

## ABSTRACT

Allelic variation of the immunodominant T-cell epitopes, Th2R and Th3R on *Plasmodium falciparum* circumsporozoite protein (CSP) gene has been determined by sequencing in 40 malaria isolates from malaria endemic areas, west of Thailand. These malaria isolates contained the total of nine and seven allelic types of variants on Th2R and Th3R regions, respectively. Comparison of T-cell epitope sequences on Th2R and Th3R region of CSP with published sequences from endemic areas of different geographical locations, five new allelic types of Th2R regions (CSP-Th2R\*16, CSP-Th2R\*17, CSP-Th2R\*18, CSP-Th2R\*20, and CSP-Th2R\*21) and three new allelic types of Th3R regions (CSP-Th3R\*14, CSP-Th3R\*15, and CSP-Th3R\*16) were demonstrated. These new allelic types were not cluster together with those found previously in Thailand. The results also revealed that the nucleotide substitutions occurred predominantly in the first codon position and less frequently in the second codon position resulting in the amino-acid changes, which were non-synonymous mutation. Moreover, the genetic diversity in such regions were also analyzed using the combination of Polymerase Chain Reaction (PCR) and sequence specific oligonucleotide probes (PCR-SSOP) in 144 and 43 falciparum malaria isolates from falciparum malaria endemic areas, Ratchaburi and Kanchanaburi provinces along Thai-Myanmar border, respectively. Three (CSP-Th2R\*05, CSP-Th2R\*12 and CSP-Th2R\*13) and five (CSP-Th2R\*05, CSP-Th2R\*12, CSP-Th2R\*13, CSP-Th2R\*14 and CSP-Th2R\*15) allelic variants from 15 reported allelic types of CSP-Th2R region were identified in Ratchaburi and Kanchanaburi provinces, respectively. For CSP-Th3R regions, four allelic types were found in Ratchaburi province (CSP-Th3R\*01, CSP-Th3R\*04, CSP-Th3R\*08 and CSP-Th3R\*09) while three allelic types (CSP-Th3R\*01, CSP-Th3R\*04 and CSP-Th3R\*10) were found in Kanchanaburi provinces. Interestingly, the patterns of hybridization of some isolates in both provinces did not correspond to any known alleles, suggesting new allelic types in CSP gene. The similar finding in both provinces was CSP-Th2R\*05 and CSP-Th3R\*01 which were the majority of CSP-Th2R and CSP-Th3R allelic types. In addition, the CSP-Th2R\*05 and CSP-Th3R\*01 were the most common allelic types in both all age groups and numbers of falciparum malaria exposure. Comparison of the CSP-Th2R and CSP-

Th3R allelic types from dot blot hybridization with DNA sequencing showed that most isolates gave similar results of CSP-Th2R and CSP-Th3R allelic types. These results revealed that the genetic diversity of CSP varied among *P. falciparum* isolates from different geographical areas. In the present study, the genotyping of *P. falciparum* isolates were also determined in the repeat regions of CSP gene and the three highly polymorphic regions encoding for merozoite surface protein-1 (MSP-1), -2 (MSP-2), and glutamine-rich protein (GLURP) by nested PCR. The results of amplification of the repeat regions in CSP gene showed three allelic sizes of variants (600, 700, and 900 bp). While the genetic diversity in MSP-1, MSP-2 and GLURP were large, with 12, 15 and 10 sizes of variants detected, respectively. The MSP-2 locus was usually highly polymorphic and FC27 and 3D7/IC families were both highly represented. These situations were similar to those found in Senegal and Trad province of Thailand. These results together with other similar studies are essential for further study to determine whether the CSP antigen polymorphism represents in fact a major obstacle for the development of anti-sporozoite vaccine against malaria parasite.

## บทคัดย่อ

การศึกษาความหลากหลายของล่าดับเบลสบันที่เซลล์อิพิทอปในส่วนที่เป็น CSP-Th2R และ CSP-Th3R ของโปรดีนแอนติเจนชนิด circumsporozoite ของเชื้อมาลาเรียชนิดพื้นเมืองพารันที่เก็บได้จากผู้ป่วยในจังหวัดราชบุรีและกาญจนบุรีซึ่งอยู่บริเวณภาคตะวันตกของประเทศไทยที่มีชายแดนติดต่อกับประเทศไทยและเป็นแหล่งที่มีการระบาดของไข้มาลาเรียค่อนข้างสูง จำนวน 40 ราย ด้วยเทคนิค DNA sequencing พบว่ามีความหลากหลายของล่าดับเบลสบันที่เซลล์อิพิทอปในส่วนที่เป็น CSP-Th2R จำนวน 9 ชนิด (CSP-Th2R\*05, CSP-Th2R\*12, CSP-Th2R\*13, CSP-Th2R\*16, CSP-Th2R\*17, CSP-Th2R\*18, CSP-Th2R\*19, CSP-Th2R\*20, CSP-Th2R\*21) ซึ่งห้าชนิดความหลากหลายได้แก่ CSP-Th2R\*16, CSP-Th2R\*17, CSP-Th2R\*18, CSP-Th2R\*20, และ CSP-Th2R\*21 เป็นความหลากหลายของล่าดับเบลสบันนิคใหม่เมื่อเปรียบเทียบกับความหลากหลายของล่าดับเบลสบันนิคต่างๆที่พบในแหล่งพันธุกรรมติดเชื้อมาลาเรียค่อนข้างสูงในภูมิภาคต่างๆในโลก สำหรับความหลากหลายของล่าดับเบลสบันที่เซลล์อิพิทอปในส่วนที่เป็น CSP-Th3R พบว่ามีความหลากหลายของล่าดับเบลสเป็นจำนวน 7 ชนิด (CSP-Th3R\*01, CSP-Th3R\*04, CSP-Th3R\*08, CSP-Th3R\*09, CSP-Th3R\*14, CSP-Th3R\*15, CSP-Th3R\*16) โดย 3 ชนิด (CSP-Th3R\*14, CSP-Th3R\*15, CSP-Th3R\*16) เป็นความหลากหลายของล่าดับเบลสบันนิคใหม่ เช่นเดียวกัน นอกจากนี้ยังพบว่าความหลากหลายของล่าดับเบลสเกิดจากการเปลี่ยนแปลงของแบบトイด์เฉพาะตำแหน่งแรกซึ่งอาจพบบ้างในตำแหน่งที่สองและบนน้อย ในตำแหน่งที่สาม トイด์การเปลี่ยนแปลงดังกล่าวมีผลทำให้เกิดการเปลี่ยนแปลงของชนิดโปรดีน ซึ่งเรียกการเปลี่ยนแปลงนี้ว่าเป็นชนิดแบบ non-synonymous นอกจากนี้ยังทำการวิเคราะห์ความหลากหลายด้วยเทคนิค PCR และ dot blot hybridization กับ DNA probe ที่มีความจำเพาะ ในการผู้ป่วยติดเชื้อจากจังหวัดราชบุรีและกาญจนบุรี จำนวน 144 และ 43 รายตามล่าดับ ผลจากการศึกษาพบว่ามีความหลากหลายในส่วนที่เป็น CSP-Th2R จำนวน 3 ชนิด (CSP-Th2R\*05, CSP-Th2R\*12, CSP-Th2R\*13) และ 5 ชนิด (CSP-Th2R\*05, CSP-Th2R\*12, CSP-Th2R\*13, CSP-Th2R\*14, CSP-Th2R\*15) สำหรับส่วนที่เป็น CSP-Th3R พบว่ามีจำนวน 4 ชนิด (CSP-Th3R\*01, CSP-Th3R\*04, CSP-Th3R\*08, CSP-Th3R\*09) และ 3 ชนิด (CSP-Th3R\*01, CSP-Th3R\*04, CSP-Th2R\*10) ในจังหวัดราชบุรีและกาญจนบุรีตามล่าดับ นอกจากนี้ยังพบว่ามีเชื้อมาลาเรียบางสายพันธุ์ที่ไม่สามารถจำแนกชนิดได้ ซึ่งอาจจะเป็นความหลากหลายชนิดใหม่ เมื่อเปรียบเทียบผลการศึกษาโดยวิธี dot blot hybridization

และ DNA sequencing พบว่าทั้งสองวิธีให้ผลของลำดับเบนทับที่เชอล็อกอิพิโทปในส่วนที่เป็น CSP-Th2R และ CSP-Th3R เหมือนกันเมื่อศึกษา ในเชื้อมากาเรียของผู้ป่วยคนเดียวกัน นอกจากนี้ยังได้ทำการศึกษาว่าเชื้อมากาเรียชนิดพิลชิพารัมที่มีของลำดับเบนทับที่เชอล็อกอิพิโทปในส่วนที่เป็น CSP-Th2R และ CSP-Th3R ชนิดเดียวกันจะเป็นเชื้อมากาเรียสายพันธุ์เดียวกันหรือไม่ โดยทำการวิเคราะห์แยกชนิดของเชื้อ MSP-1, MSP-2, GLURP และ repeat region ของโปรตีนแอนติเจนชนิด circumsporozoite ด้วยวิธี nested PCR พบว่าเชื้อมากาเรียดังกล่าวมีรูปแบบของเชื้อ MSP-1, MSP-2, GLURP และ repeat region แตกต่างกันในเชื้อที่มีลำดับเบนทับที่เชอล็อกอิพิโทปในส่วนที่เป็น CSP-Th2R และ CSP-Th3R ชนิดเดียวกัน ซึ่งบ่งบอกว่าการติดเชื้ออาจเกิดจากเชื้อสายพันธุ์ จากผลการศึกษาทั้งหมดสามารถสรุปได้ว่าความหลากหลายของลำดับเบนทับขึ้นอยู่กับความแตกต่างของแหล่งที่ตั้งของพื้นที่ที่มีการระบายน้ำของเชื้อมากาเรียชนิดดังกล่าว เมื่อจากผลการศึกษาดังกล่าวพบว่าความหลากหลายของลำดับเบนทับที่เชอล็อกอิพิโทปในส่วนที่เป็น CSP-Th2R และ CSP-Th3R ของโปรตีนแอนติเจนชนิด circumsporozoite เหล่านี้มีผลต่อการตอบสนองทางภูมิคุ้มกันต่อเชื้อมากาเรียชนิดพิลชิพารัม ขณะนี้ความหลากหลายของลำดับเบนทับเหล่านี้จะมีผลกระทบต่อการผลิตและพัฒนาวัคซีนที่ใช้ป้องกัน โรคมาเรียชนิดพิลชิพารัมหากการศึกษาในลำดับต่อไปพบว่าความหลากหลายของลำดับเบนทับดังเหล่านี้มีผลต่อการกระตุ้นภูมิคุ้มกันต่อเชื้อมากาเรีย

## EXECUTIVE SUMMARY

The genetic diversity displayed by *P. falciparum* field isolates, the occurrence of variant forms of the parasite at different frequencies in different geographic areas, and the complexity of the infections represent major obstacles for the development of effective malaria control measures such as the development of a vaccine against falciparum malaria. In the present study, we proposed to determine the genetic variation in T-helper cell epitopes (Th2R and Th3R) of *P. falciparum* CSP gene in parasite isolates obtained from malaria endemic areas with high rate of malaria transmission in Thailand using DNA sequencing and Polymerase Chain Reaction (PCR) and sequence specific oligonucleotide probes (PCR-SSOP). Thirty-one, four, four and one *P. falciparum* isolates obtained from Ratchaburi, Kanchanburi, Tak and Phetchaburi provinces, malaria endemic areas of Thailand were sequenced. These isolates contained the total of nine and seven allelic types of variants on Th2R and Th3R regions, respectively. Comparison of these sequences of T-cell epitopes on Th2R and Th3R region of CSP with published sequences from endemic areas of different geographical locations, five new allelic types of Th2R regions (CSP-Th2R\*16, CSP-Th2R\*17, CSP-Th2R\*18, CSP-Th2R\*20, and CSP-Th2R\*21) and three new allelic types of Th3R regions (CSP-Th3R\*14, CSP-Th3R\*15, and CSP-Th3R\*16) were demonstrated. These new allelic types were not clustered together with those found previously in Thailand. However, the predominant allelic types of CSP-Th2R and CSP-Th3R regions were CSP-Th2R\*05 and CSP-Th3R\*01. The results revealed that the nucleotide substitutions occurred predominantly in the first codon position and less frequently in the second codon position resulting in the amino-acid changes, which were non-synonymous mutation. The genetic diversity was also analyzed using the combination of Polymerase Chain Reaction (PCR) and sequence specific oligonucleotide probes (PCR-SSOP) in 144 and 43 falciparum malaria isolates from falciparum malaria endemic areas, Ratchaburi and Kanchanaburi provinces along Thai-Myanmar border. Three (CSP-Th2R\*05, CSP-Th2R\*12 and CSP-Th2R\*13) and five (CSP-Th2R\*05, CSP-Th2R\*12, CSP-Th2R\*13, CSP-Th2R\*14 and CSP-Th2R\*15) allelic variants from 15 reported allelic types of CSP-Th2R region were identified in Ratchaburi and Kanchanaburi provinces, respectively. For CSP-Th3R regions, four (CSP-Th3R\*01, CSP-Th3R\*04, CSP-Th3R\*08 and CSP-Th3R\*09) and

three allelic types (CSP-Th3R\*01, CSP-Th3R\* CSP-Th3R\*04 and CSP-Th3R\*10) were found in Ratchaburi and Kanchanaburi provinces, respectively. Furthermore, the patterns of hybridization of some isolates in both provinces did not correspond to any known alleles, suggesting new allelic types in CSP gene. The similar finding in both provinces was CSP-Th2R\*05 and CSP-Th3R\*01 which were the majority of CSP-Th2R and CSP-Th3R allelic types. Furthermore, the CSP-Th2R\*05 and CSP-Th3R\*01 were the most common allelic types in both all age groups and they did not change with the number of falciparum malaria attacks. Comparison of the CSP-Th2R and CSP-Th3R allelic types from dot blot hybridization with DNA sequencing had shown that most isolates gave similar results. These results revealed that geographic variation in genetic diversity of CSP. In addition, genotyping of *P. falciparum* isolates with the same Th types were also determined for the repeat region of the CSP gene and for three highly polymorphic regions of genes encoding for merozoite surface protein-1 (MSP-1), -2 (MSP-2), and glutamine-rich protein (GLURP) by nested PCR. The results of amplification of the repeat regions in CSP gene showed six allelic variants of different sizes while the MSP-1, MSP-2 and GLURP loci showed high polymorphism with 12, 15 and 10 sizes of variants detected, respectively. The MSP-2 locus was usually highly polymorphic and FC27 and 3D7/IC families were both highly represented. However, the RO33 family of MSP-1 was poorly polymorphic, with only one allele detected. These situations were similar to that of Senegal and Trad province of Thailand. These results and other similar studies are essential to determine whether the CSP antigen polymorphism represents in fact a major obstacle for the development of anti-sporozoite vaccine against the malaria parasite.