

OPEN

Genome-scale exploration of transcriptional regulation in the nisin Z producer *Lactococcus lactis* subsp. *lactis* IO-1

Naghmeh Poorinmohammad^{1,2}, Javad Hamedi^{1,2*} & Ali Masoudi-Nejad^{3*}

Transcription is of the most crucial steps of gene expression in bacteria, whose regulation guarantees the bacteria's ability to adapt to varying environmental conditions. Discovering the molecular basis and genomic principles of the transcriptional regulation is thus one of the most important tasks in cellular and molecular biology. Here, a comprehensive phylogenetic footprinting framework was implemented to predict maximal regulons of *Lactococcus lactis* subsp. *lactis* IO-1, a lactic acid bacterium known for its high potentials in nisin Z production as well as efficient xylose consumption which have made it a promising biotechnological strain. A total set of 321 regulons covering more than 90% of all the bacterium's operons have been elucidated and validated according to available data. Multiple novel biologically-relevant members were introduced amongst which *arsC*, *mtlA* and *mtl* operon for BusR, MtlR and XylR regulons can be named, respectively. Moreover, the effect of riboflavin on nisin biosynthesis was assessed *in vitro* and a negative correlation was observed. It is believed that understandings from such networks not only can be useful for studying transcriptional regulatory potentials of the target organism but also can be implemented in biotechnology to rationally design favorable production conditions.

Regulation at transcription level is one of the main mechanisms by which bacteria adapt their metabolism to changing environments. In response to these environmental or internal dynamics, controlling the gene expression is mainly exerted using transcription factors (TFs), proteins that specifically binds TF binding sites (TFBSs) or *cis*-regulatory motifs and in turn change the expression pattern of the downstream gene(s)^{1,2}.

Bacterial genomes are organized in the form of *operons*, a group of genes in tandem sharing a terminator and a promoter each containing one or more TFBSs. The maximal number of co-regulated operons under the direct control of certain TF(s) is then called a regulon, a term which was firstly introduced by Maas *et al.* in 1964³. Studying the coordination of cellular behavior in response to internal or external signals in transcription level as a crucial aspect in microbial genomics depends on our ability to elucidate correctly all these regulons or by other means, the global transcriptional regulatory network (TRN) of a bacterium.

Although the most desired way of elucidating the entire set of regulons in a bacterium is experiment, serious challenges exist. Apart from being highly expensive and time-consuming⁴, laboratory-based elucidation of all regulons requires researchers to know exactly what conditions will trigger which regulons. Therefore, all conditions that will lead to activating all regulons must be known, a procedure which seems to be yet impossible⁵. Moreover, many bacterial TRN elucidation studies use gene expression profiles to predict the interactions of TFs and their target TFBSs^{6–8}. However, there are two problems in this assumption: (a) not always a co-expressed situation denotes co-regulation⁹; and (b) post-transcriptional regulation of TFs can also occur and in this situation the cellular level of the TF is constant however the activity is altered. These issues can also compromise the whole predictions.

To this end, computational tools have been designed to elucidate all the co-regulated operons encoded in bacterial genomes with their main aim being to find the maximal sets of operons that share conserved *cis*-regulatory

¹Department of Microbial Biotechnology, School of Biology and Centre of Excellence in Phylogeny of Living Organisms, College of Science, University of Tehran, Tehran, Iran. ²Microbial Technology and Products (MTP) Research Center, University of Tehran, Tehran, Iran. ³Laboratory of Systems Biology and Bioinformatics (LBB), Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran. *email: jhamedi@ut.ac.ir; amasoudin@ut.ac.ir

motifs (or in short, motifs)^{10–12}. For this *ab initio* regulon inference, motifs must be scanned for all operons via phylogenetic footprinting technique. Since being introduced in 1988¹³, phylogenetic footprinting has greatly improved the computational motif identification process via using a comparative genomics analysis of closely related organisms¹⁴. Finding motifs of all operons in bacterial genomes then allows for further analysis to cluster conserved and similar motifs with the opinion that they will show similar regulation pattern, thus being classified as a regulon. Regulon prediction or TRN reconstruction is a key to understand gene function and evolution of the bacterium being studied.

Lactococcus lactis is one of the best-studied members of lactic acid bacteria (LAB) and can be classified into different subspecies. Of these, *L. lactis* subsp. *lactis* and *L. lactis* subsp. *cremoris* shows high biotechnological potentials ranging from their wide applications in production of fermented dairy products¹⁵ as well as flavor formation¹⁶ and lactic acid production (subsp. *lactis* and *cremoris*)¹⁷ to their implementation as drug delivery systems for many therapeutic proteins (subsp. *lactis*)^{18,19}. Moreover, different strains of *L. lactis* subsp. *lactis* are known as bacteriocin producers. Lantibiotic nisin, the most intensively studied and used bacteriocin, is the first antimicrobial peptide approved by the FDA to be used as food preservative and is produced by many strains of *L. lactis* subsp. *lactis*²⁰. According to other researches, new applications have been also proposed for nisin in biomedical field, such as its potential in treatment of colon intestinal infections²¹, bovine subclinical mastitis²² as well as its synergistic role in treatment of head and neck tumors and colorectal cancer^{23,24} and pan drug-resistant infections^{25–27}. Due to the aforementioned potentials, lots of studies have been conducted to reveal the molecular basis of nisin production as well as other important features of this bacterium to both understand the biological basics and improvement of these potentials such as studies on nisin production or the bacterial cell growth improvement in different conditions ambient to various industries.

L. lactis subsp. *lactis* IO-1 is a highly potent strain in producing L-lactate from xylose as well as glucose being suggested as a promising candidate to be utilized in green plastic industry for the production of polylactic acid (PLA)²⁸. The strain also produces a natural variant of nisin, nisin Z²⁹, which differs from the typical nisin A peptide in only one amino acid which results in more solubility of nisin Z at physiological pH making it a more suitable candidate for biomedical applications.

Since the production of antibiotics is also proved to be regulated tightly at the transcriptional level³⁰, in the present work we have considered reconstructing the global TRN of *L. lactis* subsp. *lactis* IO-1 by a successful phylogenetic footprinting-based approach which has provided an opportunity to better understand and study the regulatory behaviors of the organism by predicting all its regulons from the genomic data.

Results and Discussion

Computational TRN inference. *Reference genome selection and operon identification.* Among the phylum firmicutes, 74 bacterial genomes were finally selected as reference genomes set (Supplementary Table S1). Our analysis indicated that the addition of more genomes to this list will provide negligible benefit to the regulon predictions. Although a set containing as few as 25 genomes were shown to be able to cover near 89% of regulons inferred with the 74 genomes (Fig. S1), using a larger selection of organisms was considered as beneficial rather than the most appropriate organisms.

Accurate curation of predicted operons can greatly increase the regulon prediction. This is since the promoter upstream the first gene in an operon will be considered for motif discovery and mispredicted operons may result whether in false positive or false negative motifs. For instance, if *gene A* naturally forms a single membered operon within a bacterium's genome and the prediction algorithm put it wrongly in the middle of a multi-membered operon, the motif of *gene A* will be missed and will not be taken into account in next steps of regulon prediction. The final set of 1373 operons of *L. lactis* subsp. *lactis* IO-1 are listed in Supplementary Table S2.

Motif finding. In a phylogenetic footprinting framework for motif discovery, finding sufficient orthologous operons for *all* operons in the target genome is another key step since an effective regulon inference relies directly on a sufficient and appropriate number of promoters containing corresponding motifs. Therefore, selecting the maximum number of non-redundant reference genomes (here 74) can benefit the regulon inference in this approach by giving more informative promoters for motif discovery.

Based on OrthoFinder results, the average number of orthologous operons for all operons of *L. lactis* subsp. *lactis* IO-1 is 34 with 81.3% of them having more than 6 orthologous operons. More than 95% of operons possessed over 2 orthologous operons, which were further explored to find conserved motifs. The total of over 5000 motifs was predicted while 89% of operons were predicted with at least 1 significant motif (p -value < 0.05). Therefore, genes without orthology or promoters showing no conservation at sequence level is underestimated which is regarded as an internal limitation of phylogenetic footprinting approach. However, here a relatively high percentage of operons were predicted with motifs. This is even higher than the percentage gained by the MP³ authors while evaluating the algorithm on *E. coli*³¹.

Regulon prediction and evaluation. Co-regulation relationship between a pair of operons has been predicted via similarity comparison of their predicted motifs based on co-regulation score (CRS) (see methods) and as a result, a co-regulated graph and generated operon clusters for regulon prediction was built. Consequently, the reconstructed TRN of *L. lactis* subsp. *lactis* IO-1 includes 321 regulons covering 90.3% of all operons (Tables S3 and S4) with 232 of the regulons having more than 3 operons. Most of the predicted regulons include 2–4 operons (mostly 3) and 21 regulons were single-membered which does not comply with the definition for a regulon and thus a set of 300 final regulons were predicted (Fig. 1). However, although with lower scores, these operons (or single-membered regulons) are connected to other regulons' members using their lower-score motifs (Table S5). To this end, they can be used for network analysis and thus cannot be permanently omitted from the final TRN.

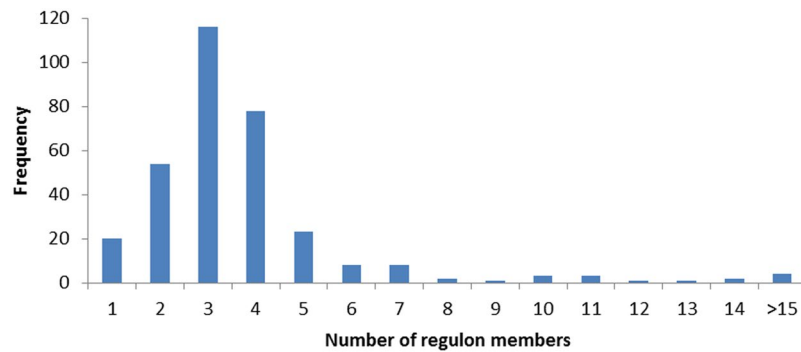


Figure 1. Frequency of predicted regulons with different size.

Regulon ID	Number of operons	Functions enriched	Predicted TF	Tomtom P-value
Regulon_309	30	Global catabolite repression	CcpA	1.02e-6
Regulon_91	15	Amino acid metabolism	CodY	5.54e-07
Regulon_323	16	Purine metabolism	PurR	1.37e-04

Table 1. Regulons with more than 15 members and their predicted TFs/TFBSs.

Most of the regulons having more than 15 operons were predicted to be regulons regulated by global TFs such as CodY, CcpA, PurR. The analysis was performed by creating the positional weight matrices (PWM) of the consensus motifs for operons of these regulons and further finding the most statistically significant TF for each PWM by Tomtom motif comparison tool³². Table 1 shows the global TF regulated regulons and details on target operons can be found in Table S2. The presence of such regulons can be a positive point for evaluation of the present regulon prediction framework showing the success of the implemented procedure in clustering a large number of operons believed to be regulated by global TFs.

The presence of other orthologous lactococcal TF regulons was also studied to further evaluate the regulon prediction. Such data were mostly available for *L. lactis* subsp. *lactis* IL1403 from literature and RegPrecise database and to a lesser extent some literature data were used for *L. lactis* subsp. *lactis* KF147³³ and *L. lactis* subsp. *lactis* F44³⁴. Focusing on the known TF regulons in these organisms, motifs of orthologous operons in *L. lactis* subsp. *lactis* IO-1 were used to verify that they can be governed by the same TF as in the other *L. lactis* strains under study, using Tomtom (p -value < 0.05).

As a result, from 31 known lactococcal regulons studied, 24 regulons were assigned with TFs similar to those regulating the orthologous regulon. In most cases, the orthologous regulon in *L. lactis* subsp. *lactis* IO-1 were significantly similar to the known regulons meaning that in some cases there were some differences in terms of the number of constituting operons. Accordingly, the IO-1 orthologous regulons often possessed a similar or higher number of operons which can both be regarded to as either a false positive or a new member of a previously-unknown regulon, the verification of which needs whether experimental verification or having sufficient biological basis.

Finding novel potential target operons/genes or generally, novel regulons using computational regulon prediction is an accepted routine. For instance, using a knowledge-driven comparative-genomics approach, Leyn *et al.* have found about 140 new target genes for 93 previously studied TF regulons in *B. subtilis* genome which provides useful data with potentials to be readily used in genetic experiments, metabolic modeling, and evolutionary analysis³⁵. Here, totally 30 new regulon members have been predicted for BirA, BusR, CmbR, CodY, CopR, FacT, hrcA, MtlR, NrdR, Rex, and XylR orthologous TF regulons in *L. lactis* subsp. *lactis* TRN (Table 2).

BusR is a repressor of the expression of osmotically inducible glycine-betaine uptake system at low osmolarity in *L. lactis*, thus having a key role in osmoregulation³⁶. Available studies on BusR regulon of lactococci introduce it as a small regulon consisting of *busA* and *busB* genes which encodes transporter for glycine betaine³⁶ which is one of the most important of compatible solutes and osmoprotectants in *L. lactis*³⁷.

Among novel members of the BusR regulon, we found *arsC* which codes for arsenic reductase and is potentially involved in arsenic detoxification³⁸. It is reported that in a different organism such as plants and microorganisms, changes in osmolytes such as glycine betaine content occurs accompanying the exposure of the organism to arsenic toxicity^{39,40}. On the other hand, the biosynthesis of arsenobetaine which is homologue glycine betaine and known as an osmoprotectant, also starts with the *arsC* activity in microorganisms but the exact enzymatic and genetic data on its biosynthesis is not clear⁴¹.

On the other hand, Cheung *et al.* have studied the expression pattern of genes in *E. coli* under osmotic stress amongst which *hom(thrA)* and *thrB* expressions showed 6-fold change denoting their apparent role in osmotic stress⁴². Since there are basic overlaps in osmotic stress responses among bacteria, further study of the role of mentioned genes in osmotic stress response in *L. lactis* can be advantageous in better understanding its physiology. Moreover, *hom(thrA)* and *thrB* participate in the biosynthesis of amino acids such as threonine, they can also

Reference strain*	TF assigned	RegulonID	Biological process	Ratio of the number operons** (in IO-1/ in reference strain)	Novel predicted target genes
IL1403	AdcR	Regulon_212	Zinc homeostasis	2/2	—
	ArgR	Regulon_308	Arginine biosynthesis; Arginine degradation	5/6	—
	BirA	Regulon_155	Biotin biosynthesis	<u>2/2</u>	<i>yseC, yseD, fabG2, yseE</i>
	BusR	Regulon_52	Betaine utilization	4/1	<i>arsC, hom, thrB</i>
	CcpA	Regulon_309	Global catabolite repression	30/31	—
	CmbR	Regulon_310		10/9	<u><i>ymhA</i></u>
	CodY	Regulon_91	Amino acid metabolism	15/12	<i>optA, dppB, lilo_0313, optD, optF, yuaE, sdaB, gltA, citB, icd</i>
	CopR	Regulon_312	Copper homeostasis	<u>5/5</u>	<i>yuhE, yihD</i>
	FabT	Regulon_313	Fatty acid biosynthesis	4/3	<i>ymdE</i>
	GlnR	Regulon_314	Nitrogen assimilation	3/3	—
	HomR	Regulon_315	Methionine metabolism; Cysteine metabolism	2/2	—
	HrcA	Regulon_316	Heat shock response	4/3	<i>yudG, yudH</i>
	MleR	Regulon_297	Malate utilization	1/1	—
	MntR	Regulon_318	Manganese homeostasis	2/2	—
	MtlR	Regulon_319	Mannitol utilization	<u>1/1</u>	<i>mtlA</i>
	NagR	Regulon_320	N-acetylglucosamine utilization	4/3	
	NrdR	Regulon_321	Deoxyribonucleotide biosynthesis	<u>2/2</u>	<i>rpsP</i>
	F44	YthA	Regulon_325	(Acid) stress response	10/10
KF147	XylR	Regulon_10	Xylose metabolism	4/3	<i>mtlA, mtlR, mtlE, mtlD,</i>

Table 2. Orthologous TF regulons in *L. lactis* subsp. *lactis* IO-1 compared to known lactococcal regulons. *IL1403: *L. lactis* subsp. *lactis* IL1403; F44: *L. lactis* subsp. *lactis* F44, *L. lactis* subsp. *lactis* KF147. **The operons in reference strains' regulons are assumed to include same set of genes as in IO-1 except when underlined. If there are novel genes in the regulons with same number of operons due to difference in operon structure, they are shown in the leftmost column.

directly affect osmoprotectants production. This is due to some evidence that threonine metabolites have been also shown to play roles in defenses against salt stress⁴³. Accordingly, the presence of both *arsC* and the operon of *thrAB* can be meaningful in the BusR regulon.

For CodY regulon which includes mainly the amino acid metabolism components, genes encoding oligopeptide transporter (*optA, optD, optF, dppB, lilo_0313*), L-serine dehydratase (*sdaB*), putative aspartate protease (*yuaE*) and enzymes of citrate cycle (*gltA, citB, icd*) all of which are apparently participated in amino acid metabolism were predicted as new targets. *gltA, citB* and *icd* are also reported to be a member of CodY regulon of *Bacillus subtilis*⁴⁴.

Among other new predicted targets, *yuhE* (encoding a copper transporter), *yudG* (encoding a heat shock protein), *mtlA* (encoding a mannitol-specific PTS system component) and *nadE* (encoding NAD-synthetase) in CopR, HrcA, MtlR and NrdR regulons, respectively, are genes whose potential annotated functions are related to their corresponding regulon's biological process.

Generally, determination of TFs and their specificity to a target (TFBS or motif) is highly complicated by the layers of complexity due to the existence of many features modulating this relationship beyond sequence motifs⁴⁵. Of the most important complexities are: relative incapability of experimental techniques to characterize the specificities of all TFs due to the need of expressing soluble and active TFs of interest, possibility of multiple specificities for an individual transcription factor⁴⁶, possibility of cooperative binding by multiple TFs as well as TFs-small molecules or TFs-proteins which in turn can also give rise to new motifs⁴⁵, and finally the effects of contextual and epigenetic information of DNA that may affect TF binding⁴⁷. Moreover, the direct transposition of a model organism's (e.g. using *Bacillus subtilis* in this case) regulatory information does not always allow building a reliable regulatory network in the target genome since many regulators have functions that could not be proposed by transposition of the knowledge currently available in other bacteria⁴⁸.

Consequently, computational assignment of a known TF to predicted regulons usually is not straightforward and more data is needed to evaluate and analysis the regulons. However, elucidating all regulons of a genome can give useful insight into potential cooperating operons/genes. Noteworthy in this respect is that not all the predicted regulons should be regarded as being available at the same time in a bacterial cell, rather a regulon or a subset of the whole predicted TRN is triggered in a specific environmental or physiological condition. Such data

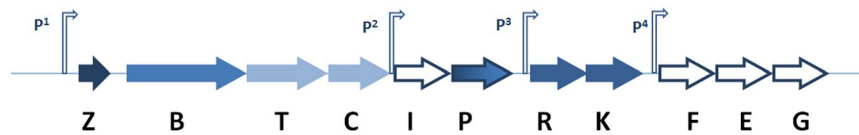


Figure 2. Schematic presentation of nisin Z biosynthesis gene cluster with four characterized operons: nisZBTC, nisIP, nisRK and nisFEG. (Promoters numbers are shown as p1 to p4, these also represents operons numbers).

on potential transcriptional responses in different bacteria can also be used in comparative omics analyses to differentiate close taxa.

Findings on xylose metabolism regulation at the transcriptional level. Very few strains of *L. lactis* from non-dairy sources (plant-based habitats) can ferment xylose; and to-date the most potent one is *L. lactis* subsp. *lactis* IO-1⁴⁹. In biotechnological productions, using cheap raw material as carbon sources greatly increase the economic feasibility of the production. Xylose can be derived from the hemicellulose abundant in agricultural wastes and could be an alternative carbon source for potent biotechnological strains. *L. lactis* subsp. *lactis* IO-1, well-known for its potential in producing nisin Z and lactic acid is reported to perform better than the type strain *L. lactis* in its industrial application in *L-lactate* fermentation⁵⁰, thus the xylose fermentation in this bacteria have been studied mostly via fermentation analysis⁵¹. The data from the present study can be complementary with experimental results which can help better understand the xylose metabolism in this organism.

According to Tables 2 and S3, apart from *xylA*, *xylB* and *xylR* as the main genes in xylose metabolism, the predicted XylR regulon in *L. lactis* subsp. *lactis* IO-1, is shown to possess genes involved in mannitol transport and metabolism which shows a potential overlapping genetic responses to xylose and mannitol. Accordingly, in the plant-associated *Pseudomonas fluorescens* SBW25 which is also able to grow on xylose as a sole carbon source, expression of the xylulokinase was shown to be mediated by the mannitol-responsive regulator *MtlR* by xylulose being the inducer and therefore both *MtlR* and *xylR* were shown to regulate xylose utilization⁵². The mechanism seems to be reasonable since as shown in the present study, there is the possibility of such novel mechanism in which two AraC-type TFs (*XylR* and *MtlR*) being involved in xylose utilization in *L. lactis* subsp. *lactis* IO-1 in terms of the mode of regulation. The fact that xylose and mannitol are mostly co-present in the plant environment can also help the idea as these will help the colonization of plant-associated bacteria⁵².

Although catabolite repression by glucose is a known feature for all xylose metabolizers, the genes of *xyl* operon was not clustered in CcpA regulon, even when tested the regulon prediction with lower motif similarity cut off. This is in line with a study by Erlandson *et al.* in which they could not find a *cis*-acting sites upstream of *xylA* or *xylB* and suggested that the strain might have a homologue of it⁴⁹. To further verify this, we have implemented AME, a motif enrichment tool from MEME suit⁵³ and analyzed the promoter of xylose operon to find any homologue CcpA motif with motif of *xyl* operon from *Lactobacillus pentosus*⁵⁴, *Staphylococcus xylosus*⁵⁵ and *Bacillus subtilis*⁵⁶. Our results also show no homologue CcpA-specific motif upstream of these genes.

Findings on nisin biosynthesis regulation at the transcriptional level. Another important feature of *L. lactis* subsp. *lactis* IO-1 is the production of nisin Z. Due to its widespread use mainly as an antimicrobial agent in the food industry, the research on its biology and biotechnology is of considerable interest. Unlike a large amount of studies on the effect of culture medium formulations, culture conditions and basic genetics study on nisin biosynthesis gene cluster, few reports exist on the molecular basis of nisin biosynthesis in relation with other components of the genome and its transcriptional regulation. Generally, *nisZBTCIPRKFEG* cluster encoding nisin Z peptide, modifying enzymes, transport, regulatory and immunity-related proteins which are organized in 4 operons (Fig. 2), is known to be regulated by a two-component system *nisRK* which triggers P¹ and P⁴ promoters by autoinduction.

There is no information on the regulatory potential of P² and P³ promoters and while the expression of NisZBTC and NisFEG are thought to be regulated by NisRK and since no specific regulator is encoded by the *nis* operon, other operons are believed to be under the control of a general regulator⁵⁷. The role of alternative σ factors has been studied for subtilin, a lantibiotic produced by *B. subtilis* which is very closely related to nisin with respect to structure, activity and genetic traits. Subtilin biosynthesis was reported to be controlled by an alternative σ factor (σ H)-mediated transcriptional regulation other than the known autoinduction regulatory mechanism³⁰. Accordingly, nisin promoters were studied to see if the mechanism is potentially available. However, no orthologs for σ H was found in *L. lactis* subsp. *lactis* IO-1. Furthermore, no hits were found when Biopropector⁵⁸ was used to search for any enriched motif known of the two sigma factors of *L. lactis* subsp. *lactis* IO-1, RpoD (or σ 39)⁵⁹ and ComX⁶⁰ within P² and P³ sequences. Although the presence of (alternative) sigma factors is generally proposed for *L. lactis* (Araya-Kojima, Ishibashi *et al.* 1995), further experimental analysis is required to verify the presence or absence of such regulation on nisin operons.

Generally, from the regulon prediction results, 6 regulons (filtered with motif similarity cut off of more than 0.5) have been illustrated within which one or more nisin operons were present (Table 3). According to Table 3, nisin operons are shown to have potential co-regulatory relationships with operons responsible for cell wall, aminoacid, vitamin, and DNA metabolism. Except for a few operons without yet-meaningful explanation of their

RegulonID	Members (operons)	Biological Function	Relevance to nisin [*] (reference)
Regulon_303	Operon_362	Nisin biosynthesis	—
	Operon_1254	Arginine metabolism	Less acidic environment as a result of ammonium production from arginine metabolism increases nisin immunity ⁶²
	Operon_1378	Arginine metabolism	Less acidic environment as a result of ammonium production from arginine metabolism increases nisin immunity ⁶²
Regulon_304	Operon_362	Nisin biosynthesis	—
	Operon_364	Nisin biosynthesis	—
	Operon_620	Unclassified (hypothetical protein)	NYK
	Operon_586	Riboflavin metabolism	NYK
Regulon_305	Operon_364	Nisin biosynthesis	—
	Operon_362	Nisin biosynthesis	—
	Operon_745	Lipoteichoic acid (LTA) biosynthesis	Cell wall biosynthesis has positive correlation with resistance to cationic antimicrobial peptides ⁶³
	Operon_608	Pyrimidine biosynthesis	Improves cell growth by being involved in DNA recombination and repair. Nisin production is growth-dependent.
Regulon_306	Operon_362	Nisin biosynthesis	—
	Operon_366	Nisin biosynthesis	—
	Operon_1254	Arginine metabolism	Less acidic environment as a result of ammonium production from arginine metabolism increases nisin immunity ⁶²
	Operon_32	Unclassified	NYK
Regulon_307	Operon_362	Nisin biosynthesis	—
	Operon_29	Cysteine and methionine metabolism	Many amino acids (like cysteine) highly stimulate nisin production by playing a precursor role ⁸⁹
	Operon_1292	Glycine, serine and threonine metabolism	Many amino acids (like serine and threonine) highly stimulate nisin production by playing a precursor role ⁸⁹
	Operon_1201	Unclassified (hypothetical protein)	NYK
Regulon_326	Operon_362	Nisin biosynthesis	—
	Operon_586	Riboflavin biosynthesis	NYK

Table 3. Predicted nisin regulons and their properties. *NYK: Not Yet Known.

relationship to nisin operon, the predicted operons with potential co-regulation relations with nisin have been directly or indirectly introduced by other studies to be related to nisin biosynthesis.

It has been long known that the level of nisin produced could be affected by the level of the producer's immunity to nisin and there have been studies showing the improvement of nisin production by the overexpression of nisin immunity genes (*nisI*, *nisFEG*) in the producer strain⁶¹. Using transcriptome analysis, Kramer *et al.* reported different potential mechanisms, other than the direct effect of the nisin immunity genes, by which *L. lactis* acquires nisin resistance⁶². Many of their discovered mechanisms such as the role of arginine metabolism genes (*arc* operon) and LTA biosynthesis genes (*dlt* operon) have been also predicted to be correlated with nisin in our study. Accordingly, genes of *dlt* operon encodes enzymes for LTA biosynthesis, thus they are a part of cell wall biogenesis which is a deterministic factor in resistance to cationic antibacterial peptides^{63,64} and can provide the producer strain better immunity and in turn a higher threshold for production.

arc genes, on the other hand, are involved in the breakdown of arginine via the arginine deiminase pathway⁶⁵ and their involvement in nisin resistance/biosynthesis could be due to the activity of enzymes in completely degrading arginine to ammonium which will result in a locally less acidic pH at the outer side of the cytoplasmic membrane and hinder its binding to its anchor on membrane⁶².

As another regulon prediction result, nisin and riboflavin operons has been shown to cluster in a regulon in multiple cases. Bacteria supply their versatile demands for riboflavin by encoding both biosynthetic enzymes and importer proteins. *L. lactis* subsp. *lactis* IO-1 possess both riboflavin biosynthesis and transport genes encoding in operons such as *ribGBAH* as well as other monocistronic genes like *ribC*, *ynaE* and *ypaA*. *ribGBAH* and *ribC* encodes biosynthetic enzyme to produce riboflavin, FMN and FAD from GTP and ribulose-5-phosphate while *ynaE* and *ypaA* code for transporter similar to riboflavin transporters known in *B. subtilis*⁶⁶. Generally, it is proposed that in situations when riboflavin is sufficiently available in the environment, bacteria prefer to uptake rather than biosynthesis the vitamin and therefore they shut down the biosynthetic genes, thus avoiding waste of metabolic energy⁶⁷. Here, regulon prediction results shows that *ribGBAH*, but not other *rib* genes have been in many cases clustered with *nis* operons in a regulon. Thus the TRN analysis suggests that riboflavin biosynthesis and nisin production may have regulatory correlation. There has been no report before on the probability of their correlation.

The most well-studied regulation mechanism on some of riboflavin biosynthesis genes is the FMN riboswitch whose alternative secondary structures can repress or induce the expression of downstream genes by sensing the flavins. Studies in different bacteria have later shown that the regulation is affected by role of different

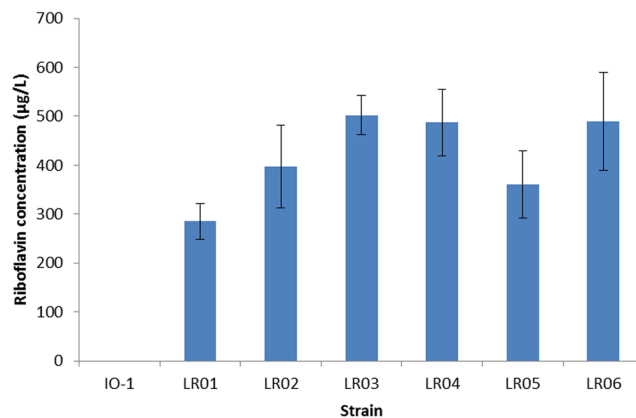


Figure 3. Riboflavin concentration produced by of roseoflavin-resistant strains (*L. lactis* subsp. *lactis* IO-1 is shown as control).

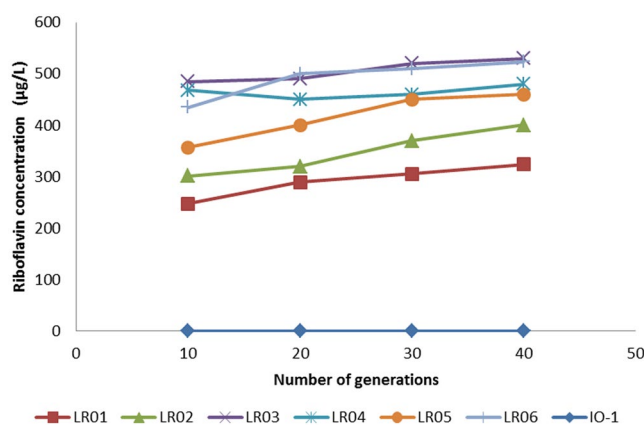


Figure 4. Stability of riboflavin production of the mutants after 40 generations (*L. lactis* subsp. *lactis* IO-1 is shown as control and concentration values are the average of three independent experiments).

proteins^{68–70}. These findings open the possibility for divergent regulatory mechanisms and suggest that there is still a long way to fully characterize riboflavin biosynthesis regulation mechanisms.

Moreover, there have been proteins such as RibR in *B. subtilis* found whose C-terminus shows regulatory roles on the FMN riboswitch, but no orthologs have been shown in any other bacteria (and also in IO-1). Therefore, search for similar regulatory proteins in other bacteria can be an interesting line of study. Although no statistically significant TF could be assigned to the similar motifs based on which *rib* and *nis* operons were clustered in a regulon, such regulatory proteins may elucidate the mechanism of this correlation.

On the other hand, there are a few reports about the role of transcription factors in regulating LAB's lantibiotics^{71,72}, one of which (LasX in *Lactobacillus sakei*) shows a bifunctional regulatory pattern by being both activator and repressor for different genes. Since there are many uncharacterized genes which are annotated as putative transcriptional regulators in the genome of IO-1 and other nisin-producing lactococci, similar mechanisms could also potentially describe the correlations such as what we have shown for *rib* and *nis* genes.

More relevant explanations on other co-clustered operons with nisin are summarized in Table 3.

Effect of riboflavin on nisin Z production by *L. lactis* IO-1. Fermentation-based experiments were implemented to study the effect of riboflavin on nisin Z biosynthesis. To this end, the effect of varying concentrations of riboflavin in the culture medium as well as riboflavin over producer *L. lactis* subsp. *lactis* IO-1 strains were studied on nisin production.

HPLC assay of the cell-free supernatant of 6 roseoflavin-resistant mutants in riboflavin-free CDM after 8 hours showed the isolates are riboflavin overproducers (Fig. 3), ranging from about 280 to 500 µg/L riboflavin. All 6 strains were shown to be stable in overproducing riboflavin (Fig. 4). These strains were used to analyze their nisin production. The fermentation experiments were performed in two conditions: a. Different riboflavin concentration (from 0–1 mg/L) in culture media and effect on nisin production; b. Overproducing riboflavin mutant strains and effects on their nisin production (Fig. 5 and Table S6).

As can be seen in Fig. 5, in the media with decreasing concentrations of riboflavin, no significant difference can be observed until the near-zero riboflavin concentrations (0.1 mg/L). Since in sufficient concentrations of

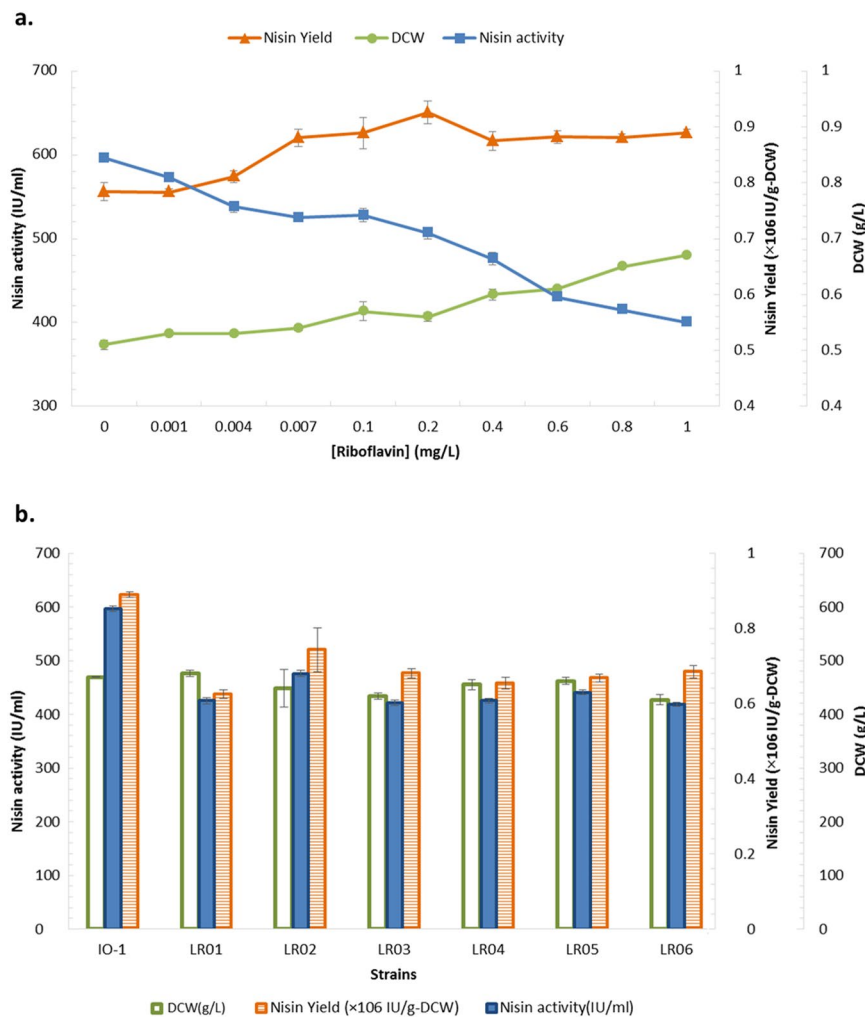


Figure 5. Effect of riboflavin on nisin Z production of *L. lactis* IO-1, studied in two conditions: (a) Different riboflavin concentrations in culture medium of *L. lactis* IO-1 and (b) assay using riboflavin overproducer mutant strains. (DCW: Dry cell weight).

riboflavin, uptake is preferred by bacteria, it is believed that the biosynthetic genes have low or no activity until near zero riboflavin concentrations where these enzymes will be derepressed⁶⁷. In this situation, there is a statistically significant decrease in nisin yield (p -value < 0.05) which can show the role of riboflavin biosynthesis on nisin production. The correlation seems to be negative as nisin biosynthesis decreased with increased riboflavin biosynthetic potential. To further study this negative correlation, stable riboflavin overproducers were isolated from roseflavin resistant mutants described in the methods and materials section.

All six riboflavin overproducers were shown to have lower nisin Z yield in comparison to the wild *L. lactis* subsp. *lactis* IO-1 under the same culture conditions. It is found that almost in all of the studied LABs, roseflavin mutants with the riboflavin overproduction phenotype contain mutations in the regulatory region upstream of the *ribGBAH* operon and it is proposed that these mutations will disturb terminator structure stability which in turn allows for continued transcription of the *rib* operon⁷³. Moreover, there have been also mutations found in *ribC* gene which play roles in converting riboflavin to FMN. Having continuously-active riboflavin biosynthetic genes, again nisin biosynthesis was confirmed with a negative correlation as seen with condition “a” (Fig. 5).

Conclusions

In the present work, a phylogenetic footprinting-based approach was implemented to fully elucidate all potential regulons in a nisin Z producing *Lactococcus lactis* with the aim of providing a systems-level insight able to aid in understanding the underlying regulatory mechanisms in this bacterium. According to a comparison with the existing knowledge on the transcription regulation in its close strains, the TRN was validated to be accurate and reasonable. For many of the predicted regulons, TFs could be assigned and multiple new targets were also introduced with meaningful correlations to their corresponding regulons. Even for regulons without assigned TFs such as nisin and riboflavin-including regulons, experimental tests showed the accuracy of the prediction.

As of important findings of inherent properties of *L. lactis* subsp. *lactis* IO-1 such as its nisin Z production potential, we have found for the first time the correlation of riboflavin and nisin biosynthesis which was shown

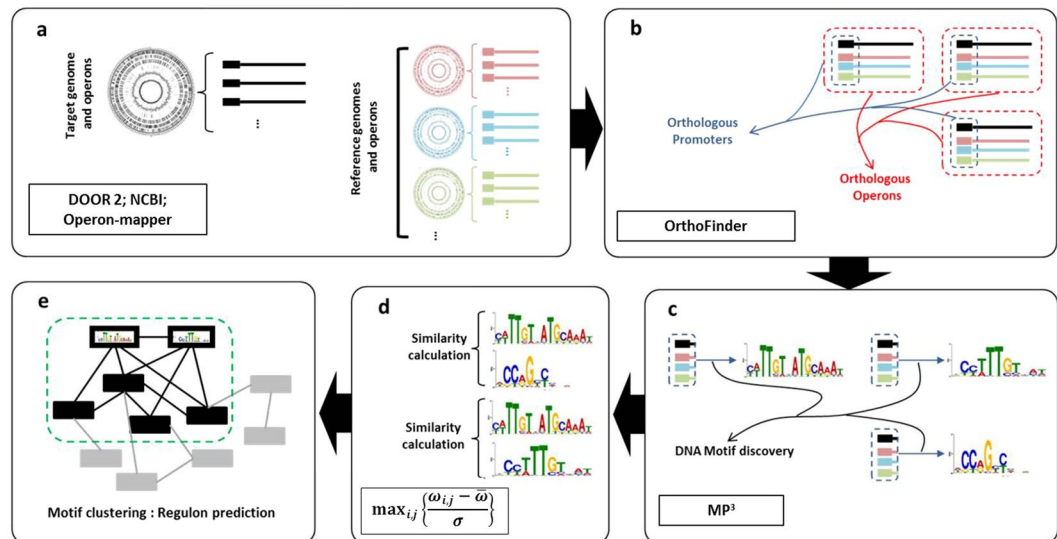


Figure 6. The comparative genomics-based regulon prediction framework used. **(a)** Reference genome selection and data retrieval, operons of the target and reference genome can be retrieved or predicted by proper tools. **(b)** Orthology identification for genes in the target genome. Relatively all of the genes in the target genome should have orthology among reference genomes to make phylogenetic footprinting feasible. Operons from target and reference genome sharing at least one orthologous gene is designated as an orthologous operon. The promoter sequence of these operons are used for motif finding **(c)**. Phylogenetic footprinting's final step is to find DNA cis-regulatory motifs **(d)**. Similarity score calculation (co-regulation score: see text) based on which interactions between motifs of operons were made **(e)**. Regulons prediction by clustering the resulting network, each cluster is a regulon.

to be negative in fermentation-based studies. The predicted regulatory link proposed by the TRN, proposed a more complex regulation on nisin or riboflavin biosynthesis. Accordingly, further evaluations can fully elucidate the exact mechanism behind this relationship which is of interest to both microbiologists and microbial biotechnologists.

In conclusion, the collection of regulons for the *L. lactis* subsp. *lactis* IO-1 obtained in this work not only can be used to better understand the unexplored physiology of this potent bacterium and rational implementation of the findings in biotechnology, but can be implemented as a template to reconstruct the TRN of other *L. lactis* strains with modifications based on comparative genomics.

Methods

TRN reconstruction. The reconstruction of a genome-based global TRN for *L. lactis* subsp. *lactis* IO-1 was performed using a four-step comparative genomics-based approach which is described stepwise below and is depicted in Fig. 6.

Target and reference genome set preparation. Our phylogenetic footprinting-based workflow for TRN reconstruction initiates with the exploitation of genomic information from closely-related strains to *L. lactis* subsp. *lactis* IO-1. Accordingly, genome and protein sequences, as well as annotation files of *L. lactis* IO-1 and other 74 selected Firmicutes, were downloaded from the NCBI RefSeq database at (<ftp://ftp.ncbi.nlm.nih.gov/genomes>) to be used as target and reference organisms, respectively.

Selection of the reference genome set directly affects the phylogenetic footprinting procedure primarily in terms of the quantity and quality of the resulting orthologous genes⁷⁴. Hence, four factors were considered for the selection of reference genomes: (i) selected genomes should be relatively close to *L. lactis* subsp. *lactis* IO-1 as phylogenetically-close organisms are more likely to share conserved regulatory mechanisms; (ii) Too closely-related organism should be excluded since analysis of upstream sequences will become less informative and returns many false positives due to the high sequence similarity; (iii) Although relatively close species are favorable, metabolic and physiological diversity should at the same time be maintained in the reference genome set which will omit any bias which can negatively affect the upstream sequences analysis; (iv) Although time-consuming, building the reference genome set should be iterative and genomes can be added or removed from the dataset to reach the improved final results (*i.e.* until reaching significant quantity and quality for the resulting orthologous genes to be used for the motif prediction). Therefore, it is assured that the species which are related enough to detect orthology, but divergent enough to maximize non-alignment noise are picked which will, in turn, reduce the false positive rate of motif discovery.

Orthologous promoters identification. Orthologous genes of *L. lactis* subsp. *lactis* IO-1 in the reference genomes were inferred using OrthoFinder2 which has been proved to achieve higher accuracy than all other orthology

inference methods⁷⁵. Since genes are organized in operons in prokaryotes, the orthologous relationships inferred between genes were then extended to operon level in a way that operons which share at least one orthologous gene are designated as orthologous operons.

Operons of the target and reference genomes were predicted by both DOOR2 and Operon-mapper tools and the overlapping predicted operons by the two tools were accepted as final result. For those operons being differently predicted by the two tools, the corresponding literature data from the very close strains to *L. lactis* subsp. *lactis* IO-1 were used. DOOR2 predicts operons using a back-end prediction algorithm which is based on the features of intergenic distance, neighborhood conservation, short DNA motifs as well as length ratio between gene pairs and shows a prediction accuracy of 95%^{76,77} and Operon-mapper, on the other hand, takes into account the functional relationships of their protein-coding products⁷⁸, thus, two different and complementary strategies were used for our operon prediction. Further, the 300-bp upstream sequences of all orthologous operons were extracted (orthologous promoters) to be used as the input for the motif discovery step.

Identification of phylogenetically conserved cis-regulatory motifs. Phylogenetic footprinting is a computational procedure of cis-regulatory motif identification in orthologous promoter regions of multiple genomes and is based on the principle that regulatory sequences in promoters are more conserved than the nonfunctional sequences since they evolve at a lower rate⁷⁹. Generally, during this process the set of orthologous promoters are aligned and this alignment is then scanned for conserved motif finding.

Motif discovery was carried out using the motif prediction by phylogenetic footprinting (MP³) method³¹. MP³ includes a promoter scoring and pruning method and it integrates the six well-known *de novo* motif finding algorithms (Biopropector⁵⁸, BOBRO⁸⁰, CONSENSUS⁸¹, CUBIC⁸², MDscan⁸³, MEME⁸⁴) as basic motif search engines and is evaluated as very accurate in predicting motifs.

The 300 nucleotides upstream each operon were then used as MP3 inputs with default parameters being used and the top five significant motifs (*p*-value < 0.05) were identified.

Regulon prediction. For the clustering of motifs to infer regulons the concept introduced in³¹ was used. According to them, a new scoring system, namely co-regulation score (CRS), is designed based on similarity relationship on operon level. For any pair of operons, A and B, in the target genome, CRS is defined as follows,

$$CRS(A, B) = \max_{i,j} \left\{ \frac{\omega_{i,j} - \bar{\omega}}{\sigma} \right\}$$

where $\bar{\omega}$ and σ represent the average and variance of similarity scores between any pair of motifs from the two sets of identified motifs in their regulatory promoters. Larger CRS denote the higher probability of co-regulatory relationship these two operons A and B.

With this score, regulons were further inferred via clustering through a heuristic graph model instead of conventionally calculating a pair-wise similarity score for predicted motifs and then identifying regulons through motif clusters¹².

According to authors, calculating CRS between any pair of operons, shows more accuracy for regulon prediction than other similar scores which are defined through various functional and evolutionary relationships, more details on the approach can be found in³¹.

Available data on regulons of the best-studied closely related strain to *L. lactis* subsp. *lactis* IO-1, namely *L. lactis* subsp. *lactis* IL1403 as well as scattered data on other *L. lactis* subsp. *lactis* strains were used to evaluate the regulon prediction. The concept is that for the most conserved regulation mechanisms such as amino acid biosynthesis, stress response as well as cell wall biosynthesis there should be conserved regulatory mechanisms among the members of same subspecies. Duly, RegPrecise database⁸⁵ which collects manually regulated regulons of prokaryotes, as well as literature data, were implemented. The orthologous *L. lactis* IO-1 genes with all the genes and operons comprising the regulons of other *L. lactis* subsp. *lactis* were inferred. Where significant number of regulons in other *L. lactis* subsp. *lactis* was present with similar members in the final predicted *L. lactis* subsp. *lactis* IO-1 TRN, their orthologous TF in *L. lactis* IO-1 TF repertoire was assigned to corresponding IO-1 regulons (Tables 1 and S4).

The final set of elucidated regulons was used to majorly analyze the special features of *L. lactis* subsp. *lactis* IO-1 such as nisin Z production potential and xylose metabolism.

Experimental procedure. **Strains and culture conditions.** *L. lactis* strain IO-1 (JCM7638), obtained from the Japan Collection of Microorganisms was grown on its specific chemically defined media (CDM)⁸⁶ supplemented with 1% glucose and incubation temperature of 37 °C in flask in all fermentation experiments as it is the optimum for of *L. lactis* subsp. *lactis* IO-1 without aeration and pH control. When appropriate riboflavin was omitted. *Micrococcus luteus* PTCC 1169 (ATCC 10240), the indicator strain used in nisin bioassay was obtained from Persian Type Culture Collection (PTCC) and grown in nutrient broth/agar medium and grown at 30 °C.

Isolation of riboflavin overproducing mutants and analysis of the phenotype's stability. Roseoflavin is a toxic analogue of riboflavin exposure to which has been reported to cause *L. lactis* strains with spontaneous resistance and this often coincides with a riboflavin-overproducing phenotype⁸⁷. Therefore, exposure to roseoflavin was used to isolate riboflavin overproducing mutants.

L. lactis subsp. *lactis* IO-1 was plated at mid exponential growth phase on CDM agar while riboflavin was omitted and supplemented with 100 mg/L roseoflavin to isolate spontaneous roseoflavin-resistant IO-1 mutants.

Stability of riboflavin overproduction phenotype was determined by subcloturing the selected mutant strains for 40 generations with periodically measuring extracellular riboflavin concentration.

Analysis of riboflavin in culture medium. In stationary phase, the cell free supernatant of the roseoflavin resistant isolates which were grown in CDM (with no riboflavin) was analyzed for riboflavin content. Accordingly, riboflavin concentrations were measured by reverse phase HPLC with fluorescent detection (excitation and emission wavelengths of 440 and 520 nm, respectively) and riboflavin was eluted with a linear gradient of acetonitrile from 3.6% to 30% at pH 3.2⁸⁵. Commercially obtained riboflavin was used for preparing a standard curve for quantitative purposes (Sigma-Aldrich, Germany).

Nisin bioassay. For the assay of nisin activity, the growth inhibition of the indicator strain *Micrococcus flavus* ATCC 10240 was determined by the agar well diffusion method described by Tramer *et al.*⁸⁸. Briefly, according to this method, Nutrient agar was inoculated with an appropriate amount of the indicator strain and after being solidified, holes of 7 mm were made to be filled with standard nisin solution and the pretreated broth being studied for its nisin bioactivity. Overnight incubated plates at 30 °C were measured in terms of diameters of the zones of inhibition. The measured diameter was then converted to unit IU/ml according to the standard curve. For making standard nisin solutions, nisin Z (TOKU-E, Osaka, Japan) was used.

Statistical analysis. The mean values obtained for nisin Z yields obtained throughout the fermentation analysis were compared using one-way ANOVA in SPSS software (version 19, IBM Co.). Probability (p) values < 0.05 were considered significant.

Data availability

All data generated or analysed during this study are included in this published article (and its Supplementary Information Files).

Received: 27 August 2019; Accepted: 13 January 2020;

Published online: 02 March 2020

References

- Balleza, E. *et al.* Regulation by transcription factors in bacteria: beyond description. *FEMS microbiology reviews* **33**, 133–151 (2008).
- Cases, I., de Lorenzo, V. & Ouzounis, C. A. Transcription regulation and environmental adaptation in bacteria. *Trends in microbiology* **11**, 248–253 (2003).
- Maas, W. K. & Clark, A. Studies on the mechanism of repression of arginine biosynthesis in *Escherichia coli*: II. Dominance of repressibility in diploids. *Journal of molecular biology* **8**, 365–370 (1964).
- Xiang, J. & Kim, S. Identifying functional targets from transcription factor binding data using SNP perturbation. *bioRxiv*, 412841 (2018).
- Chen, Y.-H., Yang, C.-D., Tseng, C.-P., Huang, H.-D. & Ho, S.-Y. GeNOSA: inferring and experimentally supporting quantitative gene regulatory networks in prokaryotes. *Bioinformatics* **31**, 2151–2158 (2015).
- Segal, E. *et al.* Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nature genetics* **34**, 166 (2003).
- Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PLoS one* **5**, e12776 (2010).
- De Smet, R. & Marchal, K. Advantages and limitations of current network inference methods. *Nature Reviews Microbiology* **8**, 717 (2010).
- Yeung, K. Y., Medvedovic, M. & Bumgarner, R. E. From co-expression to co-regulation: how many microarray experiments do we need? *Genome biology* **5**, R48 (2004).
- de Jong, A., Pietersma, H., Cordes, M., Kuipers, O. P. & Kok, J. PePPER: a webserver for prediction of prokaryote promoter elements and regulons. *BMC genomics* **13**, 299 (2012).
- Münch, R. *et al.* Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* **21**, 4187–4189 (2005).
- Liu, B. *et al.* Bacterial regulon modeling and prediction based on systematic cis regulatory motif analyses. *Scientific reports* **6**, 23030 (2016).
- Tagle, D. A. *et al.* Embryonic ϵ and γ globin genes of a prosimian primate (*Galago crassicaudatus*): Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. *Journal of molecular biology* **203**, 439–455 (1988).
- Blanchette, M. & Tompa, M. Discovery of regulatory elements by a computational method for phylogenetic footprinting. *Genome research* **12**, 739–748 (2002).
- Leroy, F. & De Vuyst, L. Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends in Food Science & Technology* **15**, 67–78 (2004).
- Ardö, Y. Flavour formation by amino acid catabolism. *Biotechnology advances* **24**, 238–242 (2006).
- Ramchandran, L. *et al.* Improving cell yield and lactic acid production of *Lactococcus lactis* ssp. *cremoris* by a novel submerged membrane fermentation process. *Journal of membrane science* **403**, 179–187 (2012).
- Yoo, J.-W., Irvine, D. J., Discher, D. E. & Mitragotri, S. Bio-inspired, bioengineered and biomimetic drug delivery carriers. *Nature reviews Drug discovery* **10**, 521 (2011).
- Chiabai, M. J. *et al.* Mucosal delivery of *Lactococcus lactis* carrying an anti-TNF scFv expression vector ameliorates experimental colitis in mice. *BMC biotechnology* **19**, 38 (2019).
- De Vuyst, L. & Vandamme, E. J. In *Bacteriocins of lactic acid bacteria* 151–221 (Springer, 1994).
- Millette, M. *et al.* Capacity of human nisin- and pediocin-producing lactic acid bacteria to reduce intestinal colonization by vancomycin-resistant enterococci. *Appl. Environ. Microbiol.* **74**, 1997–2003 (2008).
- Wu, J., Hu, S. & Cao, L. Therapeutic effect of nisin Z on subclinical mastitis in lactating cows. *Antimicrobial agents and chemotherapy* **51**, 3131–3135 (2007).
- Kamarajan, P. *et al.* Nisin ZP, a bacteriocin and food preservative, inhibits head and neck cancer tumorigenesis and prolongs survival. *PLoS one* **10**, e0131008 (2015).
- Norouzi, Z., Salimi, A., Halabian, R. & Fahimi, H. Nisin, a potent bacteriocin and anti-bacterial peptide, attenuates expression of metastatic genes in colorectal cancer cell lines. *Microbial pathogenesis* **123**, 183–189 (2018).
- Shin, J. M. *et al.* Biomedical applications of nisin. *Journal of applied microbiology* **120**, 1449–1465 (2016).

26. Thomas, V. M., Brown, R. M., Ashcraft, D. S. & Pankey, G. A. Synergistic effect between nisin and polymyxin B against pandrug-resistant and extensively drug-resistant *Acinetobacter baumannii*. *International journal of antimicrobial agents* **53**, 663–668 (2019).
27. Chikindas, M. L., Weeks, R., Drider, D., Chistyakov, V. A. & Dicks, L. M. Functions and emerging applications of bacteriocins. *Current opinion in biotechnology* **49**, 23–28 (2018).
28. Shimizu-Kadota, M. *et al.* Genomic features of *Lactococcus lactis* IO-1, a lactic acid bacterium that utilizes xylose and produces high levels of L-lactic acid. *Bioscience, biotechnology, and biochemistry*, 130080 (2013).
29. Matsusaki, H., Endo, N., Sonomoto, K. & Ishizaki, A. Lantibiotic nisin Z fermentative production by *Lactococcus lactis* IO-1: relationship between production of the lantibiotic and lactate and cell growth. *Applied Microbiology and Biotechnology* **45**, 36–40 (1996).
30. Kleerebezem, M. Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis. *Peptides* **25**, 1405–1414 (2004).
31. Liu, B. *et al.* An integrative and applicable phylogenetic footprinting framework for cis-regulatory motifs identification in prokaryotic genomes. *BMC genomics* **17**, 578 (2016).
32. Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L. & Noble, W. S. Quantifying similarity between motifs. *Genome biology* **8**, R24 (2007).
33. Ercan, O., Wels, M., Smid, E. J. & Kleerebezem, M. Genome-wide transcriptional responses to carbon starvation in nongrowing *Lactococcus lactis*. *Appl. Environ. Microbiol.* **81**, 2554–2561 (2015).
34. Wu, H. *et al.* Contribution of YthA, a PspC family transcriptional regulator of *Lactococcus lactis* F44 acid tolerance and nisin yield: a transcriptomic approach. *Appl. Environ. Microbiol.* **84**, e02483–02417 (2018).
35. Leyn, S. A. *et al.* Genomic reconstruction of the transcriptional regulatory network in *Bacillus subtilis*. *Journal of bacteriology* **195**, 2463–2473 (2013).
36. Romeo, Y. *et al.* Osmoregulation in *Lactococcus lactis*: BusR, a transcriptional repressor of the glycine betaine uptake system BusA. *Molecular microbiology* **47**, 1135–1147 (2003).
37. Molenaar, D., Hagting, A., Alkema, H., Driessen, A. & Konings, W. N. Characteristics and osmoregulatory roles of uptake systems for proline and glycine betaine in *Lactococcus lactis*. *Journal of Bacteriology* **175**, 5438–5444 (1993).
38. Martin, P. *et al.* Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme. *Structure* **9**, 1071–1081 (2001).
39. Garg, N. & Singla, P. Arsenic toxicity in crop plants: physiological effects and tolerance mechanisms. *Environmental Chemistry Letters* **9**, 303–321 (2011).
40. Hoffmann, T. *et al.* Arsenobetaine: an ecophysiological important organoarsenical confers cytoprotection against osmotic stress and growth temperature extremes. *Environmental microbiology* **20**, 305–323 (2018).
41. Popowich, A., Zhang, Q. & Le, X. C. Arsenobetaine: the ongoing mystery. *National Science Review* **3**, 451–458 (2016).
42. Cheung, K. J., Badarinarayana, V., Selinger, D. W., Janse, D. & Church, G. M. A microarray-based antibiotic screen identifies a regulatory role for supercoiling in the osmotic stress response of *Escherichia coli*. *Genome research* **13**, 206–215 (2003).
43. Muthuramalingam, P. *et al.* Global analysis of threonine metabolism genes unravel key players in rice to improve the abiotic stress tolerance. *Scientific reports* **8**, 9270 (2018).
44. Brinsmade, S. R. *et al.* Hierarchical expression of genes controlled by the *Bacillus subtilis* global regulatory protein CodY. *Proceedings of the National Academy of Sciences* **111**, 8227–8232 (2014).
45. Inukai, S., Kock, K. H. & Bulyk, M. L. Transcription factor–DNA binding: beyond binding site motifs. *Current opinion in genetics & development* **43**, 110–119 (2017).
46. Siggers, T. & Gordana, R. Protein–DNA binding: complexities and multi-protein codes. *Nucleic acids research* **42**, 2099–2111 (2013).
47. Zhu, H., Wang, G. & Qian, J. Transcription factors as readers and effectors of DNA methylation. *Nature Reviews Genetics* **17**, 551 (2016).
48. Guédon, E., Jamet, E. & Renault, P. In *Lactic Acid Bacteria: Genetics, Metabolism and Applications* 93–112 (Springer, 2002).
49. Erlanson, K. A. *et al.* Dissolution of Xylose Metabolism in *Lactococcus lactis*. *Appl. Environ. Microbiol.* **66**, 3974–3980 (2000).
50. Ishizaki, A. *et al.* Biochemical characterization of *Lactococcus lactis* IO-1 whose optimal temperature is as high as 37°C. *The Journal of General and Applied Microbiology* **36**, 1–6 (1990).
51. Tanaka, K. *et al.* Two different pathways for D-xylose metabolism and the effect of xylose concentration on the yield coefficient of L-lactate in mixed-acid fermentation by the lactic acid bacterium *Lactococcus lactis* IO-1. *Applied Microbiology and Biotechnology* **60**, 160–167 (2002).
52. Liu, Y., Rainey, P. B. & Zhang, X. X. Molecular mechanisms of xylose utilization by *Pseudomonas fluorescens*: overlapping genetic responses to xylose, xylulose, ribose and mannitol. *Molecular microbiology* **98**, 553–570 (2015).
53. McLeay, R. C. & Bailey, T. L. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC bioinformatics* **11**, 165 (2010).
54. Lokman, B. C. *et al.* Regulation of expression of the *Lactobacillus pentosus* xylAB operon. *Journal of bacteriology* **179**, 5391–5397 (1997).
55. Sizemore, C., Wieland, B., Götz, F. & Hillen, W. Regulation of *Staphylococcus xylosus* xylose utilization genes at the molecular level. *Journal of bacteriology* **174**, 3042–3048 (1992).
56. Jacob, S., Allmansberger, R., Gärtner, D. & Hillen, W. Catabolite repression of the operon for xylose utilization from *Bacillus subtilis* W23 is mediated at the level of transcription and depends on a cis site in the xylA reading frame. *Molecular and General Genetics* **229**, 189–196 (1991).
57. Lubelski, J., Rink, R., Khusainov, R., Moll, G. & Kuipers, O. Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin. *Cellular and molecular life sciences* **65**, 455–476 (2008).
58. Liu, X., Brutlag, D. L. & Liu, J. S. In *Biocomputing 2001* 127–138 (World Scientific, 2000).
59. Araya-Kojima, T., Ishibashi, N., Shimamura, S., Tanaka, K. & Takahashi, H. Identification and molecular analysis of *Lactococcus lactis* rpoD operon. *Bioscience, biotechnology, and biochemistry* **59**, 73–77 (1995).
60. Mulder, J., Wels, M., Kuipers, O. P., Kleerebezem, M. & Bron, P. A. Unleashing natural competence in *Lactococcus lactis* by induction of the competence regulator ComX. *Appl. Environ. Microbiol.* **83**, e01320–01317 (2017).
61. Kim, W., Hall, R. & Dunn, N. Improving nisin production by increasing nisin immunity/resistance genes in the producer organism *Lactococcus lactis*. *Applied microbiology and biotechnology* **50**, 429–433 (1998).
62. Kramer, N. E., van Hijum, S. A., Knol, J., Kok, J. & Kuipers, O. P. Transcriptome analysis reveals mechanisms by which *Lactococcus lactis* acquires nisin resistance. *Antimicrobial Agents and Chemotherapy* **50**, 1753–1761 (2006).
63. Fisher, N. *et al.* The dltABCD operon of *Bacillus anthracis* Sterne is required for virulence and resistance to peptide, enzymatic, and cellular mediators of innate immunity. *Journal of bacteriology* **188**, 1301–1309 (2006).
64. Peschel, A. *et al.* Inactivation of the dlt Operon in *Staphylococcus aureus* Confers Sensitivity to Defensins, Protegrins, and Other Antimicrobial Peptides. *Journal of Biological Chemistry* **274**, 8405–8410 (1999).
65. Poolman, B., Driessen, A. & Konings, W. Regulation of arginine-ornithine exchange and the arginine deiminase pathway in *Streptococcus lactis*. *Journal of Bacteriology* **169**, 5597–5604 (1987).
66. Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic acids research* **30**, 3141–3151 (2002).
67. García-Angulo, V. A. Overlapping riboflavin supply pathways in bacteria. *Critical reviews in microbiology* **43**, 196–209 (2017).

68. Pedrolli, D. B. *et al.* A dual control mechanism synchronizes riboflavin and sulphur metabolism in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences* **112**, 14054–14059 (2015).
69. Higashitsuji, Y., Angerer, A., Berghaus, S., Hobl, B. & Mack, M. RibR, a possible regulator of the *Bacillus subtilis* riboflavin biosynthetic operon, *in vivo* interacts with the 5'-untranslated leader of rib mRNA. *FEMS microbiology letters* **274**, 48–54 (2007).
70. Sepúlveda Cisternas, I., Salazar, J. C. & García-Angulo, V. A. Overview on the bacterial iron-riboflavin metabolic axis. *Frontiers in microbiology* **9**, 1478 (2018).
71. Iwatani, S. *et al.* LntR, a TetR-family transcriptional regulator, positively regulates lactacin Q production in *Lactococcus lactis* QU 5. *FEMS microbiology letters* **363**, fnw200 (2016).
72. Rawlinson, E. L. A., Nes, I. F. & Skaugen, M. LasX, a transcriptional regulator of the lactocin S biosynthetic genes in *Lactobacillus sakei* L45, acts both as an activator and a repressor. *Biochimie* **84**, 559–567 (2002).
73. Burgess, C. M., Smid, E. J., Rutten, G. & Van Sinderen, D. A general method for selection of riboflavin-overproducing food grade micro-organisms. *Microbial cell factories* **5**, 24 (2006).
74. McCue, L. A., Thompson, W., Carmack, C. S. & Lawrence, C. E. Factors influencing the identification of transcription factor binding sites by cross-species comparison. *Genome research* **12**, 1523–1532 (2002).
75. Emms, D. M. & Kelly, S. OrthoFinder2: fast and accurate phylogenomic orthology analysis from gene sequences. *BioRxiv*, 466201 (2018).
76. Dam, P., Olman, V., Harris, K., Su, Z. & Xu, Y. Operon prediction using both genome-specific and general genomic information. *Nucleic acids research* **35**, 288–298 (2006).
77. Cao, H., Ma, Q., Chen, X. & Xu, Y. DOOR: a prokaryotic operon database for genome analyses and functional inference. *Briefings in bioinformatics* (2017).
78. Taboada, B., Estrada, K., Ciria, R. & Merino, E. Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. *Bioinformatics* **34**, 4118–4120 (2018).
79. Katara, P., Grover, A. & Sharma, V. Phylogenetic footprinting: a boost for microbial regulatory genomics. *Protoplasma* **249**, 901–907 (2012).
80. Li, G., Liu, B., Ma, Q. & Xu, Y. A new framework for identifying cis-regulatory motifs in prokaryotes. *Nucleic acids research* **39**, e42–e42 (2010).
81. Hertz, G. Z. & Stormo, G. D. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics (Oxford, England)* **15**, 563–577 (1999).
82. Olman, V., Xu, D. & Xu, Y. CUBIC: identification of regulatory binding sites through data clustering. *Journal of Bioinformatics and Computational Biology* **1**, 21–40 (2003).
83. Liu, X. S., Brutlag, D. L. & Liu, J. S. An algorithm for finding protein–DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. *Nature biotechnology* **20**, 835 (2002).
84. Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic acids research* **37**, W202–W208 (2009).
85. Novichkov, P. S. *et al.* RegPrecise 3.0—a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC genomics* **14**, 745 (2013).
86. Machii, M. *et al.* Chemically defined media and auxotrophy of the prolific l-lactic acid producer *Lactococcus lactis* IO-1. *Journal of bioscience and bioengineering* **115**, 481–484 (2013).
87. Burgess, C., O'Connell-Motherway, M., Sybesma, W., Hugenholtz, J. & Van Sinderen, D. Riboflavin production in *Lactococcus lactis*: potential for *in situ* production of vitamin-enriched foods. *Appl. Environ. Microbiol.* **70**, 5769–5777 (2004).
88. Tramer, J. & Fowler, G. Estimation of nisin in foods. *Journal of the Science of Food and Agriculture* **15**, 522–528 (1964).
89. De Vuyst, L. Nutritional factors affecting nisin production by *Lactococcus lactis* subsp. *lactis* NIZO 22186 in a synthetic medium. *Journal of Applied Bacteriology* **78**, 28–33 (1995).

Author contributions

J.H., A.M.N. and N.P. designed the study. N.P. conducted the research and wrote the draft manuscript. J.H. and A.M.N. supervised the whole research procedure and edited the manuscript text. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-59731-8>.

Correspondence and requests for materials should be addressed to J.H. or A.M.-N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020