J. Microbiol. Biotechnol. (2014), 24(9), 1170–1177
http://dx.doi.org/10.4014/jmb.1311.11101

jmb

Toxic Pyrene Metabolism in *Mycobacterium gilvum* PYR-GCK Results in the Expression of Mammalian Cell Entry Genes as Revealed by Transcriptomics Study

Abimbola Comfort Badejo[1], Won Hyong Chung[2], Nam Shin Kim[2], Se Kye Kim[1], Jin Choul Chai[1], Young Seek Lee[1], Kyoung Hwa Jung[1], Hyo Joon Kim[1*], and Young Gyu Chai[1,3*]

[1]*Department of Molecular and Life Science, Hanyang University, Ansan 426-791, Republic of Korea*
[2]*Korean Bioinformation Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, Republic of Korea*
[3]*Department of Nanobiotechnology, Hanyang University, Seoul 133-791, Republic of Korea*

Received: December 2, 2013
Revised: May 14, 2014
Accepted: June 9, 2014

First published online
June 9, 2014

*Corresponding authors
Y.G.C.
Phone: +82-31-400-5513;
Fax: +82-31-436-8173;
E-mail: ygchai@hanyang.ac.kr
H.J.K.
Phone: +82-31-400-5515;
E-mail:kimhj104@hanyang.ac.kr

Supplementary data for this paper are available on-line only at http://jmb.or.kr.

pISSN 1017-7825, eISSN 1738-8872

Mycobacterium gilvum PYR-GCK is a bacterial strain under study for its bioremediation use on heavy hydrocarbon pollutants in the environment. During the course of our study, mammalian cell entry (*mce*) genes, known to facilitate pathogenicity in *M. tuberculosis*, were highly expressed during a comparative and substrate-related cultural global transcriptomic study. RNA sequencing of the global transcriptome of the test strain in two different substrates, pyrene and glucose, showed high expression of the *mce* genes based on the differential results. After validating the expression of these genes with quantitative real-time PCR, we arrived at the conclusion that the genes were expressed based on the pyrene substrate (a phytosterol compound), and sterol metabolism is said to activate the expression of the *mce* genes in some actinomycetes bacteria, *M. gilvum* PYR-GCK in this case. This study is believed to be important based on the fact that some mycobacterial strains are undergoing a continuous research as a result of their use in practical bioremediation of anthropogenic exposure of toxic organic wastes in the environment.

Keywords: Mycobacteria, mammalian cell entry (*mce*) genes, pyrene

Introduction

The actinomycetes group of bacteria is known for diverse catabolic versatilities, and these characteristics have been observed in their annotated genome sequences. Their robust range of metabolic abilities has made them very important in numerous technological fields such as the environmental, pharmaceuticals, and chemical and energy production. The mycobacteria are of special interest owing to their unique cell wall characteristic of complex waxes and glycolipids. *Mycobacterium gilvum* PYR-GCK has been under intense study owing to its ability to degrade recalcitrant high molecular weight polycyclic aromatic hydrocarbons (PAHs) in the process of crude oil bioremediation. This ability is confirmed possible as a result of its cell walls and superb armamentarium of metabolic enzymes, aiding it in exceptional survival skills (such as the ability to resist unpleasant environmental conditions: extreme salinity and pH, dehydration, and exposure to toxic substances). These abilities have also been observed in its pathogenic relatives, such as *Mycobacterium tuberculosis* and *Mycobacterium leprae*. In the pathogenic mycobacterial strains, one of their major goals is to gain entrance into their respective hosts and resist the plethora of hostile compounds therein [8]. The mammalian cell entry (*mce*) genes have been characterized as important in this invasive purpose [13]. Annotated bacterial genome data have shown *mce* genes present in a variety of nonpathogenic actinomycetes bacteria, including mycobacteria [6]. The *M. tuberculosis* genome contains four *mce* operons (*mce1* to *mce4*) [13]; their structure and

evolution have been described by Tekaia *et al*. [26]. The pioneering *mce* genes study revealed *mce1A*, in the first operon, as an essential virulence property in *M. tuberculosis*, based on the conferment of this property to *E. coli*, and a subsequent invasion into macrophages [1, 8]. With no information on the exact function of these *mce* genes in the pathogenic and nonpathogenic mycobacterial strains, studies have suggested that the presence of these putative virulence genes may not be an indicator for pathogenicity [6, 10, 13]. Further studies have shown the expression of the *mce* genes during cholesterol and steroid metabolism in the Rhodococcal species [13, 28], with the *mce4* operon encoding a steroid transporter gene. Nevertheless, mutants of this uptake system were unaffected for growth on steroid compounds having shorter polar side-chains [16].

Polycyclic aromatic hydrocarbons (PAHs) are organic compounds released into the environment mostly as anthropogenic particles/compounds [18], and various effects of these compounds on human and wildlife health have been documented widely [4, 7, 23]. Studies on *Mycobacterium gilvum* PYR-GCK and other mycobacterial strains have been undertaken as biological removal agents of PAHs from the environment [3, 25], and pyrene has been used often as one of the PAH test substrates in many studies based on its simple form as a high molecular weighted PAH compound. With the aim of finding a relationship between the *mce* genes and toxic pyrene metabolism, and also to characterize the genes using a gene expression technique, *M. gilvum* PYR-GCK was subjected to a global transcriptomic analysis. A previous proteomic study on pyrene metabolism mechanics in the same strain and conditions revealed numerous expressed proteins, but we believe that a "supplementing" transcriptomics study will reveal the expression of more gene products possibly omitted during the tricky protein analysis and quantification.

Materials and Methods

Reagents, Media, and Bacterial Strain Cultivation

Mycobacterium gilvum PYR-GCK was acquired from the American Type Culture Collection (ATCC) under the code name *M. flavescens* ATCC 700033 and maintained in Bacto Brain Heart Infusion Media (BD Laboratories, Sparks, USA) at 29ºC or stock preserved in the same medium supplemented with 28% glycerol, at –80ºC. Pyrene substrate (confirmed >98.0% pure by Aldrich Company) and other chemicals used were purchased from Sigma-Aldrich Company (St. Louis, MO, USA) and Tokyo Chemical Industry (Tokyo, Japan). *M. gilvum* PYR-GCK cells were prepared and grown in 500 ml flasks containing 200 ml of basal medium, as described in Badejo *et al.* [3].

RNA Preparation, Processing, and Sequencing

Bacterial cells were harvested from all six bacterial cultures by centrifugation at 10,000 ×*g* at 4ºC for 1 min, after the addition of RNAprotect Bacteria reagent (Qiagen, USA) to the culture broth in the ratio of 2:1. RNAiso (Takara, Japan) lysing solution was added to the cells, along with 10 µl of β-mercaptoethanol and 0.6 g of 0.1 mm zirconia/silica beads (Biospec, USA). The mix was run in mini Bead-beater (Biospec, USA) for 45 sec and immediately placed on ice. Chloroform (200 µl) was added to the solution and the lysing mix was inverted gently to mix for 5 min. The mix was centrifuged at 12,000 ×*g* for 15 min at 4ºC and the clear top solution was gently collected into a new tube. Then 500 µl of isopropanol was added and the tubes were gently inverted to mix once again before it was finally incubated on ice for 1 h. After incubation, the lysed mix was centrifuged at 12,000 ×*g* for 10 min at 4ºC and the isopropanol was discarded. Ice-cold 70% ethanol was added to the RNA pellet for gentle washing. After another round of centrifuging at the same speed for 10 min, the ethanol was carefully removed. RNA pellets were dried at room temperature for 3–5 min before reconstitution in 20 µl of RNase-free water. Reconstituted RNA was treated with RNase-free DNase (Promega, USA) to remove residual DNA and then purified using an RNeasy MinElute Cleanup Kit (Qiagen, USA). The quantity and quality of the purified RNA were determined using the Agilent 2100 Bioanalyzer (Agilent Technologies, USA) with an RNA Integrity number (RIN) value ≥8 [21] and were afterwards aliquoted for separate analysis. A set of total RNA sample triplicates, with common induction, were combined (to avoid biased results), purified, and sequenced at Macrogen Korea using the Illumina Genome Analyzer. Prior to sequencing, total RNA was processed using a Ribo-Zero rRNA removal kit (Epicentre Biotechnologies, USA) to remove the 23S and 16S ribosomal RNAs. Enrichment was assessed using the Agilent 2100 Bioanalyzer and RNA 6000 chip. The remaining mRNA was fragmented into small pieces using divalent cations under elevated temperature and converted into a library of template molecules suitable for subsequent cluster generation using the reagents provided in the Illumina TruSeq RNA Sample Preparation Kit (Illumina Inc., USA), following the manufacturer's instructions. The cleaved RNA fragments were copied into first-strand cDNA using reverse transcriptase and random primers. Second-strand cDNA was then synthesized using DNA polymerase I and RNase H. The cDNA fragments were then processed through an end repair process by the addition of a single "A" base, and then ligation of the adapters. The products were then purified and enriched with PCR to create the final cDNA library. The cDNA fragments were sequenced with HiSeq2000 (Illumina Inc., USA).

Alignment and Filtering of Sequence Reads

The alignment and sequenced fragment read bioinformatics processing was done at the Korean Bioinformation Center (KOBIC), Daejeon. A total of 40,012,820 and 21,440,720 fragments were generated from the glucose-induced and pyrene-induced RNA

samples, respectively. Differential gene and transcript expression analyses of the sequence reads were made with TopHat and Cufflinks. Data were normalized by calculating the "fragments per kilobase per million map reads" (FPKM) for each gene [17]. Briefly, data were treated with low-quality filtering and reads removal before CuffDiff analysis. Exact duplicated reads were removed using Prinseq ver. 0.19.3, with average quality >=Q20. Filtered reads were mapped with reference genomes (GenBank CP000656, CP000657, CP000658, and CP000659) using Bowtie2 ver. 2.0.0, default option. The expression levels of chromosome and plasmids were analyzed separately. Afterwards, differential expression analysis was performed using Cuffdiff (default option) of the Cufflinks 2.0.2 package [27]. Results were manually curated to identify pyrene degradation related transcripts by comparing it with the glucose-induced transcripts. The data acquired have been deposited in the Gene Expression Omnibus database (GEO: GSE44536).

Gene Expression Analysis by Quantitative Real-Time PCR

Total RNA was isolated and purified as described earlier and the second set (10 μg), from the same template as the first set, was reverse transcribed separately, making six samples in all. The reverse transcription step was carried out using random hexamer primers and the PrimeScript 1st strand cDNA synthesis Kit (Takara, Japan), according to manufacturer's instructions. Briefly, random hexamers and RNA templates were mixed and denatured at 65°C for 5 min followed by cooling for 2 min on ice. Then 5× Primescript buffers, RTase, and RNAse inhibitor were added to the cooled template mix and incubated for 1 h at 50°C before enzyme inactivation at 70°C for 15 min. Negative controls for reverse transcription were performed to test for the presence of genomic DNA contamination in the RNA samples. Complementary DNA samples were diluted 1.5-fold and relative quantification real-time PCR was carried out in a standard fashion using SYBR Premix Ex-Taq II (Takara Bio, Japan) according to the manufacturer's directions. The AB-7500 Real-Time PCR System (Applied Biosystems Inc., USA) was employed for the real-time PCR run. The primer3 software [19] and NCBI primer designing tool were used to design primers that would amplify a product of approximately 200 base-pairs. Amplicon expected sizes and the absence of nonspecific products were confirmed by analysis of the PCR products in 2% agarose gels in TAE buffer, stained with ethidium bromide and visualized under UV light. PCRs were set up according to the manufacturer's instructions, and three technical replicates for each sample were included. The PCR volume was 20 μl, consisting of 0.4 μM of each primer (Table 1). Each PCR run included a no-template control with water instead of cDNA, as well as a RT negative control for each gene. The amplification conditions were 95°C for 15 sec; and 40 cycles of 95°C for 15 sec and 60°C for 1 min. The specificity of the reaction was confirmed by obtaining a melting curve from 55–95°C. The efficiency values of the reactions were automatically calculated by the PCR machine. To analyze the differential expressions, the obtained mRNA levels were compared in the pyrene-induced genes with respect to the glucose-induced genes as a control measure. In all cases, the expression of all genes were quantitated after normalization of their RNA levels to the expression of the *rpoB* gene coding for the β-subunit of bacterial RNA polymerase (Mflv_5097), as previously described [2]. The fold change was calculated using the relative quantification ($2\text{-}\Delta\Delta^{CT}$) method [14]. Statistical analysis was performed using the SPSS ver. 21.0 software package for Windows (SPSS Inc., USA), applying the Student's *t*-test. When the differential *p*-value was less than 0.05, it was regarded as statistically significant.

Results and Discussion

Mammalian cell entry proteins are found in diverse Actinobacteria (*Mycobacterium*, *Nocardia*, *Rhodococcus*, and *Streptomyces*), as components of ATP-binding cassette transporter systems, based on bioinformatics data systems [6, 15]; and as an essential virulence property in *M. tuberculosis*, based on gene cloning experiments [1, 8]. The analysis of the complete sequence of the *Mycobacterium tuberculosis* H37Rv genome revealed the presence of four paralogous

Table 1. Sequences of primers used in the analysis of the expression profile of *mce* genes expressed during the substrates metabolism study.

Gene tag	Forward primer (5′–3′)	Reverse primer (5′–3′)
Mflv_0009	GAATACGGCACCTTCTCCCC	GGAGCCCTGATGAAGTACCG
Mflv_0653	CGAAGGAGAAGAACACGGCT	TCGCCTACCTAGCCTCAACT
Mflv_0709	GGACACCCATGATCTGGACC	ACGTCCGAAACATGAAGGGG
Mflv_1556	TACTACGGCTACAACGCCTC	TGTTCCTTCCCTTACCCGGA
Mflv_2104	AAACTGTCCCAGGCGTTCTC	TTCTGGTTGAGCTGACGCAT
Mflv_4164	GATCCCGGCTATGAGACGTT	ACCGATGCTATGGCGAAGAG
Mflv_4536	GTCGGGTTTGAGTACCAGCA	AACGAGGTGAAGGTGTCAGG
Mflv_2509	CCGTCATCCAGGGTCTGAAC	TCTTGGCGAACAGTGAGTCC
Mflv_5097	TCTCGTGCTCTTCGATGTGG	GCATCAACACTGGGAGGGTG

mce genes, all encoded in an operon structure, consisting of eight genes (*yrbEA*, *yrbEB*, *mceA*, *mceB*, *mceC*, *mceD*, *mceE*, and *mceF*). We identified seven complete and one incomplete operons from the sequenced RNA bioinformatics analysis, based on the annotated genome sequence of our test strain (GenBank Accession No. CP000656, CP000657, CP000658, and CP000659) (Table 2, Fig. 1). With unclear information on the functions of these genes other than the reported mycobacterial mammalian cell entry facilitation, rhodococcal cholesterol metabolism, and actinomycetal envelope biosynthesis process/structure maintenance, in this work, we evaluated the gene expression pattern of *M. gilvum* PYR-GCK during toxic pyrene metabolism as a molecular study on PAH bioremediation. Our results have suggested the possible involvement of the *mce* operons or genes in the toxic hydrocarbon metabolism. With the different gene expression activities in the various operons, we had cause to believe the various operons had different functions in pyrene metabolism. No gene expression activity was observed in one operon, and minimal activity in some operons, ranging from the expression of one gene to total operon activity in others (Fig. 1). The expression activities of some of these genes were confirmed in a subsequent qRT-PCR analysis (Fig. 2). The eight genes, chosen from the identified gene clusters, were based on their expression activity and orientation in the various operons. Two genes (Mflv_4536 and Mflv_0709) were of similar orientation but different expression activities in the pyrene-induced cells. Likewise, Mflv_2509 and Mflv_0009 were of different orientation and different expression activities (Figs. 1 and 2). Three genes were negatively expressed, whereas five were expressed positively to pyrene induction. Information on the quantity of *mce* operons in various mycobacterial strains differ between strains, with four and six operons documented in *M. tuberculosis* and *M. smegmatis*, respectively [11]. Seven complete *mce* operons have been identified in our test strain, with yet another cluster of *mce* genes highly expressed in response to pyrene, but these failed the "complete operon" cut-off as a result of the absence of the first two integral membrane protein genes of the operon (*yrbEA*, *yrbEB*). These operons with homologous set of genes have been reported to have varying effects, either increasing or decreasing virulence [9, 20], but their specific roles are still unclear and under further study. Owing to the varying homology of these operons in the different mycobacterial species, it could not be decided or confirmed

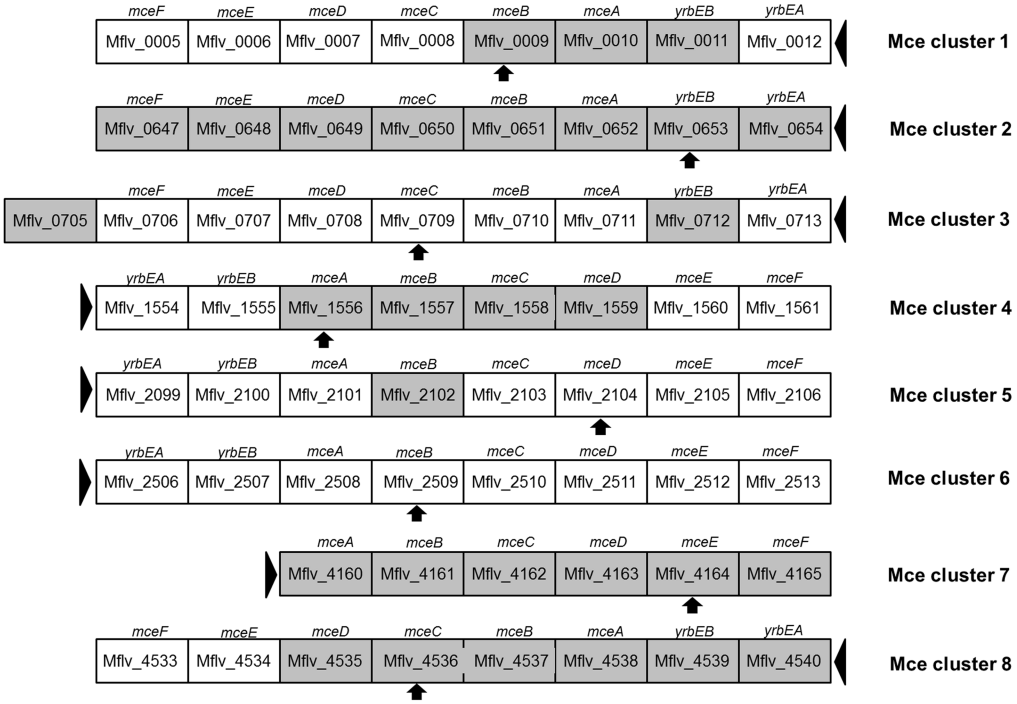

Fig. 1. The *mce* gene clusters of *Mycobacterium gilvum* PYR-GCK.

All positively expressed genes as a result of pyrene induction/metabolism are in the shaded gray boxes. Averagely expressed genes are situated in the unshaded boxes. The corresponding gene codes are located at the top of the respective boxes with Entrez gene IDs. Arrowed boxes indicate transcribed genes detected in the qRT-PCR validation assay.

Table 2. List of genes present in the *mce* gene clusters in *Mycobacterium gilvum* PYR-GCK.

Gene ID	Identification	Code	Pyrene/glucose expression ratio
Mflv_0005	Virulence factor Mce family protein	*mceF*	1.88
Mflv_0006	Mammalian cell-entry-related domain protein	*mceE*	1.86
Mflv_0007	Virulence factor Mce family protein	*mceD*	1.81
Mflv_0008	Virulence factor Mce family protein	*mceC*	1.93
Mflv_0009	Virulence factor Mce family protein	*mceB*	**2.38**
Mflv_0010	Mammalian cell-entry-related domain protein	*mceA*	**2.32**
Mflv_0011	Protein of unknown function DUF140	*yrbEB*	**2.22**
Mflv_0012	Protein of unknown function DUF140	*yrbEA*	1.09
Mflv_0647	Virulence factor Mce family protein	*mceF*	**4.53**
Mflv_0648	Virulence factor Mce family protein	*mceE*	**6.00**
Mflv_0649	Virulence factor Mce family protein	*mceD*	**3.49**
Mflv_0650	Virulence factor Mce family protein	*mceC*	**2.50**
Mflv_0651	Virulence factor Mce family protein	*mceB*	**3.51**
Mflv_0652	Virulence factor Mce family protein	*mceA*	**4.85**
Mflv_0653	Protein of unknown function DUF140	*yrbEB*	**14.01**
Mflv_0654	Protein of unknown function DUF140	*yrbEA*	**6.33**
Mflv_0705	Putative conserved MCE-associated membrane protein		**2.67**
Mflv_0706	Virulence factor Mce family protein	*mceF*	1.01
Mflv_0707	Virulence factor Mce family protein	*mceE*	0.76
Mflv_0708	Virulence factor Mce family protein	*mceD*	1.19
Mflv_0709	Virulence factor Mce family protein	*mceC*	1.56
Mflv_0710	Virulence factor Mce family protein	*mceB*	1.34
Mflv_0711	Virulence factor Mce family protein	*mceA*	1.01
Mflv_0712	Protein of unknown function DUF140	*yrbEB*	**3.55**
Mflv_0713	Protein of unknown function DUF140	*yrbEA*	1.20
Mflv_1554	Protein of unknown function DUF140	*yrbEA*	0.94
Mflv_1555	Protein of unknown function DUF140	*yrbEB*	1.78
Mflv_1556	Virulence factor Mce family protein	*mceA*	**3.01**
Mflv_1557	Virulence factor Mce family protein	*mceB*	**2.60**
Mflv_1558	Virulence factor Mce family protein	*mceC*	**3.02**
Mflv_1559	Virulence factor Mce family protein	*mceD*	**2.80**
Mflv_1560	Virulence factor Mce family protein	*mceE*	1.89
Mflv_1561	Virulence factor Mce family protein	*mceF*	1.35
Mflv_2099	Protein of unknown function DUF140	*yrbEA*	1.66
Mflv_2100	Protein of unknown function DUF140	*yrbEB*	1.24
Mflv_2101	Virulence factor Mce family protein	*mceA*	1.22
Mflv_2102	Virulence factor Mce family protein	*mceB*	**2.04**
Mflv_2103	Virulence factor Mce family protein	*mceC*	1.28
Mflv_2104	Virulence factor Mce family protein	*mceD*	1.04
Mflv_2105	Virulence factor Mce family protein	*mceE*	0.99
Mflv_2106	Virulence factor Mce family protein	*mceF*	0.81
Mflv_2506	Protein of unknown function DUF140	*yrbEA*	1.90
Mflv_2507	Protein of unknown function DUF140	*yrbEB*	1.24

Table 2. Continued.

Gene ID	Identification	Code	Pyrene/glucose expression ratio
Mflv_2508	Virulence factor Mce family protein	*mceA*	1.41
Mflv_2509	Virulence factor Mce family protein	*mceB*	1.50
Mflv_2510	Virulence factor Mce family protein	*mceC*	1.14
Mflv_2511	Virulence factor Mce family protein	*mceD*	1.64
Mflv_2512	Virulence factor Mce family protein	*mceE*	1.22
Mflv_2513	Virulence factor Mce family protein	*mceF*	1.18
Mflv_4160	Virulence factor Mce family protein	*mceA*	**4.22**
Mflv_4161	Virulence factor Mce family protein	*mceB*	**4.55**
Mflv_4162	Virulence factor Mce family protein	*mceC*	**3.26**
Mflv_4163	Virulence factor Mce family protein	*mceD*	**4.47**
Mflv_4164	Virulence factor Mce family protein	*mceE*	**3.85**
Mflv_4165	Virulence factor Mce family protein	*mceF*	**2.23**
Mflv_4533	Virulence factor Mce family protein	*mceF*	0.78
Mflv_4534	Virulence factor Mce family protein	*mceE*	1.98
Mflv_4535	Virulence factor Mce family protein	*mceD*	**3.75**
Mflv_4536	Virulence factor Mce family protein	*mceC*	**3.55**
Mflv_4537	Virulence factor Mce family protein	*mceB*	**2.99**
Mflv_4538	Virulence factor Mce family protein	*mceA*	**2.09**
Mflv_4539	Protein of unknown function DUF140	*yrbEB*	**4.17**
Mflv_4540	Protein of unknown function DUF140	*yrbEA*	**3.78**

The genes were identified based on GenBank genome annotation, and the gene IDs are of Entrez origin. The differential expression values (pyrene/glucose expression ratio) were calculated based on the NGS RNA-seq data results curated using CuffDiff from the Cufflinks 2.0.2 software package. Highly expressed fold change values are emboldened. Differential cut-off values were <0.5 and ≥2 for lowly expressed and highly expressed gene transcripts, respectively.

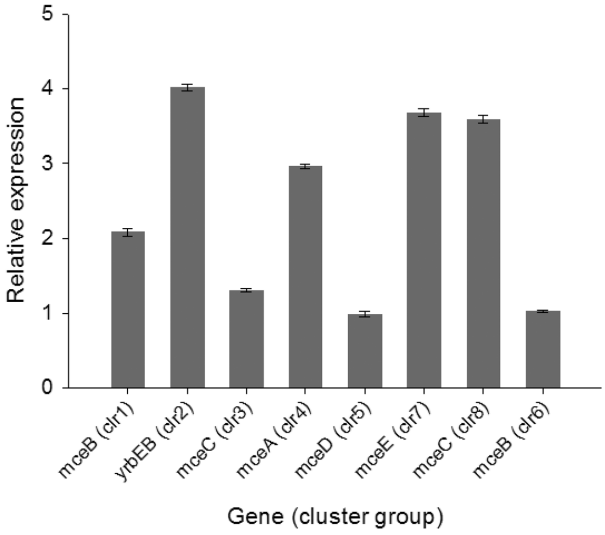

Fig. 2. Transcriptional profile of eight *mce* genes in *Mycobacterium gilvum* PYR-GCK.

The relative expressions of the gene transcripts were based on pyrene-induced cells, using the glucose-induced cells as control ($n \geq 3$ independent bacterial cultures; bars indicate standard deviation).

about the specific function of each operon, as done with the *mce1* and *mce4* operons [16, 22]. Nevertheless, we observed pyrene-induced expression in some genes, irrespective of their gene cluster orientations (Fig. 1). The pathogenic strains equipped with the *mce* operons have been reported to function under the control of the transcription factor *mce1R*, but no significant homolog of the transcriptional factor in any of the sought nonpathogenic mycobacterial species (*M. smegmatis* and *M. gilvum* PYR-GCK) were found, except in the pathogenic strains *M. tuberculosis*, *M. bovis*, *M. avium*, *M. africanuum*, and *M. canetti*, with nucleotide sequence identities ranging between 81% and 100%.

A recent study of the *mce* genes in a nonpathogenic soil bacterium, *Streptomyces coelicolor*, revealed the importance of these genes in enhanced access to soil resources and niche specialization [5]. It was proposed that over time, the genes became specialized into strong invasive tools in the pathogenic actinomycetes strains, shifting their niche from the soil to eukaryotes. PAHs (pyrene inclusive) occur naturally in coal, crude oil, and anthropogenically in fossil fuel products. Owing to the fact that all these compounds

have their origin in lignified plant material [12], the ability of *M. gilvum* PYR-GCK and some other members of this genus to degrade PAH efficiently may be as a result of the presence of these *mce* genes in their genome. Nevertheless, further experiments involving a gene-knockout study will be required to confirm their roles in this capacity. In our study, the *mce* genes were upregulated only in the pyrene-induced mycobacterial cells, and not a single observation was found of high expression in the glucose-induced mycobacterial cells (Supplementary material).

Moreover, based on preceding reports [13, 16, 28], *mce* genes have been observed expressed during cholesterol and other sterol metabolism in rhodococcal species. The only report in mycobacterial species, made by Wipperman *et al*. [29], stated that only two genes in the *mce3* operon (*fadE17-fadE18*) and the *mce3* transcriptional regulator (*mce3R*) were expressed in response to cholesterol metabolism and virulence in *M. tuberculosis*. In our results, no *mce3R* homolog was found, and likewise in the most commonly utilized nonpathogenic strain, *M. smegmatis*. Since we observed the expression of numerous *mce* genes in response to pyrene metabolism, coupled with the fact that our strain was a nonpathogenic strain, we did not expect the same result as Wipperman *et al*.'s report. Of great importance, however, was the fact that PAHs are a form of phytosterol, based on their plant origin, but differ from most sterols owing to their different polar properties [24]. As a result of this, we chose to integrate the idea of the *mce* genes activation in the rhodococcal species due to sterol metabolism, with the same in our mycobacterial strain.

Although no study has been done yet on the effect of PAH degradation on *mce* genes, save for this, a further study is only advisable to shed more light on the level of activity that can result in a rapid evolution of the non-existing *mce* transcription factors *mce1R* and *mce3R*; and the possible conversion of the *mce* operons to full virulence activity in the nonpathogenic strains. This idea was suggested based on the observation of some other genes highly expressed during the global transcriptomics study, such as carbon starvation protein CstA (Mflv_4322), two copper resistance D domain proteins (Mflv_2826, Mflv_1126), heat shock protein Hsp15 (Mflv_3238), stress responsive alpha-beta barrel domain protein (Mflv_4555), two UspA domain proteins (Mflv_0390, Mflv_4854), and two competence-related genes (Mflv_2662, Mflv_2663 coding for *comEA* and *comEC* genes), in response to pyrene induction (Supplementary material).

In conclusion, mammalian cell entry genes are important for host invasiveness in pathogenic mycobacteria. Based on the limited information on the functions of these genes, we sought for its expression in a nonpathogenic bacterium intended for use in toxic substrate bioremediation purposes. Our results showed a high expression of the *mce* genes during pyrene metabolism compared with its expression during glucose metabolism. We proposed that the pyrene-induced expression was as a result of pyrene, being a phytosterol, emulating the result of sterol metabolism in the rhodococcal species, thereby resulting in a similar expression of the *mce* genes; and we also repudiate the possibility of these genes functioning in the reported pathogenic invasion, based on the absence of the transcription factors required for this purpose in the *M. gilvum* PYR-GCK strain.

Acknowledgments

This work was supported by Hanyang University (ERICA) under the contract number HY-2013-P.

References

1. Arruda S, Bomfim G, Knights R, Huima-Byron T, Riley LW. 1993. Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells. *Science* **261:** 1454-1457.
2. Badejo AC, Badejo AO, Shin KH, Chai YG. 2013. A gene expression study of the activities of aromatic ring-cleavage dioxygenases in *Mycobacterium gilvum* PYR-GCK to changes in salinity and pH during pyrene degradation. *Plos One* **8:** e58066
3. Badejo AC, Choi CW, Badejo AO, Shin KH, Hyun JH, Lee YG, *et al.* 2013. A global proteome study of *Mycobacterium gilvum* PYR-GCK grown on pyrene and glucose reveals the activation of glyoxylate, shikimate and gluconeogenetic pathways through the central carbon metabolism highway. *Biodegradation* **24:** 741-752.
4. Boonchan S, Britz ML, Stanley GA. 2000. Degradation and mineralization of high-molecular-weight polycyclic aromatic hydrocarbons by defined fungal-bacterial cocultures. *Appl. Environ. Microbiol.* **66:** 1007-1019.
5. Clark CL, Seipke RF, Prieto P, Willemse J, van Wezel GP, Hutchings MI, Hoskisson PA. 2013. Mammalian cell entry genes in *Streptomyces* may provide clues to the evolution of bacterial virulence. *Sci. Rep.* **3:** 1109.
6. Casali N, Riley LW. 2007. A phylogenomic analysis of the Actinomycetales *mce* operons. *BMC Genom.* **8:** 60.
7. Cerniglia CE. 1984. Microbial metabolism of polycyclic aromatic hydrocarbons. *Adv. Appl. Microbiol.* **30:** 31-71.
8. Chitale S, Ehrt S, Kawamura I, Fujimura T, Shimono N, Anand N, *et al.* 2001. Recombinant *Mycobacterium tuberculosis* protein associated with mammalian cell entry. *Cell. Microbiol.*

3: 247-254.
9. Gioffre A, Infante E, Aguilar D, Santangelo MP, Klepp L, Amadio A, *et al.* 2005. Mutation in *mce* operons attenuates *Mycobacterium tuberculosis* virulence. *Microb. Infect.* **7:** 325-334.
10. Haile Y, Caugant DA, Bjune G, Wiker HG. 2002. *Mycobacterium tuberculosis* mammalian cell entry operon (*mce*) homologs in *Mycobacterium* other than tuberculosis (MOTT). *FEMS Immunol. Med. Microbiol.* **33:** 125-132.
11. Klepp LI, Forrellad MA, Osella AV, Blanco FC, Stella EJ, Bianco MV, *et al.* 2012. Impact of the deletion of the six *mce* operons in *Mycobacterium smegmatis*. *Microb. Infect.* **14:** 590-599.
12. Krauss M, Wilcke W, Martius C, Bandeira AG, Garcia MV, Amelung W. 2005. Atmospheric versus biological sources of polycyclic aromatic hydrocarbons (PAHs) in a tropical rain forest environment. *Environ. Pollut.* **135:** 143-154.
13. Kumar A, Bose M, Brahmachari V. 2003. Analysis of expression profile of mammalian cell entry (*mce*) operons of *Mycobacterium tuberculosis*. *Infect. Immun.* **71:** 6083-6087.
14. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25:** 402-408.
15. McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, *et al.* 2006. The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc. Natl. Acad. Sci. USA* **103:** 15582-15587.
16. Mohn WW, van der Geize R, Stewart GR, Okamoto S, Liu J, Dijkhuizen L, Eltis LD. 2008. The actinobacterial *mce4* locus encodes a steroid transporter. *J. Biol. Chem.* **283:** 35368-35374.
17. Mortazavi A, Williams BA, Mccue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5:** 621-628.
18. Peng RH, Xiong AS, Xue Y, Fu XY, Gao F, Zhao W, *et al.* 2008. Microbial biodegradation of polyaromatic hydrocarbons. *FEMS Microbiol. Rev.* **32:** 927-955.
19. Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **132:** 365-386.
20. Sassetti CM, Rubin EJ. 2003. Genetic requirements for mycobacterial survival during infection. *Proc. Natl. Acad. Sci. USA* **100:** 12989-12994.
21. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, *et al.* 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol. Biol.* **7:** 3.
22. Shimono N, Morici L, Casali N, Cantrell S, Sidders B, Ehrt S, Riley LW. 2003. Hypervirulent mutant of *Mycobacterium tuberculosis* resulting from disruption of the *mce1* operon. *Proc. Natl. Acad. Sci. USA* **100:** 15918-15923.
23. Shuttleworth KL, Cerniglia CE. 1995. Environmental aspects of PAH biodegradation. *Appl. Biochem. Biotechnol.* **54:** 291-302.
24. Slaytor M, Bloch K. 1965. Metabolic transformation of cholestenediols. *J. Biol. Chem.* **240:** 4598-4602.
25. Stingley RL, Brezna B, Khan AA, Cerniglia CE. 2004. Novel organization of genes in a phthalate degradation operon of *Mycobacterium vanbaalenii* PYR-1. *Microbiology* **150:** 3749-3761.
26. Tekaia F, Gordon SV, Garnier T, Brosch R, Barrell BG, Cole ST. 1999. Analysis of the proteome of *Mycobacterium tuberculosis in silico*. *Tuber. Lung Dis.* **79:** 329-342.
27. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, *et al.* 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28:** 511-515.
28. Van der Geize R, Yam K, Heuser T, Wilbrink MH, Hara H, Anderton MC, *et al.* 2007. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. *Proc. Natl. Acad. Sci. USA* **104:** 1947-1952.
29. Wipperman MF, Yang M, Thomas ST, Sampson NS. 2013. Shrinking the FadE proteome of *Mycobacterium tuberculosis*: insights into cholesterol metabolism through identification of an alpha2beta2 heterotetrameric acyl coenzyme A dehydrogenase family. *J. Bacteriol.* **195:** 4331-4341.