บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 ชีววิทยาของชา

ชาเป็นพืชที่มีใบสีเขียวตลอดปี ชาถูกจัดอยู่ในวงศ์ (Family) Theaceae และสกุล (Genus) *Camellia* สายพันธุ์ที่นิยมนำมาบริโภคมี 2 สายพันธุ์ได้แก่ 1) ชาจีน (*C. sinensis* var. *sinensis*) มีต้นกำเนิดมาจากประเทศจีน ถูกค้นพบมานานมากกว่า 3,000 ปี เมื่อเจริญเติบโตเต็มที่สูงถึง 4 เมตร ทนต่อความหนาวเย็น และเจริญเติบโตอย่างช้า ๆ ชาจีนนี้มีใบดกแต่มีขนาดเล็กและมีกลิ่นหอมมาก และ 2) ชาอัสสัม (*C. sinensis* var. *assamica*) ที่มีต้นกำเนิดจากรัฐอัสสัม ประเทศอินเดีย ชาชนิดนี้ความสูงถึง 15-20 เมตร จัดว่าเป็นชนิดที่โตเร็วกว่าชาจีนมีใบขนาดใหญ่และรสชาดเข้มข้นมากกว่าชาจีน นอกจากนี้มีชาพันธุ์ผสม (hybrid) ที่เกิดจากการผสมพันธุ์ระหว่างชาสองสายพันธุ์ข้างต้นเกิดเป็นสายพันธุ์ใหม่ที่มีคุณสมบัติทนทานต่อสภาพอากาศ โดยทั่วไปชาใช้ระยะเวลาในการเจริญเติบโตประมาณ 6-18 เดือน ขึ้นอยู่กับสถานที่เพาะปลูกและสภาพอากาศ โดยการเพาะพันธุ์เริ่มจากเพาะให้เป็นหน่ออ่อนก่อนนำไปปลูกในไร่ การปลูกชามีความหนาแน่นของต้นเฉลี่ย 15,000-20,000 ต้นต่อ 10,000 ตารางเมตร และเมื่อเจริญเติบโตจนมีอายุประมาณ 2-3 ปี เกษตรกรจะทำการตัดพุ่มชา เมื่อต้นสูงประมาณ 1 เมตร การผลิตชาจะใช้ใบอ่อน 2-3 ใบบนสุด หรือยอดอ่อน ใบชาจะถูกเก็บในช่วงฤดูใบไม้ผลิ และฤดูใบไม้ร่วง ซึ่งขึ้นอยู่กับสถานที่ตั้ง อากาศ และชนิดของชา ทั้งนี้รสชาด และคุณภาพของชาได้รับอิทธิพลจากปัจจัยต่าง ๆ เช่น ดิน ระดับความสูง อุณหภูมิ ความชื้น วิธีการเก็บและเวลาเก็บ เป็นต้น ชาที่ปลูกและเจริญเติบโตในที่ระดับความสูงมากจะมีรสชาดและคุณภาพดีกว่าชาที่ปลูกในระดับความสูงน้อยกว่า แต่ทั้งนี้ต้นชาที่เกิดจากกิ่งชำหรือกิ่งตอนจะมีอายุได้ประมาณ 60-80 ปี ส่วนต้นชาที่เกิดจากการเพาะเมล็ดมีอายุได้มากกว่า 100 ปี (ประสานพร มณฑลธรรม, 2537)

มีการศึกษาความหลากหลายของชาพันธุ์พื้นเมืองในเขตพื้นที่ภาคเหนือของประเทศไทย โดยชวลิต กอสัมพันธ์ วราพงษ์ บุญมา และกนกวรรณ ศรีงาม (2553) พบว่า มีชาสายพันธุ์พื้นเมืองทั้งที่เกิดขึ้นเองตามธรรมชาติและนำมาปลูกกระจายอยู่ในหลายพื้นที่ตามแนวเหนือ-ใต้ ตั้งแต่ อ. แม่ฟ้าหลวง จ. เชียงราย ถึง อ. พบพระ จ. ตาก และในแนวตะวันออก-ตะวันตก ใน อ. ท่าวังผา จ. น่าน ถึง อ. เมือง จ. แม่ฮ่องสอน โดยพบที่ระดับความสูงจากระดับน้ำทะเลตั้งแต่ 450 เมตร ถึง 2,285 เมตร สภาพพื้นที่ที่พบชาพันธุ์พื้นเมืองพบทั้งป่าธรรมชาติ ป่าเหล่า ไร่ชา และ สวนปลูกที่ปล่อยให้รกร้าง พบทั้งที่มีและไม่มีการเก็บเกี่ยวใบอ่อนของชา โดยแบ่งกลุ่มชาที่พบออกเป็น

มหาวิทยาลัยบูรพา Burapha University

3 กลุ่ม คือ 1) ชาเมี่ยงหรือชาอัสสัม (*Camellia sinensis* var. *assamica*) 2) ชาป่า (*C. taliensis*) และ 3) ชาเมี่ยงและชาป่า ซึ่งเป็นพืชกลุ่มชาติพันธุ์ท้องถิ่นที่ถูกนำมาบริโภคแบบชา ส่วนพืชคล้ายชานั้นไม่ได้ใช้บริโภคแต่เกษตรกรรู้จักเป็นอย่างดี เช่น เมี่ยงอาม (*C. oleifera*) และ *C. connata* เป็นต้น และมีพืชที่มีลำต้น และใบ คล้ายชาที่ไม่ได้จำแนกอีกจำนวนหนึ่ง โดยพบว่าชาอัสสัมเป็นชนิดชาที่คนนำมาปลูกเพื่อทำเมี่ยง

ปัจจุบันการบริโภคชาได้รับความนิยมเพิ่มขึ้นอย่างรวดเร็ว โดยมีผลิตภัณฑ์แปรรูปจากชาหลากหลายรูปแบบ เช่น ชาชงพร้อมดื่ม ชาซอง และชาผง มีงานวิจัยที่ยืนยันถึงสารที่มีประโยชน์ในเครื่องดื่มจากชาต่อสุขภาพ ได้แก่ คาเทชิน จากการศึกษาในวิถีการสังเคราะห์สารฟลาโวนอยด์ในชา *C. sinensis* โดยพบเอนไซม์ที่จำเป็นในการสังเคราะห์สารหลายชนิด ได้แก่ อีพิกาโลคาเทชิน (epigallocatechin) อีพิคาเทชิน (epicatechin) และกาโลคาเทชิน (gallocatechin) ที่นำไปสร้างสารคาเทชินที่มีประโยชน์ต่อสุขภาพ เช่น ต่อต้านอนุมูลอิสระ ต่อต้านการเกิดมะเร็ง ต้านความดันโลหิตสูง สารต้านทานโรคเกี่ยวกับหลอดเลือด และสารต้านกระบวนการอักเสบ เป็นต้น (Punyasiri et al., 2004) มีรายงานการศึกษาเกี่ยวกับการสะสมและการแสดงออกของยีนสำหรับการสร้างสารคาเทชินในชา *C. sinensis* var. *sinensis* โดยพบปริมาณสารคาเทชินที่สะสมในใบอ่อนของชามีมากกว่าใบชาที่แก่เต็มที่และส่วนลำต้นรวมถึงการแสดงออกของยีนที่ควบคุมการสร้างสารคาเทชินมีมากที่ปลายยอดอ่อนของชามากกว่าส่วนอื่น นอกจากนี้ยังพบการแสดงออกของยีนอื่น ๆ ที่ปลายยอดอ่อนเช่นกัน เช่น *phenylalanine ammonia-lyase* 1 *(PAL*1*), chalcone synthase (ChS), dihydroflavonol 4-reductase (DFR), leucoanthocyanidin reductase (LCR)* และ *flavanone 3-hydroxylase (F3H)* (Eungwanichayapant & Popluechai, 2009)

2.2 ลักษณะสัณฐานวิทยาของชา (Shu, 2007)

ชา เป็นพืชไม้พุ่มที่มีใบสีเขียวตลอดปี ลักษณะของใบเป็นใบเดี่ยวเรียงสลับใบเรียงเวียนระนาบเดียว ขอบใบหยักแบบฟันเลื่อย ใบมันวาว มีหูใบชัดเจน แผ่นใบมีขนาดบางมีทั้งขอบใบหยักและขอบใบเรียบ ดอกออกตามซอกกิ่ง แต่ละดอกมีกลีบดอก 5 กลีบ แต่ละกลีบมีความแตกต่างกันจัดเป็นดอกเดี่ยวหรือออกเป็นพวง รูปทรงของดอกได้สัดส่วน เป็นดอกสมบูรณ์เพศ มีกลีบประดับรองรับดอกอยู่ 2 กลีบ มีกลีบเลี้ยง 5 กลีบ หรือติดกันที่ฐาน มีเกสรตัวผู้จำนวนมากติดกันที่ฐาน หรือติดกับโคนกลีบเกสรตัวเมีย มีรังไข่ 1 อัน อยู่เหนือฐานรองดอก มี 3 คาร์เพล และพลาเซนตารอบแกนร่วม ลักษณะทางสัณฐานวิทยาแสดงในภาพที่ 2-1

มหาวิทยาลัยบูรพา Burapha University

1 = เกสรตัวผู้
2 = เกสรตัวเมีย
3 = รังไข่
4 = ดอกแบบสมบูรณ์
5 = ลักษณะใบ กิ่ง ดอก ของชา
6 = คาร์เพล
7 = ออร์วุล
8 = เอมบริโอ

ภาพที่ 2-1 ลักษณะทางสัณฐานวิทยาของชา แสดงลักษณะใบ ดอก และรังไข่
ที่มา: http://www.pfaf.org/user/Plant.aspx?LatinName=Camellia+sinensis

2.3 เทคนิคชีววิทยาเชิงโมเลกุล

2.3.1 เทคนิค PCR (วีระพงศ์ ลุลิตานนท์ และนิภาภรณ์ แสนคุณท้าว, 2551)

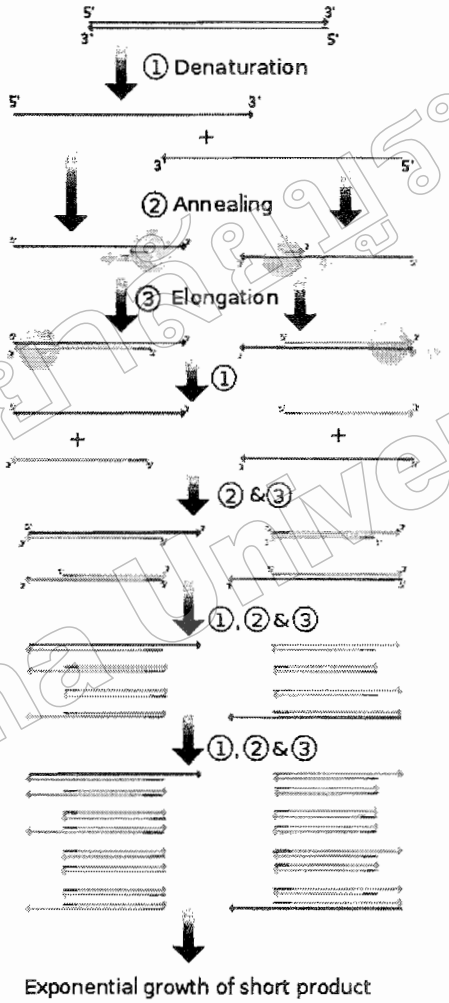
Polymerase Chain Reaction (PCR) หรือเรียกว่า *in vitro* enzymatic gene amplification เป็นเทคนิคการเพิ่มขยายปริมาณชิ้นส่วนของดีเอ็นเอในหลอดทดลองทั้งนี้ในการเพิ่มขยายปริมาณดีเอ็นเอ จำเป็นต้องอาศัยองค์ประกอบต่าง ๆ ดังนี้ คือ ดีเอ็นเอแม่พิมพ์ (template DNA) เอนไซม์ DNA polymerase ชนิดทนความร้อน (thermostable DNA polymerase) ดีออกซีนิวคลีโอไทด์ไตรฟอสเฟต (deoxyribonucleotide triphosphates; dNTPs) 4 ชนิด ไพรเมอร์ (oligonucleotide primers) และบัฟเฟอร์ที่เหมาะสม ปฏิกิริยาการสังเคราะห์จะเกิดต่อเนื่องซ้ำกันเป็นวงจรลูกโซ่ ในแต่ละรอบ (cycle) ประกอบด้วย 3 ขั้นตอน คือ

1) ขั้นตอน Denaturation เป็นขั้นตอนการแยกสายคู่ของดีเอ็นเอแม่พิมพ์ให้เป็นสายเดี่ยว โดยใช้อุณหภูมิประมาณ 90-95°C

2) ขั้นตอน Annealing เป็นขั้นตอนลดอุณหภูมิลงมาที่ 50-55°C เพื่อให้ไพรเมอร์สามารถเกาะติดกับดีเอ็นเอแม่พิมพ์สายเดี่ยวตรงบริเวณลำดับนิวคลีโอไทด์คู่สม

มหาวิทยาลัยบูรพา Burapha University

3) ขั้นตอน Elongation หรือ Extension เป็นขั้นตอนการสร้างสายดีเอ็นเอสายใหม่ต่อออกจากไพรเมอร์ในทิศทางจาก 5' ไป 3' อุณหภูมิในขั้นนี้จะอยู่ในช่วง 70-75 °C การสังเคราะห์จะดำเนินตามลำดับ 3 ขั้นตอนซ้ำกันเป็นจำนวน 20-30 รอบ ทำให้ได้ผลผลิต PCR หรือ amplified products เป็นดีเอ็นเอสายใหม่เพิ่มขึ้นเป็นจำนวนมาก ดังภาพที่ 2-2



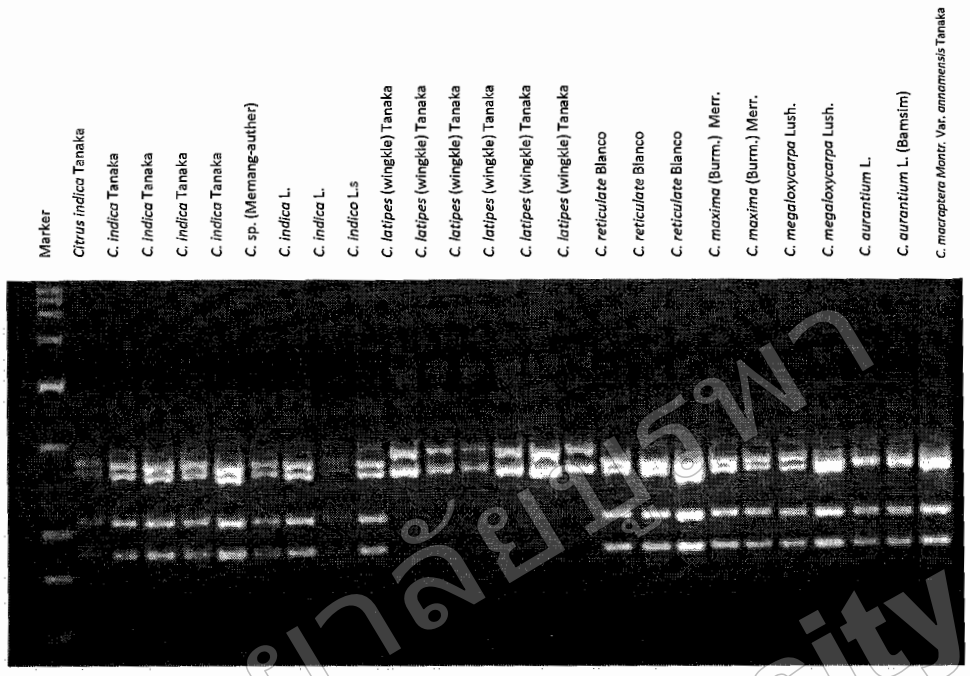
ภาพที่ 2-2 กระบวนการเพิ่มจำนวนดีเอ็นเอด้วยเทคนิค PCR มี 3 ขั้นตอนหลัก คือ denaturation, annealing และ elongation (Bartlett & Stirling, 2003)

มหาวิทยาลัยบูรพา Burapha University

ประสิทธิผลของการทำ PCR สามารถวัดได้จาก 3 ปัจจัย คือ ความจำเพาะ ประสิทธิภาพ หรือผลิตผลที่ได้ และความถูกต้อง การทำ PCR ที่มีความจำเพาะสูงมากจะทำให้ได้ผลิตภัณฑ์ PCR เกิดขึ้นจากส่วนของแม่พิมพ์ที่จำเพาะเท่านั้น ในขณะที่การทำ PCR ที่มีประสิทธิภาพสูง จะทำให้ได้ผลิตภัณฑ์ PCR ในประมาณที่สูงกว่า โดยใช้จำนวนรอบการทำน้อยกว่า ส่วนการทำ PCR ที่มีความถูกต้องแม่นยำสูง ก็จะทำให้ได้ผลิตภัณฑ์ PCR ที่มีความถูกต้องในลำดับเบสสูง หรือผิดพลาดน้อยมาก มีองค์ประกอบและภาวะต่าง ๆ จำนวนมากที่มีผลกระทบต่อปัจจัยทั้งสามที่กล่าวมา และโดยทั่วไปพบว่าการปรับองค์ประกอบ และภาวะต่าง ๆ ของปฏิกิริยา PCR เพื่อให้ได้ความจำเพาะหรือความถูกต้องแม่นยำสูงสุดมักจะมีผลให้ผลิตผลที่ได้ลดลง ในทางกลับกันการปรับปฏิกิริยา PCR เพื่อให้ได้ผลิตผลสูงสุด มักจะมีผลให้ความจำเพาะ หรือความถูกต้องแม่นยำลดลง ดังนั้นในการทำ PCR ระบบใดระบบหนึ่ง จำเป็นอย่างยิ่งที่จะต้องทราบว่าปัจจัยใดที่สำคัญที่สุดของงาน PCR แล้วปรับองค์ประกอบ และภาวะต่าง ๆให้เหมาะสมสอดคล้องที่สุด (วีระพงศ์ ลูลิตานนท์ และนิภาภรณ์ แสนคุณท้าว, 2551)

2.3.2 เทคนิค RFLP (Restriction Fragment Length Polymorphism)

เป็นเทคนิคที่อาศัยความแตกต่าง หรือความหลากหลายของขนาดดีเอ็นเอที่เกิดจากการตัดด้วยเอนไซม์ตัดจำเพาะ (restriction enzyme) ข้อมูลทางพันธุกรรมของสิ่งมีชีวิตที่เก็บอยู่ในนิวเคลียส และออร์แกเนลล์บางชนิด ได้แก่ คลอโรพลาสต์ และไมโทคอนเดรีย ในรูปของโมเลกุลดีเอ็นเอที่อยู่ในเซลล์ของสิ่งมีชีวิตนี้มีความสามารถที่จะจำลองโมเลกุลได้อย่างถูกต้องแม่นยำ แต่บางครั้งก็มีการเปลี่ยนแปลงของเบสภายในดีเอ็นเอได้ เนื่องจากสิ่งแวดล้อมหรือข้อผิดพลาดของเซลล์เองหรือเหตุอื่น โดยการเปลี่ยนแปลงดังกล่าวนี้ทำให้เกิดความหลากหลายของสิ่งมีชีวิตแต่ละชนิด ความหลากหลายดังกล่าวนี้สามารถตรวจพบได้โดยการตัดสายดีเอ็นเอด้วยเอนไซม์ตัดจำเพาะเกิดขนาดของชิ้นส่วนของดีเอ็นเอที่หลากหลายบนเจลอะกาโรสภายหลังทำอิเล็คโทรโฟริซีสแล้วเปรียบเทียบกันหรือสร้างเป็นลายพิมพ์ดีเอ็นเอ (สุรินทร์ ปิยะโชคณากุล, 2552) ตัวอย่างเช่น ลายพิมพ์ดีเอ็นเอของส้มอินเดียสกุล *Citrus* L.ในรายงานการศึกษา Jena, Kumar, and Nair (2009) วิเคราะห์ในลำดับนิวคลีโอไทด์ของคลอโรพลาสต์ดีเอ็นเอบริเวณอินทรอน โดยศึกษากับตัวอย่างส้มอินเดียทั้งหมด 50 ตัวอย่าง เพิ่มจำนวนโดยใช้เทคนิค PCR และตัดผลผลิต PCR ด้วยเอนไซม์ *Hinf* I, *Msp*I, *Alu*I และ *Hae*III เกิดเป็นลายพิมพ์ดีเอ็นเอ (แสดงบางตัวอย่างในภาพที่ 2-3)



ภาพที่ 2-3 ตัวอย่างรูปแบบลายพิมพ์ดีเอ็นเอ (DNA fingerprint) ในส้มอินเดีย *Citrus* L.บาง สายพันธุ์ที่ศึกษาจากส่วนของอินทรอน (trnD-trnT) ในคลอโรพลาสต์ดีเอ็นเอ (cpDNA) ผลผลิต PCR นำมาตัดด้วยเอนไซม์ *Alu* I (Jena et al., 2009)

2.4 ชีวสารสนเทศศาสตร์ (Bioinformatics) (วีระพงศ์ ลุลิตานนท์, 2544)

ชีวสารสนเทศศาสตร์จัดเป็นศาสตร์ที่พัฒนาขึ้นมาเพื่อใช้ในการบริหารจัดการกับข้อมูลด้านชีววิทยาอย่างเป็นระบบฐานข้อมูล เพื่อการค้นข้อมูลรวมถึงการวิเคราะห์ และแปรผลข้อมูลอย่างสะดวก และรวดเร็ว จึงจำเป็นที่ต้องนำชีวสารสนเทศศาสตร์มาช่วยในการบริหารจัดการข้อมูลเหล่านี้ เช่น การคิดค้นพัฒนาโปรแกรมเพื่อช่วยในการเก็บและสืบค้นข้อมูล การออกแบบระบบฐานข้อมูลรวมถึงโปรแกรมในการวิเคราะห์ข้อมูลรูปแบบต่าง ๆ ที่สะดวกต่อการเรียนรู้และการใช้งาน เป็นต้น การบริหารจัดการข้อมูลดังกล่าวในรูปแบบที่เหมาะสมด้วยชีวสารสนเทศศาสตร์สามารถนำไปสู่ความรู้ความเข้าใจในกระบวนการต่าง ๆ ที่เกิดขึ้นในสิ่งมีชีวิต

มหาวิทยาลัยบูรพา Burapha University

2.4.1 การนำเครือข่ายอินเทอร์เน็ตมาประยุกต์ใช้กับงานทางอณูชีววิทยา

(สิทธิรักษ์ รอยตระกูล, 2544)

ปัจจุบันวิทยาการด้านเทคโนโลยีสารสนเทศมีความเจริญรุดหน้าและได้รับความสนใจอย่างกว้างขวาง เครือข่ายอินเทอร์เน็ตมีความสำคัญกับงานทางอณูชีววิทยา ทำให้นักวิจัยทั่วโลกสามารถแลกเปลี่ยนความรู้ความคิดเห็นผ่านเครือข่ายอินเทอร์เน็ตได้อย่างสะดวก รวดเร็ว ด้วยเหตุนี้หน่วยงานดูแลฐานข้อมูลด้านอณูชีววิทยาต่าง ๆ จึงนำฐานข้อมูลที่มีอยู่แล้วในระบบคอมพิวเตอร์มาเชื่อมต่อกับเครือข่ายอินเทอร์เน็ต เรียกว่า public domain ซึ่งผู้ที่สนใจใช้ข้อมูลใด ๆ ในฐานข้อมูลดังกล่าวสามารถเรียก (access) และทำการคัดลอก (copy) ข้อมูลที่ต้องการ เพื่อใช้ในงานวิจัยที่เกี่ยวข้อง ฐานข้อมูลทางอณูชีววิทยาบนเครือข่ายอินเทอร์เน็ตทั่วโลกมีมากกว่า 400 ฐานข้อมูล แต่ละฐานข้อมูลจะเก็บรักษาข้อมูลที่มีเอกลักษณ์เฉพาะตัว ดังนั้นผู้วิจัยต้องเลือกฐานข้อมูลให้ตรงกับความต้องการของการใช้งาน เพื่อช่วยเก็บและวิเคราะห์ข้อมูลจะทำให้ประสิทธิภาพในการเรียกใช้ฐานข้อมูลได้อย่างมีประสิทธิภาพซึ่งสามารถแบ่งฐานข้อมูลทางอณูชีววิทยาออกเป็น 3 กลุ่ม คือ

1) Generalized database บรรจุข้อมูลทางอณูชีววิทยาที่เป็นข้อมูลดิบหรือเป็นข้อมูลที่ได้รับการวิเคราะห์ในเบื้องต้นแล้ว พร้อมกับลักษณะที่สำคัญทางชีววิทยา สัณฐานวิทยา และข้อมูลที่สำคัญ เช่น accession number ชื่อเจ้าของข้อมูลวารสารที่ถูกอ้างถึงข้อมูลนี้ สิ่งมีชีวิตที่กำลังศึกษาข้อมูล เป็นต้น ส่วนใหญ่วารสารทางวิทยาศาสตร์จะตีพิมพ์เฉพาะบทความที่อ้างอิงถึงลำดับนิวคลีโอไทด์ หรือลำดับกรดอะมิโนที่มี accession number ปรากฏในฐานข้อมูลประเภทนี้แล้วเท่านั้นตัวอย่างของฐานข้อมูลลำดับเบส ได้แก่

GenBank	http://www.ncbi.nlm.nih.gov/
EMBL	http://www.ebi.ac.uk/
DDBJ	http://www.ddbj.nig.ac.jp/

ฐานข้อมูลลำดับเบสทั้ง 3 ฐานข้อมูล เป็นฐานข้อมูลหลักที่มีการอ้างอิงถึงเสมอ แต่ละฐานข้อมูลตั้งอยู่ตามแหล่งข้อมูลสำคัญของโลก (สหรัฐอเมริกา อังกฤษ และญี่ปุ่น) แต่มีการแลกเปลี่ยนข้อมูลลำดับเบสระหว่าง 3 ฐานข้อมูลเป็นประจำ

2) Specialized database ฐานข้อมูลประเภทนี้แบ่งออกได้เป็น 3 จำพวก ได้แก่

2.1 ฐานข้อมูลที่ระบุเฉพาะลงไปว่าจะเก็บข้อมูลชนิดใดชนิดหนึ่งเท่านั้น เช่น เฉพาะข้อมูลที่เป็นเนื้อยีนควบคุมการสร้างโปรตีน (Express Sequence Tag หรือ EST)

DBEST	http://www.ncbi.nlm.nih.gov/dbEST/index.html

เฉพาะข้อมูลที่เป็นกรดนิวคลีอิกชนิดอาร์เอ็นเอ

มหาวิทยาลัยบูรพา Burapha University

RDP http://rdp.life.uiuc.edu/

rRNA WWW http://rrna.uia.ac.be/

2.2 ฐานข้อมูลที่จัดเก็บเฉพาะจีโนมของสิ่งมีชีวิต หรือของออร์แกเนลล์ เช่น

GDB http://www.gdb.org/

MITOMAP http://www.gen.emory.edu/mitomap.html

2.3 ฐานข้อมูลที่เก็บข้อมูลทางอณูชีววิทยา ชีววิทยา และสาขาที่เกี่ยวข้อง เพื่อใช้เป็นแหล่งข้อมูลที่สำคัญเมื่อต้องศึกษาวิเคราะห์ข้อมูลในระดับสูงต่อไป เช่น ข้อมูลการใช้รหัสทางพันธุกรรม (genetic codon) ในยีนของสิ่งมีชีวิตกว่า 6,048 ชนิด

CUD http://www.dna.affrc.go.jp/~nakamura//cutg.html

ข้อมูลเกี่ยวกับเอนไซม์ตัดจำเพาะ (restriction enzyme)

REBASE http://www.neb.com/rebase/rebase.html

ข้อมูลโครงร่างสามมิติของโปรตีนกรดนิวคลีอิกหรือมหโมเลกุล

PDB http://www.pdb.bnl.gov/

ข้อมูลเกี่ยวกับลายพิมพ์โปรตีน

ฯลฯ

3) Integrated database ซอฟต์แวร์ ที่สามารถรับคำสั่งหาข้อมูลที่ซับซ้อน แล้วไปค้นหาข้อมูลที่จำเพาะหรือใกล้เคียงกับคำสั่งของข้อมูลนั้นมากที่สุดที่ได้จากฐานข้อมูลใหญ่ ๆ ทั่วโลก ตัวอย่างฐานข้อมูลประเภทนี้ได้แก่

ENTREZ http://www.ncbi.nlm.nih.gov/entrez/

PUBMED http://www.ncbi.nlm.nih.gov/pubmed/

DBGET http://www.genome.ad.jp/dbget/dbget.html

ทั้ง 3 ข้อมูลมีลักษณะพิเศษ คือรับคำสั่งหาข้อมูล ที่ซับซ้อนได้ไม่ว่าจะเป็นข้อมูลเกี่ยวกับลำดับนิวคลีโอไทด์ โปรตีน ลำดับกรดอะมิโนหรือส่วนที่เกี่ยวข้องอื่น ๆ เช่น keywords, accession number, ชื่อเจ้าของข้อมูล ตำแหน่งบนโครโมโซม ชนิดของสิ่งมีชีวิต หรือบทความในวารสารวิทยาศาสตร์ฐานข้อมูลประเภทนี้จะทำการค้นหาข้อมูลในฐานข้อมูล 2 ประเภทแรก และแสดงผลข้อมูลที่ตรงหรือใกล้เคียงกับคำสั่ง หรือแสดงผลในรูปแบบที่สามารถเชื่อมโยงกับฐานข้อมูลที่เกี่ยวข้องกับการค้นหาตำแหน่งของข้อมูลบนเครือข่ายอินเทอร์เน็ตที่เกี่ยวข้อง เห็นได้ว่าจากคำสำคัญเพียงคำเดียวฐานข้อมูลประเภทนี้จะเชื่อมโยงไปยังข้อมูลที่เกี่ยวข้องได้จำนวนมาก ทั้งที่เกี่ยวข้องทั้งทางตรงและทางอ้อม

มหาวิทยาลัยบูรพา Burapha University

ส่วนการบริการวิเคราะห์ข้อมูลทางอณูชีววิทยาโดยใช้ซอฟต์แวร์บนเครือข่ายอินเทอร์เน็ต ต่อไปนี้ เป็นตัวอย่างซอฟต์แวร์ที่นิยมใช้ในการวิเคราะห์ข้อมูลทางอณูชีววิทยา ซึ่งถูกจัดหมวดหมู่ตามลักษณะการทำงาน ดังนี้

1) บริการกลับทิศสายนิวคลีโอไทด์ (reverse) และบริการหาเส้นคู่สม (complementary) ของสายนิวคลีโอไทด์ เมื่อต้องการปรับเปลี่ยนลำดับนิวคลีโอไทด์จากการทำ DNA sequencing ทั้งสองสายให้เป็นทิศทางเดียวกัน (5' to 3' หรือ 3' to 5') เพื่อให้ง่ายต่อการวิเคราะห์เปรียบเทียบความเหมือนระหว่างสายนิวคลีโอไทด์ หรือการต่อสายดีเอ็นเอยาว ๆ (contig assembly) เป็นต้น

Genetic Computer Group Wisconsin Package:-

http://www.embl-heidelberg.de/~toldo/Reverse.html

Human Genome Center, Baylor College of Medicine:-

http://dot.imgen.bcm.tmc.edu:9331/seq-util/option/revcomp.html

2) บริการเปรียบเทียบความเหมือน (identity) หรือความคล้ายคลึง (similarity) ระหว่างสายดีเอ็นเอ หรือสายโปรตีนที่สนใจกับฐานข้อมูล (Homology or similarity search) เมื่อทราบลำดับเบสบนสายดีเอ็นเอที่กำลังศึกษา หรือลำดับกรดอะมิโนบนสายโปรตีนแล้ว สิ่งต่อไปที่ควรวิเคราะห์ คือ วิเคราะห์หาว่าสายดีเอ็นเอหรือสายโปรตีนที่ต้องการศึกษามีความเหมือนหรือคล้ายคลึงกับยีนหรือโปรตีนชนิดใดในฐานข้อมูลบ้างเพื่อจะได้ทำนายลักษณะของสายดีเอ็นเอหรือสายโปรตีน

3) บริการเปรียบเทียบความเหมือน (identity) หรือความคล้ายคลึง (similarity) ระหว่างกลุ่มของสายดีเอ็นเอ หรือสายโปรตีนที่กำลังศึกษา (sequence alignment) โดยทั่วไปแล้วการเปรียบเทียบความเหมือนของกลุ่มสายดีเอ็นเอ หรือสายโปรตีนที่มีอยู่แล้วเพื่อหาช่วงลำดับอนุรักษ์ (consensus หรือ conserved region) ของสายดีเอ็นเอหรือสายโปรตีนกลุ่มนั้น มีประโยชน์สำหรับการออกแบบ Primers สำหรับปฏิกิริยาลูกโซ่พีซีอาร์และใช้ในการทำนายสายดีเอ็นเอหรือสายโปรตีน ที่สำคัญควรศึกษาวิเคราะห์ต่อไป หลังจากผ่านการคัดเลือก ผลลัพธ์จากการทำ sequence alignment นำไปสู่การทำนายลักษณะโครงร่างทุติยภูมิ ตติยภูมิของโปรตีน หรือดีเอ็นเอได้ นอกจากนี้ผลลัพธ์การทำนายโครงร่างดังกล่าว สามารถใช้แบ่งกลุ่มโปรตีน ศึกษาวิเคราะห์ลำดับจีโนม และศึกษาความสัมพันธ์ทางวิวัฒนาการของสิ่งมีชีวิตได้อีกด้วย

3.1 Pair sequence alignment เปรียบเทียบระหว่างความเหมือนระหว่างสายดีเอ็นเอ หรือสายโปรตีน 2 เส้น

LALIGN	http://www.ch.embnet.org/software/LALIGN_form.html
	http://fasta.bioch.virginia.edu/fasta/lalign.html
LFASTA	http://www2.igh.cnrs.fr/fasta/fasta-query.html

มหาวิทยาลัยบูรพา Burapha University

http://bioinfo.ncc.go.jp/FASTA/fasta.html199612217

3.2 Multiple sequence alignment เปรียบเทียบความเหมือนของสายดีเอ็นเอ หรือสายโปรตีนมากกว่า 2 เส้น

ClustalW 1.6	http://dot.imgen.bcm.tmc.edu:9331/multi-align/Option/clustalw.html
MAP	http://dot.imgen.bcm.tmc.edu:9331/seq-search/protein-search.html
MSA	http://dot.imgen.bcm.tmc.edu:9331/seq-search/protein-search.html
MultAlin	http://www.toulouse.inra.fr/multalin.html
Molsoft	http://molsoft.com/serv/walignm.htm

4) บริการสร้าง Phylogenetic tree จากลำดับเบส หรือลำดับกรดอะมิโน หรือจากค่า genetic distance มีนักวิทยาศาสตร์หลายกลุ่มที่พยายามศึกษาเกี่ยวกับการหาความสัมพันธ์ทางลำดับวิวัฒนาการโดยอาศัยข้อมูลจากซากบรรพชีวิน (fossil) หรือจากลักษณะสัณฐานวิทยา (morphology) ซึ่งได้ข้อสรุปที่ให้ความสนใจกับข้อมูลระดับโมเลกุล จากเซลล์สิ่งมีชีวิตหรือข้อมูลจากสายดีเอ็นเอ และสายโปรตีน สิ่งมีชีวิตที่มีบรรพบุรุษ ร่วมกันควรจะมีข้อมูลระดับโมเลกุลคล้ายกัน และสิ่งมีชีวิตชนิดเดียวกันอาจมีความแตกต่างของข้อมูลระดับ โมเลกุลได้จากวิวัฒนาการที่ต่างกันของสิ่งมีชีวิตแต่ละสายพันธุ์

เนื่องจากข้อมูลความหลากหลายทางชีวภาพของสิ่งมีชีวิตมีจำนวนมากจึงเก็บรักษาไว้เป็นฐานข้อมูลจำเพาะ นอกจากนี้ยังมีการพัฒนาซอฟต์แวร์สำหรับใช้วิเคราะห์ข้อมูลหาความสัมพันธ์ทางลำดับวิวัฒนาการ และ การจัดจำแนกสายพันธุ์ของสิ่งมีชีวิต การเลือกใช้ข้อมูลขึ้นอยู่กับลักษณะของข้อมูลดิบ วิธีวิเคราะห์ข้อมูล (parsimony หรือ distance matrix หรือ maximum likelihood) และรูปแบบของผลลัพธ์ที่ต้องการ (ค่า distance หรือ phylogenetic tree) เป็นสำคัญ ซึ่งบางครั้งซอฟต์แวร์เพียงโปรแกรมเดียวอาจไม่เพียงพอสำหรับการวิเคราะห์ข้อมูลจนได้ผลลัพธ์ที่ต้องการ และที่สำคัญควรเลือกใช้กลุ่มโปรแกรมที่ถูกต้องเหมาะสมเพื่อใช้วิเคราะห์ข้อมูลอย่างมีประสิทธิภาพ ตัวอย่างฐานข้อมูลความหลากหลายทางชีวภาพของสิ่งมีชีวิต ได้แก่

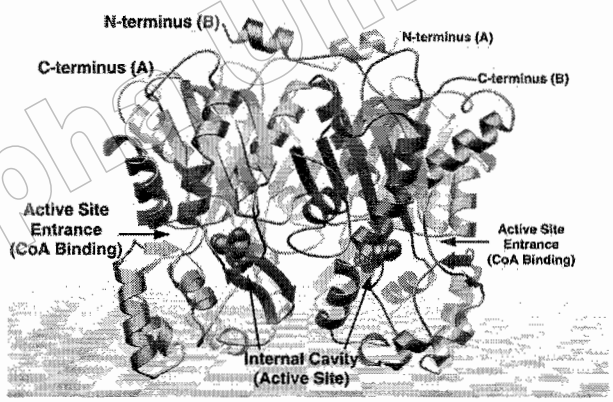
TreeBase (database)	http://www.herbaria.harvard.edu/treebase/
Dst (Repeat Sequence Analysis)	http://alces.med.umm.edu/newdst.html
The Tree of Life	http://phylogeny.arizona.edu/tree/phylogeny.html
PhyloTree/RootedTree	http://cbrg.inf.ethz.ch/subsection3_1_1.html

มหาวิทยาลัยบูรพา Burapha University

TreeGen (สร้าง phylogenetic tree)	http://cbrg.inf.ethz.ch/subsection3_1_6.html
Phylip program	http://bioweb.pasteur.fr/seqanal/phylogeny/phylip-uk.html
Phylogenetic Tree prediction	http://www.genebee.msu.su/seervices/phtree_full.html
Phylogeny Software	http://evolution.genetics.washington.edu/phylip/software.html

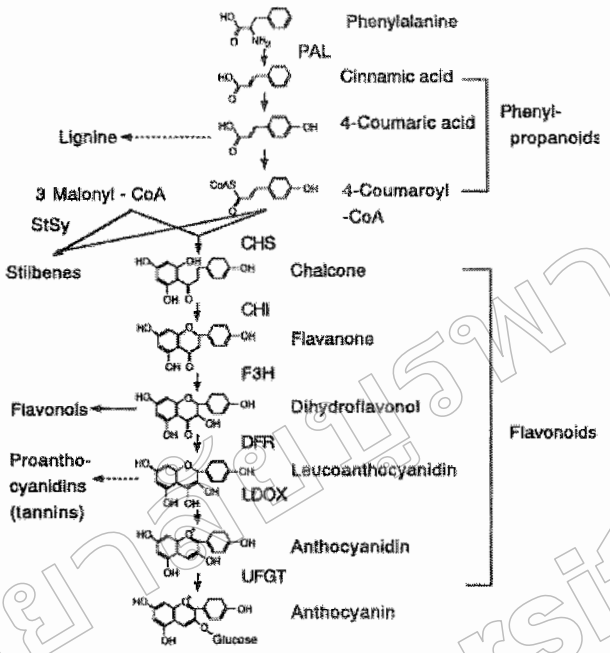
2.6 เอนไซม์ Chalcone synthase (CHS)

เอนไซม์ Chalcone synthase มีโครงสร้างทางเคมีแสดงในภาพที่ 2-4 ทำหน้าที่เป็นตัวเร่งปฏิกิริยาในวิถีการสังเคราะห์ฟลาโวนอยด์ในพืช จากการศึกษากลไกการทำงานของ Chalcone synthase (CHS) ของ Ferrer et al. (1999) ในวิถีการสังเคราะห์ polyketide พบว่าเอนไซม์ Chalcone synthase ทำหน้าที่เร่งปฏิกิริยาควบแน่นระหว่าง 3 Malonyl-CoA กับ 4-Coumaroyl-CoA ไปเป็น Chalcone ซึ่งเป็นสารตัวกลางในวิถีการสังเคราะห์ฟลาโวนอยด์ วิถีการเกิดปฏิกิริยาการสังเคราะห์ฟลาโวนอยด์แสดงในภาพที่ 2-5



ภาพที่ 2-4 โครงสร้างสามมิติของเอนไซม์ CHS บริเวณปลาย N (A) และปลาย C (B) ของแต่ละมอนอเมอร์ แสดงตำแหน่งจับของ CoA บนบริเวณเร่งของเอนไซม์ CHS (Ferrer et al., 1999)

มหาวิทยาลัยบูรพา Burapha University



ภาพที่ 2-5 วิถีการเกิดปฏิกิริยาเคมีของเอนไซม์ Chalcone synthase ในวิถีการสังเคราะห์ฟลาโวนอยด์ (Goto, Wana, Masakib, & Kobayashi, 2002)

2.7 งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยที่เกี่ยวข้องกับการนำเทคนิค PCR-RFLP ไปใช้ประโยชน์อย่างหลากหลายทั้งในสัตว์ พืช และจุลินทรีย์ เช่น นำมาใช้ตรวจสอบการปลอมปนของผลิตภัณฑ์ที่มาจากสิ่งมีชีวิต จำแนกและระบุชนิดของสัตว์ ตัวอย่างเช่น การศึกษาใน Boonphakdee and Sawangwong (2008) นำเทคนิค PCR-RFLP ประยุกต์ใช้จำแนกชนิดของปลาการ์ตูนในสกุล *Amphiprion* และ *Premnas* เนื่องจากในวัยอ่อนของปลาสกุลดังกล่าวไม่สามารถจำแนกชนิดโดยใช้ลักษณะทางสัณฐานวิทยาได้ โดยพิจารณาเพิ่มจำนวนบางส่วนของยีนในไมโทคอนเดรียดีเอ็นเอแล้วตัดด้วยเอนไซม์ตัดจำเพาะ *Bfu*CI กับ *Rsa*I และ *Hinf*I กับ *Rsa*I ปรากฏรูปแบบของแถบดีเอ็นเอ ที่สามารถจำแนกปลาการ์ตูนที่ศึกษาได้เป็น 6 ชนิด และ 7 สายพันธุ์ รวมถึงมีรายงานวิจัยเกี่ยวกับการตรวจสอบผลิตภัณฑ์จากสิ่งมีชีวิตด้วยใช้เทคนิค PCR-RFLP เช่น ใช้จำแนกชนิดของผลิตภัณฑ์จากเนื้อ เนื่องจากราคาเนื้อสัตว์แต่ละชนิดแตกต่างกัน จึงอาจพบการปลอมปนของเนื้อสัตว์ที่มีราคาต่ำกว่าในผลิตภัณฑ์ได้โดยนำยีนในไมโทคอนเดรีย (12S *rRNA*) พบว่าภายหลังจากใช้เอนไซม์ *Alu*I, *Hha*I, *Apo*I และ *Bsp*TI เกิดรูปแบบแถบดีเอ็นเอที่แตกต่างกัน สามารถแยกผลิตภัณฑ์ที่ทำมาจากเนื้อวัว ควาย และแกะ ออกจากกันได้ (Girish et al., 2005)

มหาวิทยาลัยบูรพา Burapha University

นอกจากนั้นยังพบตัวอย่างรายงานการนำเทคนิค PCR-RFLP ไปใช้ในจุลินทรีย์ *Arcobacter* ซึ่งเป็นแบคทีเรียก่อโรคที่พบว่ามีการกลายพันธุ์ และเกิดชนิดใหม่เสมอ โดยศึกษาบริเวณยีน 16S *rRNA* ร่วมกับการตัดด้วยเอนไซม์ *Mse*I พบรูปแบบแถบดีเอ็นเอที่แตกต่างกัน สามารถแยก *Arcobacter* ได้ 6 ชนิด ที่จะเป็นประโยชน์ต่อการพัฒนาด้านการแพทย์ต่อไป (Figueras, Collado, & Guarro, 2008) นอกจากนี้ยังพบรายงานการศึกษาที่จำแนกชนิดของราที่คัดแยกจากพริกขี้หนู ซึ่งราบางชนิดสร้างสารพิษทำให้เกิดเจ็บป่วยได้ โดยใช้เทคนิค PCR-RFLP ศึกษาบริเวณ 5.8S *rRNA*-ITS (internal transcribed spacer) และนำมาตัดด้วยเอนไซม์ *Sma*I, *Sau3*AI, *Hinf*I, *Taq*I และ *Eco*RI สามารถจำแนกราสกุล *Fusarium*, *Aspergillus*, *Penicillium*, *Cladosporium*, *Mucor* และ *Phlebia* ออกเป็น 16 ชนิด ซึ่งมีความสำคัญต่อการพัฒนายาที่มีความจำเพาะต่อเชื้อราที่จำแนกในลำดับต่อไปได้ (Moyano et al., 2009) และมีตัวอย่างการวิเคราะห์ลำดับนิวคลีโอไทด์บางส่วนของยีนที่สังเคราะห์โปรตีน Heat-shock ของ *Leishmania* spp. ซึ่งเป็นเชื้อที่ทำให้เกิดโรคแก่คนในหลายประเทศ พบแถบดีเอ็นเอที่แตกต่างกัน 9 รูปแบบ ภายหลังการตัดบริเวณของยีนดังกล่าวด้วยเอนไซม์ *Hae*III สามารถจำแนกเชื้อได้ 2 ชนิด ได้แก่ *L. lainsoni* และ *L. major* แต่อีก 4 ชนิด ไม่สามารถจำแนกได้ แต่เมื่อเลือกใช้เอนไซม์ *Hae*III, *Mlu*I, *Bsa*HI, *Bcc*I และ *Rsa*I จะสามารถจำแนกเชื้อแต่ละชนิดได้ (Montalvo et al., 2010)

สำหรับการบ่งชี้ชนิดของพืชด้วยวิธี PCR-RFLP จากลำดับนิวคลีโอไทด์บนดีเอ็นเอของคลอโรพลาสต์ นิวเคลียร์ดีเอ็นเอ และไมโทคอนเดรียดีเอ็นเอ เช่น Jena et al. (2009) ศึกษาความสัมพันธ์ทางวิวัฒนาการของส้มอินเดียสกุล *Citrus* L. โดยวิเคราะห์จากลำดับนิวคลีโอไทด์บริเวณ ITS ของนิวเคลียร์ดีเอ็นเอ และบริเวณอินทรอน (*trn*L และ *trn*-F) ของคลอโรพลาสต์ดีเอ็นเอได้ผลผลิต PCR ขนาด 3,100 และ 1,600 คู่เบส ตามลำดับ ในส้มอินเดียสกุล *Citrus* spp. ทั้งหมด 50 ตัวอย่าง สามารถระบุได้ว่า *C. maxima, C. medica* และ *C. reticulate* เป็นส้มอินเดียสายพันธุ์ต้นกำเนิด นอกจากนั้นมีรายงานการศึกษาของ Qing, Qing, Hong, Chen, and Zheng (2009) ใช้เทคนิค PCR-RFLP จำแนกชนิดของลูกแพร์ (*Pyrus pyrifolia* Nakai) ซึ่งจัดเป็นพืชที่ไม่มีการผสมตัวเอง ลักษณะของลูกไม่เหมือนกับต้นพ่อแม่จึงยากที่จะจำแนกชนิดโดยใช้ลักษณะภายนอกแต่ด้วยเทคนิค PCR-RFLP สามารถจำแนกชนิดของลูกแพร์ที่มีแหล่งกำเนิดจากประเทศจีน 7 ชนิด และญี่ปุ่น 2 ชนิด ออกจากกันได้

มีรายงานการประยุกต์ใช้ข้อมูลชีวสารสนเทศบนปฏิบัติการคอมพิวเตอร์ก่อนเพื่อนำมาใช้ประโยชน์ในการจัดจำแนกหรือบ่งชี้ชนิดของสิ่งมีชีวิตในห้องปฏิบัติการ ดังเช่น การจำแนกแบคทีเรียในสกุล *Bacillus* โดยใช้ลำดับนิวคลีโอไทด์ของยีน 16S *rRNA* จากฐานข้อมูล GenBank ขนาด 463 คู่เบส นำมาเทียบเคียงความเหมือนและความแตกต่างของลำดับนิวคลีโอไทด์โดยใช้

มหาวิทยาลัยบูรพา Burapha University

โปรแกรม ClustalW คัดเลือกเอนไซม์ตัดจำเพาะได้เป็น *Nsp*BII, *Sac*II และ *Cfr*I จากนั้นนำมาวิเคราะห์ความสัมพันธ์ทางวิวัฒนาการ พบว่าสามารถบ่งชี้ชนิดของแบคทีเรียได้จำนวน 34 ชนิด (Chukeatirote, Wisitrasamewong, & Jongmahasavat, 2008) ยังพบรายงานการนำข้อมูลสารสนเทศไปใช้ประโยชน์ในการจำแนกเชื้อไฟโตพลาสมาที่เป็นสาเหตุให้เกิดโรคใบเลี้ยงขนาดยักษ์ (Giant calyx) ในพืชวงศ์ Solanaceae (*Solanum melongena* L.) วิเคราะห์ส่วนของยีน 16S *rRNA* ขนาด 1,247 คู่เบส แล้วตัดด้วยเอนไซม์ตัดจำเพาะจำนวน 17 ชนิด ได้แก่ *Alu*I, *Bam*HI, *Bfa*I, *Bst*UI (*Tha*I), *Dra*I, *Eco*RI, *Hae*III, *Hha*I, *Hinf*I, *Hpa*I, *Hpa*II, *Kpn*I, *Mbo*I, *Mse*I, *Rsa*I, *Ssp*I และ *Taq*I จำแนกเชื้อไฟโตพลาสมา 16 SrIII ออกเป็นกลุ่มย่อย 16 SrIII-J ซึ่งพบว่าเชื้อในกลุ่มย่อย 16 SrIII-J เป็นเชื้อใหม่ที่เข้ามาอาศัยอยู่ในพืชวงศ์ Solanaceae (*S. melongena* L.) (Mello, Eckstein, Flores, Kreyci, & Bedendo, 2011) นอกจากนั้นยังพบการประยุกต์ใช้รูปแบบของแถบดีเอ็นเอที่เกิดจากการวิเคราะห์ข้อมูลลำดับนิวคลีโอไทด์ของบริเวณ ITS, ยีน TUB (b-tubulin), ACT (actin) และ GPDH (glyceraldehyde-3-phosphate dehydrogenase) ตัดด้วยเอนไซม์ 13 ชนิด ได้แก่ *Alu*I, *Apa*I, *Bam*HI, *Dpn*I, *Hae*III, *Hind*III, *Hin*fI, *Mse*I, *Pst*I, *Pvu*II, *Rsa*I, *Sma*I และ *Taq*I โดยจำลองการตัดของเอนไซม์บนปฏิบัติการคอมพิวเตอร์ที่ให้ผลผลิต PCR ขนาด 275-300 คู่เบส ของเชื้อรา *Colletotrichum gloeosporioides sensu lato* ที่ก่อโรคแอนแทรคโนส (Anthracnose) ในมะละกอ พบว่า เกิดรูปแบบของแถบดีเอ็นเอที่มีความแตกต่างกัน สามารถบ่งชี้ชนิดของเชื้อราดังกล่าวได้ ซึ่งเป็นประโยชน์ต่อการปรับปรุงพันธุ์มะละกอได้ต่อไป (Ramdeen & Rampersad, 2013)

มหาวิทยาลัยบูรพา Burapha University