

Highly multiplexed mammalian metabolic engineering with a shotgun approach

Corresponding Author: Professor Jef Boeke

Parts of this Peer Review File have been redacted as indicated as we could not obtain permission to publish the reports from one of the peer reviewers.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

SUMMARY

This manuscript by Trolle et al. introduces Shotgun Genetic Engineering (SGE), a novel high-throughput approach for metabolic engineering in mammalian cells, though it is almost certainly generalizable to other engineering objectives. SGE addresses a fundamental challenge in mammalian cell engineering: the difficulty of screening large, multi-gene pathways due to the inverse relationship between DNA construct size and delivery efficiency. SGE accomplishes this by via pooled delivery of individually barcoded transcription units (TUs) at high multiplicity of infection, turning each cell in the pool into an independent experiment that tests a unique pathway combination. This is followed by a selective pressure to enrich productive solutions — in this case cells are grown in amino acid deficient media to identify auxotrophic phenotypes. The authors show engineering biosynthesis of valine (an essential AA), achieving near-WT growth rates for valine prototrophy in CHO cells, outperforming previous rationally engineered solutions. The authors also demonstrate isoleucine prototrophy in mammalian cells, which is previously unreported. Finally, the authors extend their demonstration to Jurkat cells, enabling cells to better survive in nutrient poor and stressful conditions such as solid tumor microenvironments, which has future implications for therapeutic cell engineering. It is notable that the “winning” solutions from the authors’ screens typically required 23-52kb of integrated DNA, a scale that is impractical to screen using most current rational engineering approaches.

The manuscript is very well written and structured, and the results are technically rigorous and clearly demonstrate the utility of SGE. The extension of their approach to Jurkat cells demonstrates the adaptability to other cell types, and the authors do a nice job of highlighting a broad array of potential applications that their approach may enable for metabolic engineering, bioproduction, and cell therapy. I believe the manuscript is well suited for publication in NBT, and fits within the scope of the journal. The comments below address several important concerns regarding the design choices that are made in their libraries, clearly describing their methods, and more rigorous validation of their engineered solutions, as highlighted in detail below

MAJOR COMMENTS

1. Key design considerations for the SGE method need to be described in more detail. Since the methods aspect of this work is central to this manuscript, key considerations regarding the delivery vector design, number of cells to target for integration, the target MOI etc., need to be more clearly detailed in order for readers to better understand its generalizability.

a. For instance, in their first library, on page 6, lines 147-148, the authors mention “These components - CDS, promoter, and optionally OLS - are assembled in a single-pot Golden Gate reaction into lentiviral-compatible expression vectors.” What was the rationale for selecting lentiviral transductions over other random integration methods, such as piggyBac? Is this due to the broader host-range for lentiviral transductions over the latter, or was the choice made for a different reason? The authors mention that the method is compatible with piggyBac, but it is unclear to a reader looking to adopt SGE if lentivirus should be a starting point for a first screen. The authors should clearly state which design parameters for their platform are central for maximizing the likelihood of success in their screens, and which options are more flexible, contrasting the strengths and weaknesses of possible options.

b. Similarly, on Page 8, lines 218-219, the authors deliver their library to 2.42M CHO cells at an MOI of 8.8. It is unclear if these were calculated based on the size of their library and hypothesized solution space, or if they represent an attempt to

get as high an MOI as possible. It would be beneficial to readers to have a clearer understanding of the target MOI range and number of transduced cells given that cell type, packaging efficiency, and any other considerations that may influence these numbers.

c. The authors use Monte Carlo simulations to estimate the proportion of cells that will contain a given number of genetic variant combinations as a function of MOI (e.g. lines 236-248, Extended Data 1C). These simulations could be a powerful means to estimate the numbers mentioned above and could be helpful in giving readers a clear starting point for designing their own experiments.

2. The design criteria used by the authors for library construction needs to be more thoroughly described. While the authors convincingly demonstrate the ability to screen for novel prototrophic phenotypes, a priori considerations for the genetic parts selected for these screens, which I assume would be a critical consideration, are lacking. Some examples of this include:

a. Organelle localization signals (OLSs). The authors included a single mitochondrial localization signal, or no signal, as design choices. Across all three of their engineered prototrophy demos, the optimized solutions all contained the mitochondrial localization signal. However, the authors do not provide a rationale for the inclusion of the mito OLS, and even refer to the inclusion of OLSs as optional (line 147), despite papering to a part of the solution. Could more OLSs for other organelles have been useful to include in the libraries, or were the mito OLSs selected given the hypothesized evolutionary origins of the mitochondria from bacteria, as the authors discuss later? Regardless, the basis of these design choices should be explained in more detail.

b. Promoter choices. The authors decided to use a high expression (Ef1a) and medium expression (PGK) in their screens, without apparent justification or regard to the genes being tested. Could lower expression promoters be helpful if any of the candidate genes in the screen may be important for pathway function but potentially toxic to the cells?

c. Number of combinatorial variants to screen. The authors chose a set of 2 promoters, 2 OLS signals, and 10 genes, for a total of 40 variants. How did the authors determine whether this library size would generate sufficient combinatorial and functional diversity when randomly integrated at their chosen MOI? Could the capacity of the screen have accommodated a larger library size, given other considerations (such as the number of transduced cells at their given MOI)? The authors should discuss the strategy they used for optimizing the target library size, and how transduced cell numbers, MOIs, and transduction efficiency of the target cell type factor into this strategy.

d. The candidate genes included in the screen. In lines 197-199, the inclusion of gene *ilvA* is adequately justified, as it is required to facilitate 2-ketobutyrate production. However, more justification is needed for why two particular variants of *ilvA* were chosen, beyond what is included in Table 1. For instance, were these specific feedback-insensitive variants (L447F with complete insensitivity and L481F with partial insensitivity) selected to test whether different degrees of feedback inhibition would affect pathway performance? Were they chosen based on prior literature demonstrating their effectiveness in other systems? If so, the results section should discuss these findings and contextualize their impact on the screen.

3. The genotypic and biological characterization of the isolated clones should be expanded. While the authors convincingly demonstrate the phenotypes of their engineered clones across all three cell types, showing robust growth in amino acid-free media and confirming endogenous amino acid production through isotope labeling, the underlying genetic and molecular characterization remains incomplete. It is important to more comprehensively characterize the genotype-to-phenotype relationship of their engineered strains in order to form an understanding of the mechanistic basis for how the strains achieve the observed growth rates and aa production levels. Key information currently lacking from their analysis include:

a. Integration site analysis. While authors report high-copy random lentiviral integration of 23-52 kb of integrated DNA, they provide no information about where these integrations occurred in the genome, particularly in their clonal isolates that grow and near wild-type levels (e.g., in their valine demonstration in CHO cells). Analysis of integration sites is critical because (1) disruption of active genes could contribute to the observed phenotype, (2) integration near highly expressed genes could benefit from local chromatin accessibility while integration in heterochromatin regions could lead to lower expression, and silencing over time, (3) Clustering of integrations could create local chromatin disruptions disrupting function in contexts outside of simple growth in complete media, and (4) the stability and heritability of the phenotype may depend on integration site context. The authors should perform integration site analysis for, at minimum, the top 1-2 clones from each screen given its central importance here, specifically in the context of their high MOI transductions.

b. Global transcriptomic characterization. The authors should consider performing bulk RNA-seq to compare their engineered strains (particularly the top performers like CHO-Val-D1, CHO-Ile-10-H1) against WT parental cells. This analysis would reveal in engineered cells either altered regulation or larger-scale compensate for the increased metabolic burden from expressing 6-13 foreign genes that potentially contributes to the remarkable growth phenotypes observed by the authors. Further, it would be helpful to characterize the expression levels of the integrated synthetic genes from their pathway. Finally, it may help understand whether mitochondrial-targeted enzymes affect mitochondrial function, since their winning solutions frequently are localized to the mitochondria.

MINOR COMMENTS

1. The resolution on some of the figures is quite low, making it hard to read and interpret results from, such as Fig. 4J. Some figures also contain faint, stray bounding boxes overlapping with labels within the figure panel, such as Fig. 3H and Fig. 4G-I (only visible when zoomed in).

2. On Page 4, lines 86-87, the statement "some metabolic engineering approaches incorporate in silico predictions of metabolic networks to prioritize genetic interventions before testing in the lab" is unclear. Consider rephrasing it for clarity to something along the lines of "...enable computational screening of candidate interventions".

3. On Page 4, lines 94-95, the authors mention the limitations of current modelling approaches for entirely synthetic TUs, and should consider also mentioning entire TU arrays; a further layer of complexity above synthetic TUs that is essential for metabolic engineering applications.

4. On Page 5, line 200, the authors mention the addition of specific gene variants in order to combat the limitations highlighted in the previous paragraph. However, it should be specified which genes were added in order to combat which of

the corresponding challenges, in order to improve the readability of the manuscript.

5. From the statement on Page 5, lines 210-211, ie "Each promoter appeared either standalone or paired with a mitochondrial targeting sequence (MTS), and each version was assigned 3 unique barcodes", it is unclear whether each promoter and MTS had 3 barcodes each, or if each promoter-MTS pair had three barcodes. This is clear from their discussion of the results in Fig. 3H later, but this sentence should be rephrased to clarify this ambiguity.
6. In Extended Data 1A and accompanying results on Page 9, lines 229-232, the authors attribute the underrepresentation of multiple Ef1a containing solutions to lentiviral interference between strong promoters. However, the papers they cite demonstrate this effect only with multiple promoters in a multi-gene array packaged into a single lentivirus, while in their own experiments, every gene was (presumably) packaged as a single gene in a virus, and as such, may not be the explanation for the results they are seeing here. They authors should clarify, or come up with other hypotheses, such as the instability from the expression of multiple genes at a high expression level within the same cell.
7. Does the nomenclature for the clones isolated in Figs. 3 and 4 represent something specific, e.g. Clones C1-C8 vs D1-D8? This should be clarified in either the main text or figure legend, or the authors should consider simplifying the notation to clones 1-16, as is done in Fig. 5.
8. In Fig. 3F, and accompanying description of the results on Page 11, lines 308-312, the authors find that ilvN is expressed in 94% of all clones, "suggesting that it supports valine biosynthetic function without being strictly necessary". They further report that this may be due to the fact that ilvN is reported to be a regulatory unit that "boosts ilvB catalytic rates and substrate specificities but that ilvB can perform its catalytic function in the absence of ilvN". This is an interesting finding that the authors should elaborate further on. Given its high occurrence in their data (94%), it seems that ilvN may be providing a significant boost to ilvB's catalytic rate. Is this supported by literature? What are the numbers, if they are known?
9. The legend for Fig. 3G is lacking and needs to be expanded. What do the "median normalized read counts" refer to? What are they normalized to? A percentage of total reads, or internally normalized read values for each clone? Expanding the figure legend will help improve clarity and also provide a means for how to compare the authors' numbers across the different combinations (ie, is a score of 10 twice as good as a score of 5? How does that compare to the prototrophy score previously defined? Or are these results independent of those, since this is performed on only high prototrophy variants?)
10. The results in Extended Data 6A (dual prototrophic variants grown in ile-/val- media) should be contrasted with their results from Figs. 3 and 4 respectively, in more detail. On Pages 16-17, lines 467-470, the authors state "we noticed that CHO-Val clones and CHO-Ile clones exhibited differential underlying CDS usage suggesting that analyzing both populations on aggregate might conflate signal given that each population was not specifically selected to be doubly prototrophic (Extended Data 6C)". While this may be true, some insights could still be gleaned about the utility of genes that enable single vs dual prototrophy. A description of these differences, along with a hypothesis for what genes help enable these functional phenotypes could be beneficial for the readers.
11. Page 18, lines 526-527. Potential typo "...proceeded to selection for..." should be replaced with "proceeded with selection for..." or "proceeded to select for..."
12. The analysis in Fig. 5E is done by removing the duplicate "siblings", in order to generate a strong composite barcode fingerprint for prototrophic solutions. However, the presence of these siblings may also be confounding the analysis of CDS content presented in Fig. 5D, and should potentially be re-done by removing duplicate siblings as well.
13. The description of results from the ML model presented in Fig. 5G, H requires more clarification and elaboration. Specifically:
 - a. What were the inputs to the ML model, and how were the data pre-processed and presented to the model? What was the model's output (I'm assuming probability distribution for the two output classes)?
 - b. In total, the model accurately classified 112/119, which is a convincing demonstration of the model's accuracy. However, the authors should analyze the handful of incorrect variants, and generate some hypotheses for why the model may have failed to classify them accurately, given its exceptional performance on the entire dataset otherwise. For instance, it may be beneficial to know if the clones that were incorrectly predicted contained features that didn't match those with the highest importance score.

(Remarks on code availability)

The files the authors cite in their manuscript appear to be in the Github repository and in order.

Reviewer #3

(Remarks to the Author)

In the manuscript entitled "A shotgun approach for highly multiplexed mammalian metabolic engineering" by Trolle et al. the authors present an innovative, systems level approach to reconstitute a valine biosynthetic capacity in mammalian cells. This enables a metabolic bypass potentially simplifying culture conditions, while also exposing a new experimental system to study essential amino acid metabolism in mammalian cells. The experimental framework is thoughtfully constructed, and the focus on screening for "winner" clones capable of essential amino acid biosynthesis is well-motivated. We commend the overall ambition and conceptual direction of the study, but we feel there are items, critical to the claims of this work, that should be considered and likely addressed prior to publication.

Major concerns:

The current screening library appears to be preselected and limited in complexity (4 per promoter/OLS, 3 barcodes per CDS across 10 genes), which seems to contradict the authors' claim of "screened millions" (Lines 28, 639). This overstatement is misleading and undermines the credibility of the study. The limited library size constrains scalability, compromises the generalizability of the findings, and raises concerns about the overall robustness of the screening platform.

Moreover, viral vectors are inherently prone to recombination, particularly in complex pooled libraries, as demonstrated in early versions of Perturb-seq (PMID: 29457792). Such recombination events can disrupt transcriptional unit (TU) and barcode pairing, thereby reducing screen sensitivity. It is therefore critical to incorporate mitigation strategies. For example, long-read sequencing can verify expression cassette and barcode integrity, or 5' plus 3' barcoding schemes can improve accuracy. The authors should also follow established best practices from the CLASSIC (PMID: 36993481) and MPRA literatures (PMID: 37103023).

The manuscript introduces machine learning (ML) as part of the analytical framework, but there are concerns here. First, the model and dataset were not adequately described to understand if the dataset was large enough and structured to avoid overfitting, data leakage, or if it is generalizable.

The barcodes of EF1a-driven transcriptional units (TUs) appear to suffer from insufficient recovery, likely representing a technical bottleneck caused by suboptimal PCR enrichment. Notably, over 50% of clones containing the GFP transcriptional unit exhibited high basal expression, as shown in Figure 3F. This suggests an elevated signal-to-noise ratio and prompts a key question: does *ilvA*-L481F reflect a genuine biological signal, or is it merely background noise? Fine-tuning the PCR enrichment primer set may substantially improve barcode capture efficiency, thereby enabling more scalable and sensitive screening.

The current version of the manuscript lacks sufficient system-wide, functional validation of the SGE methodology. Incorporating RNA-seq profiling of the selected clones would significantly strengthen the study. This approach would not only confirm transgene expression and complement genomic copy number analysis but also offer valuable insights into the biological effects of the perturbations, thereby deepening our understanding of essential amino acid biosynthesis in mammalian cells.

Minor concerns:

Provide full details on all reporting materials, including plasmid construction, and CDS/gene sequences.

Show library distribution using a histogram where X-axis is Log2-transformed read counts (barcode abundance) and Y-axis is frequency (number of barcodes).

Provide raw barcode counts across LV virus, pre-infection, and post-harvest samples (e.g., Extended Data 1 and 7).

Report individual barcode counts per TU and display library distribution as a histogram (X-axis: log2 barcode read count, Y-axis: frequency).

Please clarify the actual number of promoter/OLS used in the screening library (ie. 3 or 4).

There was an inconsistency between GFP signal presentation in Figures 3F and 3G. We suggest including a side-by-side heatmap of GFP TUs with valine prototrophy scores (e.g., Val Score >0.75) to help disentangle functional hits from background expression.

(Remarks on code availability)

Author Rebuttal letter:

Response to Reviewers

We thank the two reviewers for their careful and constructive reading of our manuscript, and for their general enthusiasm regarding the content and its presentation. In the point-by-point response that follows, reviewer comments are replicated in full in blue text, with our responses provided in-line in black text.

Reviewer #1

This manuscript by Trolle et al. introduces Shotgun Genetic Engineering (SGE), a novel high-throughput approach for metabolic engineering in mammalian cells, though it is almost certainly generalizable to other engineering objectives. SGE addresses a fundamental challenge in mammalian cell engineering: the difficulty of screening large, multi-gene pathways due to the inverse relationship between DNA construct size and delivery efficiency. SGE accomplishes this by via pooled delivery of individually barcoded transcription units (TUs) at high multiplicity of infection, turning each cell in the pool into an independent experiment that tests a unique pathway combination. This is followed by a selective pressure to enrich productive solutions — in this case cells are grown in amino acid deficient media to identify auxotrophic phenotypes. The authors show engineering biosynthesis of valine (an essential AA), achieving near-WT growth rates for valine prototrophy in CHO cells, outperforming previous rationally engineered solutions. The authors also demonstrate isoleucine prototrophy in mammalian cells, which is previously unreported. Finally, the authors extend their demonstration to Jurkat cells, enabling cells to better survive in nutrient poor and stressful conditions such as solid tumor

microenvironments, which has future implications for therapeutic cell engineering. It is notable that the “winning” solutions from the authors’ screens typically required 23-52kb of integrated DNA, a scale that is impractical to screen using most current rational engineering approaches.

The manuscript is very well written and structured, and the results are technically rigorous and clearly demonstrate the utility of SGE. The extension of their approach to Jurkat cells demonstrates the adaptability to other cell types, and the authors do a nice job of highlighting a broad array of potential applications that their approach may enable for metabolic engineering, bioproduction, and cell therapy. I believe the manuscript is well suited for publication in NBT, and fits within the scope of the journal. The comments below address several important concerns regarding the design choices that are made in their libraries, clearly describing their methods, and more rigorous validation of their engineered solutions, as highlighted in detail below.

We thank this reviewer for their excellent summary of the original manuscript, as well as their kind words regarding the content, structure and relevance of the study.

MAJOR COMMENTS

1. Key design considerations for the SGE method need to be described in more detail. Since the methods aspect of this work is central to this manuscript, key considerations regarding the delivery vector design, number of cells to target for integration, the target MOI etc., need to be more clearly detailed in order for readers to better understand its generalizability.

This is a terrific point, and we agree that these details should be included in the manuscript to facilitate broad adoption. We have now provided the annotated vector sequence in the github repository associated with the paper: <https://github.com/julietrolle/shotgun-genetic-engineering>. We also include all promoter and coding sequences in Supplementary Table 8. We also address the specific comments regarding experimental details in each of the sections below.

a. For instance, in their first library, on page 6, lines 147-148, the authors mention “These components - CDS, promoter, and optionally OLS - are assembled in a single-pot Golden Gate reaction into lentiviral-compatible expression vectors.” What was the rationale for selecting lentiviral transductions over other random integration methods, such as piggyBac? Is this due to the broader host-range for lentiviral transductions over the latter, or was the choice made for a different reason? The authors mention that the method is compatible with piggyBac, but it is unclear to a reader looking to adopt SGE if lentivirus should be a starting point for a first screen. The authors should clearly state which design parameters for their platform are central for maximizing the likelihood of success in their screens, and which options are more flexible, contrasting the strengths and weaknesses of possible options.

As the reviewer suggests, we initially selected lentivirus for its broad host range and reasoned that it would allow us to target multiple cell types with the same library. However, there are a few downsides to using lentivirus, including the need to perform packaging after library cloning, limited packaging capacity, and the potential for barcode swapping. For these reasons, a transposon-based strategy otherwise based exactly on the SGE strategy may be more suitable in some contexts. We have added the following text to the paper to help readers navigate this decision.

“We initially selected lentivirus because of the ease of titrating MOI and its broad host range, which allowed us to deploy the same libraries across multiple cell types. However, lentiviral delivery also comes with several practical limitations, including the need to perform packaging after library cloning, constraints on cargo size, and potential template switching that can lead to barcode swapping²⁷. For some applications, especially when working with larger constructs or when maintaining barcode linkage is essential, a transposon-based strategy (e.g., piggyBac) may therefore be more suitable.”

b. Similarly, on Page 8, lines 218-219, the authors deliver their library to 2.42M CHO cells at an MOI of 8.8. It is unclear if these were calculated based on the size of their library and hypothesized solution space, or if they represent an attempt to get as high an MOI as possible. It would be beneficial to readers to have a clearer understanding of the target MOI range and number of transduced cells given that cell type, packaging efficiency, and any other considerations that may influence these numbers.

Our target of ~2.5M infected CHO cells and the achieved MOI of 8.8 were chosen to ensure adequate sampling of the combinatorial design space in an experimentally feasible manner rather than simply maximizing MOI. We also aimed to exceed an MOI of 4 to ensure that the majority of cells received ≥ 4 integrations, thereby allowing effective sampling of 4-TU integration genotypes, which we knew was required for valine prototrophy. As might be expected, our simulations indicated that higher MOIs substantially improve coverage of multi-integration designs. We add the following text to the revision to clarify these considerations for readers.

“Our choice of 10 CDSs, 2 promoters, and 2 OLSs (40 total TUs) reflected a balance between biological relevance and practical sampleability. As we knew 4 genes to be sufficient to establish a valine prototrophic phenotype, we at minimum wanted to sample all possible 4-TU combinations of which there are 123,410 from a 40-member library. However, we were interested in sampling variation in copy number as well,

motivating a target of sampling all 5-TU combinations of which there are 962,598. This design space can be sampled within a single 10cm² dish, which we reasoned was appropriate for a pilot-scale demonstration of the approach. Even modest expansions to the library, for example, adding one additional promoter or OLS to yield a 60-TU library, would increase the number of possible 5-TU combinations to ~7M, making comprehensive sampling substantially more challenging in practice. The effective extent to which this combinatorial space is explored can be modeled using a Monte Carlo simulation, as described below.”

c. The authors use Monte Carlo simulations to estimate the proportion of cells that will contain a given number of genetic variant combinations as a function of MOI (e.g. lines 236-248, Extended Data 1C). These simulations could be a powerful means to estimate the numbers mentioned above and could be helpful in giving readers a clear starting point for designing their own experiments.

We thank the reviewer for this important point. Monte Carlo simulations are indeed a powerful way to estimate the key parameters that govern SGE performance and provide a useful starting point for others when designing their own experiments. To make this explicit for readers, we have added the following paragraph to the Results section:

“More broadly, these simulations illustrate how key experimental parameters such as library size, number of infected cells, and MOI, jointly determine the fraction of the combinatorial design space that can be sampled in a single screen. Therefore, the achievable coverage can be predicted for any SGE experiment using the same simulation framework, providing a useful starting point for selecting library complexity, MOI, and culture scale when designing SGE screens.”

We have also added text to the Methods section to clarify the general utility of this modeling framework and have made the simulation code publicly available:

“These simulations were used not only to quantify diversity in our study but also to provide a general framework for designing SGE experiments. By varying the library size, MOI, and number of simulated cells, one can estimate the achievable coverage of TU combinations under their specific experimental parameters. The analysis code used to run these simulations is available in the GitHub repository associated with this paper: [https://github.com/julietrolle/shotgun-genetic-engineering/tree/main/SGE_Paper_Finalized/Monte Carlo Simulations](https://github.com/julietrolle/shotgun-genetic-engineering/tree/main/SGE_Paper_Finalized/Monte_Carlo_Simulations)”

2. The design criteria used by the authors for library construction needs to be more thoroughly described. While the authors convincingly demonstrate the ability to screen for novel prototrophic phenotypes, a priori considerations for the genetic parts selected for these screens, which I assume would be a critical consideration, are lacking. Some examples of this include:

a. Organelle localization signals (OLSs). The authors included a single mitochondrial localization signal, or no signal, as design choices. Across all three of their engineered prototrophy demos, the optimized solutions all contained the mitochondrial localization signal. However, the authors do not provide a rationale for the inclusion of the mito OLS, and even refer to the inclusion of OLSs as optional (line 147), despite papering to a part of the solution. Could more OLSs for other organelles have been useful to include in the libraries, or were the mito OLSs selected given the hypothesized evolutionary origins of the mitochondria from bacteria, as the authors discuss later? Regardless, the basis of these design choices should be explained in more detail.

This is a fantastic suggestion. We have added the following two passages to the text to better clarify the importance of sub-cellular compartmentalization for metabolic engineering broadly and for BCAA biosynthesis in mammalian cells more specifically.

“Further, it has been demonstrated that when expressing heterologous genes in new host contexts, directing them to the appropriate subcellular compartment can be important for optimal function^{1,25}. In particular, localizing metabolic enzymes to mitochondria has been shown to increase local enzyme concentrations and improve pathway performance²⁵. Peroxisomes have also been harnessed in yeast as insulated compartments for synthetic metabolic pathways²⁶. Given these observations, subcellular localization represents another important parameter to vary. In SGE, we can investigate these effects by pairing promoters with N-terminal organellar localization signals (OLSs), which receive additional identifying barcodes. These components—CDS, promoter, and OLS—are assembled in a single-pot Golden Gate reaction into lentiviral-compatible expression vectors.”

“We chose to include the MTS because previous studies have demonstrated that organellar targeting can enhance pathway function by locally concentrating metabolic enzymes and because BCAA biosynthesis in extant eukaryotic prototrophs such as fungi are localized to mitochondria, suggesting that mitochondria may provide an advantageous biochemical or biophysical setting for BCAA biosynthesis ^{25,36} (Fig. 2C).”

b. Promoter choices. The authors decided to use a high expression (Ef1a) and medium expression (PGK) in their screens, without apparent justification or regard to the genes being tested. Could lower expression promoters be helpful if any of the candidate genes in the screen may be important for pathway function but potentially toxic to the cells?

It is difficult to know a priori whether any given gene will be toxic, but our initial choice of EF1a and PGK was motivated by the goal of optimizing pathway flux, consistent with approaches in metabolic engineering in *E. coli* and yeast. That said, we agree that lower-expression promoters could be advantageous in cases where specific genes or intermediates are toxic, and we have added clarifying text to help readers make these choices:

“These CDSs are synthesized with flanking Type IIS restriction sites to enable directional insertion into a pooled set of pre-assembled vectors carrying a diversity of barcoded promoters. These promoters can drive varying levels of expression, an important parameter to explore given that tuning promoter strength to modulate the expression of individual pathway components is a well-established approach for balancing flux and improving pathway function in metabolic engineering across various organisms^{20–24}. Lower-expression promoters may also be beneficial in cases where specific pathway genes or intermediates are toxic to the host.”

c. Number of combinatorial variants to screen. The authors chose a set of 2 promoters, 2 OLS signals, and 10 genes, for a total of 40 variants. How did the authors determine whether this library size would generate sufficient combinatorial and functional diversity when randomly integrated at their chosen MOI? Could the capacity of the screen have accommodated a larger library size, given other considerations (such as the number of transduced cells at their given MOI)? The authors should discuss the strategy they used for optimizing the target library size, and how transduced cell numbers, MOIs, and transduction efficiency of the target cell type factor into this strategy.

We thank the reviewer for pointing this out - this is an important discussion point.

As discussed in our response to point 1b above, our choice of 10 CDSs, 2 promoters, and 2 OLSs (40 total TUs) reflected both biological considerations and practical constraints on combinatorial sampling. Biologically, the selected 10 CDSs were motivated by prior knowledge of valine biosynthesis together with our aim to explore isoleucine biosynthetic routes, resulting in a focused but diverse set of pathway components. From a combinatorial perspective, given a requirement of at least four integrations to confer valine prototrophy, a 40-TU library contains 123,410 possible 4-TU combinations. At the achieved MOI (>4) and ~2.5M infected cells, this design space can be sampled with sufficient coverage for discovery in a pilot-scale experiment. We note that even modest expansions of the library (e.g., adding one additional promoter or OLS, increasing the library to 60 TUs) increase the number of possible 4-TU combinations to ~7M, which becomes significantly more challenging to sample adequately given the achievable ranges of MOI, transduction efficiency, and cell culture scalability.

We have added text to the manuscript clarifying how library size, cell number, MOI, and transduction efficiency influence the design of SGE screens, and describing how Monte Carlo simulations can be used to estimate these parameters and guide future experiments.

“Our choice of 10 CDSs, 2 promoters, and 2 OLSs (40 total TUs) reflected a balance between biological relevance and practical sampleability. As we knew 4 genes to be sufficient to establish a valine prototrophic phenotype, we at minimum wanted to sample all possible 4-TU combinations of which there are 123,410 from a 40-member library. However, we were interested in sampling variation in copy number as well, motivating a target of sampling all 5-TU combinations of which there are 962,598. This design space can be sampled within a single 10cm² dish, which we reasoned was appropriate for a pilot-scale demonstration of the approach. Even modest expansions to the library, for example, adding one additional promoter or OLS to yield a 60-TU library, would increase the number of possible 5-TU combinations to ~7M, making comprehensive sampling substantially more challenging in practice. The effective extent to which this combinatorial space is explored can be modeled using a Monte Carlo simulation, as described below.”

d. The candidate genes included in the screen. In lines 197-199, the inclusion of gene *ilvA* is adequately justified, as it is required to facilitate 2-ketobutyrate production. However, more justification is needed for why two particular variants of *ilvA* were chosen, beyond what is included in Table 1. For instance, were these specific feedback-insensitive variants (L447F with complete insensitivity and L481F with partial insensitivity) selected to test whether different degrees of feedback inhibition would affect pathway performance? Were they chosen based on prior literature demonstrating their effectiveness in other systems? If so, the results section should discuss these findings and contextualize their impact on the screen.

In our previous paper (Trolle et al., *eLife* 2022), we observed that reducing the isoleucine concentration increased the growth rate of rationally engineered valine-prototrophic CHO cells. We reasoned that this effect might reflect feedback inhibition within the pathway and therefore decided to include feedback-resistant mutants in the library. The *ilvA* variants L447F and L481F were selected based on prior literature (as cited in Table 1), which demonstrated that these mutations confer complete or partial insensitivity, respectively, to feedback inhibition by isoleucine. Because this inhibition occurs at the protein level, we expected these altered regulatory properties to transfer into the mammalian context. We modified the relevant section of the paper to better explain this choice.

“Specifically, we added *E. coli*-derived *ilvA* to address the hypothesized limitation in the availability of the isoleucine-specific substrate 2-oxobutanoate. To address the second challenge of feedback inhibition we included two mutated variants of *ilvA* with partial (*ilvA*-L481F) or complete (*ilvA*-L447F) insensitivity to negative feedback inhibition mediated by isoleucine itself^{31,32,33}. In our previous work, we observed that rationally engineered valine-prototrophic CHO cells (pMTIV) grew faster in media with reduced isoleucine concentrations, which we reasoned reflected strong feedback inhibition within the pathway⁵. Including these variants allowed us to test whether modulating feedback inhibition in a heterologous context could improve pathway function. Because this inhibition operates at the protein level, we reasoned that these variants would retain their altered regulatory properties in the mammalian context. Finally, to potentially improve the catalytic properties of the pathway, we added *ilvG* and *ilvM*, isozymes of *ilvB* and *ilvN*, respectively.”

3. The genotypic and biological characterization of the isolated clones should be expanded. While the authors convincingly demonstrate the phenotypes of their engineered clones across all three cell types, showing robust growth in amino acid-free media and confirming endogenous amino acid production through isotope labeling, the underlying genetic and molecular characterization remains incomplete. It is important to more comprehensively characterize the genotype-to-phenotype relationship of their engineered strains in order to form an understanding of the mechanistic basis for how the strains achieve the observed growth rates and aa production levels. Key information currently lacking from their analysis include:

a. Integration site analysis. While authors report high-copy random lentiviral integration of 23-52 kb of integrated DNA, they provide no information about where these integrations occurred in the genome, particularly in their clonal isolates that grow and near wild-type levels (e.g., in their valine demonstration in CHO cells). Analysis of integration sites is critical because (1) disruption of active genes could contribute to the observed phenotype, (2) integration near highly expressed genes could benefit from local chromatin accessibility while integration in heterochromatin regions could lead to lower expression, and silencing over time, (3) Clustering of integrations could create local chromatin disruptions disrupting function in contexts outside of simple growth in complete media, and (4) the stability and heritability of the phenotype may depend on integration site context. The authors should perform integration site analysis for, at minimum, the top 1-2 clones from each screen given its central importance here, specifically in the context of their high MOI transductions.

We thank the reviewer for this important suggestion. We agree that integration site context can, in principle, influence gene expression, genomic stability, and phenotypic outcomes in high-MOI lentiviral integrations.

However, the primary goal of this study is to use SGE as a discovery framework to identify functional pathway solutions within a large combinatorial search space, which can subsequently be reconstituted on fully defined constructs in a clean genomic background (as now demonstrated for the valine pathway in Fig. 3J-K). As such, we do not rely on SGE-derived clones themselves as final engineered strains.

To address the reviewer’s concern, we have now performed integration site analysis on representative clones across all screens (CHO-Val-D1, CHO-Val-D3, CHO-Ile-10-H1, CHO-Ile-10-H2, and Jurkat-Val-31) using targeted long-read nanopore sequencing with adaptive sampling. In parallel, we performed RNA-seq analysis to assess the impact of TU integration on gene expression. Experimental and analytical details have been added to the revised Methods section, and integration site annotations are provided in Supplementary Table 3 and Supplementary Table 6 for CHO and Jurkat clones respectively. We have also added relevant text to Supplementary Note 1 and reproduce it below for ease of reference:

1) CHO cells:

“We mapped TU integration sites across CHO-Val-D1, CHO-Val-D3, CHO-Ile-10-H1, and CHO-Ile-10-H2 using long-read DNA sequencing. Reads supporting these sites frequently contained target site duplications (TSDs), consistent with lentiviral integration (Supplementary Fig. 1C). Across both valine and isoleucine prototrophic clones, integrations occurred predominantly in intergenic and intronic regions (Supplementary Fig. 1D-E; Supplementary Table 3). Importantly, integration of TUs within endogenous genes was not associated with substantial changes to target gene expression. The only exceptions were *Brinp3*, a mitochondrially localized protein involved in neuronal differentiation⁶¹, and *Ccdc112*, a protein implicated in centriolar organization and sperm mitochondrial sheath formation^{62,63}. *Brinp3* was upregulated in CHO-Val-D1 (in which a TU insertion occurred), but not in CHO-Val-D3, which had no insertion. In CHO-Val-D1, reads mapping to *Brinp3* were observed exclusively downstream of the integrated TU, in the same orientation as the TU transcript, and were restricted to exon 7 rather than spanning the full gene. This pattern suggests that the apparent increase in *Brinp3* expression may reflect transcriptional readthrough from the integrated TU captured in RNA-seq, rather than a bona fide upregulation of the endogenous gene. (Supplementary Fig. 1F-G)

Ccdc112 was marginally downregulated in CHO-Ile-10-H1 (in which a TU insertion occurred) but marginally upregulated in CHO-Ile-10-H2, which had no insertion, despite similar prototrophic phenotypes, further suggesting that these effects reflect local genomic disruptions rather than primary drivers of the respective phenotypes (Supplementary Fig. 1H).”

2) Jurkat cells:

“For Jurkat-Val-31, TU integrations mapped to 7 genomic loci, all within introns. At two loci, we observed tandem TU integrations, a previously documented phenomenon for high MOI lentiviral screens; however, none of the intronic disruptions correlated to notable changes in expression of affected genes (Supplementary Table 6-7)”

b. Global transcriptomic characterization. The authors should consider performing bulk RNA-seq to compare their engineered strains (particularly the top performers like CHO-Val-D1, CHO-Ile-10-H1) against WT parental cells. This analysis would reveal in engineered cells either altered regulation or larger-scale compensate for the increased metabolic burden from expressing 6-13 foreign genes that potentially contributes to the remarkable growth phenotypes observed by the authors. Further, it would be helpful to characterize the expression levels of the integrated synthetic genes from their pathway. Finally, it may help understand whether mitochondrial-targeted enzymes affect mitochondrial function, since their winning solutions frequently are localized to the mitochondria.

We thank the reviewer for this suggestion. While the primary goal of this study is to use SGE as a discovery framework to identify functional pathway solutions rather than to extensively characterize individual clones, we agree that transcriptomic profiling provides useful insight into the effects of TU integration.

To address this, we performed bulk RNA-seq (polyA capture) on representative prototrophic clones across both systems (CHO-Val-D1, CHO-Ile-10-H1, and Jurkat-Val-31), each profiled under nutrient-replete and amino acid dropout conditions (with biological replicates), alongside parental control cells. We have added text describing these experiments to Supplementary Note 1.

Across all comparisons, transcriptomic differences relative to parental cells were modest under both nutrient-replete and dropout conditions (Supplementary Fig. 2A-E; Supplementary Fig. 3A-E; Supplementary Fig. 4A-E; Supplementary Tables 4, 5, 7). Gene ontology (GO) enrichment analysis did not reveal strong or consistent signatures of stress, metabolic compensation, global regulatory rewiring or perturbation of mitochondrial function. Consistent with this, we do not observe a growth defect in engineered clones under nutrient-replete conditions (Extended Data Figs. 2A; 5A; 10B), suggesting that expression of multiple heterologous genes imposes limited metabolic burden in nutrient replete conditions.

We further examined expression of SGE-derived transgenes. As these transcripts are not polyadenylated (to enable lentiviral delivery), their detection by conventional polyA-based RNA-seq aimed at the lentiviral polyA is expected to be reduced relative to endogenous transcripts. Despite this limitation, transgene-derived reads were detected and were consistent with TU barcode assignments from amplicon-seq (Extended Data Figs. 3B,E; 6B,E; 11C,G), supporting the accuracy of barcode-based genotyping and confirming expression of pathway components.

MINOR COMMENTS

1. The resolution on some of the figures is quite low, making it hard to read and interpret results from, such as Fig. 4J. Some figures also contain faint, stray bounding boxes overlapping with labels within the figure panel, such as Fig. 3H and Fig. 4G-I (only visible when zoomed in).

We appreciate the reviewer for their attention to detail - we have fixed all of these issues in the revised figures.

2. On Page 4, lines 86-87, the statement “some metabolic engineering approaches incorporate in silico predictions of metabolic networks to prioritize genetic interventions before testing in the lab” is unclear. Consider rephrasing it for clarity to something along the lines of “...enable computational screening of candidate interventions”.

We have edited this sentence to:

“To overcome the limitations of experimental screening in mammalian systems, some metabolic engineering approaches incorporate computational screening of candidate interventions before testing in the lab”

3. On Page 4, lines 94-95, the authors mention the limitations of current modelling approaches for entirely synthetic TUs, and should consider also mentioning entire TU arrays; a further layer of complexity above synthetic TUs that is essential for metabolic engineering applications.

We have modified the text to the following to incorporate this suggestion:

“As a result, they are poorly suited for generative design, including scenarios involving the introduction of heterologous or entirely synthetic TUs. The design of TU arrays that comprise synthetic metabolic pathways adds yet another layer of complexity, beyond what these models can currently handle.”

4. On Page 5, line 200, the authors mention the addition of specific gene variants in order to combat the limitations highlighted in the previous paragraph. However, it should be specified which genes were added in order to combat which of the corresponding challenges, in order to improve the readability of the manuscript.

We thank the reviewer for this suggestion and have explicitly stated the rationale behind the choice of gene variants for inclusion in the library. We have reproduced the modified sections of the text below.

“However, this same gene set was insufficient to confer isoleucine biosynthesis to CHO cells despite its overlap in enzymatic steps with valine biosynthesis (Fig. 2B). We hypothesized that this might be due to (i) insufficient availability of an isoleucine pathway-specific substrate, 2-oxobutanoate, or (ii) feedback inhibition of early steps in the pathway by pathway products, or (iii) suboptimal catalytic properties of the acetohydroxy acid synthase I (AHAS I) complex encoded by catalytic subunit, *ilvB*, and regulatory subunit, *ilvN*.”

“Specifically, we added *E. coli*-derived *ilvA* to address the hypothesized limitation in the availability of the isoleucine-specific substrate 2-oxobutanoate. To address the second challenge of feedback inhibition we included two mutated variants of *ilvA* with partial (*ilvA*-L481F) or complete (*ilvA*-L447F) insensitivity to negative feedback inhibition mediated by isoleucine itself^{31,32,33}. In our previous work, we observed that rationally engineered valine-prototrophic CHO cells (pMTIV) grew faster in media with reduced isoleucine concentrations, which we reasoned reflected strong feedback inhibition within the pathway⁵. Including these variants allowed us to test whether modulating feedback inhibition in a heterologous context could improve pathway function. Because this inhibition operates at the protein level, we reasoned that these variants would retain their altered regulatory properties in the mammalian context. Finally, to potentially improve the catalytic properties of the pathway, we added *ilvG* and *ilvM*, isozymes of *ilvB* and *ilvN*, respectively.”

5. From the statement on Page 5, lines 210-211, ie “Each promoter appeared either standalone or paired with a mitochondrial targeting sequence (MTS), and each version was assigned 3 unique barcodes”, it is unclear whether each promoter and MTS had 3 barcodes each, or if each promoter-MTS pair had three barcodes. This is clear from their discussion of the results in Fig. 3H later, but this sentence should be rephrased to clarify this ambiguity.

We have edited that sentence and included an example to remove ambiguity:

“We included two constitutive promoters in the library, EF1a (strong) and PGK (medium), to explore varying expression levels of each pathway component as discussed above. Each promoter appeared either standalone or paired with a mitochondrial targeting sequence (MTS), and each promoter-MTS combination was assigned three unique barcodes (e.g., EF1a-no-MTS: 3 barcodes; EF1a-MTS: 3 barcodes).”

6. In Extended Data 1A and accompanying results on Page 9, lines 229-232, the authors attribute the underrepresentation of multiple EF1a containing solutions to lentiviral interference between strong promoters. However, the papers they cite demonstrate this effect only with multiple promoters in a multi-gene array packaged into a single lentivirus, while in their own experiments, every gene was (presumably) packaged as a single gene in a virus, and as such, may not be the explanation for the results they are seeing here. They authors should clarify, or come up with other hypotheses, such as the instability from the expression of multiple genes at a high expression level within the same cell.

We appreciate the reviewer’s point that the studies we cited examined promoter interference within multi-gene arrays packaged into a single lentiviral construct, whereas in our system each TU is delivered on a separate viral particle. We agree that the underlying mechanism may therefore not be identical. Importantly, direct head-to-head comparisons of PGK and EF1a in pooled lentiviral libraries have not, to our knowledge, been systematically explored before, so the underrepresentation of EF1a-linked TUs is itself a novel observation.

Several alternative hypotheses could explain this effect. Because EF1a is a stronger promoter than PGK, recruitment of transcriptional machinery to the EF1a promoter could interfere with transcription from the viral packaging promoter, thereby reducing titer. Another possibility relates to the presence of an intron in the EF1a promoter that is not present in PGK. In our vector, the transgene was placed in the forward orientation relative to the viral transcript despite the theoretical risk of splicing, based on previous work demonstrating that the EF1a intron is typically retained during lentiviral packaging (<https://doi.org/10.1093/nar/gku1312>). Nevertheless, it remains possible that interactions between the splicing machinery and viral packaging machinery contribute to lower titers. Finally, as the reviewer suggests, cellular toxicity or reduced fitness associated with expressing multiple pathway genes at high levels in the same cell could also lead to depletion of EF1a-driven constructs.

To address these points, we have added text to the manuscript outlining these potential mechanisms and noting that further work will be required to determine the precise cause of EF1a underrepresentation.

“Post-infection, we performed amplicon-seq to monitor TU distribution and observed underrepresentation of EF1a promoters both with and without MTS (Extended Data Fig. 1A). Because head-to-head

comparisons of PGK and EF1a in pooled lentiviral libraries have not been systematically examined before, the basis of this effect remains unclear. In our pooled setting, several mechanisms may contribute to the underrepresentation of EF1a-driven TUs. First, because EF1a is a stronger promoter than PGK, its recruitment of transcriptional machinery could interfere with transcription of the viral genome during packaging, thereby reducing viral titer^{37–39}. Second, EF1a contains an intron that is absent from PGK, and although this intron is typically retained during lentiviral packaging, interactions between the intron's splice sites and the viral splicing/packaging machinery could still modestly impair packaging efficiency⁴⁰. Finally, high-level expression driven by EF1a may impose a fitness cost when multiple pathway components are present in the same cell, leading to depletion of EF1a-linked TUs upon infection of the target cells.”

7. Does the nomenclature for the clones isolated in Figs. 3 and 4 represent something specific, e.g. Clones C1-C8 vs D1-D8? This should be clarified in either the main text or figure legend, or the authors should consider simplifying the notation to clones 1-16, as is done in Fig. 5.

CHO clone names such as C1-C8 and D1-D8 refer to the clone's position within a 96 well plate. In order to allow sampling of all infected cells while minimizing the risk of merging with other clones, infected cells were split across 2 x 10-cm² plates prior to selection on low valine or low isoleucine medium. Clones from the two plates were then assigned to separate rows (e.g. C and D) in the 96 well plate.

For CHO-Ile clones, each clone carries an added numerical label (10, 5, or 2.5), which refers to the selection condition they were subject to i.e. Clone-Ile-10-XX refers to a clone subject to 10% Ile RPMI or RPMI with 6 nM of Ile; Clone-Ile-5-XX was subject to 5% Ile RPMI or RPMI with 3 nM of Ile; and Clone-Ile-2.5-XX to RPMI with 1.5 nM Ile.

We have added the following text to the Methods section to clarify these naming schemes, and referenced this section in the relevant sections of the main text and legends to help direct readers:

“Clonal Isolation and Amino Acid Prototrophy Screens

After selection on the indicated amino acid dropout medium, CHO clones were manually isolated using a P200 pipette tip, transferred to a 96-well plate, and expanded before phenotypic screening. For each selection condition (CHO-Val in 4.25 nM valine RPMI; CHO-Ile in 1.5, 3, and 6 nM isoleucine RPMI), clones were distributed across two separate 10 cm² dishes to allow sampling of all infected cells but minimize risk of clones merging. Clones from each dish were assigned to a separate row of the 96-well plate and named according to plate position (for example, C1-C8 and D1-D8). CHO-Ile clones were further labeled with numeric identifiers corresponding to the isoleucine concentration used for selection (2.5, 5, and 10% of 0.06 mM, respectively). For instance, Clone-Ile-10-XX refers to a clone subject to 10% Ile RPMI or RPMI with 6 nM of Ile. For Jurkat cells, clones were isolated by serial dilution, and numbered sequentially.”

8. In Fig. 3F, and accompanying description of the results on Page 11, lines 308-312, the authors find that *ilvN* is expressed in 94% of all clones, “suggesting that it supports valine biosynthetic function without being strictly necessary”. They further report that this may be due to the fact that *ilvN* is reported to be a regulatory unit that “boosts *ilvB* catalytic rates and substrate specificities but that *ilvB* can perform its catalytic function in the absence of *ilvN*”. This is an interesting finding that the authors should elaborate further on. Given its high occurrence in their data (94%), it seems that *ilvN* may be providing a significant boost to *ilvB*'s catalytic rate. Is this supported by literature? What are the numbers, if they are known?

We have modified the following text to report the impact of *ilvN* on *ilvB* activity:

ilvN appeared in 94% (15/16) of clones after selection, suggesting that it strongly supports valine biosynthetic function. This is consistent with prior reports that *ilvN* acts as a regulatory subunit that enhances *ilvB* catalytic activity and substrate specificity; *ilvB* alone retains activity but operates at ~15% of the *ilvBN* holoenzyme³⁹.

9. The legend for Fig. 3G is lacking and needs to be expanded. What do the “median normalized read counts” refer to? What are they normalized to? A percentage of total reads, or internally normalized read values for each clone? Expanding the figure legend will help improve clarity and also provide a means for how to compare the authors' numbers across the different combinations (ie, is a score of 10 twice as good as a score of 5? How does that compare to the prototrophy score previously defined? Or are these results independent of those, since this is performed on only high prototrophy variants?)

We have clarified the language used in the legends of Figures 3G, 4G, and 5E as follows:

“Composite barcode ‘fingerprint’ across strongly valine prototrophic clones. For each TU within a sample, the percentage of barcodes assigned to that TU (relative to total reads for that sample) was calculated. The median percentage representation of each TU across all clones is displayed as a composite heatmap.”

We have also added the following clarification to the Methods section:

“For generation of per-clone heatmaps, the percentage of reads assigned to a given TU within a sample

was calculated, normalized to total reads per sample. For generation of heatmaps to identify optimal pathways, the median percentage representation of each TU across all indicated clones is displayed as a composite heatmap.”

In other words, values in per-clone heatmaps represent the fraction of reads assigned to each TU within that clone, while composite heatmaps report the median of these values across clones. These metrics reflect the relative usage of each TU within and across clones, and are used to identify enriched promoter and localization configurations among strongly prototrophic variants.

Importantly, these values are independent of the prototrophy score and should not be interpreted as measures of functional effect size. The previously defined prototrophy scores represent phenotypic readouts derived from the resazurin-based assay conducted prior to barcode sequencing and are not used in the generation of barcode-derived heatmaps. Rather, higher values indicate more frequent representation of a TU among selected clones, which, when combined with comparisons to baseline representation and simulation-based expectations, enables identification of functional pathway components.

10. The results in Extended Data 6A (dual prototrophic variants grown in ile-/val- media) should be contrasted with their results from Figs. 3 and 4 respectively, in more detail. On Pages 16-17, lines 467-470, the authors state “we noticed that CHO-Val clones and CHO-Ile clones exhibited differential underlying CDS usage suggesting that analyzing both populations on aggregate might conflate signal given that each population was not specifically selected to be doubly prototrophic (Extended Data 6C)”. While this may be true, some insights could still be gleaned about the utility of genes that enable single vs dual prototrophy. A description of these differences, along with a hypothesis for what genes help enable these functional phenotypes could be beneficial for the readers.

We have modified the language used in this section to center the comparison of CDS usage between singly and doubly prototrophic clones. We further hypothesize how differences in CDS usage may connect to each subphenotype:

“Among CHO-Val clones, 6 clones were prototrophic for both valine and isoleucine (Val+Ile+), 10 were prototrophic for valine only (Val+Ile-) and none were prototrophic for isoleucine only (Val-Ile+). Among CHO-Ile clones, 12 clones were Ile+Val+, 7 clones were Ile+Val- and 21 clones were Ile-Val+ (Extended Data Fig. 8D). When looking at the gene content underlying each phenotype, CHO-Val clones looked to have similar CDS representation irrespective of phenotype. On the other hand, CHO-Ile cells showed notable differences in *ilvG*, *ilvA*, and *ilvA-L481F* content across phenotypes (Extended Data Fig. 8E). In particular, *ilvG* was seen in 100% of Ile+Val- clones, 75% of Ile+Val+ clones and in <10% of Ile-Val+ clones indicating it is a strong determinant of isoleucine prototrophic function and possibly detrimental to valine prototrophic function. This is consistent with the known catalytic function of *ilvG*, which exhibits 57-180X greater preference for isoleucine biosynthetic intermediates over valine biosynthetic intermediates (Table 1). *ilvA* was most likely to be found in the Ile+Val- population suggesting it supports isoleucine prototrophic function but not valine prototrophic function, while *ilvA-L481F* was most used in the Ile+Val+ population suggesting that a feedback mechanism may be constraining the manifestation of the doubly prototrophic phenotype.”

11. Page 18, lines 526-527. Potential typo “...proceeded to selection for...” should be replaced with “proceeded with selection for...” or “proceeded to select for...”

Thank you, we have corrected this typo.

12. The analysis in Fig. 5E is done by removing the duplicate “siblings”, in order to generate a strong composite barcode fingerprint for prototrophic solutions. However, the presence of these siblings may also be confounding the analysis of CDS content presented in Fig. 5D, and should potentially be re-done by removing duplicate siblings as well.

We appreciate the reviewer’s concern and have added the requested plot as Extended Data Fig. 12F.

13. The description of results from the ML model presented in Fig. 5G, H requires more clarification and elaboration. Specifically:

a. What were the inputs to the ML model, and how were the data pre-processed and presented to the model? What was the model’s output (I’m assuming probability distribution for the two output classes)?

We thank the reviewer for this important suggestion.

The inputs to the model were the proportions of reads assigned to each TU within a given clone (analogous to the per-clone heatmaps shown in Extended Data Fig 3B-C), while the model outputs the estimated probability of being prototrophic or non-prototrophic, representing the averaged class likelihood produced across all random forest decision trees. Discrete class predictions are obtained by assigning clones to the class with the higher predicted probability.

We have modified the relevant section of the results for clarification:

“As input features, we used the proportion of reads assigned to each TU within a clone (analogous to the clone-level TU abundance heatmaps). Each clone’s phenotypic performance was binarized as prototrophic or auxotrophic using the Val Score threshold of 0.23.”

We have further modified the methods section to incorporate details on model inputs/outputs:

“To identify transcription unit (TU) features predictive of valine prototrophy, we trained a Random Forest classifier using relative barcode read counts as input features. Relative barcode read counts were calculated as the absolute read count for each TU per clone normalized by the total barcode read count for that clone. 119 samples represented by 40 features (i.e. TUs) were used to train the model. A binary phenotype was defined by thresholding the Val Score at 0.23, classifying samples as either prototrophic or non-prototrophic; phenotypic scores were used solely to generate this binary label and were not included among the predictive features, thereby preventing data leakage. For each clone, the Random Forest classifier output an estimated probability of being prototrophic or non-prototrophic, representing the averaged class likelihood produced across all decision trees. Discrete class predictions were obtained by assigning clones to the class with the higher predicted probability.”

All preprocessing steps, code, and the full analysis pipeline are available on GitHub:

[https://github.com/julietrolle/shotgun-genetic-engineering/blob/main/SGE_Paper_Finalized/Clonal Analysis/Jur_Val/9_Jur_Val_RandomForest.ipynb](https://github.com/julietrolle/shotgun-genetic-engineering/blob/main/SGE_Paper_Finalized/Clonal%20Analysis/Jur_Val/9_Jur_Val_RandomForest.ipynb)

b. In total, the model accurately classified 112/119, which is a convincing demonstration of the model’s accuracy. However, the authors should analyze the handful of incorrect variants, and generate some hypotheses for why the model may have failed to classify them accurately, given its exceptional performance on the entire dataset otherwise. For instance, it may be beneficial to know if the clones that were incorrectly predicted contained features that didn’t match those with the highest importance score.

We thank the reviewer for this suggestion and have included a new Extended Data Fig. 16 as well as the following text to explain the model’s misclassifications:

“5 auxotrophic clones were misclassified as prototrophic: 2 (Jurkat-Val-55 and Jurkat-Val-104) carried all 4 CDS that form our discovered ‘pathway’ solution and likely reflect examples of gene silencing or barcode swapping; 2 clones (Jurkat-Val-79 and Jurkat-Val-113) lacked ilvN only, potentially reflecting insufficient AHAS activity for prototrophic function; while the final clone (Jurkat-Val-110) lacked both ilvB and ilvD (Extended Data Fig. 16A). On the other hand, 2 prototrophic clones were misclassified as auxotrophic: Clone Jurkat-Val-60 (Val score: 0.31) appears to be clonally impure, carrying very lowly abundant ilvC and ilvD, which the model may not consider sufficient to generate a prototrophic phenotype, and Jurkat-Val-68 (Val score: 1.05) carried ilvN, ilvC, and ilvD under control of PGK-MTS as found in our discovered optimal pathway solution while ilvB is under control of EF1a-MTS, which is a much less represented possibility in our starting SGE library and may therefore not have been classified as contributing to the valine prototrophic phenotype (Extended Data Fig. 16B).”

Reviewer #1 (Remarks on code availability):

The files the authors cite in their manuscript appear to be in the Github repository and in order.

Reviewer #3 (Remarks to the Author):

In the manuscript entitled "A shotgun approach for highly multiplexed mammalian metabolic engineering" by Trolle et al. the authors present an innovative, systems level approach to reconstitute a valine biosynthetic capacity in mammalian cells. This enables a metabolic bypass potentially simplifying culture conditions, while also exposing a new experimental system to study essential amino acid metabolism in mammalian cells. The experimental framework is thoughtfully constructed, and the focus on screening for “winner” clones capable of essential amino acid biosynthesis is well-motivated. We commend the overall ambition and conceptual direction of the study, but we feel there are items, critical to the claims of this work, that should be considered and likely addressed prior to publication.

Major concerns:

1. The current screening library appears to be preselected and limited in complexity (4 per promoter/OLS, 3 barcodes per CDS across 10 genes), which seems to contradict the authors’ claim of “screened millions” (Lines 28, 639). This overstatement is misleading and undermines the credibility of the study. The limited library size constrains scalability, compromises the generalizability of the findings, and raises concerns about the overall robustness of the screening platform.

We appreciate the reviewer’s concern and agree that it is important to clearly distinguish between library size and the effective combinatorial diversity explored in the screen.

While the library is composed of 40 TUs, the combinatorial diversity arises during high-MOI lentiviral delivery, in which each cell stochastically assembles a distinct subset of TUs. Each infected cell therefore constitutes an independent combinatorial experiment, enabling the exploration of a much larger design space than is implied by library size alone.

To quantify the scale of diversity achieved in our screen, we performed a Monte Carlo simulation using the empirically measured barcode distribution after infection in CHO cells. This analysis showed that ~1.56M unique TU combinations were realized across the 2.42M infected CHO cells, which clarifies the basis for our statement that the screen assays “millions” of pathway configurations. We have revised the text to make this clearer in the manuscript.

“Our choice of 10 CDSs, 2 promoters, and 2 OLSs (40 total TUs) reflected a balance between biological relevance and practical sampleability. As we knew 4 genes to be sufficient to establish a valine prototrophic phenotype, we at minimum wanted to sample all possible 4-TU combinations of which there are 123,410 from a 40-member library. However, we were interested in sampling variation in copy number as well, motivating a target of sampling all 5-TU combinations of which there are 962,598. This design space can be sampled within a single 10cm² dish, which we reasoned was appropriate for a pilot-scale demonstration of the approach.”

“We simulated randomized choices of TUs based on the barcode distribution measured for CHO-Div at a range of MOIs between 2 and 14 assuming a Poisson distribution for integration events. At an MOI of 8, the simulation indicated that we would observe 1.56M different pathway combinations across the 2.42M infected cells with 12.4% of cells encompassing the gene content of our valine biosynthetic pathway (ilvN, ilvB, ilvC, ilvD) while 2.0% of cells would have all 4 genes localized to the cytoplasm and 1.6% would have all 4 genes localized to the mitochondria (Extended Data Fig. 1C).”

2. Moreover, viral vectors are inherently prone to recombination, particularly in complex pooled libraries, as demonstrated in early versions of Perturb-seq (PMID: 29457792). Such recombination events can disrupt transcriptional unit (TU) and barcode pairing, thereby reducing screen sensitivity. It is therefore critical to incorporate mitigation strategies. For example, long-read sequencing can verify expression cassette and barcode integrity, or 5' plus 3' barcoding schemes can improve accuracy. The authors should also follow established best practices from the CLASSIC (PMID: 36993481) and MPRA literatures (PMID: 37103023).

We thank the reviewer for raising this important point. We agree that lentiviral recombination that disrupts TU-barcode linkage (barcode-swapping) can in principle reduce the sensitivity of SGE screens.

We initially selected lentivirus for its broad host range and the ability to precisely control MOI, enabling us to target multiple cell types with the same library. However, as the reviewer notes, there are a few limitations to using lentivirus, including the potential for barcode swapping. For these reasons, a transposon-based strategy may be more suitable in some contexts. We also suggest that borrowing mitigation strategies from the MPRA literatures such as dual 5' and 3' barcodes can help address this concern if lentivirus is to be used. We have added the following text to the paper to help readers navigate this decision.

“We initially selected lentivirus because of the ease of titrating MOI and its broad host range, which allowed us to deploy the same libraries across multiple cell types. However, lentiviral delivery also comes with several practical limitations, including the need to perform packaging after library cloning, constraints on cargo size, and potential template switching that can lead to barcode swapping. For some applications, especially when working with larger constructs or when maintaining barcode linkage is essential, a transposon-based strategy (e.g., piggyBac) may therefore be more suitable. Additionally, the incorporation of dual 5' plus 3' barcoding strategies as employed in Massively Parallel Reporter Assays (MPRAs), or direct long-read readouts may help mitigate barcode swapping concerns if lentivirus is to be used.”

Importantly, we directly assessed the extent of barcode swapping in our system using long-read sequencing and RNA-seq, which are fully described in Supplementary Note 1. In brief, across four CHO clones (CHO-Val-D1, CHO-Val-D3, CHO-Ile-10-H1, and CHO-Ile-10-H2), 95.1% and 93.7% of reads carried the expected promoter and CDS barcodes, respectively (Supplementary Fig. 1A), corresponding to approximately one barcode swap per clone (Supplementary Table 2). Residual discordance is consistent with low-level clonal heterogeneity rather than widespread recombination. Specifically, TUs classified as background by amplicon-seq were detectable at low abundance by long-read sequencing (Extended Data Fig. 3B-D; Extended Data Fig. 6B-D). In Jurkat-Val-31, we observed complete concordance between long-read and amplicon-seq TU calls (Extended Data Fig. 11C-E).

We also assessed the expression of SGE transgenes via RNA-seq. As these transcripts are not polyadenylated (to enable lentiviral delivery), their detection by polyA-based RNA-seq is expected to be reduced relative to endogenous transcripts. Despite this limitation, measurements of transgene expression were largely consistent with amplicon-seq barcode assignments, further supporting the integrity of barcode-CDS linkage (Extended Data Fig. 3B,E; 6B,E; 11C,G).

Together, these results indicate that barcode swapping is infrequent in our system and does not substantially impact genotype assignment at the population level. Moreover, because SGE identifies functional pathway

solutions based on aggregate enrichment across many independent clones, it is inherently robust to low levels of barcode misassignment. This is further supported by our ability to accurately reconstruct functional pathway designs from the screen (Fig. 3J-K).

3. The manuscript introduces machine learning (ML) as part of the analytical framework, but there are concerns here. First, the model and dataset were not adequately described to understand if the dataset was large enough and structured to avoid overfitting, data leakage, or if it is generalizable.

Random Forest models are well established as robust to overfitting as each decision tree is trained on a bootstrap sample of the data and evaluates only a random subset of features at each split, reducing variance and preventing memorization of individual samples.

To avoid data leakage, only relative barcode read counts were supplied as model inputs, while the experimentally measured Val Score was used solely to derive binary phenotype labels and not included among the predictive features. All steps of feature processing and model fitting were performed independently within each training fold.

Generalizability and overfitting were assessed using stratified 4-fold cross-validation, which preserved the proportion of prototrophic and non-prototrophic clones across folds. Model performance was highly consistent, with comparable ROC AUC, precision-recall AUC, and confusion matrices across folds. Feature importance rankings were similarly stable. These indicate that the model generalizes well to unseen data and is not overfitting to any subset of clones.

We have updated the results and methods as follows to alleviate these concerns:

“As input features, we used the proportion of reads assigned to each TU within a clone (analogous to the clone-level TU abundance heatmaps). Each clone’s phenotypic performance was binarized as prototrophic or auxotrophic using the Val Score threshold of 0.23.”

“To identify transcription unit (TU) features predictive of valine prototrophy, we trained a Random Forest classifier using relative barcode read counts as input features. 119 samples represented by 40 features (i.e. TUs) were used to train the model. Relative barcode read counts were calculated as the absolute read count for each TU per clone normalized by the total barcode read count for that clone. 119 samples represented by 40 features (i.e. TUs) were used to train the model. A binary phenotype was defined by thresholding the Val Score at 0.23, classifying samples as either prototrophic or non-prototrophic; phenotypic scores were used solely to generate this binary label and were not included among the predictive features. For each clone, the Random Forest classifier output an estimated probability of being prototrophic or non-prototrophic, representing the averaged class likelihood produced across all decision trees. Discrete class predictions were obtained by assigning clones to the class with the higher predicted probability.”

Finally, all preprocessing steps, code, and the full analysis pipeline are available on GitHub:

[https://github.com/julietrolle/shotgun-genetic-engineering/blob/main/SGE_Paper_Finalized/Clonal Analysis/Jur_Val/9_Jur_Val_RandomForest.ipynb](https://github.com/julietrolle/shotgun-genetic-engineering/blob/main/SGE_Paper_Finalized/Clonal%20Analysis/Jur_Val/9_Jur_Val_RandomForest.ipynb)

4. The barcodes of EF1a-driven transcriptional units (TUs) appear to suffer from insufficient recovery, likely representing a technical bottleneck caused by suboptimal PCR enrichment. Notably, over 50% of clones containing the GFP transcriptional unit exhibited high basal expression, as shown in Figure 3F. This suggests an elevated signal-to-noise ratio and prompts a key question: does *ilvA*-L481F reflect a genuine biological signal, or is it merely background noise? Fine-tuning the PCR enrichment primer set may substantially improve barcode capture efficiency, thereby enabling more scalable and sensitive screening.

We agree with the reviewer that both barcode recovery and the distinction between true biological signal and background are important considerations in pooled screening approaches.

We note that this comment raises two distinct issues: (i) potential PCR bias affecting recovery of EF1a-driven TUs, and (ii) whether highly represented TUs such as *ilvA*-L481F reflect true biological signal or background.

First, regarding PCR bias - insufficient PCR enrichment is unlikely to explain the underrepresentation of EF1a containing TUs. All TUs are amplified using the same primer pair and identical amplicon architecture, and in the cloned bacterial library the EF1a barcodes were recovered at comparable levels to PGK-driven TUs, indicating no intrinsic PCR bias against EF1a barcodes (Extended Data Fig. 1A). As discussed in our response to Reviewer 1, we believe the drop in EF1a representation is most plausibly explained by events occurring during lentiviral delivery or subsequent cellular selection, rather than by PCR recovery. We have now elaborated on these various possibilities in the manuscript:

“Post-infection, we performed amplicon-seq to monitor TU distribution and observed underrepresentation of EF1a promoters both with and without MTS (Extended Data Fig. 1A). Because head-to-head comparisons of PGK and EF1a in pooled lentiviral libraries have not been systematically examined before,

the basis of this effect remains unclear. In our pooled setting, several mechanisms may contribute to the underrepresentation of EF1a-driven TUs. First, because EF1a is a stronger promoter than PGK, its recruitment of transcriptional machinery could interfere with transcription of the viral genome during packaging, thereby reducing viral titer. Second, EF1a contains an intron that is absent from PGK, and although this intron is typically retained during lentiviral packaging, interactions between the intron's splice sites and the viral splicing/packaging machinery could still modestly impair packaging efficiency. Finally, high-level expression driven by EF1a may impose a fitness cost when multiple pathway components are present in the same cell, leading to depletion of EF1a-linked TUs upon infection of the target cells."

Second, regarding the separation of noise from true biological signals, we can evaluate the ilvA-L481F case as an example. We note that this TU is already overrepresented in the starting CHO-Div population (Extended Data Fig. 1A). Monte Carlo simulations based on the measured barcode distribution and at an MOI of 8 predict that we would find PGK-MTS ilvA-L481F occurring in 59% of clones when picking TUs at random (Extended Data Fig. 4B), which closely matches its observed frequency in selected prototrophic clones (Fig. 3F). This indicates that its enrichment can be explained by baseline representation rather than selection.

Consistent with this, ilvA-L481F appears in the composite barcode "fingerprint" (median TU representation across strongly prototrophic clones) as expected for a highly abundant TU, even in the absence of functional contribution. Importantly, ilvA-L481F is absent from the strongly prototrophic clones CHO-Val-D1 and CHO-Val-D3, further supporting that it is not required for valine prototrophy. Based on these observations, we did not include ilvA-L481F in the inferred pathway solution.

We can contrast this with PGK-MTS ilvB, which is even more overrepresented in the starting CHO-Div library (Extended Data Fig. 1A) and is similarly predicted to appear in our composite fingerprint by chance (Fig. 3G). However, the universality of ilvB across all prototrophic clones (Fig. 3F) combined with the overrepresentation of its PGK expressed counterpart prompted us to consider it part of the prototrophy solution.

We have added a panel, Extended Data Fig. 4B, and the following text to clarify this reasoning: "While ilvA-L481F expressed from a mitochondrially localized PGK promoter was also present in our composite heatmap, we did not consider this part of the pathway solution considering its overrepresentation in the starting CHO-Div population (Extended Data Fig. 1A). When simulating randomized choices of TUs at MOI=8, we find PGK-MTS ilvA-L481F in 59% of clones (Extended Data Fig. 4B), which correlates well to its presence across all valine prototrophic clones, indicating that it is not overrepresented following selection in reduced valine RPMI. When additionally considering its absence from strongly valine prototrophic clones CHO-Val-D1 and CHO-Val-D3 (Extended Data Fig. 3B), we did not find reason to implicate ilvA-L481F in contributing to the valine prototrophy solution."

Finally, we validated this framework by reconstructing an optimized pathway consisting of mitochondrially-localized ilvB, ilvN, ilvC, and ilvD. This SGE-informed design substantially improved valine biosynthetic performance in valine free medium (1.25 days/doubling) compared to our previous pathway iteration, pMTIV (3.77 days/doubling), and closely matched the performance of strongly prototrophic clones CHO-Val-D1 and CHO-Val-D3 (at 1.14 and 1.05 days/doubling, respectively). We added panels Fig. 3J-K to reflect this data as well as the following text:

Based on these results, we constructed an optimized valine biosynthetic pathway using rational design. We matched all design choices used for our previous pathway, pMTIV, except that CDSes were localized to mitochondria as informed by the SGE screen. The pathway was synthesized in ~3 kilobase fragments, assembled using homologous recombination in *S. cerevisiae*, and delivered to CHO cells using Flp-InTM recombination (Fig. 3J). The new SGE-informed design enabled CHO cells to grow in valine-free RPMI at 1.25 doublings/day, considerably outperforming the growth of pMTIV cells at 3.77 doublings/day (Fig. 3K).

5. The current version of the manuscript lacks sufficient system-wide, functional validation of the SGE methodology. Incorporating RNA-seq profiling of the selected clones would significantly strengthen the study. This approach would not only confirm transgene expression and complement genomic copy number analysis but also offer valuable insights into the biological effects of the perturbations, thereby deepening our understanding of essential amino acid biosynthesis in mammalian cells.

While the primary goal of this study is to use SGE as a discovery framework to identify functional pathway solutions rather than to extensively characterize individual clones, we agree that system-level transcriptomic profiling provides an important complementary readout for understanding the effects of SGE perturbations.

To address this, we performed bulk RNA-seq (polyA capture) on representative prototrophic clones across both systems (CHO-Val-D1, CHO-Ile-10-H1, and Jurkat-Val-31), each profiled under nutrient-replete and amino acid dropout conditions (with biological replicates), alongside parental control cells. We have added text describing these experiments to Supplementary Note 1.

First, we confirmed the expression of SGE-derived transgenes. As these transcripts are not polyadenylated (to enable lentiviral delivery), their detection by polyA-based RNA-seq is expected to be reduced relative to endogenous transcripts. Despite this limitation, measurements of transgene expression were largely consistent

with amplicon-seq barcode assignments, further supporting the integrity of barcode-CDS linkage (Extended Data Fig. 3B,E; 6B,E; 11C,G).

Second, across all transcriptomic comparisons, differences relative to parental cells were modest under both nutrient-replete and dropout conditions (Supplementary Fig. 2A-E; Supplementary Fig. 3A-E; Supplementary Fig. 4A-E; Supplementary Tables 4, 5, 7). Gene ontology (GO) enrichment analysis did not reveal strong or consistent signatures of stress, metabolic compensation, global regulatory rewiring or perturbation of mitochondrial function. Consistent with this, we do not observe a growth defect in engineered clones under nutrient-replete conditions (Extended Data Figs. 2A; 5A; 10B), suggesting that expression of multiple heterologous genes imposes limited metabolic burden in nutrient replete conditions.

Minor concerns:

6. Provide full details on all reporting materials, including plasmid construction, and CDS/gene sequences.

We have uploaded a sequence file for the base vector to the Github and have provided all library sequences as Supplementary Table 8.

7. Show library distribution using a histogram where X-axis is Log2-transformed read counts (barcode abundance) and Y-axis is frequency (number of barcodes).

We have plotted read counts by individual barcodes as well as read counts aggregated by TU as histograms as requested. For ease of comparison, we have placed Z-score distribution histograms alongside.

Raw and TU aggregated barcode read counts have also been made available for all bulk populations as well as all individual clones as Supplementary Table 9 and Supplementary Table 10, respectively, to allow readers the option to make equivalent plots for any individual clone.

Shown in the figure below are:

1. DNA library prior to viral packaging
 2. The infected CHO-Div population prior to selection on reduced amino acid medium
 3. A consolidated plot featuring all sequenced CHO-Val clones
 4. A consolidated plot featuring all sequenced CHO-Ile clones
 5. The infected Jurkat-Div population prior to selection on reduced valine medium
 6. A consolidated plot featuring all sequenced Jurkat-Val clones
- We have furthermore generated per-sample plots that details usage of all 9 barcodes per TU for each control (DNA Library, pre-selection populations) as well as clones that we highlight in the manuscript for a more detailed view of barcode distribution:
8. Provide raw barcode counts across LV virus, pre-infection, and post-harvest samples (e.g., Extended Data 1 and 7).
- Raw SGE read counts for all sequenced samples have been added as Supplementary Table 9 while TU aggregated barcode counts have been added as Supplementary Table 10.

9. Report individual barcode counts per TU and display library distribution as a histogram (X-axis: log2 barcode read count, Y-axis: frequency).

Please refer to the above response to Comment #7.

10. Please clarify the actual number of promoter/OLS used in the screening library (ie. 3 or 4).

We have modified the following text to clarify the number of promoter/OLS used:

“We included two constitutive promoters in the library, EF1a (strong) and PGK (medium), to explore varying expression levels of each pathway component. Each promoter appeared either standalone or paired with a mitochondrial targeting sequence (MTS), and each promoter-OLS combination was assigned three unique barcodes (e.g., EF1a-no-MTS: 3 barcodes; EF1a-MTS: 3 barcodes).”

11. There was an inconsistency between GFP signal presentation in Figures 3F and 3G. We suggest including a side-by-side heatmap of GFP TUs with valine prototrophy scores (e.g., Val Score >0.75) to help disentangle functional hits from background expression.

We appreciate the reviewer’s concern and agree that it is important to distinguish functional signals from background expression. To address this, we have added the following text to clarify the differences in the data depicted in Figures 3F and 3G:

3F: “Percentage of all valine prototrophic clones (Val Score > 0.28) and all strongly valine prototrophic clones (Val Score > 0.75) that contain each depicted CDS as measured by amplicon-seq”

3G: “Composite barcode ‘fingerprint’ across strongly valine prototrophic clones. For each TU within a

sample, the percentage of barcodes assigned to that TU (relative to total reads for that sample) was calculated. The median percentage representation of each TU across all clones is displayed as a composite heatmap."

In other words, 3F represents the frequency of a particular CDS across CHO-Val clones, whereas 3G depicts the relative usage of each TU within CHO-Val clones with Val Score >0.75 (strongly prototrophic). As a result, a CDS such as GFP may be present in many clones (Fig. 3F) but lack a consistently enriched promoter/OLS configuration, and therefore not exhibit a strong signal in the composite heatmap (Fig. 3G).

While we agree that a side-by-side visualization could be informative, the two panels represent fundamentally different quantities (presence vs relative abundance), and combining them directly into a single heatmap would risk conflating these distinct measures. Instead, to enable deeper inspection of the relationship between TU usage and prototrophy, we have made per-clone heatmaps and corresponding prototrophy scores available on github (<https://github.com/julietrolle/shotgun-genetic-engineering>), allowing readers to directly compare barcode usage and phenotype at the level of individual clones.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I think the authors did a thorough job of addressing all my comments with their edits and additional experiments. The manuscript is very strong, and I have no further concerns.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>