The Journal of Antibiotics (2016) 69, 712–718

www.nature.com/ja

NOTE

Genome-based survey of nonribosomal peptide synthetase and polyketide synthase gene clusters in type strains of the genus *Microtetraspora*

Hisayuki Komaki[1], Natsuko Ichikawa[2], Tomohiko Tamura[1], Akio Oguchi[2], Moriyuki Hamada[1] and Nobuyuki Fujita[2]

The Journal of Antibiotics (2016) **69,** 712–718; doi:10.1038/ja.2015.139; published online 17 February 2016

Since the discovery of streptomycin from *Streptomyces griseus*, many additional antibiotics have been identified from cultures of the genus *Streptomyces*.[1,2] Consequently, the likelihood of discovering novel secondary metabolites from *Streptomyces* members has dwindled over time and the focus of screening has moved to less studied genera such as non-*Streptomyces* called rare actinomycetes. For example, members of genera within the family *Streptosporangiaceae*, such as *Microbispora*, *Nonomuraea*, *Planobispora* and *Streptosporangium,* have been sources for many novel compounds.[3] The genus *Microtetraspora*, belonging to the family *Streptosporangiaceae*, was once comprised of 21 species, but 16 of them were reclassified into other genera such as *Nonomuraea*, *Thermopolyspora* and *Actinomadura* (Supplementary Figure S1).[4–7] Consequently, the genus *Microtetraspora* now contains only five valid species, *Microtetraspora fusca*, *M. glauca*, *M. malaysiensis*, *M. niveoalba* and *M. tyrrhenii*. However, it was later suggested that *M. tyrrhenii*, based on its morphological characteristics, should be transferred to the genus *Nonomuraea,* but nomenclatural change for the species has not been proposed because the type strain, ATCC 53931, is no longer available at the American Type Culture Collection and there is not alternative collection.[8] Although several bioactive compounds were reportedly discovered from *Microtetraspora* strains, the producing strains have either been reclassified into other genera or are not valid *Microtetraspora* species.[9–11] Therefore, no bioactive compound has been discovered from any valid species in the genus *Microtetraspora*, and so it remains uncertain whether *Microtetraspora* members are useful sources of novel bioactive compounds.

Recent genomic studies of actinomycetes revealed that each actinomycete genome encodes various biosynthetic gene clusters for secondary metabolites and about half to three quarters of these clusters are associated with nonribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) pathways. This suggests that nonribosomal peptides and polyketide compounds are the major secondary metabolites of actinomycetes.[12] Nonribosomal peptides and polyketide compounds often show pharmaceutically useful bioactivities, and many have been developed into various drugs including antibiotics, anticancer agents and immunosuppressants. Therefore, actinomycete strains are often surveyed for the presence, abundance and novelty of NRPS and PKS genes to evaluate these strains as potential secondary metabolite producers.[13–15] In the present study, we sequenced the whole genomes of all type strains from the genus *Microtetraspora* available for public use because no *Microtetraspora* genome sequences have been registered in public databases. Next, we searched for NRPS, type-I PKS (PKS-I) and type-II PKS (PKS-II) gene clusters in the genome sequences, predicted their metabolites and generated information on the novelty and diversity of the NRPS and PKS pathways to assess the potential of members of the genus as secondary metabolite producers.

Genomic DNAs of *M. glauca* NBRC 14761^T, *M. fusca* NBRC 13915^T, *M. malaysiensis* NBRC 100735^T and *M. niveoalba* NBRC 102641^T were prepared directly from liquid-dried cells in ampoules provided from the NBRC (Biological Resource Center, National Institute of Technology and Evaluation, Chiba, Japan) culture collection, using a Qiagen EZ1 tissue kit and an EZ1 Advanced instrument (Qiagen, Germantown, MD, USA). The prepared DNA was sequenced using paired-end sequencing with a MiSeq (Illumina, San Diego, CA, USA). The average sequence redundancy for the four draft genomes ranged from 56.0 to 66.8. The sequence reads were assembled using Newbler v2.6 (454 Life Sciences, Branford, CT, USA) and subsequently finished using GenoFinisher.[16] Coding regions in the draft genome sequences were predicted by Prodigal v2.6.[17] NRPS and PKS gene clusters were determined as previously reported.[13,14] NRPS and PKS-I genes containing only a single domain were excluded from the present analysis, because we considered them atypical in the assembly line theory and focused on multi-domain genes. A BLASTP search was performed using the NCBI Protein BLAST program against the non-redundant protein sequence database. We considered genes of distinct strains to be orthologous when their closest homologs in the BLASTP search were the same, and also when their domain

[1]Biological Resource Center, National Institute of Technology and Evaluation (NBRC), Chiba, Japan and [2]Biological Resource Center, National Institute of Technology and Evaluation (NBRC), Tokyo, Japan
Correspondence: Dr H Komaki, Biological Resource Center, National Institute of Technology and Evaluation (NBRC), 2-5-8 Kazusakamatari, Kisarazu, Chiba 292-0818, Japan.
E-mail: komaki-hisayuki@nite.go.jp
Received 3 August 2015; revised 30 November 2015; accepted 6 December 2015; published online 17 February 2016

Table 1 Genome sequencing and numbers of NRPS and PKS gene clusters in *Microtetraspora* strains

Strain	Source of isolation	Reads (Mb)	No. of scaffolds	Genome size (bp)	G+C content (%)	Accession no.	Number of gene clusters				
							NRPS	PKS/NRPS hybrid	PKS-I	PKS-II	Total
M. glauca NBRC 14761^T	Soil	608.0	62	9,275,754	70.15	BBYL01000000	5	1	0	0	6
M. fusca NBRC 13915^T	Rubber plantation soil	659.6	109	9,024,887	70.93	BBYK01000000	6	2	1	1	10
M. malaysiensis NBRC 100735^T	Soil from primary lowland dipterocarp forest	772.3	53	8,822,209	70.73	BBYJ01000000	4	3	0	0	7
M. niveoalba NBRC 15239^T	Soil	810.5	98	8,469,978	71.65	BBYM01000000	5	2	3	0	10

Abbreviations: NBRC, Biological Resource Center, National Institute of Technology and Evaluation; NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase.

organizations were identical or almost the same. AntiSMASH[18] was used to predict substrates for adenylation (A) domains. Average nucleotide identity according to BLAST (ANIb) was calculated using the JSpecies program.[19]

Phylogenetic positions of the strains studied here, along with strains previously misidentified as *Microtetraspora*, are shown in Supplementary Figure S1. During this study, the draft genome sequence of *M. glauca* NRRL B-3735^T was published (accession no. JOFO00000000.1). However, it is questionable whether the sequences are from *M. glauca* NRRL B-3735^T, because the 16S rRNA gene sequence in JOFO01000118 showed <90% similarity to the sequence of the *M. glauca* type strain (D85490) and, instead, showed high (99%) similarity to *Streptomyces* strains. Therefore, in the present study, we did not analyze the NRPS and PKS gene clusters of JOFO01000000 and focused on only our data. Genome sizes of the four *Microtetraspora* type strains ranged from 8.5 to 9.3 Mb (Table 1), which places them in the middle of the range seen for *Streptomyces* strains (5.0–11.9 Mb) and strains in the family *Streptosporangiaceae* (5.5–13 Mb). We calculated ANIb values among the four strains using their genome sequences in order to confirm that each strain is an independent species, because these species show high 16S rRNA gene sequence similarities (98.02–99.86%) (Supplementary Table S1). Since the ANIb values (86.5–93.4%) were below the threshold (95–96%) corresponding to DNA relatedness value of 70% recommended as the cutoff point for the assignment of bacterial strains to the same species,[19,20] each strain is definitely independent species. The four strains each possessed 6–10 gene clusters for NRPS, PKS/NRPS hybrid, PKS-I and PKS-II pathways. The numbers of gene clusters were smaller (by about one-third to one-half) than the numbers found in *Streptomyces* genomes.[12,14,21] The number of each of the four types of gene clusters found in each strain is listed in Table 1. Unlike in the published genome sequences of typical actinomycetes, no PKS-I gene cluster was present in *M. glauca* NBRC 14761^T or *M. malaysiensis* NBRC 100735^T, and only one PKS-I gene cluster was present in *M. fusca* NBRC 13915^T. This suggests that PKS-I pathways are not abundant in the genus *Microtetraspora*, as compared with other genera of actinomycetes. Table 2 shows details of all clusters found in each genome. Orthologous genes and gene clusters are aligned in the same row of the table. These orthologous genes had the same domain organization and, therefore, their gene clusters should synthesize the same products (shown in the 'Predicted product' column of Table 2). Among the 18 gene clusters (*nrps-1* to *-11*, *pks/nrps-1* to *-3*, *pks-1* to *-3*, *pksII*) identified from the four strains, three (*nrps-1, nrps-2, pks/nrps-1*) were conserved in all strains, five (*nrps-3* to *-6*, *pks-1*) were shared by two or three strains, and ten (*nrps-7* to *-11*, *pks/nrps-2* and *-3*, *pks-2* and *-3*, *pksII*) were strain-specific.

Gene clusters conserved in all four strains. The data in Table 2 suggested that three predicted products (*nrps-1, nrps-2* and *pks/nrps-1*) are common among the four type strains. Based on their module numbers and A domain substrates, *nrps-1* and *-2* were predicted to synthesize a tetrapeptide containing glycine (Gly), lysine (Lys) and aspartic acid (Asp) residues, and a dipeptide containing one Asp residue, respectively. *Pks/nrps-1* is a PKS/NRPS hybrid gene encoding five NRPS modules and one PKS module, suggesting a product composed of five amino acid residues and one polyketide unit. Because the NRPS and PKS-I genes in the *pks/nrps-1* cluster showed high sequence homologies (70–88% identity) with those of *Streptosporangium roseum,* and the domain organization of *Microtetraspora* strains and *S. roseum* is the same, their products are likely to be similar.

Table 2 Open reading frames (ORFs) encoding modular NRPSs and PKSs in NRPS and PKS gene clusters of *Microtetraspora* strains

	ORF no.[a] (size in aa, % identity/similarity to closest homolog)						*Closest homolog*	
Gene cluster	M. glauca *NBRC 14761^T*	M. fusca *NBRC 13915^T*	M. malaysiensis *NBRC 100735^T*	M. niveoalba *NBRC 15239^T*	*Domain organization*	*Predicted product*	*Accession no.*	*Origin*
nrps-1	32-55 (1843, 79/84) 32-57 (526, 70/79) 32-62 (650, 71/78) 32-63 (1047, 75/80)	04-337 (1884, 79/84) 04-339 (528, 70/79) 04-344 (642, 68/75) 04-345 (1052, 73/79)	06-250 (1870, 79/85) 06-252 (532, 70/78) 06-257 (641, 68/75) 06-258 (1059, 73/79)	05-158 (1885, 79/85) 05-160 (525, 68/78) 05-165 (692, 68/75) 05-166 (1090, 71/77)	C/A_{Asp}/T-TE C/T A_{Gly}/T C/A_{Lys}/T	Gly-z-Lys-Asp	WP_030507016 WP_030507014 ACZ89541 WP_036405664	*Microbispora rosea* *M. rosea* *Streptosporangium roseum* DSM 43021 *M. rosea*
nrps-2	13-96[b] (558, 67/76) 13-98 (1855, 58/68)	29-31[b] (514, 67/76) 29-29 (1787, 61/70)	07-152[b] (519, 66/75) 07-154 (1761, 62/71)	19-224[b] (537, 65/75) 19-226 (1786, 60/69)	C/T C/A_{Asp}/T-TE	z-Asp	WP_036322383 WP_012890029	*Microbispora* sp. ATCC PTA-5024 *S. roseum*
nrps-3	15-124 (1637, 58/68) 15-123 (2918, 60/69)	33-23 (1614, 58/67) 33-22 (2831, 62/71)	—	24-83 (1662, 58/67) 24-84 (3045, 56/65)	A/T-C/A/T C/A/T-C/A/T-TE	x-x-x-x	WP_026413360 WP_034485599	*Actinomadura oligospora* *A. oligospora*
nrps-4	—	40-44 (1614, 85/89) 40-45 (1778, 88/92)	10-34 (1614, 85/89) 10-33 (1778, 88/92)	—	A_{Ala}/T-C/A_{Ala}/T C/A/T-TE	Ala-Ala-x	WP_031161416 ACZ86273	*S. roseum* *S. roseum* DSM 43021
nrps-5	01-500 (1121, 60/71) 01-511 (1086, 58/71)	—	—	33-54 (1133, 60/71) 33-43 (1092, 58/71)	C/A_{Cys}/T C/A/T	Cys-x	WP_019872681 WP_030845862	*Salinispora pacifica* *Streptomyces* sp. NRRL F-4474
mrps-6	—	18-24 (1279, 73/81)	—	04-225 (1279, 73/81)	C/A/T	x	ELS56639	*Streptomyces viridochromogenes* Tue57
nrps-7	24-30 (1600, 83/88) 24-29 (2724, 88/92)	—	—	—	A/T-C/A/T C/A_{Asp}/T-C/A/T-TE	x-x-Asp-x	ACZ87027 WP_031168611	*S. roseum* DSM 43021 *S. roseum*
nrps-8	—	*19-1 (>698, 49/57)* 19-3 (2127, 48/59) 19-4 (3297, 45/56) 19-5 (1064, 45/58) 19-6 (2494, 48/59)	—	—	…A/T C/A/T-C/A_{Ser}/T C/A_{Thr}/T-C/A_{Asn}/T-C/A_{Asn}/T C/A/T C/A_{Thr}/T-C/A_{Asn}/T-TE	…x-x-Ser-Thr-Asn-Asn-x-Thr-Asn	EPH43046 EPH43048 BAB69333 WP_045933577 WP_026410862	*Streptomyces aurantiacus* JA 4570 *S. aurantiacus* JA 4570 *Streptomyces avermitilis* *Streptomyces* sp. NRRL B-1568 *A. oligospora*
nrps-9	—	—	10-117[b] (1513, 51/66) 10-126[b] (600, 53/67) 10-131 (1155, 51/63)	—	C/A_{Cys}/MT/T A_{Phe}/T C/A_{Cys}/T	Phe-mCys-Cys	WP_031170325 WP_045704794 WP_044575336	*S. roseum* *Streptomyces rubellomurinus* *Saccharopolyspora spinosa*
nrps-10	—	—	10-9 (1646, 75/82) 10-10 (1807, 91/94)	—	A/T-C/A_{Val}/T C/A/T-TE	x-Val-x	ACZ88334 WP_012892072	*S. roseum* DSM 43021 *S. roseum*
nrps-11	—	—	—	14-149 (2090, 43/56) 14-150 (2093, 45/59) 14-153 (1730, 38/55) 14-155 (3194, 42/56) 14-156 (1094, 45/60) 14-157 (1292, 44/58)	C/A_{Gly}/T-C/A_{Thr}/T C/A_{Trp}/T-C/A/T CoL/T-C/A_{Gly}/T C/A_{Asp}/T-C/A_{Val}/T-C/A/T C/A_{Ser}/T C/A_{Gly}/T-TE	y-Gly-Gly-Thr-Trp-x-Asp-Val-x-Ser-Gly	WP_020668485 WP_031061687 WP_041038272 WP_043410518 AFZ17608 WP_033284790	*Amycolatopsis nigrescens* *Streptomyces ochraceiscleroticus* *Tolypothrix campylonemoides* *Cystobacter violaceus* *Microcoleus* sp. PCC 7113 *Streptomyces* sp. NRRL F-525

Table 2 (Continued)

	ORF no.[a] (size in aa, % identity/similarity to closest homolog)						*Closest homolog*	
Gene cluster	M. glauca *NBRC 14761^T*	M. fusca *NBRC 13915^T*	M. malaysiensis *NBRC 100735^T*	M. niveoalba *NBRC 15239^T*	*Domain organization*	*Predicted product*	*Accession no.*	*Origin*
	36-45 (1207, 80/85)	39-24 (1147, 82/87)	40-57 (1170, 81/87)	12-187 (1193, 81/86)	A/T-KS		WP_012889152	*S. roseum*
	36-48 (1208, 81/86)	39-20 (1175, 78/85)	48-2 (1175, 76/83)	59-3 (1177, 79/84)	A/T-C/T		WP_012889147	*S. roseum*
pks/nrps-1	36-49 (1877, 84/88)	39-19 (1828, 84/88)	48-3 (1822, 83/87)	59-4 (1840, 83/87)	KS/AT/KR/DH/ACP	x-x-z-pk-Asn-Ser	WP_012889146	*S. roseum*
	36-51 (1211, 70/75)	39-17 (1075, 75/83)	48-5 (1081, 75/81)	59-7 (1101, 75/81)	C/A_{Asn}/T		WP_012889144	*S. roseum*
	36-52 (1861, 88/93)	39-16 (1836, 87/92)	48-6 (1832, 87/92)	59-8 (1858, 86/91)	C/A_{Ser}/T-TE		WP_031168074	*S. roseum*
		55-6 (1819, 55/67)			KS/AT/KR/DH/ACP		EPH45771	*S. aurantiacus* JA 4570
pks/nrps-2	—	55-3 (1863, 47/56)	—	—	A/T-C/A_{Orn}/T	x-Orn-x-pk…	EPH45774	*S. aurantiacus* JA 4570
		55-2 (1226, 38/52)			C/A/T		WP_044889155	*Myxococcus* sp. (contaminant ex DSM 436)
		…			…			
			51-14 (1535, 53/62)		A/T-C/A_{Glu}/T		WP_020542770	*Nonomuraea coxensis*
			51-15 (1788, 55/68)		KS/AT/KR/DH/ACP		WP_040392143	*Catelliglobosispora koreensis*
			51-17 (1042, 49/61)		C/A_{Ser}/T	x-Glu-pk-Ser-x-Ser-Gly-Ser	WP_043407564	*Cystobacter violaceus*
pks/nrps-3	—	—	51-18 (1033, 51/61)	—	C/A/T		WP_031162978	*S. roseum*
			51-19 (990, 47/60)		C/A_{Ser}/T		WP_016333617	*Amycolatopsis orientalis*
			51-20 (1187, 56/66)		C/A_{Ser}/T-TE		WP_033408846	*N. coxensis*
			51-24[b] (1315, 49/60)		C/A_{Gly}/T		WP_036391030	*Micromonospora chokoriensis*
pks-1	—	14-100 (2484, 58/68)	—	04-258 (2543, 86/89)	KS/AT/DH/MT/ER/KR/ACP	Unpredictable	WP_036407059	*M. rosea*
				60-5 (>3538, 60/69)	…KR/ACP-KS/AT/KR/ACP-KS/AT/DH/KR/ACP-TE		CAJ88175	*Streptomyces ambofaciens* ATCC 23877
				60-3 (5476, 56/66)	KS/AT/DH/ER/KR/ACP-KS/AT/KR/ACP-KS/AT/DH/KR/ACP		CAJ88187	*S. ambofaciens* ATCC 23877
				60-2 (1589, 59/68)	KS/AT/KR/ACP		CAJ88175	*S. ambofaciens* ATCC 23877
				60-1 (>534, 67/77)	KS…	Stambomycin-like polyketide	WP_018835241	*Streptomyces* sp. CNQ766
pks-2	—	—	—	*66-11 (>2378, 57/66)*	…KR/ACP-KS/AT/DH/ER/KR/ACP		CAJ88175	*S. ambofaciens* ATCC 23877
				66-1 (>3154, 56/67)	KS/AT/ACP-KS/AT/DH/KR/ACP-KS…		AAZ77693	*Streptomyces antibioticus*
				81-1 (>1725, 68/76)	…ACP-KS/AT/KR/ACP		CAJ88176	*S. ambofaciens* ATCC 23877
				81-2 (>715, 78/84)	KS/AT…		CAJ88175	*S. ambofaciens* ATCC 23877

Table 2 (Continued)

Gene cluster	ORF no.[a] (size in aa, % identity/similarity to closest homolog)				Domain organization	Predicted product	Closest homolog	
	M. glauca *NBRC 14761*^{T}	M. fusca *NBRC 13915*^{T}	M. malaysiensis *NBRC 100735*^{T}	M. niveoalba *NBRC 15239*^{T}			Accession no.	Origin
				86-1 (>2891, 62/71)	...AT/KR/ACP-KS/AT/KR/ ACP-KS/AT...		CAJ88175	*S. ambofaciens* ATCC 23877
				74-1 (>4210, 71/79)	...AT/DH/ER/KR/ACP-KS/AT/KR/ACP-KS/AT/DH...		CAJ88175	*S. ambofaciens* ATCC 23877
				79-1 (>2844, 64/73)	...KR/ACP-KS/AT/KR/ACP-KS/ AT...		CAJ88175	*S. ambofaciens* ATCC 23877
				93-1 (>1278, 69/77)	...AT/KR/ACP-KS...		CAJ88175	*S. ambofaciens* ATCC 23877
				388-1 (>447, 75/82)	...KS/AT...		CAJ88175	*S. ambofaciens* ATCC 23877
				33-93 (>1887, 67/76)	...KR/ACP-KS/AT/KR/ACP		CAJ88175	*S. ambofaciens* ATCC 23877
				75-1 (>1983, 58/67)	KS/AT/DH/ER/KR...		WP_042389815	*Streptacidiphilus melanogenes*
				90-1 (>1300, 58/68)	...KS/AT/DH...		WP_014060786	*Streptomyces violaceusniger*
pks-3	—	—	—	*76-1 (>4931, 59/69)*	...KR/ACP-KS/AT/DH/ER/KR/ ACP-KS/AT/DH/KR/ACP-KS...	Unpredictable	WP_045867303	*Streptomyces* sp. NBRC 110027
				80-3 (>1024, 53/63)	...DH/KR/ACP		AJC56294	*Streptomyces* sp. 769
				80-2 (1633, 54/66)	KS/AT/KR/ACP		WP_032769932	*Streptomyces* sp. CNS654
				80-1 (>496, 66/78)	KS...		WP_044580835	*Streptomyces iranensis*
				91-1 (>1713, 57/68)	...KR/ACP-KS/AT/DH...		WP_014060786	*Streptomyces violaceusniger*
		10-107[b] (84, 57/73)			ACP		CTQ88430	*Kibdelosporangium* sp. MJ126-NF4
pksII	—	10-116 (422, 78/86)	—	—	KSα	Xantholipin-like polyketide	AJD20010	Uncultured bacterium
		10-117 (418, 65/75)			KSβ (CLF)		AJD20009	Uncultured bacterium

Abbreviations: A, adenylation; ACP, acyl carrier protein; AT, acyltransferase; C, condensation; CoL, CoA ligase; CLF, chain length factor; DH, dehydratase; E, epimerization; ER, enoylreductase; KR, ketoreductase; KS, ketosynthase; mCys, methyl cysteine; MT, methyltransferase; NBRC, Biological Resource Center, National Institute of Technology and Evaluation; NRPS, nonribosomal peptide synthetase; pk, polyketide unit; PKS, polyketide synthase; T, thiolation; TE, thioesterase; x, unpredicted amino acid; y, unpredicted starter molecule; z, lack of A domain.
Strain-specific gene clusters are in bold. Open reading frames not completely sequenced are shown in italics and the undetermined domains are shown as '...'.
[a]ORF numbers are shown in combination with scaffold/contig numbers.
[b]Encoded in the complementary strand.
Predicted amino acid substrates of A domains are shown in subscripted letters.

Gene clusters shared among two to three strains. Four NRPS gene clusters (*nrps-3* to *-6*) and one PKS gene cluster (*pks-1*) were shared by two to three strains. *Nrps-3*, *nrps-4* and *nrps-5* were predicted to synthesize a tetrapeptide, a tripeptide containing two alanine (Ala) residues and a dipeptide containing one cysteine (Cys) residue, respectively. Since the NRPS genes in the *nrps-4* cluster showed high sequence homologies (85–88% identity) with those of *S. roseum*, and the domain organizations in *Microtetraspora* strains and *S. roseum* are equivalent, these strains will produce similar tripeptide compounds. We were not able to predict the peptide compound product of *nrps-6* because the gene cluster encoded only a single module. *Pks-1* also encoded only a single module. The PKS-I resembled that of *Microbispora rosea*, but the product of *M. rosea* has not been identified.

Strain-specific gene clusters. *M. glauca* NBRC 14761^T, *M. fusca* NBRC 13915^T, *M. malaysiensis* NBRC 100735^T and *M. niveoalba* NBRC 15239^T had one, three, three and three strain-specific NRPS and/or PKS gene clusters, respectively. *M. glauca* NBRC 14761^T possessed one strain-specific NRPS gene cluster, *nrps-7*. This cluster comprised four modules encoded by two NRPS genes and one of the modules was predicted to code for Asp, suggesting that the product is a tetrapeptide containing an Asp. Because this cluster is homologous to that of *S. roseum*, *M. glauca* NBRC 14761^T likely produces a tetrapeptide similar to that of *S. roseum*. *M. fusca* NBRC 13915^T possessed one specific NRPS gene cluster, *nrps-8*; one specific PKS/NRPS gene cluster, *pks/nrps-2*; and one specific PKS-II gene cluster, *pksII*. Though *nrps-8* was not completely sequenced, this cluster encoded at least nine modules and the predicted product was a peptide composed of nine or more amino acid residues including serine (Ser), threonine (Thr) and asparagine (Asn). *Pks/nrps-2* also may not have been completely sequenced, but this cluster encoded at least one PKS-I module and three NRPS modules, suggesting the product contains at least three amino acid residues and a polyketide unit. *PksII* was predicted to synthesis a xantholipin-like polyketide, because its PKSs show 63–78% sequence identities with synthases of xantholipin and related compound.[22,23] *M. malaysiensis* NBRC 100735^T possessed two specific NRPS gene clusters, *nrps-9* and *nrps-10* and one specific PKS/NRPS gene cluster, *pks/nrps-3*. The products of *nrps-9* and *nrps-10* were predicted to be a tripeptide composed of one phenylalanine (Phe) and two Cys residues, and a tripeptide containing valine (Val), respectively. *Pks/nrps-3* encoded seven NRPS modules and one PKS module, suggesting that its product is a polyketide-nonribosomal peptide hybrid compound composed of seven amino acid residues and one polyketide unit. Substrate prediction for the A domains suggested that this hybrid compound contains glutamate (Glu), Ser and Gly residues. *M. niveoalba* NBRC 15239^T possessed one specific NRPS gene cluster, *nrps-11*, and two specific PKS-I gene clusters, *pks-2* and *pks-3*. *Nrps-11* encoded 11 NRPS modules, one of which had a CoA-ligase domain instead of an A domain. Substrate prediction for the ten A domains indicated a product composed of a starter molecule and ten amino acid residues such as Gly, Thr, tryptophan (Trp), Asp, Val and Ser. *Pks-2* was not completely sequenced, but encoded at least 22 modules, and 12 of its putative 14 PKS-Is showed sequence homologies (56–78% similarity) with those of *Streptomyces ambofaciens* ATCC 23877. Because these closest homologs are reportedly involved in the synthesis of stambomycin, a 51-membered macrolide,[24] we predict that *pks-2* will synthesize a stambomycin-like large polyketide compound. *Pks-3* was also not completely sequenced but encoded at least seven PKS modules. We predict that its product is a polyketide compound including a C_{14} or longer polyketide chain.

Ayuso-Sacido *et al*.[25] searched for PKS-I genes in actinomycete strains by PCR using degenerate primers and could not detect any PKS-I genes in type strains of *M. fusca* and *M. niveoalba*. However, our present study indicates that these two strains harbor PKS-I genes, suggesting the usefulness of genome analysis to survey for PKS-I genes in rare actinomycetes. This is the first report on *Microtetraspora* genome sequences, revealing novel and diverse NRPS and PKS pathways whose products have yet to be identified. Among the 18 types of NRPS and PKS gene clusters found in the four *Microtetraspora* type strains, seven (*nrps-1*, *nrps-2*, *nrps-4*, *nrps-7*, *nrps-10*, *pks/nrps-1* and *pks-1*) are similar to those present in strains belonging to the genera *Streptosporangium* and *Microbispora*, both phylogenetically close to the genus *Microtetraspora*. This suggests that these products are not specific to *Microtetraspora*. In contrast, the remaining 11 types of gene clusters do not show high sequence similarities to those registered in GenBank/EMBL/DDBJ databases, suggesting they are specific to *Microtetraspora*. Therefore, these products are potentially novel. We suggest that *Microtetraspora* strains might be an untapped, rich source of secondary metabolites and should be studied further for the discovery of novel bioactive compounds.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to Ms Yuko Kitahashi for annotating the NRPS and PKS genes. We also thank Ms Tomoko Hanamaki for assistance with whole-genome sequencing.

1 Berdy, J. Bioactive microbial metabolites. *J. Antibiot. (Tokyo)* **58**, 1–26 (2005).
2 Watve, M. G., Tickoo, R., Jog, M. M. & Bhole, B. D. How many antibiotics are produced by the genus *Streptomyces*? *Arch. Microbiol.* **176**, 386–390 (2001).
3 Lazzarini, A., Cavaletti, L., Toppo, G. & Marinelli, F. Rare genera of actinomycetes as potential producers of new antibiotics. *Antonie Van Leeuwenhoek* **78**, 399–405 (2000).
4 Zhang, Z., Wang, Y. & Ruan, J. Reclassification of *Thermomonospora* and *Microtetraspora*. *Int. J. Syst .Bacteriol.* **48**, 411–422 (1998).
5 Thiemann, J. E., Pagani, H. & Beretta, G. A new genus of the *Actinomycetales*: *Microtetraspora* gen. nov. *J. Gen. Microbiol.* **50**, 295–303 (1968).
6 Miyadoh, S., Anzai, H., Amano, S. & Shomura, T. *Actinomadura malachitica* and *Microtetraspora viridis* are synonyms and should be transferred as *Actinomadura viridis* comb. nov. *Int. J. Syst. Bacteriol.* **39**, 152–158 (1989).
7 Goodfellow, M. Maldonado, L. A. & Quintana, E. T. Reclassification of *Nonomuraea flexuosa* (Meyer 1989) Zhang *et al.* 1998 as *Thermopolyspora flexuosa* gen. nov., comb. nov., nom. rev. *Int. J. Syst. Evol. Microbiol.* **55**, 1979–1983 (2005).
8 Nakajima, Y., Ho, C. C. & Kudo, T. *Microtetraspora malaysiensis* sp. nov., isolated from Malaysian primary dipterocarp forest soil. *J. Gen. Appl. Microbiol.* **49**, 181–189 (2003).
9 Kawamura, N. *et al.* Pyralomicins, novel antibiotics from *Microtetraspora spiralis*. I. Taxonomy and production. *J. Antibiot. (Tokyo)* **49**, 706–709 (1996).
10 Toda, S. *et al.* A new neuritogenetic compound BU-4514N produced by *Microtetraspora* sp. *J. Antibiot. (Tokyo)* **46**, 875–883 (1993).
11 Tomita, K. *et al.* Taxonomy of the antibiotic Bu-2313-producing organism. *Microtetraspora caesia* sp. nov. *J. Antibiot. (Tokyo)* **33**, 1491–1501 (1980).
12 Nett, M., Ikeda, H. & Moore, B. S. Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat. Prod. Rep.* **26**, 1362–1384 (2009).
13 Komaki, H. *et al.* Genome based analysis of type-I polyketide synthase and non-ribosomal peptide synthetase gene clusters in seven strains of five representative *Nocardia* species. *BMC Genomics* **15**, 323 (2014).
14 Komaki, H., Ichikawa, N., Oguchi, A., Hanamaki, T. & Fujita, N. Genome-wide survey of polyketide synthase and nonribosomal peptide synthetase gene clusters in *Streptomyces turgidiscabies* NBRC 16081. *J. Gen. Appl. Microbiol.* **58**, 363–372 (2012).
15 Komaki, H. *et al.* Genome-based analysis of type-I polyketide synthase and nonribosomal peptide synthetase gene clusters in a novel strain taxonomically close to the genus *Salinispora*. *J. Antibiot. (Tokyo)* **68**, 767–770 (2015).
16 Ohtsubo, Y., Maruyama, F., Mitsui, H., Nagata, Y. & Tsuda, M. Complete genome sequence of *Acidovorax* sp. strain KKS102, a polychlorinated-biphenyl degrader. *J. Bacteriol.* **194**, 6970–6971 (2012).

17 Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W. & Hauser, L. J. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
18 Blin, K. *et al.* antiSMASH 2.0–a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **41**, W204–W212 (2013).
19 Richter, M. & Rossello-Mora, R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl Acad. Sci. USA* **106**, 19126–19131 (2009).
20 Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P. & Tiedje, J. M. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* **57**, 81–91 (2007).
21 Doroghazi, J. R. & Metcalf, W. W. Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. *BMC Genomics* **14**, 611 (2013).
22 Kang, H. S. & Brady, S. F. Mining soil metagenomes to better understand the evolution of natural product structural diversity: pentangular polyphenols as a case study. *J. Am. Chem. Soc.* **136**, 18111–18119 (2014).
23 Zhang, W. *et al.* Unveiling the post-PKS redox tailoring steps in biosynthesis of the type II polyketide antitumor antibiotic xantholipin. *Chem. Biol.* **19**, 422–432 (2012).
24 Laureti, L. *et al.* Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. *Proc. Natl Acad. Sci. USA* **108**, 6258–6263 (2011).
25 Ayuso-Sacido, A. & Genilloud, O. New PCR primers for the screening of NRPS and PKS-I systems in actinomycetes: detection and distribution of these biosynthetic gene sequences in major taxonomic groups. *Microb. Ecol.* **49**, 10–24 (2005).

Supplementary Information accompanies the paper on The Journal of Antibiotics website (http://www.nature.com/ja)