SIT760 and the other to SIT134. The SIT134 "H3" harbored the *recC*$^{1491\ (ttc\text{-}ttg)}$ X lineage associated SNP and the two SIT316 "H3" clinical isolates harbored the mutation on *recR*$^{94\ (ttg\text{-}ctg)}$ hypothesized to be linked to the T2 lineage isolates from Central African Republic clinical isolates (Dos Vultos et al., 2008) (see also Table 3 and Supplemental Table 3).

Within LAM-labeled clinical isolates (n = 72), we found that 47 did not harbor the expected mutation on *ligB*$^{1212\ (tcc\text{-}tcg)}$, forty of these were LAM7_TUR and the seven others were LAM3-S (Table 3). All others LAM sublineages tested so far (LAM9, LAM5, LAM3, and LAM2) did harbor the expected *ligB* mutation.

Within the T lineage, there are a lot of sublineages defined by spoligotype signatures (Brudey et al., 2006), among which one is T2. *recR*$^{94\ (ttg\text{-}ctg)}$ SNP was found associated with Central African Republic isolates, all of them carrying a T2 signature (absence of spacer 40) (Dos Vultos et al., 2008). Here, 83 DNA samples belonged to the T lineage (sublineages: either T1, T1-Tuscany variant, T1–T2 undefined, T2, T4_CEU1, T5 and T5_RUS1), out of 32 T2 isolates, six only harbored the mutation on *recR*$^{94\ (ttg\text{-}ctg)}$ (Table 3). These results show that the *recR*^{94} mutation cannot be considered as being linked to the absence of spacer 40, which defines the T2 sublineage.

Table 4
Lineage identification of spoligotype that lack of it, the so-called "U" spoligotypes in SpolDB4 (Brudey et al., 2006), either because they were not reported in the database, or because this identification was not clear. Lineage assignment by SNP and by SPOTCLUST (Vitol et al., 2006) is shown.

			Family			
Pos[a]	SIT	Spoligotype	SpolDB4	SNP	No. of strains[c]	SPOTCLUST[b]
1	NR		U	EAI	2	EAI5 (0.95)
2	NR		U	EAI	1	EAI5 (1.00)
3	NR		U	H	1	Family 34 (1.00)
4	NR		U	H	3	Family 35 (1.00)
5	NR		U	H	1	H1 (1.00)
6	NR		U	H	2	H3 (0.77)
7	NR		U	H	1	H3 (0.96)
8	1274		U	H	1	H1 (1.00)
9	NR		U	H	1	H3 (0.77)
10	NR		U	H	1	H1 (1.00)
11	NR		U	H	1	H1 (1.00)
12	NR		U	H	2	H3 (0.77)
13	NR		U	H	1	H1 (1.00)
14	NR		U	LAM	2	T1 (0.80)
15	NR		U	LAM	1	LAM3 (0.93)
16	NR		U	LAM	4	LAM9 (1.00)
17	NR		U	LAM	1	LAM8 (0.94)
18	NR		U	LAM	1	LAM3 (1.00)
19	NR		U	LAM	1	LAM8 (0.99)
20	NR		U	LAM	1	LAM3 (1.00)
21	NR		U	LAM	1	LAM9 (0.96)
22	NR		U	LAM	1	LAM9 (1.00)
23	NR		U	LAM	1	T4 (1.00)
24	NR		U	LAM	2	LAM (1.00)
25	105		U (LAM3?)	LAM	2	LAM7 (1.00)
26	NR		U	LAM	1	LAM8 (1.00)
27	NR		U	LAM	1	LAM9 (0.97)
28	NR		U	LAM	1	T4 (1.00)
29	1531		U	X	1	X1 (0.65)
30	NR		U	X	2	X3 (1.00)
31	NR		U	X	1	H3 (0.77)
32	2125[d]		U	TUR- T3- Osa	1	LAM9 (1.00)
33	NR		U	TUR- T3- Osa	1	LAM9 (0.99)
34	NR		U	TUR- T3- Osa	1	T1 (1.00)
35	NR		U	TUR- T3- Osa	1	T1 (1.00)

[a]This number is provided just to locate the pattern in this list.
[b]SPOTCLUST Identification according the spoligotype pattern and the algorithm probability of the result of being true in parenthesis.
[c]Number of strains related to the frequency of the spoligotype pattern.
[d]SIT number from Aktas et al. (2008).
To read and interpret the spoligotype, spacers are placed in order from 1 to 43 (left to right).

An interesting secondary result found in this study is the confirmation of the phylogenetic link between the clinical isolates from the T3_Osaka sublineage and the ones from the LAM7_TUR. These two groups of strains share the same MIRU-VNTR[1] 12 loci signature (MIRU-VNTR International Type, MIT310) and harbor divergent spoligotypes (SIT41 and SIT627). Here, using a large set of DNA samples from Zonguldak – Turkey (LAM7_TUR) and from Japan (T3_Osaka and other variants), our results show that they shared the same SNP on *ligC*$^{809\ (tgg\ [trp]\ to\ ttg\ [leu])}$. Moreover, we observed that 2/2 SIT78 (T1–T2) and 2/25 SIT53 (T1) also had this SNP.

Our study also included 84 spoligotypes for which the lineage information was unknown, the so-called "U" isolates (for Unknown), *i.e.* isolates for which the spoligotype signature either did not allow assignation to a lineage ($n = 7$) (Brudey et al., 2006) or had not been reported in the international spoligotype database, SpolDB4 ($n = 77$, referred here as "NR"). With our SNP-based scheme, 47 clinical isolates (56%, with same frequency for strict "U" and "NR") could now be assigned at a lineage level (Table 4). The identification of these clinical isolates is distributed as follows: 21 isolates were assigned to the LAM-lineage, 15 to Haarlem, 4 to X and 3 to the EAI lineage. We also tested the congruence between our SNP classification and SPOTCLUST, a Probabilistic Multivariate Bernoulli Mixture Model that uses Naïve Bayes assumptions, as reported previously (Vitol et al., 2006). This algorithm automatically provides a probability for a spoligotype pattern to be part of a lineage using previous assignations reported in SpolDB4 (Brudey et al., 2006). Interestingly, we found a good correlation (72%) between SPOTCLUST assignation and our SNP identification (Table 4). However, we also observed discrepancies (positions 3, 4, 14, 23, 28 and 31 in Table 4). Indeed, SPOTCLUST classified these clinical isolates as: Lineage 34, Lineage 35, T1, T4, T4 and H3, respectively with probabilities of higher than 0.77 whereas our results classify these samples as H, H, LAM, LAM, LAM and X, respectively. Note that TUR_T3_Osaka strains were identified by SPOTCLUST as LAM9 or T1 with probabilities close to 1.00 (Table 4).

4. Discussion

We aimed to confirm the spoligotype-based classification of MTC using a restricted SNP-based scheme designed to be performed in a high-throughput way. Most commonly used classification of MTC relies on spoligotype patterns. Lineages and sublineages have been defined based on recurrent so-called spoligo-signatures (absence and/or presence of specific spacers) that have proven to be phylogeographically meaningful (Brudey et al., 2006). The existence of these lineages was confirmed using independent techniques such as LSP (Gagneux et al., 2006), MIRU-VNTR typing (Allix-Beguec et al., 2008b; Wirth et al., 2008)), or via sequencing projects (Comas et al., 2009).

To provide phylogenetically reliable information, markers should not be prone to converge. SNPs most of the time are unlikely to convergence and also are less prone to distortion by selective pressures (Schork et al., 2000). Notwithstanding drug resistance genes, there is just one evidence of convergence evolution due to a nonsense variation in ada/alkA gene observed in *M. tuberculosis* and in *M. bovis* that could confer a selective advantage (Nouvel et al., 2007), a SNP that is not included in our data set.

In *M. tuberculosis*, and maybe other clonal organisms, the genes involved in replication, repair and recombination (3R) seemed to play a key role in adaptation (Dos Vultos et al., 2008) so they might be prone to experience selection. However, in our study, four of the targeted SNPs are synonymous (sSNPs) so they are unlikely to be selected for. In addition, the three other non-synonymous mutations chosen are conservative so the replaced amino acid is biochemically similar to the pre-existing one. Altogether, the SNPs we provide are likely to be neutral. Regarding classification, we confirmed that EAI, BOVIS and X lineages are monophyletic, *i.e.* that the signature that defines them is phylogenetically relevant: EAI lineage is defined by the absence of spacers 29–32, absence of 34 and the *recO*$^{606\ (ggc\text{-}ggt)}$ mutation; the BOVIS lineage is both defined by the absence of spacers 39–43 and the *alkA* $^{807\ (ctg\text{-}cta)}$ mutation; the X lineage is defined by the absence of spacer 18, 33–36 group of spacers and by the *recC*$^{1491\ (ttc\text{-}ttg)}$ mutation. The results we provide come in addition to previous results obtained (Dos Vultos et al., 2008) so that we can say that there is strong support for these correlations.

This study in contrast provides some evidence of mis-labeling in LAM, T and Haarlem lineages due to the used marker itself. Regarding Haarlem lineage, H4 sublineage did not carry the expected *mgtC*$^{545\ (cgc\text{-}cac)}$ mutation. Kovalev et al. (2005) showed that according to their MIRU-VNTR signatures (especially MIRU10 with 7–10 copies) this lineage was unlikely to be phylogenetically related to Haarlem so it was renamed as URAL. Here, we confirm that the loss of spacer 29–31 in URAL is a genetic event that is independent from the loss of spacer 26–31 in H1 and H2, or of spacer 31 in H3. For the time being we have not identified a SNP that correlates with this lineage. Interestingly, all the strains that carried this signature in our scheme came from Turkey that is geographically close to Russia.

Two SITs (134 and 316) that belong to the H3 lineage (according to their spoligotype signatures) lacked the *mgtC*^{545} SNP Haarlem-specific mutation but instead harbored the *recC*^{1491}-X and *recR*^{94}-T2 lineage-specific mutations, respectively that may indicate that the deletion of spacer 31 could have occurred several times independently so that it could be used cautiously as a phylogenetic marker. We suggest that a "Haarlem" lineage assignment for strains classified as H3 could therefore be confirmed by the SNP method described here.

The X lineage was first defined by the absence of spacer 18 (Sebban et al., 2002; Filliol et al., 2002) in addition to the absence of spacers 33–36. Here, we found a strain that was not labeled as X according to its spoligotype, but in fact it was according to its SNP pattern. This suggests that the deletion of spacer 18 occurred after the *recC*^{1491} SNP event. Hence, we suggest to broaden X lineage to all strains carrying this SNP.

Within the T2 sublineage (defined by the absence of spacer 40, in addition to the absence of 33–36 group of spacers) many clinical isolates sharing SIT52 or SIT736 spoligotype did not harbor the SNP on *recR*^{94}. In addition a relatively low number of *recR*^{94} mutations were found in the whole set of strains, suggesting that this mutation occurred too recently. For the time being, we are not able to well describe the T lineage and resolve all the relationships between modern strains carrying the 33–36 spacer deletion, because there is no single SNP available to do that by now and due to the big genetic diversity inside this main lineage, it remains "ill-defined" or poorly defined (Brudey et al., 2006). Also, Vitol et al. (2006) were unable to create models that discriminate well among the members of the T lineage based on their spoligotype signatures.

Another contribution to the T sublineage markers is the confirmation (with strong support, $n = 62$) of the existence of a TUR_T3_Osaka lineage based on the *ligC*^{809} SNP. Spoligotypes that harbors this SNP are very different between them (Table 3 and Supplementary Table 1). Millet et al. (2007) reported before a shared MIRU12-VNTR pattern called MIT310 (215125113322) between strains from T3_Osaka lineage (Takashima and Iwamoto, 2006) and the previously reported LAM7_TUR lineage (Zozio et al., 2005; Millet et al., 2007). All the LAM7_TUR strains tested so far lacked the mutation on the LAM-associated *ligB*^{1212} SNP, consequently they are

[1] Mycobacterial Interspersed Repetitive Unit-Variable Number of Tandem Repeat.

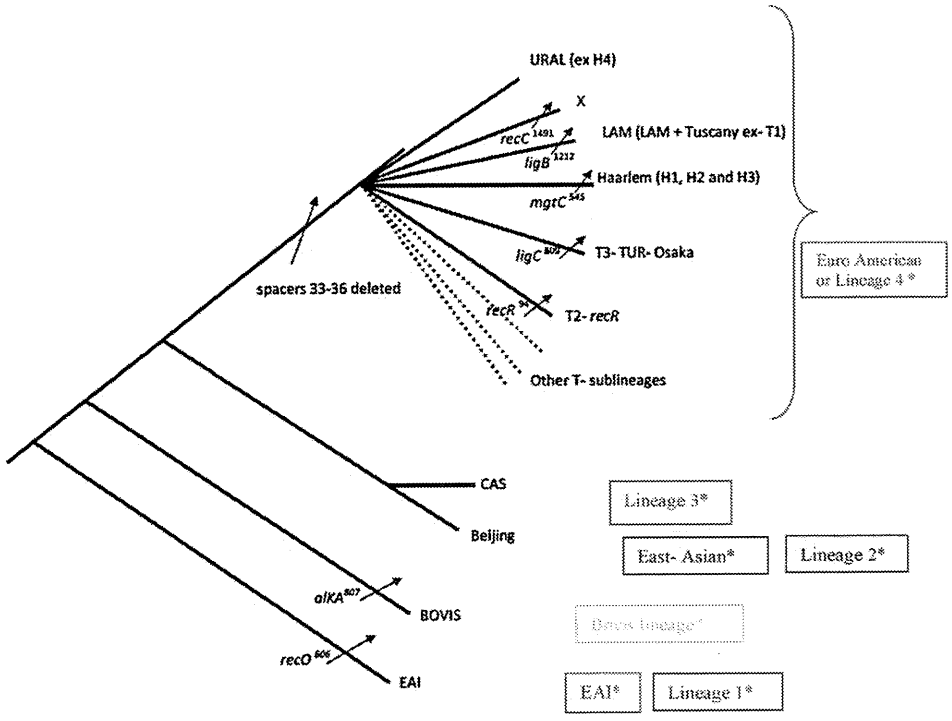

Fig. 2. Evolutionary proposal to identify *M. tuberculosis* families based on the SNP used in this study. Small arrows show the genes in which the SNP is present to allow the identification. The big arrow shows the deletion of spacers 33–36 that is common to all the modern strains. It is shown the correspondence between major lineages defined by Spoligotyping, our SNP proposal and inside boxes the one related to LSP plus a recently work done in 89 sequenced genes. *The LSP classification is related to the work of (Gagneux et al., 2006) and recently to an SNP work of (Comas et al., 2009).

not a sublineage of LAM so we give further support to renaming them TUR (Table 3). We also observed that strains carrying SIT78 T1–T2, SIT53 T1 (2/25), or SIT553 T2–T3 and 4 NR strains also harbored the TUR-T3-Osaka-*ligC*^{809} mutation. This suggests that this mutation could be useful digging inside a scheme targeting T_lineage genetic diversity.

Another result is that one SIT53 T1 strain, which is both wt for *recR*^{94} and for *ligC*^{809}, harbors the Haarlem-specific mutation for *mgtC*^{545} (Table 3). This finding could corroborate the hypothesis made before by others (Duchene et al., 2004) about the relation between T and Haarlem strains. In this case, one possibility is that strains harboring the SIT53 T1 spoligo-signature may have evolved at a genome level to Haarlem in spite of retaining the CRISPR region of the progenitor. This could be also the case for SIT120 T1 for which we found that it also harbors the mutation for Haarlem on *mgtC*^{545}, the only difference at a spoligo level between SIT53 and 120 is one spacer missing, spacer 20 (Table 3).

With our SNP-signature panel we also provide evidence favoring the renaming of several spoligotypes previously assigned to T-clade based on spoligo rules, these are: SIT159 T1 (Tuscany variant), SIT254 (T5_RUS1) and SIT370 (T1) that harbored the *ligB*^{1212} SNP-LAM-associated mutation so we assigned the right lineage to these strains. The LAM assignation rule is based on the simultaneous absence of spacers 21–24 and 33–36 that is not accomplished by SIT159, SIT 254 T5_RUS1or SIT370. Taken together, this finding could be an indication that spacers 21 and 22 are not so informative to define the LAM lineage as is the absence of spacer 23.

One of the most important features of our new SNP proposal relies on the identification of those spoligotypes that lack the lineage identification. Emerging clones continuously will appear as a result of genetic variation of pre-existing ones (Brudey et al., 2006). The SNP assignment is unambiguous, providing a precise tool that targets several sites (7 in our actual proposal) with a high precision performing the lineage assignation. We could successfully assign a lineage to 47 out of 84 U strains. One interesting lineage assignment in the group of "U" strains was the SIT 105 that had been responsible for an MDR-TB outbreak in Spain. This SIT was thought to be "LAM3?" in SpolDB4 with doubts and now we could confirm it because it was found to harbor the *ligB*^{1212}-LAM-specific mutation.

Even though Spoligotyping remains a first line tool to delineate the molecular ecology of the circulating strains, we have shown that sometimes the spoligotype signature will not always reflect the real lineage of the MTC strains. With this work, we have solved some intraspecies taxonomic issues and we provide an accurate lineage assignation using a minimal set of SNPs almost as informative as large sequencing projects similar to those of Comas et al. (2009). However, we still need to include other SNPs for other lineages, because there are still spoligotype patterns that we were not able to identify with this proposal. This is also true to reach the level of spoligotyping sublineage discrimination. Nevertheless we could correctly identify strains belonging to the major MTC lineages or families such as EAI, BOVIS, X, LAM, Haarlem as well as some T sublineages, such as T2-*recR*^{94} and TUR_T3_OSAKA. We do not have one specific SNP for the T or S strain lineage for the time being. In Fig. 2 we show an evolutionary scenario built with the set of SNPs used in this study and a hypothesis of when the deletion of spacers 33–36 may have occurred.

Some SNP-based studies have already been performed in *M. tuberculosis* to date (Alland et al., 2003; Baker et al., 2004; Filliol et al., 2006; Gutacker et al., 2006; Gutacker et al., 2002; Sreevatsan et al., 1997). However, the composition and number of SNP cluster groups (SCGs) within *M. tuberculosis* have remained unclear since it also used some inadequate genes and SNPs (i.e. involved in drug resistance) (Baker et al., 2004) whereas others were selected from a non-representative set of available genomes (Filliol et al., 2006; Gutacker et al., 2002, 2006). These previous studies had not taken full advantage of the power of SNP-based methods. Here we show that SNP typing could serve as a "gold standard" for DNA typing. We think that in the near future the inclusion of more SNP information will become an important parameter to efficiently

classify a given MTC clone, either for molecular epidemiological or for evolutionary purposes. High-throughput multiplexing among various techniques and platforms will be one of the best ways to achieve this with reasonable economic costs (Dunbar, 2006; Bergval et al., 2008).

Acknowledgments

This study was supported by the "Excellency Chair in microbiology of the Université Paris-Sud 11" granted to CS in September 2007, through a 12-month post-doctoral fellowship grant given to EA. EA is a senior mycobacteriology scientist of the IVIC (Instituto Venezolano de Investigaciones Científicas, Caracas, Venezuela). JZ is a PhD fellow of the Université Paris-Sud 11 through the "Ecole Doctorale Gènes Génomes Cellules". We are grateful to François Topin, Luminex BV, The Netherlands, and to Luminex Corporation, Austin, Texas for their support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.meegid.2010.07.006.

References

Abadia, E., Sequera, M., Ortega, D., Mendez, M.V., Escalona, A., Da Mata, O., Izarra, E., Rojas, Y., Jaspe, R., Motiwala, A.S., Alland, D., de Waard, J., Takiff, H.E., 2009. *Mycobacterium tuberculosis* ecology in Venezuela: epidemiologic correlates of common spoligotypes and a large clonal cluster defined by MIRU-VNTR-24. BMC Infect. Dis. 9, 122.

Achtman, M., 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu. Rev. Microbiol. 62, 53–70.

Aktas, E., Zozio, T., Comert, F.B., Kulah, C., Aydin, O., Rastogi, N., Sola, C., 2008. A first insight into the genetic diversity and population structure of *Mycobacterium tuberculosis* in Zonguldak, Turkey. Clin. Microbiol. Infect. 55–59.

Alix, E., Godreuil, S., Blanc-Potard, A.B., 2006. Identification of a Haarlem genotype-specific single nucleotide polymorphism in the mgtC virulence gene of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 44, 2093–2098.

Alland, D., Whittam, T.S., Murray, M.B., Cave, M.D., Hazbon, M.H., Dix, K., Kokoris, M., Duesterhoeft, A., Eisen, J.A., Fraser, C.M., Fleischmann, R.D., 2003. Modeling bacterial evolution with comparative-genome-based marker systems: application to *Mycobacterium tuberculosis* evolution and pathogenesis. J. Bacteriol. 185, 3392–3399.

Allix-Beguec, C., Fauville-Dufaux, M., Supply, P., 2008a. Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 46, 1398–1406.

Allix-Beguec, C., Harmsen, D., Weniger, T., Supply, P., Niemann, S., 2008b. Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. J. Clin. Microbiol. 46, 2692–2699.

Baker, L., Brown, T., Maiden, M.C., Drobniewski, F., 2004. Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. Emerg. Infect. Dis 10, 1568–1577.

Baranov, A.A., Mariandyshev, A.O., Mannsaker, T., Dahle, U.R., Bjune, G.A., 2009. Molecular epidemiology and drug resistance of widespread genotypes of *Mycobacterium tuberculosis* in northwestern Russia. Int. J. Tuberc. Lung Dis. 13, 1288–1293.

Bergval, I.L., Vijzelaar, R.N., Dalla Costa, E.R., Schuitema, A.R., Oskam, L., Kritski, A.L., Klatser, P.R., Anthony, R.M., 2008. Development of multiplex assay for rapid characterization of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 46, 689–699.

Brudey, K., Driscoll, J., Rigouts, L., Prodinger, W.M., Gori, A., Al-Hajoj, S.A.M., Allix, C., Aristimuno, L., Arora, J., Baumanis, V., Binder, L., Cafrune, P., Cataldi, A., Cheong, S., Diel, R., Ellermeier, C., Evans, J.T., Fauville-Dufaux, M., Ferdinand, S., Garcia de Viedma, D., Garzelli, C., Gazzola, L., Gomes, H.M., Gutierrez, M.C., Hawkey, P.M., van Helden, P.D., Kadival, G.V., Kreiswirth, B.N., Kremer, K., Kubin, M., Kulkarni, S.P., Liens, B., Lillebaek, T., Ly, H.M., Martin, C., Martin, C., Mokrousov, I., Narvskaia, O., Ngeow, Y.F., Naumann, L., Niemann, S., Parwati, I., Rahim, M.Z., Rasolofo-Razanamparany, V., Rasolonavalona, T., Rossetti, M.L., Rüsch-Gerdes, S., Sajduda, A., Samper, S., Shemyakin, I.G., Singh, U.B., Somoskovi, A., Skuce, R., Van Soolingen, D., Streicher, E.M., Suffys, P.N., Tortoli, E., Tracevska, T., Vincent, V., Victor, T.C., Warren, R., Yap, S.F., Zaman, K., Portaels, F., Rastogi, N., Sola, C., 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics, and epidemiology. BMC Microbiol. 6, 23.

Buchmeier, N., Blanc-Potard, A., Ehrt, S., Piddington, D., Riley, L., Groisman, E.A., 2000. A parallel intraphagosomal survival strategy shared by *Mycobacterium tuberculosis* and *Salmonella enterica*. Mol. Microbiol. 35, 1375–1382.

Chuang, P.C., Liu, H., Sola, C., Chen, Y.M., Jou, R., 2008. Spoligotypes of *Mycobacterium tuberculosis* isolates of a high tuberculosis burden aboriginal township in Taiwan. Infect. Genet. Evol. 8, 553–557.

Comas, I., Homolka, S., Niemann, S., Gagneux, S., 2009. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. PLoS ONE 4, e7815.

Corpet, F., 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 16, 10881–10890.

Cowan, L.S., Diem, L., Brake, M.C., Crawford, J.T., 2004. Transfer of a *Mycobacterium tuberculosis* genotyping method, Spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system. J. Clin. Microbiol. 42, 474–477.

Dos Vultos, T., Mestre, O., Rauzier, J., Golec, M., Rastogi, N., Rasolofo, V., Tonjum, T., Sola, C., Matic, I., Gicquel, B., 2008. Evolution and diversity of clonal bacteria: the paradigm of *Mycobacterium tuberculosis*. PLoS ONE 3, e1538.

Duchene, V., Ferdinand, S., Filliol, I., Guégan, J.F., Rastogi, N., Sola, C., 2004. Phylogenetic reconstruction of the *Mycobacterium tuberculosis* complex within four settings of the Caribbean region: tree comparative analysis and first appraisal on their phylogeography. Infect. Gen. Evol. 4, 5–14.

Dunbar, S.A., 2006. Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. Clin. Chim. Acta 363, 71–82.

Filliol, I., Driscoll, J.R., Van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valetudie, G., Anh, D.D., Barlow, R., Banerjee, D., Bifani, P.J., Brudey, K., Cataldi, A., Cooksey, R.C., Cousins, D.V., Dale, J.W., Dellagostin, O.A., Drobniewski, F., Engelmann, G., Ferdinand, S., Gascoyne-Binzi, D., Gordon, M., Gutierrez, M.C., Haas, W.H., Heersma, H., Kallenius, G., Kassa-Kelembho, E., Koivula, T., Ly, H.M., Makristathis, A., Mammina, C., Martin, G., Mostrom, P., Mokrousov, I., Narbonne, V., Narvskaya, O., Nastasi, A., Niobe-Eyangoh, S.N., Pape, J.W., Rasolofo-Razanamparany, V., Ridell, M., Rossetti, M.L., Stauffer, F., Suffys, P.N., Takiff, H., Texier-Maugein, J., Vincent, V., De Waard, J.H., Sola, C., Rastogi, N., 2002. Global distribution of *Mycobacterium tuberculosis* spoligotypes. Emerg. Infect. Dis. 8, 1347–1349.

Filliol, I., Motiwala, A.S., Cavatore, M., Qi, W., Hernando Hazbon, M., Bobadilla Del Valle, M., Fyfe, J., Garcia-Garcia, L., Rastogi, N., Sola, C., Zozio, T., Guerrero, M.I., Leon, C.I., Crabtree, J., Angiuoli, S., Eisenach, K.D., Durmaz, R., Joloba, M.L., Rendon, A., Sifuentes-Osornio, J., Ponce de Leon, A., Cave, M.D., Fleischmann, R., Whittam, T.S., Alland, D., 2006. Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. J. Bacteriol. 188, 759–772.

Frothingham, R., Meeker-O'Connell, W.A., 1998. Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. Microbiology 144 (Pt 5), 1189–1196.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C., Small, P.M., 2006. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U.S.A. 103, 2869–2873.

Groenen, P.M., Bunschoten, A.E., van Soolingen, D., van Embden, J.D., 1993. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. Mol. Microbiol. 10, 1057–1065.

Gutacker, M.M., Mathema, B., Soini, H., Shashkina, E., Kreiswirth, B.N., Graviss, E.A., Musser, J.M., 2006. Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites. J. Infect. Dis. 193, 121–128.

Gutacker, M.M., Smoot, J.C., Migliaccio, C.A., Ricklefs, S.M., Hua, S., Cousins, D.V., Graviss, E.A., Shashkina, E., Kreiswirth, B.N., Musser, J.M., 2002. Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains. Genetics 162, 1533–1543.

Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for indows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Hanekom, M., van der Spuy, G.D., Gey van Pittius, N.C., McEvoy, C.R., Hoek, K.G., Ndabambi, S.L., Jordaan, A.M., Victor, T.C., van Helden, P.D., Warren, R.M., 2008. Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism genotyping for analysis of *Mycobacterium tuberculosis* Beijing strains in a setting of high incidence of tuberculosis. J. Clin. Microbiol. 46, 3338–3345.

Helal, Z.H., Ashour, M.S., Eissa, S.A., Abd-Elatef, G., Zozio, T., Babapoor, S., Rastogi, N., Khan, M.I., 2009. Unexpectedly high proportion of ancestral Manu genotype *Mycobacterium tuberculosis* strains cultured from tuberculosis patients in Egypt. J. Clin. Microbiol. 47, 2794–2801.

Hershberg, R., Lipatov, M., Small, P.M., Sheffer, H., Niemann, S., Homolka, S., Roach, J.C., Kremer, K., Petrov, D.A., Feldman, M.W., Gagneux, S., 2008. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. PLoS Biol. 6, e311.

Hirsh, A.E., Tsolaki, A.G., DeRiemer, K., Feldman, M.W., Small, P.M., 2004. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. Proc. Natl. Acad. Sci. U.S.A. 101, 4871–4876.

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M., van Embden, J., 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. 35, 907–914.

Kapur, V., Whittam, T.S., Musser, J.M., 1994. Is *Mycobacterium tuberculosis* 15,000 years old? J. Infect. Dis. 170, 1348–1349.

Kovalev, S.Y., Kamaev, E.Y., Kravchenko, M.A., Kurepina, N.E., Skorniakov, S.N., 2005. Genetic analysis of *Mycobacterium tuberculosis* strains isolated in Ural region, Russian federation, by MIRU-VNTR genotyping. Int. J. Tuberc. Lung Dis. 9, 746–752.

Kwok, P.Y., 2001. Methods for genotyping single nucleotide polymorphisms. Annu. Rev. Genomics Hum. Genet. 2, 235–258.

Le Fleche, P., Fabre, M., Denoeud, F., Koeck, J.L., Vergnaud, G., 2002. High resolution, on-line identification of strains from the *Mycobacterium tuberculosis* complex based on tandem repeat typing. BMC Microbiol. 2, 37.

Mazars, E., Lesjean, S., Banuls, A.L., Gilbert, M., Vincent, V., Gicquel, B., Tibayrenc, M., Locht, C., Supply, P., 2001. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. Proc. Natl. Acad. Sci. U.S.A. 98, 1901–1906.

Millet, J., Miyagi-Shiohira, C., Yamane, N., Sola, C., Rastogi, N., 2007. Assessment of mycobacterial interspersed repetitive unit-QUB markers to further discriminate the Beijing genotype in a population-based study of the genetic diversity of *Mycobacterium tuberculosis* clinical isolates from Okinawa, Ryukyu Islands, Japan. J. Clin. Microbiol. 45, 3606–3615.

Mostowy, S., Cousins, D., Brinkman, J., Aranaz, A., Behr, M.A., 2002. Genomic deletions suggest a phylogeny for the *Mycobacterium tuberculosis* complex. J. Infect. Dis. 186, 74–80.

Musser, J.M., Amin, A., Ramaswamy, S., 2000. Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: evidence of limited selective pressure. Genetics 155, 7–16.

Nakajima, H., 1993. Tuberculosis: a global emergency. World Health Forum 14, 438.

Nouvel, L.X., Dos Vultos, T., Kassa-Kelembho, E., Rauzier, J., Gicquel, B., 2007. A nonsense mutation in the putative anti-mutator gene ada/alkA of *Mycobacterium tuberculosis* and *M. bovis* isolates suggests convergent evolution. BMC Microbiol. 7, 39.

Ralph, A.P., Anstey, N.M., Kelly, P.M., 2009. Tuberculosis into the 2010s: is the glass half full? Clin. Infect. Dis. 49, 574–583.

Rohani, M., Farnia, P., Nasab, M.N., Moniri, R., Torfeh, M., Amiri, M.M., 2009. Beijing genotype and other predominant *Mycobacterium tuberculosis* spoligotypes observed in Mashhad city, Iran. Indian J. Med. Microbiol. 27, 306–310.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Schork, N.J., Fallin, D., Lanchbury, J.S., 2000. Single nucleotide polymorphisms and the future of genetic epidemiology. Clin. Genet. 58, 250–264.

Sebban, M., Mokrousov, I., Rastogi, N., Sola, C., 2002. A data-mining approach to spacer oligonucleotide typing of *Mycobacterium tuberculosis*. Bioinformatics 18, 235–243.

Skuce, R.A., McCorry, T.P., McCarroll, J.F., Roring, S.M., Scott, A.N., Brittain, D., Hughes, S.L., Hewinson, R.G., Neill, S.D., 2002. Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets. Microbiology 148, 519–528.

Sorek, R., Kunin, V., Hugenholtz, P., 2008. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat. Rev. Microbiol. 6, 181–186.

Sreevatsan, S., Pan, X., Stockbauer, K.E., Connell, N.D., Kreiswirth, B.N., Whittam, T.S., Musser, J.M., 1997. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. Proc. Natl. Acad. Sci. U.S.A. 94, 9869–9874.

Stavrum, R., Mphahlele, M., Ovreas, K., Muthivhi, T., Fourie, P.B., Weyer, K., Grewal, H.M., 2009. High diversity of *Mycobacterium tuberculosis* genotypes in South Africa and preponderance of mixed infections among ST53 isolates. J. Clin. Microbiol. 47, 1848–1856.

Styblo, K., Bumgarner, J., 1991. Tuberculosis can be controlled with existing technologies: evidence. Tuberculosis Surveillance Research Unit of the IUATLD Progress Report 2, 60–72.

Supply, P., Allix, C., Lesjean, S., Cardoso-Oelemann, M., Rusch-Gerdes, S., Willery, E., Savine, E., de Haas, P., van Deutekom, H., Roring, S., Bifani, P., Kurepina, N., Kreiswirth, B., Sola, C., Rastogi, N., Vatin, V., Gutierrez, M.C., Fauville, M., Niemann, S., Skuce, R., Kremer, K., Locht, C., van Soolingen, D., 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 44, 4498–4510.

Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D., Locht, C., 2001. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. J. Clin. Microbiol. 39, 3563–3571.

Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B., Locht, C., 2000. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol. Microbiol. 36, 762–771.

Supply, P., Warren, R.M., Banuls, A.L., Lesjean, S., Van Der Spuy, G.D., Lewis, L.A., Tibayrenc, M., Van Helden, P.D., Locht, C., 2003. Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. Mol. Microbiol. 47, 529–538.

Takashima, T., Iwamoto, T., 2006. New era in molecular epidemiology of tuberculosis in Japan. Kekkaku 81, 693–707.

van Embden, J.D., Cave, M.D., Crawford, J.T., Dale, J.W., Eisenach, K.D., Gicquel, B., Hermans, P., Martin, C., McAdam, R., Shinnick, T.M., et al., 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. 31, 406–409.

van Embden, J.D.A., van Gorkom, T., Kremer, K., Jansen, R., van der Zeijst, B.A.M., Schouls, L.M., 2000. Genetic variation and evolutionary origin of the Direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. J. Bacteriol. 182, 2393–2401.

Vitol, I., Driscoll, J., Kreiswirth, B., Kurepina, N., Bennett, K.P., 2006. Identifying *Mycobacterium tuberculosis* complex strain families using spoligotypes. Infect. Genet. Evol. 6, 491–504.

Warren, R.M., Streicher, E.M., Sampson, S.L., Van Der Spuy, G.D., Richardson, M., Nguyen, D., Behr, M.A., Victor, T.C., Van Helden, P.D., 2002. Microevolution of the direct repeat region of *Mycobacterium tuberculosis*: implications for interpretation of spoligotyping data. J. Clin. Microbiol. 40, 4457–4465.

Wirth, T., Hildebrand, F., Allix-Beguec, C., Wolbeling, F., Kubica, T., Kremer, K., van Soolingen, D., Rusch-Gerdes, S., Locht, C., Brisse, S., Meyer, A., Supply, P., Niemann, S., 2008. Origin, spread and demography of the *Mycobacterium tuberculosis* complex. PLoS Pathog. 4, e1000160.

Zhang, J., Abadia, E., Refregier, G., Tafaj, S., Boschiroli, M.L., Guillard, B., Andremont, A., Ruimy, R., Sola, C., 2010. *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay. J. Med. Microbiol. 59, 285–294.

Zozio, T., Allix, C., Günal, S., Saribas, S., Alp, A., Durmaz, R., Fauville-Dufaux, M., Rastogi, N., Sola, C., 2005. The previously identified "LAM7" genotype typifies a Turkish-phylogeographically specific clonal complex of *Mycobacterium tuberculosis*, 26th Annual Congress of the European Society for Mycobacteriology June 26th–29th 2005. Abstract Book, O-8 *presenting author, Istanbul, Turkey, 26–29 June 2005, p. 31.

Radiol Phys Technol (2011) 4:13–18
DOI 10.1007/s12194-010-0099-5

Investigation of optimal viewing size for detecting nodular ground-glass opacity on high-resolution computed tomography with cine-mode display

Michihiro Yamaguchi · Yuichi Bessho · Tatsuro Inoue · Yoshiyuki Asai · Tomoshige Matsumoto · Kenya Murase

Received: 6 May 2010 / Revised: 31 July 2010 / Accepted: 3 August 2010 / Published online: 19 August 2010

Abstract We evaluated the effect of the displayed image sizes on observers' ability to detect nodular ground-glass opacity (n-GGO) on CT and investigated the optimal viewing size for soft-copy reading at CT screening for lung cancer. A total of 46 patients' high-resolution computed tomography (HRCT) images (22 patients with one GGO; 24 without GGO) were displayed on a monochromatic liquid crystal display monitor at a resolution of 1,200 × 1,600. HRCT was presented on the screen with cine-mode display. We compared two viewing sizes (original size, i.e., the image displayed with a zoom factor of 1 in which each pixel value in the image is displayed as one pixel on the display: 13 cm × 13 cm; fit size, i.e., by zooming the captured image until it occupies the entire screen: 30 cm × 30 cm) in terms of radiologists' performance for detecting n-GGO on HRCT and the viewing times required for soft-copy reading decisions. Observer performance was analyzed in terms of the receiver operating characteristic (ROC) curve. A statistically significant improvement was found with the original size in the average area-under-the-ROC curve values for the accuracy of diagnosis and the viewing times compared to the fit size ($P < 0.05$). The original size with cine-mode display leads to increased lung GGO detection at CT screening for lung cancer, and the reduced time spent performing the diagnosis offers cost savings.

Keywords CT, high resolution · Computer applications, detection · Lung · Observer performance · Screening

M. Yamaguchi (✉) · Y. Bessho · T. Inoue
Department of Radiology, Osaka Prefectural Medical Center for Respiratory and Allergic Disease, Osaka Prefectural Hospital Organization, 3-7-1 Habikino, Habikino, Osaka 583-8588, Japan
e-mail: yamami@iris.eonet.ne.jp

Y. Bessho
e-mail: betsushiyoy@opho.jp

T. Inoue
e-mail: inouetat@opho.jp

Y. Asai
Department of Radiology, Kinki University Hospital, 377-2, Ono-Higashi, Osaka-Sayama, Osaka 589-8511, Japan
e-mail: asai-mail@muf.biglobe.ne.jp

T. Matsumoto
Department of Clinical Research and Development, Osaka Prefectural Medical Center for Respiratory and Allergic Disease, Osaka Prefectural Hospital Organization, 3-7-1 Habikino, Habikino, Osaka 583-8588, Japan
e-mail: matsumototo@opho.jp

K. Murase
Division of Medical Technology and Science, Department of Medical Physics and Engineering, Course of Health Science, Graduate School of Medicine, Osaka University, 1-7 Yamadaoka, Suita, Osaka 565-0871, Japan
e-mail: murase@sahs.med.osaka-u.ac.jp

1 Introduction

With the increasing popularity of multi-detector computed tomography (MDCT), viewing and reporting of CT images from monitors has become common practice in medicine. In particular, chest CT examination is, in most cases, performed by use of MDCT, which produces the very large data sets per patient. A large number of images produced by lung MDCT are usually presented on the screen in cine viewing.

When radiologists observe CT images on the monitor by using cine mode, viewing sizes are changed according to their preferences. For example, radiologists frequently use

the following two different formats as usual display sizes for soft-copy reading in our hospital: the original size, which means that the image is displayed with a zoom factor of 1, whereby each pixel value in the image is displayed as one pixel on the monitor screen; and an enlarged size, obtained when the captured image is zoomed until it occupies the entire screen (referred to as "fit size" hereinafter). The reason why radiologists in our hospital tend to use the two viewing sizes is that the functions of displaying the original size and the fit size are preset as toolbar icons or as the main menu on general viewer software. Therefore, only by clicking the relevant icon do radiologists observe a CT image in the center of the monitor with the cine-mode display. According to our recent investigation, either of the two display methods was used in the daily work by the reviewers of our prefectural hospital organization. Although viewing sizes of cine mode vary according to radiologists' taste, they require the optimum viewing size for soft-copy reading for efficient detection and to avoid spending unnecessary time and energy.

According to a study conducted by Li et al. [1], radiologists frequently failed to detect lung cancers with GGO upon screening CT. Nakata et al. [2] reported that GGO persisting for several months is an indicator of early adenocarcinoma or its precursor. This was the reason why we selected GGO lesions in our study.

Several studies demonstrated the influence of the image display size on observer performance [3–6]. In a previous paper [7, 8], we reported the usefulness of the reduced display size in soft-copy diagnosis by using clinical chest X-ray images and a digital contrast detail phantom. However, to date, the optimal size for viewing of lung HRCT images that are displayed in cine mode on the LCD monitor has not been clarified. Our purpose in the present study was to investigate the effect of the image size on the accuracy in detecting ground-glass opacities and differences in time for reviewing with cine-mode display at CT screening for lung cancer.

2 Materials and methods

2.1 Case selection

Case selection was done retrospectively with the consensus of two chest radiologists who had more than 10 and 15 years of CT imaging experience and who did not participate in this observer performance study. HRCT images of 46 patients were selected based on the presence (22 patients) of nodular GGO (n-GGO) lesions (17 nodules had pure GGO, nonsolid; 5 had mixed GGO, part solid) and the absence of GGO lesions (24 patients) from chest CT examinations data that were obtained at the Osaka Prefectural Medical Center for Respiratory and Allergic Diseases during a 2.2-year period from April 2005 to June 2007. The 46 patients consisted of 21 men and 25 women. The mean age of the men was 58.4 years (age range 18–82 years), and the mean age of the women was 51.8 years (age range 26–70 years). The 22 patients who were suspected of having GGO lesions at HRCT were followed by CT for at least 3 months (follow-up period, 3 months to 2 years). On HRCT, the diameters of these 22 GGO lesions at the time of discovery ranged from 4 mm× 5 mm to 11 mm × 19 mm. The mean size of these 22 GGO lesions was 11.6 ± 4.5 mm in the largest diameter (range 5–19 mm). The demographic characteristics of the 22 patients with GGO are presented in Table 1.

All chest CT examinations were performed with use of a Sensation-16 scanner (Siemens Medical Solutions, Forchheim, Germany) without administration of contrast medium. A dose-reduction system (Care Dose 4D) was available on the Sensation-16 scanner. The mean values ± SD of the CT dose index volume (CTDIvol) for all standard-dose chest CT were 9.86 ± 1.76 mGy lower than that in the proposed "ICRP Draft Report on MDCT Posted

Table 1 Patient demographics

Age (years)	Gender	Location (lobe)	Size (mm^2)	Nodule type
61	F	S4 (RM)	10 × 15	PGGO
40	F	S4 (RM)	6 × 6	PGGO
62	M	S1 + 2 (LU)	9 × 10	PGGO
70	F	S1 (RU)	14 × 15	MGGO
64	F	S1 (RU)	9 × 9	PGGO
52	F	S1 (RU)	10 × 12	PGGO
67	F	S6 (RL)	8 × 15	PGGO
51	M	S8 (LL)	8 × 9	PGGO
56	M	S1 (RU)	9 × 10	PGGO
62	F	S1 (RU)	9 × 10	PGGO
54	M	S2 (RU)	14 × 16	MGGO
46	M	S4 (RM)	6 × 9	PGGO
67	M	S9 (RL)	7 × 9	PGGO
68	F	S1 + 2 (LU)	5 × 5	PGGO
54	F	S9 (RL)	11 × 19	MGGO
61	F	S1 + 2 (LU)	9 × 16	PGGO
69	M	S1 (RU)	11 × 17	MGGO
78	M	S6 (RL)	14 × 18	MGGO
42	F	S1 (RU)	4 × 6	PGGO
62	M	S9 (RL)	4 × 5	PGGO
64	F	S6 (RL)	9 × 12	PGGO
57	F	S1 (RU)	5 × 6	PGGO

PGGO pure GGO, *MGGO* mixed GGO, *M* male, *F* female, *S* segment, *RU* right upper, *RM* right middle, *RL* right lower, *LU* left upper, *LL* left lower

for Public Consultation" reference value of 10.1 mGy (routine chest) [9].

Helical CT images were reconstructed through the entire GGO with a 1–2 cm margin of normal lung parenchyma by use of a detector row width of 0.75 mm, volume pitch of 0.9, 2.0 mm section thickness and intervals, 0.75-s rotation time, 120 kVp, and an automatic tube current modulation technique. The image matrix size was 512 × 512 pixels. The field of view was adjusted to be optimized for the size of the patient and ranged from 300 to 365 mm. The images were displayed at a window level of −650 HU and a window width of 1,800 HU with a sharp kernel (B70f) for image reconstruction of the lung.

The observers read the extracted images with a monochrome LCD monitor that had a native resolution of 1,200 × 1,600 pixels and an active display size of 298.9 mm × 398.4 mm (Radiforce G22; Eizo, Ishikawa, Japan) with use of a Digital Imaging and Communications in Medicine viewer (Pop-net Essential; Image One, Tokyo, Japan). The observation conditions were as follows: the actual maximum luminance of the LCD monitor was set at 400 cd/m^2, ambient lighting conditions were 20–50 lx, and the characteristic curve of the monitor was set according to part 14 of the DICOM standard (calibrated with the Grayscale Standard Display Function: GSDF) [10]. We confirmed that the LCD monitor for observation conformed to the standards proposed by the American Association of Physicists in Medicine Task Group 18 (AAPM-TG18) [11]. When CT images were displayed, the surrounding screen was black, i.e., the caption for the CT image was not displayed. The hard/software of our viewing system performed equally fast when displaying the two sizes. Consequently, each image was displayed on the monitor in real time.

2.2 Observer performance study

Our institutional review board does not require approval or patient informed consent for retrospective review of previously obtained image data. Patient confidentiality, however, was protected. Informed consent for the observer performance study was obtained from all observers. Seven board-certified radiologists who specialized in chest radiology participated in our study as observers. The board-certified radiologists had 9–28 years of experience (mean 19.6 years).

To expedite the observer performance test, we chose positive samples that were obtained from a few HRCT images (10–30 images per case) containing the entire GGO and a margin of normal parenchyma. On the other hand, negative samples that almost agreed with the segmentation of GGO tumors were chosen. Consequently, 719 images of 46 patients were prepared for ROC study. Each observer read thin slice CT images independently by using cine mode for displaying and rated his or her confidence in determining the presence or absence of a GGO lesion. The speed or sequence of the image display for cine mode was controlled manually by the observer. All observers read the test image in order to become familiarized with the appearance of abnormalities before ROC examination. The observers were not allowed to adjust the contrast, brightness, window width, or window level of CT images because the display conditions were adjusted properly so that all observers might minimize the influence of observation conditions. No limit was imposed on the reading time and viewing distance, although the viewing time was recorded. To reduce learning effects, we required that there was an interval of 4 weeks at minimum between the two interpretation sessions (observations of original size and fit size). We decided by using a table of random numbers which size to observe first.

The observers were required to indicate only the presence or absence of a GGO lesion without characterizing the lesion. The observers' viewing image was adjusted as original size (13 cm × 13 cm), i.e., the image displayed with a zoom factor of 1 in which each pixel value in the image is displayed as one pixel on the display, and fit size (30 cm × 30 cm), i.e., by zooming of the captured image until it occupies the entire screen. Each observer interpreted the cases (original size and fit size) in a different order to avoid bias. Next, observers used a continuous rating scale to rate the confidence level by marking on a line 5 cm in length.

2.3 Statistical analysis

For ROC analysis, we used the computer program DBM MRMC 2.1 Beta 3 (Metz CE, The University of Chicago) [12], which uses analyses of variance in pseudo-values of the area under the best-fit binormal ROC curve (AUC) calculated from all rating scores of all radiologists. The viewing time data were analyzed by Wilcoxon signed-rank test, which uses the computer program SPSS 11.0 for Windows (SPSS Inc., Chicago, Illinois, USA).

3 Results

Figure 1 shows the averaged ROC curve of seven radiologists for both original and fit size images. The AUC values for the seven observers with the two sizes of images (original size, 13 cm × 13 cm; fit size, 30 cm × 30 cm) are shown in Table 2. The estimated mean ± SD of AUC values was 0.988 ± 0.006 for original size images and 0.930 ± 0.059 for fit size images. All of the AUCs of the original size images appeared to be larger than those of the

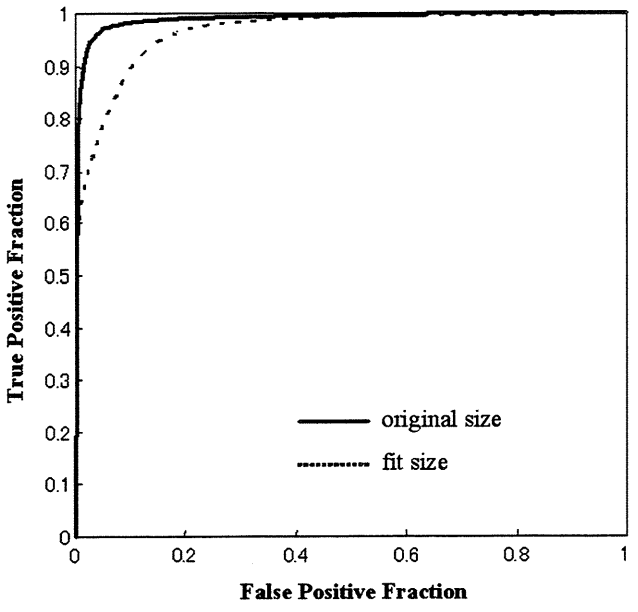

Fig. 1 *Graph* shows averaged ROC curves of seven board-certified radiologists for both sizes. All of the areas under the ROC curve of original size images appeared to be larger than those of the fit size images

Table 2 Area under the curve (AUC) values for the seven observers with the two sizes of images (original size, 13 cm × 13 cm; fit size, 30 cm × 30 cm)

Observer	Original	Fit
1	0.995	0.989
2	0.994	0.991
3	0.982	0.919
4	0.977	0.948
5	0.992	0.819
6	0.990	0.941
7	0.987	0.901
Mean	0.988	0.930
SD	0.006	0.059

The estimated mean ± SD of AUC values was 0.988 ± 0.006 for original size images and 0.930 ± 0.059 for fit size images. A statistically significant difference was found between the two groups ($P = 0.048$)

fit size images. A statistically significant difference was found between the two groups ($P = 0.048$).

Table 3 shows the viewing times for the two groups. The average total viewing time was 876 s for original size images and 2,268 s for fit size images. The viewing time of the original size images was significantly shorter (about one-third) than that of the fit size images. A statistically significant difference was also found in viewing times between the two categories with the Wilcoxon signed-rank test ($P = 0.016$).

Table 3 Total viewing time (seconds) of the seven observers

Observer	Original	Fit
1	464	1,280
2	678	2,000
3	714	1,670
4	840	1,652
5	1,192	3,925
6	804	1,510
7	1,437	3,837
Mean	876	2,268
SD	1,040	306

The viewing time for original size images was significantly shorter (about one-third) than that for fit size images. A statistically significant difference was also found in viewing times between the two groups with Wilcoxon signed-rank test ($P = 0.016$)

4 Discussion

All observers in our study emphasized that the original size image was easier to interpret than the fit size image. To add to this opinion, the radiologists insisted that it took a longer time to observe a large image than a small image because the viewing area of a fit image became wider than that of an original image. In our study, a statistically significant difference was found between the two groups (original and fit sizes) in terms of observer performance. This result suggests that the original size is not as much affected adversely by the technique due to zooming and interpolation as is the fit size.

To add to this fact, the viewing time for the original size images was significantly shorter (about one-third) than that for fit size images. Krupinski et al. [13] reported that shortening of the diagnostic viewing time is very important for avoiding radiologists' inconvenience. Incidentally, not a few radiologists think, on the basis of their experience, that the magnified image size is more useful than the original image size.

Several studies demonstrated the influence of image display size on observer performance with cine-mode display [14–16]. Gur et al. [5] who reported the effect of image display size on observer performance showed that sequential viewing of enlarged CT images for the detection of abdominal masses did not improve the observer performance and increased the reader variability. They discussed the relationship between visual angle and fixation. Seltzer et al. [16] performed a comparative study to find lung nodules with both cine and conventional film-based viewing. The authors reported that cine viewing (24.1 cm × 27.9 cm) of chest CT images improved radiologists' ability to detect nodules. They did not discuss the speed or sequence of the image display for cine mode;

however, the speed of cine mode might influence the observer performance.

On the other hand, the viewing distance was shown by Seltzer et al. to affect the detection performance. They reported the observer's eye compensation of the smaller image by adjusting the viewing distance [6].

Flynn [17] described a typical distance for viewing of medical images. He reported that devices with 0.250 mm pixels were shown to be well matched to visual performance when a 60-cm viewing distance was used. The distance of observation was unrestricted in our study, but the observers' average distance from the eye to the LCD monitor was 60 cm (55–65). When we observe a CT image on a display device with 0.25 mm pixels, the active original image size is 12.8 cm × 12.8 cm. This size is close to the size (13 cm × 13 cm) on a 2 Mega-pixels (MP) display.

Our study demonstrated that the original size was superior to the fit size for detection of n-GGOs in lung cancer screening with cine mode. For radiologists to work in lung cancer screening on 1, 2, and 3 MP LCD monitors, the original size should be recommended as the proper size for viewing because one pixel of a CT image in its original size equals one pixel on the monitor screen. However, we should adjust the active original image to the proper size when we observe at 5-MP resolution or more because the scale becomes smaller than the vertical size of the human effective field of vision (9.9 cm) [18]. Radiologists should use the original size first in lung cancer screening, and, if necessary, the fit size or magnified size should be used next according to their preference. This reading method is useful for the detection of small n-GGOs in chest CT for lung cancer screening, where double reading is recommended [19].

It should be noted that this study had some limitations. First, although we chose GGO lesions as true-positive samples, radiologists must detect not only GGO lesions, but also severe intestinal disease and pulmonary fibrosis, diffuse bronchiectasis, and inflammatory scars; however, it is very important to find hazy GGO lesions in lung cancer screening for radiologists. In our study, the case selection of pure GGOs may have led to a statistically significant difference. Second, we used only two sizes (original size, 13 cm × 13 cm; fit size, 30 cm × 30 cm) as viewing images; however, other sizes should be evaluated. We must investigate the optimal viewing size for detecting n-GGOs from the viewpoint of the human visual system in the future. Third, other factors were not evaluated such as the display luminance and display function, which also influence observer performance [20–23]; however, GSDF is recommended as a standard display function for soft-copy diagnosis [24]. Fourth, we did not use a low-dose protocol for lung cancer CT screening. We must consider this problem in the future. Fifth, we did not evaluate the relationship of viewing size to asthenopia. The relationship between reading time and asthenopia is important, too. Sixth, the observers were not allowed to adjust the display conditions in our study; however, they might do the evaluation differently by under other observation conditions. Finally, we must evaluate the viewing image size for other modalities [magnetic resonance imaging (MRI), positron emission tomography (PET), etc.] in the future. Niimi et al. [14] reported on the effectiveness of display methods for magnetic resonance image interpretation by using an eye-tracking device. From the viewpoint of the radiologist's reading process, any other browsing methods may be necessary for different imaging modalities. Nonetheless, the present study demonstrated the benefit of original size on the accuracy in detecting ground-glass opacities.

5 Conclusion

The present study has shown a statistically significant difference in AUC values for the accuracy of diagnosis and reading times between original and fit size HRCT images with ROC analysis of radiologists' performance for n-GGO.

Original size with cine-mode display leads to increased detection of lung nodular GGO at CT screening for lung cancer: interpretation of n-GGO is easier because the pixel value of CT image retains its original size on the monitor screen, and the reduced time that is spent performing detection offers cost savings.

Acknowledgments The authors sincerely thank Drs Fukuda H, Higami T, Oonishi M, Ninoi T, Tsumura M, Daikokuya H, Sasada S, Awai K, Kobayashi S, and Tanaka S, for participating as reviewers and observers. This study was partly supported by a grant from the Japanese Ministry of Health, Labour and Welfare, no. H18-Shinko-011.

References

1. Li F, Sone S, Abe H, MacMahon H, Armato SG 3rd, Doi K. Lung cancers missed at low dose helical CT screening in a general population: comparison of clinical, histopathologic, and imaging findings. Radiology. 2002;225:673–83.
2. Nakata M, Saeki H, Takata I, Segawa Y, Mogami H, Mandai K, Eguchi K. Focal ground-glass opacity detected by low-dose helical CT. Chest. 2002;121:1464–7.
3. Schaefer CM, Prokop M, Oestmann JW, Wiesmann W, Haubitz B, Meschede A, et al. Impact of hard-copy size on observer performance in digital chest radiography. Radiology. 1992;184: 77–81.
4. Fisher PD, Brauer GW. Impact of image size on effectiveness of digital imaging systems. J Digital Imaging. 1989;2:39–41.
5. Gur D, Klym AH, King JL, Maitz GS, Mello-Thoms C, Rockette HE, et al. The effect of image display size on observer

performance an assessment of variance components. Acad Radiol. 2006;13:409–13.
6. Seltzer SE, Judy PF, Feldman U, Scarff L, Jacobson FL. Influence of CT image size and format on accuracy of lung nodule detection. Radiology. 1998;206:617–22.
7. Bessho Y, Yamaguchi M, Fujita H, Azuma M. Usefulness of reduced image display size in softcopy reading: evaluation of lung nodules in chest screening. Acad Radiol. 2009;16:940–6.
8. Yamaguchi M, Fujita H, Bessho Y, Inoue T, Asai Y, Murase K. Investigation of optimal display size for detecting ground-glass opacity on high resolution computed tomography using a new digital contrast-detail phantom. Eur J Radiol (Ireland). May 10 2010. (epub ahead of print) (record supplied by publisher)
9. International Commission on Radiological Protection. Managing patient dose in multi-detector computed tomography (MDCT). ICRP Publication 2007; 102.
10. Digital Imaging and Communications in Medicine (DICOM) part 14: Grayscale standard display function. Rosslyn, VA: National Electric Manufacturers Association. 1999; pp. 1–14.
11. Tanaka H. Guideline for quality assurance of medical imaging display monitors (In Japanese). Jpn J Radiol Technol. 2006;62: 508–51.
12. Roe CA, Metz CE. The Dorfman–Berbaum–Metz method for statistical analysis of multi-reader, multi-modality ROC data: validation by computer simulation. Acad Radiol. 1997;4:298.
13. Krupinski EA, Maloney K, Bessen SC, Capp MP, Graham K, Hunt R, et al. Receiver operating characteristic evaluation of computer display of adult portable chest radiographs. Invest Radiol. 1994;29:147–9.
14. Niimi R, Shimamoto K, Sawaki A, Ishigaki T, Takahashi Y, Sugiyama N, et al. Eye-tracking device comparisons of three methods of magnetic resonance image series displays. J Digit Imaging. 1997;10:147–51.
15. Beard DV, Molina PL, Muller KE, Denelsbeck KM, Hemminger BM, Perry JR, et al. Interpretation time of serial chest CT examinations with stacked-metaphor workstation versus film alternator. Radiology. 1995;197:753–8.
16. Seltzer SE, Judy PF, Adams DF, Jacobson FL, Stark P, Kikinis R, et al. Spiral CT of the chest: comparison of cine and film-based viewing. Radiology. 1995;197:73–8.
17. Flynn MJ. Visual requirements for high-fidelity display. Advances in digital radiography: RSNA categorical course in diagnostic radiology physics 2003; pp. 103–7.
18. Hatada T. Eyeball movement and glasses. Jpn J Glasses Sci. 1977. http://www.kokuyo.co.jp/press/2005/09/433.html.
19. Wormanns D, Ludwig K, Beyer F, Heindel W, Diederich S. Detection of pulmonary nodules at multirow-detector CT: effectiveness of double reading to improve sensitivity at standard-dose and low-dose chest CT. Eur Radiol. 2005;15:14–22.
20. Fujita H, Yamaguchi M, Bessho Y, Uemura M, Abe S, Nakahira S, et al. Analysis of the display function on viewing monitors for the digital diagnostic imaging modality. Med. Imaging Inf. Sci. 2005;22:153–9. (in Japanese with English abstract)
21. Uemura M, Asai Y, Yamaguchi M, Fujita H, Shintani Y, Sanada S. Psychophysical evaluation of calibration curve for diagnostic LCD monitor. Radiat Med. 2006;24:653–8.
22. Goo JM, Choi JY, Im JG, Lee HJ, Chung MJ, Han D, et al. Effect of monitor luminance and ambient light on observer performance in soft-copy reading of digital chest radiographs. Radiology. 2004;232:762–6.
23. Yamaguchi M, Fujita H, Asai Y, Uemura M, Ookura Y, Matsumoto M, et al. Psychophysical analysis of monitor display functions affecting observer diagnostic performance of CT image on liquid crystal display monitors. Eur Radiol. 2005;15:2487–96.
24. Samei E, Badano A, Chakraborty D, Compton K, Cornelius C, Corrigan K, et al. Assessment of display performance for medical imaging systems: executive summary of AAPM TG18 report. Med Phys. 2005;32:1205–25.

ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Sept. 2011, p. 4352–4360 Vol. 55, No. 9
0066-4804/11/$12.00 doi:10.1128/AAC.01274-10

Genetic Organization and Mode of Action of a Novel Bacteriocin, Bacteriocin 51: Determinant of VanA-Type Vancomycin-Resistant *Enterococcus faecium*▽†

Hitoshi Yamashita,[1] Haruyoshi Tomita,[1,2] Takako Inoue,[1] and Yasuyoshi Ike[1,2]*

Department of Bacteriology[1] and Laboratory of Bacterial Drug Resistance,[2] Gunma University Graduate School of Medicine, Maebashi, Gunma 371-8511, Japan

Received 21 September 2010/Returned for modification 3 November 2010/Accepted 17 June 2011

Bacteriocin 51 (Bac 51) is encoded on the mobile plasmid pHY (6,037 bp), which was isolated from vancomycin-resistant *Enterococcus faecium* VRE38. Bacteriocin 51 is active against *E. faecium*, *E. hirae*, and *E. durans*. Sequence analysis of pHY showed that it encodes nine open reading frames (ORFs) from ORF1 to ORF9 (in that order). Genetic analysis suggested that ORF1 and ORF2, which were designated *bacA* and *bacB*, respectively, are the bacteriocin and immunity genes. *bacA* encodes a 144-amino-acid protein. The deduced BacA protein has a typical signal sequence at its amino terminus, and a potential signal peptidase-processing site corresponding to the V-E-A sequence is located between the 37th and 39th amino acids. The predicted mature BacA protein consists of 105 amino acids. A potential promoter sequence was identified upstream of the start codon. *bacB* encodes a 55-amino-acid protein. No obvious promoter or terminator sequence was identified between *bacA* and *bacB*. Northern blot analysis of *bacA* and *bacB* with a *bacA* RNA probe produced a transcript of approximately 700 nucleotides, which corresponded to the combined nucleotide sizes of *bacA* and *bacB*, indicating that transcription was initiated from the promoter upstream of *bacA*, continued through *bacB*, and was terminated at the terminator downstream of *bacB*. The transcription start site was determined to be the T nucleotide located 6 nucleotides downstream from the −10 promoter sequence. These results indicate that *bacA* and *bacB* constitute an operon and that *bacA* is the bacteriocin structural gene while *bacB* is the immunity gene. The purified C-terminally His tagged BacA protein of Bac 51 showed bacteriostatic activity against the indicator strain. The purified C-terminally His tagged BacA protein of Bac 32 (whose mature BacA protein has 54 amino acids) and the culture filtrates of the Bac 31- and Bac 43-producing *E. faecalis* strain FA2-2 showed bactericidal activity. Bac 31 and Bac 43 are pore-forming bacteriocins, unlike the newly characterized bacteriocin Bac 51.

Bacteriocins are produced by a wide variety of Gram-positive and Gram-negative bacteria. They are bacterial proteins that inhibit the growth of bacteria closely related to the producer strain, and they usually exhibit a relatively narrow spectrum of activity. Bacteriocins are thought to provide the producer strain with an ecological or selective advantage over other strains. Bacteriocin production has been described for several genera of lactic acid bacteria (LAB) (10). LAB bacteriocins can be divided into two main classes. Class I consists of modified bacteriocins (the lantibiotics), and class II consists of the small heat-stable nonlantibiotics (27). Class II bacteriocins are further divided into subgroup IIa, comprising pediocin-like bacteriocins with strong antilisterial effects, and subgroup IIb, comprising non-pediocin-like bacteriocins composed of two peptides and requiring the complementary activity of both peptides for full antimicrobial activity. In the genus *Enterococcus*, *Enterococcus faecalis* and *E. faecium* bacteriocins have been well characterized genetically and biochemically. *E. faecalis* bacteriocins include the beta-hemolysin/bacteriocin (cytolysin) (8, 9, 17–22), the peptide antibiotic AS-48 (26), bacteriocin 21 (Bac 21) (16, 33), and bacteriocin 31 (Bac 31) (32). These *E. faecalis* bacteriocins have been identified in clinical isolates (8, 26, 31–33). The well-characterized *E. faecium* bacteriocins have been identified in food-grade organisms (7). These include enterocins A (2), B (4), P (6), I (13), and L50A and L50B (5). These bacteriocins belong to LAB class II bacteriocins and are active against *Listeria monocytogenes* (27). Enterocins A and P are pediocin-like bacteriocins (27).

Little is known about the bacteriocins present in *E. faecium* clinical isolates. Previously, we were the first to report the isolation and characterization of bacteriocin 32 (Bac 32), which is encoded on mobile plasmid pTI1 (12.5 kbp) of VRE200, a vancomycin-resistant *E. faecium* (VRE) clinical isolate (23). Bac 32 is active against *E. faecium*, *E. hirae*, and *E. durans* but has shown no activity against *L. monocytogenes*, unlike the bacteriocins identified in food-grade *E. faecium* strains. A high frequency of clinical isolates (about 40%) encode Bac 32-type bacteriocins (23). In this study, we describe the genetic analysis of a novel bacteriocin, designated bacteriocin 51 (Bac 51), that was identified in a VRE clinical isolate. We show that Bac 51 is active against *E. faecium*, *E. hirae*, and *E. durans* but not against *L. monocytogenes*.

* Corresponding author. Mailing address: Department of Bacteriology and Laboratory of Bacterial Drug Resistance, Gunma University Graduate School of Medicine, 3-39-22 Showa-Machi, Maebashi, Gunma 371-8511, Japan. Phone: 81-27-220-7990. Fax: 81-27-220-7996. E-mail: yasuike@med.gunma-u.ac.jp.

† Supplemental material for this article may be found at http://aac.asm.org/.

▽ Published ahead of print on 27 June 2011.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Genotype or phenotype	Description	Reference or source
Strains			
Enterococcus faecalis			
FA2-2	Rif^r Fus^r	Derivative of JH2	31
OG1S	Str^r	Derivative of OG1	11
Enterococcus faecium			
BM4105RF	Rif^r Fus^r	Derivative of plasmid-free *E. faecium* BM4105	35
BM4105SS	Str^r Spc^r	Derivative of plasmid-free *E. faecium* BM4105	35
VRE38	pHY (Bac 51) Van^r Gm^r Em^r Tc^r	Bacteriocinogenic clinical isolate	This study
Enterococcus hirae ATCC 9790			
Enterococcus durans ATCC 49135			
Enterococcus raffinosus JCM8733			
Enterococcus gallinarum BM4174			
Staphylococcus aureus FDA209P			
Escherichia coli			
DH5α	*recA1 endA1 gyrA96 thi-1 relA1 hsdR17 supE44* Δ(*argE-lacZYA*)*U169*		Bethesda Research Laboratories
TH688	CSH57b *thr*::Tn*5*		
BL21	F$^-$ *ompT hsdS*(r_B^- m_B^-) *gal dcm* (DE3)	Protein expression	Novagen
Plasmids			
pHY	Bac 51	Mobilizable plasmid (6.0 kb)	This study
pTI1	Bac 32	Mobilizable plasmid (12.5 kb)	23
pYI17	Bac 31	Conjugative plasmid (57.5 kb)	32
pMG502	Bac 43	pAM401 containing ORF1 and ORF2 of pDT1	30
pAM401	Cm^r Tc^r	*E. coli-E. faecalis* shuttle vector	36
pMW119	Amp^r *lacZ*	*E. coli* cloning vector; low copy number	Nippon Gene Co.
pUC19	Amp^r *lacZ*	*E. coli* cloning vector	Nippon Gene Co.
pET22b(+)	Amp^r; His_6 affinity tag	Overexpression vector	Novagen

MATERIALS AND METHODS

Bacteria, media, and reagents. The laboratory strains and plasmids used in this study are listed in Table 1. Six VanA-type vancomycin-resistant *Enterococcus faecium* (VRE) clinical isolates were obtained from different patients who had been admitted to a hospital in Japan. A total of 138 vancomycin-sensitive *E. faecium* and 120 vancomycin-sensitive *E. faecalis* clinical isolates were included in the study. These strains were obtained from different patients at the Gunma University School of Medicine Hospital, Maebashi, Gunma, Japan. In addition, 87 VRE clinical isolates were obtained from different patients in Japanese hospitals in 2002, and 662 VRE isolates were obtained at the University of Michigan Medical School Hospital, Ann Arbor, between 1994 and 1999 (35). Enterococcal strains were grown in Todd-Hewitt broth (THB; Difco, Detroit, MI). *Escherichia coli* strains were grown in Luria-Bertani (LB) medium. Solid or soft medium was prepared by the addition of 1.5 or 0.75% (wt/vol) agar, respectively. All cultures were grown at 37°C. The following antibiotics were used at the indicated concentrations: ampicillin, 100 μg/ml; rifampin, 25 μg/ml; fusidic acid, 25 μg/ml; streptomycin, 500 μg/ml; spectinomycin, 500 μg/ml for enterococci and 50 μg/ml for *E. coli*; kanamycin, 40 μg/ml; tetracycline, 12.5 μg/ml; vancomycin, 5 μg/ml; and gentamicin, 200 μg/ml.

Antimicrobial susceptibility testing. The MICs of the antibiotics were determined by the agar dilution method. A pure overnight culture of each strain grown in Mueller-Hinton broth (Nissui, Tokyo, Japan) was diluted 100-fold with fresh broth. An inoculum of approximately 5×10^5 cells was plated on a series of Mueller-Hinton agar (Eiken, Tokyo, Japan) plates containing a range of concentrations of the test drug. The plates were incubated at 37°C, and the susceptibility results were finalized after 24 h of incubation. Susceptibility testing and interpretation of the results were in compliance with the standards of the Clinical and Laboratory Standards Institute (formerly NCCLS). *E. hirae* ATCC 9790 was used as a control strain.

Detection of bacteriocin production and immunity. To detect antimicrobial activity, a soft-agar assay was performed as described previously (20). The test for immunity to the bacteriocin was performed essentially as described previously (20). To examine the bacteriocin activities of liquid samples (e.g., cell-free supernatants or purified bacteriocin protein solutions), the metal-cup method was used. A 50-μl volume of the sample was dropped into a metal cup on the surface of soft (0.75%) THB agar inoculated with the indicator strain. After incubation, the plate was examined for inhibition zones below and around the metal cup.

Conjugative transfer and mobilization experiments. Solid-surface mating was performed on agar plates (19, 23, 34). The mating mixture of the donor and recipient was made with a donor/recipient ratio of 1:10, and 10 μl of the mixed culture was spotted onto THB agar without antibiotics. The plates were then incubated overnight at 37°C. After incubation, the bacteria that had grown on the agar plates were scraped off and transferred to 1 ml of fresh THB, and then 0.1 ml of the mixture was spread onto appropriate selective agar plates. The colonies were counted after 48 h of incubation at 37°C. The frequency of conjugative transfer was calculated as the ratio of the number of transconjugants to the number of donors. The frequency of mobilized transfer was calculated as the ratio of the number of transconjugants that showed bacteriocin activity to the number of donors.

Isolation and manipulation of plasmid DNA. Plasmid DNA was isolated by the alkaline lysis method (28). Plasmid DNA was treated with restriction enzymes and was subjected to agarose gel electrophoresis for analysis of DNA fragments, etc. Restriction enzymes were obtained from Nippon Gene (Toyama, Japan), New England Biolabs, Inc., and Takara (Tokyo, Japan) and were used in accordance with the suppliers' specifications. Agarose was obtained from Wako Chemicals, Osaka, Japan. Electrophoresis on 0.8% agarose gels was used to determine the sizes of DNA fragments larger than 0.5 kb. The eluted fragments were ligated to dephosphorylated, restriction enzyme-digested vector DNA with

TABLE 2. Oligonucleotides used in this study

Primer	Sequence (5′–3′)	Description
M13F	GTTTTCCCAGTCACGACGTT	Analysis of pHY sequence
M13R	GGAAACAGCTATGACCATGA	
D1	AAACAGATTTAACGGAGTAC	Analysis of pHY sequence
D2	CAAAGAAACGGAAAGAAGCAAC	
D3	AAAGACATTCATTGCATCAAC	
D4	GATTGCCTGGGCAATGTCTG	
D5	TAGCCACTTTGTAGCTCGTTG	
D6	TACAACAGTCCTAACGAAGA	
Tn5	AATTGGGCGGCGACGTTAAC	Analysis of Tn5 insertion mutant
T1	ATTTTAGCTGTAATTATTTC	Analysis of Tn5 insertion mutant
T2	TGAAAAAGCCAACCAACTAG	
T3	GGCTCATCTATAAAATTTTC	
T4	AGGATTTTTATCCAAGACTG	
T5	TCCTAAAAGTTATAAGCAAC	
T6	TTCTAGGCTTCAGTCCGATG	
T7-1F	ATATATTAATACGACTCACTATAGGGAAGGAGGAAGTAACAATGAA[a]	*bacA*-specific primer for RNA probe
T7-1R	TTATTTTATTACATAACGAG	
T7-2F	ATATATTAATACGACTCACTATAGGGTCAGAAAGTGAAACTATCTC[a]	*bacB*-specific primer for RNA probe
T7-2R	CTATTTAGAGTTTTTATTTC	
RT377	[Phos]ACATAATAGTTC[b]	Reverse transcription primer for 5′-RACE[c] method
A1-246	ACCGAACAAAAATAGACAAG	5′-RACE method
A2-201	CATTGTTACTTCCTCCTTAG	
S1-260	CAAGTACTCCGTTAAATCTG	
S2-288	TAGTGCTAGAGGTATAAAAG	
EP1	ACTTGTCTATTTTTGTTCGG	*bacA*-specific primer
EP2	GTTGTATTGCTAAAGTGAGC	
pHY-316F-*Nde*I	TTTTTTCATATGGCAAGTTCACGATATAATCATAATC[d]	Cloning into pET22b(+)
pHY-630R-*Xho*I	TTTTTTCTCGAGTTTTATTACATAACGAGCGTAATC[d]	
pTI1-183F-*Nde*I	TTTTTTCATATGGCTGCTCAAAGAGGATATATC[d]	Cloning into pET22b(+)
pTI1-344R-*Xho*I	TTTTTTCTCGAGAATGTAGTAATAATATTTGGC[d]	

[a] The T7 promoter sequence is underlined.
[b] [Phos], phosphorylated 5′ end.
[c] RACE, rapid amplification of cDNA ends.
[d] Restriction enzyme cloning sites are underlined.

T4 DNA ligase and were then introduced into *E. coli* by electrotransformation (15). Transformants were selected on Luria-Bertani agar containing the appropriate antibiotics.

PFGE analysis. Pulsed-field gel electrophoresis (PFGE) was carried out on a 1% agarose gel with 0.5% Tris-borate-EDTA buffer. The following settings were applied: 1 to 23 s, 6 V/cm, and 22 h (with the CHEF Mapper system [Bio-Rad]).

PCR methodology. The PCR program, with *Ex Taq* DNA polymerase (Takara), comprised 2 min at 95°C, followed by 30 cycles of 30 s at 95°C, 30 s at 56°C, and 1 min at 72°C, and a final incubation at 4°C with a GeneAmp PCR system 9700 thermal cycler (Perkin-Elmer).

DNA sequence analysis. Nucleotide sequence analysis was carried out as described previously (28). To determine the entire sequence of pHY, shotgun sequencing was performed. Fragmented DNA libraries were constructed by sonication of pHY, followed by ligation into the SmaI-digested pUC19 vector plasmid. pUC19 plasmids containing 0.5- to 1.0-kb inserts were used to transform *E. coli* DH5α. The resulting constructs were sequenced in both orientations with an ABI Prism 310 genetic analyzer. The BigDye Terminator cycle sequencing kit (version 1.1; Applied Biosystems) and primers M13F and M13R were used for the sequencing reaction (Table 2). Open reading frame (ORF) analysis was performed with Genetyx, version 6.1 (Genetyx Corp., Tokyo, Japan). The DNA Data Bank of Japan (DDBJ; National Institute of Genetics, Mishima, Japan) and a BLAST search (http://www.ncbi.nlm.nih.gov/BLAST) were used for homology analysis of nucleotide and amino acid sequences.

Cloning of pHY plasmid DNA. To identify the bacteriocin determinant and the immunity determinant of the bacteriocin, the entire pHY plasmid DNA was digested with EcoRI (since there was a single EcoRI site in ORF9) and was cloned into vector pMW119. The spectinomycin resistance gene, which is expressed in both *E. coli* and *E. faecalis*, was cloned into the pMW119 XmaI site from pDL278 (12, 25). The cloned plasmid DNA was prepared from *E. coli* DH5α and was used either to transform *E. faecalis* OG1S in order to test for bacteriocin activity or to transform *E. hirae* ATCC 9790 in order to test for immunity to the bacteriocin.

Generation of transposon (Tn*5*) insertion mutants. Tn*5* (Km[r]) was inserted into the cloned plasmid DNA as described elsewhere (29). The target plasmid pMW119 (Spc[r])::pHY was introduced into *E. coli* K-12 TH688 (with Tn*5* in the *thr* locus) by electrotransformation. Transformants were spread onto selective plates containing kanamycin and spectinomycin, and the plates were left at room temperature for 10 days. The bacteria that grew on the selective plates were pooled, and the plasmid DNA was then isolated and used to transform *E. coli* DH5α. The transformants were selected on plates containing kanamycin and spectinomycin for the selection of Tn*5*-mediated kanamycin resistance and plasmid-mediated spectinomycin resistance, respectively. The transformants were purified and examined to determine the location of Tn*5* within the plasmid. The precise locations of Tn*5* insertion were determined by DNA sequence analysis using a synthetic primer that hybridized to the end of Tn*5* (Table 2).

Northern blotting of *bacA* and *bacB* transcripts. The method for Northern blotting of *bacA* and *bacB* transcripts is described elsewhere (28). The *bacA* and *bacB* probes were prepared with a DIG Northern Starter kit (Roche Diagnostics GmbH, Mannheim, Germany). T7-1F and T7-1R were used as primers with *bacA*; T7-2F and T7-2R were used as primers with *bacB*; and pHY plasmid DNA was used as the template (Table 2). Total RNAs were prepared with a FastRNA Pro Red kit (MP Biomedicals, LLC). Northern blot analysis involved electrophoresis in a 1.2% agarose-morpholinepropanesulfonic acid (MOPS)-formaldehyde system. Hybridization and detection were performed with a DIG Northern Starter kit. The chemiluminescent substrate CDP-Star (Boehringer Mannheim) was used for visualization of the RNA bands. Chemiluminescence was detected using Lumi-Film (Boehringer Mannheim).

Identification of the bacteriocin 51 transcriptional initiation site. To determine the initiation site of the transcript, rapid amplification of cDNA ends (RACE) was used. Total RNA was isolated from *E. faecium* BM4105SS(pHY). The transcriptional initiation site was determined using the 5′-Full RACE core set (Takara Bio Inc.) according to the manufacturer's instructions. First-strand cDNA synthesis was carried out in a reverse transcription reaction using the 5′-phosphorylated primer RT377, which is specific to the target RNA, and 5 μg

TABLE 3. Drug resistance of VanA-type vancomycin-resistant enterococci

E. faecium strain	Level of resistance (MIC [μg/ml]) to the following antimicrobial drug[a]:									
	VAN	TEC	GEN	KAN	STR	ERY	CHL	TET	AMP	CIP
VRE34	1,024	32	≥1,024	≥1,024	16	≥1,024	8	32	512	8
VRE35	1,024	16	1,024	≥1,024	16	≥1,024	4	64	512	8
VRE36	1,024	128	4	128	16	≥1,024	4	64	512	128
VRE37	1,024	32	≥1,024	≥1,024	≥1,024	≥1,024	8	32	512	8
VRE38	1,024	128	≥1,024	≥1,024	16	≥1,024	4	128	512	256
VRE39	1,024	32	≥1,024	≥1,024	16	≥1,024	4	64	256	8

[a] VAN, vancomycin; TEC, teicoplanin; GEN, gentamicin; KAN, kanamycin; STR, streptomycin; ERY, erythromycin; CHL, chloramphenicol; TET, tetracycline; AMP, ampicillin; CIP, ciprofloxacin.

of total RNA (Table 2). The resulting purified cDNA was circularized with RNA ligase. The sequence of the unknown region containing the transcription initiation site was determined by two rounds of PCR; the first round used the circular cDNA as a template with primers A1-246 and S1-260, and the second round used the first PCR product as a template with primers A2-201 and S2-288 (Table 2).

DNA-DNA hybridization. Southern hybridization was performed with the digoxigenin-based nonradioisotope system of Boehringer GmbH (Mannheim, Germany). All procedures were based on the manufacturer's manual and standard protocols (28). The plasmid DNA was isolated and digested with EcoRI. Hybridization was performed overnight at 42°C in the presence of 50% formamide. The probes were generated by using a DIG probe synthesis kit (Roche Diagnostics GmbH, Mannheim, Germany). The signals were detected with a nitroblue tetrazolium–5-bromo-4-chloro-3-indolylphosphate stock solution (Roche Diagnostics GmbH).

Construction of the plasmid producing C-terminally His tagged bacteriocin protein. The plasmid producing C-terminally His tagged bacteriocin (Bac) protein was constructed using the pET-22b(+) vector plasmid according to the manufacturer's instructions (Novagen). The *E. coli* vector pET-22b(+) (Amp^r), containing the C-terminal 6-amino-acid His tag sequence just downstream from the XhoI sequence, was used to introduce a His tag into the C terminus of the mature Bac protein. The DNA sequence of the predicted mature Bac protein was amplified by PCR using *E. faecium* BM4105SS(pHY) for Bac 51 and *E. faecalis* FA2-2(pTI1) for Bac 32. The PCR product obtained was cloned between the NdeI and XhoI sites of the pET-22b(+) vector. The resultant cloned plasmid was transformed into *E. coli* BL21. Ampicillin-resistant transformants (Amp^r) were selected on LB plates containing ampicillin, and the transformant was designated *E. coli* BL21[pET-22b(+)::*bac*].

Purification of His-tagged Bac 51 protein from *E. coli* BL21[pET22b(+)::*bac*]. The *bac* gene was expressed as a His-tagged fusion protein, which was purified on a Ni-nitrilotriacetic acid (NTA) column according to the manufacturer's protocol (Novagen, Madison, WI). LB broth (100 ml) containing 100 μg of ampicillin per ml was inoculated with 1 ml of an overnight culture of *E. coli* BL21[pET22b(+)::*bac*]. Cells were grown aerobically at 30°C to an optical density at 600 nm (OD_{600}) of 0.5, and then isopropyl-β-D-thiogalactoside (IPTG) was added to a concentration of 0.5 mM for 4 h at 30°C to induce protein expression. Cells were harvested from liquid medium by centrifugation for 10 min at 10,000 × *g*. The medium was decanted, and the pellet was allowed to drain so as to remove as much liquid as possible.

The pellet was completely resuspended in 10 ml lysis buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl) to give a 10× concentration factor (100 ml of culture to a 10-ml buffer volume), and lysozyme was then added to a 100-μg per ml concentration. Phenylmethylsulfonyl fluoride (PMSF) to a final concentration of 1 mM was also added. The mixture was incubated at 25°C for 30 min prior to sonication. The resuspended pellet was mixed by swirling and was sonicated on ice using a sonicator (Ultrasonic Disruptor UD-201; Tomy) set at power level 6, at 40% duty, for 20 min. The entire lysate was centrifuged for 10 min at 14,000 × *g* to separate the soluble and insoluble fractions. The soluble supernatant was used for the isolation of His-tagged Bac. His-tagged Bac was isolated using the Ni-NTA nickel chromatography resin (Ni-NTA purification system; Invitrogen, CA) according to the manufacturer's instructions. The bacteriocin protein was eluted from the column by first loading 10 ml of imidazole at concentrations of 10 mM, 20 mM, 35 mM, and 50 mM and then loading 4 ml of imidazole at concentrations of 100 mM, 200 mM, 300 mM, and 400 mM.

SDS-PAGE. The purity of each His-tagged BacA (BacA-His) sample was determined using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), which was carried out as described previously (28). Electrophoresis was conducted on 18% (wt/vol) polyacrylamide separating gels at a constant current of 30 mA for 60 min. To visualize the proteins, the gel was stained with Silver Stain Plus (Bio-Rad Laboratories, CA) or a SeePico CBB stain kit (Benebiosis Co., Seoul, South Korea) according to the manufacturer's instructions. Precison Plus protein dual-color standards (Bio-Rad Laboratories, CA), consisting of a mixture of proteins with molecular masses ranging from 10 to 250 kDa, were electrophoresed beside the bacteriocin samples in each gel.

Preparation of a culture filtrate from *E. faecalis* FA2-2 carrying a bacteriocin plasmid. *E. faecalis* FA2-2 carrying a bacteriocin plasmid was grown to late-exponential phase. The culture was centrifuged at 14,000 × *g* for 10 min at 4°C. The supernatant was filtered through a Millipore filter (pore size, 0.45 μm), and the filtrate was used to test for bacteriocin activity.

Bacteriocin activity. The bacteriocin activities of the eluates obtained throughout the purification process or of the culture filtrate were calculated by microtube assays. The test samples were 2-fold serially diluted with microtubes containing 5 μl imidazole. A 5-μl volume of each dilution was spotted onto the soft-agar plate containing the indicator strain, and the plates were incubated at 37°C for 18 h. Growth inhibition was examined, and the bacteriocin activity was calculated in bacteriocin units (BU). One bacteriocin unit was defined as the reciprocal of the highest dilution of bacteriocin causing growth inhibition. For example, if the highest dilution was 128-fold, the bacteriocin activity was 25 BU/μl (i.e., 128/5 μl).

Liquid bacteriocin assay. The bacteriocin indicator strain *E. hirae* ATCC 9790 was grown overnight in THB at 37°C. An overnight culture of the strain was diluted 10^2-fold with fresh THB. A diluted culture of the indicator strain containing about 10^7 cells per ml was incubated for 90 min at 37°C with gentle shaking. The culture was diluted 10^4-fold with fresh THB. The diluted culture, which contained about 10^3 bacterial cells per ml, was used as the indicator strain in the liquid bacteriocin assay. One portion (0.2 ml) of the test sample was added to nine portions (1.8 ml) of the diluted culture of the indicator strain. These cultures were incubated at 37°C with gentle shaking. Samples were removed at various intervals and were appropriately diluted, and the dilution was plated onto a THB agar plate to determine cell survival.

Effect of temperature on bacteriocin activity. Purified bacteriocin from the BacA-His samples was exposed to temperatures of 60°C, 70°C, 80°C, 90°C, and 100°C for 5 min or 10 min. After heat treatment, the samples were cooled rapidly, and then 5 μl of the sample was spread onto soft agar containing the indicator strain in order to carry out the soft-agar assay. Based on the initial results, BacA-His samples were then exposed to 70°C, 72°C, 74°C, 76°C, 78°C, and 80°C for 10 min, and the residual bacteriocin activities of the heat-treated samples was examined using the soft-agar assay. Untreated samples were used as controls in all cases.

Nucleotide sequence accession number. The nucleotide sequence data reported in this article are available from the DDBJ, EMBL, and GenBank nucleotide sequence databases under accession number AB570326.

RESULTS

Bacteriocinogenic strains and the plasmid encoding the bacteriocin. Drug resistance of six VanA-type vancomycin-resistant *E. faecium* (VRE) isolates recovered in a hospital is shown in Table 3. Two strains, VRE34 and VRE35, were found to be identical, and the other four were found to be different, by PFGE analysis of their chromosomal DNAs (data not shown). Of these VRE strains, five (i.e., VRE34, -36, -37, -38, and -39)

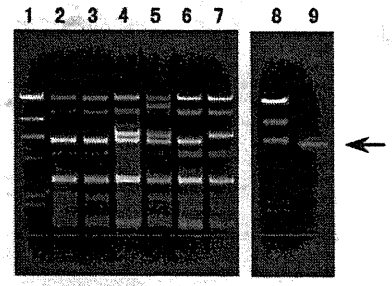

FIG. 1. Agarose gel electrophoresis of EcoRI-digested plasmid DNAs of wild-type VRE strains and pHY. Lanes: 1 and 8, HindIII-digested lamda DNA; 2 to 7, strains VRE34, VRE35, VRE36, VRE37, VRE38, and VRE39, respectively; 9, pHY harbored by the bacteriocinogenic VRE38–BM4105SS transconjugant. The arrow indicates the 6.0-kbp plasmid DNA fragment.

exhibited bacteriocin activity against *E. faecalis*, *E. faecium*, *E. hirae*, *E. durans*, *E. raffinosus*, *E. gallinarum*, and *L. monocytogenes*, and one strain, designated VRE35, exhibited bacteriocin activity against *E. faecium*, *E. hirae*, and *E. durans*. These strains contained several plasmids (Fig. 1). Since it was possible that the bacteriocin determinant was encoded on a plasmid, the cotransferability of the bacteriocin activities with drug resistance was examined. The transferability of the gentamicin resistance of all strains except VRE36 to the recipient strain *E. faecium* BM4105RF was examined by filter mating. The gentamicin resistance of four strains was transferred to the recipient strain at a frequency of about 10^{-4} per donor cell. Fifty gentamicin resistance transconjugants from each of the four strains were examined for bacteriocin activity. About 75% to 95% of the gentamicin transconjugants showed bacteriocin activity against *E. faecium*, *E. hirae*, and *E. durans*, suggesting that the bacteriocin activity was cotransferred. A DNA plasmid of about 6.0 kbp was found in the bacteriocinogenic transconjugants of each strain and was not found in nonbacteriocinogenic transconjugants. The experiment to transfer bacteriocin production was repeated by using the initial bacteriocinogenic *E. faecium* BM4105RF transconjugant of each strain to transfer the bacteriocin to *E. faecium* BM4105SS by filter mating. About 80 to 90% of the gentamicin-resistant transconjugants were bacteriocinogenic and showed bacteriocin activity identical to that of the *E. faecium* BM4105RF donor strain. The bacteriocinogenic transconjugants harbored only the 6.0-kbp plasmid, implying that this plasmid might encode the bacteriocin. The 6.0-kbp plasmid itself is not conjugative, but it is presumed to be mobilized by a resident plasmid of unknown identity. The transconjugant of *E. faecium* VRE38 was used as the representative strain for further analyses. The 6.0-kbp plasmid DNA identified in the bacteriocinogenic transconjugants of *E. faecium* VRE38 was designated pHY.

DNA sequence of the pHY plasmid. The DNA sequence of the pHY plasmid was determined and was found to be 6,037 bp long. Computer analysis revealed the presence of nine open reading frames (ORFs) in pHY, all of which were oriented in the same direction. Figure 2a shows the ORFs that had a good ribosome binding site within a 20-base region upstream of the predicted start codon. Homology analyses of the deduced amino acid sequences and the nucleotide sequence of each ORF of pHY were performed. The deduced proteins encoded on ORF3, ORF4, ORF5, ORF7, and ORF8 showed significant homology to those encoded by the reported genes *mobC* (1, 23), *mobA* (1, 23), ORF6 of plasmid pTI1 (23), *repA* (23, 30), and ORF10 of plasmid pTI1 (23), respectively. Mobilization plasmids usually carry mobilization genes (*mob*) that encode specific relaxasome components, as well as the origin of transfer (*oriT*) (3, 14). The nicking of *oriT* is an essential initial step in the mobilization of plasmid DNA. The MobA relaxase makes a reversible site- and strand-specific nick at a specific sequence within *oriT* (3, 14). MobC is an accessory protein of MobA for DNA-nicking activity. Rep protein has replication origin binding and nicking-closing activities (24). The deduced proteins of ORF1, ORF2, ORF6, and ORF9 showed no significant homology with reported proteins. The analysis of ORF1 and ORF2 is described below.

Cloning of pHY into pMW119 (Spc^r). The attempt to clone pHY into the cloning vector pAM401 failed. The *E. coli* cloning vector pMW119 was tagged with the spectinomycin resistance gene (Spc^r) for use as a selective marker in enterococci as described in Materials and Methods, and the resultant plasmid was designated pMW119 (Spc^r). pHY has a single EcoRI site, which was cut and used to clone it into the EcoRI site of pMW119 (Spc^r). *E. faecalis* OG1S or *E. hirae* ATCC 9790 was transformed by the cloned pMW119 (Spc^r)::pHY plasmid. The

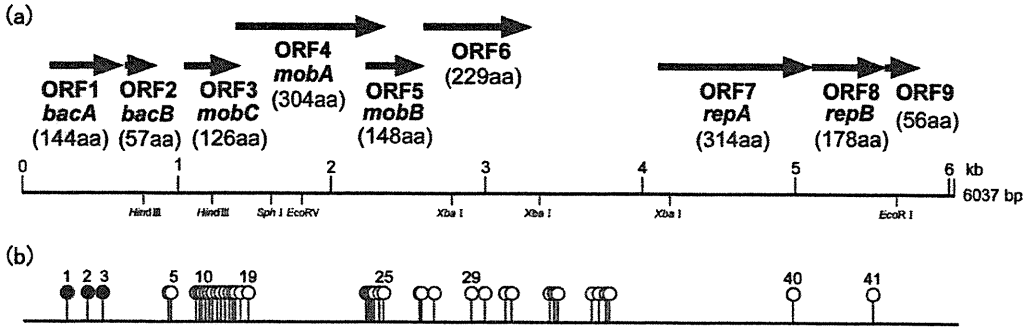

FIG. 2. Physical map of pHY (6.0 kb) showing the deduced ORFs and mapping of the transposon insertions. (a) Physical map of pHY and deduced ORFs. Thick horizontal arrows indicate the deduced ORFs encoded on pHY and the direction of transcription. (b) Map of Tn5 insertions into pMW119 (Spc^r)::pHY. Lollipops indicate the points of Tn5 mutant insertion: filled circles, mutants that did not express bacteriocin; open circles, bacteriocinogenic mutants.

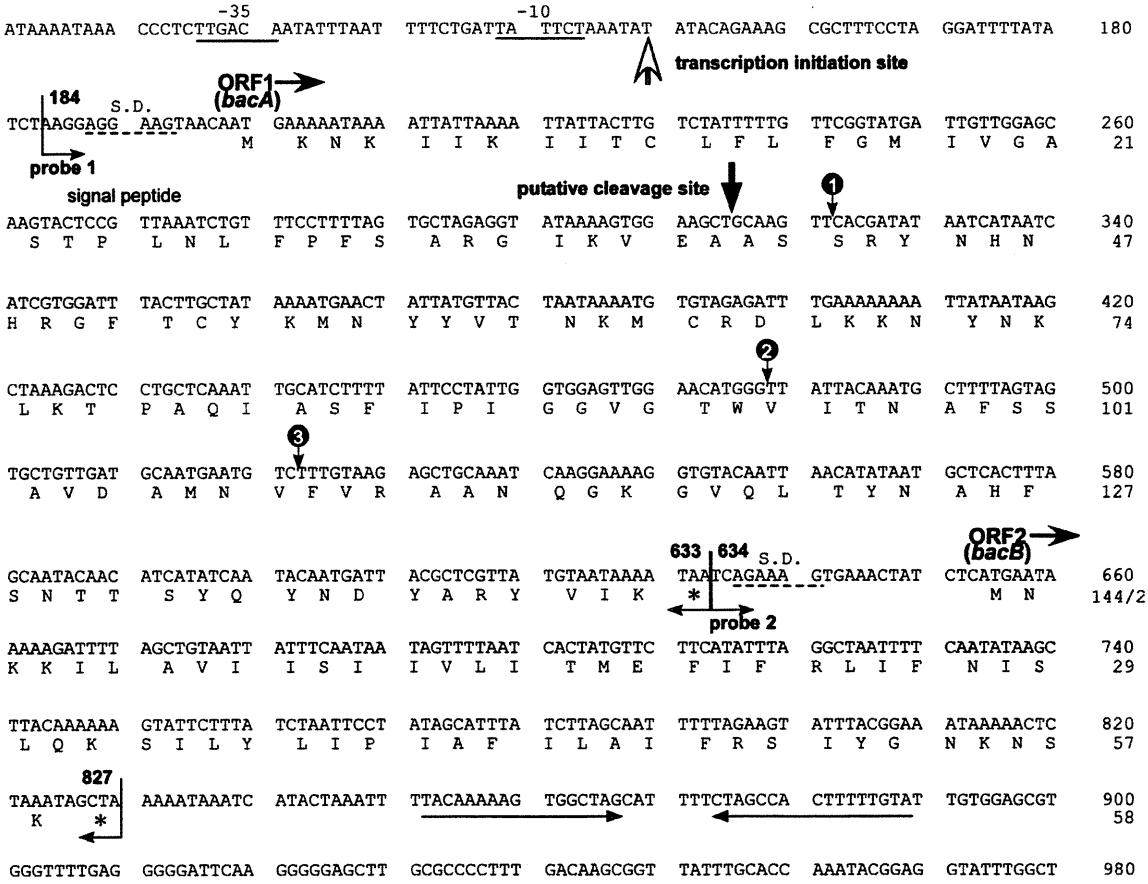

FIG. 3. Nucleotide sequences of *bacA* and *bacB* for bacteriocin 51 and deduced amino acid sequences. Potential promoters (−10 and −35) are underlined, and the Shine-Dalgarno (S.D.) ribosome binding sequences are marked by dashed underlines. The potential signal peptidase cleavage site is indicated by a filled vertical arrow. The transcription initiation site of bacteriocin 51 is indicated by an open vertical arrow. Horizontal arrows under the nucleotide sequence indicate inverted repeats, which are explained in the text. Three numbered circles indicate the points of Tn*5* insertion that abolished bacteriocin expression. Probe 1, which corresponds to a 450-base sequence in *bacA* (bases 184 to 633), and probe 2, which corresponds to a 194-base sequence in *bacB* (bases 634 to 827), were used for Northern hybridization analysis.

spectinomycin-resistant transformants of *E. faecalis* OG1S and *E. hirae* ATCC 9790 showed bacteriocin activity against *E. faecium*, *E. hirae*, and *E. durans*, indicating that pMW119 (Spc^r)::pHY transformed both *E. faecalis* OG1S and *E. hirae* ATCC 9790. *E. hirae* ATCC 9790 was sensitive to the bacteriocin of the bacteriocinogenic transformant *E. faecalis* OG1S. However, the bacteriocinogenic *E. hirae* ATCC 9790 strain was resistant to the bacteriocin activity of the bacteriocinogenic *E. faecalis* OG1S strain (data not shown). These results indicate that plasmid pHY encodes the bacteriocin and the immunity function for its own bacteriocin. The bacteriocin encoded on pHY was designated bacteriocin 51 (Bac 51).

Analysis of Tn*5* insertion mutants. Tn*5* insertion mutants of pMW119 (Spc^r)::pHY were generated in *E. coli* TH688::Tn*5*. A total of 41 insertions into pHY were examined; the location of each insertion is shown in Fig. 2b. Three of the insertions resulted in the loss of bacteriocin activity, while activity was retained by the remaining insertion mutants. The three non-bacteriocinogenic mutants, designated pMG701, pMG702, and pMG703, were mapped at bp 322, bp 478, and bp 522 from the first adenine nucleotide of the pHY map within ORF1 (Fig. 3). *E. hirae* ATCC 9790 harboring pMG701, pMG702, or pMG703 did not show bacteriocin activity or resistance (immunity) to the bacteriocin activity of *E. faecalis* OG1S or *E. hirae* ATCC 9790 harboring pMW119 (Spc^r)::pHY. These results suggested that ORF1 and ORF2, which we designated *bacA* and *bacB*, were bacteriocin determinants.

Detailed sequence analysis of ORF1 (*bacA*) and ORF2 (*bacB*). *bacA* encoded a 144-amino-acid protein. The ATG start codon was preceded by a potential ribosome binding site located 5 bp upstream. Computer analysis suggested that the deduced *bacA* protein had a signal peptide sequence, and a potential signal peptidase-processing site corresponding to the V-E-A sequence was located at positions 37 to 39 (Fig. 3). The predicted mature BacA protein consisted of 105 amino acids. BacB encoded a 55-amino-acid protein. The ATG start codon was preceded by a potential ribosome binding site located 12 bp upstream. There was no obvious promoter sequence upstream of the ribosome binding site. A putative transcription terminator signal was identified downstream of *bacB*. There was no obvious promoter sequence between *bacA* and *bacB*. These results suggested that *bacA* was the structural gene and *bacB* was the immunity gene for bacteriocin 51.

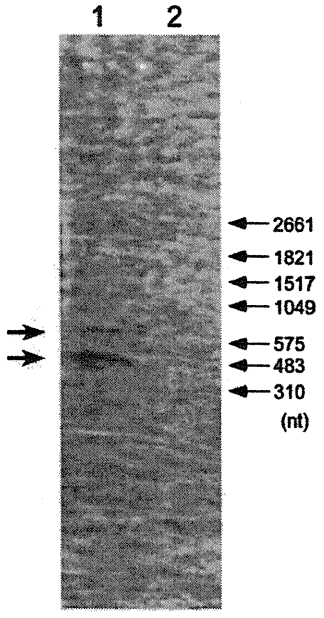

FIG. 4. Northern blot analysis of the *bacA* and *bacB* genes. Lanes; 1, *E. faecium* BM4105SS cells harboring pHY; 2, BM4105SS plasmid-free cells. Arrows on the left indicate mRNAs of about 500 nt and 700 nt. The positions of the RNA markers are noted on the right.

Analysis of transcripts of *bacA* and *bacB*. DNA sequence analysis revealed that *bacA* and *bacB* were transcribed in the same direction. The transcripts of these genes were analyzed by Northern hybridization (Fig. 4). Probe 1, which corresponds to a 450-nucleotide (nt) sequence in *bacA* (Fig. 3), detected two transcripts of about 700 and 500 nucleotides, corresponding to the molecular sizes of *bacA* and *bacB* combined and *bacA* alone, respectively. The quantity of transcript corresponding to the combined *bacA* and *bacB* nucleotide sequence was much smaller than that corresponding to the *bacA* sequence.

Probe 2, which corresponds to a 194-nucleotide sequence in *bacB* (Fig. 3), did not detect any transcript (data not shown) from *bacB* alone, implying that the small quantity of transcript that was detected corresponded to combined *bacA* and *bacB* (Fig. 4). The transcription start site upstream of *bacA* was examined by the RACE method and was determined to be the T nucleotide located 6 nucleotides downstream from the −10 promoter sequence (Fig. 3). These results indicated that transcription initiated from the transcription start site upstream of *bacA*.

Construction of the plasmid producing C-terminally His tagged mature Bac 51 protein. Plasmid DNA from *E. faecium* BM4105SS(pHY) was used to amplify the DNA sequence of the predicted mature BacA protein of Bac 51 by PCR. The PCR product was cloned into the *E. coli* vector pET22b(+). The resulting cloned plasmid, pET22b(+)::*bac51*, was used to transform *E. coli* BL21 as described in Materials and Methods.

Purification of His-tagged Bac 51 protein from *E. coli* BL21[pET22b(+)::*bac51* (*bacA*)]. The procedure for the purification of His-tagged Bac 51 is described in Materials and Methods. Analysis by SDS-PAGE showed that elution with either 35 mM, 50 mM, 100 mM, or 200 mM imidazole gave rise to one band, which corresponded to a molecular mass of approximately 12 kDa (see Fig. S1 in the supplemental material). The eluates obtained using a 100 mM, 200 mM, or 300 mM concentration of imidazole were examined for bacteriocin activity against *E. hirae* ATCC 9790 by the metal-cup method as described in Materials and Methods. The eluates obtained with 100 mM or 200 mM imidazole gave rise to a 12-kDa protein band by SDS-PAGE analysis and produced a complete bacterial growth inhibition zone with *E. faecium*, *E. hirae*, or *E. durans* below and around the metal cup (see Fig. S2 in the supplemental material). However, no inhibition zone was observed with the eluate obtained with 300 mM imidazole, which did not gave rise to a 12-kDa protein band in SDS-PAGE analysis. These results indicated that the mature His-tagged Bac 51 protein was purified, showed the same activity as the Bac 51 protein of the wild-type strain, and had a molecular mass of about 12 kDa.

Construction of the plasmid producing C-terminally His tagged Bac 32 protein. *bac32* consists of the *bacA* and *bacB* genes (23). *bacA* is the structural gene for Bac 32, and *bacB* is the immunity gene for its own bacteriocin. The *bacA* protein is composed of 89 amino acids with a putative signal peptide at the N terminus. The predicted signal peptide could be cleaved after the V-E-A residues (amino acids 33 and 35) by computer analysis using the SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP/), and the *bacA* product would consist of a 35-amino-acid signal peptide and a 54-amino-acid mature BacA protein. Plasmid DNA from FA2-2(pTI1) was used for PCR amplification of the DNA sequence of the predicted mature BacA protein of Bac 32. The plasmid producing C-terminally His tagged Bac 32 protein was constructed by using pET-22b(+) as described above.

Purification of His-tagged Bac 32 protein from *E. coli* BL21[pET22b(+)::*bac32* (*bacA*)]. SDS-PAGE analysis showed that elution with imidazole at concentrations between 35 mM and 400 mM gave rise to a single band corresponding to a molecular mass of approximately 6.4 kDa (data not shown). Each of these eluates showed complete inhibition of bacterial growth by *E. faecium*, *E. hirae*, and *E. durans* below and around the metal cup containing the eluate. These results indicated that the mature His-tagged Bac 32 protein had been purified, that it showed the same activity as the Bac 32 protein of the wild-type strain, and that its molecular mass was about 6.4 kDa.

Liquid bacteriocin assay. The results of the liquid bacteriocin assay are shown in Fig. 5. The bacteriocin activity of purified Bac 51 BacA-His was examined using the liquid bacteriocin assay as described in Materials and Methods. The number of cells of the indicator strain in the culture containing Bac 51 BacA-His remained unchanged in the 5-h period of the experiment (data not shown). On the other hand, the survival of indicator strain cells in cultures containing the bacteriocin 31 culture filtrate or purified Bac 32 BacA-His was reduced to <1/10,000 after 30 min. The survival of indicator strain cells in the culture containing the bacteriocin 43 culture filtrate was also reduced logarithmically. The cell counts of the indicator strain in the culture without bacteriocin increased logarithmically during incubation.

Thermostability of the bacteriocin. Purified BacA-His was examined for bacteriocin activity after heat treatment. The

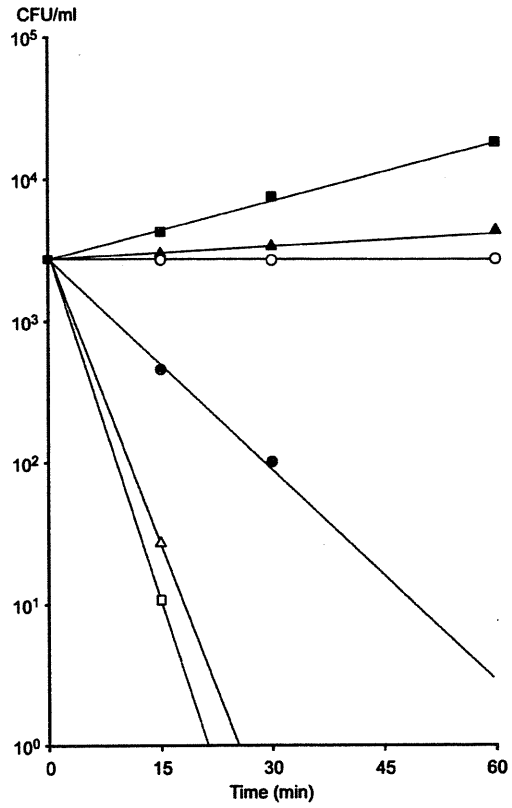

FIG. 5. Liquid bacteriocin assay. The assay was performed by adding 0.2 ml of the test sample to 1.8 ml of the *E. hirae* indicator strain. Symbols for test samples: ○, purified His-tagged Bac 51 (2.5×10^3 BU/ml) from elution with 200 mM imidazole; △, FA2-2(pYI17 *bac31*) culture filtrate (8.0×10^2 BU/ml); □, purified His-tagged Bac 32 (2.5×10^3 BU/ml) from elution with 200 mM imidazole; ●, FA2-2(pMG502 *bac43*) culture filtrate (8.0×10^2 BU/ml); ▲, *E. faecalis* FA2-2 culture filtrate; ■, fresh THB broth.

bacteriocin was stable at temperatures below 100°C for 5 min or at 70°C for 10 min. It was inactivated when incubated at 72°C for 10 min. These results indicated that the bacteriocin was heat stable.

Identification of Bac 51 in *E. faecium* isolates. One hundred thirty-eight vancomycin-sensitive *E. faecium* isolates, 87 vancomycin-resistant *E. faecium* isolates, 120 vancomycin-sensitive *E. faecalis* isolates recovered in Japanese hospitals, and 662 vancomycin-resistant *E. faecium* isolates recovered in the University of Michigan hospital in Ann Arbor (23, 35) were examined for the presence of the Bac 51 determinant by PCR analysis with a primer specific for Bac 51 *bacA*. Two (1.4%) of the 138 vancomycin-sensitive *E. faecium* strains, 2 (2.3%) of the 87 VRE strains isolated in Japanese hospitals, and 2 (0.3%) of the VRE strains isolated in the University of Michigan hospital gave rise to the expected 365-bp PCR product with the *bacA*-specific primer. No *bacA*-specific PCR products were detected in *E. faecalis* isolates. Plasmid DNAs isolated from the *bacA*-positive strains were examined for the presence of the Bac 51 gene by PCR analysis with the primer and by Southern blot analysis with the *bacA* probe. The results showed that the Bac 51 determinant was encoded on the plasmids of these strains.

DISCUSSION

A new bacteriocin, designated Bac 51, was identified in a VanA-type vancomycin-resistant *E. faecium* strain, VRE38. Bac 51 exhibited a relatively narrow spectrum of activity and was active against *E. faecium*, *E. hirae*, and *E. durans* strains. Bac 51 was encoded on plasmid pHY (6.0 kbp), which was efficiently mobilized for transfer to a recipient *E. faecium* strain at a frequency of 10^{-5} to 10^{-7} per donor cell with the coresident conjugative gentamicin resistance (35). The Bac 51 determinant consisted of the bacteriocin structural gene, *bacA*, and the immunity gene, *bacB*. The *bacA* gene encoded a deduced 144-amino-acid protein with a putative signal sequence of 39 amino acid residues at the N terminus that was predicted to give rise to a 105-amino-acid mature protein. The *bacB* gene encoded a deduced 55-amino-acid protein without a putative signal sequence. The deduced BacA and BacB proteins had no homology to known bacteriocins, indicating that Bac 51 was a new type of bacteriocin from *E. faecium*. The results of Northern blot analysis of *bacA* and *bacB* suggest that there is significant transcription termination after *bacA*, but with some extension through *bacB*. The transcription start site was determined to be the T nucleotide located 6 nucleotides downstream from the −10 promoter sequence. These results indicated that *bacA* and *bacB* constituted an operon; *bacA* was the bacteriocin structural gene, and *bacB* was the immunity gene.

The purified C-terminally His tagged BacA (BacA-His) of Bac 51 inhibited bacterial cell growth in *E. faecium*, *E. hirae*, and *E. durans* but did not inhibit cell growth in *L. monocytogenes*, as shown by the metal-cup bacteriocin assay. Bac 31, originally isolated from *E. faecalis* (32), and Bac 32 (23) and Bac 43 (30), originally isolated from *E. faecium* clinical isolates, are pore-forming bacteriocins. Purified Bac 32 BacA-His and the culture filtrates of *E. faecalis* strain FA2-2, which produces each of these pore-forming bacteriocins, were active against the indicator strains and inhibited cell growth by the metal-cup method. In the liquid bacteriocin assay, purified Bac 51 BacA-His did not kill the indicator strain. On the other hand, purified Bac 32 BacA-His and the culture filtrates of the Bac 31- and Bac 43-producing *E. faecalis* strain FA2-2 killed the indicator strain shortly after it was exposed to the bacteriocin. These results indicated that in contrast to the bactericidal effect produced by the pore-forming bacteriocins Bac 31, Bac 32, and Bac 43, the mode of action of Bac 51 is bacteriostatic.

In our previous study of Bac 32, we showed that the Bac 32 *bacA* product comprised 89 amino acids with a putative signal peptide at the N terminus, based on genetic and DNA sequence analysis (23). We carried out computer analysis using the SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP/) in the present study. The *bacA* product may consist of a 35-amino-acid signal peptide and a 54-amino-acid mature BacA protein, and the signal peptide may be cleaved after the V-E-A residues (positions 33 to 35). The DNA primer for amplification of the mature *bacA* protein was designed to produce a DNA fragment for the 54-amino-acid mature BacA protein. The purified BacA-His protein was active against the indicator strain. These data indicated that the signal peptide could be cleaved after the V-E-A residues (positions 33 to 35). Computer analysis using Genetyx, version 7, was described in our previous report (23), where we predicted that the deduced

bacA product consisted of a 19-amino-acid signal peptide and a 70-amino-acid mature protein, and that the signal peptide was cleaved after the L-L-A residues (positions 17 to 19). The results of the present study of the bacteriocin activity of the purified 54-amino-acid BacA mature protein of Bac 32 showed that the previous prediction of the signal peptidase cleavage site could be wrong (23).

The well-characterized *E. faecium* bacteriocins (i.e., enterocins) are produced by food-grade organisms and show strong bacteriocin activity against *L. monocytogenes* (27). In contrast to the bacteriocins of food-grade *E. faecium* strains, Bac 32 or Bac 32-type bacteriocins, which are frequently (about 40% of clinical isolates) identified in *E. faecium* clinical isolates, are active against *E. faecium*, *E. hirae*, and *E. durans* but show no activity against *L. monocytogenes* (23). Bac 51 was also identified in VRE clinical isolates and was active against *E. faecium*, *E. hirae*, and *E. durans* but not against *L. monocytogenes*. However, there was no sequence homology between the predicted Bac 51 and Bac 32 proteins. Bac 51-type bacteriocins were identified in *E. faecium* clinical isolates at a frequency of about 2% but were not identified in *E. faecalis* clinical isolates, indicating that Bac 51 is specific to *E. faecium* clinical isolates.

ACKNOWLEDGMENTS

This work was supported by grants from the Japanese Ministry of Education, Culture, Sports, Science and Technology (Tokutei-ryoiki [Matrix of Infection Phenomena], Kiban [B], Kiban [C]) and the Japanese Ministry of Health, Labor and Welfare (H21-Shinkou-Ippan-008).

We thank Elizabeth Kamei and Masatomo Mori for helpful advice.

REFERENCES

1. **Apisiridej, A., A. Leelaporn, C. D. Scaramuzzi, R. A. Skurray, and N. Firth.** 1997. Molecular analysis of a mobilizable theta-mode trimethoprim resistance plasmid from coagulase-negative staphylococci. Plasmid **38:**13–24.
2. **Aymerich, T., et al.** 1996. Biochemical and genetic characterization of enterocin A from *Enterococcus faecium*, a new antilisterial bacteriocin in the pediocin family of bacteriocins. Appl. Environ. Microbiol. **62:**1676–1682.
3. **Caryl, J. A., and C. D. Thomas.** 2006. Investigating the basis of substrate recognition in the pC221 relaxosome. Mol. Microbiol. **60:**1302–1318.
4. **Casaus, P., et al.** 1997. Enterocin B, a new bacteriocin from *Enterococcus faecium* T136 which can act synergistically with enterocin A. Microbiology **143:**2287–2294.
5. **Cintas, L. M., et al.** 1998. Enterocins L50A and L50B, two novel bacteriocins from *Enterococcus faecium* L50, are related to staphylococcal hemolysins. J. Bacteriol. **180:**1988–1994.
6. **Cintas, L. M., P. Casaus, L. S. Havarstein, P. E. Hernandez, and I. F. Nes.** 1997. Biochemical and genetic characterization of enterocin P, a novel *sec*-dependent bacteriocin from *Enterococcus faecium* P13 with a broad antimicrobial spectrum. Appl. Environ. Microbiol. **63:**4321–4330.
7. **Cleveland, J., T. J. Montville, I. F. Nes, and M. L. Chikindas.** 2001. Bacteriocins: safe, natural antimicrobials for food preservation. Int. J. Food Microbiol. **71:**1–20.
8. **Clewell, D. B., et al.** 1982. Mapping of *Streptococcus faecalis* plasmids pAD1 and pAD2 and studies relating to transposition of Tn*917*. J. Bacteriol. **152:** 1220–1230.
9. **Cox, C. R., P. S. Coburn, and M. S. Gilmore.** 2005. Enterococcal cytolysin: a novel two component peptide system that serves as a bacterial defense against eukaryotic and prokaryotic cells. Curr. Protein Pept. Sci. **6:**77–84.
10. **De Vuyst, L., and E. J. Vandamme.** 1994. Bacteriocins of lactic acid bacteria: microbiology, genetics, and applications, p. 91–142. Blackie Academic & Professional, London, United Kingdom.
11. **Dunny, G. M., B. L. Brown, and D. B. Clewell.** 1978. Induced cell aggregation and mating in *Streptococcus faecalis*: evidence for a bacterial sex pheromone. Proc. Natl. Acad. Sci. U. S. A. **75:**3479–3483.
12. **Dunny, G. M., L. N. Lee, and D. J. LeBlanc.** 1991. Improved electroporation and cloning vector system for gram-positive bacteria. Appl. Environ. Microbiol. **57:**1194–1201.
13. **Floriano, B., J. L. Ruiz-Barba, and R. Jimenez-Diaz.** 1998. Purification and genetic characterization of enterocin I from *Enterococcus faecium* 6T1a, a novel antilisterial plasmid-encoded bacteriocin which does not belong to the pediocin family of bacteriocins. Appl. Environ. Microbiol. **64:**4883–4890.
14. **Francia, M. V., et al.** 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. **28:**79–100.
15. **Fujimoto, S., H. Hashimoto, and Y. Ike.** 1991. Low cost device for electrotransformation and its application to the highly efficient transformation of *Escherichia coli* and *Enterococcus faecalis*. Plasmid **26:**131–135.
16. **Fujimoto, S., H. Tomita, E. Wakamatsu, K. Tanimoto, and Y. Ike.** 1995. Physical mapping of the conjugative bacteriocin plasmid pPD1 of *Enterococcus faecalis* and identification of the determinant related to the pheromone response. J. Bacteriol. **177:**5574–5581.
17. **Gilmore, M. S., et al.** 1994. Genetic structure of the *Enterococcus faecalis* plasmid pAD1-encoded cytolytic toxin system and its relationship to lantibiotic determinants. J. Bacteriol. **176:**7335–7344.
18. **Haas, W., B. D. Shepard, and M. S. Gilmore.** 2002. Two-component regulator of *Enterococcus faecalis* cytolysin responds to quorum-sensing autoinduction. Nature **415:**84–87.
19. **Ike, Y., and D. B. Clewell.** 1984. Genetic analysis of the pAD1 pheromone response in *Streptococcus faecalis*, using transposon Tn*917* as an insertional mutagen. J. Bacteriol. **158:**777–783.
20. **Ike, Y., D. B. Clewell, R. A. Segarra, and M. S. Gilmore.** 1990. Genetic analysis of the pAD1 hemolysin/bacteriocin determinant in *Enterococcus faecalis*: Tn*917* insertional mutagenesis and cloning. J. Bacteriol. **172:**155–163.
21. **Ike, Y., H. Hashimoto, and D. B. Clewell.** 1984. Hemolysin of *Streptococcus faecalis* subspecies *zymogenes* contributes to virulence in mice. Infect. Immun. **45:**528–530.
22. **Ike, Y., H. Hashimoto, and D. B. Clewell.** 1987. High incidence of hemolysin production by *Enterococcus* (*Streptococcus*) *faecalis* strains associated with human parenteral infections. J. Clin. Microbiol. **25:**1524–1528.
23. **Inoue, T., H. Tomita, and Y. Ike.** 2006. Bac 32, a novel bacteriocin widely disseminated among clinical isolates of *Enterococcus faecium*. Antimicrob. Agents Chemother. **50:**1202–1212.
24. **Khan, S. A.** 2005. Plasmid rolling-circle replication: highlights of two decades of research. Plasmid **53:**126–136.
25. **LeBlanc, D. J., L. N. Lee, and J. M. Inamine.** 1991. Cloning and nucleotide base sequence analysis of a spectinomycin adenyltransferase AAD(9) determinant from *Enterococcus faecalis*. Antimicrob. Agents Chemother. **35:** 1804–1810.
26. **Martínez-Bueno, M., A. Galvez, E. Valdivia, and M. Maqueda.** 1990. A transferable plasmid associated with AS-48 production in *Enterococcus faecalis*. J. Bacteriol. **172:**2817–2818.
27. **Nes, I. F., and H. Holo.** 2000. Class II antimicrobial peptides from lactic acid bacteria. Biopolymers **55:**50–61.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. **Tanimoto, K., and D. B. Clewell.** 1993. Regulation of the pAD1-encoded sex pheromone response in *Enterococcus faecalis*: expression of the positive regulator TraE1. J. Bacteriol. **175:**1008–1018.
30. **Todokoro, D., H. Tomita, T. Inoue, and Y. Ike.** 2006. Genetic analysis of bacteriocin 43 of vancomycin-resistant *Enterococcus faecium*. Appl. Environ. Microbiol. **72:**6955–6964.
31. **Tomich, P. K., F. Y. An, S. P. Damle, and D. B. Clewell.** 1979. Plasmid-related transmissibility and multiple-drug resistance in *Streptococcus faecalis* subsp. *zymogenes* DS16. Antimicrob. Agents Chemother. **15:**828–830.
32. **Tomita, H., S. Fujimoto, K. Tanimoto, and Y. Ike.** 1996. Cloning and genetic organization of the bacteriocin 31 determinant encoded on the *Enterococcus faecalis* pheromone-responsive conjugative plasmid pYI17. J. Bacteriol. **178:** 3585–3593.
33. **Tomita, H., S. Fujimoto, K. Tanimoto, and Y. Ike.** 1997. Cloning and genetic and sequence analyses of the bacteriocin 21 determinant encoded on the *Enterococcus faecalis* pheromone-responsive conjugative plasmid pPD1. J. Bacteriol. **179:**7843–7855.
34. **Tomita, H., and Y. Ike.** 2005. Genetic analysis of transfer-related regions of the vancomycin resistance *Enterococcus* conjugative plasmid pHTβ: identification of *oriT* and a putative relaxase gene. J. Bacteriol. **187:**7727–7737.
35. **Tomita, H., C. Pierson, S. K. Lim, D. B. Clewell, and Y. Ike.** 2002. Possible connection between a widely disseminated conjugative gentamicin resistance (pMG1-like) plasmid and the emergence of vancomycin resistance in *Enterococcus faecium*. J. Clin. Microbiol. **40:**3326–3333.
36. **Wirth, R., F. Y. An, and D. B. Clewell.** 1986. Highly efficient protoplast transformation system for *Streptococcus faecalis* and a new *Escherichia coli-S. faecalis* shuttle vector. J. Bacteriol. **165:**831–836.