Table 2. Locations of oligonucleotide probes designed to cover *M. tuberculosis gyrA*

Probe	Amino acid region detected by each probe	Nucleotide sequence*
S1	88–92	GATCGACGCGTCGCC
S2	92–97	CACCAGGSTGTCGTAGAT
R1	A90V mutation	GATCGAC**A**CGTCGCC
R2	D94G mutation	CACCAGGCTG**C**CGTAGAT
R3	D94A mutation	CACCAGGCTG**G**CGTAGAT
R4	D94G-S95T mutations	CACCAGG**G**TG**C**CGTAGAT
R5	D94A-S95T mutations	CACCAGG**G**TG**G**CGTAGAT

*S represents C or G; bold letters indicate mutations.

had an A90V mutation, 8 were positive for R4 (D94G mutation) and 9 were positive for R5 (D94A mutation). However, one isolate, NCGM2803, with ΔS2 was negative for R1–R5 indicating that the LiPA could not identify a mutation associated with FQ resistance. As shown in Table 1, the DNA sequencing data were fully consistent with results obtained by culture-based susceptibility testing. The FQ-resistant NCGM2803 isolate had a mutation of g280t (D94Y). In addition, it was reported that FQ-resistant isolates had mutations in *gyrB* but not in *gyrA* (Aubry *et al.*, 2006). Therefore, the present LiPA will be improved to detect the mutations.

The number of the MDR isolates used in this study is enough to allow estimation of the rate of FQ resistance in Japan. We tested 33 of 60 MDR isolates collected in the national surveillance study. JATA reported that there were 11 933 cases positive for sputum smear in Japan in 2002 (www.jata-org.jp). Whereas, 3122 isolates were collected for 6 months in 2002 during the national surveillance study (TRC, 2007). It is estimated that the number of isolates collected in the study covered more than 50 % of cases positive for sputum smear. Of the 3122 isolates, 60 isolates were reported to be MDR (TRC, 2007).

Occasional monitoring of the drug susceptibility of patients with MDR TB before and during chemotherapy is essential as MDR isolates can easily acquire resistance to additional antituberculosis drugs. We reported previously that more than 50 % of MDR isolates have already acquired PZA resistance (Ando *et al.*, 2010). In the present study, we found that nearly half of our Japanese MDR *M. tuberculosis* isolates were resistant to FQ. However, little information is available regarding resistance to other second-line drugs, and it will be necessary in future to monitor susceptibility to these drugs.

We strongly suggest that FQ susceptibility needs to be monitored as soon as patients are diagnosed with TB, because FQs are widely used for the treatment of bacterial infections. A population of patients with TB who received FQ treatment have been reported by Wang *et al.* (2006). Some other TB patients were initially treated as having community-acquired pneumonia and were administered FQs (Yoon *et al.*, 2005). We could not obtain information

about the TB patients whose isolates were tested in the present study, especially about their previous treatments with antituberculosis drugs, including FQs. Nevertheless, it is very likely that these patients received FQs. In Japan, FQs were used for patients with TB, especially MDR TB, before 2008 but LVFX, MFLX, GFLX, SPFX and CPFX were not officially recommended as antituberculosis drugs by the Japanese Society for Tuberculosis until 2008.

It is also necessary to develop a rapid and inexpensive diagnostic method to determine the drug susceptibility of *M. tuberculosis*. The whole procedure of the LiPA described here takes only 9 h, and the estimated cost per sample is £22 (US $35). A DNA sequencing-based method (Sekiguchi *et al.*, 2007a) is also rapid, but is more expensive than the LiPA. Therefore, the LiPA is suitable for this purpose. Clinical trials for *in vitro* diagnosis including non-MDR TB are in progress in Japan.

ACKNOWLEDGEMENTS

REFERENCES

Ando, H., Mitarai, S., Kondo, Y., Suetake, T., Sekiguchi, J. I., Kato, S., Mori, T. & Kirikae, T. (2010). Pyrazinamide resistance in multidrug-resistant *Mycobacterium tuberculosis* isolates in Japan. *Clin Microbiol Infect* 16, 1164–1168.

Aubry, A., Veziris, N., Cambau, E., Truffot-Pernot, C., Jarlier, V. & Fisher, L. M. (2006). Novel gyrase mutations in quinolone-resistant and -hypersusceptible clinical isolates of *Mycobacterium tuberculosis*: functional analysis of mutant enzymes. *Antimicrob Agents Chemother* 50, 104–112.

Aziz, M., Laszlo, A., Raviglione, M. C., Rieder, H. L., Espinal, M. A. & Wright, A. (2003). *Guidelines for Surveillance of Drug Resistance in Tuberculosis.* Geneva: World Health Organization.

Caminero, J. A. (2010). Multidrug-resistant tuberculosis: epidemiology, risk factors and case finding. *Int J Tuberc Lung Dis* 14, 382–390.

Conde, M. B., Efron, A., Loredo, C., De Souza, G. R., Graca, N. P., Cezar, M. C., Ram, M., Chaudhary, M. A., Bishai, W. R. & other authors (2009). Moxifloxacin versus ethambutol in the initial treatment of tuberculosis: a double-blind, randomised, controlled phase II trial. *Lancet* 373, 1183–1189.

Fujiki, A. (2001). *TB Bacteriology Examination to Stop TB*. Tokyo: The Research Institute of Tuberculosis.

Guillemin, I., Jarlier, V. & Cambau, E. (1998). Correlation between quinolone susceptibility patterns and sequences in the A and B subunits of DNA gyrase in mycobacteria. *Antimicrob Agents Chemother* 42, 2084–2088.

Johnston, J. C., Shahidi, N. C., Sadatsafavi, M. & Fitzgerald, J. M. (2009). Treatment outcomes of multidrug-resistant tuberculosis: a systematic review and meta-analysis. *PLoS ONE* 4, e6914.

Sekiguchi, J., Miyoshi-Akiyama, T., Augustynowicz-Kopec, E., Zwolska, Z., Kirikae, F., Toyota, E., Kobayashi, I., Morita, K., Kudo, K. & other authors (2007a). Detection of multidrug resistance in *Mycobacterium tuberculosis*. *J Clin Microbiol* 45, 179–192.

Sekiguchi, J., Nakamura, T., Miyoshi-Akiyama, T., Kirikae, F., Kobayashi, I., Augustynowicz-Kopec, E., Zwolska, Z., Morita, K., Suetake, T. & other authors (2007b). Development and evaluation of a line probe assay for rapid identification of *pncA* mutations in pyrazinamide-resistant *Mycobacterium tuberculosis* strains. *J Clin Microbiol* 45, 2802–2807.

Sotgiu, G., Ferrara, G., Matteelli, A., Richardson, M. D., Centis, R., Ruesch-Gerdes, S., Toungoussova, O., Zellweger, J. P., Spanevello, A. & other authors (2009). Epidemiology and clinical management of XDR-TB: a systematic review by TBNET. *Eur Respir J* 33, 871–881.

TRC (2007). Drug-resistant *Mycobacterium tuberculosis* in Japan: a nationwide survey, 2002. *Int J Tuberc Lung Dis* 11, 1129–1135.

Van Deun, A., Martin, A. & Palomino, J. C. (2010a). Diagnosis of drug-resistant tuberculosis: reliability and rapidity of detection. *Int J Tuberc Lung Dis* 14, 131–140.

Van Deun, A., Maug, A. K., Salim, M. A., Das, P. K., Sarker, M. R., Daru, P. & Rieder, H. L. (2010b). Short, highly effective, and inexpensive standardized treatment of multidrug-resistant tuberculosis. *Am J Respir Crit Care Med* 182, 684–692.

Wang, J. Y., Hsueh, P. R., Jan, I. S., Lee, L. N., Liaw, Y. S., Yang, P. C. & Luh, K. T. (2006). Empirical treatment with a fluoroquinolone delays the treatment for tuberculosis and is associated with a poor prognosis in endemic areas. *Thorax* 61, 903–908.

Yew, W. W., Lange, C. & Leung, C. C. (2010). Treatment of tuberculosis update 2010. *Eur Respir J* (in press). doi: 10.1183/09031936.00033010

Yoon, Y. S., Lee, H. J., Yoon, H. I., Yoo, C. G., Kim, Y. W., Han, S. K., Shim, Y. S. & Yim, J. J. (2005). Impact of fluoroquinolones on the diagnosis of pulmonary tuberculosis initially treated as bacterial pneumonia. *Int J Tuberc Lung Dis* 9, 1215–1219.

Zhang, Y. & Telenti, A. (2000). *Genetics of Drug Resistance in Mycobacterium tuberculosis*. Washington, DC: American Society for Microbiology.

Downregulation of *katG* expression is associated with isoniazid resistance in *Mycobacterium tuberculosis*

Hiroki Ando,[1] Tomoe Kitao,[1]
Tohru Miyoshi-Akiyama,[1] Seiya Kato,[2] Toru Mori[2]
and Teruo Kirikae[1]*
[1]National Center for Global Health and Medicine, 1-21-1
Toyama, Shinjuku, Tokyo 162-8655, Japan.
[2]Research Institute of Tuberculosis, Japan
Anti-Tuberculosis Association, Matsuyama 3-1-24,
Kiyose, Tokyo 204-8533, Japan.

Summary

Isoniazid (INH) is a key agent in the treatment of tuberculosis. In *Mycobacterium tuberculosis*, INH is converted to its active form by KatG, a catalase-peroxidase, and attacks InhA, which is essential for the synthesis of mycolic acids. We sequenced *furA–katG* and *fabG1–inhA* in 108 INH-resistant (INH^r) and 51 INH-susceptible (INH^s) isolates, and found three mutations in the *furA–katG* intergenic region (Int^g-7a, Int^a-10c and Int^g-12a) in four of 108 INH^r isolates (4%), and the *furA* ^c41t mutation with an amino acid substitution in 18 INH^r isolates (17%). These mutations were not found in any of 51 INH^s isolates tested. We reconstructed these mutations in isogenic strains to determine whether they conferred INH resistance. We found that the Int^g-7a, Int^a-10c and Int^g-12a single mutations in the *furA–katG* intergenic region decreased *katG* expression and conferred INH resistance. In contrast, the *furA* ^c41t mutation was not sufficient to confer INH resistance. These results suggested that downregulation of *katG* is a mechanism of INH resistance in *M. tuberculosis* and that mutations in the *furA–katG* intergenic region play a role in this resistance mechanism.

Introduction

Isoniazid (INH) is one of the most effective and specific antituberculosis drugs, which has been a key to treatment since its introduction in 1952 (Bernstein *et al.*, 1952). However, the continued utility of INH is being

jeopardized by the emergence of INH-resistant (INH^r) and multidrug-resistant *Mycobacterium tuberculosis* (Espinal *et al.*, 2001). Furthermore, INH resistance is often the first type of resistance that occurs on the way to multidrug resistance (Caminero, 2010). Thus, there is considerable interest in determining the molecular basis of INH resistance in clinical *M. tuberculosis* isolates.

INH enters the mycobacterial cell by passive diffusion (Bardou *et al.*, 1998). INH is a prodrug that is activated by the mycobacterial enzyme KatG (Zhang *et al.*, 1992). Although mutations in *katG* have been shown to be responsible for INH resistance (Zhang and Telenti, 2000), it is not clear whether the regulation of *katG* expression plays a role in INH resistance. The *katG* gene encodes a bifunctional catalase-peroxidase that converts INH to its active form (Zhang *et al.*, 1992). Activated INH inhibits the synthesis of essential mycolic acids by inactivating the NADH-dependent enoyl-acyl carrier protein reductase encoded by *inhA* (Banerjee *et al.*, 1994; Vilcheze *et al.*, 2006). We have developed a DNA sequencing-based method to detect mutations in eight genome regions associated with drug resistance in *M. tuberculosis*, including *katG* (Sekiguchi *et al.*, 2007). This led to the identification of 18 novel *katG* mutations, which caused loss of catalase and INH oxidation activities (Sekiguchi *et al.*, 2007; Ando *et al.*, 2010). However, no mutations in *katG* were found during screening of a significant population of INH^r *M. tuberculosis* isolates.

We describe here the identification of three novel mutations in the *furA–katG* intergenic region, which are associated with INH resistance in *M. tuberculosis*. To our knowledge, no previous reports have identified mutations in the *furA–katG* intergenic region related to INH resistance in *M. tuberculosis*. A polypurine sequence in the *furA–katG* intergenic region complementary to 16S rRNA has been found to act as the *katG* ribosome binding site (Sala *et al.*, 2008). Mutations altering complementarity to 16S rRNA decreased the level of transcription (Sala *et al.*, 2008). We found that single mutations in these polypurine sequences decreased *katG* expression and conferred INH resistance in *M. tuberculosis*.

Results

Genetic mutations in INH^r isolates

We sequenced *furA–katG* and *fabG1–inhA* in 108 INH-resistant (INH^r) and 51 INH-susceptible (INH^s) isolates. The details of the results for all INH^r isolates have been reported previously (Ando *et al.*, 2010). These regions included *katG*, the upstream region of *fabG1–inhA* (P$_{inhA}$), and *inhA*, which are responsible for INH resistance (Zhang and Telenti, 2000). We found three mutations in the *furA–katG* intergenic region (Int^{g-7a}, Int^{a-10c} and Int^{g-12a}) in four (4%) of the 108 INH^r isolates, a *furA*^{c41t} mutation with an amino acid substitution in 18 INH^r isolates (17%), and a *furA*$^{\Delta34c}$ mutation with a frameshift in one INH^r isolate (1%) (Table 1). None of these mutations was found in any of the 51 INH^s isolates tested (data not shown). In addition, we sequenced *ndh*, *ahpC* and *kasA–kasB* in all the INH^r isolates tested. Mutations in *ndh* (Zhang and Telenti, 2000; Lee *et al.*, 2001), *ahpC* (Zhang and Telenti, 2000) and *kasA* (Lee *et al.*, 1999; Zhang and Telenti, 2000) have been found in INH^r clinical isolates of *M. tuberculosis*. Mutations in *ndh* have been found to mediate INH resistance by increasing NADH cellular concentration in mycobacteria (Vilcheze *et al.*, 2005), whereas the roles of *ahpC* and *kasA* in INH resistance have not yet been determined (Zhang and Telenti, 2000). Of the four isolates containing mutations in the *furA–katG* intergenic region (Int^{g-7a}, Int^{a-10c} and Int^{g-12a}), the isolate with the Int^{g-12a} mutation also had a mutation in the P$_{inhA}$ (P$_{inhA}$$^{c-15t}$) (Table 1). Of the 18 isolates containing the *furA*^{c41t} mutation, seven had no other mutation, six had a P$_{inhA}$$^{c-15t}$ mutation, one had a *katG*^{g1624c} mutation and a mutation in the region upstream of *ahpC*, three had a *katG*^{g944c} mutation, and one had *katG*^{a290g} and *fabG1*^{g609a} mutations. The isolate with the *furA*$^{\Delta34c}$ mutation also had three other mutations, *katG*^{g368a}, *katG*^{g895a} and P$_{inhA}$$^{t-8c}$. Two drug-susceptible control strains, H37Rv and NCGM2898, had no mutation (Table 1). An INH^r strain of NCGM2828 had the P$_{inhA}$$^{c-15t}$ mutation. An INH^r strain of NCGM2836 had a large-scale deletion of *furA–katG*, as described previously (Ando *et al.*, 2010). All isolates tested except H37Rv had a g1388t mutation in *katG*, which is a neutral mutation that is not associated with INH resistance (Zhang and Telenti, 2000).

IS6110-probed restriction fragment length polymorphism (RFLP) analysis of isolates harbouring mutations in the furA–katG intergenic region or furA

The results of IS6110-probed fingerprinting of the 159 isolates, the 108 INH^r isolates and the 51 INH^s isolates, are shown in Fig. S1. Mutations in the *furA–katG* intergenic region and *furA* are also shown in Fig. S1. Six clusters were detected by restriction fragment length poly-

morphism (RFLP) analysis, with 65 INH^r isolates (60%) belonging to these clusters. Of the four isolates with mutations in the *furA–katG* intergenic region, one belonged to cluster III, one to cluster IV, and the remaining two did not belong to any cluster. Among the 18 isolates with *furA*^{c41t}, 12 (67%) belonged to cluster II, and the remaining six did not belong to any cluster. Of the 12 isolates belonging to cluster II, five (NCGM2806, 2807, 2833, 2853 and 2945) had another mutation, P$_{inhA}$$^{c-15t}$; four (NCGM2823, 2829, 2846 and 2881) had a *katG* mutation; and three (NCGM2790, 2812 and 2821) did not have any other mutation (Table 1). These findings suggest that the majority of INH^r isolates expanded in a clonal manner in Japan. The isolates with Int^{g-7a}, Int^{a-10c} and Int^{g-12a} expanded sporadically in Japan; about two-thirds of the isolates with *furA*^{c41t} mutations expanded clonally, while the others did so sporadically.

Construction of isogenic strains with Int^{g-7a}, Int^{a-10c}, Int^{g-12a} and furA^{c41t}

The NCGM2836 strain was resistant to INH (1.0 µg ml^{-1}) and had deletions of *furA* and *katG* (Ando *et al.*, 2010). This strain had a deletion of 5603 bp, extending from position 2 152 072 to 2 157 675 of the H37Rv sequence. This was predicted to result in the complete loss of *furA*, *katG*, and the adjacent genes *rv1906c*, *rv1907c* and *rv1910c*, as well as the disruption of *aao* and *lppC* (Fig. 1A).

We assembled a series of constructs to complement the Δ(*furA–katG*) mutant strain NCGM2836 with pMV306 (Fig. 1B and C), by site-specific integration into the chromosomal *attB* site (Stover *et al.*, 1991; Kong and Kunimoto, 1995). The constructs had no *furA–katG* (NCGM2836/Vector); or *furA–katG* harbouring no mutation (NCGM2836/WT), *furA*^{c41t} (NCGM2836/*furA*^{c41t}), Int^{g-7a} (NCGM2836/Int^{g-7a}), Int^{a-10c} (NCGM2836/Int^{a-10c}) or Int^{g-12a} (NCGM2836/Int^{g-12a}).

### Reduced katG expression in the complemented strains with a mutation in the furA–katG intergenic region

We found that H37Rv expressed KatG, whereas the Δ(*furA–katG*) mutants NCGM2836 and NCGM2836/ Vector did not (Fig. 2A). NCGM2836/WT and NCGM2836/ *furA*^{c41t} expressed KatG at levels similar to that of H37Rv, whereas NCGM2836/Int^{g-7a} and NCGM2836/Int^{a-10c} expressed KatG at levels 17% and 19%, respectively, that of NCGM2836/WT. The strain NCGM2836/Int^{g-12a} produced KatG, but its expression level was slightly lower (66%) than that of NCGM2836/WT.

The INH oxidase activities of these strains were correlated with their levels of KatG expression (Fig. 2B). H37Rv showed INH oxidase activity, while NCGM2836

Table 1. The drug susceptibility and mutation(s) of *M. tuberculosis* clinical isolates.

Isolate	Drug susceptibility — Agar proportion method (by Vit Spectrum-SR)	MIC of INH (µg ml⁻¹) (by BrothMIC MTB-I)	furA–katG Up	furA	Int	katG	fabG1–inhA Up	fabG1	Int	inhA	ahpC Up	ahpC	ndh Up	ndh	kasA	kasA–kasB Int	kasB
NCGM2874	INH(0.2), SM	1.0	–	–	g-7a	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2875	INH(0.2), SM	1.0	–	–	g-7a	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2930	INH(1.0), RIF, SM, KM, EB, TH, PAS, LVFX	16.0	–	–	a-10c	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2934	INH(0.2), RIF, SM, EB	2.0	–	c41t	g-12a	–	c-15t	–	–	–	–	–	–	–	–	–	–
NCGM2790	INH(0.2), SM, KM, PAS	2.0	–	c41t	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2795	INH(0.2), SM	1.0	–	c41t	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2812	INH(0.2)	1.0	–	c41t	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2821	INH(1.0)	1.0	–	c41t	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2906	INH(0.2), LVFX	0.5	–	c41t	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2907	INH(0.2)	1.0	–	c41t	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2911	INH(0.2)	1.0	–	c41t	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2806	INH(0.2), SM	2.0	–	c41t	–	–	c-15t	–	–	–	–	–	–	–	–	–	–
NCGM2807	INH(0.2), SM	2.0	–	c41t	–	–	c-15t	–	–	–	–	–	–	–	–	–	–
NCGM2833	INH(0.2), SM	2.0	–	c41t	–	–	c-15t	–	–	–	–	–	–	–	–	–	–
NCGM2853	INH(0.2), SM	2.0	–	c41t	–	–	c-15t	–	–	–	–	–	–	–	–	–	–
NCGM2858	INH(0.2), SM	2.0	–	c41t	–	–	c-15t	–	–	–	–	–	–	–	–	–	–
NCGM2945	INH(0.2)	1.0	–	c41t	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2823	INH(1.0)	>32.0	–	c41t	–	g1624c	–	–	–	–	c-81t	–	–	–	–	–	–
NCGM2829	INH(1.0)	>32.0	–	c41t	–	g944c	–	–	–	–	–	–	–	–	–	–	–
NCGM2846	INH(1.0)	>32.0	–	c41t	–	g944c	–	–	–	–	–	–	–	–	–	–	–
NCGM2881	INH(0.2), KM, EVM, TH	16.0	–	c41t	–	a290g	–	g609a	–	–	–	–	–	–	–	–	–
NCGM2895	INH(1.0)	>32.0	–	c41t	–	g944c	–	–	–	–	–	–	–	–	–	–	–
NCGM2835	INH(1.0), RIF, SM, EB, CS, TH, PAS, LVFX	>32.0	–	Δ34c	–	g368a, g895a	t-8c	–	–	–	–	–	–	–	–	–	–
(Controls)																	
H37Rv	Pan-sensitive	0.125	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2898	Pan-sensitive	0.125	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
NCGM2828	INH(0.2)	2.0	–	–	–	–	c-15t	–	–	–	–	–	–	–	–	–	–
(A host for re-construction of a mutation)																	
NCGM2836	INH(1.0), PAS	>32.0	Deletion														

INH(0.2), resistant to INH (0.2 µg ml⁻¹) and susceptible to INH (1.0 µg ml⁻¹); INH(1.0), resistant to INH (1.0 µg ml⁻¹); RIF, rifampin; EB, ethambutol; KM, kanamycin; PAS, p-aminosalicylic acid; SM, streptomycin; TH, ethionamide; EVM, enviomycin; CS, cycloserine; LVFX, levofloxacin; Up, upstream region; Int, intergenic region; –, no mutation. Excluding the *katG* g1388t (KatG R463L) polymorphism.
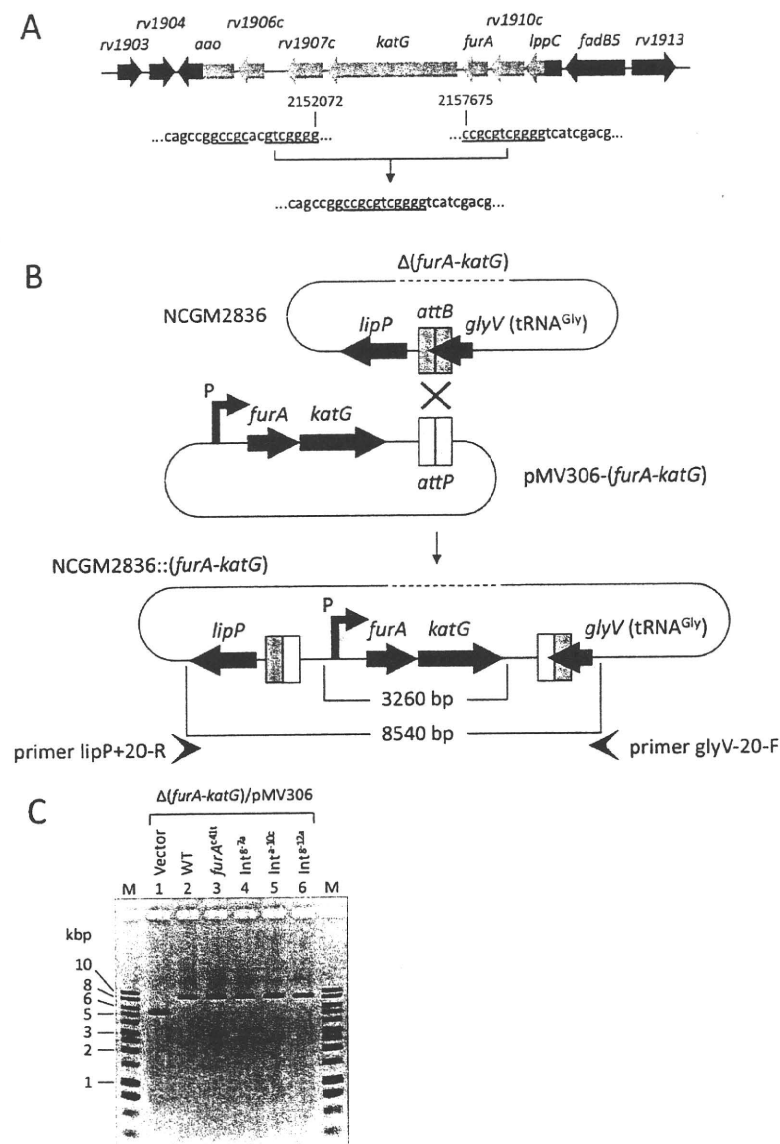
Fig. 1. A. Maps of the deleted regions adjacent to *furA–katG* in INH[r] *M. tuberculosis* strain NCGM2836. Bold arrows indicate the open reading frames annotated in the H37Rv genome sequence (http://tuberculist.epfl.ch/). The grey bold arrows correspond to the deleted regions, with the end sequences and H37Rv genome co-ordinates given below. Underlined sequences are possible substrates for recombination.
B. Schematic representation of the construction of complementary strains. An integration-proficient vector pMV306 containing *furA–katG* with the Int[g-7a], Int[a-10c], Int[g-12a] or *furA*[c41t] mutation was used to transform NCGM2836. Site-specific integration between *attB* and *attP* sites generated complementary isogenic strains.
C. The complementary isogenic strains were examined for the correct chromosomal structures by PCR using locus-specific primers (glyV-20-F/lipP+20-R). M, size marker.

and NCGM2836/Vector did not. NCGM2836/WT and NCGM2836/*furA*[c41t] showed similar levels of activity as H37Rv, whereas NCGM2836/Int[g-7a], NCGM2836/Int[a-10c] and NCGM2836/Int[g-12a] showed significantly lower levels of activity than NCGM2836/WT.

These experiments using complementary strains demonstrated that the mutations Int[g-7a], Int[a-10c] and Int[g-12a] were associated with reductions in KatG expression and INH oxidase activities, but that the *furA*[c41t] mutation was not.

INH susceptibility of the complementary strains

The complementary strains NCGM2836/Int[g-7a], NCGM2836/Int[a-10c] and NCGM2836/Int[g-12a] showed low levels of INH resistance (Fig. 3). NCGM2836/WT was resistant to 0.05 μg ml⁻¹ but susceptible to 0.1 μg ml⁻¹ INH, whereas NCGM2836/Vector was resistant to 1.0 μg ml⁻¹ INH. NCGM2836/*furA*[c41t] showed the same INH susceptibility as NCGM2836/WT, being resistant to 0.05 μg ml⁻¹ but susceptible to 0.1 μg ml⁻¹ INH. NCGM2836/Int[g-7a] and NCGM2836/Int[a-10c] were resistant to 0.2 μg ml⁻¹ but susceptible to 0.4 μg ml⁻¹ INH, whereas NCGM2836/Int[g-12a] was resistant to 0.1 μg ml⁻¹ but susceptible to 0.15 μg ml⁻¹ INH.

Three other INH susceptibility tests showed similar results (Table 2). The three complementary strains with mutations in the *furA–katG* intergenic region showed low levels of INH resistance. NCGM2836/Int[g-7a] and

A
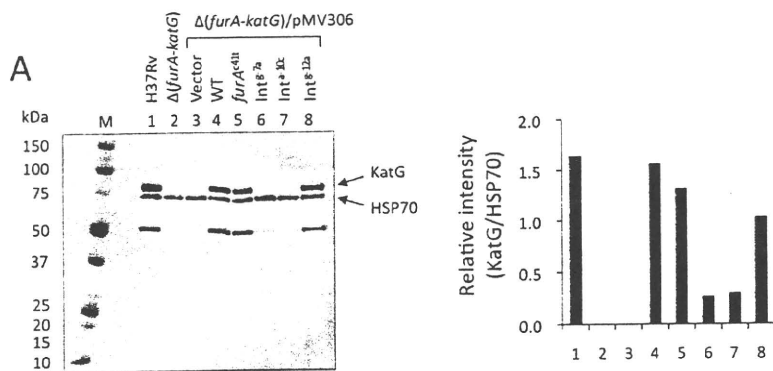
Fig. 2. KatG expression of complementary strains with mutations in *furA–katG*.
A. KatG and HSP70 expression were analysed in complementary strains by SDS-5%–20%-gradient-PAGE and Western blotting. The intensity of KatG was normalized relative to HSP70. This experiment was repeated three times with similar results. M, size marker.
B. INH oxidase activity of KatG in complementary strains. Each transformant was incubated in triplicate, and INH oxidase activity was measured. Data are shown as the means ± SD (*$P < 0.01$).

B
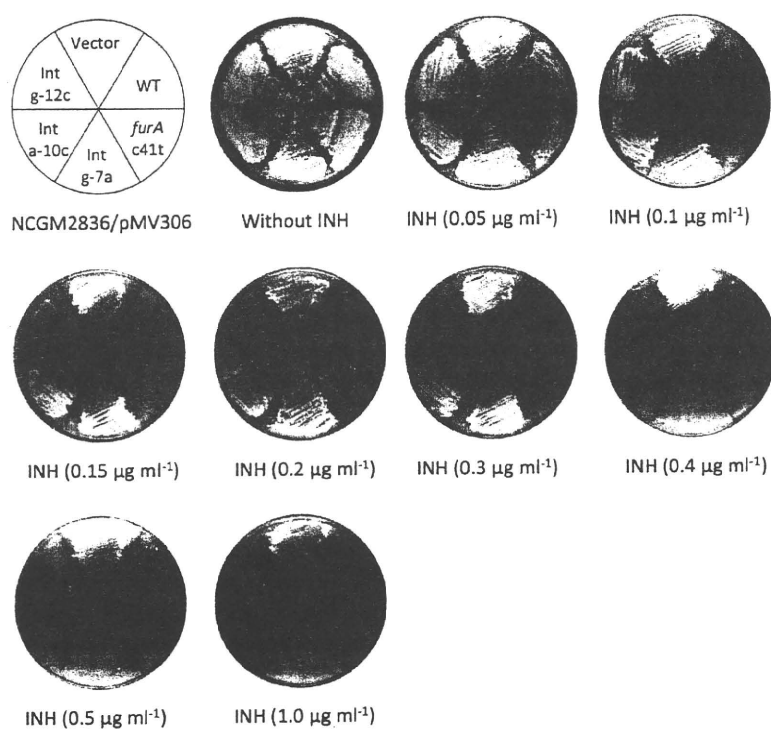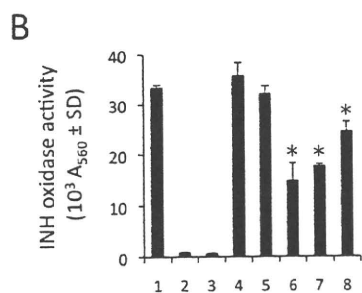
Fig. 3. INH susceptibility of complementary strains. *M. tuberculosis* strains were streaked on plates containing increasing concentrations of INH as indicated. The plates were incubated at 37°C for 2 weeks. Two clones of transformants were obtained by picking colonies in each complementary strain. INH susceptibility testing was performed twice using different clones. The same MICs were obtained in the two experiments.

Table 2. INH susceptibility of *M. tuberculosis* strain NCGM2836 Δ(*furA–katG*) transformants.

| | Genotype | | | Proportion method | | | | Dilution method [MIC (µg ml⁻¹)] | |
| | | | | 7H10 agar[a] | | Egg-based Ogawa medium[b] | | | |
Strain	*furA*	Int[e]	*katG*	0.2[f]	1.0[g]	0.2[f]	1.0[g]	7H10 agar[c]	7H9 broth[d]
H37Rv	wt	wt	wt	S	S	S	S	0.05	0.125
Δ(*furA–katG*)/Vector	Deletion			R	R	R	R	>1.0	>32
Δ(*furA–katG*)/WT	wt	wt	wt	S	S	S	S	0.1	0.125
Δ(*furA–katG*)/furA^c41t	c41t	wt	wt	S	S	S	S	0.1	0.25
Δ(*furA–katG*)/Int^g-7a	wt	g-7a	wt	R	S	R	S	0.4	2.0
Δ(*furA–katG*)/Int^a-10c	wt	a-10c	wt	R	S	R	S	0.4	2.0
Δ(*furA–katG*)/Int^g-12a	wt	g-12a	wt	S	S	S	S	0.15	0.5

a. An agar proportion method using Middlebrook 7H10 agar medium recommended by Clinical and Laboratory Standards Institute. The susceptibility testing was performed twice with the same results.
b. A proportion method using egg-based Ogawa medium modified by WHO protocol and recommended by the Japanese Society of Tuberculosis. The susceptibility testing was performed twice with the same results.
c. A dilution method using Middlebrook 7H10 agar medium. The susceptibility testing was performed twice with the same results.
d. A microdilution method using Middlebrook 7H9 broth medium. The susceptibility testing was performed twice with the same results.
e. Int: intergenic region.
f. 0.2: resistant to INH (0.2 µg ml⁻¹) and susceptible to INH (1.0 µg ml⁻¹).
g. 1.0: resistant to INH (1.0 µg ml⁻¹).

NCGM2836/Int^a-10c were more resistant than NCGM2836/Int^g-12a. H37Rv and NCGM2836/WT were susceptible to INH, whereas NCGM2836/Vector was resistant. NCGM2836/furA^c41t was susceptible to INH (0.2 µg ml⁻¹) by proportion methods, with MICs of 0.1 µg ml⁻¹ as determined by the 7H10 agar dilution method and 0.25 µg ml⁻¹ as determined by the 7H9 broth dilution method. Although furA^c41t showed twofold higher INH resistance than NCGM2836/WT by the 7H9 broth dilution method, this difference was not significant relative to the error associated with this type of experiment. NCGM2836/Int^g-7a and NCGM2836/Int^a-10c had identical INH susceptibility profiles, i.e. they were resistant to INH (0.2 µg ml⁻¹) and their MICs were 0.4 and 2.0 µg ml⁻¹ respectively. NCGM2836/Int^g-12a was susceptible to INH (0.2 µg ml⁻¹), with MICs of 0.15 and 0.5 µg ml⁻¹ respectively. These data indicated that Int^g-7a and Int^a-10c mutations confer resistance to INH (0.2 µg ml⁻¹) and that the Int^g-12a mutation causes a slight decrease in INH susceptibility.

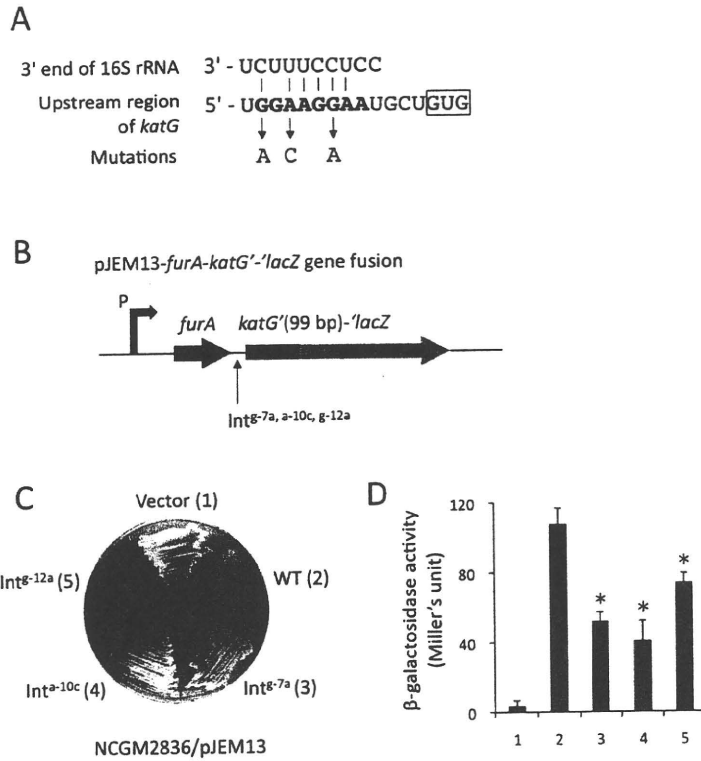### Regulation of katG expression by the furA–katG intergenic region

A polypurine sequence (GGAAGGAA) was identified in the *furA–katG* intergenic region, which is complementary to the 3′ end of the 16S rRNA sequence in *M. tuberculosis* (Fig. 4A) (Sala *et al.*, 2008). All three mutations in the intergenic region identified in INH^r clinical isolates, i.e. Int^g-7a, Int^a-10c and Int^g-12a, were located within this polypurine sequence (Fig. 4A). A reporter assay using *furA–katG′-′lacZ* gene fusions (Fig. 4B) showed that these mutations significantly repressed *lacZ* expression

(Fig. 4C and D). These results suggested that the intergenic region plays a critical role in *katG* expression and that these mutations are responsible for INH resistance by decreasing *katG* expression.

Reduced katG expression in INHʳ clinical isolates with a mutation in the furA–katG intergenic region

The four clinical isolates with mutations in the *furA–katG* intergenic region (Table 1) were tested for their levels of *katG* expression by Western blotting (Fig. 5A) and KatG activity by INH oxidation assays (Fig. 5B). The INH^r Δ(*furA–katG*) mutant NCGM2836 did not produce *katG*, whereas H37Rv, INH^s NCGM2898 and INH^r NCGM2828 with P_inhA^c-15t expressed *katG* at similar levels. The isolates containing the Int^g-7a (NCGM2874 and NCGM2875) and Int^a-10c (NCGM2930) mutations expressed little *katG*. An isolate with Int^g-12a (NCGM2934) expressed *katG*, but at a level lower than that of INH^s controls (68% of H37Rv and 76% of INH^s NCGM2898). The levels of INH oxidase activities were correlated with those of *katG* expression (Fig. 5B). Four isolates with Int^g-7a, Int^a-10c and Int^g-12a mutations showed significantly lower levels of INH oxidase activity than the two INH^s controls and the one INH^r isolate with the P_inhA^c-15t mutation; these levels were 21%, 21%, 29% and 57%, respectively, when compared to H37Rv (Fig. 5B). These results suggested that these mutations in the *furA–katG* intergenic region decrease the levels of *katG* expression and INH oxidase activity. Reduced *katG* expression is not a general feature of INH^r isolates, because an INH^r NCGM2828 with P_inhA^c-15t expressed *katG* at similar levels to INH^s isolates.

A

3' end of 16S rRNA   3' - UCUUUCCUCC
                            | | | | | |
Upstream region   5' - **UGGAAGGAA**UGCU$\boxed{\text{GUG}}$
of *katG*
                              ↓ ↓   ↓
Mutations   A C   A

B

pJEM13-*furA-katG'-'lacZ* gene fusion

P
furA *katG'*(99 bp)-*'lacZ*

Int$^{\text{g-7a, a-10c, g-12a}}$

C

Vector (1)

Int$^{\text{g-12a}}$ (5)   WT (2)

Int$^{\text{a-10c}}$ (4)   Int$^{\text{g-7a}}$ (3)

NCGM2836/pJEM13

D

Fig. 4. β-Galactosidase activity of the *furA–katG'–'lacZ* fused gene carrying the Int$^{\text{g-7a}}$, Int$^{\text{a-10c}}$ or Int$^{\text{g-12a}}$ mutation.
A. Complementarities between the 3' end of 16S rRNA and the polypurine sequence in the *furA–katG* intergenic region. The polypurine sequence is shown in bold and the translational initiation codon of *katG* is boxed.
B. Schematic representation of the construction of *furA–katG'–'lacZ* gene fusion. A fragment from 138 bp upstream of the first nucleotide of *furA* to the 99th base of *katG* with Int$^{\text{g-7a}}$, Int$^{\text{a-10c}}$ or Int$^{\text{g-12a}}$ mutation was fused to the *lacZ* gene. pJEM13-derived vectors were used to transform the NCGM2836 Δ(*furA–katG*) strain.
C. β-Galactosidase activity was monitored by streaking strains on 7H10 plates containing X-gal. The plates were incubated at 37°C for 2 weeks. Two clones of transformants were obtained by picking colonies in each strain. Duplicate experiments with different clones yielded similar results.
D. Quantitative β-galactosidase assay. Each transformant was incubated in triplicate, and β-galactosidase activity was measured. Data are shown as the means ± SD (*$P < 0.01$).

A

B
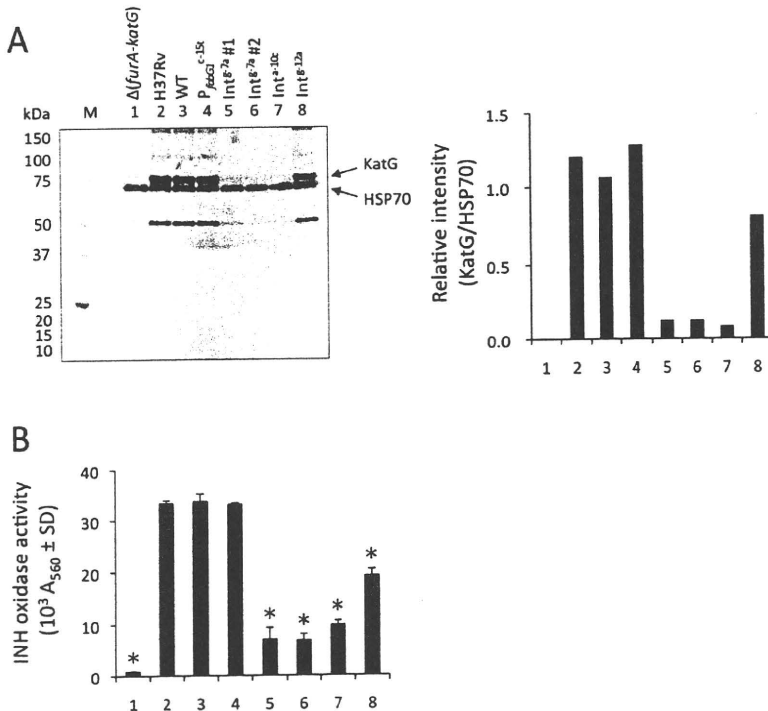
Fig. 5. KatG expression of clinical isolates with Int$^{\text{g-7a}}$, Int$^{\text{a-10c}}$ or Int$^{\text{g-12a}}$ mutation.
A. KatG and HSP70 expression in clinical isolates were analysed by SDS-5%–20%-gradient-PAGE and Western blotting. The intensity of KatG was normalized relative to HSP70. This experiment was repeated twice with similar results. M, size marker.
B. INH oxidase activity of KatG in clinical isolates. Each transformant was incubated in triplicate, and INH oxidase activity was measured. Data are shown as the means ± SD (*$P < 0.01$).
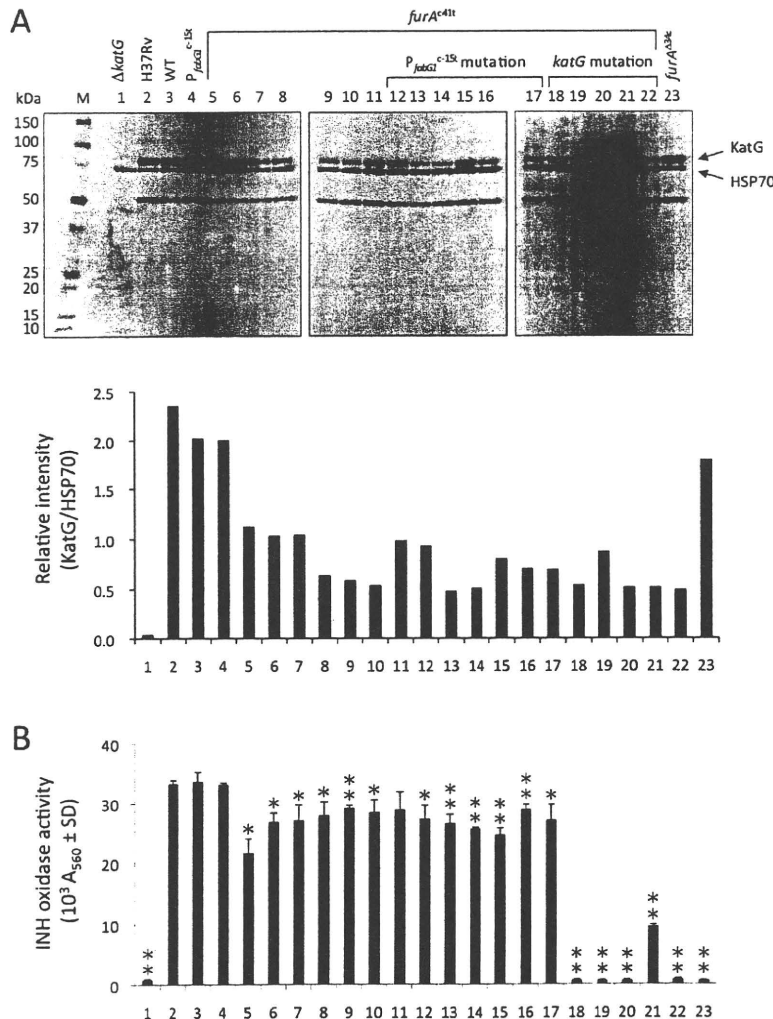
Reduced katG *expression in INH* *clinical isolates with* furA *c41t mutation*

We also tested 18 clinical isolates with *furA* c41t mutation and the one clinical isolate with a frameshift of *furA* (*furA*Δ34c) (Table 1) for their levels of *katG* expression by Western blotting (Fig. 6A) and for KatG activity by INH oxidation assays (Fig. 6B). The INHʳ Δ(*furA–katG*) mutant did not produce *katG*. H37Rv, INHˢ NCGM2898 and INHʳ NCGM2828 with PinhA c-15t expressed *katG* at similar levels. All 18 INHʳ isolates with *furA* c41t tested produced smaller quantities of KatG than did the INHˢ isolates and the INHʳ isolates with PinhA c-15t, whereas the *furA*Δ34c mutant produced similar amounts of KatG as those of INHˢ isolates (Fig. 6A). Quantification of these findings showed that the levels of KatG production in INHʳ isolates with *furA* c41t were 20–48% those of H37Rv and 23–55% those of NCGM2898. All of the clinical isolates with *furA* c41t without *katG* mutations showed significantly lower levels of INH

oxidase activity than did the two INHˢ controls and the one INHʳ isolate with PinhA c-15t, but these levels were only 64–87% those of the INHˢ controls (Fig. 6B). Four of the five isolates with *furA* c41t and *katG* mutations showed no INH oxidase activity, whereas the fourth, NCGM2881, had significantly lower activity, corresponding to 28% of the activity in INHˢ controls (Fig. 6B). The *furA*Δ34c mutant with two *katG* mutations (Table 1) did not show any INH oxidase activity (Fig. 6B).

*FurA*A14V *binding to the region upstream of* furA–katG

FurA is a negative transcriptional regulator of the *furA–katG* gene (Zahrt *et al.*, 2001). In addition, FurA may regulate other genes involved in pathogenesis (Pym *et al.*, 2001). In *Pseudomonas aeruginosa* Fur (PaFur), the loop between the N-terminal helices H1 and H2 is thought to be involved in DNA recognition (Pohl *et al.*,

Fig. 6. KatG expression of clinical isolates with the *furA* c41t mutation.
A. KatG and HSP70 expression in clinical isolates were analysed by SDS-5%–20%-gradient-PAGE and Western blotting. Relative intensity of KatG was normalized against HSP70. This experiment was repeated twice with similar results. M, size marker.
B. INH oxidase activity of KatG in clinical isolates. Each transformant was incubated in triplicate, and INH oxidase activity was measured. Data are shown as the means ± SD (*$P < 0.05$, **$P < 0.01$).
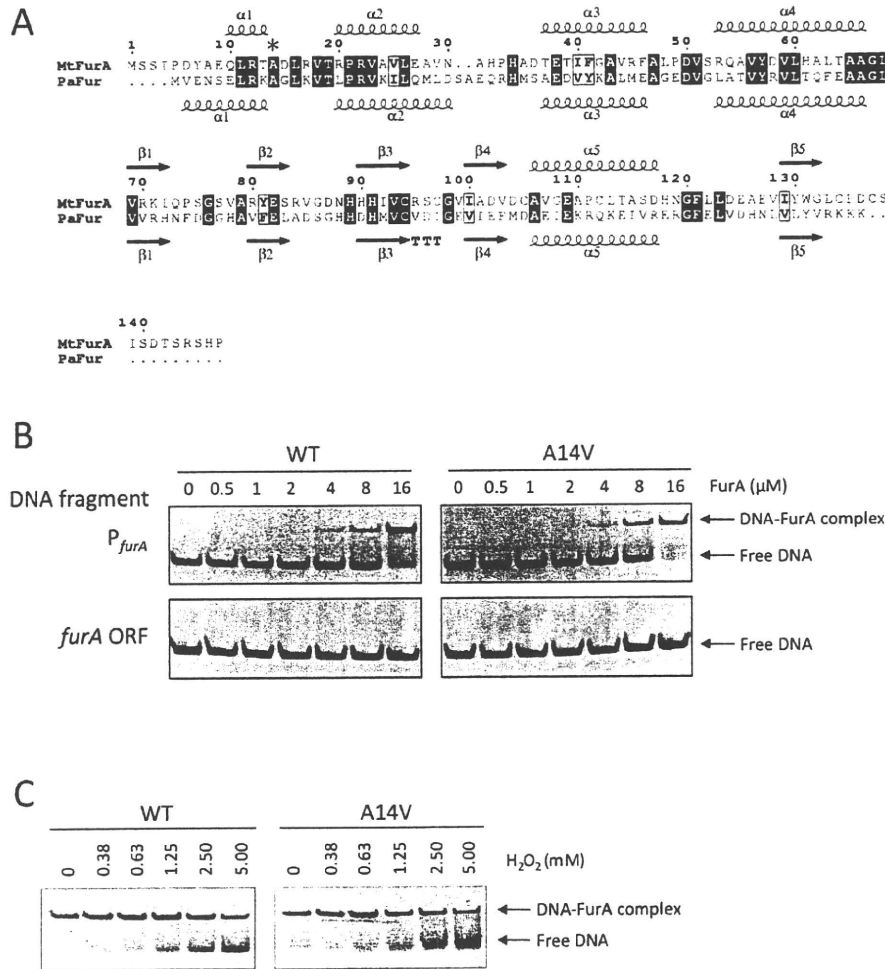
Fig. 7. Binding of FurA^{A14V} to the P$_{furA}$.
A. Sequence alignment of FurA (MtFurA) and a ferric uptake regulator of *P. aeruginosa* (PaFur). Secondary structures of MtFurA and PaFur are shown above and below the alignment respectively. Conserved residues are highlighted in black, and residues with highly conservative substitutions are boxed. The asterisk indicates the amino acid residue at the 14th position in MtFurA that was found to be mutated to valine in this study.
B. Effects of H$_2$O$_2$ on the binding of FurA^{A14V}. Purified FurA or FurA^{A14V} was mixed with DNA fragments and analysed by electrophoresis on 7.5% polyacrylamide gels, which were stained with ethidium bromide. This experiment was repeated three times with similar results.
C. Purified FurA or FurA^{A14V} was incubated with H$_2$O$_2$ at various concentrations for 10 min, mixed with DNA fragments and analysed by electrophoresis. This experiment was repeated three times with similar results.

2003). PaFur^{A10G} mutant fails to protect the *pvdS* promoter region (Barton *et al.*, 1996), suggesting that residue Ala-10, located at the C-terminal end of the helix H1, is critical for DNA recognition. As shown by the structure-based alignment (Fig. 7A), the A14V mutation in *M. tuberculosis* FurA (MtFurA) was located at the C-terminal end of helix H1. We therefore performed electrophoretic mobility shift assays (EMSA) to determine whether the FurA^{A14V} mutant affects DNA binding ability (Fig. 7B). EMSA showed that FurA^{A14V} as well as FurA bound to the upstream region of the *furA–katG* (P$_{furA}$) in a specific, protein dose-dependent manner.

Treatment of FurA with hydrogen peroxide (H$_2$O$_2$) has been reported to inhibit its binding to P$_{furA}$ (Sala *et al.*, 2003). Since H$_2$O$_2$ treatment may remove iron from the FurA and/or oxidize cysteine residues, we assessed whether H$_2$O$_2$ treatment of FurA^{A14V} affects DNA binding (Fig. 7C). Treatment of FurA and FurA^{A14V} with H$_2$O$_2$ inhibited the binding of P$_{furA}$ in a H$_2$O$_2$ dose-dependent manner. H$_2$O$_2$ had similar effects on binding of P$_{furA}$ to FurA and FurA^{A14V}.

Since the binding of FurA to P$_{furA}$ has been reported strong in the presence of Ni^{2+} ions (Sala *et al.*, 2003), we assessed whether its binding to FurA and FurA^{A14V} dif-

fered in the presence or absence of Ni^{2+}, Mn^{2+} or Zn^{2+}. We found that Ni^{2+}, Mn^{2+} and Zn^{2+} did not affect the binding pattern of P_{furA} to FurA and $FurA^{A14V}$ (Fig. S2). In the presence of Ni^{2+}, almost all the P_{furA} added to the binding buffer bound to FurA (Fig. S2). In the absence of Ni^{2+} or in the presence of Mn^{2+} or Zn^{2+}, however, the binding of FurA to P_{furA} was incomplete (Fig. S2). Similar results were observed when we assessed the binding of P_{furA} to $FurA^{A14V}$ (Fig. S2).

Discussion

To our knowledge, this is the first report showing that downregulation of *katG* expression causes INH resistance in clinical isolates of *M. tuberculosis*. These isolates had mutations in the *furA–katG* intergenic region. We constructed these mutations in isogenic strains and demonstrated that mutations in the *furA–katG* intergenic region (g-7a, a-10c and g-12a) are associated with INH resistance. The mechanism responsible for this downregulation involves a decrease in translational efficiency caused by base substitution in the *furA–katG* intergenic region. The polypurine sequence (GGAAGGAA), complementing the 3′ end of the 16S rRNA sequence, in the *furA–katG* intergenic region may act as a ribosome binding site in *M. tuberculosis* (Sala *et al.*, 2008). Int^{a-10c} had an A to C mutation at the third nucleotide, and Int^{g-7a} had a G to A mutation at the sixth nucleotide of the polypurine sequence (Fig. 4A). We found that these mutations in the polypurine sequence play a critical role in *katG* expression and showed that the mutations are responsible for INH resistance by decreasing *katG* expression (Figs 2–4). In the polypurine sequence, A at the third nucleotide and G at the sixth nucleotide are especially important for *katG* expression. Using reporter assays, we found that these two mutations significantly repressed *lacZ* expression to 48% and 38%, respectively, of that in the wild-type control. Similar *lacZ* repression was observed when the polypurine sequence was completely substituted with a polypyrimidine sequence (CCTC-CCTC), suggesting that the polypurine sequence is necessary for the full expression of *katG* (Sala *et al.*, 2008). The Int^{g-7a} and Int^{a-10c} mutations conferred resistance to INH (0.2 μg ml^{-1}) and the Int^{g-12a} mutation caused a slight decrease in INH susceptibility. We confirmed that the Δ(*furA–katG*) mutant, with no *katG* production and no INH oxidase activity, was highly resistant to INH. In contrast, the isogenic strains with Int^{g-7a} and Int^{a-10c} had low but detectable production and activity, with low-level resistance to INH. The levels of *katG*, and probably the levels of INH activated by *katG*, may therefore correlate with the levels of INH susceptibility.

It is unclear whether the *furA*^{c41t} mutation is associated with INH resistance. Of the 18 INH^r clinical isolates with

furA^{c41t} mutations, seven had no other mutations, suggesting an association between the mutation and INH resistance. A role of FurA in INH resistance was suggested, and two strains carrying $FurA^{L68P}$ or $FurA^{C97Y}$ mutations were reported previously (Pym *et al.*, 2001). These isolates had additional mutations in *katG*. Therefore, the involvement of these both FurA mutants in *katG* expression and INH resistance is unclear. We found that the results of drug susceptibility testing with complementary strains differed from those with the clinical isolates, and EMSA showed no differences in the binding patterns of FurA and $FurA^{A14V}$, suggesting that this mutation is not associated with INH resistance. One explanation for the discrepancy is that these clinical isolates may have an unknown additional mutation(s) associated with INH resistance. RFLP analysis showed that 67% of the INH^r isolates carrying *furA*^{c41t} belonged to cluster II, suggesting that these isolates expanded in a clonal manner. Therefore, these INH^r isolates carrying *furA*^{c41t} may have unknown mutation(s). Alternatively, *furA*^{c41t} may be associated with INH resistance, but the effect of the mutation may be observed in clinical isolates but not in complementary strains. Similar discrepancies between the drug MICs of clinical isolates and the complementary strains were reported in ethambutol-resistant clinical isolates and laboratory mutants harbouring an *embB* mutation (Safi *et al.*, 2010).

We found that FurA will not function as a negative regulator in *M. tuberculosis* under some culture conditions. For example, FurA did not effectively repress *katG* expression when *M. tuberculosis* isolates were cultured in MycoBroth (modified Middlebrook 7H9-ADC) for 2 weeks, since the *furA*^{Δ34c} mutant, which did not produce functional FurA, expressed KatG at the same level as INH^s isolates with *furA*. Deletion of *furA* has been shown to result in derepression of *katG* in *Mycobacterium smegmatis* (Zahrt *et al.*, 2001). In addition, complementation of the *M. tuberculosis* Δ(*furA–katG*) strain with *katG*, but not *furA* resulted in high levels of KatG production and catalase activity (Pym *et al.*, 2001). These discrepancies in the function of FurA may have been due to differences in host strains used or phases of culture examined. In addition, the *furA*^{Δ34c} mutant also had mutations in *katG* [g368a (G123E), g895a (G299S) and g1388t (R463L)] and in *inhA* (t-8c), suggesting that the Δ34c mutation may not be involved in INH resistance. *katG*^{g1388t} (R463L) is a neutral mutation and P_{inhA}^{t-8c} is responsible for INH resistance (Zhang and Telenti, 2000).

The mutations of *katG*, the regulatory region of *fabG1–inhA*, and *inhA* have been associated with INH resistance, although a significant population of INH^r clinical isolates in Japan showed no mutations in these regions (Ando *et al.*, 2010). The mutations Int^{g-7a}, Int^{a-10c}, Int^{g-12a} and *furA*^{c41t} characterized in the present

study, may yield important insight into INH resistance in *M. tuberculosis*.

Experimental procedures

Bacterial strains and plasmids

One hundred and eight INH[r] *M. tuberculosis* clinical isolates were obtained from single patients with INH[r] tuberculosis hospitalized at the National Center for Global Health and Medicine (formerly International Medical Center of Japan) and National Hospital Organization Tokyo National Hospital from 2003 to 2008. Fifty-one INH[s] isolates were chosen using a random number list from INH[s] isolates obtained from single patients hospitalized in the two hospitals from 2003 to 2008. The *M. tuberculosis* and *Escherichia coli* strains and plasmids used in this study are listed in Table S1. *M. tuberculosis* H37Rv, NCGM2828, NCGM2836 and NCGM2898 were used as controls for drug susceptibility testing, Western blotting analysis and INH oxidase assays. NCGM2836 was also used for β-galactosidase and complementation assays. *E. coli* DH5α and TOP10 (Invitrogen), and BL21(DE3) (Stratagene) were used for cloning and for protein overexpression studies respectively.

Growth conditions

Unless otherwise specified, *M. tuberculosis* clinical isolates and strains were pre-cultured in 125 ml Erlenmeyer flasks (Corning) containing 20 ml of MycoBroth [modified Middlebrook 7H9 broth (BD), Kyokuto] (500 mg of ammonium sulphate, 500 mg of L-glutamic acid, 100 mg of sodium citrate, 1 mg of pyridoxine, 0.5 mg of biotin, 2.5 g of disodium phosphate, 1 g of monopotassium phosphate, 50 mg of magnesium sulphate, 500 mg of calcium chloride, 1 mg of zinc sulphate, 850 mg of NaCl, 5 g of BSA, 2 g of dextrose, 3 mg of catalase and 500 mg of Tween 80 per litre, pH 6.6 ± 0.2) for 14 days at 37°C. The bacterial pre-cultures were inoculated into fresh MycoBroth or 7H10 agar (BD) supplemented with 10% OADC enrichment (BD) and 0.5% glycerol (Nacalai Tesque) and further cultured at 37°C. *E. coli* was grown in Luria–Bertani (LB) medium (BD). When required, kanamycin (KM) (Sigma) was added at 20 μg ml^{-1} to cultures of *M. tuberculosis* or 50 μg ml^{-1} to cultures of *E. coli*.

Drug susceptibility testing

All clinical isolates and *M. tuberculosis* strains were tested for drug susceptibility using the agar proportion method with egg-based Ogawa medium (Vit Spectrum-SR, Kyokuto; or Wellpack, Japan BCG Laboratory) according to the manufacturer's instructions, which were based on a slightly modified WHO protocol and have been recommended by the Japanese Society for Tuberculosis (Fujiki, 2001; WHO, 2003). The medium contained INH (0.2 μg ml^{-1} and 1.0 μg ml^{-1}), rifampin (RIF) (40 μg ml^{-1}), ethambutol (EB) (2.5 μg ml^{-1}), KM (20 μg ml^{-1}), *p*-aminosalicylic acid (PAS) (0.5 μg ml^{-1}), streptomycin (SM) (10 μg ml^{-1}), ethionamide (TH) (20 μg ml^{-1}), enviomycin (EVM) (20 μg ml^{-1}), cycloserine (CS) (30 μg ml^{-1}) and levofloxacin (LVFX) (1.0 μg ml^{-1}). The results are shown

in Table 1 and Fig. S1. Clinical isolates harbouring mutations in *furA* or the *furA–katG* intergenic region, including H37Rv, NCGM2828, NCGM2836, NCGM2898 and NCGM 2836-derived strains, were also tested for INH susceptibility using an agar proportion method with 7H10 agar plates according to the protocols of the Clinical and Laboratory Standards Institute (NCCLS, 2003) and a broth dilution method (BrothMIC MTB-I, Kyokuto). The results are shown in Tables 1 and 2.

Isolation of genomic DNA

Genomic DNA from bacteria was extracted as described previously (Otsuka *et al.*, 2004).

DNA sequencing of INH resistance-related genes

The DNA sequences of the oligonucleotide primers used in the present study are listed in Table S2. The *furA–katG*, *fabG1–inhA*, *ndh*, *ahpC* genes and their upstream regions, and *kasA–kasB*, were amplified by two-temperature PCR with the same conditions, i.e. reaction mixtures contained 0.5 U of *Z-Taq* polymerase (Takara), 5 μl of 10× *Z-Taq* buffer (Takara), 4 μl of 2.5 mM dNTP mixture (Takara), 0.5 μl of each primer at 25 mM (Invitrogen), 20 ng of genomic DNA and sterile distilled water to 50 μl. Thermal cycling was performed on a GeneAmp PCR system 9700 thermocycler (Applied Biosystems) with 30 cycles of 98°C for 1 s and 68°C for 30 s. PCR products were purified with Microcon YM-30 centrifugal filter devices (Millipore) and used as templates for direct DNA sequencing. DNA sequencing was performed with BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). DNA sequences were compared with *M. tuberculosis* H37Rv using Genetyx-Mac (Genetyx Corporation).

Restriction fragment length polymorphism (RFLP)

IS*6110*-probed RFLP was performed as described previously (Otsuka *et al.*, 2004). The fingerprinting patterns were analysed with Fingerprinting II software (Bio-Rad). Each dendrogram was calculated with the unweighted pair group method with average linkage according to the manufacturer's instructions. Patterns with more than 70% similarity were postulated to form a cluster.

Construction of plasmids

The *furA* gene and its variant *furA*^{c41t} were amplified by PCR with the BamHI-furA-F/furA-EcoRI-R primer set (Table S2) from *M. tuberculosis* H37Rv and a clinical isolate harbouring *furA*^{c41t} mutation. The PCR products were digested with BamHI and EcoRI, and ligated into the pGEX-2T vector (GE Healthcare) digested with the same restriction enzymes. Recombinant plasmids were used to transform *E. coli* DH5α and selected on LB plates containing 100 μg ml^{-1} ampicillin (AMP). We utilized pJEM13 (Timm *et al.*, 1994) carrying the fused P$_{furA}$-*furA-katG'–'lacZ* gene for β-galactosidase assays in *M. tuberculosis*. This plasmid contained a DNA fragment

from 138 bp upstream of the initiation codon of *furA* to the 99th base of *katG*. DNA fragments with or without a mutation in the *furA–katG* intergenic region were prepared by PCR using primer sets ApaI-P_{furA}-F/katG100-KpnI-R (Table S2). These fragments were digested with ApaI and KpnI and cloned into the corresponding site of pJEM13. Plasmids were used to transform *E. coli* TOP10 and selected on LB plates containing KM. For complementation assays of *M. tuberculosis* Δ(*furA–katG*), pMV306-derived plasmids were constructed. The entire *furA–katG* and 500 bp upstream region with or without a mutation was obtained by PCR with primer sets XbaI-(-500)furA-F/katG+40-HindIII-R (Table S2). These PCR products were digested with XbaI and HindIII and cloned into the corresponding site of pMV306. Plasmids were used to transform *E. coli* TOP10 and selected on LB plates containing KM. The DNA sequences of all clones were confirmed by sequencing.

Transformation of M. tuberculosis

INH^r *M. tuberculosis* NCGM2836 was grown in 20 ml of MycoBroth for 14 days as described above. Two-millilitre aliquots of 2 M glycine were added and the cultures were incubated for 24 h at 37°C. Bacteria were harvested by centrifugation at 3000 r.p.m. for 15 min at room temperature (RT), washed twice with 20 ml of 10% glycerol at RT and resuspended in 400 µl of 10% glycerol. Bacteria were electroporated with 1–4 mg of plasmids using a Bio-Rad Gene Pulser with settings of 2.5 kV, 25 mF and 1000 W. After electroporation, the bacteria were added to 4 ml of MycoBroth, incubated for 24 h at 37°C, harvested by centrifugation at 3000 r.p.m. for 15 min at RT and resuspended in 300 µl of MycoBroth. Transformants were selected by plating out on 7H10 agar plates with KM. To check the correct chromosomal structures of the complementary isogenic strains, KM-resistant colonies were examined by PCR with GC buffer, i.e. reaction mixtures contained 2.5 U of *LA Taq* polymerase (Takara), 25 µl of 2× GC buffer I (Takara), 8 µl of 2.5 mM dNTP mixture (Takara), 0.5 µl of each primer (glyV-20-F/lipP+20-R) at 25 mM (Invitrogen), 20 ng of genomic DNA and sterile distilled water to 50 µl. Thermal cycling was performed on a GeneAmp PCR system 9700 thermocycler with 30 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 10 min.

β-Galactosidase assay

The pJEM13-derived gene fusions were used to transform NCGM2836 and β-galactosidase activity was measured as described (Alland *et al.*, 2000). The transformed *M. tuberculosis* clones were streaked onto plates containing X-gal to monitor *lacZ* expression.

Preparation of total protein extracts

Bacteria were grown in 20 ml of MycoBroth for 14 days as described above, harvested by centrifugation at 3000 r.p.m. for 15 min at RT, washed twice with 50 mM phosphate buffer (pH 7.0) and resuspended in 500 µl of the same buffer. Bacteria were then lysed by shaking in a FastPrep FP100A homogenizer (Savant) (speed: 6.5 m s^{-1}; time: 20 s, twice)

with 70 mg of Lysing Matrix B (Qbiogene). The supernatant obtained after centrifugation at 12 000 r.p.m. for 1 min was added to Spin-X centrifuge tube filters with cellulose acetate membranes of pore size 0.22 µm (Costar) and centrifuged at 12 000 r.p.m. for 5 min at RT. Total protein extracts were quantified using Protein Assay CBB Solution (Nacalai Tesque).

Western blotting analysis

Proteins separated by SDS-5%–20%-gradient-PAGE were transferred onto Immun-Blot PVDF membranes (Bio-Rad). The membranes were incubated simultaneously with anti-KatG polyclonal antibody (diluted 1:10 000) (Sekiguchi *et al.*, 2007) and anti-HSP70 monoclonal antibody (diluted 1:500) (Santa Cruz Biotechnology), followed by incubation with horseradish peroxidase-conjugated donkey anti-rabbit IgG (diluted 1:10 000) (Amersham Biosciences) and goat anti-mouse IgG (diluted 1:1000) (Santa Cruz Biotechnology). Proteins were visualized using SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific). Quantity One (Bio-Rad) was used to quantify the KatG and HSP70 protein levels.

INH oxidase assay

Total protein extracts were prepared and quantified as above. The INH oxidase activity of KatG was assayed as described previously (Wei *et al.*, 2003; Sekiguchi *et al.*, 2007; Ando *et al.*, 2010). To 1 ml of 50 mM phosphate buffer (pH 7.0) were added 300 mg of total protein extracts, 0.04 mM nitroblue tetrazolium (NBT), 9 mM INH, 0.5 mg of glucose oxidase and 0.4 mM glucose. KatG activity was measured spectrophotometrically by monitoring the reduction of NBT at A_{560}. The absorbance was read 200 s after initiation of the reaction. All assays were performed at 25°C. NBT reduction in the absence of INH was subtracted from that in the presence of INH.

Purification of recombinant FurA and FurA^{A14V}

To purify FurA and FurA^{A14V}, *E. coli* BL21(DE3) carrying the plasmids pGEX-*furA* and pGEX-*furA*^{c41t} were grown in LB medium containing 200 µg ml^{-1} AMP at 37°C. Induction and purification of the GST-fused FurA and FurA^{A14V} were performed according to the manufacturer's instructions (GE Healthcare). Thrombin (GE Healthcare) was used to cleave the GST-Tag at the N-termini of the FurA proteins. The final concentration of protein was determined using a bicinchoninic acid protein assay kit (Pierce).

Electrophoretic mobility shift assay (EMSA)

Binding between purified *M. tuberculosis* FurA or FurA^{A14V} and P_{furA} was assessed by EMSA as described (Sala *et al.*, 2003). DNA fragments for EMSA were amplified by PCR using specific primers, -129furA and furA33-R for the upstream region of *furA*, and -5furA-F and furA166-R as a negative control. Both PCR products were 171 bp in length.

Binding reaction mixtures in 20 µl of binding buffer [20 mM Tris-HCl (pH 8.0), 1 mM DTT, 50 mM KCl, 5 mM MgCl$_2$, 10% glycerol, 50 µg of BSA per ml and 200 µM NiSO$_4$] containing 180 ng DNA fragment were incubated with purified FurA protein for 20 min at RT. Reaction mixtures were analysed by electrophoresis on 7.5% polyacrylamide gels in 40 mM Tris-acetate buffer at RT. DNA was visualized by ethidium bromide staining. For exposure of proteins to H$_2$O$_2$ (Santoku), 16 µM FurA protein was mixed with various concentrations of H$_2$O$_2$ in reaction mixtures without DNA fragments and incubated for 10 min at RT. DNA fragments were then added and incubation was continued for a further 10 min.

Structure-based sequence alignment

Amino acid sequences of *M. tuberculosis* FurA were aligned using CLUSTALW (Thompson *et al.*, 1994) and edited using ESPript (Gouet *et al.*, 1999). The secondary structure of FurA was determined by ESPript based on the crystal structure of *P. aeruginosa* Fur (PDB ID: 1MZB), a template predicted by automated modelling with the SWISS MODEL server.

References

Alland, D., Steyn, A.J., Weisbrod, T., Aldrich, K., and Jacobs, W.R., Jr (2000) Characterization of the *Mycobacterium tuberculosis iniBAC* promoter, a promoter that responds to cell wall biosynthesis inhibition. *J Bacteriol* **182:** 1802–1811.

Ando, H., Kondo, Y., Suetake, T., Toyota, E., Kato, S., Mori, T., and Kirikae, T. (2010) Identification of *katG* mutations associated with high-level isoniazid resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **54:** 1793–1799.

Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K.S., Wilson, T., *et al.* (1994) *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science* **263:** 227–230.

Bardou, F., Raynaud, C., Ramos, C., Laneelle, M.A., and Laneelle, G. (1998) Mechanism of isoniazid uptake in *Mycobacterium tuberculosis*. *Microbiology* **144:** 2539–2544.

Barton, H.A., Johnson, Z., Cox, C.D., Vasil, A.I., and Vasil, M.L. (1996) Ferric uptake regulator mutants of *Pseudomonas aeruginosa* with distinct alterations in the iron-

dependent repression of exotoxin A and siderophores in aerobic and microaerobic environments. *Mol Microbiol* **21:** 1001–1017.

Bernstein, J., Lott, W.A., Steinberg, B.A., and Yale, H.L. (1952) Chemotherapy of experimental tuberculosis. V. Isonicotinic acid hydrazide (nydrazid) and related compounds. *Am Rev Tuberc* **65:** 357–364.

Caminero, J.A. (2010) Multidrug-resistant tuberculosis: epidemiology, risk factors and case finding. *Int J Tuberc Lung Dis* **14:** 382–390.

Espinal, M.A., Laszlo, A., Simonsen, L., Boulahbal, F., Kim, S.J., Reniero, A., *et al.* (2001) Global trends in resistance to antituberculosis drugs. World Health Organization-International Union against Tuberculosis and Lung Disease Working Group on Anti-Tuberculosis Drug Resistance Surveillance. *N Engl J Med* **344:** 1294–1303.

Fujiki, A. (2001) *TB Bacteriology Examination to Stop TB*. Tokyo: The Research Institute of Tuberculosis.

Gouet, P., Courcelle, E., Stuart, D.I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15:** 305–308.

Kong, D., and Kunimoto, D.Y. (1995) Secretion of human interleukin 2 by recombinant *Mycobacterium bovis* BCG. *Infect Immun* **63:** 799–803.

Lee, A.S., Lim, I.H., Tang, L.L., Telenti, A., and Wong, S.Y. (1999) Contribution of *kasA* analysis to detection of isoniazid-resistant *Mycobacterium tuberculosis* in Singapore. *Antimicrob Agents Chemother* **43:** 2087–2089.

Lee, A.S., Teo, A.S., and Wong, S.Y. (2001) Novel mutations in *ndh* in isoniazid-resistant *Mycobacterium tuberculosis* isolates. *Antimicrob Agents Chemother* **45:** 2157–2159.

NCCLS (2003) *Susceptibility Testing of Mycobacteria, Nocardiae, and Other Aerobic Actinomycetes; Approved Standard*. NCCLS document M24-A, Wayne, PA: NCCLS.

Otsuka, Y., Parniewski, P., Zwolska, Z., Kai, M., Fujino, T., Kirikae, F., *et al.* (2004) Characterization of a trinucleotide repeat sequence (CGG)$_5$ and potential use in restriction fragment length polymorphism typing of *Mycobacterium tuberculosis*. *J Clin Microbiol* **42:** 3538–3548.

Pohl, E., Haller, J.C., Mijovilovich, A., Meyer-Klaucke, W., Garman, E., and Vasil, M.L. (2003) Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator. *Mol Microbiol* **47:** 903–915.

Pym, A.S., Domenech, P., Honore, N., Song, J., Deretic, V., and Cole, S.T. (2001) Regulation of catalase-peroxidase (KatG) expression, isoniazid sensitivity and virulence by *furA* of *Mycobacterium tuberculosis*. *Mol Microbiol* **40:** 879–889.

Safi, H., Fleischmann, R.D., Peterson, S.N., Jones, M.B., Jarrahi, B., and Alland, D. (2010) Allelic exchange and mutant selection demonstrate that common clinical *embCAB* gene mutations only modestly increase resistance to ethambutol in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **54:** 103–108.

Sala, C., Forti, F., Di Florio, E., Canneva, F., Milano, A., Riccardi, G., and Ghisotti, D. (2003) *Mycobacterium tuberculosis* FurA autoregulates its own expression. *J Bacteriol* **185:** 5357–5362.

Sala, C., Forti, F., Magnoni, F., and Ghisotti, D. (2008) The *katG* mRNA of *Mycobacterium tuberculosis* and *Mycobac-*

terium smegmatis is processed at its 5′ end and is stabilized by both a polypurine sequence and translation initiation. *BMC Mol Biol* **9**: 33.

Sekiguchi, J., Miyoshi-Akiyama, T., Augustynowicz-Kopec, E., Zwolska, Z., Kirikae, F., Toyota, E., *et al.* (2007) Detection of multidrug resistance in *Mycobacterium tuberculosis*. *J Clin Microbiol* **45**: 179–192.

Stover, C.K., Cruz, V.F., Fuerst, T.R., Burlein, J.E., Benson, L.A., Bennett, L.T., *et al.* (1991) New use of BCG for recombinant vaccines. *Nature* **351**: 456–460.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.

Timm, J., Lim, E.M., and Gicquel, B. (1994) *Escherichia coli*-mycobacteria shuttle vectors for operon and gene fusions to *lacZ*: the pJEM series. *J Bacteriol* **176**: 6749–6753.

Vilcheze, C., Weisbrod, T.R., Chen, B., Kremer, L., Hazbon, M.H., Wang, F., *et al.* (2005) Altered NADH/NAD⁺ ratio mediates coresistance to isoniazid and ethionamide in mycobacteria. *Antimicrob Agents Chemother* **49**: 708–720.

Vilcheze, C., Wang, F., Arai, M., Hazbon, M.H., Colangeli, R., Kremer, L., *et al.* (2006) Transfer of a point mutation in *Mycobacterium tuberculosis inhA* resolves the target of isoniazid. *Nat Med* **12**: 1027–1029.

Wei, C.J., Lei, B., Musser, J.M., and Tu, S.C. (2003) Isoniazid activation defects in recombinant *Mycobacterium tuberculosis* catalase-peroxidase (KatG) mutants evident in InhA inhibitor production. *Antimicrob Agents Chemother* **47**: 670–675.

WHO (2003) *Guidelines for Surveillance of Drug Resistance in Tuberculosis*. Geneva: WHO.

Zahrt, T.C., Song, J., Siple, J., and Deretic, V. (2001) Mycobacterial FurA is a negative regulator of catalase-peroxidase gene *katG*. *Mol Microbiol* **39**: 1174–1185.

Zhang, Y., and Telenti, A. (2000) Genetics of drug resistance in *Mycobacterium tuberculosis*. In *Molecular Genetics of Mycobacteria*. Hatful, G.F., and Jacobs, W.R., Jr (eds). Washington, DC: ASM Press, pp. 235–254.

Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. (1992) The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* **358**: 591–593.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

BMC
Genomics

Open Access

Complete genome sequencing and analysis of a Lancefield group G *Streptococcus dysgalactiae* subsp. *equisimilis* strain causing streptococcal toxic shock syndrome (STSS)

Yumi Shimomura[1], Kayo Okumura[1,4], Somay Yamagata Murayama[2], Junji Yagi[3], Kimiko Ubukata[2], Teruo Kirikae[1], Tohru Miyoshi-Akiyama[1*]

Abstract

Background: *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) causes invasive streptococcal infections, including streptococcal toxic shock syndrome (STSS), as does Lancefield group A *Streptococcus pyogenes* (GAS). We sequenced the entire genome of SDSE strain GGS_124 isolated from a patient with STSS.

Results: We found that GGS_124 consisted of a circular genome of 2,106,340 bp. Comparative analyses among bacterial genomes indicated that GGS_124 was most closely related to GAS. GGS_124 and GAS, but not other streptococci, shared a number of virulence factor genes, including genes encoding streptolysin O, NADase, and streptokinase A, distantly related to SIC (DRS), suggesting the importance of these factors in the development of invasive disease. GGS_124 contained 3 prophages, with one containing a virulence factor gene for streptodornase. All 3 prophages were significantly similar to GAS prophages that carry virulence factor genes, indicating that these prophages had transferred these genes between pathogens. SDSE was found to contain a gene encoding a superantigen, streptococcal exotoxin type G, but lacked several genes present in GAS that encode virulence factors, such as other superantigens, cysteine protease *speB*, and hyaluronan synthase operon *hasABC*. Similar to GGS_124, the SDSE strains contained larger numbers of clustered, regularly interspaced, short palindromic repeats (CRISPR) spacers than did GAS, suggesting that horizontal gene transfer via streptococcal phages between SDSE and GAS is somewhat restricted, although they share phage species.

Conclusion: Genome wide comparisons of SDSE with GAS indicate that SDSE is closely and quantitatively related to GAS. SDSE, however, lacks several virulence factors of GAS, including superantigens, SPE-B and the *hasABC* operon. CRISPR spacers may limit the horizontal transfer of phage encoded GAS virulence genes into SDSE. These findings may provide clues for dissecting the pathological roles of the virulence factors in SDSE and GAS that cause STSS.

Background

Since Lancefield group G streptococci (GGS) have been considered components of the normal flora of the human throat, skin, and genitourinary tract, the contributions of GGS to streptococcal disease have often been overlooked [1]. Over the last decade, however, infections by pathogenic GGS have been reported, including life-threatening invasive β-hemolytic streptococcal disease [1-7], making it necessary to expand our knowledge of the pathogenesis of GGS infections, especially invasive infections. Several species of streptococci can carry group C and G antigens, including *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE), *S. canis*, *S. dysgalactiae* subsp. *dysgalactiae*, *S. equi* subsp. *equi* (SESE), *S. equi* subsp. *zooepidemicus* (SESZ), and *S. anginosus* group bacteria. SDSE, which consists of Lancefield group G and C bacteria, in a ratio of about 4:1 [3,8,9], has been isolated from patients at higher frequency than

* Correspondence: takiyam@ri.ncgm.go.jp
[1] Department of Infectious Diseases, National Center for Global Health and Medicine, 1-21-1, Toyama, Shinjuku-ku, Tokyo 162-8655, Japan
Full list of author information is available at the end of the article

other GGS and GCS species. For example, of 313 strains of GCS and GGS isolated from clinical samples in Southern India between 2006 and 2007, 254 (81.1%) were SDSE [9], as were 80% of the 266 invasive non-A and non-B β-hemolytic streptococcal isolates in the USA [3]. The spectrum and clinical courses of SDSE infection, including pharyngitis, cellulitis, infective arthritis, vertebral osteomyelitis, and streptococcal toxic shock syndrome (STSS), show substantial overlap with those of GAS [10-16]. Despite the increased clinical importance of SDSE, however, the entire SDSE genome has not yet been sequenced. Knowledge of its entire genome sequence is fundamental to gain insights into the pathogenicity of SDSE. We describe here the entire genome sequence of a Lancefield group G SDSE strain, GGS_124, which had been isolated from a patient with STSS.

Results
Selection of an SDSE isolate for genome sequencing
We chose a clinical isolate of SDSE, strain GGS_124, for genome sequence determination for several reasons. First, GGS_124 belongs to Lancefield group G, to which most clinical isolates of SDSE also belong [3,8,9]. Second, GGS_124 caused STSS in a patient. Third, GGS_124 was the most virulent strain among 8 group G SDSE isolates, as determined by their LD_{50} values in a mouse infection model (Table 1).

Overview of the SDSE GGS_124 genome sequence
We found that, similar to other streptococcal genomes, the SDSE GGS_124 genome consists of a single circular chromosome of 2,106,340 bp (Additional file 1) and has a G+C content of 39.6% (Figure 1). Based on its location in the intergenic region upstream of the *dnaA* gene (SDEG_0001), the GC skew, and the clustering of dnaA box motifs, the start point of the SDSE GGS_124 genome was assigned to the putative origin of replication

Table 1 *emm* types and mouse LD_{50} values of 8 SDSE isolates used in this study

Strain	Origin	Symptom STSS/Non-STSS	LD_{50} value (CFU/head)	*emm* type
GGS_124	human	STSS	2.1×10^6	*stG480.0*
168	human		4.6×10^6	*stG480.0*
GGS_117	human	STSS	5.6×10^6	*stG4974.1*
170	human		5.6×10^6	*stC36.0*
164	human		1.9×10^7	*stG485*
GGS_118	human	STSS	2.0×10^7	*stG67920*
169	human		4.4×10^7	*stG11*
163	human		4.5×10^7	*stG643*

LD_{50} values of the isolates were determined as described in MATERIALS AND METHODS.

(*oriC*). An AT-rich 13-mer (AGTCTGTTTTTTT), located in the intergenic region upstream of the *dnaA* gene [17], was selected as the starting point for nucleotide numbering. The GGS_124 genome was shown to contain 2095 predicted coding sequences (CDS), which account for 1.83 Mbp (86.9%) of the genome. In addition, this genome was shown to harbor 3 prophage-like elements, designated ΦGGS_124.1, ΦGGS_124.2, and ΦGGS_124.3. Moreover, there were 27 insertion sequence (IS) elements throughout the genome.

Genome sequence homology analysis of GGS_124 with the other 11 sequenced streptococcal species and subspecies showed that GGS_124 was closest in sequence to GAS, with 72% similarity (Additional file 1). GGS_124 was less similar to SESZ and SESE, with 65% and 64% coverage. Although *S. agalactiae* is the closest-relative of SDSE based on 16S rRNA analysis, the *S. agalactiae* strains were less similar to SDSE than GAS based on the genome wide comparison (Additional file 1). In addition, we constructed a phylogenetic tree of all sequenced *Streptococcus* species based on the neighbor-joining method (Additional file 2). Although neighbor-joining methods are less accurate than the other methods such as most likelihood methods, SDSE is clustered with the GAS strains as their closest relative.
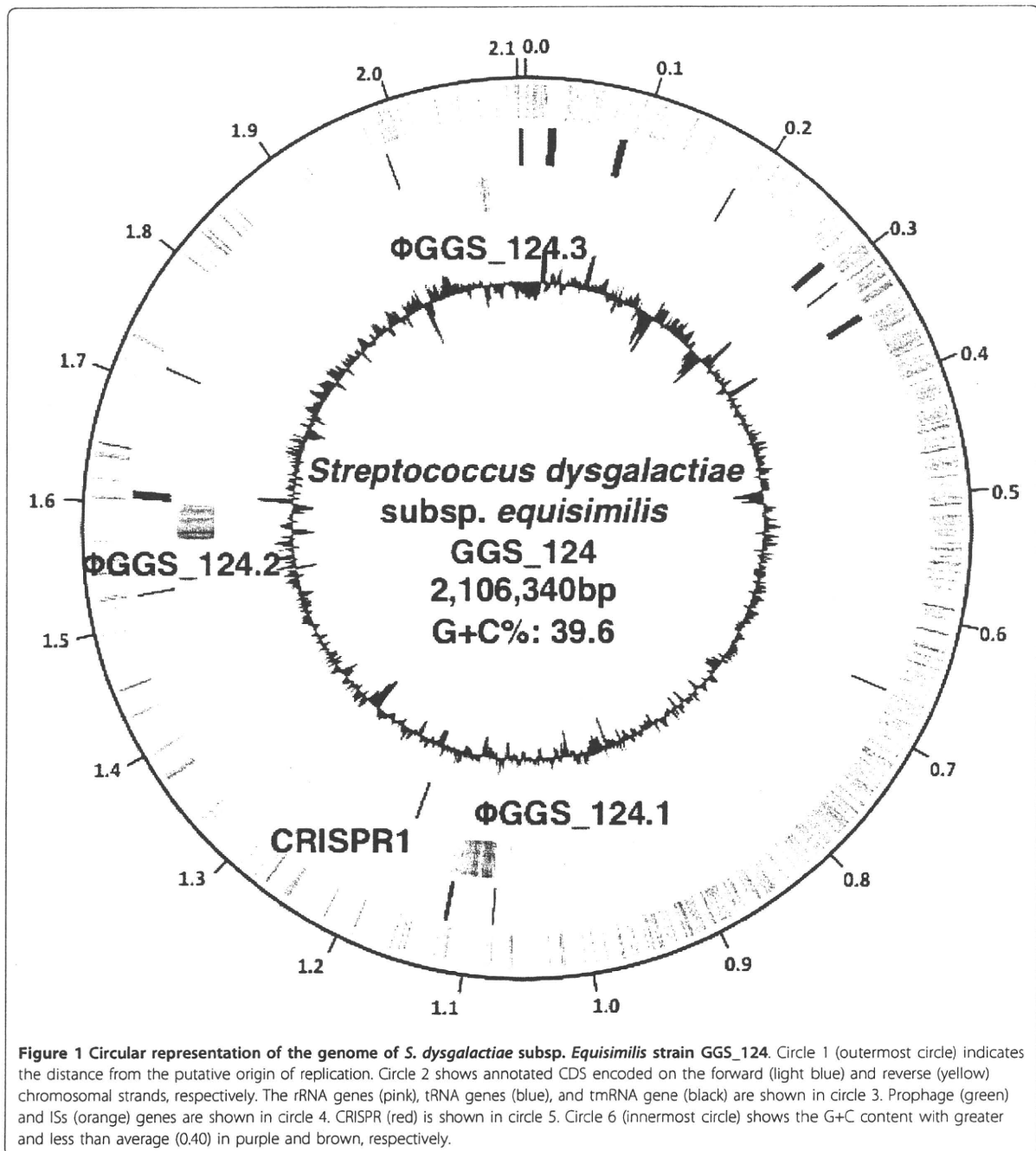
We compared the gene organization of GGS_124 with that of GAS by genomic rearrangement analyses (Figure 2 and Additional file 3). GAS could be divided roughly into 2 groups according to the orientation of the genes [18,19]. Therefore, SSI-1 and MGAS315, both of which are M3 serotype strains and have opposite gene orientations from each other, were chosen for the analysis. We found that the GGS_124 genome was organizationally more similar to that of GAS strain MGAS315 than GAS strain SSI-1 (Figure 2). Interestingly, the colinearity of GGS_124 and *S. uberis* genomes was quite remarkable but the percent amino acid identity was lower than that of the GAS strains (Additional file 3). The gene structure of GGS_124 was similar to the structures of GAS strain SSI-1, SESZ strain MGCS10565, and SESE strain 4047, although the GGS_124 genome contains large-scale genomic rearrangements. The GGS_124 genome differed markedly in gene organization from the genome of GBS strain A909.

When we compared genes from GGS_124 and two relatively homologous species, GAS (MGAS315) and SESZ (MGCS10565) (Figure 3), we found that these three streptococcal genomes contain more than 1,200 orthologous genes, accounting for 59% of the total CDSs of GGS_124. GGS_124 shares 282 genes with MGAS315 and 153 genes with MGCS10565. Moreover, 71.6% of the genes of GGS_124 were homologous to GAS genes, with 88.5% amino acid identity, whereas 66.5% of GGS_124 genes were homologous to MGCS10565

Figure 1 Circular representation of the genome of *S. dysgalactiae* subsp. *Equisimilis* strain GGS_124. Circle 1 (outermost circle) indicates the distance from the putative origin of replication. Circle 2 shows annotated CDS encoded on the forward (light blue) and reverse (yellow) chromosomal strands, respectively. The rRNA genes (pink), tRNA genes (blue), and tmRNA gene (black) are shown in circle 3. Prophage (green) and ISs (orange) genes are shown in circle 4. CRISPR (red) is shown in circle 5. Circle 6 (innermost circle) shows the G+C content with greater and less than average (0.40) in purple and brown, respectively.

genes, with 79.9% amino acid identity. These findings indicate that SDSE is closely related to GAS in both nucleotide and amino acid sequences.

We also analyzed the distribution of genes shown to be more homologous to genes derived from bacteria other than GAS (Additional file 4). We found that 299 genes showed higher similarity to genes from *Streptococci* other than GAS and 92 genes showed higher similarity to genes from a genus other than *Streptococcus*. In addition, we identified 11 genes that did not show significant homology to any genes in the databases. These genes were scattered throughout the entire GGS_124 genome, suggesting that they had not been acquired by massive genome recombination.

Figure 2 Genome rearrangement maps of *S. dysgalactiae* subsp. *equisimilis* GGS_124 with five species in the pyogenic group. Sequences were aligned from the predicted replication origin of each genome. The colored bars separating each genome (red and green) represent similarity matches identified by *in silico* Molecular Cloning. Links shown in green match in the same orientation, while those in red match in the reverse orientation. Prophages are highlighted as pale blue boxes.

Putative prophages and CRISPR/Cas

We found that all three prophage-like elements of GGS_124 were homologous to previously sequenced GAS prophages, and that they were integrated at sites similar to those of GAS strains, with the same upstream and downstream genes (Figure 4).

(i) Prophage GGS_124.1

We found that the ΦGGS_124.1 prophage is 35.593 bp in length with a G+C content of 38.04% and carries 60 CDS. Ninety-seven percent of the CDS in ΦGGS_124.1 have homologues, with more than 40% identity to GAS prophages, suggesting that ΦGGS_124.1 is a chimeric