAAGATTAAGACTAACTA
CTTCATTGTTTCCTTGGCTTTCGCTGATTTGTTGGTTTCTGTCTTGGTTATGCCATTCGGTGCTATTGAATTGGTTCAAGATATCTGGAT
TTACGGTGAAGTCTTTTGTTTGGTTAGAACTTCTTTGGATGTCTTGTTGACTACTGCTTCTATCTTCCATTTGTGTTGTATTTCCTTGGA
TAGATATTACGCTATCTGTTGTCAACCATTGGTTTACAGAAACAAGATGACTCCATTGAGAATTGCTTTGATGTTGGGTGGTTGTTGGG
TCATTCCTACTTTCATCTCTTTCTTGCCAATCATGCAAGGTTGGAACAACATTGGTATCATCGATTTGATTGAAAAGAGAAAGTTCAAC
CAAAACTCTAACTCTACTTACTGTGTCTTTATGGTTAACAAGCCATACGCTATTACTTGTTCCGTCGTTGCTTTCTACATTCCATTCCTA
TTGATGGTCTTGGCTTACTACAGAATCTACGTTACTGCTAAGGAACATGCTCATCAAATCCAAATGTTGCAAAGAGCTGGTGCTTCCT
CTGAATCTAGACCACAATCTGCTGATCAACATTCTACTCATAGAATGAGAACTGAAACTAAGGCTGCTAAGACCTTGTGTATCATCATG
GGATGTTTCTGTTTGTGTTGGGCTCCATTCTTCGTTACTAACATCGTTGATCCATTCATCGATTACACTGTTCCAGGTCAAGTTTGGAC
TGCTTTCTTGTGGTTGGGTTACATCAACTCTGGTTTGAATCCATTCTTGTACGCATTCTTGAACAAGTCTTTCAGAAGAGCTTTCTTGAT
CATCTTGTGTTGTGATGATGAAAGATACAGAAGGCCATCTATCCTAGGTCAAACTGTTCCTTGTTCTACTACTACTATCAACGGTTCCA
CACATGTCTTGAGAGATGCTGTTGAATGTGGTGGACAATGGGAATCTCAATGTCATCCACCAGCTACTTCTCCATTGGTTGCAGCTCA
ACCATCTGATACCGG**ATCC**TCGAGACC

pWS1679.- CrtE [3] CamR-ColE1

GGTCTCCT**ATG**GATTACGCGAACATCCTCACAGCAATTCCACTCGAGTTTACTCCTCAGGATGATATCGTGCTCCTTGAACCGTATCA
CTACCTAGGAAAGAACCCTGGAAAAGAAATTCGATCACAACTCATCGAGGCTTTCAACTATTGGTTGGATGTCAAGAAGGAGGATCTC
GAGGTCATCCAGAACGTTGTTGGCATGCTACATACCGCTAGCTTATTAATGGACGATGTGGAGGATTCATCGGTCCTCAGGCGTGGG
TCGCCTGTGGCCCATCTAATTTACGGGATTCCGCAGACAATAAACACTGCAAACTACGTCTACTTTCTGGCTTATCAAGAGATCTTCA
AGCTTCGCCCAACACCGATACCCATGCCTGTAATTCCTCCTTCATCTGCTTCGCTTCAATCATCAGTCTCCTCTGCATCCTCCTCCTC
CTCGGCCTCGTCTGAAAACGGGGGCACGTCAACTCCTAATTCGCAGATTCCGTTCTCGAAAGATACGTATCTTGATAAAGTGATCACA
GACGAGATACTTTCCCTCCATAGAGGGCAGGGCCTGGAGCTATTCTGGAGAGATAGTCTGACGTGTCCTAGCGAAGAGGAATATGT
GAAAATGGTTCTTGGAAAGACGGGAGGTTTGTTCCGTATAGCGGTCAGATTGATGATGGCAAAGTCAGAATGTGACATAGACTTTGTC
CAGCTTGTCAACTTGATCTCAATATACTTCCAGATCAGGGATGACTATATGAACCTTCAGTCATCTGAGTATGCCCATAATAAGAATTT
TGCAGAGGACCTCACAGAAGGGAAATTCAGTTTTCCCACTATCCACTCGATTCATGCCAACCCCTCATCGAGACTCGTCATCAATACG
TTGCAGAAGAAATCGACCTCTCCTGAGATCCTTCACCACTGTGTAAACTACATGCGCACAGAAACCCACTCATTCGAATATACTCAGG
AAGTCCTCAACACCTTGTCAGGTGCACTCGAGAGAGAACTAGGAAGGCTTCAAGGAGAGTTCGCAGAAGCTAACTCAAGGATGGATC
TTGGTGACGTAGATTCGGAAGGAAGAACGGGGAAGAACGTCAAATTGGAAGCGATCCTGAAAAAGCTAGCCGATATCCCTCTGGG**A**
TCCTGAGACC

pWS1680 – CrtI [3] CamR-ColE1

GGTCTCCT**ATG**GGAAAAGAACAAGATCAGGATAAACCCACAGCTATCATCGTGGGATGTGGTATCGGTGGAATCGCCACTGCCGCTC
GTCTTGCTAAAGAAGGTTTCCAGGTCACGGTGTTCGAGAAGAACGACTACTCCGGAGGTCGATGCTCTTTAATCGAGCGAGATGGTT
ATCGATTCGATCAGGGGCCCAGTTTGCTGCTCTTGCCAGATCTCTTCAAGCAGACATTCGAAGATTTGGGGGAGAAGATGGAAGATT
GGGTCGATCTCATCAAGTGTGAACCCAACTATGTTTGCCACTTCCACGATGAAGAGACTTTCACTTTTTCAACCGACATGGCGTTGCT
CAAGCGGGAAGTCGAGCGTTTTGAAGGCAAAGATGGATTTGATCGGTTCTTGTCGTTTATCCAAGAAGCCCACAGACATTACGAGCT
TGCTGTCGTTCACGTCCTGCAGAAGAACTTCCCTGGCTTCGCAGCATTCTTACGGCTACAGTTCATTGGCCAAATCCTGGCTCTTCAC
CCCTTCGAGTCTATCTGGACAAGAGTTTGTCGATATTTCAAGACCGACAGATTACGAAGAGTTTTCTCGTTTGCAGTGATGTACATGG
GTCAAAGCCCATACAGTGCGCCCGGAACATATTCCTTGCTCCAATACACCGAATTGACCGAGGGCATCTGGTATCCGAGAGGAGGCT
TTTGGCAGGTTCCTAATACTCTTCTTCAGATCGTCAAGCGCAACAATCCCTCAGCCAAGTTCAATTTCAACGCTCCAGTTTCCCAGGTT
CTTCTCTCTCCTGCCAAGGACCGAGCGACTGGTGTTCGACTTGAATCCGGCGAGGAACATCACGCCGATGTTGTGATTGTCAATGCT
GACCTCGTTTACGCCTCCGAGCACTTGATTCCTGACGATGCCAGAAACAAGATTGGCCAACTGGGTGAAGTCAAGAGAAGTTGGTGG
GCTGACTTAGTTGGTGGAAAGAAGCTCAAGGGAAGTTGCAGTAGTTTGAGCTTCTACTGGAGCATGGACCGAATCGTGGACGGTCTG
GGCGGACACAATATCTTCTTGGCCGAGGACTTCAAGGGATCATTCGACACAATCTTCGAGGAGTTGGGACTCCCAGCCGATCCTTCC
TTTTACGTGAACGTTCCCTCGCGAATCGATCCTTCTGCCGCTCCCGAAGGCAAAGATGCTATCGTCATTCTTGTGCCGTGTGGCCATA
TCGACGCTTCGAACCCTCAAGATTACAACAAGCTTGTTGCTCGGGCAAGGAAGTTTGTGATCCAAACGCTTTCCGCCAAGCTTGGAC
TTCCCGACTTTGAAAAAATGATTGTGGCAGAGAAGGTTCACGATGCTCCCTCTTGGGAGAAAGAATTTAACCTCAAGGACGGAAGCAT
CTTGGGACTGGCTCACAACTTTATGCAAGTTCTTGGTTTCAGGCCGAGCACCAGACATCCCAAGTATGACAAGTTGTTCTTTGTCGGG
GCTTCGACTCATCCCGGAACTGGGGTTCCCATCGTCTTGGCTGGAGCCAAGTTAACTGCCAACCAAGTTCTCGAATCCTTTGACCGA
TCCCCAGCTCCAGATCCCAATATGTCACTCTCCGTACCATATGGAAAACCTCTCAAATCAAATGGAACGGGTATCGATTCTCAGGTCC
AGCTGAAGTTCATGGATTTGGAGAGATGGGTATACCTTTTGGTGTTGTTGATTGGGGCCGTGATCGCTCGATCCGTTGGTGTTCTTG
CTTTCGG**ATCC**TGAGACC

pWS1681 – CrtYB [3] CamR-ColE1

GGTCTCCT**ATG**ACGGCTCTCGCATATTACCAGATCCATCTGATCTATACTCTCCCAATTCTTGGTTTGCTCGGCCTGCTCACTTCCCC
GATTTTGACAAAATTTGACATCTACAAAATATCGATCCTCGTATTTATTGCGTTTAGTGCAACCACACCATGGGACTCATGGATCATCA
GAAATGGCGCATGGACATATCCATCAGCGGAGAGTGGCCAAGGCGTGTTTGGAACGTTTCTAGATGTTCCATATGAAGAGTACGCTT
TCTTTGTCATTCAAACCGTAATCACCGGCTTGGTCTACGTCTTGGCAACTAGGCACCTTCTCCCATCTCTCGCGCTTCCCAAGACTAG
ATCGTCCGCCCTTTCTCTCGCGCTCAAGGCGCTCATCCCTCTGCCCATTATCTACCTATTTACCGCTCACCCCAGCCCATCGCCCGA
CCCGCTCGTGACAGATCACTACTTCTACATGCGGGCACTCTCCTTACTCATCACCCCACCTACCATGCTCTTGGCAGCATTATCAGGC
GAATATGCTTTCGATTGGAAAAGTGGCCGAGCAAAGTCAACTATTGCAACAATCATGATCCCGACGGTGTATCTGATTTGGGTAGATT
ATGTTGCTGTCGGTCAAGACTCTTGGTCGATCAACGATGAGAAGATTGTAGGGTGGAGGCTTGGAGGTGTACTACCCATTGAGGAAG
CTATGTTCTTCTTACTGACGAATCTAATGATTGTTCTGGGTCTGTCTGCCTGCGATCATACTCAGGCCCTATACCTGCTACACGGTCG
AACTATTTATGGCAACAAAAAGATGCCATCTTCATTTCCCCTCATTACACCGCCTGTGCTCTCCCTGTTTTTTAGCAGCCGACCATACT
CTTCTCAGCCAAAACGTGACTTGGAACTGGCAGTCAAGTTGTTGGAGAAAAAGAGCCGGAGCTTTTTTGTTGCCTCGGCTGGATTTC
CTAGCGAAGTTAGGGAGAGGCTGGTTGGACTATACGCATTCTGCCGGGTGACTGATGATCTTATCGACTCTCCTGAAGTATCTTCCA
ACCCGCATGCCACAATTGACATGGTATCCGATTTTCTTACCCTACTATTTGGGCCCCCGCTACACCCTTCGCAACCTGACAAGATCCT
TTCTTCGCCTTTACTTCCTCCTTCGCACCCTTCCCGACCCACGGGAATGTATCCCCTCCCGCCTCCTCCTTCGCTCTCGCCTGCCGA
GCTCGTTCAATTCCTTACCGAAAGGGTTCCCGTTCAATACCATTTCGCCTTCAGGTTGCTCGCTAAGTTGCAAGGGCTGATCCCTCGA
TACCCACTCGACGAACTCCTTAGAGGATACACCACTGATCTTATCTTTCCCTTATCGACAGAGGCAGTCCAGGCTAGAAAGACGCCTA
TCGAGACAACAGCTGACTTGCTGGACTATGGTCTATGTGTAGCAGGCTCAGTCGCCGAGCTATTGGTCTATGTCTCTTGGGCAAGTG
CACCAAGTCAGGTCCCTGCCACCATAGAAGAAAGAGAAGCTGTGTTAGTGGCAAGCCGAGAGATGGGAACTGCCCTTCAGTTGGTG
AACATTGCTAGGGACATTAAAGGGGACGCAACAGAAGGGAGATTTTACCTACCACTCTCATTCTTTGGTTTGCGGGATGAATCAAAGC
TTGCGATCCCGACTGATTGGACGGAACCTCGGCCTCAAGATTTCGACAAACTCCTCAGTCTATCTCCTTCGTCCACATTACCATCTTC
AAACGCCTCAGAAAGCTTCCGGTTCGAATGGAAAACGTACTCGCTTCCATTAGTCGCCTACGCAGAGGATCTTGCCAAACATTCTTAT
AAGGGAATTGACCGACTTCCTACCGAGGTTCAAGCGGGAATGCGAGCGGCTTGCGCGAGCTACCTACTGATCGGCCGAGAGATCAA
AGTCGTTTGGAAAGGAGATGTCGGAGAGAGAAGGACAGTTGCCGGATGGAGGAGAGTACGGAAAGTCTTGAGTGTGGTCATGAGCG
GATGGGAAGGGCAGGG**ATCC**TGAGACC

pWS470 -tHMG1 [3] CamR-ColE1

GGTCTCCT**ATG**GACCAATTGGTGAAAACTGAAGTCACCAAGAAGTCTTTTACTGCTCCTGTACAAAAGGCTTCTACACCAGTTTTAAC
CAATAAAACAGTCATTTCTGGATCGAAAGTCAAAAGTTTATCATCTGCGCAATCGAGCTCATCAGGACCTTCATCATCTAGTGAGGAA
GATGATTCCCGCGATATTGAAAGCTTGGATAAGAAAATACGTCCTTTAGAAGAATTAGAAGCATTATTAAGTAGTGGAAATACAAAACA
ATTGAAGAACAAAGAGGTCGCTGCCTTGGTTATTCACGGTAAGTTACCTTTGTACGCTTTGGAGAAAAAATTAGGTGATACTACGAGA
GCGGTTGCGGTACGTAGGAAGGCTCTTTCAATTTTGGCAGAAGCTCCTGTATTAGCATCTGATCGTTTACCATATAAAAATTATGACTA
CGACCGCGTATTTGGCGCTTGTTGTGAAAATGTTATAGGTTACATGCCTTTGCCCGTTGGTGTTATAGGCCCCTTGGTTATCGATGGT
ACATCTTATCATATACCAATGGCAACTACAGAGGGTTGTTTGGTAGCTTCTGCCATGCGTGGCTGTAAGGCAATCAATGCTGGCGGT
GGTGCAACAACTGTTTTAACTAAGGATGGTATGACAAGAGGCCCAGTAGTCCGTTTCCCAACTTTGAAAAGATCTGGTGCCTGTAAGA
TATGGTTAGACTCAGAAGAGGGACAAAACGCAATTAAAAAAGCTTTTAACTCTACATCAAGATTTGCACGTCTGCAACATATTCAAACT
TGTCTAGCAGGAGATTTACTCTTCATGAGATTTAGAACAACTACTGGTGACGCAATGGGTATGAATATGATTTCTAAAGGTGTCGAATA
CTCATTAAAGCAAATGGTAGAAGAGTATGGCTGGGAAGATATGGAGGTTGTCTCCGTTTCTGGTAACTACTGTACCGACAAAAAACCA
GCTGCCATCAACTGGATCGAAGGTCGTGGTAAGAGTGTCGTCGCAGAAGCTACTATTCCTGGTGATGTTGTCAGAAAAGTGTTAAAA
AGTGATGTTTCCGCATTGGTTGAGTTGAACATTGCTAAGAATTTGGTTGGATCTGCAATGGCTGGGTCTGTTGGTGGATTTAACGCAC
ATGCAGCTAATTTAGTGACAGCTGTTTTCTTGGCATTAGGACAAGATCCTGCACAAAATGTTGAAAGTTCCAACTGTATAACATTGATG
AAAGAAGTGGACGGTGATTTGAGAATTTCCGTATCCATGCCATCCATCGAAGTAGGTACCATCGGTGGTGGTACTGTTCTAGAACCA
CAAGGTGCCATGTTGGACTTATTAGGTGTAAGAGGCCCGCATGCTACCGCTCCTGGTACCAACGCACGTCAATTAGCAAGAATAGTT
GCCTGTGCCGTCTTGGCAGGTGAATTATCCTTATGTGCTGCCCTAGCAGCCGGCCATTTGGTTCAAAGTCATATGACCCACAACAGG
AAACCTGCTGAACCAACAAAACCTAACAATTTGGACGCCACTGATATAAATCGTTTGAAAGATGGGTCCGTCACCTGCATTAAATCCG
G**ATCC**TGAGACC

pWS2064 – LbCpf1 [3] CamR-ColE1

GGTCTCATATGTCTGACCCAAAGAAGAAGAGAAAGGTTGGTTCTGGTTCCAAGTTGGAAAAGTTTACCAACTGCTACTCTCTGTCTAA
GACCTTGAGATTCAAAGCTATTCCTGTTGGTAAGACCCAAGAAAACATTGACAACAAGAGGTTGTTGGTTGAGGACGAAAAAAGAGCT
GAAGATTACAAGGGTGTCAAGAAGTTGCTGGACAGATATTACTTGTCCTTCATCAACGATGTCCTGCACTCTATTAAGCTGAAGAACT
TGAACAACTACATCTCCCTGTTCAGAAAAAAGACCAGAACCGAGAAAGAGAACAAAGAGTTAGAAAACTTGGAGATCAACCTGAGGAA
AGAAATTGCTAAGGCTTTCAAGGGTAACGAGGGTTACAAATCTTTGTTCAAGAAGGACATCATCGAAACCATCTTGCCAGAATTCTTG
GATGATAAGGACGAAATTGCCTTGGTCAATTCTTTCAACGGTTTCACTACTGCTTTCACTGGTTTCTTTGACAACAGGGAAAACATGTT
CTCTGAAGAAGCTAAGTCTACCTCCATTGCTTTCAGATGCATTAACGAAAACTTGACCAGGTACATCTCCAACATGGATATCTTCGAAA
AGGTTGATGCCATCTTCGATAAGCACGAAGTCCAAGAAATCAAAGAGAAGATCTTGAACTCCGATTACGACGTCGAAGATTTTTTCGA
AGGTGAGTTCTTCAACTTCGTTTTGACTCAAGAAGGTATCGACGTTTACAACGCTATTATTGGTGGTTTCGTTACCGAATCTGGTGAAA
AGATTAAGGGTTTGAACGAGTACATCAACCTGTACAATCAAAAGACCAAGCAGAAGTTGCCAAAGTTTAAGCCCTTGTACAAACAAGT
CTTGTCCGACAGAGAATCCTTGTCTTTTTATGGTGAAGGTTACACCTCCGATGAAGAAGTTTTGGAAGTTTTCAGAAACACCCTGAAC
AAGAACTCCGAAATCTTCTCATCCATTAAGAAGTTGGAGAAGCTGTTTAAGAACTTCGACGAATATTCCTCCGCTGGTATCTTTGTCAA
AAATGGTCCAGCTATTTCCACCATCTCCAAGGATATTTTTGGTGAATGGAACGTCATCAGAGATAAGTGGAATGCTGAATACGATGAC
ATCCACTTGAAAAAGAAGGCTGTTGTTACCGAAAAGTACGAGGATGATAGACGTAAGTCCTTCAAGAAGATTGGCTCTTTCTCTCTGG
AACAATTGCAAGAATATGCTGATGCCGATTTGTCCGTTGTCGAAAAATTGAAAGAAATCATCATCCAGAAGGTCGACGAAATCTACAA
AGTTTACGGTTCCTCTGAAAAGTTGTTCGATGCTGATTTCGTTTTGGAGAAGTCGTTGAAGAAAAACGATGCTGTTGTTGCCATCATGA
AGGATTTGTTGGATTCCGTTAAGAGCTTCGAGAACTACATTAAGGCTTTTTTCGGTGAGGGCAAAGAAACTAACAGAGATGAATCATT
CTACGGCGATTTTGTTTTGGCCTACGATATCTTGTTGAAGGTTGATCATATCTACGACGCCATCAGAAACTACGTTACTCAAAAACCAT
ACTCCAAAGACAAGTTCAAGCTGTACTTTCAAAACCCACAATTCATGGGTGGTTGGGACAAAGACAAAGAAACAGATTACAGAGCCAC
CATCTTGAGATATGGTTCTAAGTACTACTTGGCCATCATGGATAAGAAATACGCTAAGTGCTTGCAAAAGATCGATAAGGATGATGTC
AACGGTAACTACGAGAAGATCAACTACAAATTATTGCCCGGTCCAAACAAGATGTTGCCTAAGGTTTTCTTCAGCAAAAAGTGGATGG
CTTACTACAACCCATCCGAAGATATTCAAAAGATCTACAAGAACGGCACCTTCAAAAAGGGTGATATGTTCAACTTGAACGACTGCCA
TAAGCTGATCGATTTCTTCAAGGATTCCATCTCTAGATACCCAAAGTGGTCTAATGCTTACGATTTCAACTTCTCTGAAACCGAGAAGT
ACAAAGATATTGCCGGTTTTTACAGGGAAGTCGAAGAACAAGGTTACAAGGTTTCATTCGAATCCGCCTCTAAAAAAGAAGTTGACAA
GTTGGTTGAAGAGGGCAAGTTGTACATGTTCCAGATCTATAACAAGGACTTCTCCGATAAGTCTCATGGTACTCCAAACTTGCATACC
ATGTACTTCAAGTTGTTGTTCGACGAAAACAATCACGGTCAGATTAGATTGTCTGGTGGTGCTGAGTTGTTTATGAGAAGGGCTTCTTT
GAAGAAAGAAGAGTTGGTTGTTCATCCAGCCAATTCTCCAATTGCTAACAAGAATCCAGACAACCCAAAAAAGACTACCACCTTGTCT
TACGATGTTTACAAGGATAAGAGATTCTCCGAGGACCAATACGAATTGCATATTCCAATTGCCATTAACAAGTGCCCCAAGAACATTTT
CAAGATCAACACCGAAGTTAGGGTGTTGTTGAAGCACGATGATAACCCATACGTTATTGGTATTGATAGGGGTGAAAGAAACTTGTTG
TACATCGTTGTTGTTGATGGTAAGGGTAACATCGTGGAACAATACTCATTGAACGAGATTATCAACAACTTCAACGGCATCAGAATCAA
GACCGATTACCATTCTCTGCTTGACAAGAAAGAAAAAGAAAGATTCGAAGCCAGACAGAACTGGACCTCTATTGAAAACATCAAAGAG
TTGAAGGCCGGTTACATTTCACAAGTTGTTCATAAGATCTGCGAGCTGGTTGAAAAGTATGATGCAGTTATTGCTTTGGAGGACTTGA
ACTCTGGTTTCAAGAATTCCAGAGTTAAGGTCGAAAAGCAGGTCTACCAAAAGTTCGAAAAGATGTTGATCGACAAGCTGAACTACAT
GGTCGACAAAAAGTCTAATCCATGTGCTACAGGTGGTGCTTTGAAAGGTTATCAAATTACGAACAAGTTCGAGAGCTTCAAGTCTATG
TCTACTCAGAATGGTTTCATCTTCTACATTCCAGCTTGGTTGACCTCTAAGATTGATCCATCTACTGGTTTCGTCAACCTGCTGAAAAC
TAAGTACACTTCTATTGCCGACTCCAAGAAGTTCATCTCTTCATTCGATAGAATCATGTACGTTCCAGAAGAGGACTTATTCGAATTCG
CTTTGGATTACAAGAACTTCTCTAGAACTGATGCTGACTACATCAAGAAGTGGAAGTTGTACTCTTACGGTAACAGAATCAGGATCTTC
AGAAACCCTAAGAAGAACAACGTTTTCGATTGGGAAGAAGTCTGTTTGACCTCTGCTTACAAAGAACTGTTTAACAAGTACGGCATCA
ACTACCAACAAGGTGATATTAGAGCTTTGTTGTGCGAACAATCTGATAAGGCCTTCTACTCATCTTTTATGGCCTTGATGTCACTGATG
TTGCAGATGAGAAATTCCATTACTGGTAGAACCGATGTCGACTTTTTGATCTCTCCAGTTAAGAACTCTGACGGTATTTTCTACGACTC
TAGAAACTATGAAGCTCAAGAAAACGCCATTTTGCCAAAAAACGCTGATGCTAATGGTGCTTACAACATTGCTAGAAAAGTTTTGTGG
GCTATCGGCCAATTCAAAAAAGCAGAAGATGAGAAGCTGGACAAGGTTAAGATTGCCATCTCAAACAAAGAATGGTTGGAATACGCT
CAGACTTCTGTCAAACATGGTTCTGGTTCTCCACCAAAGAAGAAAAGAAAGGTTTAAGGATCCTGAGACC

pWS805 – RI7 [3] CamR-ColE1

GGTCTCATATGGAAAGAAGAAATCATTCTGGTAGAGTTTCAGAATTCGTTTTGTTGGGTTTTCCAGCTCCAGCACCATTGAGAGTTTTG
TTGTTTTTCTTGTCTTTGTTGGCTTATGTTTTGGTTTTGACAGAAAACATGTTGATCATCATCGCTATTAGAAACCATCCAACTTTGCAT
AAGCCAATGTATTTCTTTTTGGCTAACATGTCATTTTTAGAAATCTGGTATGTTACTGTTACAATTCCAAAAATGTTGGCAGGTTTTATT
GGTTCTAAGGAAAACCATGGTCAATTGATCTCATTCGAAGCTTGTATGACACAATTGTATTTCTTTTTGGGTTTAGGTTGTACTGAATG
TGTTTTGTTAGCTGTTATGGCTTATGATAGATACGTTGCAATTTGTCATCCATTGCATTACCCAGTTATTGTTTCTTCAAGATTGTGTGT
TCAAATGGCTGCAGGTTCTTGGGCTGGTGGTTTCGGTATCTCAATGGTTAAGGTTTTCTTGATCTCTAGATTGTCATACTGTGGTCCA
AACACAATTAATCATTTCTTTTGTGATGTTTCTCCATTGTTGAATTTGTCTTGTACTGATATGTCAACAGCAGAATTGACTGATTTCGTTT
TGGCTATTTTCATTTTGTTGGGTCCATTGTCTGTTACAGGTGCTTCATACATGGCAATTACTGGTGCTGTTATGAGAATTCCATCTGCT
GCAGGTAGACATAAAGCATTTTCTACATGTGCTTCACATTTGACTGTTGTTATTATTTTCTATGCTGCATCAATTTTTATCTATGCAAGA
CCAAAGGCTTTGTCTGCATTCGATACTAATAAGTTGGTTTCAGTTTTGTACGCTGTTATTGTTCCATTGTTTAATCCAATCATCTATTGT
TTGAGAAACCAAGATGTTAAGAGAGCTTTGAGAAGAACATTGCATTTGGCACAAGATCAAGAAGCTAACACTAATAAGGGTTCTAAAA
ATGGATCCTGAGACC

pWS338 – Olfr154 [3] CamR-ColE1

GGTCTCATATGATGCATAGGAACCAGACCGTGGTTACCGAATTTTTTTTTACCGGACTCACTTCATCATTTCACCTACAGATAGTACTA
TTTTTGACATTCCTATGCGTGTACCTAGCCACCCTTTTGGGAAATTTGGGAATGATTATACTTATACATTTGGATACCCGTCTACATATT
CCTATGTATTTCTTCCTATCTCATCTTTCATTCGTTGACGCTTGTTCATCATCAGTAATCTCCCCTAAAATGCTTTCAGATATGTTCGTT
GACAAGAAAGTTATTTCATTTCTCGGCTGCGCAATTCAACTTTGCCTCTTCTCTCAATTCGTGGTTACAGAGTGCTTTCTATTGGCTTC
AATGGCATACGACAGGTACGTGGCTATTTGCAAACCTCTTTTGTACACCCTAATTATGTCACAACGTGTATGCGTGCAACTCGTTATA
GGACCTTATTCTATCGGATTCGTGTCTACCATGGTACACATAATATCAGCCTTCGTACTTCCTTATTGCGGCCCTAACCTAATTAACCA
TTTCTTTTGCGATCTTCTTCCTGTACTATCCTTGGCCTGCGCCAACACCCAAATGAAGAAGCGACTTCTTTTTATTGTTGCAGGAATTC
TCGGCGTGTTTTCCGGCATCATTATCCTTGTTTCATACGTGTATATAGCTATTACTATTCTCAAAATTTCCTCAGCAGACGGAAGGCGT
AAGGCTTTTTCAACTTGTTCATCCCATTTGACAGCCGTATCAATACTTTATGGAACCCTATTTTTCATTTACGTGCGACCTTCTTCATCT
TTTTCATTGGACATTAACAAAGTGGTTTCCCTCTTCTACACTACCGTGATTCCTATGCTTAATCCTTTTATATACTCTTTGAGGAACAAA
GAGGTAAAGGACGCCCTCATTAGGACATTCGAGAAACAATTCTGTTATTCTTTCCAGGACAAGATCCTCGGATCCTGAGACC

pWS339 – Olfr73 [3] CamR-ColE1

GGTCTCATATGACCTTGTCCGACGGAAACCATTCCGGAGCAGTTTTTACTCTACTTGGATTTTCCGACTACCCTGAGCTTACAATTCC
TCTATTCCTTATCTTCCTAACTATTTATTCTATTACCGTTGTGGGAAACATAGGAATGATAGTAATCATTAGGATAAACCCTAAGTTGCA
TATTCCTATGTATTTTTTCCTATCTCATCTATCATTCGTTGACTTCTGCTACTCATCAATAGTGGCACCTAAAATGTTGGTGAACCTCGT
GACAATGAATAGGGGAATCTCCTTCGTGGGATGTCTCGTTCAGTTTTTCTTCTTTTGCACCTTTGTTGTGACCGAGTCATTCCTCCTCG
GAGTTATGGCATACGATAGGTTCGTTGCTATTCGTAATCCTCTCCTATATACAGTTGCTATGTCACAAAGGCTATGCGCTATGTTGGTG
CTTGGATCATACGCATGGGGAGTTGTGTGTTCACTTATCTTGACTTGTTCAGCTCTTAACCTTTCATTCTACGGCTTTAACATGATTAAT
CATTTCTTCTGCGAATTTTCATCACTACTTTCCCCTTTCCCGTTCAGATACATCAGTATCCCAGTTGCTTCTTTTTGTGTTCGCTACCTTC
AACGAAATATCTACACTACTTATTATACTACTTTCATACGTTCTAATAGTGGTGACCATATTGAAAATGAAATCCGCCTCCGGACGTCG
TAAGGCTTTTTCAACCTGCGCATCACACTTGACAGCAATCACTATTTTTCACGGAACAATTCTCTTTCTCTATTGCGTGCCTAATTCAAA
AAATTCAAGGCACACTGTGAAGGTTGCTTCAGTTTTCTATACAGTTGTTATTCCTATGTTGAACCCTCTCATCTATTCCTTGAGGAACA
AAGACGTAAAAGATACAGTGAAGAAGATCATCGGCACCAAGGTATACTCATCAGG**ATCC**TGAGACC

pWS340 – OR2W1 [3] CamR-ColE1

GGTCTCATATGGATCAGTCTAACTACTCCAGTCTCCACGGCTTCATATTGCTAGGATTTAGTAATCACCCTAAGATGGAAATGATTTTG
TCCGGAGTGGTAGCTATTTTTTATCTCATAACACTTGTTGGCAATACAGCTATTATACTAGCCAGTCTATTGGACTCACAACTACACAC
ACCTATGTATTTCTTTCTAAGGAACCTCAGTTTTCTCGACCTCTGCTTTACAACTTCTATTATTCCTCAAATGTTGGTAAATCTTTGGGG
ACCTGACAAAACTATTTCTTACGTTGGCTGCATTATTCAGCTATACGTGTATATGTGGCTTGGATCCGTGGAATGTCTACTATTGGCAG
TGATGTCATACGACCGATTCACAGCAATCTGCAAACCTCTTCACTACTTCGTGGTAATGAATCCTCACCTCTGCCTCAAAATGATAATT
ATGATTTGGTCCATATCCCTTGCTAACAGTGTGGTGCTCTGCACACTAACCTTGAACCTTCCTACATGCGGAAATAATATATTGGACCA
CTTTCTTTGCGAACTTCCTGCATTGGTAAAAATCGCATGCGTGGATACTACAACAGTGGAGATGAGTGTGTTTGCACTCGGAATAATC
ATAGTACTAACACCTCTAATTCTAATACTAATCTCATACGGATATATAGCTAAGGCAGTTTTGAGGACGAAATCCAAGGCCTCTCAACG
TAAGGCCATGAACACTTGCGGAAGTCACCTAACCGTGGTTAGTATGTTTTACGGAACCATAATTTATATGTATTTGCAGCCTGGCAATA
GGGCATCAAAGGATCAAGGAAAATTTCTAACTCTATTCTATACTGTAATTACCCCTTCCCTAAATCCGCTAATATATACTCTCAGGAAC
AAAGATATGAAAGACGCTTTGAAAAAGTTGATGAGGTTCCATCATAAGAGTACAAAGATCAAAAGGAACTGTAAATCCGG**ATCC**TGAG
ACC

pWS1155 – OR3A1 [3] CamR-ColE1

GGTCTCATATGCAACCAGAATCTGGTGCTAATGGTACTGTTATTGCCGAATTCATTTTGTTGGGTTTGTTGGAAGCTCCAGGTTTACAA
CCAGTTGTTTTCGTTTTGTTCTTGTTCGCTTACTTGGTTACCGTTAGAGGTAACTTGTCTATTTTGGCTGCTGTTTTGGTTGAACCTAAG
TTGCATACTCCCATGTACTTTTTCTTGGGCAATTTGTCCGTTTTGGATGTTGGTTGTATTTCCGTTACTGTCCCATCTATGTTGTCTAGG
TTGTTGAGTAGAAAAAGGGCTGTTCCATGTGGTGCTTGTTTGACTCAATTATTCTTCTTCCACTTGTTCGTTGGTGTTGACTGTTTTTTG
TTGACTGCTATGGCTTACGATAGATTCTTGGCTATATGTAGACCATTGACCTACTCTACCAGAATGTCTCAAACTGTCCAAAGAATGTT
AGTTGCTGCTTCTTGGGCTTGTGCTTTTACTAATGCTTTGACTCATACCGTTGCTATGTCTACCTTGAATTTTTGTGGTCCAAACGTCA
TCAACCACTTCTATTGTGATTTGCCACAGTTGTTCCAACTGTCTTGTTCTTCTACACAACTGAACGAGTTGTTGTTGTTTGCCGTTGGT
TTCATTATGGCTGGTACTCCAATGGCTTTGATCGTTATTTCCTACATTCATGTTGCTGCTGCCGTTTTGAGAATCAGATCTGTTGAAGG
TAGAAAAAAGGCCTTCTCTACTTGTGGTTCTCATTTGACTGTTGTTGCCATTTTTTACGGTTCCGGTATCTTCAACTATATGAGATTGG
GTTCTACCAAGTTGTCCGATAAGGATAAGGCTGTTGGTATTTTCAACACCGTTATCAACCCAATGTTGAACCCCATTATCTACTCATTC
AGAAACCCAGATGTTCAGTCTGCTATTTGGAGAATGTTGACTGGTAGAAGATCTTTGGCTGG**ATCC**TGAGACC

pWS1156 – OR1A1 [3] CamR-ColE1

GGTCTCATATGCGTGAGAACAATCAATCTTCTACCTTGGAGTTCATCTTGTTGGGTGTTACTGGTCAACAAGAACAAGAGGATTTCTT
CTACATCCTGTTCCTATTCATCTACCCAATTACCTTGATCGGCAACTTGTTGATCGTTTTGGCTATTTGCTCTGATGTCAGATTGCATAA
CCCCATGTATTTCTTGTTGGCTAACTTGTCCTTGGTCGACATTTTTTTCTCATCCGTTACCATTCCAAAGATGTTGGCCAATCATTTGTT
GGGCTCTAAGTCTATTTCTTTCGGTGGTTGTTTGACCCAGATGTACTTTATGATTGCTTTGGGTAACACCGACTCCTATATTTTGGCTG
CTATGGCTTATGATAGAGCCGTTGCTATTTCTAGACCATTGCATTACACTACCATCATGTCTCCAAGATCCTGCATTTGGTTGATTGCT
GGTTCTTGGGTTATTGGTAATGCTAATGCTTTGCCACATACTTTGTTGACTGCCTCTTTGTCTTTCTGCGGTAATCAAGAAGTTGCTAA
CTTTTACTGCGATATCACCCCTTTGTTGAAGTTGTCCTGTTCTGATATTCACTTCCACGTCAAGATGATGTACTTAGGTGTTGGTATTTT
CTCTGTGCCTTTGTTGTGCATCATCGTGTCTTACATTAGAGTTTTCTCCACCGTTTTCCAAGTTCCATCTACAAAAGGTGTTCTGAAGG
CTTTTTCTACCTGTGGTTCTCATTTGACCGTCGTGTCCTTGTATTATGGTACTGTTATGGGTACTTACTTCAGGCCATTGACCAACTATT
CTTTGAAGGATGCTGTTATCACCGTTATGTACACTGCTGTTACTCCAATGTTGAACCCTTTCATCTACAGCTTGAGAAACAGAGATATG
AAGGCTGCTTTGAGGAAGTTGTTCAACAAGAGAATCTCCTCCGG**ATCC**TGAGACC

pWS1157 – OR2J2 [3] CamR-ColE1

GGTCTCATATGATGATTAAGAAGAACGCCTCTTCCGAGGATTTCTTCATTTTGTTGGGTTTTTCTAACTGGCCACAATTGGAAGTTGTT
TTGTTCGTTGTCATCCTGATCTTCTACTTGATGACTTTGACTGGCAACCTGTTCATCATCATCTTGTCTTATGTTGACTCTCACTTGCAT
ACGCCCATGTACTTTTTTCTGTCTAACTTGTCCTTCTTGGATTTGTGTCATACCACCTCTTCCATTCCACAGTTGTTGGTTAATTTGAGA
GGTCCAGAAAAGACCATTTCTTACGCTGGTTGTATGGTTCAGTTGTACTTTGTTTTGGCTTTGGGTATTGCTGAATGCGTTTTGTTGGT
CGTTATGTCCTATGATAGATACGTTGCTGTATGCAGACCATTGCATTACACTGTTTTGATGCATCCAAGATTCTGCCATTTGTTGGCTG
CTGCTTCATGGGTTATTGGTTTTACCATTTCTGCCTTGCATTCCTCTTTTACTTTTTGGGTTCCATTGTGCGGTCACAGATTGGTTGATC
ATTTCTTTTGTGAAGTTCCAGCCTTGTTGAGATTGTCTTGCGTTGATACTCATGCTAACGAATTGACCTTGATGGTCATGTCCTCTATC
TTCGTTTTGATCCCATTGATCTTGATTTTGACTGCTTACGGTGCTATTGCTAGAGCTGTTTTGTCTATGCAATCTACTACCGGCTTGCA
AAAGGTTTTTAGAACTTGTGGTGCTCACCTGATGGTTGTTTCTTTGTTTTTCATTCCAGTCATGTGCATGTACTTGCAACCACCATCTG
AAAATTCTCCAGACCAGGGTAAATTCATTGCTTTGTTCTACACTGTTGTCACGCCATCTTTGAATCCTTTGATCTACACCTTGAGAAAC
AAGCACGTTAAGGGTGCTGCTAAGAGATTATTAGGTTGGGAATGGGGTAAGGG**ATCC**TGAGACC

pWS1158 – Olfr168 [3] CamR-ColE1

GGTCTCATATGGAAAAGTGGAATCAATCCTCCTCCGATTTCATTTTGTTGGGTTTGTTGCCACAAAATCAGACCGGTTTGTTGTTGATG
ATGTTGATCATCCTGGTTTTTTTCTTGGCCTTGTTTGGTAACTCCGCCATGATTCATTTGATCAGAGTTGATCCAAGATTGCACACCCC
AATGTACTTTTTGTTGTCCCAATTGTCCTTGATGGACCTGATGTACATTTCTACTACTGTTCCAAAGATGGCCTTCAACTTTTTGTCTGG
TCAAAAGAACATCTCGTTCTTAGGTTGTGGTGTCCAATCTTTCTTCTTCTTGACTATGGCTGGTTCTGAAGGTTTGTTATTGGCTTCTAT
GGCTTACGATAGATTCGTTGCTATTTGCCATCCATTGCATTACCCAATCAGAATGTCTAAGATCATGTGCCTGAAGATGATTATCGGTT
CTTGGATTTTGGGCTCCATCAATTCTTTGGCTCATTCTATCTACGCCTTGCATATTCCATACTGCCATTCCAGATCTATCAACCACTTTT
TCTGTGATGTTCCAGCTATGTTGCCATTGGCATGTATGGATACTTGGGTTTATGAGTACATGGTGTTCGTTTCTACCTCCTTGTTCTTG
TTGTTACCATTCTTGGGTATTACCGCTTCTTACGGTAGAGTTTTGTTTGCTGTTTTTCACATGAGGTCCAAAGAGGGTAAAAAGAAGGC
TTTCACTACTTGCTCTACTCACTTGACTGTTGTCACTTTTTACTACGCTCCATTCGTCTACACTTATTTGAGGCCAAGATCTTTGAGATC
TCCAACCGAAGATAAGATCTTGACCGTTTTCTACACTATCTTGACCCCTATGTTGAACCCAATCATCTACAGCTTGAGAAACAAAGAAG
TTTTGGGTGCTATGACTAGAGTCTTGGGTACTTTTTCATCTATGAAGCCCGGATCCTGAGACC

pWS1159 – Olfr556 [3] CamR-ColE1

GGTCTCATATGTTGGGTCCATCTTACAATCATACCATGGAATCTCCAGGTACGTTTTTCTTGTTGGGTATTCCAGGTTTCCAGTCATCT
TATTTGTGGTTGGCTATTTCCTTGTCCACCATGTATTCTATTGCTTTGTTGGGCAACATGCTGATCATTATCGTTATCTGTATGGACTCT
ACCTTGCAAGAACCTATGTACTTCTTCTTGTGTGTTTTGGCTGCTGTCGATATAGTTATGGCTTCTTCTGTTGTTCCCAAGATGGTTTC
TATTTTCTCCTCTGGTGATTCCTCCATTTCTTTCAATGCTTGTTTCACCCAAATGTACTTCGTTCATGCTGCTACTGCTGTTGAAACTGG
TTTGTTGTTAGCTATGGCTTTCGATAGATACGTTGCTATCTGTAAGCCATTGCACTACATGAGAATTTTGACCAGACACGTTATGTTGG
GCATTTCTGTTACTATTACTGTTAGAGCCGTTATCTTCATGACTCCATTGTCTTGGATGTTGTCTCATTTGCCATTTTGCGCTTCTAACG
TTGTTCCACATTCTTACTGTGAACATATGGCTGTTGCTAAATTGGCTTGTGCTGATCCAATGCCATCTTCCTTGTATTCCTTGATCTTCT
CCTCTATCATCGTTGGTTCTGATGTTGCTTTTATCTCCGCTTCCTACTCCTTGATTTTGAAGGCTGTTTTTGGTCTGTCCTCTAGAAATG
CTCAATGGAAAGCTTTGTCTACCTGTGGTTCACATGTTGGTGTTATGGCCTTGTATTATTTGCCAGGTATGGCTTCAATCTACGTTGCT
TGGTTGGGTCAAGATAGAGTTCCATTGCATACTCAAGTTTTGTTGGCTGACTTGTACTTGATTATTCCACCAACTTTGAACCCAATCAT
CTACGGTATTAGAACCAGACAGATCAGAGAAAGAATCTGGTCTTTGTTGACTCACTGCTTCTTCTCACAATGTACTCAAGGTTCTGGATCCTGAGACC

pWS1160 – Olfr609 [3] CamR-ColE1

GGTCTCATATGTCCTACTCTAACCATTCTTCTACCTCATTCTTCTTGACTGGTTTGCCAGGTTTGGAAACTGTTTATTTGTGGTTGTCTA
TTCCCCTGTGTACCATGTATATTGCTTCATTGGCTGGTAACGGTTTGATTTTGTGGGTTGTTAAGTCTGAACCATCCTTGCATCAACCT
ATGTACTACTTCTTGTCTATGTTGGCTGTTACCGATTTGGGTTTGTCTGTTTCTACTTTGCCAACTATGTTGACCATCTACATGATGGGT
GTTTCTGAAGTTGCTTTGGATATGTGCTTGGCTCAGTTGTTCTTCATTCATACCTTCTCCATCATGGAATCCTCTGTCTTGTTGACTATG
GCTTTCGATAGAGTTGTTGCCATTTCTTCTCCATTGCATTACGCTACCATTTTGACTAATCCAAGAGTTGCCTCTTTAGGCATGGTTAT
CTTGGTTAGATCTATCGGCTTGCATATTCCAGCTCCAATTATGTTGAAAAAGTTGCCATACTGCCAGAAGAGACATTTGTCTCATTCTT
ATTGCTTGCACCCAGATGTTATGAAGTTGGCTTGTACTGATACCAGAATCAATTCTGCTTACGGTTTGTTCGTTGTCTTGTCTACTTTA
GGTGTTGACTCCGTTTTGATCGTTTTGTCCTATGGTTTGATCTTGTACACCGTTTTGTCTATTGCTTCCAAGACCGAAAGATTGAAGGC
TTTGAATACCTGCGTTTCCCATATTTGCTCTGTGTTGTTGTTTTACACCCCAATGATTGGCTTGTCCATGATTCATAGATTTGGTAAATG
GGCTTCTCCATGCTCTAGAGTTTTGTTGTCTTACTTGCATTTCTTGACCCCACCAGTTTTGAATCCAGTTGTTTACACTATCAAGACCA
AGCAGATCAGACAAAGAATTTGGTGCATTTTCAGATGCGGTGGTAGATCCATTGGTCATATTCAAGGTCATGGATCCTGAGACC

pWS1604 - A2BR-ΔC-tail [3-4a] CamR-ColE1

GGTCTCATATGTTGTTGGAAACACAAGATGCGTTGTATGTTGCGTTGGAATTGGTAATTGCTGCGCTATCGGTTGCGGGAAATGTTTT
GGTTTGTGCTGCGGTTGGAACGGCGAATACCTTGCAAACGCCTACTAATTATTTTTTGGTTTCATTGGCAGCGGCTGATGTTGCTGTT
GGACTATTTGCTATTCCTTTCGCTATTACTATTTCTTTGGGATTTTGTACCGATTTTTATGGATGTCTATTTCTAGCTTGTTTTGTTTTGG
TTCTAACGCAATCTTCAATTTTTTCTCTATTGGCTGTTGCCGTAGATAGGTATTTGGCTATTTGCGTACCGCTAAGGTACAAGTCCCTT
GTAACGGGAACTCGTGCCAGGGGAGTAATAGCAGTACTATGGGTACTAGCTTTCGGAATTGGACTTACCCCTTTTTTGGGATGGAAT
TCCAAGGATTCCGCTACTAATAATTGTACAGAGCCTTGGGACGGAACTACGAACGAGTCTTGTTGTCTAGTTAAATGCCTATTCGAAA
ACGTTGTACCTATGTCTTATATGGTGTACTTTAACTTTTTCGGATGCGTGTTGCCTCCTTTGCTAATCATGTTGGTTATTTATATAAAAA
TTTTTTTGGTTGCTTGTAGGCAACTACAACGTACCGAATTGATGGATCATTCGAGGACTACTCTACAAAGGGAAATTCACGCCGCTAA
ATCCTTGGCTATGATAGTTGGAATATTCGCTTTGTGTTGGCTCCCTGTTCACGCAGTGAATTGCGTAACCCTATTTCAACCTGCACAA
GGCAAGAACAAACCTAAATGGGCCATGAACATGGCTATACTATTGTCCCACGCTAACTCCGTGGTAAACCCTATAGTATACGCATATA
GGTAACTCGAGTGGCTGAGACC

pWS1037 - GPA1 [3-4a] CamR-ColE1

GGTCTCATATGGGGTGTACAGTGAGTACGCAAACAATAGGTGACGAAAGTGATCCTTTTCTACAGAACAAAAGAGCCAATGATGTCAT
CGAGCAATCGTTGCAGCTGGAGAAACAACGTGACAAGAATGAAATAAAACTGTTACTATTAGGTGCCGGTGAGTCAGGTAAATCAAC
GGTTTTAAAACAATTAAAATTATTACATCAAGGCGGTTTCTCCCATCAAGAAAGGTTACAGTATGCTCAAGTGATATGGGCAGATGCCA
TACAATCAATGAAAATTTTGATTATTCAGGCCAGAAAACTAGGTATTCAACTTGACTGTGATGATCCGATCAACAATAAAGATTTGTTTG
CATGCAAGAGAATACTGCTAAAGGCTAAAGCTTTAGATTATATCAACGCCAGTGTTGCCGGTGGTTCTGATTTTCTAAATGATTATGTA
CTGAAGTACTCAGAAAGGTATGAAACTAGGAGGCGTGTTCAGAGTACCGGACGAGCAAAAGCTGCTTTCGATGAGGACGGAAATATT
TCTAATGTCAAAAGTGACACTGACAGAGATGCTGAAACGGTGACGCAAAATGAGGATGCTGATAGAAACAACAGTAGTAGAATTAACC
TACAGGATATTTGCAAGGACTTGAACCAAGAAGGCGATGACCAGATGTTTGTTAGAAAAACATCAAGGGAAATTCAAGGACAAAATAG
ACGAAATCTTATTCACGAGGACATTGCTAAGGCAATAAAGCAACTTTGGAATAACGACAAAGGTATAAAGCAGTGTTTTGCACGTTCTA
ATGAGTTTCAATTGGAGGGCTCAGCTGCATACTACTTTGATAACATTGAGAAATTTGCTAGTCCGAATTATGTCTGTACGGATGAGGA
CATTTTGAAGGGCCGTATAAAGACTACAGGCATTACAGAAACCGAATTTAACATCGGCTCGTCCAAATTCAAGGTTCTCGACGCTGGT
GGGCAGCGTTCTGAACGTAAGAAGTGGATTCATTGTTTCGAAGGAATTACAGCAGTTTTATTTGTTTTAGCAATGAGTGAATACGACC
AGATGTTGTTTGAGGATGAAAGAGTGAACAGAATGCATGAATCAATAATGCTATTTGACACGTTATTGAACTCTAAGTGGTTCAAAGAT
ACACCGTTTATTTTGTTTTTAAATAAAATTGATTTGTTCGAGGAAAAGGTAAAAAGCATGCCCATAAGAAAGTACTTTCCTGATTACCAG
GGACGTGTCGGCGATGCAGAAGCGGGTCTAAAATATTTTGAGAAGATATTTTTGAGCTTGAATAAGACAAACAAACCAATCTACGTGA
AACGAACCTGCGCTACCGATACCCAAACTATGAAGTTCGTATTGAGTGCAGTCACCGATCTAATCATCCAGCAAAACCTGAAAAAAAT
TGGTATTATTTAACTCGAGTGGCTGAGACC

pWS936 – GPA1 C-terminal Truncation Dropout [3-4a] CamR-ColE1

GGTCTCATATGGGGTGTACAGTGAGTACGCAAACAATAGGTGACGAAAGTGATCCTTTTCTACAGAACAAAAGAGCCAATGATGTCAT
CGAGCAATCGTTGCAGCTGGAGAAACAACGTGACAAGAATGAAATAAAACTGTTACTATTAGGTGCCGGTGAGTCAGGTAAATCAAC
GGTTTTAAAACAATTAAAATTATTACATCAAGGCGGTTTCTCCCATCAAGAAAGGTTACAGTATGCTCAAGTGATATGGGCAGATGCCA
TACAATCAATGAAAATTTTGATTATTCAGGCCAGAAAACTAGGTATTCAACTTGACTGTGATGATCCGATCAACAATAAAGATTTGTTTG
CATGCAAGAGAATACTGCTAAAGGCTAAAGCTTTAGATTATATCAACGCCAGTGTTGCCGGTGGTTCTGATTTTCTAAATGATTATGTA
CTGAAGTACTCAGAAAGGTATGAAACTAGGAGGCGTGTTCAGAGTACCGGACGAGCAAAAGCTGCTTTCGATGAGGACGGAAATATT
TCTAATGTCAAAAGTGACACTGACAGAGATGCTGAAACGGTGACGCAAAATGAGGATGCTGATAGAAACAACAGTAGTAGAATTAACC
TACAGGATATTTGCAAGGACTTGAACCAAGAAGGCGATGACCAGATGTTTGTTAGAAAAACATCAAGGGAAATTCAAGGACAAAATAG
ACGAAATCTTATTCACGAGGACATTGCTAAGGCAATAAAGCAACTTTGGAATAACGACAAAGGTATAAAGCAGTGTTTTGCACGTTCTA
ATGAGTTTCAATTGGAGGGCTCAGCTGCATACTACTTTGATAACATTGAGAAATTTGCTAGTCCGAATTATGTCTGTACGGATGAGGA
CATTTTGAAGGGCCGTATAAAGACTACAGGCATTACAGAAACCGAATTTAACATCGGCTCGTCCAAATTCAAGGTTCTCGACGCTGGT
GGGCAGCGTTCTGAACGTAAGAAGTGGATTCATTGTTTCGAAGGAATTACAGCAGTTTTATTTGTTTTAGCAATGAGTGAATACGACC
AGATGTTGTTTGAGGATGAAAGAGTGAACAGAATGCATGAATCAATAATGCTATTTGACACGTTATTGAACTCTAAGTGGTTCAAAGAT
ACACCGTTTATTTTGTTTTTAAATAAAATTGATTTGTTCGAGGAAAAGGTAAAAAGCATGCCCATAAGAAAGTACTTTCCTGATTACCAG
GGACGTGTCGGCGATGCAGAAGCGGGTCTAAAATATTTTGAGAAGATATTTTTGAGCTTGAATAAGACAAACAAACCAATCTACGTGA
AACGAACCTGCGCTACCGATACCCAAACTATGAAGTTCGTATTGAGTGCAGTCACCGATCTAATCATCCAGCAAAACCT**GAAA**CGAG
ACGGAAAGTGAAACGTGATTTCATGCGTCATTTTGAACATTTTGTAAATCTTATTTAATAATGTGTGCGGCAATTCACATTTAATTTATG
AATGTTTTCTTAACATCGCGGCAACTCAAGAAACGGCAGGTTCGGATCTTAGCTACTAGAGAAAGAGGAGAAATACTAGATGCGTAAA
GGCGAAGAGCTGTTCACTGGTGTCGTCCCTATTCTGGTGGAACTGGATGGTGATGTCAACGGTCATAAGTTTTCCGTGCGTGGCGAG
GGTGAAGGTGACGCAACTAATGGTAAACTGACGCTGAAGTTCATCTGTACTACTGGTAAACTGCCGGTTCCTTGGCCGACTCTGGTA
ACGACGCTGACTTATGGTGTTCAGTGCTTTGCTCGTTATCCGGACCATATGAAGCAGCATGACTTCTTCAAGTCCGCCATGCCGGAA
GGCTATGTGCAGGAACGCACGATTTCCTTTAAGGATGACGGCACGTACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGATACCCTG
GTAAACCGCATTGAGCTGAAAGGCATTGACTTTAAAGAGGACGGCAATATCCTGGGCCATAAGCTGGAATACAATTTTAACAGCCACA
ATGTTTACATCACCGCCGATAAACAAAAAAATGGCATTAAAGCGAATTTTAAAATTCGCCACAACGTGGAGGATGGCAGCGTGCAGCT
GGCTGATCACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCACTATCTGAGCACGCAAAGCGTTCT
GTCTAAAGATCCGAACGAGAAACGCGATCATATGGTTCTGCTGGAGTTCGTAACCGCAGCGGGCATCACGCATGGTATGGATGAACT
GTACAAATGACCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTC
TCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGTCTCTTGGCTGAGACC

pWS1837 – MF(ALPHA)1 [3-4a] CamR-ColE1

GGTCTCATATGAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAA
GATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGTTACTTAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTCCA
ACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTTTGGATAAAAGAGAG
GCTGAAGCTTGGCATTGGTTGCAACTAAAACCTGGCCAACCAATGTACAAGAGAGAAGCCGAAGCTGAAGCTTGGCATTGGCTGCAA
CTAAAGCCTGGCCAACCAATGTACAAAAGAGAAGCCGACGCTGAAGCTTGGCATTGGCTGCAACTAAAGCCTGGCCAACCAATGTAC
AAAAGAGAAGCCGACGCTGAAGCTTGGCATTGGTTGCAGTTAAAACCCGGCCAACCAATGTACTAACTCGAG**TGGC**TGAGACC

pWS627 – STE4-2A-STE18 [3-4a] CamR-ColE1

GGTCTCATATGGCAGCACATCAGATGGACTCGATAACGTATTCTAATAATGTCACCCAACAGTATATACAACCACAAAGTCTACAGGA
TATCTCTGCAGTGGAGGATGAAATTCAAAATAAAATAGAGGCCGCCAGACAAGAGAGTAAACAGCTTCATGCTCAAATAAATAAAGCA
AAACACAAGATACAAGATGCAAGCTTATTCCAGATGGCCAACAAAGTTACTTCGTTGACCAAAAATAAGATCAACTTAAAGCCAAATAT
CGTGTTGAAAGGCCATAATAATAAAATCTCAGATTTTCGGTGGAGTCGAGATTCAAAACGTATTTTGAGTGCAAGTCAAGATGGCTTTA
TGCTTATATGGGACAGTGCTTCAGGTTTAAAACAGAACGCTATTCCATTAGATTCTCAATGGGTTCTTTCCTGCGCTATTTCGCCATCG
AGTACTTTGGTAGCAAGCGCAGGATTAAACAATAACTGTACCATTTATAGAGTTTCGAAAGAAAACAGAGTAGCGCAAAACGTTGCGT
CAATTTTCAAAGGACATACTTGCTATATTTCTGACATTGAATTTACAGATAACGCACATATATTGACAGCAAGTGGGGATATGACATGT
GCCTTGTGGGATATACCGAAAGCAAAGAGGGTGAGAGAATATTCTGACCATTTAGGTGATGTTTTGGCATTAGCTATTCCTGAAGAGC
CAAACTCAGAAAATTCTTCGAACACATTCGCTAGCTGTGGATCAGACGGGTATACTTACATATGGGATAGCAGATCTCCGTCCGCTGT
ACAAAGCTTTTACGTTAACGATAGTGATATTAATGCACTTCGTTTTTTCAAAGACGGGATGTCGATTGTTGCAGGAAGTGACAATGGTG
CGATAAATATGTATGATTTAAGGTCGGACTGTTCTATTGCTACTTTTTCTCTTTTTCGAGGTTATGAAGAACGTACCCCTACCCCTACTT
ATATGGCAGCTAACATGGAGTACAATACCGCGCAATCGCCACAAACTTTAAAATCAACAAGCTCAAGCTATCTAGACAACCAAGGCGT
TGTTTCTTTAGATTTTAGTGCATCTGGAAGATTGATGTACTCATGCTATACAGACATTGGTTGTGTTGTGTGGGATGTATTAAAAGGAG
AGATTGTTGGAAAATTAGAAGGTCATGGTGGCAGAGTCACTGGTGTGCGCTCGAGTCCAGATGGGTTAGCTGTATGTACAGGTTCAT
GGGACTCAACCATGAAAATATGGTCACCAGGTTATCAAGAAGCTAGACATAAACAAAAGATTGTTGCTCCAGTTAAACAAACTTTGAA
CTTTGATTTGTTGAAATTGGCTGGTGATGTTGAATCTAATCCAGGGCCCACATCAGTTCAAAACTCTCCACGCTTACAACAACCTCAG
GAACAGCAACAGCAACAGCAACAGCTTTCCTTAAAGATAAAACAATTGAAGTTAAAAAGAATCAACGAACTTAACAATAAACTGAGGAA
AGAACTCAGCCGTGAAAGAATTACTGCTTCAAATGCATGTCTTACAATAATAAACTATACCTCGAATACAAAAGATTATACATTACCAG
AACTATGGGGCTACCCCGTAGCAGGATCAAATCATTTTATAGAGGGTTTGAAAAATGCTCAAAAAAATAGCCAAATGTCAAACTCAAAT
AGTGTTTGTTGTACGCTTATGTAACTCGAG**TGGC**TGAGACC

pWS926 – LexA-NLS [3a] CamR-ColE1

GGTCTCATATGAAGGCTTTAACTGCAAGACAACAGGAAGTTTTTGATTTGATAAGAGATCATATATCTCAAACTGGAATGCCACCAACA
AGAGCTGAAATAGCTCAAAGATTAGGATTTAGATCTCCAAACGCTGCAGAAGAACATTTGAAAGCTTTGGCAAGGAAAGGTGTCATCG
AAATTGTCTCTGGTGCTTCAAGAGGTATTAGACTATTACAAGAAGAGGAAGAAGGATTGCCATTGGTTGGTAGAGTTGCTGCTGGTGA
ACCATTATTAGCACAGCAACATATTGAGGGTCATTATCAAGTAGATCCATCTTTATTCAAACCAAATGCTGATTTTTTGTTGAGAGTAAG
TGGTATGTCTATGAAAGATATTGGTATCATGGATGGTGACTTATTGGCTGTTCATAAAACTCAAGATGTTAGAAACGGTCAAGTTGTCG
TTGCCAGAATTGATGATGAAGTCACTGTTAAAAGATTGAAGAAACAAGGTAATAAAGTGGAATTGTTACCAGAAAATTCTGAATTTAAG
CCAATTGTTGTTGATTTAAGGCAACAATCTTTTACTATTGAAGGTTTAGCTGTCGGTGTTATCAGAAATGGTGATTGGTTAGGTTCAGG
TTCTCCTCCAAAAAAAAAGAGAAAAGTTGGATCTGG**TTCT**TGAGACC

pWS928 – dCas9 [3a] CamR-ColE1

GGTCTCCT**ATG**GACAAGAAGTATTCTATCGGACTGGCTATCGGGACTAATAGCGTCGGGTGGGCCGTGATCACTGACGAGTACAAG
GTGCCCTCTAAGAAGTTCAAGGTGCTCGGGAACACCGACCGGCATTCCATCAAGAAAAATCTGATCGGAGCTCTCCTCTTTGATTCA
GGGGAAACCGCTGAAGCAACCCGCCTCAAGCGGACTGCTAGACGGCGGTACACCAGGAGGAAGAACCGGATTTGTTACCTTCAAGA
GATATTCTCCAACGAAATGGCAAAGGTCGACGACAGCTTCTTCCATAGGCTGGAAGAATCATTCCTCGTGGAAGAGGATAAGAAGCA
TGAACGGCATCCCATCTTCGGTAATATCGTCGACGAGGTGGCCTATCACGAGAAATACCCAACCATCTACCATCTTCGCAAAAAGCT
GGTGGACTCAACCGACAAGGCAGACCTCCGGCTTATCTACCTGGCCCTGGCCCACATGATCAAGTTCAGAGGCCACTTCCTGATCG
AGGGCGACCTCAATCCTGACAATAGCGATGTGGATAAACTGTTCATCCAGCTGGTGCAGACTTACAACCAGCTCTTTGAAGAGAACC
CCATCAATGCAAGCGGAGTCGATGCCAAGGCCATTCTGTCAGCCCGGCTGTCAAAGAGCCGCAGACTTGAGAATCTTATCGCTCAG
CTGCCGGGTGAAAAGAAAAATGGACTGTTCGGGAACCTGATTGCTCTTTCACTTGGGCTGACTCCCAATTTCAAGTCTAATTTCGACC
TGGCAGAGGATGCCAAGCTGCAACTGTCCAAGGACACCTATGATGACGATCTCGACAACCTCCTGGCCCAGATCGGTGACCAATAC
GCCGACCTTTTCCTTGCTGCTAAGAATCTTTCTGACGCCATCCTGCTGTCTGACATTCTCCGCGTGAACACTGAAATCACCAAGGCCC
CTCTTTCAGCTTCAATGATTAAGCGGTATGATGAGCACCACCAGGACCTGACCCTGCTTAAGGCACTCGTCCGGCAGCAGCTTCCGG
AGAAGTACAAGGAAATCTTCTTTGACCAGTCAAAGAATGGATACGCCGGCTACATCGACGGAGGTGCCTCCCAAGAGGAATTTTATA
AGTTTATCAAACCTATCCTTGAGAAGATGGACGGCACCGAAGAGCTCCTCGTGAAACTGAATCGGGAGGATCTGCTGCGGAAGCAG
CGCACTTTCGACAATGGGAGCATTCCCCACCAGATCCATCTTGGGGAGCTTCACGCCATCCTTCGGCGCCAAGAGGACTTCTACCCC
TTTCTTAAGGACAACAGGGAGAAGATTGAGAAAATTCTCACTTTCCGCATCCCCTACTACGTGGGACCCCTCGCCAGAGGAAATAGC
CGGTTTGCTTGGATGACCAGAAAGTCAGAAGAAACTATCACTCCCTGGAACTTCGAAGAGGTGGTGGACAAGGGAGCCAGCGCTCA
GTCATTCATCGAACGGATGACTAACTTCGATAAGAACCTCCCCAATGAGAAGGTCCTGCCGAAACATTCCCTGCTCTACGAGTACTTT
ACCGTGTACAACGAGCTGACCAAGGTGAAATATGTCACCGAAGGGATGAGGAAGCCCGCATTCCTGTCAGGCGAACAAAAGAAGGC
AATTGTGGACCTTCTGTTCAAGACCAATAGAAAGGTGACCGTGAAGCAGCTGAAGGAGGACTATTTCAAGAAAATTGAATGCTTCGAC
TCTGTGGAGATTAGCGGGGTCGAAGATCGGTTCAACGCAAGCCTGGGTACTTACCATGATCTGCTTAAGATCATCAAGGACAAGGAT
TTTCTGGACAATGAGGAGAACGAGGACATCCTTGAGGACATTGTCCTGACTCTCACTCTGTTCGAGGACCGGGAAATGATCGAGGAG
AGGCTTAAGACCTACGCCCATCTGTTCGACGATAAAGTGATGAAGCAACTTAAACGGAGAAGATATACCGGATGGGGACGCCTTAGC
CGCAAACTCATCAACGGAATCCGGGACAAACAGAGCGGAAAGACCATTCTTGATTTCCTTAAGAGCGACGGATTCGCTAATCGCAAC
TTCATGCAACTTATCCATGATGATTCCCTGACCTTTAAGGAGGACATCCAGAAGGCCCAAGTGTCTGGACAAGGTGACTCACTGCAC
GAGCATATCGCAAATCTGGCTGGTTCACCCGCTATTAAGAAGGGTATTCTCCAGACCGTGAAAGTCGTGGACGAGCTGGTCAAGGTG
ATGGGTCGCCATAAACCAGAGAACATTGTCATCGAGATGGCCAGGGAAAACCAGACTACCCAGAAGGGACAGAAGAACAGCAGGGA
GCGGATGAAAAGAATTGAGGAAGGGATTAAGGAGCTCGGGTCACAGATCCTTAAAGAGCACCCGGTGGAAAACACCCAGCTTCAGA
ATGAGAAGCTCTATCTGTACTACCTTCAAAATGGACGCGATATGTATGTGGACCAAGAGCTTGATATCAACAGGCTCTCAGACTACGA
CGTGGACGCAATCGTCCCTCAGAGCTTCCTCAAAGACGACTCAATTGACAATAAGGTGCTGACTCGCTCAGACAAGAACCGGGGAAA
GTCAGATAACGTGCCCTCAGAGGAAGTCGTGAAAAAGATGAAGAACTATTGGCGCCAGCTTCTGAACGCAAAGCTGATCACTCAGCG
GAAGTTCGACAATCTCACTAAGGCTGAGAGGGGCGGACTGAGCGAACTGGACAAAGCAGGATTCATTAAACGGCAACTTGTGGAGA
CTCGGCAGATTACTAAACATGTCGCCCAAATCCTTGACTCACGCATGAATACCAAGTACGACGAAAACGACAAACTTATCCGCGAGGT
GAAGGTGATTACCCTGAAGTCCAAGCTGGTCAGCGATTTCAGAAAGGACTTTCAATTCTACAAAGTGCGGGAGATCAATAACTATCAT
CATGCTCATGACGCATATCTGAATGCCGTGGTGGGAACCGCCCTGATCAAGAAGTACCCAAAGCTGGAAAGCGAGTTCGTGTACGG
AGACTACAAGGTCTACGACGTGCGCAAGATGATTGCCAAATCTGAGCAGGAGATCGGAAAGGCCACCGCAAAGTACTTCTTCTACAG
CAACATCATGAATTTCTTCAAGACCGAAATCACCCTTGCAAACGGTGAGATCCGGAAGAGGCCGCTCATCGAGACTAATGGGGAGAC
TGGCGAAATCGTGTGGGACAAGGGCAGAGATTTCGCTACCGTGCGCAAAGTGCTTTCTATGCCTCAAGTGAACATCGTGAAGAAAAC
CGAGGTGCAAACCGGAGGCTTTTCTAAGGAATCAATCCTCCCCAAGCGCAACTCCGACAAGCTCATTGCAAGGAAGAAGGATTGGG
ACCCTAAGAAGTACGGCGGATTCGATTCACCAACTGTGGCTTATTCTGTCCTGGTCGTGGCTAAGGTGGAAAAAGGAAAGTCTAAGA
AGCTCAAGAGCGTGAAGGAACTGCTGGGTATCACCATTATGGAGCGCAGCTCCTTCGAGAAGAACCCAATTGACTTTCTCGAAGCCA
AAGGTTACAAGGAAGTCAAGAAGGACCTTATCATCAAGCTCCCAAAGTATAGCCTGTTCGAACTGGAGAATGGGCGGAAGCGGATGC
TCGCCTCCGCTGGCGAACTTCAGAAGGGTAATGAGCTGGCTCTCCCCTCCAAGTACGTGAATTTCCTCTACCTTGCAAGCCATTACG
AGAAGCTGAAGGGGAGCCCCGAGGACAACGAGCAAAAGCAACTGTTTGTGGAGCAGCATAAGCATTATCTGGACGAGATCATTGAG
CAGATTTCCGAGTTTTCTAAACGCGTCATTCTCGCTGATGCCAACCTCGATAAAGTCCTTAGCGCATACAATAAGCACAGAGACAAAC
CAATTCGGGAGCAGGCTGAGAATATCATCCACCTGTTCACCCTCACCAATCTTGGTGCCCCTGCCGCATTCAAGTACTTCGACACCA
CCATCGACCGGAAACGCTATACCTCCACCAAAGAAGTGCTGGACGCCACCCTCATCCACCAGAGCATCACCGGACTTTACGAAACTC
GGATTGACCTCTCACAGCTCGGAGGGGATGGTTCAGGTTCTCCTCCAAAAAAAAAGAGAAAAGTTGGATCTGG**TTCT**TGAGACC

pWS925 – Gal4AD [3a] CamR-ColE1

GGTCTCAT**ATG**AAGTTGTTGTCTAGTATAGAACAGGCGTGTGATATTTGCAGATTAAAGAAATTGAAATGCTCTAAAGAAAAACCGAAG
TGTGCAAAATGTCTAAAAAATAACTGGGAATGTAGATATTCCCCGAAAACAAAAAGATCACCACTAACACGTGCTCATCTAACAGAAGT
CGAGAGTAGATTGGAAAGATTAGAACAGTTATTTCTTCTAATATTCCCACGTGAAGATCTTGATATGATATTGAAAATGGATAGTCTAC
AGGATATTAAAGCATTGTTAACTGGTCTATTTGTTCAAGACAATGTAAACAAAGATGCTGTCACTGATAGACTAGCTTCTGTTGAAACC
GATATGCCTCTAACCTTGCGTCAACATAGAATCAGTGCTACATCCTCTTCTGAAGAATCTAGTAACAAGGGTCAAAGACAACTAACAG
TCTCAGG**TTCT**TGAGACC

pWS931 – TetR-NLS [3a] CamR-ColE1

GGTCTCAT**ATG**TCTAGATTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGCTGCTTAATGAGGTCGGAATCGAAGGTTTAACAAC
CCGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCCTACATTGTATTGGCATGTAAAAAATAAGCGGGCTTTGCTCGACGCCTTAGC
CATTGAGATGTTAGATAGGCACCATACTCACTTTTGCCCTTTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAACGCTAAAAGTT
TTAGATGTGCTTTACTAAGTCATCGCGATGGAGCAAAAGTACATTTAGGTACACGGCCTACAGAAAAACAGTATGAAACTCTCGAAAA
TCAATTAGCCTTTTTATGCCAACAAGGTTTTTCACTAGAGAATGCATTATATGCACTCAGCGCTGTGGGGCATTTTACTTTAGGTTGCG
TATTGGAAGATCAAGAGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATGCCGCCATTATTACGACAAGCTAT
CGAATTATTTGATCACCAAGGTGCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGCGGATTAGAAAAACAACTTAAATGTG
AAAGTGGATCAGGTTCTCCTCCAAAAAAAAAGAGAAAAGTTGGATCTGG**TTCT**TGAGACC

pWS1808 – TetA-NLS [3a] CamR-ColE1

GGTCTCATATGTCCAGATTGGATAAGTCCAAAGTTATTAACGGTGCCTTGGAGTTGTTGAACGGTGTTGGTATTGAAGGTTTGACTAC
TAGAAAGTTGGCTCAAAAGTTGGGTGTTGAACAACCTACATTATACTGGCACGTTAAGAACAAAAGAGCTTTGTTGGACGCTTTGCCA
ATCGAAATGTTGGATAGACATCATACCCATTTCTGTCCATTGGAAGGTGAATCATGGCAAGATTTCTTGAGAAACAACGCCAAGTCTTT
CAGATGTGCTTTGTTGTCTCATAGAGATGGTGCTAAAGTTCACTTGGGTACTAGACCAACTGAAAAGCAATACGAAACTTTGGAAAAC
CAGTTGGCTTTCTTGTGTCAACAAGGTTTCTCTTTAGAAAACGCCTTGTATGCTTTGTCTGCTGTTGGTCATTTTACCTTGGGTTGTGT
TTTGGAAGAACAAGAACATCAAGTCGCCAAAGAAGAAAGGGAAACTCCAACTACTGATTCTATGCCACCATTATTGAGACAAGCCATT
GAGTTGTTTGATAGACAAGGTGCTGAACCAGCTTTTTTGTTTGGTTTGGAATTGATCATCTGCGGTTTGGAGAAACAATTGAAGTGTG
AATCTGGTTCTGGTTCTCCACCAAAAAAGAAAAGAAAGGTCGGTTCCGGTTCTTGAGACC

pWS930 – Z3E [3a] CamR-ColE1

GGTCTCATATGAGGCCCTACGCTTGCCCTGTAGAAAGTTGTGATAGGAGATTCAGCAGATCTGACGAGCTGACTAGGCATATTCGTA
TTCACACCGGCCAAAAGCCTTTCCAATGTAGAATCTGTATGAGAAACTTCAGCAGAAGTGACCATTTGACTACGCACATTAGAACCCA
CACCGGCGAGAAGCCGTTCGCGTGTGACATTTGTGGAAGAAAGTTCGCCAGATCTGATGAAAGAAAGAGACATACAAAGATCCATAC
AGGAGGTGGGGGCACTCCGGCAGCAGCCTCAACCCTAGAAGATCCAAGCGCTGGAGACATGAGGGCAGCCAACTTGTGGCCAAGC
CCATTAATGATCAAAAGATCTAAGAAGAATTCTCTTGCACTGTCTTTGACCGCTGATCAAATGGTTTCCGCACTACTTGATGCTGAACC
ACCTATATTGTACAGTGAATACGACCCAACTCGTCCTTTTTCCGAGGCTTCCATGATGGGTCTTTTGACAAATCTAGCAGACAGGGAA
TTAGTTCATATGATTAATTGGGCCAAAAGAGTTCCCGGCTTTGTTGACTTGACATTACATGACCAGGTACATCTTCTTGAATGCGCGTG
GTTAGAAATATTGATGATCGGATTAGTGTGGAGAAGTATGGAGCACCCTGTAAAATTATTATTTGCACCTAATTTATTATTGGACAGGA
ACCAAGGTAAATGCGTAGAAGGTATGGTTGAGATTTTTGATATGCTTCTTGCCACTAGCTCCAGGTTCAGAATGATGAATTTGCAAGG
AGAAGAATTCGTTTGCCTGAAATCAATTATTTTGTTAAATAGCGGTGTCTATACATTTTTAAGCTCAACTCTGAAAAGCCTGGAAGAGA
AAGACCATATACATAGAGTGTTGGACAAGATCACTGATACATTGATTCACCTTATGGCGAAAGCAGGATTGACTCTGCAGCAACAGCA
TCAGAGGCTAGCTCAGTTACTACTAATTTTGTCACACATTAGACATATGTCTAACAAGGGTATGGAGCACCTGTATTCTATGAAGTGTA
AAAACGTCGTTCCTCTGTATGATTTATTACTAGAAATGCTGGACGCTCACAGGTTGCATGCACCAACTTCTAGAGGCGGAGCCTCCGT
TGAAGAAACCGACCAATCCCACCTAGCGACAGCAGGGTCTACATCTAGCGGTTCTTGAGACC

pWS958 – SpyTag-NLS [3a] CamR-ColE1

GGTCTCATATGGCTCACATCGTTATGGTTGACGCTTACAAGCCAACTAAGGGTTCTGGTTCTCCACCAAAGAAGAAGAGAAAGGTTG
GTTCTGGTTCTTGAGACC

pWS932 – PRD [3b] CamR-ColE1

GGTCTCATTCTAGACCATCTAGTACAACAAAATCAGATAATTCGCCTCCAAAATTAGAAAGCGAGAATTTTAAGGATAATGAGTTGGTA
ACAGTAACTAATCAGCCGCTTTTAGGCGTTGGCCTCATGGATGACGATGCGCCAGAATCCCCCTCTCAAATTAATGATTTTATTCCTC
AGAAATTGATTATAGAACCCAATACTCTCGAATTGAATGGTCTAACAGAAGAAACGCCTCATGACTTACCCAAGAATACCGCTAAGGG
CAGAGATGAAGAAGATTTTCCTCTCGACTATTTTCCTGTATCTGTTGAATACCCTACGGAGGAAAATGCGTTTGATCCGTTCCCTCCAC
AGGCTTTTACGCCAGCTGCCCCTTCCATGCCTATTTCCTATGATAACGTGAATGAAAGGGATTCTATGCCCGTTAATTCTCTTCTTAAT
AGATACCCCTATCAGTTATCAGTGGCACCCACTTTCCCAGTGCCACCATCATCATCGAGGCAACATTTTATGACAAATCGGGATTTTTA
TTCATCTAACAATAACAAGGAAAAATTGGTATCTCCTAGCGACCCTACCAGCTACATGAAGTATGACGAACCAGTTATGGATTTTGATG
AATCTCGGCCAAATGAAAACTGTACAAATGCAAAATCTCACAACTCTGGCCAGCAAACTAAACAACACCAATTATATTCTAACAACTTC
CAGCAATCTTACCCAAACGGAATGGTTCCAGGATACTACCCAAAAATGCCGTATAATCCCATGGGGGGGGATCCTCTACTCGATCAA
GCCTTTTATGGCGCGGACGATTTTTTCTTTCCACCAGAAGGATGTGATAACAATATGCTGTATCCACAAACTGCAACTTCATGGAATGT
TTTGCCCCCTCAAGCTATGCAACCAGCTCCAACCTATGTTGGGAGGCCATACACACCGAATTATAGATCGACACCAGGTTCCGCGAT
GTTCCCATACATGCAAAGTTCAAATTCCATGCAGTGGAACACTGCTGTTTCACCTTATAGTTCGAGAGCACCATCTACAACTGCTAAAA
ACTATCCTCCTAGCACATTTTATTCTCAAAATATAAATCAATACCCACGGCGAAGAACTGTGGGAATGAAGTCATCACAAGGAAATGTT
CCAACAGGTAATAAACAATCTGTGGGCAAGTCTGCAAAAATTTCAAAGCCTCTACATATTAAGACAAGTGCTTATCAGAAGCAATACAA
AATCAACTTGGAAACGAAAGCCAGGCCAAGTGCTGGTGACGAAGATTCTGCTCATCCTGATAAGAACAAAGAAATTTCGATGCCTACT
CCGGATTCCAATACTTTGGTGGTCCAGTCAGAAGAAGGTGGAGCTCATTCACTTGAGGTAGATACCAATCGAAGGTCCGATAAAAAC
CTTCCAGATGCAACCGGATCCTGAGACC

pRC003 – GAL4AD [3b] CamR-ColE1 (Kindly provided by Robert Chen)

GGTCTCATTCTGCCAATTTTAATCAAAGTGGGAATATTGCTGATAGCTCATTGTCCTTCACTTTCACTAACAGTAGCAACGGTCCGAAC
CTCATAACAACTCAAACAAATTCTCAAGCGCTTTCACAACCAATTGCCTCCTCTAACGTTCATGATAACTTCATGAATAATGAAATCAC
GGCTAGTAAAATTGATGATGGTAATAATTCAAAACCACTGTCACCTGGTTGGACGGACCAAACTGCGTATAACGCGTTTGGAATCACT
ACAGGGATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATGATGAAGATACCCCACCAAACCCAAAAAAAGAGG
GATCCTGAGACC

pWS934 – B42AD [3b] CamR-ColE1

GGTCTCATTCTGGTATTAACAAGGACATTGAAGAGTGCAACGCTATAATCGAACAGTTTATTGACTATTTACGTACTGGTCAAGAAATG
CCTATGGAAATGGCAGATCAGGCTATCAACGTCGTCCCCGGCATGACACCAAAAACTATCCTGCATGCTGGTCCTCCTATACAGCCA
GATTGGTTAAAATCAAATGGCTTCCATGAGATAGAGGCTGACGTGAATGACACTTCTTTATTGTTGAGCGGCGATGCATCAGGATCCT
GAGACC

pWS935 – VP16AD [3b] CamR-ColE1

GGTCTCATTCTGAACTGCATTTGGACGGTGAGGACGTTGCCATGGCCCACGCTGACGCTTTAGACGATTTTGATTTAGATATGTTGG
GTGACGGGGACTCTCCAGGGCCTGGCTTCACTCCACATGATTCAGCACCTTACGGTGCGCTGGATATGGCTGACTTCGAATTTGAGC
AGATGTTCACCGATGCTCTGGGTATTGATGAGTATGGTGGATCCTGAGACC

pRC004 – VP64AD [3b] CamR-ColE1 (Kindly provided by Robert Chen)

GGTCTCATTCTGCTGATGCATTGGACGATTTTGATTTGGATATGCTGGGTAGTGACGCATTAGATGATTTCGATTTAGACATGCTTGGT
TCTGATGCCCTAGACGACTTTGACCTAGATATGTTAGGATCCGACGCCTTGGATGATTTCGACTTGGACATGTTGGGATTGGGATCCT
GAGACC

pGPY107 – MxiI [3b] CamR-ColE1 **(Kindly provided by Georgios Pothoulakis)**

GGTCTCA**TTCT**ATGGAACGTGTGAGAATGATTAATGTGCAAAGGCTGTTAGAAGCCGCAGAGTTTTTAGAAAGAAGAGAAAGAGAATG
CGAACACGGGTATGCCAGTTCTTTCCCTAGCATGCCCTCTCCCAGAGG**ATCC**TGAGACC

pWS957 – SpyCatcher [3b] CamR-ColE1

GGTCTCA**TTCT**GACTCAGCTACACATATCAAATTTTCTAAGAGAGATGAGGATGGTAAGGAATTAGCAGGAGCTACTATGGAGCTAAG
GGATAGTTCCGGCAAGACCATCTCAACTTGGATCTCTGATGGACAAGTTAAAGACTTTTACCTGTACCCTGGCAAGTATACTTTTGTC
GAGACAGCTGCACCTGATGGTTATGAAGTAGCAACAGCTATTACGTTCACGGTTAACGAGCAAGGACAAGTCACAGTTAACGGTGG**A
TCC**TGAGACC

pWS1546 – sfGFP BpiI Dropout Marker [234] CamR-ColE1

GGTCTCA**AACG**TAGTCTTCCAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCC
TAACTTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATATATAA
CATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAA
AGAATCCCAGCACCAAAATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCAC
AAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATG
TATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGA
AATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTA
CTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAACAAAAGATCTATGTCCAAG
GGTGAAGAGCTATTTACTGGGGTTGTACCCATTTTGGTAGAACTGGACGGAGATGTAAACGGACATAAATTCTCTGTTAGAGGTGAG
GGCGAAGGCGATGCCACCAATGGTAAATTGACTCTGAAGTTTATATGCACTACGGGTAAATTACCTGTTCCTTGGCCAACCCTAGTAA
CAACTTTGACATATGGTGTTCAATGTTTCTCAAGATACCCAGACCATATGAAAAGGCATGATTTCTTTAAAAGTGCTATGCCAGAAGGC
TACGTGCAAGAGAGAACTATCTCCTTTAAGGATGACGGTACGTATAAAACACGAGCAGAAGTGAAATTCGAAGGGGATACACTAGTTA
ATCGCATCGAATTAAAGGGTATAGACTTTAAGGAAGATGGTAATATTCTCGGCCATAAACTTGAGTATAATTTCAACTCGCATAATGTG
TACATTACAGCTGACAAACAAAAGAACGGAATTAAAGCGAATTTTAAAATCAGGCACAACGTCGAAGATGGGTCTGTTCAACTTGCCG
ATCATTATCAGCAAAACACCCCTATTGGTGATGGTCCAGTCTTGTTACCCGATAATCACTACTTAAGCACACAGTCTAGATTGTCAAAA
GATCCGAATGAAAAGCGTGATCACATGGTTTTATTGGAATTTGTCACCGCTGCAGGAATAACTCACGGAATGGACGAGCTTTATAAGG
GATCCTAACTCGAGATAAAGCAATCTTGATGAGGATAATGATTTTTTTTTGAATATACATAAATACTACCGTTTTTCTGCTAGATTTTGT
GATGACGTAAATAAGTACATATTACTTTTTAAGCCAAGACAAGATTAAGCATTAACTTTACCCTTTTCTTTCTAAGTTTCAATATTAGTTA
TCACTGTTTAAAAGTTATGGCGAGAACGTCGGCGGTTAAAATATATTACCCTGAACGGAAGACTG**GCTG**TGAGACC

pWS1547 – mScarlet BpiI Dropout Marker [234] CamR-ColE1

GGTCTCA**AACG**TAGTCTTCCAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCC
TAACTTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATATATAA
CATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAA
AGAATCCCAGCACCAAAATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCAC
AAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATG
TATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGA
AATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTA
CTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAACAAAAGATCTATGGTTTCT
AAAGGTGAAGCAGTTATTAAGGAATTCATGAGATTCAAGGTACACATGGAAGGTAGTATGAACGGTCACGAATTTGAAATTGAAGGTG
AAGGTGAAGGTAGACCATATGAAGGTACTCAAACTGCTAAGTTGAAGGTTACTAAAGGTGGTCCATTGCCATTTTCTTGGGATATCTT
GTCTCCACAATTCATGTACGGTTCTAGAGCTTTTACAAAACATCCAGCAGATATCCCAGATTACTACAAGCAATCATTCCCAGAAGGTT
TTAAATGGGAAAGAGTTATGAACTTCGAAGATGGTGGTGCAGTTACTGTTACACAAGATACTTCTTTGGAAGATGGTACATTGATCTAT
AAGGTTAAGTTGAGAGGTACTAATTTTCCACCAGATGGTCCAGTTATGCAAAAGAAAACTATGGGTTGGGAAGCTTCAACAGAAAGAT
TGTACCCAGAAGATGGTGTTTTGAAGGGTGACATCAAGATGGCATTGAGATTGAAGGATGGTGGTAGATATTTGGCTGATTTCAAGAC
TACATACAAGGCTAAGAAACCAGTTCAAATGCCAGGTGCTTACAACGTTGATAGAAAGTTGGATATCACTTCTCATAATGAAGATTACA
CAGTTGTTGAACAATATGAAAGAAGTGAAGGTAGACACTCAACTGGTGGTATGGACGAATTATACAAGTAACTCGAGATAAAGCAATC
TTGATGAGGATAATGATTTTTTTTTGAATATACATAAATACTACCGTTTTTCTGCTAGATTTTGTGATGACGTAAATAAGTACATATTACT
TTTTAAGCCAAGACAAGATTAAGCATTAACTTTACCCTTTTCTTTCTAAGTTTCAATATTAGTTATCACTGTTTAAAAGTTATGGCGAGAA
CGTCGGCGGTTAAAATATATTACCCTGAACGGAAGACTA**GCTG**TGAGACC

pWS1548 – AmilCP BpiI Dropout Marker [234] CamR-ColE1

GGTCTCA**AACG**TAGTCTTCCAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCC
TAACTTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATATATAA
CATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAA
AGAATCCCAGCACCAAAATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCAC
AAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATG
TATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGA
AATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTA
CTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAACAAAAGATCTATGAGTGTG
ATCGCTAAACAAATGACTTATAAGGTCTACATGTCAGGAACAGTGAATGGACACTATTTCGAAGTAGAGGGTGACGGTAAAGGTAAAC
CGTATGAGGGTGAACAGACGGTCAAACTTACCGTAACGAAGGGTGGACCCCTGCCTTTTGCCTGGGATATTCTTAGCCCACAGTGCC
AGTATGGTTCTATTCCTTTCACCAAATATCCCGAAGATATTCCAGACTATGTTAAACAAAGTTTTCCTGAGGGCTACACATGGGAGAGA
ATTATGAATTTTGAGGATGGAGCAGTATGTACAGTATCCAACGACTCCAGCATACAAGGAAATTGCTTTATTTATCATGTGAAGTTTAG
TGGTCTGAATTTCCCGCCGAACGGCCCTGTAATGCAGAAAAAAACGCAAGGCTGGGAGCCAAACACGGAGAGACTTTTCGCCCGTG
ATGGAATGCTGTTGGGGAACAACTTCATGGCCCTGAAGTTAGAAGGAGGAGGCCATTACTTATGTGAGTTTAAGACTACGTATAAAGC
TAAAAAACCTGTTAAGATGCCTGGATACCACTATGTCGATAGAAAACTTGACGTAACCAATCACAATAAAGACTATACGTCCGTCGAG
CAGTGTGAGATTAGTATTGCAAGGAAACCCGTTGTTGCGGGATCCTAACTCGAGATAAAGCAATCTTGATGAGGATAATGATTTTTTTT
TGAATATACATAAATACTACCGTTTTTCTGCTAGATTTTGTGATGACGTAAATAAGTACATATTACTTTTTAAGCCAAGACAAGATTAAG
CATTAACTTTACCCTTTTCTTTCTAAGTTTCAATATTAGTTATCACTGTTTAAAAGTTATGGCGAGAACGTCGGCGGTTAAAATATATTA
CCCTGAACGGAAGACTA**GCTG**TGAGACC

pWS2005 – sfGFP Dropout 500 bp Homology [234] CamR-ColE1

GGTCTCCAACGGGGAAGGTTTAAAGCTTATATGCGAGGCCCAACCTAGGCGTTATAATGCTATTAGGACTACACGAACGGTGTCCGA
GCGTGGTTTCGCTCAAGTGAGCACCTTAAATCTCATCACGCTTGGTAACACGGGGCAAACGCTGCCAAGCTACTATCCCAATTTATG
GACCGTGTACAGGGAGGACTGGCTGCTGAATAGTTCCCATCTATAGTCCCTGTGTTGCATCTCCGTTAGCATTCGTATGGCACTTACT
GCTGCGGACCGAAAACCCATGCAGCACCTTCACCGCCTAGCGATTATAAACGGAAAGCGTATAAGCTTTTTGTGGTAGTATCCGACT
ACAGGACAAGCGCCGGGTGGAGACTTAAATCATATGTGTTTGGGTCAACAGCTCCACTCGGCGGTCAGGGTCACTGTTCGTAAATTG
CCAAACACCCCCGCGGGGGCATAGCATCCAGCAGAGTTATTAGTTATCCCAGATGTTTTATACGGCGAGTCACCTTAGTCTTCCAGT
TCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTA
GCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACA
GTTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAAGAATCCCAGCACCAAAATATTGT
TTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACAGGCAAAAAACGGGCACAA
CCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTCTTACACCT
TCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAAGGTTGAAACCAGTTCCCTGAAATTATTCCCCTACTTGACTAATA
AGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAG
TTTTAAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAACAAAAGATCTATGTCCAAGGGTGAAGAGCTATTTACTGGGGT
TGTACCCATTTTGGTAGAACTGGACGGAGATGTAAACGGACATAAATTCTCTGTTAGAGGTGAGGGCGAAGGCGATGCCACCAATGG
TAAATTGACTCTGAAGTTTATATGCACTACGGGTAAATTACCTGTTCCTTGGCCAACCCTAGTAACAACTTTGACATATGGTGTTCAAT
GTTTCTCAAGATACCCAGACCATATGAAAAGGCATGATTTCTTTAAAAGTGCTATGCCAGAAGGCTACGTGCAAGAGAGAACTATCTC
CTTTAAGGATGACGGTACGTATAAAACACGAGCAGAAGTGAAATTCGAAGGGGATACACTAGTTAATCGCATCGAATTAAAGGGTATA
GACTTTAAGGAAGATGGTAATATTCTCGGCCATAAACTTGAGTATAATTTCAACTCGCATAATGTGTACATTACAGCTGACAAACAAAA
GAACGGAATTAAAGCGAATTTTAAAATCAGGCACAACGTCGAAGATGGGTCTGTTCAACTTGCCGATCATTATCAGCAAAACACCCCT
ATTGGTGATGGTCCAGTCTTGTTACCCGATAATCACTACTTAAGCACACAGTCTAGATTGTCAAAAGATCCGAATGAAAAGCGTGATC
ACATGGTTTTATTGGAATTTGTCACCGCTGCAGGAATAACTCACGGAATGGACGAGCTTTATAAGGGATCCTAACTCGAGATAAAGCA
ATCTTGATGAGGATAATGATTTTTTTTTGAATATACATAAATACTACCGTTTTTCTGCTAGATTTTGTGATGACGTAAATAAGTACATATT
ACTTTTTAAGCCAAGACAAGATTAAGCATTAACTTTACCCTTTTCTTTCTAAGTTTCAATATTAGTTATCACTGTTTAAAAGTTATGGCGA
GAACGTCGGCGGTTAAAATATATTACCCTGAACGGAAGACTGATACAGATCGTGAGCACCACTTGATTCTCTACGATGCCTCACGTC
CACACGAACTATTACGCTATCCTCTCTTACAGAGCCTAAGGCAGTTGAGAATCAATCTCAAGATGGATCGGCGCTATCTACCGTGGGT
GGGGAAACTCCTACTGACGATTCAATGTGTGACTGCCTGCCATGCAACTGTCTACCAGATCTGTGTCGGGTTCATATTATTACTCCCC
TCGTATACTTGCGCCACGGGAGGACGCACGGTCATCCGACTCCTGAAATCGTCGGCGAATATCTGCTAGAGGTCGTCTGTTGTGCAA
TATCGAAGAAGAGAGACACCCAACACATTGGTATCGTGAAACATTGCTGTTGGGTAGGAGGCGAAGAGCTACGCATGATTTCCGGCA
ATCACTATACCTTCGTACAGACAATGTGGTGAGTGACTTCCATGCCTCAGGTGTAAATTGGGACCTCTCGCGGGGAACGATCGGCGA
CTAGAAATGGTCGATTGCGCTGTGAGACC

pWS2006 – mScarlet Dropout 500 bp Homology [234] CamR-ColE1

GGTCTCCAACGGGGAAGGTTTAAAGCTTATATGCGAGGCCCAACCTAGGCGTTATAATGCTATTAGGACTACACGAACGGTGTCCGA
GCGTGGTTTCGCTCAAGTGAGCACCTTAAATCTCATCACGCTTGGTAACACGGGGCAAACGCTGCCAAGCTACTATCCCAATTTATG
GACCGTGTACAGGGAGGACTGGCTGCTGAATAGTTCCCATCTATAGTCCCTGTGTTGCATCTCCGTTAGCATTCGTATGGCACTTACT
GCTGCGGACCGAAAACCCATGCAGCACCTTCACCGCCTAGCGATTATAAACGGAAAGCGTATAAGCTTTTTGTGGTAGTATCCGACT
ACAGGACAAGCGCCGGGTGGAGACTTAAATCATATGTGTTTGGGTCAACAGCTCCACTCGGCGGTCAGGGTCACTGTTCGTAAATTG
CCAAACACCCCCGCGGGGGCATAGCATCCAGCAGAGTTATTAGTTATCCCAGATGTTTTATACGGCGAGTCACCTTAGTCTTCCAGT
TCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTA
GCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACA
GTTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCACCAAAATATTGT
TTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACAGGCAAAAAACGGGCACAA
CCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTCTTACACCT
TCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAAGGTTGAAACCAGTTCCCTGAAATTATTCCCCTACTTGACTAATA
AGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAG
TTTTAAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAACAAAAGATCTATGGTTTCTAAAGGTGAAGCAGTTATTAAGGAA
TTCATGAGATTCAAGGTACACATGGAAGGTAGTATGAACGGTCACGAATTTGAAATTGAAGGTGAAGGTGAAGGTAGACCATATGAAG
GTACTCAAACTGCTAAGTTGAAGGTTACTAAAGGTGGTCCATTGCCATTTTCTTGGGATATCTTGTCTCCACAATTCATGTACGGTTCT
AGAGCTTTTACAAAACATCCAGCAGATATCCCAGATTACTACAAGCAATCATTCCCAGAAGGTTTTAAATGGGAAAGAGTTATGAACTT
CGAAGATGGTGGTGCAGTTACTGTTACACAAGATACTTCTTTGGAAGATGGTACATTGATCTATAAGGTTAAGTTGAGAGGTACTAATT
TTCCACCAGATGGTCCAGTTATGCAAAAGAAAACTATGGGTTGGGAAGCTTCAACAGAAAGATTGTACCCAGAAGATGGTGTTTTGAA
GGGTGACATCAAGATGGCATTGAGATTGAAGGATGGTGGTAGATATTTGGCTGATTTCAAGACTACATACAAGGCTAAGAAACCAGTT
CAAATGCCAGGTGCTTACAACGTTGATAGAAAGTTGGATATCACTTCTCATAATGAAGATTACACAGTTGTTGAACAATATGAAAGAAG
TGAAGGTAGACACTCAACTGGTGGTATGGACGAATTATACAAGTAACTCGAGATAAAGCAATCTTGATGAGGATAATGATTTTTTTTTG
AATATACATAAATACTACCGTTTTTCTGCTAGATTTTGTGATGACGTAAATAAGTACATATTACTTTTTAAGCCAAGACAAGATTAAGCA
TTAACTTTACCCTTTTCTTTCTAAGTTTCAATATTAGTTATCACTGTTTAAAAGTTATGGCGAGAACGTCGGCGGTTAAAATATATTACC
CTGAACGGAAGACTGATACAGATCGTGAGCACCACTTGATTCTCTACGATGCCTCACGTCCACACGAACTATTACGCTATCCTCTCTT
ACAGAGCCTAAGGCAGTTGAGAATCAATCTCAAGATGGATCGGCGCTATCTACCGTGGGTGGGGAAACTCCTACTGACGATTCAATG
TGTGACTGCCTGCCATGCAACTGTCTACCAGATCTGTGTCGGGTTCATATTATTACTCCCCTCGTATACTTGCGCCACGGGAGGACG
CACGGTCATCCGACTCCTGAAATCGTCGGCGAATATCTGCTAGAGGTCGTCTGTTGTGCAATATCGAAGAAGAGAGACACCCAACAC
ATTGGTATCGTGAAACATTGCTGTTGGGTAGGAGGCGAAGAGCTACGCATGATTTCCGGCAATCACTATACCTTCGTACAGACAATGT
GGTGAGTGACTTCCATGCCTCAGGTGTAAATTGGGACCTCTCGCGGGGAACGATCGGCGACTAGAAATGGTCGATTGCGCTGTGAGC
ACC

pWS2007 – AmilCP Dropout 500 bp Homology [234] CamR-ColE1

GGTCTCCAACGGGGAAGGTTTAAAGCTTATATGCGAGGCCCAACCTAGGCGTTATAATGCTATTAGGACTACACGAACGGTGTCCGA
GCGTGGTTTCGCTCAAGTGAGCACCTTAAATCTCATCACGCTTGGTAACACGGGGCAAACGCTGCCAAGCTACTATCCCAATTTATG
GACCGTGTACAGGGAGGACTGGCTGCTGAATAGTTCCCATCTATAGTCCCTGTGTTGCATCTCCGTTAGCATTCGTATGGCACTTACT
GCTGCGGACCGAAAACCCATGCAGCACCTTCACCGCCTAGCGATTATAAACGGAAAGCGTATAAGCTTTTTGTGGTAGTATCCGACT
ACAGGACAAGCGCCGGGTGGAGACTTAAATCATATGTGTTTGGGTCAACAGCTCCACTCGGCGGTCAGGGTCACTGTTCGTAAATTG
CCAAACACCCCCGCGGGGGCATAGCATCCAGCAGAGTTATTAGTTATCCCAGATGTTTTATACGGCGAGTCACCTTAGTCTTCCAGT
TCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTA
GCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGGCGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACA
GTTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCACCAAAATATTGT
TTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACAGGCAAAAAACGGGCACAA
CCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTCTTACACCT
TCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAAGGTTGAAACCAGTTCCCTGAAATTATTCCCCTACTTGACTAATA
AGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAG
TTTTAAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAACAAAAGATCTATGAGTGTGATCGCTAAACAAATGACTTATAAG
GTCTACATGTCAGGAACAGTGAATGGACACTATTTCGAAGTAGAGGGTGACGGTAAAGGTAAACCGTATGAGGGTGAACAGACGGTC
AAACTTACCGTAACGAAGGGTGGACCCCTGCCTTTTGCCTGGGATATTCTTAGCCCACAGTGCCAGTATGGTTCTATTCCTTTCACCA
AATATCCCGAAGATATTCCAGACTATGTTAAACAAAGTTTTCCTGAGGGCTACACATGGGAGAGAATTATGAATTTTGAGGATGGAGC
AGTATGTACAGTATCCAACGACTCCAGCATACAAGGAAATTGCTTTATTTATCATGTGAAGTTTAGTGGTCTGAATTTCCCGCCGAACG
GCCCTGTAATGCAGAAAAAAACGCAAGGCTGGGAGCCAAACACGGAGAGACTTTTCGCCCGTGATGGAATGCTGTTGGGGAACAAC
TTCATGGCCCTGAAGTTAGAAGGAGGAGGCCATTACTTATGTGAGTTTAAGACTACGTATAAAGCTAAAAAACCTGTTAAGATGCCTG
GATACCACTATGTCGATAGAAAACTTGACGTAACCAATCACAATAAAGACTATACGTCCGTCGAGCAGTGTGAGATTAGTATTGCAAG
GAAACCCGTTGTTGCGGGATCCTAACTCGAGATAAAGCAATCTTGATGAGGATAATGATTTTTTTTTGAATATACATAAATACTACCGT
TTTTCTGCTAGATTTTGTGATGACGTAAATAAGTACATATTACTTTTTAAGCCAAGACAAGATTAAGCATTAACTTTACCCTTTTCTTTCT
AAGTTTCAATATTAGTTATCACTGTTTAAAAGTTATGGCGAGAACGTCGGCGGTTAAAATATATTACCCTGAACGGAAGACTGATACAG
ATCGTGAGCACCACTTGATTCTCTACGATGCCTCACGTCCACACGAACTATTACGCTATCCTCTCTTACAGAGCCTAAGGCAGTTGAG
AATCAATCTCAAGATGGATCGGCGCTATCTACCGTGGGTGGGGAAACTCCTACTGACGATTCAATGTGTGACTGCCTGCCATGCAAC
TGTCTACCAGATCTGTGTCGGGTTCATATTATTACTCCCCTCGTATACTTGCGCCACGGGAGGACGCACGGTCATCCGACTCCTGAA
ATCGTCGGCGAATATCTGCTAGAGGTCGTCTGTTGTGCAATATCGAAGAAGAGAGACACCCAACACATTGGTATCGTGAAACATTGCT
GTTGGGTAGGAGGCGAAGAGCTACGCATGATTTCCGGCAATCACTATACCTTCGTACAGACAATGTGGTGAGTGACTTCCATGCCTC
AGGTGTAAATTGGGACCTCTCGCGGGGAACGATCGGCGACTAGAAATGGTCGATTGCGCTGTGAGACC

pWS2061 – SpCas9 sgRNA Dropout [234] CamR-ColE1

GGTCTCCAACGATGTGCTTCAGTATTACATTTTTTGCCTTCAACGCCTTGATTGTTCTATTTTTGCTAATAATAAATCTATTTCATCGGA
CTAAAAGTCCATTAGTTGTAAGCGGATTTAGCTCAGTTGGGAGAGCGCCAGACTGAAGAAAAACTTCGGTCAAGTCATCTGGAGGTC
CTGTGTTCGATCCACAGAATTCGCAATCAATACTTATCTGGCTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCCCAGATTGA
GACGGAAAGTGAAACGTGATTTCATGCGTCATTTTGAACATTTTGTAAATCTTATTTAATAATGTGTGCGGCAATTCACATTTAATTTAT
GAATGTTTTCTTAACATCGCGGCAACTCAAGAAACGGCAGGTTCGGATCTTAGCTACTAGAGAAAGAGGAGAAATACTAGATGCGTAA
AGGCGAAGAGCTGTTCACTGGTGTCGTCCCTATTCTGGTGGAACTGGATGGTGATGTCAACGGTCATAAGTTTTCCGTGCGTGGCGA
GGGTGAAGGTGACGCAACTAATGGTAAACTGACGCTGAAGTTCATCTGTACTACTGGTAAACTGCCGGTTCCTTGGCCGACTCTGGT
AACGACGCTGACTTATGGTGTTCAGTGCTTTGCTCGTTATCCGGACCATATGAAGCAGCATGACTTCTTCAAGTCCGCCATGCCGGA
AGGCTATGTGCAGGAACGCACGATTTCCTTTAAGGATGACGGCACGTACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGATACCCT
GGTAAACCGCATTGAGCTGAAAGGCATTGACTTTAAAGAGGACGGCAATATCCTGGGCCATAAGCTGGAATACAATTTTAACAGCCA
CAATGTTTACATCACCGCCGATAAACAAAAAAATGGCATTAAAGCGAATTTTAAAATTCGCCACAACGTGGAGGATGGCAGCGTGCAG
CTGGCTGATCACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCACTATCTGAGCACGCAAAGCGTT
CTGTCTAAAGATCCGAACGAGAAACGCGATCATATGGTTCTGCTGGAGTTCGTAACCGCAGCGGGCATCACGCATGGTATGGATGAA
CTGTACAAATGACCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGC
TCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGTCTCAGTTTTAGAGCTAGAAATAGCAAGTTAAA
ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCG
GCTGGGCAACACCTTCGGGTGGCGAATGGGACTTTTATTTTTTGTCACTATTGTTATGTAAAATGCCACCTCTGACAGTATGGAACGC
AAACTTCTGTCTAGTGGATAGCTGTGAGACC

pWS2063 – LbCpf1 sgRNA Dropout [234] CamR-ColE1

GGTCTCCAACGATGTGCTTCAGTATTACATTTTTTGCCTTCAACGCCTTGATTGTTCTATTTTTGCTAATAATAAATCTATTTCATCGGA
CTAAAAGTCCATTAGTTGTAAGCGGATTTAGCTCAGTTGGGAGAGCGCCAGACTGAAGAAAAACTTCGGTCAAGTCATCTGGAGGTC
CTGTGTTCGATCCACAGAATTCGCAATCAATACTTGAAATTCTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCAATTTCTACT
AAGTGTAGATTGAGACGGAAAGTGAAACGTGATTTCATGCGTCATTTTGAACATTTTGTAAATCTTATTTAATAATGTGTGCGGCAATT
CACATTTAATTTATGAATGTTTTCTTAACATCGCGGCAACTCAAGAAACGGCAGGTTCGGATCTTAGCTACTAGAGAAAGAGGAGAAA
TACTAGATGCGTAAAGGCGAAGAGCTGTTCACTGGTGTCGTCCCTATTCTGGTGGAACTGGATGGTGATGTCAACGGTCATAAGTTTT
CCGTGCGTGGCGAGGGTGAAGGTGACGCAACTAATGGTAAACTGACGCTGAAGTTCATCTGTACTACTGGTAAACTGCCGGTTCCTT
GGCCGACTCTGGTAACGACGCTGACTTATGGTGTTCAGTGCTTTGCTCGTTATCCGGACCATATGAAGCAGCATGACTTCTTCAAGTC
CGCCATGCCGGAAGGCTATGTGCAGGAACGCACGATTTCCTTTAAGGATGACGGCACGTACAAAACGCGTGCGGAAGTGAAATTTG
AAGGCGATACCCTGGTAAACCGCATTGAGCTGAAAGGCATTGACTTTAAAGAGGACGGCAATATCCTGGGCCATAAGCTGGAATACA
ATTTTAACAGCCACAATGTTTACATCACCGCCGATAAACAAAAAAATGGCATTAAAGCGAATTTTAAAATTCGCCACAACGTGGAGGAT
GGCAGCGTGCAGCTGGCTGATCACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCACTATCTGAGC
ACGCAAAGCGTTCTGTCTAAAGATCCGAACGAGAAACGCGATCATATGGTTCTGCTGGAGTTCGTAACCGCAGCGGGCATCACGCAT
GGTATGGATGAACTGTACAAATGACCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTT
GTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGTCTCAGTTTTGGCCGGCATGG
TCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACACCTTCGGGTGGCGAATGGGACTTTTATTTTTTGTCACTATTGTTATGTAAAATG
CCACCTCTGACAGTATGGAACGCAAACTTCTGTCTAGTGGATAGCTGTGAGACC

pWS2011 – S/1 Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**TGGT**AGAGCCACAAACAGCCGGTACAAGCAACGATCTCCAGGACCATCTGA
ATCATGCGCGGATGACACGAACTCACGACGGCGATCACAGACATTAACCCACAGTACAGACACTGCGACAACGTGGCAATTCGTCG
CAATACAACGCGTCTCA**CTGA**TTTGCCGATAATTGCAGACGAACGCGGAATTTAAGTCAAATC**CCAA**TGAGACGACGGGGTCATCAC
GGCTCATCATGCGCCAAACAAATGTGTGCAATACACGCTCGGATGACTGCATGATGACCGCACTGACTGGGGACAGCAGATCCACC
TAAGCCTGTGAGAGAAGCAGACACCCGACAGATCAAGGCA**GTTA**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS2012 – 1/E Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**ACGG**GGTCATCACGGCTCATCATGCGCCAAACAAATGTGTGCAATACACGC
TCGGATGACTGCATGATGACCGCACTGACTGGGGACAGCAGATCCACCTAAGCCTGTGAGAGAAGCAGACACCCGACAGATCAAGG
CAGTTACGTCTCA**CCAA**TTTGATGTCAACACAGCTACAACGCGGAATTTTTAGTGGTCA**AGCA**TGAGACGGCTGGAAATCTGCTCGT
CAGTGGTGCTCACACTGACGAATCATGTACAGATCATACCGATGACTGCCTGGCGACTCACAACTAAGCAAGACAGCCGGAACCAGC
GCCGGCGAACACCACTGCATATATGGCATATCACAACAGT**CCAT**AGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC
pWS2013 – 1/2 Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**ACGG**GGTCATCACGGCTCATCATGCGCCAAACAAATGTGTGCAATACACGC
TCGGATGACTGCATGATGACCGCACTGACTGGGGACAGCAGATCCACCTAAGCCTGTGAGAGAAGCAGACACCCGACAGATCAAGG
CAGTTACGTCTCA**CCAA**TTTGATGTCAACACAGCTACAACGCGGAATTTTTAGTGGTCA**GATG**TGAGACGATAAATCGCAGCCAAGTG
AGTGAATAGATGACGCACCACGGTCAGACACGGCCACATCGTATCTCACAGGAGCAAGCGCGATAGGAGCACTCACACATAGTACG
GTGATCCGCTGACTCCTTTGCCCAAATAAGACGTG**AGCC**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS2014 – 2/E Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**ATAA**ATCGCAGCCAAGTGAGTGAATAGATGACGCACCACGGTCAGACACGG
CCACATCGTATCTCACAGGAGCAAGCGCGATAGGAGCACTCACACATAGTACGGTGATCCGCTGACTCCTTTGCCCAAATAAGACGT
GAGCCCGTCTCA**GATG**TTTGACTGGCTTAAGATGACAACGCGGAATTGTGCAACCGAT**AGCA**TGAGACGGCTGGAAATCTGCTCGT
CAGTGGTGCTCACACTGACGAATCATGTACAGATCATACCGATGACTGCCTGGCGACTCACAACTAAGCAAGACAGCCGGAACCAGC
GCCGGCGAACACCACTGCATATATGGCATATCACAACAG**TCCA**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS2015 – 2/3 Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**ATAA**ATCGCAGCCAAGTGAGTGAATAGATGACGCACCACGGTCAGACACGG
CCACATCGTATCTCACAGGAGCAAGCGCGATAGGAGCACTCACACATAGTACGGTGATCCGCTGACTCCTTTGCCCAAATAAGACGT
GAGCCCGTCTCA**GATG**TTTGACTGGCTTAAGATGACAACGCGGAATTGTGCAACCGAT**GTTC**TGAGACGGCAGACCTAGACCACATG
AGGCTGATGTAGGACAGCCACCAGTGGCAGCTAATCAACACCGCAAGATGCCGATGCACGCTCATATCATCGTCGTCAGCCTGGTA
GCCATTCGACATACGGATCAGGGAACTCGAACCCA**GTAC**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS2016 – 3/E Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**GCAG**ACCTAGACCACATGAGGCTGATGTAGGACAGCCACCAGTGGCAGCTA
ATCAACACCGCAAGATGCCGATGCACGCTCATATCATCGTCGTCAGCCTGGTAGCCATTCGACATACGGATCAGGGAACTCGAACCC
AGTACCGTCTCA**GTTC**TTTGGGGCACAGACAACCTAAACGCGGAATTATCTTCCTGCT**AGCA**TGAGACGGCTGGAAATCTGCTCGTC
AGTGGTGCTCACACTGACGAATCATGTACAGATCATACCGATGACTGCCTGGCGACTCACAACTAAGCAAGACAGCCGGAACCAGC
GCCGGCGAACACCACTGCATATATGGCATATCACAACAG**TCCA**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS2017 – 3/4 Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**GCAG**ACCTAGACCACATGAGGCTGATGTAGGACAGCCACCAGTGGCAGCTA
ATCAACACCGCAAGATGCCGATGCACGCTCATATCATCGTCGTCAGCCTGGTAGCCATTCGACATACGGATCAGGGAACTCGAACCC
AGTACCGTCTCA**GTTC**TTTGGGGCACAGACAACCTAAACGCGGAATTATCTTCCTGCT**GGTA**TGAGACGGCCTGCACTAGACGAACT
AGGCAAGATGCGTCCAATCCGTCTAAACATGGTGACAACGCTGGACAGATGACGTAACACCGAGCCACATCCTGAAATCGAGGCAG
GCTAACCGAAACCGTGACAATGCAAAGAGACAGCC**TGAC**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS2018 – 4/E Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**GCCT**GCACTAGACGAACTAGGCAAGATGCGTCCAATCCGTCTAAACATGGT
GACAACGCTGGACAGATGACGTAACACCGAGCCACATCCTGAAATCGAGGCAGGCTAACCGAAACCGTGACAATGCAAAGAGACAG
CCTGACCGTCTCA**GGTA**TTTGCCCATGAGTCACAATGAACGCGGAATTTTGCTAGCCAG**AGCA**TGAGACGGCTGGAAATCTGCTCGT
CAGTGGTGCTCACACTGACGAATCATGTACAGATCATACCGATGACTGCCTGGCGACTCACAACTAAGCAAGACAGCCGGAACCAGC
GCCGGCGAACACCACTGCATATATGGCATATCACAACAG**TCCA**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS2019 – 4/5 Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**GCCT**GCACTAGACGAACTAGGCAAGATGCGTCCAATCCGTCTAAACATGGT
GACAACGCTGGACAGATGACGTAACACCGAGCCACATCCTGAAATCGAGGCAGGCTAACCGAAACCGTGACAATGCAAAGAGACAG
CCTGACCGTCTCA**GGTA**TTTGCCCATGAGTCACAATGAACGCGGAATTTTGCTAGCCAG**AAGT**TGAGACGACGCACAAGGTCAGGG
CACTCATGCGACAATCAACTCGATGCATGATCCGCACCATTGTCGAGGGGCCAGCGTCAATAGTGCCGATGACCACAGACCCGGTT
AAGACATAGCCGAATGGAGCCGCGCCGACCACAGAAT**GATA**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS2020 – 5/E Spacer [12345] AmpR-ColE1
GCGGCCGCCCCTGAATTCGCATCTAGAGAAGACTA**ACGC**ACAAGGTCAGGGCACTCATGCGACAATCAACTCGATGCATGATCCGC
ACCATTGTCGAGGGGCCAGCGTCAATAGTGCCGATGACCACAGACCCGGTTAAGACATAGCCGAATGGAGCCGCGCCGACCACAG
AATGATACGTCTCA**AAGT**TTTGAAGCTCCACACAGTCGAACGCGGAATTCCTGATAAGTA**AGCA**TGAGACGGCTGGAAATCTGCTCG
TCAGTGGTGCTCACACTGACGAATCATGTACAGATCATACCGATGACTGCCTGGCGACTCACAACTAAGCAAGACAGCCGGAACCAG
CGCCGGCGAACACCACTGCATATATGGCATATCACAACAG**TCCA**TAGTCTTCACTAGTGCACTGCAGTACAGCGGCCGC

pWS1421 – ConR1 (BpiI) [5] CamR-ColE1

GGTCTCA**GCTGCCAA**TGAGACGACGGGGTCATCACGGCTCATCATGCGCCAAACAAATGTGTGCAATACACGCTCGGATGACTGCA
TGATGACCGCACTGACTGGGGACAGCAGATCCACCTAAGCCTGTGAGAGAAGCAGACACCCGACAGATCAAGGCA**GTTA**TAGTCTT
CACTAGTGCACTGCAG**TACA**TGAGACC

pWS1422 – ConR2 (BpiI) [5] CamR-ColE1

GGTCTCA**GCTGGATG**TGAGACGATAAATCGCAGCCAAGTGAGTGAATAGATGACGCACCACGGTCAGACACGGCCACATCGTATCT
CACAGGAGCAAGCGCGATAGGAGCACTCACACATAGTACGGTGATCCGCTGACTCCTTTGCCCAAATAAGACGTG**AGCC**TAGTCTTC
ACTAGTGCACTGCAG**TACA**TGAGACC

pWS1423 – ConR3 (BpiI) [5] CamR-ColE1

GGTCTCA**GCTGGTTC**TGAGACGGCAGACCTAGACCACATGAGGCTGATGTAGGACAGCCACCAGTGGCAGCTAATCAACACCGCAA
GATGCCGATGCACGCTCATATCATCGTCGTCAGCCTGGTAGCCATTCGACATACGGATCAGGGAACTCGAACCCA**GTAC**TAGTCTTC
ACTAGTGCACTGCAG**TACA**TGAGACC

pWS1424 – ConR4 (BpiI) [5] CamR-ColE1

GGTCTCA**GCTGGGTA**TGAGACGGCCTGCACTAGACGAACTAGGCAAGATGCGTCCAATCCGTCTAAACATGGTGACAACGCTGGAC
AGATGACGTAACACCGAGCCACATCCTGAAATCGAGGCAGGCTAACCGAAACCGTGACAATGCAAAGAGACAGCC**TGAC**TAGTCTTC
ACTAGTGCACTGCAG**TACA**TGAGACC

pWS1425 – ConR5 (BpiI) [5] CamR-ColE1

GGTCTCA**GCTGAAGT**TGAGACGACGCACAAGGTCAGGGCACTCATGCGACAATCAACTCGATGCATGATCCGCACCATTGTCGAGG
GGCCAGCGTCAATAGTGCCGATGACCACAGACCCGGTTAAGACATAGCCGAATGGAGCCGCGCCGACCACAGAAT**GATA**TAGTCTT
CACTAGTGCACTGCAG**TACA**TGAGACC

pWS1426 – ConRE (BpiI) [5] CamR-ColE1

GGTCTCA**GCTGAGCA**TGAGACGGCTGGAAATCTGCTCGTCAGTGGTGCTCACACTGACGAATCATGTACAGATCATACCGATGACTG
CCTGGCGACTCACAACTAAGCAAGACAGCCGGAACCAGCGCCGGCGAACACCACTGCATATATGGCATATCACAACAG**TCCA**TAGT
CTTCACTAGTGCACTGCAG**TACA**TGAGACC

pRC163 – Yeast Marker Spacer [6] CamR-ColE1 **(Kindly provided my Robert Chen)**

GGTCTCA**TACA**ACTAGAGATCTATGTGAGGATCCTAACTCGAGATCGA**GAGT**TGAGACC

pWS1208 – URA3 3' Homology (BpiI) [7] CamR-ColE1

GGTCTCA**GAGT**AGAGCACTTGAATCCACTGCCCCGGGAATCTCGGTCGTAATGATTTCTATAATGACGAAAAAAAAAAAAATTGGAAAG
AAAAAGCTTCATGGCCTTTATAAAAAGGAACTATCCAATACCTCGCCAGAACCAAGTAACAGTATTTTACGGGGCACAAATCAAGAAC
AATAAGACAGGACTGTAAAGATGGACGCATTGAACTCCAAAGAACAACAAGAGTTCCAAAAAGTAGTGGAACAAAAGCAAATGAAGG
ATTTCATGCGTTTGTACTCTAATCTGGTAGAAAGATGTTTCACAGACTGTGTCAATGACTTCACAACATCAAAGCTAACCAATAAGGAA
CAAACATGCATCATGAAGTGCTCAGAAAAGTTCTTGAAGCATAGCGAACGTGTAGGGCAGCGTTTCCAAGAACAAAACGCTGCCTTG
GGACAAGGCTTGGGCCGATAAGGTGTACTGGCGTATATATATCTAATTATGTATCTCTGGTGTAGCCCATTTTTAGCATGTAAATATA**A
AGA**TAGTCTTC**CCGA**TGAGACC

pWS1209 – LEU2 3' Homology (BpiI) [7] CamR-ColE1

GGTCTCA**GAGT**ACTCGTATCGCATGTCGGTGCGACACGAAATTACAAAATGGAATATGTTCATAGGGTAGACGAAACTATATACGCAA
TCTACATACATTTATCAAGAAGGAGAAAAAGGAGGATGTAAAGGAATACAGGTAAGCAAATTGATACTAATGGCTCAACGTGATAAGG
AAAAAGAATTGCACTTTAACATTAATATTGACAAGGAGGAGGGCACCACACAAAAAGTTAGGTGTAACAGAAAATCATGAAACTATGAT
TCCTAATTTATATATTGGAGGATTTTCTCTAAAAAAAAAAAAATACAACAAATAAAAAACACTCAATGACCTGACCATTTGATGGAGTTT
AAGTCAATACCTTCTTGAACCATTTCCCATAATGGTGAAAGTTCCCTCAAGAATTTTACTCTGTCAGAAACGGCCTTAACGACGTAGTC
GACCTCCTCTTCAGTACTAAATCTACCAATACCAAATCTGATGGAAGAATGGGCTAATGCATCATCCTTACCC**AGCG**TAGTCTTC**CCG
A**TGAGACC

pWS1451 – HO 3' Homology (BpiI) [7] CamR-ColE1

GGTCTCAGAGTTATCGTGTTGCATCTGCGGCTTTAAATTGATGTATCTCATCGCAGGCACGGGCAGTACAGTGCCCTGAGCGTAGGG
AAAAATGAAAAAAAGGATGTAACTTTTAACATAATTCCAGCACGCAGCGATTGGGTATAATGAAGATTGTTAAGTTCAACAACATGTGA
AATTCTTAGCTCCGCAACTTGCTTGAATTATGAGCTCTAAGATTCAAGAAGTAAATAGGTCATATAATGGAATGATAGTAGTACTTGAT
AAAGGAACCATGTGATCTTACGTTGATATGAAAATAATTCTCTCACAGAAGTCAGTTTGTATGACTAACATAGAGAGTAATTGCTTTTC
GGATAAAATCGCACCTAACAAAGATTACCTTTTGTAAAGCCTCCAGAACAGCTATGAATGTTCGTTATTCTACGCACTTTCTTTCATAC
ATCTTGATCCGCCTTTTTTGATGTAGGTTTTAACGCGGCAAAACATATTAATTTTTCTGATTTGGAAAAATAAAGAGTTAGTTC**TGCC**TA
GTCTTC**CCGA**TGAGACC

pWS1212 – URA3 5' Homology [8b] CamR-ColE1

GGTCTCA**CAAT**GAAGAGTA**GCTA**AATTCGAGTGAAACACAGGAAGATCAGAAAATCCTCATTTCATCCATATTAACAATAATTTCAAAT
GTTTATTTGCATTATTTGAAACTAGGCAAGACAAGCAACGAAACGTTTTTGAAAATTTTGAGTATTTTCAATAAATTTGTAGAGGACTCA
GATATTGAAAAAAAGCTACAGCAATTAATACTTGATAAGAAGAGTATTGAGAAGGGCAACGGTTCATCATCTCATGGATCTGCACATG
AACAAACACCAGAGTCAAACGACGTTGAAATTGAGGCTACTGCGCCAATTGATGACAATACAGACGATGATAACAAACCGAAGTTATC
TGATGTAGAAAAGGATTAAAGATGCTAAGAGATAGTGATGATATTTCATAAATAATGTAATTCTATATATGTTAATTACCTTTTTTGCGA
GGCATATTTATGGTGAAGGATAAGTTTTGACCATCAAAGAAGGTTAATGTGGCTGTGGTTTCAGGGTCCATAAAGCCCACATGGATAA
CATTAC**CCCT**TGAGACC

pWS1212 – URA3 5' Homology [8b] CamR-ColE1

GGTCTCA**CAAT**GAAGACTA**ACCC**AGAAGAACAGTAAAATAAAGCAAGGTACGTGAAATTAATATTTTTAAATGGTTCTAACCGATGCC
GAAGAACTGCGCAGTCCGGTTATAACGTCTGACATGTCCTTTTTTGATTTGGAATCCAACCACTCAAGTGACTCTGTTCATTTACTTTG
CGAAAAATATACCCACAAATTGCCCATCGAAAGTGAATCGCAAACCACCTTCAGACTGGCACCGACAAAGCAAAGATTATACAGACAG
AGTACTTTATACGTACCGTTAAGTCTCAAGCAAAGGGTTTTCTTATTTACTGAACGGGTAAAGAGTATCTGGGCCGGCTTGCCAAGAT
GCAAACCGAATAAGTATTTCAAAGTTGCATTTGCCTTAGCCGTCCTGACACCATTGGCTATTTGGATATTTTATATTGACTTTCGTGTA
CATTGATCACATCGACTGTTCTATTGGCAAATGAACCACGGGCATTGACTATTTTTCAGGTTACTACTATATATTATCATCACGGGCAA
GGATTGTA**CCCT**TGAGACC

pWS1452 – HO 5' Homology [8b] CamR-ColE1

GGTCTCA**CAAT**GAAGACTA**TCTA**ATTGTATCGAGATCACTTTTCGTGATCCGCTAATCAGCGACGGTCACATTAGGTTTGCCAAGTCA
GGGTATGAACCATACGATCAGTTTTCGTGAACCTGGTACGTATATTGTGGCGTTTGTGTATATTTTCATTCTTTGACAACAATCAATAC
CAACCTCAAATAGGAAAAGTAATAAGTTTGGCGTTACACCCCAAAAGACGCCAAACGGATCGAACTTACTCAATAGCAATTAGCGAGA
CAAAACCTACGTTAAGACCTGTAACCGATTTATCAAAGCACTCTGCGGTTCTTTCTTGGGAATATTACCTGGACATTTTGTGCCCTCAA
GAAACGAGGCTCTACGAGCCTGTTGGAGCCCCTCAGACATTAGCCGCCACGAATCAAACTTTTTACGCGATTCGGCCCAAATCAGTT
TCTCACAGATCATTCGTAGAGTGAAAAAGCACATCGATTATTTGATACCCCTTTGGGTTAATTACTGTTGAGGTCTTTGCCGTAAGTTC
CATCGCTGT**CCCT**TGAGACC

pWS2024 – 500 bp Gap Repair Vector [cas] AmpR-ColE1

GCGGCCGCGAAGACTA**GGGA**AGGTTTAAAGCTTATATGCGAGGCCCAACCTAGGCGTTATAATGCTATTAGGACTACACGAACGGT
GTCCGAGCGTGGTTTCGCTCAAGTGAGCACCTTAAATCTCATCACGCTTGGTAACACGGGGCAAACGCTGCCAAGCTACTATCCCAA
TTTATGGACCGTGTACAGGGAGGACTGGCTGCTGAATAGTTCCCATCTATAGTCCCTGTGTTGCATCTCCGTTAGCATTCGTATGGCA
CTTACTGCTGCGGACCGAAAACCCATGCAGCACCTTCACCGCCTAGCGATTATAAACGGAAAGCGTATAAGCTTTTTGTGGTAGTATC
CGACTACAGGACAAGCGCCGGGTGGAGACTTAAATCATATGTGTTTGGGTCAACAGCTCCACTCGGCGGTCAGGGTCACTGTTCGT
AAATTGCCAAACACCCCCGCGGGGGCATAGCATCCAGCAGAGTTATTAGTTATCCCAGATGTTTTATACGGCGAGTCACCT**AACG**TC
AGACCAGTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACGGATCTGAAAGAGGAGAAAGGATCTA
TGGTTAGCAAAGGCGAGGCGGTTATCAAGGAGTTTATGCGTTTTAAGGTTCACATGGAGGGTAGCATGAATGGTCACGAGTTCGAGA
TCGAGGGTGAAGGCGAGGGTCGTCCGTACGAAGGCACCCAGACCGCGAAGCTGAAAGTGACCAAGGGTGGCCCGCTGCCGTTCAG
CTGGGACATCCTGAGCCCGCAGTTCATGTATGGCAGCCGTGCGTTTACCAAACACCCGGCGGACATTCCGGATTACTATAAGCAAAG
CTTCCCGGAAGGTTTTAAATGGGAGCGTGTTATGAACTTCGAAGATGGTGGCGCGGTGACCGTTACCCAGGACACCAGCCTGGAGG
ATGGCACCCTGATTTACAAGGTGAAACTGCGTGGCACCAACTTTCCGCCGGATGGTCCGGTTATGCAGAAGAAAACGATGGGTTGG
GAAGCGAGCACCGAGCGTCTGTATCCGGAAGATGGCGTGCTGAAGGGTGATATCAAAATGGCGCTGCGTCTGAAGGACGGTGGCC
GTTACCTGGCGGATTTTAAGACCACCTATAAAGCGAAGAAACCGGTGCAAATGCCGGGTGCGTACAACGTTGACCGTAAACTGGATA
TTACCAGCCACAACGAGGATTATACCGTGGTTGAGCAATATGAGCGTAGCGAGGGTCGCCACAGCACCGGCGGCATGGACGAACTG
TATAAGTAATAAGGATCTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGT
GAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAAGTCGGTCTCA**GCTG**ATACAGATCGTGA
GCACCACTTGATTCTCTACGATGCCTCACGTCCACACGAACTATTACGCTATCCTCTCTTACAGAGCCTAAGGCAGTTGAGAATCAAT
CTCAAGATGGATCGGCGCTATCTACCGTGGGTGGGGAAACTCCTACTGACGATTCAATGTGTGACTGCCTGCCATGCAACTGTCTAC
CAGATCTGTGTCGGGTTCATATTATTACTCCCCTCGTATACTTGCGCCACGGGAGGACGCACGGTCATCCGACTCCTGAAATCGTCG
GCGAATATCTGCTAGAGGTCGTCTGTTGTGCAATATCGAAGAAGAGAGACACCCAACACATTGGTATCGTGAAACATTGCTGTTGGG
TAGGAGGCGAAGAGCTACGCATGATTTCCGGCAATCACTATACCTTCGTACAGACAATGTGGTGAGTGACTTCCATGCCTCAGGTGT
AAATTGGGACCTCTCGCGGGGAACGATCGGCGACTAGAAATGGTCGA**TTGC**TAGTCTTCGCGGCCGC

8.5.2 sgRNAs used in this study

Supplementary Table S9. sgRNA targets used in this study.

Target	Guide	CRISPR/Cas system
GFP (Marker cycling)	CTAGTAACAACTTTGACATA	SpCas9
BFP (Marker cycling)	CTAGTAACAACTTTATCTCA	SpCas9
***SST2* 5'**	AAGATAGAGTTGTAAGATGG	SpCas9/LbCpf1
***SST2* 3'**	AATGAAATTAGCACTTTTCT	SpCas9/LbCpf1
***FAR1* 5'**	TTTTCAAACGAAACTCTTGT	SpCas9/LbCpf1
***FAR1* 3'**	CTAGAGGTTGGGAACTTCCA	SpCas9/LbCpf1
***BAR1* 5'**	TCTTTGTTTGAAACTTATTT	SpCas9
***BAR1* 3'**	GTTATGACTGTCTTATGAGT	SpCas9
***STE2* 5'**	TTGCAACTCATCGAAAGTGA	SpCas9
***STE2* 3'**	GCTGATGCAAGTTACAAAGA	SpCas9
***STE12* 5'**	AATAACCAATAGTAGAACAG	SpCas9
***STE12* 3'**	GTTTTTATCGGACCTTCGAT	SpCas9
***GPA1* 5'**	ACAGTGAGTACGCAAACAAT	SpCas9
***GPA1* 3'**	GTTTTGCTGGATGATTAGAT	SpCas9
***MF(ALPHA)1* 5'**	AAAACTGCAGTAAAAATTGA	SpCas9
***MF(ALPHA)1* 3'**	CATTGGTTGCAGTTAAAACC	SpCas9
***MF(ALPHA)2* 5'**	CGCTAAAATAAAAGTGAGAA	SpCas9
***MF(ALPHA)2* 3'**	CACTGGTTGCAACTCAAGCC	SpCas9
MFA1	GGTCTTTTCTTTTGGAGCGG	SpCas9
MFA2	ATAGTTGTCTTTCTTTTCAG	SpCas9
***STE3* 5'**	CATACAAGTCAGCAATAATA	SpCas9
***STE3* 3'**	CATAGTTCAGAAAATACTGC	SpCas9
***GPR1* 5'**	TTCAACGCGTTTAAATTCGG	SpCas9
***GPR1* 3'**	AATGATAGTAGTGATAGTAG	SpCas9
***GPA2* 5'**	TGCGCATCTTCAGAAAAGAA	SpCas9
***GPA2* 3'**	GAAAATACATTGAAAGACTC	SpCas9
URA3	TCAGGGTCCATAAAGCTCCC	SpCas9
***LEU2* 1/2**	TTACTACTATATATTATTGT	SpCas9
***LEU2* 2/2**	CCTCTAATCATGAATGTTCT	SpCas9
HO	GCTCCAGCATTATAGCATGC	SpCas9
***URA3* LP**	ATATTATTGTACACCTACCG	SpCas9
***LEU2* LP**	GCATCAGGTGGACTAGCATG	SpCas9
***HO* LP**	ATGGACGAAATGCTTCACCA	SpCas9
***STE2* KO LP**	CTAGCTTTCGTGTTAGTACG	SpCas9
***STE12* KO LP**	CATCGCTTCCTACTTCCGCT	SpCas9
***GPA1* KO LP**	TAGCATGGTGACACAAGCAG	SpCas9
CRISPR UAS	CGGTTGTTACACATCTACCG	dCas9-PRD
***ALD6* g1**	TATAAATGTAATAAGAAGTT	dCas9-PRD
***ALD6* g2**	AAGAACTTGTTAAACACGCC	dCas9-PRD
***ALD6* g3**	ACGCCAGGCTTGACCTCGAA	dCas9-PRD
***ALD6* g4**	ACACCGTTCGAGGTCAAGCC	dCas9-PRD
***ALD6* g5**	ACACATCAAAACACCGTTCG	dCas9-PRD
***ALD6* g6**	CAGCTCAAACAGCGATTTAA	dCas9-PRD
***ALD6* g7**	GCTGTTTGAGCTGACTAACA	dCas9-PRD
***ALD6* g8**	GTAATAAATTCGGGGTGAGG	dCas9-PRD

8.5.3 Gpa1 C-terminal Gα transplants

To generate the Gpa1-Gα chimeras, annealed oligos were assembled into the Gpa1 c-terminal truncation entry vector, pWS936, following the small fragment assembly protocol in section **7.2.3.4**.

Supplementary Table S10. Gpa1-Gα C-terminal transplants.

G protein	C-terminal AAs	DNA sequence	Sense oligo	Antisense oligo
Gpa1	KIGII	aaaattggtattatt	GAAAaaaattggtattattTAACTCGAG	GCCACTCGAGTTAaataataccaatttt
Gαs/olf	QYELL	caatatgagctactt	GAAAcaatatgagctacttTAACTCGAG	GCCACTCGAGTTAaagtagctcatattg
Gα12	DIMLQ	gatattatgttgcaa	GAAAgatattatgttgcaaTAACTCGAG	GCCACTCGAGTTAttgcaacataatatc
Gα13	QLMLQ	caattgatgttgcaa	GAAAcaattgatgttgcaaTAACTCGAG	GCCACTCGAGTTAttgcaacatcaattg
Gαi1/2 (Gnat)	DCGLF	gattgtggtttgttt	GAAAgattgtggtttgtttTAACTCGAG	GCCACTCGAGTTAaaacaaaccacaatc
Gαi3	ECGLY	gaatgtggtttgtat	GAAAgaatgtggtttgtatTAACTCGAG	GCCACTCGAGTTAatacaaaccacattc
Gαz	YIGLC	tatattggtttgtgt	GAAAtatattggtttgtgtTAACTCGAG	GCCACTCGAGTTAacacaaaccaatata
Gα15/16	EINLL	gaaattaatttgttg	GAAAgaaattaatttgttgTAACTCGAG	GCCACTCGAGTTAcaacaaattaatttc
Gαq/11	EYNLV	gagtataacttggtc	GAAAgagtataacttggtcTAACTCGAG	GCCACTCGAGTTAgaccaagttatactc
Gα14	EFNLV	gaatttaatttggtt	GAAAgaatttaatttggttTAACTCGAG	GCCACTCGAGTTAaaccaaattaaattc
Gαo	GCGLY	ggttgtggtttgtat	GAAAggttgtggtttgtatTAACTCGAG	GCCACTCGAGTTAatacaaaccacaacc
Sp gpa1	QSLMF	caatctttgatgttt	GAAAcaatctttgatgtttTAACTCGAG	GCCACTCGAGTTAaaacatcaaagattg

8.5.4 Gene KO and landing pad sequences

100 bps up- and downstream of CRISPR-Cas9 landing pad at gene KOs and plasmid integration sites. Yellow – CRISPR landing pad. Bold – PAM sequence. Red highlight – native stop codon of deleted gene. Sequences shown in orientation to the deleted ORF.

***SST2* KO**
TAACTTTGGAGGTGTTACTGTCGTACGTTCCTTCTAGGTTTTGCACGCACTATCTGAGGCGTTATAGGTTCAATTTGGTAATTAAAGAT
AGAGTTGTAAGAATGCAATCGTAGTCCACCT**CGG**TTAATTTCATTGAGAGTCTTACTCATCTTCAGGTACAATTGCACAAACAGTCCTT
TTTTTTTTCTTTAGTTCTCCTAACCTAATATGTCTTGATACCCATA

***FAR1* KO**
TCAAAAAATTTCTATTTACTTTTATATTTCTTGACCATCCTTTACACAAAGTCTATAGATCCACTGGAAAGCTTCGTGGGCGTAAGAAGG
CAATCTATTAGATCGTACTTAGAAATGAGG**CGG**TTAGTAGTTCGGGAATCGAGGCCCGTATTTCGAGGCTTTTGCTTTTCCTTTTTTTT
TTTTCGTTTCTCCACGTCTATACTACGCAATGACTGAATATATAT

***BAR1* KO**
CATGATGAATTCTTTAATGATCTTCGCGTGATTTAATTCTAGTGGTTCGTATCGCCTAAAATCATACCAAAATAAAAAGAGTGTCTAGAA
GGGTCATATAAATGGGGTTAGCAAGTCGCA**CGG**TTAAGAAATCTGGAGTACAATTTTTTTATAGCATATAAATATCAAATATATAGTCA
TTTTTAATACATGGAAAGCATAATAAAAAAACAAGGGGAGTTTTA

***STE2* KO**
TTCAAAGCAATACGATACCTTTTCTTTTCACCTGCTCTGGCTATAATTATAATTGGTTACTTAAAAATGCACCGTTAAGAACCATATCCA
AGAATCAAAACTAGCTTTCGTGTTAGTACG**CGG**TTGATCAAAATTTACGGCTTTGAAAAAGTAATTTCGTGACCTTCGGTATAAGGTTA
CTACTAGATTCAGGTGCTCATCAGATGCACCACATTCTCTATAAA

***GPA1* KO**
AGACCAAACTGAGTAGAAGCTATTCATACTGTAAATTGGTATTTTAGCATCACATCAATAATCCAGAGGTGTATAAATTGATATATTAAG
GTAGGAAATATAGCATGGTGACACAAGCAG**CGG**TTGAAGGAACTGTATAATTAAAGTAGTGTTTAGATACGTAAATTCTGTTTCCGAA
GATGCAAGAAGGAGCAGCAGCACCAGAAAAAATTACTATTTTTCTT

***STE12* KO**
CTAGAGTGGATATTGATATTTCTCAAACAAGACTCGTCGAAGAAAACACACTTTTATAGCGGAACCGCTTTCTTTATTTGAATTGTCTT
GTTCACCAAGGCATCGCTTCCTACTTCCGCT**CGG**TTGATATAATATAATTTTTGAATTTATGATACAAGAATTAAAAATGCGGGCCAGA
ATTTAATATTAAACAATACTCAGAAGAAAACAACAAGGACAATCTG

***STE3* KO**
TATTCATTTTGTGCAGTATTCACATATTCTATTTTATTGCTTTTTAACTTTAGAGGCAATTAAATTTGTGTAGGAAAGGCAAAATACTATC
AAAATTTTCAATGTTTCTTGTCCAAGCGG**CGG**TTAACACAAGAGTGTCGCATTATATTTACTGGACTAGGAGTATTTTATTTTTACAGGA
CTAGGATTGAAATACTGCTTTTTAGTGAATTGTGGCTCAAATA

***MF(ALPHA)1* KO**
AAAATGTTACTGTTCTTACGATTCATTTACGATTCAAGAATAGTTCAAACAAGAAGATTACAAACTATCAATTTCATACACAATATAAAC
GATTAAAAGAACACGAGTTCCCAAAACCAG**CGG**TTAAGCCCGACTGATAACAACAGTGTAGATGTAACAAAGTCGACTTTGTTCCCAC
TGTACTTTTAGCTCGTACAAAATACAATATACTTTTCATTTCTCCG

***MF(ALPHA)2* KO**
GTTTGAGGTGTCCTTCCTATATCTGTTTTTATATTCTATATAATGGATAATTACTACCATCACCTGCATCAAATTCCAGTAAATTCACATA
TTGGAGAAAGTTCCGATAGGCCAGCATAT**CGG**TTGAAAAATGACCCTAAACTACTTCTAAACCCTCTCGATTTCTTTCACGTTCATACA
ACACCTAGTTTTATTTATTTTCTTTTCAATCTGAGTAGTTGAGT

***MFA1* KO**
TAGAGTCTTCATATATAAACCGCCAGAAATGAATTAATGAGAGGGATCTGTAACTGTTTCTCGGATAAAACCAAAATAAGTACAAAGCC
ATCGAATAGAAGCAGTAACGCTCATCAGCTA**CGG**TTAGTTTCTGCGTACAAAAACGTTGTTCTCCCTCCTTTATCTTCCTTTTCCGCTA
CACCAATATATCATGTTTGTTCGTAATATTTCTTTTTAGACCTAAT

***MFA2* KO**
TTTTATTTCCATCCACTTCTTCTGTCGTTCATCCGTTCATTGACATCACTAGAGACACCAGCGAGCTATCATCTTCATACAACAATAACT
ACCAACCTTACTTCTCCTGGAGATCAAGGA**CGG**TTAATTTTTGACGACAACCAAGAGGTCAAATCAATATCTACCCTTTCATTTATTAC
GTGTTGCTGGCAAACTAATTTATTCCAATTCTCTCATCATTAGCT

***GPR1* KO**
TGTGTGTGTGTCTATAAAAAGCAGTAAGAGTCACCAAAAAAAAAAAAACGACAAACAAGTGATCCGAAGTGTGACGAATAAAGCAAACT
CTCCAACTCAAATCTAACCGTCGACTTTGGCG**CGG**TTAAAGTTTTTGTATCGCGATGTTTGAAAATGGAAAGTAAGGAACGTAATACA
AATTGACAAGTAGCCGACATGAATGACGCTCACTTCTCTTATATATGT

***GPA2* KO**
ACTACCCAAAGAGCAATCGATAGGTATAAAAGTGAGCAATTGCTATCACAGCGAGCCTTATTGTTACAGCACAAATCACGCGTATTTT
CAAGCAAATATCGCTGTTATCCTGCATCGGAA**CGG**TTGAATGCACAGCTAAAACAGAGACAAAACTGCATGCCTCTTCTCCCCTTTAT
TATCACCTTTAAAAAAGATAAAAAAAGAAACTGGAAAAAAGGTAAAAA

***URA3* LP**
ATATATGTTAATTACCTTTTTTGCGAGGCATATTTATGGTGAAGGATAAGTTTTGACCATCAAAGAAGGTTAATGTGGCTGTGGTTTCA
GGGTCCATAAAATATTATTGTACACCTACCG**CGG**TCCCGGGAATCTCGGTCGTAATGATTTCTATAATGACGAAAAAAAAAAAATTGG
AAAGAAAAAGCTTCATGGCCTTTATAAAAAGGAACTATCCAATACCT

***LEU2* LP**
TATTTTATATTGACTTTCGTGTACATTGATCACATCGACTGTTCTATTGGCAAATGAACCACGGGCATTGACTATTTTTCAGGTTACTAC
TATATATTATGCATCAGGTGGACTAGCATG**CGG**TCGACACGAAATTACAAAATGGAATATGTTCATAGGGTAGACGAAACTATATACG
CAATCTACATACATTTATCAAGAAGGAGAAAAAGGAGGATGTAAAG

***HO* LP**
CGCGATTCGGCCCAAATCAGTTTCTCACAGATCATTCGTAGAGTGAAAAAGCACATCGATTATTTGATACCCCTTTGGGTTAATTACTG
TTGAGGTCTTTATGGACGAAATGCTTCACCA**CGG**TTTTAAATTGATGTATCTCATCGCAGGCACGGGCAGTACAGTGCCCTGAGCGTA
GGGAAAAATGAAAAAAAGGATGTAACTTTTAACATAATTCCAGCACG

8.5.5 Primers used in this study

Supplementary Table S11. Primers used in this study.

Primer description	Sequence
SST2 KO seq./colony PCR FP	CGAATTTTTGAAGGTCTTTCC
SST2 KO colony PCR RP	CGAGGTGGACTACGATTGC
SST2 KO seq. RP	ATCAAAGAATCACCCAATTCC
FAR1 KO seq./colony PCR FP	CATTCATACGATGGTGAACAG
FAR1 KO colony PCR RP	CCGCCTCATTTCTAAGTACG
FAR1 KO seq. RP	AGTACACGCTGACCCGTT
BAR1 KO seq./colony PCR FP	TGGAAGGTCGTAGCAAGG
BAR1 KO colony PCR RP	CGTGCGACTTGCTAACC
BAR1 KO seq. RP	AATACCGGGGTGTCTTGAC
STE12 KO seq./colony PCR FP	AACGAACGTTAAGGAACCC
STE12 KO colony PCR validation	CGGAAGTAGGAAGCGATG
STE12 KO seq. RP	GCTTTTTCCTTTGGCTTACC
STE2 KO seq./colony PCR FP	ATCTTCCCTTCCCAGAGAG
STE2 KO colony PCR validation	CCGCGTACTAACACGAAAG
STE2 KO seq. RP	AATATGACGTTGCTTCTGCTT
GPA1 KO seq./colony PCR FP	GCTGTTTTACTCGACTCAACG
GPA1 KO colony PCR validation	CTGCTTGTGTCACCATGC
GPA1 KO seq. RP	CTTGAGGAGTGGTCGTACTG
GPR1 KO seq./colony PCR FP	GACGATAGAGTCCTTGGGAG
GPR1 KO colony PCR RP	GCCAAAGTCGACGGTTAG
GPR1 KO seq. RP	AAAGGTCCTGTACGTAAAATGG
GPA2 KO seq./colony PCR FP	AGCTGCCATTCTTATGATACTG
GPA2 KO colony PCR RP	TTCCGATGCAGGATAACAG
GPA2 KO seq. RP	TCAAAAGCTCCTGGTTCCT
MF(ALPHA)1 KO seq./colony PCR FP	CTTCTTTTCTTGAGGAGAGATCC
MF(ALPHA)1 KO colony PCR RP	TGGTTTTGGGAACTCGTG
MF(ALPHA)1 KO seq. RP	TCAGAACATGAGATCAACGG
MF(ALPHA)2 KO seq. FP	TGTTGTTCTTATTTGAGCGAAC
MF(ALPHA)2 ORF KO colony PCR RP	GATATGCTGGCCTATCGG
MF(ALPHA)2 KO seq. RP	CACAAGTCGGAGGAGAGC
MFA1 KO seq./colony PCR FP	AGATGCTGTACCGTTCACG
MFA1 colony PCR RP	AGCTGATGAGCGTTACTGC
MFA1 KO seq. RP	TGCATATTTCACAATAAAGACAGTC
MFA2 KO seq./colony PCR FP	ATTCATACTCATAATGTTGATATTTCG
MFA2 colony PCR RP	ACCGTCCTTGATCTCCAG
MFA2 KO seq. RP	AATCAATCCAGTAACGATTCG
STE3 KO seq./colony PCR FP	CTAGTAGACCGGCCAAGC
STE3 colony PCR RP	CCGCTTGGACAAGAAACA
STE3 KO seq. RP	CCCAAGATAATATCATTTGTTACG
URA3 LP seq./colony PCR FP	CGTGTAAGCAGATAAGTGAATTTG
URA3 LP colony PCR RP	CGCGGTAGGTGTACAATAATATT
URA3 LP seq. RP	CCCATATCCAACTTCCAATTTA

LEU2 LP seq./colony PCR FP	CTATGTGGTATTCGATTATGCG
LEU2 LP colony PCR RP	ATGCTAGTCCACCTGATGC
LEU2 LP seq. RP	CGTGGAAGGAGAATCTTTATTG
HO LP seq./colony PCR FP	ATTCACATCATTTTCGTGGATC
HO LP colony PCR RP	GGTGAAGCATTTCGTCCAT
HO LP seq. RP	CCTTTGGACTTAAAATGGCG
Validation of CRISPR plasmid loss FP	CCGATAATTGCAGACGAAC
Validation of CRISPR plasmid loss RP	CCCGACGCTATTAGTCCC
sfGFP FP for Ste2 donor	TTAAAAATGCACCGTTAAGAACCATATCCAAGAATCAAAAATGTCCAAGGGTGAAGAG
sfGFP RP for Ste2 donor	GGTCACGAAATTACTTTTTCAAAGCCGTAAATTTTGATCACTTATAAAGCTCGTCCATTC
sfGFP FP for Gpa1 donor	ATCCAGAGGTGTATAAATTGATATATTAAGGTAGGAAATAATGTCCAAGGGTGAAGAG
sfGFP RP for Gpa1 donor	TTACGTATCTAAACACTACTTTAATTATACAGTTCCTTCACTTATAAAGCTCGTCCATTC
sfGFP FP for Ste12 donor	GGAACCGCTTTCTTTATTTGAATTGTCTTGTTCACCAAGGATGTCCAAGGGTGAAGAG
sfGFP RP for Ste12 donor	TTAATTCTTGTATCATAAATTCAAAAATTATATTATATCACTTATAAAGCTCGTCCATTC
STE2 qPCR FP	CTCAAGCAGTTCGATAGTTTCC
STE2 qPCR RP	GACATCTGTTCCCTGGTTTG
GPA1 qPCR FP	GACTTGAACCAAGAAGGCG
GPA1 qPCR RP	GCCTTAGCAATGTCTTCGTG
STE12 qPCR FP	CCACCATCATCATCGAGG
STE12 qPCR RP	CATACTTCATGTAGCTGGTAGGGT
DIG1 qPCR FP	CTGGGAAAACCAGACGATC
DIG1 qPCR RP	CTGCTGGGCGTAGATCC
DIG2 qPCR FP	GGGTCAGATATATGGGTAGGATG
DIG2 qPCR RP	GGTGGCAGTCCAGTATATGG
HTB2 qPCR FP	AAGCAAACTCACCCAGACAC
HTB2 qPCR RP	AGCGGCCAATTTAGAAGC
FIG1 qPCR FP	AAGAATTTAAGCTCGGTTCCC
FIG1 qPCR RP	TCAAAATGTTCAAGACAACGC
PRM2 qPCR FP	CATTCCACAAAAGAGCGG
PRM2 qPCR RP	TTTTCACACAATAATGATACGCTG
CIK1 qPCR FP	AAAGACCTACAAGACACCCATG
CIK1 qPCR RP	TCACTTCTGCCTCTCCAATC
sfGFP qPCR FP	ATTCTCGGCCATAAACTTGAG
sfGFP qPCR RP	CTTCGACGTTGTGCCTG

8.5.6 Plasmids used in this study

Plasmids listed here are organized in tables by figure number and then by their name within the figure. Each row corresponds to a particular cassette or multigene cassette, indicating the integration site, selection marker, and strain transformed into. Note, some strains are composed of multiple plasmids, and so will include a (1/2), (2/2), (1/3) … etc. to show which plasmids they are partnered with. All plasmids are integration cassettes integrating at either the *URA3*, *LEU2* or *HO* locus. Parts created in this study, not included in the YTK (bold) and are listed in section **8.5.1.**

Figure 22B+C

Name	Cassette Position			LP	Marker	Strain
	1					
sfGFP	*pRPL18B*	***sfGFP***	*tPGK1*	*URA3*	*URA3*	yWS677
mRuby2	*pHHF1*	*mRuby2*	*tENO2*	*LEU2*	*LEU2*	yWS677
mTagBFP2	*pTDH3*	***mTagBFP2***	*tTDH1*	*HO*	*HIS3*	yWS677

Figure 24A

Name	Cassette Position			LP	Marker	Strain
	1					
BY4741	***pFUS1***	***sfGFP***	*tTDH1*	*URA3*	*URA3*	BY4741
Quasi-WT	***pFUS1***	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT

Figure 24B

Name	Cassette Position			LP	Marker	Strain
	1					
Quasi-WT-*pFUS1*	***pFUS1***	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
Quasi-WT-*pFIG1*	***pFIG1***	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT

Figure 26B+C

Name	Cassette Position			LP	Marker	Strain
	1					
pREV1	*pREV1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pPSP2	*pPSP2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pPOP6	*pPOP6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRAD27	*pRAD27*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRNR2	*pRNR2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pSAC6	*pSAC6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRET2	*pRET2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pPAB1	*pPAB1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pALD6	*pALD6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRNR1	*pRNR1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRPL18B	*pRPL18B*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pHTB2	*pHTB2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pPGK1	*pPGK1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pHHF1	*pHHF1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pTEF1	*pTEF1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pHHF2	*pHHF2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pTEF2	*pTEF2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pTDH3	*pTDH3*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pCCW12	*pCCW12*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT

Figure 27C

Name	Cassette Position			LP	Marker	Strain
	1					
STE2						*STE2-GFP*
GPA1						*GPA1-GFP*
STE12						*STE12-GFP*
pREV1	*pREV1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pPSP2	*pPSP2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pPOP6	*pPOP6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRAD27	*pRAD27*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRNR2	*pRNR2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pSAC6	*pSAC6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRET2	*pRET2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pPAB1	*pPAB1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pALD6	*pALD6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRNR1	*pRNR1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pRPL18B	*pRPL18B*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pHTB2	*pHTB2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pPGK1	*pPGK1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pHHF1	*pHHF1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pTEF1	*pTEF1*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pHHF2	*pHHF2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pTEF2	*pTEF2*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pTDH3	*pTDH3*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
pCCW12	*pCCW12*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT

Figure 27D

Name	Cassette Position			LP	Marker	Strain
	1					
STE2						*STE2-GFP*
GPA1						*GPA1-GFP*
STE12						*STE12-GFP*
P1T1	*pSAC6*	***sfGFP***	*tENO1*	*URA3*	*URA3*	Quasi-WT
P1T2	*pSAC6*	***sfGFP***	*tSSA1*	*URA3*	*URA3*	Quasi-WT
P1T3	*pSAC6*	***sfGFP***	*tADH1*	*URA3*	*URA3*	Quasi-WT
P1T4	*pSAC6*	***sfGFP***	*tPGK1*	*URA3*	*URA3*	Quasi-WT
P1T5	*pSAC6*	***sfGFP***	*tENO2*	*URA3*	*URA3*	Quasi-WT
P1T6	*pSAC6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
P2T1	*pPOP6*	***sfGFP***	*tENO1*	*URA3*	*URA3*	Quasi-WT
P2T2	*pPOP6*	***sfGFP***	*tSSA1*	*URA3*	*URA3*	Quasi-WT
P2T3	*pPOP6*	***sfGFP***	*tADH1*	*URA3*	*URA3*	Quasi-WT
P2T4	*pPOP6*	***sfGFP***	*tPGK1*	*URA3*	*URA3*	Quasi-WT
P2T5	*pPOP6*	***sfGFP***	*tENO2*	*URA3*	*URA3*	Quasi-WT
P2T6	*pPOP6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
P3T1	*pPOP6*	***sfGFP***	*tENO1*	*URA3*	*URA3*	Quasi-WT
P3T2	*pPOP6*	***sfGFP***	*tSSA1*	*URA3*	*URA3*	Quasi-WT
P3T3	*pPOP6*	***sfGFP***	*tADH1*	*URA3*	*URA3*	Quasi-WT
P3T4	*pPOP6*	***sfGFP***	*tPGK1*	*URA3*	*URA3*	Quasi-WT
P3T5	*pPOP6*	***sfGFP***	*tENO2*	*URA3*	*URA3*	Quasi-WT
P3T6	*pPOP6*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT
P4T1	*pRAD27*	***sfGFP***	*tENO1*	*URA3*	*URA3*	Quasi-WT
P4T2	*pRAD27*	***sfGFP***	*tSSA1*	*URA3*	*URA3*	Quasi-WT
P4T3	*pRAD27*	***sfGFP***	*tADH1*	*URA3*	*URA3*	Quasi-WT
P4T4	*pRAD27*	***sfGFP***	*tPGK1*	*URA3*	*URA3*	Quasi-WT
P4T5	*pRAD27*	***sfGFP***	*tENO2*	*URA3*	*URA3*	Quasi-WT
P4T6	*pRAD27*	***sfGFP***	*tTDH1*	*URA3*	*URA3*	Quasi-WT

Figure 28D

Name	Cassette Position												LP	Marker	Strain
	1			2			3			4					
Quasi-WT	***pFUS1***	***sfGFP***	*tTDH1*										*URA3*	*URA3*	Quasi-WT
Design 1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 29B

Name	Cassette Position										LP	Marker	Strain
	1		2			3			4				
pSTE2-sfGFP-tSTE2	***Spacer 1***		*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	*STE2-GFP*
pGPA1-sfGFP-tGPA1	***Spacer 1***		*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	*GPA1-GFP*
pSTE12-sfGFP-tSTE12	***Spacer 1***		*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	*STE12-GFP*

Figure 29E

Name	Cassette Position												LP	Marker	Strain
	1			2			3			4					
Quasi-WT	***pFUS1***	***sfGFP***	*tTDH1*										*URA3*	*URA3*	Quasi-WT
Design 1 w/o feedback	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
Design 1 w/ feedback (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
Design 1 w/ feedback (2/2)	***pFUS1***	***STE2***	*tTDH1*	***pFUS1***	***GPA1***	*tTDH1*	***pFUS1***	***STE12***	*tTDH1*				*LEU2*	*LEU2*	yWS677

Figure 31B+D

Name	Cassette Position												LP	Marker	Strain
	1			2			3			4					
pPSP2	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPSP2*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pREV1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pREV1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pPOP6	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR2	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pRNR2*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pRET2	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pRET2*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pALD6	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pALD6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pHTB2	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pHTB2*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pPAB1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPAB1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pRNR1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pHHF1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pHHF1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pHHF2	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pHHF2*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pPGK1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pRPL18B	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pRPL18B*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pTEF2	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pTEF2*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pTEF1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pTEF1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pTDH3	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pTDH3*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pCCW12	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pCCW12*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 32B

Name	Cassette Position												LP	Marker	Strain
	1			2			3			4					
pCCW12	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pHHF2	***pFUS1***	***sfGFP***	*tTDH1*	*pHHF2*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pPOP6	***pFUS1***	***sfGFP***	*tTDH1*	*pPOP6*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pPSP2	***pFUS1***	***sfGFP***	*tTDH1*	*pPSP2*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pREV1	***pFUS1***	***sfGFP***	*tTDH1*	*pREV1*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 33A

Name	Cassette Position												LP	Marker	Strain
	1			2			3			4					
pREV1	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pREV1*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pRPL18B	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRPL18B*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
pTDH3	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pTDH3*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 34

Name	Cassette Position												LP	Marker	Strain
	1			2			3			4					
1 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pTDH3*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
2 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pTEF1*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
3 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pTEF2*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
4 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pHHF2*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
5 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pHTB2*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
6 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pHHF1*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
7 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRNR1*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
8 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRPL18B*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
9 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pPAB1*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
10 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pALD6*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
11 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRET2*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
12 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pSAC6*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
13 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pPOP6*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
14 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRNR2*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
15 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
16 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pPSP2*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
17 (1/2)	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pREV1*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
Dig1- Dig2- (2/2)	***Spacer 1***			***Spacer 2***									*LEU2*	*LEU2*	yWS677
Dig1+ Dig2- (2/2)	*pCCW12*	***DIG1***	*tADH1*	***Spacer 2***									*LEU2*	*LEU2*	yWS677
Dig1- Dig2+ (2/2)	***Spacer 1***			*pTDH3*	***DIG2***	*tTDH1*							*LEU2*	*LEU2*	yWS677
Dig1+ Dig2+ (2/2)	*pCCW12*	***DIG1***	*tADH1*	*pTDH3*	***DIG2***	*tTDH1*							*LEU2*	*LEU2*	yWS677

Figure 35

Name	Cassette Position												LP	Marker	Strain
	1			2			3			4					
Quasi-WT	***pFUS1***	***sfGFP***	*tTDH1*										*URA3*	*URA3*	Quasi-WT
Design 1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
Design 2	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
Design 3	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 36D+E

Name	Cassette Position													LP	Marker	Strain
	1				2			3			4					
Quasi-WT	***pFUS1***	***sfGFP***		*tTDH1*										*URA3*	*URA3*	Quasi-WT
Design 3	***pFUS1***	***sfGFP***		*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
STF1	***Gal4BS(5x)-pLEU2m***		***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STF1***	*tENO1*	*URA3*	*URA3*	yWS677
STF2	***LexO(4x)-pLEU2m***		***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STF2***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 37C+D

Name	Cassette Position													LP	Marker	Strain
	1				2			3			4					
Quasi-WT	***pFUS1***	***sfGFP***		*tTDH1*										*URA3*	*URA3*	Quasi-WT
Design 3	***pFUS1***	***sfGFP***		*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
LexA-PRD	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 38B

Name	Cassette Position													LP	Marker	Strain
	1				2			3			4					
LexO (1x)	***LexO(1x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (2x)	***LexO(2x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (3x)	***LexO(3x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (4x)	***LexO(4x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (6x)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (8x)	***LexO(8x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
pCCW12	***pCCW12***		***sfGFP***	*tTDH1*											*URA3*	*URA3*	*y*WS677

Figure 38C

Name	Cassette Position														LP	Marker	Strain
	1				2			3			4						
pRNR2m	***LexO(6x)***	***pRNR2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pCUP1m	***LexO(6x)***	***pCUP1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pCYC1m	***LexO(6x)***	***pCYC1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPHO5m	***LexO(6x)***	***pPHO5m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pGAL1m	***LexO(6x)***	***pGAL1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pCCW12m	***LexO(6x)***	***pCCW12m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pLEU2m	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPGK1m	***LexO(6x)***	***pPGK1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pALD6m	***LexO(6x)***	***pALD6m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pTDH3m	***LexO(6x)***	***pTDH3m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pTEF2m	***LexO(6x)***	***pTEF2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR1m	***LexO(6x)***	***pRNR1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 39B+C

Name	Cassette Position														LP	Marker	Strain
	1				2			3			4						
TetR-PRD	**TetO (6x)**	***pLEU2m***	**sfGFP**	*tTDH1*	*pCCW12*	**STE2**	*tSSA1*	*pPGK1*	**GPA1**	*tENO2*	*pRAD27*	**TetR**	**PRD**	*tENO1*	*URA3*	*URA3*	yWS677

Figure 39E+F

Name	Cassette Position														LP	Marker	Strain
	1				2			3			4						
Z3E-PRD	***pZ3***	**sfGFP**		*tTDH1*	*pCCW12*	**STE2**	*tSSA1*	*pPGK1*	**GPA1**	*tENO2*	*pRAD27*	**Z3EV**	**PRD**	*tENO1*	*URA3*	*URA3*	yWS677

Figure 40D

Name	Cassette Position												LP	Marker	Strain
	1			2			3			4					
Synthetic promoter (1/2)	*pALD6*	*sfGFP*	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***dCas9***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Synthetic promoter (2/2)	***gCRISPR-UAS***													*LEU2*	*LEU2*	yWS677

Figure 40E

Name	Cassette Position										LP	Marker	Strain
	1	2			3			4					
All (1/2)	***Spacer 1***	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***dCas9***	***PRD***	*tENO1*	*URA3*	*URA3*	*ALD6 GFP*
g8 (2/2)	***ALD6* g1**											*LEU2*	*LEU2*	*ALD6 GFP*
g7 (2/2)	***ALD6* g2**											*LEU2*	*LEU2*	*ALD6 GFP*
g6 (2/2)	***ALD6* g3**											*LEU2*	*LEU2*	*ALD6 GFP*
g5 (2/2)	***ALD6* g4**											*LEU2*	*LEU2*	*ALD6 GFP*
g4 (2/2)	***ALD6* g5**											*LEU2*	*LEU2*	*ALD6 GFP*
g3 (2/2)	***ALD6* g6**											*LEU2*	*LEU2*	*ALD6 GFP*
g2 (2/2)	***ALD6* g7**											*LEU2*	*LEU2*	*ALD6 GFP*
g1 (2/2)	***ALD6* g8**											*LEU2*	*LEU2*	*ALD6 GFP*
Control (2/2)	***Spacer 1***											*LEU2*	*LEU2*	*ALD6 GFP*
g8+5 (2/2)	***ALD6* g8**	***ALD6* g5**										*LEU2*	*LEU2*	*ALD6 GFP*

Figure 40F

Name	Cassette Position										LP	Marker	Strain
	1	2			3			4					
g8+5 (1/2)	***Spacer 1***	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***dCas9***	***PRD***	*tENO1*	*URA3*	*URA3*	*ALD6 GFP*
g8+5 (2/2)	***ALD6* g8**	***ALD6* g5**										*LEU2*	*LEU2*	*ALD6 GFP*

Figure 41

Name	Cassette Position												LP	Marker	Strain
	1				2			3			4				
Quasi-WT	***pFUS1***	***sfGFP***	*tTDH1*										*URA3*	*URA3*	Quasi-WT
Design 1	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
Design 2	***pFUS1***	***sfGFP***	*tTDH1*	*pSAC6*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
Design 3	***pFUS1***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***	*tENO1*	*URA3*	*URA3*	yWS677
Design 4	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 42D+E

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
pPSP2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPSP2*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pREV1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pREV1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pSAC6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pSAC6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPOP6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pRNR2*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRET2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pRET2*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pALD6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pALD6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPAB1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPAB1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pRNR1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRPL18B (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pRPL18B*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Receptor (2/2)	*pTDH3*	***A2BR***		*tSSA1*									*LEU2*	*LEU2*	yWS677
No receptor (2/2)	***LEU2 Spacer***												*LEU2*	*LEU2*	yWS677

Figure 43B

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
pSAC6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pSAC6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPOP6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pRNR2*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRET2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pRET2*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pALD6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pALD6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPAB1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPAB1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pRNR1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRPL18B (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pRPL18B*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
A2BR (2/2)	*pTDH3*	***A2BR***		*tSSA1*									*LEU2*	*LEU2*	yWS677
A2BR-ΔC-tail (2/2)	*pTDH3*	***A2BR-ΔC-tail***		*tSSA1*									*LEU2*	*LEU2*	yWS677
No receptor	***LEU2 Spacer***												*LEU2*	*LEU2*	yWS677

Figure 43C

Name	Cassette Position												LP	Marker	Strain
	1				2			3			4				
pSAC6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pSAC6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPOP6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPOP6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pRNR2*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRET2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pRET2*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pALD6 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pALD6*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPAB1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPAB1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pRNR1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRPL18B (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pRPL18B*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Ste2 (2/2)	*pTDH3*		***STE2***	*tSSA1*											*LEU2*	*LEU2*	yWS677
No receptor (2/2)	***Spacer 1***														*LEU2*	*LEU2*	yWS677

Figure 44A

Name	Cassette Position												LP	Marker	Strain
	1				2			3			4				
Quasi-WT	***pFUS1***		***sfGFP***	*tTDH1*									*URA3*	*URA3*	Quasi-WT
Design 4	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 44B

Name	Cassette Position														LP	Marker	Strain
	1				2			3			4						
Ste12 (1/2)	***Spacer 1***				*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***STE12***		*tENO1*	*URA3*	*URA3*	yWS677
LexA (1/2)	***Spacer 1***				*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
TetR (1/2)	***Spacer 1***				*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***TetR***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Z3E (1/2)	***Spacer 1***				*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***Z3E***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pFUS1 (2/2)	***pFUS1***		***sfGFP***	*tTDH1*											*LEU2*	*LEU2*	yWS677
LexO(6x)-pLEU2m (2/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*											*LEU2*	*LEU2*	yWS677
TetO(6x)-pLEU2m (2/2)	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*											*LEU2*	*LEU2*	yWS677
pZ3 (2/2)	***pZ3***		***sfGFP***	*tTDH1*											*LEU2*	*LEU2*	yWS677

Figure 45B

Name	Cassette Position												LP	Marker	Strain
	1				2			3			4				
Mam2 - yWS677	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***Mam2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677

Figure 46D+E

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
GPA1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gαs (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gαs***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gα12 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gα12***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gαi1/2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gαi1/2***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gαi3 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gαi3***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gαz (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gαz***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gα15/16 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gα15/16***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gαq/11 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gαq/11***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gα14 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gα14***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
GPA1-Gαo (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gαo***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
tGPA1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***tGPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
A2BR (2/2)	*pTDH3*	***A2BR***		*tSSA1*									*LEU2*	*LEU2*	yWS677
MTNR1A (2/2)	*pTDH3*	***MTNR1A***		*tSSA1*									*LEU2*	*LEU2*	yWS677

Figure 47A

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
A2BR sensor (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
A2BR sensor (2/2)	*pTDH3*	***A2BR***		*tSSA1*									*LEU2*	*LEU2*	yWS677

Figure 47B

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
MTNR1A sensor (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
MTNR1A sensor (2/2)	*pTDH3*	***MTNR1A***		*tSSA1*									*LEU2*	*LEU2*	yWS677

Figure 48B-D

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
All conditions (1/3)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
All conditions (2/3)	*pHHF2*	***A2BR***		*tTDH1*									*LEU2*	*LEU2*	yWS677
No feedback (3/3)	***HO Spacer***												*HO*	*HIS3*	yWS677
Gpa1 feedback (3/3)	***LexO(6x)***	***pLEU2m***	***GPA1***	*tTDH1*									*HO*	*HIS3*	yWS677
Sst2 feedback (3/3)	***LexO(6x)***	***pLEU2m***	***SST2***	*tTDH1*									*HO*	*HIS3*	yWS677
Msg5 feedback (3/3)	***LexO(6x)***	***pLEU2m***	***MSG5***	*tTDH1*									*HO*	*HIS3*	yWS677
Dig1 feedback (3/3)	***LexO(6x)***	***pLEU2m***	***DIG1***	*tTDH1*									*HO*	*HIS3*	yWS677

Figure 48F-H

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
All conditions (1/3)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
All conditions (2/3)	*pHHF2*	***MTNR1A***	*tTDH1*										*LEU2*	*LEU2*	yWS677
No feedback (3/3)	***HO Spacer***												*HO*	*HIS3*	yWS677
MTNR1A feedback (3/3)	***LexO(6x)***	***pLEU2m***	***MTNR1A***	*tTDH1*									*HO*	*HIS3*	yWS677
Ste4-2A-Ste18 feedback (3/3)	***LexO(6x)***	***pLEU2m***	***STE4-2A-STE18***	*tTDH1*									*HO*	*HIS3*	yWS677
Ste50 feedback (3/3)	***LexO(6x)***	***pLEU2m***	***STE50***	*tTDH1*									*HO*	*HIS3*	yWS677

Figure 49A+B

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
High sensitivity (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
High sensitivity (2/2)	*pCCW12*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677
Mid sensitivity (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid sensitivity (2/2)	*pHHF2*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677
Low sensitivity (1/2)	***LexO(6x)***	***pLEU2m***	***SST2***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Low sensitivity (2/2)	*pRPL18B*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677

Figure 50B

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
High sensitivity (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
High sensitivity (2/2)	*pCCW12*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677
Mid sensitivity (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid sensitivity (2/2)	*pHHF2*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677
Low sensitivity (1/2)	***LexO(6x)***	***pLEU2m***	***SST2***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Low sensitivity (2/2)	*pRPL18B*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677

Figure 50C

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
Mid 2.01/ High 2.01 (1/2)	***LexO(1x)***	***pRNR2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.02/ High 2.02 (1/2)	***LexO(1x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.03/ High 2.03 (1/2)	***LexO(1x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.04/ High 2.04 (1/2)	***LexO(2x)***	***pRNR2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.05/ High 2.05 (1/2)	***LexO(2x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.06/ High 2.06 (1/2)	***LexO(2x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.07/ High 2.07 (1/2)	***LexO(3x)***	***pRNR2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.08/ High 2.08 (1/2)	***LexO(3x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.09/ High 2.09 (1/2)	***LexO(3x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.10/ High 2.10 (1/2)	***LexO(4x)***	***pRNR2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.11/ High 2.11 (1/2)	***LexO(4x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 2.12/ High 2.12 (1/2)	***LexO(4x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid 1.0/ High 1.0 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid (2/2)	*pCCW12*	***A2BR***		*tTDH1*									*LEU2*	*LEU2*	yWS677
High (2/2)	*pHHF2*	***A2BR***		*tTDH1*									*LEU2*	*LEU2*	yWS677
Low 1.0 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Low 1.0 (2/2)	*pRPL18B*	***A2BR***		*tTDH1*									*LEU2*	*LEU2*	yWS677

Figure 50D

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
High sensitivity (1/2)	***LexO(2x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
High sensitivity (2/2)	*pCCW12*	***A2BR***		*tTDH1*									*LEU2*	*LEU2*	yWS677
Mid sensitivity (1/2)	***LexO(3x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Mid sensitivity (2/2)	*pHHF2*	***A2BR***		*tTDH1*									*LEU2*	*LEU2*	yWS677
Low sensitivity (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Low sensitivity (2/2)	*pRPL18B*	***A2BR***		*tTDH1*									*LEU2*	*LEU2*	yWS677

Figure 51A

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
High sensitivity (1/2)	***LexO(2x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Low sensitivity (2/2)	*pHHF2*	***A2BR***		*tTDH1*									*LEU2*	*LEU2*	yWS677

Figure 51B

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
Cell 1 (1/2)	***LexO(2x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	*pCCW12*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677
Cell 2 (1/2)	***LexO(3x)***	***pPHO5m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 2 (2/2)	*pHHF2*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677
Cell 3 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 3 (2/2)	*pRPL18B*	***A2BR***	*tTDH1*										*LEU2*	*LEU2*	yWS677

Figure 53B+C

Name	Cassette Position												LP	Marker	Strain
	1				2			3			4				
Cell 1 (1/2)	***LexO(6x)***	***pLEU2m***	***MF(ALPHA)1***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	*pHHF2*	***MTNR1A***	*tTDH1*												*LEU2*	*LEU2*	yWS677
All conditions Cell 2 (1/2)	***LexO(6x)***	***pCUP1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pCCW12 Cell 2 (2/2)	*pCCW12*	***BAR1***	*tTDH1*												*HO*	*HIS3*	yWS677
pTDH3 Cell 2 (2/2)	*pTDH3*	***BAR1***	*tTDH1*												*HO*	*HIS3*	yWS677
pPGK1 Cell 2 (2/2)	*pPGK1*	***BAR1***	*tTDH1*												*HO*	*HIS3*	yWS677
pRPL18B Cell 2 (2/2)	*pRPL18B*	***BAR1***	*tTDH1*												*HO*	*HIS3*	yWS677
pRNR2 Cell 2 (2/2)	*pRNR2*	***BAR1***	*tTDH1*												*HO*	*HIS3*	yWS677
No Bar1 Cell 2 (2/2)	***HO Spacer***														*HO*	*HIS3*	yWS677

Figure 54A

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
Cell 1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	*pHHF2*	***MTNR1A***	*tTDH1*										*LEU2*	*LEU2*	yWS677

Figure 54B

Name	Cassette Position												LP	Marker	Strain
	1				2			3			4				
Cell 1 (1/2)	***LexO(6x)***	***pLEU2m***	***MF(ALPHA)1***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	*pHHF2*	***MTNR1A***	*tTDH1*												*LEU2*	*LEU2*	yWS677
Cell 2 (1/2)	***LexO(6x)***	***pCUP1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 2 (2/2)	*pRPL18B*	***BAR1***	*tTDH1*												*HO*	*HIS3*	yWS677

Figure 54C

Name	Cassette Position													LP	Marker	Strain
	1				2			3			4					
Cell 1 (1/2)	***LexO(6x)***	***pLEU2m***	***MF(ALPHA)1***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	*pALD6*	***MTNR1A***		*tTDH1*											*LEU2*	*LEU2*	yWS677
Cell 2 (1/2)	***LexO(6x)***	***pCUP1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 2 (2/2)	*pRPL18B*	***BAR1***		*tTDH1*											*HO*	*HIS3*	yWS677

Figure 55A

Name	Cassette Position													LP	Marker	Strain
	1				2			3			4					
Cell 1 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	*pHHF2*	***MTNR1A***		*tTDH1*											*LEU2*	*LEU2*	yWS677

Figure 55B

Name	Cassette Position													LP	Marker	Strain
	1				2			3			4					
Cell 1 (1/2)	***LexO(6x)***	***pLEU2m***	***MF(ALPHA)1***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	*pALD6*	***MTNR1A***		*tTDH1*											*LEU2*	*LEU2*	yWS677
Cell 2 (1/2)	***LexO(6x)***	***pCUP1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 2 (2/2)	*pRPL18B*	***BAR1***		*tTDH1*											*HO*	*HIS3*	yWS677

Figure 56B

Name	Cassette Position													LP	Marker	Strain
	1				2			3			4					
MTNR1A sensor (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
MTNR1A sensor (2/2)	*pHHF2*	***MTNR1A***		*tTDH1*											*LEU2*	*LEU2*	yWS677

Figure 56C+E

Name	Cassette Position													LP	Marker	Strain
	1				2			3			4					
Cell 1 (1/2)	***LexO(2x)***	***pRNR2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	*pCCW12*	***MTNR1A***		*tTDH1*											*LEU2*	*LEU2*	yWS677
Cell 2 (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 2 (2/2)	*pALD6*	***MTNR1A***		*tTDH1*											*LEU2*	*LEU2*	yWS677

Figure 57A

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
HTR4 sensor (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
HTR4 sensor (2/2)	*pCCW12*	***HTR4***	*tTDH1*										*LEU2*	*LEU2*	yWS677

Figure 57B

Name	Cassette Position												LP	Marker	Strain
	1				2	3			4						
All conditions (1/2)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***	*pPGK1*	***GPA1-Gα(olf)***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Control (2/2)	***LEU2 Spacer***												*LEU2*	*LEU2*	yWS677
Olfr609 (2/2)	*pTDH3*	***Olfr609***	*tSSA1*										*LEU2*	*LEU2*	yWS677
Olfr556 (2/2)	*pTDH3*	***Olfr556***	*tSSA1*										*LEU2*	*LEU2*	yWS677
Olfr168 (2/2)	*pTDH3*	***Olfr168***	*tSSA1*										*LEU2*	*LEU2*	yWS677
OR2J2 (2/2)	*pTDH3*	***OR2J2***	*tSSA1*										*LEU2*	*LEU2*	yWS677
OR1A1 (2/2)	*pTDH3*	***OR1A1***	*tSSA1*										*LEU2*	*LEU2*	yWS677
OR3A1 (2/2)	*pTDH3*	***OR3A1***	*tSSA1*										*LEU2*	*LEU2*	yWS677
OR2W1 (2/2)	*pTDH3*	***OR2W1***	*tSSA1*										*LEU2*	*LEU2*	yWS677
Olfr73 (2/2)	*pTDH3*	***Olfr73***	*tSSA1*										*LEU2*	*LEU2*	yWS677
Olfr154 (2/2)	*pTDH3*	***Olfr154***	*tSSA1*										*LEU2*	*LEU2*	yWS677
RI7 (2/2)	*pTDH3*	***RI7***	*tSSA1*										*LEU2*	*LEU2*	yWS677

Figure 58D

Name	Cassette Position																LP	Marker	Strain
	1				2				3				4						
Cell 1 (1/2)	***LexO(2x)***	***pRNR2m***	***MF(ALPHA)1***	*tTDH1*	*pRPL18B*	***A2BR***	*tSSA1*		*pPGK1*	***GPA1***	*tENO2*		*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 1 (2/2)	***pZ3***	***LexA***	***MxiI***	*tTDH1*	*pREV1*	***Z3E***	***VP16***	*tENO2*									*LEU2*	*LEU2*	yWS677
Cell 2 (1/2)	***LexO(2x)***	***pRNR2m***	***Z3E***	***MxiI***	*tTDH1*	*pRPL18B*	***A2BR***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*		*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 2 (2/2)	***pZ3***	***MF(ALPHA)1***	*tTDH1*		*pREV1*	***Z3E***	***VP16***	*tTDH1*									*LEU2*	*LEU2*	yWS677
Cell 3	***LexO(2x)***	***pRNR2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***A2BR***	*tSSA1*		*pPGK1*	***GPA1***	*tENO2*		*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 4 (1/2)	*pREV1*	***Z3E***	***VP16***	*tTDH1*	*pRPL18B*	***A2BR***	*tSSA1*		*pPGK1*	***GPA1***	*tENO2*		*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 4 (2/2)	***pZ3***	***TetR***	***SpyCatcher***	*tPGK1*	***LexO(6x)***	***pCYC1m***	***SpyTag***	***Gal4$_{AD}$*** *tENO2*	***TetO(6x)***	***pLEU2m***	***BAR1***	*tENO2*					*LEU2*	*LEU2*	yWS677
Cell 5 (1/2)	*pREV1*	***Z3E***	***VP16***	*tTDH1*	*pRPL18B*	***A2BR***	*tSSA1*		*pPGK1*	***GPA1***	*tENO2*		*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Cell 5 (2/2)	***pZ3***	***TetR***	***SpyCatcher***	*tPGK1*	***LexO(6x)***	***pCYC1m***	***SpyTag***	***Gal4$_{AD}$*** *tENO2*	***TetO(6x)***	***pLEU2m***	***mScarlet***	*tENO2*					*LEU2*	*LEU2*	yWS677

Supplementary Figure S2A

Name	Cassette Position														LP	Marker	Strain
	1				2			3			4						
LexO (1x)	***LexO(1x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (2x)	***LexO(2x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (3x)	***LexO(3x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (4x)	***LexO(4x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (6x)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677
LexO (8x)	***LexO(8x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	*y*WS677

Supplementary Figure S2B

Name	Cassette Position														LP	Marker	Strain
	1				2			3			4						
pRNR2m	***LexO(6x)***	***pRNR2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pCUP1m	***LexO(6x)***	***pCUP1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pCYC1m	***LexO(6x)***	***pCYC1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPHO5m	***LexO(6x)***	***pPHO5m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pGAL1m	***LexO(6x)***	***pGAL1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pCCW12m	***LexO(6x)***	***pCCW12m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pLEU2m	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pPGK1m	***LexO(6x)***	***pPGK1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pALD6m	***LexO(6x)***	***pALD6m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pTDH3m	***LexO(6x)***	***pTDH3m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pTEF2m	***LexO(6x)***	***pTEF2m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pRNR1m	***LexO(6x)***	***pRNR1m***	***sfGFP***	*tTDH1*	*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677

Supplementary Figure S3A+B

Name	Cassette Position											LP	Marker	Strain
	1			2			3			4				
All (1/2)	***Spacer 1***			*pCCW12*	***STE2***	*tSSA1*	*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***dCas9***	***PRD***	*tENO1*	*URA3*	*URA3*	*ALD6 GFP*
g8-7-5 (2/2)	***ALD6* g8**			***ALD6* g7**			***ALD6* g5**							*LEU2*	*LEU2*	*ALD6 GFP*
g8-7-4 (2/2)	***ALD6* g8**			***ALD6* g7**			***ALD6* g4**							*LEU2*	*LEU2*	*ALD6 GFP*
g8-7-3 (2/2)	***ALD6* g8**			***ALD6* g7**			***ALD6* g3**							*LEU2*	*LEU2*	*ALD6 GFP*
g8-6-4 (2/2)	***ALD6* g8**			***ALD6* g6**			***ALD6* g4**							*LEU2*	*LEU2*	*ALD6 GFP*
g8-6-3 (2/2)	***ALD6* g8**			***ALD6* g6**			***ALD6* g3**							*LEU2*	*LEU2*	*ALD6 GFP*
g8-7 (2/2)	***ALD6* g8**			***ALD6* g7**										*LEU2*	*LEU2*	*ALD6 GFP*
g8-6 (2/2)	***ALD6* g8**			***ALD6* g6**										*LEU2*	*LEU2*	*ALD6 GFP*
g8-5 (2/2)	***ALD6* g8**			***ALD6* g5**										*LEU2*	*LEU2*	*ALD6 GFP*
g8-4 (2/2)	***ALD6* g8**			***ALD6* g4**										*LEU2*	*LEU2*	*ALD6 GFP*
g8-3 (2/2)	***ALD6* g8**			***ALD6* g3**										*LEU2*	*LEU2*	*ALD6 GFP*
g7-5 (2/2)	***ALD6* g7**			***ALD6* g5**										*LEU2*	*LEU2*	*ALD6 GFP*
g7-4 (2/2)	***ALD6* g7**			***ALD6* g4**										*LEU2*	*LEU2*	*ALD6 GFP*
g7-3 (2/2)	***ALD6* g7**			***ALD6* g3**										*LEU2*	*LEU2*	*ALD6 GFP*
g6-5 (2/2)	***ALD6* g6**			***ALD6* g5**										*LEU2*	*LEU2*	*ALD6 GFP*
g6-4 (2/2)	***ALD6* g6**			***ALD6* g4**										*LEU2*	*LEU2*	*ALD6 GFP*
g6-3 (2/2)	***ALD6* g6**			***ALD6* g3**										*LEU2*	*LEU2*	*ALD6 GFP*
Control (2/2)	***LEU2 Spacer***													*LEU2*	*LEU2*	*ALD6 GFP*

Supplementary Figure S4C

Name	Cassette Position														LP	Marker	Strain
	1				2				3			4					
pZ3-STE4-2A-STE18 (1/3)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pHHF2*	***A2BR***	*tTDH1*		*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
pZ3-STE4-2A-STE18 (2/3)	*pHHF2*	***A2BR***		*tTDH1*												*LEU2*	*LEU2*	yWS677
pZ3-STE4-2A-STE18 (3/3)	*pRAD27*	***Z3E***	***VP16***	*tTDH1*	***pZ3***	**STE4-2A-STE18**	*tENO2*									*HO*	*HIS3*	yWS677
Control (1/3)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pHHF2*	***A2BR***	*tTDH1*		*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*	*URA3*	*URA3*	yWS677
Control (2/3)	*pHHF2*	***A2BR***		*tTDH1*												*LEU2*	*LEU2*	yWS677
Control (3/3)	*pRAD27*	***Z3E***	***VP16***	*tTDH1*	***Spacer 2***											*HO*	*HIS3*	yWS677

Supplementary Figure S5B+C

Name	Cassette Position												LP	Marker	Strain
	1				2				3						
TetA-Gal4	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRAD27*	***TetA***	*Gal4$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
TetA-B42	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRAD27*	***TetA***	*B42$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
TetA-VP16	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRAD27*	***TetA***	*VP16$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
TetA-VP64	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRAD27*	***TetA***	*VP64$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Z3E-Gal4	***pZ3***	***sfGFP***		*tTDH1*	*pRAD27*	***Z3E***	*Gal4$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Z3E-B42	***pZ3***	***sfGFP***		*tTDH1*	*pRAD27*	***Z3E***	*B42$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Z3E-VP16	***pZ3***	***sfGFP***		*tTDH1*	*pRAD27*	***Z3E***	*VP16$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Z3E-VP64	***pZ3***	***sfGFP***		*tTDH1*	*pRAD27*	***Z3E***	*VP64$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677

Supplementary Figure S6B-E

Name	Cassette Position																LP	Marker	Strain
	1				2			3			4								
pCCW12 (1/3)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*			*URA3*	*URA3*	yWS677
pCCW12 (2/3)	*pCCW12*	***A2BR***		*tTDH1*													*LEU2*	*LEU2*	yWS677
pCCW12 (3/3)	***HO Spacer***																*HO*	*HIS3*	yWS677
pHHF2 (1/3)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*			*URA3*	*URA3*	yWS677
pHHF2 (2/3)	*pHHF2*	***A2BR***		*tTDH1*													*LEU2*	*LEU2*	yWS677
pHHF2 (3/3)	*pRPL18B*	*mRuby2*		*tTDH1*													*HO*	*HIS3*	yWS677
pRPL18B (1/3)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*			*URA3*	*URA3*	yWS677
pRPL18B (2/3)	*pRPL18B*	***A2BR***		*tTDH1*													*LEU2*	*LEU2*	yWS677
pRPL18B (3/3)	*pTDH3*	***mTagBFP2***		*tTDH1*													*HO*	*HIS3*	yWS677
pALD6 (1/3)	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	***Spacer 2***			*pPGK1*	***GPA1***	*tENO2*	*pRAD27*	***LexA***	***PRD***	*tENO1*			*URA3*	*URA3*	yWS677
pALD6 (2/3)	*pALD6*	***A2BR***		*tTDH1*													*LEU2*	*LEU2*	yWS677
pALD6 (3/3)	*pTDH3*	***mTagBFP2***		*tTDH1*	*pRPL18B*	*mRuby2*	*tTDH1*										*HO*	*HIS3*	yWS677

Supplementary Figure S8B

Name	Cassette Position																LP	Marker	Strain
	1				2				3				4						
Gal4-Gal4	***Gal4BS(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***Gal4$_{DBD}$***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*Gal4$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Gal4-B42	***Gal4BS(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***Gal4$_{DBD}$***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*B42$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Gal4-VP16	***Gal4BS(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***Gal4$_{DBD}$***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*VP16$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Gal4-VP64	***Gal4BS(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***Gal4$_{DBD}$***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*VP64$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
LexA-Gal4	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***LexA***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*Gal4$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
LexA-B42	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***LexA***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*B42$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
LexA-VP16	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***LexA***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*VP16$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
LexA-VP64	***LexO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***LexA***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*VP64$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Z3E-Gal4	***pZ3***	***sfGFP***	*tTDH1*		*pRPL18B*	***Z3E***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*Gal4$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Z3E-B42	***pZ3***	***sfGFP***	*tTDH1*		*pRPL18B*	***Z3E***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*B42$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Z3E-VP16	***pZ3***	***sfGFP***	*tTDH1*		*pRPL18B*	***Z3E***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*VP16$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
Z3E-VP64	***pZ3***	***sfGFP***	*tTDH1*		*pRPL18B*	***Z3E***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*VP64$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
TetR-Gal4	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***TetR***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*Gal4$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
TetR-B42	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***TetR***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*B42$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
TetR-VP16	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***TetR***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*VP16$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677
TetR-VP64	***TetO(6x)***	***pLEU2m***	***sfGFP***	*tTDH1*	*pRPL18B*	***TetR***	***SpyCatcher***	*tPGK1*	*pHTB2*	***SpyTag-NLS***	*VP64$_{AD}$*	*tENO2*					*URA3*	*URA3*	yWS677

Supplementary Figure S8C

Name	Cassette Position																LP	Marker	Strain
	1				2					3				4					
Gal4-Gal4	*pREV1*	***Z3E***	***VP16***	*tTDH1*	*pRPL18B*	***A2BR***		*tSSA1*		*pPGK1*	***GPA1***	*tENO2*		*pRAD27*	***LexA***	***PRD***	*tENO1*			
TetR-VP64	***pZ3***	***TetR***	***SpyCatcher***	*tPGK1*	***LexO(6x)***	***pCYC1m***	***SpyTag***	***Gal4$_{AD}$***	*tENO2*	*TetO(6x)*	***pLEU2m***	*sfGFP*	*tTDH1*					*URA3*	*URA3*	yWS677

9 Bibliography

1. World Economic Forum. The global risks report 2018, 13th edition. (2018).
2. Hoekstra, A. Y. & Wiedmann, T. O. Humanity's unsustainable environmental footprint. *Science (80-.).* **344,** 1114–1117 (2014).
3. Holdren, J. P. Science and Technology for Sustainable Well-Being. *Science (80-.).* **319,** 424–434 (2008).
4. Andrianantoandro, E., Basu, S., Karig, D. K. & Weiss, R. Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.* **2,** 2006.0028 (2006).
5. Endy, D. Foundations for engineering biology. *Nature* **438,** 449–453 (2005).
6. Way, J. C., Collins, J. J., Keasling, J. D. & Silver, P. A. Integrating Biological Redesign: Where Synthetic Biology Came From and Where It Needs to Go. *Cell* **157,** 151–161 (2014).
7. Lee, J. W. *et al.* Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* **8,** 536–546 (2012).
8. Keasling, J. D. Manufacturing Molecules Through Metabolic Engineering. *Science (80-.).* **330,** 1355–1358 (2010).
9. Lopes, M. S. G. Engineering biological systems toward a sustainable bioeconomy. *J. Ind. Microbiol. Biotechnol.* **42,** 813–838 (2015).
10. Venkata Mohan, S. *et al.* Waste biorefinery models towards sustainable circular bioeconomy: Critical review and future perspectives. *Bioresour. Technol.* **215,** 2–12 (2016).
11. Carlson, R. Estimating the biotech sector's contribution to the US economy. *Nat. Biotechnol.* **34,** 247–255 (2016).
12. Chao, R., Mishra, S., Si, T. & Zhao, H. Engineering biological systems using automated biofoundries. *Metab. Eng.* **42,** 98–108 (2017).
13. Smolke, C. D. Building outside of the box: iGEM and the BioBricks Foundation. *Nat. Biotechnol.* **27,** 1099–1102 (2009).
14. Kamens, J. The Addgene repository: an international nonprofit plasmid and data resource. *Nucleic Acids Res.* **43,** D1152–D1157 (2015).
15. Weber, T. *et al.* antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **43,** W237–W243 (2015).
16. Stanton, B. C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10,** 99–105 (2014).
17. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26,** 1135–1145 (2008).
18. Huttenhower, C. & Hofmann, O. A Quick Guide to Large-Scale Genomic Data Mining. *PLoS Comput. Biol.* **6,** e1000779 (2010).
19. Nielsen, A. A. K., Segall-Shapiro, T. H. & Voigt, C. A. Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Curr. Opin. Chem. Biol.* **17,** 878–892 (2013).
20. Carbonell, P. *et al.* Bioinformatics for the synthetic biology of natural products: integrating across the Design–Build–Test cycle. *Nat. Prod. Rep.* **33,** 925–932 (2016).
21. Nielsen, A. A. K. *et al.* Genetic circuit design automation. *Science (80-.).* **352,** aac7341-aac7341 (2016).
22. Hsu, P. D., Lander, E. S. & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157,** 1262–1278 (2014).
23. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science (80-.).* **346,** 1258096–1258096 (2014).
24. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460,** 894–898 (2009).
25. Hughes, R. A. & Ellington, A. D. Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology. *Cold Spring Harb. Perspect. Biol.* **9,** a023812 (2017).

26. Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **11,** 499–507 (2014).

27. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

28. Gibson, D. G. *et al.* Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science (80-.).* **329,** 52–56 (2010).

29. Richardson, S. M. *et al.* Design of a synthetic yeast genome. *Science (80-.).* **355,** 1040–1044 (2017).

30. Lajoie, M. J. *et al.* Genomically Recoded Organisms Expand Biological Functions. *Science (80-.).* **342,** 357–360 (2013).

31. Engler, C., Kandzia, R. & Marillonnet, S. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS One* **3,** e3647 (2008).

32. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS One* **6,** e16765 (2011).

33. Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* **4,** 975–986 (2015).

34. Moore, S. J. *et al.* EcoFlex: A Multifunctional MoClo Kit for E. coli Synthetic Biology. *ACS Synth. Biol.* **5,** 1059–1069 (2016).

35. Iverson, S. V., Haddock, T. L., Beal, J. & Densmore, D. M. CIDAR MoClo: Improved MoClo Assembly Standard and New E. coli Part Library Enable Rapid Combinatorial Design for Synthetic and Traditional Biology. *ACS Synth. Biol.* **5,** 99–103 (2016).

36. Engler, C. *et al.* A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synth. Biol.* **3,** 839–843 (2014).

37. Casini, A., Storch, M., Baldwin, G. S. & Ellis, T. Bricks and blueprints: methods and standards for DNA assembly. *Nat. Rev. Mol. Cell Biol.* **16,** 568–576 (2015).

38. Casini, A. *et al.* A Pressure Test to Make 10 Molecules in 90 Days: External Evaluation of Methods to Engineer Biology. *J. Am. Chem. Soc.* **140,** 4302–4316 (2018).

39. Chambers, S., Kitney, R. & Freemont, P. The Foundry: the DNA synthesis and construction Foundry at Imperial College. *Biochem. Soc. Trans.* **44,** 687–688 (2016).

40. Jeschek, M., Gerngross, D. & Panke, S. Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. *Nat. Commun.* **7,** 11163 (2016).

41. Biggs, B. W., De Paepe, B., Santos, C. N. S., De Mey, M. & Kumaran Ajikumar, P. Multivariate modular metabolic engineering for pathway and strain optimization. *Curr. Opin. Biotechnol.* **29,** 156–162 (2014).

42. Nielsen, J. & Keasling, J. D. Engineering Cellular Metabolism. *Cell* **164,** 1185–1197 (2016).

43. Rogers, J. K., Taylor, N. D. & Church, G. M. Biosensor-based engineering of biosynthetic pathways. *Curr. Opin. Biotechnol.* **42,** 84–91 (2016).

44. Williams, T. C., Pretorius, I. S. & Paulsen, I. T. Synthetic Evolution of Metabolic Productivity Using Biosensors. *Trends Biotechnol.* **34,** 371–381 (2016).

45. Zhang, J., Jensen, M. K. & Keasling, J. D. Development of biosensors and their application in metabolic engineering. *Curr. Opin. Chem. Biol.* **28,** 1–8 (2015).

46. Rohman, M. & Wingfield, J. High-Throughput Screening Using Mass Spectrometry within Drug Discovery. in *Methods in Molecular Biology* **1439,** 47–63 (2016).

47. Yan, C. *et al.* Real-Time Screening of Biocatalysts in Live Bacterial Colonies. *J. Am. Chem. Soc.* **139,** 1408–1411 (2017).

48. Venil, C. K., Zakaria, Z. A. & Ahmad, W. A. Bacterial pigments and their applications. *Process Biochem.* **48,** 1065–1079 (2013).

49. Yadav, V. G., De Mey, M., Giaw Lim, C., Kumaran Ajikumar, P. & Stephanopoulos, G. The future of metabolic engineering and synthetic biology: Towards a systematic practice. *Metab. Eng.* **14,** 233–241 (2012).

50. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440,** 940–943 (2006).

51. Paddon, C. J. *et al.* High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496,** 528–532 (2013).

52. Lee, M. E., Aswani, A., Han, A. S., Tomlin, C. J. & Dueber, J. E. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Res.* **41,** 10668–10678 (2013).

53. Xu, P., Rizzoni, E. A., Sul, S.-Y. & Stephanopoulos, G. Improving Metabolic Pathway Efficiency by Statistical Model-Based Multivariate Regulatory Metabolic Engineering. *ACS Synth. Biol.* **6,** 148–158 (2017).

54. Carpenter, A., Paulsen, I. & Williams, T. Blueprints for Biosensors: Design, Limitations, and Applications. *Genes (Basel).* **9,** 375 (2018).

55. Ehrenworth, A. M., Claiborne, T. & Peralta-Yahya, P. Medium-Throughput Screen of Microbially Produced Serotonin via a G-Protein-Coupled Receptor-Based Sensor. *Biochemistry* **56,** 5471–5475 (2017).

56. Terekhov, S. S. *et al.* Microfluidic droplet platform for ultrahigh-throughput single-cell screening of biodiversity. *Proc. Natl. Acad. Sci.* **114,** 2550–2555 (2017).

57. Slomovic, S., Pardee, K. & Collins, J. J. Synthetic biology devices for in vitro and in vivo diagnostics. *Proc. Natl. Acad. Sci.* **112,** 14429–14435 (2015).

58. Kojima, R., Aubel, D. & Fussenegger, M. Toward a world of theranostic medication: Programming biological sentinel systems for therapeutic intervention. *Adv. Drug Deliv. Rev.* **105,** 66–76 (2016).

59. Alon, U. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8,** 450–461 (2007).

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

61. Briat, C., Gupta, A. & Khammash, M. Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Biomolecular Networks. *Cell Syst.* **2,** 15–26 (2016).

62. Shong, J., Jimenez Diaz, M. R. & Collins, C. H. Towards synthetic microbial consortia for bioprocessing. *Curr. Opin. Biotechnol.* **23,** 798–802 (2012).

63. Brenner, K., You, L. & Arnold, F. H. Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol.* **26,** 483–489 (2008).

64. Scott, S. R. & Hasty, J. Quorum Sensing Communication Modules for Microbial Consortia. *ACS Synth. Biol.* **5,** 969–977 (2016).

65. Kylilis, N., Tuza, Z. A., Stan, G.-B. & Polizzi, K. M. Tools for engineering coordinated system behaviour in synthetic microbial consortia. *Nat. Commun.* **9,** 2677 (2018).

66. Kong, W., Meldgin, D. R., Collins, J. J. & Lu, T. Designing microbial consortia with defined social interactions. *Nat. Chem. Biol.* **14,** 821–829 (2018).

67. Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nat. Rev. Genet.* **11,** 367–379 (2010).

68. Ye, H., Baba, M. D.-E., Peng, R.-W. & Fussenegger, M. A Synthetic Optogenetic Transcription Device Enhances Blood-Glucose Homeostasis in Mice. *Science (80-.).* **332,** 1565–1568 (2011).

69. Zhao, Y. *et al.* An Expanded Palette of Genetically Encoded Ca2+ Indicators. *Science (80-.).* **333,** 1888–1891 (2011).

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

71. Adeniran, A., Sherer, M. & Tyo, K. E. J. Yeast-based biosensors: design and applications. *FEMS Yeast Res.* **15,** 1–15 (2014).

72. Pu, J., Zinkus-Boltz, J. & Dickinson, B. C. Evolution of a split RNA polymerase as a versatile biosensor platform. *Nat. Chem. Biol.* **13,** 432–438 (2017).

73. Brown, M. *et al.* Lac repressor can regulate expression from a hybrid SV40 early promoter containing a lac operator in animal cells. *Cell* **49,** 603–612 (1987).

74. Ausländer, S. & Fussenegger, M. From gene switches to mammalian designer cells: present and future prospects. *Trends Biotechnol.* **31,** 155–168 (2013).

75. Tigges, M. & Fussenegger, M. Recent advances in mammalian synthetic biology—design of synthetic transgene control networks. *Curr. Opin. Biotechnol.* **20,** 449–460 (2009).

76. Blazeck, J. & Alper, H. S. Promoter engineering: Recent advances in controlling transcription at the most

fundamental level. *Biotechnol. J.* **8,** 46–58 (2013).

77. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci.* **89,** 5547–5551 (1992).

78. Urlinger, S. *et al.* Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci.* **97,** 7963–7968 (2000).

79. Gitzinger, M., Kemmer, C., El-Baba, M. D., Weber, W. & Fussenegger, M. Controlling transgene expression in subcutaneous implants using a skin lotion containing the apple metabolite phloretin. *Proc. Natl. Acad. Sci.* **106,** 10638–10643 (2009).

80. Gitzinger, M. *et al.* The food additive vanillic acid controls transgene expression in mammalian cells and mice. *Nucleic Acids Res.* **40,** e37–e37 (2012).

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

82. Weber, W. *et al.* A synthetic mammalian electro-genetic transcription circuit. *Nucleic Acids Res.* **37,** e33–e33 (2008).

83. Müller, K. *et al.* Multi-chromatic control of mammalian gene expression and signaling. *Nucleic Acids Res.* **41,** e124–e124 (2013).

84. D'Ambrosio, V. & Jensen, M. K. Lighting up yeast cell factories by transcription factor-based biosensors. *FEMS Yeast Res.* **17,** 1–12 (2017).

85. Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* **30,** 354–359 (2012).

86. Hector, R. E. & Mertens, J. A. A Synthetic Hybrid Promoter for Xylose-Regulated Control of Gene Expression in Saccharomyces Yeasts. *Mol. Biotechnol.* **59,** 24–33 (2017).

87. Li, S., Si, T., Wang, M. & Zhao, H. Development of a Synthetic Malonyl-CoA Sensor in Saccharomyces cerevisiae for Intracellular Metabolite Monitoring and Genetic Screening. *ACS Synth. Biol.* **4,** 1308–1315 (2015).

88. Binder, S. *et al.* A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biol.* **13,** R40 (2012).

89. Mahr, R. *et al.* Biosensor-driven adaptive laboratory evolution of L-valine production in Corynebacterium glutamicum. *Metab. Eng.* **32,** 184–194 (2015).

90. Leavitt, J. M. *et al.* Biosensor-Enabled Directed Evolution to Improve Muconic Acid Production in Saccharomyces cerevisiae. *Biotechnol. J.* **12,** 1600687 (2017).

91. Esvelt, K. M., Carlson, J. C. & Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **472,** 499–503 (2011).

92. Brödel, A. K., Jaramillo, A. & Isalan, M. Intracellular directed evolution of proteins from combinatorial libraries based on conditional phage replication. *Nat. Protoc.* **12,** 1830–1843 (2017).

93. Eddy, S. R. Non–coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2,** 919–929 (2001).

94. Mattick, J. S. & Makunin, I. V. Non-coding RNA. *Hum. Mol. Genet.* **15,** R17–R29 (2006).

95. Mironov, A. S. *et al.* Sensing Small Molecules by Nascent RNA. *Cell* **111,** 747–756 (2002).

96. Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C. & Breaker, R. R. Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. *Cell* **113,** 577–586 (2003).

97. Liang, J. C., Bloom, R. J. & Smolke, C. D. Engineering Biological Systems with Synthetic RNA Molecules. *Mol. Cell* **43,** 915–926 (2011).

98. Isaacs, F. J., Dwyer, D. J. & Collins, J. J. RNA synthetic biology. *Nat. Biotechnol.* **24,** 545–554 (2006).

99. Henkin, T. M. Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev.* **22,** 3383–3390 (2008).

100. Good, L. Translation repression by antisense sequences. *Cell. Mol. Life Sci.* **60,** 854–861 (2003).

101. Bayer, T. S. & Smolke, C. D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat. Biotechnol.* **23,** 337–343 (2005).

102. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428,** 281–286 (2004).

103. Yen, L. *et al.* Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. *Nature* **431,** 471–476 (2004).

104. McKeague, M., Wang, Y.-H. & Smolke, C. D. In Vitro Screening and in Silico Modeling of RNA-Based Gene Expression Control. *ACS Chem. Biol.* **10,** 2463–2467 (2015).

105. Townshend, B., Kennedy, A. B., Xiang, J. S. & Smolke, C. D. High-throughput cellular RNA device engineering. *Nat. Methods* **12,** 989–994 (2015).

106. Yang, J. *et al.* Synthetic RNA devices to expedite the evolution of metabolite-producing microbes. *Nat. Commun.* **4,** 1413 (2013).

107. Lynch, S. A. & Gallivan, J. P. A flow cytometry-based screen for synthetic riboswitches. *Nucleic Acids Res.* **37,** 184–192 (2009).

108. Muranaka, N., Sharma, V., Nomura, Y. & Yokobayashi, Y. An efficient platform for genetic selection and screening of gene switches in Escherichia coli. *Nucleic Acids Res.* **37,** e39–e39 (2009).

109. Kiel, C., Yus, E. & Serrano, L. Engineering Signal Transduction Pathways. *Cell* **140,** 33–47 (2010).

110. Kollmann, M., Løvdok, L., Bartholomé, K., Timmer, J. & Sourjik, V. Design principles of a bacterial signalling network. *Nature* **438,** 504–507 (2005).

111. Ravikumar, S., Baylon, M. G., Park, S. J. & Choi, J. Engineered microbial biosensors based on bacterial two-component systems as synthetic biotechnology platforms in bioremediation and biorefinery. *Microb. Cell Fact.* **16,** 62 (2017).

112. Casino, P., Rubio, V. & Marina, A. The mechanism of signal transduction by two-component systems. *Curr. Opin. Struct. Biol.* **20,** 763–771 (2010).

113. Maruthamuthu, M. kannan, Ganesh, I., Ravikumar, S. & Hong, S. H. Evaluation of zraP gene expression characteristics and construction of a lead (Pb) sensing and removal system in a recombinant Escherichia coli. *Biotechnol. Lett.* **37,** 659–664 (2015).

114. Ganesh, I., Ravikumar, S., Yoo, I. & Hong, S. H. Construction of malate-sensing Escherichia coli by introduction of a novel chimeric two-component system. *Bioprocess Biosyst. Eng.* **38,** 797–804 (2015).

115. Ganesh, I., Ravikumar, S., Lee, S. H., Park, S. J. & Hong, S. H. Engineered fumarate sensing Escherichia coli based on novel chimeric two-component system. *J. Biotechnol.* **168,** 560–566 (2013).

116. Levskaya, A. *et al.* Engineering Escherichia coli to see light. *Nature* **438,** 441–442 (2005).

117. Clarke, E. J. & Voigt, C. A. Characterization of combinatorial patterns generated by multiple two-component sensors in E. coli that respond to many stimuli. *Biotechnol. Bioeng.* **108,** 666–675 (2011).

118. Shalaeva, D. N., Galperin, M. Y. & Mulkidjanian, A. Y. Eukaryotic G protein-coupled receptors as descendants of prokaryotic sodium-translocating rhodopsins. *Biol. Direct* **10,** 63 (2015).

119. Wickstrand, C., Dods, R., Royant, A. & Neutze, R. Bacteriorhodopsin: Would the real structural intermediates please stand up? *Biochim. Biophys. Acta - Gen. Subj.* **1850,** 536–553 (2015).

120. Nussinov, R. The spatial structure of cell signaling systems. *Phys. Biol.* **10,** 045004 (2013).

121. Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. Computational design of receptor and sensor proteins with novel functions. *Nature* **423,** 185–190 (2003).

122. Öling, D. *et al.* Large Scale Synthetic Site Saturation GPCR Libraries Reveal Novel Mutations That Alter Glucose Signaling. *ACS Synth. Biol.* **7,** 2317–2321 (2018).

123. Brocchieri, L. Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Res.* **33,** 3390–3400 (2005).

124. Dueber, J. E. Reprogramming Control of an Allosteric Signaling Switch Through Modular Recombination. *Science (80-.).* **301,** 1904–1908 (2003).

125. Pierce, K. L., Premont, R. T. & Lefkowitz, R. J. Seven-transmembrane receptors. *Nat. Rev. Mol. Cell Biol.* **3,** 639–650 (2002).

126. Fredriksson, R. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Mol. Pharmacol.* **63,** 1256–1272 (2003).

127. Heng, B. C., Aubel, D. & Fussenegger, M. G Protein–Coupled Receptors Revisited: Therapeutic Applications Inspired by Synthetic Biology. *Annu. Rev. Pharmacol. Toxicol.* **54,** 227–249 (2014).

128. Katritch, V., Cherezov, V. & Stevens, R. C. Diversity and modularity of G protein-coupled receptor structures. *Trends Pharmacol. Sci.* **33,** 17–27 (2012).

129. de Mendoza, A., Sebé-Pedrós, A. & Ruiz-Trillo, I. The Evolution of the GPCR Signaling System in Eukaryotes: Modularity, Conservation, and the Transition to Metazoan Multicellularity. *Genome Biol. Evol.* **6,** 606–619 (2014).

130. Dorsam, R. T. & Gutkind, J. S. G-protein-coupled receptors and cancer. *Nat. Rev. Cancer* **7,** 79–94 (2007).

131. Xie, M. & Fussenegger, M. Designing cell function: assembly of synthetic gene circuits for cell biology applications. *Nat. Rev. Mol. Cell Biol.* **19,** 507–525 (2018).

132. Xue, S. *et al.* A Synthetic-Biology-Inspired Therapeutic Strategy for Targeting and Treating Hepatogenous Diabetes. *Mol. Ther.* **25,** 443–455 (2017).

133. Saxena, P. *et al.* A programmable synthetic lineage-control network that differentiates human IPSCs into glucose-sensitive insulin-secreting beta-like cells. *Nat. Commun.* **7,** 11247 (2016).

134. Airan, R. D., Thompson, K. R., Fenno, L. E., Bernstein, H. & Deisseroth, K. Temporally precise in vivo control of intracellular signalling. *Nature* **458,** 1025–1029 (2009).

135. Muir, A. J., Sanders, L. L., Wilkinson, W. E. & Schmader, K. Reducing medication regimen complexity. *J. Gen. Intern. Med.* **16,** 77–82 (2001).

136. Bai, P. *et al.* A synthetic biology-based device prevents liver injury in mice. *J. Hepatol.* **65,** 84–94 (2016).

137. Spehr, M. & Munger, S. D. Olfactory receptors: G protein-coupled receptors and beyond. *J. Neurochem.* **109,** 1570–1583 (2009).

138. Fleischer, J. Mammalian olfactory receptors. *Front. Cell. Neurosci.* **3,** 1–10 (2009).

139. Du, L., Wu, C., Liu, Q., Huang, L. & Wang, P. Recent advances in olfactory receptor-basedbiosensors. *Biosens. Bioelectron.* **42,** 570–580 (2013).

140. Glatz, R. & Bailey-Hill, K. Mimicking nature's noses: From receptor deorphaning to olfactory biosensing. *Prog. Neurobiol.* **93,** 270–296 (2011).

141. Lu, D., Lu, F., Geng, L. & Pang, G. Recent Advances in Olfactory Receptor (OR) Biosensors and Cell Signaling Cascade Amplification Systems. *Sensors Mater.* **30,** 67 (2018).

142. Panula, P. *et al.* International Union of Basic and Clinical Pharmacology. XCVIII. Histamine Receptors. *Pharmacol. Rev.* **67,** 601–655 (2015).

143. Nichols, D. E. & Nichols, C. D. Serotonin Receptors. *Chem. Rev.* **108,** 1614–1641 (2008).

144. Kruse, A. C. *et al.* Muscarinic acetylcholine receptors: novel opportunities for drug development. *Nat. Rev. Drug Discov.* **13,** 549–560 (2014).

145. Tfelt-Hansen, J. & Brown, E. M. THE CALCIUM-SENSING RECEPTOR IN NORMAL PHYSIOLOGY AND PATHOPHYSIOLOGY: A Review. *Crit. Rev. Clin. Lab. Sci.* **42,** 35–70 (2005).

146. Niswender, C. M. & Conn, P. J. Metabotropic Glutamate Receptors: Physiology, Pharmacology, and Disease. *Annu. Rev. Pharmacol. Toxicol.* **50,** 295–322 (2010).

147. Adams, M. N. *et al.* Structure, function and pathophysiology of protease activated receptors. *Pharmacol. Ther.* **130,** 248–282 (2011).

148. Authier, F. & Desbuquois, B. Glucagon receptors. *Cell. Mol. Life Sci.* **65,** 1880–1899 (2008).

149. Simoni, M. The Follicle-Stimulating Hormone Receptor: Biochemistry, Molecular Biology, Physiology, and Pathophysiology. *Endocr. Rev.* **18,** 739–773 (1997).

150. Choi, J. W. *et al.* LPA Receptors: Subtypes and Biological Actions. *Annu. Rev. Pharmacol. Toxicol.* **50,** 157–186 (2010).

151. Rosen, H., Stevens, R. C., Hanson, M., Roberts, E. & Oldstone, M. B. A. Sphingosine-1-Phosphate and Its Receptors: Structure, Signaling, and Influence. *Annu. Rev. Biochem.* **82,** 637–662 (2013).

152. Fredholm, B. B., IJzerman, a P., Jacobson, K. a, Klotz, K. N. & Linden, J. International Union of Pharmacology. XXV. Nomenclature and Classification of Adenosine Receptors. *Pharmacol. Rev.* **53,** 527–552 (2001).

153. Zhou, X. E., Melcher, K. & Xu, H. E. Structure and activation of rhodopsin. *Acta Pharmacol. Sin.* **33,** 291–299 (2012).

154. Shen, W. L. *et al.* Function of Rhodopsin in Temperature Discrimination in Drosophila. *Science (80-.).* **331,** 1333–

1336 (2011).

155. Pertwee, R. G. The pharmacology of cannabinoid receptors and their ligands: an overview. *Int. J. Obes.* **30,** S13–S18 (2006).

156. Waldhoer, M., Bartlett, S. E. & Whistler, J. L. Opioid Receptors. *Annu. Rev. Biochem.* **73,** 953–990 (2004).

157. Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science (80-.).* **349,** 1095–1100 (2015).

158. Volpe, D. A. *et al.* Uniform assessment and ranking of opioid Mu receptor binding constants for selected opioid drugs. *Regul. Toxicol. Pharmacol.* **59,** 385–390 (2011).

159. Horn, F. GPCRDB: an information system for G protein-coupled receptors. *Nucleic Acids Res.* **26,** 275–279 (1998).

160. Pándy-Szekeres, G. *et al.* GPCRdb in 2018: adding GPCR structure models and ligands. *Nucleic Acids Res.* **46,** D440–D446 (2018).

161. Palczewski, K. G Protein–Coupled Receptor Rhodopsin. *Annu. Rev. Biochem.* **75,** 743–767 (2006).

162. Macfarlane, S. R., Seatter, M. J., Kanke, T., Hunter, G. D. & Plevin, R. Proteinase-Activated Receptors. *Pharmacol. Rev.* **53,** 245 LP-282 (2001).

163. Dong, S., Rogan, S. C. & Roth, B. L. Directed molecular evolution of DREADDs: a generic approach to creating next-generation RASSLs. *Nat. Protoc.* **5,** 561–573 (2010).

164. Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S. & Roth, B. L. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl. Acad. Sci.* **104,** 5163–5168 (2007).

165. Urban, D. J. & Roth, B. L. DREADDs (Designer Receptors Exclusively Activated by Designer Drugs): Chemogenetic Tools with Therapeutic Utility. *Annu. Rev. Pharmacol. Toxicol.* **55,** 399–417 (2015).

166. Adeniran, A., Stainbrook, S., Bostick, J. W. & Tyo, K. E. J. Detection of a Peptide Biomarker by Engineered Yeast Receptors. *ACS Synth. Biol.* **7,** 696–705 (2018).

167. Molina Grima, E., Belarbi, E.-H., Acién Fernández, F. ., Robles Medina, A. & Chisti, Y. Recovery of microalgal biomass and metabolites: process options and economics. *Biotechnol. Adv.* **20,** 491–515 (2003).

168. Ledesma-Amaro, R., Dulermo, R., Niehus, X. & Nicaud, J.-M. Combining metabolic engineering and process optimization to improve production and secretion of fatty acids. *Metab. Eng.* **38,** 38–46 (2016).

169. Lundstrom, K. Present and future approaches to screening of G-protein-coupled receptors. *Future Med. Chem.* **5,** 523–538 (2013).

170. Zhang, R. & Xie, X. Tools for GPCR drug discovery. *Acta Pharmacol. Sin.* **33,** 372–384 (2012).

171. Bradley, S. A. *et al.* Fermentanomics: Monitoring Mammalian Cell Cultures with NMR Spectroscopy. *J. Am. Chem. Soc.* **132,** 9531–9533 (2010).

172. Barrett, T. A., Wu, A., Zhang, H., Levy, M. S. & Lye, G. J. Microwell engineering characterization for mammalian cell culture process development. *Biotechnol. Bioeng.* **105,** 260–275 (2010).

173. Phelan, M. C. Basic Techniques in Mammalian Cell Tissue Culture. *Curr. Protoc. Cell Biol.* **36,** 1.1.1-1.1.18 (2007).

174. Kenakin, T. New concepts in pharmacological efficacy at 7TM receptors: IUPHAR Review 2. *Br. J. Pharmacol.* **168,** 554–575 (2013).

175. Neves, S. R. G Protein Pathways. *Science (80-.).* **296,** 1636–1639 (2002).

176. Karathia, H., Vilaprinyo, E., Sorribas, A. & Alves, R. Saccharomyces cerevisiae as a Model Organism: A Comparative Study. *PLoS One* **6,** e16015 (2011).

177. Hedges, S. B. The origin and evolution of model organisms. *Nat. Rev. Genet.* **3,** 838–849 (2002).

178. Ankeny, R. A. & Leonelli, S. What's so special about model organisms? *Stud. Hist. Philos. Sci. Part A* **42,** 313–323 (2011).

179. Bardwell, L. A walk-through of the yeast mating pheromone response pathway. *Peptides* **25,** 1465–1476 (2004).

180. Ladds, G., Goddard, A. & Davey, J. Functional analysis of heterologous GPCR signalling pathways in yeast. *Trends Biotechnol.* **23,** 367–373 (2005).

181. Versele, M., Lemaire, K. & Thevelein, J. M. Sex and sugar in yeast: two distinct GPCR systems. *EMBO Rep.* **2,** 574–579 (2001).

182. Nielsen, J. Yeast cell factories on the horizon. *Science (80-.).* **349,** 1050–1051 (2015).

183. Nielsen, J., Larsson, C., van Maris, A. & Pronk, J. Metabolic engineering of yeast for production of fuels and chemicals. *Curr. Opin. Biotechnol.* **24,** 398–404 (2013).

184. Nevoigt, E. Progress in Metabolic Engineering of Saccharomyces cerevisiae. *Microbiol. Mol. Biol. Rev.* **72,** 379–412 (2008).

185. Moqtaderi, Z. & Geisberg, J. V. Construction of Mutant Alleles in Saccharomyces cerevisiae without Cloning: Overview and the Delitto Perfetto Method. in *Current Protocols in Molecular Biology* 13.10C.1-13.10C.17 (John Wiley & Sons, Inc., 2013). doi:10.1002/0471142727.mb1310cs104

186. Hammer, S. K. & Avalos, J. L. Harnessing yeast organelles for metabolic engineering. *Nat. Chem. Biol.* **13,** 823–832 (2017).

187. Lian, J., Mishra, S. & Zhao, H. Recent advances in metabolic engineering of Saccharomyces cerevisiae : New tools and their applications. *Metab. Eng.* 0–1 (2018). doi:10.1016/j.ymben.2018.04.011

188. Bai Flagfeldt, D., Siewers, V., Huang, L. & Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast* **26,** 545–551 (2009).

189. Curran, K. a. *et al.* Short Synthetic Terminators for Improved Heterologous Gene Expression in Yeast. *ACS Synth. Biol.* **4,** 824–832 (2015).

190. Curran, K. a., Karim, A. S., Gupta, A. & Alper, H. S. Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab. Eng.* **19,** 88–97 (2013).

191. Schrader, E. K., Harstad, K. G. & Matouschek, A. Targeting proteins for degradation. *Nat. Chem. Biol.* **5,** 815–822 (2009).

192. Chen, R. *et al.* A Barcoding Strategy Enabling Higher-Throughput Library Screening by Microscopy. *ACS Synth. Biol.* **4,** 1205–1216 (2015).

193. Idiris, A., Tohda, H., Kumagai, H. & Takegawa, K. Engineering of protein secretion in yeast: strategies and impact on protein production. *Appl. Microbiol. Biotechnol.* **86,** 403–417 (2010).

194. Wang, Y., San, K.-Y. & Bennett, G. N. Cofactor engineering for advancing chemical biotechnology. *Curr. Opin. Biotechnol.* **24,** 994–999 (2013).

195. Ravikumar, A., Arrieta, A. & Liu, C. C. An orthogonal DNA replication system in yeast. *Nat. Chem. Biol.* **10,** 175–177 (2014).

196. Ravikumar, A., Arzumanyan, G. A., Obadi, M. K. A., Javanpour, A. A. & Liu, C. C. Scalable continuous evolution of genes at mutation rates above genomic error thresholds. *bioRxiv* (2018).

197. Lian, J., HamediRad, M., Hu, S. & Zhao, H. Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* **8,** 1688 (2017).

198. Jovicevic, D., Blount, B. A. & Ellis, T. Total synthesis of a eukaryotic chromosome: Redesigning and SCRaMbLE-ing yeast. *BioEssays* **36,** 855–860 (2014).

199. Blount, B. A. *et al.* Rapid host strain improvement by in vivo rearrangement of a synthetic yeast chromosome. *Nat. Commun.* **9,** 1932 (2018).

200. Blount, B. A., Weenink, T. & Ellis, T. Construction of synthetic regulatory networks in yeast. *FEBS Lett.* **586,** 2112–2121 (2012).

201. Jarque, S., Bittner, M. & Hilscherová, K. Freeze-drying as suitable method to achieve ready-to-use yeast biosensors for androgenic and estrogenic compounds. *Chemosphere* **148,** 204–210 (2016).

202. Alvaro, C. G. & Thorner, J. Heterotrimeric G Protein-coupled Receptor Signaling in Yeast Mating Pheromone Response. *J. Biol. Chem.* **291,** 7788–7795 (2016).

203. Cullen, P. J. & Sprague, G. F. The Regulation of Filamentous Growth in Yeast. *Genetics* **190,** 23–49 (2012).

204. Hohmann, S. An integrated view on a eukaryotic osmoregulation system. *Curr. Genet.* **61,** 373–382 (2015).

205. Paliwal, S. *et al.* MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. *Nature* **446,** 46–51 (2007).

206. Moore, T. I., Chou, C.-S., Nie, Q., Jeon, N. L. & Yi, T.-M. Robust Spatial Sensing of Mating Pheromone Gradients by Yeast Cells. *PLoS One* **3,** e3865 (2008).

207. Pearson, G. *et al.* Mitogen-Activated Protein (MAP) Kinase Pathways: Regulation and Physiological Functions 1. *Endocr. Rev.* **22,** 153–183 (2001).

208. King, K., Dohlman, H., Thorner, J., Caron, M. & Lefkowitz, R. Control of yeast mating signal transduction by a mammalian beta 2-adrenergic receptor and Gs alpha subunit. *Science (80-.).* **250,** 121–123 (1990).

209. Beukers, M. W. Random Mutagenesis of the Human Adenosine A2B Receptor Followed by Growth Selection in Yeast. Identification of Constitutively Active and Gain of Function Mutations. *Mol. Pharmacol.* **65,** 702–710 (2004).

210. Weston, C., Poyner, D., Patel, V., Dowell, S. & Ladds, G. Investigating G protein signalling bias at the glucagon-like peptide-1 receptor in yeast. *Br. J. Pharmacol.* **171,** 3651–3665 (2014).

211. Kokkola, T. *et al.* Mutagenesis of Human Mel1aMelatonin Receptor Expressed in Yeast Reveals Domains Important for Receptor Function. *Biochem. Biophys. Res. Commun.* **249,** 531–536 (1998).

212. Erlenbach, I. *et al.* Functional expression of M1, M3 and M5 muscarinic acetylcholine receptors in yeast. *J. Neurochem.* **77,** 1327–1337 (2001).

213. Yoo, J. I. & O'Malley, M. A. Tuning Vector Stability and Integration Frequency Elevates Functional GPCR Production and Homogeneity in Saccharomyces cerevisiae. *ACS Synth. Biol.* **7,** 1763–1772 (2018).

214. McCusker, E. C., Bane, S. E., O'Malley, M. A. & Robinson, A. S. Heterologous GPCR Expression: A Bottleneck to Obtaining Crystal Structures. *Biotechnol. Prog.* **23,** 540–547 (2008).

215. Miret, J. J., Rakhilina, L., Silverman, L. & Oehlen, B. Functional Expression of Heteromeric Calcitonin Gene-related Peptide and Adrenomedullin Receptors in Yeast. *J. Biol. Chem.* **277,** 6881–6887 (2002).

216. Weston, C. *et al.* Modulation of Glucagon Receptor Pharmacology by Receptor Activity-modifying Protein-2 (RAMP2). *J. Biol. Chem.* **290,** 23009–23022 (2015).

217. Brown, A. J. *et al.* Functional coupling of mammalian receptors to the yeast mating pathway using novel yeast/mammalian G protein α-subunit chimeras. *Yeast* **16,** 11–22 (2000).

218. Hagen, D. C., McCaffrey, G. & Sprague, G. F. Pheromone response elements are necessary and sufficient for basal and pheromone-induced transcription of the FUS1 gene of Saccharomyces cerevisiae. *Mol. Cell. Biol.* **11,** 2952–2961 (1991).

219. Dowell, S. J. & Brown, A. J. Yeast Assays for G Protein-Coupled Receptors. in *G Protein-Coupled Receptors in Drug Discovery* (ed. Leifert, W. R.) **552,** 213–229 (Humana Press, 2009).

220. Price, L. a, Kajkowski, E. M., Hadcock, J. R., Ozenberger, B. a & Pausch, M. H. Functional coupling of a mammalian somatostatin receptor to the yeast pheromone response pathway. *Mol. Cell. Biol.* **15,** 6188–6195 (1995).

221. Erickson, J. R. *et al.* Edg-2/Vzg-1 Couples to the Yeast Pheromone Response Pathway Selectively in Response to Lysophosphatidic Acid. *J. Biol. Chem.* **273,** 1506–1510 (1998).

222. Ishii, J. *et al.* Yeast-Based Fluorescence Reporter Assay of G Protein-coupled Receptor Signalling for Flow Cytometric Screening: FAR1-Disruption Recovers Loss of Episomal Plasmid Caused by Signalling in Yeast. *J. Biochem.* **143,** 667–674 (2008).

223. Crowe, M. L., Perry, B. N. & Connerton, I. F. G Olf Complements A Gpa1 Null Mutation in Olf Saccharomyces Cerevisiae and Functionally Couples to the Ste2 Pheromone Receptor. *J. Recept. Signal Transduct.* **20,** 61–73 (2000).

224. Pajot-Augy, E., Crowe, M., Levasseur, G., Salesse, R. & Connerton, I. Engineered Yeasts as Reporter Systems for Odorant Detection. *J. Recept. Signal Transduct.* **23,** 155–171 (2003).

225. Minic, J. *et al.* Functional expression of olfactory receptors in yeast and development of a bioassay for odorant screening. *FEBS J.* **272,** 524–537 (2005).

226. Fukutani, Y. *et al.* The N-terminal replacement of an olfactory receptor for the development of a Yeast-based biomimetic odor sensor. *Biotechnol. Bioeng.* **109,** 205–212 (2012).

227. Fukutani, Y. *et al.* Improving the odorant sensitivity of olfactory receptor-expressing yeast with accessory proteins. *Anal. Biochem.* **471,** 1–8 (2015).

228. Mukherjee, K., Bhattacharyya, S. & Peralta-Yahya, P. GPCR-Based Chemical Biosensors for Medium-Chain Fatty Acids. *ACS Synth. Biol.* **4,** 1261–1269 (2015).

229. Ostrov, N. *et al.* A modular yeast biosensor for low-cost point-of-care pathogen detection. *Sci. Adv.* **3,** e1603221 (2017).

230. Kar, S., Paglialunga, S. & Islam, R. Cystatin C Is a More Reliable Biomarker for Determining eGFR to Support Drug Development Studies. *J. Clin. Pharmacol.* **58,** 1239–1247 (2018).

231. Ang, J., Harris, E., Hussey, B. J., Kil, R. & McMillen, D. R. Tuning Response Curves for Synthetic Biology. *ACS Synth. Biol.* **2,** 547–567 (2013).

232. Landry, B. P., Palanki, R., Dyulgyarov, N., Hartsough, L. A. & Tabor, J. J. Phosphatase activity tunes two-component system sensor detection threshold. *Nat. Commun.* **9,** 1433 (2018).

233. De Paepe, B., Maertens, J., Vanholme, B. & De Mey, M. Modularization and Response Curve Engineering of a Naringenin-Responsive Transcriptional Biosensor. *ACS Synth. Biol.* **7,** 1303–1314 (2018).

234. Dueber, J. E., Mirsky, E. A. & Lim, W. A. Engineering synthetic signaling proteins with ultrasensitive input/output control. *Nat. Biotechnol.* **25,** 660–662 (2007).

235. Regot, S. *et al.* Distributed biological computation with multicellular engineered networks. *Nature* **469,** 207–211 (2011).

236. Macia, J. *et al.* Implementation of Complex Biological Logic Circuits Using Spatially Distributed Multicellular Consortia. *PLOS Comput. Biol.* **12,** e1004685 (2016).

237. Bhattacharyya, R. P. The Ste5 Scaffold Allosterically Modulates Signaling Output of the Yeast Mating Pathway. *Science (80-.).* **311,** 822–826 (2006).

238. Hao, N. *et al.* Regulation of Cell Signaling Dynamics by the Protein Kinase-Scaffold Ste5. *Mol. Cell* **30,** 649–656 (2008).

239. Bashor, C. J., Helman, N. C., Yan, S. & Lim, W. A. Using Engineered Scaffold Interactions to Reshape MAP Kinase Pathway Signaling Dynamics. *Science (80-.).* **319,** 1539–1543 (2008).

240. Ingolia, N. T. & Murray, A. W. Positive-Feedback Loops as a Flexible Biological Module. *Curr. Biol.* **17,** 668–677 (2007).

241. Galloway, K. E., Franco, E. & Smolke, C. D. Dynamically Reshaping Signaling Networks to Program Cell Fate via Genetic Controllers. *Science (80-.).* **341,** 1235005–1235005 (2013).

242. Ritter, S. L. & Hall, R. A. Fine-tuning of GPCR activity by receptor-interacting proteins. *Nat. Rev. Mol. Cell Biol.* **10,** 819–830 (2009).

243. Reiter, E. & Lefkowitz, R. J. GRKs and β-arrestins: roles in receptor silencing, trafficking and signaling. *Trends Endocrinol. Metab.* **17,** 159–165 (2006).

244. Chan, L. Y., Kosuri, S. & Endy, D. Refactoring bacteriophage T7. *Mol. Syst. Biol.* **1,** E1–E10 (2005).

245. Temme, K., Zhao, D. & Voigt, C. A. Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. *Proc. Natl. Acad. Sci.* **109,** 7085–7090 (2012).

246. Fowler, M., Beck, K., Brant, J., Opdyke, W. & Roberts, D. *Refactoring: improving the design of existing code*. (Addison-Wesley Professional, 1999).

247. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12,** 381–390 (2014).

248. Kofahl, B. & Klipp, E. Modelling the dynamics of the yeast pheromone pathway. *Yeast* **21,** 831–850 (2004).

249. Bridge, L. J., Mead, J., Frattini, E., Winfield, I. & Ladds, G. Modelling and simulation of biased agonism dynamics at a G protein-coupled receptor. *J. Theor. Biol.* **442,** 44–65 (2018).

250. Houser, J. R., Ford, E., Nagiec, M. J., Errede, B. & Elston, T. C. Positive roles for negative regulators in the mating response of yeast. *Mol. Syst. Biol.* **8,** 1–10 (2012).

251. Alexander, W. G. A history of genome editing in Saccharomyces cerevisiae. *Yeast* **35,** 355–360 (2018).

252. Bush, A. *et al.* Yeast GPCR signaling reflects the fraction of occupied receptors, not the number. *Mol. Syst. Biol.* **12,** 898 (2016).

253. Bakker, R. A., Schoonus, S. B. J., Smit, M. J., Timmerman, H. & Leurs, R. Histamine H1-Receptor Activation of Nuclear Factor-κB: Roles for Gβγ- and Gα q/11 -Subunits in Constitutive and Agonist-Mediated Signaling. *Mol. Pharmacol.* **60,** 1133–1142 (2001).

254. Burstein, E. S., Spalding, T. a & Brann, M. R. Pharmacology of Muscarinic Receptor Subtypes Constitutively Activated by G Proteins. *Mol. Pharmacol.* **51,** 312–319 (1997).

255. Brewster, R. C. *et al.* The Transcription Factor Titration Effect Dictates Level of Gene Expression. *Cell* **156,** 1312–1323 (2014).

256. Pi, H., Chien, C. T. & Fields, S. Transcriptional activation upon pheromone stimulation mediated by a small domain of Saccharomyces cerevisiae Ste12p. *Mol. Cell. Biol.* **17,** 6410–6418 (1997).

257. Blazeck, J. & Alper, H. S. Promoter engineering: Recent advances in controlling transcription at the most fundamental level. *Biotechnol. J.* **8,** 46–58 (2013).

258. Rantasalo, A. *et al.* Synthetic Transcription Amplifier System for Orthogonal Control of Gene Expression in Saccharomyces cerevisiae. *PLoS One* **11,** e0148320 (2016).

259. Roberts, C. J. *et al.* Signaling and Circuitry of Multiple MAPK Pathways Revealed by a Matrix of Global Gene Expression Profiles. *Science (80-.).* **287,** 873–880 (2000).

260. Ryan, O. W. *et al.* Selection of chromosomal DNA libraries using a multiplex CRISPR system. *Elife* **3,** e03703 (2014).

261. Gnügge, R. & Rudolf, F. Saccharomyces cerevisiae Shuttle vectors. *Yeast* **34,** 205–221 (2017).

262. Agmon, N. *et al.* Yeast Golden Gate (yGG) for the Efficient Assembly of S. cerevisiae Transcription Units. *ACS Synth. Biol.* **4,** 853–859 (2015).

263. Guo, Y. *et al.* YeastFab: the design and construction of standard biological parts for metabolic engineering in Saccharomyces cerevisiae. *Nucleic Acids Res.* **43,** e88–e88 (2015).

264. Shaw, W. M. *et al.* Engineering a model cell for rational tuning of GPCR signaling. *BioRxiv* 1–59 (2018).

265. Storici, F. & Resnick, M. A. The Delitto Perfetto Approach to In Vivo Site-Directed Mutagenesis and Chromosome Rearrangements with Synthetic Oligonucleotides in Yeast. in *Methods in Enzymology* **409,** 329–345 (2006).

266. Jakočiūnas, T., Jensen, M. K. & Keasling, J. D. CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* **34,** 44–59 (2016).

267. Baker Brachmann, C. *et al.* Designer deletion strains derived fromSaccharomyces cerevisiae S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14,** 115–132 (1998).

268. Winston, F., Dollard, C. & Ricupero-Hovasse, S. L. Construction of a set of convenientsaccharomyces cerevisiae strains that are isogenic to S288C. *Yeast* **11,** 53–55 (1995).

269. Engel, S. R. *et al.* The Reference Genome Sequence of Saccharomyces cerevisiae : Then and Now. *G3* **4,** 389–398 (2014).

270. Elion, E. a, Brill, J. a & Fink, G. R. FUS3 represses CLN1 and CLN2 and in concert with KSS1 promotes signal transduction. *Proc. Natl. Acad. Sci.* **88,** 9392–9396 (1991).

271. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34,** 184–191 (2016).

272. Casini, A. *et al.* R2oDNA Designer: Computational Design of Biologically Neutral Synthetic DNA Sequences. *ACS Synth. Biol.* **3,** 525–528 (2014).

273. Glaser, A., McColl, B. & Vadolas, J. GFP to BFP Conversion: A Versatile Assay for the Quantification of CRISPR/Cas9-mediated Genome Editing. *Mol. Ther. - Nucleic Acids* **5,** e334 (2016).

274. Dohlman, H. G. & Thorner, J. Regulation of G Protein–Initiated Signal Transduction in Yeast: Paradigms and Principles. *Annu. Rev. Biochem.* **70,** 703–754 (2001).

275. Aldridge, B. B., Burke, J. M., Lauffenburger, D. A. & Sorger, P. K. Physicochemical modelling of cell signalling pathways. *Nat. Cell Biol.* **8,** 1195–1203 (2006).

276. Kholodenko, B. N. Cell-signalling dynamics in time and space. *Nat. Rev. Mol. Cell Biol.* **7,** 165–176 (2006).

277. Fukuda, N., Kaishima, M., Ishii, J. & Honda, S. Positive Detection of GPCR Antagonists Using a System for Inverted Expression of a Fluorescent Reporter Gene. *ACS Synth. Biol.* **6,** 1554–1562 (2017).

278. Turner, J. J., Ewald, J. C. & Skotheim, J. M. Cell Size Control in Yeast. *Curr. Biol.* **22,** R350–R359 (2012).

279. Bashor, C. J., Horwitz, A. A., Peisajovich, S. G. & Lim, W. A. Rewiring Cells: Synthetic Biology as a Tool to Interrogate the Organizational Principles of Living Systems. *Annu. Rev. Biophys.* **39,** 515–537 (2010).

280. Dolan, J. W. & Fields, S. Overproduction of the yeast STE12 protein leads to constitutive transcriptional induction. *Genes Dev.* **4,** 492–502 (1990).

281. Blazeck, J., Garg, R., Reed, B. & Alper, H. S. Controlling promoter strength and regulation in Saccharomyces cerevisiae using synthetic hybrid promoters. *Biotechnol. Bioeng.* **109,** 2884–2895 (2012).

282. Redden, H. & Alper, H. S. The development and characterization of synthetic minimal yeast promoters. *Nat. Commun.* **6,** 7810 (2015).

283. Rantasalo, A., Kuivanen, J., Penttilä, M., Jäntti, J. & Mojzita, D. Synthetic Toolkit for Complex Genetic Circuit Engineering in Saccharomyces cerevisiae. *ACS Synth. Biol.* **7,** 1573–1587 (2018).

284. Mohana-Borges, R. *et al.* LexA Repressor Forms Stable Dimers in Solution. *J. Biol. Chem.* **275,** 4708–4712 (2000).

285. Zhang, A. P. P., Pigli, Y. Z. & Rice, P. A. Structure of the LexA–DNA complex and implications for SOS box measurement. *Nature* **466,** 883–886 (2010).

286. Su, T.-C., Tamarkina, E. & Sadowski, I. Organizational constraints on Ste12 cis-elements for a pheromone response in Saccharomyces cerevisiae. *FEBS J.* **277,** 3235–3248 (2010).

287. Neher, S. B. Latent ClpX-recognition signals ensure LexA destruction after DNA damage. *Genes Dev.* **17,** 1084–1089 (2003).

288. Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. A short amino acid sequence able to specify nuclear location. *Cell* **39,** 499–509 (1984).

289. Jack, B. R. *et al.* Predicting the Genetic Stability of Engineered DNA Sequences with the EFM Calculator. *ACS Synth. Biol.* **4,** 939–943 (2015).

290. McIsaac, R. S., Gibney, P. A., Chandran, S. S., Benjamin, K. R. & Botstein, D. Synthetic biology tools for programming gene expression without nutritional perturbations in Saccharomyces cerevisiae. *Nucleic Acids Res.* **42,** e48–e48 (2014).

291. Gilbert, L. A. *et al.* CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* **154,** 442–451 (2013).

292. Gander, M. W., Vrana, J. D., Voje, W. E., Carothers, J. M. & Klavins, E. Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. *Nat. Commun.* **8,** 15459 (2017).

293. Kipniss, N. H. *et al.* Engineering cell sensing and responses using a GPCR-coupled CRISPR-Cas system. *Nat. Commun.* **8,** 2212 (2017).

294. Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nat. Biotechnol.* **32,** 1262–1267 (2014).

295. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16,** 129–143 (2015).

296. Milligan, G. Constitutive Activity and Inverse Agonists of G Protein-Coupled Receptors: a Current Perspective. *Mol. Pharmacol.* **64,** 1271–1276 (2003).

297. Seifert, R. & Wenzel-Seifert, K. Constitutive activity of G-protein-coupled receptors: cause of disease and common property of wild-type receptors. *Naunyn. Schmiedebergs. Arch. Pharmacol.* **366,** 381–416 (2002).

298. Sommers, C. M. *et al.* A Limited Spectrum of Mutations Causes Constitutive Activation of the Yeast α-Factor Receptor. *Biochemistry* **39,** 6898–6909 (2000).

299. Vecchio, E. A. *et al.* Ligand-Independent Adenosine A2B Receptor Constitutive Activity as a Promoter of Prostate Cancer Cell Proliferation. *J. Pharmacol. Exp. Ther.* **357,** 36–44 (2016).

300. Knight, A. *et al.* Discovery of Novel Adenosine Receptor Agonists That Exhibit Subtype Selectivity. *J. Med. Chem.* **59,** 947–964 (2016).

301. Ryzhov, S. *et al.* Differential role of the carboxy-terminus of the A2B adenosine receptor in stimulation of adenylate cyclase, phospholipase Cβ, and interleukin-8. *Purinergic Signal.* **5,** 289–298 (2009).

302. Wootten, D., Christopoulos, A., Marti-Solano, M., Babu, M. M. & Sexton, P. M. Mechanisms of signalling and biased agonism in G protein-coupled receptors. *Nat. Rev. Mol. Cell Biol.* (2018). doi:10.1038/s41580-018-0049-3

303. Zheng, W., Zhao, H., Mancera, E., Steinmetz, L. M. & Snyder, M. Genetic analysis of variation in transcription factor binding in yeast. *Nature* **464,** 1187–1191 (2010).

304. Hoffman, C. S. Except in Every Detail: Comparing and Contrasting G-Protein Signaling in Saccharomyces cerevisiae and Schizosaccharomyces pombe. *Eukaryot. Cell* **4,** 495–503 (2005).

305. Croft, W. *et al.* A Physiologically Required G Protein-coupled Receptor (GPCR)-Regulator of G Protein Signaling

(RGS) Interaction That Compartmentalizes RGS Activity. *J. Biol. Chem.* **288,** 27327–27342 (2013).

306. Edelstein, S. J. & Le Novère, N. Cooperativity of Allosteric Receptors. *J. Mol. Biol.* **425,** 1424–1432 (2013).

307. Vallée-Bélisle, A., Ricci, F. & Plaxco, K. W. Engineering Biosensors with Extended, Narrowed, or Arbitrarily Edited Dynamic Range. *J. Am. Chem. Soc.* **134,** 2876–2879 (2012).

308. Kang, D., Vallée-Bélisle, A., Porchetta, A., Plaxco, K. W. & Ricci, F. Re-engineering Electrochemical Biosensors To Narrow or Extend Their Useful Dynamic Range. *Angew. Chemie Int. Ed.* **51,** 6717–6721 (2012).

309. Buchler, N. E. & Louis, M. Molecular Titration and Ultrasensitivity in Regulatory Networks. *J. Mol. Biol.* **384,** 1106–1119 (2008).

310. Ferrell, J. E. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr. Opin. Cell Biol.* **14,** 140–148 (2002).

311. Groß, A., Rödel, G. & Ostermann, K. Application of the yeast pheromone system for controlled cell-cell communication and signal amplification. *Lett. Appl. Microbiol.* **52,** 521–526 (2011).

312. Julius, D., Blair, L., Brake, A., Sprague, G. & Thorner, J. Yeast alpha factor is processed from a larger precursor polypeptide: the essential role of a membrane-bound dipeptidyl aminopeptidase. *Cell* **32,** 839–52 (1983).

313. Germann, S. M. *et al.* Glucose-based microbial production of the hormone melatonin in yeast Saccharomyces cerevisiae. *Biotechnol. J.* **11,** 717–724 (2016).

314. Gika, H. G., Wilson, I. D. & Theodoridis, G. A. LC–MS-based holistic metabolic profiling. Problems, limitations, advantages, and future perspectives. *J. Chromatogr. B* **966,** 1–6 (2014).

315. Ni, J., Tao, F., Du, H. & Xu, P. Mimicking a natural pathway for de novo biosynthesis: natural vanillin production from accessible carbon sources. *Sci. Rep.* **5,** 13670 (2015).

316. Lee, D., Lloyd, N. D. R., Pretorius, I. S. & Borneman, A. R. Heterologous production of raspberry ketone in the wine yeast Saccharomyces cerevisiae via pathway engineering and synthetic enzyme fusion. *Microb. Cell Fact.* **15,** 49 (2016).

317. Modena, D., Trentini, M., Corsini, M., Bombaci, A. & Giorgetti, A. OlfactionDB: A Database of Olfactory Receptors and Their Ligands. *Adv. Life Sci.* **1,** 1–5 (2012).

318. Davis, R. M., Muller, R. Y. & Haynes, K. A. Can the Natural Diversity of Quorum-Sensing Advance Synthetic Biology? *Front. Bioeng. Biotechnol.* **3,** 1–10 (2015).

319. Urrios, A. *et al.* A Synthetic Multicellular Memory Device. *ACS Synth. Biol.* **5,** 862–873 (2016).

320. Williams, T. C., Nielsen, L. K. & Vickers, C. E. Engineered quorum sensing using pheromone-mediated cell-to-cell communication in saccharomyces cerevisiae. *ACS Synth. Biol.* **2,** 136–149 (2013).

321. Hennig, S., Rödel, G. & Ostermann, K. Artificial cell-cell communication as an emerging tool in synthetic biology applications. *J. Biol. Eng.* **9,** 13 (2015).

322. Ward, L. D. & Kellis, M. Interpreting noncoding genetic variation in complex traits and human disease. *Nat. Biotechnol.* **30,** 1095–1106 (2012).

323. Hauser, A. S. *et al.* Pharmacogenomics of GPCR Drug Targets. *Cell* **172,** 41–54.e19 (2018).

324. Good, M., Tang, G., Singleton, J., Reményi, A. & Lim, W. A. The Ste5 Scaffold Directs Mating Signaling by Catalytically Unlocking the Fus3 MAP Kinase for Activation. *Cell* **136,** 1085–1097 (2009).

325. Yu, R. C. *et al.* The Alpha Project: a model system for systems biology research. *IET Syst. Biol.* **2,** 222–233 (2008).

326. Boeke, J. D. *et al.* The Genome Project-Write. *Science (80-.).* **353,** 126–127 (2016).

327. Prezeau, L. *et al.* Functional crosstalk between GPCRs: with or without oligomerization. *Curr. Opin. Pharmacol.* **10,** 6–13 (2010).

328. Chung, C. T., Niemela, S. L. & Miller, R. H. One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci.* **86,** 2172–2175 (1989).

329. Gietz, R. D. & Schiestl, R. H. Microtiter plate transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2,** 5–8 (2007).

330. Lõoke, M., Kristjuhan, K. & Kristjuhan, A. Extraction of genomic DNA from yeasts for PCR-based applications. *Biotechniques* **50,** 325–328 (2011).

331. Yan, P., Gao, X., Shen, W., Zhou, P. & Duan, J. Parallel assembly for multiple site-directed mutagenesis of plasmids. *Anal. Biochem.* **430,** 65–67 (2012).

332. Mitchell, L. A. *et al.* Versatile genetic assembly system (VEGAS) to assemble pathways for expression in S. cerevisiae. *Nucleic Acids Res.* **43,** 6620–6630 (2015).

333. Liu, Z. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci. Rep.* **7,** 2193 (2017).

334. Obst, U., Lu, T. K. & Sieber, V. A Modular Toolkit for Generating Pichia pastoris Secretion Libraries. *ACS Synth. Biol.* **6,** 1016–1025 (2017).

335. Leonard, S. P. *et al.* Genetic Engineering of Bee Gut Microbiome Bacteria with a Toolkit for Modular Assembly of Broad-Host-Range Plasmids. *ACS Synth. Biol.* **7,** 1279–1290 (2018).

336. Storici, F., Durham, C. L., Gordenin, D. a & Resnick, M. a. Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. *Proc. Natl. Acad. Sci.* **100,** 14994–14999 (2003).

337. DiCarlo, J. E. *et al.* Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res.* **41,** 4336–4343 (2013).

338. Yuen, G. *et al.* CRISPR/Cas9-mediated gene knockout is insensitive to target copy number but is dependent on guide RNA potency and Cas9/sgRNA threshold expression level. *Nucleic Acids Res.* **45,** 12039–12053 (2017).

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

340. Bao, Z. *et al.* Homology-Integrated CRISPR–Cas (HI-CRISPR) System for One-Step Multigene Disruption in Saccharomyces cerevisiae. *ACS Synth. Biol.* **4,** 585–594 (2015).

341. Jakočiūnas, T. *et al.* Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. *Metab. Eng.* **28,** 213–222 (2015).

342. Ronda, C. *et al.* CrEdit: CRISPR mediated multi-loci gene integration in Saccharomyces cerevisiae. *Microb. Cell Fact.* **14,** 97 (2015).

343. Stovicek, V., Borodina, I. & Forster, J. CRISPR–Cas system enables fast and simple genome editing of industrial Saccharomyces cerevisiae strains. *Metab. Eng. Commun.* **2,** 13–22 (2015).

344. Tsai, C.-S. *et al.* Rapid and marker-free refactoring of xylose-fermenting yeast strains with Cas9/CRISPR. *Biotechnol. Bioeng.* **112,** 2406–2411 (2015).

345. Mans, R. *et al.* CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. *FEMS Yeast Res.* **15,** 1–15 (2015).

346. Horwitz, A. A. *et al.* Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas. *Cell Syst.* **1,** 88–96 (2015).

347. Nishimasu, H. *et al.* Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156,** 935–949 (2014).

348. Gao, Y. & Zhao, Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J. Integr. Plant Biol.* **56,** 343–349 (2014).

349. Walter, J. M., Chandran, S. S. & Horwitz, A. A. CRISPR-Cas-Assisted Multiplexing (CAM): Simple Same-Day Multi-Locus Engineering in Yeast. *J. Cell. Physiol.* **231,** 2563–2569 (2016).

350. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163,** 759–771 (2015).

351. Verwaal, R., Buiting-Wiessenhaan, N., Dalhuijsen, S. & Roubos, J. A. CRISPR/Cpf1 enables fast and simple genome editing of Saccharomyces cerevisiae. *Yeast* **35,** 201–211 (2018).

352. Liu, Y. *et al.* Engineering cell signaling using tunable CRISPR–Cpf1-based transcription factors. *Nat. Commun.* **8,** 2095 (2017).

353. Arimbasseri, A. G., Rijal, K. & Maraia, R. J. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription* **5,** e27369 (2014).