MOLECULAR CHARACTERIZATION OF THE HOPPER TRANSPOSABLE ELEMENT IN *BACTROCERA TAU* (WALKER)

CHATCHAWAL SANGUANSILP

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(ENVIRONMENTAL BIOLOGY)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2001

ISBN 974-665-302-4

MOLECULAR CHARACTERIZATION OF THE HOPPER TRANSPOSABLE ELEMENT IN *BACTROCERA TAU* (WALKER)

CHATCHAWAL SANGUANSILP

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF SCIENCE

(ENVIRONMENTAL BIOLOGY)

FACULTY OF GRADUATE STUDIES

MAHIDOL UNIVERSITY

2001

ISBN 974-665-302-4

Thesis

entitled

MOLECULAR CHARACTERIZATION OF THE HOPPER TRANSPOSABLE ELEMENT IN *BACTROCERA TAU* (WALKER)

Chatchawal Sanguansilp
.....
Mr.Chatchawal Sanguansilp
Candidate

Sujinda Thanaphum
.....
Lect.Sujinda Thanaphum, Ph.D.
Major-Advisor

Apinunt Udomkit
.....
Asst.Prof.Apinunt Udomkit, Ph.D.
Co-advisor

Liangchai Limlomwongse
.....
Prof.Liangchai Limlomwongse, Ph.D.
Dean
Faculty of Graduate Studies

Praneet Damrongphol
.....
Assoc.Prof.Praneet Damrongphol, Ph.D.
Chairman
Master of Science Programme
in Environmental Biology
Faculty of Science

Thesis

entitled

MOLECULAR CHARACTERIZATION OF THE HOPPER TRANSPOSABLE ELEMENT IN *BACTROCERA TAU* (WALKER)

was submitted to the Faculty of Graduate Studies, Mahidol University
for the degree of Master of Science (Environmental Biology)

on

February 13, 2001

Chatchawal Sanguansilp
.....
Mr.Chatchawal Sanguansilp
Candidate

Sujinda Thanaphum
.....
Lect.Sujinda Thanaphum, Ph.D.
Chairman

Apinunt Udomkit
.....
Asst.Prof.Apinunt Udomkit, Ph.D.
Member

.....
Prof.Vithaya Meevootisom, Ph.D.
Member

.....
Prof.Liangchai Limlomwongse, Ph.D.
Dean
Faculty of Graduate Studies
Mahidol University

.....
Prof.Amaret Bhumiratana, Ph.D.
Dean
Faculty of Science
Mahidol University

ACKNOWLEDGEMENT

I would like to express my sincere gratitude and deep appreciation to my advisor, Dr.Sujinda Thanaphum for his kindness, valuable advises throughout this work, critically reading my manuscript and for giving me the opportunity to do this thesis. I am very grateful to Prof.Vithaya Meevootisom and Asst.Prof.Apinunt Udomkit for participating in the members of examination committee and their valuable advise to improve this study. I am grateful to Assoc.Prof.Saen Tigvattananont for providing the fruit fly specimens.

I was partially financial support by the Thailand Research Fund (grant no. RSA/06/38 of Dr.Sujinda Thanaphum and grant no. RTA3880008 of Prof.Visut Baimai) and by the TRF/BIOTEC Special Program for Biodiversity Research and Training (grant no. BRT542007 of Dr.Sujinda Thanaphum). I also thank the Department of Biology for partially support during this study.

Finally, my sincere appreciation is expressed to all of the students and technician in N404 laboratory for their help and sincerity. I would like to express my deepest appreciation to my family especially my parents and my grand mother for their love, understanding, concern and insight throughout the period of hard work. I do it for you, grand mother.

Chatchawal Sanguansilp

3936824 SCEB/M : MAJOR: ENVIRONMENTAL BIOLOGY ; M.Sc. (ENVIRONMENTAL BIOLOGY)
KEY WORDS : HOPPER TRANSPOSABLE ELEMENT/ *hAT* FAMILY / TRANSPOSASE / *BACTROCERA TAU* (WALKER)
CHATCHAWAL SANGUANSILP: MOLECULAR CHARACTERIZATION OF THE HOPPER TRANSPOSABLE ELEMENT IN *BACTROCERA TAU* (WALKER) THESIS ADVISOR: SUJINDA THANAPHUM, Ph.D., APINUNT UDOMKIT, Ph.D. 65 p. ISBN 974-665-302-4

A genetic phenomenon that often induces mutations is a transposition of genetic materials called transposable elements. The genome of the cucurbit-attacking fruit fly, *Bactrocera tau* (Walker), contains a putative functional transposable element called *hopper* element. Using PCR and inverse PCR, we amplified and sequenced *hopper* element and its target site duplication sequences. This *hopper* element is 3172 bp in length. The termini of *hopper* element are composed of 19 bp inverted terminal repeats (ITRs) with single mismatch and the perfect 8 bp direct duplication of the target sites. The comparison between the ITRs of the *hopper* element to those of the *hobo* element of *Drosophila melanogaster,* the *Ac* element of *Zea mays*, and the *Tam3* element of *Antirrhinum majus*, as well as several other plant and insect elements, revealed a conserved terminal sequence pattern.

The *hopper* element contains a single long open reading frame (ORF) encoding a putative transposase protein of 651 amino acids. The comparisons between the entire length of *hopper* amino acids sequences with the transposase amino acid sequences of the *hobo, Ac,* and *Tam3* elements showed that these transposase protein sequences have a similar level of distance ranging from 22-30 % identity. The genomic representation of *hopper* element in genome of *B. tau* from different geographical areas in Thailand showed differences in banding patterns. These results showed that this *hopper* element corresponds to a putative functional transposase gene and may infer a recently active transposition.

3936824 SCEB/M : สาขาวิชา : ชีววิทยาสภาวะแวดล้อม ; วท.ม. (ชีววิทยาสภาวะแวดล้อม)

ชัชวาล สงวนศิลป์ : การศึกษาอณูพันธุศาสตร์ของ *hopper* transposable element จากจีโนมของแมลงวันแตงกวา (*Bactrocera tau* (WALKER)) (MOLECULAR CHARACTERIZATION OF THE HOPPER TRANSPOSABLE ELEMENT IN *BACTROCERA TAU* (WALKER).). คณะกรรมการควบคุมวิทยานิพนธ์ : สุจินดา ธนะภูมิ, Ph.D., อภินันท์ อุดมกิจ, Ph.D. 65หน้า. ISBN 974-665-302-4

กลไกทางพันธุกรรมอย่างหนึ่งที่ก่อให้เกิดการกลายพันธุ์คือการเคลื่อนย้ายตำแหน่งบนจีโนมของยีนที่เรียกว่า Transposable element พบว่าจีโนมของแมลงวันแตงกวา (*Bactrocera tau* (Walker)) ประกอบด้วย putative functional transposable element ชนิดหนึ่ง ที่มีชื่อว่า *hopper* element เทคนิค PCR และ inverse PCR ได้ถูกนำมาใช้ในการเพิ่มจำนวน และหารหัสทางพันธุกรรมของ *hopper* element และ บริเวณ target site duplication ของยีนชนิดนี้ พบว่า *hopper* element มีขนาดความยาว 3172 คู่เบส และที่บริเวณปลายทั้งสองด้านของ *hopper* element ประกอบด้วยส่วน inverted terminal repeats (ITRs) ที่มีขนาดความยาว 19 คู่เบส ที่มีลักษณะ single mismatch และส่วนของ target site duplication ขนาด 8 คู่เบส ที่มีลักษณะ perfect direct duplication เมื่อทำการเทียบเรียงรหัสพันธุกรรมในส่วนบริเวณ ITRs ของ *hopper* element กับรหัสพันธุกรรมบริเวณเดียวกันของ *hobo* element ที่พบใน *Drosophila melanogaster*, *Ac* element ที่พบใน *Zea mays*, และ *Tam3* element ที่พบใน *Antirrhinum majus* และ transposable element ชนิดอื่นๆที่พบในพืชและแมลง ผลที่ได้พบว่ายังคงมีรูปแบบของรหัสพันธุกรรมที่คล้ายกันของรหัสพันธุกรรมบริเวณนี้

ยีน *hopper* element ที่พบ ประกอบ ด้วย single long open reading frame (ORF) ซึ่งสามารถแปลเป็นรหัสของ putative transposase โปรตีน ที่มีขนาดความยาว 651 กรดอะมิโน เมื่อ นำ รหัส กรด อะมิโน ทั้งหมด ที่ ได้ จาก *hopper* element ไป เทียบเรียง กับ รหัสกรดอะมิโน transposase โปรตีนของ *hobo*, *Ac*, และ *Tam3* elements พบว่ามีค่า % identity ที่คล้ายกันอยู่ในช่วง 22-30 % การศึกษา genomic representation ของ *hopper* element ในจีโนม *B. tau* จากภูมิภาคต่างๆในประเทศไทย พบว่ามีความแตกต่างของ banding patterns ที่เกิดขึ้น ผลการศึกษาที่ได้เหล่านี้แสดงให้เห็นว่า *hopper* element ที่พบในครั้งนี้มีคุณสมบัติสอดคล้องกับลักษณะ putative functional ของยีน transposase และอาจแสดงถึงการเคลื่อนย้ายตำแหน่งของยีนชนิดนี้ที่เพิ่งจะเกิดขึ้นได้

CONTENTS

CONTENTS (cont.)

CONTENTS (cont.)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (cont.)

LIST OF ABBREVIATIONS

BLAST	=	Basic Local Alignment Search Tool
BSA	=	bovine serum albumin
bp	=	base pair (s)
°C	=	degree celsius
cDNA	=	complementary DNA
dATP	=	deoxyadenosine-5'-triphosphate
dCTP	=	deoxycytidine-5'-triphosphate
dGTP	=	deoxyguanosine-5'-triphosphate
dTTP	=	deoxythymidine-5'-triphosphate
dNTP	=	dATP, dCTP, dGTP, and dTTP
DNA	=	deoxyribonucleic acid
E. coli	=	*Escherichia coli*
EDTA	=	ethylenediamine tetraacetic acid
EtBr	=	ethidium bromide
et al.	=	and others
g	=	gram
hAT	=	*hobo, Ac,* and *Tam3*
hr	=	hour (s)
HTH	=	helix-turn-helix
IPCR	=	inverse polymerase chain reaction
ITRs	=	inverted terminal repeats
kb	=	kilobase (s)

LIST OF ABBREVIATIONS (cont.)

LB	=	Luria-Bertani (medium)
LTR	=	long terminal repeat
M	=	molar
mM	=	millimolar
mg	=	milligram
min	=	minute (s)
ml	=	milliliter
μg	=	*microgram*
μl	=	microliter
ng	=	nanogram
ORF	=	open reading frame
PCR	=	polymerase chain reaction
pmole	=	picomole
RNA	=	ribonucleic acid
RT	=	reverse transcriptase
SDS	=	sodium dodecyl sulfate
sec	=	second
SSC	=	*sodium-sodium citrate*
TE	=	transposable element
Tris-HCl	=	Tris-(hydroxymethyl)-aminoethane hydrochloric acid
U	=	unit (s)

LIST OF ABBREVIATIONS (cont.)

UV	=	ultraviolet
V	=	volt
X-Gal	=	5-bromo-4chloro-3indolyl-β-D-galactopyranoside

CHAPTER I

INTRODUCTION

One of the foundations underlying organic diversity is a kind of variation, preserved by natural selection and provided by mutations. One of the natural phenomenon that induces mutations is the movement of genetic material from one chromosomal location to another. DNA sequences that possess a capability to change their genomic location are called transposable elements.

Transposable elements (TE) are discrete segments of DNA that are distinguished by their ability to move and replicate within genome. Since the discovery of transposable elements by Barbara McClintock in the 1940s, these elements have been found to be ubiquitous in most living organisms (1). Transposable elements are interesting not just for the mechanisms involved in the manipulation of DNA, but also for the evolutionary consequences of their mobility.

Transposable elements can be classified into two major classes according to their transposition mechanism (1). Class I elements are retroelements that use reverse transcriptase to transpose by means of an RNA intermediate. They include retrotransposons such as *copia*-like elements in *Drosophila melanogaster*, *Ty* element in *Saccharomyces cerevisae*, and *DIR-1* in *Dictyostelium discoidium*. For class II elements, such as *P* and *hobo* elements in *D. melanogaster*, *Ac/Ds* in *Zea mays*, and *Tam3* in *Antirrhinum majus*, they transpose directly as DNA from one location to another in the genome. Moreover, transposable elements are classified into families according to their sequence similarity such as the *hAT* family (2).

The *hAT* stands for the *hobo* element from *D. melanogaster*, the *Ac* element from *Z. mays*, and the *Tam3* element from *A. majus*. All three elements generate 8 bp. target site duplication different from those observed for other transposable elements (3). The members of this superfamily are widely distributed among both plants and animals (3,4,5).

One of the transposable elements in the *hAT* family, called *hopper* element, was previously isolated from genome of the oriental fruit fly, *Bactrocera dorsalis* (Hendel) (6). It was approximately 3.1 kb in length with 19 bp. inverted terminal repeat sequences having a single mismatch. The finding that this element described does not contain an 8 bp insertion site direct repeat which is consistent in all other *hAT* elements and the interrupted transposase ORF is not supportive of autonomous function. The present work is intended to isolate and characterize a putative functional *hopper* transposable element gene from genome of *Bactrocera tau* (Walker) in Thailand.

The tephritidae family, the true fruit flies, includes about 4,000 species arranged in 500 genera. As a result, it is one of the largest families of Diptera (true flies) and one of the most economically important fruit flies in the world, especially in the Tropical America continent, South Pacific, and Southeast Asia (7). Since Thailand is one of the agricultural countries in the tropic zone, there are several species of fruit flies in this region. The important economic pests of fruits consist of *Bactrocera dorsalis* (Hendel), *Bactrocera correcta* (Bezzi), *Bactrocera cucurbitae* (Coquillett), and *Bactrocera tau* (Walker).

In 1986, Tigvattananont S. reported that *B. tau* (Walker) attacked very young melon fruit (8). As shown in figure 1.1, the adult flies of this species have orange brown scutum which is marked with black and with lateral and medial yellow stripes. They have facial spots, anterior supra-alar setae, prescutellar acrostichal setae, 4 scutellar setae, and wings with a costal band expanded into an apical spot (9). This cucurbit-attacking dacine fruit fly is widely distributed throughout Thailand and prefer to oviposit in cucurbitaceous fruits such as cucumber, bitter gourd, and pumpkin.

Figure 1.1 The morphology of the adult *Bactrocera tau* (Walker). This figure shows the morphology of the adult male *B. tau* (Walker). *B. tau* is in Order Diptera, Family Tephritidae, Subfamily Dacinae, Genus *Bactrocera*, Subgenus *Zeugodacus*. For the adult identification, it consists of scutum orange brown and marked with black, and with lateral and medial yellow stripes; with facial spots, anterior supra-alar setae, prescutellar acrostichal setae, 4 scutellar setae, and wing with a costal band expanded into an apical spot. This description was taken from White IM (9).

For this thesis, a gene of hobo-related transposable element in the genome of *B. tau* (Walker) was isolated by using the PCR approach. The structural similarities suggest that this *hobo*-related element is a member of the *hAT* family, which was called *hopper* element. Unlike the previously isolated from *B. dorsalis* (Hendel), this copy of *hopper* element contains a new putative-functional transposase gene that consists of an uninterrupted ORF for 651 amino acids beginning with a start codon. When we compared the deduced *hopper* transposase protein sequences to the transposase of the *hobo*, *Ac*, and *Tam3* elements, we found that the sequence alignment was conserved in three regions in the middle and C-terminal regions of these transposase proteins. Moreover, the genomic representation of the *hopper* transposable element in genome of *B. tau* (Walker) from different geographical areas in Thailand shows differences in banding patterns. The result shows that the *hopper* element is located in different genomic locations that may infer a recently active transposition.

CHAPTER II

LITERATURE REVIEW

2.1. Transposition of transposable elements

A great part of classical genetic analysis has been devoted to the localization of genes on chromosomes. Most genes occupy fixed sites on the chromosomes, and the overall structure of the genetic map is practically invariant . However, beginning in the 1940s, researchers have found that some DNA sequences can change their genomic location. In the late 1940s, Barbara McClintock developed the concept of what she termed "controlling elements" which was based on the variegation patterns of the maize. These "controlling elements" were genetic units associated with a gene, for example one involved in kernel pigmentation of *Zea mays*, thus controlling the expression of that gene (10). Through out her genetic analysis, the current picture of transposable elements ultimately evolved.

Transposable elements (TEs), also known as mobile or parasitic DNA, are moderately repetitive sequences that can move about in the genome of an organism. Transposable elements are widespread in all genomes out of which they have been sought (11). Mobile elements are major constituents of eukaryotic genomes, making upto 10 % of the *Drosophila* genome and 35 % of the human genome (12). TEs occur in many shapes and sizes but they share the common ability to insert into new sites in genomes. This common ability is called transposition.

Transposition is a recombination process in which DNA sequences termed transposable elements move from an original site on a DNA molecule to a new site on the same or different DNA molecule (13). The mechanism of DNA transposition follows a common scheme with some variation in details depending on the element (14). There are two types of transpositions which are distinguished by whether the transposable element is replicated or not. In replicative transposition, the mobile element is replicated and a copy is inserted to a target site while the other copy remains at the original site as shown in figure 2.1. Phage Mu and TnA family elements, such as Tn3 and Tn1000, use this type of transposition (15,16).

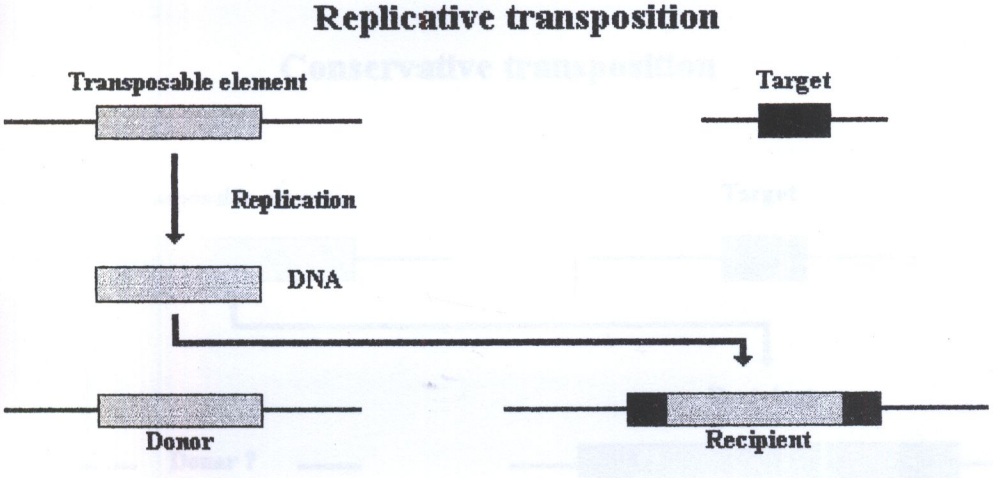

Figure 2.1 The mechanisms in the replicative transposition. This figure shows the mechanisms in the replicative transposition. The transposable element is replicated and a copy is inserted at a target site, while the other copy remains at the original site.

The other type of transposition is conservative transposition. This strategy is used by transposable elements that move by a cut and paste mechanism. The mobile element is transposed from the donor site to a target site and is conserved as shown in figure 2.2. It leaves a break at the original site. It will be deleterious to the host genome unless it can be repaired. This paradigm was first established for the bacterial transposon Tn10 and Tn7, but it is now clear that numerous elements, including *P* element in *Drosophila*, *Tc* element in *Caenorhabditis elegans* and other elements, use this mechanism (17).

Although some transposons use only one type of pathways for transposition, others, such as IS903 and IS1, predominantly form conservative transposition but also appear to form replicative transposition at a significant rate (18).

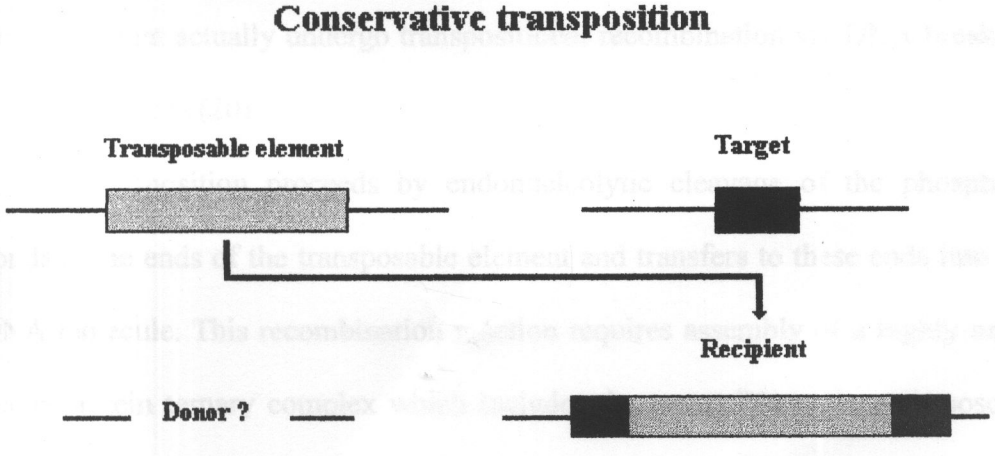

Figure 2.2 The mechanisms in the conservative transposition. This figure shows the mechanisms in the conservative transposition. The transposable element is transposed from the donor to a target site. What will happen to the donor site is unclear. The donor molecule may be destroyed or repaired by the host.

In the above types of transpositions, transposable element translocation results from DNA breakage and joining reaction (14). This DNA-mediated transposition, the transposase enzyme is positioned at the transposon termini and assembled into an active complex. It executes the DNA breakage reactions that cut the tips of the transposon away from flanking DNA at the donor site, and then joins these exposed ends to the target DNA (19).

It is known that genetic transformation can also be transposed through RNA. In order to distinguish between the DNA-mediated and RNA-mediated modes, the RNA-mediated mode has been termed retrotransposition. In this RNA-mediated mode of transposition, mobile elements called retrotransposons transpose via an RNA transcript that is reverse transcribed to DNA before insertion into the target site. The DNA forms of retrotransposons, generated by reverse transcription of their RNAs, are the molecules that actually undergo transpositional recombination via DNA breakage and joining reactions (20).

Transposition proceeds by endonucleolytic cleavage of the phosphodiester bonds at the ends of the transposable element and transfers to these ends into a target DNA molecule. This recombination reaction requires assembly of a highly organized nucleoprotein ternary complex which includes the target DNA, the transposon ends, and transposase (21).

Although it is convenient to analyze transposase function in terms of domain structure, it should be kept in mind that individual functions are not necessarily accomplished by a single isolated domain but may be assumed by several different regions of a single polypeptide or by more than one polypeptide in a multimeric complex (21). However, a general pattern for the functional organization of transposase is emerging: sequence specific DNA-binding activities are generally located in the N-terminal region, whereas the catalytic domain is often localized toward the C-terminal end. This catalytic domain is most likely responsible for a DNA cleavage and strand transfer during transposition process (22). This arrangement has been observed for *Tc1*, *P*, and *Ac* elements (23,24,25).

A key feature of transposases is their capacity to specifically recognize and bind the transposon ends. A common helix-turn-helix (HTH) motif is often found at the sequence-specific DNA-binding domain both in prokaryotic and eukaryotic elements. This motif is potentially able to provide binding specificity (21). The domain may be simple, as appears to be the case for transposases of IS elements, or bipartite and able to recognize different DNA sequences, as found in the *Tc1* and *Ac* transposases (23,25).

2.2. Classifications of transposable elements

Transposable elements are widespread in all major phylogenetic groups (2). Mobile elements have been found in all genomes where they have been sought (26). They seem to be ubiquitous, and represent a quantitatively important component of genomes. Studies of these agents of genomic instability, spanning more than 50 years, have revealed much information about their mode of transposition and structures. Transposable elements can be classified according to their mechanism of transposition, and fall into two main classes. They consist of elements that transpose by reverse transcription of an RNA intermediate, Class I elements, and elements that transpose directly from DNA to DNA, Class II elements (27).

Class I elements or retrotransposons are a major class of eukaryotic mobile elements and transpose via an RNA transcript that is reverse transcribed to DNA before insertion (28). Depending on their structure, retrotransposons are distinguished into those flanked by long terminal repeats (LTR) and the non- LTR retrotransposons. The structure of LTR retrotransposons resemble that of retroviruses with the major difference that retrotransposons are not infectious and do not have the envelope (*env*) gene (20). According to their sequence homology and the organization of coding domains, LTR retrotransposons can be classified into two groups. They consist of the *Ty1-copia* group, in which the endonuclease domain is found upstream of the reverse transcriptase domain, and the *gypsy* group, in which the organization of the coding domains is similar to that of retrovirus (29) as shown in figure 2.3.

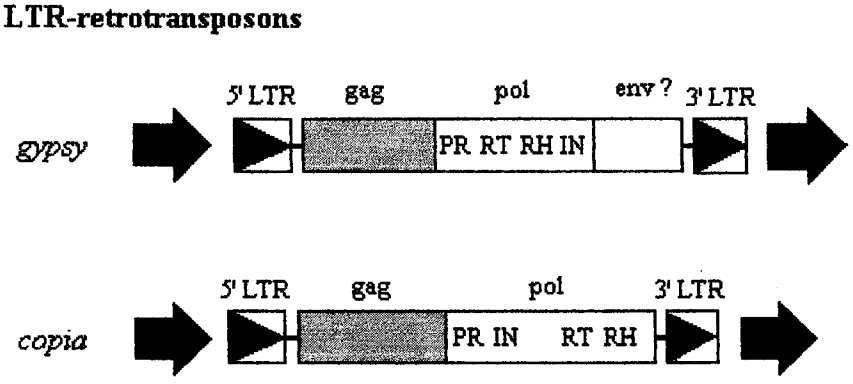

Figure 2.3 The structures of LTR-retrotransposons. This figure shows the structures of LTR-retrotransposons that can be devided into two groups, the gypsy-type and copia-type, on the basis of the order of the pol products.
(PR, protease; RT, reverse transcriptase; RH, RnaseH; IN, integrase).

For the non-LTR retrotransposons, these elements also contain retroviral-like *gag* and *pol* genes but do not contain LTRs. This group of retrotransposon has also been called the LINE-like element or the poly (A) – type retrotransposons (29) as shown in figure 2.4. Typical features of these elements include a 3' poly (A), variable length target size duplications, and multiple copies, which are truncated at their 5' ends (20). Several of these elements have been reported from *D. melanogaster* such as *F, G, I* and *jockey* elements (20). The diversity of non-LTR retrotransposons which is revealed by the deep branches on the tree in this group and their wider distribution than another classes of retrotransposons supports the suggestion that the non-LTR elements are the oldest group of retroelements (29).

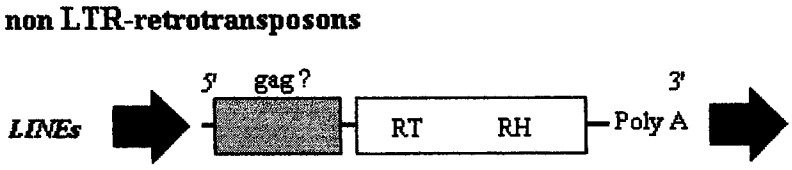

Figure 2.4 The structures of non LTR-retrotransposon. This figure shows the structure of non LTR-retrotransposon that contains RT (reverse transcriptase) and RH (RnaseH).

Class II elements are transposable elements that have short terminal inverted repeats and transpose directly from DNA to DNA as shown in figure 2.5. They generally move by an excision-insertion mechanism and encode proteins, transposase, required for their own transposition (30). During transposition the ends of an element must be brought together, released from flanking sequences and joined to target DNA. These steps are mediated by transposase acting alone or in concert with other proteins (30). If the sites at which the target is cleaved are staggered, as always appears to be the case, then there will be a short single-stranded region at the junction between the transposon and the flanking DNA. Repair of these regions by host factors generates target site duplications. These reaction have been studied *in vitro* using the transposase of the *Tc3* elements of *Caenorhabditis elegans* (31), the *P* element of *D. melanogaster* (32) and *Himar 1* of *Haemotobia irritans* (33).

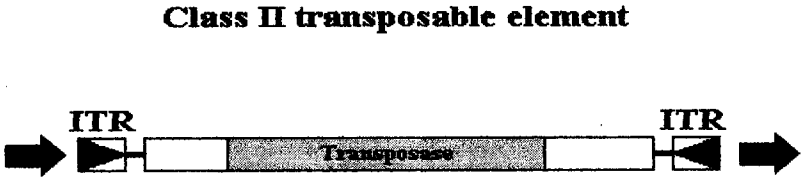

Figure 2.5 The structure of class II transposable element. This figure shows the structure of class II transposable element. Class II elements have short inverted terminal repeats (ITR) and contain a gene encoding transposase.

Since eukaryotic transposable elements were first identified and characterized at the molecular level, a lot of information has been discovered about their DNA sequences. According to their sequence similarity, mobile elements can be classified into families such as *Tgm* family, *Tc1/mariner* family and *hAT* family (2,23,34).

2.3. The *hAT* transposable elements family

Transposable elements are widespread in all major phylogenetic groups. They appear to be ubiquitous genomic component of most organisms, with many sharing common ancestors. Recently, a structural relationship has been established among mobile elements found in both plant and insect species. These transposable elements are considered to be members of the *hAT* transposon family. The *hAT* superfamily is a group of class II transposable elements from animals and plants that share structural and functional characteristics (35). The *hAT* stands for the *hobo* element from *D. melanogaster*, the *Ac* element from *Z. mays*, and the *Tam3* element from *A. majus* (2).

The *hobo* element belongs to the *hAT* family of elements. It produces a transposase which is thought to catalyze a DNA-mediated transposition reaction, and create an 8 bp. duplication of host DNA upon insertion.The first probable full size element was the $hobo_{108}$ element identified by Streck, MacGaffey and Beckendorf (36). It is 3017 bp. in length and has 12 bp. inverted terminal repeats. More recently Calvi et al. have reported the sequence of a hobo element, *HFL1*, which encodes a functional transposase (2). It differs slightly from $hobo_{108}$, and has one short open reading frame (ORF0) and a long one (ORF1) of 1983 bp. The ORFs are separated by a single in frame stop codon.

Ac element is 4565 bp. in length and has 11 bp. long terminal inverted repeats. It creates 8 bp. target site duplications upon insertion (37). The first ATG codon on the *Ac* mRNA opens a 2421 bp. ORF_a which encodes the transposase protein (37). In Western blotting experiments, a protein with an apparent molecular weight of 112 kD was detected exclusively in nuclear endosperm extracts from maize lines containing a

functional *Ac* element. The N- terminal 200 amino acids of the *Ac* transposase have no homology to any known protein, whereas the C-terminal 600 amino acids are on average about 30 % homologous to the putative transposase of the *Tam3* element from *A. majus* (38).

Tam3 transposable element was cloned from the *nivea* locus of line 98 (*niv-98*) encoding the chalcone synthase in genome of *A. majus* or snapdragon (39). It is 3629 bp. long and has 12 bp. inverted terminal repeat (39). The *Tam3* element contains an open reading frame of 2.48 kb. long. A 3.0 kb. long cDNA clone that is colinear with the genomic sequence confirms that this open reading frame is transcribed and the element has no intron (38).

In 1991, Feldmer and Kunze (40) have shown that these three elements exhibit levels of coding sequence similarity that span several hundred codons as shown in figure 2.6. All three element generate 8 bp target site duplications upon insertion and leave similar empty excision site that distinctly different from those observed for other transposable elements. These data suggest that *hobo*, *Ac*, and *Tam3* belong to a family of related transposable elements (41). The *hAT* transposable elements are widespread in both plants and animals. The identification of related, functional element from other species would allow us to understand better the evolutionary history of the *hAT* family of elements.

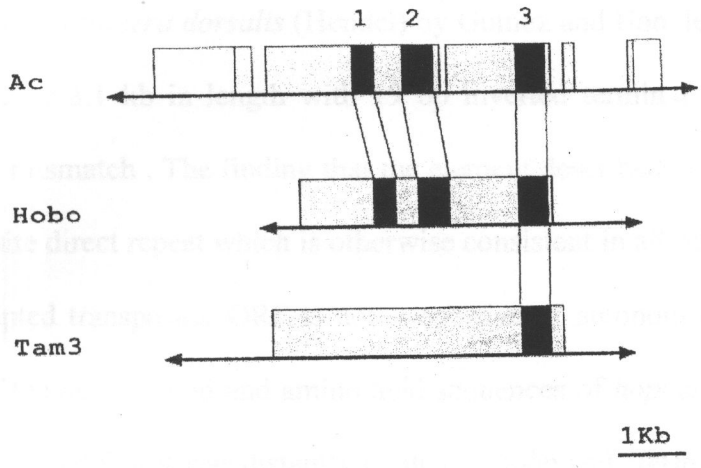

Figure 2.6 The schematic of similar regions among *Ac*, *hobo*, and *Tam3* transposase proteins. This figure shows the schematic representation of the position of the regions that are similar among *Ac*, *hobo*, and *Tam3* amino acids sequences. Double arrowhead lines represent the complete element. For *Ac*, boxes represent the five exons of the element. Solid boxes represent the regions of high similarity among these three elements. Taken from Calvi, BR. (2).

There is one of the transposable elements in the *hAT* family, called *hopper* transposable element. This element was previously isolated from genome of the oriental fruit fly, *Bactrocera dorsalis* (Hendel) by Gomez and Handler in 1997 (42). It was approximately 3.1 kb in length with 19 bp inverted terminal repeat sequences having a single mismatch . The finding that the element described does not contain an 8 bp insertion site direct repeat which is otherwise consistent in all other *hAT* elements and the interrupted transposase ORF is not supportive of autonomous function (42). Comparison of the nucleic acid and amino acid sequences of *hopper* element to other *hAT* elements showed that it was distantly related to *hobo* and *Hermes* and to putative *hAT* elements from *B. cucurbitae* and *C. capitata* (43). It suggests that the *hopper* element may represent an ancient link to the non-insect *hAT* element.

CHAPTER III

OBJECTIVE

3.1 Molecular characterizing the gene of *hopper* transposable element from genome of *Bactrocera tau* (Walker).

3.2 Structurally analyzing of a putative-functional *hopper* transposase gene in *Bactrocera tau* (Walker).

3.3 Studying the genomic representation of *hopper* transposable element in genome of *Bactrocera tau* (Walker) in Thailand.

CHAPTER IV

MATERIALS AND METHODS

4.1. Fly stocks and DNA isolation

All the species and strains used in the present study are listed in Table 4.1; each listing includes the locations of fruit fly collection. Genomic DNA of the adult flies were extracted by the Lifton's method as described by Thanaphum and Haymer in 1998 (44).

Table 4.1 *Bactrocera* strains used in this study.

Species	Strains	Locations [1]
Bactrocera tau	UB NA RN	Ubol Ratchatani Nan Ranong
Bactrocera dorsalis	-	Bangkok
Bactrocera cucurbitae	-	Ratchaburi

[1] Provinces of Thailand.

4.2. Designing oligonucleotide primers

The oligonucleotide primers were designed by using the computer program, OLIGO, with the DNA sequences data. All of the oligonucleotide primers that were used in this thesis are listed in table 4.2. These primers were used as PCR primers and primers for sequencing both strands of the entire *hopper* gene as shown in figure 4.1.

Table 4.2. Length (bp) and position of primers used in this study.

Primer	Length (bp)	Strand
HpI	20	Plus
HpII	21	Minus
ITR	19	Plus
HpRI	20	Minus
HpRII	20	Plus
SEQ I	17	Plus
SEQ II	17	Plus
SEQ III	17	Minus
hop I	17	Minus
hop II	17	Minus
hop III	17	Minus
hop IV	17	Plus
hop V	17	Minus
hop VI	17	Minus
hop VII	17	Plus
hop VIII	17	Plus
hop IX	20	Minus
hop X	21	Plus

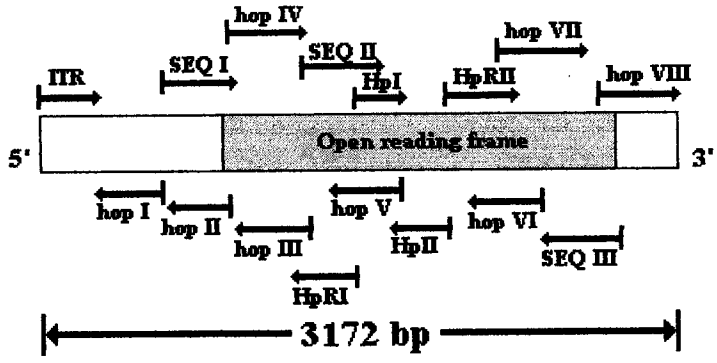

Figure 4.1. The positions and directions of primers used for both strands sequencing of *hopper* element. Diagram shows the positions and directions of primers that were used for DNA sequencing in both strands of *hopper* transposable element gene. Solid box represents the putative open reading frame of the *hopper* transposase gene.

4.3. PCR amplification of the *hobo*-related transposase fragment in *B. tau*

The oligonucleotide primers, HpI and HpII were used to amplify the coding region of *hobo*-related elements from genomic DNA prepared from the adults of *B. tau* (Walker) specimens. Each PCR reaction contained 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 20 pM of each primers, 10 ng of template DNA, and 1 unit of *Taq* polymerase (AmpliTaq, Perkin-Elmer) in a total volume of 50 ul. PCR amplification was performed in a Perkin-Elmer 2400 thermocycles programed at 94 °C (2 min) then 35 cycles of 93 °C (1 min), 50 °C (1 min), 72 °C (2 min) and a final extension step at 72 °C (10 min). The amplified products were then size fractionated by electrophoresis on 1.0 % agarose gel, stained with ethidium bromide and visualized on a UV transilluminator.

4.4. Isolation of the *hopper* transposable element gene from genome of *B. tau*

The oligonucleotide primer ITR was used to amplify the *hopper* gene from genomic DNA of adults *B. tau* (Walker) which were collected from Ubol Ratchathani province, Thailand. The PCR reaction contained chemicals as described above. The PCR profile was at 94 °C (2 min) then 30 cycles of 93 °C (1 min), 50 °C (1 min), and 72 °C (4 min) and a final extension step at 72 °C (10 min). A 3.1 kb amplified PCR product was then size fractionated by electrophoresis in 1.3 % agarose with ethidium bromide stain and visualized by UV luminescence. This PCR product was cloned in pCR II (Invitrogen) plasmid vectors and transformed to *Escherichia coli* host cells for DNA sequencing.

4.5. Inverse PCR (IPCR) analysis

To isolate the regions that flank the insertion sites of the *hopper* element, 2 ug of genomic DNA was digested with MspI (Promega), which do not have restriction sites within this *hopper* element. The digested DNA was extracted with phenol:chloroform and precipitated in ethanol. 50 ng of the digested DNA directly self-circularized in a total volume of 100 ul at 14 °C by incubating with 3 units of T4 ligase (Promega) for 20 hrs. After self-ligation, DNA was precipitated and resuspended in ddH_2O. A 10 ng of DNA was nicked at 94 °C, 30 sec and used as template in IPCR. The oligonucleotides hop IX and hop X were used as inverse primers. PCR conditions were 10 min at 94 °C followed by 30 cycles at 93 °C for1 min, 50 °C for 1 min, and 72 °C for 4 min and a final elongation step at 72 °C for 10 min.

4.6. Agarose gel electrophoresis

Agarose gel electrophoresis is a standard method used to separate, identify and purify DNA fraction. Agarose gel can be prepared in various concentration according to the size of DNA to be analyzed by dissolving agarose in TBE buffer. The melted solution was then added with ethidium bromide solution and poured into a gel chamber set with a comb of desired thickness and number of wells and allow to harden. DNA markers and DNA sample were mixed with 1/6 volumes of loading dye before being loaded into the wells. Electrophoresis was performed at 50-100 Volts for 1-16 hr by using TBE buffer. After electrophoresis, the gel was visualized under UV light and photographed for analyses.

4.7. Cloning strategy

Cloning in plasmid vectors consists of the cleavage of plasmid DNA with a restriction enzyme and joining *in vitro* to foreign DNA. The resulting recombinant plasmids are then used to transform bacteria. Plasmid pCR II (Invitrogen) and pUC19 were used in cloning procedures. The DH5α strain of *Escherichia coli* was used for bacterial transformation and plasmid propagation. The bacteria carrying recombinant plasmids therefore form white colonies over the surface of LB agar plate containing ampicillin and X-gal. All of the clones were kept in glycerol stock at –20 °C.

4.8. Purification of DNA fragment from agarose gel by GENECLEAN Kit (BIO101)

After gel electrophoresis, the desired DNA band was excised from ethidium bromide- stained TBE agarose gel under the UV light, cut and transferred to a microcentrifuge tube and determined its weight (0.1 g or 100 ul). Subsequently, 3 volumes of NaI stock solution and 0.5 volumes of TBE modifier (BIO 101) were added and the tube was incubated at 50 °C for 5 min. When agarose gel was completely dissolved, the solution was mixed with 5 ul of the GLASSMILK suspension and kept on ice for 5 min. The silica matrix with the bound DNA was collected by centrifugation for 15 sec and washed 3 times with the NEW WASH solution. The DNA was eluted from the matrix by adding of ddH_2O and incubating at 50 °C for 3 min. After spinning down for 1 min, the supernatant containing the eluted DNA was transferred to a new tube.

4.9. Southern blot analysis

4.9.1 Preparation of DNA probes by DIG labeling

Clone of 470 bp. *hopper* transposase fragment and clone of a complete *hopper* gene from *B. tau* were used to prepare DNA probes using DIG DNA Labeling system from Boehringer Mannheim. DNA fragments were denatured by heating for 10 min in a boiling water bath and quickly chilled on ice. The reaction mixture consisted of 2 ul of hexanucleotide, 2 ul of dNTP mixture, sterile water to a final volume of 19 ul and 1 ul of Klenow enzyme. The reaction was incubated for 20 hrs at 37 °C and then stopped by adding 2 ul of 0.2 M EDTA pH 8.0. The labeled DNA was precipitated by adding 2.5 ul of 4 M LiCl and 60 ul of prechilled absolute ethanol, incubated for 2 hr at –20 °C. The solution was spun in a microcentrifuge for 2 min and pellet was washed with 50 ul of cold 70% ethanol. The DNA pellet was dried and dissolved in 50 ul of TE buffer.

4.9.2 Gel treatment

In order to transfer DNA from agarose gel to nylon membrane, the agarose gel from protocol 4.6 was treated with the denaturation solution (0.5 M NaOH and 1.5 M NaCl) for 1 hr with shaking at room temperature. After that, the gel was shaking with the neutralization solution pH 8.0 (1.0 M Tris-Base, 1.0 M Tris-HCl, and 1.5 M NaCl) for 1 hr at the room temperature.

4.9.3 Transfer DNA to membrane

The transfer set illustrated in figure 4.2 comprised of 2X SSC-soaked Whatman paper, the treated gel, nylon membrane, stack of dry paper towel and a light weight at the top. The transfer step was complete after overnight.

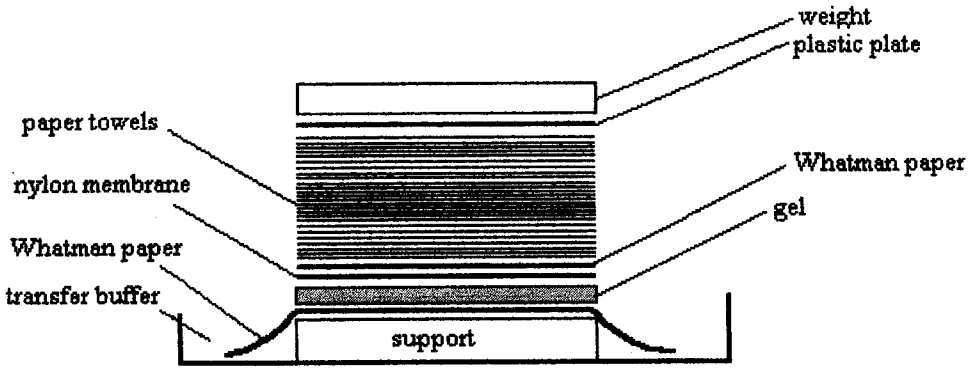

Figure 4.2. Capillary transfer of DNA from agarose gel. Buffer is drawn from a reservoir and passed through the gel into a stack of paper towels. The DNA is eluted from the gel by the moving stream of buffer and is deposited on a nylon membrane. A weight applied to the top of the paper towels helps to ensure a tight connection between the layers of material used in the transfer system.

4.9.4 Hybridization and washing

The nylon membrane from protocal 4.9.3 was prehybridized with prehybridization solution at 42 °C for overnight with gentle agitation. DIG-labeled DNA probe was denatured by boiling at 95 °C for 10 min and rapidly added to the prehybridized membrane and the membrane was then incubated for 16 hrs. The membrane was non-stringently washed twice with 2 x SSC, 0.1% SDS at room temperature for 5 min each. Stringent wash was performed twice with solution of 0.5 x SSC, 0.1% SDS at 60 °C for 15 min each. Gentle agitation was necessary during all steps of washing. The membrane was then color detected.

4.9.5 Chromogenic detection

Genomic DNAs of the *Bactrocera* strains were digested with the restriction enzymes. The detection and analysis of *hopper* homologous sequences were done by the kit protocal . The probes were labeled with the nonradioactive digoxygenin systems of Boehringer Mannheim . The final wash of the nylon membranes was 0.5 x SSC, 43 % formamide at 60 °C.

4.10. Sequencing and data analysis

All of the cloned PCR products were sequenced by the Bio Service Unit (BSU) of National Science and Technology Development Agency of Thailand. For the data analysis, DNA sequences were analyzed by the computer programe : CLUSTAL W is used for sequences alignment (both of amino acid sequences and nucleotide sequences).

CHAPTER V

RESULTS

5.1. Amplification and cloning of the *hobo*-related fragment in *Bactrocera tau.*

Degenerated oligonucleotide primers (Hp I and Hp II) based on the conserved region in the transposase coding regions of the *hobo* element of *D. melanogaster* and the *Ac* element of *Zea mays* as shown in figure 5.1 were designed and used in polymerase chain reaction (PCR) to determine if the *hobo*-related sequences were present in the genome of *B. tau.* An expected fragment of approximately 470 bp in length was observed and appeared to be consistent in size with PCR product expected from a *hobo* element template as shown in figure 5.2. The PCR fragment was subsequently cloned and sequenced. A search of the sequence database using the BLAST program revealed a high scoring DNA identity with one of the *hAT* elements called *hopper* transposable element of *Bactrocera dorsalis* (accession number U70428) (42). DNA sequences analysis by the alignment of this 441 bp. fragment as shown in figure 5.3 revealed 94% DNA identity of these sequences. The DNA sequence alignment shows that there are deletions in the *hopper* transposase fragment from *B. tau* when compared with the previous isolated *hopper* element of *B. dorsalis.* This result also indicates that our 441 bp DNA fragment is a fragment of *hopper* transposase but might not be a functional copy.

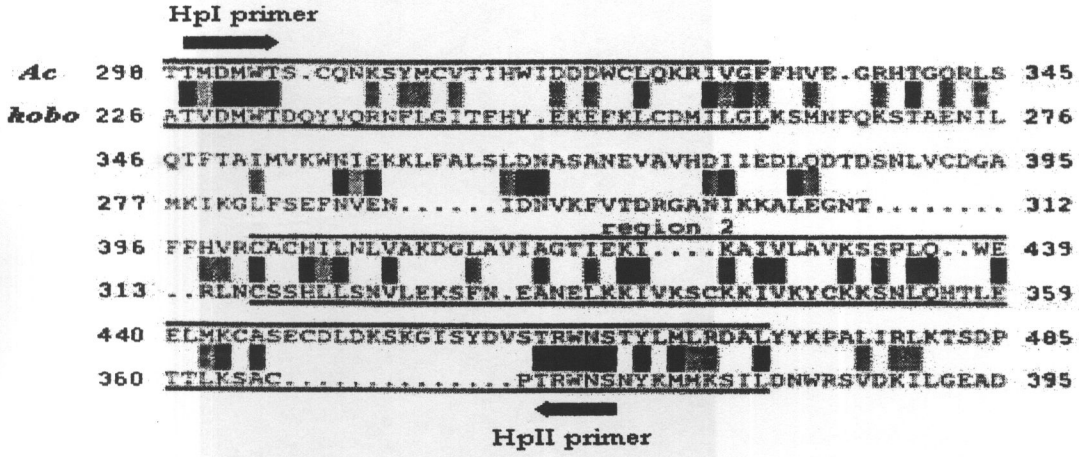

Figure 5.1 Amino acid sequences alignment of *hobo* and *Ac* transposase fragments. The amino acid sequence of *hobo* and the *Ac* transposase fragments were aligned. The forward primer was designated as HpI corresponding to the T (V/M) DMWT amino acid sequence consensus between hobo and Ac, and the reverse primer was designated as HpII corresponding to the TRWNS consensus. Solid boxes between the sequences indicate identities, and stippled boxes indicate conservative changes. Dots within the sequences indicate gaps created to optimize the alignment.

(Modified from Calvi et al (2))

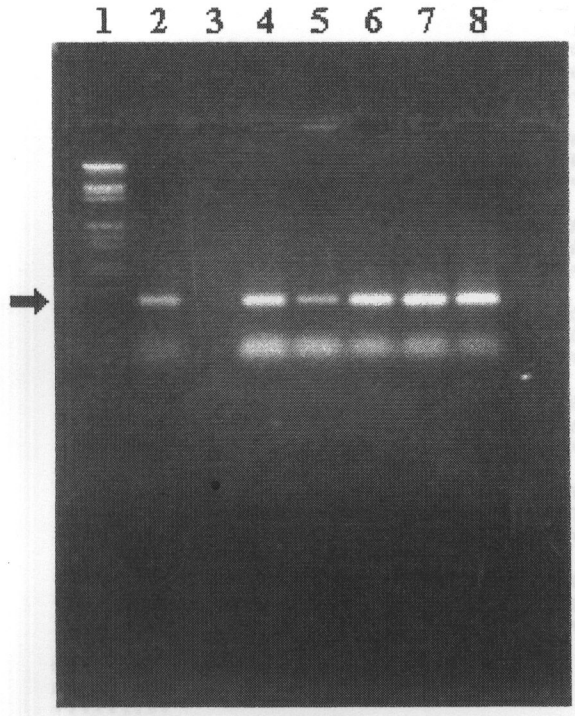

Figure 5.2 PCR products of *hobo*-related fragments from *B. tau*. This figure shows the amplified PCR products of the *hobo*-related fragments from the genome of *B. tau* . Oligonucleotide primers HpI and HpII were used for this PCR reaction. The λ-*Hind* III-*Eco*RI DNA markers are shown in lane number 1. Lane number 4 is a positive control from the initial PCR product template. At the beginning of PCR reaction, the initial PCR product in lane number 2 was just a little amount of the amplified product. For the PCR optimization, the concentrations of genomic DNA templates were varied from 1, 10, 25, and 100 ng. The PCR products of these variations are shown in the lane number 5-8 respectively. The 441 bp. PCR products were indicated by the arrow and nearly consistent with the predicted 470 bp. fragment of the hobo-related fragment from genome of *B. dorsalis.*

Figure 5.3 Nucleotide sequence alignment. Nucleotide sequences alignment based on Clustal W (1.74) analysis between the DNA sequences of 441 bp. PCR product amplified by using HpI and HpII primers with genomic DNA from *B. tau* (Btau-HP) and the corresponding sequences from the *hopper* transposable element that was previously isolated from genome of *B. dorsalis* (hopper1) (42). The identical sequences are indicated by asterisks and the stop codons of each fragment are underlined.

5.2 Structural features of *hopper* transposable element from *Bactrocera tau*.

An oligonucleotide primer specific to the inverted terminal repeats of *hopper* transposable element published by Handler and Gomez, ITR primer, was used in a PCR reaction to investigate the *hopper* transposable element gene from genome of *B. tau*. As shown in figure 5.4 the largest band of amplification products is consistent with the previous data from *Bactrocera dorsalis* (42) which predict that this primer would generate a 3.1 kb. PCR product from a full-length *hopper* element. The variation in the size and number of internal deleted *hopper* element was also observed. The amplified products were cloned and the 3.1 kb. band was selected for both strands sequencing.

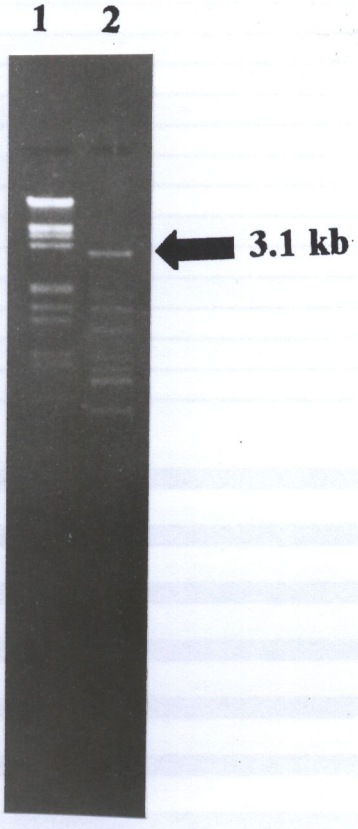

Figure 5.4 PCR products of *hopper* transposable element gene from *B. tau*. Ethidium bromide stained agarose gel containing putative *hopper* related sequence amplified by PCR from genomic DNA of *B. tau* with the ITR primer. The λ-*Hind* III-*Eco*RI DNA markers are in lane 1 and the arrow at right indicates the 3.1 kb PCR product.

TAGTGTTGGGGACTATCGA----------

61 ----------
121 ----------
181 ----------
241 ----------
301 ----------
361 ----------
421 ----------
481 ----------
541 ----------
601 ----------
661 ----------
721 ----------
781 ----------
841 ----------
901 907

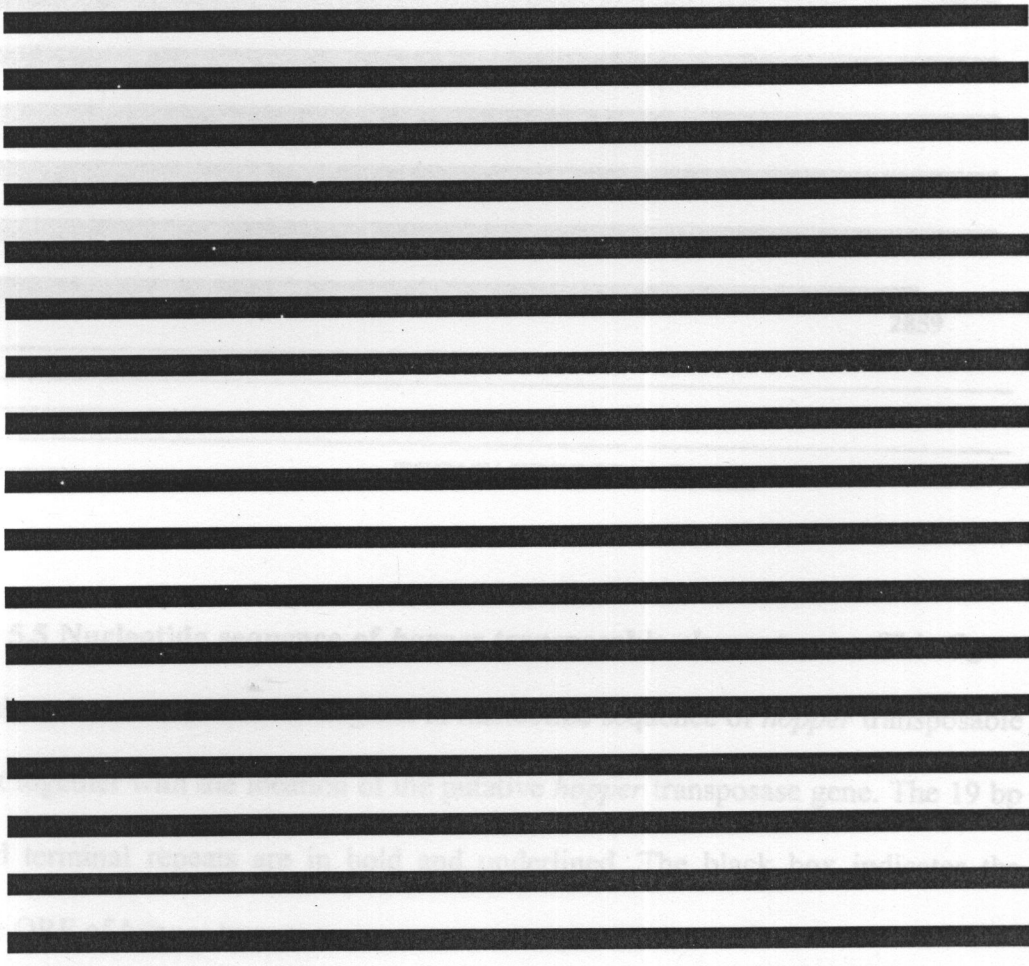

Figure 5.5. For legend see next page.

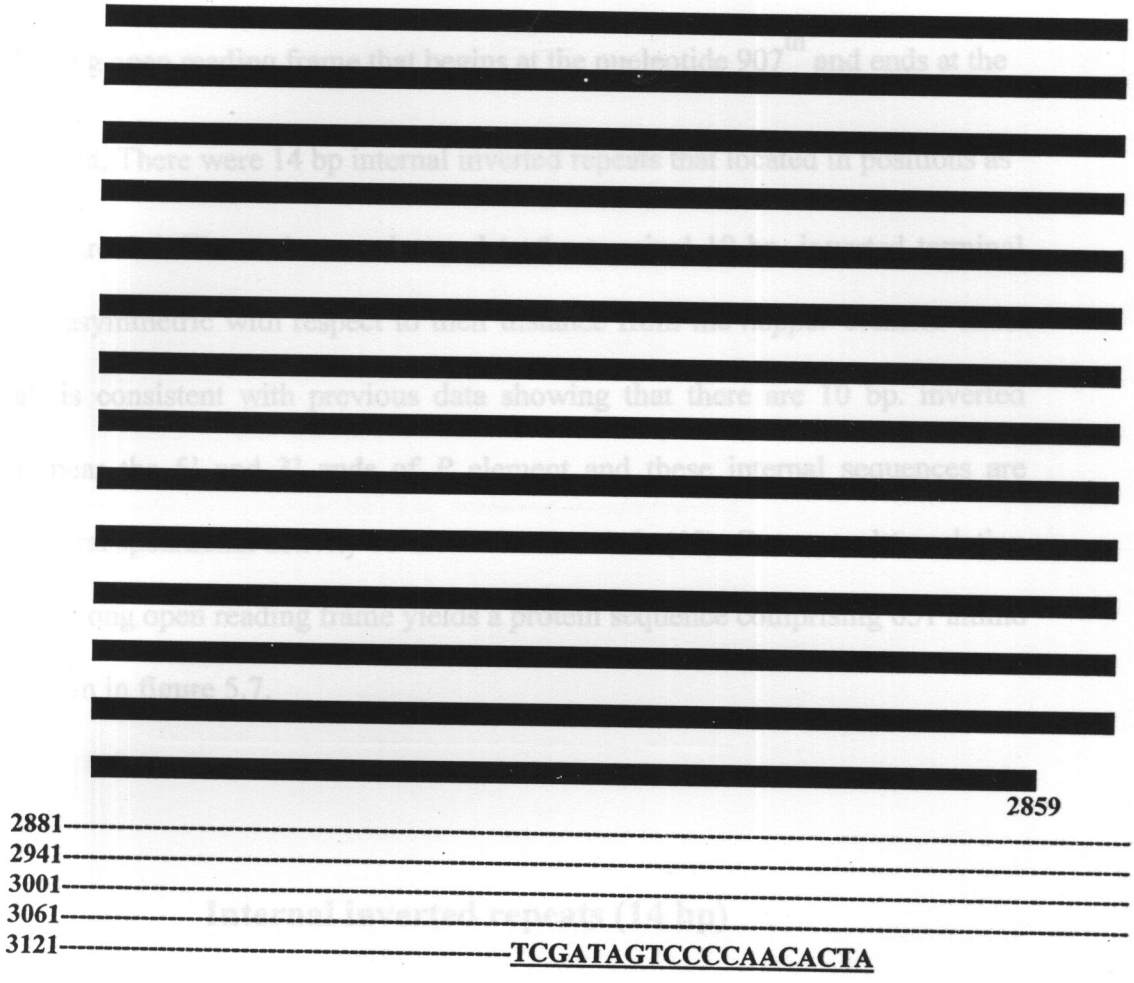

Figure 5.5 Nucleotide sequence of *hopper* transposable element gene. This figure shows the schematic structural diagram of nucleotide sequence of *hopper* transposable element together with the location of the putative *hopper* transposase gene. The 19 bp inverted terminal repeats are in bold and underlined. The black box indicates the putative ORF of *hopper* transposase.

Examination of the DNA sequences shown in figure 5.5 revealed the presence of a single long open reading frame that begins at the nucleotide 907th and ends at the 2859th position. There were 14 bp internal inverted repeats that located in positions as shown in figure 5.6. These sites are internal to the terminal 19 bp. inverted terminal repeats and asymmetric with respect to their distance from the *hopper* element ends. This result is consistent with previous data showing that there are 10 bp. inverted sequences near the 5' and 3' ends of *P* element and these internal sequences are required for transpositional activity of the *P* element ends (45). Conceptual translation of this single long open reading frame yields a protein sequence comprising 651 amino acids as shown in figure 5.7.

Internal inverted repeats (14 bp)

Nucleotide Sequences	Position (5'-3')
A T T T T T T C G T T C A T	96-109
A T G A A C G A A A A A A T	3088-3101

Figure 5.6 Nucleotide sequences of internal inverted repeats of *hopper* element. Nucleotide sequences of the 14 bp internal inverted repeats of *hopper* transposable element from genome of *B. tau* are shown with their positions.

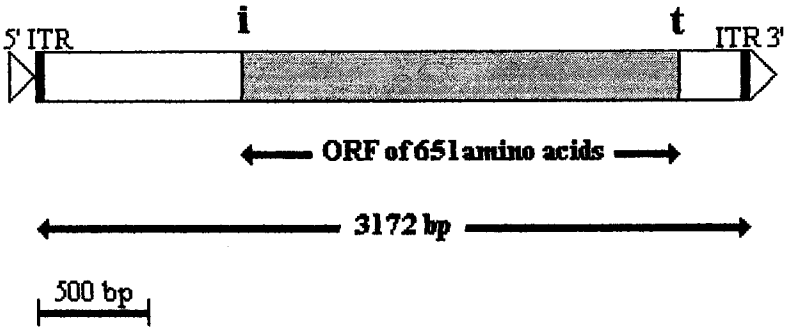

Figure 5.7 Schematic structural diagram of *hopper* element. The 19 bp inverted terminal repeats are blacked. The boxed area represents the element itself, which is flanked by target site duplications represented by white triangles. The element carries a unique long open reading frame (ORF) potentially encoding a putative transposase protein of 651 amino acids. The i and t letters indicate the position of *initiation* and *termination* codons of this putative transposase protein.

5.3. The deduced amino acid sequence of *hopper* transposase predicts a protein similar to *hobo*, *Ac*, and *Tam3* transposase.

Using the deduced *hopper* transposase sequence, the ORF sequence within the *hopper* element was compared with the known *hobo*, *Ac*, and *Tam3* transposase sequence. Each of these three elements encodes a transposase that catalyzes its mobilization. The ORF1 of *hobo* is 661 amino acid in length. *Ac* transposase is thought to be 807 amino acids in length and the *Tam3* transposase is encoded by an ORF of 748 amino acids beginning from the first AUG (2). Comparisons of the entire length *hopper* amino acid sequences with these three elements show that these transposase protein sequences have a similar level of distant relationship ranging from 22% to 30% identity as shown in table 5.1.

Table 5.1. Pairwise comparisons of the amino acid sequence identities between putative transposase protein of *hobo*, *Ac*, *Tam3*, and *hopper* elements.

	hobo	***Ac***	***Tam3***	***hopper***
hobo				
Ac	22			
Tam3	23	30		
hopper	25	29	25	

The amino acid sequence identities analyzed from the entire length of amino acid sequences of *hAT* transposase ORF are shown. These data were based on Clustal W analysis between these amino acid sequences.

To better define the relationship among these *hAT* elements, the amino acid sequences of their transposase proteins were aligned by using the Clustal W program. Each of three elements encodes a transposase that catalyzes its transposition. The ORF1 of *hobo* transposase is 661 amino acids in length (2). *Ac* transposase, encoded within 5 exons, is thought to be 807 amino acids in length (37). *Tam3* transposase is encoded within a single exon, with an ORF of 748 amino acids (38). Our putative ORF of 651 amino acids was aligned with these three amino acid sequences. There are three conserved regions of these transposase sequences which were designated to A, B and C region respectively (modified from Feldmar and Kunze, 1991 (40)) as shown in figure5.8. The position of these amino acids regions located in the middle and C-terminal regions of the transposase proteins of each element are shown in table 5.2. The amino acid sequence identity of A, B, and C regions are shown in table 5.3. The highest level of identity is found in the C region with ranging from 32% to 60% identity of amino acid sequences. This result also indicates that our putative functional *hopper* transposase sequences has a similar level of genetic distant relationship to *hobo*, *Ac*, and *Tam3* transposase sequences. The secondary structure prediction analysis by using the PROSIS program shows that there were several of the helix-turn-helix structures among these three regions of these transposase sequences (data not shown).

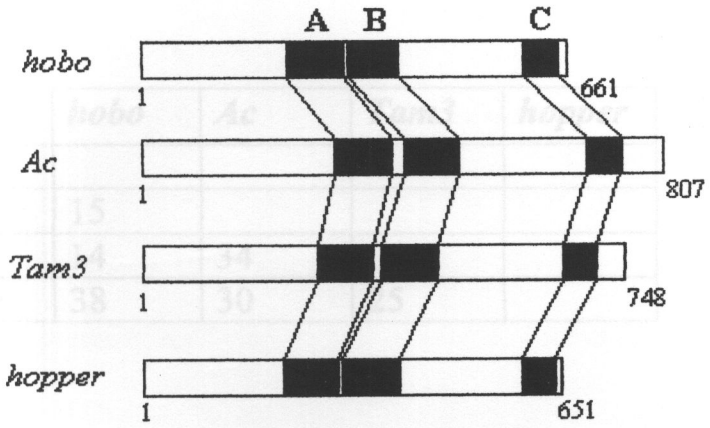

Figure 5.8 Schematic representation of similar regions among *hobo*, *Ac*, *Tam3* and *hopper* elements are shown. The boxes area represents the ORF of the transposase sequences of each element. The conserved regions among these sequences are designated to A, B and C and indicated with blacked regions.

Table 5.2. The amino acid positions in A, B, and C fragments of *hobo*, *Ac*, *Tam3*, and *hopper* transposase sequences.

	hobo	***Ac***	***Tam3***	***hopper***
A	227-313	297-384	270-354	218-301
B	315-392	400-486	368-453	306-394
C	587-639	684-736	645-697	582-634

Table 5.3. The identities of A, B, and C fragments of *hopper* and *hAT* transposase sequences.

(A)	***hobo***	***Ac***	***Tam3***	***hopper***
hobo				
Ac	15			
Tam3	14	34		
hopper	38	30	25	

(B)	***hobo***	***Ac***	***Tam3***	***hopper***
hobo				
Ac	18			
Tam3	14	37		
hopper	24	26	24	

(C)	***hobo***	***Ac***	***Tam3***	***hopper***
hobo				
Ac	36			
Tam3	40	60		
hopper	42	34	32	

These data were based on Clustal W analysis between these transposase amino acid sequences. The most identity is the C region and seems likely that this region has a functional role in the transposition of *hAT* elements.

Comparisons of the deduced *hopper* transposase sequence with those of the *hobo*, *Ac* and *Tam3* elements show that the putative *hopper* transposase protein sequence may be one of the members of the *hAT* element family. Moreover, the putative *hopper* ORF contains the conserved region between *hopper* and the other *hAT* elements. As shown in figure 5.9, the region exhibits the highest levels of sequence conservation among these *hAT* element transposases and is located near their C-termini. Given their conservation, it seems likely that this region has a functional role in the transposition of *hAT* elements. However, the function of this conserved region has yet to be established.

Figure 5.9 Amino acid sequence alignment of C-terminal for the *hAT* elements. Alignment of the C - terminal amino acid sequences for the *hAT* elements and *hopper* is shown. identical residues are indicated by white letters on a black background. Data presented here are obtained from the following GenBank accession numbers: *hobo* (*D. melanogaster*), M69216; *Hermes* (*M. domestica*), L34807; *homer* (*B. tryoni*), AF110403; *hermit* (*L. cuprina*), U22467; *Ac* (*Z. mays*), X05424 and X05425; *tag2* (*A. thaliana*), AF120335; *Slide* (*N. tabacum*), X97569 and *Tam3* (*A. majus*), X55078 respectively.

5.4. DNA sequence determining of target site duplications and inverted terminal repeats (ITRs) of the *hopper* transposable element.

Inverse PCR, a variation of polymerase chain reaction that permit the amplification of regions of unknown sequence that flank a known sequence, was used to amplify and isolate the target site duplications and the inverted terminal repeats of *hopper* element from genome of *B. tau* as shown in figure 5.10. Initially, genomic DNA was digested with *Msp* I, the resulting fragment were circularized and those containing *hopper* sequences were amplified by using PCR approach with oligonucleotide primers, hop IX and hop X, based upon the sequence data of *hopper* element from *B. tau*. Prior to amplify the circularized DNA was nicked by heating at 94 °C for 2 min. The expected 1.4 kb. PCR product as shown in figure 5.11 was cloned and sequenced respectively. This expected 1.4 kb size of IPCR product was estimated from the positions of IPCR primers (hop VIII and hop IX) and 3.5 kb fragment from Southern blot analysis data. This Southern blot consists of the *Msp* I digested genomic DNA fragments of *B. tau* which were hybridized with the 3.1 kb HPTAU probe as described in chapter 4.5. Figure 5.12 shows the DNA sequences result of the target site duplications and the inverted terminal repeats of *hopper* transposable element in genome of *B. tau.*

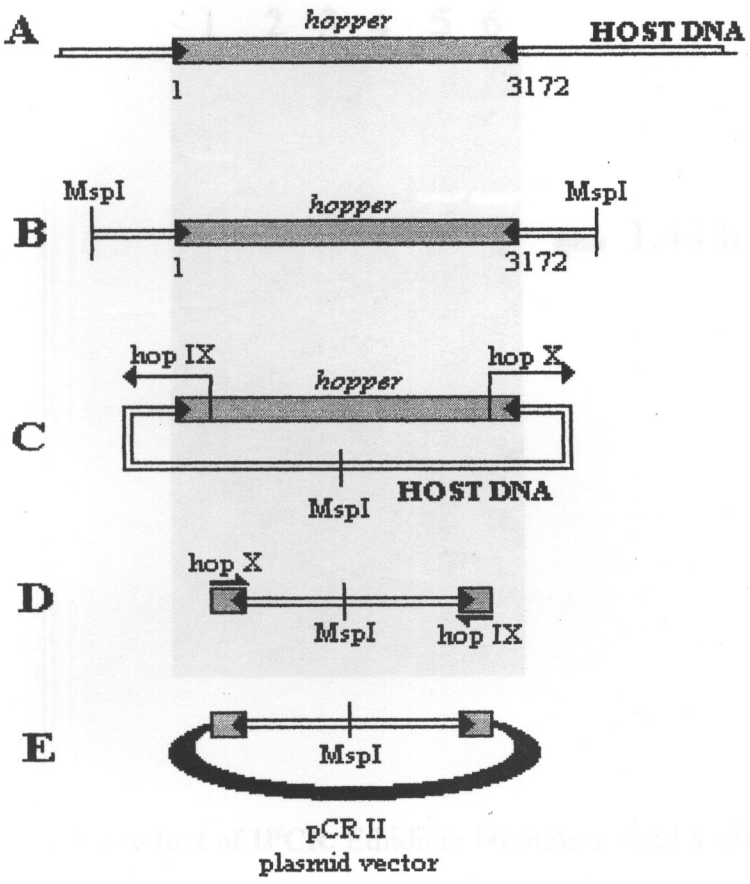

Figure 5.10 IPCR strategy. This figure shows inverse PCR strategy that was used to amplify and isolate the target site duplications and the inverted terminal repeats of *hopper* element. (A) The 3172 bp *hopper* element within the *B. tau* genome is shown. (B) Genomic DNA was digested with Msp I restriction enzyme. (C) Resulting fragments were circularized by self-ligation. (D) The oligonucleotide primers, hop IX and hop X, were used to amplify the flanking region of *hopper* element. (E) The expected PCR product was cloned into a pCR II plasmid vector for sequencing. The inverted terminal repeats are indicated with the black arrowheads.

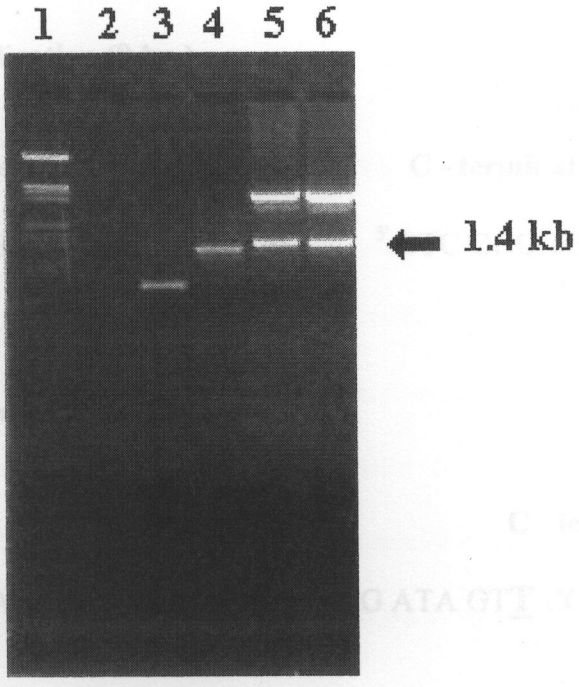

Figure 5.11 PCR product of IPCR. Ethidium bromide stained agarose gel containing the insertion sites and ITRs sequences of *hopper* element. These sequences were amplified by inverse PCR method with hop VIII and hop IX oligonucleotide primers. The λ-*Hind* III-*Eco*RI DNA markers are in lane number 1 and lane number 3-6 are the Msp I digested PCR product, PCR product of IPCR, and two clones of this PCR product in the pCR II plasmid vectors respectively. The arrow indicates the expected 1.4 kb PCR products.

A Target site duplication (8 bp.)

N - terminal	C - terminal
$^{5'}$ GTC GGC TG $^{3'}$	$^{5'}$ GTC GGC TG $^{3'}$

B Inverted terminal repeat (19 bp.)

N - terminal	C - terminal
$^{5'}$ TAG TGT TGG G<u>G</u>A CTA TCG A $^{3'}$	$^{5'}$ TCG ATA GT<u>T</u> CCC AAC ACT A $^{3'}$

Figure 5.12 Nucleotide sequences of target site duplications and ITRs of *hopper* element. (A) Nucleic acid sequences of the 8 bp target site duplications of *hopper* transposable element from genome of *B. tau.* (B) Nucleotide sequence of the left and right inverted terminal repeats were identified by alignment of complementary distal sequences, which yielded 19 bp containing a single mismatch nucleotide (underlined).

Comparison of the left and right terminal sequences of *hopper* element reveals that it is composed of 19 bp. inverted terminal repeats with one single mismatch and 8 bp. of target site duplications. When the inverted terminal sequences of other members of the *hAT* element family are aligned with those of *hopper*, they share a conserved A and G at positions 2 and 5 respectively, in their left inverted terminal sequences and a complementary C and T in their right terminal sequences as shown in figure 5.13. As one of the functions of the transposase is to interact with the inverted terminal repeats it should be interesting to determine the evolutionary relationship between the ITRs sequences of this superfamily elements. The longest 19 bp ITRs of this *hopper* element may represent the most informative data for evolutionary study. Moreover, all these elements generate 8 bp. target site duplications upon transposition. These observations suggest that *hopper* element may excise and transpose using an enzymology similar to that employed by *hAT* elements and may perhaps be evolutionarily related.

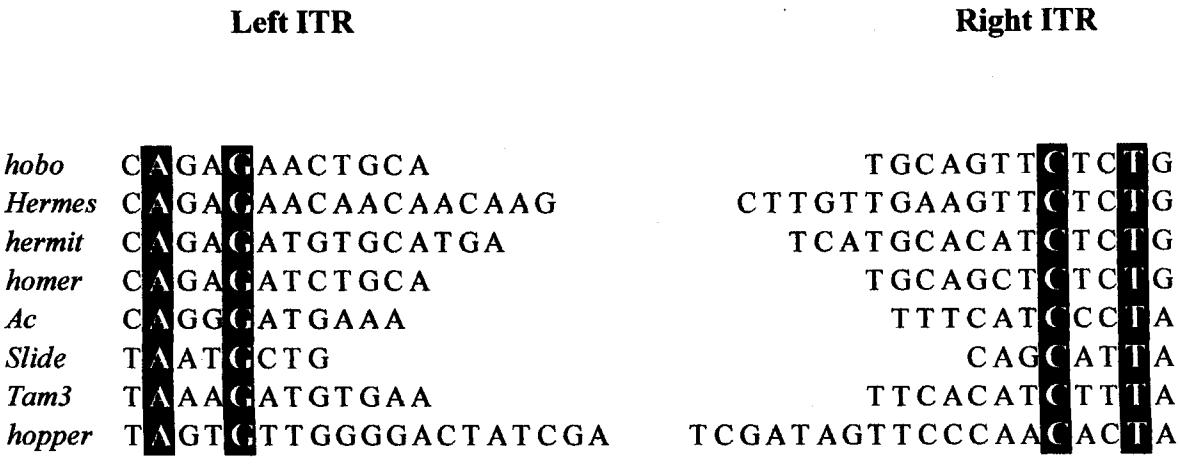

Figure 5.13. Nucleotide sequence of the left and right inverted terminal repeats (ITRs) of member the *hAT* element family and *hopper* element. Bases identical (A2:G5) in all elements are in bold. Data presented here were obtained from the GenBank database as described in figure 5.9.

5.5. Studying the genomic representation of *hopper* element in *Bactrocera tau.*

To determine the genomic representation of the *hopper* element, Southern blot of restriction digests genomic DNA was prepared from *B. tau* strains that were collected from different geographical areas in Thailand and the other *Bactrocera* species. An initial goal is to study the presence and the distribution of *hopper* element in the genome of *B. tau*. The cloned 441 bp fragment of *hopper* transposase amplified by using Hp I and Hp II oligonucleotide primers was labeled with the nonradioactive labeling system and used as DNA probe. This probe was termed TETAU and used for this hybridization. Figure 5.14 shows the results obtained with *B. tau* DNAs, representing strains from natural populations of different areas in Thailand, and the other *Bactrocera* species. Although the hybridization and washing conditions used were highly stringent (0.5x SSC and 60°C wash), *hopper* element are clearly present in all strains tested and the multiple banding pattern shows different genomic location.

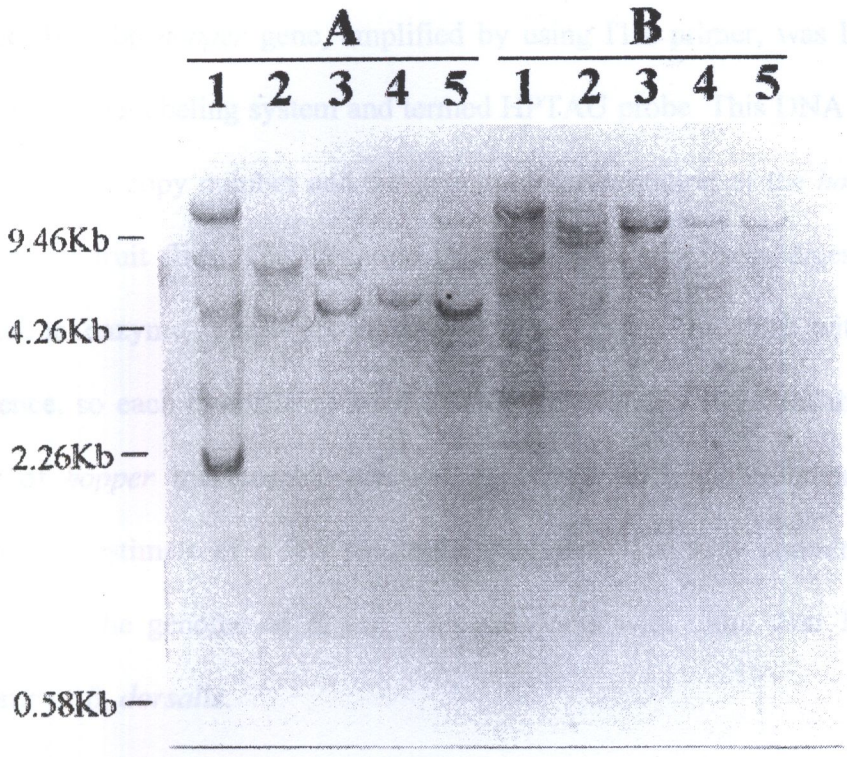

Figure 5.14 Southern blot analysis with TETAU probe. This figure shows the Southern blot hybridization of the *hobo*-related in *B. tau* and the other *Bactrocera* species genomic DNA. In panel A, the genomic DNAs were digested with *Hind* III restriction enzyme and in panel B, the genomic DNAs were digested with *Eco* RI. Lane 1-5 of each panel contain the genomic DNA of *Bactrocera nongensis*, *B. tau* from Ubol ratchatani, *B. tau* from Nan, *B. tau* from Ranong, and *B. tau* from Chumporn province respectively. Size standard markers are shown leftmost. These fragments were compared with the single band of heat shock 70 kD to confirm the complete digestion (data not shown).

To detect the copy number and the genomic representation of this element, the cloned 3172 bp *hopper* gene, amplified by using ITR primer, was labeled with the nonradioactive labeling system and termed HPTAU probe. This DNA probe was used to present the copy number and the genomic representation of the *hopper* element in genome of fruit flies. The Genomic DNAs of fruit flies were digested with *Msp* I restriction enzyme. There are no known *Msp*I restriction sites within the *hopper* sequence, so each hybridizing *Msp*I fragment potentially indicates the presence of a copy of *hopper* transposable element. In figure 5.15, the Southern analysis also permits an estimate of a few restriction fragments that may contain *hopper*-related element in the genome of *B. tau.* This compares with more than 10 copies in the genome of *B. dorsalis*.

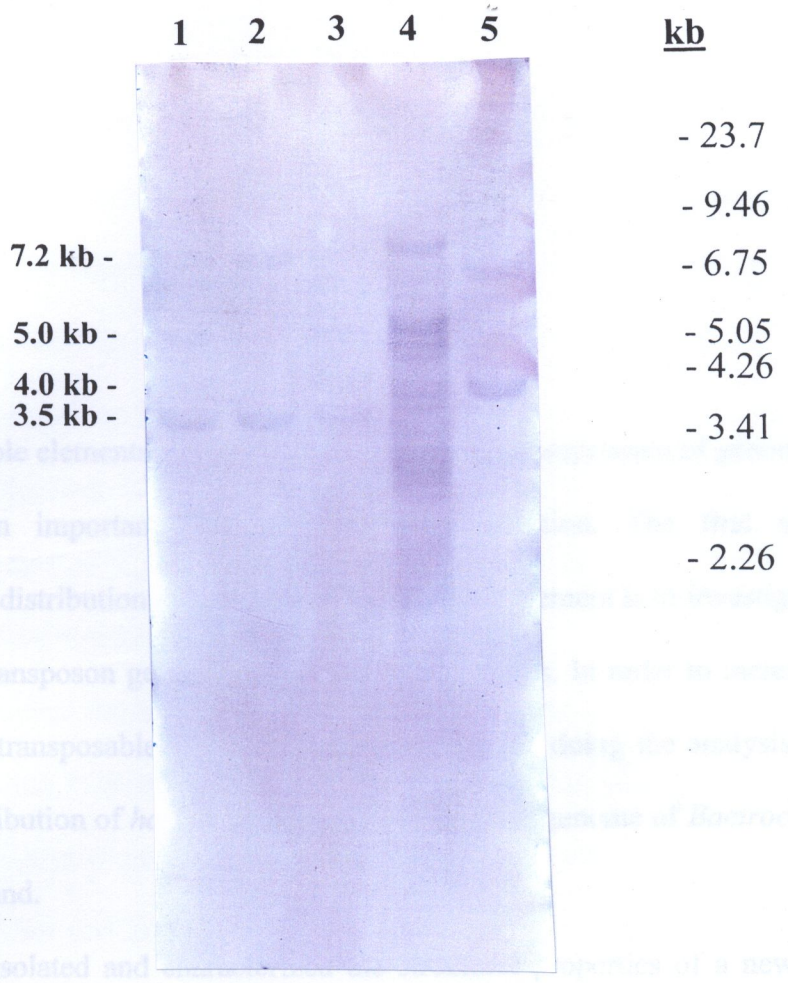

Figure 5.15 Southern blot analysis with HPTAU probe. Genomic representation of *hopper* in *Bactrocera* species. Msp I digested of genomic DNAs from different populations of *B. tau* were used; these were strain trapped from : Ubol Ratchathani (lane number 1); Nan (lane number 2); Ranong (lane number3) provinces and the other *Bactrocera* species: *B. dorsalis* (lane number 4) and *B. cucurbitae* (lane number 5) respectively. Genomic DNAs were probed with nonradioactive labeled HPTAU probe. Lambda DNAs digested with *Hind* III and *Hind* III-*Eco* RI were used as the molecular weight standards and marked at the rightmost.

CHAPTER VI

DISCUSSION

Transposable elements are known to be ubiquitous components of genome, and probably play an important role in eukaryotic evolution. The first step in understanding the distribution of a particular transposable element is to investigate and characterize the transposon gene from genome of organisms. In order to increase our knowledge about transposable element distribution, we are doing the analysis of the structure and distribution of *hopper* transposable element in genome of *Bactrocera tau* (Walker) in Thailand.

We have isolated and characterized the structural properties of a new insect *hAT* transposable element, the *hopper* element of the tephritid, *B. tau* . The *hopper* transposable element is 3172 bp in length and contains 19 bp ITRs with one single mismatch. It possesses a single long ORF that encodes a putative 651 amino acids in length. Based on amino acid comparisons, the *hopper* element is as related to the *hobo*, *Ac*, and *Tam3* elements.

6.1. The putative functional *hopper* transposase protein sequences

Our sequence analysis suggests that there is a putative transposase protein coding gene in 3.1 kb *hopper* transposable element. This transposable element has previously been isolated from *B. dorsalis* (42). Of interest is the finding that the element that was described by Handler and Gomez contains the transposase ORF discontinuity and the lack of 8 bp insertion site direct repeat which is consistently present in all other *hAT* elements. These suggest that they do not possess characteristic of an autonomous functional transposable element.

We have isolated the *hopper* transposable element from genome of *Bactrocera tau* (Walker) by using PCR method. Unlike the previously isolated *hopper* element, this copy contains a new putative-functional transposase gene that consists of an uninterrupted ORF for 651 amino acids beginning with a start codon. The insertion sites of this element have a perfect 8 bp. target site duplication. We found that the short inverted terminal repeat termini is 19 bp. with a single mismatch.

The genomic representation of *hopper* transposable element in genome of the *B. tau* from different geographical areas in Thailand shows differences in banding patterns. This suggests that *hopper* elements are vary in copy number in *B. tau* from different geographical areas. These results show that this *hopper* element from *B. tau* would correspond to a putative functional transposase gene and may infer a recently active transposition within the *B. tau* genome.

6.2. The *hopper* and *hAT* transposable elements

Although some sequence similarity among the ClassII short inverted terminal repeat element *hobo, Ac,* and *Tam3* was first pointed out by Streck et al. (36) the extent of this similarity was too weak to infer significant homology. More recently Calvi et al. (2) have presented compelling evidence for a common evolution origin of *hobo* from *D. melanogaster*, *Ac* from maize, and *Tam3* from snapdragon, on the basis of sequence similarity among their respective transposase genes which are incongruous with the phylogenies of their host species.

When we compare the deduced *hopper* transposase protein sequences to the transposase of *hobo, Ac,* and *Tam3* elements, we found that the sequence alignment was conserved in three regions in the middle and C-terminal regions of the transposase proteins as shown in table 5.3. Out of these regions, the most identical one is ranging between 32-60 %. This protein sequences alignment shows that a putative functional *hopper* transposase has a similar distance with no more than 42 % identity from other transposase protein sequences. These sequence identities appear to be considerably closer than would be expected from strictly vertical transmission and independent divergence from a common ancestral element present at the time of splitting of plant and animal kingdoms. Taken together with these transposase sequences comparisons, they are supporting that the *hopper* element may be an ancient predecessor to other insect *hAT* elements, perhaps being a like or early branch from the non-insect elements.

6.3. Inverse PCR (IPCR)

The IPCR method was developed for amplification of regions lying outside known sequences (46). This method has been used to isolate genes tagged by DNA transposons in plants (47), to trap promoters with retroviruses in cultured mammalian cells and to identify *Tnt1* insertions in transgenic *Arabidopsis* (48). We have employed IPCR here to clone the inverted terminal repeats termini and insertion sites sequences of *hopper* transposable element from genome of *Bactrocera tau*. These sequences may contain important functional regions, notably the priming sites for transposase enzyme, as well as the host sites targeted during transposition.

6.4. *hAT* elements inverted terminal repeats

A comparison of the inverted terminal repeats (ITRs) of *hopper* with those of the other members of the *hAT* family revealed a sequence similarity. These elements, although have inverted repeats of various lengths and sequence compositions, all have an A at position 2 and a G at position 5 of their left termini, and complementary bases at the corresponding positions in their right termini. This observation suggests that these nucleotides play a central role in the biochemistry of transposition in this family of elements. This A2G5 pattern is unlikely to be fortuitous. All these elements share the common feature of generating an 8 bp insertion site duplication. These similarities suggest that these *hAT* elements may share common enzymologies in their transposition mechanisms.

CHAPTER VII

CONCLUSION

1. A gene of *hopper* transposable element in the genome of *Bactrocera tau* (Walker) was isolated by using the PCR approach.
2. This *hopper* element is 3172 bp in length and consists of 19 bp inverted terminal repeats with single mismatch.
3. The insertion sites of *hopper* element have a perfect 8 bp target site duplication.
4. Unlike the previously isolated *hopper* element from *Bactrocera dorsalis* by Handler A.M. and Gomez S.P. in 1997 (42), this copy of *hopper* contains a new putative-functional gene that consists of an uninterrupted ORF for 651 amino acids in length.
5. The comparison of the deduced *hopper* transposase protein sequences to those of *hobo, Ac,* and *Tam3* elements shows that the sequence was conserved in three regions in the middle and C-terminal regions of these elements. Out of these regions, *hopper* transposase has a similar distance with no more than 42% identity from other transposase protein sequences.
6. The genomic representation of the *hopper* transposable element in the genome of the *B. tau* (Walker) from different geographical areas in Thailand shows differences in banding patterns.
7. We suggested that this *hopper* transposable element would correspond to a putative-functional transposase gene and may infer a recently active transposition.

REFERENCES

1. Kidwell MG, Lisch D. Transposable elements as sources of variation in animals and plants. Proc Natl Acad Sci USA 1997;94:7704-11.
2. Calvi BR, Hong TJ, Findley SD, Gelbart WM. Evidence for a common evolutionary origin of inverted terminal repeat transposons in Drosophila and plants: hobo, Activator and Tam3. Cell 1991;66:465-71.
3. Grappin P, Audeon C, Chupeau MC, Grandbastien MA. Molecular and functional characterization of Slide, an Ac-like autonomous transposable element from tobacco. Mol Gen Genet 1996;252:386-97.
4. Warren WD, Atkinson PW, O'Brochta DA. The Australian bushfly Musca vetustissima contains a sequence related to transposons of the hobo, Ac, and Tam3 family. Gene 1995;154:133-4.
5. Warren WD, Atkinson PW, O'Brochta DA. The Hermes transposable element from the house fly, Musca domestica, is a short inverted repeat-type element of the hobo, Ac, and Tam3 (hAT) element family. Genet Res 1994;64:87-97.
6. Handler AM, Gomez SP. A new hobo, Ac, Tam3 transposable element, hopper, from Bactrocera dorsalis is distantly related to hobo and Ac. Gene 1997;185:133-5.
7. Loreto EL, da Silva LB, Zaha A, Valente VL. Distribution of transposable elements in neotropical species of Drosophila. Genetica 1998;101:153-65.

8. Meksongsee B, Liewvanich A, and Jirasuratana M. Fruit flies in Thailand. Proceedings of the First International Symposium Fruit Flies in the Tropics; 1988 March 14-16; Kualalumpur. Malaysian Agriculture Research and Development Institute; 1991. p. 83-98.

9. White IM, and Elson-Haris MM. Fruit flies of economic significance: Their identification and binomics. 1st ed. Melksham: Red wood Press Ltd.; 1992.

10. Shapiro JA. Mobile genetic elements. 1st ed. New York: Academic press; 1983.

11. Lai C. Genetic applications of transposable elements in eukaryotes. Genome 1994;37:519-525.

12. Labrador M, Corces VG. Transposable element-host interaction: regulation of insertion and excision. Annu Rev Gen 1997;31:381-404.

13. Reznikoff WS. The Tn5 transposon. Annu Rev Microbiol 1993;47:945-63.

14. Mizuuchi K. Transpositional recombination: mechanistic insights from studies of Mu and other elements. Annu Rev Biochem 1992;61:1011-51.

15. Lavoie BD, Chaconas G. Transposition of phage Mu DNA. Curr Top Microbiol Immunol 1996;204:83-102.

16. Lewin B. GeneV. New York: Oxford University Press; 1994.

17. Craig NL. Unity in transposition reactions. Science 1995;270:253-4.

18. May EW and Craig NL. Switching from cut-and-paste to replicative Tn7 transposition. Science 1996;272:401-4.

19. Craig NL.Target site selection in transposition. Annu Rev Biochem 1997;66:437-74.

20. Boeke JD and Corces V. Transcription and reverse transcription of retrotransposons. Annu Rev Microbiol 1989;43:403-34.

21. Haren L, Ton-Hoang B, Chandler M. Integrating DNA: transposases and retroviral integrases. Annu Rev Microbiol 1999;53:245-81.

22. Doak TG, Doerder FP, Jahn CG, Herick AG. A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. Proc Natl Acad Sci 1994;91:942-46.

23. Plasterk RH.The Tc1/mariner transposon family. Curr Top Microbiol Immunol 1996;204:125-43.

24. Beall EL, Rio DC. Transposase makes critical contacts with, and is stimulated by, single-strand DNA at the P element termini in vitro. EMBO J 1998;17:2122-36.

25. Becker HA, Kunze R. Maize Activator transposase has a bipartite DNA bindingdomain that recognizes subterminal sequences and the terminal inverted repeats. Mol Gen Genet 1997;254:219-30.

26. Quesneville H, Anxolabehere D. A simulation of P element horizontal transfer in Drosophila. Genetica 1997;100:295-307.

27. Finnegan DJ. Eukaryotic transposable elements and genome evolution. Trends Genet 1989;5:103-7.

28. Boeke JD, Garfinkel DJ, Styles CA, Fink GR. Ty elements transpose through an RNA intermediate. Cell 1985;40:491-500.

29. Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcripase sequences. EMBO J 1990;9:3353-62.

30. Wang H, Hartswood E, Finnegan DJ. Pogo transposase contains a putative helix-turn-helix DNA binding domain that recognises a 12 bp sequence within the terminal inverted repeats. Nucleic Acids Res 1999;27:455-61.

31. van Luenen HG, Colloms SD, Plasterk RH. The mechanism of transposition of Tc3 in C. elegans. Cell 1994;79:293-301.

32. Kaufman PD, Rio DC. P element transposition in vitro proceeds by a cut and paste mechanism and use GTP as a cofactor. Cell 1992;69:27-39 .

33. Lampe DJ, Churchill ME, Robertson HM. A purified mariner transposase is sufficient to mediate transposition in vitro. EMBO J 1996;15:5470-79.

34. Rhodes PR, Vodkin LO. Organization of the *Tgm* family of transposable elements in soybean. Genetics 1988;120:597-604.

35. Pinkerton AC, Whyard S, Mende HA, Coates CJ, O'Brochta DA, Atkinson PW. The Queensland fruit fly, Bactrocera tryoni, contains multiple members of the hAT family of transposable elements. Insect Mol Biol 1999;8:423-34.

36. Streck RD, MacGaffey JE, Beckendorf SK. The structure of hobo transposable elements and their insertion. EMBO J 1986;5:3615-23.

37. Kunze R, Starlinger P. The putative transposase of transposable element Ac from Zea mays L. interacts with subterminal sequences of Ac. EMBO J 1987;8:3177-85.

38. Hehl R, Nacken WK, Krause A, Saedler H, Sommer H. Structure analysis of Tam3, a transposable element from Antirrhinum majus, reveals homologies to the Ac element from maize. Plant Mol Biol 1991;16:369-71.

39. Sommer H, Hehl R, Harrison BJ, Saedler H. The transposable element Tam3 of Antirrhinum majus generates a novel type of sequence alteration upon excision. Mol Gen Genet 1985;199:225-31.

40. Feldmar S and Kunze R. The ORFa protein, the putative transposase of maize transposable element Ac, has a basic DNA binding domain. EMBO J 1991;10:4003-10.

41. O'Brochta DA, Warren WD, Saville KJ, Atkinson PW. Hermes, a functional non-Drosophilid insect gene vector from Musca domestica. Genetics 1996;142:907-14.

42. Handler AM, Gomez SP. A new hobo, Ac, Tam3 transposable element, hopper, from *Bactrocera dorsalis* is distantly related to hobo and Ac. Gene 1997;185:133-5.

43. Handler AM, Gomez SP. The *hobo* transposable element excised and has related elements in tephritid species. Genetics 1996;143:1339-1347.

44. Thanaphum S, Haymer DS. A member of the hsp70 gene family from the Mediterranean fruit fly, Ceratitis capitata. Insect Mol Biol 1998;7:63-72.

45. Kaufman PK, Doll RF, Rio DC. *Drosophila* P element transposase recognizes internal P element DNA sequences. Cell 1989;59:359-71.

46. Ochman H, Gerber AS, Hartl DL. Genetic applications of an inverse polymerase chain reaction. Genetics 1988;120:621-3.

47. Souer E, Quattrocchio F, deVetten N, Mol J, Koes R. A general method to isolate genes tagged by a high copy number transposable element. Plant J 1995;7:677-85.

48. Lucas H, Feuerbach F, Kunert K, Grandbastien MA, Caboche M. RNA-mediated transposition of the tobacco retrotransposon Tnt1 in Arabidopsis thaliana. EMBO J 1995;14:2364-73.

BIOGRAPHY

NAME	Mr. Chatchawal Sanguansilp
DATE OF BIRTH	26 November 1974
PLACE OF BIRTH	Lopburi, Thailand
INSTITUTIONS ATTENDED	**Chulalongkorn University, 1992-1995:** **Bachelor of Science (Medical Technology)** **Mahidol University, 1996-2000:** Master of Science (Environmental Biology)