Building cellular pathways and programs enabled by the genetic diversity of allo- and metagenomes

Kyle A. Zingaro and Eleftherios Terry Papoutsakis *

Molecular Biotechnology Laboratory, Dept. of Chemical & Biomolecular Engineering, the
Delaware Biotechnology Institute, University of Delaware, Newark, DE 19711, USA

* Corresponding author: Molecular Biotechnology Laboratory, Dept. of Chemical &
Biomolecular Engineering, and the Delaware Biotechnology Institute, University of Delaware,
15 Innovation Way, Newark, DE 19711. Phone: (302) 831-8376; Fax: (302) 831-4841; E-mail:

epaps@udel.edu

(ABSTRACT)

Engineering pathways, programs and traits that require interactions of many, often unknown, genes requires advanced engineering strategies in the context of synthetic biology. Such strategies derive from three basic requirements: a suitably enlarged gene pool compared to the parent strain; a method of identification and incorporation of genetic-element interactions to generate the multigenic pathway or trait; and a process of selection of individuals from diverse strain populations that benefit the desirable pathway or trait. We review potential methods utilized in such advanced engineering strategies, emphasizing methods that explore the genomic diversity of allogeneic DNA (the allogenome) or the metagenome. We also propose a modular iterative approach for developing multigenic cellular traits.

Introduction: Rational versus selection-based cell engineering

Cellular phenotypes (or traits) are inherently complex, stemming from the orchestrated actions of many, frequently unknown, genes often located distally on a chromosome [1,2]. As such, strategies to engineer complex multigenic traits must engage approaches, which are both able to identify and apply multiple genes and systems in a single cell. Two broad categories for engineering complex cellular traits can be logically identified: rational design and selection-based genome engineering. Both constitute different facets of synthetic biology.

Rational design (Case A; Fig. 1) is driven by prior knowledge in seeking to develop a multigenic trait through engagement of new genes or pathways or structural modifications to individual enzymes. Addition or removal of genes or pathways, or changes to protein or nucleic acid sequences, is made based on established knowledge or modeling indicating a strong likelihood of developing or improving a desirable trait. In the case of pathway engineering, this method allows for specific generation of desired products with reduction in, or elimination of, undesired side-products. Rational design approaches have proven very successful but may be inherently limited by the state of understanding of biological systems. If we do not know the specific mutation, gene, or combination of genes to produce the desired result, we have no means of generating a new trait using rational design. Significantly, most such approaches are based on well-annotated genomes and hosts, thus restricting the effort to a limited genomic space [1-4].

Although rational design offers a means of making significant modifications to a cellular pathway or program, the genes involved in complex cellular traits are often unknown or

incompletely known making rational development of such phenotypes difficult and some times impossible. It is therefore logical to seek to utilize selection-based strategies to identify genes capable of imparting improvements to an engineered trait or in generating an altogether novel trait. Selection-based cell engineering (Case B; Fig. 1) employs selective pressure or assay-based detection of a cellular trait to identify genetic elements that enable the desired trait. The success of this approach depends, first, on the diversity of the genomic space to be screened, with a more diverse space being more likely to contain the genes or DNA determinants that enable the desirable trait. For multigenic traits, it also requires methods, which are able to identify beneficial interactions between genetic determinants for developing the target trait. Finally, a method of enrichment for the desired trait is required to select beneficial genetic determinants and then combinations of genetic determinants. We review the techniques and technologies applied for each of these aspects of cell engineering. As a capstone, we propose a modular, iterative approach towards future developments of desirable complex cellular traits.

Generating and using genetic diversity within a genome

There exists a variety of ways to generate genetic diversity in an organism in order to identify beneficial mutations (Fig. 2). Development of complex cellular traits will likely require combinations of these approaches to identify and optimize complex synthetic or semi-synthetic pathways and programs.

Accumulation of mutations in the host organism

Single mutations. By simply applying a selective pressure to an organism, it is possible, through accumulation of random mutations, to generate changes to an organism that benefit a phenotype

of interest [5,6]. This process, frequently referred to as strain evolution, is time consuming due to the low mutation rate (0.0037 average mutations per base pair per replication in *E. coli* [7]). For example, in generating tolerance to toxic chemicals, a widely examined complex trait, this process typically involves growing an organism in continuous culture or in a batch culture where the biomass is recycled to allow exposure of the cells to the toxic chemical for the long time periods necessary for the accumulation of mutations. This process can be accelerated by employing one or more of a variety of methods for increasing the mutation rate. These include irradiation (e.g, ultraviolet light exposure) and chemical mutagens. The mutation rate can also be significantly enhanced by removing or disabling DNA repair mechanisms within the cell [8].

Whole genome shuffling (WGS) (Fig. 2) is a method of recursive genomic recombination that is typically employed in combination with a mutagenesis technique[9,10]. Following mutagenesis of an organism, the diversified cell population is subjected to binary protoplast fusions, allowing for homologous recombination to generate genomic material containing combinations of the mutations that have been generated from the mutagenesis process. Through repeated protoplast fusions, faster accumulation of genetic mutations can be achieved thus making possible to screen a more diverse and presumably superior pool of clones. WGS is typically limited to Gram⁺ prokaryotes due to the requirement for efficient protoplast generation. However, WGS has also been demonstrated in the Gram⁻ *E. coli*, although the efficiency was lower than in Gram⁺ organisms [11,12]. Additionally, WGS has been used for engineering yeast strains for improved tolerance and fermentation productivities [9,13,14]. WGS is still limited to mutations accumulated in the original parent strain and it serves to more rapidly combine those mutations.

Thus, WGS does not expand the genomic capabilities of the organism to include novel biocatalytic or biophysical functions.

Hybrid rational/selection-based pathway and program engineering

With some knowledge of the genes or pathways involved in a phenotype, it is possible to combine rational gene targets with selection-based genome engineering techniques to enhance or optimize a desirable complex trait. These hybrid approaches use genome-engineering techniques targeted to specific genes or regulons to affect global cellular processes. Such strategies allow for optimization of metabolic pathways through simultaneous modification of gene dosages for multiple targets or global transcriptomic changes to generate complex traits.

Randomized mutant libraries. Libraries are collections of DNA fragments stored and propagated in a chosen host organism. Randomized mutant libraries, as opposed to genomic and cDNA libraries, are synthesized *de novo* and are generally used for examination of mutations to a single gene. These libraries are typically used for examining the effects of mutations on protein activity, protein-protein interactions, or protein stability. An application of the randomized, mutant library for genome-wide engineering is global transcription machinery engineering (gTME)[15]. gTME uses error prone PCR amplified global transcriptional regulators as the source of diversity in the library. These mutated regulators then cause significant diversity among the transcriptomes of the library population.

Recombineering technologies. Several technologies have been established for generating genetic diversity in a population using *in vivo* recombination-based genetic engineering

(Recombineering) [16,17]. These powerful technologies allow for rapid changes to a genome but are largely limited to sequenced and annotated genomes. Employing homologous recombination between linear synthetic oligos targeted to specific locations on an organism's chromosome, desirable modifications to specific genetic sequences can be made *in vivo*.

Multiplex automated genetic engineering (MAGE; Fig. 2) is an application of recombineering that allows for rapid, large scale cellular reprogramming [18]. The MAGE system employs an automated process using multiple cycles of recombination with a pool of linear oligos. After several cycles, the resulting strains represent a diverse pool of genotypes, which can be screened for a desirable trait. The MAGE technology has been applied to *E. coli* using oligos designed to modify the ribosome binding sites of a several genes simultaneously. In one example, a population of clones having variable expression levels for the genes of a synthetic pathway for lycopene biosynthesis was generated [18]. The MAGE technology was then extended to manipulate an entire genome using a so-called hierarchical Conjugative Assembly Genome Engineering (CAGE) [19]. Utilizing specific positive and negative selection markers in bacterial conjugation between a donor and recipient strain, modifications developed in one strain can be transferred and incorporated into another. Using this approach hierarchically, one can then make edits to a limited region of a chromosome and then combine those changes into a single clone using a series of assembly steps.

Screening genomic libraries for more than single genes: small regulatory RNAs, natural and synthetic

Genomic libraries (Fig. 2) allow for enrichment of beneficial determinants typically from the genomic space of a single organism [2,20-24]. A typical DNA insert will be 4-5 kb, and typically one gene or small operon is identified as having an impact by overexpression. Interestingly, though homologous genomic libraries based on fragmented DNA offer the possibility to search for genes whose downregulation benefits the screenable phenotype. This is accomplished via antisense transcripts (in *cis*), naturally coded or artificially generated by expression in the opposite orientation from promoters on the library vector driving expression from the inserted DNA fragment. A case in point is the identification of non-coding RNAs, antisense to the promoter regions of the rRNA loci of *Clostridium acetobutylicum*, whose expression imparts tolerance to butyric acid stress [20]. Similar approaches can be employed to search for *cis* or *trans* noncoding RNAs downregulating detrimental gene expression or promoting regulation beneficial to the desirable phenotype. A case of the latter is the identification, via screening of genomic libraries, of the small non-coding RNA (sRNA) ArcZ in *E. coli* that promotes tolerance to acid stress by interacting with the 5' UTR of the *rpoS* transcript thus benefiting the expression of the stress sigma factor RpoS [25]. There are very few other options in searching for such small RNA molecules that are involved in the development of a trait.

Engagement of more diverse gene pools: heterologous genomic spaces (the allo-genome) and the metagenome

As discussed, a variety of techniques have been employed to diversify the genome of a single organism, however these methods are all limited to the genomic space of the host organism and therefore can only explore an organism's intrinsic diversity [3]. By introducing genetic material from an outside source, be that a single heterologous genome, a consortium of organisms, or the

metagenome, the diversity of the screenable genomic space can be greatly expanded to explore a much larger and more diverse source of enzymatic and biophysical traits [26,27].

Allogeneic or metagenomic libraries make possible to search a much broader genomic space for novel genetic material in a host strain. However, there is reduced expression of genes from heterologous sources as the host organism's sigma factors have a limited capacity to recognize heterologous promoters. This effect can be lessened for libraries with small genomic inserts by using external promoters that can be recognized by the screening host. For large-insert libraries, which maximize the genomic space that can be explored, there will be minimal expression from any internal genes on the insert thus limiting the effectiveness of the library. To address this problem, we have recently demonstrated [28] that through expression of heterologous sigma factors, one can increase the fraction of genes expressed in a host from a heterologous library, including genes inaccessible because of internal transcription terminators (Fig. 3). Expression of the heterologous sigma factor RpoD from *Lactobacillus plantarum* was shown to allow expression of about 90% of all *L. plantarum* promoters in *E. coli*, a 3.5-fold increase from an *E. coli* host without the *L. plantarum* RpoD expression as well as increased expression from other genomic libraries including libraries from *C. acetobutylicum*, *C. pasteurianum*, *Deinococcus radiodurans*, *Bacillus subtilis*, and from soil metagenomic DNA [28]. Such strategies provide access to a larger genomic space which can then be assayed for desirable traits, although there has been relatively limited work towards the application of heterologous and particularly metagenomic libraries in the development of complex cellular traits, all for now focused on developing tolerance to toxic chemicals [28-30]. Also, until recently [28], metagenomic studies have focused almost entirely on identification of novel enzymatic activities [3,31,32].

Capturing the combinatorial complexity

Due to the required interactions among several genetic determinants to generate complex traits, genome-engineering efforts which target only single genes or a few genes in a host's genomic space are unlikely to yield new or dramatically enhanced complex phenotypes. Although sequential construction methods provide the ability to combine beneficial targets identified through rational design or genome engineering, they cannot capture interactions between distant genomic loci that may be beneficial only when both are expressed without *a priori* knowledge of at least one of the interacting members. This limitation can be overcome using the coexisting/coexpressing genomic library technology (CoGeL) [2]. The CoGeL technology relies on expression of two genomic libraries in a single cell (Fig. 4). This process allows for screening of the beneficial impact of co-expressing two genetic determinants providing access to potential combinatorial interactions. Using the CoGeL system, it becomes possible to assay for improved phenotypes granted only through interactions of two separate, and in the case of genomic libraries distant, genetic determinants. Combining this technology with heterologous and metagenomic libraries introduces a diverse space of genomic targets that can be assayed in pairs for identifying interacting genetic determinants that enhance cellular traits.

Enriching and screening based on growth and cell viability

The method of selection plays a critical role in the identification of improved or novel cellular traits. One cellular trait of both industrial and scientific importance that has been the target of many genome engineering attempts is microbial tolerance to toxic chemicals [1,4,20,24,33]. Traditionally, an engineered population is selected based on its growth or survival in the

presence of a toxic chemical or in the presence of a specific and unusual nutrient or metabolizable chemical. However, the specific methods engaged for enrichment of individual strains from a genetically diverse population for engineering chemical tolerance have been extremely varied. A summary of some selection methods used previously is provided in Table 1.

As with the development of assays to measure tolerance of microbes to toxic chemicals, the effects of chemical concentration, exposure time, temperature, and medium used in the selection can be significant. The nature of these effects has been recently reviewed in great detail [4]. Briefly, those assay parameters should be directly related to the desired use of the engineered strain. Of particular note are the effects of toxic chemical concentration and exposure time on enriched mutations or genetic determinants and the resulting selected traits. Low concentrations and short exposures may select for improved cell growth while long exposures and high concentrations enrich for increased cell survival [4,6,25].

Often, in attempting to engineer a complex trait, the selection imparted by a single exposure to selective pressure is not enough to reduce the population to the desired subset of improved clones. By repeating the exposure of the strain to the toxic chemical, it is possible to apply a more stringent selection without exceeding tolerance limits of the strain. These exposure cycles also allow for increased stringency in the selection process. The transfer process also allows the cells to grown in fresh media so that depletion of essential nutrients does not become an additional selection parameter.

The repeated exposure however also places significant burden on the cells. Repeated ethanol exposure has been shown to negatively affect the tolerance response in *E. coli* during subsequent exposures [34]. Enrichments of *E. coli* cells containing heterologous *L. plantarum* libraries for improved ethanol tolerance have also shown decreased growth and survival across multiple ethanol exposures [25].

Tracking the screening process

In addition to basic screening approaches, several technologies have recently emerged which allow for more detailed examinations of cellular populations. These technologies typically employ a molecular barcode system [35], a unique nucleotide sequence capable of identifying individuals in a population using a microarray analysis. Scalar analysis of library enrichments (SCALEs) is a method for assessment of individual cellular fitness during enrichments of a genomic libraries [23]. Through application of the barcode technology to each plasmid in a library, individual genetic determinants can be tracked and used to assess the fitness of strains bearing specific genetic determinants. Trackable multiplex recombineering (TRMR) is a method of generating and evaluating large numbers of genomic modifications simultaneously [36]. By labeling the oligos used to recombine with genomic targets in multiplex genome engineering, modifications to the chromosome are trackable during enrichment of the diversified population. Individual modifications to the genome can then be assessed for their effect on cellular fitness.

Beyond growth and survival: Selecting and screening for complex cellular traits

When one is selecting for genetic determinants that impart improved growth or survival under a set culture conditions, such as in the presence of specific nutrients or under stress from or more

toxic chemicals, screening for and enriching cell populations is relatively straightforward . Most other desirable traits are harder to screen and select for. A typical trait is the ability of the engineered strain to produce a metabolite (within the cell or secreted) at higher rate and/or titer. Due to limitations in detecting metabolites, it is often difficult to develop or implement assays to select for improved productivity [4,37]. Here we discuss some of the methods used for enriching for improved metabolite production.

In vitro colorimetric or fluorescence based assays

Detection of a colored chemical or fluorophore by visual inspection, microscopy, or spectrophotometry provides an easy route to identify improved production. These methods however require a fluorophore or detectable pigment as the desired product or a downstream product of a coupled assay. Easily detectable, natural molecules are an ideal product choice and increased productivity in such systems can be easily assayed. The number of naturally detectable products is however relatively small. Therefore many chemical and enzymatic assays have been developed to generate detectable fluorescing chemicals from desirable products. Fluorescence derivatization of valuable products including carboxylic acids, amino acids, and alcohols [38-42] previously developed for improving detection in high performance liquid chromatography can be used for *in vitro* screening in the context of strain engineering. Enzymatic conversion of intermediates or end-products to detectable molecules has been used for improving production of thiols [43] and degradation of polycyclic aromatic hydrocarbons [44]. Detection of alcohols produced by *C. acetobutylicum* has also been accomplished using the nitro blue tetrazolium (NBT) assay for detection of alcohol dehydrogenase activity. This was accomplished by oxidizing alcohol products into aldehydes using a commercial alcohol dehydrogenase which

generates NADH which then reduces nitroblue tetrazolium chloride into the detectable formazan [45].

In vivo assays

By integrating a synthetic detection system into the host organism or in a co-cultured strain, it is possible to significantly increase the throughput of the screening method. *In vivo* assays can be broadly classified into two categories. The first uses enzymatic assays, which produce a fluorescent or detectable small molecule from a desirable product or an intermediate of the production pathway. The second relies on biosensors, which detect a small molecule by using an enzyme, which then controls the production of a detectable product such as green fluorescent protein or an easily assayed enzyme such as β -galactosidase. Biosensors, or microbial-based molecular reporters, rely on naturally occurring or engineered protein or RNA molecules to detect a specific ligand [37,46]. These molecular reporters generate a detectable signal, which provides a means for measuring the ligand concentration *in vivo*. Applications of these biosensors include improved production of mevalonate via a mutated AraC promoter modified to be activated by the presence of mevalonate [47], identifying new enzymes capable of producing biologically active small molecules from metagenomic libraries [48], optimizing fatty acid production via detection of fatty acid intermediates [49], and identifying novel amidases from metagenomic libraries using a transcriptional regulator activated by benzoate [50]. Growth dependency using a TetA-GFP fusion protein under the regulation of an alcohol-responsive transcriptional regulator from *Thauera butanivorans* sp. nov. was also utilized to improve 1-butanol production in *E. coli* [51].

Single cell screening technologies

Single cell screening technologies such as Fluorescent Assisted Cell Sorting (FACS) offers a means to dramatically increase the throughput of these screening methods. In addition to detectable fluorescent products generated either as detectable small molecules or as part of biosensor platforms, FACS is also capable of detecting cells with the desirable tolerance and production capabilities [32,46]. FACS-compatible dyes capable of labeling nucleic acids, determining membrane potential and fluidity, measuring intracellular pH, and identifying toxic intracellular chemicals including oxygen species and radicals have been developed [32,52-54]. FACS provides a high-throughput approach to assay both features of tolerance and production in the generation of new strains for microbial production of chemicals, and can be employed in all aspects of strain engineering.

Using auxotrophic strains: single and multiple deletions to identify novel genes and develop multigenic traits

The rate of product formation can be linked to growth of a strain through auxotrophy for the desired product or an intermediate of its production. Auxotrophic strains have been used in engineering improved productivity of amino acids [55-57]. Strains have also been engineered to increased utilization of non-natively consumed metabolites using auxotrophic strains [58,59]. Finally, productivities of alcohols including 1-propanol and 1-butanol have been improved in *E. coli* by screening in a strain auxotrophic for 2-ketobutyrate, a metabolic precursor to both alcohols [60]. In virtually all cases of genetically well-defined auxotrophs, the auxotrophy derives from a single gene deletion or inactivation. Such well defined auxotrophic mutants can be used to identify superior gene homologues or orthologs from other genomes or the

metagenome using the sigma-factor technology [28] discussed above, which makes possible to express many or most of the genes from a allogenome or the metagenome, thus expanding the genomic space that can be searched. This concept can be generalized to capture beneficial genomic-element interactions using the CoGeL technology [2] and auxotrophs with multiple mutations. This is discussed next.

Through exploration of the metagenome and application of technologies capable of simultaneously examining multiple genetic determinants it may be possible to engineer novel biochemical pathways. One can envision the use of the CoGeL technology to produce a multistep enzymatic pathway. Production of desired intermediates or final chemicals could then be screened using biosensor based FACS or through development of auxotrophic mutants. The basis for this approach has been examined previously through the development of a *lysA* and *dapA* knockout *E. coli* strains which was auxotrophic for lysine [2]. Using the CoGeL technology Nicolaou, et. al. were able successfully restore growth in media lacking lysine to that double knockout strain through identification of distally located genes on co-expressed plasmid and fosmid libraries. The application of this and other multiple genetic determinant screening technologies may allow for the rapid development and improvement of multigenic traits.

Stepwise construction of multigenic cellular pathways and traits

Technologies such as MAGE/CAGE and WGS offer the ability to rapidly accumulate changes to a host genomic space but do not expand the genomic space beyond the host's genome [18,19]. Ideally, for developing complex cellular traits the effects of genetic determinants from

metagenomic sources, including interactions from combinations of determinants, must be engaged.

A logical approach towards development of complex traits is a step-wise or sequential, semi-synthetic approach which first identifies engineering targets, either through rational design or genome engineering, and combines those beneficial changes to identify combinations that generate desired improvements. These combinations of targets are then amenable to modifications to gene dosage within the system using the hybrid rational/selection-based approaches described above, including systems such as MAGE [18] and TIGR [61], and more recently tools like the ePathbrick [62,63], which enable optimization of expresses of large sets of genes in a pathway of program to achieve optimal or improved outcomes. Sequential construction of complex cellular traits can then be further enhanced by introducing additional modifications using processes such as mutagenesis with WGS where possible, further screening with genomic or cDNA libraries including metagenomic libraries, and further exploration of rational targets. These step-wise methods have been employed to develop alcohol tolerance in *E. coli* using rationally identified overexpression targets [64] in combination with heterologous libraries from more tolerant organisms [30].

Based on these technologies, both alone and in combination with the recombineering approaches and WGS, one can thus envision host organisms expressing multiple inserts from metagenomic sources providing the capability to screen for interactions of several distant genetic loci simultaneously. This process in combination with a step-wise engineering approach makes possible both the identification of individual members and the construction of complex semi-

synthetic systems for generating complex traits. Using the diversity present in metagenomic libraries, these systems could capture the most beneficial interactions from a broad range of organisms generating hybrid responses to improve desired traits. The resulting strains could then be combined utilizing WGS to generate organisms and an even greater genetic diversity from which to screen for improved traits (Fig. 5).

Conclusions

In the development of complex microbial phenotypes, limitations to our understanding of biological processes require us to employ genome engineering approaches to identify new genetic determinants as targets for engineering. For advancing genome engineering technologies, we must consider both the methods for generating diversity and how to identify individual improvements from the diversified populations. The variety of methods for generation of diversity provides a powerful toolset allowing the user to target sequences from complex metagenomes down to changes to individual amino acids within a protein. This wealth of diversity then requires reliable methods for isolating new genes and mutations, which enhance desirable phenotypes. These selection methods are often the limiting factor in the development of strains capable of new and more efficient chemical conversions. Next generation technologies employing advanced biosensors in combination with high throughput screening tools such as FACS will provide new opportunities for engineering microbes. These technologies form the foundation for developing novel processes for producing chemicals, for degradation of toxins, and for improved productivities of biological hosts.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation Grant CBET-1033926.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest

** of outstanding interest

1. Nicolaou SA, Gaida SM, Papoutsakis ET: **A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation.** *Metabolic Engineering* 2010, **12**:307-331.

2. Nicolaou SA, Gaida SM, Papoutsakis ET: **Coexisting/Coexpressing Genomic Libraries (CoGeL) identify interactions among distantly located genetic loci for developing complex microbial phenotypes. *Nucleic Acids Research* 2011, **39**:e152
110.1093/nar/gkr1817.

The method that enables to screen two and possibly three genomic or metagenomic libraries aiming to capture genetic-loci interactions that enable desirable cellular traits.

3. Nettleton JA, Hivert MF, Lemaitre RN, McKeown NM, Mozaffarian D, Tanaka T, Wojczynski MK, Hruby A, Djousse L, Ngwa JS, et al.: **Meta-analysis investigating associations between healthy diet and fasting glucose and insulin levels and modification by loci associated with glucose homeostasis in data from 15 cohorts.** *Am J Epidemiol* 2013, **177**:103-115.

4. Zingaro KA, Nicolaou SA, Papoutsakis ET: **Dissecting the assays to assess microbial tolerance to toxic chemicals in bioprocessing.** *Trends Biotechnol* 2013, **31**:643-653.

5. Fernandez-Sandoval MT, Huerta-Beristain G, Trujillo-Martinez B, Bustos P, Gonzalez V, Bolivar F, Gosset G, Martinez A: **Laboratory metabolic evolution improves acetate tolerance and growth on acetate of ethanologenic *Escherichia coli* under non-aerated conditions in glucose-mineral medium.** *Applied Microbiology and Biotechnology* 2012, **96**:1291-1300.

6. Minty JJ, Lesnefsky AA, Lin FM, Chen Y, Zaroff TA, Veloso AB, Xie B, McConnell CA, Ward RJ, Schwartz DR, et al.: **Evolution combined with genomic study elucidates genetic bases of isobutanol tolerance in *Escherichia coli*.** *Microbial Cell Factories* 2011, **10**:article 18.

7. Drake JW: **Avoiding Dangerous Missense: Thermophiles Display Especially Low Mutation Rates.** *Plos Genetics* 2009, **5**:e1000520.

*8. Fabret C, Poncet S, Danielsen S, Borchert TV, Ehrlich SD, Janniere L: **Efficient gene targeted random mutagenesis in genetically stable *Escherichia coli* strains.** *Nucleic Acids Research* 2000, **28**.

Describes an *in vivo* random mutagenesis method for generating localized mutant variants in a host chromosome several orders of magnitude higher than in previously used mutagenesis methods.

9. Hou LH: **Improved Production of Ethanol by Novel Genome Shuffling in *Saccharomyces cerevisiae***. *Applied Biochemistry and Biotechnology* 2010, **160**:1084-1093.
10. Zhang YX, Perry K, Vinci VA, Powell K, Stemmer WPC, del Cardayre SB: **Genome shuffling leads to rapid phenotypic improvement in bacteria**. *Nature* 2002, **415**:644-646.
11. Reyes LH, Almario MP, Winkler J, Orozco MM, Kao KC: **Visualizing evolution in real time to determine the molecular mechanisms of n-butanol tolerance in *Escherichia coli***. *Metabolic Engineering* 2012, **14**:579-590.
- *12. Dai MH, Ziesman S, Ratcliffe T, Gill RT, Copley SD: **Visualization of protoplast fusion and quantitation of recombination in fused protoplasts of auxotrophic strains of *Escherichia coli***. *Metabolic Engineering* 2005, **7**:45-52.

Demonstration of successful protoplast fusion method for *E. coli*, a Gram-negative bacterium producing 3 – 4 orders of magnitude increase in fused protoplasts.

13. Ge JP, Sun HB, Song G, Ling HZ, Ping WX: **A genome shuffling-generated *Saccharomyces cerevisiae* isolate that ferments xylose and glucose to produce high levels of ethanol**. *Journal of Industrial Microbiology & Biotechnology* 2012, **39**:777-787.
14. Shi DJ, Wang CL, Wang KM: **Genome shuffling to improve thermotolerance, ethanol tolerance and ethanol productivity of *Saccharomyces cerevisiae***. *Journal of Industrial Microbiology & Biotechnology* 2009, **36**:139-147.
15. Alper H, Moxley J, Nevoigt E, Fink GR, Stephanopoulos G: **Engineering yeast transcription machinery for improved ethanol tolerance and production**. *Science* 2006, **314**:1565-1568.
16. Zhang YM, Buchholz F, Muyrers JPP, Stewart AF: **A new logic for DNA engineering using recombination in *Escherichia coli***. *Nature Genetics* 1998, **20**:123-128.
17. Sharan SK, Thomason LC, Kuznetsov SG, Court DL: **Recombineering: a homologous recombination-based method of genetic engineering**. *Nature Protocols* 2009, **4**:206-223.
- *18. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM: **Programming cells by multiplex genome engineering and accelerated evolution**. *Nature* 2009, **460**:894-898.

Method for multiplex automated genome engineering (MAGE) used to simultaneously modify multiple targets across loci in a single genome or population. This method was applied to increase lycopene production by 5-fold in *E. coli*.

19. Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, Tolonen AC, Gianoulis TA, Goodman DB, Reppas NB, et al.: **Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement**. *Science* 2011, **333**:348-353.
20. Borden JR, Jones SW, Indurthi D, Chen YL, Papoutsakis ET: **A genomic-library based discovery of a novel, possibly synthetic, acid-tolerance mechanism in *Clostridium***

acetobutylicum involving non-coding RNAs and ribosomal RNA processing.
Metabolic Engineering 2010, **12**:268-281.

21. Nicolaou SA, Gaida SM, Papoutsakis ET: **Exploring the combinatorial genomic space in *Escherichia coli* for ethanol tolerance.** *Biotechnology Journal* 2012, **7**:1337-1345.

*22. Reyes LH, Almario MP, Kao KC: **Genomic Library Screens for Genes Involved in n-Butanol Tolerance in *Escherichia coli*.** *Plos One* 2011, **6**.

Identified 14 genes that decrease the growth inhibition of n-butanol through screening of an *E. coli* genomic library. Genes identified indicate that iron transport, metabolism, and decreased acid resistance enhance n-butanol tolerance.

23. Sandoval NR, Mills TY, Zhang M, Gill RT: **Elucidating acetate tolerance in *E. coli* using a genome-wide approach.** *Metabolic Engineering* 2011, **13**:214-224.

24. Borden JR, Papoutsakis ET: **Dynamics of genomic-library enrichment and identification of solvent tolerance genes for *Clostridium acetobutylicum*.** *Applied and Environmental Microbiology* 2007, **73**:3061-3068.

25. Gaida SM, Al-Hinai MA, Indurthi D, Nicolaou SA, Papoutsakis ET: **Synthetic tolerance: three non-coding small RNAs, DsrA, ArcZ and RprA, acting supra-additively against acid stress.** *Nucl. Acids Res.* 2013, **41**: 8726-8737.

26. Sommer MOA, Church GM, Dantas G: **A functional metagenomic approach for expanding the synthetic biology toolbox for biomass conversion.** *Molecular Systems Biology* 2010, **6**:article 360.

27. Papoutsakis ET: **Exploring the Capabilities of the Geobiosphere's Microbial Genome.** *Aiche Journal* 2013, **59**:688-698.

28. Gaida SM, Sandoval NR, Nicolaou SA, Y. C, Venkataramanan KP, Papoutsakis ET: **Expression of heterologous sigma factors enables functional screening of metagenomic and heterologous genomic libraries. *Nature Communications* 2015, **6**:7045 doi: 7010.1038/ncomms8045.

Method and technology to solve an old-standing problem, namely the ability to functionally screen metagenomic and heterologous genomic libraries in *E. coli*. Can be likely extended to other hosts.

29. Yuan YB, Bi CH, Nicolaou SA, Zingaro KA, Ralston M, Papoutsakis ET: **Overexpression of the *Lactobacillus plantarum* peptidoglycan biosynthesis murA2 gene increases the tolerance of *Escherichia coli* to alcohols and enhances ethanol production.** *Applied Microbiology and Biotechnology* 2014, **98**:8399-8411.

*30. Zingaro KA, Nicolaou SA, Yuan YB, Papoutsakis ET: **Exploring the Heterologous Genomic Space for Building, Stepwise, Complex, Multicomponent Tolerance to Toxic Chemicals.** *Acs Synthetic Biology* 2014, **3**:476-486.

Demonstration of method for building multigenic traits based on a combination of rational and library-screening that makes possible to identify and engage heterologous genetic elements.

*31. Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, et al.: **Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms.** *Applied and Environmental Microbiology* 2000, **66**:2541-2547.

Description of the production of a bacterial artificial chromosome (BAC) library from soil metagenomic DNA. Library screening in *E. coli* indicated successful expression of multiple enzymatic activities and phylogenetic analysis indicated the presence of DNA from multiple microbial phyla.

32. Tracy BP, Gaida SM, Papoutsakis ET: **Flow cytometry for bacteria: enabling metabolic engineering, synthetic biology and the elucidation of complex phenotypes.** *Current Opinion in Biotechnology* 2010, **21**:85-99.
33. Peabody V GL, Winkler J, Kao KC: **Tools for developing tolerance to toxic chemicals in microbial systems and perspectives on moving the field forward and into the industrial setting.** *Current Opinion in Chemical Engineering* 2014, **6**:9-17.
34. Zingaro KA, Papoutsakis ET: **GroESL overexpression imparts *Escherichia coli* tolerance to i-, n-, and 2-butanol, 1,2,4-butanetriol and ethanol with complex and unpredictable patterns.** *Metabolic Engineering* 2013, **15**:196-205.
35. Shoemaker DD, Lashkari DA, Morris D, Mittmann M, Davis RW: **Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy.** *Nature Genetics* 1996, **14**:450-456.
- **36. Warner JR, Reeder PJ, Karimpour-Fard A, Woodruff LBA, Gill RT: **Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides.** *Nature Biotechnology* 2010, **28**:856-U138.

Method description of trackable multiplex recombineering (TRMR). Using synthetic DNA cassettes and molecular barcodes thousands of genetic modifications can be made to a host genome and tracked simultaneously allowing for quantification of population dynamics.

37. Dietrich JA, McKee AE, Keasling JD: **High-Throughput Metabolic Engineering: Advances in Small-Molecule Screening and Selection.** *Annual Review of Biochemistry*, Vol 79 2010, **79**:563-590.
38. Mukherjee PS, Karnes HT: **Ultraviolet and fluorescence derivatization reagents for carboxylic acids suitable for high performance liquid chromatography: A review.** *Biomedical Chromatography* 1996, **10**:193-204.
39. Takadate A, Irikura M, Suehiro T, Fujino H, Goya S: **New Labeling Reagents for Alcohols in Fluorescence High-Performance Liquid-Chromatography.** *Chemical & Pharmaceutical Bulletin* 1985, **33**:1164-1169.
40. Toyo'oka T: **Fluorescent tagging of physiologically important carboxylic acids, including fatty acids, for their detection in liquid chromatography.** *Analytica Chimica Acta* 2002, **465**:111-130.
41. Yu H, Tyo K, Alper H, Klein-Marcuschamer D, Stephanopoulos G: **A High-Throughput Screen for Hyaluronic Acid Accumulation in Recombinant *Escherichia coli* Transformed by Libraries of Engineered Sigma Factors.** *Biotechnology and Bioengineering* 2008, **101**:788-796.
42. Lindroth P, Mopper K: **High-Performance Liquid-Chromatographic Determination of Subpicomole Amounts of Amino-Acids by Precolumn Fluorescence Derivatization with Ortho-Phthaldialdehyde.** *Analytical Chemistry* 1979, **51**:1667-1674.
43. Wang CH, Leffler S, Thompson DH, Hrycyna CA: **A general fluorescence-based coupled assay for S-adenosylmethionine-dependent methyltransferases.** *Biochemical and Biophysical Research Communications* 2005, **331**:351-356.

44. Alcalde M, Bulter T, Arnold FH: **Colorimetric assays for biodegradation of polycyclic aromatic hydrocarbons by fungal laccases.** *Journal of Biomolecular Screening* 2002, **7**:547-553.
45. Scheel M, Lutke-Eversloh T: **New options to engineer biofuel microbes: Development and application of a high-throughput screening system.** *Metabolic Engineering* 2013, **17**:51-58.
46. Gredell JA, Frei CS, Cirino PC: **Protein and RNA engineering to customize microbial molecular reporting.** *Biotechnology Journal* 2012, **7**:477-499.
- **47. Tang SY, Cirino PC: **Design and Application of a Mevalonate-Responsive Regulatory Protein.** *Angewandte Chemie-International Edition* 2011, **50**:1084-1086.

Method description for engineering the *E. coli* AraC protein to respond to mevalonate to act as a biosensor for improvement of mevalonate production. Used to screen mutants to the mevalonate-synthesis operon to identify improved mevalonate titer.

48. Williamson LL, Borlee BR, Schloss PD, Guan CH, Allen HK, Handelsman J: **Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor.** *Applied and Environmental Microbiology* 2005, **71**:6335-6344.
- **49. Zhang FZ, Carothers JM, Keasling JD: **Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids.** *Nature Biotechnology* 2012, **30**:354-U166.

This article describes creation of a biosensor for a key intermediate of biodiesel production in *E. coli*. This dynamic sensor-regulator system (DSRS) is applied to improve biodiesel yield by 28%.

- **50. Uchiyama T, Miyazaki K: **Product-Induced Gene Expression, a Product-Responsive Reporter Assay Used To Screen Metagenomic Libraries for Enzyme-Encoding Genes.** *Applied and Environmental Microbiology* 2010, **76**:7029-7035.

Application of a benzoate-responsive transcriptional activator based biosensor to improve benzoate production using a metagenomic library in *E. coli*.

- *51. Dietrich JA, Shis DL, Alikhani A, Keasling JD: **Transcription Factor-Based Screens and Synthetic Selections for Microbial Small-Molecule Biosynthesis.** *Acs Synthetic Biology* 2013, **2**:47-58.

Description of a screening method using a tetracycline resistance gene expressed under regulation of small molecule inducible promoters. This method is applied to increase 1-butanol production by 35% and a 120-fold enrichment for 1-butanol production.

52. Cronin UP, Wilkinson MG: **Monitoring growth phase-related changes in phosphatidylcholine-specific phospholipase C production, adhesion properties and physiology of *Bacillus cereus* vegetative cells.** *Journal of Industrial Microbiology & Biotechnology* 2008, **35**:1695-1703.
53. Nielsen TH, Sjöholm OR, Sørensen J: **Multiple physiological states of a *Pseudomonas fluorescens* DR54 biocontrol inoculant monitored by a new flow cytometry protocol.** *Fems Microbiology Ecology* 2009, **67**:479-490.

54. Rault A, Bouix M, Beal C: **Fermentation pH Influences the Physiological-State Dynamics of *Lactobacillus bulgaricus* CFL1 during pH-Controlled Culture.** *Applied and Environmental Microbiology* 2009, **75**:4374-4381.
55. Ju J, Misano H, Ohnishi K: **Directed evolution of bacterial alanine racemases with higher expression level.** *Journal of Bioscience and Bioengineering* 2005, **100**:246-254.
56. Neuenschwander M, Butz M, Heintz C, Kast P, Hilvert D: **A simple selection strategy for evolving highly efficient enzymes.** *Nature Biotechnology* 2007, **25**:1145-1147.
57. Oue S, Okamoto A, Yano T, Kagamiyama H: **Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues.** *Journal of Biological Chemistry* 1999, **274**:2344-2349.
58. Parales RE, Bruce NC, Schmid A, Wackett LP: **Biodegradation, biotransformation, and biocatalysis (B3).** *Applied and Environmental Microbiology* 2002, **68**:4699-4709.
59. Singh S, Kang SH, Mulchandani A, Chen W: **Bioremediation: environmental clean-up through pathway engineering.** *Current Opinion in Biotechnology* 2008, **19**:437-444.
60. Atsumi S, Liao JC: **Directed Evolution of *Methanococcus jannaschii* Citramalate Synthase for Biosynthesis of 1-Propanol and 1-Butanol by *Escherichia coli*.** *Applied and Environmental Microbiology* 2008, **74**:7802-7808.
- **61. Pfleger BF, Pitera DJ, D Smolke C, Keasling JD: **Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes.** *Nature Biotechnology* 2006, **24**:1027-1032.

Method description for tunable intergenic regions (TIGRs) which have been shown to modulate reporter gene expression over a 100-fold range and can be used to optimize natural and synthetic operons in both prokaryotes and eukaryotes.

- **62. Xu P, Gu Q, Wang WY, Wong L, Bower AGW, Collins CH, Koffas MAG: **Modular optimization of multi-gene pathways for fatty acids production in *E. coli*.** *Nature Communications* 2013, **4**:1409.

Method and technology for pathway optimization of multigenic pathways using a generalizable modular approach. Application demonstrates exceptional levels of fatty-acid production in *E. coli*.

- **63. Xu P, Vansiri A, Bhan N, Koffas MAG: **ePathBrick: A Synthetic Biology Platform for Engineering Metabolic Pathways in *E. coli*.** *Acs Synthetic Biology* 2012, **1**:256-266.

A general strategy for optimizing coordinated expression of synthetic or semi-synthetic multigenic pathways.

64. Zingaro KA, Papoutsakis ET: **Toward a Semisynthetic Stress Response System To Engineer Microbial Solvent Tolerance.** *Mbio* 2012, **3**:e00308-00312.
65. Yomano LP, York SW, Ingram LO: **Isolation and characterization of ethanol-tolerant mutants of *Escherichia coli* KO11 for fuel ethanol production.** *Journal of Industrial Microbiology & Biotechnology* 1998, **20**:132-138.
66. Heer D, Sauer U: **Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain.** *Microbial Biotechnology* 2008, **1**:497-506.
67. Williams TI, Combs JC, Lynn BC, Strobel HJ: **Proteomic profile changes in membranes of ethanol-tolerant *Clostridium thermocellum*.** *Applied Microbiology and Biotechnology* 2007, **74**:422-432.

68. Patnaik R, Louie S, Gavrilovic V, Perry K, Stemmer WPC, Ryan CM, del Cardayre S: **Genome shuffling of Lactobacillus for improved acid tolerance.** *Nature Biotechnology* 2002, **20**:707-712.
69. Zheng HB, Wang X, Yomano LP, Geddes RD, Shanmugam KT, Ingram LO: **Improving Escherichia coli FucO for Furfural Tolerance by Saturation Mutagenesis of Individual Amino Acid Positions.** *Applied and Environmental Microbiology* 2013, **79**:3202-3208.
70. Hou LH: **Novel methods of genome shuffling in Saccharomyces cerevisiae.** *Biotechnology Letters* 2009, **31**:671-677.
71. Brennan TC, Williams TC, Schulz BL, Palfreyman RW, Kromer JO, Nielsen LK: **Evolutionary Engineering Improves Tolerance for Replacement Jet Fuels in Saccharomyces cerevisiae.** *Appl Environ Microbiol* 2015, **81**:3316-3325.
72. Wang HK, Sun Y, Chen C, Sun Z, Zhou YC, Shen FD, Zhang HP, Dai YJ: **Genome shuffling of Lactobacillus plantarum for improving antifungal activity.** *Food Control* 2013, **32**:341-347.
73. Winkler J, Rehmann M, Kao KC: **Novel Escherichia coli hybrids with enhanced butanol tolerance.** *Biotechnology Letters* 2010, **32**:915-920.
74. Chong H, Geng H, Zhang H, Song H, Huang L, Jiang R: **Enhancing E. coli isobutanol tolerance through engineering its global transcription factor cAMP receptor protein (CRP).** *Biotechnology and Bioengineering* 2014, **111**:700-708.
75. Zhu L, Cai Z, Zhang Y, Li Y: **Engineering stress tolerance of Escherichia coli by stress-induced mutagenesis (SIM)-based adaptive evolution.** *Biotechnology Journal* 2014, **9**:120-127.
76. Hong ME, Lee KS, Yu BJ, Sung YJ, Park SM, Koo HM, Kweon DH, Park JC, Jin YS: **Identification of gene targets eliciting improved alcohol tolerance in Saccharomyces cerevisiae through inverse metabolic engineering.** *Journal of Biotechnology* 2010, **149**:52-59.
77. Dunlop MJ, Dossani ZY, Szmidsztal HL, Chu HC, Lee TS, Keasling JD, Hadi MZ, Mukhopadhyay A: **Engineering microbial biofuel tolerance and export using efflux pumps.** *Molecular Systems Biology* 2011, **7**:487.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Fig 1: Module based approach towards development of complex cellular traits. Case A outlines a rational approach where a few known genes are assembled into a strain and optimized for a desired trait. Case B encompasses cellular phenotypes which require a large number of largely unknown genes. By expanding the searchable genomic space through metagenomic or specialized libraries in combination with generation of genetic diversity an enrichment protocol identifies beneficial individual and combinations of genetic determinants which can then be assembled and optimized as in the rational approach.

Fig. 2: Methods for generation of genetic diversity for cell engineering of desirable traits. Techniques used in genome engineering can be broadly subdivided into methods which rely on accumulation of mutations, hybrid rational/selection-based approaches, and methods which introduce new genomic material. By accumulating random mutations through mutagenesis or by selective pressure positive traits can develop and be enriched in a strain. Combinations of these mutations can be combined via protoplast fusion in whole genome shuffling. Hybrid rational/selection-based approaches target specific sequences in a genome and attempt to modify them. The recombination based Multiplex Automated Genome Engineering (MAGE) uses recombination based editing of a chromosome. In combination with Conjugative Assembly Genome Engineering (CAGE) it can accumulate large numbers of precise changes to a target genome. Genomic libraries including those bearing heterologous or metagenomic DNA can be used to expand the genomic space beyond that of a single species.

Fig. 3: Improved expression of heterologous genes and genetic elements through introduction of heterologous sigma factors. By expressing heterologous sigma factors it becomes possible for the autologous RNA polymerase complex in the screening strain (such as *E. coli*) to recognize promoters from heterologous DNA (including metagenomic DNA), significantly improving transcription of heterologous genetic elements from genomic or metagenomic libraries.

Fig. 4: Exploration of combinatorial interactions using Coexisting/Coexpressing Genomic Libraries (CoGeLs). The CoGeL technology makes possible to screen for genetic-element interactions without *a priori* knowledge of interacting partners. This powerful technology is able to screen for new traits or trait improvements when binary genetic-element interactions are necessary or important.

Fig. 5: Flow chart for sequential, semi-synthetic construction of complex phenotypic traits. By utilizing multiple cell-engineering technologies in tandem with rational design approaches, where possible, one can sequentially identify new genetic-element targets and develop functioning cellular systems for engineering desirable traits.

Table 1

Table 1: Selection methods for development of chemical tolerance phenotypes.

Organism	Genome Engineering Technique	Stressor	Media – Complex or Defined	Enrichment Conditions	Selection Method	Ref.
<i>E. coli</i> KO11	Random Mutation	Ethanol	Complex (media + glucose)	Incubated 24 h at 35°C then transferred to fresh media; initial ethanol concentration of 35 g/L increased incrementally during transfers to 50 g/L	Culture was plated every 3-4 transfers; largest colonies were selected	[10]
<i>E. coli</i> JU01	Random Mutation	Acetate	Defined (AM1 media + xylose)	Mid-exponential culture was used to inoculate a mini fermenter, serial transfers were performed (after 9 th serial transfer initial xylose concentration was increased to 120 g/L)	Clonal strain was isolated from the 15 th transfer	[6]
<i>E. coli</i> EcNR1	Random Mutation	Isobutanol	Defined (NG50 media + glucose or xylose)	Three independent populations were grown in 3 mL of media containing 0.75% isobutanol in 15 mL Falcon tubes over 180 days. 96 well microplates containing isobutanol were then inoculated and specific growth rate was monitored at varying isobutanol concentrations	Clones identified to have highest specific growth rates were selected from microwell plates	[8]
<i>S. cerevisiae</i> TMB3400	Random Mutation	Furfural	Defined media	Cells were cultured in shake flasks containing 3 mM furfural. During late stationary phase 10 ⁷ cells were transferred to fresh media	Single colonies were isolated from plated culture after 30 th transfer	[7]
<i>C. thermocellum</i> ATCC 27405	Random Mutation	Ethanol	Complex (Basal media + cellobiose)	Incubated at 55°C with initial ethanol concentration of 0.5%; transferred to fresh media when maximum OD was reached.	Individual clones were isolated from cultures reliably growing in 5% ethanol	[9]

			(4 g/L))	Ethanol concentration increased by 0.5% every 2 transfers		
<i>Lactobacillus</i> LB-WT	Random Mutation, Mutagenesis, and Whole Genome Shuffling	Low pH	Complex (YE media)	Draw-fill semicontinuous chemostat adaption with controlled decrease in pH or NTG mutagenesis with plating on pH gradient plates, cultures then used for protoplast fusion and were enriched using pH gradient plates	Individual colonies selected from low pH section of pH gradient plates	[20]
<i>E. coli</i> XW92	Saturation Mutagenesis	Furfural	Defined (AM1 media + xylose)	100 colonies were used to inoculate separate wells in a microwell plate containing 12.5 mM furfural; growth was monitored via spectrophotometer	Cultures showing improved growth were selected	[93]
<i>S. cerevisiae</i> TH-AADY	Mutagenesis and Whole Genome Shuffling	Ethanol	Complex (enzyme hydrolyzed corn powder)	EMS mutated cells were plated on 8% ethanol plates, 183 colonies were selected and used as initial pool for genome shuffling through 3 rounds of fermentation with increasing glucose concentrations (up to 30%)	Culture was plated on 15% ethanol plates and 18 colonies were selected	[19]
<i>S. cerevisiae</i> W303	Mutagenesis and Whole Genome Shuffling	Ethanol	Complex (YPD media + glucose)	EMS mutated cells were plated on 6% ethanol plates, 162 colonies were selected and used as initial pool for genome shuffling; process was then repeated in 9, 12, and 15% ethanol	Thirteen individual colonies were isolated from 15% ethanol plates	[18]
<i>S. cerevisiae</i> 5288C	Random mutagenesis	Limonene,	Defined (CBS media)	Repeated culture in limonene containing media over 63 passages with limonene concentrations increasing from 7.6 to 222 mg/L. Frozen stock were saved from each passage.	Clones were selected by plating of the frozen stock on LM plates and selection of the largest colonies	[94]

<i>Lactobacillus rhamnosus</i> ATCC 11443	Mutagenesis and Whole Genome Shuffling	Low pH	Complex (YE media)	Mutagenesis with either UV irradiation or NTG; cultures then plated in 5% CO ₂ for 5 days at 37°C. Colonies with large transparent circles were selected and tested for increased lactic acid productivity. Two UV mutagenized strains and three NTG mutagenized strains were isolated and used in three rounds of protoplast fusion. Each round was then screened on successively lower pH plates.	Protoplast fusions were plated on low pH plates. Six colonies were selected from low pH plates after first shuffling, four colonies from the second, and four additional colonies from the third.	[95]
<i>E. coli</i> BW25113/ <i>Lactobacillus brevis</i> ATCC 367	Whole Genome Shuffling (two different organisms)	1-Butanol	Complex (LB media)	Protoplast fusion between <i>E. coli</i> and <i>L. brevis</i> to produce hybrid organisms. Individual strain measurement on 96-well plates.	Growth characteristics of each strain measured on 96-well plates.	[43]
<i>E. coli</i> DH5 α Δ crp	Error prone PCR of global regulator, crp	Isobutanol	Complex (LBGMg Media)	Error prone PCR library for crp expressing cells were cultured in isobutanol containing media over 3 – 5 passages with concentrations increasing with successive passages from 0.8% to 1.1% isobutanol.	Cells from subcultures were plated and individual colonies were selected.	[96]
<i>E. coli</i> FC40 Δ mutL/ <i>E. coli</i> FC40 Δ mutS	Stress induced mutagenesis on MutL and MutS knockout strains	Butanol, Salt, High Temperature	Complex (LB Media)	Growth at sub-inhibitory concentrations of desired stressant followed by plating at inhibitory concentrations of stressant.	Colonies isolated from growth inhibitory concentration plates	[97]
<i>C. acetobutylicum</i> ATCC 824	Genomic Library (<i>C. acetobutylicum</i> ATCC 824)	Butyric Acid	Complex (CGM media)	10% inoculum into media with 0.6% butyric acid with transfers every 24 hours into increasing butyric acid concentrations up to 1.3%	Final culture was plated and 18 clones were selected	[32]
<i>E. coli</i> MG1655	Genomic Library	Low pH	Complex (LB media)	10% inoculum into pH 2.0 culture for 1 hour at 37°C	Colonies selected from serial dilutions of	[2]

	(<i>E. coli</i>)		+ glutamate)		library enrichment	
<i>E. coli</i> BW25113 Δ recA::Kan	Genomic Library (<i>E. coli</i>)	Acetate	Defined (MOPS media)	1% inoculum into neutral pH media with 1.75 g/L acetate with serial transfers into fresh media at 24, 48 and 60 hours, culture was plated at 72 hours	10 individual colonies were selected and SCALES was performed on enriched population	[35]
<i>E. coli</i> BW25113	Genomic Library (<i>E. coli</i>)	1-butanol	Complex (LB media)	Mid-exponential cultures used to inoculate enrichment culture containing 0.9% 1-butanol; at OD ₆₀₀ ~ 0.7 culture was used to inoculate fresh media containing increased concentration of 1-butanol	Library plasmids were isolated during each transfer and used for comparative genome hybridization microarray to identify enriched genes	[34]
<i>E. coli</i> MG1655	Genomic Library (<i>E. coli</i>)	Ethanol	Complex (LB media)	Incubated 24 h at 37°C then transferred to fresh media; initial ethanol concentration of 2.5% increased incrementally during transfers to 5.25%	10 colonies isolated from 5.25% ethanol culture	[33]
<i>S. cerevisiae</i> CEN.PK2-1D	Genomic Library (<i>S. cerevisiae</i> SC288C)	Isobutanol	Defined (YSC media)	Library was plated on isobutanol gradient plates, colonies in high concentration zone were selected and serially cultured in 1% isobutanol media 10 times	Plated on 1% isobutanol plates and largest colonies were selected	[98]
<i>E. coli</i> DH1 Δ acrAB	Efflux Pump Library	1-butanol, isopentanol, geranyl acetate, geraniol, α -pinene, limonene, farnesyl hexanoate	Defined (M9 media)	Cultures were grown and solvent was added at concentration shown to inhibit growth (v/v: 0.5% 1-butanol; 0.25% isopentanol; 2% geranyl acetate; 0.05% geraniol; 2% α -pinene; 0.025% limonene; 2.5% farnesyl hexanoate); Every 10-14 hours cultures were diluted 1:100 in fresh media for 6 cycles	Library plasmids were extracted from selected cultures after 72 hours and were analyzed on a microarray to determine concentration of each library plasmid	[99]

<i>S. cerevisiae</i> BY4741	Library (gTME)	Ethanol/ Glucose	Defined (YSC- URA media)	Inoculated media with 100 g/L glucose and 5% ethanol at 30°C with two transfers to fresh media or same process with final transfer into 120 g/L glucose and 6% ethanol	20 colonies isolated from streaked plates from final cultures	[26]
<i>E. coli</i> MG1655	Multiplex Recombineering (TRMR)	Lignocellulosic hydrolysate	Complex (Low salt LB media)	Media containing 15-20% hydrolysate was inoculated to OD ₆₀₀ of 0.01, growth was monitored to OD ₆₀₀ = 0.2 where the culture was used to inoculate fresh media with 1% lower hydrolysate concentration; this process was repeated one additional time	TRMR modifications were tracked using microarray detection of molecular barcodes	[51]