

ANALYSIS OF THE POLYMORPHIC *SER* GENE FAMILY IN *TETRAHYMENA THERMOPHILA*

PATRATH PONSUWANNA 5037194 SCBC / D

Ph.D.(BIOCHEMISTRY)

THESIS ADVISORY COMMITTEE: THANAT CHOOKAJORN, Ph.D., SUMALEE TUNGPRADABKUL, Ph.D., WILAI NOONPAKDEE, Ph.D.

ABSTRACT

Tetrahymena thermophila, a free-living ciliate, displays immobilization antigens (i-ag) on its surface. These proteins are cysteine-rich and are linked to the membrane by a glycosylphosphatidylinositol (GPI) anchor. They are encoded by a family of polymorphic *Ser* genes. Thirteen *Ser* genes were characterized so far, representing only a portion of the reported i-ag. Characterization of *Ser* gene family is impeded by its sequence polymorphism. In this study an algorithm to select *Ser* candidate from *T. thermophila* MAC genome was developed. The *Ser* prediction algorithm exploited two characteristic features of *Ser*—repetitive cysteine pattern and GPI anchor signal. The algorithm successfully selected *Ser* candidates including known *Ser*. *Ser* candidates were compared to known *Ser* and classified into subtypes by phylogenetic analysis. *Ser* candidates were found to be located as gene tandem array, suggesting that *Ser* gene family expands via gene duplication. *Ser* genes located on the same tandem array may have similar gene expression profile. The *Ser* prediction algorithm can select *Ser* candidates that did not belong to previously characterized subtype. Because the algorithm does not rely on sequence similarity, it can perform *Ser* identification more extensively than the sequence homology approach.

KEY WORDS: IMMOBILIZATION ANTIGEN / *TETRAHYMENA THERMOPHILA*
/ POLYMORPHIC GENE FAMILY

133 pages

การวิเคราะห์กลุ่มยีน *SER* ซึ่งมีความหลากหลายทางพันธุกรรมใน *TETRAHYMENA THERMOPHILA*

ANALYSIS OF THE POLYMORPHIC *SER* GENE FAMILY IN *TETRAHYMENA THERMOPHILA*

พัทธรัต พลสุวรรณ 5037194 SCBC / D

ปร.ค. (ชีวเคมี)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์: ชนรรต สุขจร, Ph.D., สุมาลี ตั้งประดับกุล, Ph.D., วิไล หนูนภักดี, Ph.D.

บทคัดย่อ

Tetrahymena thermophila เป็นสิ่งมีชีวิตเซลล์เดียวประเภทดำรงชีวิตอิสระ (free-living) ชนิดหนึ่งซึ่งแสดงโปรตีน immobilization antigen (i-ag) บนผิวของมัน โปรตีนที่ผิวเหล่านี้มีกรดอะมิโน cysteine เป็นส่วนประกอบมากและยึดโยงอยู่กับเยื่อหุ้มเซลล์ด้วย glycosylphosphatidyl-inositol (GPI) โปรตีนเหล่านี้ถูก encode โดยกลุ่มยีน *Ser* ซึ่งเป็นกลุ่มยีนที่มีความหลากหลายทางพันธุกรรม จวบปัจจุบันนี้มีการศึกษาและอธิบายลำดับนิวคลีโอไทด์ (characterize) ของยีน *Ser* ไว้ทั้งสิ้น 13 ยีนซึ่งนับเป็นเพียงส่วนหนึ่งของ i-ag ทั้งหมดที่เคยมีการรายงานไว้ การศึกษาลักษณะยีน *Ser* ประสบปัญหาจากความหลากหลายทางพันธุกรรมในลำดับนิวคลีโอไทด์และกรดอะมิโน การศึกษานี้ได้พัฒนาอัลกอริทึมเพื่อค้นหาและคัดเลือกยีน *Ser* จากมาโครนิวเคลียร์จีโนมของ *T. thermo-philus* อัลกอริทึมสำหรับทำนายยีน *Ser* อาศัยลักษณะเด่นสองประการของโปรตีนที่ยีนนี้ encode ได้แก่ลำดับกรดอะมิโนที่ประกอบด้วย cysteine ซ้ำ (repetitive) และ GPI anchor signal อัลกอริทึมนี้สามารถคัดยีนกลุ่มหนึ่งเป็นตัวเลือกยีน *Ser* ซึ่งในกลุ่มนี้มียีน *Ser* ที่เคยมีการ characterize เอาไว้รวมอยู่ด้วย ตัวเลือกยีน *Ser* ถูกนำมาศึกษาเปรียบเทียบและจัดหมวดหมู่โดยการวิเคราะห์ด้วย phylogenetic การศึกษานี้พบยีน *Ser* ซึ่งไม่ถูกจัดอยู่ในกลุ่ม (subtype) เดียวกับยีน *Ser* ซึ่งเคยมีการ characterize เอาไว้และพบว่ายีนในกลุ่มตัวเลือกยีน *Ser* อยู่เป็นกลุ่มของยีนบนโครโมโซมในลักษณะเป็น tandem array ซึ่งบ่งชี้ว่ากลุ่มยีน *Ser* อาจขยายขนาดกลุ่มโดยการเพิ่มชุดของยีน (gene duplication) ยีน *Ser* ใน tandem array เดียวกันมักมีรูปแบบของการแสดงออกคล้ายคลึงกัน อัลกอริทึมนี้สามารถค้นหาและระบุยีน *Ser* ได้อย่างครอบคลุมกว่าวิธีการค้นหาโดยอาศัยความคล้ายคลึงของลำดับนิวคลีโอไทด์ หรือกรดอะมิโน (sequence homology)