บทนำ

ประเทศไทยตั้งอยู่ในเขตร้อนชื้น และมีสภาพทางภูมิศาสตร์ที่เอื้อให้มีความหลากหลายทางชีวภาพสูง ซึ่งรวมถึงทรัพยากรทางทะเล ประเทศไทยน่าจะมีสิ่งมีชีวิตอยู่ประมาณ 124,003 ชนิด หรือร้อยละ 8.9 ของสิ่งมีชีวิตอยู่ในโลกทั้งหมด ซึ่งรวมถึงชนิดพันธุ์ที่ทราบชื่อแล้ว และที่ยังไม่พบ (วิสุทธิ์ ใบไม้, 2548) อย่างไรก็ตาม กิจกรรมของมนุษย์กำลังคุกคามความหลากหลายทางชีวภาพดังกล่าว ซึ่งเปรียบเสมือนต้นทุนทางทรัพยากรของประเทศ ดังนั้นด้วยสายพระเนตรกว้างและยาวไกล พระบาทสมเด็จพระเจ้าอยู่หัว ทรงเห็นถึงความสำคัญของการอนุรักษ์พันธุกรรมพืช โดยทรงเริ่มดำเนินงานพัฒนาและอนุรักษ์ทรัพยากรธรรมชาติ และความหลากหลายทางชีวภาพ ตั้งแต่ปี 2503 เป็นต้นมา โดยมีพระราชดำริให้ดำเนินการสำรวจรวบรวมปลูกดูแลรักษาพรรณพืชต่างๆ ที่หายากและกำลังจะหมดไป ต่อมาในปี พ.ศ. 2535 สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี ได้ทรงประสานพระราชปณิธานต่อโดยมีพระราชดำริกับนายแก้วขวัญ วัชโรทัย เลขาธิการพระราชวัง ให้ดำเนินการอนุรักษ์พืชพรรณของประเทศโดยพระราชทานให้โครงการส่วนพระองค์ฯ สวนจิตรลดาเป็นผู้ดำเนินการจัดสร้างธนาคารพืชพรรณขึ้น ตั้งแต่ปี พ.ศ. 2536 เป็นต้นมา

สัตว์กลุ่มกุ้ง ปู และกั้งนับเป็นสัตว์ที่มีคุณค่าทางเศรษฐกิจ (กรมประมง, 2544) และสำคัญต่อระบบนิเวศน์ โดยเป็นส่วนหนึ่งของห่วงโซ่อาหาร โดยเฉพาะอย่างยิ่งกลุ่มที่มีขนาดเล็ก อย่างไรก็ตาม การจัดจำแนกหมวดหมู่ของกุ้ง ปู และกั้งยังมีความสับสนอยู่มาก และมีการปรับเปลี่ยนการเรียกชื่ออยู่เสมอ อีกทั้งยังมีสัตว์ชนิดใหม่ๆ ที่เพิ่งจะถูกค้นพบ จึงมีความจำเป็นที่จะต้องมีการรวบรวมและเรียบเรียงให้การจัดหมวดหมู่สัตว์กลุ่มนี้ไปในทิศทางเดียวกัน แม้ว่าการจัดจำแนกโดยใช้ลักษณะภายนอก (morphology) ยังมีความจำเป็นในการแยกชนิดสิ่งมีชีวิต และเป็นวิธีมาตรฐานที่นักอนุกรมวิธานใช้ทั่วโลก แต่เครื่องมือดังกล่าวก็มีข้อจำกัดในบางกรณี เนื่องจากบางลักษณะที่ใช้แยกจีนัส หรือสปีชีส์อาจไม่ได้สะท้อนถึงความสัมพันธ์ทางวิวัฒนาการ ปัจจุบันจึงมีการพัฒนาเครื่องมือใหม่ๆ ในระดับดีเอ็นเอ เพื่อช่วยในการจัดจำแนกและจัดลำดับความสัมพันธ์ทางวิวัฒนาการของสัตว์กลุ่มกุ้ง ปู และกั้งทะเล (เช่น Kim and Abele, 1990; Tudge and Cunningham, 2002; Schubart et al., 2006) นอกจากนี้ประชาคมวิทยาศาสตร์ระดับโลก ได้เล็งเห็นความจำเป็นในการพัฒนาฐานข้อมูลดีเอ็นของสิ่งมีชีวิตประเภทต่างๆ ที่ผ่านการทดสอบและเทียบความถูกต้อง (verify) กับตัวอย่างอ้างอิง (voucher specimen) เพื่อให้การจำแนกชนิดเป็นไปอย่างรวดเร็ว สำหรับการใช้งานที่อาจจะหาผู้เชี่ยวชาญได้ยาก เช่น การตรวจสอบชนิดพืชและสัตว์ที่นำเข้า หรือ การคัดกรองศัตรูพืช หรือ เชื้อก่อโรค (เช่น Consortium on the barcode of life; http://www.barcoding.si.edu/)

การศึกษานี้ใช้ตัวแทนยีนของทั้งจากไมโตคอนเดรีย (16SrRNA) และจากในนิวเคลียส (18SrRNA) ในการสะท้อนรูปแบบความสัมพันธ์ทางวิวัฒนาการของสัตว์กลุ่มกุ้ง ปู และกั้ง

เนื่องจากยีนแต่ละประเภทมีความโดดเด่นเฉพาะตัวที่อาจให้รูปแบบความสัมพันธ์ทางวิวัฒนาการที่เหมือนหรือแตกต่างกัน โดยยีนบนไมโตคอนเดรียดีเอ็นเอ มักมีอัตราการเกิดการแทนที่นิวคลีโอไทด์สูงกว่าดีเอ็นเอในนิวเคลียส (ประมาณ 5-10 เท่า) ไม่มีการรวมยีนหรือการแลกเปลี่ยนสารพันธุกรรม (Recombination) และมีการถ่ายทอดสารพันธุกรรรมจากแม่สู่ลูก (Avise, 2004) ซึ่งต่างจากยีนในนิวเคลียสที่มี recombination และ ถ่ายทอดสู่รุ่นลูกจากทั้งทางพ่อและแม่ การใช้เครื่องหมายพันธุกรรมทั้งสองประเภทน่าจะสามารถสะท้อนภาพความสัมพันธ์ทางวิวัฒนาการของกลุ่มอนุกรมวิธาน (taxonomic unit) ที่มีความใกล้เคียงกัน ได้แม่นยำกว่าการเครื่องหมายพันธุกรรมประเภทเดียว การวิเคราะห์ที่ดีที่สุดคือการใช้ข้อมูลหลายๆ ประเภทรวมถึง สัณฐาน ดีเอ็นในนิวเคลียส และดีเอ็นเอในไมโตคอนเดรีย

รายงานฉบับนี้จะเสนอผลการวิจัยที่แล้วเสร็จในปีที่ 1 ของการวิจัย นั่นคือ การวิเคราะห์ความหลากหลายทางพันธุกรรมของปูน้ำเค็มทั้งสิ้น 14 สปีชีส์

วัตถุประสงค์ของโครงการวิจัย (ปีที่ 1)

1. พัฒนาเครื่องมือทางชีวโมเลกุลในการแยกชนิดปู ที่พบบริเวณเกาะแสมสาร เกาะแรด เกาะฉางเกลือ และเกาะขาม อำเภอสัตหีบ จังหวัดชลบุรี
2. พัฒนาฐานข้อมูลพันธุกรรมระดับชนิดของปู ที่พบบริเวณเกาะแสมสาร เกาะแรด เกาะฉางเกลือ และเกาะขาม อำเภอสัตหีบ จังหวัดชลบุรี
3. วิเคราะห์ความสัมพันธ์ทางวิวัฒนาการของบางกลุ่มอนุกรมวิธานของปู ที่พบบริเวณเกาะแสมสาร เกาะแรด เกาะฉางเกลือ และเกาะขาม อำเภอสัตหีบ จังหวัดชลบุรีโดยใช้ลำดับนิวคลีโอไทด์ของยีนบนไมโตคอนเดรีย และยีนในนิวเคลียส

กรอบแนวคิดงานวิจัย และเอกสารที่เกี่ยวข้อง

ความสัมพันธ์เชิงวิวัฒนาการของสิ่งมีชีวิต เป็นความรู้พื้นฐานในการเข้าใจความสัมพันธ์ระหว่างสิ่งมีชีวิต ในอดีตนักอนุกรมวิธานมักพึ่งพาสัณฐาน (ซึ่งในสัตว์กลุ่มกุ้ง ปู และกั้ง ลักษณะที่สำคัญจะรวมถึงสีลำตัว ระยางค์ รูปร่าง และพื้นผิวของเปลือก) ในจัดจำแนกชนิด ซึ่งการอนุมานความสัมพันธ์ทางวิวัฒนาการจากลักษณะสัณฐานเพียงอย่างเดียว อาจทำให้แผนภูมิความสัมพันธ์ที่ได้คลาดเคลื่อนไปจากความเป็นจริง เนื่องจากลักษณะสัณฐานที่ใช้ในการแยกชนิดบางอย่างอาจไม่ได้สะท้อนความสัมพันธ์เชิงวิวัฒนาการ หรืออัตราการเปลี่ยนแปลงทางวิวัฒนาการ ดั้งนั้นจึงมีการนำเครื่องหมายพันธุกรรมมาช่วยใช้ในการจัดจำแนกและจัดลำดับความสัมพันธ์ทางวิวัฒนาการของหน่วยอนุกรมวิธาน (เช่นจีนัส หรือ สปีชีส์) ของสิ่งมีชีวิต

เครื่องหมายพันธุกรรมและความสัมพันธ์ทางวิวัฒนาการ

เครื่องหมายพันธุกรรม ซึ่งหมายถึงสารพันธุกรรมที่บ่งชี้ความแตกต่างระหว่างชนิด ประชากร หรือ บุคคล ที่ใช้กันอย่างแพร่หลายในปัจจุบันมีหลากหลายประเภท เช่น โมเลกุลของโปรตีน ชิ้นส่วนดีเอ็นเอที่สร้างโปรตีน (ยีน) และ ชิ้นส่วนดีเอ็นเอที่ไม่สร้างโปรตีน ทั้งที่อยู่ในในนิวเคลียส และ ออร์กาแนล เช่นไมโตคอนเดรีย เครื่องหมายพันธุกรรมที่เหมาะสำหรับการแยกครอบครัว สกุล และสปีชีส์ ควรเป็นเครื่องหมายฯ ที่มีอัตราการกลายที่เหมาะสม ไม่ช้า หรือเร็วเกินไป ข้อสำคัญในการศึกษาความสัมพันธ์ทางวิวัฒนาการของสิ่งมีชีวิตในระดับโมเลกุลนั้นคือ การเลือกยีนหรือชิ้นส่วนดีเอ็นเอที่ใช้ให้มีอัตราการเกิดวิวัฒนาการ (Rate of Evolution) เหมาะสมกับระดับความสัมพันธ์ทางวิวัฒนาการระหว่างกลุ่มสิ่งมีชีวิตที่สนใจ โดยยีนที่มีอัตราวิวัฒนาการช้า เช่นยีน 16S rRNA และ COI บนไมโตคอนเดรีย จะเหมาะกับการตรวจสอบความสัมพันธ์ทางวิวัฒนาการของสิ่งมีชีวิตระหว่างชนิดและระหว่างสกุล (Billington, 2003; Baldwin et al., 1998; Chu et al., 1999; Maggioni et al., 2001)

เครื่องหมายฯ ที่นิยมใช้ในการจัดความสัมพันธ์ทางวิวัฒนาการได้แก่ ยีน 12SrRNA (Robel et al., 2007), 16S rRNA (เช่น Baldwin et al., 1998; Lavery et al., 2004; Maggioni et al., 2001) และ COI (Consortium on the barcode of life, Hebert et al. 2004; วนิสา กัณฑมณี, 2550) บนไมโตคอนเดรีย และ ยีน 18S rRNA ในนิวเคลียส (เช่น Kim and Abele, 1990; Tudge and Cunningham, 2002) การศึกษานี้ได้ใช้ตัวแทนยีนของทั้งจากไมโตคอนเดรียและจากในนิวเคลียสเนื่องจาก ยีนแต่ละประเภทมีความโดดเด่นเฉพาะตัวที่อาจให้รูปแบบความสัมพันธ์ทางวิวัฒนาการที่เหมือนหรือแตกต่างกัน

ไมโตคอนเดรียดีเอ็นเอในสัตว์เป็นดีเอ็นเอรูปวงแหวนขนาดเล็ก (Double-Stranded Loop DNA) ขนาดประมาณ 14,000-26,000 คู่เบส ประกอบด้วยยีนที่สังเคราะห์โปรตีน 13 ยีน ยีนที่สร้าง rRNA 2 ยีน และ t rRNA 22 ยีน ตามลำดับ โดยมีการกลายพันธุ์สูงสุดในส่วนของ Control Region (D-Loop) และต่ำสุดในส่วนของยีนที่ควบคุมการสร้าง RNA ไมโตคอนเดรียดีเอ็นเอมีอัตราการเกิดการแทนที่นิวคลีโอไทด์สูงกว่าดีเอ็นเอในนิวเคลียส (ประมาณ 5-10 เท่า) ไม่มีการรวมยีนหรือการแลกเปลี่ยนสารพันธุกรรม (Recombination) และมีการถ่ายทอดสารพันธุกรรรมจากแม่สู่ลูก คุณสมบัติต่าง ๆ เหล่านี้ทำให้สารพันธุกรรมในไมโตคอนเดรียสามารถเป็นเครื่องหมายพันธุกรรมที่เหมาะสมในการประเมินความสัมพันธ์ทางวิวัฒนาการระหว่างสิ่งมีชีวิตที่มีบรรพบุรุษร่วมกันได้ (Avise, 2004; Billington, 2003; Lowe et al., 2004)

ส่วนในนิวเคลียส มีความหลากหลายของรูปแบบดีเอ็นเอให้เลือกค่อนข้างมาก โดยมีทั้งยีนและส่วนที่ไม่ใช่ยีน แต่ในการศึกษาวิวัฒนาการระดับสปีชีส์ หรือ หมวดหมู่ที่ใหญ่กว่านี้ มักใช้ยีน

เนื่องจาก มีระดับการกลายที่ไม่สูงจนกลบความแตกต่างระหว่างชนิด ทั้งนี้ศักยภาพของยีนแต่ละยีนในการจัดหมวดหมู่ค่อนข้างจะเฉพาะเจาะจงกับสายทางวิวัฒนาการหนึ่งๆ (lineage) และจำนวนกลุ่มอนุกรมวิธานที่อยู่ในฐานข้อมูลมีจำนวนน้อยกว่าข้อมูลยีนบนไมโตคอนเดรีย อย่างไรก็ตาม ยีนในนิวเคลียสก็มีคุณสมบัติที่น่าสนใจ กล่าวคือเป็นยีนที่สามารถเกิด recombination และการถ่ายทอดจากทั้งทางพ่อและแม่ ยีนในนิวเคลียสที่มีการใช้ในกลุ่มปู ได้แก่ยีน 18SrRNA (Ahyong et al. 2007), Arginine Kinase (Mahon and Neigel, 2008) และ ยีน Histone3 (Schubart and Reuchel, 2010)

กลไกการเปลี่ยนแปลงระดับโมเลกุลที่ทำให้เกิดความหลากหลายทางพันธุกรรมในสิ่งมีชีวิต ได้แก่การกลายพันธุ์ (Mutation) โดยแบ่งรูปแบบของการกลายพันธุ์เป็น 2 รูปแบบ คือ ระดับโครโมโซม และระดับนิวคลีโอไทด์ โดย การกลายพันธุ์ระดับโครโมโซม รวมถึง

1. การย้ายสลับที่ระหว่างชิ้นส่วนโครโมโซมต่างคู่ (Translocation) ส่งผลกระทบต่อหน้าที่ของยีน
2. การกลับทิศของชิ้นส่วนโครโมโซม (Inversion)
3. การมีบางส่วนของโครโมโซมขาดหายไป (Deletion) หรือเกินมา (Insertion or Duplication)
4. การเปลี่ยนแปลงจำนวนโครโมโซม

ในขณะที่ การกลายพันธุ์ระดับยีน เป็นการเปลี่ยนแปลงรูปแบบของยีนจาก อัลลีลหนึ่งไปเป็นอีกอัลลีลหนึ่งซึ่งเป็นผลมาจากการเปลี่ยนแปลงนิวคลีโอไทด์ในดีเอ็นเอ ซึ่งได้แก่ การแทนที่คู่เบส (Nucleotide or Base Substitution) และ การสูญเสียหรือการเพิ่มนิวคลีโอไทด์ (deletion and insertion) การแทนที่คู่เบส คือ การแทนที่เบสในสายโพลีนิวคลีโอไทด์ของดีเอ็นเอ แบ่งออกเป็น 2 ชนิด คือ

1. การแทนที่ระหว่างเบสกลุ่มเดียวกัน (Transition) คือ การเกิดการย้ายโปรตอนหรืออิเล็กตอนของไฮโดรเจนภายในโมเลกุลของเบสเอง ทำให้เบส อะดีนีน (Adenine - A) สลับแทนที่กับ กัวนีน (Guanine - G) ในกลุ่มเบสพิวรีน (Purine) หรือ ไซโตซีน (Cytosine - C) สลับแทนที่กับไทมีน (Thymine - T) ในกลุ่มเบสไพริมิดีน (Pyrimidine)
2. การแทนที่ระหว่างเบสต่างกลุ่ม (Transversion) มักเกิดในช่วงการซ่อมแซมตัวเองของดีเอ็นเอ เป็นการแทนที่เบสในกลุ่มพิวรีน (A, G) ด้วยเบสกลุ่มไพริมิดีน (C, T) หรือเกิดในทางกลับกัน

การกลายพันธุ์อาจเปลี่ยนแปลงหรือไม่เปลี่ยนแปลงการแสดงออกของยีน อย่างไรก็ตาม การเกิดการกลายพันธุ์ที่มีการเปลี่ยนแปลงโครงสร้าง เช่น การเพิ่ม การสูญเสีย และการกลับทิศของสายดีเอ็นเอเกิดขึ้นได้ยากกว่าการเกิดการแทนที่คู่เบส (Lowe et al., 2004)

กุ้ง และปูบางชนิดที่พบที่บางเกาะของหมู่เกาะแสมสาร ได้แก่กุ้งดีดขัน (ครอบครัว Caridae) ปูในครอบครัว Portunidae, Grapsidae และ Xanthidae (นงนุช ตั้งเกริกโอฬาร, ติดต่อส่วนตัว) สำหรับสัตว์กลุ่มดังกล่าวได้มีรายงานข้อมูลความสัมพันธ์ทางวิวัฒนาการโดยใช้เครื่องหมายพันธุกรรมอยู่บ้าง ซึ่งงานทั้งหมดเป็นงานของต่างประเทศ Mathews and Anker (2009) ได้รายงานความสัมพันธ์ของ species complex ของกุ้งดีดขัน (snapping shrimp, ครอบครัว Alpheidae) จีนัส Alpheidae บริเวณ ทะเลคาริเบียน, อ่าวเมซิโก, ฟลอริด้า, บราซิล และทะเลแปซิฟิคฝั่งตะวันออกโดยใช้ลำดับนิวคลีโอไทด์ของยีน 16S rRNA และ COI บนไมโตคอนเดรีย และยีน myosin heavy chain ในนิวเคลียส กุ้งเหล่านี้มีความแตกต่างกันเพียงแค่สีของลำตัว และการมี/ไม่มีแถบบริเวณก้ามเท่านั้น ผู้วิจัยพบว่า ใน species complex นี้ประกอบไปด้วยอย่างน้อย 19 สปีชีส์

ตัวอย่างการศึกษาความสัมพันธ์ของปูในครอบครัว Portunidae ได้มีรายงานในประเทศจีน (เช่น Chu et al. 1999) ทวีปอเมริกา (เช่น Robels et al. 2006) และ บริเวณอ่าวเมซิโกและมหาสมุทรแอตแลนติคฝั่งตะวันตก (เช่น Mantelatto et al., 2007) Chu et al. (1999) ได้วิเคราะห์ความสัมพันธ์ทางวิวัฒนาการของปูจีนัส *Charybdis* 4 ชนิดโดยใช้ยีน COI บนไมโตคอนเดรีย ซึ่งผู้วิจัยพบการแบ่งแยกชนิดที่สอดคล้องกับสัณฐาน โดย *C. a*ḥ*nis* และ *C. japonica* มีความสัมพันธ์ใกล้ชิดกันมากกว่าชนิดอื่น (*C. acuta* และ *C. feriatus*) ในทวีปอเมริกา Robels et al. (2006) ใช้ 16S และ 12S rRNA ในการแยกสปีชีส์ในจีนัส *Callinectes* ซึ่งประกอบไปด้วย 16 ชนิดในบริเวณที่ศึกษา (แยกโดยสัณฐาน) และพบว่ามีเพียง 12 ชนิดที่น่าจะเป็นสปีชีส์ที่แท้จริง

บ่อยครั้งที่การศึกษาทางพันธุกรรม ให้ผลสอดคล้องกับ การแยกชนิดโดยใช้สัณฐาน เช่น การศึกษาของ Chu et al. (1999) อย่างไรก็ตาม ข้อมูลพันธุกรรมบางชุด ก็ให้ผลไม่สอดคล้องกับลักษณะทางสัณฐาน เช่น ความสัมพันธ์ของปูบางสปีชีส์ภายในจีนัส *Callinectes* หรือ จีนัส *Portunus* นอกจากนี้ยีนบนไมโตคอนเดรีย อาจให้ผลที่แตกต่างจากยีนในนิวเคลียส ดังนั้นเพื่อให้ได้ภาพความสัมพันธ์ที่ใกล้เคียงกับความเป็นจริง จึงควรต้องใช้ข้อมูลหลายชุดไม่ว่าจะเป็นข้อมูลสัณฐาน ยีนบนไมโตคอนเดรีย และ ยีนในนิวเคลียส

<u>การวิเคราะห์ความสัมพันธ์ทางวิวัฒนาการ</u>

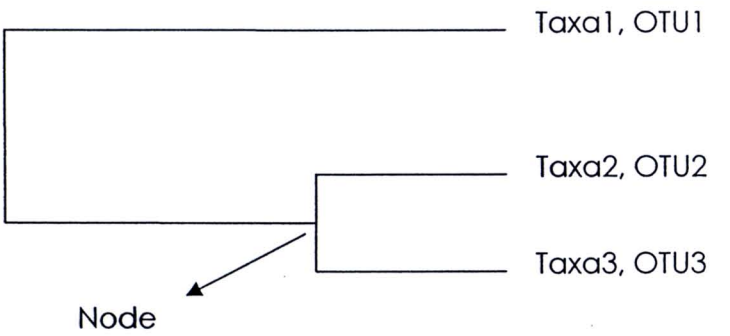
การจัดหมวดหมู่ของกลุ่มทางอนุกรมวิธาน (operating taxon unit หรือ OUT) หรือ taxa (เช่น ครอบครัว สกุล และ สปีชีส์) ให้สอดคล้องกับความสัมพันธ์ทางวิวัฒนาการ มักจะ

แสดงการจัดกลุ่ม (cluster) ในลักษณะของแผนภูมิแผนภูมิต้นไม้ (phylogenic tree, ภาพที่ 1) โดยกลุ่มอนุกรมวิธานที่อยู่ใน cluster เดียวกันจะมีความสัมพันธ์มากกว่า กลุ่มอนุกรมวิธานที่อยู่กลุ่มอื่น แผนภูมิต้นไม้จะประกอบด้วยส่วนที่เป็นปม (Nodes) โดยยอด (Tips) หรือปมภายนอก (Terminal Nodes) แสดงถึง OTUs และกิ่ง (Branches) ที่เชื่อมระหว่างปมจะแสดงถึงความสัมพันธ์ทางวิวัฒนาการ ที่มีบรรพบุรุษร่วมกันระหว่าง OTUs (ภาพที่ 1) โดยรูปแบบของแผนภูมิต้นไม้ทางวิวัฒนาการในลักษณะนี้ว่า Topology

การสร้างแผนภูมิต้นไม้ จำเป็นต้องใช้วิธีการทางคณิตศาสตร์ที่ค่อนข้างจำเพาะ (เช่น Hillis and Moritz 1990; Lemey et al. 2009) โดยใช้ข้อมูล 2 ประเภทคือ

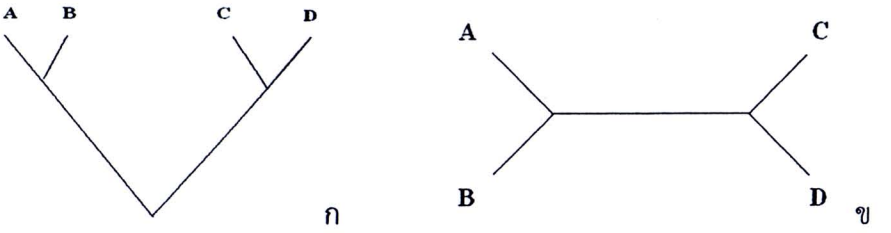
1. ข้อมูลที่จัดเป็นหมวดหมู่ได้ชัดเจน (discrete data) ซึ่งอาจเป็นใช้ได้ทั้งข้อมูลสัณฐาน และ ข้อมูลสารพันธุกรรม เช่น การมี/ไม่มีลักษณะหนึ่งๆ (หนามที่ก้ามหนีบของปู หรือ รูปร่างคาราเปส ของปูต่างชนิด) การมี/ไม่มีนิวคลีโอไทด์เบสหนึ่งๆ บนสายดีเอ็นเอ และการมีลำดับเบสที่แตกต่างกันที่ตำแหน่งเดียวกันระหว่างสายนิวคลีโอไทด์ โดย OTUs ที่มีความคล้ายคลึงกันมากจะมีความสัมพันธ์ทางวิวัฒนาการใกล้ชิดกันมากกว่า OTUs อื่น

2. ดัชนีความเหมือน/แตกต่าง (distance data) โดยการแปลงชุดข้อมูล ระหว่าง คู่กลุ่มอนุกรมวิธาน เป็นตารางของค่าความห่างทางพันธุกรรม (Distance Matrix) ทั้งนี้การคำนวณระยะห่างทางพันธุกรรมอาจใช้โมเดลทางคณิตศาสตร์ที่สะท้อนอัตราทางวิวัฒนาการของยีนหนึ่งๆ OTUs ที่มีค่าความห่างทางพันธุกรรมน้อยที่สุดเป็นกลุ่มที่มีความสัมพันธ์ทางวิวัฒนาการใกล้ชิดกันมากที่สุด

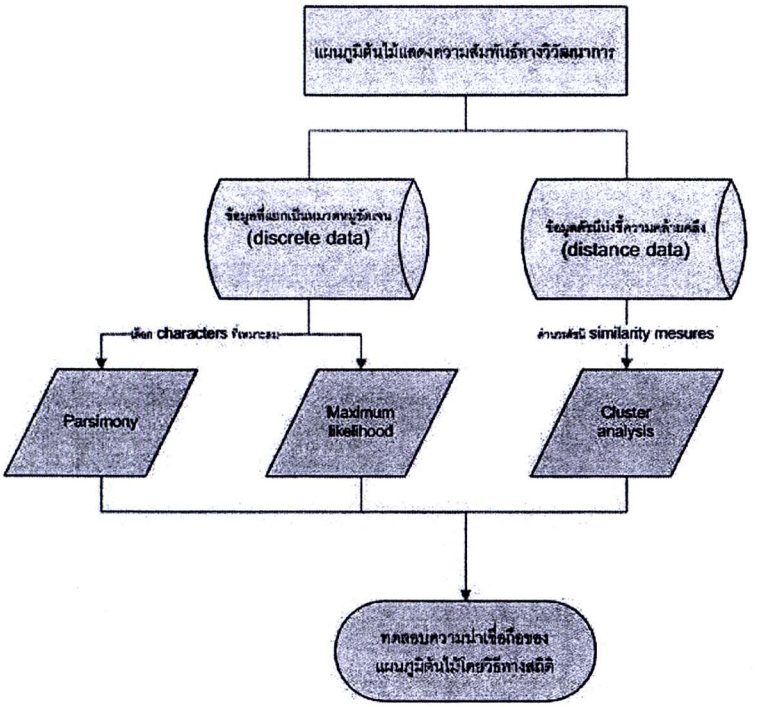


ภาพที่ 1 ส่วนประกอบของแผนภูมิต้นไม้แสดงความสัมพันธ์ทางวิวัฒนาการ (phylogenetic tree)

Phylogenetic Tree สามารถแบ่งออกเป็น 2 ประเภท คือ Rooted Tree (ภาพที่ 2ก.) และ Unrooted Tree (ภาพที่ 2ข.) โดยการสร้าง Phylogenetic Tree ส่วนใหญ่มักสร้างในรูป Unrooted Tree ก่อน ซึ่งจะแสดงความสัมพันธ์ระหว่าง OTUs ที่ทำการศึกษาแต่ไม่สามารถระบุตำแหน่งของการมีบรรพบุรุษร่วมได้ ส่วน Rooted Tree เป็นการตรึงราก Unrooted Tree กับกลุ่มอนุกรมวิธานที่มีความห่างทางพันธุกรรม(Outgroup) กับ OTUs เพื่อตรึงแผนภูมิแสดงความสัมพันธ์ทางวิวัฒนาการ (ศิราวุธ กลิ่นบุหงา, 2544; Halliburton, 2004)



ภาพที่ 2 ลักษณะแผนภูมิต้นไม้ทางวิวัฒนาการแบบ ก. Rooted Tree ข. แบบ Unrooted Tree



ภาพที่ 3 แนวการสร้างแผนภูมิต้นไม้โดยใช้วิธีทางคณิตศาสตร์ และ ประเภทของข้อมูลที่ใช้

เมื่อพิจาณาข้อมูลแบบใดแบบหนึ่ง (สัณฐาน ดีเอ็นในนิวเคลียส หรือ ดีเอ็นเอบนไมโตคอนเครีย) อาจได้แผนภูมิต้นไม้หลายรูปแบบที่อาจสอดคล้อง หรือ ขัดแย้งกับการใช้ข้อมูลอีกชุดหนึ่ง การศึกษานี้จะทดสอบสมมติฐานโดยการยอมรับ หรือ ปฏิเสธ รูปแบบแผนภูมิต้นไม้ที่สร้างจากข้อมูลชุดหนึ่งๆ (สัณฐาน หรือ พันธุกรรม) และ ข้อมูลทั้งสองชุดรวมกัน

การสร้างแผนภูมิต้นไม้ มักใช้วิธีการทางคณิตศาสตร์ และ สถิติ 3 แบบ (Hillis and Moritz 1990; ภาพที่ 3) คือ (1) maximum parsimony คือ การสร้างต้นไม้ที่ใช้จำนวนการเปลี่ยนแปลงของลักษณะน้อยที่สุด (2) maximum likelihood คือ การหาต้นไม้ที่มีค่าความน่าจะเป็นมากที่สุดโดยคิดจากอัตราการเปลี่ยนแปลงของลักษณะ เช่น อัตราการแทนที่ของนิวคลีโอไทด์ ณ ตำแหน่งหนึ่งๆ และ (3) การจัดกลุ่ม (cluster analysis) โดยใช้ข้อมูลระยะห่างทางพันธุกรรม

1. Maximum Parsimony Method

วิธีนี้ประเมินความสัมพันธ์ทางวิวัฒนาการจากข้อมูลที่เป็น Discrete Data โดยใช้จำนวนการเปลี่ยนแปลงของลักษณะ ซึ่งในการวิเคราะห์ความแตกต่างลำดับนิวคลีโอไทด์ของยีนตำแหน่งเดียวกันในสิ่งมีชีวิตนั้น ตำแหน่งนิวคลีโอไทด์ ที่สามารถนำมาสร้างแผนภูมิต้นไม้ทางวิวัฒนาการได้ต้องเป็น Informative Site นั่นคือต้องมีเบสแตกต่างกันอย่างน้อย 2 เบส และพบเบสที่แตกต่างในอย่างน้อย 2 OTUs โดยแผนภูมิต้นไม้ที่ดีที่สุดจะมีจำนวนการเปลี่ยนแปลงของนิวคลีโอไทด์น้อยที่สุดในทุก ๆ ตำแหน่งที่ใช้ในการวิเคราะห์ (Halliburton, 2004; Swofford and Sullivan, 2009) โดยวิธีการหา Tree แบบ Heuristic Search จะสร้าง Core Tree เริ่มต้นก่อนขึ้นด้วยวิธีที่เรียกว่า Stepwise Addition Algorithm จากนั้นจะทำการสลับตำแหน่งของกิ่งเพื่อหา Tree ที่ดีที่สุดและเพิ่มตัวอย่างเข้าไปใน Core Tree แบบสุ่มตัวอย่าง (Random) โดยพิจารณาตามหลักของ Maximum Parsimonyโดยต้นไม้ที่ดีที่สุด คือการจัดกลุ่มที่มีการเปลี่ยนแปลงทางวิวัฒนาการที่น้อยที่สุด

2. Maximum Likelihood Method

วิธีนี้ประเมินความสัมพันธ์ทางวิวัฒนาการของสิ่งมีชีวิตจากการคำนวนค่าความน่าจะเป็นของอัตราการเปลี่ยนแปลงของลักษณะ (เช่น การแทนที่นิวคลีโอไทด์) เพื่อหาแผนภูมิต้นไม้แสดงความสัมพันธ์ทางวิวัฒนาการที่มีค่าความน่าจะเป็นสูงสุด ซึ่งส่วนใหญ่มักใช้ข้อมูลลำดับนิวคลีโอไทด์ คำนวนค่าความน่าจะเป็นของ Phylogenetic Tree จากผลรวมของค่าความน่าจะเป็นของการแทนที่นิวคลีโอไทด์ ณ ตำแหน่งนั้น (Lowe et al., 2004) โดยอัตราการเกิดการแทนที่นิวคลีโอไทด์ต่างรูปแบบ (Transition/ Transversion) จะมีค่าเท่ากันหรือแตกต่างขึ้นอยู่กับแบบจำลองการ

แทนที่นิวคลีโอไทด์ (Model of Nucleotide Substitution) เช่น Jukes-Cartor One Parameter และ Kimura's Two Parameter เป็นต้น (Halliburton, 2004; Lowe et al., 2004)

3. Distance Method

วิธีนี้ประเมินความสัมพันธ์ทางวิวัฒนาการโดยการจัดหมวดหมู่ OTUs ตามระยะห่างทางพันธุกรรม (Genetic Distance) โดยคู่ OTUs ที่มีค่าความห่างทางพันธุกรรมน้อยสุดจะมีความสัมพันธ์ทางวิวัฒนาการใกล้ชิดกันมากที่สุดและจะถูกจัดเป็น OTUs ใหม่ 1 OTUs เพื่อนำมาคำนวนหาค่าความห่างทางพันธุกรรมระหว่างคู่ OTUs ในรูป Distance Matrix ใหม่อีกครั้ง ทำเช่นนี้จนครบทุก OTUs จะได้ความสัมพันธ์ทางวิวัฒนาการระหว่าง OTUs ที่ทำการศึกษาทั้งหมด โดยมักนิยมใช้วิธีการจัดกลุ่มแบบ Neighbor-Joining (Saitou & Nei, 1987) ในการสร้าง Phylogenetic Tree (Halliburton, 2004)

การประเมินค่าความเชื่อมั่นในการประเมินความสัมพันธ์ทางวิวัฒนาการนั้น นิยมวิธี Bootstrapping ซึ่งเป็นการสุ่มจากข้อมูลลำดับนิวคลีโอไทด์เดิม โดยเป็นการสุ่มและแทนที่นิวคลีโอไทด์ในแต่ละตำแหน่งในชุดข้อมูลเดิมจนได้เป็นชุดข้อมูลลำดับนิวคลีนิวคลีโอไทด์ใหม่ เพื่อนำไปสร้าง Phylogenetic Tree ใหม่ด้วยวิธีการสร้างแผนภูมิเดิม ขบวนการนี้จะทำซ้ำหลายครั้งจนได้แผนภูมิต้นไม้หลากหลาย Topology ซึ่งท้ายที่สุดจะได้แผนภูมิต้นไม้ที่แสดงค่าความเชื่อมั่นในรูปร้อยละของรูปแบบแผนภูมิที่สร้างขึ้นใหม่ ซึ่งจะสัมพันธ์กับข้อมูลลำดับนิวคลีโอไทด์เดิม โดยกิ่งที่มีค่าความเชื่อมั่น 100% น่าจะแสดงแม่นยำของการจัดกลุ่ม (Felsenstein, 1985)

การเลือกหารูปแบบการจัดกลุ่มที่พบใน Phylogenetic Tree เป็นผลสรุปจาก Tree ทั้งหมด ซึ่งอาจเกิดจากการทดสอบ Bootstrapping หรือเป็นผลมาจากการสร้าง Tree จากวิธีต่าง ๆ เช่น Maximum Parsimony แล้วได้ Tree จำนวนหนึ่งที่มีสัณฐานของ Phylogenetic Tree แตกต่างกัน แต่มีจำนวนการกลายพันธุ์เท่ากันและเป็นการกลายพันธุ์ที่น้อยที่สุดเช่นกัน วิธีการสร้าง Consensus Tree ที่นิยมใช้กันอยู่มี 2 วิธี คือ วิธี Strict ซึ่งจะสร้าง Consensus Tree ด้วยข้อกำหนดว่ากิ่งใด ๆ ที่ปรากฏใน Consensus Tree จะต้องปรากฏใน Phylogenetic Tree ทั้งหมด (100%) ที่นำมาพิจารณา และวิธี Majority Rule ซึ่งจะสร้าง Consensus Tree ซึ่งมีข้อกำหนดว่ากิ่งที่ปรากฏใน Consensus Tree จะต้องปรากฏเป็นส่วนใหญ่ของ Phylogenetic Tree ที่นำมาพิจารณา