Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR): ระบบภูมิคุ้มกันแบบจำเพาะของแบคทีเรีย

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR): The Specific Immune System of Bacteria

มณฑล เลิศวรปรีชา[1,2]* และ ศุภารัตน์ สุทธิมุสิก[2]

Monthon Lertworapreecha[1,2]* and Suparat Sutthimusik[2]

[1]*หน่วยวิจัยการจัดการทรัพยากรจุลินทรีย์ สาขาชีววิทยา คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ (วิทยาเขตพัทลุง)*

[2]*สาขาชีววิทยา คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ (วิทยาเขตพัทลุง)*

[1]*Microbial Resource Management Research Unit, Department of Biology, Faculty of Science, Thaksin University (Phatthalung campus)*

[2] *Department of Biology, Faculty of Science, Thaksin University (Phatthalung campus)*

บทคัดย่อ

แบคทีเรียมีกลไกที่ใช้สำหรับป้องกันตัวเองจากการบุกรุกของสารพันธุกรรมแปลกปลอม เช่น แบคเทอริโอฟาจ และพลาสมิด กลไกที่สำคัญอย่างหนึ่งคือ ระบบ Clustered Regularly Interspaced Short Palindromic Repeats และโปรตีน Cas (CRISPR-Cas system) ซึ่งการทำงานของระบบ CRISPR-Cas เปรียบได้กับระบบภูมิคุ้มกันแบบจำเพาะของแบคทีเรีย กลไกการทำงานคือแบคทีเรียสามารถนำเอาชิ้นส่วนของสารพันธุกรรมของแบคเทอริโอฟาจ หรือพลาสมิดที่บุกรุกเซลล์เข้าสู่โครโมโซมของตัวเอง ในบริเวณของ CRISPR ทำให้แบคทีเรียนั้นสามารถจดจำแบคเทอริโอฟาจหรือพลาสมิดที่เคยบุกรุกเข้ามาได้ ดังนั้นเมื่อแบคเทอริโอฟาจหรือพลาสมิดชนิดเดิมบุกรุกกลับเข้ามาอีกในครั้งต่อไป แบคทีเรียจะสร้าง RNA ที่เรียกว่า CRISPR–RNA หรือ crRNA ซึ่งมีความจำเพาะกับแบคเทอริโอฟาจ หรือพลาสมิดนั้น และ crRNA ที่สร้างขึ้นจะทำงานร่วมกับโปรตีน Cas ในการทำลายสารพันธุกรรมของแบคเทอริโอฟาจ หรือพลาสมิดนั้น ๆ ปัจจุบันความรู้เรื่องระบบ CRISPR-Cas ได้นำไปประยุกต์ใช้ในงานด้านวิทยาศาสตร์การแพทย์ และเทคโนโลยีชีวภาพ เช่น การพัฒนา CRISPR เพื่อใช้เป็นโมเลกุลเครื่องหมายสำหรับการศึกษาทางระบาดวิทยาของแบคทีเรียก่อโรค การพัฒนาสายพันธุ์แบคทีเรียที่สามารถทนต่อการบุกรุกของแบคเทอริโอฟาจและการใช้เป็นเครื่องมือสำหรับการแก้ไขปรับปรุงพันธุกรรมระดับจีโนมในสัตว์และพืช

คำสำคัญ : CRISPR ยีน Cas ภูมิคุ้มกันของแบคทีเรีย

*Corresponding author. E-mail : worapreecha@gmail.com

Abstract

Bacteria have a defense mechanism to protect themselves from invading bacteriophage or plasmid. The importance mechanism is the Clustered Regularly Interspaced Short Palindromic Repeats, and Cas protein (CRISPR-Cas system), which working as specific immune system of bacterial cell. The mechanism involved with the ability to incorporate bacteriophage or plasmid DNA into bacterial chromosome, which act as a memory card of bacteria to recognize of previously invaded bacteriophage or plasmid. Therefore, when the bacterial cell are invaded again by the previously known bacteriophage or plasmid that will activates the CRISPR and Cas system to create a small interference RNA called “crRNA”. This crRNA are worked together with Cas protein to destroy the bacteriophage and plasmid DNA. Nowadays, the knowledge of the fascinating CRISPR-Cas system have applied to biomedical and biotechnology. For example, developing as a molecular marker for epidemiological study of pathogenic bacteria, constructing the bacterial strains that resistant to bacteriophage infection, and using as a tool for genome engineering in animal and plant.

Keywords : CRISPR, Cas gene, adaptive immune in Bacteria

บทนำ

สิ่งมีชีวิตทุกชนิดล้วนมีศัตรูตามธรรมชาติ ในแบคทีเรียก็เช่นกัน ศัตรูตามธรรมชาติของแบคทีเรีย ก็คือ แบคเทอริโอฟาจ หรือ “ฟาจ” ซึ่งเป็นไวรัสชนิดหนึ่ง มีความหลากหลายของชนิดและสายพันธุ์ที่สูงมากพบได้ในสิ่งแวดล้อมที่หลากหลาย เมื่อฟาจบุกรุกเข้าสู่เซลล์แบคทีเรีย ฟาจจะเพิ่มจำนวนมากขึ้นโดยอาศัยกลไกภายในของเซลล์แบคทีเรีย เพื่อสร้างองค์ประกอบต่าง ๆ ของฟาจ และทำลายแบคทีเรียให้เซลล์แตกสลายไป อย่างไรก็ตามแบคทีเรียก็มีกลไกที่สามารถป้องกันตัวเองจากฟาจ และสารพันธุกรรมแปลกปลอมที่บุกรุกเข้ามา เช่นการเปลี่ยนแปลงโมเลกุลที่เป็นตัวรับ (receptor) บนผิวเซลล์ ป้องกันการยึดเกาะของฟาจ และฉีดสารพันธุกรรมเข้าสู่เซลล์ อีกกลไกหนึ่งซึ่งเป็นที่รู้จักดี คือ เอนไซม์ตัดจำเพาะ (restriction endonuclease enzymes) ซึ่งจะตัดทำลาย DNA แปลกปลอมที่บุกรุกเข้าสู่เซลล์ ในขณะที่แบคทีเรียจะป้องกัน DNA ของตนเองจากเอนไซม์ตัดจำเพาะด้วยการดัดแปลงโมเลกุลของ DNA ตัวเองโดยเติมหมู่ methyl ไปที่ตำแหน่งจำเพาะที่เอนไซม์ตัด เรียกกระบวนการนี้ว่า methylation (Pingoud *et al*., 2005) กลไกนี้เปรียบเสมือนเป็นระบบภูมิคุ้มกันที่มีมาแต่กำเนิด (innate immune system) ที่พบได้ในสัตว์มีกระดูกสันหลังทั่วไป และเมื่อเร็ว ๆ นี้ มีการค้นพบกลไกอีกรูปแบบหนึ่ง คือ Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)

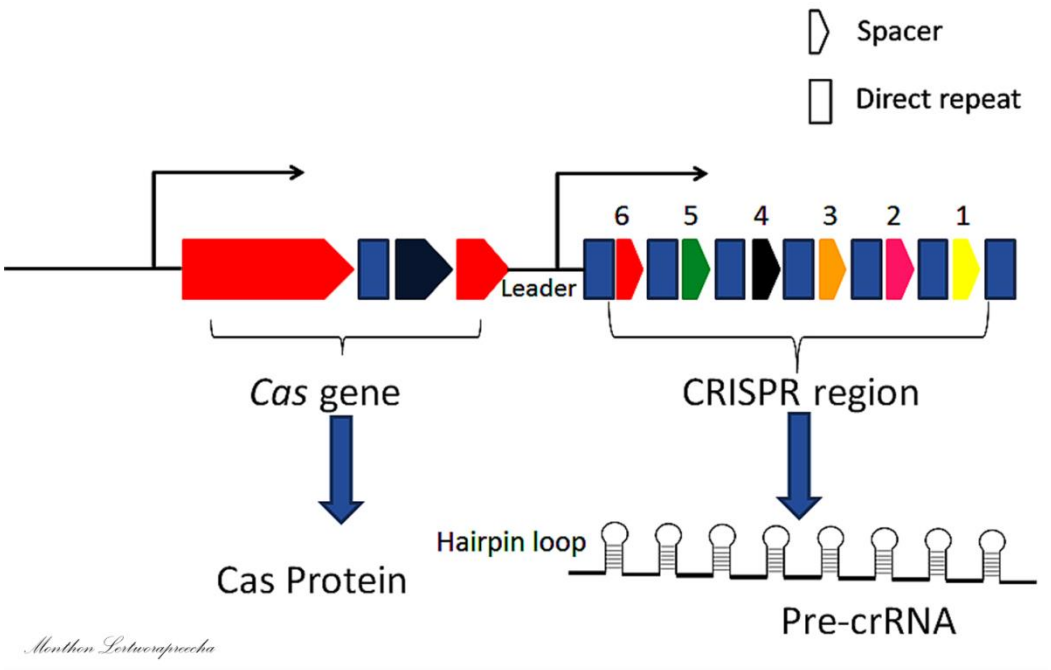
CRISPR มีรายงานครั้งแรกในเชื้อ *E.coli* ในปี 1987 (Ishino *et al*., 1987) การศึกษาต่อมาพบว่าบริเวณดังกล่าวเกี่ยวข้องกับกระบวนการป้องกันตัวเองของแบคทีเรีย โดยเปรียบได้กับภูมิคุ้มกันแบบจำเพาะ (specific or adaptive immune system) ที่พบในสัตว์ชั้นสูง CRISPR เป็นตำแหน่งของนิวคลีโอไทด์ที่จัดเรียงตัวแบบพิเศษแบ่งเป็นสองบริเวณ โดยบริเวณแรกเป็นนิวคลีโอไทด์ที่มีขนาดสั้น ๆ เป็นบริเวณอนุรักษ์ (conserve) ที่มีการเรียงตัวซ้ำ ๆ กัน เรียกว่า direct repeat และพบว่าลำดับนิวคลีโอไทด์ของ direct repeat นั้นมีลักษณะเป็น palindrome ทำให้สาย RNA ที่ถอดรหัสจากบริเวณของ direct repeat นั้นสามารถจับกันเองเป็นโครงสร้างทุติยภูมิในลักษณะของ hairpin loop ได้ บริเวณที่สองเป็นนิวคลีโอไทด์ที่มีความหลากหลาย (variable) ไม่ซ้ำกัน เรียกว่า spacer โดย spacer มีขนาดประมาณ

24-28 คู่เบส จะแทรกตัวอยู่ระหว่าง direct repeat มีลักษณะดังนี้ คือ repeat—spacer—repeat (ภาพที่ 1) พบว่าส่วนที่เป็น spacer มีลำดับนิวคลีโอไทด์เหมือนกันกับลำดับนิวคลีโอไทด์บางส่วนของฟาจ และ พลาสมิดบางชนิดที่แบคทีเรียนั้นเคยได้รับหรือถูกบุกรุกมาก่อน (Bolotin *et al*., 2005; Godde & Bickerton, 2006) และพบว่าแบคทีเรียมีกลไกที่นำเอาชิ้นส่วน DNA บางส่วนของฟาจหรือ พลาสมิดเข้าไปแทรกอยู่ในบริเวณ direct repeat โดยชิ้นส่วน DNA ที่นำเข้าไปเปรียบได้กับเป็นหน่วยความจำ (memory card) ที่แบคทีเรียใช้จดจำชนิดของฟาจที่เคยบุกรุกเข้ามา เมื่อมีการบุกรุกของฟาจหรือได้รับพลาสมิดชนิดเดิมกลับเข้ามา กลไกของ CRISPR จะสร้าง CRISPR–RNA (crRNA) ออกมาและไปทำลายฟาจหรือ พลาสมิดนั้นอย่างจำเพาะ ซึ่งกลไกลดังกล่าวคล้ายกับ RNAi ที่พบใน eukaryotic cells (Wiedenheft *et al*., 2012)

ความรู้เรื่อง CRISPR นับว่ายังคงเป็นเรื่องที่ค่อนข้างใหม่ นักวิทยาศาสตร์มีความเข้าใจกลไกการทำงานของ CRISPR ได้ไม่นานนัก ปัจจุบัน CRISPR เป็นประเด็นหนึ่งที่นักวิทยาศาสตร์จำนวนมากให้ความสนใจศึกษาเพื่อที่จะเข้าใจกลไกการทำงานให้มากขึ้นและหาแนวทางการประยุกต์ใช้ประโยชน์จาก CRISPR ในบทความนี้ได้รวบรวมความรู้พื้นฐานในเรื่องดังกล่าว และแนวทางการวิจัยรวมถึงนำไปประยุกต์ใช้ เพื่อเป็นแนวทางสำหรับผู้สนใจได้ศึกษาต่อไป

การค้นพบ CRISPR และสมมุติฐานหน้าที่ของ CRISPR

CRISPR มีรายงานครั้งแรกโดย Nakata และคณะ ในปี 1987 ได้ศึกษาลำดับนิวคลีโอไทด์ของยีน *iap* ซึ่งเป็นยีนควบคุมการสร้างเอนไซม์ alkaline phosphatase พบว่าลำดับนิวคลีโอไทด์บางส่วนมีลักษณะการจัดเรียงตัวแบบพิเศษอยู่ที่บริเวณด้านข้าง (flanking region) ไปทางด้านปลาย 3' ของยีน *iap* โดยพบว่าบริเวณดังกล่าวมีลำดับนิวคลีโอไทด์ที่เป็นบริเวณอนุรักษ์ขนาดประมาณ 29 นิวคลีโอไทด์จัดเรียงตัวตัวแบบ direct repeat สลับกับบริเวณที่มีความหลากหลายประมาณ 32 นิวคลีโอไทด์ โดยมีการเรียงตัวในแบบดังกล่าวสลับกันประมาณ 5 ชุด แต่ยังไม่ทราบว่าลำดับนิวคลีโอไทด์ดังกล่าวทำหน้าที่อย่างไร (Ishino *et al*., 1987) หลังจากนั้นก็มีรายงานการค้นพบบริเวณ DNA ที่มีลักษณะดังกล่าวในเชื้อแบคทีเรียและอาร์เคีย (archaea) อื่น ๆ อีกเป็นจำนวนมาก (Hermans *et al*., 1991; Groenen *et al*., 1993; Bult *et al*., 1996; Horvath *et al*., 2008)



ภาพที่ 1 ลักษณะการจัดเรียงตัวของลำดับนิวคลีโอไทด์บริเวณ *Cas* gene และ CRISPR; บริเวณของ CRISPR ประกอบด้วยนิวคลีโอไทด์สองส่วนคือ direct repeat และ spacer โดยส่วน CRISPR จะถูกถอดรหัสเป็น pre-crRNA และจะถูกตัดด้วยโปรตีน Cas เพื่อให้ได้เป็น crRNA

การทดลองที่สนับสนุนสมมุติฐานที่ว่า CRISPR อาจเกี่ยวข้องกับการจดจำ และป้องกันการบุกรุกของฟาจนั้นมาจากรายงานการศึกษาโดย 3 ทีมวิจัยที่ไม่มีความเกี่ยวข้องกัน แต่รายงานผลการศึกษาที่มีความสอดคล้องกัน คือพบว่าลำดับนิวคลีโอไทด์ในบริเวณ spacer ที่แทรกอยู่ระหว่าง direct repeat นั้นมีส่วนที่คล้ายกับนิวคลีโอไทด์บางส่วนในฟาจหรือพลาสมิดบางชนิด (Bolotin *et al*., 2005; Mojica *et al*., 2005; Pourcel *et al*., 2005) นอกจากนี้รายงานการศึกษาในเชื้อ *Streptococcus thermophilus* พบว่าความไวของแบคทีเรียต่อการถูกฟาจบุกรุกนั้นสัมพันธ์กับจำนวนซ้ำของ spacer ที่พบใน CRISPR โดยเชื้อที่มีจำนวนซ้ำของ spacer น้อยจะมีแนวโน้มที่จะถูกฟาจบุกรุกได้ง่ายกว่าเชื้อที่ตรวจพบว่ามีจำนวนซ้ำของ spacer มากกว่า (Bolotin *et al*., 2005) การทดลองที่สนับสนุนสมมุติฐานอีกเรื่องหนึ่งคือการศึกษาในเชื้อ *Staphylococcus epidermidis* พบว่าบริเวณของ spacer ของเชื้อ *S. epidermidis* บางสายพันธุ์มีลำดับนิวคลีโอไทด์ที่เหมือนกับบางส่วนของยีน nickase (*nes*) ที่อยู่ในคอนจูเกทีพพลาสมิด (conjugative plasmid) ในเชื้อ *S. epidermidis* การทดลองนั้นพบว่าหากเปลี่ยนแปลงลำดับนิวคลีโอไทด์บริเวณของยีน nickase ที่พบบนพลาสมิดและนำกลับเข้าไปใส่ในเชื้อ *S. epidermidis* ทั้งสายพันธุ์ที่มีและไม่ spacer ต่อยีน nickase ผลปรากฏว่าพลาสมิดที่มีการเปลี่ยนแปลงลำดับนิวคลีโอไทด์ของยีน nickase สามารถเข้าสู่เชื้อ *S. epidermidis* ทั้งสายพันธุ์ที่มีและไม่ spacer ต่อยีน nickase ได้ ในขณะที่พลาสมิดที่มียีน nickase ปกติจะไม่สามารถเข้าสู่เชื้อ *S. epidermidis* ที่มี spacer ต่อยีน nickase ได้ แต่จะเข้าสู่ *S. epidermidis* สายพันธุ์ที่ไม่มี CRISPR ได้เป็นปกติ ซึ่งพิสูจน์ให้เห็นว่า spacer ที่พบนั้นจดจำลำดับนิวคลีโอไทด์ของพลาสมิด และเกี่ยวข้องกับการป้องกันการบุกรุกของพลาสมิด (Marraffini & Sontheimer, 2008) การศึกษาในรายงานอื่น ๆ ก็มีผลที่สอดคล้องกัน เช่นกรณีศึกษาในเชื้อ *S. pyogenes* ซึ่งเป็นแบคทีเรียที่มีระดับความรุนแรงของการก่อโรคแตกต่างกัน โดยคุณสมบัติการก่อโรคนั้นขึ้นอยู่กับการได้รับยีนก่อโรค

(virulence genes) จากฟาจ โดยฟาจจะสอดแทรก DNA ของตัวเองเข้าสู่โครโมโซมของแบคทีเรีย และพบว่า CRISPR สามารถจำกัดการแทรก DNA ของฟาจเข้าสู่โครโมโซมของแบคทีเรียได้ ซึ่งในการศึกษานั้นได้ตั้งข้อสันนิษฐานไว้ว่า ข้อมูลในส่วน spacer ที่มีอยู่อย่างจำกัดหรือแตกต่างกันอาจจะเกี่ยวข้องกับการก่อโรคที่แตกต่างกันในเชื้อ *S. pyogenes* (Nozawa *et al*., 2011) และเช่นเดียวกันกับการศึกษาในเชื้อ *S. thermophilus* ซึ่งพบว่าเชื้อที่สามารถนำชิ้นส่วนของพลาสมิดที่มียีนดื้อยาเข้าสู่ CRISPR จะมีความสามารถทำลายพลาสมิดชนิดเดียวกันที่อยู่ภายในเซลล์ได้ (Garneau *et al*., 2010) จากผลการศึกษาในหลาย ๆ รายงาน และในแบคทีเรียที่แตกต่างกันให้ผลที่สอดคล้องกัน ทำให้นักวิทยาศาสตร์มีความเข้าใจบทบาทหน้าที่ของ CRISPR มากขึ้นและนำไปสู่ข้อสรุปว่า CRISPR มีความเกี่ยวข้องกับระบบการป้องกันการบุกรุกของสารพันธุกรรมแปลกปลอม

กลไกการทำงานของ CRISPR

กระบวนการทำงานของ CRISPR นั้นแบ่งออกได้เป็น 3 ขั้นตอน ขั้นตอนแรกเรียกว่า “adaptation” หรือ “immunization” เป็นกระบวนการนำชิ้นส่วน DNA ของฟาจหรือพลาสมิดเข้าสู่ในบริเวณ CRISPR กระบวนการนี้อาศัยโปรตีน Cas 1 และ 2 โดยโปรตีน Cas จะเป็นตัวรับรู้ลำดับนิวคลีโอไทด์ที่จำเพาะของฟาจ เรียกว่า Protospacer Associated Motif (PAM) ซึ่งอยู่ใกล้กับบริเวณ protospacer ที่จะถูกตัดและนำเข้าไปต่อใน CRISPR ทางด้านปลายของ leader เสมอ (ภาพที่ 1) ดังนั้นลำดับของ spacer ที่พบจึงสามารถบ่งบอกถึงลำดับก่อนหลังของการได้รับ spacer นั้นได้ ขั้นตอนที่สอง คือ “CRISPR expression” เป็นการแสดงออกของ CRISPR โดยลำดับนิวคลีโอไทด์ทั้งหมดของ CRISPR จะถูกถอดรหัส (transcription) เป็น RNA เรียกว่า pre-crRNA จากนั้น pre-crRNA จะถูกโปรตีน Cas บางตัวที่มีคุณสมบัติเป็นเอนไซม์ endoribonucleases ที่จำเพาะตัดให้เป็น small CRISPR RNA เรียกว่า crRNA ซึ่งส่วน crRNA จะเป็นบริเวณ spacer ที่มีความจำเพาะกับ DNA ของฟาจ crRNA จับกับโปรตีน Cas และนำพาโปรตีน Cas เข้าไปจับกับ DNA ของฟาจ และตัดทำลายสาย DNA ของฟาจนั้น เรียกขั้นตอนสุดท้ายนี้ว่า “CRISPR interference” เนื่องจาก crRNA ทำหน้าที่เป็น silencing RNA จึงเรียกอีกชื่อหนึ่งว่า prokaryotic silencing RNA (psiRNA)

ความหลากหลายและหน้าที่ของระบบ CRISPR-Cas

แบคทีเรียแต่ละจีนัส และสปีชีส์ จะมีจะมีจำนวน direct repeat ที่แตกต่างกัน และมีความยาวของนิวคลีโอไทด์แตกต่างกันตั้งแต่ 24-47 คู่เบส (Grissa *et al*., 2007) ลำดับนิวคลีโอไทด์ของ direct repeat ที่พบในแบคทีเรียสปีชีส์เดียวกันส่วนใหญ่จะมีลำดับนิวคลีโอไทด์ที่เหมือนกันทั้งจำนวนคู่เบสและลำดับนิวคลีโอไทด์ ในขณะที่แบคทีเรียต่างสปีชีส์อาจจะพบลำดับนิวคลีโอไทด์ของ direct repeat แตกต่างกันได้ มีรายงานการศึกษาซึ่งได้วิเคราะห์ลำดับนิวคลีโอไทด์ของ direct repeat ในแบคทีเรียจำนวน 195 ชนิด พบว่า direct repeat นั้นสามารถแบ่งออกได้อย่างน้อย 12 กลุ่ม ตามความเหมือนกัน (homology) ของลำดับนิวคลีโอไทด์ โดย direct repeat ทั้ง 12 กลุ่มนั้นมีทั้งกลุ่มที่มีลักษณะของ palindrome และไม่เป็น palindrome (Kunin *et al*., 2007) เช่นเดียวกับ direct repeat บริเวณ spacer ก็มีความแตกต่างกันทั้งจำนวนและลำดับนิวคลีโอไทด์ ผลการวิเคราะห์ลำดับนิวคลีโอไทด์ของ spacer ที่พบในแบคทีเรียจีนัส และสปีชีส์ต่าง ๆ จำนวน 67 เชื้อ พบว่ามีจำนวน spacer รวมกันมากกว่า 4,000 spacer เมื่อวิเคราะห์ลำดับนิวคลีโอไทด์ของ spacer นั้นกับลำดับนิวคลีโอไทด์จากฟาจ และพลาสมิดที่มีในฐานข้อมูลของ GenBank พบว่ามีเพียงร้อยละ 2 ของ spacer เท่านั้นที่ลำดับนิวคลีโอไทด์มีความคล้ายกับลำดับนิวคลีโอไทด์ของฟาจ และพลาสมิดที่สืบค้นได้จากฐานข้อมูล GenBank ในขณะที่ spacer ส่วนที่เหลือนั้นไม่มีความคล้ายกับลำดับนิวคลีโอไทด์ของฟาจและพลาสมิดใด ๆ ที่สืบค้นได้จากฐานข้อมูล

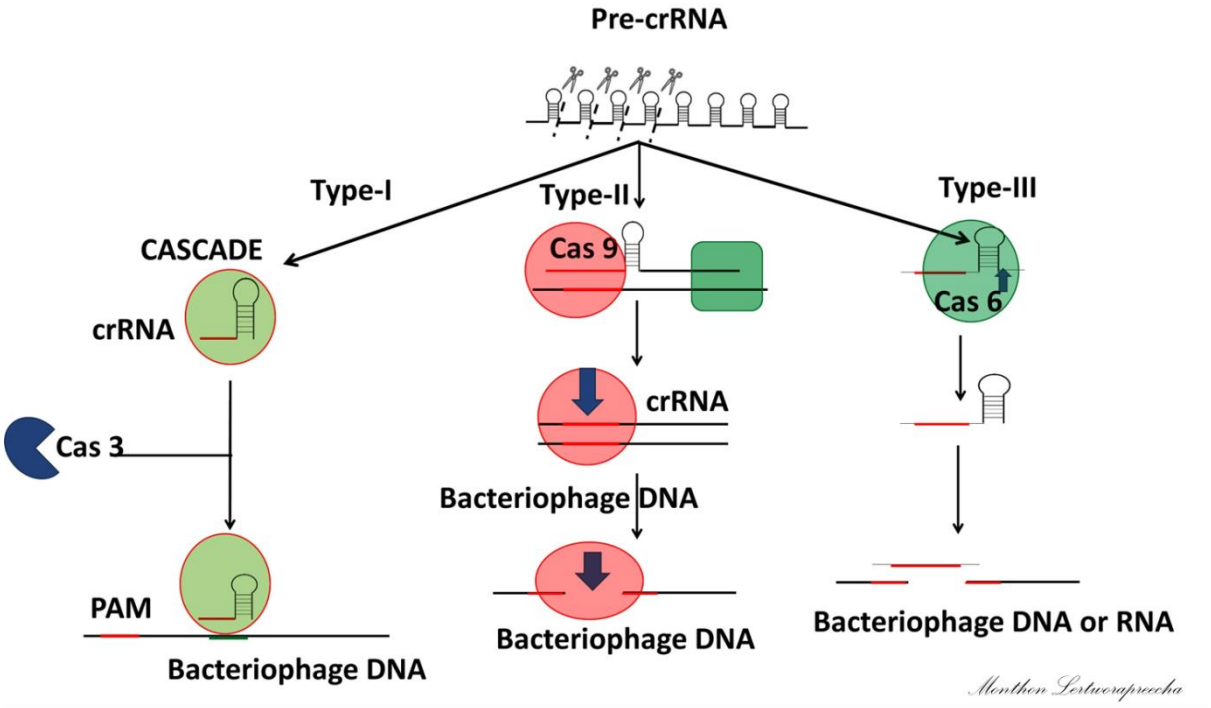
ผลการศึกษานี้แสดงให้เห็นว่าบริเวณของ spacer นั้นมีความหลากหลายที่สูงมาก นอกจากนี้ยังอาจจะเป็นข้อมูลสนับสนุนที่บ่งชี้ว่าฟาจ และพลาสมิด มีความหลากหลายที่สูงมากในสิ่งแวดล้อม และยังมีฟาจอีกเป็นจำนวนมากที่ยังไม่รู้จักและไม่มีข้อมูล (Mojica *et al*., 2005)

นอกจากบริเวณของ CRISPR แล้ว การศึกษาต่อมายังพบตำแหน่งของยีนที่เรียกว่า CRISPR-associated (*Cas*) genes อยู่ใกล้กับ CRISPR เสมอ โดยจะอยู่ห่างออกไปจาก CRISPR ทางด้าน upstream ยีน Cas จะควบคุมการสร้างโปรตีนเรียกว่าโปรตีน Cas พบว่าแบคทีเรียที่ไม่มี CRISPR ก็จะไม่พบกลุ่มของยีน *Cas* เช่นกัน ทำให้นักวิทยาศาสตร์สันนิษฐานว่า CRISPR และยีน *Cas* นั้นอาจจะเกี่ยวข้อง และทำงานร่วมกัน (Ruud *et al*., 2002) จึงเรียกระบบนี้ว่า CRISPR-Cas system ยีน *Cas* นั้นมีอยู่ด้วยกันหลายชนิด และเป็นยีนที่นำไปใช้สำหรับแบ่งกลุ่มของ CRISPR ด้วย ยีน *Cas* มีหน้าที่ควบคุมการสร้างเอนไซม์ในกลุ่ม helicase, nuclease และ RNase ซึ่งเกี่ยวข้องกับการนำชิ้นส่วนของ crRNA ไปยังเป้าหมาย กระบวนการตัดทำลาย DNA ของฟาจ และการนำพลาสมิดหรือฟาจนั้นเข้าไปแทรกระหว่าง short palindromic direct repeat (Bolotin *et al*., 2005; Godde & Bickerton, 2006; Pougacha *et al*., 2012) นอกจากโปรตีน Cas แล้วยังพบโปรตีนอีกกลุ่มหนึ่งเรียกว่า “Repeat-associated mysterious proteins” (RAMPs) ซึ่งจัดเป็นโปรตีนในกลุ่ม Cas superfamily เช่นกัน มีคุณสมบัติเป็น RNA binding protein โครงสร้างพื้นฐานของ RAMPs เหมือนกับโปรตีนที่จับกับ RNA ทั่ว ๆ ไป โดยมีส่วนปลายด้าน N-terminal เป็นสายโปรตีนที่พันทับกันไปมาระหว่างสาย α และ β เรียกว่า “ferredoxin fold” RAMPs จะทำงานร่วมกับโปรตีน Cas ในการรวมตัวกับ crRNA และนำไปจับกับเป้าหมายคือ DNA หรือ RNA ของฟาจ (Wang & Li, 2012)

ปัจุบันพบยีน *Cas* ประมาณ 45 ชนิด (Haft *et al*., 2005) ยีน *Cas* ที่พบได้บ่อยในแบคทีเรียคือยีน *Cas* 1 ถึง *Cas* 6 โดยพบว่าใน CRISPR ทุก type จะพบยีน *Cas* 1 และ 2 เสมอ โปรตีน Cas 1 เป็นเอนไซม์ nuclease ที่ทำงานโดยต้องพึ่งพาไอออนของโลหะ (metal ion dependent) สามารถตัดสาย DNA ได้แบบจำเพาะ ทำให้ได้ DNA สายคู่มีขนาดยาวประมาณ 80 คู่เบส ในขณะที่โปรตีน Cas 2 นั้นยังไม่ทราบหน้าที่อย่างแน่ชัด แต่สันนิษฐานว่าโปรตีน Cas 1 และ Cas 2 อาจจะเกี่ยวข้องกับการตัด DNA และการนำ spacer ใหม่เข้าสู่ CRISPR (Wiedenheft *et al*., 2009; Karginov & Hannon, 2010) สำหรับโปรตีน Cas 4 นั้นพบว่าเป็นเอนไซม์ 5’-3’ exonucelase จากผลการศึกษาในอาร์เคีย *Sulfolobus solfataricus* พบว่าโปรตีน Cas 4 เป็น iron sulfur ที่มีคุณสมบัติของ 5’-3’ exonucelase สามารถตัด DNA ที่เป็นสายเดี่ยวได้ และเชื่อว่าเป็นโปรตีนที่มีบทบาทเกี่ยวข้องกับการจับสารพันธุกรรมของฟาจที่แทรกเข้ามาภายในเซลล์ (Zhang *et al*., 2012) CRISPR-Cas system สามารถแบ่งออกได้เป็น 3 type ตามรูปแบบของการทำงาน โดยแบคทีเรียแต่ละชนิดอาจจะพบ CRISPR ในโครโมโซมได้มากกว่า 1 type (Makarova *et al*., 2011; Wiedenheft *et al*., 2012)

Type-I CRISPR-Cas system เป็นชนิดที่พบได้มากที่สุดในแบคทีเรีย ลักษณะสำคัญของ type-I คือจะพบยีน *Cas* 3 เป็นยีนหลัก (Makarova, Haft *et al*., 2011; Sinkunas *et al*., 2011; Ivancic-Bace *et al*., 2013) กลไกในการทำลายสาย DNA เป้าหมายโดย type-I นั้นอาศัยโปรตีน Cas หลายชนิดมาจับกันและร่วมกันทำงานเป็นกลุ่ม (multiple protein complex) เรียกว่า CRISPR-associated complex for antiviral defense หรือ “CASCADE” โดย CASCADE จะประกอบด้วยโปรตีน Cas6e และ Cas6f โปรตีน Cas 6 ทำหน้าที่ตัดชิ้นส่วน Pre-CrRNA ที่ถอดรหัสมาจาก CRISPR ให้เป็น crRNA และ จะรวมกลุ่มอยู่กับ crRNA นั้น complex ที่เกิดขึ้นจะไปจับกับ DNA เป้าหมายของฟาจหรือพลาสมิด โดยกระบวนการนี้จะมีโปรตีน Cas 3 เข้ามาร่วมกลุ่มใน CASCADE นั้นด้วย โปรตีน Cas 3 เป็นโปรตีนที่มีส่วน N-terminal domain เป็นเอนไซม์ HD phosphohydrolase และปลายด้าน C-terminal เป็น helicase

(ATP-dependent helicase) ซึ่งทำหน้าที่แยกสาย DNA สายคู่ออกจากกัน (Sinkunas *et al*., 2011) โปรตีน Cas 3 จะนำ crRNA ไปยังเป้าหมาย (crRNA-guide) ที่เป็นสารพันธุกรรมของฟาจหรือ พลาสมิด และจะแยก DNA สายคู่ของฟาจหรือพลาสมิดออกจากกันเพื่อให้ crRNA สามารถจับกับเป้าหมายได้ ก่อนที่จะทำลายสารพันธุกรรมของฟาจหรือพลาสมิดนั้น (ภาพที่ 2)



ภาพที่ 2 รูปแบบการทำงานของระบบ CRISPR-Cas type I-III

Type-II CRISPR-Cas system จะพบยีน *Cas* ได้ 4 ยีน ได้แก่ *cas* 9, *cas* 1, *cas* 2, และ *cas* 4 โดยจะพบยีน *Cas* 9 เป็นหลัก ยีน *Cas* 9 ควบคุมการสร้างโปรตีน Cas 9 ซึ่งมีโมเลกุลขนาดใหญ่ และมี nuclease โดเมนจำนวน 2 โดเมน อยู่ทางด้าน N-terminal และที่บริเวณใจกลางโมเลกุล กลไกการทำงานของ type-II CRISPR นั้นต่างจาก type-I คือไม่มี CASCADE กระบวนการที่เกิดขึ้นตั้งแต่กระบวนการตัด pre-crRNA (RNA processing) จนได้ crRNA และนำไปจับกับ DNA เป้าหมายรวมทั้งการตัดทำลาย DNA ของฟาจ เป็นผลมาจากการทำงานของโปรตีน Cas 9 ซึ่งมีคุณสมบัติเป็น multifunctional protein (Bhaya *et al*., 2011; Makarova, Haft *et al*., 2011) (ภาพที่ 2)

Type-III CRISPR-Cas system พบได้มากในกลุ่มของอาร์เคีย แบ่งออกได้เป็น 2 subtype คือ type-IIIA และ IIIB (Hrle *et al*., 2013; Richter *et al*., 2013; Staals *et al*., 2013) ใน type-III จะพบโปรตีน Cas 6 และ Cas 10 เป็นหลัก จากการศึกษาในเชื้อ *Pyrococcus furiosus* พบว่า Cas 6 มีคุณสมบัติเป็น endoribonuclease ซึ่งจะตัด crRNA ที่ repeat sequence ทำให้ได้ส่วนของ crRNA ของ spacer ที่จะไปจับกับเป้าหมาย และ active site ของ Cas 6 มีความเหมือนกับ tRNA splicing endonuclease ซึ่งการทำงานนั้นไม่ต้องพึ่งพาไอออนของโลหะหนัก (metal independent) (Carte *et al*., 2008) สำหรับโปรตีน Cas 10 มีคุณสมบัติคล้ายกับ Cas 3 คือมีส่วนโดเมนที่เป็น HD nuclease จึงคาดว่า Cas 10 อาจจะมีบทบาทในการตัดทำลายสารพันธุกรรมเป้าหมาย (Richter *et al*., 2013) นอกจากนี้

พบว่าเป้าหมายของ type-IIIA และ IIIB แตกต่างกัน กล่าวคือ type-IIIA CRISPR-Cas system ที่พบในเชื้อ *P. furiosus* จะจับกับเป้าหมายและทำลายสารพันธุกรรมของฟาจที่เป็น mRNA (Hale *et al*., 2009) ในขณะที่ type-IIIB CRISPR-Cas system ที่พบในเชื้อ *S. epidermidis* จับกับเป้าหมายและทำลายสารพันธุกรรมที่เป็น DNA (Marraffini & Sontheimer, 2008)

การประยุกต์ CRISPR สำหรับงานวิจัยทางจุลชีววิทยาด้านต่าง ๆ

การศึกษา CRISPR ในปัจจุบันให้ความสนใจใน 3 ประเด็นหลักคือ

1. การประยุกต์ใช้ความหลากหลายของ CRISPR เพื่อเป็นเครื่องมือสำหรับการจำแนกสายพันธุ์ของแบคทีเรีย สำหรับการศึกษาทางด้านระบาดวิทยา เนื่องจากบริเวณ CRISPR สามารถถ่ายทอดต่อไปยังเซลล์ลูกได้ ทำให้แบคทีเรียที่เป็นสายพันธุ์เดียวกันจะตรวจพบลักษณะของ CRISPR ที่เหมือนกัน เช่นการประยุกต์ใช้ CRISPR ในการจำแนกเชื้อกลุ่ม *Mycobacterium tuberculosis* complex (MTBC) ออกเป็นสปีชีส์ต่าง ๆ ได้แก่ *M. tuberculosis, M. africanum, M. bovis,* และ *M. canettii* ได้ (Groenen *et al*., 1993; Kamerbeek *et al*., 1997)

เชื้อ *Yersinia pestis* ซึ่งเป็นแบคทีเรียก่อโรครุนแรง และมีความหลากหลายของสายพันธุ์ต่ำ จัดเป็น monomorphic species แต่จากการศึกษาเชื้อ *Y. pestis* จากตัวอย่างที่เป็น clinical isolate จำนวนมาก ซึ่งได้มาจากแหล่งที่มีการระบาดที่เดียวกัน พบว่าเชื้อที่แยกได้นั้นมี variation ของ CRISPR ซึ่งแสดงให้เห็นว่าเชื้อที่เป็น monomorphic อาจจะมีความหลากหลายอยู่ในระดับ intra-species ดังนั้นจึงมีความเป็นไปได้ที่จะนำ CRISPR ไปประยุกต์ใช้เป็นเครื่องมือเพื่อการศึกษาทางด้านระบาดวิทยา (Pourcel *et al*., 2004) นอกจากนี้กรณีของเชื้อ *Salmonella enterica* ซึ่งเป็นแบคทีเรียก่อโรคที่สำคัญในระบบทางเดินอาหารและพบแพร่กระจายค่อนข้างสูงในประเทศไทย เชื้อ *S. enterica* มีความหลากหลายที่สูงมาก ถึงแม้ว่าจะจัดอยู่ในซีโรวาร์เดียวกัน ดังนั้นการศึกษาทางด้านระบาดวิทยาเพื่อให้ทราบถึงสายพันธุ์ที่แท้จริงที่เป็นสาเหตุการระบาดจึงมีความสำคัญมาก ปัจจุบันพบมีรายงานการศึกษาในต่างประเทศที่มีการใช้ CRISPR เพื่อจำแนกสายพันธุ์ของเชื้อ *S. enterica* โดยเปรียบเทียบ CRISPR กับ serotype และ multi-locus sequence typing พบว่าให้ผลสอดคล้องกัน แต่ที่น่าสนใจคือ variation ของ spacer นั้นสามารถแยกความแตกต่างภายในซีโรวาร์เดียวกันได้ (discriminated between subtypes within prevalent serovars) ทำให้แนวโน้มการนำ CRISPR ไปประยุกต์ใช้สำหรับการศึกษาทางระบาดวิทยามีความละเอียดมากขึ้น (Fabre *et al*., 2012)

2. การศึกษากลไกการทำหน้าที่ป้องกันตนเองของแบคทีเรีย ซึ่งได้มีการศึกษายืนยันชัดเจนว่า แบคทีเรียสามารถนำเอาชิ้นส่วน DNA จากฟาจหรือพลาสมิดเข้าสู่ CRISPR และเชื้อที่ได้รับ spacers นั้นจะสามารถป้องกันตนเองจากการบุกรุกของฟาจในครั้งต่อไปได้ ความเข้าใจในกลไกดังกล่าวมีความสำคัญมาก เนื่องจากสามารถนำไปประยุกต์ใช้เพื่อการสร้างสายพันธุ์ของเชื้อที่รู้จักฟาจหรือพลาสมิดที่เป็นพาหะนำพา virulence gene หรือ antimicrobial resistance gene เข้าสู่แบคทีเรีย ซึ่งการศึกษาในเชื้อ *S. thermophilus* ได้พิสูจน์ให้เห็นว่าเชื้อ *S. thermophilus* สามารถนำชิ้นส่วนของพลาสมิดที่มียีนดื้อยาและเพิ่มจำนวนได้เอง (antibiotic resistance self-replicating plasmid) เข้าสู่ CRISPR ได้ และเชื้อ *S. thermophilus* ได้รับ spacer ของพลาสมิดสามารถทำลายพลาสมิดที่อยู่ภายในเซลล์ได้ ทำให้เชื้อนั้นไม่ดื้อต่อยาต้านจุลชีพอีกต่อไป และไม่ว่าจะใส่พลาสมิดกลับเข้าไปใหม่ พลาสมิดนั้นก็ไม่สามารถอยู่ และเพิ่มจำนวนใน *S. thermophilus* ได้อีก (Garneau *et al*., 2010) นอกจากนี้ยังมีการศึกษาในเชื้อ *S. pyogenes* เป็นแบคทีเรียที่ก่อโรคได้หลายแบบ คุณสมบัติการก่อโรคนั้นขึ้นอยู่กับการได้รับยีนก่อโรครุนแรง (virulence gene) จากฟาจ โดยฟาจจะสอดแทรก

DNA ของตัวเองเข้าสู่โครโมโซมของแบคทีเรีย จากการศึกษาพบว่า CRISPR สามารถจำกัดการแทรกตัวของฟาจเข้าสู่โครโมโซมของแบคทีเรียได้ และได้ตั้งข้อสันนิษฐานไว้ว่า ข้อมูลในส่วน spacer ที่มีอยู่อย่างจำกัดหรือแตกต่างกันอาจจะเกี่ยวข้องกับการก่อโรคที่แตกต่างกันของเชื้อ *S. pyogenes* (Nozawa *et al*., 2011) ความเข้าใจการทำงานของ CRISPR สามารถนำไปประยุกต์ใช้ในงานอุตสาหกรรม โดยการคัดเลือกสายพันธุ์ที่ทนต่อการบุกรุกของฟาจ เช่นกรณีเชื้อ *Lactococcus lactis* เป็นแบคทีเรียที่ใช้ประโยชน์ในในอุตสาหกรรมการทำนมเปรี้ยว ปัญหาที่ผู้ผลิตพบคือแบคทีเรียมักจะถูกทำลายด้วยฟาจ ซึ่งจะส่งผลต่อกระบวนการผลิต ดังนั้นนักวิจัยจึงพยายามค้นหาสายพันธุ์ของ *L. lactis* ที่ทนต่อการทำลายด้วยฟาจ ซึ่งมีการค้นพบ type-III CRISPR-Cas ชนิดใหม่ซึ่งอยู่บนพลาสมิดของแบคทีเรียดังกล่าว และพบว่าพลาสมิดดังกล่าวสามารถส่งผ่านตามแนวราบ (horizontal transfer) ไปยัง *L. lactis* เซลล์อื่น ๆ ได้ ซึ่งจะช่วยให้ได้แบคทีเรียที่ทนต่อการทำลายด้วยฟาจในกระบวนการหมักอาหารได้ (Millen *et al*., 2012)

3. การศึกษาเพื่อใช้เป็นเครื่องมือสำหรับการวิจัยด้านการปรับปรุงพันธุกรรมในระดับจีโนม (genome editing or genome engineering) ความเข้าใจกลไกของระบบ CRISPR-Cas โดยเฉพาะระบบ CRISPR-Cas9 ช่วยให้นักวิทยาศาสตร์นำไปประยุกต์ใช้สำหรับการแก้ไขหรือเปลี่ยนแปลงพันธุกรรมของสิ่งมีชีวิตหลายชนิด รวมทั้งใช้เป็นเครื่องมือสำหรับสร้างสัตว์ทดลองกลายพันธุ์ (transgenic animal) เพื่อใช้เป็นโมเดลสำหรับศึกษาด้านต่าง ๆ หลักการสำคัญของการประยุกต์ใช้ระบบ CRISPR-Cas9 คือโปรตีน Cas9 ซึ่งมีคุณสมบัติเป็น endonuclease สามารถตัด DNA และทำให้เกิด double strand break (DSB) โดยตำแหน่งที่โปรตีน Cas9 ตัด DNA มีความจำเพาะ (sequence specific) ซึ่งขึ้นอยู่กับลำดับนิวคลีโอไทด์ของ spacer ที่เป็นต้นแบบสำหรับสร้าง crRNA และจะไปจับกับ DNA ของฟาจ ซึ่งจะนำโปรตีน Cas9 เข้ามาและตัดทำลาย DNA ของฟาจ (รูปที่ 2) โดยหลักการนี้หากสังเคราะห์ลำดับนิวคลีโอไทด์แทนบริเวณ spacer โดยให้ลำดับนิวคลีโอไทด์นั้นเหมือนกับบริเวณของยีนที่ต้องการตัด ลำดับนิวคลีโอไทด์นั้นจะเป็นต้นแบบและคัดลอกรหัสเป็น RNA เรียกว่า synthetic single guide RNA (sgRNA) ซึ่งทำหน้าที่เปรียบเหมือนกับ crRNA ไปจับกับยีนที่มีลำดับนิวคลีโอไทด์ที่เป็นคู่สมกัน และจะชักนำให้โปรตีน Cas9 เข้ามาจับและตัด DNA ในบริเวณดังกล่าวอย่างจำเพาะ ทำให้เกิด DSB ของสาย DNA (Bassett & Liu, 2014) เมื่อเกิด DSB ขึ้น เซลล์จะมีกระบวนการที่จะซ่อมแซม DNA ตัวเองซึ่งอาจจะเกิดจากกระบวนการ homologous recombination (HR) หรือ nonhomologous end joining (NHEJ) การใช้ระบบ CRISPR-Cas9 จะมีข้อดี คือใช้เวลาที่สั้น สามารถกำหนดบริเวรที่จะให้เกิด DSB ได้อย่างจำเพาะมากกว่าการชักนำให้เกิด DSB โดยการใช้สารเคมีหรือรังสี ซึ่งช่วยให้การทำให้เกิดการกลายพันธุ์หรือการเปลี่ยนแปลงยีนทำได้อย่างจำเพาะมากขึ้น (Sokolov *et al*., 2005; Dmitrieva *et al*., 2011) ตัวอย่างเช่นการทดลองทำให้เกิดการกลายพันธุ์ในแมลงหวี่ (*Drosophila*) ด้วยการฉีดโปรตีน Cas9 และพลาสมิดที่ควบคุมการสร้าง sgRNA ที่จำเพาะต่อยีน *yellow* ซึ่งควบคุมการสร้าง melanin ในแมลงหวี่ตัวเต็มวัยเข้าสู่เซลล์ของแมลงหวี่ที่อยู่ในระยะ blastoderm ผลการศึกษาพบว่าระบบ CRISPR-Cas9 ทำให้เกิด DSB ส่งผลให้เกิดการกลายพันธุ์ของยีนดังกล่าวได้อย่างจำเพาะ (Gratz *et al*., 2013) และยังมีการทดลองการกลายพันธุ์ที่ประสบผลสำเร็จในสัตว์อื่น ๆ อีกหลายชนิด เช่น ปลาม้าลาย (Zebra fish) ไส้เดือนฝอย (*Caenorhabditis elegans*) รวมทั้งในพืชด้วย (Auer *et al*., 2014; Friedland *et al*., 2013; Li *et al*., 2013) สำหรับสัตว์เลี้ยงลูกด้วยน้ำนม มีรายงานการศึกษาในเซลล์ของลิง (Cynomolgus monkey) โดยการฉีด sgRNA ที่จำเพาะกับยีน *Ppar-γ* และ *Rag-1* และฉีด mRNA ที่ควบคุมการสร้างโปรตีน Cas9 เข้าสู่เซลล์เอ็มบริโอของลิงที่อยู่ในระยะ 1 เซลล์ (on-cell-stage embryo) ซึ่งพบว่า mRNA ทั้งสองนั้นสามารถทำให้เกิดระบบ CRISPR-Cas9 ขึ้นภายในเซลล์และชักนำให้เกิด DSB ในจีโนมในเซลล์ตัวอ่อนของลิงได้อย่างจำเพาะ (Niu *et al*., 2014) นอกจากนี้ยังมีการทดลองใช้ระบบ CRISPR-Cas9 สำหรับการแก้ไขความผิดปกติของยีนใน stem cell เพื่อการรักษาโรคทางพันธุกรรม เช่น โรค cystic fibrosis (CF)

โดยการใช้ intestinal stem cell จากผู้ป่วยโรค CF ซึ่งยีนควบคุมการสร้างโปรตีน CF transmembrane conductance regulator (CFTR) ผิดปกติ และใช้ระบบ CRISPR-Cas9 ตัดสาย DNA บริเวณของยีนดังกล่าวให้เกิด DSB และใส่ยีนที่ปกติกลับเข้าไปใหม่โดยกระบวนการ homologous recombination ด้วยวิธีการนี้พบว่าสามารถแก้ไขยีนที่ผิดปกติที่เกิดขึ้นใน stem cell ได้อย่างมีประสิทธิภาพ ซึ่งในอานาคตวิธีการนี้อาจจะเป็นวิธีการใหม่สำหรับการรักษาด้วย stem cell (Schwank *et al*., 2013)

สรุป

ระบบ CRISPR-Cas เป็นกลไกระบบภูมิคุ้มกันแบบหนึ่งของแบคทีเรีย ลักษณะการทำงานของระบบนี้คล้ายคลึงกับการทำงานของ RNA interference (RNAi) ที่พบได้ในเซลล์สัตว์และพืชชั้นสูง ดังนั้นจึงเรียกชื่อระบบ CRISPR-Cas อีกแบบหนึ่งว่า prokaryotic silencing RNA (psiRNA) ปัจจุบันมีรายงานการวิจัยที่เกี่ยวข้องกับการทำงานของระบบ CRISPR-Cas ที่เกียวข้องกับกลไกการดื้อยาต้านจุลชีพของแบคทีเรียเพิ่มขึ้น ซึ่งหากมีความเข้าใจในกลไกการทำงานของระบบภูมิคุ้มกันนี้มากขึ้น อาจจะเป็นแนวทางสำคัญที่จะช่วยลดปัญหาการแพร่กระจายของแบคทีเรียดื้อยาต้านจุลชีพได้ และจากการศึกษาซึ่งพบว่าระบบ CRISPR-Cas สามารถถ่ายทอดไปยังเซลล์แบคทีเรียเซลล์ลูก (daughter cell) ได้ทำให้ยีนกลุ่มนี้สามารถนำไปประยุกต์ใช้ในการศึกษาด้านระบาดวิทยาของแบคทีเรียได้ ซึ่งในอนาคตอาจจะมีการนำเทคนิคนี้ไปใช้เพื่อทดแทนเทคนิคเดิม เช่นการทำ Pulse Field Gel Electrophoresis (PFGE) ที่ต้องใช้เครื่องมือที่มีราคาสูงได้ นอกจากนี้ความเข้าใจเกี่ยวกับกระบวนการทำงานของระบบ CRISPR-Cas ยังมีประโยชน์อย่างมากสำหรับการนำไปประยุกต์ใช้ในการศึกษาทางด้านการแก้ไขหรือปรับปรุงทางพันธุกรรมในระดับจีโนม ซึ่งเป็นงานวิจัยที่จะมีผลกระทบ (impact) อย่างมากต่อความก้าวหน้าทางวิทยาศาสตร์ต่อไป อย่างไรก็ตามองค์ความรู้เกี่ยวกับระบบ CRISPR-Cas ยังมีอยู่จำกัด อีกทั้งยังมีแบคทีเรียที่สำคัญอีกหลายชนิดที่ยังไม่ได้มีการศึกษา และยังคงรอให้นักวิทยาศาสตร์ได้ช่วยกันศึกษาวิจัยต่อไป

เอกสารอ้างอิง

Auer, T. O., Duroure, K., De Cian, A., Concordet, J. P.& Del Bene, F. (2014). Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Resrearch*, *24*, 142-153.

Bassett, A. R.& Liu, J. L. (2014). CRISPR/Cas9 and genome editing in Drosophila. *Journal of Genetics and Genomics*, *41*, 7-19.

Bhaya, D., Davison, M.& Barrangou, R. (2011). CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annual Review of Genetics*, *45*, 273-297.

Bolotin, A., Quinquis, B., Sorokin, A.& Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*,2551-2561.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L.,

Geoghagen, N. S.& Venter, J. C. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science*, *273*, 1058-1073.

Carte, J., Wang, R., Li, H., Terns, R. M.& Terns, M. P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes & Development*, *22*, 3489-3496.

Dmitrieva, N. I., Cui, K., Kitchaev, D. A., Zhao, K.& Burg, M. B. (2011). DNA double-strand breaks induced by high NaCl occur predominantly in gene deserts. *Proceedings of the National* Academy *of Sciences*, *108*, 20796-20801.

Fabre, L., Zhang, J., Guigon, G., Le Hello, S., Guibert, V., Accou-Demartin, M., de Romans, S., Lim, C., Roux, C., Passet, V., Diancourt, L., Guibourdenche, M., Issenhuth-Jeanjean, S., Achtman, M., Brisse, S., Sola, C.& Weill, F. X. (2012). CRISPR typing and subtyping for improved laboratory surveillance of *Salmonella* infections. *PLoS One*, *7,* e36995.

Friedland, A. E., Tzur, Y. B., Esvelt, K. M., Colaiacovo, M. P., Church, G. M.& Calarco, J. A. (2013). Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nature Methods*, *10*, 741-743.

Garneau, J. E., Dupuis, M.-E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A. H.& Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*, 67-71.

Godde, J. S.& Bickerton, A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *Journal of Molecular Evolution*, *62*, 718-729.

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J.& O'Connor-Giles, K. M. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics*, *194*, 1029-1035.

Grissa, I., Vergnaud, G.& Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*, *8*, 172.

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

Haft, D. H., Selengut, J., Mongodin, E. F.& Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computational Biology*, *1*, e60.

Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., Terns, R. M.& Terns, M. P. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell*, *139*, 945-956.

Hermans, P. W., van Soolingen, D., Bik, E. M., de Haas, P. E., Dale, J. W.& van Embden, J. D. (1991). Insertion element IS987 from Mycobacterium bovis BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. *Infection and Immunity*, *59*, 2695-2705.

Horvath, P., Romero, D. A., Coute-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C.& Barrangou, R. (2008). Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *Journal of Bacteriology*, *190*, 1401-1412.

Hrle, A., Su, A. A., Ebert, J., Benda, C., Randau, L.& Conti, E. (2013). Structure and RNA-binding properties of the type III-A CRISPR-associated protein Csm3. *RNA Biology*, *10*, 1670-1678.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M.& Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journla of Bacteriology, 169*, 5429-5433.

Ivancic-Bace, I., Radovcic, M., Bockor, L., Howard, J. L.& Bolt, E. L. (2013). Cas3 stimulates runaway replication of a ColE1 plasmid in Escherichia coli and antagonises RNaseHI. *RNA Biology, 10*, 770-778.

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M.& van Embden, J. (1997). Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. *Journal of Clinical Microbiology*, *35,* 907-914.

Karginov, F. V.& Hannon, G. J. (2010). The CRISPR system: small RNA-guided defense in bacteria and archaea. *Molecular Cell*, *37*, 7-19.

Kunin, V., Sorek, R.& Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biololgy*, *8*, R61.

Li, J. F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G. M.& Sheen, J. (2013). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nature Biotechnology*, *31*, 688-691.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J.& Koonin, E. V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nature Review Microbiology*, *9,* 467-477.

Marraffini, L. A.& Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, *322*, 1843-1845.

Millen, A. M., Horvath, P., Boyaval, P.& Romero, D. A. (2012). Mobile CRISPR/Cas-mediated bacteriophage resistance in *Lactococcus lactis*. *PLoS One*, *7*, e51663.

Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J.& Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, *60,* 174-182.

Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W., Xiang, A. P., Zhou, J., Guo, X., Bi, Y., Si, C., Hu, B., Dong, G., Wang, H., Zhou, Z., Li, T., Tan, T., Pu, X., Wang, F., Ji, S.,

Zhou, Q., Huang, X., Ji, W.& Sha, J. (2014). Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell*, *156*, 836-843.

Nozawa, T., Furukawa, N., Aikawa, C., Watanabe, T., Haobam, B., Kurokawa, K., Maruyama, F.& Nakagawa, I. (2011). CRISPR inhibition of prophage acquisition in *Streptococcus pyogenes*. *PLoS One*, 6, e19543.

Pingoud, A., Fuxreiter, M., Pingoud, V.& Wende, W. (2005). Type II restriction endonucleases: structure and mechanism. *Cell and Molecular Life Science*, *62*, 685-707.

Pougacha K. S., Lopatina, A. V.& Severinov K. V. (2012). CRISPR adaptive immunity systems of procaryotes. Molecular Biology, *46*, 195-203.

Pourcel, C., Andre-Mazeaud, F., Neubauer, H., Ramisse, F.& Vergnaud, G. (2004). Tandem repeats analysis for the high resolution phylogenetic analysis of *Yersinia pestis*. *BMC Microbiology*, *4*, 22.

Pourcel, C., Salvignol, G.& Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*, *151*, 653-663.

Richter, H., Randau, L.& Plagens, A. (2013). Exploiting CRISPR/Cas: interference mechanisms and applications. *International Journal of Molecular Science*, *14*, 14518-14531.

Ruud, J., Jan, D. A. v. E., Wim, G.& Leo, M. S. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, *43*, 1565-1575.

Schwank, G., Koo, B. K., Sasselli, V., Dekkers, J. F., Heo, I., Demircan, T., Sasaki, N., Boymans, S., Cuppen, E., van der Ent, C. K., Nieuwenhuis, E. E., Beekman, J. M.& Clevers, H. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell*, *13*, 653-658.

Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P.& Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO Journal*, *30*, 1335-1342.

Sokolov, M. V., Smilenov, L. B., Hall, E. J., Panyutin, I. G., Bonner, W. M.& Sedelnikova, O. A. (2005). Ionizing radiation induces DNA double-strand breaks in bystander primary human fibroblasts. *Oncogene*, *24*, 7257-7265.

Staals, R. H., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J. J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P. J., Doudna, J. A., Heck, A. J., Yonekura, K., van der Oost, J.& Shinkai, A. (2013). Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of *Thermus thermophilus*. *Molecular Cell*, *52,* 135-145.

Wang, R.& Li, H. (2012). The mysterious RAMP proteins and their roles in small RNA-based immunity. *Protein Science*, *21,* 463-470.

Wiedenheft, B., Sternberg, S. H.& Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, *482*, 331-338.

Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S. M., Ma, W.& Doudna, J. A. (2009). Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure*, *17*, 904-912.

Zhang, J., Kasciukovic, T.& White, M. F. (2012). The CRISPR associated protein Cas4 Is a 5' to 3' DNA exonuclease with an iron-sulfur cluster. *PLoS One, 7,* e47232.