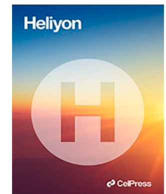
Heliyon 9 (2023) e22986

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



ArdA genes from pKM101 and from *B. bifidum* chromosome have a different range of regulated genes

M.V. Gladysheva-Azgari [a], F.S. Sharko [a,h], M.A. Evteeva [a], A.P. Kuvyrchenkova [a], E.S. Boulygina [a], S.V. Tsygankova [a], N.V. Slobodova [a], K.S. Pustovoit [c], O.E. Melkina [c], A.V. Nedoluzhko [b], A.A. Korzhenkov [a], A.A. Kudryavtseva [d,*], A.A. Utkina [d], I.V. Manukhov [d,e,f], S.M. Rastorguev [a,g], G.B. Zavilgelsky [c,1]

[a] *National Research Center "Kurchatov Institute", 123182, Moscow, Russia*
[b] *European University at Saint Petersburg, 191187, Saint-Petersburg, Russia*
[c] *State Research Institute of Genetics and Selection of Industrial Microorganisms of the National Research Center "Kurchatov Institute", Moscow, Russia, 115454*
[d] *Research Center for Molecular Mechanisms of Aging and Age-Related Diseases, Moscow Institute of Physics and Technology, 141701, Dolgoprudny, Russia*
[e] *Faculty of Physics, HSE University, 109028, Moscow, Russia*
[f] *Laboratory for Microbiology, BIOTECH University, 125080, Moscow, Russia*
[g] *Pirogov Russian National Research Medical University, Ostrovityanova Str. 1, Moscow, 117997, Russia*
[h] *Research Center of Biotechnology of the Russian Academy of Sciences, Moscow 119071, Russia*

ARTICLE INFO

Keywords:
ArdA
Antirestriction
Gene regulation

ABSTRACT

The *ardA* genes are present in a wide variety of conjugative plasmids and play an important role in overcoming the restriction barrier. To date, there is no information on the chromosomal *ardA* genes. It is still unclear whether they keep their antirestriction activity and why bacterial chromosomes contain these genes. In the present study, we confirmed the antirestriction function of the *ardA* gene from the *Bifidobacterium bifidum* chromosome. Transcriptome analysis in *Escherichia coli* showed that the range of regulated genes varies significantly for *ardA* from conjugative plasmid pKM101 and from the *B. bifidum* chromosome. Moreover, if the targets for both *ardA* genes match, they often show an opposite effect on regulated gene expression. The results obtained indicate two seemingly mutually exclusive conclusions. On the one hand, the pleiotropic effect of *ardA* genes was shown not only on restriction-modification system, but also on expression of a number of other genes. On the other hand, the range of affected genes varies significantly for *ardA* genes from different sources, which indicates the specificity of *ardA* to inhibited targets.

Author Summary. Conjugative plasmids, bacteriophages, as well as transposons, are capable to transfer various genes, including antibiotic resistance genes, among bacterial cells. However, many of those genes pose a threat to the bacterial cells, therefore bacterial cells have special restriction systems that limit such transfer.

Antirestriction genes have previously been described as a part of conjugative plasmids, and bacteriophages and transposons. Those plasmids are able to overcome bacterial cell protection in the presence of antirestriction genes, which inhibit bacterial restriction systems.

* Corresponding author.
E-mail address: kudryavtseva@phystech.edu (A.A. Kudryavtseva).

[1] Deceased.

https://doi.org/10.1016/j.heliyon.2023.e22986
Received 17 April 2023; Received in revised form 18 September 2023; Accepted 23 November 2023
Available online 29 November 2023
2405-8440/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

This work unveils the antirestriction mechanisms, which play an important role in the bacterial life cycle. Here, we clearly show that antirestriction genes, which are able to inhibit cell protection, exist not only in plasmids but also in the bacterial chromosomes themselves.

Moreover, antirestrictases have not only an inhibitory function but also participate in the regulation of other bacterial genes. The regulatory function of plasmid antirestriction genes also helps them to overcome the bacterial cell protection against gene transfer, whereas the regulatory function of genomic antirestrictases has no such effect.

1. Introduction

Ard genes were first detected in plasmids of the N- and I-incompatibility groups (pKM101*,* ColIb-P9) [1] [2] and then were determined in I complex (IncB, I1 and K), the F complex (IncFV) and the IncN group of plasmids [3] as well as in bacterial chromosomes, transposons, bacteriophages [4]. The genes *ardA* and *ardB* encode antirestriction proteins, which allow plasmids to overcome restriction barrier [5, 6] and appears to play an important role in horizontal gene transfer in bacteria. While the exact antirestriction mechanism of ArdB has not been determined yet [7,8], it is known that the inhibition of RM systems by ArdA proteins is occurred by substrate-induced suppression.

The mechanism of RMI inhibition by ArdA could be confirmed by the atomic structure which was studied by Dryden's group [9]. According to this work ArdA protein looks like DNA molecule with surface negative charges which mimic phosphate groups. It was also shown in the work mentioned that ArdA is sterically like the DNA cleavage site and irreversibly binds to the restriction complex in the active site. It is known that ArdA could regulate the activity of some other proteins besides the restriction/modification systems [10, 11].

A wide variety of *ardA* genes were found in the nucleotide sequence databases, suggesting that antirestriction genes play an important role in the bacterial life. However, the experimentally confirmed antirestriction function has been described only for a small group of highly similar *ardA* genes from conjugative plasmids [5, 9, 11]. But most chromosomal *ardA* genes were identified only by their homology to the plasmid *ardA*. Antirestriction activity of chromosomal *ardA* was unexplored.

Here we compared using heterologous *E. coli* system the antirestriction activity of *ardA* gene from the conjugative plasmid pKM101 [12] and the antirestriction activity of *ardA* gene from *B. bifidum* Ac1784 chromosome. To verify the hypothesis that DNA-mimicking proteins ArdA could regulate gene expression we performed transcriptome analysis. We compared the impact of *ardA* genes from *B. bifidum* and pKM101 on the level of gene transcription in *E. coli*.

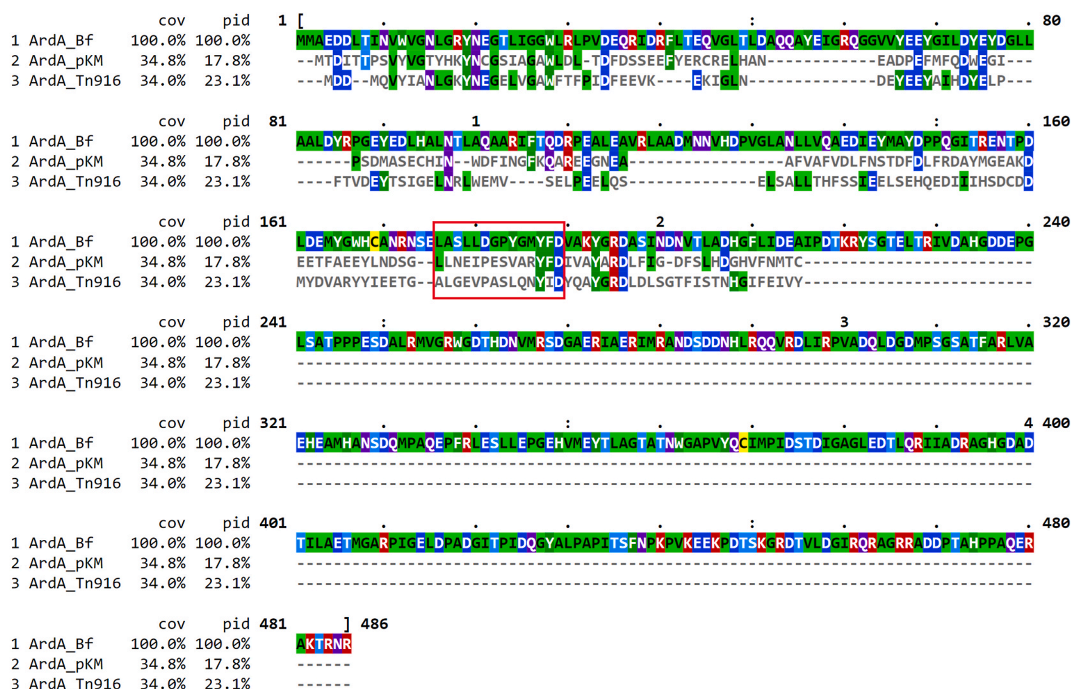


Fig. 1. ArdA sequence alignment. Aligned amino acid sequences of ArdA proteins: chromosomal ArdA_Bf, plasmid ArdA_pKM and transposal ArdA_Tn916. The dash (-) sign indicates the absence of a homologous amino acid at this position in the second sequence. The Cov value presents the percent of coverage of each sequence and Pid value is the percent of identical amino acids. The red frame indicates the antirestriction motif.

2. Results

2.1. Comparison of ArdA from the B. bifidum chromosome with plasmid and transposonal antirestrictases

As a chromosomal homologue of *ardA* pKM101 (hereinafter *ardA*_pKM) the *ardA* gene (MBH8618360.1) from *B. bifidum* Ac1784 (hereinafter *ardA*_Bf) was chosen. The resulting annotation for the region containing the *ardA* gene is provided as a GenBank file. When we applied the IslandViewer program (available at https://www.pathogenomics.sfu.ca/islandviewer/browse/), it did not identify this region as a genomic island. Investigation using the CONJscan tool also failed to reveal the presence of genomic islands in this specific region.

Here we believe that the *ardA* gene in the *B. bifidum* chromosome may have originated from mobile genetic elements, but this origin appears to be quite ancient. This is supported by the presence of the same gene in other *Bifidobacterium* species within the same genomic region. The fact that this gene has persisted in bacterial chromosomes over an extended period suggests its utility to bacteria. Therefore, it likely serves a different function within bacterial chromosomes compared to its role within mobile genetic elements. This longevity and retention in bacterial genomes indicate its significance within bacterial biology.

The choice of *ardA*_Bf is due to this gene is surrounded by genes that are typical for chromosomal DNA: ABC transporter, histidine kinase, *parB* etc. Blast analysis revealed presence of the *B.bifidum ardA* gene with high identity (from 100 % to 89.44 %) in the chromosome of other species of the genus *Bifidobacterium*, such as *B. breve, B. saguini* and *B.longum*. At the same time, this analysis did not reveal the presence of the *ardA* gene on transmissible plasmids. On the contrary, the other antirestriction gene (*ardA*_pKM) is widely presented in various transmissible plasmids according to the Blast analysis. In this regard, we considered the *ardA* gene from *B. bifidum* (*ardA*_Bf) to be a chromosomal gene, whereas the *ardA*_pKM is extrachromosomal one.

Comparing the sequences of the chromosomal ArdA_Bf with the sequences of "classical" antirestrictases from conjugatve plasmid ArdA_pKM and conjugative transposon Tn916 (hereinafter ArdA_Tn916) demonstrated that the absolute value of the sequence similarity is quite low – just 17,8 % and 23,1 % identity of amino acid positions and the length is very different and amounts 487 versus 169 and 166 amino acids, respectively (Fig. 1). However, certain sophisticated Hidden Markov model (HMM) search algorithms for calculating the sequence similarity level show that these sequences are statistically significantly similar to each other. Well-conserved amino acids generally match for studied proteins. Antirestriction motif [13], which represents the interface between two ArdA subunits, is marked in a red frame.

To date the 3D-structures of the ArdA_Bf and ArdA_pKM proteins are unknown. To compare these proteins, we used 3D-structures predicted by AlphaFold [14]. AlphaFold structures of ArdA_Bf and ArdA_pKM were studied and aligned with PyMOL (Fig. 2). The known structure of ArdA_Tn916 protein was used for comparison [8].

Fig. 2 shows that the 3D-structure of ArdA_pKM is highly similar to that of ArdA_Tn916. However, the structure of ArdA_Bf protein contains two distinct domains. The N-terminal domain is structurally similar to the ArdA_pKM and ArdA_Tn916 proteins, while the C-terminal domain of ArdA_Bf is located separately.

The structures shown in Fig. 2A and B shows the similarity of the spatial arrangement of a number of structural elements of ArdA_pKM and the N-domain of ArdA_Bf. The antirestriction motif, marked in red in Figs. 1 and 2, is located in the region of the dimer interface; however, a significant difference between ArdA_Bf and ArdA_Tn916 (and ArdA_pKM) makes the dimerization questionable. Negatively charged amino acids of the N-terminal domain of the ArdA_Bf protein are distributed over the protein surface (Fig. 2D) in a

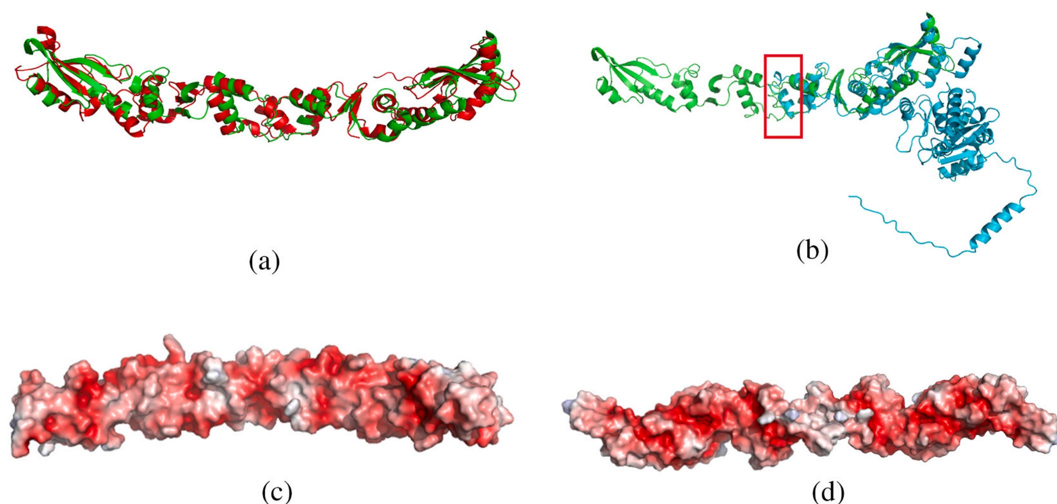


Fig. 2. 3D-structures of ArdA proteins. A – alignment of the predicted structure of ArdA_pKM protein (red) and the well-known ArdA_Tn916 (green). B – alignment of the predicted structure of ArdA_Bf protein (blue) and the well-known ArdA_Tn916 (green). The red frame indicates the antirestriction motif. C – the surface of ArdA_Tn916 protein with acidic residues colored red. D – the surface of the N-domain (Met1-Gly240) of the ArdA_Bf protein with acidic residues colored red. The structure is presented as a hypothetical dimer.

similar way to ArdA_pTn916 (Fig. 2C) and are obviously necessary for the DNA mimicry.

Analysis of the putative structures of ArdA_Bf using the Predictor of Natural Disordered Regions (PONDR) showed a high internal disorder of the C-domain (figures S1, S2), which may indicate flexible regions of the protein. The presence of the C-domain in the ArdA_Bf structure can also lead to additional protein functionality compared to the plasmid ArdA_pKM.

2.2. Antirestriction effect of ardA from B. bifidum

The antirestriction activity was evaluated by counting the negative colonies of the unmodified bacteriophage λ (λ_0) on bacterial strain AB1157, which contains EcoKI restriction-modification system. TG1 strain (deleted for EcoKI) was used as a negative control.

To compare the effect of ArdA_pKM and ArdA_Bf on the RMI-system, we used *E. coli* K12 AB1157 cells containing hybrid plasmids pAB7 [2] and pArdA_Bf (this work) with the corresponding *ardA* genes. Promoter P_{lac} with no IPTG induction was used for expression of both genes. P_{lac} leakage in AB1157 strain was enough for complete antirestriction effect.

The antirestriction activity of the *ardA* from *B. bifidum* is shown in Fig. 3. The efficiency of plating (EOP) was estimated as described previously [15].

$$EOP_X = \frac{N_X}{N_{TG1}} \quad (1)$$

where N_X-the number of λ.0 phage plaques on the *E. coli* cells carrying genes 'X' affecting the plaque forming, N_{TG1} – the number of λ.0 phage plaques on *E. coli* TG1 (without any additional restriction or antirestriction genes).

Fig. 3 represents that AB1157 strain has a fully active EcoKI restriction-modification system (about four orders of magnitude in phage plaquing), and in the absence of antirestriction genes, it plays the role of negative control. However, in the presence of *ardA_pKM*, the phage titer is "restored" to the level of the "restriction-free" TG1 strain. It could be seen that *ardA*_Bf behaves in the same way, almost completely removing the effect of restriction protection of the bacterial cell.

In addition, we tested antirestriction activity of pIRDPAL-ArdA_Bifi_cutG240 construct. It determines the expression the N-domain (Met1-Gly240) of the ArdA_Bf protein proposed to be the main DNA-mimic part of the protein. Fig. 3 demonstrates that the N-domain (Met1-Gly240) of the ArdA_Bf works as a complete antirestriction protein.

The data obtained in the antirestriction test demonstrate that *pArdA_Bf* и *pArdA_pKM* plasmids provide sufficient gene expression for complete antirestriction effect in R+M+ strains. Also it could be assumed that disordered C-domain of ArdA-bifi (downstream G240) could be used for some other process but not for antirestriction.

2.3. Transcriptome analysis of ardA impact on bacterial gene expression

To compare the effect of ArdA_pKM and ArdA_Bf on the regulation of bacterial gene expression, we used *E. coli* K12 AB1157 cells containing hybrid plasmids pAB7 and pArdA_Bf with the corresponding *ardA* genes. *E. coli* AB1157 with pBluescript(KS+) and pKAN-T vectors, respectively, were used as negative controls. RNA extraction from overnight culture, cDNA library preparation and Illumina sequencing were done for three clones for each of the four choices. Totally, more than 600 million paired reads were produced for these twelve libraries (table S0). The nucleotide reads were mapped to the *E. coli* K12 genome, and the number of the reads mapped to it, was calculated for each gene. For all the genes, a differential analysis was performed and the genes whose expression significantly differs

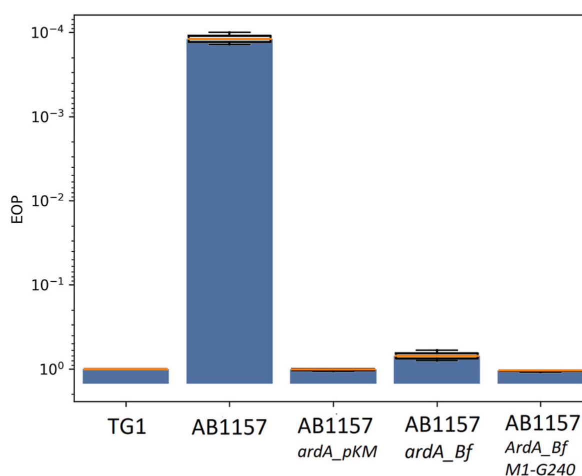


Fig. 3. Results of the λ.0 phage plaquing (EOP) on a lawn of *E. coli* cells containing EcoKI restriction-modification system and antirestriction constructs. The averaged results of three independent experiments are shown. Columns TG1, AB1157 – plain strains, AB1157 *ardA*_pKM – AB1157 contains *pAB7* plasmid, AB1157 ardA_Bf - AB1157 contains pKAN-T-ArdA_Bifi plasmid, AB1157 ardA_Bf M1-G240 – AB1157contains plasmid pIRDPAL-ArdA_Bifi_cutG240.

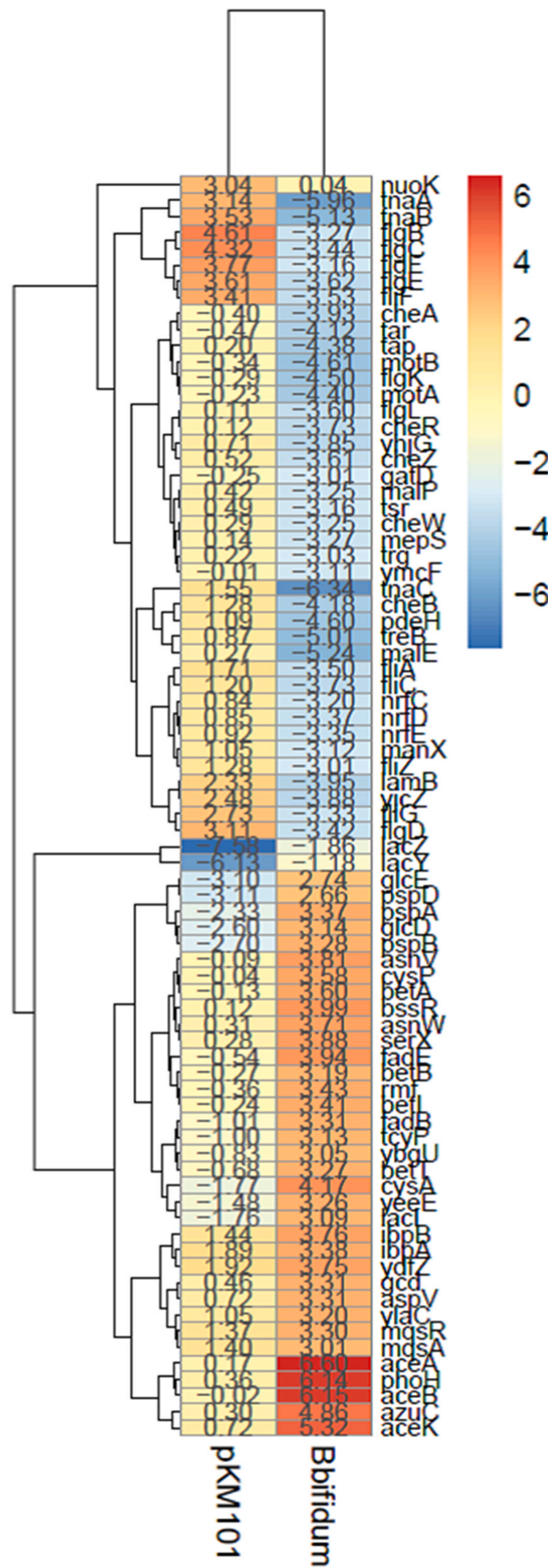


Fig. 4. Heatmap of significantly altered expressed genes with presence of *ardA* genes from pKM101 plasmid and *B. bifidum* chromosome. The logFC values is specified in the cells. The colors in the heat map indicated the log transformed values of each expression changes.

Table 1
Some DE genes, related to interspecies transfer of genetic material and biofilm formation.

Gene	Gene function and regulation	*ardA* from B. bifidum			*ardA* from pKM101		
		logFC	logCPM	FDR	logFC	logCPM	FDR
ansB	The glutaminase-asparaginase AnsB were detected in anaerobically grown biofilms [16].	−1,64	4,38	2,60E-04	1,61	4,80	3,49E-09
cydX	As part of the *cydABX* operon, was obtained by *E. coli* cells by horizontal transfer [17].	−2,53	2,06	2,40E-28	1,93	2,99	4,34E-14
nrdD	Anaerobic ribonucleoside-triphosphate reductase is essential for strict anaerobic growth of *E. coli* [18].	−1,93	2,72	1,06E-15	1,53	2,02	1,36E-07
yjjI	Uncharacterized gene, which participates in anaerobic respiration [19].	−2,11	2,86	5,37E-14	1,73	2,14	7,56E-09
lamB	Involved in the transport of maltose and maltodextrins [20] and receptor for phage lambda [21,22].	−3,95	2,31	2,19E-57	2,33	0,03	3,43E-05
crfC	Regulation of diguanylate cyclase (RdcA + RdcB/DgcE - ubiquitous second messenger c-di-GMP)	−2,86	1,42	4,35E-35	2,26	2,77	8,08E-08
yjcZ		−3,88	1,78	1,40E-55	2,48	2,77	2,93E-06
fau	Transcription of *ygfA* (fau) is induced upon biofilm formation [23].	0,09	1,43	7,51E-01	2,20	3,09	4,19E-04
yjhQ	Expression of *yjhQ* is induced in biofilms of a *tqsA* mutant strain [24].	−2,03	−1,24	7,68E-05	2,58	−0,96	2,73E-03
fimA	Transcription Unit: *fimAICDFGH*, *E. coli* type 1 fimbrial (pili) structural and regulation genes [25].	−0,69	7,05	3,48E-01	2,94	4,93	9,36E-38
fimC	IscR, controls biofilm formation in response to changes in cellular Fe–S homeostasis. IscR regulates the FimE recombinase to control expression of type I fimbriae in *E. coli*.	−1,45	3,49	1,25E-02	2,51	1,43	2,18E-10
fimD		−0,69	3,50	3,24E-01	2,60	1,58	7,09E-14
fimE		0,38	−0,12	3,92E-01	0,79	2,24	3,60E-02
fimG		0,28	0,66	7,34E-01	6,50	−1,24	5,34E-05
fimH		0,31	1,27	7,00E-01	2,35	−0,21	7,02E-05
fimI		−1,04	4,44	1,34E-01	2,69	2,22	1,64E-17
fliA	σ^{28} is a minor sigma factor that is responsible for initiation of transcription of a number of genes involved in motility and flagellar synthesis. Flagellin, is the basic	−3,50	2,27	1,86E-46	1,71	2,50	1,77E-07
fliC	subunit that polymerizes to form the rigid flagellar filament of *E. coli*. Activators: FlhDC, H-Ns, GadE	−3,73	6,58	3,73E-34	1,20	4,54	1,03E-05
fliD	FliM is one of three components of the flagellar motor's "switch complex" [26].	−4,55	3,41	2,67E-47	1,73	−0,55	1,38E-02
fliE		−2,40	−1,69	1,33E-06	1,90	−1,03	5,81E-02
fliF		−3,53	1,04	2,46E-36	3,41	0,11	3,47E-10
fliM		−3,47	1,40	8,61E-48	1,51	−0,25	1,38E-02
fliN		−4,05	0,34	8,23E-36	1,68	−1,24	1,17E-01
fliO		−3,10	−0,23	9,51E-21	2,26	−1,82	6,11E-02
gadE	Acid resistance system. H-NS –repressor, SdiA -activator	−0,80	−3,38	5,05E-01	2,96	−1,42	7,37E-03
cspH	CspA - cold shock protein [27].	−2,28	0,30	3,00E-12	2,33	−1,23	3,21E-02
pdeH	c-di-GMP phosphodiesterase/Regulation of the flagellar switching end exponential growth [28].	−4,60	2,58	1,97E-55	1,09	0,05	1,03E-01
pspA	The phenotypes that have been reported for cells in which the *psp* operon is deleted are as follows: lowered survival at alkaline pH in stationary phase [29], slower	3,37	2,90	9,43E-34	−2,33	6,12	6,12E-14
pspB	protein translocation [30,31], greater motility, and the loss of membrane potential when cells are subjected to a specific, proton-motive force (pmf)-depleting stress	3,28	0,83	4,12E-38	−2,70	4,23	1,55E-19
pspC	[31]. activator: IHF	2,95	1,01	3,00E-31	−2,86	4,38	1,43E-18
pspD		2,66	0,08	1,34E-19	−3,11	3,47	2,25E-23
pspG		2,86	−0,84	2,09E-16	−2,44	2,68	7,06E-11
iscR	Regulator of iron-sulfur cluster [32,33], biofilm formation, FimE recombinase [34], repressor of *hya*-operon	0,32	2,45	2,00E-01	−1,28	4,82	2,61E-05

(*continued on next page*)

Table 1 (*continued*)

Gene	Gene function and regulation	*ardA* from B. bifidum			*ardA* from pKM101		
		logFC	logCPM	FDR	logFC	logCPM	FDR
hyaA	Hydrogenase 1, iscR - repressor	−0,41	−1,02	3,22E-01	3,01	1,63	1,25E-15
hyaB		−1,22	−0,72	1,83E-04	2,81	1,40	5,36E-15
hdeA	HdeA is an energy-independent chaperone that protects periplasmic proteins from acid-induced aggregation. *hns*-dependent expression HdeA and HdeB are structural homologues	−1,67	−2,41	5,50E-03	2,47	−0,68	1,28E-03
hdeB		−3,18	−2,84	3,86E-04	2,97	−1,74	2,20E-02
hdeD		−2,81	−3,14	6,49E-03	3,35	−1,52	3,58E-02
ariR	Expression of *ariR* is regulated by the repressor BluR [35] and the alternative sigma factor RpoS [35,36]. Deletion of ariR increases biofilm formation and motility, reduces acid resistance [37]. AI-2 repress *ariR* expression [23].	4,07	−2,83	8,22E-10	−0,82	−0,96	4,30E-01
tnaA	Indole by tryptophanase [38]. Indole is a signal [39,40] that inhibits *E. coli* biofilms [41] and works in a quorum sensing fashion [37]. Activator - CRP,	−5,96	6,21	1,27E-122	3,14	5,34	3,51E-15
tnaB		−5,13	2,39	1,44E-72	3,53	2,05	1,28E-13

between the samples with the *ardA* gene and the controls were identified.

As a result, it was shown that in response to the appearance of *ardA_pKM* in the cell, the expression of 396 genes significantly (95 % confidence interval) increased and 398 significantly decreased. For *ardA_Bf*, 893 genes significantly increased their expression and 910 of them were down-regulated. The list of *E. coli* genes, with their expression changed under acting of the *ardA_Bf* and *ardA_pKM* genes is presented in Supplementary Tables S1 and S2, respectively. For genes whose expression significantly differs between the experimental groups, the GO analysis was performed to find certain groups of genes (gene categories, gene ontologies) that are enriched with differential genes. The results of this analysis are presented in supplementary tables S3 and S4. Fig. 4 shows genes with a logFC value greater than 3 (transcriptional changes more than 8-fold) when the *ardA_Bf* or *ardA_pKM* genes are added to the cell. The transcription changes for these genes are compared.

The data presented in Tables S3 and S4 and on Fig. 4 show that the introduction of the *ardA_Bf* or *ardA_pKM* genes into *E. coli* leads to a change in the transcription of different genes. There are genes whose expression is affected by both *ardA_Bf* and *ardA_pKM*, but nine times out of ten the impact is opposite: if in one case activation, then in another repression.

The comparison of the impact of *ardA* genes from *B. bifidum* and pKM101 on the gene transcription in *E. coli*.

An important topic of this study is to compare the impact on gene transcription in *E. coli* of structurally different ardA genes: *B. bifidum* and pKM101. We obtained 782 differentially expressed genes (DEGs) for the plasmid ardA variant and 1552 for the genomic ardA ones (two-sided p-value <0.01), Moreover, 400 DEGs from these lists overlapped between groups. It seems that they are regulated by both plasmid as well as genomic antirestrictases.

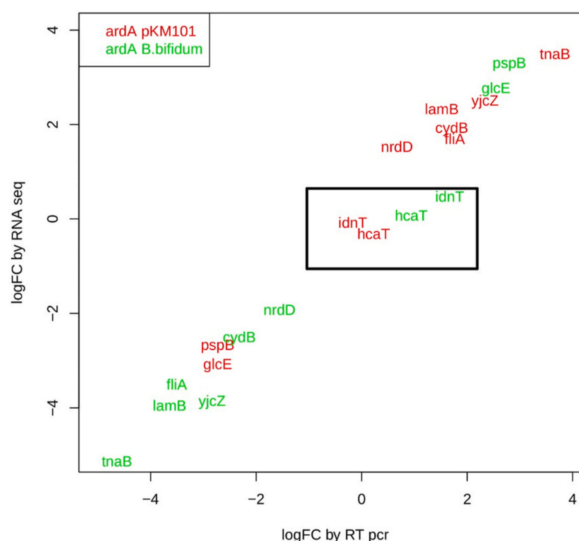


Fig. 5. Gene expression validation. The plot of logFC values, which were obtained by RNA-seq and qRT-PCR methods, for ten DEGs and housekeeping genes. Red colored captions show logFC for DEGs in presence of plasmid ardA variant from pKM101, green colored captions show logFC for DEGs in the presence ardA variant from B. bifidum genome. The black rectangle in the center of the plot denotes housekeeping gene logFC values.

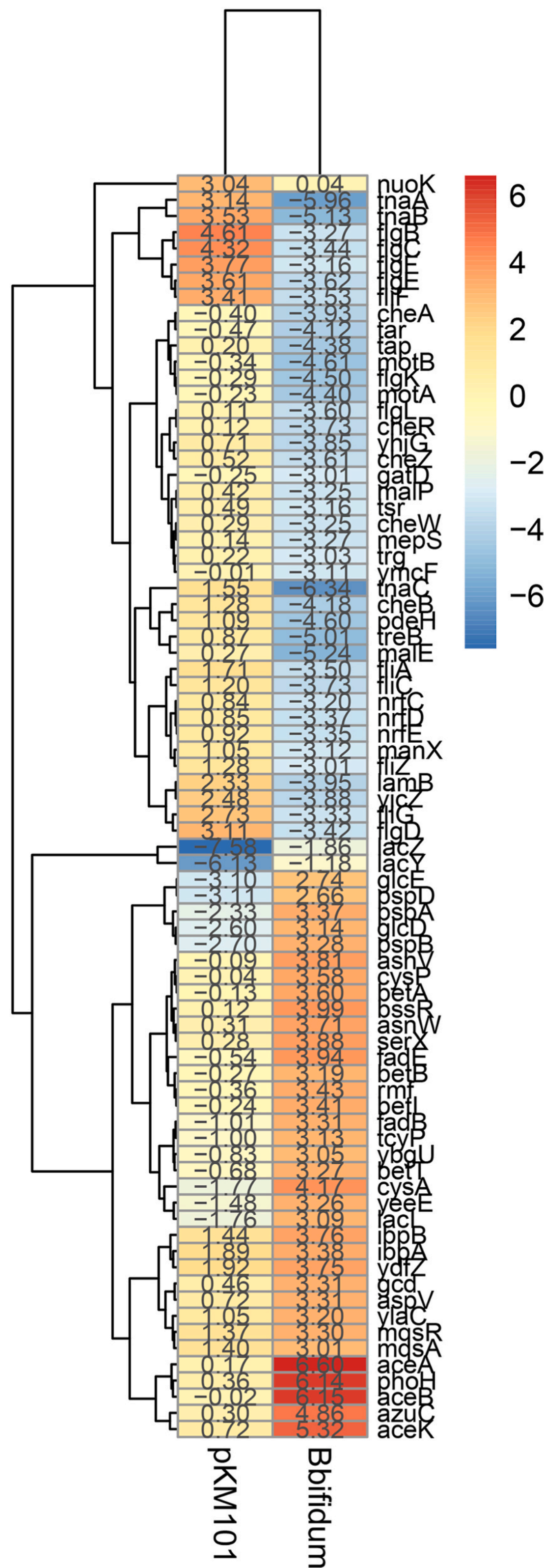


Fig. 6. Biofilm formation. Optical density (OD) values distribution for *E. coli* cells, containing *ardA*_Bf and *ardA*_pKM genes and empty cloning vector. The values for 600 nm wavelength correspond to planktonic cell concentration, while values for 570 nm correspond to attached cells, biofilm formation. The measurement is presented on a logarithmic scale.

To estimate the significance of the deviation of the obtained values in 400 common DEGs from the theoretical ones, we used the binomial test, with the theoretical frequency of coincidence as 0.0621913 ((782/4419) × (1553/4419) – multiple frequencies of both gene lists). The test result revealed that true probability of intersection is higher than theoretical ones. The two-sided p-value (1.325e-13) shows the high reliability of this result. Thus, we determined that antirestrictases from different sources (conjugative plasmids or bacterial genome) regulate the gene expression to a large extent of the same genes.

At the next step, we performed a correlation analysis of changes in gene expression (logFC changes) for genes regulated by plasmid and genomic ardA gene variants. We found a significant negative correlation in gene expression regulation between these two variants (statistical power −0.1846632, with two-sided p-value = 7.648e-11). This means that antirestrictases from the bacterial genome and from the plasmids have a opposite function in the bacterial cell and regulate the same genes differently.

The analysis of gene list with this opposite antirestrictase regulation effect, we obtained a lot of genes, which are involved in horizontal gene transfer (HGT), and flagellum and biofilms formation. The list of genes, which could be responsible for lateral genetic material transfer and biofilm formation, with their expression data and statistical test results, is present in Table 1.

Verification of opposite regulation of *E. coli* genes by ArdA_Bf and ArdA_pKM antirestrictases using qRT-PCR.

To verify RNA-seq results, we conducted qRT-PCR for several of the genes with opposite gene expression changes in presence ardA from pKM101 and from B. bifidum genome. The values of the gene expression (logFC), obtained by the RNA-seq and qRT-PCR methods had a high degree of correlation (correlation value - 0.980 for the plasmid ardA gene variant, and 0.984 for the genomic one). Fig. 5 shows results of the comparison of logFC values, for the ten genes analyzed. In the figure no data are provided for both ardA genes, because they don't valid for expression shift, since their expression is absent in controls. Also, the data of cysG gene expression changes is not shown in the figure, because its expression value was used as control data for normalizing expression of other genes.

The data obtained using qRT-PCR (Fig. 5) strongly support the opposite regulation of a number of DEGs and housekeeping genes by antirestrictases ArdA_Bf and ArdA_pKM and confirm the data obtained using transcriptome analysis.

2.4. The influence of ardA genes on biofilm formation

Because of the complexity of the biofilm formation regulation it is difficult to unambiguously predict the influence of *ard*-gene introduction on the final effect. We conducted a biofilm formation assay for cells with *ardA_Bf*, *ardA_pKM* and controls with pBluescript KS + vector. The result of the assay is present in Fig. 6.

We showed that in presence of the *ardA* genes the concentration of attached cells significantly increases. Moreover, *ardA*_pKM shows more significant effect than chromosomal *ardA*_Bf. The two-sided p-*value* of the ratio of OD600/OD570 for vector vs *ardA*_pKM is 8.238e-06, for vector vs *ardA*_Bf 1.406e-05.

According to transcriptome data *ardA*_Bf and *ardA*_pKM genes affect on a number of biofilm formation genes (*ansB, fau, yjhQ, fim, iscR, ariR, tna* in Table 1). This regulation is often opposite. However the data from Fig. 6 demonstrate that both *ardA*_Bf and *ardA*_pKM genes enhance the biofilm formation although chromosomal *ardA*_Bf seems to be less effective.

2.5. Putative mechanism of gene regulation

The mechanism of gene regulation by antirestrictases remains unclear, however, it can be assumed that if these proteins work as DNA mimetic molecules, then they can interact with some affinity to different DNA-binding proteins, such as transcription factors (TF). Direct inhibition of transcription factors may lead to a change in the expression of these TF targeted genes. In this case, the shift in gene expression, in the presence of *ardA*, will be correlated with the regulation of TFs — genes that are regulated by some TF will change their expression synchronously.

To determine whether this is the case, we assigned all the genes to their individual TF, which regulates the genes, and conducted ANOVA tests for these groups using the logFC value in the presence of *ardA*. Statistical tests are presented in Table 2.

According to Table 2 the expression of genes, more likely, changes synchronously for groups of TF targeted genes — the intra-group average values significantly differ from the intergroup one. Thus, for both tested *ard* genes, the same mechanism of the effect on the gene expression takes place, which is associated with the modulation of TF activity. The effect on TF may be opposite and the range of inhibited TF may be different, however, the mechanism of interaction seems to be the same.

Table 2
ANOVA analysis of the values of expression changes (logFC) in the presence of *ardA*_Bf and *ardA*_pKM genes, grouped according to the regulation by individual transcription factors.

logFC_pKM101					
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
TF	210	1024	4.878	3.588	<2e-16
Residuals	4391	5969	1.359		
logFC_bifidum					
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
TF	210	1914	9.113	4.963	<2e-16
Residuals	4391	8063	1.836		

3. Discussion

Previously studied *ardA* genes were described in bacterial plasmids ([1–3]1, 2, 3). Here, for the first time, the antirestriction function of *B. bifidum* chromosome variant of the *ardA* gene was confirmed. Chromosomal *ardA* product also has the antirestriction effect in bacterial cells despite the weak sequence similarity and different length. The necessity of this antirestriction activity in the cell itself is not entirely clear, since it was previously supposed that antirestriction genes protect plasmids from host cell RM systems and help them to spread through bacterial cells. However, these *ardA* genes, located directly in the bacterial chromosome, should be useful for the bacteria themselves. Significant differences in the structure of ArdA_Bf and ArdA_pKM, including the presence of an additional C-domain in the chromosomal ArdA_Bf, suggest the differences in specificity when interacting with DNA-binding proteins.

Perhaps some cells "intentionally" allow mobile genetic elements to enter the cell, since the exchange of genetic material in certain cases can be useful – for example, cells have to change fast to adapt against external stresses. This is the basis for the SOS mutagenesis phenomenon, which is based on inaccurate DNA synthesis using a special "mutagenic" UmuC polymerase. This mechanism is activated when a bacterial cell gets into stressful conditions and makes a lot of mutations during the cellular DNA synthesis [[42], 4,42]. The mutants can also reverse to the wild type with a high frequency, if a stressful component, which assimilation affected by the mutation, disappears from the environment. Perhaps, it is a good idea to increase the exchange of genetic material in stressful conditions, given that natural microbial communities are quite diverse and the probability that suitable genes can be found in the environment is relatively high. Also, accepting external genetic material could be useful sometimes, e.g. in the case of the appearance of antibiotics in the environment. Mobile genetic elements often contain genes, that confer resistance to antibiotics, and such acceptance should be an adaptive process.

The synthesis of unmodified chromosomal DNA in the cell is possible under certain conditions. This can occur, for example, when cells are irradiated with UV or during cell growth in the presence of 2-aminopurine or 5-bromuracil and their analogs [43–46]. In this case, the cell RM system can pose a threat to the cell itself and an inhibitor of this system can help the cells survive. It is known that another plasmid antirestrictase, *ardB*, which is not a DNA mimetic, artificially introduced into a bacterial cell on a molecular vector, can protect the cell from damage by its own restrictases in the presence of 2-aminopurine, which induces the synthesis of unmodified DNA [8]. It is possible that the *ardA* cell antirestrictase from *B. bifidum* also protects cell from its own endonucleases of the RM system.

In this study, we present a large-scale experiment of *E. coli* gene expression changes under the influence of *ardA* antirestriction genes. For the first time, these investigations were conducted for the new genomic variant of *ardA* gene and for the well-known plasmid *ardA* variant. Intriguing results were obtained regarding the regulatory function of *ardA* and it was shown that plasmid and bacterial genome antirestrictases regulate the activity of genes which can affect the transfer of genetic elements into a bacterial cell. However, antirestrictases from different sources regulate bacteria gene activities in their own way. Plasmid *ardA* regulates the genes facilitate HGT, while the genomic version of *ardA*, on the contrary, complicates this process. According to the obtained results, regulation activity can impact on some cell property in different ways. For example, biofilm formation, induced by plasmid *ardA* regulator, increased through c-di-GMP way, by inducing of *pdeH* phosphodiesterase, and through Fe/S biogenesis way, by down-regulating of *iscR* and *fim* gene cluster. Along with the regulation of biofilm formation, Fe/S biogenesis pathway changes, are involved in phage infection control. That is revealed the polygenetic and pleiotropic effects of *ardA* regulation. It concerns regulation activity of plasmid antirestrictase *ardA* from pKM101, but bacterial *ardA* often acts differently. However, we should be careful in reasoning when making a conclusion about the action of *ardA* from *B. bifidum* genome in this experiment, because we don't know whether the *ardA* would regulate genes of *B. bifidum* in the same way as genes of *E.coli*. That is quite difficult to estimate because *B. bifidum* is an obligate anaerobe and special equipment is required for its cultivation. However, it could be stated from our studies that the proteins of the ArdA family differing in sequence and structure, are able to regulate the expression of a number of chromosomal genes. Moreover, the set of regulated genes varies for two studied genes *ardA*_Bf and *ardA*_pKM, i.e. they demonstrate specificity to targets.

Our experiment shows that plasmid *ardA* gene variant influence on more active the biofilm formation. At the same time, in general, the growth rate of the bacteria in the presence of *ardA* is higher. This phenomenon requires further research and explanation. The question of the mechanisms of regulation of *ardA* genes remains open. The results of ANOVA test show that genes, regulated by the same TF, change their expression, in the presence of *ardA*, to a large extent, synchronously. Indirectly, this suggests that the regulatory function of *ardA* is associated with interaction to TF regulator — there is probably a direct inhibition of TF by the antirestrictase and, as a result, all genes regulated by this TF change their expression. This would be the most obvious explanation of the regulatory action of antirestrictases, especially since there is evidence that antirestrictases interact not only with type I restriction-modification systems but also with RNA polymerase, partially inhibiting it [11] and participate in H-NS related silencing [10]. H-NS plays an important role in the transcription suppression of AT-rich genes because AT enrichment is a trait of many genes involved in the HGT [47,48]. Suppression of the foreign gene expression by H-NS, allows bacteria to safely acquire new genetic material without compromising their genomic integrity [49–51].

> We can conclude that the regulatory activity of ardA from the plasmid leads to the activation of genes that promote HGT, whereas ardA from the bacterial chromosome induces genes that protect cell from the HGT.

This work could have a natural applied continuation: DNA-mimic proteins regulate the transcription from the specific set of promoters, apparently due to interaction with regulatory DNA-binding proteins. The data obtained open the prospect for the development of drugs based on DNA-mimic proteins targeted at site-specific DNA-binding regulators.

4. Materials and methods

4.1. Bacterial strains, plasmids and phages

The strains of *E. coli* K12 used in the work are: strain AB1157 (F- thr-1, leu-6, proA2, his-4, thi-1, argE3, lacY1, galK2, ara14, xyl-5, mtl-1, tsx-33, rpsL31, supE44, r + m+); strain TG1 glnV44 thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5($r_K^- m_K^-$) F′ traD36 proAB $^+$ $lacI^q$ lacZΔM15, the bacteria B. bifidum strain Ac-1784 [52], which was obtained from the Russian National Collection of Industrial Microorganisms (VKPM, https://vkpm.genetika.ru/).

The plasmid pAB7, containing the ardA gene from conjugative plasmid pKM101 in the pBluescript (KS+) vector, was provided by Bacterial Genetics Laboratory of the Federal Institution "State Research Institute of Genetics and Selection of Industrial Microorganisms of the National Research Center" Kurchatov Institute" (GENETIKA) (Belogurov et al., 1992).

The bacteriophage λ was provided by the Bacterial Genetics Laboratory (GENETIKA). We used unmodified phages λ_0 and modified phages λ_k grown on *E. coli* K12, strain TG-1 and *E. coli* Kl2, strain AB1157, respectively.

The ardA_Bf gene was amplified using B. bifidum chromosomal DNA and following primers:

direct: 5’ - CGC CAT ATG GCG GAA GAC GAT CTG – 3’ and

reverse: 5’ - GGC CTG CAG GGC GTA TGC CGT CGA GCA – 3’.

Cleaned resulting PCR product was ligated to the pKAN-T cloning vector with T4 ligase (New England Biolabs, USA). Resulting construct was verified using the Applied Biosystems 3730xl genetic analyzer.

The ardA_Bf_cutG240 gene was amplified using following primers:

Dir_Bifi_pIRDPAL 5′-TCACCATCACCACCATATGGCGGAAGACGAT – 3′ and.

4.1.1. Rev_Bifi_pIRDPAL_Cut_G240 5′-CAGCGGTGTCATTATTCCTTAGCCGGGCTCGT-3′

Cleaned resulting PCR product was ligated to pIRDPAL cloning vector [53] using Gibson Assembly with NEBuilder HiFi master mix (NEB).

The resulting constructs named pKAN-T-ArdA_Bifi and pIRDPAL-ArdA_Bifi_cutG240 respectively were transformed into the *E. coli* strain AB1157 for further manipulations.

4.2. Antirestriction activity measure by phage infection

The antirestriction value was estimated as described previously [54] by comparing λ_0 bacteriophage titers in AB1157 cells with *ardA* gene and in AB1157 cells without it, using the “double agar layer” method (or method Gratia). The control strain was *E. coli* K-12 TG-1, which is usually used for phage infection experiments, since it does not have a RMI system.

To determine the λ_0 titer, a single bacterial colony of AB1157 was inoculated into falcons with 3 ml of Luria-Bertani (LB) culture medium with appropriate antibiotics and the culture was grown for 16 h at 37 °C and 190 rpm in a shaking incubator. After that, 20 ml of the overnight culture was transferred to 2 ml of a new LB broth and incubated at 37 °C and 190 rpm in a shaking incubator for 1.5 h to obtain bacterial cultures in the logarithmic growth phase. Then, the suspension of the bacteriophage lambda in a specific dilution was mixed with bacterial culture in a ratio of 1:2, introduced into a low concentration LB agar (0.7 %) and layered on the surface of the previously prepared 1.5 % LB agar in a Petri dish. Then the cells were incubated at 37 °C for 12–16 h. The amount of phage to be added was estimated experimentally so that no more than several hundred phage plaques were formed on the cup. It is possible to determine the effectiveness of anti-restriction by comparing the number of plaques on individual cultures.

4.3. RNA-sequencing analysis

For each *ardA* gene, we prepared six RNA libraries: three *ardA* + clones and three *ardA*-clones. *E. coli* cell cultures with the appropriate constructs were grown overnight at 37 °C in Luria-Bertani medium with the appropriate antibiotic, and 0.8 ml of overnight culture was taken for RNA isolation. The cells were centrifuged at 5000 rcf for 2 min, the precipitate was resuspended in ExtractRNA reagent (Evrogen, Russia) and incubated for 20 min at 55 °C. The resulting mixture was centrifuged for 10 min at 13,000 rcf, extracted with chloroform, precipitated with isopropyl alcohol, and washed with 80 % ethanol. The concentration and purity of the extracted RNA were evaluated using a Qubit4 fluorometer (ThermoFisher Scientific, USA).

The cDNA libraries were prepared using the NEBNext Ultra II kit for Illumina (New England Biolabs, USA) according to the protocol for use with purified mRNA or rRNA, depleted RNA" (Chapter 4) of the kit manual. The libraries were multiplexed using the NEBNext oligos set (96-well tablet format, both forward and reverse). The quality and size of the library were determined using a Bioanalyzer 2100 (Agilent, USA) using a high-sensitivity DNA Concentration Measurement Kit. The concentration of the library was determined by the Qubit4 fluorometer (ThermoFisher Scientific, USA).

The cDNA libraries were sequenced on the Illumina Novaseq6000 genome analyzer (Illumina, USA) with paired-end reads of 150 bp in length.

4.4. ArdA structure analysis

The prediction of the 3D structure of ArdA_Bf was performed using AlphaFold2 0. Predicted structure was further validated using

protein structure validation software (PSVS) tool. The structures were aligned and analyzed using PyMOL (The PyMOL Molecular Graphics System, Version 2.5.2, Schrödinger, LLC). Analysis of the putative structures of ArdA proteins was performed using the Predictor of Natural Disordered Regions (PONDR®)

4.5. qRT-PCR validation of differentially expressed genes

Total RNA was extracted as described previously. For each experiment (*ardA*+) and each control (*ardA*-) three different colonies were analyzed as biological replicas for each gene. Accordingly, for thirteen genes, seventy eight reactions were done. Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) was performed on CFX96 qPCR instrument (Biorad, USA), using OneTube RT-PCR SYBR reagent kit (Evrogen, Russia) with following conditions:

55 °C–15 min, reverse transcription stage; 95 °C–1 min reverse transcriptase deactivation and DNA polymerase activation; then 40 cycles of amplification with 95 °C–15 s; 55 °C–20 s and 72 °C–20 s.

Primer sequences were designed with web version of Primer3 software [55]. Primer sequences for each selected gene are presented in Table S5.

Primer sequence for RT PCR confirmation of the gene expression changes. *ardA* and housekeeping gene are marked. The housekeeping genes were chosen according work [56].

From the qRT-PCR data, an Cq values were calculated for each gene and were normalized then to the *cysG* house-keeping gene Cq values with pcr_analyze function of 'pcr' R package [57]. The obtained relative expression data were used for comparing with RNA-sequencing data.

4.6. Biofilm formation assay

Biofilm formation assay was conducted as described in Ref. [34]. Overnight cultures grown in LB medium were diluted and normalized to an OD600 of 0.05 in fresh LB medium. Aliquots of 200 μl of diluted fresh culture were added to 96-well microtiter plates and cells were grown at 30 °C for 48 h without shaking. The level of planktonic cell growth was determined by measuring the final OD600 using the plate reader Clariostar (BMG Labtech, Germany) with a path length of 0.6 cm. Planktonic cells were removed, and wells were washed with distilled water two times to remove unattached cells.

A total of 220 μl of 0.1 % (Weight/Volume) crystal violet (Sigma-Aldrich, USA) was used to stain the attached cells for 10 min. Unattached dye was rinsed away by washing three times with distilled water, the plate was dried for 30 min, and stained biomass was dissolved with 1:4 (vol/vol) mixture of acetone and ethanol. After 10 min, the OD570 was measured to quantify biofilm cells.

4.7. Differential expression gene analysis of RNA sequencing data

Genomic sequences of the *B. bifidum* Ac-1784 strain, as well as other strains of the VKPM collection, were obtained during another project [52]. Search of *ardA* genes was performed using PF07275 model from Pfam database [58] and hmmsearch tool from HMMER software package [59] with all default settings but E-value =< 1E-6 and with all heuristic filters turned off.

Amino acid sequence alignment of the ArdA proteins, shown in Fig. 1, was done with Mview web service at EMBL-EBI site [60].

To analyze the gene expression, we mapped the Illumina data to the reference genome of *E. coli* K12 (NCBI: NC_000913. 3) using the bowtie2 software package version 2.3.4.1 [61] with a set of parameters "– very sensitive". Nucleotide reads in the *.sam file format, were converted to a binary file -*. bam, sorted and indexed using the samtools package version 1.7 [62]. The coverage of each gene was evaluated using "BamToBed" command of Bedtools package version 2.26.0 [63]. The number of reads for each gene was combined into a single table using a custom perl script, and differential analysis was performed with the edgeR [64] using Fisher's exact test to differentially expressed (DE) gene estimating.

For ANOVA analysis of changes in transcription factor (TF) target gene expression, we obtained TFs and their targets from the RegulonDB database [65]. Dataset with name "Datasets supported by literature with experimental evidence" were used. For each gene, expression differences were determined in the presence of *ardA* genes, as described above. The ANOVA test was performed in an software environment for statistical computing and graphics - R, using the aov function. The logarithm of gene expression fold change was used as estimated variable, TF regulon membership was used as grouping variable.

Gene ontology (GO) analysis of DE genes with FDR<0.05 was done with David web service (v.6.8) [66],[67]).

Data availability

Data are available at the NCBI SRA database: NCBI accession numbers are also presented in Supplementary Table S0.

Additional information

The work of A.A. Kudryavtseva, related to structural studies of protein interfaces, the development of approaches and methods for the analysis of folding and formation of spatial structures of soluble antirestriction proteins, was supported by the Russian Science Foundation (project No. 22-74-00027).

Funding statement

This work has been carried out using computing resources of the federal collective usage center Complex for Simulation and Data Processing for Mega-science Facilities at NRC “Kurchatov Institute”, http://ckp.nrcki.ru/. The work of I.V. Manukhov, related to work administration and paper writing and edition was supported by Ministry of Science and Higher Education of the Russian Federation (agreement 075-03-2022-107, project FSMG-2021-0002, project FSMF-2022-0007).

Conflict of interest

Conflict of interest The authors declare no conflict of interest.

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.

Acknowledgements

Professor Gennady Zavilgelsky passed away before the submission of the final version of this manuscript. He was the first who described the *ardA* genes and their antirestriction activity. This work is dedicated to his bright memory.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.heliyon.2023.e22986.

Cloning of *ardA* from *B.bifidum* genome, RNA seq analysis and DE gene analysis were supported by an internal grant of the National Research Center Kurchatov Institute No. 591 of March 26, 2020 (Artificial evolution of restriction-modification enzyme inhibitors, connection with natural polymorphism of the antirestrictases).

References

[1] G.B. Zavil’gel’skiĭ, V.I. Mershavka, T.N. Iusifov, A.A. Belogurov, [Weakening of bacteriophage lambda EcoK DNA restriction in the presence of plasmid pKM101 ard+. I. General characteristics and genetic localization], Mol. Biol. (Mosk). 18 (1984) 1590–1596.

[2] A.A. Belogurov, E.P. Delver, O.V. Rodzevich, IncN plasmid pKM101 and IncI1 plasmid ColIb-P9 encode homologous antirestriction proteins in their leading regions, J. Bacteriol. 174 (1992), https://doi.org/10.1128/jb.174.15.5079-5085.1992.

[3] P.M. Chilley, B.M. Wilkins, Distribution of the ardA family of antirestriction genes on conjugative plasmids, Microbiology 141 (1995), https://doi.org/10.1099/13500872-141-9-2157.

[4] G.B. Zavilgelsky, Antirestriction, Mol. Biol. 34 (2000) 724–732, https://doi.org/10.1007/BF02759614.

[5] N.J. Althorpe, P.M. Chilley, A.T. Thomas, W.J. Brammar, B.M. Wilkins, Transient transcriptional activation of the Incl1 plasmid anti-restriction gene (ardA) and SOS inhibition gene (psiB) early in conjugating recipient, Mol. Microbiol. 31 (1999), https://doi.org/10.1046/j.1365-2958.1999.01153.x.

[6] A.A. Belogurov, E.P. Delver, O. V Rodzevich, Plasmid pKM101 encodes two nonhomologous antirestriction proteins (ArdA and ArdB) whose expression is controlled by homologous regulatory sequences, J. Bacteriol. 175 (1993) 4843–4850, https://doi.org/10.1128/jb.175.15.4843-4850.1993.

[7] A.A. Kudryavtseva, I.S. Okhrimenko, V.S. Didina, G.B. Zavilgelsky, I.V. Manukhov, Antirestriction protein ArdB (R64) interacts with DNA, Biochemist 85 (2020) 318–325, https://doi.org/10.1134/S0006297920030074.

[8] D. Serfiotis-Mitsa, A.P. Herbert, G.A. Roberts, D.C. Soares, J.H. White, G.W. Blakely, D. Uhrín, D.T.F. Dryden, The structure of the KlcA and ArdB proteins reveals a novel fold and antirestriction activity against type I DNA restriction systems in vivo but not in vitro, Nucleic Acids Res. 38 (2009) 1723–1737, https://doi.org/10.1093/nar/gkp1144.

[9] S.A. McMahon, G.A. Roberts, K.A. Johnson, L.P. Cooper, H. Liu, J.H. White, L.G. Carter, B. Sanghvi, M. Oke, M.D. Walkinshaw, G.W. Blakely, J.H. Naismith, D.T. F. Dryden, Extensive DNA mimicry by the ArdA anti-restriction protein and its role in the spread of antibiotic resistance, Nucleic Acids Res. 37 (2009) 4887–4897, https://doi.org/10.1093/nar/gkp478.

[10] O.E. Melkina, I.I. Goryanin, G.B. Zavilgelsky, The DNA–mimic antirestriction proteins ArdA ColIB-P9, Arn T4, and Ocr T7 as activators of H-NS-dependent gene transcription, Microbiol. Res. 192 (2016) 283–291, https://doi.org/10.1016/j.micres.2016.07.008.

[11] F. Ye, I. Kotta-Loizou, M. Jovanovic, X. Liu, D.T.F. Dryden, M. Buck, X. Zhang, Structural basis of 1 transcription inhibition by the DNA mimic protein OCR of bacteriophage t7, Elife 9 (2020), https://doi.org/10.7554/eLife.52125.

[12] J. McCann, N.E. Spingarn, J. Kobori, B.N. Ames, Detection of carcinogens as mutagens: bacterial tester strains with R factor plasmids, Proc. Natl. Acad. Sci. U.S. A. 72 (1975), https://doi.org/10.1073/pnas.72.3.979.

[13] S.M. Rastorguev, G.B. Zavilgelsky, Role of “antirestriction” motif in functional activity of antirestriction protein ArdA pKM101 (IncN), Russ. J. Genet. 39 (2003), https://doi.org/10.1023/A:1022444212370.

[14] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021), https://doi.org/10.1038/s41586-021-03819-2.

[15] A.A. Kudryavtseva, E. Cséfalvay, E.Y. Gnuchikh, D.D. Yanovskaya, M.A. Skutel, A.B. Isaev, S. V Bazhenov, A.A. Utkina, I. V Manukhov, Broadness and specificity: ArdB, ArdA, and Ocr against various restriction-modification systems, Front. Microbiol. 14 (2023), https://doi.org/10.3389/fmicb.2023.1133144.

[16] P. Tielen, N. Rosin, A.K. Meyer, K. Dohnt, I. Haddad, L. Jänsch, J. Klein, M. Narten, C. Pommerenke, M. Scheer, M. Schobert, D. Schomburg, B. Thielen, D. Jahn, Regulatory and metabolic networks for the adaptation of Pseudomonas aeruginosa biofilms to urinary tract-like conditions, PLoS One 8 (2013), https://doi.org/10.1371/journal.pone.0071845.

[17] R.J. Allen, E.P. Brenner, C.E. VanOrsdel, J.J. Hobson, D.J. Hearn, M.R. Hemm, Conservation analysis of the CydX protein yields insights into small protein identification and evolution, BMC Genom. 15 (2014), https://doi.org/10.1186/1471-2164-15-946.
[18] X. Garriga, R. Eliasson, E. Torrents, A. Jordan, J. Barbé, I. Gibert, P. Reichardf, nrdD and nrdG genes are essential for strict anaerobic growth of Escherichia coli, Biochem. Biophys. Res. Commun. 229 (1996), https://doi.org/10.1006/bbrc.1996.1778.
[19] Y. Kang, K.D. Weber, Y. Qiu, P.J. Kiley, F.R. Blattner, Genome-wide expression analysis indicates that FNR of Escherichia coli K-12 regulates a large number of genes of unknown function, J. Bacteriol. 187 (2005), https://doi.org/10.1128/JB.187.3.1135-1160.2005.
[20] H.G. Heine, J. Kyngdon, T. Ferenci, Sequence determinants in the lamB gene of Escherichia coli influencing the binding and pore selectivity of maltoporin, Gene 53 (1987), https://doi.org/10.1016/0378-1119(87)90018-7.
[21] L. Randall Hazelbauer, M. Schwartz, Isolation of the bacteriophage lambda receptor from Escherichia coli, J. Bacteriol. 116 (1973), https://doi.org/10.1128/jb.116.3.1436-1446.1973.
[22] D. Skliros, P.G. Kalatzis, C. Kalloniati, F. Komaitis, S. Papathanasiou, E.D. Kouri, M.K. Udvardi, C. Kokkari, P. Katharios, E. Flemetakis, The development of bacteriophage resistance in vibrio alginolyticus depends on a complex metabolic adaptation strategy, Viruses 13 (2021), https://doi.org/10.3390/v13040656.
[23] D. Ren, L.A. Bedzyk, S.M. Thomas, R.W. Ye, T.K. Wood, Gene expression in Escherichia coli biofilms, Appl. Microbiol. Biotechnol. 64 (2004), https://doi.org/10.1007/s00253-003-1517-y.
[24] M. Herzberg, I.K. Kaye, W. Peti, T.K. Wood, YdgG (TqsA) controls biofilm formation in Escherichia coli K-12 through autoinducer 2 transport, J. Bacteriol. 188 (2006), https://doi.org/10.1128/JB.188.2.587-598.2006.
[25] P.E. Orndorff, P.A. Spears, D. Schauer, S. Falkow, Two modes of control of pilA, the gene encoding type 1 pilin in Escherichia coli, J. Bacteriol. 164 (1985), https://doi.org/10.1128/jb.164.1.321-330.1985.
[26] L.J. Lee, J.A. Barrett, R.K. Poole, Genome-wide transcriptional response of chemostat-cultured Escherichia coli to zinc, J. Bacteriol. 187 (2005), https://doi.org/10.1128/JB.187.3.1124-1134.2005.
[27] V. Hancock, P. Klemm, Global gene expression profiling of asymptomatic bacteriuria Escherichia coli during biofilm growth in human urine, Infect. Immun. 75 (2007), https://doi.org/10.1128/IAI.01748-06.
[28] C. Pesavento, G. Becker, N. Sommerfeldt, A. Possling, N. Tschowri, A. Mehlis, R. Hengge, Inverse regulatory coordination of motility and curli-mediated adhesion in Escherichia coli, Genes Dev. 22 (2008), https://doi.org/10.1101/gad.475808.
[29] L. Weiner, P. Model, Role of an Escherichia coli stress-response operon in stationary-phase survival, Proc. Natl. Acad. Sci. U.S.A. 91 (1994), https://doi.org/10.1073/pnas.91.6.2191.
[30] M. Kleerebezem, J. Tommassen, Expression of the pspA gene stimulates efficient protein export in Escherichia coli, Mol. Microbiol. 7 (1993), https://doi.org/10.1111/j.1365-2958.1993.tb01186.x.
[31] M. Kleerebezem, W. Crielaard, J. Tommassen, Involvement of stress protein PspA (phage shock protein A) of Escherichia coli in maintenance of the protonmotive force under stress conditions, EMBO J. 15 (1996), https://doi.org/10.1002/j.1460-2075.1996.tb00344.x.
[32] C.J. Schwartz, J.L. Giel, T. Patschkowski, C. Luther, F.J. Ruzicka, H. Beinert, P.J. Kiley, IscR, an Fe-S cluster-containing transcription factor, represses expression of escherichia coli genes encoding Fe-S cluster assembly proteins, Proc. Natl. Acad. Sci. U.S.A. 98 (2001), https://doi.org/10.1073/pnas.251550898.
[33] J.L. Giel, D. Rodionov, M. Liu, F.R. Blattner, P.J. Kiley, IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O2-regulated genes in Escherichia coli, Mol. Microbiol. 60 (2006), https://doi.org/10.1111/j.1365-2958.2006.05160.x.
[34] Y. Wu, F.W. Outten, IscR controls iron-dependent biofilm formation in Escherichia coli by regulating type I fimbria expression, J. Bacteriol. 191 (2009), https://doi.org/10.1128/JB.01086-08.
[35] N. Tschowri, S. Busse, R. Hengge, The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of Escherichia coli, Genes Dev. 23 (2009), https://doi.org/10.1101/gad.499409.
[36] H. Weber, T. Polen, J. Heuveling, V.F. Wendisch, R. Hengge, Genome-wide analysis of the general stress response network in Escherichia coli: σS-dependent genes, promoters, and sigma factor selectivity, J. Bacteriol. 187 (2005), https://doi.org/10.1128/JB.187.5.1591-1603.2005.
[37] J. Lee, R. Page, R. García-Contreras, J.M. Palermino, X.S. Zhang, O. Doshi, T.K. Wood, W. Peti, Structure and function of the Escherichia coli protein YmgB: a protein critical for biofilm formation and acid-resistance, J. Mol. Biol. 373 (2007), https://doi.org/10.1016/j.jmb.2007.07.037.
[38] W.A. Newton, E.E. Snell, Formation and interrelationships of tryptophanase and tryptophan synthetases in Escherichia coli, J. Bacteriol. 89 (1965), https://doi.org/10.1128/jb.89.2.355-364.1965.
[39] P. Di Martino, R. Fursy, L. Bret, B. Sundararaju, R.S. Phillips, Indole can act as an extracellular signal to regulate biofilm formation of Escherichia coli and other indole-producing bacteria, Can. J. Microbiol. 49 (2003), https://doi.org/10.1139/w03-056.
[40] D. Wang, X. Ding, P.N. Rather, Indole can act as an extracellular signal in Escherichia coli, J. Bacteriol. 183 (2001), https://doi.org/10.1128/JB.183.14.4210-4216.2001.
[41] J. Lee, A. Jayaraman, T.K. Wood, Indole is an inter-species biofilm signal mediated by SdiA, BMC Microbiol. 7 (2007), https://doi.org/10.1186/1471-2180-7-42.
[42] N.B. Reuven, G. Arad, A. Maor-Shoshani, Z. Livneh, The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication, J. Biol. Chem. 274 (1999), https://doi.org/10.1074/jbc.274.45.31763.
[43] I. Ivančić-Baće, I. Vlašić, G. Čogelja-Čajo, K. Brčić-Kostić, E. Salaj-Šmic, Roles of PriA protein and double-strand DNA break repair functions in UV-induced restriction alleviation in Escherichia coli, Genetics 174 (2006), https://doi.org/10.1534/genetics.106.063750.
[44] J.E. Kelleher, E.A. Raleigh, Response to UV damage by four Escherichia coli K-12 restriction systems, J. Bacteriol. 176 (1994), https://doi.org/10.1128/jb.176.19.5888-5896.1994.
[45] S. Makovets, V.A. Doronina, N.E. Murray, Regulation of endonuclease activity by proteolysis prevents breakage of unmodified bacterial chromosomes by type I restriction enzymes, Proc. Natl. Acad. Sci. U.S.A. 96 (1999), https://doi.org/10.1073/pnas.96.17.9757.
[46] B. Thoms, W. Wackernagel, Genetic control of damage-inducible restriction alleviation in Escherichia coli K12: an SOS function not repressed by lexA, MGG Mol. Gen. Genet. 197 (1984), https://doi.org/10.1007/BF00330977.
[47] C.J. Dorman, H-NS-like nucleoid-associated proteins, mobile genetic elements and horizontal gene transfer in bacteria, Plasmid 75 (2014), https://doi.org/10.1016/j.plasmid.2014.06.004.
[48] Y.A. Purtov, O.A. Glazunova, S.S. Antipov, V.O. Pokusaeva, E.E. Fesenko, E.V. Preobrazhenskaya, K.S. Shavkunov, M.N. Tutukina, V.I. Lukyanov, O.N. Ozoline, Promoter islands as a platform for interaction with nucleoid proteins and transcription factors, J. Bioinf. Comput. Biol. 12 (2014), https://doi.org/10.1142/S0219720014410066.
[49] C.J. Dorman, The genome sentinel, Nat. Rev. Microbiol. 5 (2007), https://doi.org/10.1038/nrmicro1598.
[50] F.C. Fang, S. Rimsky, New insights into transcriptional regulation by H-NS, Curr. Opin. Microbiol. 11 (2008), https://doi.org/10.1016/j.mib.2008.02.011.
[51] W.W. Navarre, M. McClelland, S.J. Libby, F.C. Fang, Silencing of xenogeneic DNA by H-NS - facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA, Genes Dev. 21 (2007), https://doi.org/10.1101/gad.1543107.
[52] A.A. Korzhenkov, A.V. Tepliuk, K.V. Sidoruk, K.E. Voyushin, M.V. Patrushev, I.V. Kublanov, S.V. Toshchakov, A dataset of four probiotic Bifidobacterium strains genome assemblies, Data Brief 34 (2021), https://doi.org/10.1016/j.dib.2020.106710.
[53] S.V. Bazhenov, E.S. Scheglova, A.A. Utkina, A.A. Kudryavtseva, R. Al Ebrahim, I.V. Manukhov, New temperature-switchable acyl homoserine lactone-regulated expression vector, Appl. Microbiol. Biotechnol. 107 (2022) 807–818, https://doi.org/10.1007/s00253-022-12341-y.
[54] A.A. Kudryavtseva, V.A. Alekhin, M.D. Lebedeva, E. Cséfalvay, Anti-restriction activity of ArdB protein against, EcoAI Endonuclease 57 (2023) 101–104, https://doi.org/10.1134/S0026893323010053.
[55] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm, S.G. Rozen, Primer3-new capabilities and interfaces, Nucleic Acids Res. 40 (2012), https://doi.org/10.1093/nar/gks596.
[56] K. Zhou, L. Zhou, Q. Lim, R. Zou, G. Stephanopoulos, H.P. Too, Novel reference genes for quantifying transcriptional responses of Escherichia coli to protein overexpression by quantitative PCR, BMC Mol. Biol. 12 (2011), https://doi.org/10.1186/1471-2199-12-18.

[57] M. Ahmed, D.R. Kim, Pcr: an R Package for Quality Assessment, Analysis and Testing of qPCR Data, PeerJ vol. 2018 (2018), https://doi.org/10.7717/peerj.4473.
[58] S. El-Gebali, J. Mistry, A. Bateman, S.R. Eddy, A. Luciani, S.C. Potter, M. Qureshi, L.J. Richardson, G.A. Salazar, A. Smart, E.L.L. Sonnhammer, L. Hirsh, L. Paladin, D. Piovesan, S.C.E. Tosatto, R.D. Finn, The Pfam protein families database in 2019, Nucleic Acids Res. 47 (2019), https://doi.org/10.1093/nar/gky995.
[59] S.R. Eddy, Accelerated profile HMM searches, PLoS Comput, Biol. 7 (2011), https://doi.org/10.1371/journal.pcbi.1002195.
[60] F. Madeira, Y.M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A.R.N. Tivey, S.C. Potter, R.D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019, Nucleic Acids Res. 47 (2019), https://doi.org/10.1093/nar/gkz268.
[61] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat. Methods 9 (2012), https://doi.org/10.1038/nmeth.1923.
[62] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, Bioinformatics 25 (2009), https://doi.org/10.1093/bioinformatics/btp352.
[63] A.R. Quinlan, I.M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features, Bioinformatics 26 (2010), https://doi.org/10.1093/bioinformatics/btq033.
[64] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR, A Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics 26 (2009), https://doi.org/10.1093/bioinformatics/btp616.
[65] A. Santos-Zavaleta, H. Salgado, S. Gama-Castro, M. Sánchez-Pérez, L. Gómez-Romero, D. Ledezma-Tejeida, J.S. García-Sotelo, K. Alquicira-Hernández, L. J. Muñiz-Rascado, P. Peña-Loredo, C. Ishida-Gutiérrez, D.A. Velázquez-Ramírez, V. Del Moral-Chávez, C. Bonavides-Martínez, C.F. Méndez-Cruz, J. Galagan, J. Collado-Vides, RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in E. coli K-12, Nucleic Acids Res. 47 (2019), https://doi.org/10.1093/nar/gky1077.
[66] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (2009), https://doi.org/10.1038/nprot.2008.211.
[67] D.W. Huang, B.T. Sherman, R.A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, Nucleic Acids Res. 37 (2009), https://doi.org/10.1093/nar/gkn923.