Metabolic Engineering 85 (2024) 159–166

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

Genetic heterogeneity of engineered *Escherichia coli* Nissle 1917 strains during scale-up simulation

Lara P. Munkler [a], Elsayed T. Mohamed [a], Ruben Vazquez-Uribe [a,e], Victoria Visby Nissen [a], Peter Rugbjerg [b], Andreas Worberg [a], John M. Woodley [c], Adam M. Feist [a,d], Morten O. A. Sommer [a,*]

[a] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark
[b] Enduro Genetics ApS, Copenhagen, Denmark
[c] Department of Chemical and Biochemical Engineering, Technical University of Denmark, Lyngby, Denmark
[d] Department of Bioengineering, University of California, San Diego, CA, USA
[e] Vlaams Instituut voor Biotechnologie, Center for Microbiology, Leuven, Belgium

ARTICLE INFO

Keywords:
Escherichia coli Nissle 1917
Advanced microbiome therapeutic (AMT)
Genetic stability
Scale-up

ABSTRACT

Advanced microbiome therapeutics have emerged as a powerful approach for the treatment of numerous diseases. While the genetic instability of genetically engineered microorganisms is a well-known challenge in the scale-up of biomanufacturing processes, it has not yet been investigated for advanced microbiome therapeutics. Here, the evolution of engineered *Escherichia coli* Nissle 1917 strains producing Interleukin 2 and Aldafermin were investigated in two strain backgrounds with and without the three error-prone DNA polymerases polB, dinB, and umuDC, which contribute to the mutation rate of the host strain. Whole genome short-read sequencing revealed the genetic instability of the pMUT-based production plasmid after serial passaging for approximately 150 generations using an automated platform for high-throughput microbial evolution in five independent lineages for six distinct strains. While a reduction of the number of mutations of 12%–43% could be observed after the deletion of the error-prone DNA polymerases, the interruption of production-relevant genes could not be prevented, highlighting the need for additional strategies to improve the stability of advanced microbiome therapeutics.

1. Introduction

A promising biomedical application of synthetic biology lies in an approach where engineered microorganisms are being harnessed for therapeutic and diagnostic applications (Charbonneau et al., 2020; Chen et al., 2014; Iannone et al., 2023; Praveschotinunt et al., 2019; Tumas et al., 2023). Within this context, the human microbiome has become an interesting therapeutic target, as it plays a key role in human health (Juarez et al., 2022). Among the different approaches that target the gut microbiome, advanced microbial therapeutics (AMTs) have emerged as a powerful tool to treat numerous diseases (Arora et al., 2016; Chen et al., 2014; Steidler et al., 2003). Specifically, engineered strains of the probiotic *Escherichia coli* Nissle 1917 (DSM 6601, O6:K5:H1, *Ec*N) have demonstrated remarkable potential in serving as a chassis for treating conditions such as cancer (Tumas et al., 2023; Yu et al., 2020) and metabolic diseases (Armetta et al., 2021; Charbonneau et al., 2020; Iannone et al., 2023; Praveschotinunt et al., 2019). *Ec*N is an extensively studied microorganism and has been implemented as a probiotic since its isolation by Alfred Nissle in 1917 for its ability to inhibit *Salmonella* growth (Sonnenborn, 2016). *Ec*N is now commercially available as Mutaflor® and widely used as a probiotic (Jacobi and Malfertheiner, 2011; Scaldaferri et al., 2016; Schultz, 2008). Despite the considerable potential of AMTs, numerous challenges must be addressed, including the genetic stability and maintenance of viability of the chassis organism during the manufacturing process as well as during further replication within the patient's microbiome (Charbonneau et al., 2020).

The stability of genetically engineered microorganisms during the fermentation process scale-up is a broadly investigated challenge in biomanufacturing. Throughout large-scale cultivations, the cells are exposed to numerous stress factors. In a large tank, shear forces and mixing gradients result in zones with variable or limited oxygen and nutrient supply or gradients in pH and substrate/product concentrations

* Corresponding author.
E-mail address: msom@bio.dtu.dk (M.O.A. Sommer).

https://doi.org/10.1016/j.ymben.2024.08.001
Received 1 May 2024; Received in revised form 25 July 2024; Accepted 5 August 2024
Available online 5 August 2024

(Takors, 2012; Xia et al., 2015). Besides those physical stress factors, the engineering of a host organism for the expression of foreign genes stresses the cell by adding an additional metabolic burden on the cell. This metabolic burden mostly entails a reduction of cell fitness, and hence a reduced growth rate of the production host compared to the wild type (Glick, 1995). There are several emerging techniques available to improve the genetic stability of the production host. Next to synthetic addiction (Rugbjerg et al., 2018b) or genome-reduced strains (Csörgo et al., 2012; Sharma et al., 2007), a significant reduction of the mutation rate has been achieved by the deletion of the error-prone DNA polymerases in *E. coli* K12 (Csörgo et al., 2012). Error-prone DNA polymerases are involved in the SOS response and activated under stressful conditions, such as substrate or oxygen limitations, pH shift, or the additional metabolic load (Reuven et al., 1999).

In this study, the evolution of two engineered *Ec*N strains and one *Ec*N control strain were investigated with two different genetic backgrounds (with and without error-prone DNA polymerases). The two therapeutic proteins were: Interleukin 2 (IL-2) (Tumas et al., 2023) and Aldafermin (Iannone et al., 2023), an analog of Fibroblast Growth Factor 19 (FGF19). A strain without bioactive molecule production was utilized as a control. Utilizing the serial passaging approach presented by Rugbjerg et al. (2018) combined with a fully automated liquid handling platform customized for high-throughput microbial evolution studies (LaCroix et al., 2015; Rugbjerg et al., 2018a), allowed for accurate transfer of grown cultures to fresh high-density culture medium according to their optical density, independent of temporal changes in their growth rate. To assess the genetic alterations of the strains during the experiment, short-read whole genome sequencing was performed on the endpoint-derived lineages. Comparing the sequencing reads to the initial strain led to the identification of the genetic changes during the course of the experiment. The experiment accounted for the number of generations in large-scale fermentation (~60–80 generations) and subsequent application as AMT, estimating an additional ~50–100 generations after the fermentation process ended. Thus, the study extended until 150 generations were reached. Our findings revealed that the AMT strains escaped the production burden of the therapeutic compounds IL-2 and Aldafermin by targeted mutations. The stability of AMT strains was enhanced by deleting the error-prone DNA polymerases. However, despite these improvements, an interruption in the therapeutic genes on the pMUT-based production plasmid was consistently observed. This underscores the necessity for further investigation into optimizing AMT production processes.

2. Results

To study the genetic stability of different probiotic *Ec*N populations, the large number of generations occurring during large-scale fermentation processes was mimicked by serial passaging over an average of 10 days (Fig. 1A). The investigated *Ec*N strains are potential candidates for advanced microbiome therapeutics (AMT), which entails additional cell divisions after the production fermentation process has ended. Hence, the commonly investigated span of 60–80 generations (Heieck et al., 2023; Rugbjerg et al., 2018a), was extended to 140–160 generations. By combining a fully automated online measurement of the OD_{600} with an automated liquid handling robot, the cells were transferred to a new cultivation medium in several replicate experiments (n = 5). Besides enabling an accurate, OD-triggered transfer, the utilized methodology effectively avoided starvation stress. In addition, it enabled the automated adjustment of the transfer time point with an increasing growth rate of the population.

As previously demonstrated in Csörgö et al. (2012), the deletion of the SOS-inducible DNA polymerases resulted in a reduction of mutation rate without impacting the cells growth in *E. coli* K12 (Csörgo et al., 2012). In this study, this strategy was implemented by knocking out the genes of the three error-prone DNA polymerases Pol II, Pol IV, and PolV (Δ*polB*, Δ*dinB*, Δ*umuDC*, summarized as ΔEPP) by scarless lambda Red integrations of dsDNA using CRISPR/Cas9-mediated counterselection and compared the resulting to the original background strain. Subsequent to the performed triple KO, the background strains were transformed to carry either an empty plasmid (pMUT_empty), used as control to exclude mutagenesis caused by the pMUT plasmid itself, or one of the production plasmids pMUT_IL-2 or pMUT_Aldaf. (Fig. 1B).

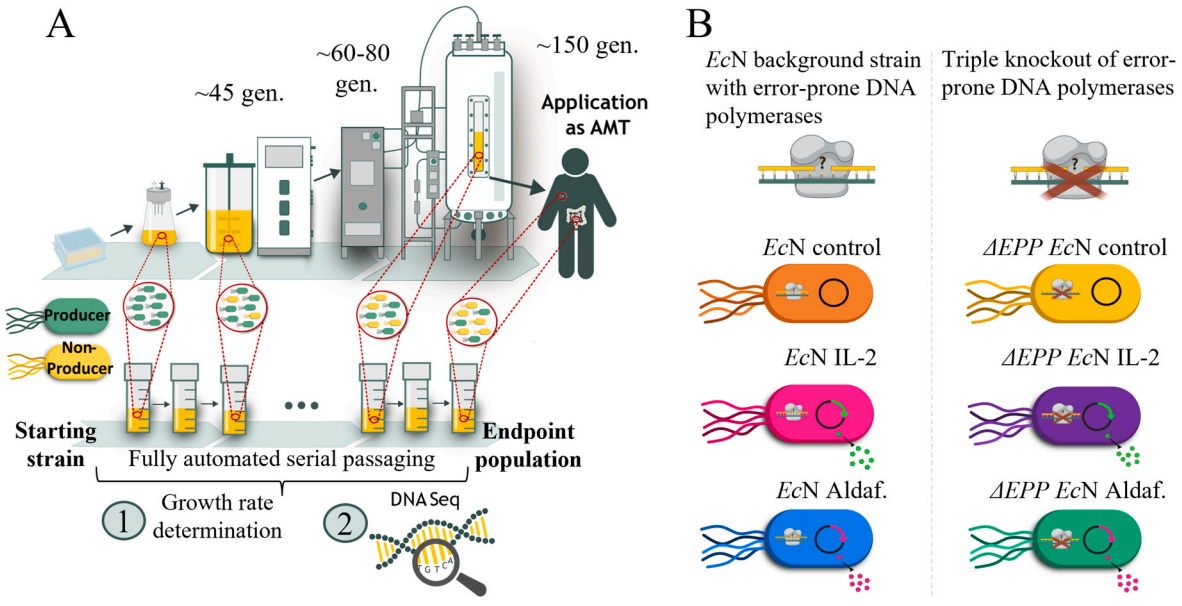

Fig. 1. Simulation of the number of generations occurring during large-scale long-term fermentations and application as AMT by fully automated serial passaging (adapted from Rugbjerg et al. (2018) (Rugbjerg et al., 2018a)) **A**: By fully automated serial passaging using the liquid handling platform represented in LaCroix et al. (2015) (LaCroix et al., 2015), the number of generations during large-scale bioreactors and during the application as advanced living therapeutic (AMT) were mimicked. Daily sampling was performed until 140–160 generations were achieved, followed by growth rate determination and whole genome sequencing. All experiments were carried out in 5 independent replicates. **B**: The six investigated EcN strains originated from two different background strains (ΔclbP EcN or ΔclbP ΔEPP EcN) carrying one of the three pMUT-based plasmids (pMUT_empty, pMUT_IL-2, or pMUT_Aldaf.). The three DNA error-prone polymerases were knocked out by scarless lambda Red integrations of dsDNA using CRISPR/Cas9-mediated counterselection.

2.1. Comparison of maximal growth rates over increasing number of generations reveals differences in cell fitness recovery

The fitness of a cell population directly relates to the ability of an organism to survive and thrive in its environment. In this experiment, the *growth rate* of the different strains was evaluated and compared as a measure of the fitness of the cell populations. The growth rate of the starting strains (Fig. 2A, light-colored bars) was compared to the endpoint populations (140–155 generations) (Fig. 2A, colored bars). A marginal increase in growth rate from 100% to 104 ± 1% was observed for *Ec*N control populations (orange) and 97% to 110 ± 2% for ΔEPP *Ec*N control (yellow bars). In contrast, for the growth rates of the four protein-producing populations *Ec*N IL-2 (pink bars), ΔEPP *Ec*N IL-2 (violet bars), *Ec*N Aldaf. (blue bars), and ΔEPP *Ec*N Aldaf. (turquoise bars), a considerable improvement of 29 ± 3%, 32 ± 4%, 32 ± 6%, or 33 ± 3%, respectively, was observed (Fig. 2A). This suggested that the production populations exhibited significant enhancements in their growth rates compared to the control populations over time.

The dynamics of growth rate improvement across multiple generations are shown in Fig. 2B–D. The IL-2-producing strains *Ec*N IL-2 (pink symbols) and ΔEPP *Ec*N IL-2 (violet symbols) exhibited a relatively stable growth rate for the initial 13 generations at approx. 79–81% (Fig. 2C). Subsequently, there was a notable surge in growth rate after approximately 30–35 generations, reaching 101 ± 4% and 105 ± 4%, respectively. This was followed by a period of stagnation between 50 and 60 generations, after which a slight increase was observed, peaking at 103 ± 5% for *Ec*N IL-2 and 106 ± 3% for ΔEPP *Ec*N IL-2. Notably, from the 30th generation onwards, the *Ec*N IL-2 populations (pink symbols) exhibited a slightly lower relative population growth rate compared to ΔEPP *Ec*N IL-2 (violet symbols).

The Aldafermin-producing strains *Ec*N Aldaf. (blue symbols) and ΔEPP *Ec*N Aldaf. (turquoise symbols) showed a significant growth rate increase from 76% to 94 ± 3% for *Ec*N Aldaf. and from 77% to 104 ± 2% for ΔEPP *Ec*N Aldaf. within the initial 13 generations (Fig. 2D). Subsequently, the growth rate showed only a slight upward trend, reaching 108 ± 6% and 110 ± 3% after 150 generations for *Ec*N Aldaf. and ΔEPP *Ec*N Aldaf., respectively.

2.2. Whole genome sequencing revealed the production plasmid expression cassettes as the main mutation targets

To investigate the genetic stability of the *Ec*N strains in this study, and to reveal potential mutational hot spots and mechanisms, samples of endpoint populations were sequenced and compared with starting cultures. Focusing on the mutation type in the IL-2-producing strains, the triple KO of the error-prone DNA polymerases (EPP) led to a reduction of SNP events (Fig. 3A), from 14 to 2. The deletion of the EPP in the Aldafermin-producing strain changed the total mutation counts from eight to seven. Additionally, a change in the mutation type from deletion to SNPs could be observed.

Focusing on the landscape of mutations, occurring in two or more replicates (single replicates in Supplement Fig. S1), for some of the investigated *Ec*N populations, mutations in the same genes are present, including *pMUT IL-2*, *pMUT (putative HTH)*, *pMUT (ompA/IL-2)*, *hyrR/ipbA*, *pMUT (ompA)*, or *pMUT Aldaf.* (Fig. 3B, right-hand side black nodes). Besides the mutation in *hyxR/ipbA*, all other shared mutations occur in production-related genes on the pMUT-based plasmid. By a more detailed investigation of the long insertions, the IS-like element ISEc52 family transposase was identified as the most frequently inserted mobile element (Supplement Table S1). An overview of the gene name in which the mutation occurred, the mutation type including unique counts, the gene annotation, and the *Ec*N population in which the

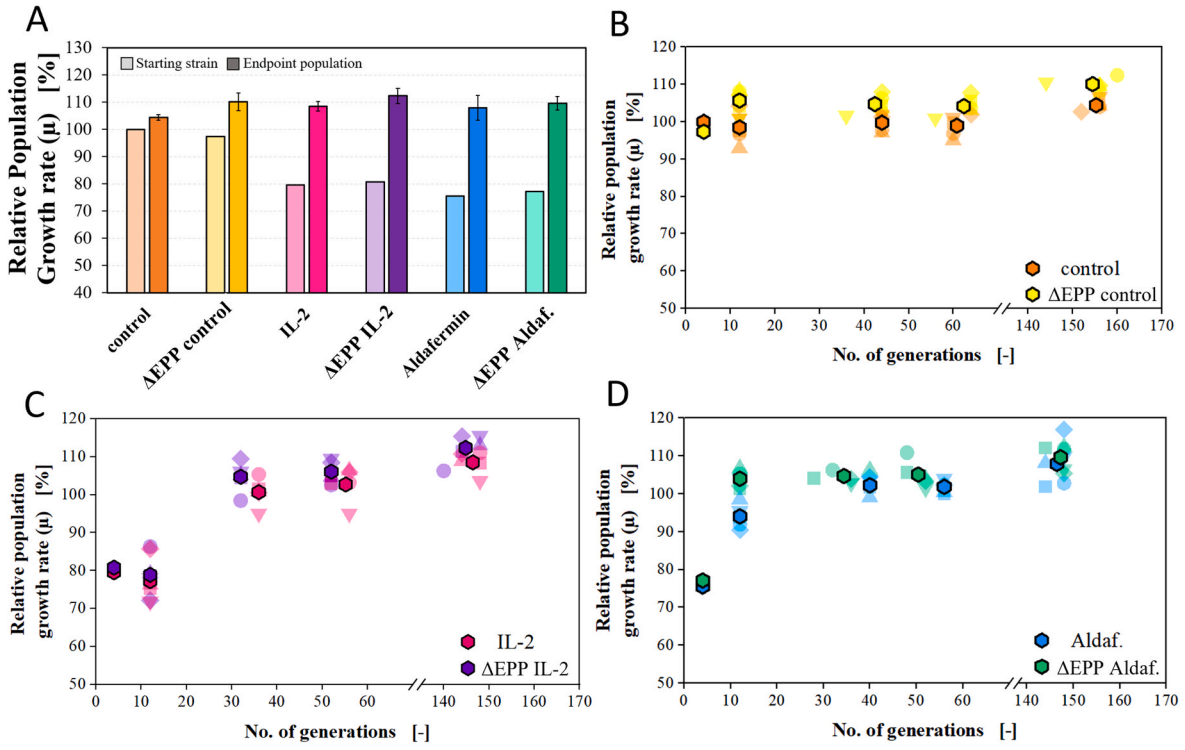

Fig. 2. Recovery of cell fitness of *Ec*N production populations. A: Comparison of relative maximal growth rate of the starting strains (light-colored bars) with the endpoint populations (colored bars). Biological replicates of the endpoint populations are depicted in the error bars (starting strains: n = 2, endpoint population: n = 5). **B-D:** Progression of the relative maximal growth rate of the *Ec*N populations (B: EcN control (orange), ΔEPP EcN control (yellow), C: *Ec*N IL-2 (pink), ΔEPP *Ec*N IL-2 (violet), D: *Ec*N Aldafermin (blue), ΔEPP *Ec*N Aldafermin (turquoise)) with dependence of the number of generations. The single biological replicates are shown in light colors (variable no. of generations for the different replicates due to differences in the growth rates), average in bright colors. All growth rates are relative to the highest maximal growth rate ($P1_{Aldaf.}$ = 0.995 h^{-1}).

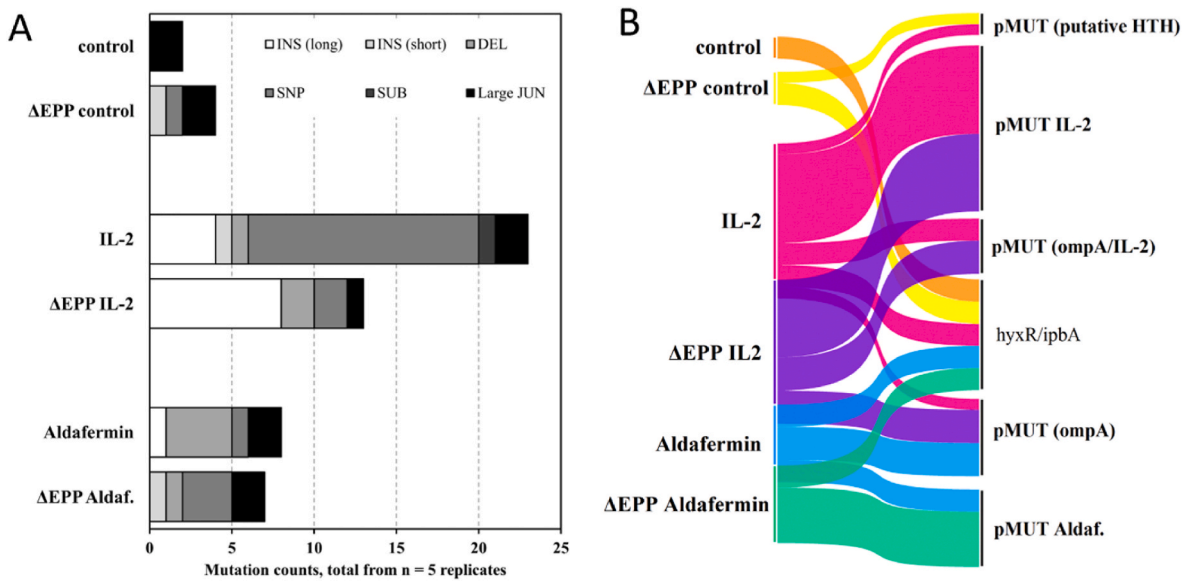

Fig. 3. Overview of mutations. **A**: Number of mutation counts when a type of mutation was observed in one of the five replicates of the investigated endpoint populations (130–150 generations). Mutation type abbreviations: INS = insertion, DEL = deletion, SNP = single nucleotide polymorphism, SUB = substitution, JUN = junction. **B**: Landscape of the mutations identified in endpoint populations of *Ec*N control (orange), ΔEPP *Ec*N control (yellow), *Ec*N IL-2 (pink), ΔEPP *Ec*N IL-2 (violet), *Ec*N Aldaf. (blue), and ΔEPP *Ec*N Aldaf. (turquoise). Each population group contains five replicates, mutations are only depicted if they are shared between two replicates. Filter for mutation frequency threshold = 30%. The width of the link correlates with the number of unique occurrences of mutations in the right-hand side listed gene or genetic loci. Right-hand bold gene names indicate plasmid mutations.

mutations were identified is given in Table 1.

2.3. *Production cassette in pMUT-based production plasmids as the main target for mutations*

In this study, the two *Ec*N background strains *Ec*N and ΔEPP *Ec*N were modified by engineering one of the two pMUT plasmids. Generally, the pMUT-based production vector was previously described as stable (Lan et al., 2021). However, most of the frequently shared mutations occurred in genes related to the pMUT-based production plasmid (Fig. 4).

All investigated populations carried an engineered pMUT-based plasmid, either with no additional features (*Ec*N control, ΔEPP *Ec*N control in Supplement Fig. S2) or carrying the IL-2 (*Ec*N IL-2, Fig. 4A, ΔEPP EcN IL-2, Fig. 4B) or Aldafermin (EcN Aldaf., Fig. 4C, ΔEPP EcN Aldaf., Fig. 4D) production gene after a constitutive promoter (J23107)

Table 1
Mutation overview.

Gene name	Mutation type* (unique counts)	Gene annotation**	EcN populations in which mutations were identified
ecpR	SNP (1)	ECP biosynthesis operon DNA-binding transcriptional regulator EcpR	*Ec*N Aldafermin
yqeB	SNP (3)	selenium-dependent molybdenum cofactor biosynthesis protein YqeB	*Ec*N IL-2
oxyR	SNP (1)	DNA-binding transcriptional regulator OxyR	*Ec*N IL-2
CIW80_RS13115	SNP (1)	Hypothetical protein	*Ec*N IL-2
dnaK	SNP (1)	molecular chaperone DnaK	*Ec*N IL-2
selB	SNP (1)	selenocysteine-specific translation elongation factor	*Ec*N IL-2
pykF	SNP (1)	pyruvate kinase PykF	ΔEPP *Ec*N Aldafermin
gatY, fbaB	SNP (1)	tagatose-bisphosphate aldolase subunit GatY/class I fructose-bisphosphate aldolase	ΔEPP *Ec*N control
papX	SNP (1)	HTH-type transcriptional regulator PapX	*Ec*N IL-2
CIW80_RS04770	SNP (1)	integrase domain-containing protein	*Ec*N IL-2
hyxR/ipbA	large JUN (30)	LuxR family transcriptional regulator HyxR/tyrosine-type DNA invertase IpbA	*Ec*N control, *ΔEPP Ec*N control, *Ec*N IL-2, *ΔEPP Ec*N Aldafermin, EcN Aldafermin
agn43, CIW80_RS23450	SNP (1)	autotransporter adhesin Ag43/hypothetical protein	EcN IL-2
pMUT J23107	long INS (1)	Promoter Region	EcN IL-2
pMUT Aldaf.	DEL (3), SNP (2), short INS (2)	Aldafermin production gene	*Ec*N Aldafermin, ΔEPP *Ec*N Aldafermin
pMUT ompA	DEL (3), long INS (3), large JUN (1)	ompA secretion signal	*Ec*N Aldafermin, EcN IL-2, ΔEPP *Ec*N IL-2
pMUT IL-2	SNP (8), SUB (1), DEL (2), long INS (4)	IL-2 production gene	*Ec*N IL-2, ΔEPP *Ec*N IL-2
pMUT HTH	short INS (2)	putative HTH domain on pMUT-based production plasmid	*Ec*N IL-2, *ΔEPP Ec*N control
pMUT ompA/IL-2	long INS (5)	Interface between ompA secretion signal and IL-2	*Ec*N IL-2, ΔEPP *Ec*N IL-2

*Mutation type abbreviations: DEL = deletion, INS = Insertion, SNP = single nucleotide polymorphism, SUB = substitution. Indication of the unique number of counts of mutation types in the populations in parenthesis. **Annotations assigned from NCBI Reference Genome NZ_CP022686.

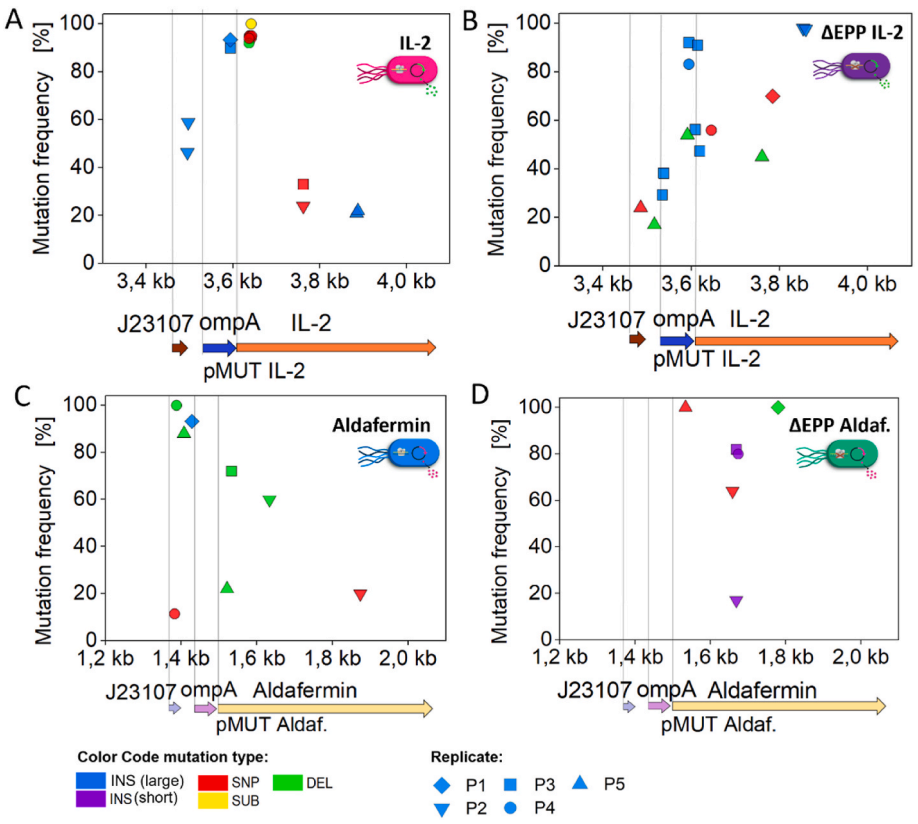

Fig. 4. Genetic pathway stability of pMUT-based production plasmids. A-D: The different mutation mechanisms are color-coded: large INS (blue), short INS (purple), DEL (green), SNP (red), or SUB (yellow). The five replicates are shown in the different symbols (P1 = diamond, P2 = inverted triangle, P3 = square, P4 = circle, P5 = triangle). In C-D, it was zoomed in to the production cassette, as the predominant region of mutations. Filter for mutation frequency threshold = 15%.

and an ompA secretion signal.

In the pMUT IL-2 plasmid of the *Ec*N IL-2 population, one short insertion at 4.8 kb was found (full plasmid in Supplement Fig. S2). This region is annotated as *hok/sok,* a toxin anti-toxin system for enhanced plasmid stability (Gerdes, 1988). Besides this, all mutations occurred in the region of the production cassette consisting of the promoter J23107, an ompA secretion signal, and the IL-2 gene (Fig. 4A).

In each of the five replicates P1–P5 of the four investigated production strains, the population tried to escape production by either introducing targeted short mutations in the J23107 promoter region, between the promoter region and *ompA* secretion signal, *ompA* secretion signal itself, or the production gene, or by large insertion events. Often, the short mutations resulted in the introduction of a stop codon, as summarized in Table 2.

Table 2
Plasmid mutations resulting in stop codons in *Ec*N production populations.

Population	Population replicates No.	Stop-codon mutation	Position [kb]	Frequency [%]
*Ec*N IL-2	P3	E52* (GAA→TAA)	3762	33
	P4	Q11* (CAG→TAG)	3639	95
ΔEPP *Ec*N IL-2	P1	L59* (TTA→TGA)	3784	70
	P4	Q13* (CAG→TAG)	3645	56
*Ec*N Aldaf.	P2	Q125* (CAG→TA)	1874	20
ΔEPP *Ec*N Aldaf.	P2	L54* (TTA→TGA)	1659	64
	P5	W12* (TGG→TGA)	1534	100

3. Discussion

The application of engineered *Ec*N as an advanced microbiome therapeutic (AMT) has been harnessed as a potential treatment for various diseases (Iannone et al., 2023; Tumas et al., 2023). However, engineering of microorganisms often results in an additional metabolic burden, entailing a reduction of cell fitness. Hence, a reduction of the production host's growth rate compared to the wild type can be observed. The selective pressure to escape this additional burden often leads to the emergence of strains with mutations. In this study, the genetic stability of the promising microbiome therapeutic *Ec*N strain (Iannone et al., 2023; Tumas et al., 2023), was investigated during serial passaging over 150 generations. Additionally, the stabilizing effect of the deletion of the three diversity-generating, error-prone DNA polymerases was examined, as previously shown for *E. coli* K12 (Csörgo et al., 2012).

A recovery of the initially reduced cell fitness due to the additional metabolic burden was indicated. At first, the growth rates of four *Ec*N production strains - *Ec*N IL-2, *Ec*N Aldafermin, ΔEPP *Ec*N IL-2, and ΔEPP *Ec*N Aldafermin were lower compared to two control strains *Ec*N control and ΔEPP *Ec*N control (Fig. 2A). Additionally, the control strains marginally improved their growth rate, which potentially can be attributed to adaptations to their environment, as previously described for *E. coli* MG1655 (Knöppel et al., 2018; LaCroix et al., 2015). However, none of these possible mechanisms for medium adaptation were found in this study. The occurring new junction between the LuxR family transcriptional regulator *hyxR* and tyrosine-type DNA invertase *ipbA* was identified mainly at low frequencies (Fig. 3B). It requires further investigation to understand its significance. Potentially, utilizing long-read sequencing techniques such as Nanopore sequencing could have covered more large structural changes.

The driving force for the substantial enhancement in growth rates in

the production strains shown in Fig. 2 was presumably the interruption of the production of IL-2 or Aldafermin by targeted mutations (Fig. 4), which presumably resulted in the loss of production (Supplement Fig. S3). Generally, the native cryptic pMUT plasmids pMUT1 and pMUT2 are reported to be stable within *Ec*N both *in vitro* and *in vivo* (Kan et al., 2021). However, introducing a production cassette into pMUT would appear to lead to mutations being identified which silence the expression of the heterologous constructs. The expression of production genes was either interrupted by smaller changes resulting in stop codons, or by larger insertions of mainly the IS3-like element ISEc52 family transposase. ISEc52 was found to be the most abundant IS element within the genome of *Ec*N (Supplement Table S2) and therefore also prone to be the most frequently inserted mobile element. Notably, the interference by IS elements, represents a commonly observed challenge during the scale-up of fermentation processes (Heieck et al., 2023; Rugbjerg et al., 2018a, 2021). Consequently, the disruption of the expression of heterologous genes allowed the cell to allocate resources towards growth. Additionally, the observed mutation in the hok/sok locus for the IL-2 producing population was presumably due to the cell counteracting its stabilizing effect to reduce the additional production load.

In this study, the primary objective was to address the high mutational rate by eliminating the endogenous error-prone DNA polymerases in two AMT strains. This approach led to a reduction of mutation counts of 12.5–43.5% within the two production strains. Notably, the ΔEPP *Ec*N IL-2 strain exhibited a significant decrease in SNPs compared to the *Ec*N IL-2 counterpart. Nevertheless, this approach proved insufficient, as residual mutations were identified in production-related loci situated on the pMUT-based plasmid. Consistent with previous studies (Lan et al., 2021), the plasmid remained stable for the two control *Ec*N strains.

While the deletion of the three error-prone DNA polymerases has previously been shown to reduce the mutation rate in *E. coli* (Csörgo et al., 2012), the production stability of the selected AMT host could not be improved by deleting the SOS-inducable DNA in this study. However, there are several other opportunities for improvement. In this study, a plasmid-based expression system was used. While this system is easy to manipulate, it comes with the disadvantage of higher genetic instability, as plasmids always pose an additional metabolic burden on the host both by plasmid replication and recombinant gene expression (Bentley et al., 1990; Friehs, 2004; Glick, 1995). To enable the stable and safe development and manufacturing of AMTs, chromosomal integration is likely preferred. However, chromosomally integrated genes would still likely be a target for the mutations observed in this plasmid-based study if there is a large tradeoff in growth vs. production. Another option to stabilize the production is synthetic product addiction (Rugbjerg et al., 2018b) or overlapping the production genes with essential genes (Blazejewski et al., 2019). In addition, creating a genome-reduced *Ec*N host could offer a further option to improve and ensure the genetic stability of AMTs, as previously shown for *E. coli* MS56 (Park et al., 2014). However, for cultivations of strains with minimal genomes, unpredictable phenotypes were reported, which could potentially influence the behavior of the AMT within the human microbiome and therefore require additional investigation (Choe et al., 2019). Further improvement of the strain stability could potentially be achieved through the use of inducible expression systems to shorten the phase where selection is acting on an expressing strain (Vaaben et al., 2022). In this study, a constitutive promoter was used, which will result in the production of the therapeutic active compound throughout the manufacturing process as well as in a patient. However, for the application as AMT, production during the manufacturing process is not required. Hence, a promoter that is only active once the AMT is dosed to the treated patient would likely be beneficial for maintaining the genetic stability of the AMT during manufacturing.

For the application as AMT, it should be further considered that there will be variations between the cultivation conditions during the manufacturing process compared to growth in the human microbiome. During the fermentation process, the system is well-aerated and mixed with a controlled pH, while the microbiome environment is very different from the applied cultivation conditions. The human microbiome is anaerobic, can fluctuate in pH, provides complex substrates, and other microorganisms are present. Therefore, further investigation is recommended to understand the impact of changing environmental conditions on the stability of AMTs.

In conclusion, although AMTs represent a promising new modality of living medicine, they require a better understanding of their vulnerability to mutations during both manufacturing and in the patient host environment. This work contributes essential insights into the need for ongoing research to refine and optimize the genetic stability of AMTs, ensuring their efficacy, safety, and stability across diverse settings. This calls for continued innovation in order to address the unique challenges associated with moving AMTs from a laboratory setting to a clinical environment.

4. Materials and methods

4.1. Strains, plasmids, and oligonucleotide primers

All investigated strains were modified streptomycin-resistant sorts of the wild-type *Escherichia coli* Nissle 1917 (DSM 6601, O6:K5:H1, *Ec*N), cured from any native pMUT1 plasmid. A deletion of the gene *clbP* encoding for an essential peptidase in the maturation of the genotoxic colibactin was introduced (Nougayrède et al., 2021). Knockout (KO) of the error-prone DNA polymerase genes *dinB*, *polB*, and *umuDC* was individually performed by scarless lambda Red integrations of dsDNA using CRISPR/Cas9-mediated counterselection. The specific primers used for each deletion are detailed in Table 3.

The two background strains (*Ec*N ΔclbP and *Ec*N ΔclbP ΔdinB ΔpolB ΔumuDC) were transformed by standard electroporation and three different pMUT1-based plasmids, presented previously in Tumas et al., (2023) and Iannone et al., 2023 (Iannone et al., 2023; Tumas et al., 2023) (see Table 4). The corresponding pMUT-based plasmid maps are shown in Supplementary Material (Supplement Fig. S4). The deletions and production plasmids were verified using PCR (Supplement Fig. S5) and Sanger sequencing using flanking primers (Supplement Table S3). An overview of the six investigated strains is given in Table 5.

4.2. Media and cultivation setup

For all serial passaging cultivations, liquid minimal medium with the following composition was used: 7.6 g/L $(NH_4)_2SO_4$, 30 g/L KH_2PO_4, 68 g/L Na_2HPO_4, 1.86 g/L Citric acid monohydrate, 4 g/L Glucose, 0.1 g/L Thiamine hydrochloride, 1.5 mL/L Trace metal solution (Supplement Table 2), 1 mL/L 2M $MgSO_4$, 50 mg/L Kanamycin, 50 mg/L Streptomycin. Precultures of the six *Ec*N strains were inoculated in 15 mL minimal medium from three single isolates on LB-Kanamycin agar plates (50 mg/L kanamycin). Cultures were incubated overnight at 37 °C and stirred at 1200 rpm. These pre-cultures were used to inoculate five biological replicates in 15 mL minimal medium. Automated at-line optical density measurements at 600 nm were performed periodically during the cultivation on a Sunrise plate reader (Tecan Group Ltd., Switzerland). Cells were periodically transferred once OD_{600} reached 0.60 (equal to approximately $OD_{600} = 2.4$ for benchtop plate readers), guided by a predictive script to maintain cells under constant exponential growth and non-limiting substrate conditions. Subsequently, an aliquot of 800 μL of culture was transferred into flasks with fresh minimal medium using an automated liquid handling robot system (LaCroix et al., 2017). The serial passaging experiments were generally terminated after an average duration of 9–10 days, aligning with the predetermined criterion of achieving 150 generations. The number of generation (N) was calculated based on the growth rate (μ) and the cultivation time (t) based on the following formula:

Table 3
Genomic coordinates of the deletions.

Gene (Polymerase ID[a])	Primer Sequence	Description
polB (P21189)	AAATGGTATCTGGCGAACTCTTTTTTTTGCCTGCGCCACGCTGAAAATCC	polB LFR forward primer
	AGGCATTCCCGTCTATTGGC	polB LFR reverse primer
	CGGTTTGCTGAACACCAGTT	*polB* RFR forward primer
	GCAAAAAAAAGAGTTCGCCAG	*polB* RFR reverse primer
dinB (Q47155)	CCAGTGTTGAGAGGTGAGCAAACTGGTGCTGGGATTATGAATAAC	*dinB* LFR forward primer
	GCAGTTTGTGGAATAAGAAAACGC	*dinB* LFR reverse primer
	CCAGAACCGCAATATGTTTGAAC	*dinB* RFR forward primer
	TGCTCACCTCTCAACACTGG	*dinB* RFR reverse primer
umuDC (P04152)	GTTGTTTATCAAGCCTGCGGATCTACATAGCGGCAGGAAAAAGC	*umuDC* LFR forward
	TCTGGCACTGGAACTGACGG	*umuDC* LFR reverse
	CATTTTTTGGGCCGCTTCTTTTAG	*umuDC* RFR forward
	AGATCCGCAGGCTTGATAAACAAC	*umuDC* RFR reverse

[a] IDs obtained from the UniProt database www.uniprot.org.

Table 4
Plasmid overview.

Plasmid	Relevant features	Product	Reference
pMUT hok empty	kan^R, J23107	–	Armetta et al. (2021)
IL-2	kan^R, ompA secretion signal, His6, J23107,	IL-2	Tumas et al. (2023)
Aldafermin	kan^R, ompA secretin signal, J23107	Aldafermin (NGM282)	Iannone et al. (2023)

Table 5
Overview of the different *Ec*N strains investigated in this study.

Full Strain name	Abbreviation
E. coli Nissle 1917 ΔclbP control	*Ec*N control
E. coli Nissle 1917 ΔclbP IL-2	*Ec*N IL-2
E. coli Nissle 1917 ΔclbP Aldafermin	*Ec*N Aldaf.
E. coli Nissle 1917 ΔclbP ΔdinB ΔpolB ΔumuDC control	ΔEPP *Ec*N control
E. coli Nissle 1917 ΔclbP ΔdinB ΔpolB ΔumuDC IL-2	ΔEPP *Ec*N IL-2
E. coli Nissle 1917 ΔclbP ΔdinB ΔpolB ΔumuDC aldafermin	ΔEPP *Ec*N Aldaf.

$$N = \frac{\mu * t}{\ln(2)}$$

This approach ensured consistent experimental conditions and comparability of results among different experiments. Approximately once a day, samples were withdrawn for cryopreservation. A 25% glycerol cryo stock was prepared by mixing 500 μL cell suspension with 500 μL 50% glycerol and stored at −70 °C. After centrifugation (5 min, 13.000 rpm), the supernatant and pellet were stored separately at −20 °C.

4.3. Offline growth kinetics

In addition to the previously described OD_{600} at-line measurement, more detailed growth kinetics were recorded using the Growth Profiler 960 (Enzyscreen, Heemstede, Netherlands). Therefore, 300 μL of culture medium were inoculated with the glycerol cryo stocks in 96 square deep well MTP. All cultures were grown overnight at 37 °C and 250 rpm using the Duetz system to avoid evaporation. Afterwards, 300 μL of new minimal medium were inoculated using a replicator, to ensure identical starting conditions for all strains. For the cultivation, polystyrene greyish-white square 96-half-deepwell microplates with a transparent bottom (-CR1496dg) were used. The green value readings were converted into OD_{600} values following the manufacturer's instructions. Growth rates were subsequently calculated from the OD_{600} trajectories using the QurvE 1.1 software (Wirth et al., 2023).

4.4. Whole genome sequencing and data analysis

Whole genome sequencing was performed using the entire endpoint populations. Genomic DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen), and the libraries were prepared using the Plexwell 384 Library Prep kit (SeqWell). Sequencing was performed using a NextSeq machine (Illumina) using NextSeq high Output kit V2 a 300 cycle (150bp x 2) from Illumina (San Diego, CA) with an average coverage of 150X for each sample. The sequencing files were analyzed using a previously described mutation identification and databasing workflow (V Phaneuf et al., 2019) to align reads to reference genomes (NZ_CP022686.1, NZ_CP058219) using bowtie2 (Langmead and Salzberg, 2012) and BreSeq (Deatherage and Barrick, 2014) to identify mutations. BreSeq employs a statistical approach to determine mutation frequencies in the population samples, finding the maximum likelihood of frequencies over an entire range from 0% to 100%. The mutation tables can be found on the ALEdb database (aledb.org) under the project name "Nissle_ALE_sim.".

CRediT authorship contribution statement

Lara P. Munkler: Writing – original draft, Visualization, Methodology, Investigation, Data curation. **Elsayed T. Mohamed:** Writing – review & editing, Resources, Methodology, Data curation. **Ruben Vazquez-Uribe:** Writing – review & editing, Supervision, Investigation, Data curation, Conceptualization. **Victoria Visby Nissen:** Writing – review & editing, Investigation, Data curation. **Peter Rugbjerg:** Writing – review & editing, Supervision, Conceptualization. **Andreas Worberg:** Writing – review & editing, Supervision, Funding acquisition. **John M. Woodley:** Writing – review & editing, Supervision, Funding acquisition. **Adam M. Feist:** Writing – review & editing, Validation, Supervision, Conceptualization. **Morten O.A. Sommer:** Writing – review & editing, Supervision, Project administration, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

The authors gratefully acknowledge the scientific support of Josefin Johnsen (Novo Nordisk Foundation Center for Biosustainability, Lyngby, Denmark). This work was funded by the Novo Nordisk Foundation within the framework of the Fermentation-based

Biomanufacturing Initiative (FBM), grant number: NNF17SA0031362, and The Novo Nordisk Foundation under NNF grant number: NNF20CC0035580.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymben.2024.08.001.

References

Armetta, J., Schantz-Klausen, M., Shepelin, D., Vazquez-Uribe, R., Bahl, M.I., Laursen, M. F., Licht, T.R., Sommer, M.O.A., 2021. *Escherichia coli* promoters with consistent expression throughout the murine gut. ACS Synth. Biol. 10, 3359–3368.

Arora, T., Wegmann, U., Bobhate, A., Lee, Y.S., Greiner, T.U., Drucker, D.J., Narbad, A., Bäckhed, F., 2016. Microbially produced glucagon-like peptide 1 improves glucose tolerance in mice. Mol Metab 5, 725–730.

Bentley, W.E., Mirjalili, N., Andersen, D.C., Davis, R.H., Kompala, D.S., 1990. Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. Biotechnol. Bioeng. 35, 668–681.

Blazejewski, T., Ho, H.-I., Wang, H.H., 2019. Synthetic sequence entanglement augments stability and containment of genetic information in cells. Science 365, 595–598 (1979).

Charbonneau, M.R., Isabella, V.M., Li, N., Kurtz, C.B., 2020. Developing a new class of engineered live bacterial therapeutics to treat human diseases. Nat. Commun. 11, 1738.

Chen, Z., Guo, L., Zhang, Y., Walzem, R.L., Pendergast, J.S., Printz, R.L., Morris, L.C., Matafonova, E., Stien, X., Kang, L., Coulon, D., McGuinness, O.P., Niswender, K.D., Davies, S.S., 2014. Incorporation of therapeutically modified bacteria into gut microbiota inhibits obesity. J. Clin. Invest. 124, 3391–3406.

Choe, D., Lee, J.H., Yoo, M., Hwang, S., Sung, B.H., Cho, S., Palsson, B., Kim, S.C., Cho, B.-K., 2019. Adaptive laboratory evolution of a genome-reduced *Escherichia coli*. Nat. Commun. 10, 935.

Csörgo, B., Fehér, T., Tímár, E., Blattner, F.R., Pósfai, G., 2012. Low-mutation-rate, reduced-genome *Escherichia coli*: an improved host for faithful maintenance of engineered genetic constructs. Microb. Cell Factories 11.

Deatherage, D.E., Barrick, J.E., 2014. Identification of Mutations in Laboratory-Evolved Microbes from Next-Generation Sequencing Data Using Breseq, pp. 165–188.

Friehs, K., 2004. Plasmid Copy Number and Plasmid Stability, pp. 47–82.

Gerdes, K., 1988. The parB (hok/sok) locus of plasmid R1: a general purpose plasmid Stabilization system. Nat. Biotechnol. 6, 1402–1405.

Glick, B.R., 1995. Metabolic load and heterologous gene expression. Biotechnol. Adv. 13, 247–261.

Heieck, K., Arnold, N.D., Brück, T.B., 2023. Metabolic stress constrains microbial L-cysteine production in *Escherichia coli* by accelerating transposition through mobile genetic elements. Microb. Cell Factories 22.

Iannone, V., Babu, A.F., Lok, J., Gómez-Gallego, C., D'Auria, G., Vazquez-Uribe, R., Vaaben, T.H., Bongers, M., Mikkonen, S., Vaittinen, M., Tikkanen, I., Kettunen, M., Klåvus, A., Sehgal, R., Kaminska, D., Pihlajamaki, J., Hanhineva, K., El-Nezami, H., Sommer, M.O.A., Kolehmainen, M., 2023. Changes in liver metabolic pathways demonstrate efficacy of the combined dietary and microbial therapeutic intervention in MASLD mouse model. Mol Metab 78, 101823.

Jacobi, C.A., Malfertheiner, P., 2011. *Escherichia coli* Nissle 1917 (mutaflor): new insights into an old probiotic bacterium. Dig. Dis. 29, 600–607.

Juarez, V.M., Montalbine, A.N., Singh, A., 2022. Microbiome as an immune regulator in health, disease, and therapeutics. Adv. Drug Deliv. Rev. 188, 114400.

Kan, A., Gelfat, I., Emani, S., Praveschotinunt, P., Joshi, N.S., 2021. Plasmid vectors for in vivo selection-free use with the probiotic *E. coli* Nissle 1917. ACS Synth. Biol. 10, 94–106.

Knöppel, A., Knopp, M., Albrecht, L.M., Lundin, E., Lustig, U., Näsvall, J., Andersson, D. I., 2018. Genetic adaptation to growth under laboratory conditions in *Escherichia coli* and *Salmonella enterica*. Front. Microbiol. 9.

LaCroix, R.A., Sandberg, T.E., O'Brien, E.J., Utrilla, J., Ebrahim, A., Guzman, G.I., Szubin, R., Palsson, B.O., Feist, A.M., 2015. Use of adaptive laboratory evolution to discover key mutations enabling Rapid growth of *Escherichia coli* K-12 MG1655 on glucose minimal medium. Appl. Environ. Microbiol. 81, 17–30.

LaCroix, R.A., Palsson, B.O., Feist, A.M., 2017. A model for designing adaptive laboratory evolution experiments. Appl. Environ. Microbiol. 83, e03115. -16.

Lan, Y.-J., Tan, S.-I., Cheng, S.-Y., Ting, W.-W., Xue, C., Lin, T.-H., Cai, M.-Z., Chen, P.-T., Ng, I.-S., 2021. Development of *Escherichia coli* Nissle 1917 derivative by CRISPR/Cas9 and application for gamma-aminobutyric acid (GABA) production in antibiotic-free system. Biochem. Eng. J. 168, 107952.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Nougayrède, J.-P., Chagneau, C.V., Motta, J.-P., Bossuet-Greif, N., Belloy, M., Taieb, F., Gratadoux, J.-J., Thomas, M., Langella, P., Oswald, E., 2021. A toxic Friend: genotoxic and mutagenic Activity of the probiotic strain *Escherichia coli* Nissle 1917. mSphere 6.

Park, M.K., Lee, S.H., Yang, K.S., Jung, S.-C., Lee, J.H., Kim, S.C., 2014. Enhancing recombinant protein production with an *Escherichia coli* host strain lacking insertion sequences. Appl. Microbiol. Biotechnol. 98, 6701–6713.

Praveschotinunt, P., Duraj-Thatte, A.M., Gelfat, I., Bahl, F., Chou, D.B., Joshi, N.S., 2019. Engineered *E. coli* Nissle 1917 for the delivery of matrix-tethered therapeutic domains to the gut. Nat. Commun. 10, 5580.

Reuven, N.B., Arad, G., Maor-Shoshani, A., Livneh, Z., 1999. The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. J. Biol. Chem. 274, 31763–31766.

Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K., Sommer, M.O.A., 2018a. Diverse genetic error modes constrain large-scale bio-based production. Nat. Commun. 9.

Rugbjerg, P., Sarup-Lytzen, K., Nagy, M., Sommer, M.O.A., 2018b. Synthetic addiction extends the productive life time of engineered Escherichia coli populations. Proc. Natl. Acad. Sci. USA 115, 2347–2352.

Rugbjerg, P., Dyerberg, A.S.B., Quainoo, S., Munck, C., Sommer, M.O.A., 2021. Short and long-read ultra-deep sequencing profiles emerging heterogeneity across five platform *Escherichia coli* strains. Metab. Eng. 65, 197–206.

Scaldaferri, F., Gerardi, V., Mangiola, F., Lopetuso, L.R., Pizzoferrato, M., Petito, V., Papa, A., Stojanovic, J., Poscia, A., Cammarota, G., Gasbarrini, A., 2016. Role and mechanisms of action of *Escherichia coli* Nissle 1917 in the maintenance of remission in ulcerative colitis patients: an update. World J. Gastroenterol. 22, 5505.

Schultz, M., 2008. Clinical use of *E. coli* Nissle 1917 in inflammatory bowel disease. Inflamm. Bowel Dis. 14, 1012–1018.

Sharma, S.S., Blattner, F.R., Harcum, S.W., 2007. Recombinant protein production in an *Escherichia coli* reduced genome strain. Metab. Eng. 9, 133–141.

Sonnenborn, U., 2016. *Escherichia coli* strain Nissle 1917—from bench to bedside and back: history of a special Escherichia coli strain with probiotic properties. FEMS Microbiol. Lett. 363, fnw212.

Steidler, L., Neirynck, S., Huyghebaert, N., Snoeck, V., Vermeire, A., Goddeeris, B., Cox, E., Remon, J.P., Remaut, E., 2003. Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10. Nat. Biotechnol. 21, 785–789.

Takors, R., 2012. Scale-up of microbial processes: Impacts, tools and open questions. J. Biotechnol. 160, 3–9.

Tumas, S., Meldgaard, T.S., Vaaben, T.H., Suarez Hernandez, S., Rasmussen, A.T., Vazquez-Uribe, R., Hadrup, S.R., Sommer, M.O.A., 2023. Engineered *E. coli* Nissle 1917 for delivery of bioactive IL-2 for cancer immunotherapy. Sci. Rep. 13, 12506.

V Phaneuf, P., Gosting, D., Palsson, B.O., Feist, A.M., 2019. ALEdb 1.0: a database of mutations from adaptive laboratory evolution experimentation. Nucleic Acids Res. 47, D1164–D1171.

Vaaben, T.H., Vazquez-Uribe, R., Sommer, M.O.A., 2022. Characterization of eight bacterial Biosensors for microbial diagnostic and therapeutic applications. ACS Synth. Biol. 11, 4184–4192.

Wirth, N.T., Funk, J., Donati, S., Nikel, P.I., 2023. QurvE: user-friendly software for the analysis of biological growth and fluorescence data. Nat. Protoc. 18, 2401–2403.

Xia, J., Wang, G., Lin, J., Wang, Y., Chu, J., Zhuang, Y., Zhang, S., 2015. Advances and Practices of Bioprocess Scale-Up, pp. 137–151.

Yu, X., Lin, C., Yu, J., Qi, Q., Wang, Q., 2020. Bioengineered *Escherichia coli* Nissle 1917 for tumour-targeting therapy. Microb. Biotechnol. 13, 629–636.

Glossary

AMT: Advanced Microbiome Therapeutics
EcN: Escherichia coli Nissle 1917
EPP: Error-prone DNA polymerases
IL-2: Interleukin 2
IS: Insertion sequence
KO: Knockout
NGS: Next generation sequencing
SNP: Single-nucleotide polymorphism
ΔEPP: Knockout of dinB, polB, umuDC