

Electronic Supplementary Materials

Identification of Putative Novel Class-I Lanthipeptides in Firmicutes: A Combinatorial *In Silico* Analysis Approach Performed on Genome Sequenced Bacteria and a Close Inspection of Z-geobacillin Lanthipeptide Biosynthesis Gene Cluster of the Thermophilic *Geobacillus* sp. Strain ZGt-1

Rawana N. Alkhalili ^{1,*} and Björn Canbäck ²

¹ Biotechnology, Department of Chemistry, Lund University Lund, SE-221 00, Sweden

² Department of Biology, Lund University, Lund SE-221 00, Sweden; Bjorn.Canback@biol.lu.se

* Correspondence: Rawana.Alkhalili@biotek.lu.se

Table S.1. Amino acid sequences of the unmodified antiSMASH-detected lanthipeptides and the corresponding nucleotide sequences. Sequences in lower case correspond to the leader peptide. The presented gene positions are based on those predicted by antiSMASH and confirmed by BLAST analysis. When the gene start position value is larger than that of the end position, the gene is located on the reverse strand. Entries typed in bold represent data related to the putative novel lanthipeptides.

Bacterial species, strain (lanthipeptide reference number), unmodified aa sequence of lanthipeptide	Bacterial species, strain (lanthipeptide reference number), lanthipeptide nucleotide sequence	RefSeq genome accession number	Start position	End position
> <i>Bacillus clausii</i> KSM-K16 mekafddldlevvhtkakdvqpdFTSVSFCTPGCGETGS FNSFCC	> <i>Bacillus clausii</i> KSM-K16 ATGGAAAAAGCTTTTGACTTGGATTTGGAA GTAGTACACACGAAAGCGAAGGATGTACA ACCAGATTTCAACAGCGTCAGTTTCTGTAC TCCTGGTTGCGGGGAAACGGGTAGCTTTAA TAGCTTCTGCTGCTAA	NC_006582	3706418	3706284
> <i>Bacillus megaterium</i> QM B1551 (I) mekafddldlevvhtkakdvqpdFTSVSFCTPGCGETGS FNSFCC	> <i>Bacillus megaterium</i> QM B1551 (I) ATGAATAACGTAAAAACCTTTTGATTGA GATGTACAAGTAACAACCTCAAGTGAT GTAGATCCACAAATTACAAGTGATCTTTA TGTACACCAGGTTGTGGAGATACAGGTAGT TGGAACAGCTTCTGCTGCTAA	NC_014023	137678	137538
> <i>Bacillus megaterium</i> QM B1551 (II) mnnvknfldldvqvtassdvdppqITSVSLCTPGCGDT GSWNSFCC	> <i>Bacillus megaterium</i> QM B1551 (II) ATGAATAATGTAAAAACCTTTTGATTGA GATGTACAAGTAACAACCTCAAGTGAT GTAGATCCACAAATTACAAGTGATCTTTA TGTACACCAGGTTGTGGAGATACAGGTAGC TGGAACAGCTTCTGCTGCTAA	NC_014023	137893	137753
> <i>Bacillus subtilis</i> BSn5 meknnifddinkmestsevsaqTWTATIGKTIVQSVK KCRFTFCGCSLGSNSCN	> <i>Bacillus subtilis</i> BSn5 ATGGAGAAGAATAATTTTTGACCTAGAC ATTAATAAAAAGATGGAATCTACTAGTGA GGTTTCTGCTCAAACCTGGGCTACAATCGG AAAAACTATTGTGCAGTCTGTGAAAAAAT GTAGAACATTTACTTGTGGATGTTGCTAG GCTCTGCACTAAGTAACTAA	NC_014976	2428804	2428974
> <i>Bacillus subtilis spizizenii</i> TU-B-10 mskfddfdldvkvskqdskitpqWKSESVCTPGCVT GLLQTCFLQITCNCKISK	> <i>Bacillus subtilis spizizenii</i> TU-B-10 ATGTCAAAGTTCGATGATTTTCGATTTGGAT GTTGTGAAAGTCTCTAAGCAAGACTCAA AATCACTCCTCAATGGAAAAGTGAGTCAGT TTGTACACCAGGTTGTGTAACCTGGTCTATT GCAAACCTGCTTCCTTCAAACAATCACTTG TAACTGCAAAATCTCTAAATAA	NC_016047	3414730	3414560
> <i>Bacillus subtilis spizizenii</i> W23 mskfddfdldvkvskqdskitpqWKSESLCTPGCVT GALQTCFLQITCNCKISK	> <i>Bacillus subtilis spizizenii</i> W23 ATGTCAAAGTTCGATGATTTTCGATTTGGAT GTTGTGAAAGTCTCTAAACAAGACTCAA AATCACTCCGCAATGGAAAAGTGAATCAC TTGTACACCAGGATGTGTAACCTGGTGCAT TGCAAACCTGCTTCCTTCAAACAATCACTTG GTAACCTGCAAAATCTCTAAATAA	NC_014479	3256972	3256802
> <i>Bacillus thuringiensis serovar finitimus</i> YBT-020 mdnksqfedfdlnlyeedysedgsvqsySSWGGTNS STSIACSLTNVTEGCNWC	> <i>Bacillus thuringiensis serovar finitimus</i> YBT-020 ATGGACAACAAATCGCAATTTGAAGATT TTGATTTAAACTATATGAAGAAATTGAT TATAGTGAAGATGGATCAGTGCAATCAT ACTCATCTTGGGGGGGAACGAATAGTTC TACATCAATTGCATGTTCTTTATTAATA ATGTTACTGAAGGATGTAATTGGTGCTA A	NC_017200	1205213 ¹	1205383 ¹

> <i>Bacillus thuringiensis</i> serovar IS5056 (I) mnkelfdldinkmetptemtaqTWTIVKVS KAVCK TGTCICTTSCSNCK	> <i>Bacillus thuringiensis</i> serovar IS5056 (I) ATGAATAAAGA AACTATTTGATTTAGATATT AATAAGAAAATGGAAACACCTACTGAAAT GACAGCACAAACTTGGACA AACTATTGTTAA AGTCTCAAAGGCTGTTGTAAAACAGGAA CTTGTATTTGACTACTTCATGCTCTAATTG TAAGTAA	NC_020394	132195	132040
> <i>Bacillus thuringiensis</i> serovar IS5056 (II) mnkelfdldinkmetptemtaqTWTIVKVS KAVCK TGTCICTTSCSNCK	> <i>Bacillusthuringiensis</i> serovar IS5056 (II) ATGAATAAAGA AACTATTTGATTTAGATATT AATAAGAAAATGGAAACACCTACTGAAAT GACAGCACAAACTTGGACA AACTATTGTTAA AGTCTCAAAGGCTGTTGTAAAACAGGAA CTTGTATTTGACTACTTCATGCTCTAATTG TAAGTAA	NC_020394	132712	132557
> <i>Bacillus thuringiensis</i> serovar IS5056 (III) mdkelfdldinkmetptemtaqTWT SILKVS RAPCKS GTCLCTTSCSNCK	> <i>Bacillus thuringiensis</i> serovar IS5056 (III) ATGGATAAAGA AACTATTTGATTTAGATATT AATAAGAAAATGGAAACACCTACTGAAAT GACAGCACAAACTTGGACTAGTATTTTAAA GGTGTCAAGGGCTCCTGTAAATCAGGAAC TTGTCTTTGACTACTTCATGCTCTAATTGT AAATAA	NC_020394	133239	133084
> <i>Bacillus thuringiensis</i> serovar IS5056 (IV) mnkelfdldinkmeaptemvtqwGTVVKVSKAICK TGTCIGTISCTNCK	> <i>Bacillus thuringiensis</i> serovar IS5056 (IV) ATGAATAAAGA AACTATTTGATTTAGATATT AATAAGAAAATGGAAGCACCTACTGAAAT GACTGTACAAACTTGGGGAACGGTTGTTAA AGTATCAAAGGCTATTTGTAAAACAGGAA CTTGTATTGGCACTATTCATGCTACTAATTG TAAGTAA	NC_020394	133674	133519
> <i>Bacillus thuringiensis</i> serovar IS5056 (V) mnkelfdldinkmetptemtaqvPTTIFVSRVCKTL TCICTISCSNCK	> <i>Bacillus thuringiensis</i> serovar IS5056 (V) ATGAATAAAGA ATTATTTGATTTAGATATT AATAAGAAAATGGAAACACCTACTGAAAT GACTGCACAAACTGTGCCAACAACTATTTT TGTATCAAGAAGTGTGTTGTAAAACATTAAC TTGTATTTGACTATTTTCATGCTCTAATTGT AAGTAA	NC_020394	134225	134070
> <i>Bacillus thuringiensis</i> YBT 1518 (I) mineknfldldvqvttatgdvdpqITRISACTPGCGNTG SFNSFCC	> <i>Bacillus thuringiensis</i> YBT 1518 (I) ATGATTAATGAAAAAAACCTTTTTGATTTA GATGTACAAGTAACAACAGCTACAGGCGA TG TAGATCCACAAATTACAAGTATCTCAGC TTGTACGCCCTGGTTGTGGAAACACAGGTTT ATTCAATAGCTTCTGTGCTAA	NC_022873	3953881	3953741
> <i>Bacillus thuringiensis</i> YBT 1518 (II) mineknfldldvqvttatgdvdpqITRISACTPGCGNTG SFNSFCC	> <i>Bacillus thuringiensis</i> YBT 1518 (II) ATGATTAATGAAAAAAACCTTTTTGATTTA GATGTACAAGTAACAACAGCTACAGGTTA TG TAGATCCACAAATTACAAGAATCTCAGC TTGTACGCCCTGGTTGTGGAAACACAGGTTT ATTCAATAGCTTCTGTGCTAA	NC_022873	3954071	3953931
> <i>Geobacillus kaustophilus</i> HTA426 (I) maklddflddivvkkqdnivqpnITSKSLCTPGCITGIL MCLTQNSCVSCNSCIRC	> <i>Geobacillus kaustophilus</i> HTA426 (I) ATGGCAAAGTTGGATGATTTTCGATTTGGAC ATTGTTGTCAAAAAACAGGACAACATTGTA CAACCTAATATTACAAGTAAATCTCTTTGT ACCCCGGGTTGCATCACAGGAATCTTGATG TGCTTAACGCAAAATTCTTGTGTATCTTGTA ACTCTTGCAATTCGATGCTAA	NC_006510	314934	315104
> <i>Geobacillus kaustophilus</i> HTA426 (II) maklddflddivvkkqdnivqpnITSKSLLYPGLHHR HLNVLNAKFLCIL	> <i>Geobacillus kaustophilus</i> HTA426 (II) ATGGCGAAGTTGGATGATTTTCGATTTGGAC ATTGTTGTCAAAAAACAGGACAACATTGTA CAACCTAATATTACAAGTAAATCTCTTTTGT	NC_006510	320984 ¹	321136 ¹

	ACCCCGGGTTGCATCACAGGCACCTTAATG TGCTTAACGCAAAATTCTTGTGTATCTTGTA A			
> <i>Geobacillus</i> sp. ZGt-1 maklddflddivvkkqdnivqpnITSKSLCTPGCITG ILMCLTQNSCVSCNSCIRC	> <i>Geobacillus</i> sp. ZGt-1 ATGGCAAAGTTGGATGATTTTCGATTTGG ACATTGTTGTCAAAAAACAGGACAACAT TGTACAACCTAATATTACAAGTAAATCTC TTTGTACCCCGGGTTGCATCACAGGAATC TTGATGTGCTTAACGCAAAATTCTTGTGT ATCTTGTAACTCTTGCAATTCGATGCTAA	LDPD01000000 ²	33624	33794
> <i>Geobacillus thermodenitrificans</i> NG80-2 maklddflddivvkkqdnivqpnVTSKSLCTPGCITG VLMCLTQNSCVSCNSCIRC	> <i>Geobacillus thermodenitrificans</i> NG80-2 ATGGCGAAGTTGATGATTTTGATTTGGAC ATTGTGGTGAAAAAACAGGACGACGTGGT ACAACCGAACGTGACGAGTAAATCTCTTTG TACGCCAGGTTGTATCACAGGTGTGTTAAT GTGTCTCACGCAAAATTCTTGTGTATCTTGT AACTCTTGCATTCGATGCTAA	NC_009328	298823	298993
> <i>Geobacillus thermoleovorans</i> CCB_US3_UF5 maklddflddivvkkqdnivqpnITSKSLYPGLHHR HLNVLNAKFLCIL	> <i>Geobacillus thermoleovorans</i> CCB_US3_UF5 ATGGCGAAGTTGGATGATTTTCGATTTGGAC ATTGTTGTCAAAAAACAGGACAACATTGTA CAACCTAATATTACAAGTAAATCTCTTTTGT ACCCCGGGTTGCATCACAGGCATCTTAATG TGCTTAACGCAAAATTCTTGTGTATCTTGTA A	NC_016593	362226 ¹	362378 ¹
> <i>Lactococcus lactis</i> CV56 mstkdfnldlvsvskkdsqasprITSISLCTPGCKTGAL MGCNMKTATCHCSIHVSK	> <i>Lactococcus lactis</i> CV56 ATGAGTACAAAAGATTTTAACTTGGATTG GTATCTGTTTGAAGAAAGATTCAGGTGCA TCACCACGCATTACAAGTATTTGCTATGT ACACCCGGTTGTAACAGGAGCTCTGAT GGTTGTAAACATGAAAACAGCAACTTGTG ATTGTAGTATTCACGTAAGCAAATAA	NC_017486	604687	604860
> <i>Lactococcus lactis</i> IO mstkdfnldlvsvskkdsqasprITSISLCTPGCKTGAL MGCNMKTATCNCISHVSK	> <i>Lactococcus lactis</i> IO ATGAGTACAAAAGATTTTAACTTGGATTG GTATCTGTTTGAAGAAAGATTCAGGTGCA TCACCACGCATTACAAGTATTTGCTATGT ACACCCGGTTGTAACAGGAGCTCTGAT GGTTGTAAACATGAAAACAGCAACTTGTG ATTGTAGTATTCACGTAAGCAAATAA	NC_020450	599615	599788
> <i>Paenibacillus polymyxa</i> CR1 mknqfdldlvaknevpkqvpaSGII CTPSCATGTLNCQVSLTFCKTC	> <i>Paenibacillus polymyxa</i> CR1 ATGAAAAATCAATTTGATCTGGACTTGCAA GTAGCTAAAAATGAGGTTGCTCCAAAAGG AGTTCAACCAGCCAGTGGTATTATCTGTAC TCCTTCATGTGCAACAGGACTTTGAATTGT CAAGTTAGTCTTACCTTCTGCAAAAACGTGT TAA	NC_023037	1748588	1748740
> <i>Paenibacillus polymyxa</i> E681 mknqfdldlvaknevasqpaSGLICTPSCATGTL NCQVSLSFCKTC	> <i>Paenibacillus polymyxa</i> E681 ATGAAAAATCAATTTGATCTGGACTTGCAA GTAGCTAAAAATGAGGTTGCTCCAAAAGG GGTTCAACCAGCCAGTGGTCTTATCTGTAC TCCTTCATGTGCAACAGGACTTTGAATTGT CAAGTTAGTCTTACCTTCTGCAAAAACGTGT TAA	NC_014483	1558420	1558572
> <i>Paenibacillus polymyxa</i> M1 (I) mknqfdldlvaknevasqpaSGLICTPCLTSIL NCYTSISHCGPC	> <i>Paenibacillus polymyxa</i> M1 (I) ATGAAAAATCAATTTGATCTGGATTGCA AGTGACCAAAAAGTGAGAGTGCTTCTAAA GAGCTTCAAGCAGACAGCGGCATTATTT GTACTCCTACTTGTAAACATCTATTTTAA ACTGCTATACTAGCATTTCACATTGTGGA CCTTGTTGA	NC_017542	1693817	1693969

> <i>Paenibacillus polymyxa</i> M1 (II) mknqfdldlvaknevpkeqpasGLICTPSCATG TLNCQVSLSFCKTC	> <i>Paenibacillus polymyxa</i> M1 (II) ATGAAAAATCAATTTGATCTGGACTTGCA AGTAGCCAAAAATGAGGTTGCTCCTAAA GAGGTTCAACCAGCTAGCGGTCTTATCT GTACTCCTTCCTGTGCAACAGGTACTTTG AATTGCCAAGTTAGCCTTTCCTTCTGCAA AACGTGTAA	NC_017542	1695699	1695851
> <i>Paenibacillus polymyxa</i> SC2 (I) mknqfdldlvqtkesesaskelqadSGIICPTCLTSIL NCYTSISHCGPC	> <i>Paenibacillus polymyxa</i> SC2 (I) ATGAAAAATCAATTTGATCTGGATTGCA AGTGACCAAAAGTGAGAGTGCTTCTAAA GAGCTTCAAGCAGACAGCGGCATTATTT GTACTCCTACTTGTTAACATCTATTTTAA ACTGCTATACTAGCATTTCACATTGTGGA CCTTGTTGA	NC_014622	1695159	1695311
> <i>Paenibacillus polymyxa</i> SC2 (II) mknqfdldlvaknevpkeqpasGLICTPSCATG TLNCQVSLSFCKTC	> <i>Paenibacillus polymyxa</i> SC2 (II) ATGAAAAATCAATTTGATCTGGACTTGCA AGTAGCCAAAAATGAGGTTGCTCCTAAA GAGGTTCAACCAGCTAGCGGTCTTATCT GTACTCCTTCCTGTGCAACAGGTACTTTG AATTGCCAAGTTAGCCTTTCCTTCTGCAA AACGTGTAA	NC_014622	1697041	1697193
> <i>Staphylococcus aureus</i> 11819-97 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> 11819-97 (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATT AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_017351	1942472	1942329
> <i>Staphylococcus aureus</i> 11819-97 (II) mekvldldvqvkannntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> 11819-97 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATT AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTAA	NC_017351	1943352	1943209
> <i>Staphylococcus aureus</i> Bmb9393 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> Bmb9393 (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATT AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_021670	2038104	2037961
> <i>Staphylococcus aureus</i> Bmb9393 (II) mekvldldvqvkannntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> Bmb9393 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATT AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTAA	NC_021670	2038984	2038841
> <i>Staphylococcus aureus</i> COL (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> COL (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATT AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_002951	1932028	1931885
> <i>Staphylococcus aureus</i> COL (II) mekvldldvqvkannntndsagdeRITSHLFCFSGCEK TVLHQMWVWYNLKNYF	> <i>Staphylococcus aureus</i> COL (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATT AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGT TCTCCACCAAATGTGGTGGGTATATAATTT AAAGAACTATTTTTAA	NC_002951	1934429	1934265

> <i>Staphylococcus aureus</i> ED133 (I) mekvldldvqvkakndtsdsGDERITFIGCTPGCG KTGSFNSFCC	> <i>Staphylococcus aureus</i> ED133 (I) ATGGAAAATGTTCTTGATTTAGACGTACAA GTAAAAGCGAAAAACGACACTTCTGATTC AGCGGGTGACGAACGTATTACTAGTTTCAT TGGTTGACTCCTGGTTGTGGTAAGACTGGT AGTTTTAATAGCTTCTGCTGTAA	NC_017337	1905359	1905216
> <i>Staphylococcus aureus</i> ED133 (II) mekvldldvqvkgnntndsaderITSHFLCSFGCGK TGSFNSFCC	> <i>Staphylococcus aureus</i> ED133 (II) ATGGAAAAAGTTCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGCTGACGAAAGAATAACTAGCCATT TTTTATGTAGCTTTGGTTGTGGAAAGACGG GTAGTTTTAACAGCTTCTGTTGTAA	NC_017337	1907901	1907758
> <i>Staphylococcus aureus</i> LGA251 (I) mekvldldvqvkakndtsdsGDERITFIGCTPGCG KTGSFNSFCC	> <i>Staphylococcus aureus</i> LGA251 (I) ATGGAAAATGTTCTTGATTTAGACGTACAA GTAAAAGCGAAAAACGACACTTCTGATTC AGCGGGTGACGAACGTATTACTAGTTTCAT TGGTTGACTCCTGGTTGTGGTAAGACTGGT AGTTTTAATAGCTTCTGCTGTAA	NC_017349	1893217	1893074
> <i>Staphylococcus aureus</i> LGA251 (II) mekvldldvqvkgnntndsaderITSHFLCSFGCGK TGSFNSFCC	> <i>Staphylococcus aureus</i> LGA251 (II) ATGGAAAAAGTTCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGCTGACGAAAGAATAACTAGCCATT TTTTATGTAGCTTTGGTTGTGGAAAGACGG GTAGTTTTAACAGCTTCTGTTGTAA	NC_017349	1894096	1893953
> <i>Staphylococcus aureus</i> M1 (I) mekvldldvqvkannsndsagdeRITSHSLCTPGCA KTGSFNSFC	> <i>Staphylococcus aureus</i> M1 (I) ATGGAAAAAGTTCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_021059	1944795	1944652
> <i>Staphylococcus aureus</i> M1 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> M1 (II) TTGGAAAAAGTTCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTAA	NC_021059	1945675	1945532
> <i>Staphylococcus aureus</i> MSSA476 (I) mekvldldvqvkannsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> MSSA476 (I) ATGGAAAAAGTTCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_002953	1896691	1896548
> <i>Staphylococcus aureus</i> MSSA476 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> MSSA476 (II) TTGGAAAAAGTTCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTAA	NC_002953	1897571	1897428
> <i>Staphylococcus aureus</i> MW2 (I) mekvldldvqvkannsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> MW2 (I) ATGGAAAAAGTTCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_003923	1917357	1917214
> <i>Staphylococcus aureus</i> MW2 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> MW2 (II) TTGGAAAAAGTTCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG	NC_003923	1918237	1918094

	TAGTTTTAACAGCTTCTGTGTAA			
> <i>Staphylococcus aureus</i> NCTC 8325 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> NCTC 8325 (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_007795	1857888	1857745
> <i>Staphylococcus aureus</i> NCTC 8325 (II) lekvlldlvqvkgnntndsagdeRITSHLFCFSGCE KTGSFNSFCC	> <i>Staphylococcus aureus</i> NCTC 8325 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA AGTTAAAGGAAATAACAACACTAATGAT TCAGCGGGTGACGAAAGAATAACTAGCC ATCTTTTTGTAGCTTTGGTTGTGAAAAG ACGGGTAGTTTTAACAGCTTCTGTTGTTA A	NC_007795	1858768	1858625
> <i>Staphylococcus aureus</i> Newman (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> Newman (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_009641	1914736	1914593
> <i>Staphylococcus aureus</i> Newman (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> Newman (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTGTAA	NC_009641	1915616	1915473
> <i>Staphylococcus aureus</i> RF122 (I) menvldldvqvkakndtsdsGDERITSFIGCTPGCG KTGSFNSFCC	> <i>Staphylococcus aureus</i> RF122 (I) ATGGAAAATGTTCTTGATTTAGACGTACAA GTTAAAGCGAAAAACGACACTTCTGATTC AGCGGGTGACGAACGTATTACTAGTTTCAT TGGTTGTACTCCTGGTTGTGTAAGACTGGT AGTTTTAATAGCTTCTGCTGTAA	NC_007622	1829509	1829366
> <i>Staphylococcus aureus</i> RF122 (II) mekvldldvqvkgnntndsagdeRITSHFLCSFGCGK TGSFNSFCC	> <i>Staphylococcus aureus</i> RF122 (II) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATT TTTTATGTAGCTTTGGTTGTGAAAAGACGG GTAGTTTTAACAGCTTCTGTGTAA	NC_007622	1830386	1830243
> <i>Staphylococcus aureus</i> T0131 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> T0131 (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_017347	1932461	1932318
> <i>Staphylococcus aureus</i> T0131 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> T0131 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTGTAA	NC_017347	1933341	1933198
> <i>Staphylococcus aureus</i> TW20 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> TW20 (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTAA	NC_017331	1978007	1977864
> <i>Staphylococcus aureus</i> TW20 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> TW20 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC	NC_017331	1978887	1978744

	AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTTAA			
> <i>Staphylococcus aureus</i> USA300 FPR3757 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> USA300 FPR3757(I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTTAA	NC_007793	1954197	1954054
> <i>Staphylococcus aureus</i> USA300 FPR3757 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> USA300 FPR3757 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTTAA	NC_007793	1955077	1954934
> <i>Staphylococcus aureus</i> USA300_TCH1516 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> USA300_TCH1516 (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTTAA	NC_010079	1954946	1954803
> <i>Staphylococcus aureus</i> USA300_TCH1516 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> USA300_TCH1516 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTTAA	NC_010079	1955826	1955683
> <i>Staphylococcus aureus</i> VC40 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> VC40 (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTTAA	NC_016912	1815518	1815375
> <i>Staphylococcus aureus</i> VC40 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> VC40 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTTAA	NC_016912	1816398	1816255
> <i>Staphylococcus aureus</i> Z172 (I) mekvldldvqvkannnsndsagdeRITSHSLCTPGCA KTGSFNSFCC	> <i>Staphylococcus aureus</i> Z172 (I) ATGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGCAAACAATAACTCAAATGATTC AGCAGGTGACGAACGTATTACAAGTCATA GTTTATGTACTCCTGGTTGTGCTAAGACTGG TAGTTTTAATAGCTTCTGCTGTTAA	NC_022604	1976108	1975965
> <i>Staphylococcus aureus</i> Z172 (II) mekvldldvqvkgnntndsagdeRITSHLFCFSGCEK TGSFNSFCC	> <i>Staphylococcus aureus</i> Z172 (II) TTGGAAAAAGTTCCTTGATTTAGACGTGCAA GTTAAAGGAAATAACAACACTAATGATTC AGCGGGTGACGAAAGAATAACTAGCCATC TTTTTTGTAGCTTTGGTTGTGAAAAGACGGG TAGTTTTAACAGCTTCTGTTGTTAA	NC_022604	1978319	1978176
> <i>Streptococcus intermedius</i> B196 msnhddfqltdnkvndndqevrGPLSWLGGTVLT SIVTGCTGDCLTRHCADNVTKNTGCTVTH PRCKR	> <i>Streptococcus intermedius</i> B196 ATGTCAAATCATGATGATTTTCAATTAGA TACAAACAAAGTTAATTTTGATAATGATC AAGAAGTAAGAGGTCCGCTTTCTTGGCT AGGCGGAACAGTATTAACCTCAATAGTT ACAGGCTGCACAGGAGATTGTTTAAACAA GGCACTGTGCCGATAATGTTACAAAAAA TACAGGCTGTACAGTTACGCATCCTAGAT	NC_022246	590584	590793

GTAAAAGGTAG

> <i>Streptococcus intermedius</i> C270 mstkdfnldlvsvsktdagasprVTSVFMCTPGCVTGA LMGCNLQTATCNCHVHISK	> <i>Streptococcus intermedius</i> C270 ATGAGTACGAAAGATTTAACTTAGATTG GTTTCTGTTTCAAAAACAGATGCTGGTGCTT CACCTCGTGTACTAGTGTGTTTATGTGTAC TCCAGGTTGTGTAAGTGGTCTTTGATGGGT TGTAAGTGCAAAACAGCTACATGTAAGTGT CATGTTACATTAGTAAATAA	NC_022237	1010608	1010435
> <i>Streptococcus pasteurianus</i> ATCC 43144 mnnedfnldlvtskennsgasprVTSKSLCTPGCKTGI LMTCAIKTATCGCHF	> <i>Streptococcus pasteurianus</i> ATCC 43144 ATGAACAATGAAGATTTAATTTGGATCTT GTCACAATCTCAAAGGAAAACAACACTCAGG AGCTTCACCTCGGGTAACTAGTAAATCATT ATGTACTCCTGGATGTAAGACGGGTATTTT GATGACCTGTGCAATCAAAAACGCAACTTG TGTTGCCATTTGGATAA	NC_015600	1148270	1148103
> <i>Streptococcus pyogenes</i> MGAS6180 mnntikdfldldlknkddatpyVGSRYLCTPGSCWKL VCFTTTVK	> <i>Streptococcus pyogenes</i> MGAS6180 ATGAATAACACAATTAAGACTTTGATCTC GATTTGAAAACAAATAAAAAAGACACTGC TACACCTTATGTTGGTAGCCGTTACCTATGT ACCCCTGGTCTTGTGGAAATTAGTTTGCT TTACAACAACACTGTAAATAA	NC_007296	804231	804371
> <i>Streptococcus pyogenes</i> MGAS9429 mnntikdfldldlknkddatpyVGSRYLCTPGSCWKL VCFTTTVK	> <i>Streptococcus pyogenes</i> MGAS9429 ATGAATAACACAATTAAGACTTTGATCTC GATTTGAAAACAAATAAAAAAGACACTGC TACACCTTATGTTGGTAGCCGTTACCTATGT ACCCCTGGTCTTGTGGAAATTAGTTTGCT TTACAACAACACTGTAAATAA	NC_008021	885676	885816
> <i>Streptococcus pyogenes</i> MGAS10270 mnntikdfldldlknkddatpyVGSRYLCTPGSCWKL VCFTTTVK	> <i>Streptococcus pyogenes</i> MGAS10270 ATGAATAACACAATTAAGACTTTGATCTC GATTTGAAAACAAATAAAAAAGACACTGC TACACCTTATGTTGGTAGCCGTTACCTATGT ACCCCTGGTCTTGTGGAAATTAGTTTGCT TTACAACAACACTGTAAATAA	NC_008022	876632	876772
> <i>Streptococcus pyogenes</i> MGAS10750 mnttikdfldldlknkddatpyVGSRYLCTPGSCWKL VCFTTTVK	> <i>Streptococcus pyogenes</i> MGAS10750 ATGACTAACACAATTAAGACTTTGATCTC GATTTGAAAACAAATAAAAAAGACACTGC TACACCTTATGTTGGTAGCCGTTACCTATGT ACCCCTGGTCTTGTGGAAATTAGTTTGCT TTACAACAACACTGTAAATAA	NC_008024	907659	907799
> <i>Streptococcus suis</i> JS14 mnnedfnldlvtskennsgasprVTSKSLCTPGCKTGI LMTCAIKTATCGCHF	> <i>Streptococcus suis</i> JS14 ATGAACAATGAAGATTTAATTTGGATCTT GTCACAATCTCAAAGGAAAACAACACTCAGG AGCTTCACCTCGGGTAACTAGTAAATCATT ATGTACTCCTGGATGTAAGACGGGTATTTT GATGACCTGTGCAATCAAAAACGCAACTTG TGTTGCCATTTGGATAA	NC_017618	646342	646509
> <i>Streptococcus suis</i> SC070731 mnnedfnldlvtskennsgasprVTSKSLCTPGCKTGI LMTCAIKTATCGCHF	> <i>Streptococcus suis</i> SC070731 ATGAACAATGAAGATTTAATTTGGATCTT GTCACAATCTCAAAGGAAAACAACACTCAGG AGCTTCACCTCGGGTAACTAGTAAATCATT ATGTACTCCTGGATGTAAGACGGGTATTTT GATGACCTGTGCAATCAAAAACGCAACTTG TGTTGCCATTTGGATAA	NC_020526	650359	650526

¹ The presented gene positions are based on those predicted by antiSMASH 3.0 and confirmed by BLAST analysis.

² The presented genome sequence accession number for *Geobacillus* sp. ZGt-1 belongs to the original genome record; the whole genome shotgun (WGS) project.

Table S.2. The antiSMASH-predicted lanthipeptides whose coding genes have either been unannotated, inaccurately annotated, or annotated as coding for “hypothetical” proteins or generally as “antibiotics”.

Bacterial species, strain (lanthipeptide reference number) ¹	Refseq genome accession number	Annotation of the predicted lanthipeptide		Identity to experimentally verified lanthipeptide ²	Reference ³
		Original genome record	RefSeq genome record		
<i>Bacillus clausii</i> KSM-K16	NC_006582	Unannotated	“Gallidermin/nisin family lantibiotic”	56% clausin	[30], and this study
<i>Bacillus megaterium</i> QM B1551 (I)	NC_014023	“Antibiotic protein, putative”	“Gallidermin/nisin family lantibiotic”	56% gallidermin	[31], and this study
<i>Bacillus megaterium</i> QM B1551 (II)	NC_014023	“Antibiotic protein, putative”	“Gallidermin/nisin family lantibiotic”	56% gallidermin	[31], and this study
<i>Bacillus subtilis</i> BSn5	NC_014976	“Hypothetical”	“Hypothetical”	100% subtilomycin	[32], and this study
<i>Bacillus subtilis spizizenii</i> W23	NC_014479	“Hypothetical”	“Gallidermin/nisin family lantibiotic”	100% subtilin	This study
<i>Bacillus thuringiensis</i> serovar finitimus YBT-020	NC_017200	“Hypothetical”	Unannotated	No hits	This study
<i>Bacillus thuringiensis</i> serovar IS5056 (I)	NC_020394	“Hypothetical”	“Hypothetical”	100% thuricin 4A-4	[31], and this study
<i>Bacillus thuringiensis</i> serovar IS5056 (II)	NC_020394	“Hypothetical”	“Hypothetical”	100% thuricin 4A-4	[31], and this study
<i>Bacillus thuringiensis</i> serovar IS5056 (III)	NC_020394	“Hypothetical”	“Hypothetical”	86% thuricin 4A-4	[31], and this study
<i>Bacillus thuringiensis</i> serovar IS5056 (IV)	NC_020394	“Hypothetical”	“Hypothetical”	84% thuricin 4A-4	[31], and this study
<i>Bacillus thuringiensis</i> serovar IS5056 (V)	NC_020394	“Hypothetical”	“Hypothetical”	82% thuricin 4A-4	[31], and this study
<i>Bacillus thuringiensis</i> YBT 1518 (I)	NC_022873	“Antibiotic protein, putative”	“Gallidermin/nisin family lantibiotic”	53% gallidermin	This study
<i>Bacillus thuringiensis</i> YBT 1518 (II)	NC_022873	“Antibiotic protein, putative”	“Gallidermin/nisin family lantibiotic”	51% gallidermin	This study
<i>Geobacillus thermoleovorans</i> CCB_US3_UF5	NC_016593	Unannotated	Partly inaccurately annotated gene	79% geobacillin I	This study
<i>Paenibacillus polymyxa</i> CR1	NC_023037	“Subtilin lantibiotic” ⁴	“Hypothetical”	94% paenilan	[36, 37], and this study
<i>Paenibacillus polymyxa</i> E681	NC_014483	“Subtilin lantibiotic” ⁴	“Gallidermin/nisin family lantibiotic”	100% paenilan	[36], and this study
<i>Paenibacillus polymyxa</i> M1 (I)	NC_017542	Unannotated	“Hypothetical”	64% paenilan	This study
<i>Paenibacillus polymyxa</i> M1 (II)	NC_017542	Unannotated	“Hypothetical”	96% paenilan	This study
<i>Paenibacillus polymyxa</i> SC2 (I)	NC_014622	Partly inaccurately annotated gene	“Hypothetical”	64% paenilan	[30], and this study
<i>Paenibacillus polymyxa</i> SC2 (II)	NC_014622	“Subtilin lantibiotic” ⁴	“Hypothetical”	96% paenilan	This study
<i>Staphylococcus aureus</i> 11819-97 (II)	NC_017351	Unannotated	“Gallidermin/nisin family lantibiotic”	83% BsaA2	This study
<i>Staphylococcus aureus</i> COL (II)	NC_002951	Unannotated	“Gallidermin/nisin family lantibiotic”	79% BsaA2	This study
<i>Staphylococcus aureus</i> ED133 (I)	NC_017337	“Hypothetical”	“Gallidermin/nisin family lantibiotic”	100% BacCH91	This study
<i>Staphylococcus aureus</i> ED133 (II)	NC_017337	“Hypothetical”	“Gallidermin/nisin family lantibiotic”	85% BsaA2	[30], and this study
<i>Staphylococcus aureus</i> M1 (II)	NC_021059	Unannotated	“Gallidermin/nisin family lantibiotic”	83% BsaA2	This study
<i>Staphylococcus aureus</i> MSSA476 (I)	NC_002953	“Putative	“Gallidermin/nisin	100% BsaA2	[8], and this

		isochorismatase"	family lantibiotic"		study
<i>Staphylococcus aureus</i> NCTC 8325 (II)	NC_007795	Unannotated	Unannotated	83% BsaA2	This study
<i>Staphylococcus aureus</i> RF122 (I)	NC_007622	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	100% BacCH91	[15], and this study
<i>Staphylococcus aureus</i> RF122 (II)	NC_007622	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	87% identical to BsaA2	[8], and this study
<i>Staphylococcus aureus</i> T0131 (II)	NC_017347	Unannotated	"Gallidermin/nisin family lantibiotic"	83% BsaA2	This study
<i>Staphylococcus aureus</i> USA300 FPR3757 (II)	NC_007793	Unannotated	"Gallidermin/nisin family lantibiotic"	83% BsaA2	This study
<i>Staphylococcus aureus</i> USA300_TCH1516 (II)	NC_010079	Unannotated	"Gallidermin/nisin family lantibiotic"	83% BsaA2	This study
<i>Staphylococcus aureus</i> Z172 (II)	NC_022604	Unannotated	"Gallidermin/nisin family lantibiotic"	83% BsaA2	This study
<i>Streptococcus intermedius</i> B196	NC_022246	Unannotated	"Hypothetical"	No hits	This study
<i>Streptococcus intermedius</i> C270	NC_022237	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	81% nisin F	This study
<i>Streptococcus pyogenes</i> MGAS9429	NC_008021	Unannotated	"Gallidermin/nisin family lantibiotic"	100% streptin	This study
<i>Streptococcus suis</i> JS14	NC_017618	Unannotated	"Gallidermin/nisin family lantibiotic"	100% suicin 90-1330	This study
<i>Streptococcus suis</i> SC070731	NC_020526	Unannotated	"Gallidermin/nisin family lantibiotic"	100% suicin 90-1330	This study

¹ The lanthipeptide aa sequence is presented in Table S1.

² Identity % to experimentally proved lanthipeptides is based on BAGEL BLAST, literature, or both.

³ The cited references represent studies where the aa sequence of the respective lanthipeptide (or its core peptide) was determined either based on an *in silico* analysis or experimental investigation.

⁴ Incorrectly annotated as subtilin. The aa sequences of the lanthipeptides do not match with that of the experimentally confirmed "lantibiotic subtilin" (AAB91589; WP_003220055) produced by *B. subtilis*.

Table S3. BLASTp results of the proteins coded by the Z-geobacillin biosynthesis gene cluster of *Geobacillus* sp. strain ZGt-1. blastp was run against the non-redundant protein sequences (nr) database.

Protein	Gene id ¹	Protein description	Accession number	Identity	Query Coverage	E-value
ZgeoA	6_34	MULTISPECIES: gallidermin/nisin family lantibiotic [Bacillaceae]	WP_011229795	100%	100%	2e-31
ZgeoB	6_44	MULTISPECIES: lantibiotic biosynthesis protein SpaB [<i>Geobacillus</i>]	WP_014194788	100%	100%	0.0
ZgeoT	6_45	MULTISPECIES: ABC transporter ATP-binding protein [<i>Geobacillus</i>]	WP_025038888	100%	35% ²	0.0
ZgeoC	6_46	MULTISPECIES: lanthionine synthetase C family protein [<i>Geobacillus</i>]	WP_020279663	100%	100%	0.0
ZgeoR	6_47	Two component transcriptional regulator, winged helix [<i>Geobacillus thermoleovorans</i> CCB_US3_UF5]	AEV17716.1 ³	100%	100%	3e-167
ZgeoK	6_48	MULTISPECIES: sensor histidine kinase [<i>Geobacillus</i>]	WP_011229815	100%	100%	0.0
ZgeoI	6_49	MULTISPECIES: NisI/SpaI family lantibiotic immunity lipoprotein [<i>Geobacillus</i>]	WP_025039376	100%	100%	8e-102
ZgeoG	6_50	MULTISPECIES: lantibiotic immunity ABC transporter MutG family permease subunit [Bacillaceae]	WP_004888811 ⁴	100%	100%	1e-170
ZgeoE	6_52	lantibiotic ABC transporter permease [<i>Geobacillus thermoleovorans</i> B23]	GAJ60336 ⁵	100%	100%	2e-175
ZgeoF	6_53	MULTISPECIES: lantibiotic ABC transporter ATP-binding protein [<i>Geobacillus</i>]	WP_008881431	100%	100%	2e-162

¹ Gene identity is presented as contig, and running number of the predicted genes within the contig. The contig numbers are the same as found in the NCBI genome record (LDPD01000000).

² The low matching identity is explained by the 3288 undetermined nucleotides in the gene sequence.

³ The protein hit has 5 more amino acids at the N-terminal as compared to the antiSMASH 3.0-predicted lanthipeptide.

⁴ The protein hit has 20 more amino acids at the N-terminal as compared to the antiSMASH 3.0-predicted lanthipeptide.

⁵ The protein hit has 3 more amino acids at the N-terminal as compared to the antiSMASH 3.0-predicted lanthipeptide.

Table S4. Annotation of the proteins coded by the lanthipeptide biosynthesis gene cluster of *G.kaustophilus* HTA426. As described by NCBI, protein accession numbers with the WP_ prefix represent “non-redundant RefSeq protein records” that are “found in RefSeq genomes from multiple species”. (~) denotes that the presented protein is similar to the corresponding lanthipeptide cluster protein. Locus-tags containing underscores represent annotations of the RefSeq genome records, while those without underscores were derived from the annotations of the original genome records.

Protein coded by a member of the lanthipeptide gene cluster	Accession number	Locus-tag
LanAI	BAD74571 WP_011229795	GK0286 GK_RS18085
LanAII	BAD74579 ~WP_008881441	GK0294 GK_RS01950
LanB	BAD74585 ~WP_014194788	GK0300 GK_RS18110
LanBII	~WP_014194788	Not annotated on the original genome In Refseq: GK_RS01970 ¹
LanC	BAD74589 WP_041467776	GK0304 GK_RS01985
LanT	BAD74588 WP_011229812	GK0303 GK_RS01980
LanK	BAD74592 WP_011229815	GK0307 GK_RS01995
LanR	BAD74591 WP_004888808	GK0306 GK_RS01990
LanI	BAD74594 ~WP_008881434	GK0309 GK_RS02005
LanF	BAD74597 WP_008881431	GK0312 GK_RS02020
LanE	BAD714596 WP_008881432	GK0311 GK_RS02015
LanG	BAD74595 WP_004888811	GK0310 GK_RS02010

¹ The position of the gene annotated on the RefSeq genome differs from that predicted by antiSMASH 3.0. On the RefSeq genome, the position of the gene is (328676..329895), while the position predicted by antiSMASH 3.0 is (328679..329833).

Table S5. Annotation of the proteins coded by the lanthipeptide biosynthesis gene cluster of *G. thermodenitrificans* NG80-2. As described by NCBI, protein accession numbers with the WP_ prefix represent “non-redundant RefSeq protein records” that are “found in RefSeq genomes from multiple species”. Locus-tags containing “RS” represent annotations of the RefSeq genome records, while the others were derived from the annotations of the original genome records.

Protein coded by a member of the lanthipeptide gene cluster	Accession number	Locus-tag
GeoAI	ABO65649	GTNG_0265
	WP_008881441	GTNG_RS18005
GeoB	ABO65650	GTNG_0266
	WP_011886696	GTNG_RS01590
GeoC	ABO65652	GTNG_0268
	WP_011886697	GTNG_RS01600
GeoTI	ABO65651	GTNG_0267
	WP_008881438	GTNG_RS01595
GeoK	ABO65654	GTNG_0270
	WP_011886698	GTNG_RS01610
GeoR	ABO65653	GTNG_0269
	WP_008881436	GTNG_RS01605
GeoI	ABO65655	GTNG_0271
	WP_011886699	GTNG_RS01615
GeoF	ABO65658	GTNG_0274
	WP_011886702	GTNG_RS01630
GeoE	ABO65657	GTNG_0273
	WP_011886701	GTNG_RS01625
GeoG	ABO65656	GTNG_0272
	WP_041264183	GTNG_RS01620 ¹

¹ The position of the gene annotated on the RefSeq genome differs from that annotated on the original genome and predicted by antiSMASH 3.0. On the RefSeq genome, the position of the gene is complement (309209..309997), while based on the original genome annotation, as well as the antiSMASH 3.0, the position of the gene is complement (309209..309823).

Table S6. The set of previously described lanthipeptides that have supported the predictions of antiSMASH.

Bacterial species, strain (lanthipeptide reference number)	Refseq genome accession number	Annotation of the predicted lanthipeptide		Matching lanthipeptide ¹ (100% identical)	Supportive reference ²
		Original genome record	RefSeq genome record		
<i>Bacillus clausii</i> KSM-K16	NC_006582	Unannotated	"Gallidermin/nisin family lantibiotic"	- ³	NCBI
<i>Bacillus megaterium</i> QM B1551 (I-II)	NC_014023	"Antibiotic protein, putative"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI, and [31]
<i>Bacillus subtilis</i> BSn5	NC_014976	"Hypothetical"	"Hypothetical"	Subtilomycin	BAGEL BLAST, and [32,33]
<i>Bacillus subtilis</i> spizizenii TU-B-10	NC_016047	Unannotated	"Gallidermin/nisin family lantibiotic"	Entianin	NCBI, BAGEL BLAST, and [47]
<i>Bacillus subtilis</i> spizizenii W23	NC_014479	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	Subtilin	NCBI, and BAGEL BLAST
<i>Bacillus thuringiensis</i> serovar IS5056 (I-V)	NC_020394	"Hypothetical"	"Hypothetical"	Lanthipeptides of thuricin 4A cluster	[31]
<i>Bacillus thuringiensis</i> YBT-1518 (I-II)	NC_022873	"Antibiotic protein, putative"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI
<i>Geobacillus kaustophilus</i> HTA426 (I)	NC_006510	"Lantibiotic precursor"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI, [15], and [34]
<i>Geobacillus kaustophilus</i> HTA426 (II)	NC_006510	"Lantibiotic precursor"	"Lantibiotic nisin-A"	- ³	NCBI, and [15]
<i>Geobacillus thermodenitrificans</i> NG80_2	NC_009328	"Lantibiotic antimicrobial precursor peptide"	"Gallidermin/nisin family lantibiotic"	100% geobacillin I	NCBI, and [34]
<i>Lactococcus lactis</i> CV56	NC_017486	"Nisin precursor NisinA"	"Gallidermin/nisin family lantibiotic"	Nisin A	NCBI, BAGEL BLAST, [15], and [35]
<i>Lactococcus lactis</i> IO	NC_020450	"Lantibiotic antimicrobial precursor"	"Gallidermin/nisin family lantibiotic"	Nisin Z	NCBI, and BAGEL BLAST
<i>Paenibacillus polymyxa</i> E681	NC_014483	"Lantibiotic paenilan precursor peptide"	"Gallidermin family protein"	Paenilan	NCBI, and [36]
<i>Staphylococcus aureus</i> 11819-97 (I)	NC_017351	"Lantibiotic gallidermin"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> Bmb9393 (I)	NC_021670	"Gallidermin/nisin family lantibiotic"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> COL (I)	NC_002951	"Lantibiotic epidermin precursor EpiA"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> ED133 (I)	NC_017337	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	BacCH91	NCBI, BAGEL BLAST, and [81]
<i>Staphylococcus aureus</i> ED133 (II)	NC_017337	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI
<i>Staphylococcus aureus</i> LGA251 (I-II)	NC_017349	"Lantibiotic precursor"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI
<i>Staphylococcus aureus</i> M1 (I)	NC_021059	"Lanthionine precursor peptide LanA"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> MSSA476 (I)	NC_002953	"Putative isochorismatase"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> MW2 (I)	NC_003923	"Hypothetical protein, similar to gallidermin"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST

		precursor"			
<i>Staphylococcus aureus</i> NCTC 8325 (I)	NC_007795	"Gallidermin superfamily epiA, putative"	"Gallidermin superfamily epiA, putative"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> Newman (I)	NC_009641	"Lantibiotic precursor"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> RF122 (I)	NC_007622	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	BacCH91	NCBI, [15], and [81]
<i>Staphylococcus aureus</i> RF122 (II)	NC_007622	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI
<i>Staphylococcus aureus</i> T0131 (I)	NC_017347	"Lantibiotic"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> TW20 (I)	NC_017331	"Lantibiotic precursor"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> USA300 FPR3757 (I)	NC_007793	"Lantibiotic epidermin biosynthesis protein EpiA"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> USA300_TCH1516 (I)	NC_010079	"Lantibiotic epidermin EpiA"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> VC40 (I)	NC_016912	"Lantibiotic epidermin biosynthesis protein EpiA"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Staphylococcus aureus</i> Z172 (I)	NC_022604	"Lanthionine precursor peptide LanA"	"Gallidermin/nisin family lantibiotic"	BsaA2	NCBI, and BAGEL BLAST
<i>Streptococcus</i> <i>intermedius</i> C270	NC_022237	"Hypothetical"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI
<i>Streptococcus</i> <i>pasteurianus</i> ATCC 43144	NC_015600	"Nisin U lantibiotic"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI
<i>Streptococcus pyogenes</i> MGAS6180	NC_007296	"Lantibiotic"	"Gallidermin/nisin family lantibiotic"	Streptin	NCBI, BAGEL BLAST, [15], and [84]
<i>Streptococcus pyogenes</i> MGAS9429	NC_008021	"Lantibiotic biosynthesis sensor protein"	"Gallidermin/nisin family lantibiotic"	Streptin	NCBI, BAGEL BLAST, and [84]
<i>Streptococcus pyogenes</i> MGAS10270	NC_008022	"Lantibiotic srtA precursor"	"Gallidermin/nisin family lantibiotic"	Streptin	NCBI, BAGEL BLAST, [15], and [84]
<i>Streptococcus pyogenes</i> MGAS10750	NC_008024	"Lantibiotic srtA precursor"	"Gallidermin/nisin family lantibiotic"	- ³	NCBI, and [15]
<i>Streptococcus suis</i> JS14	NC_017618	Unannotated	"Gallidermin/nisin family lantibiotic"	Suicin 90-1330	NCBI, and [96]
<i>Streptococcus suis</i> SC070731	NC_020526	Unannotated	"Gallidermin/nisin family lantibiotic"	Suicin 90-1330	NCBI, and [96]

¹ Identity to experimentally verified lanthipeptide. ² Supportive reference represents literature study(ies) where the respective lanthipeptide was identified, annotation(s) of the genome deposited in NCBI (RefSeq/GenBank), and/or the results of BAGEL BLAST. ³ (-) indicates that among experimentally verified lanthipeptides, there is not any that has 100% identity to the respective lanthipeptide.

S.2.3.1*. Identification of *Bacillus*-associated lanthipeptide gene clusters

S.2.3.1.1. *B. clausii* KSM-K16

The nucleotide (nt) sequence of the gene coding for the antiSMASH-predicted lanthipeptide (Table S1) is 100% identical to “ABC_RS22115” annotated on the RefSeq as coding for a class-I lanthipeptide (Table S2). Thus, the RefSeq annotation supports our analysis (Table S6). The core sequence of the predicted lanthipeptide was briefly mentioned by van Heel *et al*, 2016 [30]. The blastp analysis indicated that the predicted lanthipeptide is 56% identical to clausin (Table 2). We noticed that clausin has not been reported in BAGEL database; therefore, we recommend including it as a class-I lanthipeptide.

S.2.3.1.2. *B. megaterium* QM B1551

The aa sequences of the two antiSMASH-predicted lanthipeptides are identical, but the nt sequences of their two coding-genes are not (Table S1). The nt sequence of the gene coding for lanthipeptide (I) showed 100% identity to “BMQ_RS27575”, and that coding for lanthipeptide (II) showed 100% identity to “BMQ_RS27580”. These genes have been annotated in the RefSeq record as coding for class-I lanthipeptides (Table S2). Both lanthipeptides showed 56% identity to gallidermin as indicated by BAGEL BLAST (Table 2). It is noteworthy that using BAGEL3, Xin *et al*, 2015 identified one of the putative lanthipeptides [31] presented in the current study. Thus, the RefSeq annotation and results of Xin *et al*, 2015 support our analysis (Table S6).

S.2.3.1.3. *B. subtilis spizizenii* DSM 15029T (TU-B-10)

Our analysis has indicated points that are worth noting, as discussed below.

Using antiSMASH analysis, the gene coding for entianin was predicted (Table S1). In addition to the original and RefSeq genome sequence records of the strain, there is also a record on the entianin gene cluster that belongs to the same strain, separately deposited in NCBI under the accession number (HQ871873). We noticed that the entianin-coding gene has been annotated on the RefSeq as “GYO_RS39160” that codes for a class-I lanthipeptide. On the other hand, the coding gene in the entianin cluster record has been annotated and named as “*etnS*” coding for “EtnS” (AEK64494). The nt sequences of “GYO_RS39160” and “*etnS*” are identical and they both code for the same peptide. The two separate records related to the lanthipeptide resulted in having different tags/names for the same gene and its coded lanthipeptide. The entianin cluster (HQ871873) of the strain was sequenced by Fuchs *et al* and deposited in NCBI in 2011 [47] and the designation of the strain displayed for the record is DSM 15029^T. The antimicrobial activity of the entianin cluster has been experimentally proved [47]. On the other hand, the original genome (CP002905) was sequenced by Earl *et al* in 2012 [46], whereas the RefSeq genome was annotated in 2017, and the stated designation of the strain for both genome records is (TU-B-10). Accordingly, a re-evaluation of naming the gene and its lanthipeptide product in the RefSeq needs to be considered to match the description of the experimentally-verified entianin cluster (HQ871873). Moreover, since the RefSeq genome record shows only one of the designations of the strain; TU-B-10 while the record of the entianin cluster shows the other designation; DSM 15029^T, we recommend to present both designations in each record in order to avoid the confusion. Interestingly, the antibacterial activity of entianin has been experimentally confirmed; it is highly active against several Gram-positive bacteria [47].

S.2.3.1.4. *B. subtilis spizizenii* W23

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (Table S1) is 100% identical to “BSUW23_RS16845” annotated on the RefSeq as coding for a class-I

lanthipeptide (Table S2). The RefSeq annotation thus supports our analysis (Table S6). BAGEL BLAST indicated that the predicted lanthipeptide is 100% identical to subtilin (Table 2; Table S2) and 95% to entianin.

S.2.3.1.5. *B. thuringiensis* YBT-1518

The antiSMASH analysis indicated that the class-I lanthipeptide cluster coded on the chromosome of strain YBT-1518 has two putative genes coding for two different lanthipeptides. The nt sequence of the gene coding for the predicted lanthipeptide (I) (Table S1) is 100% identical to “YBT1518_RS19670”, and that of the gene coding for lanthipeptide (II) is 100% identical to “YBT1518_RS19675”. Both genes have been annotated on the RefSeq genome as coding for class-I lanthipeptides. The RefSeq annotation thus supports our analysis (Table S6). Based on the analysis using BAGEL BLAST, the predicted lanthipeptides I and II are 53% and 51% identical to gallidermin, respectively (Table 2).

S.2.3.2. Identification of *Geobacillus*-associated lanthipeptide gene clusters

S.2.3.2.1. *G. kaustophilus* HTA426

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “GK_RS18085” annotated on the RefSeq genome as coding for a class-I lanthipeptide. Therefore, the genome annotation supports our results (Table S6). BAGEL BLAST showed that lanthipeptide (I) is 91% identical to the experimentally characterized antimicrobially-active lanthipeptide; geobacillin I (ABO65649) produced by *G. thermodenitrificans* NG80-2 [34]. Garg *et al* 2012 reported the same lanthipeptide for strain HTA426 [34], and this in turn further supports our analysis (Table 2, Table S6).

S.2.3.3. Identification of *Lactococcus*-associated lanthipeptide gene clusters

S.2.3.3.1. *L. lactis* strains CV56 and IO-1

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide of strain CV56 (Table S1) is 100% identical to “CVCAS_RS03115”, and that coding for the predicted lanthipeptide of strain IO-1 (Table S1) is 100% identical to “LILO_RS03015”. Both genes have been annotated in the RefSeq genome records as coding for class-I lanthipeptides. The RefSeq annotations thus support our analysis (Table S6). BAGEL BLAST showed that the predicted lanthipeptide of strain CV56 is 100% identical to nisin A (Table 2). Our results agree with those of Gao *et al*, 2011 [35] and Marsh *et al*, 2010 [15] (Table S6). On the other hand, BAGEL BLAST showed that the predicted lanthipeptide of strain IO-1 is 100% identical to nisin Z (Table 2).

S.2.3.4. Identification of *Paenibacillus*-associated lanthipeptide gene clusters

S.2.3.4.1. *P. polymyxa* E681

The nt sequence of the gene coding for the predicted lanthipeptide (Table S1) is 100% identical to “PPE_RS07020”, which has been annotated on the RefSeq as coding for a class-I lanthipeptide (Table S2). The aa sequence of the characterized class-I lanthipeptide; named paenilan produced by strain E681 [36] is identical to that inferred in our study for the same strain (Table S2). This in turn supports our analysis (Table S6). The antibacterial activity of paenilan has been experimentally verified by Park *et al* [36].

S.2.3.5. Identification of *Staphylococcus*-associated lanthipeptide gene clusters

S.2.3.5.1. *S. aureus* 11819-97

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “MS7_RS09745” annotated on the RefSeq genome as coding for a class-I lanthipeptide. The annotation of the genome thus supports our analysis (Table S6). Moreover, BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BsaA2 (Table 2). The antimicrobial activity of BsaA2 has been experimentally verified in other *S. aureus* strains by Daly *et al*, 2010 [8].

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (II) (Table S1) is 100% identical to “MS7_RS09750” annotated on the RefSeq genome as coding for a class-I lanthipeptide (Table S2). However, in RefSeq and consequently also in antiSMASH, the start codon has been translated into an incorrect aa in the presented aa sequence. The nt sequence presented in RefSeq and antiSMASH showed the first codon as “TTG”, which codes for the aa leucine (L) and not the reported methionine (M). Therefore, we edited the start aa of the predicted lanthipeptide (II) into the correct one (L) (Table S1). BAGEL BLAST showed that lanthipeptide (II) is 83% identical to BsaA2 (Table 2).

S.2.3.5.2. *S. aureus* COL

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “SACOL_RS09635” annotated on the RefSeq genome as coding for a class-I lanthipeptide. The annotation of the genome thus supports our analysis (Table S6). Moreover, BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BsaA2 (Table 2). Daly *et al* 2010 reported strain COL as a putative producer of BsaA2 [8].

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (II) (Table S1) is 100% identical to “SACOL_RS09645” annotated on the RefSeq genome as coding for a class-I lanthipeptide (Table S2). Again, we made the same correction of the aa as we did for lanthipeptide (II) of strain 11819-97 (Table S1). BAGEL BLAST showed that lanthipeptide (II) is 79% identical to BsaA2 (Table 2).

S.2.3.5.3. *S. aureus* ED133

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “SAOV_RS09475”, and that of the other predicted gene coding for lanthipeptide (II) (Table S1) is 100% identical to “SAOV_RS09490”. Both genes have been annotated on the RefSeq genome as coding for a class-I lanthipeptide (Table S2), which in turn, supports our analysis (Table S6). Moreover, BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BacCH91 produced by *S. aureus* strain CH91 [81] (Table 2; Table S2). On the other hand, lanthipeptide (II) is 85% identical to BsaA2 (Table 2; Table S2). Only the core sequence of lanthipeptide (II) was very briefly mentioned by van Heel *et al*, 2016 and was inferred based on an *in silico* analysis [30].

S.2.3.5.4. *S. aureus* M1

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “BN843_RS09660” annotated on the RefSeq genome as coding for a class-I lanthipeptide. The annotation of the genome thus supports our analysis (Table S6). Moreover, BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BsaA2 (Table 2).

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (II) (Table S1) is 100% identical to “BN843_RS09665” annotated on the RefSeq as coding for a class-I lanthipeptide (Table S2). Here as well, we edited the start aa of the predicted lanthipeptide (II) of

strain M1 into the correct one (L) instead of M (Table S1). BAGEL BLAST showed that lanthipeptide (II) is 83% identical to BsaA2 (Table 2).

S.2.3.5.5. *S. aureus* MSSA476

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “SAS1746” which has been annotated on the RefSeq genome as coding for a class-I lanthipeptide (Table S2). The RefSeq genome annotation thus supports our analysis (Table S6). BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BsaA2 (Table 2). Daly *et al.*, 2010 reported strain MSSA476 as a putative producer of BsaA2 [8].

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (II) (Table S1) is 100% identical to “SAS_RS09300” annotated on the RefSeq genome as coding for a class-I lanthipeptide. Here as well, we edited the start aa of the predicted lanthipeptide (II) of strain MSSA476 into the correct one (L) instead of M (Table S1). BAGEL BLAST showed that lanthipeptide (II) is 83% identical to BsaA2 (Table 2).

S.2.3.5.6. *S. aureus* MW2

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “MW_RS09420” annotated on the RefSeq as coding for a class-I lanthipeptide. Therefore, the RefSeq annotation supports our analysis (Table S6). BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BsaA2 (Table 2). Daly *et al.*, 2010 reported strain MW2 as a putative producer of BsaA2 [8].

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (II) (Table S1) is 100% identical to “MW_RS09425” annotated on the RefSeq as coding for a class-I lanthipeptide. Here as well, we edited the start aa of the predicted lanthipeptide (II) of strain MW2 into the correct one (L) instead of M (Table S1). BAGEL BLAST showed that lanthipeptide (II) is 83% identical to BsaA2 (Table 2).

S.2.3.5.7. *S. aureus* NCTC 8325

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “SAOUHSC_01953” annotated on the RefSeq genome as coding for a class-I lanthipeptide. The annotation of the genome thus supports our analysis (Table S6). BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BsaA2 (Table 2). Daly *et al.*, 2010 reported strain NCTC 8325 as a producer of BsaA2 [8].

S.2.3.5.8. *S. aureus* RF122

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “SAB_RS08990” annotated on the RefSeq genome as coding for a class-I lanthipeptide (Table S2), which thus supports our analysis (Table S6). Based on the analysis using BAGEL BLAST, lanthipeptide (I) is 100% identical to BacCH91, and 81% identical to BsaA2. Moreover, our results agree with the findings of Marsh *et al.*, 2010 [15].

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (II) (Table S1) is 100% identical to “SAB_RS08995” annotated on the RefSeq as coding for a class-I lanthipeptide (Table S2). Thus, the RefSeq annotation supports our analysis (Table S6). Based on the analysis using BAGEL BLAST, lanthipeptide (II) is 87% identical to BsaA2. Daly *et al.*, 2010 also reported strain RF122 as a producer of variants of Bsa [8].

S.2.3.5.9. *S. aureus* T0131

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) (Table S1) is 100% identical to “SAT0131_RS09535” annotated on the RefSeq genome as coding for a class-I lanthipeptide. The annotation of the genome thus supports our analysis (Table S6). BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BsaA2 (Table 2).

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (II) (Table S1) is 100% identical to “SAT0131_RS09540” annotated on the RefSeq genome as coding for a class-I lanthipeptide. Here as well, we edited the start aa of the predicted lanthipeptide (II) of strain T0131 into the correct one (L) instead of M (Table S1). BAGEL BLAST showed that lanthipeptide (II) is 83% identical to BsaA2 (Table 2).

S.2.3.5.8.10. *S. aureus* strains USA300 FPR3757 and USA300_TCH1516

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (I) of strain USA300_FPR3757 (Table S1) is 100% identical to “SAUSA300_RS09670”. For strain USA300_TCH1516, the nt sequence of the gene coding for the predicted lanthipeptide (I) (Table S1) is 100% identical to “USA300HOU_RS09705”. Each of “SAUSA300_RS09670” and “USA300HOU_RS09705” has been annotated on the RefSeq genome of the respective strain as coding for a class-I lanthipeptide. The annotation of the genome thus supports our analysis (Table S6). BAGEL BLAST showed that each of the predicted lanthipeptides is 100% identical to BsaA2 (Table 2). Daly *et al* (2010) reported that the two strains are putative producers of BsaA2 [8].

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (II) of strain USA300_FPR3757 (Table S2) is 100% identical to “SAUSA300_RS09675”, and that of the gene coding for the predicted lanthipeptide (II) of strain “USA300_TCH1516” (Table S2) is identical to “USA300HOU_RS09710”. Here as well, we edited the start aa of the predicted lanthipeptide (II) of both strains into the correct one (L) instead of M (Table S1). BAGEL BLAST indicated that lanthipeptide (II) of each of the strains is 83% identical to BsaA2 (Table 2).

S.2.3.5.8.11. *S. aureus* Z172

The nt sequence of the gene coding for lanthipeptide (I) (Table S1) is 100% identical to “SAZ172_RS09800” annotated on the RefSeq genome as coding for a class-I lanthipeptide. The annotation of the genome thus supports our analysis (Table S6). BAGEL BLAST showed that lanthipeptide (I) is 100% identical to BsaA2 (Table 2).

The nt sequence of the antiSMASH-predicted lanthipeptide (II) (Table S1) is 100% identical to “SAZ172_RS09810” annotated on the RefSeq genome as coding for a class-I lanthipeptide (Table S2). Here as well, we edited the start aa of the predicted lanthipeptide (II) of strain Z172 into the correct one (L) instead of M (Table S1). BAGEL BLAST showed that lanthipeptide (II) is 83% identical to BsaA2 (Table 2).

S.2.3.5.8.12. other strains

The aa sequence of lanthipeptide (I) is identical in strains Bmb 9393, Newman, TW20, and VC40 (Table S1). The coding genes have been annotated in the RefSeq genome record of strain Bmb 9393, as “SABB_RS10200”; strain Newman as “NWMN_RS09635”; strain TW20, as “SATW20_RS09830”; and strain VC40, as “SAVC_RS08915”. Each of the genes has been annotated as coding for a class-I lanthipeptide. Therefore, the genome annotations support our results (Table S6). Moreover, the predicted lanthipeptides showed 100% identity to the BsaA2, as indicated by BAGEL BLAST (Table 2).

The aa sequence of lanthipeptide (II) in all these strains is identical (Table S1). The coding genes have been annotated in the RefSeq genome records as coding for class-I lanthipeptides. The

nt sequence of the predicted gene coding for lanthipeptide (II) of strain Bmb 9393 is 100% identical to "SABB_RS10205", that of strain Newman is 100% identical to "NWMN_RS09640", that of strain TW20 is 100% identical to "SATW20_RS09835", and that of strain VC40 is 100% identical to "SAVC_RS08920". Here as well, we edited the start aa of the predicted lanthipeptide (II) of each of these strains into the correct one (L) instead of M (Table S1). The predicted lanthipeptides showed 83% identity to the BsaA2, as indicated by BAGEL BLAST (Table 2).

Furthermore, our analysis indicated that *S. aureus* strain LGA251 harbours a class-I lanthipeptide cluster coding for two different lanthipeptides. The putative coding genes of lanthipeptides (I) and (II) (Table S1) have been annotated on the RefSeq genome as "SARLGA251_RS09195" and "SARLGA251_RS09200", each of which has been annotated as coding for a class-I lanthipeptide. Therefore, RefSeq annotations support our analysis (Table S6). BAGEL BLAST indicated that lanthipeptides (I) and (II) showed 81% and 85% identity, respectively, to the BsaA2 of *S. aureus* (Table 2).

S.2.3.6. Identification of *Streptococcus*-associated lanthipeptide gene clusters

S.2.3.6.1. *S. intermedius* C270

The nt sequence of the gene coding for the antiSMASH-predicted lanthipeptide (Table S1) is 100% identical to "SII_RS05025" annotated on the RefSeq as coding for a class-I lanthipeptide (Table S2). The RefSeq annotation, in turn, supports our analysis (Table S6). Based on the analysis using BAGEL BLAST, the predicted lanthipeptide is 81% identical to Nisin F which is produced by *L. lactis* (Table 2; Table S2).

S.2.3.6.2. *S. pasteurianus* ATCC 43144

The nt sequence of the gene coding for the lanthipeptide (Table S1) is 100% identical to "SGPB_RS10680" annotated on the RefSeq as coding for a class-I lanthipeptide. The genome annotation, thus, supports our analysis (Table S6) and also confirms the results of Lin *et al.*, 2011. Based on the analysis using BAGEL BLAST, the predicted lanthipeptide is 91% similar to nisin U (Table 2). These results indicate that the antimicrobial potential of the lanthipeptide in strain ATCC 43144 should be of interest to investigate.

S.2.3.6.3. *S. pyogenes* MGAS9429

The nt sequence coding for the antiSMASH-predicted lanthipeptide (Table S1) is 100% identical to "MGAS9429_RS04535" annotated on the RefSeq genome as coding for a class-I lanthipeptide (Table S2). This in turn supports our analysis (Table S6). Moreover, BAGEL BLAST indicated that the predicted lanthipeptide is 100% identical to streptin (Table 2). Therefore, we suggest considering *S. pyogenes* strain MGAS2096 as another potential streptin-producer.

S.2.3.6.4. Other strains of *S. pyogenes*

The aa sequences of the lanthipeptides of strains MGAS6180, MGAS10270 and MGAS10750 are identical (Table S1). The nt sequence of the gene coding for the lanthipeptide of strain MGAS6180 (Table S1) is 100% identical to "M28_RS03970". That of the gene coding for the lanthipeptide of strain MGAS10270 (Table S1) is 100% identical to "MGAS10270_RS04550". Finally, the nt sequence of the gene coding for the lanthipeptide of strain MGAS10750 (Table S1) is identical to "MGAS10750_RS04710". These 3 genes have been annotated in the RefSeq genome records as coding for class-I lanthipeptides. Therefore, our results are supported by the annotation of the genome records of the three strains (Table S6). BAGEL BLAST indicated that the the lanthipeptides of strains MGAS6180 and MGAS10270 are 100% identical to streptin, while

that of strain MGAS10750 is 98% identical to streptin (Table 2). The results of antiSMASH analysis agree with the findings of Marsh *et al.*, 2010 [15].

S.2.3.6.5. *S. suis* strains JS14 and SC070731

The nt sequences of the gene coding for the antiSMASH-predicted lanthipeptide of strain JS14, and that coding for the lanthipeptide of strain SC070731 are identical (Table S1). The coding genes have been annotated as “SSUJS14_RS10730” on the RefSeq genome of strain JS14, and as “NJAUSS_RS10650” on that of strain SC070731. Each of these genes codes for a class-I lanthipeptide. Therefore, the annotation of the RefSeq genome supports our analysis (Table S6). Each of the predicted lanthipeptides is 100% identical to suicin 90-1330 produced by *S. suis* strain 90-1330 [96] (Table 2). We noticed that suicin 90-1330 has not been reported in BAGEL databases. Since the lanthipeptide production potential of strains JS14 and SC070731 has not been investigated, we suggest the two strains as potential lanthipeptide producers.

* The section number corresponds to that in the main text but preceded with an “S.” for “supplementary”.

Additional files

- Software source code is available at <http://130.235.46.10/Lanthipeptides/Software/>
- Graphical illustrations of the antiSMASH output are available at <http://130.235.46.10/Lanthipeptides/>