



สำนักงานกองทุนสนับสนุนการวิจัย
THE THAILAND RESEARCH FUND

รายงานวิจัยฉบับสมบูรณ์

โครงการศึกษาอนุพันธุศาสตร์ในแมลงวันผลไม้
โดยใช้ยีนกลุ่มฮีดซ็อคเป็นโมเดล

โดย ผศ. ดร. สุจินดา ธนะภูมิ

วันที่ 24 กันยายน พ.ศ. 2546

รายงานวิจัยฉบับสมบูรณ์

โครงการศึกษาอนุพันธุศาสตร์ในแมลงวันผลไม้
โดยใช้ยีนกลุ่มฮีดซ็อคเป็นโมเดล

ผู้วิจัย ผศ. ดร. สุจินดา ชนะภูมิ
ภาควิชาเทคโนโลยีชีวภาพ, คณะวิทยาศาสตร์
มหาวิทยาลัยมหิดล

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย

(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว. ไม่จำเป็นต้องเห็นด้วยเสมอไป)

สารบัญ

	หน้า
กิตติกรรมประกาศ	3
บทคัดย่อภาษาไทย	4
บทคัดย่อภาษาอังกฤษ	5
เนื้อหางานวิจัย	6
บทนำ	6
วิธีการทดลอง	7- 9
ผลการทดลอง	9-13
บทวิจารณ์	13-15
ตารางที่ 1	16-17
ตารางที่ 2	18-23
ตารางที่ 3	24
ตารางที่ 4	25
รูปที่ 1	26
รูปที่ 2	27
รูปที่ 3	28
รูปที่ 4	29
หนังสืออ้างอิง	30-31
Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.	32-33
ภาคผนวก Reprint ที่ได้ตีพิมพ์ผลงานวิจัยในวารสารระดับนานาชาติ	เริ่ม 34

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณ สกว. ที่ได้สนับสนุนทุนให้ข้าพเจ้าได้ทำงานวิจัยนี้ เพื่อเป็นการพัฒนา
นักวิจัย ขอขอบคุณคณะวิทยาศาสตร์ มหาวิทยาลัยมหิดลที่ให้ทุนสมทบในการทำวิจัย และ
ขอขอบคุณภาควิชาเทคโนโลยีชีวภาพ คณะวิทยาศาสตร์ที่ได้ให้การสนับสนุนแก่นักวิจัยรุ่นใหม่
อย่างเป็นรูปธรรม ข้าพเจ้าขอขอบคุณผู้ทรงคุณวุฒิที่ช่วยเหลือให้คำแนะนำอันเกิดประโยชน์แก่
งานวิจัยชิ้นนี้รวมทั้งช่วยให้เกิดโอกาสที่ข้าพเจ้าสามารถฝึกฝนและปลูกฝังจิตวิญญาณให้นักวิจัย
ใหม่ร่วมไปด้วย

บทคัดย่อ

รหัสโครงการ: RSA/06/2538

ชื่อโครงการ: การศึกษาอนุพันธุศาสตร์ในแมลงวันผลไม้โดยใช้ยีนกลุ่มฮิตซ์ออกเป็นโมเดล

ชื่อนักวิจัย: นายสุจินดา ธนะภูมิ

ภาควิชาเทคโนโลยีชีวภาพ, คณะวิทยาศาสตร์

มหาวิทยาลัยมหิดล

E-mail Address: Testn@mahidol.ac.th

ระยะเวลาโครงการ: 1 กันยายน 2538 ถึง 31 สิงหาคม 2541

DNA เครื่องหมายชนิดหนึ่งได้ถูกวิจัยและพัฒนาขึ้นมาเพื่อให้ความกระจ่างในเรื่องสถานะภาพในการจัดจำพวกและเพื่อแสดงความสัมพันธ์ของฟอร์มต่างๆในกลุ่มแมลงวันผลไม้ที่เรียกรวมๆว่า *Bactrocera tau* (Walker) แมลงวันผลไม้ *B. tau* ในกลุ่มนี้ถูกแบ่งเป็นกลุ่มย่อยคือฟอร์ม *A*, *B*, *C*, *D*, *E*, *F*, *G* และ *I* โดยแบ่งตามชนิดพืชอาศัย, ตามลักษณะโครโมโซมและรายละเอียดทางสัณฐานวิทยาภายนอก DNA เครื่องหมายใหม่นี้ถูกพัฒนามาจากรหัสนิวคลีโอไทด์ของยีน *heat shock protein 70 cognate (Bthsc1)* ส่วนของยีนนี้ได้รับการวิเคราะห์จากแมลงตัวอย่างที่เป็นตัวแทนของแมลงในแต่ละฟอร์มต่างๆดังกล่าว โดยพบว่ารูปแบบของความแปรผันของรหัส DNA เมื่อคำนวณเป็นระยะห่างทางพันธุกรรมเฉลี่ยภายในกลุ่มฟอร์ม *A* ด้วยกันจะมีระยะสั้นอย่างมีนัยสำคัญเมื่อเทียบกับระยะห่างทางพันธุกรรมเฉลี่ยเทียบระหว่างกลุ่มฟอร์มต่างๆ การคำนวณความสัมพันธ์ทางพันธุกรรมแบบ unweighted pair-group method with arithmetic average และแบบ neighbor joining analyses ให้ผลการวิเคราะห์ตรงกันว่าแมลง *B. tau* ฟอร์ม *A* ที่กระจายอยู่ทุกภูมิภาคในประเทศไทยมีโครงสร้างพันธุกรรมที่แยกออกจากฟอร์ม *B*, *C*, *D*, *E*, *F*, *G* และ *I* นอกจากนี้ผลการทดลองได้แสดงให้เห็นว่า สามารถใช้ DNA เครื่องหมายชนิด *Bthsc1* เป็นผลสำเร็จในการแยกฟอร์มอื่นของแมลงได้ด้วย

คำหลัก: สปีชีส์ซับซ้อน, แมลงวันผลไม้, *Bactrocera tau*, ยีนฮิตซ์, ดี เอ็น เอ เครื่องหมาย

Abstract

Project Code: RSA/06/2538

Project Title: Molecular genetic studies of tephritid fruit fly through heat shock gene family

Investigator: Sujinda Thanaphum
Department of Biotechnology,
Faculty of Science, Mahidol University

E-mail Address: Testn@mahidol.ac.th

Project Period: September 1st, 1996 to August 31st, 1999

A new molecular marker has been developed to clarify the status and systematic relationships of forms within the *Bactrocera tau* (Walker) taxon. The *B. tau* taxon has recently been subdivided into forms *A*, *B*, *C*, *D*, *E*, *F*, *G* and *I* based on host-plant preferences, cytological differences and external morphologies. This new molecular marker is derived from the sequence of the *heat shock protein 70* cognate gene *Bthsc1*. Fragments of this gene were analyzed from *B. tau* individuals representing each of the different forms. Patterns of sequence variation revealed that the average genetic distance measurement within the *B. tau* form *A* is significantly smaller compared to the other *B. tau* forms. UPGMA and neighbor joining analyses both indicated that *B. tau* form *A* individuals derived from various geographical populations may be reliably separated from other forms *B*, *C*, *D*, *E*, *F*, *G* and *I*. Our results also show that the *Bthsc1* marker may successfully resolve other relationships among these forms.

Keywords: species complex, fruit flies, *Bactrocera tau*, *heat shock protein 70* cognate gene, molecular systematics, molecular marker

เนื้อหางานวิจัย

บทนำ

แมลงวันผลไม้ในจีนัส *Bactrocera* หลายชนิดเป็นแมลงศัตรูพืชเศรษฐกิจในภูมิภาคเอเชียตะวันออกเฉียงใต้ และ *Bactrocera tau* (Walker) ก็เป็นตัวอย่างหนึ่งของแมลงในกลุ่มดังกล่าวที่มีการกระจายระบาคไปทั่ว (Hardy 1973) โดยเริ่มกระจายตั้งแต่เอเชียใต้เช่น ศรีลังกา และภูฐาน มายังเอเชียตะวันออกเฉียงใต้ที่ประเทศไทย, มาเลเซีย, เวียดนาม, ฟิลิปปินส์, และอินโดนีเซีย เข้าไปถึง ตะวันออกไกลเช่น ไต้หวัน และตอนใต้ของประเทศจีน (White and Elson-Harris 1992) แมลง *B. tau* นี้มักจะวางไข่ในผลไม้หลายชนิดโดยทั่วไปแล้วจะเป็นพืชในตระกูลแตง (cucuritaceous) ที่มีความสำคัญทางเศรษฐกิจคือ แตงกวา, บวบ, แตงโม, พักทอง ที่ไม่มีความสำคัญทางเศรษฐกิจเช่นขี้กาต่างๆในป่า อย่างไรก็ตามในเรื่องอนุกรมวิธานการจัดจำพวกให้ชัดเจนของแมลงกลุ่มเหล่านี้ยังไม่ได้รับการพัฒนามาเป็นเวลานานหลายปี (White and Elson-Harris 1992)

เมื่อไม่นานมานี้ มีงานวิจัยที่แสดงหลักฐานเบื้องต้นว่าแมลง *B. tau* อาจประกอบด้วยสปีชีส์ซับซ้อนหลายชนิดที่มีชื่อเดียวกัน (Baimai et al. 2000) และกระจายอยู่อย่างมีรูปแบบ ปัญหาในเรื่องสปีชีส์ซับซ้อนมีความสำคัญกับการควบคุมแมลงศัตรูพืชอย่างมาก ในด้านการบริหารจัดการควบคุมแมลงนั้นหากแมลงมีหลายชนิดและไม่ผสมข้ามพันธุ์กันจะถือว่าเป็นแมลงคนละชนิดตามแนวคิด biological species concept แมลงหลายชนิดนี้จะต้องมีการบริหารจัดการเฉพาะแต่ละชนิดในทางตรงกันข้ามหากแมลงฟอร์มต่าง ๆ มิใช่เป็นคนละชนิดหากแต่ผสมพันธุ์กันได้ปรกติ การควบคุมแมลงโครงการเดียวก็จะสามารถควบคุมแมลงหลายฟอร์มไปพร้อม ๆ กันได้

ในทั่วไปแล้วเมื่อมีการนิสปีชีส์ซับซ้อนนักวิทยาศาสตร์จะอาศัยรูปลักษณะสัณฐานภายนอกหรือโครโมโซมเป็นหลักฐานอย่างเดียวกันก็มักจะไม่พอเพียงและเที่ยงตรงพอที่จะบ่งชี้สปีชีส์ต่าง ๆ ได้นอกจากนี้ยังมีความไม่แน่นอนในการแสดงความสัมพันธ์ทางพันธุกรรมของแมลงที่มีความคล้ายแบบซับซ้อนนี้ด้วย นักวิทยาศาสตร์จึงนำเอาเทคนิคทางอณูชีววิทยามาร่วมใช้งานทางด้านการจัดจำพวกสิ่งมีชีวิต (systematics) เพื่อให้มีการอบความคิดที่กว้างขึ้น (Hillis et al. 1996)

สำหรับงานวิจัยในด้านแมลงวันผลไม้มีการนำเอาเทคนิคอณูชีววิทยามาพัฒนา DNA เครื่องหมายเช่น microsatellite DNA, และ นิวเคลียส DNA หลายชนิดมาเป็นเครื่องหมาย หากแต่มุ่งไปที่การศึกษาระดับประชากรพันธุศาสตร์มิใช่ศึกษาการนิสปีชีส์ซับซ้อน (He and Haymer 1997, Bonizzoni et al. 2000, Yu et al. 2001) เมื่อนำเอาเทคนิคดังกล่าวใช้กับสปีชีส์ซับซ้อนจะพบว่าไม่สามารถใช้แยกแยะสปีชีส์ซับซ้อนได้เช่นเมื่อ McPherson and Han (1997) ใช้รหัส mitochondrial 16 rDNA วิเคราะห์การนิสปีชีส์ซับซ้อนของ *Rhagoletis pomonella* ไม่พบรหัส DNA ที่จะใช้เป็นลักษณะบ่งชี้สปีชีส์ต่าง ๆ ในระบบซับซ้อนได้ เมื่อ Morrow et al. (2000) ใช้รหัส DNA อีกสี่บริเวณวิเคราะห์การนิสปีชีส์ซับซ้อนของ *B. tryoni* ได้พบแต่ความแปรผันของ

รหัสพันธุกรรมเก่าไม่มีความจำเพาะตามชนิดของแมลงวันผลไม้เพียงหนึ่งตำแหน่งเท่านั้น นอกจากนี้ Murajin และ Nakahara (2001) ยังพบว่ารหัส mt rDNA สามารถแยกแยะสปีชีส์ต่างๆโดยทั่วไปของแมลงวันผลไม้ได้หากแต่ไม่สามารถใช้ในกลุ่มกรณีสปีชีส์ซับซ้อน *B. dorsalis* ได้

ในขณะที่กลุ่มสปีชีส์ซับซ้อน *B. tau* ได้ถูกจัดจำพวดย่อยแบบชั่วคราวเป็นฟอร์มได้ 8 ฟอร์มคือ *A, B, C, D, E, F, G* และ *I* โดยอาศัยลักษณะทางนิเวศวิทยาเช่นชนิดของพืชที่แมลงเหล่านี้ใช้วางไข่ (host preferences) ลักษณะทางเซลล์พันธุศาสตร์ (karyotypes) และภายในกลุ่มฟอร์มบางฟอร์มของ *B. tau* จะปรากฏลักษณะรายละเอียดภายนอกเช่นเส้นแถบสีเหลืองบริเวณส่วนอก หรือขนาดและรูปร่างของจุดสีดำที่บริเวณส่วนหลังกับสีข้างของส่วนท้อง (Tigavattananont 1986, Meksongsee et al. 1991, Baimai et al. 2000) ซึ่งลักษณะดังกล่าวสอดคล้องกับลักษณะ mitotic karyotypes ศึกษาโดย Baimai และคณะ (2000) อย่างไรก็ตามลักษณะดังกล่าวไม่มีความพอเพียงที่จะใช้เพื่อตอบปัญหาการแยกแยะสปีชีส์หรือปัญหาความสัมพันธ์ของแมลงวันผลไม้ฟอร์มต่างๆได้

ในการศึกษาของกลุ่มสปีชีส์ซับซ้อน *B. tau* นี้จึงจำเป็นต้องพัฒนาระบบ DNA เครื่องหมายขึ้นมาใหม่ที่สามารถแยกแยะความซับซ้อนของสปีชีส์และสามารถบ่งบอกความสัมพันธ์ทางพันธุกรรมของสปีชีส์ซับซ้อนเหล่านี้ด้วย ในงานวิจัยชิ้นนี้จึงเริ่มการวิจัยและพัฒนา DNA เครื่องหมายจากยีนฮีตช็อกชนิด *Heat shock protein 70* (Hsp70) เพื่อวัตถุประสงค์ดังกล่าว DNA เครื่องหมายที่เข้ารหัสยีน ฮีตช็อกนี้สามารถใช้วิเคราะห์ความสัมพันธ์ทางพันธุกรรมของสิ่งมีชีวิตในระดับอนุกรมวิธานที่สูงๆในหลายกลุ่มสิ่งมีชีวิต เช่นงานวิจัยที่ทำโดย Gupta และ Golding (1993), Graser et al. (1996), Borchellini et al. (1998) และ Budin and Philippe (1998) อย่างไรก็ตามศักยภาพของ DNA เครื่องหมายบริเวณยีนฮีตช็อกนี้ยังไม่มีผู้ใดทดสอบใช้เพื่อแก้ปัญหากรณีสปีชีส์ซับซ้อนหรือแก้ปัญหาสปีชีส์ที่มีความสัมพันธ์ทางพันธุกรรมที่ใกล้ชิดมากมาก่อน

วิธีทดลอง

ตัวอย่างแมลงวันผลไม้ ตัวอย่างแมลงวันผลไม้ที่เป็นตัวแทนของกลุ่มแมลงในฟอร์มต่างๆ (ดังปรากฏในตารางที่ 1) ได้ถูกจัดจำพวกโดยลักษณะพืชที่แมลงฟอร์มต่างๆวางไข่ และลักษณะของเซลล์พันธุศาสตร์ และลักษณะทางสัณฐานวิทยาภายนอกบางประการตามที่ได้การศึกษาไว้ก่อนแล้วใน (Tigavattananont 1986, Meksongsee et al. 1991, Baimai et al. 2000) แผนภูมิศาสตร์ในรูปที่ 1 แสดงสถานที่เก็บตัวอย่างแมลงในประเทศไทย แมลง *B. tau* ฟอร์ม *A* เป็นฟอร์มที่เป็นศัตรูพืชเศรษฐกิจสำคัญเพราะวางไข่ในผลไม้เชิงพานิชย์หลายชนิด (polyphagous) รวมทั้งยังมีการกระจายในทุกเขตภูมิภาคในประเทศไทย ส่วนแมลงวันผลไม้ *B. tau* ในฟอร์มอื่นๆ มักจะมีความสำคัญในทางเศรษฐกิจน้อยเนื่องจากจะเลือกวางไข่ในผลไม้เพียงชนิดเดียวหรือไม่กี่ชนิดชนิด (oligophagous หรือ monophagous) นอกจากนี้จะพบในบริเวณเฉพาะดังใน

แผนที่เป็นส่วนใหญ่ ฟอรัม *B, F, G* เป็นฟอรัมที่อยู่คนละท้องถิ่นและเป็น monophagous ส่วนฟอรัม *D, E*, และ *I* จะพบในภาคใต้เท่านั้นและเป็น oligophagous (Baimai 2000)

แมลงวันผลไม้ฟอรัม *A* จะถูกเก็บมาอย่างละหนึ่งตัวจากประชากรใน 10 ภูมิภาคประเทศแมลงวันฟอรัมอื่น ๆ จะถูกเก็บมาอย่างละหนึ่งตัวในภูมิภาคต่าง ๆ กัน งานวิจัยนี้จะใช้การวิเคราะห์แบบ bootstrap ทางสถิติเพื่อที่จะหลีกเลี่ยงการเบี่ยงเบนของข้อมูลเพราะจำนวนตัวอย่างน้อย (โปรดดูที่การวิเคราะห์รหัส DNA)

การสกัด DNA แมลงวันผลไม้จะถูกนำมาทำให้เย็นด้วยไนโตรเจนเหลวก่อนที่จะถูกบดเพื่อสกัด DNA โดยวิธีของ Lifton ตามที่ได้อธิบายไว้ในงานของ Haymer และคณะ (1992) จากนั้นตะกอน genomic DNA ที่ตกจะถูกละลายในน้ำกลั่นที่ความเข้มข้น 10 ng/ ul.

ปฏิกิริยา Polymerase Chain Reaction Amplification การขยายปริมาณ DNA บริเวณยีนฮิสทอนจะใช้เครื่อง GenAmp Polymerase Chain Reaction (PCR) system 2400 thermal cycler (Applied Biosystems [Thailand] Limited, Bangkok) ในปริมาตรรวม 50 ul โดยปฏิกิริยา PCR ประกอบด้วย 10 ng ของ genomic DNA, 1X *Taq* DNA polymerase buffer (Perkin Elmer, Branchburg, NJ), 200 uM ของ dNTP แต่ละชนิด (Gibco, Gaithersburg, MD), 20 pmol ของ primer แต่ละข้างและหนึ่งยูนิตของ *Taq* DNA polymerase (Perkin Elmer) การกำหนดอุณหภูมิของ PCR ของ primer I และ II (โปรดดูในส่วนผลการทดลอง) คือ เริ่มที่ 94 องศา (°C) เป็นเวลาสองนาที, ตามด้วย 29 รอบของ: 94 องศา หนึ่งนาที, 44 องศา หนึ่งนาที, และ 72 องศา สองนาที

การ PCR ที่ใช้ primer HST I และ HST II (โปรดดูที่ผลการทดลอง) ใช้วิธีเหมือนการ PCR ที่กล่าวไว้ข้างบนยกเว้นอุณหภูมิขณะ annealing เป็น 61 องศา primer คู่นี้ถูกใช้ในการ PCR กับ genomic DNA template จากตัวอย่างแมลงวันทั้งหมดของงานวิจัยนี้ดังปรากฏในตารางที่หนึ่ง นอกจากนี้ primer HST I และ HST II ยังใช้สำหรับแยกรหัส DNA ของยีนฮิสทอน *Bdorchsc* จากจีโนมของแมลงวันผลไม้ชนิด *B. dorsalis* ซึ่งได้ใช้เป็น กลุ่ม outgroup ในการวิเคราะห์รหัส DNA แบบ unweighted pair-group method with arithmetic average. Primer ที่กล่าวมาข้างต้นได้รับการวิเคราะห์ออกแบบโดยใช้โปรแกรม OLIGO (Pychik and Pychik 1989)

การโคลนและถอดรหัส DNA ผลผลิต DNA จากกระบวนการ PCR ที่ใช้ primer I และ II จะถูกโคลนใน DNA พาหะ pCR2.1 โดยใช้ TA cloning kit (Invitrogen, San Diego, CA) ตามวิธีแนะนำของบริษัท ส่วนผลผลิต DNA จากกระบวนการ PCR ที่ใช้ primer HST I และ HST II จะถูกย่อยด้วยเอนไซม์ตัดจำเพาะ *Eco*RI ก่อนที่จะนำไปต่อกับ DNA พาหะ pUC19 แล้ว transform เข้าใน *Escherichia coli* สายพันธุ์ DH5 alpha จากนั้น โคลนที่มี recombinant plasmid จะถูกเลือกโดยใช้ระบบสีขาว/สีน้ำเงินของโคลนนิ่ง การสกัด recombinant plasmid DNA จะทำโดยวิธี boiling preparation ของ Maniatis และคณะ (1982) รหัส DNA ของผลผลิต PCR

ที่โคลนได้จะถูก sequence โดยระบบ automated DNA sequencer (model 377, ABI PRISM, Foster City, CA) DNA เหล่านี้จะถูกถอดรหัสจำนวนหนึ่งครั้งทั้งสองสายยกเว้นกรณีที่มีการอ่านรหัสมีความน่าสงสัยจะทำการถอดรหัสซ้ำอีก

การวิเคราะห์รหัส DNA รหัส DNA ที่ได้จากการถอดรหัสในข้างต้นจะถูกนำมาเทียบเรียงซ้อนกันโดยโปรแกรม CLUSTAL W (Thomson et al. 1994) อัตราส่วนของการเกิดปรากฏการ transition และ tranversion จะถูกคำนวณที่ละคู่เทียบของรหัส DNA จากตัวอย่าง *B. tau* ทั้งหมดและกับ *B. dorsalis* ที่เป็นสปีชีส์ Outgroup ด้วย ระยะห่างทางพันธุกรรมระหว่างข้อมูลข้างต้นจะถูกคำนวณโดยโปรแกรม DNADIST ในโปรแกรม PHYLIP package version 3.5c (Felsenstein 1993) โดยใช้โมเดลระยะห่างทางพันธุกรรมของ Kimura's two parameter (โดยให้มีค่า default ของ transversion ต่อ transversion เป็น 2.0) รูปแผนภูมิ dendrogram แสดงภาพรวมของระยะห่างทางพันธุกรรมของแมลง *B. tau* ฟอรัมต่างๆถูกสร้างโดยเลือก unweighted pair-group method with arithmetic average เป็น algorithm option จากโปรแกรม NEIGHBOR ใน PHYLIP package

phylogenetic tree ชนิด Neighbor-Joining (NJ) ที่กำหนด 50% majority rule เป็น consensus ได้ถูกสร้างขึ้นโดยเริ่มจากการสุ่มตัวอย่างใหม่ในลักษณะ bootstrap ด้วยความถี่ 100 โดยใช้โปรแกรม SEQBOOT ใน PHYLIP package ข้อมูล 100 ชุดที่ได้จากการกระบวนการ bootstrap จะถูกนำมาใช้ผลิต matrix ของระยะห่างทางพันธุกรรมจำนวน 100 matrix โดยใช้ K2P โมเดลจากโปรแกรม DNADIST แผนภูมิ phylogentic ที่ได้จากการวิเคราะห์ข้างต้นจะถูกวาดแบบไม่มี scale โดยโปรแกรม TREEVIEW (Page 1996) ซึ่งจะแสดงค่า bootstrap ไว้ประจำแต่ละ node

ผลการทดลอง

การพัฒนาวิธีการโคลนรหัส DNA ของยีน Heat Shock Protein 70 Cognate จากแมลงวันผลไม้ *B. tau*

รหัส DNA ของยีน *Hsp 70 cognate gene-like* ได้ถูกแยกออกมาโดยใช้ข้อมูลพื้นฐานเดิมจากยีนที่มีลักษณะเหมือนกันคือ *Cerhsc1* ในจีโนมของแมลงวันผลไม้ชนิด Medfly, *Ceratitis capitata* (Wiedemann) ที่เคยศึกษาโดย (Thanaphum และ Haymer 1998) (รูป 2A) ยีน *Cerhsc1* มี open reading frame ที่ไม่มีรหัส codon หยุดอยู่ตรงกลาง มีความยาว 1,962 คู่เบส สามารถคำนวณความยาวของการ translation และมีรหัส amino acid ที่แสดงว่ามี โปรตีนโดเมนที่ถูกเก็บรักษาไว้ (conserve domain) อยู่ด้าน N terminus คือ ATPase โดเมน และด้าน C terminus คือ peptide binding domain งานวิจัยนี้ได้เลือกบริเวณสองบริเวณของยีน *Cerhsc1* คือ บริเวณที่ตรงกับตำแหน่ง amino acid ที่ 138 และ 591 เพื่อใช้เป็นพื้นฐานในการออกแบบ primer I และ primer II แบบ nondegenerate ดังแสดงในบรรทัดข้างล่าง

Primer I 5'-TTA CCA ACG CCG TCA TCA CTG-3'

Primer II 5'-CTA ACC AGT TGG CTG AGA AGG AGC-3'

เมื่อใช้ primer I และ primer II ในการ PCR ก็จะได้ผลผลิต PCR ที่มีขนาด 1.3 kb จากจีโนมของแมลงวันผลไม้ *B. tau* ที่มาจากจังหวัดเพชรบูรณ์ จึงตั้งชื่อ DNA ส่วนนี้ว่า *Bthsc1* และเมื่อนำไปถอดรหัส DNA จะได้ข้อมูลว่า *Bthsc1* มีความยาว 1,356 คู่เบสและเมื่อเทียบเรียงกับ *Cerhsc1* จะมีความเหมือนที่รหัส DNA ถึง 93.50% และมีความเหมือนของรหัส amino acid ที่ถอดได้ถึง 97.18% (มีได้แสดงข้อมูลในรายงาน)

รหัส DNA ของ *Bthsc1* ของแมลง *B. tau* ฟอรัม A อีกสองตัวอย่างที่มาจากประชากรต่างภูมิภาค (ได้แก่จากจังหวัดอุบลราชธานีและจังหวัดน่าน) ได้ถูกวิเคราะห์ด้วยการเทียบเรียงของรหัส DNA ของ *Bthsc1* ทั้งสามตัวอย่างและกับ *Cerhsc1* และการเทียบเรียงในระดับรหัส amino acid ด้วยได้แสดงให้เห็นว่ารหัส DNA บริเวณที่ตรงกับรหัส amino acid ที่ 236 และ 570 มีลักษณะที่ conserve มากทำให้สามารถออกแบบ degenerate PCR primer ที่ชื่อ HST I และ HST II ดังแสดงในบรรทัดข้างล่าง

HST I 5'-TCT CGT CAC CCA TTT CGT TCA (A/G)G-3'

HST II 5'-ATG GA(A/G) GTG CGG TCA GA(C/T) TCA GA-3'

รหัสนิวคลีโอไทด์ในวงเล็บหมายถึงตำแหน่งที่ degenerate

ผลผลิต PCR จากตัวอย่างทั้งหมดคงแสดงไว้ในตารางที่หนึ่งได้มาจากการใช้ PCR primer HST I และ HST II ทั้งสิ้นโดยมีความยาว 1,001 คู่เบส (โปรดรูป 2B) ผลผลิต PCR ทั้งหมดทุกตัวอย่างจะมี ตำแหน่ง *EcoRI* ที่ conserve อยู่สองตำแหน่งห่างกัน 664 คู่เบส (โปรดรูป 2C) ท่อน DNA ขนาด 664 คู่เบสดังกล่าวจากตัวอย่างแมลง *B. tau* ฟอรัม A ของจังหวัด เพชรบูรณ์ได้ถูกใช้เป็น DNA probe สำหรับการวิเคราะห์ แบบ Southern blot กับ restriction digests genomic DNA ของตัวอย่างแมลงวันผลไม้ *B. tau* ฟอรัม A ทั้งหมด ที่ stringency สูงๆและปรากฏผลว่าได้สัญญาณจาก DNA probe ดังกล่าวไว้ข้างต้นเป็น หนึ่งแถบในทุกๆตัวอย่าง (ข้อมูลส่วนนี้มิได้แสดง)

การเทียบเรียงรหัส DNA ของ *Bthsc1* จากตัวอย่างทั้งหมดของแมลงวันผลไม้ *B. tau* ในทุกๆฟอร์ม

ข้อมูลของรหัส DNA จากท่อน *EcoRI* ขนาด 664 คู่เบสข้างต้น จากทุกตัวอย่างแมลงใน ตารางที่หนึ่ง ซึ่งได้มาจากการขั้นตอนการ PCR, ตัดจำเพาะด้วย *EcoRI*, โคลน, และถอดรหัส ได้ถูก นำมาเทียบเรียงด้วยโปรแกรม CLUSTAL W และแสดงสรุปในตารางที่สองเฉพาะตำแหน่งที่มีความ แตกต่าง การเทียบเรียงมีความเชื่อถือได้สูงเนื่องจากไม่มีความจำเป็นที่จะต้องมีย่อว่างเมื่อเทียบเรียง เลข พบว่ามี 55 ตำแหน่งนิวคลีโอไทด์ที่แสดงความหลากหลายจากความยาวทั้งหมด 664 คู่เบส (ประมาณ 8.3%) ตำแหน่งที่มีความหลากหลายปรากฏในลักษณะกระจายไปตลอดความยาวของ DNA เครื่องหมาย ตำแหน่ง DNA ที่หลากหลายส่วนใหญ่เป็นตำแหน่ง synonymous substitution (46 ตำแหน่งหรือประมาณ 83%) และตำแหน่ง synonymous เหล่านี้ส่วนใหญ่แล้วจะเป็นตำแหน่ง นิวคลีโอไทด์ที่สามใน codon (44 ตำแหน่ง จาก 46 ตำแหน่ง)

ในตารางที่สองจะเห็นได้ว่าภายในกลุ่มฟอร์ม A ด้วยกันเองจะมีตำแหน่งอยู่ 16 ตำแหน่งที่ แสดงความแปรผันและมีอยู่ 16 ตำแหน่งดังกล่าวที่มีมากกว่าหนึ่งตัวอย่างที่มีความแปร ผันตำแหน่งนี้ (polymorphic positions) หากศึกษาความแปรผันในกลุ่มฟอร์ม B ถึง I จะพบ จำนวนตำแหน่งแปรผัน (variable sites) ที่มากกว่าคือ 41 ตำแหน่งในจำนวนนี้มีถึง 23 ตำแหน่งที่ polymorphic เมื่อพิจารณาตำแหน่งที่ 222 เบส adenine จะ fix ในฟอร์ม B ถึง I ในขณะที่ตำแหน่ง เดียวกันนี้ในฟอร์ม A ทุกตัวอย่างจะไม่ใช้เบส adenine

จะมีตำแหน่ง nucleotide substitution แบบ transversion ทั้งหมด 15 ตำแหน่งและ transition ทั้งหมด 40 ตำแหน่งเมื่อเปรียบเทียบตำแหน่งดังกล่าวกับ consensus sequence ของ *Bthsc1* ทำให้ค่าเฉลี่ยของ transition/transversion ratio คือ 2.17 ± 0.17 ($N=17$) ภายในฟอร์ม A จะมีตำแหน่ง transition (13 หรือ 81%) มากกว่า transversion (3 หรือ 19%) สัดส่วนของ transversion จะมากขึ้นเมื่อวิเคราะห์ในกลุ่มระหว่างฟอร์ม B ถึง I (12 หรือ 28%) และ transversion ส่วนใหญ่จะเป็น synonymous substitution และอยู่ที่ตำแหน่งที่สามในรหัส codon ยกเว้นตำแหน่งที่ 85

ระยะห่างทางพันธุกรรมและการเปรียบเทียบความเหมือนของ *Bthsc1*

การเปรียบเทียบรหัส DNA จากตัวอย่างที่ระบุไว้วิเคราะห์โดยการวัดระยะห่างทางพันธุกรรมโดยคิด จากการเกิด nucleotide substitution จากตารางที่สามค่าเฉลี่ยของจำนวน nucleotide ที่แตกต่างใน แต่ละคู่เทียบของภายในฟอร์ม A คือ 4.2 ± 1.5 และมีค่าเฉลี่ยของระยะห่างทางพันธุกรรม (ใช้ Kimura's two parameter model คือ 0.0063 ± 0.002 ($N=45$ ในทั้งสองค่าเฉลี่ย)

เมื่อเปรียบเทียบความแตกต่างของรหัส DNA จากระหว่าง ฟอรัม *A* ถึง *I* ในตารางที่ 4 จะได้ ค่าเฉลี่ยของความแตกต่างของรหัสนิวคลีโอไทด์ที่เพิ่มขึ้นเป็น 16 ± 5.98 ค่าเฉลี่ยของระยะห่างทางพันธุกรรมที่เพิ่มขึ้นเป็น 0.0247 ± 0.009 ($N=28$ ในทั้งสองค่าเฉลี่ย)

จากการวิเคราะห์สถิติ *t*-test (2-tailed) for equality of means ค่าเฉลี่ยของระยะห่างทางพันธุกรรมภายในกลุ่มฟอรัม *A* ด้วยกันจะน้อยกว่าค่าเฉลี่ยระยะห่างทางพันธุกรรมของการเปรียบเทียบระหว่างฟอรัม *A* ถึง *I* อย่างมีนัยสำคัญ ($t=-10.276$; $df=29.04$; $P<0.001$).

การเปรียบเทียบค่าเฉลี่ยของจำนวนนิวคลีโอไทด์ที่ต่างกันในรูปแบบภายในกลุ่มฟอรัม *A* และในแบบระหว่างฟอรัม *A* ถึง *I* ก็ให้ผลสอดคล้องทางสถิติกับในเรื่องระยะห่างทางพันธุกรรมข้างต้น คือ จากการวิเคราะห์สถิติ *t*-test (2-tailed) for equality of means ของจำนวนนิวคลีโอไทด์ที่แตกต่างเฉลี่ยภายในกลุ่มฟอรัม *A* จะมีค่าน้อยกว่าจำนวนนิวคลีโอไทด์แตกต่างเฉลี่ยระหว่างฟอรัม *A* ถึง *I* อย่างมีนัยสำคัญ ($t=-10.32$; $df=29.125$; $P<0.001$)

การวิเคราะห์แผนภูมิ dendrogram โดยวิธี Unweighted Pair-Group Method with Arithmetic average และการวิเคราะห์ phylogenetic tree แบบ Neighbor Joining

ข้อมูลเปรียบเทียบระยะห่างทางพันธุกรรมของรหัส DNA *Bthsc1* จากตารางที่สามและสี่ถูกนำมาใช้เป็นข้อมูลวิเคราะห์กลุ่ม (cluster analysis) โดยจะสร้างแผนภูมิ dendrogram เพื่อจัดกลุ่มตัวอย่างแมลงตามความเหมือนกันทางพันธุกรรม การสร้างแผนภูมิในส่วนนี้ใช้วิธี Unweighted Pair-Group Method with Arithmetic average (UPGMA) ผลที่ได้แสดงในรูปที่สามคือ แมลง *B. tau* ฟอรัม *A* ทั้งสิบตัวอย่างจากภูมิภาคทั่วประเทศไทยจะถูกจับกลุ่มกันอยู่แยกออกจากตัวอย่างฟอรัมอื่นๆ อย่างสิ้นเชิง (กรุณาดูรูปที่สาม) กลุ่มฟอรัม *B*, *F*, *G* จะถูกจัดแยกเป็นกลุ่มที่คล้ายกันมากอยู่อีกกลุ่มหนึ่ง (*Bt-Fhsc*, *Bt-Ghsc* and *Bt-Bhsc* จะอยู่ในกลุ่มด้วยกัน) ซึ่งแสดงถึงความใกล้ชิดกันมากทางพันธุกรรม และกลุ่มฟอรัม *C* และ *I* ก็จะอยู่เป็นกลุ่มแยกออกไปอีกกลุ่มหนึ่ง นอกจากนี้กลุ่มแมลงวันผลไม้ชนิด *B. dorsalis* (*Bdorhsc*) ซึ่งเป็นตัวอย่าง outgroup ก็จะพบห่างออกไปสุดตามที่คาดหมาย

การวิเคราะห์สร้าง phylogenetic tree แบบที่ใช้หลัก 50% majority rule เป็น consensus tree จากวิธี neighbor-joining ซึ่งเลือกใช้ Kimura's two parameter โมเดล แสดงไว้ในรูปที่ 4 ในรูป phylogenetic tree นี้มี node มากมายแต่ไม่ใช่ทุก node ที่แสดงค่าความน่าเชื่อถือสูงกว่า 50 % NJ phylogenetic tree นี้ให้ผลการวิเคราะห์เหมือนกับการวิเคราะห์ cluster ใน dendrogram ของรูปที่ 3 เป็นอย่างมากคือกลุ่มตัวอย่างแมลง *B. tau* ฟอรัม *A* ทั้งสิบตัวอย่างจากภูมิภาคทั่วประเทศไทย จะถูกจับกลุ่มกันอยู่แยกออกจากตัวอย่างฟอรัมอื่นๆ อย่างสิ้นเชิงโดยมีความน่าเชื่อถือ 76% (กรุณาดู

รูปที่ 1) กลุ่มฟอร์ม *B*, *F*, *G* จะถูกจัดแยกเป็นกลุ่มที่คล้ายกันมากอยู่อีกกลุ่มหนึ่ง (*Bt-Fhsc*, *Bt-Ghsc* and *Bt-Bhsc* จะอยู่ในกลุ่มด้วยกัน) ซึ่งแสดงถึงความใกล้ชิดกันมากทางพันธุกรรมที่ความน่าเชื่อถือที่ 88% และ 95% และกลุ่มฟอร์ม *C* และ *I* ก็จะอยู่เป็นกลุ่มแยกออกไปอีกกลุ่มหนึ่งที่ความน่าเชื่อถือ 100%

บทวิจารณ์

การออกแบบ PCR primer จากยีน heat shock cognate ได้แสดงผลการทดลองให้เห็นได้ว่าสามารถนำ primer นี้มาใช้งานได้อย่างน่าเชื่อถือโดยสามารถขยายรหัส *Bthsc1* จากตัวอย่างต่างๆ ของแมลงวันผลไม้ในประเทศไทยได้อย่างจำเพาะเจาะจงในทุกภูมิภาคของประเทศไทย เนื่องจาก primer เหล่านี้ออกแบบมาจากบริเวณที่ conserve อย่างมากในกลุ่มแมลงวันผลไม้ *B. tau*, *B. dorsalis*, และ *C. capitata* ผลผลิต PCR ที่ได้สามารถใช้เป็น DNA เครื่องหมายเพื่อใช้ในการวิเคราะห์การจัดจำพวกและตั้งสมมติฐานเรื่องความสัมพันธ์ทางพันธุกรรมของแมลงวันผลไม้ได้ ทำให้เกิดความกระจ่างในเรื่องฟอร์มต่างๆ ที่เคยไม่ชัดเจนในอดีต (Hardy 1997, White and Elson-Harris 1992, Baimai et al. 2000)

การวิเคราะห์ความสัมพันธ์ของแมลงวันผลไม้ฟอร์มต่าง ๆ นั้น ได้ใช้รหัส DNA ของท่อน *EcoRI* ขนาด 664 คู่เบสที่พบได้เสมอในแมลงวันผลไม้ที่อยู่ในบริเวณยีน heat shock cognate โดยขยายแบบ PCR จากนั้นเมื่อนำรหัส DNA ตัวอย่างทั้งหมดมาเทียบเรียงจะพบว่ามีความยาวตำแหน่งที่มีรหัสนิวคลีโอไทด์แตกต่าง (variable site) เป็นจำนวนรวม 55 ตำแหน่ง การเทียบเรียงข้างต้นไม่มีช่องว่างของรหัสนิวคลีโอไทด์ที่เกิดจากกระบวนการ deletion หรือ addition นอกจากนี้ความแปรผันของรหัสนิวคลีโอไทด์จะกระจายทั่วถึงตลอดความยาว 664 คู่เบส คุณลักษณะของ DNA เครื่องหมายที่ค้นพบข้างต้นแสดงว่าแมื่อยีน heat shock protein 70 cognate จะมีการ conserve มากหากแต่แฝงไว้ด้วยความหลากหลายในปริมาณที่เทียบได้กับ DNA เครื่องหมายมาตรฐานที่นิยมใช้กันอย่างแพร่หลายตัวอย่างเช่น Morrow และคณะ (2000) พบว่ามีจำนวนตำแหน่งที่มีรหัสนิวคลีโอไทด์ที่แตกต่างจำนวน 33 ตำแหน่งจากจำนวนรหัสนิวคลีโอไทด์ทั้งหมด 373 ตำแหน่งในยีน *cytb* จากไมโตคอนเดรีย และพบรหัสนิวคลีโอไทด์แตกต่างจำนวน 25 ตำแหน่งจากจำนวนรหัสนิวคลีโอไทด์ทั้งหมด 316 ตำแหน่งในยีน *COII* จากไมโตคอนเดรียเช่นกันเมื่อวิเคราะห์แมลงวันผลไม้ในกลุ่มสปีชีส์จับซ้อนอื่น

เมื่อใช้รหัส *Bthsc1* มาวิเคราะห์ความสัมพันธ์ทางพันธุกรรมโดยใช้วิธี unweighted pair-group method with arithmetic average (Felsenstein 1993) และวิธี Neighbor Joining (Saitou และ Nei 1987) ได้พบว่ากลุ่มแมลงตัวอย่าง *B. tau* ฟอร์ม *A* จะเข้าร่วมเป็นกลุ่มเดี่ยวแยก

ออกจากแมลงวันผลไม้ พอร์มอื่นๆทั้งหมดด้วยความน่าเชื่อถือทางสถิติ 76% (Neighbor Joining analysis) หากใช้การวิเคราะห์แบบใช้ องค์ประกอบทางพันธุกรรมของพอร์ม A ก็จะได้ผลสอดคล้องกันคือจัดกลุ่มพอร์ม A แยกออกมาต่างหาก ตัวอย่างเช่นค่าเฉลี่ยของ nucleotide substitution ที่พบในพอร์ม A จะน้อยกว่าอย่างมีนัยสำคัญทางสถิติกับค่าเฉลี่ยของระหว่าง พอร์มอื่น การวิเคราะห์ระยะห่างทางพันธุกรรมก็เช่นกันคือจะมีค่าเฉลี่ยที่น้อยกว่าเมื่อพิจารณาเทียบภายในพอร์ม A

หากวิเคราะห์รหัส DNA ในระดับละเอียดยิ่งละเอียดตำแหน่งจะพบตำแหน่งแตกต่างจากพอร์มอื่น (different) แต่คงที่ในพอร์ม A ด้วยตัวเอง (fix) เช่นตำแหน่งที่ 222 จะเป็นรหัส adenine ที่ตำแหน่งนี้ในพอร์ม B, C, D, E, F, G, และ I หากแต่ในพอร์ม A ทั้ง 10 ตัวอย่างจะเป็นเบส cytosine ดังนั้นที่ตำแหน่งนี้ในพอร์ม A จะมีลักษณะ synapomorphy

หากนำผลการวิเคราะห์ที่กล่าวมาวมเป็นหลักฐานที่แนะได้ว่า แมลงวันผลไม้ *B. tau* พอร์ม A มีพันธุกรรมที่แยกออกเป็นกลุ่มต่างหากด้วยระยะห่างอันหนึ่ง ระยะที่แยกนี้สื่อได้ว่าพอร์ม A มีสถานะเป็นสปีชีส์แบบ phylogenetic ของตัวเอง (Cracraft 1983) ลักษณะแยกนี้สอดคล้องกับข้อมูลเบื้องต้นที่แสดงลักษณะจำเพาะของโครโมโซมในพอร์ม A (Baimai 2000) อย่างไรก็ตามตัวอย่างแมลงจำนวนมากขึ้นจะช่วยยืนยันสมมติฐานข้างต้นได้

การวิเคราะห์ทาง phylogenetic ยังแนะว่า *B. tau* พอร์ม B, F, และ G เป็นกลุ่มสัมพันธ์ทางพันธุกรรมย่อยแบบใกล้ชิดอีกกลุ่มหนึ่งในกลุ่มนี้แมลงจะอยู่คนละท้องถิ่น (allopatric) หากแต่มีลักษณะทางนิเวศวิทยาเหมือนกันคือจะเลือกวางไข่ในพืชเพียงหนึ่งสปีชีส์ (monophagous) ที่มีไข่พืชตระกูลแตงและสื่อได้ว่าแมลงในกลุ่มพอร์มนี้มีได้เป็นปัญหาทางเศรษฐกิจแต่อย่างใด เมื่อพิจารณาคำแหน่งรหัสนิวคลีโอไทด์ของแมลงในกลุ่มนี้จะพบว่าในตำแหน่งที่ 48, 85, 87, และ 216 มีลักษณะ synapomorphic การพบว่ามีกลุ่มแมลงที่เปลี่ยนพืชที่วางไข่จาก ตระกูลแตงเช่นนี้จะสนับสนุนสมมติฐานที่ว่า การเปลี่ยน host plant ของแมลงมีบทบาทสำคัญที่ทำให้แมลงวันผลไม้ *B. tau* และแมลงวันผลไม้ชนิดอื่นๆ เกิดความแปลกแยกทางพันธุกรรม (Bush 1994) แมลงวันผลไม้ พอร์ม C และ I ก็น่าจะเป็นกลุ่มย่อยอีกกลุ่มหนึ่งจากการวิเคราะห์แบบ neighbor-joining อย่างไรก็ตามการวิเคราะห์ตัวอย่างแมลงเพิ่มจะช่วยพิสูจน์แนวคิดข้างต้นได้

โดยสรุป ขึ้น heat shock cognate มีรหัส DNA ที่สามารถนำมาใช้เป็น DNA เครื่องหมายในการศึกษากรณีสปีชีส์ซับซ้อนได้โดยได้ทำการศึกษาพอร์มต่างๆของแมลงวันผลไม้ *B. tau* เมื่อใช้ DNA เครื่องหมาย *Bthsc1* นี้ ได้ประสบความสำเร็จโดยการแก้ปัญหาในเรื่องความซับซ้อนของสปีชีส์ในบางพอร์มได้ดีกว่าบางพอร์ม เช่นมีความน่าเชื่อถือได้สูงกว่าพอร์ม A จะเป็นพอร์มที่แปลกแยกทางพันธุกรรมออกจากพอร์มอื่น สถานะของพอร์มอื่นๆ (B ถึง I) ยังไม่แน่นอนและสมบูรณ์เพราะมี

จำนวนตัวอย่างเพียงหนึ่งตัวอย่าง อย่างไรก็ตามผลวิจัยที่ได้ก็สอดคล้องกับข้อมูลวิจัยจากแหล่งอื่นในเรื่องพืชอยู่อาศัยและถิ่นฐานวิทยาต่างๆ

เมื่อมองจากแนวคิดทางวิวัฒนาการแล้วการศึกษาวิจัยความสัมพันธ์แบบ phylogenetic ของกลุ่มสปีชีส์ที่มีความใกล้เคียงกันมากอย่าง *B. tau* รูปแบบต่างๆที่ดูเหมือนว่าจะเพียงเกิดการแปลกแยกของสปีชีส์ขึ้น มีความท้าทายต่องานวิจัยในหลายด้าน นอกจากข้อมูลทางด้านชีวโมเลกุลแล้วความแตกต่างในเรื่องพฤติกรรมการสืบพันธุ์, ในเรื่องการคัดเลือกลักษณะของแมลงแต่ละเพศ, เรื่องการเปลี่ยนพืชอยู่อาศัยก็มีความจำเป็นต่อความเข้าใจในเรื่องความสัมพันธ์ของแมลงสปีชีส์จับซ้อนเหล่านี้

สุดท้ายนี้เนื่องจากรหัสของยีนฮิสซอคมีการสแกนไว้ในระดับที่สูง DNA primer ที่พัฒนาขึ้นมาจะนำไปใช้ประโยชน์ในวงกว้างอย่างน้อยก็สำหรับแมลงวันผลไม้สปีชีส์อื่นๆใน family tephritidae และยีนนี้ยังไม่เกิดปัญหาจับซ้อนในเรื่องการเทียบเรียงรหัส DNA ระหว่างตัวอย่างดังปรากฏใน DNA เครื่องหมาย mitochondrial rDNA ที่ใช้ในแมลงวันผลไม้ (Murajin and Nakahara 2001)

ตารางที่1. ตัวอย่างแมลง *Bactrocera tau* และ ขี้อรหัด *Bthsc* ของแต่ละตัวอย่าง

Species (Family)	Locations (Provinces)	Host plants Genus	Sequence code (GenBank accession #)
<i>Bactrocera tau</i> (Walker)			
<i>B. tau</i> form A	Ubon Ratchatani	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Aubhsc</i> (AF 474944)
<i>B. tau</i> form A	Phetchabun	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Apethsc</i> (AF 474945)
<i>B. tau</i> form A	Mae Hong Son	<i>Luffa</i> spp. (Cucurbitaceae)	<i>Bt-Amshsc</i> (AF 474946)
<i>B. tau</i> form A	Phetchaburi	<i>Momordica</i> spp. (Cucurbitaceae)	<i>Bt-Aphhsc</i> (AF 474947)
<i>B. tau</i> form A	Ranong	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Arnpsc</i> (AF 474948)
<i>B. tau</i> form A	Ratchaburi	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Arbhsc</i> (AF 474949)
<i>B. tau</i> form A	Chiang Mai	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Acmhsc</i> (AF 474950)
<i>B. tau</i> form A	Nan	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Ananhsc</i> (AF 474951)
<i>B. tau</i> form A	Tak	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Atkhsc</i> (AF 474952)

ตารางที่ 1 (ต่อ)

Species (Family)	Locations (Provinces)	Host plants Genus	Sequence code (GenBank accession #)
<i>B. tau</i> form <i>A</i>	Kanchanaburi	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-kbhsc</i> (AF 474953)
<i>B. tau</i> form <i>B</i>	Saraburi	<i>Siphonodon</i> spp. (Celastraceae)	<i>Bt-Bhsc</i> (AF 474956)
<i>B. tau</i> form <i>C</i>	Kanchanaburi	<i>Momordica</i> spp. (Cucurbitaceae)	<i>Bt-Chsc</i> (AF 474958)
<i>B. tau</i> form <i>D</i>	Ranong	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Dhsc</i> (AF 474960)
<i>B. tau</i> form <i>E</i>	Ranong	<i>Strychnos</i> spp. (Strychnaceae)	<i>Bt-Ehsc</i> (AF474957)
<i>B. tau</i> form <i>F</i>	Ranong	<i>Hydnocarpus</i> spp. (Flacourtiaceae)	<i>Bt-Fhsc</i> (AF 474954) <i>B.</i>
<i>tau</i> form <i>G</i>	Kanchanaburi	<i>Hydnocarpus</i> spp. (Flacourtiaceae)	<i>Bt-Ghsc</i> (AF 474955)
<i>B. tau</i> form <i>I</i>	Puttani	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Ihsc</i> (AF 474959)
<i>Bactrocera dorsalis</i> (Hendel) (outgroup)			•
<i>B. dorsalis</i>	Bangkok	<i>Terminalias</i> spp.	<i>Bdorhsc</i> (AF 474961)

ตารางที่2 (หน้า 18 ถึง 23) แสดงตำแหน่งรหัสนิวคลีโอไทด์เทียบเรียงโดยเลือกเฉพาะที่ปรากฏความแตกต่างระหว่างตัวอย่างของ *Bthsc1* ต่างๆ

สามแถวแรกเป็นหมายเลขตำแหน่งของรหัสนิวคลีโอไทด์ภายในท่อน *EcoRI* ซึ่งยาว 664 คู่เบสของ *Bthsc1*; แถวที่สี่เป็นตำแหน่งใน codon; แถวที่ห้าบ่งชี้ transitional (ts) หรือ transversional (tv) หรือทั้งสองชนิด (B)เมื่อเทียบแต่ละตำแหน่งกับรหัส *Bthsc1* consensus; แถวที่หกแสดงชนิดการแทนที่นิวคลีโอไทด์ (S) คือ synonymous และ (R) คือ nonsynonymous; จุดต่างๆแสดงถึงรหัสนิวคลีโอไทด์ที่เหมือนกับรหัส *Bthsc1* consensus; ตัวพิมพ์รหัสนิวคลีโอไทด์ที่หนาหมายถึง tv ส่วนตัวพิมพ์ปกติหมายถึง ts

Table 2.

Sequence position	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Sequence position	1	2	3	3	3	4	4	5	5	6	6	6	6	7	7	8	8	8	9	9	9
Sequence position	8	4	0	3	9	5	8	1	4	0	3	6	9	5	8	5	6	7	0	1	9
Codon position	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	2	3	3	1	3
Transition or transversion	ts	ts	tv	tv	ts	tv	ts	tv	ts	ts	ts	ts	tv	ts	ts	tv	ts	tv	ts	ts	ts
Synonymous or non-synonymous	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	R	S	S	S	S
Haplotypes																					
<i>Bt-Ananhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Apethsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aubhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Amshsc</i>	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aphhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arnhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arbhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Acmhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Atkhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Akbhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bthsc</i> 1(form <i>A</i>) consensus	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Fhsc</i>	.	A	.	.	.	T	.	.	.	.	.	.	.	.	.	A	A	G	.	.	.
<i>Bt-Ghsc</i>	.	A	.	.	.	T	.	.	.	.	.	.	.	.	.	A	A	G	.	.	.
<i>Bt-Bhsc</i>	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	A	.	G	.	.	.
<i>Bt-Ehsc</i>	.	A	.	A	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.
<i>Bt-Chsc</i>	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Ihsc</i>	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Dhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bthsc</i> 1 (all samples) consensus	T	G	G	C	C	T	C	A	T	G	C	C	T	G	C	C	G	T	G	C	C
<i>Bdorhsc</i> (outgroup)	.	.	T	.	T	A	T	T	C	A	T	.	A	A	T	.	.	A	A	T	T

Sequence position	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	
Sequence position	0	2	3	5	5	6	8	8	9	9	9	0	1	2	2	2	2	3	3	4	5	6
Sequence position	2	3	9	0	3	5	0	6	2	5	8	5	6	2	4	5	8	0	1	0	5	1
Codon position	3	3	1	3	3	3	3	3	3	3	3	1	3	3	2	3	3	2	3	3	3	3
Transition or transversion	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts
Synonymous or non-synonymous	S	S	R	S	S	S	S	S	S	S	S	R	S	S	R	S	S	R	S	S	S	S
Haplotypes																						
<i>Bt-Ananhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.
<i>Bt-Apethsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aubhsc</i>	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
<i>Bt-Amshsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aphhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arnhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arbhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.
<i>Bt-Acmhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Atkhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Akbhsc</i>	.	.	.	.	.	.	.	C	.	.	G	.	.	.	.	A	.	.	.	.	.	.
<i>Bthsc</i> 1(form <i>A</i>) consensus	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Fhsc</i>	.	.	A	.	.	.	A	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.
<i>Bt-Ghsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	A	.	.	.	.	.
<i>Bt-Bhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	A	.	.	.	.	.
<i>Bt-Ehsc</i>	.	.	.	.	.	.	A	.	.	.	.	.	.	A	.	.	A	.	.	.	.	.
<i>Bt-Chsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	T
<i>Bt-Ihsc</i>	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	T
<i>Bt-Dhsc</i>	C	.	.	.	.	C	.	.	.	.	.	.	.	A	.	A	A	.	.	.	G	C
<i>Bthsc</i> 1 (all samples) consensus	T	C	G	T	C	T	G	G	T	T	G	A	C	C	A	G	G	C	G	T	A	A
<i>Bdorhsc</i> (outgroup)	.	T	.	C	T	.	A	A	C	C	.	.	.	A	.	.	.	.	T	C	G	T

Sequence position	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4
Sequence position	7	0	1	2	4	5	7	8	8	9	9	1	2	3	3	4	5	5	5	6	8
Sequence position	6	6	5	8	1	7	2	1	4	3	6	4	6	5	8	2	0	3	9	5	0
Codon position	3	3	3	1	2	3	3	3	3	3	3	3	3	3	3	1	3	3	3	3	3
Transition or transversion	tv	ts	ts	ts	ts	tv	ts	ts	ts	tv	tv	ts	tv	ts	ts	ts	ts	tv	ts	tv	ts
Synonymous or non-synonymous	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
Haplotypes																					
<i>Bt-Ananhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Apethsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aubhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Amshsc</i>	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aphhsc</i>	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arnpsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arbhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.
<i>Bt-Acmhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Atkhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
<i>Bt-Akbhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bthsc</i> 1(form <i>A</i>) concensus	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Fhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	T	.	A	.	.	.
<i>Bt-Ghsc</i>	.	.	.	.	.	A	.	.	.	.	.	.	.	G	.	T	.	A	.	.	.
<i>Bt-Bhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.
<i>Bt-Ehsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	.	C	.
<i>Bt-Chsc</i>	.	.	.	.	.	C	.	.	.	.	.	.	.	G	G	.	T	.	A	.	.
<i>Bt-Ihsc</i>	.	.	.	G	.	.	.	.	.	.	.	.	.	G	G	.	.	.	A	.	.
<i>Bt-Dhsc</i>	.	.	T	.	C	A	.	.	.	C	.	G	.	.	.	.	.	.	.	.	.
<i>Bthsc</i> 1 (all samples) concensus	A	G	C	C	A	G	G	T	C	T	T	T	A	A	A	C	G	C	T	C	T
<i>Bdorhsc</i> (outgroup)	.	A	T	T	.	.	.	.	T	G	A	.	C	G	.	.	T	.	.	.	.

Sequence position	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6
Sequence position	8	9	9	0	1	3	4	4	4	5	6	6	7	7	0	2	2	3	4	4	5	
Sequence position	9	2	5	3	7	7	5	6	9	2	1	4	3	8	3	1	9	9	5	8	1	
Codon Position	3	3	3	2	1	3	2	3	3	3	3	3	2	3	3	2	3	3	3	3	3	
Transition or transversion	tv	ts	ts	ts	tv	tv	ts	B	tv	tv	ts	ts	ts	ts	ts	B	ts	B	ts	ts	ts	
Synonymous or non-synonymous	S	S	S	R	R	S	R	S	R	S	S	S	S	R	S	S	R	S	S	S	S	
Haplotypes																						
Bt-Ananhsc	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bt-Apethsc	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
Bt-Aubhsc	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
Bt-Amshsc	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A
Bt-Aphhsc	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	
Bt-Arnpsc	G	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	.	.	.	.	.	
Bt-Arbhsc	G	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A
Bt-Acmhsc	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	
Bt-Atkhsc	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	
Bt-Akbhsc	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
Bthsc 1(form A) concensus	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	
Bt-Fhsc	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
Bt-Ghsc	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	A
Bt-Bhsc	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
Bt-Ehsc	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	
Bt-Chsc	.	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	C	.	T	.	.	
Bt-Ihsc	.	T	.	.	.	.	.	.	.	.	C	.	.	G	.	.	C	.	T	.	.	
Bt-Dhsc	.	.	.	.	.	.	A	.	T	.	.	.	.	.	.	C	G	T	.	.	.	
Bthsc 1 (all samples) concensus	T	C	T	G	C	C	C	A	A	T	C	C	C	A	G	A	T	A	C	G		
Bdorhsc (outgroup)	.	.	C	A	G	A	T	T	.	.	.	T	T	.	A	.	G	.	T	.	.	

Sequence position	6 6
Sequence position	5 5
Sequence position	2 4
Codon position	1 3
Transition or transversion	ts tv
Synonymous or non-synonymous	S R
Haplotypes	
<i>Bt-Ananhsc</i>	. .
<i>Bt-Apethsc</i>	. .
<i>Bt-Aubhsc</i>	. .
<i>Bt-Amshsc</i>	. .
<i>Bt-Aphhsc</i>	. .
<i>Bt-Arnghsc</i>	. .
<i>Bt-Arbhsc</i>	. .
<i>Bt-Acmhsc</i>	. .
<i>Bt-Atkhsc</i>	. .
<i>Bt-Akbhsc</i>	. .
<i>Bthsc</i> 1(form <i>A</i>) consensus	. .
<i>Bt-Fhsc</i>	. .
<i>Bt-Ghsc</i>	. .
<i>Bt-Bhsc</i>	. .
<i>Bt-Ehsc</i>	. .
<i>Bt-Chsc</i>	. .
<i>Bt-Ihsc</i>	. .
<i>Bt-Dhsc</i>	. .
<i>Bthsc</i> 1 (all samples) consensus	T A
<i>Bdorchsc</i> (outgroup)	C T

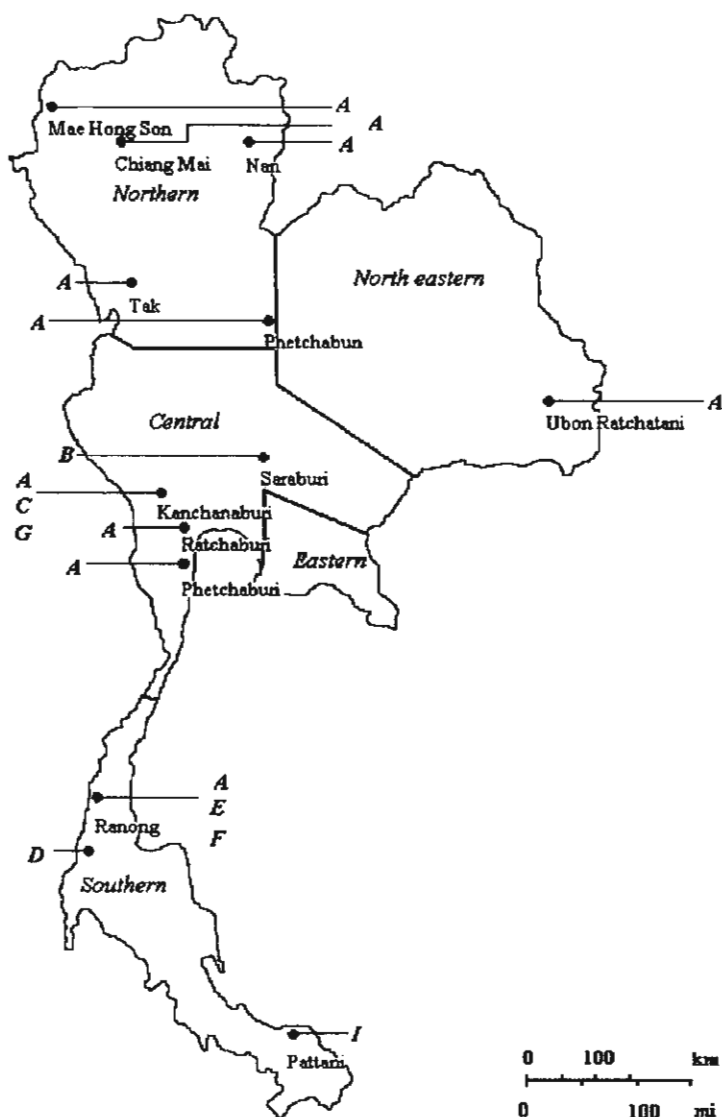
ตารางที่ 3 เมตริกแสดงระยะห่างทางพันธุกรรม (ส่วนทะเลงมุล่าง) และ แสดงจำนวนนิวคลีโอไทด์ที่แตกต่างกัน (ส่วนที่ทะเลงมุลบน) ของรหัส *Bfhscl* ต่างๆ
เมื่อเปรียบเทียบกับใน *B. fau* ฟอรัม *A*

	<i>Br-4ubhsc</i>	<i>Br-4pethsc</i>	<i>Br-4mshsc</i>	<i>Br-4pshsc</i>	<i>Br-4rnhsc</i>	<i>Br-4rbhsc</i>	<i>Br-4cmhsc</i>	<i>Br-4mahsc</i>	<i>Br-4lthsc</i>	<i>Br-4kshsc</i>
<i>Br-4ubhsc</i>		1	4	5	5	5	3	3	4	5
<i>Br-4pethsc</i>	0.0015		3	4	4	4	2	2	3	4
<i>Br-4mshsc</i>	0.0060	0.0045		3	5	5	3	5	6	7
<i>Br-4pshsc</i>	0.0076	0.0060	0.0045		4	6	2	4	5	6
<i>Br-4rnhsc</i>	0.0076	0.0061	0.0076	0.0061		4	2	4	5	6
<i>Br-4rbhsc</i>	0.0076	0.0060	0.0076	0.0091	0.0060		4	6	7	8
<i>Br-4cmhsc</i>	0.0045	0.0030	0.0045	0.0030	0.0030	0.0060		2	3	4
<i>Br-4mahsc</i>	0.0045	0.0030	0.0076	0.0060	0.0061	0.0091	0.0030		3	4
<i>Br-4lthsc</i>	0.0060	0.0045	0.0091	0.0076	0.0076	0.0106	0.0045	0.0045		5
<i>Br-4kshsc</i>	0.0076	0.0060	0.0106	0.0091	0.0091	0.0121	0.0060	0.0060	0.0076	

ตารางที่ 4 เมตริกแสดงระยะห่างทางพันธุกรรม (ส่วนที่แสดงมุมล่าง) และ แสดงจำนวนนิวคลีโอไทด์ที่แตกต่างกัน (ส่วนที่แสดงมุมบน) ของรหัส *BtHsc1* ต่างเมื่อเปรียบเทียบกับระหว่าง *B. tan* ฟอร์มต่างๆ

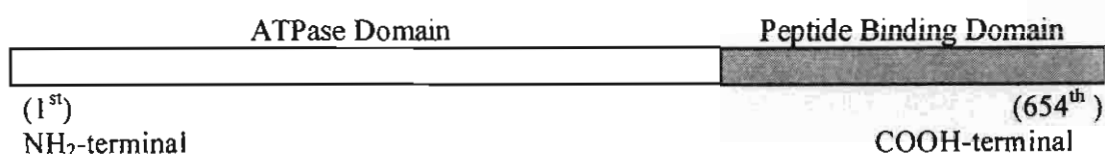
	<i>Bt-Ahsc</i>	<i>Bt-Bhsc</i>	<i>Bt-Chsc</i>	<i>Bt-Dhsc</i>	<i>Bt-Ehsc</i>	<i>Bt-Fhsc</i>	<i>Bt-Ghsc</i>	<i>Bt-Ihsc</i>
<i>Bt-Ahsc</i>		10	11	18	12	14	15	12
<i>Bt-Bhsc</i>	0.0152		17	20	14	6	5	18
<i>Bt-Chsc</i>	0.0168	0.0261		20	17	21	22	5
<i>Bt-Dhsc</i>	0.0276	0.0308	0.0307		21	26	23	21
<i>Bt-Ehsc</i>	0.0183	0.0214	0.0261	0.0323		16	17	20
<i>Bt-Fhsc</i>	0.0214	0.0091	0.0324	0.0403	0.0246		5	22
<i>Bt-Ghsc</i>	0.0230	0.0076	0.0340	0.0356	0.0261	0.0076		23
<i>Bt-Ihsc</i>	0.0183	0.0276	0.0076	0.0323	0.0307	0.0339	0.0355	

หมายเหตุ *Bt-Ahsc* คือ รหัส *BtHsc1* consensus ของ *B. tan* ฟอร์ม *A* ทั้งหมดส่วนข้อรหัสอื่นๆจะเป็นไปตามตารางที่ 1



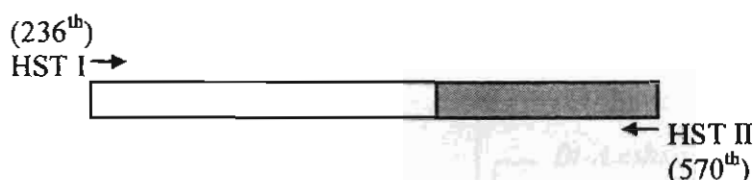
รูปที่ 1. แผนที่ประเทศไทยแสดงตำแหน่งภูมิประเทศที่เป็นที่มาของตัวอย่างแมลงและรหัสนิวคลีโอไทด์ *Bthsc1* ต่างๆในการวิเคราะห์กรณีศึกษาจีโนมของ *Bactrocera tau*

A)



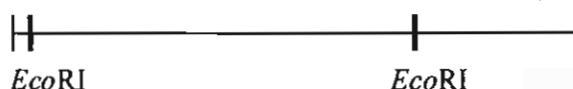
Heat shock protein 70 cognate of *C. capitata* (1,962bp = 654 amino acids)

(B)



Bthsc1 sequence (1,001bp = 334 amino acids)

(C)



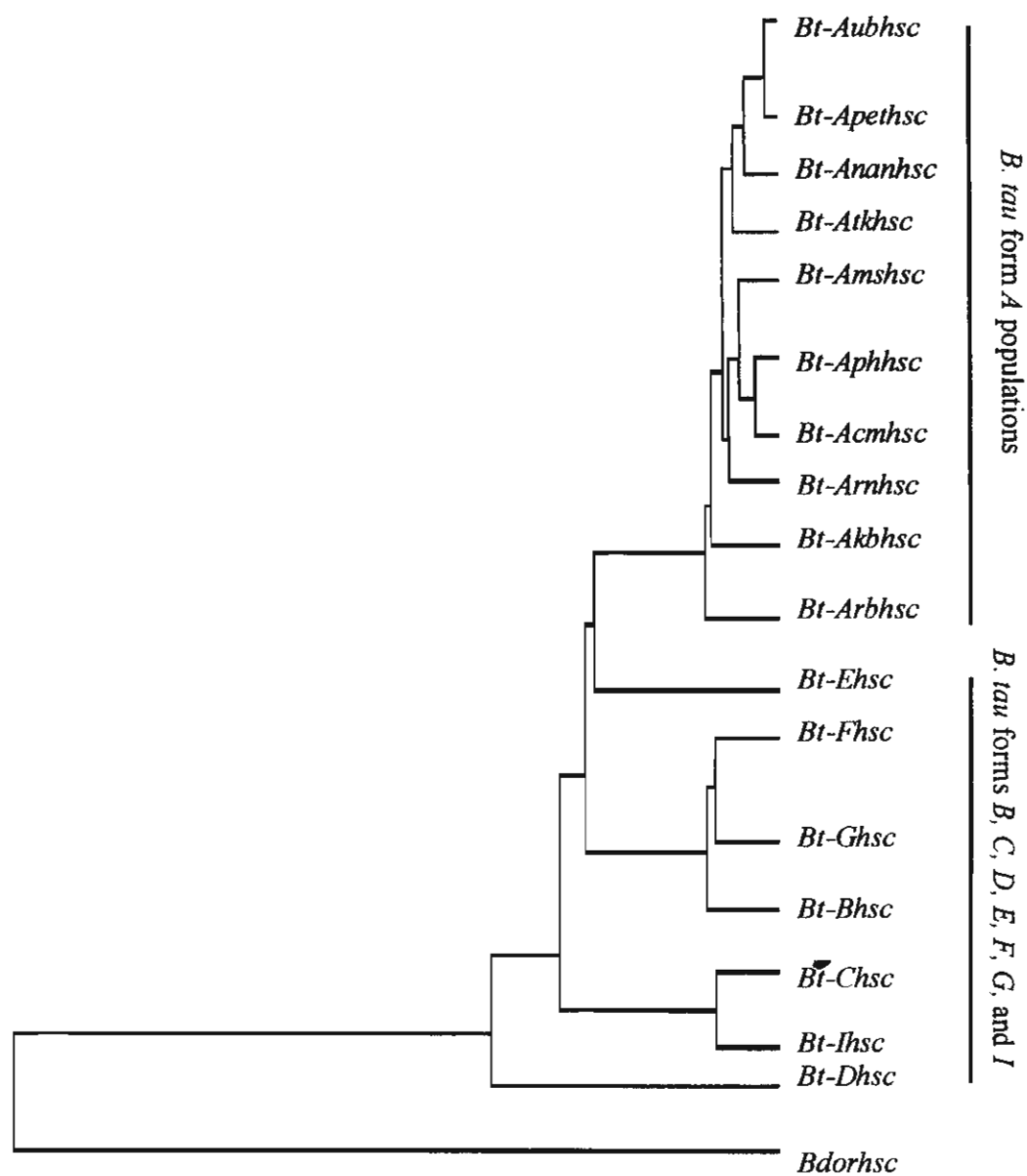
A 664bp *EcoRI* fragment of *Bthsc1* sequence

รูปที่ 2 แผนผังแสดงยีน *Cerhsc1* โดยเขียนให้ตรงตามตำแหน่งกับรหัส *Bthsc1*;

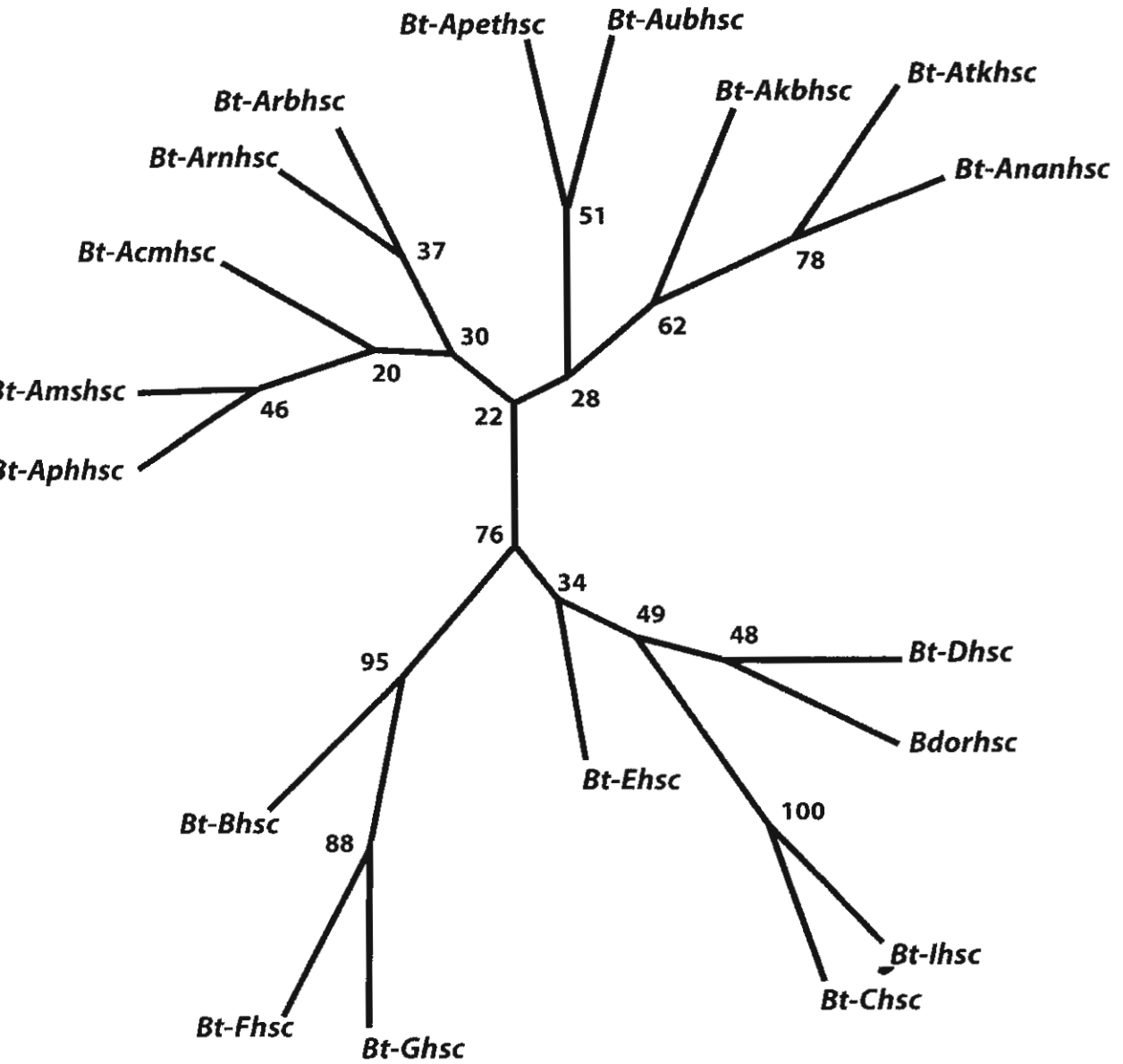
รูป 2A แสดงโปรตีนโดเมนสองบริเวณคือ ATPase โดเมนที่ด้าน N-terminus (ประมาณ 44 kDa; ส่วนที่ไม่เร่ง) และส่วน peptide binding โดเมนที่ด้าน C-terminus (ประมาณ 27 kDa; ส่วนที่เร่ง) (Lindquist และ Craig 1988) ส่วน coding region ของ *Cerhsc1* นี้ยาว 1,962 คู่เบส

รูป 2B แสดงส่วนของรหัส *Bthsc1* ที่ถูกขยายโดยวิธี PCR จาก primer HST I และ HST II จะสังเกตได้ว่าตำแหน่งของ amino acid ที่คาดจากรหัสนิวคลีโอไทด์รวมส่วนของ ATPase โดเมน และส่วนของ peptide binding โดเมนไว้ด้วย ได้ระบุตำแหน่งเกาะของ primers ไว้ด้วยซึ่งจะทำให้ผลิต PCR มีขนาด 1,001 คู่เบส (เท่ากับ 334 amino acids)

รูปที่ 2C แสดงตำแหน่งของ *EcoRI* สองตำแหน่งที่อยู่ห่างกัน 664 คู่เบสภายในผลิต PCR *Bthsc1* ที่มีขนาด 1,001 คู่เบส



รูปที่ 3 แผนภูมิ UPGMA dendrogram ของระยะห่างทางพันธุกรรมระหว่างรหัส *Bthsc1* ต่างๆ เพื่อที่จะอนุมานความสัมพันธ์ของฟอร์มและประชากรต่างๆ ; อัตราส่วนเป็น 0.001 หน่วยของระยะห่างทางพันธุกรรม



รูปที่ 4 แผนภูมิ phylogenetic tree แบบ neighbor-joining ที่มีค่า bootstrap (ทำซ้ำ 100 ครั้ง) กำกับอยู่แต่ละ node; ไม่มีค่าระยะห่างทางพันธุกรรมในแต่ละกิ่ง

หนังสืออ้างอิง

- Baimai, V., J. Phinchongsakuldit, and C. Sumrandee. 2000.** Cytological evidence for a complex of species within the taxon *B. tau* (Diptera: Tephritidae) in Thailand. *Biol. J. Linn. Soc.* 69: 399-409.
- Bonizzoni, M., A. R. Malacreda, C. R. Guglielmino, L. M. Gomsli, G. Gasperi, and L. Zheng. 2000.** Microsatellite polymorphism in the Mediterranean fruit fly, *Ceratitis capitata*. *Insect. Mol. Biol.* 9(3): 251-261.
- Borchiellini, C., N. Boury-Esnault, J. Vacelet, and Y.L. Parco. 1998.** Phylogenetic analysis of the Hsp 70 sequences reveals the monophyly of metazoa and specific phylogenetic relationships between animals and fungi. *Mol. Biol. Evol.* 15(6): 647-655.
- Budin, K., and H. Philippe. 1998.** New insights into the phylogeny of eukaryotes based on ciliate Hsp 70 sequences. *Mol. Biol. Evol.* 15(8): 943-956.
- Bush, G. L. 1994.** Sympatric speciation in animals: new wine in old bottles. *TREE* 9: 285-288.
- Cracraft, J. 1983.** Species concepts and speciation analysis, pp 159-87 *In* Current ornithology. Plenum, New York, New York.
- Drew, R.A.I., and M. C. Roving. 1997.** Overview-Tephritidae in the Pacific and Southeast Asia, pp.46-53. *In* A.J.Allwood and R.A.I. Drew (eds.), Management of fruit flies in the Pacific: a regional symposium. Australian Centre for International Agriculture Research, Canberra, Australia.
- Felsenstein, J. 1993.** PHYLIP: Phylogeny inference package, version 3.5c. The University of Washington, Seattle.
- Graser, R. T., D. Malnar-Dragojevic, and V. Vincek. 1996.** Cloning and characterization of a 70 kD heat shock cognate (*hsc70*) gene from the zebrafish (*Danio rerio*). *Genetica* 98: 273-276
- Gupta, R. S., and G. B. Golding. 1993.** Evolution of HSP70 gene and its implications regarding relationships between archaebacteria, eubacteria, and eukaryote. *J. Mol. Evol.* 37: 573-582.
- Hardy, D. E. 1973.** The fruit Flies (Tephritidae-Ditera) of Thailand and bordering countries. Pacific Insects Monograph 31, Bishop Meseum.
- Haymer, D. S., D. O. McInnis, and L. Arcangeli. 1992.** Genetic variation between strains of the Mediterranean fruit fly, *Ceratitis capitata*, detected by DNA fingerprinting. *Genome* 35: 528-533.
- He, M., and D. S. Haymer. 1997.** Polymorphic intron sequences detected within and between populations of the Oriental fruit fly (Diptera: Tephritidae). *Ann. Entomol. Soc. Am.* 90(6): 825-31.
- Hillis, D. M., C. Moritz, and B. K. Mabel. 1996.** Molecular Systematic 2nd ed. Sinauer, Sunderland, MA.
- Iwahashi, O. 1999.** Distinguishing between the two sympatric species *Bactrocera Corambolae* and *B. papayae* (Diptera: Tephritidae) based on aedeagal length. *Ann. Entomol. Soc. Am.* 92(5): 639-643.
- Lindquist, S., and E. A. Craig. 1988.** The heat shock protein. *Annu. Rev. Genet.* 163: 631-677.
- Mayr, E. 1963.** Animal Species and Evolution. Harvard Univ. Press, Cambridge, Massachusetts.

- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982.** Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- McPheron, B. A., and H-Y. Han. 1997.** Phylogenetic analysis of North America *Rhagoletis* (Diptera : Tephritidae) and related genera using mitochondrial DNA sequence data. *Mol. Phylogenet. Evol.* 7(1): 1-16.
- Meksongsee, B., A. Liewvanich, and M. Jirasuratana. 1991.** Fruit flies in Thailand, pp. 83-98. *In* S. Vijayasegaran and A.G. Ibrahim [eds]., Proceedings, 1st International Symposium on Fruit Flies in the Tropics, 14-16 March 1988, Malaysian Agricultural Research and Development Institute, Kuala Lumpur, Malaysia.
- Morrow, J., L. Scott, B. Congdon, D. Yeates, M. Frommer, and J. Sved. 2000.** Close genetic similarity between two sympatric species of tephritid fruit fly reproductively isolated by mating time. *Evolution* 54(3): 899-910.
- Muraji, M., and S. Nakahara. 2001.** Phylogenetic relationships among fruit flies, *Bactrocera* (Diptera, Tephritidae), based on the mitochondrial rDNA sequences. *Insect. Mol. Biol.* 10(6): 549-559.
- Page, R. D. M. 1996.** TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357-358.
- Psychlik, P., and W. Psychlik. 1989.** OLIGO primer analysis software, version 6.0. Molecular Biology Insights (MBI).
- Saitou, N., and M. Nei. 1987.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.
- Thanaphum, S., and D. S. Haymer. 1998.** A member of the hsp70 gene family from the Mediterranean fruit fly, *Ceratitidis capitata*. *Insect. Mol. Biol.* 7(1): 1-10.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994.** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22 :4673-4680.
- Tigvattananont, S. 1986.** The importance of *Dacus tau* (Walker) in Thailand. *Kaen Kaset* 14: 114-118.
- Yu, H., M. Frommer, M. K. Robon, A. W. Meat, D.C.A. Shearman, and J. Sved. 2001.** Microsatellite analysis of the Queensland fruit fly *Bactrocera tryoni* (Diptera: Tephritidae) indicates spatial structuring: implications for population control. *Bull. Entomol. Res.* 9: 39-148.

Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

1. ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ จำนวนหนึ่งเรื่อง

ชื่อผู้แต่ง: Sujinda Thanaphum และ Urusa Thaenkham

ชื่อเรื่อง: Relationships of Forms within the *Bactrocera tau* (Walker)
(Diptera:Tephritidae) Taxon Based on *Heat Shock Protein 70* Cognate Sequences

ชื่อวารสาร: Annals Entomological Society of America

ปี ค.ศ.: 2003 เล่มที่: 96 เลขที่ 1 หน้า 44-53

2. การนำผลงานวิจัยไปใช้ประโยชน์

เชิงวิชาการ

2.1 โครงการวิจัยนี้ได้มีส่วนร่วมสนับสนุนวิทยานิพนธ์ระดับปริญญาโท สาขาชีววิทยาศาสตร์

แวดล้อม ของนางสาวอุรุษยา แทนขำ เรื่อง Diversity and Evolutionary study of *Bactrocera tau* Species Complex and Populations By Using A 70 kDa Heat Shock Cognate-Like Sequences ปี 2544

2.2 โครงการวิจัยนี้ได้มีส่วนร่วมสนับสนุนวิทยานิพนธ์ระดับปริญญาโท สาขาชีววิทยาศาสตร์

สถานะแวดล้อม ของนายชัชวาล สงวนศิลป์ ในปริมาณที่น้อยกว่า 2.1 เรื่อง
“Molecular characterization of the *hopper* transposable element in *Bactrocera tau* (Walker)” ปี 2544

3. อื่นๆ ได้รับเชิญให้ไปบรรยายและพิมพ์ abstract ในการประชุมวิชาการระดับนานาชาติจำนวน สองครั้งคือ

3.1 บรรยายเรื่อง Molecular Ecology of the *Bactrocera tau* species complex.

ที่การประชุม The 2nd Asia- Pacific Conference on Chemical Ecology.

เมื่อ August 7th-11th, 2001 Penang, Malaysia.

3.2 บรรยายเรื่อง “Molecular phylogeny of the *Bactrocera tau* species complex.”

ที่การประชุม The XXI International Congress of Entomology.

เมื่อ August 20th-26th, 2000 Iguassu Falls, Brazil.

ได้เสนอผลงานวิจัยแบบ poster และพิมพ์ abstract ที่การประชุมระดับนานาชาติจำนวน 2 ครั้ง

3.3 เสนอ poster เรื่อง “*hopper-like transposable elements in tephritid fruitflies and anopheline mosquitoes in Thailand.*” by S. Thanaphum, V. Baimai, and P. Rongnoparut ที่การประชุม The Keystone Symposium on Molecular and Cellular Biology: Toward the Genetic Manipulation of Insects. เมื่อ January 9th–15th, 1998 Taos, New Mexico, USA.

3.4 เสนอ poster เรื่อง “*Muti-approaches toward identification and characterization of Tephritidae: Bactrocera tau species complex.*” by S. Thanaphum, S. Tigavatananont and V. Baimai ที่การประชุม the Fifth International Symposium on Fruit Flies of Economic Important. เมื่อ June 1st–5th, 1997 Penang, Malaysia.

Relationships of Forms within the *Bactrocera tau* (Walker) (Diptera: Tephritidae) Taxon Based on Heat Shock Protein 70 Cognate Sequences

SUJINDA THANAPHUM¹ AND URUSA THAENKHAM

Department of Biotechnology, Faculty of Science, Mahidol University, 272 Rama VI Road, Rajdhavac, Bangkok 10400, Thailand

Ann. Entomol. Soc. Am. 96(1): 44–53 (2003)

ABSTRACT A new molecular marker has been developed to clarify the status and systematic relationships of forms within the *Bactrocera tau* (Walker) taxon. The *B. tau* taxon, previously described as a widely distributed species, has recently been subdivided into forms A, B, C, D, E, F, G, and I based on host-plant preferences, cytological differences and external morphologies. This new molecular marker is derived from the sequence of the *heat shock protein 70* cognate gene *Bthsc1*. Fragments of this gene were analyzed from *B. tau* individuals representing each of the different forms. Patterns of sequence variation revealed that the average genetic distance measurement within the *B. tau* form A is significantly smaller compared with the other *B. tau* forms. Unweighted pair-group method with arithmetic average and neighbor joining analyses both indicated that *B. tau* form A individuals derived from various geographical populations may be reliably separated from other forms B, C, D, E, F, G, and I. Our results also show that the *Bthsc1* marker may successfully resolve other relationships among these forms. For example the *B. tau* forms B, F, and G, which are monophagous on related host plants, also cluster together as a closely related group.

KEY WORDS species complex, fruit flies, *Bactrocera tau*, *heat shock protein 70* cognate gene, molecular systematics, molecular marker

A NUMBER OF SPECIES of tephritid fruit flies are important pests of crops in Southeast Asia. One such species, *Bactrocera tau* (Walker), is widely distributed throughout this region (Hardy 1973). The distribution of this species extends from South Asian countries such as Sri Lanka and Bhutan to the Southeast Asian countries of Thailand, Malaysia, Vietnam, Philippines, and Indonesia and toward the Far East Asian region including Taiwan and South China (White and Elson-Harris 1992). *B. tau* infests a wide range of hosts, generally cucurbits such as the commercially important crops of cucumber, luffa, and melons (Drew and Roming 1997). They have been discovered both in the endemic wild (forest) cucurbitaceous fruits such as *Trichosanthes* spp. as well as in cultivated areas. However, questions regarding the taxonomic status of specimens described as *B. tau* have persisted for many years (White and Elson-Harris 1992).

Recent work on *B. tau* has suggested that this taxon actually represent a complex of widely distributed species (Baimai et al. 2000). The existence of a complex of species has a number of potentially important implications for pest species in particular. From a pest management perspective, the presence of several distinct species, in particular those that are reproduc-

tively isolated as defined by the classical biological species concept (Mayr 1963), implies that control programs may need to be tailored to the specifics of each case. If, however, the defined forms (or groups) do not represent distinct species, it may be possible to design effective control programs that can have an impact on several (or all) of the forms simultaneously.

In the case of species complexes many traditional characters, such as those based on morphological differences, are often unable to provide adequate resolution to determine the status of species and their systematic relationships. Recently, the incorporation of molecular techniques in the analysis of systematic relationships within and between species has provided a new framework for augmenting these traditional approaches (Hillis et al. 1996).

For Tephritid species, a number of DNA-based markers such as microsatellites and various nuclear DNA sequences are available from systematic studies focused at the level of population relationships (He and Haymer 1997, Bonizzoni et al. 2000, Yu et al. 2001). However, for the situation where a species complex exists, the extent to which these markers might be useful for delineating systematic relationships is not well established. McPherson and Han (1997) used mitochondrial 16S rDNA sequence data to analyze the *Rhagoletis pomonella* species complex, but did not

¹ E-mail: testn@mahidol.ac.th.

Table 1. *B. tau* samples and their corresponding *Bthsc* sequences

Species	Locations (provinces)	Host plants genus (family)	Sequence code (GenBank accession #)
<i>B. tau</i> (Walker)			
<i>B. tau</i> form A	Ubon Ratchatani	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Aubhsc</i> (AF 474944)
<i>B. tau</i> form A	Phetchabun	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Apethsc</i> (AF 474945)
<i>B. tau</i> form A	Mae Hong Son	<i>Luffa</i> spp. (Cucurbitaceae)	<i>Bt-Amshsc</i> (AF 474946)
<i>B. tau</i> form A	Phetchaburi	<i>Momordica</i> spp. (Cucurbitaceae)	<i>Bt-Aphhsc</i> (AF 474947)
<i>B. tau</i> form A	Ranong	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Arnshsc</i> (AF 474948)
<i>B. tau</i> form A	Ratchaburi	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Arbhsc</i> (AF 474949)
<i>B. tau</i> form A	Chiang Mai	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Acmhsc</i> (AF 474950)
<i>B. tau</i> form A	Nan	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Ananhsc</i> (AF 474951)
<i>B. tau</i> form A	Tak	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Atkhsc</i> (AF 474952)
<i>B. tau</i> form A	Kanchanaburi	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-kblhsc</i> (AF 474953)
<i>B. tau</i> form B	Saraburi	<i>Siphonodon</i> spp. (Celastraceae)	<i>Bt-Bhsc</i> (AF 474956)
<i>B. tau</i> form C	Kanchanaburi	<i>Momordica</i> spp. (Cucurbitaceae)	<i>Bt-Chsc</i> (AF 474958)
<i>B. tau</i> form D	Ranong	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Dhsc</i> (AF 474960)
<i>B. tau</i> form E	Ranong	<i>Strychnos</i> spp. (Strychnaceae)	<i>Bt-Elhsc</i> (AF 474957)
<i>B. tau</i> form F	Ranong	<i>Hydnocarpus</i> spp. (Flacourtiaceae)	<i>Bt-Fhsc</i> (AF 474954)
<i>B. tau</i> form C	Kanchanaburi	<i>Hydnocarpus</i> spp. (Flacourtiaceae)	<i>Bt-Ghsc</i> (AF 474955)
<i>B. tau</i> form I	Puttani	<i>Trichosanthes</i> spp. (Cucurbitaceae)	<i>Bt-Ihsc</i> (AF 474959)
<i>B. dorsalis</i> (Hendel) (outgroup)			
<i>B. dorsalis</i>	Bangkok	<i>Terminalia</i> spp.	<i>Bdorhsc</i> (AF 474961)

identify diagnostic differences among the members. Morrow et al. (2000) used four markers to study genetic differentiation of the *B. tryoni* complex. They identified a high level of shared polymorphic variation but only minor fixed differences in one marker. Mitochondrial rDNA sequences have recently been developed to resolve phylogenetic relationships among fruit flies in the genus *Bactrocera*. However, these mt rDNA sequences cannot reliably resolve closely related species in the *B. dorsalis* complex (Muraji and Nakahara 2001).

The *B. tau* complex has been temporarily systematized into eight forms designated A, B, C, D, E, F, G, and I based on studies of host-plant preferences, cytological differences and external morphologies (Tigavattananont 1986, Meksongsee et al. 1991, Baimai et al. 2000). Within the complex, members show morphological differences in the yellow stripes on the thorax and the size and shape of dark bands on the dorsal and lateral abdomen (Baimai et al. 2000). These identifications were further analyzed by comparing differences in mitotic karyotypes (Baimai et al. 2000). However to date no characteristics appear adequate to completely resolve questions about the species status and systematic relationships of all these forms.

For the case of *B. tau*, the use of a molecular genetic character may help to clarify relationships within this complex. In this study we have examined the potential of a genetic marker based on *Heat shock protein 70* (Hsp70) gene sequences to be used for this purpose. This marker has been successfully used to study phylogenetic relationships at several higher taxonomic levels, in part because it is a highly conserved gene (Gupta and Golding 1993, Graser et al. 1996, Borchellini et al. 1998, Budin and Philippe 1998). However the potential for this marker to resolve close genetic relationships, such as forms within a species or within a complex of closely related species, has not yet been established.

Materials and Methods

Fruit Fly Specimens. Specimens representing different forms (Table 1) were classified based on host-plant preferences, cytological differences and external morphologies as previously described (Tigavattananont 1986, Meksongsee et al. 1991, Baimai et al. 2000). A geographical map showing the collection locations is shown in Fig. 1. *B. tau* form A is polyphagous and is widely distributed in all regions of Thailand (Fig. 1). The other forms are relatively rare. The *B. tau* forms B, F, and G are allopatric and monophagous. The *B. tau* forms C, D, E, and I are oligophagous and are found only in the Southern region, except form C (Fig. 1; Baimai et al. 2000).

For form A, each sample was a single individual from each of 10 different geographical localities. All other forms were sampled using single individuals from individual geographical localities. We used a bootstrap statistical analysis in an attempt to alleviate sampling bias (see DNA sequence analysis).

DNA Isolation. Before genomic DNA extraction, fruit fly samples were frozen in liquid nitrogen. DNA from individual flies was extracted by using the Lifton method as described by Haymer et al. (1992). Genomic DNA was resuspended in sterile water at a concentration of 10 ng/μl.

Polymerase Chain Reaction Amplifications. Amplifications were performed on a GenAmp Polymerase Chain Reaction (PCR) system 2400 thermal cycler (Applied Biosystems [Thailand] Limited, Bangkok) in a total volume of 50 μl. Amplification reactions contained 10 ng of genomic DNA, 1x *Taq*DNA polymerase buffer (Perkin Elmer, Branchburg, NJ), 200 μM each dNTP mixed (Gibco, Gaithersburg, MD), 20 pmol of each primer, and 1U of *Taq*DNA polymerase (Perkin Elmer). The following cycling profile was used for primers I and II (see results): An initial denaturation at 94°C for 2min, followed by 29 cycles of

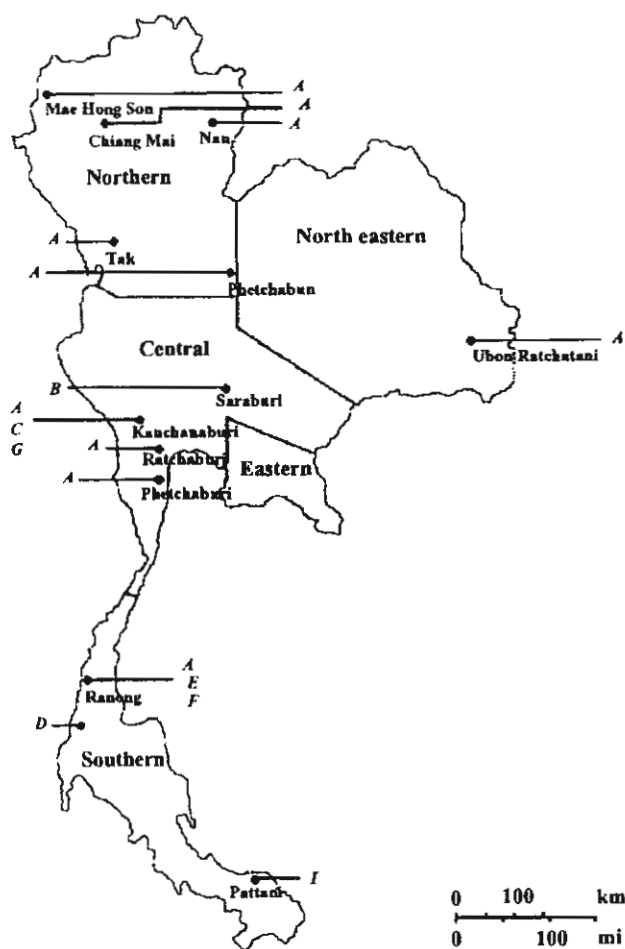


Fig. 1. Map of Thailand showing locations of samples used in the *Bthsc1* sequences analysis of the *B. tau* species complex. The scale bar of 1 cm is equal to a distance of 100 km.

denaturation at 94°C for 1 min, primer annealing at 44°C for 1 min and extension at 72°C for 2 min.

PCR amplifications using primers HST I and HST II (see Results) were carried out as described above except that primer annealing was done at 61°C for 1 min and an extension at 72°C for 2 min was used. These primers and conditions were used for amplification of the genomic DNA templates from all the samples listed in Table 1. The HST I and II primers were also used to isolate a sequence designated *Bdorhsc* from *B. dorsalis* using for use as an outgroup in the unweighted pair-group method with arithmetic average analysis. In all cases, primers were designed using the OLIGO primer analysis software (Pychlik and Pychlik 1989).

Cloning and Sequencing. PCR products generated using primers I and II were cloned into the vector pCR2.1 using the TA cloning kit (Invitrogen, San Diego, CA) according to the manufacturer's instruction. PCR products generated using the HST I and HST II primers were first digested with *EcoRI*. The resulting product was then ligated into the pUC19 vector for transformation into the DH5 α strain of *Escherichia coli*. Insert-bearing plasmids were selected by blue/white screening. Plasmids were isolated using a plasmid boiling preparation method (Maniatis et al. 1982).

Nucleotide sequences were obtained using an automated DNA sequencer (model 377, ABI PRISM, Foster City, CA). Each plasmid was sequenced once from both directions unless sequencing interpretation problems were encountered.

Sequence Analyses. Sequences from the cloned fragments were aligned using the CLUSTAL W program (Thompson et al. 1994). Transition/transversion (ti/tv) ratios were calculated from pairwise comparisons between the sequences obtained from all *B. tau* samples and the *B. dorsalis* outgroup. Genetic distances were calculated by using the DNADIST program in the PHYLIP package version 3.5c (Felsenstein 1993) with the Kimura's two parameter (K2P) model option (with a default value of ti/tv of 2.0). A dendrogram of the genetic relationships of the *B. tau* forms A, B, C, D, E, F, G, and I was constructed using the unweighted pair-group method with arithmetic average (Unweighted Pair Group Method with Arithmetic mean) algorithm option in the NEIGHBOR program in the PHYLIP package.

Neighbor-Joining (NJ) trees based on a 50% majority rule consensus were constructed. For this analysis, the sequence data set was resampled with 100 replications using a bootstrap method by SEQBOOT program in the PHYLIP package. The 100 bootstrapped data sets were used to generate 100 genetic distance matrices based on a K2P model using the DNADIST program. The 50% majority rule consensus tree was drawn as a nonscaled tree with the TREEVIEW program (Page 1996) with a bootstrap value on each node.

Results

Cloning of the Heat Shock Protein 70 Cognate Sequences from *B. tau*. *Hsp* 70 cognate gene-like sequences from *B. tau* were first obtained using information from a similar gene (*Cerhsc1*) from *Ceratitis capitata* (Wiedemann) (Thanaphum and Haymer 1998) (Fig. 2A). The *Cerhsc1* gene has an uninterrupted open reading frame 1962bp long. Conceptual translation of this open reading frame revealed two conserved domains within the protein, the N-terminal ATPase domains and the C-terminal peptide binding domain. Two regions of the *Cerhsc1* sequence, beginning at amino acid positions 138 and 591, were selected for the design of nondegenerate primers (designated I and II). The sequences of these primers are:

Primer I 5'-TTA CCA ACG CCG TCA TCA CTG-3'
Primer II 5'-CTA ACC AGT TGG CTG AGA AGG AGC-3'.

Using these primers, a 1.3-kb band was amplified from genomic DNA of an individual representing *B. tau* form A from Phetchabun. The amplification product, designated as *Bthsc1*, was cloned and sequenced. The resultant 1,356bp of sequence from *Bthsc1* was aligned with the corresponding *Cerhsc1* sequence from *C. capitata*. These sequences showed 93.50% identity at the DNA and 97.18% identity at the amino acid level (data not shown).

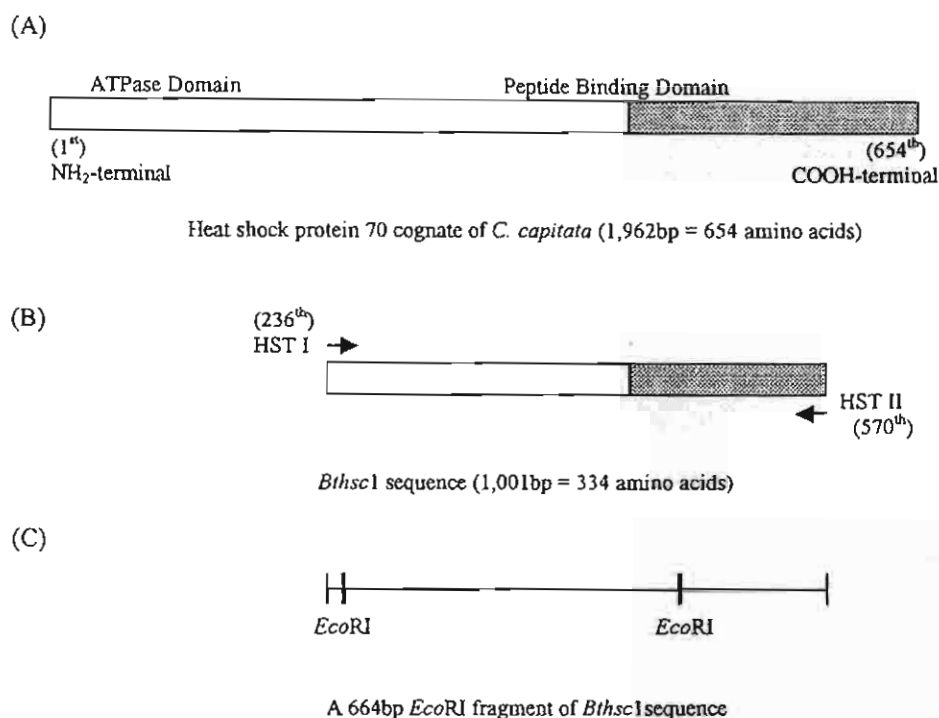


Fig. 2. Schematic representation of the *Cerhsc1* gene in corresponding to the *Bthsc1* sequences. (A) The heat shock protein 70 domains are generally of two major parts; the ATPase domain at the N-terminus (≈ 44 kDa; opened area) and the peptide binding domain at the C-terminus (≈ 27 kDa; closed area) (Lindquist and Craig 1988). This coding region of the *Cerhsc1* gene is 1,962bp. (B) Part of the *Bthsc1* sequences that are amplified by primers HST I and HST II and its corresponding conceptual amino acid may partially include the ATPase domain and the peptide-binding domain. The primer binding sites are indicated. The PCR product is 1,001bp in length (334 amino acids). (C) The *Bthsc1* /1,001bp PCR fragment has two *EcoRI* internal restriction sites generating a 664bp *EcoRI* fragments.

Two additional sequences representing form A individuals from other geographical populations (Ubon Ratchatani and Nan) were also obtained. A multiple sequence alignment of all three *Bthsc1* sequences and the *Cerhsc1* sequence, both at the level of DNA and conceptual amino acid translation, revealed two conserved regions around amino acid positions 236 and 570 of the *Cerhsc1* (data not shown). Degenerate primers were designed from these conserved regions and designated as follows:

HST I 5'-TCT CGT CAC CCA TTT CGT TCA (A/G)G-3'
 HST II 5'-ATG GA(A/G) GTG CCG TCA GA(C/T) TCA GA-3'

The nucleotides in parentheses indicate variable sites.

PCR amplification from specimens representing all the *B. tau* forms listed in Table 1 using the primers HST I and HST II produced amplification products 1,001bp in length (Fig. 2B). All of these products were found to contain two conserved *EcoRI* sites, 664bp apart (Fig. 2C). The 664bp *EcoRI* fragment from *B. tau* form A from Phetchabun was used as a probe for Southern blot analysis of samples of all the *B. tau* form A populations. At high stringency, only a single band was seen (data not shown) in each case. Similar analyses of genomic DNAs from form A populations digested with either *XbaI* or *HindIII* also showed only one band (data not shown).

Alignments of *Bthsc1* Sequences from All *B. tau* Forms. The 664bp *EcoRI* fragments obtained from the different specimens were cloned and sequenced. A summary of the variation revealed by the CLUSTAL alignment of these sequences is shown in Table 2. No gaps are required to maximize these alignments. There are 55 variable nucleotide positions out of 664 nucleotide positions (or 8.3%), and these appear to be randomly distributed throughout this region. Most of the variable positions are synonymous substitution (46 sites or 83%). Of these, almost all are the third codon position (44 sites out of 46 sites).

Table 2 also shows that within the form A samples there are 16 variable sites. Four of these polymorphisms are shared by more than one of the form A samples. Likewise, within the form B to I samples there are 41 variable sites. Of these, 23 occur in more than one of these samples. The A in position 222 appears to be fixed in forms B to I while completely absent from the form A samples.

Overall, 15 transversion and 40 transition changes were found when all samples were compared with the *Bthsc1* consensus sequence. The average ti/tv ratio was found to be 2.17 ± 0.17 ($N = 17$). Within form A, there were more transitions (13 or $\approx 81\%$) than transversions (three or $\approx 19\%$). A higher proportion of transversions was seen in the B to I forms (12 or 28%). The transversion substitutions are almost all synony-

Table 2. Sorted nucleotide sequence variable sites of the *Bthsc* sequences

Sequence position	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sequence position	1	2	3	3	3	4	4	5	5	6	6	6	6	7	7	8	8	8	9	9
Sequence position	8	4	0	3	9	5	8	1	4	0	3	6	9	5	8	5	6	7	0	1
Codon position	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	2	3	3	1
Transition or transversion	ts	ts	tv	tv	ts	tv	ts	tv	ts	ts	ts	ts	tv	ts	ts	tv	ts	tv	ts	ts
Synonymous or nonsynonymous	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	R	S	S	S
Haplotypes																				
<i>Bt-Ananhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Apeths</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aubhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Amshs</i>	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aphhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arnhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arbhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Acmhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Atkhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Akhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bthsc</i> 1 (form A) consensus	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Fhs</i>	.	A	.	.	.	.	T	.	.	.	.	.	.	.	.	A	A	G	.	.
<i>Bt-Ghs</i>	.	A	.	.	.	.	T	.	.	.	.	.	.	.	.	A	A	G	.	.
<i>Bt-Bhs</i>	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	A	.	G	.	.
<i>Bt-Ehs</i>	.	A	.	A	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.
<i>Bt-Chs</i>	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Ihs</i>	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Dhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bthsc</i> 1 (all samples) consensus	T	G	G	C	C	T	C	A	T	C	C	C	T	G	C	C	G	T	G	C
<i>Bdorhs</i> (outgroup)	.	.	T	.	T	A	T	T	C	A	T	.	A	A	T	.	A	A	T	T
Sequence position	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
Sequence position	0	2	3	5	5	6	8	8	9	9	9	0	1	2	2	2	2	3	3	4
Sequence position	2	3	9	0	3	5	0	6	2	5	8	5	6	2	4	5	8	0	1	0
Codon position	3	3	1	3	3	3	3	3	3	3	3	1	3	3	2	3	3	2	3	3
Transition or transversion	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts
Synonymous or nonsynonymous	S	S	R	S	S	S	S	S	S	S	S	R	S	S	R	S	S	R	S	S
Haplotypes																				
<i>Bt-Ananhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.
<i>Bt-Apeths</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aubhs</i>	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Amshs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aphhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arnhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arbhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.
<i>Bt-Acmhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Atkhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Akhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bthsc</i> 1 (form A) consensus	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Fhs</i>	.	.	A	.	.	A	.	.	.	.	.	.	T	A	.	.	.	.	.	.
<i>Bt-Ghs</i>	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	A	.	.	.
<i>Bt-Bhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	A	.	.	.
<i>Bt-Ehs</i>	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	.	.
<i>Bt-Chs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T
<i>Bt-Ihs</i>	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	.	T
<i>Bt-Dhs</i>	C	.	.	.	.	C	.	.	.	.	.	.	.	A	.	A	.	.	.	C
<i>Bthsc</i> 1 (all samples) consensus	T	C	G	T	C	T	G	G	T	T	G	A	C	C	A	G	G	C	G	T
<i>Bdorhs</i> (outgroup)	.	T	.	C	T	.	A	A	C	C	.	.	.	A	.	.	.	.	T	C
Sequence position	2	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4
Sequence position	7	0	1	2	4	5	7	8	8	9	9	1	2	3	3	4	5	5	5	6
Sequence position	6	6	5	8	1	7	2	1	4	3	6	4	6	5	8	2	0	3	9	5
Codon position	3	3	3	1	2	3	3	3	3	3	3	3	3	3	3	1	3	3	3	3
Transition or transversion	tv	ts	ts	ts	ts	tv	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts	ts
Synonymous or nonsynonymous	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
Haplotypes																				
<i>Bt-Ananhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Apeths</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aubhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Amshs</i>	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Aphhs</i>	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arnhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Arbhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.
<i>Bt-Acmhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Atkhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
<i>Bt-Akhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bthsc</i> 1 (form A) consensus	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
<i>Bt-Fhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	A	.
<i>Bt-Ghs</i>	.	.	.	.	.	.	A	.	.	.	.	.	.	G	.	.	T	.	A	.
<i>Bt-Bhs</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	.	.

Table 2. Continued

<i>Bt-Elhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	.	C
<i>Bt-Clhsc</i>	.	.	.	.	.	.	G	.	.	.	.	.	.	G	G	.	T	.	A	
<i>Bt-Ihsc</i>	.	.	.	.	G	.	.	.	.	.	.	.	.	G	G	.	.	.	A	
<i>Bt-Dhsc</i>	.	.	.	T	.	C	A	.	.	.	.	C	.	G	.	.	.	.	.	
<i>Bthsc</i> 1 (all samples) consensus	A	G	C	C	A	G	G	T	C	T	T	T	A	A	A	C	G	C	T	
<i>Bdorchsc</i> (outgroup)	.	A	T	T	.	.	.	T	G	A	.	C	C	A	.	C	T	.	.	
Sequence position	4	4	4	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	
Sequence position	8	9	9	0	1	3	4	4	4	5	6	6	7	7	0	2	2	3	4	
Sequence position	9	2	5	3	7	7	5	6	9	2	1	4	3	8	3	1	9	5	8	
Codon Position	3	3	3	2	1	3	2	3	3	3	3	3	3	2	3	3	2	3	3	
Transition or transversion	tv	ts	ts	ts	tv	tv	ts	B	tv	tv	ts	ts	ts	ts	ts	B	ts	B	ts	
Synonymous or nonsynonymous	S	S	S	R	R	S	R	S	R	S	S	S	S	S	R	S	S	R	S	
Haplotypes																				
<i>Bt-Ananhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
<i>Bt-Apethsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	
<i>Bt-Aubhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	
<i>Bt-Anshsc</i>	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	
<i>Bt-Aphhsc</i>	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	
<i>Bt-Arnshsc</i>	G	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	.	.	.	
<i>Bt-Arbhsc</i>	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	
<i>Bt-Acmhsc</i>	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	
<i>Bt-Atkhsc</i>	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	
<i>Bt-Akbhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
<i>Bthsc</i> 1 (form A) consensus	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	
<i>Bt-Flhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	
<i>Bt-Ghsc</i>	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	A	
<i>Bt-Bhsc</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	
<i>Bt-Elhsc</i>	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	
<i>Bt-Clhsc</i>	.	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	C	.	T	
<i>Bt-Ihsc</i>	.	T	.	.	.	.	.	.	.	.	C	.	.	.	G	.	C	.	T	
<i>Bt-Dhsc</i>	.	.	.	.	.	.	.	A	.	T	.	.	.	.	.	.	C	G	T	
<i>Bthsc</i> 1 (all samples) consensus	T	C	T	G	C	C	C	C	A	A	T	C	C	C	A	G	A	T	A	
<i>Bdorchsc</i> (outgroup)	.	.	C	A	G	A	T	T	.	.	.	.	T	T	.	A	.	G	T	
Sequence position	6	6																		
Sequence position	5	5																		
Sequence position	2	4																		
Codon position	1	3																		
Transition or transversion	ts	tv																		
Synonymous or nonsynonymous	S	R																		
Haplotypes																				
<i>Bt-Ananhsc</i>	.	.																		
<i>Bt-Apethsc</i>	.	.																		
<i>Bt-Aubhsc</i>	.	.																		
<i>Bt-Anshsc</i>	.	.																		
<i>Bt-Aphhsc</i>	.	.																		
<i>Bt-Arnshsc</i>	.	.																		
<i>Bt-Arbhsc</i>	.	.																		
<i>Bt-Acmhsc</i>	.	.																		
<i>Bt-Atkhsc</i>	.	.																		
<i>Bt-Akbhsc</i>	.	.																		
<i>Bthsc</i> 1 (form A) consensus	.	.																		
<i>Bt-Flhsc</i>	.	.																		
<i>Bt-Ghsc</i>	.	.																		
<i>Bt-Bhsc</i>	.	.																		
<i>Bt-Elhsc</i>	.	.																		
<i>Bt-Clhsc</i>	.	.																		
<i>Bt-Ihsc</i>	.	.																		
<i>Bt-Dhsc</i>	.	.																		
<i>Bthsc</i> 1 (all samples) consensus	T	A																		
<i>Bdorchsc</i> (outgroup)	C	T																		

The top three rows indicated the positions of the nucleotide residues within the 664bp *Eco* RI fragment. The fourth row indicates the position of the change within the codon. The fifth row indicates a transitional (ts) or transversional (tv) or both (B) comparing each change with the *Bthsc* 1 consensus sequence. The sixth row indicates whether the substitutions result in synonymous (S) or nonsynonymous (R) changes. Dots represent nucleotides that are identical to the *Bthsc* 1 overall consensus sequence. Normal typed and bold typed nucleotides are ts and tv, respectively.

mous and at the codon third position, except the change at position 85.

Genetic Distance and Similarity Comparisons. Pair-wise comparisons of the variation in *Bthsc*1 sequences were also investigated. Genetic distances were calcu-

lated based on the number of nucleotide substitutions. From Table 3, the average number of nucleotide differences between form A individuals is 4.2 ± 1.5 , and the average genetic distance based on Kimura's two Parameter (K2P) model is 0.0063 ± 0.002 ($N = 45$ in

Table 3. Genetic distance matrix (below diagonal) and nucleotide differences (above diagonal) of the *Bthsc1* sequences between *B. tau* form A

	<i>Bt-Aubhsc</i>	<i>Bt-Apethsc</i>	<i>Bt-Amshsc</i>	<i>Bt-Aphhsc</i>	<i>Bt-Arnshc</i>	<i>Bt-Arbhsc</i>	<i>Bt-Acnhsc</i>	<i>Bt-Ananhsc</i>	<i>Bt-Atkshc</i>	<i>Bt-Akblsc</i>
<i>Bt-Aubhsc</i>		1	4	5	5	5	3	3	4	5
<i>Bt-Apethsc</i>	0.0015		3	4	4	4	2	2	3	4
<i>Bt-Amshsc</i>	0.0060	0.0045		3	5	5	3	5	6	7
<i>Bt-Aphhsc</i>	0.0076	0.0060	0.0045		4	6	2	4	5	6
<i>Bt-Arnshc</i>	0.0076	0.0061	0.0076	0.0061		4	2	4	5	6
<i>Bt-Arbhsc</i>	0.0076	0.0060	0.0076	0.0091	0.0060		4	6	7	8
<i>Bt-Acnhsc</i>	0.0045	0.0030	0.0045	0.0030	0.0030	0.0060		2	3	4
<i>Bt-Ananhsc</i>	0.0045	0.0030	0.0076	0.0060	0.0061	0.0091	0.0030		3	4
<i>Bt-Atkshc</i>	0.0060	0.0045	0.0091	0.0076	0.0076	0.0106	0.0045	0.0045		5
<i>Bt-Akblsc</i>	0.0076	0.0060	0.0106	0.0091	0.0091	0.0121	0.0060	0.0060	0.0076	

both cases). Table 4 shows the results for all of the different *B. tau* forms. Here, the average number of nucleotide differences was found to be 16 ± 5.98 and the average genetic distance was calculated to be 0.0247 ± 0.009 ($N = 28$ in both cases).

From a *t*-test (2-tailed) for equality of means, the average genetic distance within the form A individuals is significantly smaller compared with the average genetic distances among all of the different forms ($t = -10.276$; $df = 29.04$; $P < 0.001$). Comparisons based on the average number of nucleotide differences are consistent with the results of the genetic distance estimates. Here, the *t*-test for equality of means (2-tailed) shows that the average number of nucleotide differences within form A is significantly smaller than the average number of nucleotide differences among the different forms ($t = -10.32$; $df = 29.125$; $P < 0.001$).

Unweighted Pair-Group Method with Arithmetic Average Dendrogram and Neighbor Joining Phylogenetic Analyses. The pairwise genetic distance data matrix for all of the *Bthsc1* comparisons (Tables 3 and 4 combined) was also used as input for a cluster analysis. The dendrogram resulting from the unweighted pair-group method with arithmetic average analysis groups all of the 10 form A individuals from various populations into one clade (Fig. 3), inferring a genetic separation of this form. The *Bt-Fhsc*, *Bt-Chsc*, and *Bt-Bhsc* individuals are also grouped into a distinct clade that may reflect a genetic relationship among these forms. The *Bt-Chsc* and *Bt-Ihsc* individuals also group together. The outgroup *Bdorchsc* is the most distant from all the *Bthsc1* forms.

A 50% majority-rule consensus phylogenetic tree generated by the neighbor-joining method based on

the Kimura's two parameters (K2P) model is shown in Fig. 4. In this tree many (but not all) of the nodes were supported by bootstrap analyses with relatively high confidence levels (higher than 50%). The NJ tree also groups all of the 10 form A individuals from various populations into one clade (Fig. 4) and infers a genetic separation of the A form from other forms with confidence level 76%. The *Bt-Fhsc*, *Bt-Chsc*, and *Bt-Bhsc* sequences are grouped into another distinct clade with bootstrap confidence limits of 88% and 95%. The *Bt-Chsc* and *Bt-Ihsc* sequences also formed a distinct group with a confidence level of 100%. Here, *B. dorsalis* is grouped with form D although the confidence level is below 50%. The NJ result is mostly congruent with the unweighted pair-group method with arithmetic average analysis.

Discussion

Primers designed from the DNA sequence of a heat shock cognate gene have been shown to reliably amplify sequences from *B. tau* individuals representing different forms found throughout Thailand. These primers were designed from coding regions of this gene that are conserved between *B. tau*, *B. dorsalis*, and *C. capitata*. The heat shock cognate gene amplification products obtained here from the different *B. tau* specimens provided markers for a molecular systematic approach to the analysis of relationships of these different forms. Previously, the taxonomic status and systematic relationships of the different forms of *B. tau* described have been the source of much confusion in the literature (Hardy 1973, White and Elson-Harris 1992, Baimai et al. 2000).

Table 4. Genetic distance matrix (below diagonal) and nucleotide differences (above diagonal) of the *Bthsc1* sequences between different *B. tau* forms

	<i>Bt-Ahsc</i>	<i>Bt-Bhsc</i>	<i>Bt-Chsc</i>	<i>Bt-Dhsc</i>	<i>Bt-Ehsc</i>	<i>Bt-Fhsc</i>	<i>Bt-Ghsc</i>	<i>Bt-Ihsc</i>
<i>Bt-Ahsc</i>		10	11	18	12	14	15	12
<i>Bt-Bhsc</i>	0.0152		17	20	14	6	5	18
<i>Bt-Chsc</i>	0.0168	0.0261		20	17	21	22	5
<i>Bt-Dhsc</i>	0.0276	0.0308	0.0307		21	26	23	21
<i>Bt-Ehsc</i>	0.0183	0.0214	0.0261	0.0323		16	17	20
<i>Bt-Fhsc</i>	0.0214	0.0091	0.0324	0.0403	0.0246		5	22
<i>Bt-Ghsc</i>	0.0230	0.0076	0.0340	0.0356	0.0261	0.0076		23
<i>Bt-Ihsc</i>	0.0183	0.0276	0.0076	0.0323	0.0307	0.0339	0.0355	

Bt-Ahsc is a *Bthsc1* consensus sequences of all *B. tau* form A sequences. The other sequences are as indicated in Table 1.

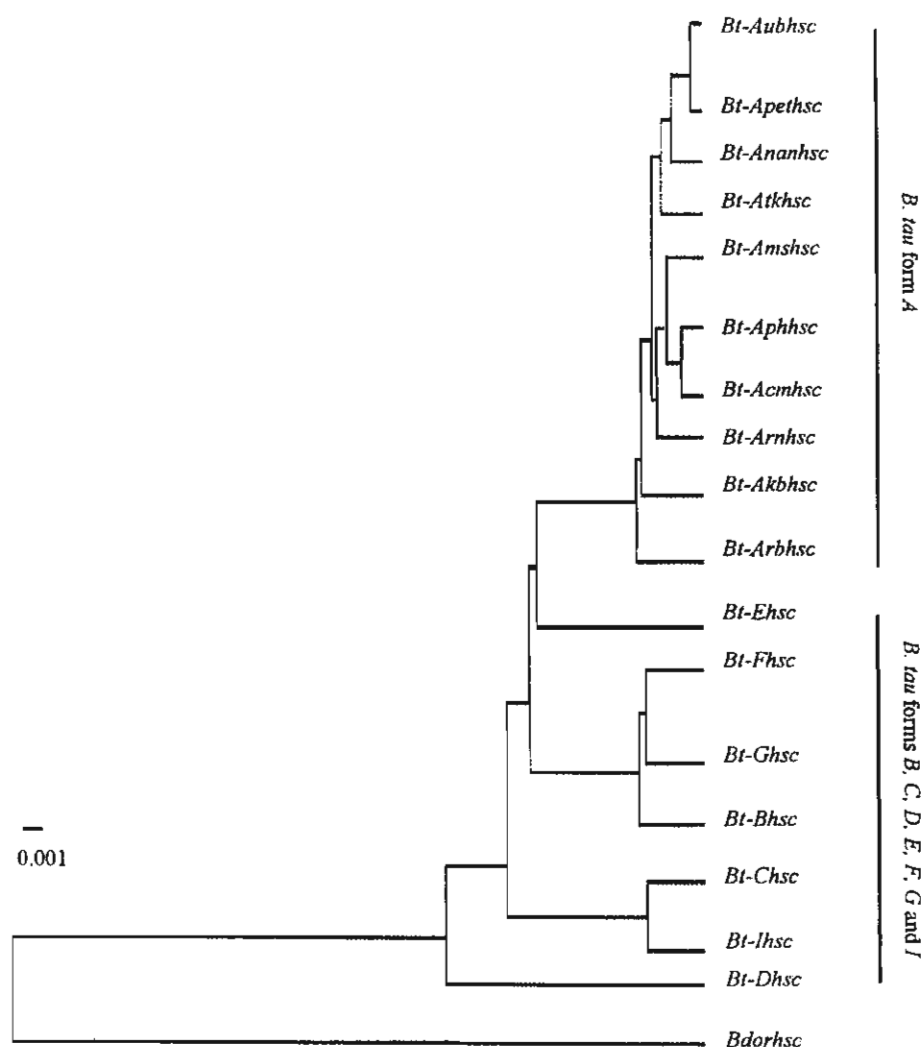


Fig. 3. Unweighted pair-group method with arithmetic average dendrogram of genetic distances from various *Bthsc1* sequences to infer relationships of forms and populations. The scale bar indicates a 0.001 U of genetic distance from Kimura's two-parameter model.

To analyze these relationships, nucleotide sequences were determined for a 664bp *EcoRI* fragment found to be present within the heat shock cognate gene amplification product obtained for all individuals sampled. Alignment of these sequences revealed that 55 of the sites were variable between the different forms. No gaps were needed for these alignments, and the variable sites identified appeared to be distributed throughout the fragment. These results show that although the *heat shock protein 70* cognate gene is conserved, it can provide a level of polymorphism comparable to other established marker systems. For example, Morrow et al. (2000) found 33 out of 373 and 25 out of 316 sites, respectively, to be variable in the *cytb* and *COII* genes they used as markers for a similar analysis of other closely related Tephritid species.

Using these sequences, phylogenetic analyses based on the unweighted pair-group method with arithmetic average (Felsenstein 1993) and Neighbor Joining (Saitou and Nei 1987) methods were carried out. Both of these analyses grouped all the samples of form A into one clade with relatively high statistical confidence (76% from New Jersey analysis). Other analyses

comparing the genetic makeup of form A individuals as a group to the other forms of *B. tau* yielded similar results. For example, the average number of substitutions found within form A was significantly lower than the level of variation among all the forms. A similar result was obtained from an analysis of genetic distances within form A individuals compared with the other forms.

At the level of direct sequence comparison, there is also a fixed difference between these forms at position 222 which may be a synapomorphy. At this position, an A nucleotide is shared between all of the individuals representing forms B, C, D, E, F, G, and I. All of the form A individuals, in contrast, have a C nucleotide at this position.

Taken together, these results suggest that the *B. tau* form A is isolated to some extent from the other *B. tau* forms. The extent of the isolation suggests that this form may have phylogenetic species status as defined by Cracraft (1983). Congruent with this, Baimai et al. (2000) had previously shown a unique chromosomal karyotype for the *B. tau* form A. However, more sam-

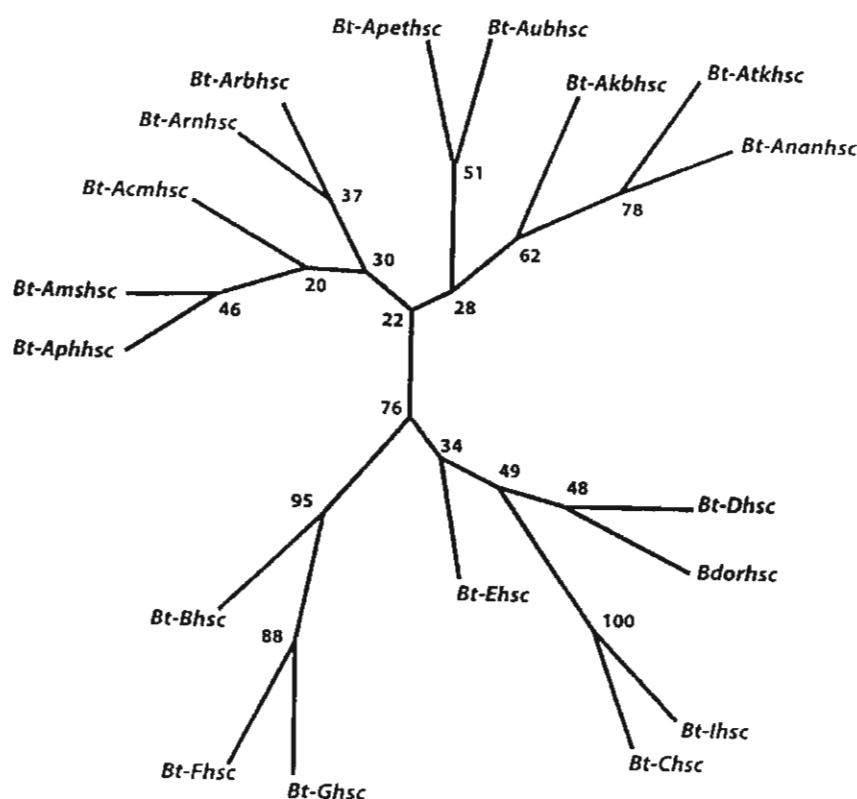


Fig. 4. A 50% majority-rule consensus neighbor-joining phylogenetic tree with bootstrap values (100 replicates) on each node. There is no genetic distance on each branch.

ples of all of *B. tau* forms represented here need to be analyzed to confirm this hypothesis.

The phylogenetic analyses also suggest that the *B. tau* forms *B*, *F*, and *G* may form another subgroup. Although they are allopatric and each is monophagous, they infest closely related hosts that are distinct from the cucurbitaceous hosts used by forms *A*, *C*, *D*, and *I* (Baimai et al. 2000). Also with the forms *B*, *F*, and *G* there are four fixed differences at positions 48, 85, 87, and 216 that could be synapomorphic. The existence of this subgroup may support a hypothesis that host shifting plays a crucial role for divergence within *B. tau* (and other Tephritids) as it has in other systems (Bush 1994). The forms *C* and *I* may be another subgroup as suggested by the neighbor-joining analysis. However, here again more samples of these forms would be required to test the validity of these ideas.

In conclusion, heat shock cognate gene sequences have been used here as a new molecular marker for the analysis of closely related forms of *B. tau*. The sequences used here successfully resolved some relationships within the *B. tau* complex better than others. We provide evidence that the *B. tau* form *A* group may be reliably isolated from the other *B. tau* forms. The status of the forms *B* to *I* cannot be completely clarified, in part because only one sample of each was included in the study. However, our results suggest relationships among these forms that are congruent with host preferences and karyotype differences described by others.

From an evolutionary perspective, the analysis of phylogenetic relationships of closely related forms (or

groups) such as *B. tau* that appear to have undergone recent expansion presents many challenges. In addition to the types of molecular data developed here, differences in mating behaviors, sexual selection, host preference shifts, and phenotypic traits important for reproductive isolation as described by Iwahashi (1999) and Morrow et al. (2000) for Tephritid species must also be taken into consideration to achieve a thorough understanding of such relationships.

Finally, because the coding regions of heat shock genes tend to be conserved across taxa (Borchiellini et al. 1998), it may also be possible to use these primers to amplify similar sequences from other *Bactrocera* species and other genera within the family Tephritidae. The use of this gene may help reduce the possibility of ambiguity in multiple sequence alignments as seen in the analysis of nineteen *Bactrocera* species using mitochondrial rDNA sequences (Muraji and Nakahara 2001).

Acknowledgments

The authors thank anonymous reviewers for constructive comments on the manuscript. The authors also thank V. Baimai and S. Tigavattananont for providing some fruit fly samples and D. S. Haymer for his useful comments. This study was supported in part by the graduate fellowship program from National Science and Technology Development Agency to U. Thaenkhom and by the Thailand Research Funds (No.RSA/06/38) to S. Thanaphum.

References Cited

- Baimai, V., J. Phinchongsakuldit, and C. Sumrandee. 2000. Cytological evidence for a complex of species within the taxon *B. tau* (Diptera: Tephritidae) in Thailand. *Biol. J. Linn. Soc.* 69: 399–409.
- Bonizzoni, M., A. R. Malacreda, C. R. Guglielmino, L. M. Gomlski, G. Gasperi, and L. Zheng. 2000. Microsatellite polymorphism in the Mediterranean fruit fly, *Ceratitis capitata*. *Insect. Mol. Biol.* 9(3): 251–261.
- Borchiellini, C., N. Boury-Esnault, J. Vacelet, and Y. L. Parco. 1998. Phylogenetic analysis of the Hsp 70 sequences reveals the monophyly of metazoa and specific phylogenetic relationships between animals and fungi. *Mol. Biol. Evol.* 15(6): 647–655.
- Budin, K., and H. Philippe. 1998. New insights into the phylogeny of eukaryotes based on ciliate Hsp 70 sequences. *Mol. Biol. Evol.* 15(8): 943–956.
- Bush, C. L. 1994. Sympatric speciation in animals: new wine in old bottles. *TREE* 9: 285–288.
- Cracraft, J. 1983. Species concepts and speciation analysis, pp. 159–187. In R. F. Johnston [ed.], *Current ornithology*. Plenum, New York.
- Drew, R.A.I., and M. C. Roving. 1997. Overview-Tephritidae in the Pacific and Southeast Asia, pp. 46–53. In A. J. Allwood and R.A.I. Drew [eds.], *Management of fruit flies in the Pacific: a regional symposium*. Australian Centre for International Agriculture Research, Canberra, Australia.
- Felsenstein, J. 1993. PHYLIP: Phylogeny inference package, version 3.5c. The University of Washington, Seattle.
- Graser, R. T., D. Malnar-Dragojevic, and V. Vincek. 1996. Cloning and characterization of a 70 kD heat shock cognate (*hsc70*) gene from the zebrafish (*Danio rerio*). *Genetica* 98: 273–276.
- Gupta, R. S., and C. B. Golding. 1993. Evolution of HSP70 gene and its implications regarding relationships between archaeobacteria, eubacteria, and eukaryote. *J. Mol. Evol.* 37: 573–582.
- Hardy, D. E. 1973. The fruit Flies (Tephritidae-Ditera) of Thailand and bordering countries. *Pacific Insects Monograph* 31, Bishop Museum.
- Haymer, D. S., D. O. McInnis, and L. Arcangeli. 1992. Genetic variation between strains of the Mediterranean fruit fly, *Ceratitis capitata*, detected by DNA fingerprinting. *Genome* 35: 528–533.
- He, M., and D. S. Haymer. 1997. Polymorphic intron sequences detected within and between populations of the Oriental fruit fly (Diptera: Tephritidae). *Ann. Entomol. Soc. Am.* 90(6): 825–831.
- Hillis, D. M., C. Moritz, and B. K. Mabel. 1996. *Molecular Systematic* 2nd ed. Sinauer, Sinauer, Sunderland, MA.
- Iwahashi, O. 1999. Distinguishing between the two sympatric species *Bactrocera* *Corambolae* and *B. papayae* (Diptera: Tephritidae) based on aedeagal length. *Ann. Entomol. Soc. Am.* 92(5): 639–643.
- Lindquist, S., and E. A. Craig. 1988. The heat shock protein. *Annu. Rev. Genet.* 163: 631–677.
- Mayr, E. 1963. *Animal species and evolution*. Harvard Univ. Press, Cambridge, MA.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- McPherson, B. A., and H.-Y. Han. 1997. Phylogenetic analysis of North America *Rhagoletis* (Diptera: Tephritidae) and related genera using mitochondrial DNA sequence data. *Mol. Phylogenet. Evol.* 7(1): 1–16.
- Meksongsee, B., A. Liewvanich, and M. Jirasuratana. 1991. Fruit flies in Thailand, pp. 83–98. In S. Vijayasegaran and A. G. Ibrahim [eds.], *Proceedings, 1st International Symposium on Fruit Flies in the Tropics*, 14–16 March 1988. Malaysian Agricultural Research and Development Institute, Kuala Lumpur, Malaysia.
- Morrow, J., L. Scott, B. Congdon, D. Yeates, M. Frommer, and J. Sved. 2000. Close genetic similarity between two sympatric species of tephritid fruit fly reproductively isolated by mating time. *Evolution* 54(3): 899–910.
- Muraji, M., and S. Nakahara. 2001. Phylogenetic relationships among fruit flies, *Bactrocera* (Diptera, Tephritidae), based on the mitochondrial rDNA sequences. *Insect. Mol. Biol.* 10(6): 549–559.
- Page, R.D.M. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357–358.
- Pychlik, P., and W. Pychlik. 1989. OLIGO primer analysis software, version 6.0. Molecular Biology Insights (MBI).
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406–425.
- Thanaphum, S., and D. S. Haymer. 1998. A member of the hsp70 gene family from the Mediterranean fruit fly, *Ceratitis capitata*. *Insect. Mol. Biol.* 7(1): 1–10.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673–4680.
- Tigvattananont, S. 1986. The importance of *Dacus tau* (Walker) in Thailand. *Kaen Kaset* 14: 114–118.
- White, I. M., and M. M. Elson-Harris. 1992. *Fruit flies of economic significance*. Center for Agriculture and Biosciences International, Wallingford, United Kingdom.
- Yu, H., M. Frommer, M. K. Robon, A. W. Meat, D.C.A. Shearman, and J. Sved. 2001. Microsatellite analysis of the Queensland fruit fly *Bactrocera tryoni* (Diptera: Tephritidae) indicates spatial structuring: implications for population control. *Bull. Entomol. Res.* 9: 39–148.

Received for publication 8 January 2002; accepted 5 September 2002.