Supplementary data file for *RSC advance*

# Complete Genome Sequencing of *Arachidicoccus ginsenosidimutans* sp. nov., and Its Application for Production of Minor Ginsenosides by Finding a Novel Ginsenoside-Transforming β-Glucosidase

Muhammad Zubair Siddiqi[1,2], Siddiqi Muhammad Shafi[3] and Wan-Taek Im[1,2]*

[1]Department of Biotechnology, Hankyong National University, 327 Chungang-no Anseong-si, Kyonggi-do 17579, Republic of Korea

[2]Center for Genetic Information, Graduate School of Bio and Information Technology, Hankyong National University, 327 Chungang-no Anseong-si, Kyonggi-do 17579, Republic of Korea

[3]Green planet Co. Ltd, Okayama ken, tsushima nishizaka2 chome 5-41-203, Okayama, Japan

*Corresponding author: Wan Taek Im; E-mail: wandra@hknu.ac.kr

Telephone: +82-31-670-5335; Fax: +82-31-670-5339

**S1 Certification. KCTC certificate of strain BS20$^{T}$.** (PDF)

**S2 Certification. LMG certificate of strain BS20$^{T}$.** (PDF)

Fig. S1

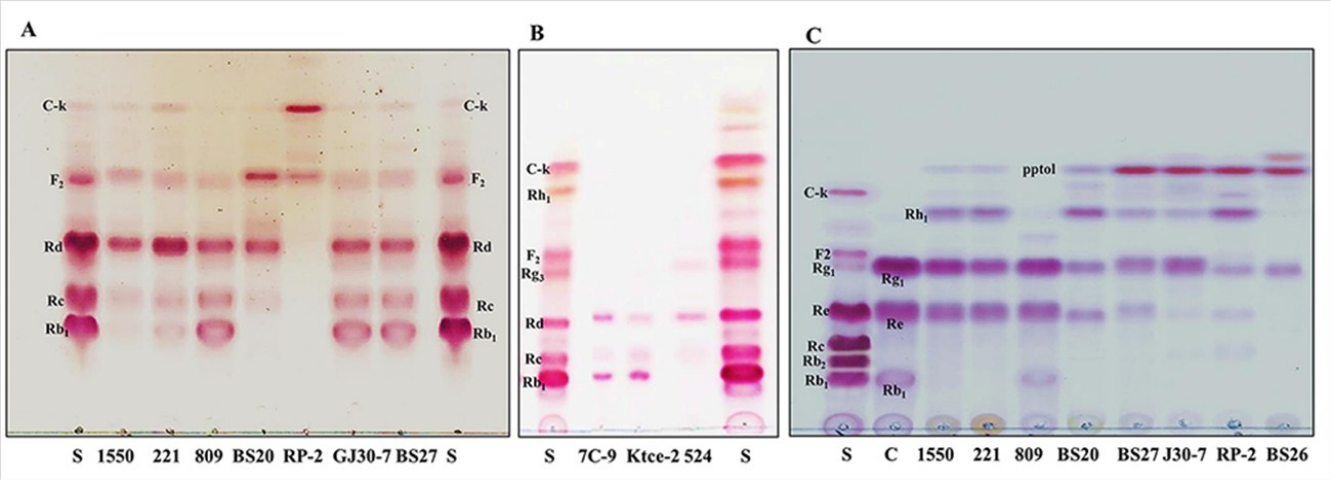


Fig. S1 (A, B, C). Thin layer chromatography [TLC] analysis shows the initial bioconversion of ginsenosides by 11 novels species. **A, C**; S, ginsenoside standards; Gsoil 1550$^{T}$, Gsoil 221$^{T}$; Gsoil 809$^{T}$; BS20$^{T}$; RP-2$^{T}$; GJ30-7$^{T}$; BS27$^{T}$. **B,** 7C-9$^{T}$; Ktce-2$^{T}$; Gsoil 524$^{T}$. C, ginsenosides controls.

Fig. S2

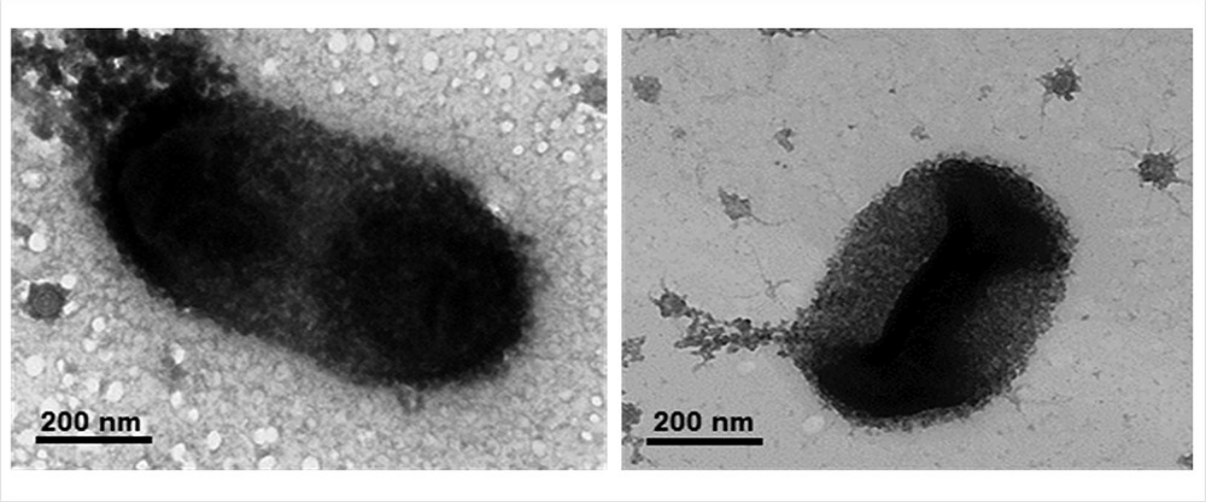


Fig. S2 Transmission electron micrograph of cells of strain BS20$^T$ after a 48-h incubation at 30 °C on R2A agar medium. Bar represents 200 nm.

Fig. S3

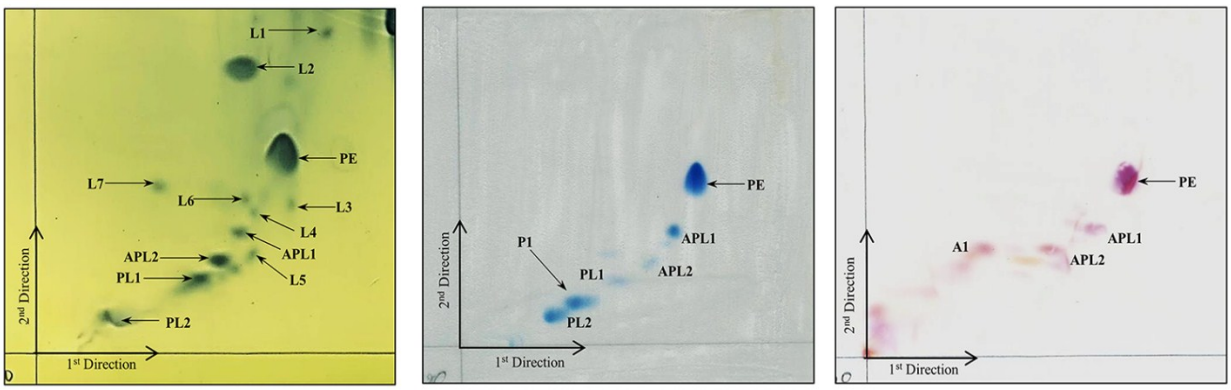


Fig. S3 (A, B, C). Two-dimensional TLC of the total polar lipids of strain BS20$^T$.

Plates were stained for total lipids (A), phospholipids (B), and aminolipids (C) using 5% ethanolic molybdophosphoric acid, molybdate, and 2% ninhydrin reagent, respectively. PE, Phosphatidylethanolamine; L, unidentified polar lipids (L-L7); PL, unidentified phospholipids (PL1 and PL2); APL, unidentified aminophospholipid (APL1-APL2).

Fig. S4

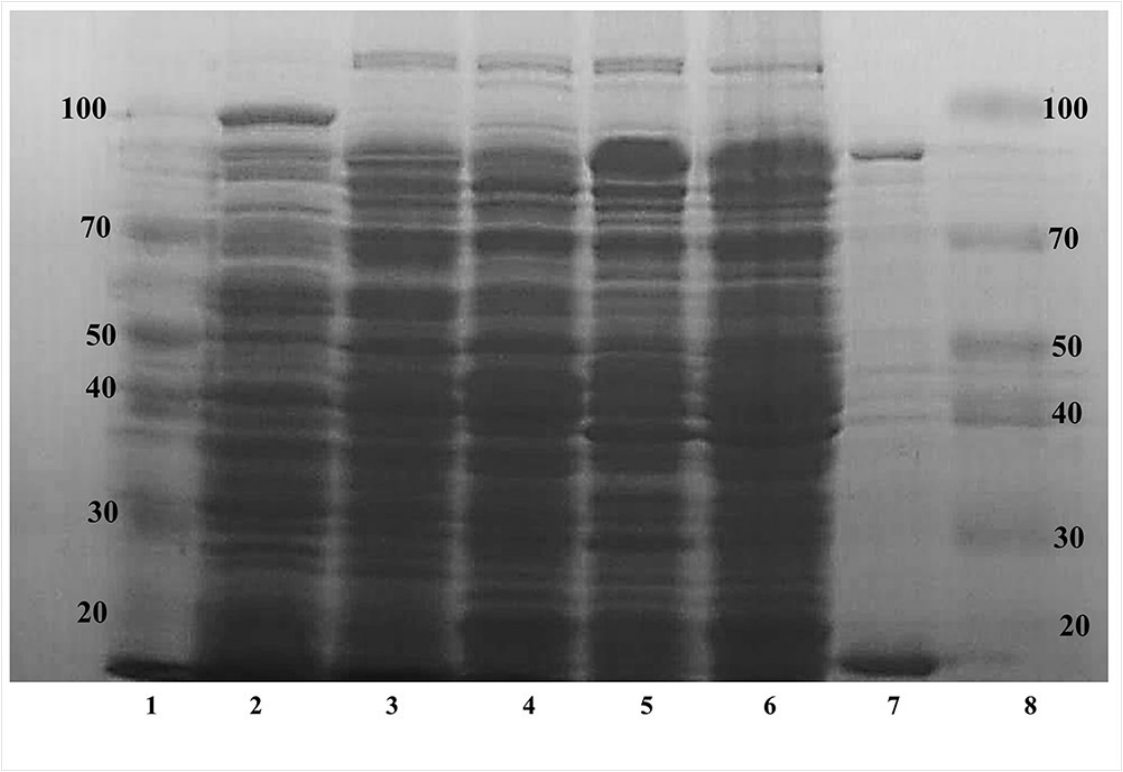


Fig. S4 Lane 1, Protein marker; 2, Crude extract of *E. coli* (BL21 DE3) cell; 3, Crude extract DH5*α* (*E. coli*) cell; 4, BL21 (DE3) harboring pGEX-BglAg-762 before induction; 5, Induced BglAg-762 total protein; 6, Induced BglAg-762 soluble protein; 7, purified protein of BglAg-762; 8, Protein marker.

Fig. S5

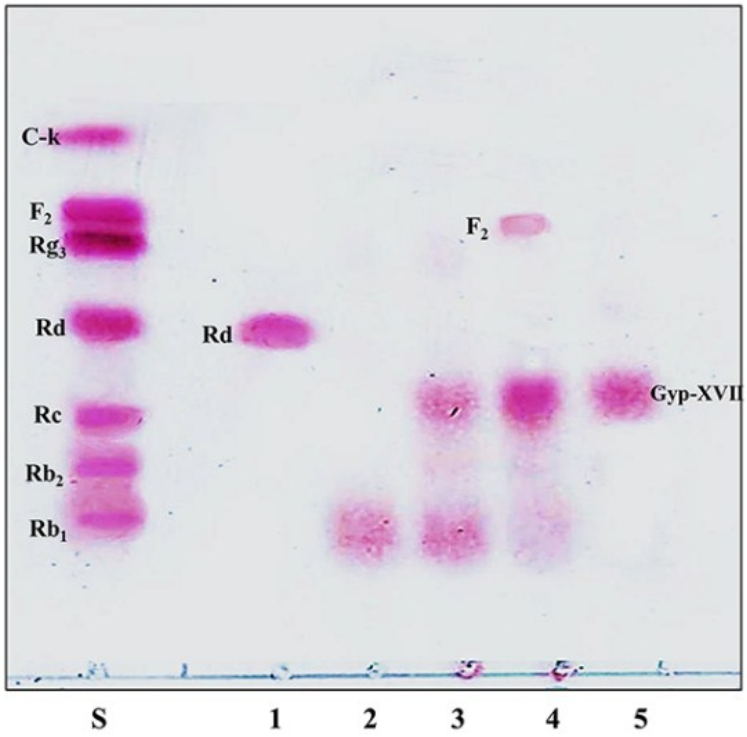


Fig. S5 TLC analyses of time course of bioconversion of ginsenoside $Rb_1$ by BglAg-762.

S, ginsenoside standards; 1, Ginsenoside Rd standard; 2, $Rb_1$ at 0 min; 3, $Rb_1$ after 10 min; 4, $Rb_1$ after 30 min; 5, Gyp-XVII standard.

Table S1 Seventeen [17] sets of clones from strain BS20$^T$ with gene names and primers used in this study (sequences 5r → 3r).

| Total clones | Recombinant Protein ID | Name of protein | Amino acids | Primer sequences (F, R) | Host/ Vector |
|---|---|---|---|---|---|
| 1 | BglAg-726 | *β*-glucosidase-726 | 726 | GGTTCCGCGTGGATCCAAAAAACATTGCATTAAAATTA | |
| | | | | GATGCGGCCGCTCGAGTTAATTAATTACCACATTCGTT | |
| 2 | BglAg-762 | *β*-glucosidase-762 | 762 | GGTTCCGCGTGGATCCAAAAAATCAATCGTAAAAATTT | |
| | | | | GATGCGGCCGCTCGAGTTATTTGCTCCAATCAATAGCT | |
| 3 | BglAg-759 | *β*-glycosidase-759 | 759 | GGTTCCGCGTGGATCCTTTAAGTGCTTCGACATGAAAA | |
| | | | | GATGCGGCCGCTCGAGCTATTGCTTCTTTCCTGCTACA | |
| 4 | BglAg-957 | *β*-glucosidase-957 | 957 | GGTTCCGCGTGGATCCGCAGGGAAAAAAGGTATTATTT | |
| | | | | GATGCGGCCGCTCGAGTCATAAATTTACTATGAGGTTT | |
| 5 | BglAg-757 | *β*-glucosidase-757 | 757 | GGTTCCGCGTGGATCCATCCTCTTCCCAATCTGTTCCT | |
| | | | | GATGCGGCCGCTCGAGTTATTTGTTCCACTCAATCATT | |
| 6 | BglAg-747 | *β*-glucosidase-747 | 747 | GGTTCCGCGTGGATCCAAAAAAATAGTTTTTTTCTTTT | |
| | | | | GATGCGGCCGCTCGAGTTAATATTCAAAATCTTTACTC | *E. coli*/ |
| 7 | BglAg-778 | *β*-glucosidase-778 | 778 | GGTTCCGCGTGGATCCAGGAAATTTTTTTTGTCGGTTA | pGEX |
| | | | | GATGCGGCCGCTCGAGTTATAAATTGATTGAGCCTTTC | 4T-1 |
| 8 | BglAg-796 | *β*-glucosidase-796 | 796 | GGTTCCGCGTGGATCCAAATTTTTCGTTGTTTTTTGTG | |
| | | | | GATGCGGCCGCTCGAGTTAATTGCCCGTTACATTAAAC | |
| 9 | AxylAg-1042 | *α*-xylosidase-1042 | 1042 | GGTTCCGCGTGGATCCCACAAAAAATTTTCAATTGCAT | |
| | | | | GATGCGGCCGCTCGAGCTAAAACTTCAAAACTATCTTT | |
| 10 | XylAg-444/abf-444 | Xylosidase and/or arabinosidase-444 | 444 | GGTTCCGCGTGGATCCAAAAAGTTTATTCTTATTTCCG | |
| | | | | GATGCGGCCGCTCGAGTTATTTAAAACTTATCCAATCA | |
| 11 | AxylAg-847 | *α*-xylosidase-847 | 847 | GGTTCCGCGTGGATCCAAAATTCTCAGCACATCTTTAA | |
| | | | | GATGCGGCCGCTCGAGCTAATATTTTAGTTCGATTGCA | |
| 12 | BxylAg-519 | *β*-xylosidase-519 | 519 | GGTTCCGCGTGGATCCACATGAATAATGCTTTATTAA | |
| | | | | GATGCGGCCGCTCGAGTTATTCGGTTCTGAACCAGTCA | |
| 13 | L-rhamAg-767 | L-rhamnosidase-767 | 767 | GGTTCCGCGTGGATCCATTTTTTCTCATAAGATATTTG | |
| | | | | GATGCGGCCGCTCGAGTCATCTATAATTGACCTGCCAA | |
| 14 | L-rhamAg-881 | L-rhamnosidase-881 | 881 | GGTTCCGCGTGGATCCTTGTTCTTATTTTTTCTTTCGA | |
| | | | | GATGCGGCCGCTCGAGTTATTTTATAACAAACTCGTAG | |
| 15 | L-abfAg-345 | L-arabinofuranosidase-345 | 345 | GGTTCCGCGTGGATCCGTCGCAAAGCAATTATCCATAG | |
| | | | | GATGCGGCCGCTCGAGTTATTCTCCTGACGGAACGGGT | |
| 16 | L-BabfAg-688 | L-*β*-arabinofuranosidase-688 | 688 | GGTTCCGCGTGGATCCAAACAAATAAAAATATTTTTCG | |
| | | | | GATGCGGCCGCTCGAGTTATTTTTTAGCTATTGTAGGG | |
| 17 | L-BabfAg-775 | L-*β*-arabinofuranosidase-775 | 775 | GGTTCCGCGTGGATCCATATGTTGCGTTACCGTTTTGG | |
| | | | | GATGCGGCCGCTCGAGTTATTTCAACAAACGGATGCCG | |

A, *α*; *B*, *β*; Ag, *Arachidicoccus ginsenosidimutans* strain BS20.

Table S2 Effects of different metal ions and chemical agents on the activity of recombinant BglAg-762.

| **Metal ions or reagents** | **Relative activity ± SD ( % ) at:** | |
|---|---|---|
| | **1mM** | **10mM** |
| NaCl | 108.8 ± 0.8 | 56.8 ± 0.6 |
| KCl | 109.8 ± 1.5 | 63.9 ± 0.4 |
| CaCl2 | 128.01 ± 5.1 | 83.1 ± 1.5 |
| MgCl2 | 107.0 ± 1.3 | 55.1 ± 0.9 |
| CoCl2 | 141.3 ± 2.6 | 108.1 ± 7.9 |
| MnSO4 | 131.4 ± 4.0 | 121.4 ± 6.8 |
| MgSO4 | 103.2 ± 1.2 | 60.3 ± 0.1 |
| β-Mercaptoethanol | 105.0 ± 1.1 | 65.2 ± 0.8 |
| EDTA | 101.2 ± 1.2 | 13.58± 0.9 |
| Control | 100.0 ± 0.5 | 100.0 ± 1.5 |

Table S3 NCBI protein blast (blastp) shows the protein similarity and related glycoside hydrolase families of 17 sets of four different glycoside hydrolase of strain BS20.–, no data available.

| S. No | Protein name | Characterized enzyme (strains) | Non-characterized enzyme (strains) | Identity (%) | Glyco_hydrolase Family | Protein named by NCBI |
|---|---|---|---|---|---|---|
| 1 | *β*-glucosidase-726 | – | *Arachidicoccus rhizosphaerae* | 67 | 3 | – |
| 2 | *β*-glucosidase-762 | – | *Mucilaginibacter gotjawali* | 63 | 3 | – |
| 3 | *β*-glucosidase-759 | – | *Chitinophaga jiangningensis* | 66 | 3 | Hypothetical |
| 4 | *β*- glucosidase-957 | – | *Dysgonomonas gadei* | 63 | 3 | Hypothetical |
| 5 | *β*-glucosidase-757 | – | *Chitinophaga jiangningensis* | 63 | 3 | Hypothetical |
| 6 | *β*-glucosidase-747 | – | *Arachidicoccus rhizosphaerae* | 66 | 3 | – |
| 7 | *β*-glucosidase-778 | – | *Pedobacter hartonius* | 58 | 3 | Hypothetical |
| 8 | *β*-glucosidase-796 | – | *Mucilaginibacter* sp. PPCGB 2223 | 71 | 3 | Putative uncharacterized |
| 9 | *α*-xylosidase-1042 | – | *Mucilaginibacter mallensis* | 66 | 31 | Hypothetical |
| 10 | Xylosidase and/or arabinosidase-444 | – | *Niastella koreensis* | 67 | 43 | Hypothetical |
| 11 | *α*-xylosidase-847 | – | *Arachidicoccus rhizosphaerae* | 70 | 31 | Hypothetical |
| 12 | *β*-xylosidase-519 | – | *Parafilimonas terrae* | 77 | 43 | Hypothetical |
| 13 | L-rhamnosidase-767 | – | *Niabella ginsenosidivorans* | 71 | – | Hypothetical |
| 14 | L-rhamnosidase-881 | – | *Mucilaginibacter* | 72 | | Putative |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | *mallensis* | | | uncharacterized |
| 15 | L-arabinofuranosidase-345 | – | *Arachidicoccus rhizosphaerae* | 66 | 43 | Putative uncharacterized |
| 16 | L-*β*-arabinofuranosidase-688 | – | *Arachidicoccus rhizosphaerae* | 69 | – | Hypothetical |
| 17 | L-*β*-arabinofuranosidase-775 | – | *Mucilaginibacter* sp. 44-25 | 60 | 127 | Hypothetical |