Biosynthetic pathways provide an enzymatic route from inexpensive renewable resources to valuable metabolic products such as pharmaceuticals and plastics. Designing these pathways is challenging due to the complexities of biology. Advances in the design and construction of genetic variants has enabled billions of cells, each possessing a slightly different metabolic design, to be rapidly generated. However, our ability to measure the quality of these designs lags by several orders of magnitude. Recent research has enabled cells to report their own success in chemical production through the use of genetically encoded biosensors. A new engineering discipline is emerging around the creation and application of biosensors. Biosensors, implemented in selections and screens to identify productive cells, are paving the way for a new era of biotechnological progress.

Address
Wyss Institute for Biologically Inspired Engineering Harvard University, 3 Blackfan Circle, Boston, MA 02115, USA

Corresponding author: Church, George M (gchurch@genetics.med.harvard.edu)
Both authors contributed equally to this work.

Metabolic engineering and the evaluation bottleneck

The enzymatic processes of cellular metabolism perform chemical conversions with exquisite specificity and speed. By engineering metabolism, we can harness these processes for human needs, such as the industrial production of organic chemicals, fuels, and polymers. Indeed, these microbial metabolic products make up a large and rapidly growing segment of the ‘bioeconomy’ [1,2] (Figure 1a).

Despite its value, metabolic engineering faces significant challenges to mature as an engineering discipline. Biological ‘parts’, such as genes that encode enzymes, or promoters that direct their expression, can be highly context-dependent. As a result, many design attempts are required before an optimal set of parts is identified. This problem is exacerbated by the slow and expensive chromatographic methods used to measure cellular chemical production during design evaluation [3]. Consequently, design-build-test cycles are painstaking, with development times measured in years or decades (Figure 1b).

While traditional engineering disciplines have had great success in a bottom-up approach to design, intracellular complexity has hampered bottom-up engineering in biology. Bottom-up design relies on well-characterized parts that follow well-defined physical rules. This, in turn, enables the design of whole systems that behave as theory and calculation would predict. In contrast, efforts to design whole organisms using bottom-up approaches are rudimentary. Regardless of the initial intention of forward engineering, metabolic engineers must rely on making many modifications to an organism’s genome before finding a design that works [4].

As a field, metabolic engineering has begun to shift toward an engineering paradigm relying on principles borrowed from Darwinian evolution. Because evolution acts top-down on biological functions rather than mechanisms, coopting evolutionary principles can yield a more efficient way to achieve biological design goals. Top-down design is achieved by defining the requirements of the complete system, without specifying the parameters of the lower-level components. If the top-level design goal of metabolite production can be evaluated through a screen or selection, rather than traditional metabolite measurement techniques, millions of designs can be evaluated rapidly. This enables an optimal combination of genetic parts to be determined without comprehensive knowledge of each individual part.

Screens and selections are enabled by biosensors that transduce intracellular metabolite concentration into gene expression changes. When metabolite production is linked to the expression of a fluorescent protein, high-throughput methods such as flow cytometry can be used to evaluate potential designs (Figure 1c). Because each individual design is evaluated, this biosensor configuration is an example of screening. Alternatively, if the biosensor actuates the expression of an antibiotic resistance gene, then only cells producing high levels of the desired compound survive an antibiotic challenge [5]. Selections can also be engineered through auxotrophy, where gene expression complements a nutrient deficiency that would otherwise inhibit growth [6]. Because the collection of designs is placed in an environment where only the best survive, and individual designs are not inspected, this is an example of selection. The choice of using a screen or selection in a specific metabolic engineering project depends on the specifics of the experiment. Screens
may be more suitable for identifying cells producing toxic compounds, whereas selections may be simpler if the design space is very large. Regardless of the paradigm employed, the single-cell resolution and high throughput evaluation provided by metabolite biosensors allow millions of variants to be assayed extremely quickly. Biosensors enable multiplexed phenotype evaluation, which transforms the engineering design-build-test cycle into an evolutionary mutate-test cycle, and finally allows engineers to test huge numbers of metabolic pathway designs in a rapid, iterative manner.

**Metabolic design and modification strategies**

To understand the need for multiplexed evaluation of metabolite production within single cells, it is instructive to inspect the metabolic design and genomic modification capabilities that are now available. Advances in these ‘design’ and ‘build’ steps, which can produce so many potential designs, require similarly high throughput tools to evaluate their success in achieving metabolic design goals.

**Pathway design and strain optimization**

It is often the case that a host organism does not naturally produce a desired compound. In these cases algorithms that rely on a catalog of known enzymes can be used to identify the shortest or most economical routes to the target compound [7**]. These heterologous enzymes can be installed within the working strain to enable production of the target compound. Such *de novo* biosynthetic pathways have enabled biological production of several valuable compounds. A notable example was the engineering of *Escherichia coli* to produce the non-natural plastic precursor, 1,4-butane-diol (BDO), requiring addition of five exogenous enzymes [8]. In this study, over 10,000 pathways were predicted, a daunting number to evaluate through traditional methods; only two were chosen for experimental analysis.

Once a strain can produce a target chemical, further optimization is necessary to achieve suitable production outcomes such as metabolite concentration, production rate and stoichiometric yield. The algorithms and techniques available for pathway prediction and optimization have been well reviewed [7**,9]. Though a great improvement over the nearly infinite space of random mutagenesis, these methods still generate thousands to millions of high quality guesses about which combinations of metabolic changes will yield the most productive strains. These metabolic designs must be cloned and experimentally tested to identify productive variants and to validate design methods for further improvement.

**Genome engineering**

Techniques to encode predicted metabolic designs by making mutations to microbial genes and genomes in a
targeted, multiplexed fashion are becoming mature. Oligonucleotide-mediated genome editing in E. coli, called multiplexed automated genome engineering (MAGE), makes use of the phage lambda beta protein to integrate oligonucleotides bearing desired changes in the place of Okazaki fragments during DNA replication \[10-12\]. CRISPR-based genome editing techniques are expanding genomic modification techniques to other bacteria \[13\] and yeast \[14,15\**\]. Multiplexed modification of genomes allows the millions of biosynthetic pathway variants generated by in silico metabolic methods to be constructed in vivo.

Analogously, advances in DNA synthesis, especially microarray-synthesized oligonucleotide pools \[16\], and the assembly of these pools into full-length genes \[17\] are enabling the construction of rationally designed gene libraries and collections of metagenomically mined orthologues. Competition among companies supplying raw oligonucleotide pools — Agilent Technologies, Custom Array and Twist Biosciences — and those supplying fully synthesized genes — Blue Heron, DNA 2.0, Integrated DNA Technologies, Genewiz, Gen9, Genscript, Twist Biosciences and more — is driving down the price of DNA \[18\] and transforming the capability of metabolic engineers to encode in silico designs into physical DNA.

**The role of biosensors in metabolic engineering**

**Chemical measurement is a screening bottleneck**

The impact of increased throughput in the design and construction of genetic elements is diminished while design evaluation remains a bottleneck. The gold standard of metabolic measurement, using liquid chromatography or mass spectrometry, is limited to around 10^3 measurements per instrument, per day, with the best equipment \[19\]. More commonly, fewer than 10^2 measurements can be made per instrument, per day. One hundred thousand carefully constructed pathway variants, or one hundred million computationally predicted enzyme active sites are wasted if only a small fraction of these can be assayed for function.

Conspicuous molecules, which are colorful, fluorescent or aid cell fitness, illustrate the power of screening multiplexed mutants: Wang, et al. were able to optimize the bioproduction of lycopene, a bright red carotenoid, by generating an estimated 15 billion unique genetic variants and visually screening almost 10^6 of these to identify mutants with the highest reported production titer in just three days \[10\]. Most molecules of interest lack such convenient spectroscopic properties and are not essential for cell growth. For this majority, a mechanism is required to couple the presence of the inconspicuous molecule to a conspicuous reporter or fitness advantage.

**Biosensors let cells make chemical measurements**

A genetically encoded biosensor propagates molecular recognition of a target molecule into biological actuation within a cell. In this way, each cell is able to ‘measure’ the concentration of an intracellular metabolite, and reports this concentration via a conspicuous gene product, providing the engineer with a multiplexed method to detect cell biosynthetic productivity. Both screens (Figure 2a) and selections (Figure 2b) are used to enrich for productive cells. Biosensors can be gauged for effectiveness via several metrics. An ideal biosensor would operate over a wide range of concentrations, have a high signal-to-noise ratio, a low false positive rate and high molecular specificity \[3\].

The range of concentrations over which the biosensor exhibits a change in output is the operational range of the biosensor. Operational range is determined by measuring the concentrations over which the biosensor shows a graded, concentration-dependent change in response (Figure 3a). The desired operational concentration varies by application: to detect new enzyme activity, nanomolar sensitivity might be ideal, but millimolar sensitivity would be more useful in optimizing a pathway to produce grams per liter of a target compound. Several strategies for modulating biosensor operational range have been developed \[5**\].

The signal-to-noise of a biosensor, also referred to as the dynamic range of the system, can be quantified as the ratio of the highest measured output of the biosensor to the lowest measured output of the biosensor (Figure 3a). Dynamic range can be affected by the number of copies of the biosensor within the cell \[20**\], or by signal amplification using an enzymatic reporter \[21\]. The larger the dynamic range, the more reliably a true signal can be discerned from noise.

False positives arise when spurious transcription, translation or protein activity create an erroneous signal that is not related to target molecule detection (Figure 3b). The false positive rate dictates the number of designs that can be interrogated when searching for rare successes. If the false positive rate is one in 1000, searching for productive variants that exist at a rate of one in every 10,000 would yield a majority of cells that are not actually productive. Multiple cycles of enrichment may help, but when the biosensor output is a fitness advantage, such as antibiotic resistance, cells that erroneously survive may take over the population. There are a number of genetic modifications and strategies for counter-selection that drastically decrease biosensor false positive rates \[5**\].

The most crucial biosensor characteristic is molecular specificity. For biosensor-directed metabolic engineering to be possible, a biosensor must be available for the target molecule. Fortunately, cells have evolved a wide array of
mechanisms to sense and respond to intracellular metabolite concentration.

**Generation of new biosensors**

Previously developed small molecule biosensors are overwhelmingly co-opted from the natural sensory machinery of cells (Table 1). Allosterically regulated transcription factors (aTFs) change their affinity for an operator DNA sequence through a conformational change enacted by ligand binding and are useful for directly controlling gene transcription [21]. Ligand-dependent protein dimerization [22] and ligand-conditional protein stability are related methods that require binding of a small molecule to stabilize a protein dimer interface or a protein monomer, respectively. Stabilization methods lead to direct changes in fluorescent reporter protein function or mediate transcriptional changes via an additional two-hybrid system.

Riboswitches are 5' untranslated regions composed of RNA that bind small molecule ligands to control the stability of the mRNA transcript, directly affecting the translation of the encoded genes [23,24]. Sensors based on fluorescence resonance energy transfer (FRET) [25] use a conformational change in a protein domain brought about by ligand binding to change the proximity of two fluorophores capable of excitation-emission photon transfer enabling direct ligand detection. The function of newly discovered members of bacterial sensory gene families can often be inferred by their proximity to the operons they regulate. Metagenomic sequence mining will continue to expand the repertoire of natural sensors that are available [26].

Where natural sensory domains have not been found, or have yet to evolve (i.e. for a synthetic target molecule),
Figure 3

![Diagram](current-opinion-in-biotechnology.png)

Defining engineering parameters for biosensors is a prerequisite for biosensor-based metabolic engineering to mature. (a) The relationship between biosensor output and product concentration is the biosensor transfer function. The range of concentrations over which the biosensor functions is the operational range. The intensity of the biosensor response is the dynamic range. (b) The false positive rate of a biosensor system determines the maximum number of designs that can be evaluated in a given experiment. A working selection results in enrichment of highly productive cells. A failed selection results in enrichment of unproductive cells that erroneously survive through mutation or other means.

Table 1

<table>
<thead>
<tr>
<th>Molecule(s)</th>
<th>Molecule type</th>
<th>Sensor</th>
<th>Sensor type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Natural biosensors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-Butanol</td>
<td>Fatty alcohol, fuel</td>
<td>BmoR</td>
<td>Allosteric TF</td>
<td>[35]</td>
</tr>
<tr>
<td>Acrylate</td>
<td>Plastic precursor</td>
<td>AcrU</td>
<td>Allosteric TF</td>
<td>[20]**</td>
</tr>
<tr>
<td>Adipate</td>
<td>Dicarboxylic acid</td>
<td>PcrR</td>
<td>Allosteric TF</td>
<td>[35]</td>
</tr>
<tr>
<td>B12</td>
<td>Vitamin</td>
<td>BtuB</td>
<td>Riboswitch</td>
<td>[50]</td>
</tr>
<tr>
<td>Benzoate, naphthalene</td>
<td>Aromatics</td>
<td>NahR</td>
<td>Allosteric TF</td>
<td>[51]</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>Macrolide</td>
<td>MphR</td>
<td>Allosteric TF</td>
<td>[52]</td>
</tr>
<tr>
<td>Fatty acids</td>
<td>Fatty acid</td>
<td>FatR</td>
<td>Allosteric TF</td>
<td>[53]</td>
</tr>
<tr>
<td>Glucarate</td>
<td>Feedstock</td>
<td>CdaR</td>
<td>Allosteric TF</td>
<td>[55]</td>
</tr>
<tr>
<td>Lysine</td>
<td>Amino acid</td>
<td>LysR</td>
<td>Allosteric TF</td>
<td>[56]</td>
</tr>
<tr>
<td>Muconate</td>
<td>Dicarboxylic acid</td>
<td>BenM</td>
<td>Allosteric TF</td>
<td>[57]</td>
</tr>
<tr>
<td>NADPH</td>
<td>Redox</td>
<td>SoxR</td>
<td>Allosteric TF</td>
<td>[58]</td>
</tr>
<tr>
<td>Naringenin</td>
<td>Flavonoid</td>
<td>TgrR</td>
<td>Allosteric TF</td>
<td>[59]</td>
</tr>
<tr>
<td>Octane</td>
<td>Alkane</td>
<td>AlkS</td>
<td>Allosteric TF</td>
<td>[60]</td>
</tr>
<tr>
<td>Succinate</td>
<td>Dicarboxylic acid</td>
<td>DcuR</td>
<td>Two-component</td>
<td>[36]</td>
</tr>
<tr>
<td>Tetracyclines</td>
<td>Polyketides</td>
<td>TetR</td>
<td>Allosteric TF</td>
<td>[21]</td>
</tr>
<tr>
<td><strong>Engineered biosensors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,4-Dihydroxybenzoxide</td>
<td>Aromatic</td>
<td>PobR</td>
<td>Allosteric TF</td>
<td>[28]</td>
</tr>
<tr>
<td>Biphenyl, nitrotoleunes</td>
<td>Aromatics</td>
<td>XylR</td>
<td>Allosteric TF</td>
<td>[61]</td>
</tr>
<tr>
<td>Mevalonate</td>
<td>Isoprenoid precursor</td>
<td>AraC</td>
<td>Allosteric TF</td>
<td>[31]</td>
</tr>
<tr>
<td>Pyruvate</td>
<td>Alpha-keto acid</td>
<td>De novo</td>
<td>FRET</td>
<td>[62]</td>
</tr>
<tr>
<td>Theophylline</td>
<td>Alkaloid</td>
<td>De novo</td>
<td>Riboswitch</td>
<td>[63]</td>
</tr>
<tr>
<td>Thiamine-pp</td>
<td>Vitamin</td>
<td>De novo</td>
<td>Riboswitch</td>
<td>[64,65]</td>
</tr>
<tr>
<td>Trehalose-6-p</td>
<td>Sugar</td>
<td>De novo</td>
<td>FRET</td>
<td>[66]</td>
</tr>
<tr>
<td>Triacetic acid lactones</td>
<td>Feedstock</td>
<td>AraC</td>
<td>Allosteric TF</td>
<td>[67]</td>
</tr>
<tr>
<td>Vanillin</td>
<td>Aromatic, flavoring</td>
<td>QacR</td>
<td>Allosteric TF</td>
<td>[29]</td>
</tr>
<tr>
<td>Zn^{2+}</td>
<td>Ion</td>
<td>De novo</td>
<td>FRET</td>
<td>[68]</td>
</tr>
</tbody>
</table>
sensor engineering via directed evolution or computational design can expand or alter the molecular specificity of existing sensors, or even create new ones de novo (Table 1). Computational approaches for engineering biosensor specificity rely on a range of strategies that include structure-based predictions [27], homology modeling [28] and mechanistic insights [29]. Computational design of ligand binding interfaces has also shown success [30**]. Random mutagenesis, or saturation mutagenesis of key positions, has shown promise in changing the specificity of allosteric TFs, including AraC [31] and LuxR [32].

**Biosensors for multiplexed phenotype evaluation**
Biosensor-based screens have been demonstrated using a number of different reporters: fluorescence, insoluble pigments, luminescence and antibiotic resistance. Biosensors have been used to screen for increased microbial production of the isopenoid precursor mevalonate [31], L-lysine [33,34*], 1-butanol [35], and triacetic acid lactones [36*]. Luciferase has been used as a reporter to screen for production of macrolides [37], or to detect tolune and related compounds [38]. Other recent examples of biosensor-based screens identifying optimized production conditions include increased production of phenol [39*], arginine and histidine [40**], 3,4 dihydroxy benzoate [41] and methionine [42].

Biosensor-based selections couple biosensor output to antibiotic resistance and use cell fitness as a proxy for target metabolite production. This strategy has been used to identify improved 1-butanol production plasmids [35] and for whole-pathway iterated selection resulting in genomes evolved for higher production of glucarate or naringenin [5**]. Other biosensor-based selections have been successful in optimizing production of N-acetyl glucosamine [43*] and lysine [44]. In theory, selection enables library sizes limited only by the size of the culture. In reality, the false positive rate of the selection imposes practical constraints that are often far lower. Fortunately, several strategies exist for attaining desired false-positive rates [5**]. Together, these works demonstrate that biosensors are a viable strategy for screening to improve metabolic pathways. Multiplexed engineering in biology is further explored in a recent perspective [45**].

**The future of biosensor-based metabolic engineering**
Biosensors offer an attractive, multiplexed phenotype screening solution with the potential to revolutionize metabolic engineering. Thus far, pathway production gains have been modest, and have not approached the grams per liter production titers mature pathways required for commercial scale [8]. As an emerging field, most studies have been proof of concept in nature, targeting a small number of genes, and using a single round of screening with a single biosensor. The strategies are clever and promising, but not yet ready for industrial use. Recent work has implemented biosensors beyond simple screens and selections: biosensors have been used to tune gene expression of biosynthetic genes in response to product concentration [46], to modulate mutagenesis rates in proportion to metabolite production [47], and for real-time observation of chemical production [48**].

To mature as a field, biosensor-directed metabolic engineering requires further characterization of biosensor systems, coupled with a push toward the discovery and creation of additional sensor domains that respond to industrially important compounds. Full biosynthetic pathways must be targeted [5**,49], and multiple sensors may be required, with graduated operational concentration ranges to avoid saturating the biosensor at high molecule titers. To enable very large libraries, of 10⁹ members and above, new interventions to improve robustness to false positives will be required, which may benefit from standardized screening chassis [20**]. Finally, proof-of-concepts developed in academic labs must be transferred to industrial partners for final optimization. Following these recommendations, metabolic engineering will benefit from a powerful application of evolutionary strategies that are ideal to solve this difficult class of biological problems.

**Acknowledgements**
This work is supported by the National Science Foundation (NSF GRFP to J.K.R.) and the US Department of Energy (DEFG02-02ER63445 to G.M.C.).

**References and recommended reading**
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- **of outstanding interest

2. Carlson R: The U.S. Bioeconomy in 2012 reached $350 billion in revenues, or about 2.5% of GDP. 2014.
5. Raman S et al.: Evolution-guided optimization of biosynthetic pathways. Proc Natl Acad Sci U S A 2014, 111:17800-17808. This study develops design rules for pathway optimization using biosensor-based selections. Biosensors for many industrially important compounds were characterized while multiplexed targeted genome engineering was demonstrated for production optimization of glucarate and naringenin.

45. Rogers JK, Church GM: Multiplexed engineering in biology. \textit{Trends Biotechnol} 2016, 34:198-206. A perspective piece that explores how multiplexed engineering in biology will shape how bioengineers approach metabolic engineering in the future. This article discusses the design-build-test cycle of bioengineering and offers insight into how the latest advances in genome engineering can best be implemented.


This study links fluorescent biosensors to several diverse metabolic pathways and shows that metabolite production can be observed as it occurs. The authors demonstrate that fluorescent readouts are a good proxy for chemical production and go on to optimize the production of 3-hydroxypropionate and acrylicate.


