Engineering Microbiology 2 (2022) 100012

Contents lists available at ScienceDirect

Engineering Microbiology

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

Original Research Article

A comparative study on the genomes, transcriptomes, and metabolic properties of *Escherichia coli* strains Nissle 1917, BL21(DE3), and MG1655

Linlin Zhao[a,b], Guobin Yin[a,b], Yonglin Zhang[a,b], Chaofan Duan[a,c], Yang Wang[a,b], Zhen Kang[a,b,*]

[a] The Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122, China
[b] The Science Center for Future Foods, Jiangnan University, Wuxi 214122, China
[c] The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, China

ARTICLE INFO

Keywords:
Escherichia coli Nissle 1917
probiotics
comparative genomics
transcriptome
synthetic biology

ABSTRACT

Escherichia coli is the most well-studied model prokaryote and has become an indispensable host for the biotechnological production of proteins and biochemicals. In particular, the probiotic status of one *E. coli strain, E. coli* Nissle 1917 (EcN) has helped it become a new favorite amongst synthetic biologists. To broaden its potential applications, here we assemble a comparative study on the genomes, transcriptomes, and metabolic properties of *E. coli* strains EcN, BL21(DE3), and MG1655. Comparative genomics data suggests that EcN possesses 1404 unique CDSs. In particular, EcN has additional iron transport systems which endow EcN with a higher tolerance to iron scarcity when compared to two other *E. coli* strains. EcN transcriptome data demonstrates that *E. coli* strains EcN, BL21(DE3), and MG1655 all have comparable activities of the central metabolic pathway, however only EcN inherits the arginine deiminase pathway. Additionally, we found that EcN displayed a lower expression of ribosomal proteins compared to BL21(DE3) and MG1655. This comparative study on *E. coli* strains EcN, BL21(DE3), and MG1655 aims to provide a reference for further engineering EcN as a biotechnological tool.

1. Introduction

Escherichia coli Nissle 1917 (EcN) has been used as an active pharmaceutical ingredient in medical products targeting intestinal health for more than 100 years (Sonnenborn, 2016). It was isolated from the stool sample of a German soldier who was not affected by Shigella-induced diarrhea (Sonnenborn and Schulze, 2009). EcN displays anti-inflammatory properties (Fabrega et al., 2017, Sarate et al., 2019), antagonistic and immunomodulatory activities, and inhibits the growth of pathogenic bacteria (Trebichavsky et al., 2010, Helmy et al., 2021, Rund et al., 2013), thus positively affecting gastrointestinal homeostasis and microbial balance. It has been used in the treatment of acute diarrhea in infants and toddlers (Henker et al., 2008, Henker et al., 2007). Owing to its probiotic properties, researchers are exploiting EcN as a platform for developing probiotic delivery system. Research on EcN-mediated tumor therapy has shown that engineered EcN can selectively colonize and replicate in solid tumors (Yu et al., 2020). Moreover, EcN has been engineered for the use of delivering drugs and producing cytotoxic compounds in tumor-bearing mice (Praveschotinunt et al., 2019, Li et al., 2019, He et al., 2019).

E. coli has the advantage of a well-characterized genetic background, short doubling period, easy cultivation, and plenty of tools for genetic engineering (Pontrelli et al., 2018). Compared to the strains MG1655 and BL21(DE3), EcN possesses the same strengths but may hold more biosafety in industrial applications. Moreover, the unique genes of EcN confer distinctive metabolic capabilities, such as the synthesis of capsular polysaccharide heparsoan, a precursor of heparin (Datta et al., 2021). EcN has many applications for treating diseases, but there are few reports evaluating the use of EcN as a platform for biotechnological applications.

In this study, we compared the genomes of EcN, BL21(DE3), and MG1655, and then identified coding sequences (CDSs) that were unique to EcN. In consideration of the fact that EcN had multiple iron uptake systems (Deriu et al., 2013, Weiss, 2013), we also compared the growth of EcN, BL21(DE3), and MG1655 in response to limited iron availability. To better understand the differences between EcN, BL21(DE3), and MG1655, we performed a basic transcriptome analysis of EcN, BL21(DE3), and MG1655. When we analyzed gene expression, we found that the arginine deiminase pathway in EcN, which generates energy through the degradation of arginine, was highly expressed. We also found that some ribosomal proteins had low levels of expression in EcN, which may have consequences for translation. Therefore, we assessed

* Corresponding author.
E-mail address: zkang@jiangnan.edu.cn (Z. Kang).

https://doi.org/10.1016/j.engmic.2022.100012
Received 17 October 2021; Received in revised form 14 February 2022; Accepted 14 February 2022
Available online 21 February 2022

Table 1
Bacterial strains and plasmids used in this study.

Strain	Features	Source
E. coli JM109	*endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB*$^+$ Δ(*lac-proAB*) *e14*$^-$ [F' *traD36 proAB*$^+$ *lacIq lacZ*ΔM15] *hsdR17*($r_K^-m_K^+$)	New England Biolabs
E. coli BL21(DE3)	*E. coli* str. B F$^-$ *ompT gal dcm lon hsdSB*($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5*-T7*p07 ind1 sam7 nin5*]) [*malB*+] K-12(λS)	New England Biolabs
E. coli MG1655	Wild type K-12 strain; F$^-$ λ$^-$ *ilvG*$^-$ *rfb*-50 *rph*-1	Lab stock
E. coli Nissle 1917	Wild type EcN; serotype O6:K5:H1	Lab stock
EcN-119	EcN derivate, carrying pCOLA-119	This study
BL21(DE3)-119	BL21(DE3) derivate, carrying pCOLA-119	This study
MG1655-119	MG1655 derivate, carrying pCOLA-119	This study
Plasmids		
pCOLADuet-1	Kan^R,T7 promoter, ColA origin, *lacI*	Novagen
pCOLA-GFP	The pCOLADuet-1 plasmid expression *gfp* gene	This study
pCOLA-119	The *gfp* gene under the control of J23119 promoter in pCOLADuet-1	This study

the expression of green fluorescence protein (GFP) in EcN, BL21(DE3), and MG1655. In order to understand the metabolic characteristics of these strains, we compared the growth curve, glucose consumption, and acetic acid production of EcN, BL21(DE3), and MG1655.

2. Materials and methods

2.1. Bacterial strains and plasmid construction

All the *E. coli* strains and plasmids used in this study are listed in Table 1. The primers used for recombinant plasmid construction are listed in Supplementary data 1. *E. coli* JM109 was used for cloning experiments. To construct a GFP-expression plasmid, the *gfp* gene was amplified by PCR with the pBBRIMCS-2-GFP plasmid as a template and GFP-F/GFP-R primers. Linearized pCOLADuet-1 was generated by PCR with primers GFP-COLA-F/GFP-COLA-R, which contain homologous terminals of *gfp*. After treatment with DpnI, the linearized vector and *gfp* were assembled with T5 exonuclease DNA assembly (Xia et al., 2019). Constitutive promoter J23119 was used to replace the T7 promoter of pCOLA-GFP by PCR with primers listed in Supplementary data 1 and *in vivo* self-cyclization of the amplified linear pCOLA-GFP plasmid.

2.2. Medium and culture conditions

E. coli strains were grown in Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) for cloning experiments, growth curve analysis and protein expression capability analysis. Mineral salt medium (2 g/L yeast extract, 15.6 g/L $Na_2HPO_4{\cdot}12H_2O$, 3 g/L KH_2PO_4, 1 g/L NH_4Cl, 0.494 g/L $MgSO_4{\cdot}7H_2O$, 0.0152 g/L $CaCl_2{\cdot}H_2O$, 0.01 g/L $FeSO_4{\cdot}7H_2O$, and 10 g/L glucose) was used to study bacterial glucose consumption and acetic acid synthesis. M9 minimal medium (7.52 g/L $Na_2HPO_4{\cdot}2H_2O$, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 0.5 g/L NH_4Cl, 0.246 g/L $MgSO_4{\cdot}7H_2O$, 0.441 g/L $CaCl_2{\cdot}2H_2O$, 0.01 g/L biotin, 0.01 g/L thiamin, and 10 mL/L 100× trace elements solution) with 83mg/L or 0.83 mg/L $FeCl_3{\cdot}6H_2O$ was used to investigate bacterial iron limitation response. The 100× trace elements solution includes 5g/L EDTA, 84 mg/L $ZnCl_2$, 13 mg/L $CuCl_2{\cdot}2H_2O$, 10 mg/L $CoCl_2{\cdot}2H_2O$, 10 mg/L H_3BO_3, and 1.6 mg/L $MnCl_2{\cdot}4H_2O$. When necessary, 50 μg/mL kanamycin was supplemented to maintain plasmid stability. All the strains were cultivated at 37°C and 220 rpm.

2.3. Genome alignment

Genome comparative analysis between *E. coli* BL21(DE3) (NC_012971.2), *E. coli* MG1655 (CP012868.1), and EcN (NZ_CP022686.1) was conducted by Mauve software with the progressiveMauve algorithm (Darling et al., 2004).

2.4. Protein function prediction

The unique proteins of EcN were classified according to the predicted function or conserved domains of the protein by InterPro (Blum et al., 2021) and NCBI Conserved Domain Database (Marchler-Bauer et al., 2015). The original classification results of the unique genes for EcN are listed in Supplementary data 2.

2.5. Transcriptome analysis

To prepare the transcriptome samples, *E. coli* strains were cultivated to early exponential phase (OD_{600} 2). Cells were collected by 10,000 *g* centrifugation 4°C, and then frozen with liquid nitrogen. The cDNA library was constructed utilizing high-quality RNA and sequenced on an Illumina HiSeq platform at Sangon Biotech (Shanghai). Raw reads were obtained and then processed by Trimmomatic according to the following rules (Bolger et al., 2014): 1) removing sequences with N bases; 2) removing the linker sequence in reads; 3) removing low-quality bases (Q value < 20); 4) removing the bases that the tail quality value was less than 20 (the window size was 5 bp); 5) removing reads less than 35nt and their paired reads. The processed reads of each sample were respectively mapped onto their reference genomes using Bowtie2 (Langdon, 2015). The expression levels of genes were quantified using featureCounts (Liao et al., 2014).

Cluster analysis was performed to evaluate the expressed genes in different strains by the transcripts per million (TPM) values (Benesty et al., 2009). It was suggested that *recA* is a stably expressed reference gene in various bacteria (Takle et al., 2007, Gomes et al., 2018, McMillan and Pereg, 2014). To eliminate the interference caused by total CDS differences, the TPM of a gene was normalized to the TPM of *recA* gene from the same strain before subjected to comparison with the transcription level of its ortholog in the other two strains. The whole transcriptome results can be found in Supplementary data 3 and the differentially expressed genes are listed in Supplementary data 4.

2.6. Fluorescence assay

The appropriately diluted cultures were added to a 96-well black plate (Corning® 96 Well Transparent Flat Bottom). The GFP fluorescence intensity was quantified with the plate reader Infinite 200 PRO (Tecan, Austria) at an excitation wavelength of 490 nm and an emission wavelength of 530 nm. The gain value was set to 55.

2.7. Measurement of acetic acid concentration

Acetic acid concentration was measured using a 1260 Infinity II liquid chromatograph system (Agilent, USA) with an organic acid column (2.1 × 100 mm, 2.7 μm). The column temperature was set to 40°C, the flow rate was 0.5 mL/min, and the mobile phase was 0.05 mM sulfuric acid (Xie et al., 2011). Acetic acid concentration quantification was carried out from integrated peak areas of the samples using the corresponding standard curves.

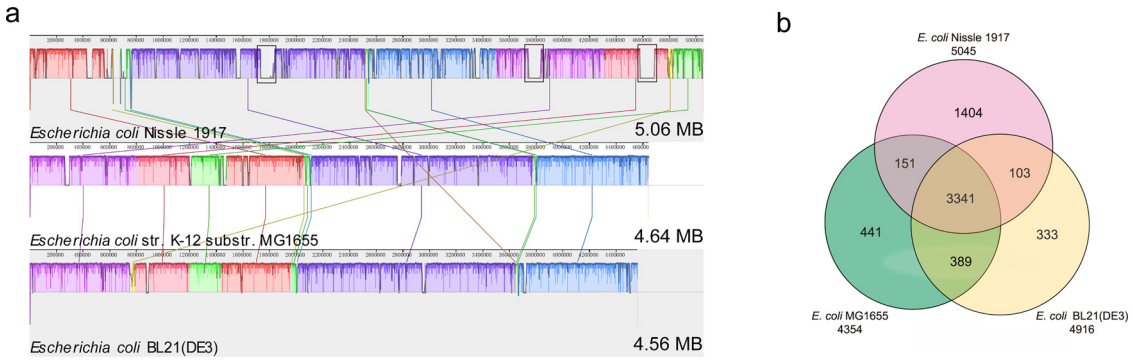

Fig. 1. Comparative genomes and growth of EcN, BL21(DE3), and MG1655. (a) Genome alignment of the *E. coli* BL21(DE3) (NC_012971.2), *E. coli* MG1655 (CP012868.1), and *E. coli* Nissle 1917 (NZ_CP022686.1) with Mauve. Colored regions represent conserved DNA regions shared by two or more genomes. The white gaps between the rectangles mean no similar fragments were found in other genomes. (b) A Venn diagram shows the number of common or unique genes in EcN, BL21(DE3), and MG1655.

Table 2
Classification of EcN's unique CDSs.

Protein categories	Number of CDS
Type VI secretion system	20
Colibactin related	38
Nuclease	78
Phage related protein	41
Fimbrial protein	29
Transposase	75
Transporter	125
Transcriptional regulator	80
Metabolic pathway/Enzyme	304
Toxin/Anti-toxin	36
Hypothetical protein	532
Unpredicted	60

3. Results and Discussion

3.1. Comparative genomics of EcN, BL21(DE3), and MG1655

We performed genome alignment of EcN, BL21(DE3), and MG1655 by Mauve (Darling et al., 2004). The genome comparison results are shown in Fig. 1a. Colored regions indicate conserved segments shared by three genomes while the gaps indicate unaligned regions. Compared to BL21(DE3) and MG1655, EcN contains more unique genomic regions, as also demonstrated by the larger genome size of EcN (Fig. 1a). The Venn diagram shows common or unique CDSs among the three bacteria (Fig. 1b). EcN contained more unique CDSs (1404) when compared to BL21(DE3) (441) and MG1655 (333).

3.2. Classification of 1404 unique EcN CDS

To better understand the genomic differences between these strains, we further classified the unique CDS of EcN. As shown in Table 2 and Supplementary data 2, 304 of the 1404 unique CDSs are related to metabolic activities. Among all three strains, only EcN has a type VI secretion system, which has been reported to contribute to the enhancement of enterohemorrhagic *E. coli* virulence (Wan et al., 2017). We predicted the type VI secretion system proteins of EcN through SecReT6 (Li et al., 2015). EcN has a complete type VI secretion system, which is mainly composed of three clusters on the genome (Fig. 2a, b). EcN also has a *pks* locus that codes for colibactin, which closely relates it to pathogenic *E. coli* strains. Moreover, EcN possesses 38 colibactin synthesis-related genes. There are 36 toxins or anti-toxins found within the EcN's unique CDS list (Table 2). The toxin/anti-toxin pairs in other bacteria were reported in response to a range of stresses and nutritional stimuli. These toxin/anti-toxin pairs went on to play a role in bacterial persistence, multidrug tolerance, and biofilm formation (Hayes and Van Melderen, 2011, Wang et al., 2015, Kamruzzaman et al., 2021). We found 125 unique CDSs that encode for transporters, with three of those being extra iron transport systems. This is consistent with reports that probiotic *E. coli* strains are equipped with more iron siderophore and sugar transporters than *E. coli* K-12 strains (Do et al., 2017). Other unique CDSs encode proteins that were predicted to be transposases, regulators, and more. (Table 2 and Supplementary data 2).

3.3. Tolerance of EcN, BL21(DE3), and MG1655 to iron scarcity

By classification of EcN's unique CDSs, we found that EcN had additional siderophores, such as salmochelin, yersiniabactin, and hydroxamate aerobactin to transport iron. Due to the high-affinity of siderophores to ferric iron, they can capture ferric iron from ferritin and transferrins. After ferric iron enters bacteria, it is reduced to divalent iron with little affinity to siderophores and then released (Fig. 3a). We further explored whether the iron transport system can make EcN more tolerant to the iron limitation. We cultivated EcN, BL21(DE3), and MG1655 in M9 minimal medium with a normal concentration of iron (31μM) or 1% of the normal concentration of iron (0.31μM). As expected, the growth of EcN was less affected by iron scarcity than the growth of both MG1655 and BL21(DE3). The latter two strains showed growth retardation at low iron concentrations. Compared to normal iron concentration conditions found in M9 minimal media, the final OD_{600} of all three bacteria was reduced at 1% of the normal iron concentration (Fig. 3b, c). This result is in accordance with the conclusion that EcN competes for iron with pathogenic bacteria and thus provides protection against pathogenicity (Deriu et al., 2013).

3.4. Comparing the transcriptomic difference between EcN, BL21(DE3), and MG1655

To better elucidate the differences between the transcriptomes of EcN, BL21(DE3), and MG1655, we conducted a basic transcriptome analysis. We first focused on the differentially expressed genes between EcN, BL21(DE3), and MG1655. We found 170 differentially expressed gene orthologs between the three strains, as indicated by the normalized TPM values. In order to make these comparisons and eliminate the impact of great differences in the total number of CDS (Fig. 1a), the TPM of any gene was normalized to the TPM of *recA* gene (Supplementary data 4). Cluster analysis based on Pearson correlation distance was performed to get more insight on the similarities between their transcription patterns. The detailed map, framed by the cyan rectangle, shows genes whose expression levels are more similar to those of BL21(DE3). We found 75 and 32 genes of EcN showed more similarity to the nor-

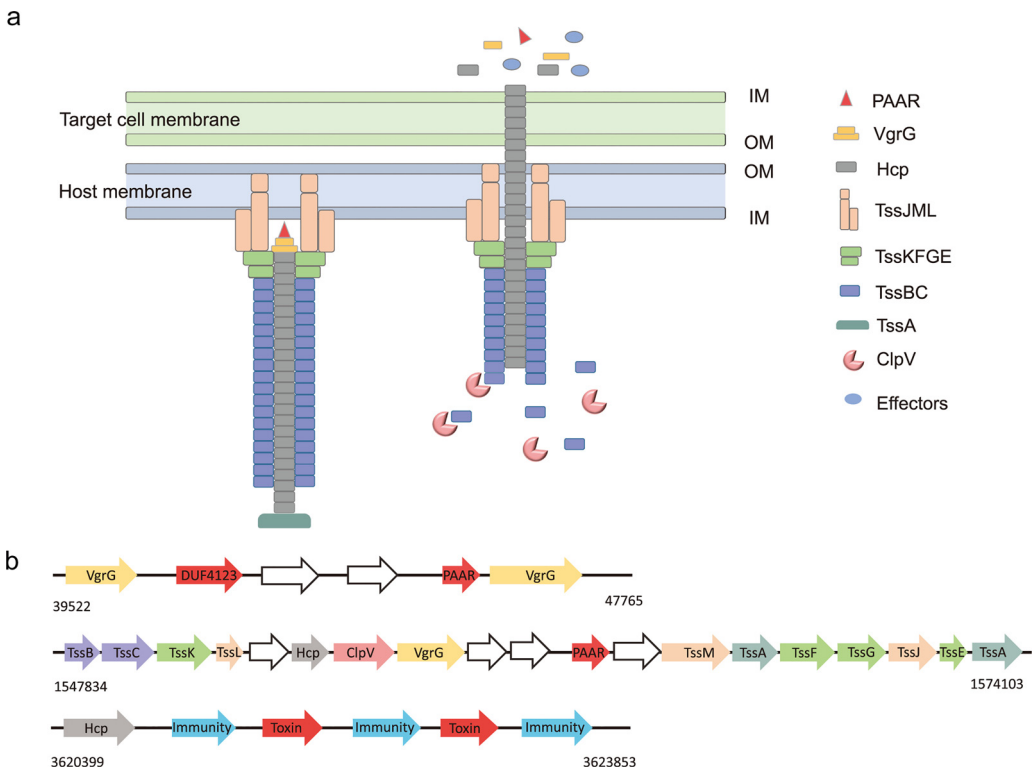

Fig. 2. The type VI secretion system. (a) Diagrammatic depiction of the type VI secretion system apparatus extension and contraction in EcN. The Type VI secretion system is composed of a membrane complex (orange color), a baseplate complex (green color), and a needle like structure (dark blue and grey color). Effectors can be directly delivered or through interaction with VgrG, PAAR, or other chaperones. (b) Genome arrangement of type VI secretion system genes in EcN.

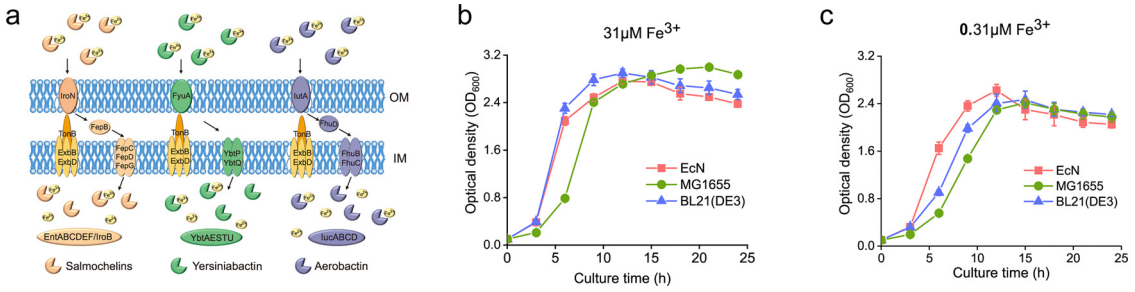

Fig. 3. Growth and acetic acid metabolism of EcN, BL21(DE3), and MG1655 at low iron concentrations. (a) Schematic of siderophore-mediated iron uptake in EcN. Salmochelins synthesis and iron transport proteins are indicated in orange. Yersiniabactin synthesis and iron transport proteins are indicated in green. Aerobactin synthesis and iron transport proteins are indicated in purple. The growth of EcN, BL21(DE3), and MG1655 in (b) M9 minimal medium or (c) M9 minimal medium with 1% of normal iron concentration. The dotted lines were references for the final OD_{600} or growth retardation of EcN, BL21(DE3), and MG1655. Data points represent mean values of three biological replicates with error-bars showing standard deviation.

malized TPM of counterparts in BL21(DE3) and in MG1655, respectively (Fig. 4a).

We then analyzed central metabolic pathways of EcN, BL21(DE3), and MG1655. In general, the central metabolic pathways of these three strains have similar expression levels (Fig. 4b). Among the 34 glycolysis and tricarboxylic acid (TCA) cycle-related genes of three strains, 25 genes belonging to MG1655 have the highest normalized TPM value (Fig. 4b and Supplementary data 5), which indicate that MG1655 has the strongest central carbon metabolism pathway. The metabolic activity level of the glycolysis pathway in BL21(DE3) was in between that of EcN and MG1655, as shown by the normalized TPM value (Fig. 4b and Supplementary data 5). In the same way, we demonstrated that the TCA cycle activity of EcN was in between that of MG1655 and BL21(DE3) (Fig. 4b and Supplementary data 5). Next, we looked at five acetic acid metabolism genes: *poxB*, *pta*, *eutD*, *ackA* and *acs*. The *acs* gene encodes the acetyl-coenzyme A synthetase, which catalyzes the conversion of acetic acid to acetyl-CoA (Fig. 4b). The others 4 genes encode enzymes that lead to the synthesis of acetic acid. MG1655 had the highest normalized TPM value of all the 5 acetic acid metabolism genes while BL21(DE3) showed the lowest normalized TPM value (Fig. 4b).

3.5. *Unique CDSs with high expression level in EcN*

We ranked the TPM values of gene expression for EcN's unique CDSs and listed the top 16 that were expressed. Seven of them encode protein orthologs within the genome of MG1655 but fail to be recognized by Mauve due to heavy synonymous mutations. For the other 9 genes, 6 of them encode proteins with unclear functions (Table 3). Therefore, we focused on the remaining 3 genes: *ompD*, *arcA* and *clbS* (Table 3). The *ompD* gene is most studied in *Salmonella enterica serovar* Typhimurium

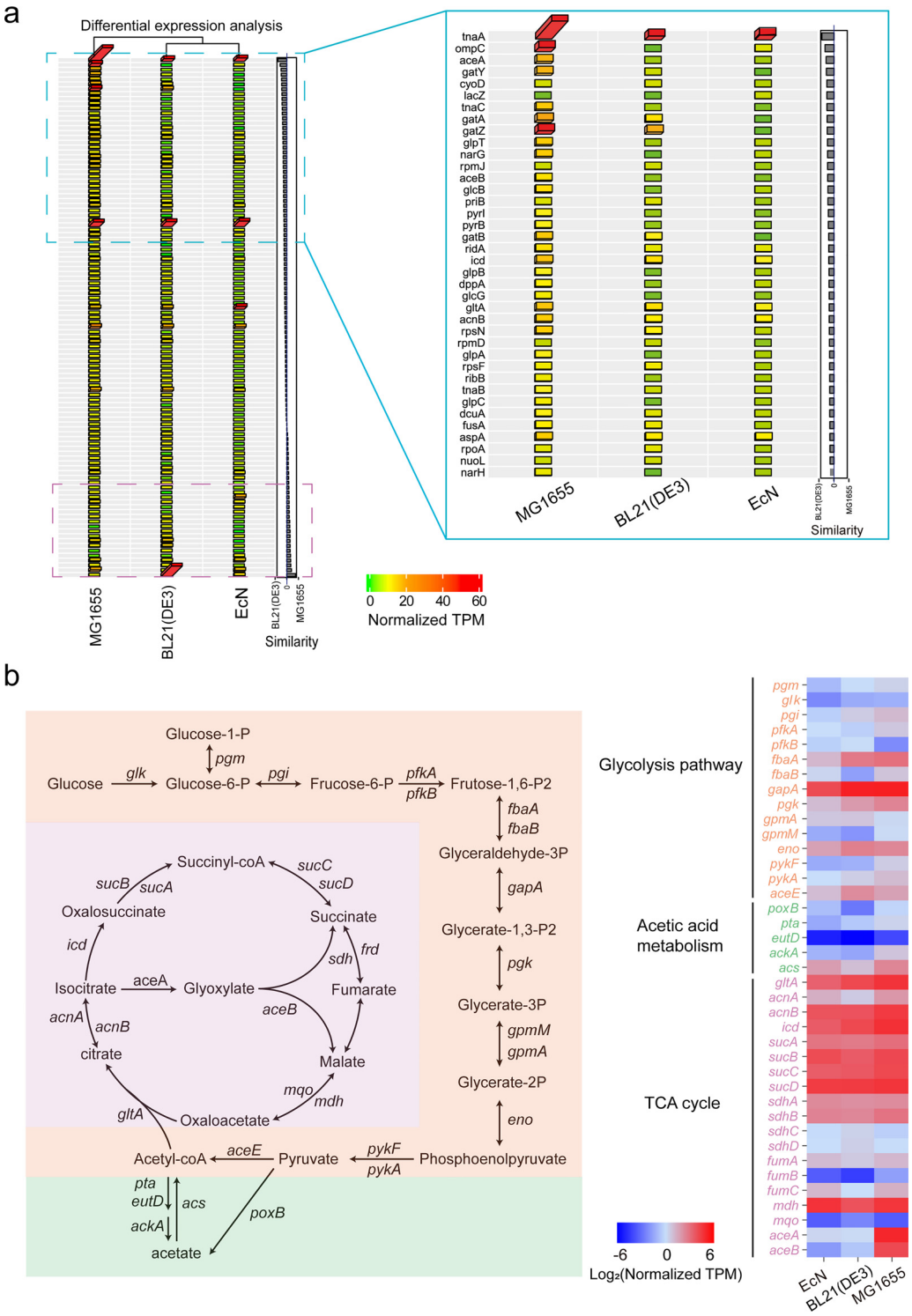

Fig. 4. Transcriptomic comparison of EcN, BL21(DE3), and MG1655. (a) Heat map of differentially expressed genes when comparing EcN, BL21(DE3), and MG1655. The heat map was generated using the normalized TPM values. The colored bar below represents their expression levels. Grey similarity bars were generated by comparing the normalized TPM of gene orthologs. (b) Expression of genes in the central metabolic pathway in the three strains. The glycolysis pathway has orange blocks as its background. *pgm*, phosphoglucomutase; *glk*, glucokinase; *pgi*, glucose-6-phosphate isomerase; *pfkA*, 6-phosphofructokinase; *pfkB*, 6-phosphofructokinase II; *fbaA*, fructose-bisphosphate aldolase; *fbaB*, fructose-bisphosphate aldolase; *gapA*, aldehyde dehydrogenase; *pgk*, phosphoglycerate kinase; *gpmA* and *gpmM*, encode phosphoglycerate mutase; *eno*, enolase; *pykA* and *pykF*, pyruvate kinase; *aceE*, pyruvate dehydrogenase. The TCA cycle has purple blocks as its background. *gltA*, citrate synthase; *acnA* and *acnB*, aconitate hydratase; *icd*, isocitrate dehydrogenase; *sucA* and *sucB*, encode 2-oxoglutarate dehydrogenase; *sucC* and *sucD*, encode succinyl-CoA ligase; *sdhA*, *sdhB*, *sdhC* and *sdhD*, encode succinate dehydrogenase; *fumA*, *fumB* and *fumC*, encode fumarate hydratase; *mdh*, malate dehydrogenase; *mqo*, malate:quinone oxidoreductase; *aceA*, isocitrate lyase; *aceB*, malate synthase. In addition, the acetic acid metabolic pathway has green blocks as its background. *poxB*, pyruvate oxidase; *pta*, phosphate acetyltransferase; *eutD*, phosphate acetyltransferase; *ackA*, acetate kinase; *acs*, acetate kinase. The heat map shows the normalized TPM of the above three pathways' genes. The colored bar below represents their expression levels.

Table 3
The top 16 EcN unique genes ranked according to the TPM value.

EcN gene locus	Gene name	TPM value	Function (encoded protein)
CIW80_02890	*ompD*	5012.12	porin OmpD
CIW80_23470	*fimA*	3995.93	type-1 fimbrial protein, C chain
CIW80_24420	*phoE*	3357.32	phosphoporin PhoE
CIW80_16610	*arcC*	1996.67	carbamate kinase
CIW80_21600	*sdhB*	1822.68	succinate dehydrogenase iron-sulfur subunit
CIW80_16615	*arcA*	1217.66	arginine deiminase
CIW80_11665	*glpD*	1129.3	glycerol-3-phosphate dehydrogenase
CIW80_16605	*argI*	864.08	ornithine carbamoyltransferase
CIW80_21575	Unpredicted	799.39	hypothetical protein
CIW80_02830	*filC*	612.66	flagellin FliC
CIW80_15385	Unpredicted	412.21	quinone oxidoreductase
CIW80_16540	Unpredicted	408.03	soluble cytochrome b562
CIW80_24980	Unpredicted	373.84	hypothetical protein
CIW80_16600	Unpredicted	363.69	YfcC family protein
CIW80_07470	Unpredicted	345.8	Hok/Gef family protein
CIW80_03385	*clbS*	335.67	colibactin self-protection protein ClbS

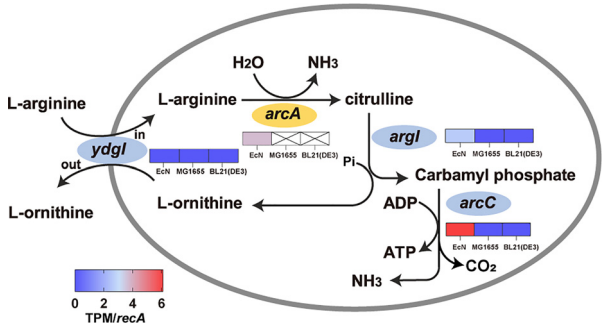

Fig. 5. The arginine deiminase pathway of EcN. Genes with a blue background are shared by these three bacteria, while the *arcA* gene was found in EcN but not in the other two strains. The heat maps below the gene name represent the expression level of the gene. *arcA*, arginine deiminase; *argI*, ornithine carbamoyltransferase; *arcC*, carbamate kinase; *ydgI*, arginine and ornithine antiporter. The colored bar below represents their expression levels.

and the abundance of OmpD increases in response to anaerobiosis and decreases in response to low pH (Santiviago et al., 2003, Sandrini et al., 2013). OmpD is involved in iron uptake and the porin proteins OmpA and OmpC in *E. coli* are utilized in the acquisition of iron (Sandrini et al., 2013, Gerken et al., 2020). Since EcN competes for iron with enteropathogenic bacteria (Deriu et al., 2013), OmpD may also contribute to the acquisition of iron in EcN. We also noticed that the colibactin self-protection protein (ClbS) gene was expressed at a high level (Table 3). ClbS is a cyclopropane hydrolase, which has been reported to protect colibactin-producing bacteria by opening its electrophilic cyclopropane ring (Tripathi et al., 2017). It was speculated that ClbS protects the bacterial genome from degradation by directly interacting with single-stranded DNA and double-stranded DNA (Molan et al., 2019).

The existence of *arcA* gene (encodes arginine deiminase) indicates EcN possesses the arginine deiminase pathway, which is missing from MG1655 and BL21(DE3). The degradation of arginine via the arginine deiminase pathway supports anaerobic metabolism in many pathogens (Zuniga and Gonzalez-Candelas, 2002, Novak et al., 2016). This pathway generates ammonia, ornithine, and carbon dioxide, and coupling the phosphorylation of ADP to ATP, and providing bacteria with energy in the absence of oxygen (Fig. 5). Ammonia and carbon dioxide can be further converted to L-arginine through the L-arginine synthesis pathway. The normalized TPM values of genes *arcC* and *argI* were considerably higher than that found in MG1655 and BL21(DE3) (Fig. 5 and Supplementary data 6). The high expression level of arginine deiminase may indicate that us of arginine degradation is used to generate energy in EcN.

3.6. Low expression of ribosomal proteins in EcN

Through transcriptome analysis, we found 63 ribosomal proteins of EcN all had the lowest relative gene expression value when compared with BL21(DE3) and MG1655 (Fig. 6a). This may lead to a lower translation level for EcN. To characterize the protein expression ability of EcN, we chose constitutive promoters J23119 to express GFP. Growth analysis showed a slight difference between EcN, BL21(DE3), and MG1655, which indicates there is little metabolic stress caused by expression of GFP. The order of fluorescence intensity was MG1655 > BL21(DE3) >EcN (Fig. 6b). Another reason for the low GFP expression level seen in EcN may be the existence of two stable cryptic plasmids, pMUT1 and pMUT2. It seemed that the presence of pMUT1 and pMUT2 in EcN limited the usefulness of other heterologous plasmids due to resource competition or plasmid incompatibility (Lan et al., 2021, Zainuddin et al., 2019). Therefore, the engineering of EcN into a biotechnological chassis may require increasing the protein expression capabilities of EcN by developing synthetic promoters, screening for native promoters with super high transcriptional strengths, or deactivating endogenous proteinases.

3.7. Growth and metabolic properties of EcN, BL21(DE3), and MG1655

As an industrial fermentation host, the glucose consumption rate and biomass of *E. coli* are the primary parameters affecting product yield. Acetic acid is a common by-product of *E. coli* fermentation, especially with glucose as carbon source. The secreted acetic acid inhibits growth at high concentrations, especially during high-cell-density cultivation (Shi et al., 2017, Eiteman and Altman, 2006). Upon sugar starvation, *E. coli* reuses the secreted acetic acid. Herein, we focused on these three indicators and made a comparison between EcN, BL21(DE3), and MG1655. When the growth rates of the three strains in LB medium were examined, we found that EcN displayed a faster growth rate in exponential phase (Fig. 7a). The final OD_{600} of EcN was between that of MG1655 and BL21(DE3).

We found EcN grew better in mineral salt medium supplemented with 10 g/L glucose, with the order of the final OD_{600} being EcN > MG1655 > BL21(DE3) (Fig. 7b). Acetic acid production of EcN was less than that of MG1655. The order of both glucose consumption and acetic acid accumulation was MG1655 > EcN > BL21(DE3) (Fig. 7c,d). The main reason for the accumulation of acetic acid is the imbalance of glucose uptake, high expression of acetic acid synthesis genes, and limited TCA cycle activities (Liu et al., 2015, Yoon et al., 2009). Consistent with the results in Fig. 4b, reasons for the low acetate accumulation may be that EcN and BL21(DE3) showed lower activity within the glycolysis pathway, while the genes related to the TCA cycle in EcN and BL21(DE3) have relatively high activities. The two main pathways for

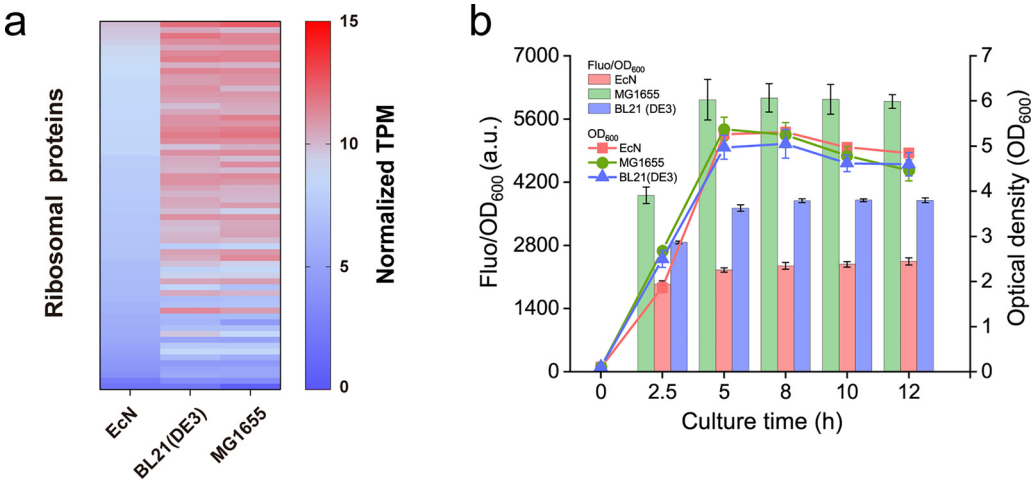

Fig. 6. Protein expression ability of EcN, BL21(DE3), and MG1655. (a) Heat map represents relative gene expression density in EcN, BL21(DE3), and MG1655. Heat map generated for the ribosomal proteins using the $\log_2$ transformed normalized TPM values. The colored bars shown on the right represent their expression levels. (b) Relative fluorescence intensity of GFP and cell growth under the control of J23119 promoter. Solid lines represent growth curve while bars represent relative fluorescence intensity. Data points represent mean values of three biological replicates with error bars showing standard deviation. The bars represent mean value of three biological replicates with error bars showing standard deviation.

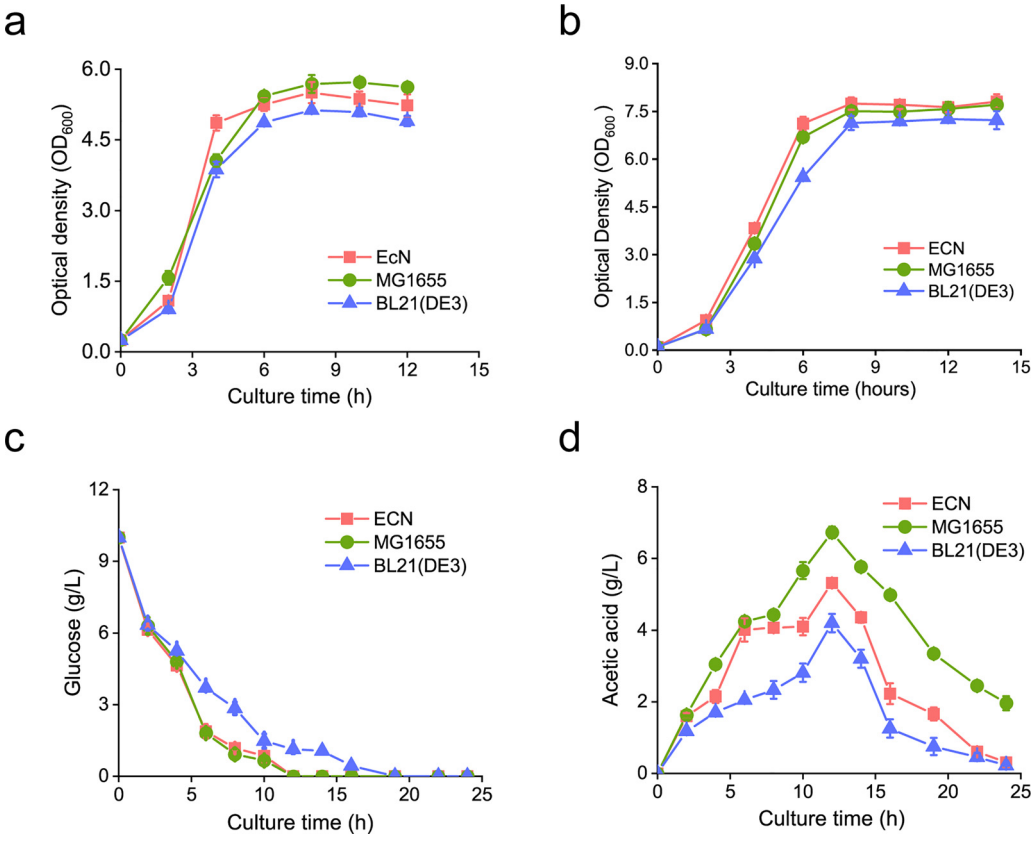

Fig. 7. Metabolic characteristics of the three *E. coli* strains in different conditions. (a) Growth curves of EcN, BL21(DE3), and MG1655 in LB medium. Data points represent mean value of three biological replicates with error bars showing standard deviation. (b) Growth curves, (c) glucose consumption, and (d) acetic acid accumulation of EcN, BL21(DE3), and MG1655 in minimal salt medium with 10 g/L glucose as a carbon source. Data points represent mean value of three biological replicates with error bars showing standard deviation.

acetate production in BL21(DE3) and EcN all have lower gene expression levels than MG1655. Another reason for the low acetate accumulation in BL21(DE3) is the absence of a regulatory mechanism and the low carbon catabolite repression (Phue et al., 2005), however the regulation mechanism in EcN is not as clear. Indeed, the metabolism of glucose and acetate in EcN requires further study.

4. Conclusions

In this report, we performed a basic study on the genomes, transcriptomes, and metabolic properties of EcN, BL21(DE3), and MG1655. Genome comparison showed that EcN has a type VI secretion system and *pks* loci, which are missing from the genomes of both BL21(DE3)

and MG1655, but are commonly found in pathogenic bacteria. The *pks* locus in pathogenic *E. coli* contributes to virulence, while in EcN it plays a role in visceral analgesia (Perez-Berezo et al., 2017, Homburg et al., 2007). According to genome comparison results (Supplementary data 2), EcN has multiple iron uptake systems. We found that EcN displays less sensitivity than BL21(DE) and MG1655 to iron scarcity (Fig. 3a-3c). The existence of the arginine deiminase pathway may provide EcN with an additional way to catabolize arginine and generate ATP (Fig. 5). This may enable EcN to be used in environments where ammonia is metabolized into exhaust gas. We cultivated EcN, BL21(DE3), and MG1655 in inorganic salt medium and found that EcN had certain growth advantages while the byproduct acetic acid production was in between that of BL21(DE3) and MG1655 (Fig. 7b-7d). This may give EcN an advantage in high-density fermentation process.

EcN has a K5 polysaccharide capsule and lipopolysaccharide, which is a potential pyrogen in other gram-negative strains of *E. coli*. It has been reported that the K5 polysaccharide capsule and lipopolysaccharide in EcN have immunomodulatory effects and protect EcN against infection with T4 phages (Nzakizwanayo et al., 2015, Hafez et al., 2010, Soundararajan et al., 2019). An adapted Ames test assessing genotoxicity showed that EcN had no DNA-damaging activity or genotoxicity (Dubbert et al., 2020). The polysaccharide capsule in EcN was non-toxic, but it may limit the efficiency of genetic manipulation in EcN. To improve genetic manipulation efficiency, knocking out key genes involved in the synthesis of the polysaccharide capsule and lipopolysaccharide should be studied in the future to improve EcN's capacity to act as a metabolic engineering host.

With the increasing attention to the biosafety of microbial products, probiotic EcN is a good prospect as a metabolic engineering host. However, as far as we know, EcN has only been applied to the production of omega-3 fatty acids, heparosan, and 5-aminolevulinic acid (Datta et al., 2021, Amiri-Jami et al., 2015, Chen et al., 2021); thus, the metabolic properties of EcN have not been elucidated. In this study, we found that the glycolysis and acetic acid synthesis activity of EcN was more similar to BL21(DE3) (Fig. 4a,b). We also found that EcN showed a lower expression level of ribosomal proteins when compared to BL21(DE3) and MG1655 (Fig. 6a,b). These results suggest that by introducing the T7 RNA polymerase-dependent transcription system, (Fiege and Frankenberg-Dinkel, 2020) and improving the expression of ribosomal proteins, EcN could be developed into a novel biotechnological expression platform.

Date availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

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.

Acknowledgments

This work was financially supported by the National Key Research and Development Program of China (2021YFC2100800), the Jiangsu Province Natural Science Fund for Distinguished Young Scholars (BK20200025), a grant from the Key Technologies R&D Program of Jiangsu Province (BE2019630), and the National First-class Discipline Program of Light Industry Technology and Engineering (LITE2018-16).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.engmic.2022.100012.

References

Amiri-Jami, M, Abdelhamid, AG, Hazaa, M, Kakuda, Y, Griffths, MW, 2015. Recombinant production of omega-3 fatty acids by probiotic *Escherichia coli* Nissle 1917. FEMS Microbiol Lett 362 (20). doi:10.1093/femsle/fnv166.

Benesty, J, Chen, J, Huang, Y, Cohen, I., 2009. Pearson correlation coefficient. Noise Reduction in Speech Processing. Springer Topics in Signal Processing, pp. 1–4.

Blum, M, Chang, HY, Chuguransky, S, Grego, T, Kandasaamy, S, Mitchell, A, et al., 2021. The InterPro protein families and domains database: 20 years on. Nucleic Acids Res 49 (D1), D344–DD54. doi:10.1093/nar/gkaa977.

Bolger, AM, Lohse, M, Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170.

Chen, J, Li, X, Liu, Y, Su, T, Lin, C, Shao, L, et al., 2021. Engineering a probiotic strain of *Escherichia coli* to induce the regression of colorectal cancer through production of 5-aminolevulinic acid. Microb Biotechnol doi:10.1111/1751-7915.13894.

Darling, AC, Mau, B, Blattner, FR, Perna, NT., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14 (7), 1394–1403. doi:10.1101/gr.2289704.

Datta, P, Fu, L, Brodfuerer, P, Dordick, JS, Linhardt, RJ., 2021. High density fermentation of probiotic *E. coli* Nissle 1917 towards heparosan production, characterization, and modification. Appl Microbiol Biotechnol 105 (3), 1051–1062. doi:10.1007/s00253-020-11079-9.

Deriu, E, Liu, JZ, Pezeshki, M, Edwards, RA, Ochoa, RJ, Contreras, H, et al., 2013. Probiotic bacteria reduce *salmonella typhimurium* intestinal colonization by competing for iron. Cell Host Microbe 14 (1), 26–37. doi:10.1016/j.chom.2013.06.007.

Do, J, Zafar, H, Saier, MH, 2017. Jr. Comparative genomics of transport proteins in probiotic and pathogenic *Escherichia coli* and *Salmonella enterica* strains. Microb Pathog 107, 106–115. doi:10.1016/j.micpath.2017.03.022.

Dubbert, S, Klinkert, B, Schimiczek, M, Wassenaar, TM, Rv, Bünau, 2020. No genotoxicity is detectable for Escherichia coli strain Nissle 1917 by standard in vitro and in vivo tests. European Journal of Microbiology and Immunology 10 (1), 11–19. doi:10.1556/1886.2019.00025.

Eiteman, MA, Altman, E., 2006. Overcoming acetate in *Escherichia coli* recombinant protein fermentations. Trends Biotechnol 24 (11), 530–536. doi:10.1016/j.tibtech.2006.09.001.

Fabrega, MJ, Rodriguez-Nogales, A, Garrido-Mesa, J, Algieri, F, Badia, J, Gimenez, R, et al., 2017. Intestinal anti-inflammatory effects of outer membrane vesicles from *Escherichia coli* Nissle 1917 in DSS-experimental colitis in mice. Front Microbiol 8, 1274. doi:10.3389/fmicb.2017.01274.

Fiege, K, Frankenberg-Dinkel, N., 2020. Construction of a new T7 promoter compatible *Escherichia coli* Nissle 1917 strain for recombinant production of heme-dependent proteins. Microb Cell Fact 19 (1), 190. doi:10.1186/s12934-020-01447-5.

Gerken, H, Vuong, P, Soparkar, K, Misra, R., 2020. Roles of the EnvZ/OmpR two-component system and porins in iron acquisition in *Escherichia coli*. mBio 11 (3). doi:10.1128/mBio.01192-20.

Gomes, AEI, Stuchi, LP, Siqueira, NMG, Henrique, JB, Vicentini, R, Ribeiro, ML, et al., 2018. Selection and validation of reference genes for gene expression studies in *Klebsiella pneumoniae* using Reverse Transcription Quantitative real-time PCR. Sci Rep 8 (1), 9001. doi:10.1038/s41598-018-27420-2.

Hafez, M, Hayes, K, Goldrick, M, Grencis, RK, Roberts, IS., 2010. The K5 capsule of Escherichia coli strain Nissle 1917 is important in stimulating expression of Toll-like receptor 5, CD14, MyD88, and TRIF together with the induction of interleukin-8 expression via the mitogen-activated protein kinase pathway in epithelial cells. Infect Immun 78 (5), 2153–2162. doi:10.1128/IAI.01406-09.

Hayes, F, Van Melderen, L., 2011. Toxins-antitoxins: diversity, evolution and function. Crit Rev Biochem Mol Biol 46 (5), 386–408. doi:10.3109/10409238.2011.600437.

He, L, Yang, H, Tang, J, Liu, Z, Chen, Y, Lu, B, et al., 2019. Intestinal probiotics *E. coli* Nissle 1917 as a targeted vehicle for delivery of p53 and Tum-5 to solid tumors for cancer therapy. J Biol Eng 13, 58. doi:10.1186/s13036-019-0189-9.

Helmy, YA, Kassem, II, Rajashekara, G, 2021. Immuno-modulatory effect of probiotic *E. coli* Nissle 1917 in polarized human colonic cells against *Campylobacter jejuni* infection. Gut Microbes 13 (1), 1–16. doi:10.1080/19490976.2020.1857514.

Henker, J, Laass, M, Blokhin, BM, Bolbot, YK, Maydannik, VG, Elze, M, et al., 2007. The probiotic *Escherichia coli* strain Nissle 1917 (EcN) stops acute diarrhoea in infants and toddlers. Eur J Pediatr 166 (4), 311–318. doi:10.1007/s00431-007-0419-x.

Henker, J, Laass, MW, Blokhin, BM, Maydannik, VG, Bolbot, YK, Elze, M, et al., 2008. Probiotic *Escherichia coli* Nissle 1917 versus placebo for treating diarrhea of greater than 4 days duration in infants and toddlers. Pediatr Infect Dis J 27 (6), 494–499. doi:10.1097/INF.0b013e318169034c.

Homburg, S, Oswald, E, Hacker, J, Dobrindt, U, 2007. Expression analysis of the colibactin gene cluster coding for a novel polyketide in *Escherichia coli*. FEMS Microbiol Lett 275 (2), 255–262. doi:10.1111/j.1574-6968.2007.00889.x.

Kamruzzaman, M, Wu, AY, Iredell, JR., 2021. Biological functions of type II toxin-antitoxin systems in bacteria. Microorganisms 9 (6). doi:10.3390/microorganisms9061276.

Langdon, WB., 2015. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. BioData Min 8 (1), 1. doi:10.1186/s13040-014-0034-0.

Li, J, Yao, Y, Xu, HH, Hao, L, Deng, Z, Rajakumar, K, et al., 2015. SecReT6: a web-based resource for type VI secretion systems found in bacteria. Environ Microbiol 17 (7), 2196–2202. doi:10.1111/1462-2920.12794.

Li, R, Helbig, L, Fu, J, Bian, X, Herrmann, J, Baumann, M, et al., 2019. Expressing cytotoxic compounds in *Escherichia coli* Nissle 1917 for tumor-targeting therapy. Res Microbiol 170 (2), 74–79. doi:10.1016/j.resmic.2018.11.001.

Liao, Y, Smyth, GK, Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30 (7), 923–930. doi:10.1093/bioinformatics/btt656.

Liu, M, Feng, X, Ding, Y, Zhao, G, Liu, H, Xian, M., 2015. Metabolic engineering of *Escherichia coli* to improve recombinant protein production. Appl Microbiol Biotechnol 99 (24), 10367–10377. doi:10.1007/s00253-015-6955-9.

Marchler-Bauer, A, Derbyshire, MK, Gonzales, NR, Lu, S, Chitsaz, F, Geer, LY, et al., 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res 43, D222–D226. doi:10.1093/nar/gku1221, Database issue.

McMillan, M, Pereg, L., 2014. Evaluation of reference genes for gene expression analysis using quantitative RT-PCR in *Azospirillum brasilense*. PLoS One 9 (5), e98162. doi:10.1371/journal.pone.0098162.

Molan, K, Podlesek, Z, Hodnik, V, Butala, M, Oswald, E, 2019. Zgur Bertok D. The *Escherichia coli* colibactin resistance protein ClbS is a novel DNA binding protein that protects DNA from nucleolytic degradation. DNA Repair (Amst) 79, 50–54. doi:10.1016/j.dnarep.2019.05.003.

Novak, L, Zubacova, Z, Karnkowska, A, Kolisko, M, Hroudova, M, Stairs, CW, et al., 2016. Arginine deiminase pathway enzymes: evolutionary history in metamonads and other eukaryotes. BMC Evol Biol 16 (1), 197. doi:10.1186/s12862-016-0771-4.

Lan, Y-J, Tan, S-I, Cheng, S-Y, Ting, W-W, Xue, C, Lin, T-H, et al., 2021. Development of *Escherichia coli* Nissle 1917 derivative by CRISPR/Cas9 and application for gamma-aminobutyric acid (GABA) production in antibiotic-free system. Biochemical Engineering Journal 168. doi:10.1016/j.bej.2021.107952.

Nzakizwanayo, J, Kumar, S, Ogilvie, LA, Patel, BA, Dedi, C, Macfarlane, WM, et al., 2015. Disruption of Escherichia coli Nissle 1917 K5 capsule biosynthesis, through loss of distinct kfi genes, modulates interaction with intestinal epithelial cells and impact on cell health. PLoS One 10 (3), e0120430. doi:10.1371/journal.pone.0120430.

Perez-Berezo, T, Pujo, J, Martin, P, Le Faouder, P, Galano, JM, Guy, A, et al., 2017. Identification of an analgesic lipopeptide produced by the probiotic *Escherichia coli* strain Nissle 1917. Nat Commun 8 (1), 1314. doi:10.1038/s41467-017-01403-9.

Phue, JN, Noronha, SB, Hattacharyya, R, Wolfe, AJ, Shiloach, J., 2005. Glucose metabolism at high density growth of E. coli B and E. coli K: differences in metabolic pathways are responsible for efficient glucose utilization in E. coli B as determined by microarrays and Northern blot analyses. Biotechnol Bioeng 90 (7), 805–820. doi:10.1002/bit.20478.

Pontrelli, S, Chiu, TY, Lan, EI, Chen, FY, Chang, P, Liao, JC, 2018. *Escherichia coli* as a host for metabolic engineering. Metab Eng 50, 16–46. doi:10.1016/j.ymben.2018.04.008.

Praveschotinunt, P, Duraj-Thatte, AM, Gelfat, I, Bahl, F, Chou, DB, Joshi, NS., 2019. Engineered *E. coli* Nissle 1917 for the delivery of matrix-tethered therapeutic domains to the gut. Nat Commun 10 (1), 5580. doi:10.1038/s41467-019-13336-6.

Rund, SA, Rohde, H, Sonnenborn, U, Oelschlaeger, TA., 2013. Antagonistic effects of probiotic *Escherichia coli* Nissle 1917 on EHEC strains of serotype O104:H4 and O157:H7. Int J Med Microbiol 303 (1), 1–8. doi:10.1016/j.ijmm.2012.11.006.

Sandrini S, Masania R, Zia F, Haigh R, Freestone P. Role of porin proteins in acquisition of transferrin iron by enteropathogens. Microbiology (Reading). 2013;159(Pt 12):2639-50. doi: 10.1099/mic.0.071928-0

Santiviago, CA, Toro, CS, Hidalgo, AA, Youderian, P, Mora, GC., 2003. Global regulation of the *Salmonella enterica serovar* typhimurium major porin. OmpD. J Bacteriol. 185 (19), 5901–5905. doi:10.1128/JB.185.19.5901-5905.2003.

Sarate, PJ, Heinl, S, Poiret, S, Drinic, M, Zwicker, C, Schabussova, I, et al., 2019. *E. coli* Nissle 1917 is a safe mucosal delivery vector for a birch-grass pollen chimera to prevent allergic poly-sensitization. Mucosal Immunol 12 (1), 132–144. doi:10.1038/s41385-018-0084-6.

Shi, X, Xie, J, Liao, S, Wu, T, Zhao, LG, Ding, G, et al., 2017. High-level expression of recombinant thermostable beta-glucosidase in *Escherichia coli* by regulating acetic acid. Bioresour Technol 241, 795–801. doi:10.1016/j.biortech.2017.05.105.

Sonnenborn, U, Schulze, J., 2009. The non-pathogenic *Escherichia coli* strain Nissle 1917 – features of a versatile probiotic. Microbial Ecology in Health and Disease 21 (3-4), 122–158. doi:10.3109/08910600903444267.

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 (19). doi:10.1093/femsle/fnw212.

Soundararajan, M, von Bunau, R, Oelschlaeger, TA., 2019. K5 Capsule and Lipopolysaccharide Are Important in Resistance to T4 Phage Attack in Probiotic E. coli Strain Nissle 1917. Front Microbiol 10, 2783. doi:10.3389/fmicb.2019.02783.

Takle, GW, Toth, IK, Brurberg, MB., 2007. Evaluation of reference genes for real-time RT-PCR expression studies in the plant pathogen *Pectobacterium atrosepticum*. BMC Plant Biol 7, 50. doi:10.1186/1471-2229-7-50.

Trebichavsky, I, Splichal, I, Rada, V, Splichalova, A., 2010. Modulation of natural immunity in the gut by *Escherichia coli* strain Nissle 1917. Nutr Rev 68 (8), 459–464. doi:10.1111/j.1753-4887.2010.00305.x.

Tripathi, P, Shine, EE, Healy, AR, Kim, CS, Herzon, SB, Bruner, SD, et al., 2017. ClbS is a cyclopropane hydrolase that confers colibactin resistance. J Am Chem Soc 139 (49), 17719–17722. doi:10.1021/jacs.7b09971.

Wan, B, Zhang, Q, Ni, J, Li, S, Wen, D, Li, J, et al., 2017. Type VI secretion system contributes to Enterohemorrhagic *Escherichia coli* virulence by secreting catalase against host reactive oxygen species (ROS). PLoS Pathog 13 (3), e1006246. doi:10.1371/journal.ppat.1006246.

Wang, Y, Wang, H, Hay, AJ, Zhong, Z, Zhu, J, Kan, B., 2015. Functional RelBE-family toxin-antitoxin pairs affect biofilm maturation and intestine colonization in *vibrio cholerae*. PLoS One 10 (8), e0135696. doi:10.1371/journal.pone.0135696.

Weiss, G., 2013. Intestinal irony: how probiotic bacteria outcompete bad bugs. Cell Host Microbe 14 (1), 3–4. doi:10.1016/j.chom.2013.07.003.

Xia, Y, Li, K, Li, J, Wang, T, Gu, L, Xun, L., 2019. T5 exonuclease-dependent assembly offers a low-cost method for efficient cloning and site-directed mutagenesis. Nucleic Acids Res 47 (3), e15. doi:10.1093/nar/gky1169.

Xie, R, Tu, M, Wu, Y, Adhikari, S., 2011. Improvement in HPLC separation of acetic acid and levulinic acid in the profiling of biomass hydrolysate. Bioresour Technol 102 (7), 4938–4942. doi:10.1016/j.biortech.2011.01.050.

Yoon, SH, Jeong, H, Kwon, S-K, Genomics, Kim JF., 2009. Biological Features, and Biotechnological Applications of *Escherichia coli* B: "Is B for better?!. Systems Biology and Biotechnology of Escherichia coli. Springer Netherlands, Dordrecht, pp. 1–17.

Yu, X, Lin, C, Yu, J, Qi, Q, Wang, Q., 2020. Bioengineered *Escherichia coli* Nissle 1917 for tumour-targeting therapy. Microbial Biotechnology 13 (3), 629–636. doi:10.1111/1751-7915.13523.

Zainuddin, HS, Bai, Y, Mansell, TJ., 2019. CRISPR-based curing and analysis of metabolic burden of cryptic plasmids in *Escherichia coli* Nissle 1917. Engineering in Life Sciences 19 (6), 478–485. doi:10.1002/elsc.201900003.

Zuniga, M., Gonzalez-Candelas, GP, F., 2002. Evolution of arginine deiminase (ADI) pathway genes. Mol Phylogenet Evolution 429–444.