LIBRARY AND INFORMATION CENTER
MAHIDOL UNIVERSITY

9 OCT 1996

CHARACTERIZATION OF TWO DNA PROBES FOR DIFFERENTIATION OF *MYCOBACTERIUM TUBERCULOSIS* STRAINS

WISES NAMWAT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE (MICROBIOLOGY)

อภินันทนาการ
จาก
บัณฑิตวิทยาลัย มหิดล

IN

FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY

1996

Copyright by Mahidol University

TH
W814c
1996

36548

ชื่อวิทยานิพนธ์	การศึกษาคุณลักษณะของ DNA probe เพื่อใช้ในการจำแนกสายพันธุ์ของเชื้อวัณโรค
ผู้วิจัย	วิเศษ นามวาท
ปริญญา	วิทยาศาสตรมหาบัณฑิต (จุลชีววิทยา)
คณะกรรมการควบคุมวิทยานิพนธ์	ประสิทธิ์ ผลิตผลการพิมพ์
	วัฒนาลัย ปานบ้านเกร็ด
	มธุรส พงษ์ลิขิตมงคล
วันที่สำเร็จการศึกษา	17 มิถุนายน พ.ศ. 2539

บทคัดย่อ

เชื้อ *Mycobacterium tuberculosis* เป็นเชื้อที่มีความสำคัญทางการแพทย์มาก เนื่องจากเป็นสาเหตุของวัณโรค ซึ่งเป็นสาเหตุการตายที่สำคัญของประชากรไทยและทั่วโลก โดยเฉพาะอย่างยิ่งพบว่าปัจจุบันมีการเพิ่มจำนวนผู้ป่วยมากขึ้นเมื่อมีการระบาดของเชื้อ Human Immunodeficiency Virus (HIV) การจำแนกสายพันธุ์ของเชื้อ *Mycobacterium tuberculosis* มีประโยชน์ทางด้านการศึกษาระบาดวิทยา การให้การรักษาผู้ป่วย และการควบคุมโรค ปัจจุบันมีการใช้วิธีทางอณูพันธุศาสตร์ Restriction Fragment Length Polymorphism (RFLP) หรือ Polymerase chain Reaction (PCR) มาจำแนกสายพันธุ์ของเชื้อ

การศึกษานี้ใช้เทคนิค PCR เพื่อจำแนกสายพันธุ์ของเชื้อวัณโรคโดยการเพิ่มขยายจำนวนดีเอ็นเอในบริเวณที่มี polymorphism จำนวน 2 บริเวณที่อยู่บน DNA probe สองชิ้น probe ทั้งสองได้มาจาก DNA ของเชื้อวัณโรคสายพันธุ์ H37Rv มีความยาว 3.5 kb และ 8.6 kb โดยโคลนอยู่ใน pUC12 เรียกว่า pTB152 และ pTB113 ตามลำดับ ได้ทำการศึกษาหาบริเวณที่มี polymorphism ในทั้งสอง probe เพื่อนำไป subclone ต่อไป

ได้พบว่าบริเวณระหว่างตำแหน่ง *Sma*I และ *Xba*I ขนาดความยาว 0.4 kb ซึ่งเรียกบริเวณนี้ว่า SX0.4 เป็นบริเวณที่มี polymorphism บน pTB152 เมื่อหาลำดับเบสพบว่ามีความยาว 441 bp ภายในบริเวณนี้มี direct repetitive sequence ขนาด 52 bp จำนวน 2 บริเวณอยู่ห่างกัน 6 bp ตำแหน่งที่ 35 ของทั้งสองบริเวณมีเบสแตกต่างกัน (T→C) เมื่อใช้โปรแกรม DNASIS version 7.00 วิเคราะห์หาบริเวณที่มีลักษณะเป็นรหัสของ

ลิขสิทธิ์มหาวิทยาลัยมหิดล Mahidol University

tRNA พบว่าบริเวณตำแหน่งเบสที่ 320 ถึง 396 มีโอกาสที่จะเป็นรหัส tRNA ของ glycine หรือ glutamine

จากการเปรียบเทียบลำดับเบสของ SX0.4 กับข้อมูลใน GenBank โดยใช้โปรแกรม BLASTn พบว่าบริเวณตำแหน่งที่ 1-196 ของปลาย 3' ของ complementary strand มีลักษณะเหมือนกับปลาย 5' ของยีน cyclopropane mycolic acid synthase 2 (*cma2*) (GenBank accession number U34637) โดยมีความเหมือน 98%(194/196)

ส่วนใน pTB113 พบบริเวณที่มี polymorphism อยู่ระหว่างตำแหน่ง *PstI* สองแห่ง มีความยาว 0.8 kb จึงเรียกบริเวณนี้ว่า PP0.8 เมื่อหาลำดับเบสพบว่ามีความยาว 766 bp โดยมี direct repetitive sequence มีขนาด 57 bp จำนวน 2 แห่งต่อเนื่องกัน นอกจากนั้น ลำดับเบสขนาด 9 bp ในบริเวณถัดไปมีลักษณะเหมือนกับ 9 bp เริ่มต้นของ repetitive sequence

เมื่อเปรียบเทียบลำดับของกรดอะมิโนที่ถอดระหัสมาจากลำดับเบสของ PP0.8 กับข้อมูลใน GenBank โดยใช้โปรแกรม BLASTx พบว่ามีลักษณะเหมือนกับยีน α-isopropylmalate synthase (α-IPMS) ของเชื้อ *Corynebacterium glutamigum* (Genbank accession number p42455 or 452382) โดยมีความเหมือน 49%(53/108) นอกจากนี้ได้ศึกษาลำดับเบสในบริเวณข้างเคียงของ PP0.8 พบว่ามีความเหมือนกับยีน α-IPMS ในลำดับที่ถูกต้อง

ได้ใช้เทคนิค PCR เพื่อตรวจวัดลักษณะ polymorphism ในบริเวณ SX0.4 และ PP0.8 ในเชื้อวัณโรค 179 ตัวอย่าง พบว่ามีลักษณะของ polymorphism 40 รูปแบบ โดยที่รูปแบบที่พบบ่อยที่สุดจะมีความยาวของ amplified products ของแต่ละบริเวณสั้นที่สุด คือ 437 bp และ 751 bp รูปแบบนี้พบในเชื้อจำนวน 80 ตัวอย่าง เชื้อในกลุ่มนี้รวมไปถึงเชื้อในสายพันธุ์ H37Rv และ Mt14323 ด้วย

ลักษณะของ polymorphism ในบริเวณ SX0.4 พบว่ามี 12 แบบโดยมีความยาวของ amplified products ตั้งแต่ 437 bp ถึง 1230 bp เมื่อใช้ SX1 และ SX2 เป็น primer

ส่วนในบริเวณ PP0.8 พบว่ามีลักษณะของ polymorphism จำนวน 18 แบบ โดยมีความยาวของ amplified products ตั้งแต่ 751 bp ถึง 1950 bp เมื่อใช้ PP3 กับ PP4 เป็น primer

ก่อนหน้านี้ได้มีการศึกษาลักษณะ RFLP ของเชื้อทั้ง 179 ตัวอย่าง ด้วยวิธีมาตรฐานคือ Southern hybridization และ probe ด้วย IS*6110* ข้อมูลลักษณะโครงสร้างประชากรจากการนี้ได้นำมาศึกษาความสัมพันธ์กับลักษณะของ polymorphism ในทั้งสองบริเวณ พบว่ามีลักษณะดังนี้

ลิขสิทธิ์มหาวิทยาลัยมหิดล Copyright by Mahidol University

เชื้อ 69 ตัวอย่างที่ถูกจำแนกเป็น Beijing family ด้วย IS*6110* มีรูปแบบของ polymorphism คล้ายกันเป็นส่วนใหญ่ คือมีลักษณะรูปแบบเหมือนเชื้อสายพันธุ์ H37Rv ยกเว้นเชื้อจำนวน 5 ตัวอย่างมีลักษณะแตกต่างไปจากกลุ่ม

เชื้อ 23 ตัวอย่างที่ถูกจำแนกเป็น Nonthaburi group มีลักษณะของ polymorphism จำนวน 12 รูปแบบที่มีขนาดความยาวของ amplified products ใกล้เคียงกัน

เชื้อ 38 ตัวอย่างที่ให้ผล hybridization เพียงหนึ่งตำแหน่งซึ่งมีอยู่ 2 กลุ่มคือ กลุ่มที่มี hybridization band ที่ตำแหน่งประมาณ 1.45 kb ซึ่งประกอบด้วยเชื้อจำนวน 29 ตัวอย่าง และ กลุ่มที่มี hybridization band ที่ตำแหน่งประมาณ 5.0 kb ซึ่งประกอบด้วยเชื้อจำนวน 9 ตัวอย่าง พบว่ากลุ่มแรกมีลักษณะของ polymorphism จำนวน 21 รูปแบบ ส่วนกลุ่มที่สองมีลักษณะของ polymorphism จำนวน 2 รูปแบบ

ส่วนเชื้ออีก 49 ตัวอย่างที่ไม่สามารถจัดกลุ่มได้ พบว่ามีลักษณะของ polymorphism จำนวน 23 รูปแบบ

นอกจากนี้พบว่าลักษณะของ polymorphism ในทั้งสองบริเวณไม่มีความสัมพันธ์กับลักษณะการดื้อยา เพศของผู้ป่วย ถิ่นที่อยู่ของผู้ป่วย หรือ ผลเลือดที่ตรวจพบการติดเชื้อ HIV ของผู้ป่วย รวมทั้งอวัยวะที่แยกเชื้อวัณโรคได้

ได้ศึกษาสาเหตุของ polymorphism ในทั้งสองบริเวณ เชื่อว่ามีสาเหตุมาจากการเพิ่มจำนวนของ repetitive sequence ขนาดความยาว 57 bp ซึ่งใน SX0.4 เชื่อว่าลำดับเบสที่เพิ่มขึ้นมาประกอบด้วยส่วนของ repetitive sequence ขนาดยาว 52 bp และแทรกด้วยลำดับเบสอื่นที่มีความยาวเฉลี่ย 5 bp ส่วนใน PP0.8 ลำดับเบสที่เพิ่มขึ้นมาน่าจะเป็น repetitive sequence ที่พบใน PP0.8 นั่นเอง

วิธีการใหม่ที่ใช้ในการจำแนกสายพันธ์ของเชื้อวัณโรคโดยใช้เทคนิค PCR จะสามารถทำได้ง่ายขึ้น รวดเร็วขึ้น และใช้ปริมาณ DNA น้อยกว่าเมื่อเปรียบเทียบกับวิธีการมาตรฐาน

Copyright by Mahidol University

Thesis Title	Characterization of two DNA Probes for Differentiation of *Mycobacterium tuberculosis* Strains
Name	Wises Namwat
Thesis Supervisory Committee	Prasit Palittapongarnpim
	Watanalai Panbangred
	Mathurose Ponglikitmongkol
Date of Graduation	17 June B.E. 2539 (1996)

ABSTRACT

Mycobacterium tuberculosis is a medically important bacterium, which is a major cause of morbidity and mortality around the world, especially when the spread of the Human Immunodeficiency virus (HIV) is rising world wide. In Thailand, tuberculosis is still a leading cause of death among infectious diseases. In the epidemiological study of *M. tuberculosis*, useful information can be obtained from a good typing system. It can provide the useful data in supporting or rejecting clinical hypothesis, such as an outbreak of infection, single or multiple infection, a relapse due to the original organism or a new infection due to a different strain.

This study involved developing a technique for a strain differentiation by using PCR technique to amplify two polymorphic segments. pTB152 and pTB113 containing 3.5 kb and 8.6 kb *Eco*RI inserted fragment respectively, were two recombinant plasmids previously cloned in our laboratories. They were previously found to contain polymorphic segments. These polymorphic segments were located and subcloned.

The polymorphic containing regions of pTB152 and pTB113 were 0.4 kb *Sma*I-*Xba*I fragment (designated SX0.4) and 0.8 kb *Pst*I-*Pst*I fragment (designated PP0.8), respectively. SX0.4 was 441 bp long. It possessed two

copies of 52 bp long direct repetitive sequence interspersed with 6 bp of nucleotides with a different nucleotide at the position 35 of the repetitive sequences (T→ C). According to analysis by DNASIS version 7.00, it possibly contained a t-RNA encoding sequence which started from nucleotide 320 to 396. The t-RNA coding sequence could have one of the two possible conformations with the anticodon specifying glycine or glutamine.

Comparison of the sequence of SX0.4 to the DNA sequence deposited in GenBank by BLASTn program revealed that nucleotide sequence from 1-196 at the 3′ end of the complementary strand of SX0.4 was 98%(194/196) homologous to the 5′ flanking sequence *M. tuberculosis* cyclopropane mycolic acid synthase 2 (*cma2*) gene, GenBank accession number U34637.

PP0.8 was 776 bp long. It possessed two adjacent direct repetitive sequences of 57 bp long. The sequence of the next adjacent 9 nucleotides was also identical to the initial 9 nucleotide of the repetitive sequences.

Comparison of the predicted amino acid sequences of PP0.8 to the sequences deposited in GenBank by BLASTx program revealed that the nucleotide fragment (46 to 525) of PP0.8 was 49%(53/108) homologous to *Corynebacterium glutamicum* isopropylmalate synthase (α-IPMS) (Genbank accession number P42455 or 452382). Moreover, another fragment flanking the PP0.8 segment was sequenced and was found to be homologous to the sequence of the corresponding position in the α-IPMS gene.

PCR condition was optimized and used for determining the polymorphism corresponding to both SX0.4 and PP0.8 segment of 179 *M. tuberculosis* isolates, by simultaneously using SX1-SX2 and PP3-PP4 as primers.

Forty patterns of polymorphic bands were found. The most common type comprising 80 isolates, was the ones with the shortests length of 437 bp

Copyright by Mahidol University

and 751 bp long. This type included *M. tuberculosis* H37Rv and Mt14323 strain.

There were 12 polymorphic types of amplified products corresponding to SX0.4 segment. Using SX1 and SX2 as primers, the length of the amplified products varied form 437 to 1230 bp long.

There were 18 polymorphic types of amplified products corresponding to PP0.8 segment. Using PP3 and PP4 as primers, the length of the amplified products varied from 751 to 1950 bp long.

The RFLP of all the 179 isolates of *M. tuberculosis* was previously determined by the standard method, that is Southern hybridization with IS*6110*. Result comparison between the standard method and the amplified fragment length polymorphism was analyzed.

Most of the 69 isolates of Beijing family showed the pattern similar to H37Rv strain, only 5 of them have 4 other different patterns.

Twenty-three isolates, belonging to the Nonthaburi group, had 12 patterns with amplified products having length close to each other. They were not belonging to the most common type.

Thirty-eigth isolates were hybridized with IS*6110* at only one position. The hybridized fragments were of only two lengths, that was either about 1.45 kb (29 isolates) or 5.0 kb (9 isolates).By using the amplified fragment length polymorphism typing, 21 and 2 patterns were found among 1.45 kb single-banded group and 5.0 kb single-banded group, respectively.

Twenty-three patterns were found among 49 isolates which could not be grouped by IS*6110*-based RFLP.

The amplified polymorphic patterns did not correlate to drug resistance, sex, location and HIV seropositivity of the patients as well as the organ sites that the isolates were obtained.

Copyright by Mahidol University

The cause of amplified fragment length polymorphism was analysed. In PP0.8 segment, the polymorphism was most likely caused by the difference in

the number of the copies of 57 bp repetitive sequence. In SX0.4 segment, the polymorphism was also possibly caused by the difference in the number of the copies of 57 bp long sequence. This sequence most likely included the 52 bp repetitive sequences which were interspersed with non-repetitive sequences with the average length of 5 bp.

These new methods for DNA polymorphism typing utilizing PCR technology will allow easier and more rapid typing by using less DNA isolated from early cultures and even uncultured specimens.

Copyright by Mahidol University