ELUCIDATION OF THE BIOSYNTHETIC PATHWAY OF B-GROUP VITAMINS VIA GENOME MINING OF FOOD-DERIVED *Bacillus velezensis* VTX20

Quach Ngoc Tung[1,2], Vu Thi Hanh Nguyen[1,2], Le Thi Thanh Xuan[1], Chu Hoang Ha[1,2], Phi Quyet Tien[1,2,*]

[1]Institute of Biotechnology, VAST, Vietnam
[2]Graduate University of Science and Technology, VAST, Vietnam

ABSTRACT

B-vitamins are micronutrients that play an important role in various cellular processes of organisms, which are only synthesized by plants, yeasts, and bacteria. Since animals and humans lack the ability to synthesize B-vitamins, supplements of vitamins from dietary and the B-vitamin producing bacteria are required. In this study, we, for the first time, shed some light on biosynthetic pathways involved in folate (vitamin B9), riboflavin (vitamin B2), and biotin (vitamin B7) production in *Bacillus velezensis* VTX20. The genome-wide comparison revealed that *B. velezensis* VTX20 shared high similarities with *B. tequilensis* KCTC 13622, *B. subtilis* 168, *B. amyloliquefaciens* DSM 7. Genomic analysis revealed the presence of a complete folate biosynthesis pathway in which some core components were not found in most *Bacillus* species. Moreover, strain VTX20 also had the metabolic pathways for riboflavin and biotin that are important probiotic traits. These results highlighted that *B. velezensis* VTX20 is a producer of B-vitamins, which can be applied further in the agricultural biotechnology industry.

Keywords: *Bacillus velezensis,* biotin, genomic analysis, folate, probiotic traits, riboflavin.

INTRODUCTION

Bacillus velezensis is a Gram-positive plant growth-promoting rhizobacterial (PGPR) strain that has been widely applied in agricultural sectors. The distinctive properties of *B. velezensis* have been reported to be the producer of phytohormones and secondary metabolites, pathogen antagonistic activity, and efficient colonization on plants (Cai et al., 2017; Kim et al., 2017). *B. velezensis,* originated from the rhizosphere, is beneficial to tomato, cucumber, cotton, tobacco, and lettuce through indole-3-acetic acid (IAA) auxin and other volatile organic compound production. In addition, this species has the ability to produce various secondary metabolites such as bacillopeptins, plipastatins, surfactins and lichenysins, iturins, and fengycins, which inhibit soilborne pathogens such as *Fusarium fujikuroi*, *Rhizoctonia solani*, and *Xanthmonas oryzae*.

Recently, *B. velezensis* has received greater attention in the probiotic industry. The use of *B. velezensis* as a feed additive for laying hens improved significantly egg production rate, egg quality and replaced flavomycin that is used frequently for livestock production (Ye et al., 2020). In 2020, the Calsporin® comprising viable spores of *B. velezensis* DSM 15544 was accepted for the first time by The European Union Food Safety Authority (EFSA) and Generally Recognized as a Safe (GRAS) status by the US Food and Drug Administration (Bampidis et al., 2020). Using the whole-genome technology, we proved that *B. velezensis* VTX9 isolated from swine faeces was a potential probiotic candidate in pig farming (Vu et al., 2021). Moreover, *B. velezensis* VTX20 derived from traditional fermented soybean pastes could produce a great amount of levan-type exopolysaccharides that is a natural alternative to a commercial antioxidant in pharmaceutical, food, and biomedical industries (Vu et al., 2021). Given that one of the distinctive characteristics of a probiotic candidate is the production of B-vitamins, the ability to produce the vitamin B group in *B. velezensis* remains unknown. In this study, we elucidated for the first time biosynthetic pathways for the entire vitamin B group, folate, riboflavin, and biotin through genomic analysis.

MATERIAL AND METHODS

Comparative genome analysis

The genome of *B. velezensis* VTX20 isolated from derived from traditional fermented soybean pastes was sequenced by using a PacBio platform as described previously (Vu et al., 2021). A comparative genome analysis between *B. velezensis* VTX20 (CP075054) and three *Bacillus* species including *B. tequilensis* KCTC 13622 (AYTO01000000), *B. subtilis* 168 (AL009126), *B. amyloliquefaciens* DSM 7 (FN597644) was carried out using the OrthoVenn web server with default parameters (E-value 1e-5 and inflation value 1.5) (Wang et al., 2015).

Identification of folate, riboflavin, and biotin biosynthetic pathways

The complete genome of strain VTX20 was annotated using Rapid Annotations via Subsystem Technology (RAST) and Pathosystems Resource Integration Center (PATRIC) platforms (Overbeek et al., 2014; Wattam et al., 2017). A cluster of Orthologous Groups (COG) was annotated based on EggNOG-mapper v1.0.3 (Huerta-Cepas et al., 2017). The genes involved in folate, riboflavin, and biotin biosynthesis were determined by using KEGG Automatic Annotation Server (KAAS) performed by an assignment method of 'bi-directional best hit' against data of any species on the KEGG database to identify functional protein sequences following KEGG Orthology (Moriya et al., 2007). All desirable open reading frames (ORFs) were blasted against the NCBI curated refseq protein database (reference proteins) by Blastp and RAST.

RESULTS AND DISCUSSION

Genome-wide comparison

B. velezensis VTX20 (3.89 Mb) had the smallest sequence size among compared

Bacillus genomes including *B. tequilensis* KCTC 13622 (3.98 Mb), *B. subtilis* 168 (4.21 Mb), *B. amyloliquefaciens* DSM 7 (3.98 Mb) (Table 1). *B. velezensis* VTX20 contained 3836 COGs, 3471 clusters, and 320 singletons, which was comparable to *B. amyloliquefaciens* DSM 7. The venn diagram denoted that a total of 2908 COGs were commonly shared by four *Bacillus* strains, accounting for more than 60% identical to each other (Fig. 1). The result proved that these *Bacillus* strains had close evolutionary relationships. Among the unique COGs observed in all genomes, *B. velezensis* VTX20 had the lowest number of 5, which could be due to its lowest genome size.

Table 1. Genomic features of the *B. velezensis* VTX20 and representative members of the genus *Bacillus*

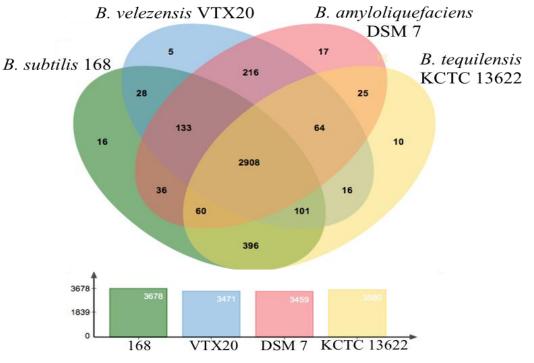| Features	*B. velezensis* VTX20	*B. tequilensis* KCTC 13622	*B. subtilis* 168	*B. amyloliquefaciens* DSM 7
Genome size (Mb)	3.89	3.98	4.21	3.98
G + C content (%)	46.1	43.9	43.5	46.1
Genes (total)	3834	4167	4537	4110
CDSs (coding)	3670	3958	4106	3858
rRNAs	27	7	10	10
tRNAs	86	74	86	94

Figure 1. Comparative genome between *B. velezensis* VTX20 and other *Bacillus* species. Venn diagram represents the numbers of unique and shared orthologous genes of each strain

Folate biosynthesis metabolic pathway

The predicted *de novo* folate synthesis pathway of strain VTX20 comprises a para-aminobenzoic acid (pABA) branch and a pterin branch (Fig. 2). In the first branch, starting with phosphoenolpyruvate and erythrose 4-phosphate, chorismite is produced through the action of 7 enzymes (*aroA, aroC,*

aroD, aroE1, aroE2, aroK, aroA1) in the shikimate pathway. Then, the presence of a *pabABC* operon including two copies *pabAB* (*orf_3894, orf_3895*) and one *pabC* (*orf_3893*) lead to the formation of pABA biosynthesis, a well-known precursor of folate. The chorismite incomplete pathway were reported in *Bacillus* probiotics, except

for *B. subtilis*, *B. pumilus,* and *B. toyonensis* (Khatri et al., 2019), As for the second branch, 9 enzymes supporting for the guanosine triphosphate (GTP) pathway are found to include *folE (orf_1800), phoAD (orf_3069, orf_3727), folB (orf_3891), sulD (orf_3890), folP (orf_3892), folC (orf_1367), folA (orf_717, orf_1903)* (Fig. 2).
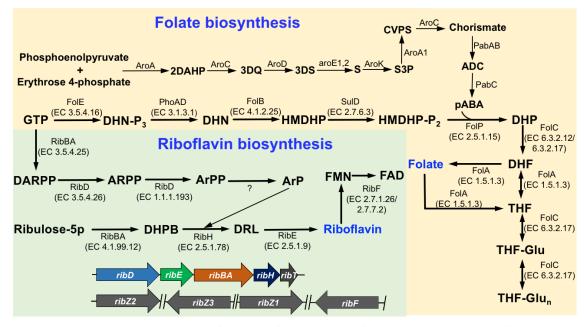

Figure 2. Schematic overview of *de novo* folate and riboflavin biosynthesis identified in *B. velezensis* VTX20. Abbreviations: GTP, guanosine triphosphate; DHN, 7,8-dihydroneopterin; HMDHP, 6-hydroxymethyl-7,8-dihydropterin; DHP, 7,8-dihydropteroate; DHF, 7,8-dihydrofolate; THF, 5,6,7,8-tetrahydrofolate; THF-Glu, tetrahydrofolate polyglutamate; -P, phosphate; -P2, pyrophosphate; -P3, triphosphate; pABA, para-aminobenzoic acid; ADC, aminodeoxychorismate; CVPS, 5-0-(1-carboxyvinyl)-shikimate 3-phosphate; S3P, shikimate 3-phosphate; S, shikimate; 3DS, 3-dehydroshikimate; 3DQ, 3-dehydroquinate; 2DAHP, 2-dehydro-3-deoxy-D-arobino-heptonate 7-phosphate; DHPB, L-3,4-dihydroxybutan-2-one 4-phosphate; DARPP, 2,5-siamino-6-(1-D-ribosylamino)pyrimidin-4(3H)-one 5′-phosphate; AFAPP, 2-amino-5-formylamino-6-(5-phospho-D-ribosylamino)pyrimidin-4(3H)-one; DARPP, 2,5-diamino-6-(1-D-ribitylamino)pyrimidin-4(3H)-one 5′-phosphate; ARPP, 5-amino-6-(ribosylamino)-2,4-(1H,3H)-pyrimidinedione 5′-phosphate; ArPP, 5-amino-2,6-dioxy-4-(5′-phospho-D-ribitylamino)pyrimidine; ArP, 5-amino-6-(1-D-ribitylamino)uracil; DrL, 6,7-dimethyl-8-(D-ribityl)lumazine; FMN, flavin mononucleotide

Folate is widely distributed in food such as fruits, vegetables, and nuts, which is in contrast to animals or humans that are unable to synthesize folate and assimilate its derivatives with a diet exploiting active transport systems (Rossi et al., 2011). The

folate biosynthetic pathway has been extensively exploited in *Lactobacillus* and *Bifidobacterium* probiotics but not much in the genus *Bacillus*, except for *B. subtilis* (Zhu et al., 2005; Rossi et al., 2011). Strains isolated from fermented foods such as

fermented soybean pastes are supposed to be folate-producing probiotics. In *Bacillus*, all genes involved in chorismate production are present in only *B. subtilis*, *B. pumilus* and *B. toyonensis*, while the other *Bacillus* strains lack some genes such as *aroD, aroK* or *aroA* (Khatri et al., 2019). Thus, *B. subtilis* is used as a host for the bioproduction of folate (Zhu et al., 2005). By contrast, the genome-wide comparison in the genus *Lactobacillus* indicated that lactobacilli are deficient in synthesizing a precursor pABA *de novo* due to the lack of genes responsible for chorismate conversion (Panda et al., 2018). In this study, all the pathway components are present and intact in *B. velezensis* VTX20, suggesting that this strain can be a novel platform for the biosynthesis of folate.

De novo production of riboflavin

Besides folate biosynthesis, *B. velezensis* VTX20 also contained the riboflavin biosynthesis pathway. The riboflavin cluster consisted of *rib* operon (*ribD, ribE, ribBA, ribH, ribT*) which is responsible for riboflavin biosynthesis from two major precursors: ribulose 5-phosphate and guanosine triphosphate (Fig. 2). Although *ribT* (*orf_1707*) was annotated as putative N-acetyltransferase, the function of *ribT* in the riboflavin biosynthetic pathway remains unknown. RibBA (*orf_1705*) catalyzes the conversion of GTP into DARPP that is subsequently exposed to corresponding enzyme RibD (*orf_1703*) to yield ArPP. ArPP is then dephosphorylated by unknown enzymes that have not been identified in plants and microorganisms. Moreover, the alternative pentose phosphate pathway takes place through the conversion of ribulose 5-phosphate into DHBP by enzyme RibBA (*orf_1705*). Thereafter, DRL is generated by the condensation of ArP and DHBP under the enzymatic activation of RibH (*orf_1706*). In the last step, DRL is dismutated by RibE (*orf_1704*) to produce riboflavin that is then converted into FMN and FAD by bifunctional riboflavin kinase/FMN adenylyltransferase RibF (*orf_2167*) (Fig. 2). FMN and FAD function as important cofactors which are required for the redox reaction of bacteria (Averianova et al., 2020). In addition, riboflavin also can be uptaken by 3 copies of transporters *ribZ*.

Despite the fact that the genus *Bacillus* contains 266 named species to date, the riboflavin biosynthetic pathway has only been exploited in *B. subtilis* and *B. tequilensis*, and *B. velezensis*. In the latest study, we proved that the probiotic *B. velezensis* VTX9 isolated from swine feces comprises the complete *rib* operon leading to the yield of 769 ± 7.5 ng/ml riboflavin (Vu et al., 2021). This present study is the second report confirming riboflavin biosynthesis in *B. velezensis*, indicating that the pathway is conserved across *B. velezensis* strains.

Identification of genes involved in biotin metabolisms

Biotin or vitamin B7 plays an important role in various enzymatic and carboxylation reactions of animals and humans. The genomic analysis revealed that biotin is synthesized by *bioWAFDBI* operon in *B. velezensis* VTX20 (Fig. 3). In this bacterium, pimelate is converted to pimeloyl-CoA through the activation of gene *bioW (orf_2000)* and *bioI (orf_2005)* encoding for 6-carboxyhexanoate--CoA ligase and cytochrome P450, respectively. Pimeloyl-CoA is preferably utilized by BioF *(orf_2002)* yielding 8-amino-7-oxononanoate, which is subsequently converted to 7,8-diaminononanote by BioA *(orf_2001)*. The resulting 7,8-diaminononanote is conjugated to CO_2 to generate dethiobiotin by ATP-dependent dethiobiotin synthetase BioD *(orf_2003)*. In the last step, the dethiobiotin is then used as a substrate for biotin synthase BioB *(orf_2004)* employing a SAM molecule in order to produce biotin. Interestingly, *bioWAFDBI* operon is conserved across *Bacillus* species (Fig. 3). By contrast, the biotin synthetic genes determined in *Escherichia coli* are located on two operons in the chromosome. The *bioH* is separated from the *bioABFCD* operon.
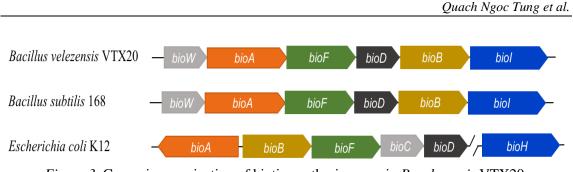
Figure 3. Genomic organization of biotin synthesis genes in *B. velezensis* VTX20,
B. subtilis 168 and *E. coli* K12

CONCLUSION

This study sheds additional light on the probiotic and biotechnology potential of *B. velezensis* VTX20 isolated from the traditional fermented soybean pastes. The most important finding here is the identification of the entire biosynthetic pathways for vitamin B groups including folate, riboflavin, and biotin through the genomic analysis. For the first time, we could show that *B. velezensis* VTX20 contains the complete *de novo* folate biosynthetic pathway, which is different from other *Bacillus* species that lack some essential genes. Moreover, the biotin cluster was found to be conserved across *B. subtilis* and *B. velezensis*. The genome analysis provided a valuable source of information for further experimental characterization of our findings and extends our knowledge of agricultural biotechnology applications made up of strain *B. velezensis* VTX20.

REFERENCES

Averianova L. A., Balabanova L. A., Son O. M., Podvolotskaya A. B., Tekutyeva L. A., 2020. Production of vitamin B2 (Riboflavin) by microorganisms: An overview. *Front. Bioeng. Biotechnol.,* 8(1172): 102−109. https://doi.org/ 10.3389/fbioe.2020.570828

Bampidis V., Azimonti G., Bastos M. d. L., Christensen H., Dusemund B., Fašmon Durjava M., Kouba M., López-Alonso M., López Puente S., Marcon F., Mayo B., Pechová A., Petkova M., Ramos F., Sanz Y., Villa R. E., Woutersen R., Cocconcelli P. S., Glandorf B., Herman L., Maradona M. P., Saarela M., Anguita M., Galobart J., Holczknecht O., Manini P., Pizzo F., Tarrés-Call J., Pettenati E., 2020. Assessment of the application for renewal of the authorisation of Calsporin® (*Bacillus velezensis* DSM 15544) as a feed additive for weaned piglets. *EFSA J.,* 18(11): e06283. https://doi.org/10.2903/ j.efsa.2020.6283

Cai X. C., Liu C. H., Wang B. T., Xue Y. R., 2017. Genomic and metabolic traits endow *Bacillus velezensis* CC09 with a potential biocontrol agent in control of wheat powdery mildew disease. *Microbiol. Res.,* 196: 89−94. https://doi.org/10.1016/j.micres.2016.12.0 07

Huerta-Cepas J., Forslund K., Coelho L. P., Szklarczyk D., Jensen L. J., von Mering C., Bork P., 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol. Biol. Evol.,* 34(8): 2115−2122. https://doi.org/10.1093/molbev/msx148

Khatri I., Sharma G., Subramanian S., 2019. Composite genome sequence of *Bacillus clausii,* a probiotic commercially available as Enterogermina®, and insights into its probiotic properties. *BMC Microbiol.,* 19(1): 307-314. https://doi.org/10.1186/ s12866-019-1680-7

Kim S. Y., Song H., Sang M. K., Weon H. Y., Song J., 2017. The complete genome sequence of *Bacillus velezensis* strain GH1-13 reveals agriculturally beneficial properties and a unique plasmid. *J. Biotechnol.,* 259: 221−227. https://doi.org/10.1016/j.jbiotec.2017.06.1 206.

Moriya Y., Itoh M., Okuda S., Yoshizawa A. C., Kanehisa M., 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.,* 35: 182−189. https://doi.org/ 10.1093/nar/gkm321

Overbeek R., Olson R., Pusch G. D., Olsen G. J., Davis J. J., Disz T., Edwards R. A., Gerdes S., Parrello B., Shukla M., Vonstein V., Wattam A. R., Xia F., Stevens R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.,* 42: D206–D214. https://doi.org/10.1093/nar/gkt1226

Panda S. H., Das S., Bal P., Panda S. K., Goli J. K., Mohanty N., 2018. Characterization of novel folate producing *Lactobacillus rhamnosus* and its appliance in fortification of ragi (*Eleusine coracana*) gruel. *Food Biosci.,* 21: 100−106. https://doi.org/10.1016/j.fbio.2017.12.006

Rossi M., Amaretti A., Raimondi S., 2011. Folate production by probiotic bacteria. *Nutrients,* 3(1): 118−134. https://doi.org/ 10.3390/nu3010118

Vu T. H. N., Quach N. T., Nguyen N. A., Nguyen H. T., Ngo C. C., Nguyen T. D., Ho P.-H., Hoang H., Chu H. H., Phi Q.-T., 2021. Genome mining associated with analysis of structure, antioxidant activity

reveals the potential production of levan-rich exopolysaccharides by food-derived *Bacillus velezensis* VTX20. *Appl. Sci.,* 11(15): 7055−7062. https://doi.org/ 10.3390/app11157055

Wang Y., Coleman-Derr D., Chen G., Gu Y. Q., 2015. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.,* 43(W1): 78−84. https://doi.org/10.1093/nar/gkv487

Wattam A. R., Davis J. J., Assaf R., Boisvert S., Brettin T., Bun C., Conrad N., Dietrich E. M., Disz T., Gabbard J. L., Gerdes S., Henry C. S., Kenyon R. W., Machi D., Mao C., Nordberg E. K., Olsen G. J., Murphy-Olson D. E., Olson R., Overbeek R., Parrello B., Pusch G. D., Shukla M., Vonstein V., Warren A., Xia F., Yoo H., Stevens R. L., 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res.,* 45(D1): D535-D542. https://doi.org/10.1093/nar/gkw1017

Ye M., Wei C., Khalid A., Hu Q., Yang R., Dai B., Cheng H., Wang Z., 2020. Effect of *Bacillus velezensis* to substitute in-feed antibiotics on the production, blood biochemistry and egg quality indices of laying hens. *BMC Vet. Res.,* 16(1): 400−409. https://doi.org/10.1186/s12917-020-02570-6

Zhu T., Pan Z., Domagalski N., Koepsel R., Ataai M. M., Domach M. M., 2005. Engineering of *Bacillus subtilis* for enhanced total synthesis of folic acid. *Appl. Environ. Microbiol.,* 71(11): 7122−7129. https://doi.org/10.1128/ AEM.71.11.7122-7129.2005