The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology

Sandberg, Troy E.; Salazar, Michael J; Weng, Liam L; Palsson, Bernhard O; Feist, Adam M

Published in:
Metabolic Engineering

Link to article, DOI:
10.1016/j.ymben.2019.08.004

Publication date:
2019

Document Version
Peer reviewed version

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology

Troy E. Sandberg, Michael J. Salazar, Liam L. Weng, Bernhard O. Palsson, Adam M. Feist

PII: S1096-7176(19)30153-3
DOI: https://doi.org/10.1016/j.ymben.2019.08.004
Reference: YMBEN 1583

To appear in: Metabolic Engineering

Received Date: 10 April 2019
Revised Date: 1 August 2019
Accepted Date: 5 August 2019

Please cite this article as: Sandberg, T.E., Salazar, M.J., Weng, L.L., Palsson, B.O., Feist, A.M., The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology, Metabolic Engineering (2019), doi: https://doi.org/10.1016/j.ymben.2019.08.004.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier Inc. on behalf of International Metabolic Engineering Society.
The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology

Authors
Troy E. Sandberg¹, Michael J. Salazar¹, Liam L. Weng¹, Bernhard O. Palsson¹,², Adam M. Feist¹,²

Affiliations:
1. Department of Bioengineering, University of California, San Diego, CA 92093, USA.
2. Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark.

*Address correspondence to Adam M. Feist, afeist@ucsd.edu

Abstract
Harnessing the process of natural selection to obtain and understand new microbial phenotypes has become increasingly possible due to advances in culturing techniques, DNA sequencing, bioinformatics, and genetic engineering. Accordingly, Adaptive Laboratory Evolution (ALE) experiments represent a powerful approach to both investigate the evolutionary forces influencing strain phenotypes, performance, and stability, and to acquire production strains that contain beneficial mutations. In this review, we summarize and categorize the applications of ALE to various aspects of microbial physiology pertinent to industrial bioproduction by collecting case studies that highlight the multitude of ways in which evolution can facilitate the strain construction process. Further, we discuss principles that inform experimental design, complementary approaches such as computational modeling that help maximize utility, and the future of ALE as an efficient strain design and build tool driven by growing adoption and improvements in automation.
1. Introduction

Adaptive Laboratory Evolution, or ALE, experiments are an increasingly popular technique for both improving microbial phenotypes and investigating biological phenomena. ALE experiments can be traced back to reports of controlled evolution studies in the earliest parts of the last century (Atwood et al., 1951; Bennett and Hughes, 2009; Novick and Szilard, 1950) and are methodologically straightforward. In their simplest form, experiments consist of prolonged culturing of cells in a chosen environment to naturally select for those which acquire beneficial mutations. Given the speed with which beneficial mutations can arise and fix, it is safe to say that many biologists have ‘adaptively evolved’ their lab microbe of choice simply via the unavoidable cycles of growth/plating/freezing involved in cell culturing. ALE works robustly in microbes due to the ease with which large populations ($10^8 - 10^{10}$ cells) of rapidly dividing (20 min - <10 hour generation time) cells can be maintained; typical mutation rates and genome sizes ensure extensive sampling of the adaptive space, providing ample genetic diversity from which beneficial mutants will be naturally enriched (Gresham and Dunham, 2014). Furthermore, the rising use of ALE has been driven by the increasing availability of low cost, high throughput DNA sequencing (Shendure et al., 2017), which, when paired with appropriate bioinformatics tools, facilitates the increasingly critical mutation-identification step in the ALE process. Similarly, genome engineering is a key complementary technology (R. Liu et al., 2015), enabling the introduction of desired mutations into strains of interest for causal determination or phenotypic alteration.

Evolution’s reliance on natural selection to enrich for mutants with increased fitness allows strain optimization to be performed without requiring a priori knowledge of the genetic alterations necessary to effect such changes (Fig. 1). With perfect information on an organism’s genotype-to-phenotype mapping, desired traits could simply be engineered in rather than evolved, but the incredible complexity of biological systems renders our current knowledgebase insufficient to inform such rational design. ALE thus serves as a powerful technique for strain construction, complementing or even replacing rational design approaches which frequently induce stressed cellular states that negatively impact phenotype. It is important to note that, despite some
interchangeable use of the terms, ALE is distinct from the field of ‘directed evolution’ (Packer and Liu, 2015; Arnold, 2018) – ALE finds whatever genome-wide mutations aid in the fitness of actively growing cultures, while directed evolution typically targets a particular gene for mutagenesis and then screens resultant variants for a phenotype of interest, often independent of fitness effects. However, both approaches fall under the category of ‘evolutionary engineering’ (Sauer, 2001; Shepelin et al., 2018) given that they generate new phenotypes via random mutation followed by selection or screening.

Key to the success of ALE experiments is the fitness advantage bestowed by adaptive mutations, allowing the mutant strains to outcompete their ancestor and dominate the population. However, it is important to note that “fitness” is not a well-defined term, but rather depends on the growth environment employed in the ALE experiment. For example, if batch cultures are kept propagating in constant exponential-phase and with excess nutrients, then fitness is essentially equivalent to growth rate (Sandberg et al., 2014). If, instead, exponential-phase growth is not constantly maintained, then fitness expands to include factors other than just growth rate, such as survival in stationary phase or decreased lag phase (Wiser and Lenski, 2015). This sensitive environmental dependence of fitness thus allows different traits to be selected for based on the specifics of the ALE procedure. In addition to batch culturing, the most popular alternative growth method is continuous culture (e.g, chemostats or turbidostats) (Gresham and Dunham, 2014). Other ALE culturing methods have also been implemented for specific purposes, but have yet to see widespread use (See Section 4).

Once a desired growth environment has been carefully selected, an ALE experiment enables the generation of a whole lineage of mutant strains with improved fitness relative to the ancestral (i.e., starting) strain (Fig. 1B) and sequencing enables causal mutation determination (Fig. 1C). Selective sweeps often occur as the result of single beneficial mutations overtaking the population, but cultures contain significant amounts of genetic diversity from which intrapopulation competition can lead to clonal interference (Maddamsetti et al., 2015). It is common practice to save frozen stocks of cells periodically over the course of an ALE experiment, allowing for characterization of both endpoint strains and evolutionary intermediates. Although midpoints are perforce
less fit than endpoints (and, in turn, might have less of a tradeoff for alternate phenotypic properties), including them in characterizations can greatly aid in understanding the adaptive steps taken by the evolving ancestor to eventually arrive at an endpoint. Furthermore, by virtue of their smaller number of genetic alterations, midpoints can allow for easier discrimination of the minimum mutational set necessary to enable a bioprocess-relevant beneficial phenotype. In concert with phenotypic characterization (or screening), which establishes the fitness improvements realized over the course of the ALE, genetic characterization allows determination of the causal mutations that enable improved fitness in the ALE environment. To aid in causal mutation identification, independent replicates starting from the same ancestral strain are often evolved simultaneously – genes or genetic regions found to mutate multiple times across the independent replicates are almost surely adaptive (Bailey et al., 2015). Such causal mutations can thus be distinguished from neutral or even deleterious hitchhiker mutations that might otherwise require categorization via site-directed mutagenesis and growth profiling of the resultant mutants. Causal establishment of ALE mutations via reintroduction to a starting strain was possible and demonstrated in the mid 2000s (Herring et al., 2006), and this process is now more efficient with high-throughput genome engineering (R. Liu et al., 2015).

This review will focus on the use of ALE in industrial biotechnology and related fields by compiling a comprehensive collection of relevant studies, analyzing their collective contributions, and highlighting several case studies of importance. An analysis of the most common methods used for ALE and the strains utilized will be included. The review will also touch on the emergence of automation and bioinformatics in the field and the impact this has had. Previous reviews relating to ALE have addressed the underlying genetic and metabolic basis for adaptive evolution experiments (Conrad et al., 2011; Gresham and Dunham, 2014; Long et al., 2015; Long and Antoniewicz, 2018; Remigi et al., 2019), the role of systems biology and in silico evolution (Hansen et al., 2017; Hindré et al., 2012; Papp et al., 2011), overall strategies to improve metabolic performance (Rabbers et al., 2015; Shepelin et al., 2018; Williams et al., 2016) and tolerance (Peabody et al., 2014), as well as other industrially relevant strain properties (Bachmann et al., 2015; Dragosits and Mattanovich, 2013; Portnoy et
This work provides an updated review and categorization of studies useful for metabolic engineering, as well as other selected foundational studies using laboratory evolution that are relevant to industrial biotechnology. As such, studies focused on the use of ALE in understanding antimicrobial resistance and other fundamental principles of evolutionary biology will not be covered here, as they have been reviewed elsewhere (Hughes and Andersson, 2015; Kawecki et al., 2012; Palmer and Kishony, 2013; Remigi et al., 2019).

Figure 1. Adaptive Laboratory Evolution

A) Microbes are cultured in a desired growth environment for an extended period of time, allowing natural selection to enrich for mutant strains (altered coloration) with improved fitness. This example depicts ALE via serial propagation of batch cultures. B) Evolved strains are characterized for phenotypic improvements relative to the ancestral strain, using whatever “fitness” metric is appropriate given the evolutionary environment. C) Evolved strains have their DNA sequenced to reveal the adaptive mutations.
enabling phenotypic improvement. This example case depicts the fixation of two successive mutations targeting the same genetic region.

2. Applications of ALE

ALE has become a valuable tool in metabolic engineering for strain development and optimization by reliably facilitating microbial fitness improvements, via both predictable and non-intuitive mechanisms. However, specific mutations themselves are rarely predictable given current knowledge, and the adaptive strategies employed are often strain-specific. Supplemental Table 1 outlines a collection of ALE-based manuscripts that will serve as case studies and a source of content for this review. The different uses of ALE can be roughly categorized into five general application areas (Figure 2): 1) growth rate optimization, 2) increasing tolerance, 3) substrate utilization, 4) increasing product yield/titer, and 5) general discovery. Although these categories and applications are not mutually exclusive and ALE experiments often improve multiple strain properties simultaneously, it serves as a useful heuristic.

Several features are revealed through a categorization of ALE studies relevant to industrial biotechnology. First, the predominant use of ALE in this collection involves increasing an organism’s ability to tolerate stresses (55 out of 159 included studies, 35%). Tolerance evolutionary engineering has significant industrial implications due to the toxicity that many biorenewables have on microorganisms’ overall production capacity and viability (Lee and Kim, 2015; Qi et al., 2019; S. Wang et al., 2018). Most starting strains and production chassis have never had to deal with high amounts of desired product chemicals in their native environments, thus improved tolerance mechanisms are almost always beneficial for a successful strain design. Second, there is significant focus on metabolite production as well as nutrient uptake, with 15% and 20% of the studies, respectively. In metabolic engineering the two concepts are frequently linked, given that sufficient nutrient uptake is a prerequisite for any metabolite biosynthesis, but there are distinct cases where microorganisms are designed and/or evolved to either utilize poorly accessible substrates or produce non-native metabolites. Nutrient uptake is particularly relevant in industrial settings where poor uptake of desired feedstocks can lead to inefficient conversion and inhibition of metabolic
pathways (Görke and Stülke, 2008; Ingram and Doran, 1995). Lastly, the general discovery category contains studies covering an array of topics that are relevant to industrial biotechnology and can be investigated with ALE, including systems biology, evolutionary modeling, and genome dynamics (Barrick and Lenski, 2013). This general category could be significantly expanded via inclusion of purely computational or theoretical rather than empirical evolution studies, but that is beyond the scope of this review.

Each category will be individually addressed in the following sections to outline a general summary of their applications, highlight case studies useful for demonstration, and summarize the lessons learned from each collection of studies.
Figure 2: A Categorization of ALE Studies

A diagram of different categories of use for laboratory evolution experiments, detailing the percent makeup of studies examined (159 total). ‘Growth rate optimization’ illustrates fitness development over the course of an ALE, with noticeable fitness jumps. ‘Increase tolerance’ illustrates an experimental schematic of an initially mixed population acquiring beneficial mutations (red cells) that promote cell survival in a constantly increasing external stress environment (e.g., pH, antibiotics, temperature, etc).

‘Substrate utilization’ and ‘Increase Product Yield/titer’ illustrate evolutionary pathways enabled via ALE that enhance the organism's ability to make use of alternative nutrient sources (colored circles), and to increase production of metabolites of interest (colored squares), respectively. Lastly, ‘General Discovery’ encompasses studies that examined ALEs at a genetic or systems level in greater detail.

2.1 Growth Rate Optimization

While “fitness” is often a complex trait with many contributing factors, growth rate (ln(2)/doubling time) is generally the most important fitness determinant in ALE experiments (Vasi et al., 1994). Indeed, although the other ALE use categories are more specialized, the enrichment for adaptive mutants would not be possible without a growth advantage relative to the starting strain, regardless of the specifics of the environment. A key reason for the use of microbes in biotech and research is their short generation time, and ALE allows for rapid improvement of this desirable trait. ALE has been applied to optimize the growth rate of various industrially relevant microbial species (Hong et al., 2011; LaCroix et al., 2014; Pfeifer et al., 2017; Yu et al., 2013), to ameliorate growth defects in engineered strains (Aguilar et al., 2018; Carroll and Marx, 2013; Radek et al., 2017; Thyer et al., 2018; Zelle et al., 2011), or to study how and why growth rate changes over the course of evolution (Barrick et al., 2009; Ferea et al., 1999; Lenski et al., 1995, 1998, 1991; Lenski and Travisano, 1994; Weikert et al., 1997; Wong and Liao, 2009).

ALE can significantly increase a strain’s growth rate, even in frequently used culturing environments for which one might expect growth to already be close to optimal. For example, a simple ~1 month ALE experiment in glucose minimal medium resulted in growth rate improvements of up to 60% for the commonly used lab strain and engineering chassis E. coli K-12 MG1655 (LaCroix et al., 2014). This rapid growth
phenotype was achieved predominantly through mutations to only three genetic regions, suggesting their easy use as allelic knock-ins to improve growth on glucose of production strains derived from MG1655, as has been demonstrated in the case of an RNA polymerase mutation (Rugbjerg et al., 2018). Such causal mutations can even have cross-phylum fitness implications – evolution on glucose minimal media led to mutations in pyruvate kinase and the CorA magnesium transporter in both C. glutamicum (Pfeifer et al., 2017) and MG1655. Moreover, the metabolic flux state of the MG1655 evolved strains was found to be remarkably unperturbed - rather than changing relative pathway usage, faster growth was enabled by proportional increases to all intracellular rates of central carbon metabolism (Long et al., 2017). ALE has been used to successfully improve growth capabilities of a variety of non-bacterial species (Supplemental Figure 1), most notably yeast, where a similar volume of ALE work has been performed (Mans et al., 2018). As with E. coli, improved growth rates in different organisms frequently come along with other beneficial traits, such as evolved yeast strains outperforming rationally designed ones in several phenotypic metrics (Hong et al., 2011), and evolved microalgae achieving improved lipid productivity (Yu et al., 2013).

In addition to optimizing industrially or scientifically relevant wild-type strains, ALE has been successfully applied to fix fitness defects in strains rationally designed for a specific metabolic phenotype. An increased growth rate, in most cases, also makes a strain more robust to perturbations resulting from genetic alterations, as mutations can restore microbes back to their wild-type physiology and activate flux through alternate energy generation pathways (McCloskey et al., 2018). In one study, an engineered M. extorquens strain with a foreign metabolic pathway (resulting in abnormal morphology and significantly decreased growth rate) was subjected to laboratory evolution (Carroll and Marx, 2013). ALE resulted in the isolation of evolved strains with up to 150% growth rate improvements, elimination of morphological abnormalities, and insight into the metabolic consequences resulting from introduction of the foreign pathway. In another study, ALE was used in concert with strain engineering to impart novel metabolic capabilities onto yeast (Zelle et al., 2011). The native malic enzyme in S. cerevisiae is unable to function anaplerotically via pyruvate carboxylation, nor is heterologous
expression of the *E. coli* malic enzyme sufficient to activate this ability, but ALE successfully selected for a mutant enzyme that enabled this trait. ALE can also reliably ameliorate fitness defects resulting from ambitious, large-scale strain alterations - a genomically recoded *E. coli* with all 300+ amber stop codons removed was evolved to achieve strains with both significantly faster (>40%) growth rates and improved incorporation of nonstandard amino acids, opening the door to novel protein production capabilities (Wannier et al., 2018). Similarly, a genome-reduced *E. coli* with more than 1 million basepairs removed from its chromosome grew 3-fold slower than the wild-type in glucose minimal media, and yet a <1,000 generation ALE experiment almost completely abolished this severe fitness deficit (Choe et al., 2019).

ALE for improvement in growth rate thus serves not only to increase the ease and speed of microbial culturing, but also as a technique for targeted selection on specific rate-limiting system components. This is especially important when a strain has been engineered and rate-limiting constraints now significantly impact fitness, which will become an increasingly frequent occurrence as genome engineering techniques continue to improve. With proven efficacy across a number of species and the frequent occurrence of concomitant desirable traits such as increased productivity/yield, growth optimization stands as a key application for ALE experiments. Although the following sections delve into more specialized uses, it is important to note that growth rate improvements underlie these ALE outcomes as well.

### 2.2 Increase Tolerance

When exposed to stress, cells respond by activating stress responses, which require a reallocation of cellular resources to enact and thus inhibit growth potential. Cellular stress can be caused by a complex growth environment that induces some form of metabolic misregulation, or by exposure to specific inhibitory chemicals or biomolecules. Common stressors of both basic and applied interest introduce a diverse range of challenges for growing cells to overcome, and ALE has been established as an efficient tool to surmount such challenges. ALE studies have investigated the adaptive responses to many of these stressful environments: different levels of pH (Fletcher et al., 2017; Kildegaard et al., 2014; Zorraquino-Salvo et al., 2014), osmotic pressure...
Compounds with potential biotechnological value can often not be efficiently metabolized or produced in living systems due to toxicity. For example, many carboxylic acids are promising biorenewables but induce membrane damage, and attempts to rationally engineer greater membrane integrity have failed to improve carboxylic acid production. Royce et al. addressed this with an ALE experiment of *E. coli* onto octanoic acid, wherein strains were obtained that had improved tolerance not just to octanoic acid but also several other carboxylic acids and butanol isomers, with concomitant 5-fold higher production titer (Royce et al., 2015). This highlights how stress mitigation can simultaneously improve production properties, and how certain classes of stressors...
trigger similar physiological responses that cause particular mutations to be beneficial across multiple environments. In another study, knowledge of *S. cerevisiae* metabolism led to an ALE experiment designed for osmotic stress tolerance - this pushed the yeast metabolic state more towards glycerol production and away from ethanol, resulting in evolved strains with improved properties for lower-alcohol wine production (Tilloy et al., 2014).

Causal mutations in tolerance-evolved strains are straightforward to identify with DNA sequencing, but interpreting the molecular causality underlying a mutation’s phenotypic impact is often more challenging. A study by Caspeta *et al.* stands as a successful example of ALE results elucidating the molecular mechanisms of adaptation – yeast was evolved for improved thermostolerance, yielding strains that grew >50% faster at 40°C (Caspeta *et al.*, 2014). Mutational convergence in independently evolved replicates revealed that a gene inactivation changed the sterol composition of the cells from ergosterol to fecosterol, optimizing membrane fluidity at the elevated growth temperatures. ALE can also be used to increase tolerance to multiple stressors at once, as demonstrated by a study in which an industrial yeast strain was evolved for improved tolerance to both biomass hydrolysate inhibitors and high temperature, creating a robust strain for ethanol production via simultaneous saccharification and fermentation (Wallace-Salinas and Gorwa-Grauslund, 2013). However, it should be noted that tolerance to one stress often comes with tradeoffs in a separate growth environment, depending on the particular adaptive mutations acquired – while strains evolved to high-temperature typically lose fitness at low-temperature, one ALE-identified mutation managed to avoid such a tradeoff (Rodriguez-Verdugo *et al.*, 2014).

A significant effort to overcome the toxicity of a nutrient stressor, specifically L-serine in high concentration, was seen in the studies performed by Mundhada *et al.* L-serine is a promising target for commercial bioproduction but induces inhibition of various cellular processes such as peptidoglycan synthesis and cell division, and the reactive byproducts such as acrylates can hinder cell growth (Zhang and Newman, 2008; Zhang *et al.*, 2010; de Lorenzo *et al.*, 2015). An initial study succeeded in achieving L-serine tolerance in *E. coli* by employing random mutagenesis and selection, but it was not able to meet the demand for extensive serine production (Mundhada et
al., 2016). By implementing ALE, E. coli lacking L-serine degradation pathways (which makes cells much more susceptible to toxicity) were evolved to increasing concentrations of L-serine that reached 100 g/L, resulting in a titer much higher than the mutagenized strain with the same pathway (Mundhada et al., 2017).

Evolutionary engineering for improved stress tolerance stands as a dominant ALE application for metabolic engineering and industrial biotechnology. As previously stated, wild-type cells are rarely well-poised to deal with the burdens caused by growth environments with high concentrations of an atypical but desired bioproduction chemical. Where rational engineering approaches can fall short due to a lack of requisite knowledge, ALE enables mechanistically naïve but reliable improvements to industrially valuable strain properties. Mutational results from ALE experiments can then be used to inform further strain design attempts, while also elucidating basic tenants of biological systems. Further, attempts to rationally design strains for a specific phenotype often induce fitness defects which can be countervailed by evolution. Even if specific tolerance mechanisms can be rationally engineered into a strain, necessary gene expression levels can be hard to infer. In contrast, in a single experiment, ALE can find both the mechanisms underlying fitness defects and the gene expression levels necessary to enable a functional and robust phenotype. ALE also enables acclimation to growth environments with complex combinations of stressors that rational design cannot currently address given our existing biological knowledge base. Importantly, when stress on strains is reduced via ALE, the result is typically not only growth rate increases, but also phenotypic improvements in industrially relevant properties such as production or degradation capabilities.

2.3 Substrate Utilization

Efficient substrate uptake is necessary to enable satisfactory production of metabolites and additional products of interest in most, if not all, industrial applications. There exists a very direct link between organism fitness and the ability to efficiently uptake and metabolize rate-limiting growth nutrients, thus ALE experiments with a properly designed selective environment readily uncover mutations conducive to robust growth on substrates of interest. Primary motivations for improving substrate utilization
of an organism are tied to current market surpluses/prices (Ji et al., 2011; Lin and Tanaka, 2006) and harnessing the energy of inaccessible or toxic byproducts to improve yield of commodity chemicals (Plácido and Capareda, 2016). The most ubiquitous organism studied to improve uptake is the yeast *Saccharomyces cerevisiae* (Cadière et al., 2011; Garcia Sanchez et al., 2010; Guimarães et al., 2008; Kim et al., 2012; Smith et al., 2014; Sonderegger and Sauer, 2003; Zha et al., 2014; Zhang et al., 2018; Zhou et al., 2012), with the CEN.PK strain and its derivatives the most common (de Kok et al., 2012; Ho et al., 2017; Jansen et al., 2004; Klimacek et al., 2014; Kuyper et al., 2005, 2004; Marques et al., 2017; Merico et al., 2011; Novy et al., 2014; Ochoa-Estopier et al., 2011; Papapetridis et al., 2018; Scalcinati et al., 2012; Strucko et al., 2018; van Rossum et al., 2016), mainly due to its widespread use in industry, robust fermentative capability, and inherent ethanol tolerance. There has been considerable interest in generating microbial strains that can utilize the fermentable sugars in lignocellulosic biomass (Clark et al., 2012; Sanderson, 2011); *S. cerevisiae* has again been the driving force in discovery to this end (Klimacek et al., 2014; Marques et al., 2017; Zha et al., 2014), with some work done in *E. coli* strains (Lee and Palsson, 2010; Rajaraman et al., 2016; Sandberg et al., 2017, 2016; Utrilla et al., 2012) and other various bacteria and yeast (Cordova et al., 2016; Latif et al., 2015; Moser et al., 2017).

*Saccharomyces cerevisiae*, due to its anaerobic fermentative capabilities and tolerance to low pH and phage infection, has been the species of interest for industrial production of specialty metabolites (Moysés et al., 2016). Although *S. cerevisiae* does not natively catabolize xylose or other pentose sugars readily available in plant biomass, there has been considerable work to engineer strains capable of co-utilizing these compounds (Chu and Lee, 2007; Van Vleet and Jeffries, 2009). Adaptive laboratory evolution has been used in tandem with other metabolic engineering techniques to further drive substrate utilization. In (Klimacek et al., 2014), previously constructed *S. cerevisiae* strains with enhanced expression of heterologous genes involved in xylose metabolism were evolved via a two phase process. ALE selection for anaerobic xylose fermentation led to a strain with faster growth but undesirable metabolic byproducts, and a subsequent ALE on this strain under batch culture conditions ameliorated these phenotypic defects. The final strain grew more than 500%
faster than the original progenitor despite fewer than 100 total generations of evolution, highlighting the impressive speed with which ALE can realize significant fitness improvements.

ALE has been used to tackle challenges in the utilization of plant biomass as a nutrient source, such as the presence of toxic or inhibitory byproducts like furfural and acetate (Bellissimi et al., 2009; Heer and Sauer, 2008). A study by Rajaraman et al. addressed this with a chemostat ALE on several E. coli strains with acetate as the sole carbon source (Rajaraman et al., 2016). Evolved strains had a specific growth rate increase of ~25% and significantly altered expression for a number of genes, enabled by a single amino acid substitution in the RNA polymerase complex subunit RpoA; unlike with rational engineering of specific genes for altered expression, ALE can find point mutations changing global expression patterns in a net beneficial way not predictable with current protein modeling techniques. In addition to improving growth rate on non-optimal nutrients, a properly designed ALE experiment can also enable growth on completely non-permissive substrates. With a passage protocol in which glycerol concentration was steadily lowered while L-1,2-propanediol was increased, E. coli strains were obtained capable of growth solely on the latter substrate, despite the ancestral strain’s inability to uptake or metabolize this nonnative carbon source (Lee and Palsson, 2010). A study that significantly expanded this approach involved automation of the process, weaning cultures off of a helper substrate (which generated diversity), and selecting for strains that could utilize a number of non-native carbon sources by directly targeting promiscuity properties of enzymes (Guzmán et al., 2019; Notebaart et al., 2018). Impressively, growth on these new substrates was enabled by the acquisition of only one or two point mutations – completely ‘novel’ phenotypes for a particular strain can readily be acquired via evolution if the growth environment provides the necessary fitness benefit for such an occurrence.

Although ALE has proven efficacy for enabling growth on non-native substrates, access to a new nutrient niche can require large jumps across the fitness landscape that render a phenotype infeasible to select for on reasonable timeframes – famously, it took 15 years for one of twelve Long Term Evolution Experiment populations to acquire a citrate-consuming phenotype, and the other eleven populations are still unable to
consume citrate after 30 total years (Leon et al., 2018). Rational design with heterologous pathways can facilitate such fitness landscape leaps, but clever use of ALE alone can sometimes suffice. Szapponos et al. utilized metabolic modeling to identify non-native growth substrates with similar catabolic pathways, then performed a multi-step ALE to select for a complex metabolic innovation (Szappanos et al., 2016).

Specifically, *E. coli* cannot natively grow on propylene glycol (PG) or ethylene glycol (EG), but culturing of a hypermutable variant resulted in isolation of PG+, but not EG+, strains. However, selecting for EG growth using PG+ mutant strains did successfully yield the EG+ phenotype, indicating an increase in adaptation rate of more than two orders of magnitude. Innovative niche expansions can thus be achieved via stepwise evolution, with complimentary technologies such as metabolic modeling providing the information necessary to choose suitable substrate targets.

ALE has seen significant use as a tool for improving substrate uptake and metabolism, often countervailing the growth-inhibitory effects of toxic byproducts or fixing metabolic network misregulation in non-optimal growth environments. In this way, the utility of ALE has been established in use cases ranging from improving wild-type strain growth on a sub-optimal substrate, to facilitating proper metabolic assimilation of heterologous pathways in engineered strains, to imparting novel growth phenotypes even in unengineered strains - in each case improving substrate uptake as a coupled property. The mutational mechanisms range from single-nucleotide polymorphisms and small indels (insertions or deletions) in specific enzymes that enable pure innovation (e.g., utilizing a novel substrate), to altered transporter activity that leads to more efficient import and routing of substrates into central metabolism, to regulatory mutations which globally rebalance the metabolic network, shutting down cellular content unnecessary for bioreactor growth and freeing up resources for growth-coupled processes like substrate uptake. ALE can thus greatly complement traditional engineering techniques for altering microbial response to various feedstocks, creating more efficient organisms with improved chemical conversion to industrially desirable products.

### 2.4 Product Titer/Yield Optimization
Maximizing the bioproduction of metabolites of interest is a difficult task, and rational approaches to strain design are frequently insufficient to achieve desired productivity (Guimarães et al., 2008; Kuyper et al., 2005, 2004; Shepelin et al., 2018). A main source of difficulty lies in the conflict between fast, robust strain growth and the fitness-counterproductive repurposing of cellular resources to produce large amounts of a desirable compound. Nevertheless, ALE studies with proper experimental design can be used to overcome this conflict and optimize metabolite production and increase titer (Basso et al., 2011; Charusanti et al., 2012; Fong et al., 2005a; Fu et al., 2013; Grabar et al., 2006; Jiang et al., 2013; Lee et al., 2016, 2014; Lu et al., 2012; Luo et al., 2019; Mahr et al., 2015; Otero et al., 2013; Pontrelli et al., 2018; Reyes et al., 2014; Royce et al., 2015; Shen et al., 2012; Smith and Liao, 2011; Vilela et al., 2015; Wang et al., 2012; Wisselink et al., 2007; Zhang et al., 2007; Zhao et al., 2013; Zhou et al., 2005, 2003).

Given that an ALE experiment’s efficacy is contingent on the fitness advantage bestowed by adaptive mutations, it is essential to tie metabolite production to fitness in some way. The most straightforward way to do this is to genetically alter strains to couple production to overall energy and biomass generation rates in a given environment, and computational studies suggest a multitude of possible growth-coupled products that go far beyond simple and intuitive strain designs (Burgard et al., 2003; Klamt and Mahadevan, 2015; Pharkya et al., 2004; von Kamp and Klamt, 2017). However, given the limited scope of molecules for which this has been executed and validated (i.e., coupling production to overall energy generation), innovative methods (e.g., using a selective environment where production of the molecule of interest provides some protective or fitness-beneficial function) and advanced selection systems (e.g., fluorescent reporters + FACS sorting) are often needed to improve production for a wide range of molecules. Success cases for multiple forms of production optimization via growth coupling will now be discussed.

Genetic alterations that modify the metabolic network of a cell can force growth-sustaining flux through desired pathways, thereby coupling cell viability to the production of a compound of interest. Significant computational work has been performed to enable prediction of specific gene knockouts or heterologous pathway insertions that result in growth-coupling (Burgard et al., 2003; Feist et al., 2010; Jensen
et al., 2019; Klamt and Mahadevan, 2015; Pharkya et al., 2004; von Kamp and Klamt, 2017), and such approaches have been experimentally validated (Alper et al., 2005; Fong et al., 2005a; Yim et al., 2011) and are growing in utility due to advances in genome-scale metabolic modeling techniques (King et al., 2017). In a study by Jantama et al., this modeling-informed rational design was combined with ALE to optimize an *E. coli* strain with gene knockouts that forced NAD+ regeneration to occur via malate- and succinate-producing pathways under anaerobic growth conditions (Jantama et al., 2008). By thus coupling production to NADH oxidation, necessary for the growth-essential maintenance of ATP generation, ALE successfully selected for strains with robust minimal medium growth and improved malate and succinate titer. Further, physiological characterization of evolved strains pointed to additional knockouts to decrease byproduct formation, ultimately yielding strains with even greater chemical yields. Properly designed growth-coupling strategies can even see utility across diverse organisms and enzyme types (Jensen et al., 2019), as exemplified by recent work coupling SAM-dependent methylation to growth (Luo et al., 2019). By tying production of the essential amino acid cysteine to the activity of methyltransferases, Luo et al. were able to use ALE to select for both *E. coli* and *S. cerevisiae* strains with mutations providing 2-fold increases in heterologous methyltransferase activity. Ensuring that the methyltransferase activity was rate-limiting for growth forced adaptive mutations to preferentially target this enzymatic bottleneck; *in vivo* enzyme engineering was thus undertaken by the evolving cells, with demonstrated success for improving activity of both N- and O-type methyltransferases.

Innovative selection strategies and systems have also been demonstrated, which expand the range of compounds whose production can be optimized using ALE beyond those that can be directly growth-coupled. One such selective approach involves utilizing interspecies competition or cooperation (i.e., syntrophy). Charusanti et al. serially propagated MRSA alongside an antibiotic-producing microorganism, harnessing competition to achieve an *S. clavuligerus* strain with constitutive holomycin expression (Charusanti et al., 2012). The production of this molecule, though energetically costly, was nevertheless selectable with ALE due to the competitive co-culturing growth environment utilized. Similarly, through the use of syntrophic coupling and co-cultures,
Lloyd et al. designed and evolved mutually reliant strains to optimize the sharing and production of non-trivial compounds (Lloyd et al., 2019). Environmental manipulation in monocultures has also been successfully employed, as in a study by Reyes et al. wherein exposure to hydrogen peroxide selected for more than 3-fold increases in carotenoid production due to these compounds’ antioxidant properties (Reyes et al., 2014). Such approaches change the meaning of “fitness” to now include production of the molecule of interest as a factor, but other selective techniques are possible – a noteworthy study in this category used FACS to apply multi-dimensional selection to a strain engineered with a biosensor for valine-induced fluorescence (Mahr et al., 2015).

Biosensors provide a powerful way to link otherwise-obscured intracellular states to a screenable output (D. Liu et al., 2015), and by screening out and serially propagating only the most highly fluorescent cells, Mahr et al. increased valine titer by ~25% alongside growth rate improvements and decreased byproduct formation, despite maximum possible growth rate being higher if the cells had acquired mutations to eliminate their energetically wasteful expression of fluorescent proteins.

When paired with properly selected experimental methodology and biological design, ALE’s ability to increase bioproduction of desired molecules has been established. Though more difficult to select for than simple growth rate improvements, fitness can still be tied to titer increases in a variety of circumstances: when the produced molecule can enable better growth or survivability in spite of production costs; when evolution can countervail phenotypic defects resulting from metabolic engineering for a novel production phenotype; or when growth rate maximization is not the sole driving force behind selection. An important factor to keep in mind is the stability of production strains – though production phenotypes can be selectable under carefully designed ALE conditions, industrial-scale bioreactor growth can open the door for population-takeover by mutant strains which genetically purge the production machinery (Rugbjerg and Sommer, 2019). Engineering techniques can improve strain stability by removing mobile genetic elements or error-prone polymerases and thus reducing mutation rate (Csorgo et al., 2012), but, fortunately, this is not a strict necessity – a 1,4-BDO bioproduction strain (Yim et al., 2011) has been used at industrial scales by Genomatica, and found to remain genetically stable over the course of a bioprocess
cycle (personal communication, John D. Trawick). As our knowledge of metabolic and regulatory networks grows and techniques for engineering and strain design continue to improve, we are poised to see an increased adoption of ALE as a complementary technology for optimizing microbial production strains.

2.5 General Discovery

By enabling observation of evolutionary outcomes in a controlled laboratory setting, ALE facilitates research into basic bioprocesses in addition to its more applied uses as an engineering and strain design/optimization tool. ALE studies have yielded insight into important evolutionary phenomena such as clonal interference and regulatory rewiring (Kao and Sherlock, 2008; Oud et al., 2013; Sniegowski and Gerrish, 2010; Maddamsetti et al., 2015), rate and mechanism of mutation development (Araya et al., 2010; Chang et al., 2013; Dunham et al., 2002; Notley-McRobb and Ferenci, 1999; Yona et al., 2012; Lind et al., 2017; Ene et al., 2018; Lauer et al., 2018), and response to genetic perturbation or sub-optimal growth environment (Charusanti et al., 2010; Conrad et al., 2010; Fong et al., 2005b; Wang et al., 2010; Giusy M. Adamo et al., 2012; Deng and Fong, 2011; Szamecz et al., 2014; Wenger et al., 2011; Wright et al., 2011). ALE has also been used to study fundamental aspects of biology such as regulatory gene network response to external stressors or a new environment (Quan et al., 2012; Zhu et al., 2015), as well as the impact of genome shuffling and recombination techniques on the construction of industrially pertinent strains (Peabody et al., 2016; Reyes et al., 2012; Winkler and Kao, 2012). When paired with in silico modeling and algorithmic approaches frequently used in systems-level analyses, ALE additionally serves as an empirical hypothesis tester (Dekel and Alon, 2005; Ibarra et al., 2002).

Several studies have been dedicated to uncovering the molecular framework of an organism’s response to a genetic perturbation, with obvious implications for strain engineering. Gene essentiality and the impact of gene knockouts has been analyzed using ALE (Charusanti et al., 2010; Tokuyama et al., 2018), wherein strain physiology immediately post knockout can be compared against the physiological state following the strain’s adaptive evolution. For example, Charusanti et al. investigated the adaptive
response of multiple *E. coli* lineages following knockout of the important phosphoglucose isomerase (*pgi*) gene (Charusanti et al., 2010), a frequent target for disruption in engineered strains (Chin and Cirino, 2011; Papapetridis et al., 2018; Shiue et al., 2015). Distinct phenotypic outcomes were reached across the different evolved endpoints, and examination of the flux state via carbon-13 metabolic flux analysis determined that the dominant mechanism of adaptation was to overcome key bottlenecks in cofactor metabolism and substrate uptake that were induced by the knockout (Long et al., 2018). Studies such as these examining evolutionary response to genetic perturbation are increasingly yielding insight into the regulatory architecture of metabolic networks (McCloskey et al., 2018).

Modifications to ALE experiment methodology have been explored for the purposes of speeding adaptation, increasing mutation rate, and maintaining beneficial mutations. Bacterial asexual reproduction leads to clonal interference, preventing many beneficial mutations from reaching fixation within the evolving populations. Chu et al. investigated induced horizontal gene transfer as a method to speed adaptation and circumvent clonal interference, and found that genetic exchange between distinct co-cultured *E. coli* strains could significantly aid in adaptation to growth on novel carbon sources, contingent on the specifics of donor vs. host strain similarity and evolutionary environment (Chu et al., 2018). Contrasting with horizontal gene transfer, another technique to increase the accessibility of adaptive mutations is simply to increase the mutation rate, which can be induced with chemical mutagens (Lee et al., 2011) or naturally result from a DNA repair gene deactivation (LaCroix et al., 2014; Lenski et al., 2015); in all cases, mutation rate positively correlated with the speed of adaptation and magnitude of fitness gains. However, Couce et al. found that, even under conditions of strong selection and with improving population fitness, hypermutability leads to steady decay of the genome and can cause ‘genomic draft,’ whereby deleterious mutations hitchhike and cannot be purged from the population (Couce et al., 2017). Care should thus be taken when evolving hypermutable strains – though hypermutability can facilitate the speed and magnitude of ALE fitness increases, and even make leaps across the fitness landscape accessible by providing a large mutational pool with
complex epistatic interactions, there are downsides in regards to genome stability and even adaptive potential, if the mutation rate is inordinately high (Sprouffske et al., 2018). The novel or improved microbial phenotypes achievable with ALE makes it a powerful engineering tool, and studying the genetic and metabolic mechanisms underlying these phenotypic changes can yield valuable biological discoveries. Moreover, modifications to the selection protocol and evolutionary environment are easily explorable with ALE, allowing investigation into methods by which adaptation can be facilitated. Insights gained through the use of ALE can then be employed in other design strategies for novel solutions in industrial settings.

3. Strain Types and Generations

As long as an organism is reproducibly culturable in laboratory conditions it can be subjected to ALE, though the majority of studies analyzed in this review utilize either bacteria or yeast, specifically *E. coli* and *S. cerevisiae* (Supplemental Figure 1A). This is due to their fast growth, genetic tractability, and status as model organisms, providing the necessary knowledge base with which to interpret evolutionary outcomes. Using different organisms such as algae can allow for evolved phenotypes not achievable with bacteria or yeast, but ALE studies focusing on less standard microbial species are comparatively scant. Although much less relevant to metabolic engineering, higher organisms like fruit flies or even mice have been selectively competed and bred under controlled conditions (Barrett et al., 2019; Folk and Bradley, 2005), though these non-microbial ALE studies are beyond the scope of this review. ALE for mammalian production systems is limited given their slow growth rates and more involved culturing requirements, but some work has been done examining the genetic changes which accumulate in CHO cell lines under various culturing conditions (Feichtinger et al., 2016). Earlier examples without genetic analysis also exist, such as adapting hybridoma cells to lower serum concentrations (Lee et al., 1991). However, larger scale DNA sequencing is needed to reliably identify the potential causal mutations from such mammalian selection experiments, and epigenetic factors can further complicate the interpretation of results.
A critical parameter in an ALE experiment is the length of time for which populations of strains are evolved. Generations is the most commonly used metric for evolutionary time, though alternatives exist that include more data on the selective protocol employed, such as Cumulative Cell Divisions, which factors in population bottlenecks resulting from serial passage (Lee et al., 2011; Sandberg et al., 2014). Most ALE studies have a duration of 100-500 generations of growth (Supplemental Figure 1B), though significantly reduced (<50 generations) and increased (>60,000 generations) timescales have also seen use. The decision to stop an ALE experiment after a certain number of generations is always somewhat arbitrary - though evolved strains get labelled ‘endpoints,’ this does not mean further fitness gains are not achievable with increased experiment duration. The Long Term Evolution Experiment (LTEE) provides the best example of this; *E. coli* has been kept adapting to glucose minimal media for over 30 years, and competition assays on strains spanning 60,000 generations of evolution revealed that fitness gains best fit a power-law model, indicating a decreasing rate of improvement over time but no asymptotic limit (Lenski et al., 2015). Despite fitness gains being achievable indefinitely, for practical applications it is important to strike an appropriate balance between acquiring a fit endpoint and not wasting resources drawing out an experiment excessively. For example, Hua et al. evolved *E. coli* onto lactate growth and found ~60% fitness improvements with less than 250 generations of ALE, while increasing experiment duration to 900 generations only resulted in a further fitness gain of less than 20% (Hua et al., 2007).

When an ALE experiment is performed in a way that facilitates easy fitness tracking over time, it is typically sufficient to set an ‘endpoint’ based on subjective valuation of when fitness starts to plateau. If fitness-tracking is not possible, the studies collected herein establish ALEs of a several hundred generation timescale as suitable for most purposes. Further, modeling approaches have been developed to optimize ALE experimentation and best decide passage protocols and when to end an experiment (Bittiin et al., 2017; LaCroix et al., 2017). Such approaches will better enable automation and optimization of ALE use in the strain design process.

4. ALE Protocols and Experimental Setup
The evolutionary environment employed in an ALE experiment determines the selective pressure guiding adaptation, and is thus a critical choice when a particular phenotypic outcome is sought. Batch culturing is both the simplest and most popular ALE method (Supplemental Figure 1C), though there are significant differences between various batch culturing methods. Often it is desirable for fitness and growth rate to be equivalent, and serial batch culture propagation is a way to enforce this if cultures are passed while still in exponential phase and with excess nutrients. Without limited resources or shifts in growth phase between lag/exponential/stationary, improvements to maximum growth rate determine which strains come to dominate the population. Unfortunately, as growth rate increases over the course of the ALE, changes must be made either to the passage frequency or passage volume used for propagation if stationary phase is to be avoided (Charusanti et al., 2010; Fong and Palsson, 2004). Increasing passage frequency can be prohibitively difficult without automation, while decreasing passage volume leads to tighter population bottlenecks and a potential loss in adaptive mutations. Given these issues, many researchers opt for batch culture propagation of fixed volumes at fixed intervals (generally once per day), as in the LTEE (Lenski et al., 1991). This is a more easily maintainable experimental setup, but by virtue of going through cycles of resource excess followed by depletion selection occurs for more than just growth rate: decreases in lag phase and survival in stationary phase also significantly impact fitness. With a dynamic environment such as this, selection can also lead to co-existing specialist strains that thrive in different phases of the daily cycle (Rozen et al., 2009).

Non-batch ALE culturing setups are dominated by chemostat use, but other methods also exist and are practiced. Chemostats are the second most popular culturing technique for ALEs, and this method differs from batch culturing in ways that alter the selective forces at play. Chemostats culture cells in a bioreactor, enabling tight control over environmental parameters like pH and oxygenation, and maintain continuous growth via a constant influx of media at a set dilution rate (Gresham and Dunham, 2014). Although this can facilitate undesirable adaptive events, such as bacterial persistence due to adhesive wall growth (Rao and Rao, 2004), and can be more difficult to maintain multiple replicates in parallel, it can also be used to select for
particular phenotypes. For example, Koppram et al. evolved yeast for improved biomass inhibitor tolerance via both batch and chemostat culture, and found that batch culture selected more for improvements in growth rate, glucose consumption, and ethanol productivity, while chemostat evolution favored improved inhibitor conversion rates (Koppram et al., 2012). Other culturing methods have seen occasional use in ALE experiments for specific purposes, such as long term culturing in a single flask to select for a growth advantage in stationary phase (Westphal et al., 2018), or growth on an oversized petri dish with antibiotic gradients to allow spatiotemporal tracking of competing lineages (Baym et al., 2016). Serial propagation of colonies streaked onto plates, by virtue of the clonal bottlenecks at each step, is not a form of ‘adaptive’ evolution without some secondary screening measure influencing the transfer process, but can be used for mutation accumulation studies (Shibai et al., 2017).

5. Automation and ALE

Although ALE experiments can be performed manually without significant difficulty and require little setup or specialized equipment, automating the process can facilitate greatly increased experimental capabilities, throughput, and clarity in outcome. In terms of expanding capabilities while running an ALE experiment, automation allows for complex culturing procedures which cannot be reasonably performed manually (e.g., propagating cultures many times a day, or splitting many independent cultures at different thresholds) as well as for dynamic alterations to passage protocols (e.g., alternating substrates or temperature from flask to flask) (Sandberg et al., 2017; Wong et al., 2018). Automation also allows for improved monitoring of physiological properties of evolving cultures, enabling researchers to better understand the evolutionary dynamics at play and to make informed decisions on when and how many times to sequence for adaptive mutations. Furthermore, eliminating repetitive manual ALE passage work saves time and effort and reduces researcher fatigue, which can lead to mistakes and premature stoppage of an experiment where a mutational jump in fitness is still possible. Perhaps the most important aspect of automated ALE is the ability to efficiently execute multiple independent replicates of the same experiment at one time.
With this ability to “play the same game” repeatedly, evolved genotypes can be more easily interpreted by leveraging cross-replicate mutational comparisons (e.g., identifying genetic hitchhikers which lack statistically significant independent recurrence), and multiple equivalent fitness-optimized clones are obtainable from which a strain with desirable properties can be selected, spanning the range of intracellular states enabling improved growth (e.g., a high uptake rate vs. a high yield clone).

Automation can speed up the time it takes to obtain a desirable strain by virtue of less stringent population bottlenecks and sustained exponential-phase growth, translating to faster mutant fixation. The tight environmental control of automated systems can also obviate an important issue intrinsic to manually performed batch culture propagation (Gresham and Dunham, 2014), namely the difficulty in maintaining a strict selection pressure that evolves cells under essentially invariant conditions. For example, automation has been used to select for growth rate improvement in a 2,000 generation ALE in which the cells never experienced nutrient limitation or left exponential growth phase, all while maintaining large passage volumes (LaCroix et al., 2014). When contrasted with the manually-performed Long Term Evolution Experiment (Lenski et al., 1991), this automated ALE saw equivalent fitness gains on a real-world timescale nearly 10 times shorter.

A number of automated ALE devices have been developed that span the range of culturing methods: large volume batch culture (Sandberg et al., 2014; LaCroix et al., 2014; Wong et al., 2018), microtiter plates (Horinouchi et al., 2014; Radek et al., 2017), chemostats/turbidostats (Blaby et al., 2012; Toprak et al., 2013), or microfluidics (Rotem et al., 2016). Such demonstrations establish the broad scope of technologies which can be part of automated strain construction workflows (Chao et al., 2017). Putting some of the current capabilities of such automated systems in context, gains in throughput enabled by automation allowed for more than 30 independent replicates to be evolved simultaneously on the same robotic system, or for large culture volumes to be propagated multiple times per day, minimizing the risk of missing out on adaptive mutations due to population bottlenecks (LaCroix et al., 2017). With real-time tracking of growth response to environmental particulars, automation allows a ‘sweet spot’ of chemical concentration to be maintained, such as keeping cultures in a constantly
growth-inhibiting but non-lethal antibiotic environment to study mechanisms of
resistance evolution (Toprak et al., 2013). This ability to modify an ALE environment on
the fly opens up a number of experimental possibilities - tolerance to certain chemicals
can be incrementally built up, resulting in optimized bioproduction strains (Mohamed et
al., 2017), or the growth environment can be rapidly alternated to yield strains with
desirable diauxic phenotypes (Sandberg et al., 2017). The information needed to
develop inexpensive, automated culturing systems is also becoming publicly available
(Wong et al., 2018), contrasting with systems using industrial-grade hardware (LaCroix
et al., 2014), and will doubtlessly lead to more widespread and innovative adoption of
ALE as a powerful technique for both biological discovery and metabolic engineering.

6. Conclusions and Future Directions

Adaptive laboratory evolution stands as a powerful tool available to strain
engineers and biologists, yielding strains with phenotypic improvements that stem from
random mutation and natural selection rather than rational engineering. This renders
ALE a multi-functional technique: 1) it enables laboratory investigation of evolutionary
processes, expanding the biological knowledge base that informs successful strain
engineering through identification of beneficial mutations associated to a specific
selection pressure, or those which counteract production goals; 2) it complements
rational design approaches by improving engineered strain properties, countervailing
the fitness defects frequently introduced by genetic manipulation; and 3) it can replace
rational design approaches when a desired phenotype is selectable due to a fitness
benefit in a particular environment. Furthermore, with a sizable and well-annotated
collection of causal mutations, phenotypes of interest could be designed by directly
engineering ALE-identified alleles into a strain to induce a particular cellular state. ALE
can thus be of great utility in the Design-Build-Test-Learn cycle employed in strain
construction (R. Liu et al., 2015), augmenting or even replacing the Design and Build
steps (Figure 3).
Figure 3: **ALE for use in the Design, Build, Test, Learn cycle.**

A) The typical Design, Build, Test, Learn cycle used in metabolic engineering to generate a strain with a desired property. B) Augmentation of the cycle where ALE is included in the Build step to rescue a strain that has decreased fitness due to a perturbation, or to optimize a strain after removal or addition of genetic content. C) Augmentation of the cycle where a collection of mutations (e.g., ALEdb (Phaneuf et al., 2018)) associated to a particular phenotype is leveraged for the Design step. D) Augmentation of the cycle where ALE can be used to completely replace the Design and Build steps and a desirable strain is acquired directly from ALE when a phenotype can be tied to selection without engineering.

ALE can serve as a powerful tool, but several principles constrain such experiments. First and foremost, ALE only has efficacy with achieving evolutionarily selectable phenotypes. Growth rate improvements are inherently selectable, making tolerance and substrate-utilization ALEs straightforward approaches for phenotypic
improvement, but desired phenotypes can be fitness-counterproductive for cells (e.g.,
overproduction of some metabolite), preventing evolution from selecting for such traits.
In these cases, approaches must be taken to alter the environment and/or genome of
the starting strain in a way that links the property of interest to a growth advantage, e.g.
via metabolic growth-coupling, or by using a growth environment where the otherwise-
wasteful metabolite production now confers a benefit. Additionally, the choice of
organism is key; model organisms are genetically tractable and have the
knowledgebase with which to interpret mutational results, but more rarely-used species
can have atypical physiological properties that make them ideal for a particular
phenotype. The overall goals of the study are paramount to consider when selecting
ALE experimental parameters, such as replicate number, evolutionary duration, serial
passage bottleneck size, etc. – e.g., if a particular strain phenotype is all that’s desired
then replicates are less important, but if one wants to decipher molecular mechanisms
then replicates greatly facilitate identifying regions of convergent evolution. Fortunately,
tools exist to aid in the selection of parameters for optimizing results from an ALE
experiment without wasting time or resources (LaCroix et al., 2017).

Advances in genetic engineering techniques are facilitating ambitious and novel
strain designs, which can benefit from ALE as a method by which to understand
limitations in these highly perturbed systems and uncover fitness-restoring mutations.
Some prominent examples of studies enabled by advanced genetic engineering include
synthetic minimal cells (Hutchison et al., 2016) and genome-wide codon repurposing
(Wannier et al., 2018). ALE can aid in these efforts by mitigating the fitness defects of
such highly altered strains, as demonstrated in the case of an E. coli strain with nearly
25% of its genome removed (Choe et al., 2019). Additionally, improvements in strain
engineering techniques allow for particular phenotypes to be linked to other properties,
such as biosensors that cause cellular fluorescence upon production of a desired
compound. Biosensors hold great promise as a complementary ALE technology, but
advancements are needed in their dynamic ranges and saturation points, compatibility
across organisms, and evolutionary robustness to ‘cheaters’ that purge the sensing
components, all of which currently limit efficacy in host strains (Williams et al., 2016).
Nevertheless, biosensors enable enrichment for traits other than simple growth rate
when the selective environment is properly designed (Mahr et al., 2015). Modifications
to the ALE environment can also select for complex phenotypes, such as the evolution
of multicellularity in a unicellular yeast as a result of predator-induced selection against
small or individual cells (Ratcliff et al., 2015). The inherent dependence of ‘fitness’ on
the particulars of the evolutionary environment makes ALE capable of selecting for
diverse and novel traits.

The growing body of ALE work and the increased throughput of studies driven by
automation necessitate bioinformatics pipelines and databases to fully realize its
potential as an efficient tool in the strain engineering process. The ability to generate
hundreds of relevant clones and populations in any given experiment (which can
contain thousands of unique mutations), and the ease and cost-efficiency with which
they can be sequenced, makes mutational analysis and data handling non-trivial. In
addition, new technologies are enabling increased clarity on genetic alterations, such as
long-read sequencing that reveals genomic rearrangements and expansions difficult to
discover with shotgun sequencing alone (Pollard et al., 2018). There is a need to
coalesce the growing body of studies (~10^5), relevant strains (~10^4), and mutations
(~10^5-6) derived from ALE to learn the overarching principles and mutational types that
drive phenotypic changes. Work towards meeting this challenge has appeared in the
form of a database which collates ALE-derived mutations across a variety of species,
growth environments, and selection schemes (Phaneuf et al., 2018). Such mutation
collections essentially represent “parts lists” for cell engineers (Fig. 3C), and are ripe for
computational probing with Big Data techniques that can extract salient evolutionary
features. Inclusion of mutations identified in environmental and clinical isolates can also
expand this rich source of allelic parts for strain engineering. The grand promise of such
collections is that ALE wet lab experiments may not be required to engineer a desirable
trait - mutations could be mined directly from a structured database with query tools to
isolate alleles which convey desired physiological features.

In conclusion, ALE has facilitated a number of strain engineering efforts, and the
rapid improvements in complementary technologies leave it poised for accelerated
adoption, improved utility, and establishment as an essential method in the metabolic
gineer’s experimental toolbox. Causal mutations identified in evolved strains serve
both as a way to interpret the biochemical method by which fitness and any associated
production or growth-coupled property was increased, and as genetic parts that can be
introduced into related strains to induce a desired phenotype; importantly, these
mutations are also patentable and can form the basis for intellectual property
generation.

Acknowledgements
We would like to thank Zachary A. King, Nathan E. Lewis, and John D. Trawick for
informative discussions and insights related to the review. Funding for this work was
provided by the Novo Nordisk Foundation through the Center for Biosustainability at the
Technical University of Denmark under Grant NNF10CC1016517. Funding was also
provided by the NIH National Institute of Allergy and Infectious Diseases under Grant
no. U01AI124316. TES was supported in part through the National Science Foundation
Graduate Research Fellowship (grant DGE1144086).

Supplementary Files
Supplemental Table 1 - A collection and categorization of ALE studies relevant to
industrial biotechnology. Excel file .xlsx
Supplemental Figure 1 - Categorization of species, timescales of evolutions studies,
and experimental culturing setups used in ALE studies.

References
Adamo, G.M., Brocca, S., Passolunghi, S., Salvato, B., Lotti, M., 2012. Laboratory evolution of
11-1
Adamo, G.M., Lotti, M., Tamás, M.J., Brocca, S., 2012. Amplification of the CUP1 gene is
associated with evolution of copper tolerance in Saccharomyces cerevisiae. Microbiology
158, 2325–2335. https://doi.org/10.1099/mic.0.058024-0


Spatiotemporal microbial evolution on antibiotic landscapes. Science 353, 1147–1151. 
https://doi.org/10.1126/science.aag0822


https://doi.org/10.1103/PhysRevLett.118.028102

https://doi.org/10.1128/AEM.05773-11

https://doi.org/10.1128/AEM.04144-14


https://doi.org/10.1016/j.ymben.2011.01.008


https://doi.org/10.1371/journal.pgen.1003427


https://doi.org/10.1371/journal.pgen.1003232


https://doi.org/10.1371/journal.pgen.1001186

https://doi.org/10.1371/journal.pone.0033727


Fu, W., Guðmundsson, O., Paglia, G., Herjólfsson, G., Andrésson, O.S., Palsson, B.O.,


https://doi.org/10.1002/bit.23346


https://doi.org/10.1111/1751-7915.12465


https://doi.org/10.1038/ng.280


https://doi.org/10.1016/j.ymben.2012.04.001


https://doi.org/10.1186/1475-2859-13-37


https://doi.org/10.1016/j.femsyr.2005.04.004


Single-cell copy number variant detection reveals the dynamics and diversity of adaptation. PloS Biol. 16, e3000069. https://doi.org/10.1371/journal.pbio.3000069


Palmer, A.C., Kishony, R., 2013. Understanding, predicting and manipulating the genotypic


1. Food Chem. 66, 9061–9069. https://doi.org/10.1021/jacs.jafc.8b01313
**Highlights**

- Rational approaches to strain design are frequently insufficient to achieve desired phenotypes and production capabilities.
- Adaptive Laboratory Evolution (ALE) leverages natural selection to optimize strain properties and obtain valuable phenotypes without necessary a priori knowledge.
- Studies utilizing ALE for metabolic engineering are collected and categorized.
- ALE experimental design, complementary technologies, and future directions are discussed.