



Multilocus Sequence Typing (MLST): เทคนิคระดับโมเลกุลเพื่อติดตามความปลอดภัยในอาหาร

ชนม์ชนก ถิระแก้ว*

บทคัดย่อ

ปัจจุบันโรคติดเชื้อทางเดินอาหารถือเป็นปัญหาสาธารณสุขทั่วโลกโดยเฉพาะในประเทศแถบเขตร้อน วิธีการตรวจสอบแหล่งที่มาของการติดเชื้อแบคทีเรียก่อโรคทางเดินอาหารสามารถทำได้โดยการศึกษาความสัมพันธ์ทางพันธุกรรมของสายพันธุ์ที่แยกได้จากผู้ป่วยโรคอุจจาระร่วงและที่แยกได้จากแหล่งต่างๆ รวมทั้งอาหาร Multilocus Sequence Typing (MLST) เป็นวิธีการวิเคราะห์ทางพันธุกรรมและการจำแนกชนิดแบคทีเรียโดยใช้หลักความสัมพันธ์ของรหัสพันธุกรรมของยีนพื้นฐาน (Housekeeping Genes) ของแบคทีเรียที่แยกได้จากแหล่งต่างๆ เพื่อวิเคราะห์ความใกล้ชิดของสายพันธุ์และใช้ในการระบุแหล่งการแพร่กระจายของเชื้อก่อโรค ข้อได้เปรียบที่สำคัญของวิธี MLST

กับวิธีทางโมเลกุลประเภทอื่นคือ สามารถเปรียบเทียบความสัมพันธ์ของสายพันธุ์ที่ศึกษาด้วยสายพันธุ์อื่นจากทั่วโลกได้โดยผ่านฐานข้อมูลทางอินเทอร์เน็ต ในปัจจุบันนำวิธี MLST มาใช้อย่างแพร่หลายในด้านวิทยาการระบาดและวิวัฒนาการของแบคทีเรียก่อโรคในอาหารที่สำคัญหลายชนิด ข้อมูลที่ได้จากวิธี MLST สามารถนำไปใช้ประโยชน์ในการตรวจตราการปนเปื้อนแบคทีเรียก่อโรคในอุตสาหกรรมอาหาร ส่งผลให้ระบบรักษาความปลอดภัยในอาหารมีมาตรฐานสูงขึ้น

คำสำคัญ: วิธีจำแนกโดยรหัสพันธุกรรมหลายตำแหน่ง, จุลชีพก่อโรคโดยอาหารเป็นสื่อ, ความสัมพันธ์ของพันธุกรรม, ระบาดวิทยาโมเลกุล

1. ความเป็นมาของวิธี MLST

Multilocus sequence typing (MLST) เป็นวิธีที่ใช้จำแนกสายพันธุ์ของแบคทีเรียโดยการเปรียบเทียบรหัสพันธุกรรมของยีนพื้นฐาน (Housekeeping Genes) ยีนเหล่านี้มีบทบาทสำคัญในการดำรงชีวิตของจุลชีพ เช่น ยีนที่ทำหน้าที่ในการสร้างเอนไซม์เกี่ยวข้องในกลไกการสร้างพลังงาน เช่น Fructose-bisphosphate Aldolase Class 2 (*fbaA*) และยีนที่ใช้ในกระบวนการซ่อมแซมสารพันธุกรรม เช่น Recombinase A (*recA*) เป็นต้น วิธี MLST เป็นที่รู้จักครั้งแรกเมื่อปี ค.ศ. 1998 โดย Martin C. J. Maiden และคณะ¹ เป็นผู้พัฒนามาจากวิธี Multilocus Enzyme Electrophoresis (MLEE) ที่คิดค้นโดย Robert K. Selander และคณะ² ในปี ค.ศ.1986 วิธี MLEE เป็นวิธีการจำแนกสายพันธุ์ของแบคทีเรียจากลักษณะทางกายภาพ (Phenotype) ของเอนไซม์ที่สำคัญในการดำรงชีวิตของจุลชีพโดยใช้เกณฑ์การเปรียบเทียบระยะทางการเคลื่อนที่บนสนามไฟฟ้าของเอนไซม์นั้น แต่เนื่องจากวิธี MLEE สามารถจำแนกสายพันธุ์แบคทีเรียได้ต่อเมื่อความต่างของสารพันธุกรรมในยีนที่สร้างเอนไซม์มีอิทธิพลมากพอที่จะทำให้คุณสมบัติการเคลื่อนที่ของเอนไซม์บนสนามไฟฟ้าต่างกัน ทำให้การจำแนกสายพันธุ์แบคทีเรียด้วยวิธี MLEE มีความละเอียดไม่มากพอที่จะจำแนกความแตกต่างของแบคทีเรียบางสายพันธุ์ได้ ในปี ค.ศ. 1998 Martin C. J. Maiden และคณะจึงริเริ่มพัฒนาวิธี MLST ขึ้นเพื่อการจำแนกสายพันธุ์แบคทีเรียให้มีความละเอียดยิ่งขึ้นในระดับที่แม้แต่ความต่างเพียงตำแหน่งเดียวของรหัสพันธุกรรมก็สามารถจำแนกประเภทของสายพันธุ์แบคทีเรียได้

นอกจากนี้ข้อได้เปรียบที่สำคัญของวิธี MLST ได้แก่ สามารถทดสอบซ้ำได้ มีความน่าเชื่อถือสูง และสามารถนำข้อมูลของแบคทีเรียที่ศึกษาไปเปรียบเทียบกับข้อมูลจากทั่วโลกโดยผ่านฐานข้อมูลทางอินเทอร์เน็ตเพื่อใช้ศึกษาการระบาดของแบคทีเรียเหล่านั้น³ อย่างไรก็ตามวิธี MLST มีข้อเสียคือต้องอาศัยอุปกรณ์ที่มีเทคโนโลยีสมัยใหม่ เช่น เครื่องถอดรหัสพันธุกรรม (DNA Sequencer) ทำให้ค่าใช้จ่ายในการวิเคราะห์สูง แต่ในปัจจุบันมีการพัฒนาเทคโนโลยีดังกล่าวอย่างต่อเนื่องทำให้ค่าใช้จ่ายในการถอดรหัสพันธุกรรมลดลงกว่าสมัยก่อนมาก และมีแนวโน้มจะลดลงอีกในอนาคตจนถึงระดับที่สามารถประยุกต์เพื่อพัฒนาเทคนิคปฏิบัติการประจำวันของการตรวจพันธุกรรมแบคทีเรียก่อโรคในอาหารได้

ขั้นตอนโดยย่อของวิธี MLST เริ่มต้นจากการแยกแบคทีเรียจากตัวอย่างอุจจาระผู้ป่วยโรคอุจจาระร่วง อาหาร หรือสิ่งแวดล้อม จากนั้นสกัดสารพันธุกรรมจากแบคทีเรียที่แยกได้ และใช้วิธี Polymerase Chain Reaction (PCR) เพิ่มจำนวนสารพันธุกรรมของยีนพื้นฐานของแบคทีเรียแต่ละสายพันธุ์เพื่อเข้าสู่กระบวนการถอดรหัสพันธุกรรม เมื่อได้ข้อมูลรหัสพันธุกรรมของแบคทีเรียที่ใช้ศึกษาแล้วจึงบรรจุข้อมูลดังกล่าวลงในฐานข้อมูลทางอินเทอร์เน็ต (<http://pubmlst.org>) โดยผ่านผู้ตรวจสอบประจำฐานข้อมูล (curator) ทำให้สามารถวิเคราะห์เปรียบเทียบความสัมพันธ์เชิงพันธุศาสตร์ของเชื้อก่อโรคจากตัวอย่างที่ศึกษาและแบคทีเรียชนิดดังกล่าวจากทั่วโลกโดยใช้โปรแกรมคอมพิวเตอร์⁴⁻⁶ (รูปที่ 1)

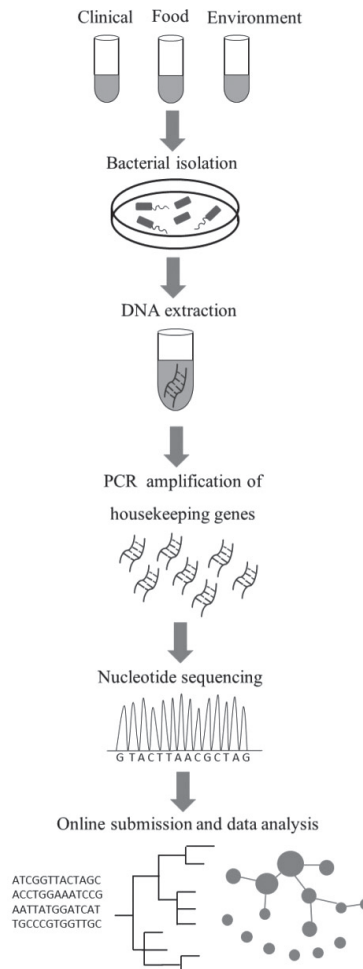


Figure 1 Diagram Shows MLST Scheme Starting from Bacterial Isolation from Clinical, Seafood and Environmental Samples. DNA extraction was subsequently performed, followed by PCR amplification of housekeeping genes for each isolate. Nucleotides of PCR fragments of the housekeeping genes are sequenced, and the nucleotide sequence data will be submitted to online database.

2. วิธีตรวจทางชีววิทยาระดับโมเลกุลที่ใช้ในการตรวจย้อนกลับเพื่อสืบหาแหล่งที่มาของแบคทีเรียก่อโรค

การศึกษาแหล่งที่มาของเชื้อก่อโรคในอุตสาหกรรมอาหาร จำเป็นต้องมีความเข้าใจกระบวนการผลิตและการขนส่งอาหารตั้งแต่แหล่งวัตถุดิบ เช่น จากทะเล

แม่น้ำ หรือ ฟาร์มเลี้ยงสัตว์ จนถึงระบบการขนส่งวัตถุดิบจากแหล่งผลิตจนถึงโรงงานแปรรูป ตลอดจนสภาพแวดล้อมของแหล่งจำหน่ายในตลาดท้องถิ่น และในห้างสรรพสินค้า จนถึงมือผู้บริโภค เนื่องจากจำนวนและชนิดของแบคทีเรียในแต่ละแหล่งที่พบมีความแตกต่างกัน ยกตัวอย่างเช่น จำนวนแบคทีเรีย

ที่พบในอาหารชนิดเดียวกันจากสถานที่ขายอาหาร บริเวณในและนอกอาคารมีปริมาณต่างกัน⁷ และการปฏิบัติด้านสุขาภิบาลอาหาร เช่น การให้ความรู้ เพื่อเพิ่มทัศนคติการปฏิบัติอย่างถูกสุขอนามัยของผู้สัมผัสอาหารส่งผลต่อระดับการปนเปื้อนแบคทีเรียในอาหาร⁸ เป็นต้น นอกจากนี้แบคทีเรียบางชนิดเป็นแบคทีเรียประจำถิ่นที่อาศัยอยู่ในร่างกายสัตว์ จึงสามารถปนเปื้อนมากับเนื้อสัตว์ที่นำมาใช้ทำอาหาร แบคทีเรียบางชนิดปนเปื้อนมากับน้ำหรือมือผู้ผลิตอาหารเอง และอาจทำให้แบคทีเรียเหล่านั้นเข้าสู่ร่างกายผู้บริโภคได้ ปัจจัยที่เกี่ยวข้องกับความเสี่ยงในการปนเปื้อนเชื้อก่อโรคในอาหารมีหลายประการ เช่น คุณภาพการขนส่งอาหารจนถึงมือผู้บริโภค ความสะอาดและอนามัยของผู้จำหน่ายอาหาร และวิธีการปรุงอาหาร เป็นต้น⁹ อย่างไรก็ตามแบคทีเรียที่ปนเปื้อนในอาหารมีทั้งสายพันธุ์ก่อโรคและไม่ก่อโรค การตรวจสอบการปนเปื้อนแบคทีเรียในอาหารจึงมีความจำเป็นต้องจำแนกถึงระดับสายพันธุ์เพื่อการประเมินความเสี่ยงที่แม่นยำ¹⁰ นอกจากนี้การศึกษาสายพันธุ์ของแบคทีเรียมีประโยชน์อย่างยิ่งในการศึกษาวิทยาการระบาด เช่น การสืบหาแหล่งที่มาของสายพันธุ์ที่ระบาดมาจากแหล่งใดโดยการเปรียบเทียบลักษณะทางพันธุกรรมของแบคทีเรียที่แยกได้ในผู้ป่วยติดเชื้อ และจากตัวอย่างอาหารหรือสภาพแวดล้อมได้โดยวิธีการจำแนกชนิดที่ระดับโมเลกุล (Molecular Typing) โดยอาศัยความแตกต่างของรหัสพันธุกรรมของแบคทีเรียต่างสายพันธุ์ วิธีดังกล่าวประยุกต์เพื่อสืบหาแหล่งที่มาของแบคทีเรียก่อโรคอุจจาระร่วงอย่างแพร่หลาย¹¹⁻¹⁴ ยกตัวอย่างเช่น วิธี Pulse Field Gel Electrophoresis (PFGE) และ Restriction Fragment Length Polymorphism (RFLP) ใช้จำแนกสายพันธุ์โดยอาศัยเอนไซม์ตัดจำเพาะ (Restriction Enzyme) กับลำดับสารพันธุกรรมในสาย DNA หลังจากที่ได้สาย

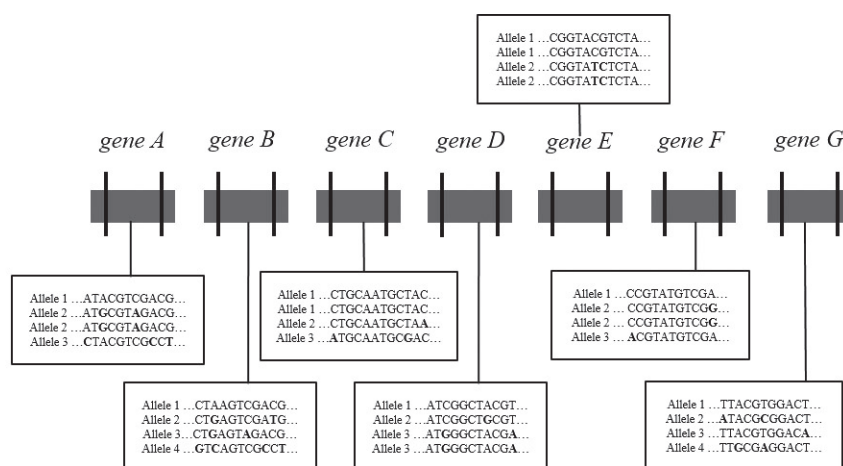
DNA ของแบคทีเรียถูกเอนไซม์ตัดแล้ว ชิ้น DNA จะเคลื่อนที่บนเจลตัวนำกระแสสลับในสนามไฟฟ้า ระยะการเคลื่อนที่ของชิ้น DNA จากแบคทีเรียที่มีรหัสพันธุกรรมแตกต่างกันจะไม่เท่ากัน ในขณะที่แบคทีเรียสายพันธุ์เดียวกันที่มีรหัสพันธุกรรมเหมือนกัน ชิ้น DNA ที่ถูกตัดจะเคลื่อนที่ในระยะทางที่เท่ากัน รูปแบบการเคลื่อนที่ของ DNA นี้เรียกว่า DNA fingerprint หรือ ลายพิมพ์ DNA ซึ่งเป็นสิ่งที่ใช้ในการจำแนกสายพันธุ์แบคทีเรียได้ สำหรับวิธี Random Amplified Polymorphism Deoxyribonucleic Acid (RAPD), Amplified Fragment Length Polymorphism (AFLP) และ Variable Number of Tandem Repeat (VNTR) Analysis ใช้จำแนกสายพันธุ์โดยอาศัยหลักการ PCR ในการเพิ่มจำนวนชิ้น DNA ที่จำเพาะกับสาย DNA ต้นแบบ (Primer) และนำชิ้น DNA ที่ถูกเพิ่มจำนวนนั้นมาวิเคราะห์การเคลื่อนที่บนเจลที่ผ่านกระแสไฟฟ้า แบคทีเรียที่มีรหัสพันธุกรรมเหมือนกันจะให้ผลการเพิ่มจำนวนชิ้น DNA ที่จำเพาะต่อ Primer เหมือนกัน ทำให้เห็นรูปแบบของ DNA ที่ปรากฏบนเจลเหมือนกัน ในขณะที่แบคทีเรียที่มีรหัสพันธุกรรมต่างกันจะแสดงรูปแบบการเคลื่อนที่ของ DNA ต่างกัน ส่วนวิธี Multilocus Sequence Typing (MLST) และ Single Nucleotide Polymorphism (SNP) Analysis เป็นวิธีทางโมเลกุลที่อาศัยเทคโนโลยีการอ่านรหัสพันธุกรรม (DNA Sequencing) และถือเป็นวิธีที่มีวิวัฒนาการทันสมัยที่สุด วิธี MLST ใช้จำแนกสายพันธุ์โดยการเปรียบเทียบรหัสพันธุกรรมของยีนพื้นฐานอย่างน้อย 7 ยีน ส่วนวิธี SNP Analysis จำแนกโดยการเปรียบเทียบรหัสพันธุกรรมหรือชนิดของเบสที่แตกต่างกันเพียงตำแหน่งเดียว (Single Polymorphic Site) ที่พบในรหัสพันธุกรรมทั้งหมดของแบคทีเรีย

วิธีเหล่านี้สามารถจำแนกสายพันธุ์ของแบคทีเรียก่อโรคได้โดยใช้คุณสมบัติที่แตกต่างกัน จึงทำให้แต่ละวิธีมีข้อดีข้อเสียต่างกัน ยกตัวอย่างเช่น วิธี PFGE เป็นวิธีที่นิยมใช้อย่างแพร่หลายในการทำวิจัย เนื่องจากมีความละเอียดในการจำแนกและมีความจำเพาะมากกว่าวิธีที่ใช้หลักการ PCR เช่น RAPD และ AFLP อย่างไรก็ตามการวิเคราะห์โดยวิธี PFGE มีข้อจำกัดคือไม่สามารถนำข้อมูลที่ได้มาวิเคราะห์เปรียบเทียบสายพันธุ์ที่ก่อให้เกิดโรคระบาดจากต่างพื้นที่ซึ่งศึกษาโดยคณะผู้ทดลองคนละกลุ่ม จึงทำให้การเชื่อมโยงข้อมูลด้านการระบาดของเชื้อจากต่างถิ่นเป็นไปได้ยากถึงแม้ว่าวิธีการจำแนกโดยการอ่านรหัสพันธุกรรม เช่น MLST และ SNP ต้องอาศัยอุปกรณ์ที่มีความจำเพาะสูงและมีราคาแพง เช่น DNA Sequencer แต่ในปัจจุบันการวิเคราะห์รหัสพันธุกรรมสามารถทำได้ง่ายขึ้นเนื่องจากเครื่อง DNA Sequencer มีราคาถูกลง และบริษัทเอกชนที่มีบริการวิเคราะห์รหัสพันธุกรรมเกิดขึ้นจำนวนมาก ทำให้ห้องปฏิบัติการตรวจความปลอดภัยในอาหารสามารถเข้าถึงบริการการวิเคราะห์รหัสพันธุกรรมได้สะดวกขึ้น วิธี MLST จึงเป็นทางเลือกที่ได้รับความนิยมในการทำวิจัยเนื่องจากสามารถจำแนกสายพันธุ์แบคทีเรียได้ละเอียดกว่าวิธี PFGE นอกจากนี้ผลที่ได้จากวิธี MLST สามารถทำซ้ำและเปรียบเทียบกับสายพันธุ์อื่นจากต่างถิ่นได้ จึงทำให้การแปลผลมีความคงที่ เป็นมาตรฐานสากลและมีประโยชน์ต่อการศึกษาเพื่อวิเคราะห์การระบาดของเชื้อก่อโรคได้อย่างมีประสิทธิภาพ

3. วิธีการวิเคราะห์รหัสพันธุกรรมโดยวิธี MLST

การวิเคราะห์ความสัมพันธ์ของรหัสพันธุกรรม

โดยวิธี MLST สามารถทำได้ 2 แบบ คือ วิธีดูลำดับสารพันธุกรรมทั้ง 7 ยีนมาเรียงต่อเนื่องกันไป (Concatenated Sequence) และวิธีกำหนดตัวเลขที่บ่งบอกลักษณะจำเพาะของรหัสพันธุกรรมในแต่ละยีน (Allele Profile) แบคทีเรียแต่ละสายพันธุ์มี Concatenated Sequence และ Allele Profile ที่จำเพาะกรณีวิเคราะห์จาก Concatenated Sequence หากรหัสพันธุกรรมมีความต่างกันหลายตำแหน่งหมายความว่าสายพันธุ์นั้นมีความต่างทางพันธุกรรมสูง ส่วนในกรณีที่ใช้การวิเคราะห์โดย Allele Profile รหัสพันธุกรรมของยีนแต่ละยีนที่มีเอกลักษณ์ไม่ซ้ำกับรหัสพันธุกรรมของยีนนั้นในสายพันธุ์อื่นจะมีตัวเลขจำเพาะ (Allele Number) สำหรับยีนที่มีรหัสพันธุกรรมซ้ำกับยีนที่ได้รับการกำหนด Allele Number แล้ว จะมี Allele Number ที่เหมือนกันกับยีนนั้น ลักษณะพันธุกรรมที่ได้จากวิธี MLST ของแบคทีเรียหนึ่งสายพันธุ์ประกอบไปด้วย 7 Allele Numbers โดยตัวเลขแต่ละตัวกำหนดมาจากรหัสพันธุกรรมที่จำเพาะของยีนทั้ง 7 ยีน กลุ่มตัวเลขทั้ง 7 ตัวนี้เรียกว่า Allele Profile เอกลักษณ์เฉพาะของแต่ละ Allele Profile จะถูกแต่งตั้งเป็นตัวเลข เรียกว่า Sequence Type (ST) โดยแบคทีเรียแต่ละสายพันธุ์จะมี ST ประจำตัวแบคทีเรียต่างสายพันธุ์อาจมี ST เหมือนกันก็ได้ หากสายพันธุ์ดังกล่าวมีรหัสพันธุกรรมของทั้ง 7 ยีนเหมือนกัน ส่วนการวิเคราะห์รหัสพันธุกรรมโดยใช้ Allele Profile ทำได้โดยการพิจารณาจากความแตกต่างของ Allele Number รูปที่ 2 แสดงให้เห็นการระบุ Allele Number (รูปที่ 2A) ในแต่ละตำแหน่งของยีนพื้นฐานและการกำหนด ST (รูปที่ 2B) จาก Allele Profile ของเชื้อที่ศึกษาโดยวิธี MLST



(A)

Isolate	Sequence type (ST)	Allelic type of individual housekeeping genes						
		gene 1	gene 2	gene 3	gene 4	gene 5	gene 6	gene 7
1	ST1	1	1	1	1	1	1	1
2	ST2	2	2	1	2	2	2	2
3	ST3	2	3	2	2	1	2	3
4	ST4	3	4	3	3	2	3	4

(B)

Figure 2 MLST Data Interpretation. (A) Fragments of seven housekeeping genes, *genes 1-7*, used for the MLST scheme are presented. Allelic types, e.g. alleles 1, 2, 3, etc., are assigned for the unique nucleotide sequences of individual gene fragments. (B) Unique combinations of seven allelic types create the ST, e.g. ST1, 2, 3, etc., for an individual isolate³.

การวิเคราะห์รหัสพันธุกรรมจาก Concatenated Sequence มีข้อได้เปรียบคือสามารถบอกความแตกต่างของรหัสพันธุกรรมจากแบคทีเรียต่างสายพันธุ์ได้ละเอียดมากกว่าการวิเคราะห์จาก Allele Profile เนื่องจากสามารถระบุตำแหน่งและจำนวนเบสที่ต่างกันได้ในขณะที่การวิเคราะห์จาก Allele Profile ไม่สามารถจำแนกระดับความแตกต่างของรหัสพันธุกรรมในยีนเดียวกันได้ บ่งบอกได้เพียงว่ารหัสพันธุกรรมในยีนที่ศึกษามีความเหมือนหรือต่างกันเท่านั้น อย่างไรก็ตาม Allele Profile มีประโยชน์ในการวิเคราะห์การระบาด

ของกลุ่มแบคทีเรียสายพันธุ์ที่อาจมาจากบรรพบุรุษเดียวกัน (Clonal Relationship) ได้โดยอาศัยโปรแกรมคอมพิวเตอร์วิเคราะห์ เช่น eBURST¹⁵ ข้อมูล MLST ของแบคทีเรียที่ศึกษาภายในห้องปฏิบัติการของผู้วิจัยสามารถนำไปเปรียบเทียบกับข้อมูลของแบคทีเรียจากแหล่งอื่นทั่วโลกได้ที่เว็บไซต์ฐานข้อมูล MLST (www.pubmlst.org) ผู้วิจัยสามารถเทียบหา Allele Profile และ ST ของแบคทีเรียที่ศึกษาจาก Concatenated Sequence โดยผ่านโปรแกรมสำเร็จรูปในเว็บไซต์ได้เช่นกัน

4. การคัดเลือกตัวอย่างแบคทีเรีย (bacterial selection) ยีนที่เหมาะสม (gene selection) และการออกแบบสารพันธุกรรมต้นแบบใช้ในการเพิ่มจำนวนยีน (primer design) ด้วยวิธี MLST

การคัดเลือกแบคทีเรียที่จะนำมาศึกษาถือเป็นปัจจัยที่มีความสำคัญในวิธี MLST ตัวแทนแบคทีเรียที่นำมาวิเคราะห์ควรคัดเลือกมาจากแบคทีเรียที่พบในแหล่งที่มาหลากหลาย เช่น แบคทีเรียที่แยกได้จากผู้ป่วยโรคอุจจาระร่วงในประเทศไทยและต่างประเทศทั่วโลกมีประโยชน์ในการวิเคราะห์จุดเริ่มต้นของสายพันธุ์ที่กำลังระบาดรวมถึงเส้นทางการแพร่กระจายของเชื้อโรคเหล่านั้น นอกจากนี้การเปรียบเทียบความใกล้ชิดของสายพันธุ์ที่แยกได้จากผู้ป่วย สิ่งแวดล้อมและปัจจัยที่เกี่ยวข้องในกระบวนการผลิตอาหาร เช่น วัตถุดิบ น้ำล้างวัตถุดิบ ภาชนะบรรจุวัตถุดิบ พาหนะขนส่งอาหาร และคนงานในโรงงานผลิตอาหาร เป็นต้น ทำให้สามารถระบุแหล่งที่มาของการติดเชื้อก่อโรคสู่คนซึ่งมีประโยชน์ต่อการประเมินความเสี่ยงด้านความปลอดภัยในการผลิตอาหาร และส่งผลให้การกำหนดมาตรการการเฝ้าระวังการติดเชื้อในผู้บริโภคมีประสิทธิภาพมากขึ้น

การพิจารณาเลือกยีนที่เหมาะสมกับวิธี MLST ประกอบไปด้วยปัจจัยหลักคือ ชนิดยีน ตำแหน่งของยีนนั้นบนโครโมโซม และจำนวนของยีนที่ใช้^{3,5,16} การจำแนกประเภทแบคทีเรียจำเป็นต้องเข้าใจถึงวิวัฒนาการทางพันธุศาสตร์ที่ปราศจากกระบวนการคัดเลือกโดยปัจจัยกระตุ้นจากสิ่งแวดล้อมภายนอก (Positive Selection) ยีนพื้นฐานทำหน้าที่สำคัญในการดำรงชีพซึ่งพบได้ในแบคทีเรียทุกสายพันธุ์ และแบคทีเรียจำเป็นต้องรักษาคุณสมบัติที่สำคัญในการดำรงชีพนั้นไว้ให้คงที่โดยหลีกเลี่ยงให้เกิดการกลายพันธุ์น้อยที่สุด

ในขณะที่ยีนประเภทอื่น เช่น ยีนที่ทำหน้าที่ในการผลิตโปรตีนที่ใช้ในการสร้างผนังเซลล์ อาจได้รับอิทธิพลจากสิ่งแวดล้อมทำให้เกิดการกลายพันธุ์เพื่อปรับตัวในสภาวะที่เปลี่ยนไปได้ง่าย หรือ ยีนที่ควบคุมการผลิตสารพิษ อาจพบเฉพาะในสายพันธุ์ที่ก่อโรค แต่อาจไม่พบในสายพันธุ์ที่ไม่ก่อโรค ดังนั้นยีนพื้นฐานจึงเป็นตัวเลือกที่เหมาะสมในการศึกษาความเกี่ยวข้องทางพันธุกรรมรวมถึงวิวัฒนาการดั้งเดิมซึ่งส่งผลให้ทราบถึงที่มาสายพันธุ์บรรพบุรุษของแบคทีเรียได้ ยีนพื้นฐานที่ใช้ควรมีจำนวนอย่างน้อย 7 ยีนจากแบคทีเรียจำนวนอย่างน้อย 100 ตัว เนื่องจากเป็นจำนวนที่มีความเชื่อถือได้ทางสถิติในการจำแนกสายพันธุ์³ นอกจากนี้ยีนที่คัดเลือกต้องมีการกระจายตัวที่เหมาะสมบนโครโมโซม กล่าวคือไม่ตั้งอยู่ในตำแหน่งที่ติดกันเกินไปเพื่อหลีกเลี่ยงการได้รับอิทธิพลจากการเปลี่ยนแปลงสารพันธุกรรมในบริเวณใกล้เคียง ปัจจัยสำคัญในการออกแบบสารพันธุกรรมต้นแบบที่ใช้เป็น primer ได้แก่ การใช้ primer คู่สมที่สามารถเพิ่มจำนวนยีนที่มีความยาว 400-600 คู่เบสต่อหนึ่งยีน³ ระบบการออกแบบ Primer ที่เหมาะสมสำหรับวิธี MLST คือ Primer ระบบคู่ (Nested Primers) การใช้ Primer ระบบคู่ คือการใช้ Primer ที่ต่างกัน 2 ชุดในการเพิ่มจำนวนชิ้นส่วนของยีนที่ต้องการในปฏิกิริยา PCR และ Primer สำหรับกระบวนการอ่านรหัสพันธุกรรม (Nucleotide Sequencing) การใช้ Primer ระบบคู่ทำให้ผลที่ได้มีความแม่นยำสูงขึ้นเนื่องจากเป็นวิธีที่ลดความเสี่ยงการเกิดความผิดพลาดในการเพิ่มจำนวนชิ้นส่วนของยีนที่ต้องการ เมื่อได้ข้อมูลรหัสพันธุกรรมของยีนพื้นฐานจากแบคทีเรียที่ใช้ศึกษาแต่ละตัวแล้ว จึงเข้าสู่กระบวนการวิเคราะห์รหัสพันธุกรรมต่อไป

5. ผลการศึกษาโดยวิธี Multilocus sequence typing ในการสืบหาแหล่งการติดเชื้อ *Campylobacter jejuni* และ *Vibrio parahaemolyticus*

ผลการศึกษาโดยวิธี MLST มีประโยชน์ในการวิจัยเพื่อสืบหาแหล่งที่มาของเชื้อก่อโรคอุจจาระร่วงที่พบได้ในอาหาร เช่น *Campylobacter jejuni* และ *Vibrio parahaemolyticus* แบคทีเรียทั้งสองเป็นเชื้อก่อโรคในคนที่พบได้ในสัตว์ต่างประเภท *C. jejuni* เป็นเชื้อประจำถิ่นที่พบในสัตว์ปีก เช่น เป็ด ไก่ นก และสัตว์บก เช่น วัว แกะ และสุกร เป็นต้น¹⁷ ส่วน *V. parahaemolyticus* เป็นเชื้อประจำถิ่นที่พบได้ในสัตว์ทะเล¹⁸ โดยเฉพาะสัตว์มีเปลือกแข็ง เช่น กุ้ง ปู และหอย เป็นต้น แบคทีเรียเหล่านี้อาจเจ็บปนมากับอาหารหากกระบวนการผลิตไม่ได้มาตรฐาน ส่งผลให้ผู้บริโภคอาหารนั้นติดเชื้อดังกล่าวและก่อให้เกิดโรคอุจจาระร่วงหรือลำไส้อักเสบ

ในประเทศนิวซีแลนด์ วิธี MLST นำมาใช้ศึกษาวิทยาการระบาดของ *C. jejuni* จากตัวอย่างที่แยกจากอุจจาระผู้ป่วย เนื้อไก่ เนื้อหมู เนื้อวัว และเนื้อแกะ¹⁹⁻²² Wilson และคณะ (2008) ศึกษารหัสพันธุกรรมจาก ST ของเชื้อที่แยกได้จากผู้ป่วย 1,231 ราย พบว่ามีความใกล้เคียงกับ ST ของเชื้อที่แยกได้จากปศุสัตว์ (96.6%) มากกว่าเชื้อที่แยกได้จากสัตว์ป่า (2.3%) และสิ่งแวดล้อม (1.1%) ความเหมือนทางพันธุกรรมของเชื้อจากผู้ป่วยและเชื้อจากปศุสัตว์ประเภทต่างๆเรียงตามลำดับได้แก่ เนื้อไก่ (56.5%) เนื้อวัว (35.0%) และเนื้อแกะ (4.3%) แสดงให้เห็นว่าสายพันธุ์ที่ก่อโรคอุจจาระร่วงในคนพบมากที่สุดไม่ใกล้เคียงกับสัตว์ชนิดอื่น¹⁹ นอกจากนี้ Mullner และคณะ (2009) ได้วิเคราะห์รหัสพันธุกรรมจาก ST ของ *C. jejuni* ทั้งหมด 969 ตัว ซึ่งประกอบด้วย

เชื้อจากผู้ป่วยจำนวน 502 ตัว และเชื้อจากเนื้อสัตว์ชนิดต่างๆอีกจำนวน 467 ตัว พบว่า สายพันธุ์ที่แยกได้จากเนื้อไก่มี ST ที่ใกล้เคียงกับสายพันธุ์ที่แยกได้จากผู้ป่วยมากที่สุดเมื่อเทียบกับเนื้อสัตว์ชนิดอื่น โดยคิดเป็น 58-76% ของสายพันธุ์ทั้งหมดที่เป็นสาเหตุของโรคอุจจาระร่วงในผู้ป่วย การศึกษานี้สนับสนุนว่าการปนเปื้อนของเชื้อ *C. jejuni* ในเนื้อไก่น่าจะเป็นสาเหตุหลักของการเกิดโรคอุจจาระร่วงในคน²⁰ (รูปที่ 3) ข้อสรุปที่ได้จากงานวิจัยเหล่านี้นำไปสู่การกำหนดนโยบายมุ่งเน้นการควบคุมรักษาความปลอดภัยในกระบวนการผลิตเนื้อไก่ภายในประเทศนิวซีแลนด์ให้มีความเข้มงวดมากขึ้น ส่งผลให้อัตราการติดเชื้อจากเดิม 15,873 รายต่อปี ในปี ค.ศ. 2006 ลดลงเป็น 6,689 รายต่อปี ในปี ค.ศ. 2008^{20,23} นอกจากนี้วิธี MLST สามารถใช้ระบุสายพันธุ์ที่มีความสำคัญต่อการระบาดของเชื้อ *C. jejuni* ซึ่งได้แก่ สายพันธุ์ ST-474 ถึงแม้พบว่า สายพันธุ์นี้มีความชุกสูงในโรงงานผลิตเนื้อไก่แห่งหนึ่งในประเทศนิวซีแลนด์ แต่กลับพบได้น้อยในประเทศอื่นและเมื่อนำข้อมูลมาเปรียบเทียบกับฐานข้อมูล MLST ของเชื้อ *C. jejuni* ทางอินเทอร์เน็ต จึงสรุปได้ว่า ST-474 เป็นสายพันธุ์ที่มีการระบาดเฉพาะพื้นที่ และอาจมีวิวัฒนาการเพื่อความอยู่รอดที่จำเพาะต่อสิ่งแวดล้อมในพื้นที่นั้น²¹ อย่างไรก็ตาม นักวิจัยจากประเทศฟินแลนด์พบว่าเชื้อ *C. jejuni* บางกลุ่มที่แยกได้จากผู้ป่วยมี ST ที่ไม่พบในเชื้อที่แยกได้จากสัตว์ปีกและสัตว์บก²⁴ ดังนั้นอาจมีความเป็นไปได้ว่าการติดเชื้อ *C. jejuni* ในบริเวณที่ศึกษาอาจมาจากแหล่งที่มาอื่นนอกเหนือจากการบริโภคอาหารที่ปนเปื้อน ต่อมา French และคณะ (2009) พบว่า เชื้อ *C. jejuni* ที่มี ST ตรงกับสายพันธุ์ก่อโรคในคนถูกแยกได้จากมูลนกที่พบในสนามเด็กเล่น²⁵ ทำให้ทราบว่าการติดเชื้อชนิดนี้

อาจจะไม่ได้เกิดขึ้นจากการบริโภคอาหารเท่านั้น การประเมินความเสี่ยงในการติดเชื้อจากการปนเปื้อนจากแหล่งกลีกรม แหล่งน้ำ หรือภายในที่อยู่อาศัย เช่น สัตว์เลี้ยง เป็นสิ่งที่ควรระวังสำหรับการติดเชื้อชนิดนี้เช่นกัน นอกจากนี้วิธี MLST ถูกนำมาใช้ศึกษา ลักษณะทางพันธุกรรมของเชื้อ *C. jejuni* ในเนื้อไก่จากบริษัทผู้ผลิตหลากหลายบริษัท พบว่าสายพันธุ์ที่แยกได้จากเนื้อไก่ของแต่ละบริษัทมี ST แตกต่างกัน

โดยแยกกลุ่มตามบริษัทที่เป็นแหล่งที่มาของเชื้อนั้น²⁰ ผลจากการศึกษานี้นำไปสู่การทำความเข้าใจเรื่อง วิวัฒนาการความอยู่รอดของเชื้อ *C. jejuni* ที่อาจจะต่างกันตามสภาพแวดล้อมของแต่ละบริษัท ส่งผลให้เกิดการตรวจสอบขั้นตอนการผลิตของแต่ละบริษัท เพื่อรับมาตรการหรือออกข้อบังคับในกระบวนการผลิตให้ได้มาตรฐานความปลอดภัยและลดอัตราการปนเปื้อนของเชื้อก่อโรคได้มีประสิทธิภาพมากขึ้น

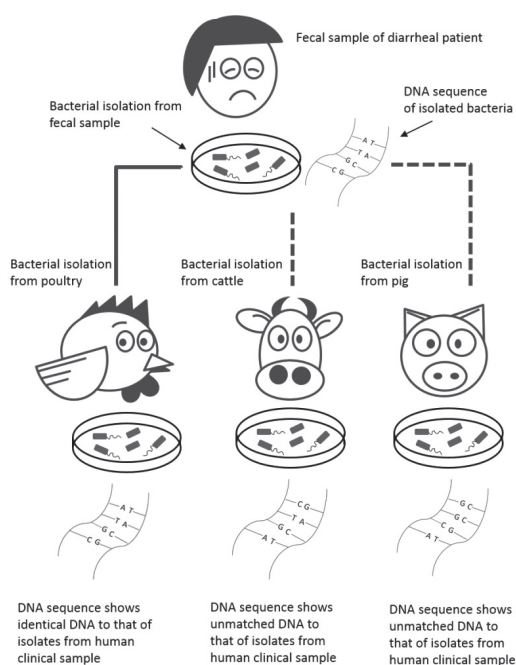


Figure 3 Process for Identifying Source of Infection of *C. jejuni* by MLST. Clinical isolates were recovered from patients, as well as other isolates were recovered from poultry, cattle and pigs. DNA extraction and MLST analysis were subsequently performed using these isolates and the genetic relationships among the strains were determined. Poultry was identified as a source of infection for human clinical strains because the MLST analysis revealed that strains recovered from poultry represent identical sequence type (ST) to those from clinical samples. Solid line represents the most genetic similarity of isolates from human clinical samples with that of poultry, and dash line represents less genetic similarity of isolates from human clinical samples with that of cattle and pig.¹⁹⁻²²

อย่างไรก็ตาม ผลวิเคราะห์โดยวิธี MLST ขึ้นอยู่กับประเภทของเชื้อที่ใช้ศึกษา ในกรณีแบคทีเรีย *V. parahaemolyticus* วิธี MLST ให้ผลวิเคราะห์ต่างจากเชื้อ *C. jejuni* การศึกษาเชื้อ *V. parahaemolyticus* ที่แยกได้จากผู้ป่วย อาหารทะเล และสิ่งแวดล้อม ไม่พบสายพันธุ์ก่อโรคที่มี ST เดียวกันในเชื้อที่แยกได้จากอาหารทะเล หรือสิ่งแวดล้อมที่นำมาศึกษา²⁶⁻²⁸ แสดงให้เห็นว่า เชื้อที่แยกได้จากผู้ป่วยมีลักษณะทางพันธุกรรมต่างจากเชื้อที่แยกจากอาหารทะเล หรือสิ่งแวดล้อม ข้อมูลเหล่านี้นำไปสู่การศึกษาเพื่อหาแหล่งที่มาของสายพันธุ์ก่อโรคที่แท้จริงหรือปัจจัยกระตุ้นที่มีผลต่อการอุบัติใหม่และอุบัติซ้ำของสายพันธุ์ก่อโรคต่อไป อย่างไรก็ตาม MLST มีประโยชน์สำหรับ *V. parahaemolyticus* กรณีศึกษาเพื่อระบุความสัมพันธ์ของเชื้อก่อโรคจากผู้ป่วยที่แยกได้ภายในประเทศและต่างประเทศ ทำให้เข้าใจที่มาของอุบัติการณ์การเกิดโรคระบาดโดยสายพันธุ์เดียวกันจากผู้ป่วยต่างทวีป²⁹⁻³¹

วิธี MLST มีประโยชน์ในการติดตามแหล่งที่มาของเชื้อก่อโรคอุจจาระร่วงโดยการจำแนกสายพันธุ์แบคทีเรีย เช่น *C. jejuni* ที่พบในตัวอย่างอุจจาระของผู้ป่วย อาหาร และสิ่งแวดล้อม ความแตกต่างของรหัสพันธุกรรมที่ได้มาจาก ST ของแต่ละสายพันธุ์สามารถนำมาคำนวณเป็นเปอร์เซ็นต์ความต่างเพื่อช่วยในการวิเคราะห์เชิงปริมาณทำให้มีความแม่นยำมากกว่า ข้อมูลที่ได้จากวิธีทางโมเลกุลประเภทอื่น ข้อสำคัญอีกประการหนึ่งคือสามารถเปรียบเทียบ ST ของเชื้อที่ศึกษากับ ST ของเชื้อที่แยกโดยคณะผู้วิจัยอื่นที่มีอยู่ในฐานข้อมูลทางอินเทอร์เน็ตเพื่อวิเคราะห์เชื่อมโยงเส้นทางการระบาดของเชื่อนั้น และสามารถศึกษาวิวัฒนาการของเชื้อที่แยกได้ภายในประเทศเมื่อเทียบ

กับเชื้อจากต่างประเทศได้ตามที่ยกตัวอย่างในกรณีเชื้อ *V. parahaemolyticus*

สรุป

เทคโนโลยีทางวิทยาศาสตร์สามารถนำมาประยุกต์ในการพัฒนาคุณภาพและความปลอดภัยของอาหารได้ การศึกษาความสัมพันธ์เชิงพันธุศาสตร์ของแบคทีเรียก่อโรคโดยวิธี MLST ทำให้เข้าใจถึงวิทยาการระบาดและแหล่งที่มาของการปนเปื้อนเชื้อก่อโรคจากอาหาร ซึ่งสามารถนำไปใช้ในการปรับปรุงกระบวนการผลิตอาหารเพื่อลดความเสี่ยงของการติดเชื้อก่อโรคในผู้บริโภค อย่างไรก็ตามการนำความรู้จากวิธีดังกล่าวไปใช้ประโยชน์จำเป็นต้องอาศัยความร่วมมือและความเชี่ยวชาญหลายด้าน ได้แก่ นักวิจัยผู้มีความรู้ด้านการสกัดและเพิ่มจำนวน DNA การวิเคราะห์รหัสพันธุกรรม บุคลากรที่มีความเข้าใจกระบวนการผลิตอาหารและสามารถเข้าถึงบริเวณเก็บตัวอย่างสำหรับการแยกแบคทีเรียได้ รวมถึงความเข้าใจสภาพแวดล้อมและวัฒนธรรมการรับประทานอาหารของคนในชุมชนนั้นจึงจะสามารถประยุกต์ความรู้ดังกล่าวในทางปฏิบัติเพื่อการลดอัตราการเกิดโรคติดเชื้อทางเดินอาหารได้อย่างมีประสิทธิภาพ

เอกสารอ้างอิง

1. Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing-A portable approach to the identification of clones within populations of clones within populations of pathogenic microorganisms. *Microbiology* 1998; 95: 3140-5.

2. Selander RK, Caugant DA, Ochman H, Musser JM, Gilmour MN, Whittam TS. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Appl Environ Microbiol* 1986; 51: 873-84.
3. Maiden MC. Multilocus sequence typing of bacteria. *Annu Rev Microbiol* 2006; 60: 561-88.
4. Spratt BG, Maiden MC. Bacterial population genetics, evolution and epidemiology. *Phil Trans R Soc Lond* 1999; 354: 701-10.
5. Feil EJ. Small change: keeping pace with microevolution. *Nat Rev Microbiol* 2004; 2: 483-95.
6. Cooper JE, Feil EJ. Multilocus sequence typing-what is resolved? *Trends Microbiol* 2004; 12: 373-7.
7. Thongprasert J, Vatanasomboon P, Luksamijarulkul P, Singhakant C. Related Factors and Contamination of Bacterial Indicators in Sushi from Shop Indoor and Outdoor Areas. *Journal of Public Health* 2015; 45: 244-55.
8. Perathornnich T, Sujirarat D, Sreesai S. The Analysis of Factors Influencing Food Sanitation in Phitsanulok Municipality, Phitsanulok Provinces. *Journal of Public Health* 2015; 45: 230-43.
9. Buckley M, Reid A. Global food safety: Keeping food safe from farm to table. *Am Acad Microbiol* 2010; 41pp. Available at http://academy.asm.org/images/stories/documents/Global_Food_Safety.pdf, accessed October 20, 2016.
10. World Health Organization, Food and Agriculture Organization of the United Nations. Risk assessment of *Vibrio parahaemolyticus* in seafood. Microbiological risk assessment series 2011; 16: 193pp. Available at <http://www.fao.org/3/a-i2225e.pdf>, accessed October 20, 2016.
11. Adzitey F, Huda N, Ali GRR. Molecular techniques for detecting and typing of bacteria, advantages and application to foodborne pathogens isolated from ducks. *3 Biotech* 2012; 3: 97-107.
12. Olive DM, Bean P. Principles and Applications of methods for DNA-based typing of microbial organisms *J Clin Microbiol* 1999; 37: 1661-9.
13. Foley SL, Lynne AM, Nayak R. Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. *Infect Genet Evol* 2009; 9: 430-40.
14. Goering RV. Pulsed field gel electrophoresis: A review of application and interpretation in the molecular epidemiology of infectious disease. *Infect Genet Evol* 2010; 10: 866-75.

15. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* 2004; 186: 1518-30.
16. Urwin R, Maiden MCJ. Multi-locus sequence typing-a tool for global epidemiology. *Trends Microbiol* 2003; 11: 479-87.
17. Young KT, Davis LM, Dirita VJ. *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat Rev Microbiol* 2007; 5: 665-79.
18. Nelapati S, Nelapati K, Chinnam BK. *Vibrio parahaemolyticus* An emerging foodborne pathogen-A Review. *Vet World* 2012; 5: 48-62.
19. Wilson DJ, Gabriel E, Leatherbarrow AJH, Cheesbrough J, Gee S, Bolton E, et al. Tracing the source of campylobacteriosis. *PLoS Genet* 2008; 4: e1000203.
20. Mullner P, Spencer SEF, Wilson DJ, Jones G, Noble AD, Midwinter AC, et al. Assigning the source of human campylobacteriosis in New Zealand: a comparative genetic and epidemiological approach. *Infect Genet Evol* 2009; 9: 1311-9.
21. Mullner P, Collins-Emerson JM, Midwinter AC, Carter P, Spencer SEF, van der Logt P, et al. Molecular epidemiology of *Campylobacter jejuni* in a geographically isolated country with a uniquely structured poultry industry. *Appl Environ Microbiol* 2010; 76: 2145-54.
22. Strachan NJC, Gormley FJ, Rotariu O, Ogden ID, Miller G, Dunn GM, et al. Attribution of *Campylobacter* infections in northeast Scotland to specific sources by use of multilocus sequence typing. *J Infect Dis* 2009; 199: 1205-8.
23. Ministry of Primary Industry. *Campylobacter* Risk Management Strategy 2013-2014. Available at http://www.foodsafety.govt.nz/elibrary/industry/Campylobacter__Risk-Comprehensive__Aimed.pdf, accessed October 20, 2016.
24. De Haan CP, Kivistö RI, Hakkinen M, Corander J, Hänninen ML. Multilocus sequence types of Finnish bovine *Campylobacter jejuni* isolates and their attribution to human infections. *BMC Microbiol* 2010; 10: 200.
25. French NP, Midwinter A, Holland B, Collins-Emerson J, Pattison R, Colles F, et al. Molecular epidemiology of *Campylobacter jejuni* isolates from wild-bird fecal material in children's playgrounds. *Appl Environ Microbiol* 2009; 75: 779-83.



26. Theethakaew C, Feil EJ, Castillo-Ramírez S, Aanensen DM, Suthienkul O, Neil DM, et al. Genetic relationships of *Vibrio parahaemolyticus* isolates from clinical, human carrier, and environmental sources in Thailand, determined by multilocus sequence analysis. *Appl Environ Microbiol* 2013; 79: 2358-70.
27. Gonzalez-Escalona N, Martinez-Urtaza J, Romero J, Espejo RT, Jaykus LA, DePaola A. Determination of molecular phylogenetics of *Vibrio parahaemolyticus* strains by multilocus sequence typing. *J Bacteriol* 2008; 190: 2831-40.
28. Yu Y, Hu W, Wu B, Zhang P, Chen J, Wang S, et al. *Vibrio parahaemolyticus* isolates from southeastern Chinese coast are genetically diverse with circulation of clonal complex 3 strains since 2002. *Foodborne Pathog Dis* 2011; 8: 1169-76.
29. Gavilan RG, Zamudio ML, Martinez-Urtaza J. Molecular epidemiology and genetic variation of pathogenic *Vibrio parahaemolyticus* in Peru. *PLoS Negl Trop Dis* 2013; 7: e2210.
30. Urmsersbach S, Alter T, Sanjeevani M, Koralage G, Sperling L, Gerdts G. Population analysis of *Vibrio parahaemolyticus* originating from different geographical regions demonstrates a high genetic diversity. *BMC Microbiol* 2014; 14: 59.
31. Han C, Tang H, Ren C, Zhu X, Han D. Sero-prevalence and genetic diversity of pandemic *V. parahaemolyticus* strains occurring at a global scale datasets utilized in the present study. *Front Microbiol* 2016; 7: 567.

Multilocus Sequence Typing (MLST): Molecular Technique for Monitoring Food Safety

Chonchanok Theethakaew*

ABSTRACT

Foodborne infectious diseases are a worldwide public health burden particularly in tropical countries. Determining potential sources of infection has mitigated transmission of bacteria. Sources of bacterial enteric infection could be identified by genetic relatedness of strains isolated from patients with diarrhea and strains isolated from various sources including food. Multilocus sequence typing (MLST) is a genetic analysis used to identify bacterial strains based on nucleotide sequences of housekeeping genes of bacteria collected from various sources. MLST analysis allows understanding of genetic relationships and sources of bacteria transmission. The advantage

of MLST over other molecular methods is the ability to compare genetic profiles of examined strains from individual laboratories to those from over the world through online database. At present, the MLST approach has been successfully applied in epidemiological and molecular evolutionary studies of several foodborne pathogenic bacteria. Generated outcomes from MLST could improve surveillance of bacterial pathogens in food industry and generate higher standard of food safety.

Keywords: Multilocus sequence typing, food-borne pathogens, genetic relationships, molecular epidemiology