สัญญาเลขที่ MRG6080358

รายงานวิจัยฉบับสมบูรณ์

โครงการการศึกษาคุณสมบัติและหน้าที่ของยีน *EgDofs* และ *EgNFY* transcription factor ซึ่งควบคุมการสังเคราะห์กรดไขมันในปาล์มน้ำมัน

ดร. อลิษา หนักแก้ว
สังกัดคณะวิทยาศาสตร์ ม.สงขลานครินทร์

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัยและต้นสังกัด (ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว.และต้นสังกัดไม่จำเป็นต้องเห็นด้วยเสมอไป)

สัญญาเลขที่ MRG6080358

รายงานวิจัยฉบับสมบูรณ์

การศึกษาคุณสมบัติและหน้าที่ของยีน *EgDofs* และ *EgNFY* transcription factor ซึ่งควบคุมการสังเคราะห์กรดไขมันในปาล์มน้ำมัน

ดร. อลิษา หนักแก้ว ผู้วิจัย
สังกัดคณะวิทยาศาสตร์ ม.สงขลานครินทร์

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัยและต้นสังกัด (ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว.และต้นสังกัดไม่จำเป็นต้องเห็นด้วยเสมอไป)

Abstract

Project Code : MRG6080358

Project Title : Characterization and functional analysis of the EgNF-Y transcription factor family genes, regulates fatty acid synthesis in *Elaeis guineensis* Jacq

Investigator : Dr.Alisa Nakkaew Faculty of Science Prince of Songkla University

E-mail Address : joy_alisa@yahoo.com,alisa.n@psu.ac.th

Project Period : April 2017- April 2019

Keywords : *Elaeis guineensis*, transcription factor, fatty acid synthesis, EgDofs, EgNF-Ys

Oil palm is an economics crop with its products currently in high demand and necessary to quickly disseminate these elite genotypes by using marker-aided selection genes that link to the yield. In our previous work, the genes encoding the heteromeric acetyl-CoA carboxylase (ht-ACCase), *EgaccA, EgaccB* and *EgaccC*, *EgaccD* were found to involve in regulation of the production of fatty acid and enhancing oil production. The promoter of *EgaccD* was sequenced and analyzed. Transcriptional activators have been shown to play important for regulation of fatty biosynthesis and oil accumulation in seed or fruit of plant oil. In this study, we identified, annotated, and further characterized of the transcription factor genes, NF-Ys and Dofs family, in oil palm. Then, *EgNF-YA, EgNF-YB1, EgNF-YB3, EgNF-YC*, *EgDof300, EgDof314 EgDof318, EgDof343* and *EgDof347* were cloned from mesocarp and leave oil palm tissue. The bioinformatic analysis of EgNF-Ys indicated it encodes NF-YA/B/C subunits and revealed a conserved domain of the NF-YA/B/C subfamilies. Phylogenetic tree analyses showed closely to date palm NF-YA/B/C subunits and the other plant HAP2/3/5 subunits. While as EgDofs were classified to Dofs protein family and closely to date palm and the other DNA-binding one zinc finger transcription factor family in plants and found zinc finger Dof-type signature and conserve domains of zinc finger Dof-type profile. Gene expression patterns showed that most EgNF-Ys were widely expressed in different tissues while as EgDofs was found have expression in a particular tissues. Moreover, we found *EgNF-YB1* have high expression levels in high accumulated oil which different from *EgDof318* that found decreased of gene expression in higher oil accumulation tissues. Moreover, this research found that *EgNF-YB1* expression level were shown high in high oil yield cultivar and low expression level of *EgDof318* in low oil yield

plants. *In silico* EgNF-Ys -DNA binding results found that EgNF-YB1 and EgDof318 have highest DNA binding domains in EgNF-Ys and EgDofs family. Thus, we were chosen EgNF-YB1 for DNA-Protein interaction enzyme linked immunosorbent assay (DPI-ELISA). The result showed the binding of EgNF-YB1 protein and *EgaccD* promoter was strongly interaction than the controls. Based on this results, enhance the expression of the *EgNF-YB1* gene possibly to be one of the methods to increase the oil content and *EgNF-YB1* and *EgDof318* uses as a marker to assist selection of possible high oil yielding cultivars in oil palm breeding programs.

บทคัดย่อ

รหัสโครงการ : MRG6080358

ชื่อโครงการ :การศึกษาคุณสมบัติและหน้าที่ของยีน *EgDofs* และ *EgNFY* transcription factor ซึ่งควบคุมการสังเคราะห์กรดไขมันในปาล์มน้ำมัน

ชื่อนักวิจัย : ดร.อลิษา หนักแก้ว คณะวิทยาศาสตร์ ม. สงขลานครินทร์

E-mail Address : joy_alisa@yahoo.com, alisa.n@psu.ac.th

ระยะเวลาโครงการ : เมษายน 2560-เมษายน 2562

คำหลัก: ปาล์มน้ำมัน, ทรานสคริปชั่นแฟคเตอร์, การสังเคราะห์กรดไขมัน, EgDofs, EgNF-Ys

ปาล์มน้ำมันเป็นพืชเศรษฐกิจที่มีความต้องการใช้น้ำมันในปริมาณที่เพิ่มสูงขึ้นอย่างสูง จึงมีความจำเป็นอย่างยิ่งที่จะต้องทำการคัดเลือกพันธุ์ปาล์มน้ำมันที่มีลักษณะทางพันธุกรรมที่ดี โดยใช้เครื่องหมายโมเลกุลที่มีความเชื่อมโยงกับยีนที่มีความเชื่อมโยงกับปริมาณผลผลิต โดยในการศึกษาก่อนหน้านี้ทางผู้วิจัยได้พบว่ากลุ่มยีนในหน่วยย่อยของเอนไซม์ heteromeric acetyl-CoA carboxylase ได้แก่ยีน *EgaccA, EgaccB, EgaccC* และ *EgaccD* นั้นมีส่วนในการควบคุมในการสร้างกรดไขมันและส่งผลให้มีการเพิ่มน้ำมัน นอกจากนั้นลำดับดีเอ็นเอของโปรโมเตอร์ของยีน *EgaccD* ได้ถูกวิเคราะห์ลำดับดีเอ็นเอและพบว่ามีตำแหน่งอนุรักษ์ของทรานสคริบชันแฟกเตอร์ที่สำคัญซึ่งแสดงให้ทราบว่ามีความสำคัญในการควบคุมวิถีการสังเคราะห์เป็นกรดไขมันและการสะสมน้ำมันในเมล็ดหรือผลของพืชให้น้ำมัน ในการศึกษาครั้งนี้ผู้วิจัยได้ทำการคัดเลือกโคลนและวิเคราะห์ข้อมูลของกลุ่มยีนแฟมิลี่ชนิด *NF-Y* และ *Dofs* ซึ่งผลการจำแนกระบุลักษณะสมบัติความจำเพาะของยีนแฟมิลี่ *NF-Y* ได้แก่ *EgNF-YA, EgNF-YB1, EgNF-YB3* และ *EgNF-YC* จากปาล์มน้ำมันด้วยการวิเคราะห์ความสัมพันธ์ทางพันธุกรรมพบว่าจัดอยู่กลุ่มยีนแฟมิลี่ชนิด NF-YA/B/C การจำแนกระบุลักษณะสมบัติความจำเพาะของยีนแฟมิลี่ *Dofs* ได้แก่ยีน *EgDof300, EgDof314 EgDof318, EgDof343* และ *EgDof347* จากเนื้อเยื่อใบและผลอ่อนปาล์มน้ำมัน เมื่อวิเคราะห์ความสัมพันธ์ทางพันธุกรรมพบว่าเป็นกลุ่มยีนแฟมิลี่ชนิด DNA-binding One Zinc Finger family พบทั้งตำแหน่งอนุรักษ์ของ Zinc finger Dof-type signature และ Zinc finger Dof-type profile โดยเมื่อศึกษาการแสดงออกของกลุ่มยีนแฟมิลี่ชนิด *EgNF-Ys* และ *EgDofs* พบว่ามีการแสดงออกได้ในทุกเนื้อเยื่อของต้นปาล์มน้ำมันและพบว่า *EgNF-YB1* มีการแสดงออกสูงในเนื้อเยื่อที่มีการสะสมน้ำมัน ในขณะที่ *EgDof318* มีการแสดงออกต่ำลงในเนื้อเยื่อที่มีการสะสมน้ำมันมากขึ้น เมื่อทำการวิเคราะห์ทาง *In silico* ของโปรตีนแฟมิลี่ชนิด EgNF-Ys และ EgDofs พบว่า EgNF-YB1 EgDof318 มีตำแหน่งที่สามารถจับกับตำแหน่งอนุรักษ์บนสายดีเอ็นเอมากที่สุด และเมื่อนำโปรตีน EgNF-YB1 ที่ผลิตได้มาทำการทดสอบการจับกับลำดับดีเอ็นเอของโปรโมเตอร์ของยีน *EgaccD* ด้วยเทคนิค DNA-Protein interaction

enzyme linked immunosorbent พบว่า สามารถจับกับตำแหน่งโปรโมเตอร์ของยีน *EgaccD* ได้อย่างแข็งแรงเมื่อเทียบกับตำแหน่งโปรโมเตอร์ในกลุ่มควบคุม

วัตถุประสงค์ของโครงการ

1) เพื่อทำการโคลนและวิเคราะห์ลักษณะของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ในปาล์มน้ำมัน
2) เพื่อศึกษาระดับการแสดงออกของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ในปาล์มน้ำมัน
3) ทำการโคลนชิ้นส่วนโปรโมเตอร์ของยีน *EgaccD* และ *EgaccB* จากปาล์มน้ำมัน
4) ศึกษาความจำเพาะของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ในการจับกับโปรโมเตอร์ของยีน *EgaccD* และ *EgaccB* จากปาล์มน้ำมันปาล์มน้ำมัน
5) ศึกษาสมบัติและหน้าที่ของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ในการควบคุมการเพิ่มน้ำมันในเอมบริโอจีนิกแคลลัส

วิธีการดำเนินงานวิจัย

2.1 ตัวอย่างที่ใช้ในการทำวิจัย

เก็บตัวอย่างใบหรือผลจากปาล์มน้ำมันพันธ์ *E. guineensis* สายพันธุ์ เทเนอรา จังหวัดสงขลาที่ทราบข้อมูลทางด้านผลผลิตเพื่อใช้สำหรับการสกัดดีเอ็นเอและอาร์เอ็นเอในการทดลองวิจัย โดยตัวอย่างเมื่อเก็บมาให้ทำความสะอาดและแยกชิ้นส่วนที่ต้องการเช่น ตัวอย่างใบ ลำต้น ราก เนื้อเยื่อ Mesocarp ที่ช่วงอายุของผลตั้งแต่ 12-26 สัปดาห์หลังการผสม ในการศึกษาปัจจัยภายนอก (Abiotic stress) ที่จะส่งผลต่อระดับการแสดงออกของยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs จะเลือกทดสอบผลอันเนื่องมาจาก NaCl, PEG, Light, Dark, sodium nitrate และ sodium acetate ที่ระยะเวลาและความเข้มข้นต่างๆ เมื่อครบตามสภาวะที่ต้องการนำตัวอย่างที่ได้นั้นทำการแช่ในไนโตรเจนเหลวทันทีก่อนเก็บ -80 °C จนกว่าจะใช้งาน ในขณะที่เอ็มบริโอจะแยกจากผลปาล์มอายุ 16-20 สัปดาห์หลังการผสมแล้วนำมาเลี้ยงในสภาวะปลอดเชื้อบน MS medium, pH 5.7 ที่มี 2 mg/L 2,4-D หรือ 1 mg/L BA และ 1 mg/L NAA เพื่อชักนำให้เกิดแคลลัสได้ภายในสัปดาห์ 4-6 และทำการย้ายเลี้ยงบน MS medium, pH 5.7 เป็นระยะเวลา 2-4 สัปดาห์ ก่อนนำไปทำการถ่ายฝากยีน

2.2 โคลนและวิเคราะห์ลักษณะของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs

2.2.1 ออกแบบไพรเมอร์สำหรับโคลนยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ออกแบบไพรเมอร์สำหรับโคลนยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ในปาล์มน้ำมันจากข้อมูล transcriptome ของปาล์มน้ำมันเพื่อโคลนยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs แต่ละ ไอโซฟอร์ม สำหรับหาลำดับนิวคลีโอไทด์เพื่อใช้ในการวิเคราะห์ด้วย เครื่องมือ

ทางด้านชีวสารสนเทศ (Bioinformatics Tools) และสังเคราะห์ไพรเมอร์เพื่อใช้ในการทำ RT-PCR/Real Time PCR เพื่อศึกษาระดับการแสดงออกของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ในปาล์มน้ำมันศึกษาระดับการแสดงออกของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ในปาล์มน้ำมัน

2.2.2 การเตรียม Total RNA ด้วย RNeasy Plant Mini Kit

นำปาล์มน้ำมันที่บดละเอียดด้วยไนโตรเจนเหลว เติมลงในบัฟเฟอร์ RLC และ β-Mercaptoethanol ดูดส่วนผสมนี้ไปยัง แล้วหมุนเหวี่ยงเพื่อเก็บส่วนใสเติม 100 % ethanol แล้วดูดส่วนใสใส่คอลัมน์หมุนเหวี่ยง ทิ้งส่วนใส เติมบัฟเฟอร์ RW1 ลงในคอลัมน์หมุนเหวี่ยงทิ้งส่วนใส เติมบัฟเฟอร์ RPE1 ลงในคอลัมน์ หมุนเหวี่ยงย้ายคอลัมน์ลงในหลอดใหม่ทำการชะอาร์เอ็นเอออกจากคอลัมน์ด้วยน้ำกลั่นนำอาร์เอ็นเอที่สกัดได้ไปวัดค่าการดูดกลืนแสงที่ความยาวคลื่น 260 นาโนเมตรและ 280 นาโนเมตร คำนวณหาความเข้มข้น และตรวจคุณภาพของอาร์เอ็นเอ ด้วย 1.0 % agarose gel electrophoresis อาร์เอ็นเอที่ได้รักษาไว้ที่ –20 องศาเซลเซียสจนกว่าจะใช้งาน

2.2.3 ทำการโคลนกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ ด้วยเทคนิค RT-PCR

ทำการโคลนยีนเส้นสมบูรณ์ของยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ด้วยการนำ Total RNA ที่สกัดได้จากปาล์มน้ำมันมาทำ RT-PCR โดยเติม 5x buffer 10 ไมโครลิตร dNTPs mix 2 ไมโครลิตร ไพรเมอร์ 5’ ที่จำเพาะกับยีนทรานสคริปชั่นแฟคเตอร์ 5 ไมโครลิตร ไพรเมอร์ 3’ ที่จำเพาะกับยีนทรานสคริปชั่นแฟคเตอร์ 5 ไมโครลิตร เติม QIAGEN Onestep RT-PCR enzyme mix 2 ไมโครลิตร ปรับปริมาตรด้วยน้ำกลั่นจนครบ 50 ไมโครลิตร ผสมให้เข้ากันบ่มให้เกิดปฏิกิริยาในเครื่อง automated thermal cycler GeneAmp®PCR system model 2400 แล้วตรวจสอบ ด้วย agarose gel electrophoresisc และทำการทำบริสุทธิ์ผลผลิตพีซีอาร์เพื่อเชื่อมดีเอ็นเอกับดีเอ็นเอพาหะ แล้วตรวจสอบหาดีเอ็นเอ ลูกผสมและศึกษาลำดับเบสว่าสามารถแปลรหัส เป็นโปรตีนที่ถูกต้องได้โดยใช้เครื่อง ABI PRISM 377 DNA sequencer (Version 3.2) และ ABI PRISM™ BigDye™ Termination kit ออกแบบไพร์เพื่อโคลนหานิวคลีโอไทด์เส้นสมบูรณ์ แต่ละไอโซฟอร์มของยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs

2.2.4 การวิเคราะห์ผลข้อมูลด้วยคอมพิวเตอร์ (Bioinformatics tools)

นำลำดับหานิวคลีโอไทด์เส้นสมบูรณ์ แต่ละไอโซฟอร์มของยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ

Dofs ที่ได้มาเปรียบเทียบกับข้อมูลของธนาคารยีนซึ่งสามารถสืบค้นได้จากธนาคารยีน Genbank (http://www.ncbi.nlm.nih. gov) แล้วนำลำดับเบสที่ได้มาเปรียบเทียบกันโดยใช้โปรแกรม ClustalX version 2.0 และโปรแกรม GENEDOC version 2.6.001 เพื่อศึกษาความเหมือนและความแตกต่างของลำดับเบสและลำดับกรดอะมิโนของยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ที่ได้จากปาล์มน้ำมันเทียบกับสิ่งมีชีวิตชนิดอื่นๆ จัดทำโครงสร้างความสัมพันธ์ทางพันธุกรรม ของประชากรโดยวิธี neighbor-joining method ด้วย PHYLIP package และประมวลผลและสรุป phylogenetic tree จากโปรแกรม PHYLIP นำลำดับกรดอะมิโนของยีนท

รานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ไปคำนายค่าพารามิเตอร์ต่างๆ ด้วยโปรแกรมทางชีวสารสนเทศได้แก่ค่า hydrophobic profiles น้ำหนักโมเลกุล และค่า pI ของโปรตีนจากยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ด้วย PROSITE และ PROSCAN วิเคราะห์ข้อมูลโครงสร้างตติยภูมิของโปรตีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs โปรแกรมในฐานข้อมูล I-TASSER และ RaptorX จากนั้นนำโมเดลของโปรตีนจากแต่ละโปรแกรมมาวิเคราะห์หาโมเดลที่เหมาะสมที่สุดโดยจะพิจารณาที่ความสมบูรณ์ครบถ้วนของโครงสร้างและวิเคราะห์ด้วย Ramachandran plot ซึ่งทำในโปรแกรมออนไลน์ rampage หรือ PROCHECK เพื่อตัดเลือกค่าการคำนวณที่ดีที่สุดแล้วทำการแสดงผลโครงสร้างตติยภูมิของโปรตีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ด้วยโปรแกรม PyMOL

2.3 โคลนชิ้นส่วนโปรโมเตอร์ของยีน *EgaccD* และ *EgaccB* จากปาล์มน้ำมัน

นำไพร์เมอร์นำผลข้อมูลลำดับนิวคลีโอไทด์ (Sequence) ที่ได้จากการศึกษาก่อนหน้านี้ (Nakkaew, et al., 2014) และฐานข้อมูลจีโนมทั้งหมดของปาล์มน้ำมันจากธนาคารยีนของ NCBI (http://www.ncbi.nlm.nih.gov) มาทำการเพิ่มจำนวนด้วยเทคนิค PCR ตามกรรมวิธีของชุดน้ำยา 5x HOT FIREPol® Blend Master Mix (Solis BioDyne) ทำบริสุทธิ์ผลผลิตพีซีอาร์ (PCR product) ด้วยชุดน้ำยา QIAquick PCR Purification นำ PCR product ที่ได้มาตรวจคุณภาพด้วย 1.0% agarose gel electrophoresis และเชื่อมดีเอ็นเอกับดีเอ็นเอพาหะ pGEM®-T Easy ด้วย T4 DNA ligase ภายใต้สภาวะที่เหมาะสม นำดีเอ็นเอลูกผสมเข้าสู่เซลล์เจ้าบ้าน *E.coli* Top 10F' ตามวิธีการของ Maniatis, et. al., 1982 แล้วทำการสกัดพลาสมิดดีเอ็นเอเพื่อตรวจหาดีเอ็นเอลูกผสมตามวิธีของ Holmes และ Quigley, 1981 ทำบริสุทธิ์ดีเอ็นเอลูกผสมด้วย QIAprep Spin Miniprep Kits (QIAGEN) เพื่อนำไปหาลำดับนิวคลีโอไทด์เพื่อตรวจสอบลำดับนิวคลีโอไทด์ วิเคราะห์แบบแผนดีเอ็นเอด้วยโปรแกรม ClustalX 2.0 (Thompson et al. 1997) และโปรแกรม GENEDOC version 2.6 (Nicholas and Nicholas 1997) เพื่อตรวจสอบลำดับนิวคลีโอไทด์ในลำดับโปรโมเตอร์ของยีน *EgaccD* และ *EgaccB* ต่อไปด้วย BLAST Genomes (https://blast.ncbi.nlm.nih.gov/Blast.cgi) และ PlantPAN2 (http://PlantPAN2.itps.ncku.edu.tw)

2.4 ศึกษาแบบแผนกระดับการแสดงออกของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ ด้วยเทคนิค RT-PCR

สังเคราะห์ cDNA จาก total RNA ที่สกัดได้จากเนื้อเยื่อส่วนต่างๆ ของปาล์มน้ำมันโดยใช้ SuperScript™III First-Strand synthesis system (Invitrogen) ตรวจสอบคุณภาพของดีเอ็นเอและความเข้มข้นของดีเอ็นเอโดยวัดการดูดกลืนแสงที่ความยาวคลื่น 260 และ 280 นาโนเมตร หลังจากนั้นจึงนำ cDNA ที่ความเข้มข้น 200 นาโนกรัมมาทำ PCR ตามกรรมวิธีของชุดน้ำยา 5x HOT FIREPol® Blend Master Mix (Solis BioDyne) ด้วยไพรเมอร์ที่จำเพาะกับยีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs และใช้ยีน 18S rDNA เป็นยีนเปรียบเทียบมาตรฐาน นำผลที่ได้ไปประมวลผลและหาค่าทางสถิติเพื่อวิเคราะห์และสรุปผล

2.5 ความจำเพาะของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ในการจับกับโปรโมเตอร์ของยีน *EgaccD* และ *EgaccB* จากปาล์มน้ำมันปาล์มน้ำมัน

2.5.1 การผลิตโปรตีนลูกผสมของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs

ทำการเพิ่มจำนวนยีนของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ที่มีตำแหน่งเอนไซม์ตัดจำเพาะที่สามารถโคลนส่วน ORF ของยีนของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ที่ต้องด้วยไพรเมอร์ที่ออกแบบให้มีจำเพาะกับเวคเตอร์ pET-28a (Novagen) หรือเวคเตอร์ pQE40 vectors (Qiagen) ที่สามารถผลิตโปรตีนในลำดับกรดอะมิโนที่ถูกต้อง สำหรับผลิตโปรตีนลูกผสมของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ซึ่งในการเพิ่มจำนวนยีนของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs จะใช้ด้วยเทคนิค Reverse transcript-polymerase chain reaction (RT-PCR) จากนั้นทำการตัดผลผลิตพีซีอาร์ด้วยเอนไซม์ตัดจำเพาะที่เฉพาะของแต่ละยีนของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs เพื่อโคลนเข้าสู่เวคเตอร์ pET-28a (Novagen) หรือเวคเตอร์ pQE40 vectors (Qiagen) ในตำแน่งของเอนไซม์ตัดจำเพาะดังกล่าว ทำการคัดเลือกและนำดีเอ็นเอลูกผสมแต่ละยีนในกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs ไปส่งหาลำดับนิวคลีโอไทด์เพื่อตรวจสอบนำเข้าของยีนบนเวคเตอร์ที่ถูกต้อง ก่อนจะนำดีเอ็นเอลูกผสมแต่ละยีนเข้าสู่เซลล์เจ้าบ้าน *E.coli* สายพันธุ์ BL21 (DE3) หรือ *E.coli* สายพันธุ์ M15 (pRep4) เพื่อชักนำให้มีการแสดงออกของโปรตีนต่อไป โดยทำการนำโคโลนีเดี่ยว เลี้ยงลงไปในอาหารเลี้ยงเชื้อเหลว (LB broth) ที่มียาปฏิชีวนะที่เหมาะสมเวคเตอร์ pET-28a (Novagen) หรือเวคเตอร์ pQE40 vectors (Qiagen) ปริมาตร 5 มิลลิลิตร บ่มในเครื่องเขย่าที่อุณหภูมิ 37 องศาเซลเซียส ความเร็ว 180 รอบต่อนาที เป็นเวลานาน 16 ชั่วโมง เพื่อทำเป็นหัวเชื้อ เมื่อครบกาหนดเวลา นำหัวเชื้อเริ่มต้นทำการย้ายเชื้อลงในอาหารอาหารเลี้ยงเชื้อเหลว (LB broth) ที่มียาปฏิชีวนะที่เหมาะสมเวคเตอร์ pET-28a (Novagen) หรือเวคเตอร์ pQE40 vectors (Qiagen) ปริมาตร 100 มิลลิลิตร บ่มในเครื่องเขย่าที่อุณหภูมิ 37 องศาเซลเซียส จนเซลล์มีค่าการดูดกลืนแสงที่ความยาคลื่น 600 นาโนเมตร เท่ากับ 0.4-0.6 จากนั้นแบ่งเก็บเชื้อไว้ในหลอดทดลองขนาด 1.5 มิลลิลิตร ปริมาตร 1 มิลลิลิตร เพื่อเป็นโปรตีนก่อนทำการชักนำ (non induce protein) ต่อมานำเซลล์แบคทีเรียที่เหลือทั้งหมดเติม isopropyl-β-D-thiogalactopyranoside (IPTG) ให้ความเข้มข้นสุดท้ายเป็น 0.05-0.1 มิลลิโมลาร์ เพื่อชักนำให้แบคทีเรียมีการสร้างโปรตีน จากนั้นเลี้ยงเชื้อต่อด้วยเครื่องเขย่าที่อุณหภูมิ 37 องศาเซลเซียส ความเร็ว 180 รอบต่อนาที อีกประมาณ 2 ชั่วโมง เมื่อครบเวลา แบ่งเก็บเชื้อไว้ในหลอดทดลองขนาด 15 มิลลิลิตร แล้วทำให้เซลล์แตกก่อนนำไปการทำบริสุทธิ์โปรตีน ด้วย Ni-NTA bead ตรวจสอบผลการแสดงออกของโปรตีนด้วย12% โพลีอะคลิลาไมด์เจลอิเล็กโตรโฟรีซิส เก็บโปรตีนลูกผสมได้ที่อุณหภูมิ -20 องศาเซลเซียส จนกว่าจะใช้งาน

2.5.2 การทำ DPI-ELISA ดูการจับของโปรตีนลูกผสมของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NFY และ Dofs กับโปรโมเตอร์

นำไพรเมอร์ที่จำเพาะกับโปรโมเตอร์ที่ต้องการใช้ในการทดสอบ (สังเคราะห์ติด 5’ biotin ไว้ที่ไพรเมอร์เส้น forward) มาเพิ่มจำนวนโปรโมเตอร์ที่ต้องการที่ได้ด้วยเทคนิค PCR ตามกรรมวิธี

ของชุดน้ำยา 5x HOT FIREPol® Blend Master Mix (Solis BioDyne) จากนั้นนำมาทำปฏิกิริยาในเครื่อง automated thermal cycler GeneAmp® PCR system model 2400 โดยมีสภาวะ คือ อุณหภูมิ 95 องศาเซลเซียส เวลา 15 นาที จำนวน 1 รอบ อุณหภูมิ 95 องศาเซลเซียส เวลา 20 วินาที อุณหภูมิ 55 องศาเซลเซียส เวลา 20 วินาที อุณหภูมิ 72 องศาเซลเซียส เวลา 2 นาที จำนวน 40 รอบ อุณหภูมิ 72 องศาเซลเซียส เวลา 10 นาทีจำนวน 1 รอบและอุณหภูมิ 8 องศาเซลเซียส จนกว่าจะนำผลผลิตออกจากเครื่องตรวจสอบผลผลิตที่ได้ด้วย 1.5 % agarose gel electrophoresis นำผลิตภัณฑ์ที่ได้ที่อุณหภูมิ -20 องศาเซลเซียส จากนั้นนำมาทำบริสุทธิ์ตามกรรมวิธีของชุดน้ำยา QIA quick PCR Purification Kit ต่อมาทำการบ่มผลิตภัณฑ์ที่ได้ใน streptavdin-coated plates โดยเติมโปรโมเตอร์ที่ทำการเพิ่มจำนวนแล้วเจืองจางให้มีความเข้มข้น 2 pmol ใน TBS-T บ่มที่อุณหภูมิ 37 องศาเซลเซียส เป็นเวลา 1 ชั่วโมง ต่อมาทำการล้าง 3 ครั้งด้วย TBS-T จำนวน 150 ไมโครลิตร บ่มต่อด้วย 5% non-fat dry milk ใน TBS-T ที่อุณหภูมิห้อง เป็นเวลา 0.50 ชั่วโมง และทำการล้างด้วย TBS-T ปริมาตร 150 ไมโครลิตร จำนวน 3 ครั้ง นำสารละลายโปรตีนลูกผสมทรานสคริปชั่นแฟคเตอร์ NFY หรือ Dofs ที่ความเข้มข้น 100 ไมโครกรัม ลงไปบ่มใน streptavdin-coated plates ที่อุณหภูมิห้องเป็นเวลา 1 ชั่วโมง ทำการล้าง 3 ครั้งด้วย TBS-T จำนวน 150 ไมโครลิตร ต่อมาเติม antibody (α-His-HRP antibody conjugate (Qiagen) เจือจาง 1:1000 ใน PBS-T) บ่มที่อุณหภูมิห้องเป็นเวลา 1 ชั่วโมง ทำการล้าง 2 ครั้งด้วย PBS-T150 ไมโครลิตร และล้างอีก 2 ครั้งด้วย PBS150 ไมโครลิตร ต่อมาตรวจสอบผลโดยการนำ plates ไปบ่มด้วย OPD-solution ในที่มืดเวลา 30 นาที หลังจากนั้นเติม stopping solution นำไปวัดค่าสารเรืองแสงที่ความยาวคลื่น 492 นาโนเมตร และ 650 นาโนเมตรเป็นพื้นหลัง ด้วยเครื่อง ELISA-reader

ผลและวิจารณผลการทดลอง

ผลการศึกษาลำดับนิวคลีโอไทด์ของของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ NF-Y

ผลการโคลนและหาลำดับนิวคลีโอไทด์ขอวยีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y จากเนื้อเยื่อใบอ่อนปาล์มน้ำมัน *E. guineensis* อายุ 2 เดือน โดยใช้เทคนิค RT-PCR เมื่อนำไปหาลำดับนิวคลีโอไทด์ และเปรียบเทียบ ลำดับนิวคลีโอไทด์โดย BLASTX พบว่าเป็นยีน NF-Y จากปาล์มน้ำมัน 4 รูปแบบคือยีน *EgNF-YA, EgNF-YB1, EgNF-YB3* และยีน *EgNF-YC* โดยมีลำดับนิวคลีโอไทด์เส้นสมบูรณ์ (full length) ของยีน NF-Y จากปาล์มน้ำมันแต่ละยีนดังแสดงด้านล่าง

>EgNF-YA

MATPVGYMMPPGHLETGHAVAQIAYPYVDPYYGGGFAGFGQTVFHPQMMGMQHPGV
PLPTEAVEEPVFVNAKQYHGILRRRQSRAKAESENKLIKTRKPYLHESRHLHALKRARGC
GGRFLNSKSEGNQQNSISNSEGNQQNSNPSSEGNQQNSNSKFEENQKNGPKAAVSDD
KEKPVSLTNSDKSSLDNKPEATS

>EgNF-YB1

MAGSGTGGDHANYGGSSDGNPAAAAGQCSVREQDRFMPIANVIRIMRKVLPTHAKISD
DAKETIQECVSEYISFITSEANERCQREQRKTITAEDVLWAMSKLGFDDYIEPLSLYLQRY
RELEGDHRGSIRGEPLPTIRHRSPNPSTDVVPFFPQPPPPSFVVPPPSGPPHLNHLYGD
GAGGYYYGMYGGGDGSGASTSQDPPMPGFDHSYAPYKQ

>EgNF-YB3

MPDSDNESGGQNNSHNNNVGEYSSSREQDRFLPIANVSRIMKKALPANAKISKDAKETV
QECVSEFISFITGEAADKCQREKRKTINGDDLLWAMTTLGFEDYVDPLKVYLHRFREME
GDKCSAGASASSQPQHKDGGDGGGGGGGGAPSMGNNVVGLGGGGGGAGGMMMM
MGQQMYATPPSYHHHMSTMSGKSSMGGGSSASSSSPGFGRQGRV

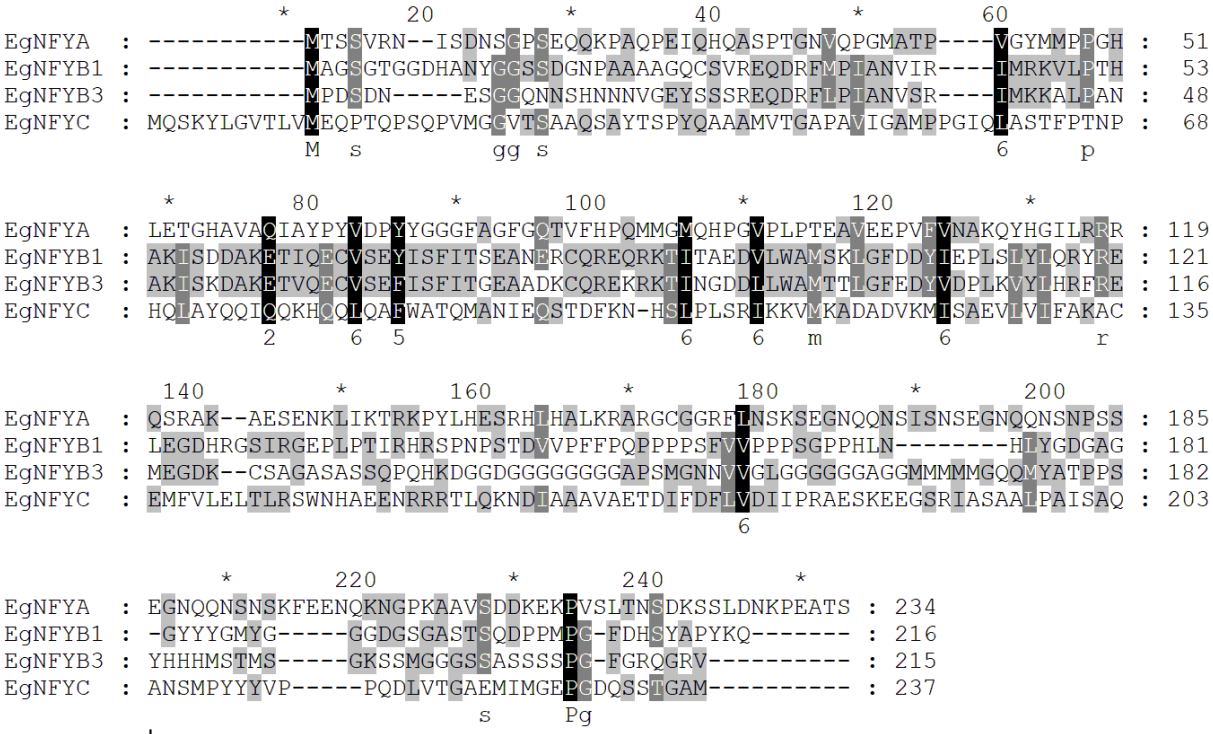
>EgNF-YC

MQSKYLGVTLVMEQPTQPSQPVMGGVTSAAQSAYTSPYQAAAMVTGAPAVIGAMPPGI
QLASTFPTNPHQLAYQQIQQKHQQLQAFWATQMANIEQSTDFKNHSLPLSRIKKVMKAD
ADVKMISAEVLVIFAKACEMFVLELTLRSWNHAEENRRRTLQKNDIAAAVAETDIFDFLVD
IIPRAESKEEGSRIASAALPAISAQANSMPYYYVPPQDLVTGAEMIMGEPGDQSSTGAM

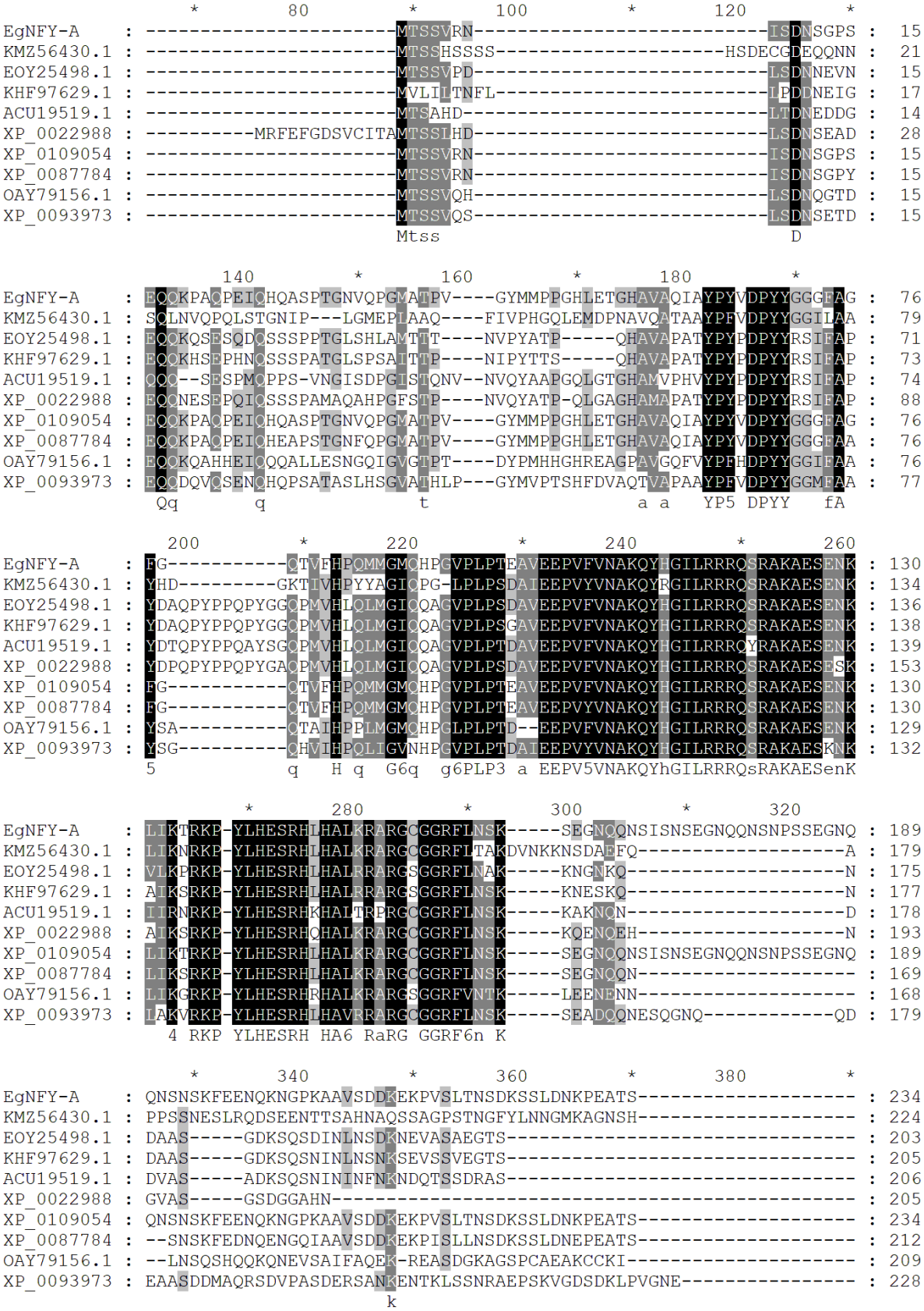
โดยจากข้อมูลลำดับกรดอะมิโนของโปรตีนNF-Y จากปาล์มน้ำมันแต่ละยีนพบว่าโปรตีน EgNF-YA (รูปที่ 2), EgNF-YB1 (รูปที่ 3), EgNF-YB3 (รูปที่ 4) และ EgNF-YC (รูปที่ 5) มีลำดับกรดอะมิโนที่แตกต่างกันเมื่อวิเคราะห์เปรียบเทียบลำดับกรดอะมิโนด้วยโปรแกรม ClustalX version 1.81 และโปรแกรม GENEDOC 2.6.001 ดังรูปที่ 1 เมื่อทำการวิเคราะห์ผลเบื้องต้น ได้แก่คำนวณค่าน้ำหนักมวลโมเลกุลของโปรตีน ค่า isoelectric point (pI) แสดงผลดังในตารางที่ 1 โดยข้อมูลลำดับกรดอะมิโนของโปรตีน NF-Y จากปาล์มน้ำมันได้ทำการวิเคราะห์ข้อมูลโครงสร้างตติยภูมิของโปรตีนทรานสคริปชั่นแฟคเตอร์ NF-Y และ Dofs โปรแกรมในฐานข้อมูล I-TASSER เรียบร้อยแล้วและอยู่ในระหว่างการวิเคราะห์หาโมเดลที่เหมาะสมที่สุด โดยจะพิจารณาที่ความสมบูรณ์ครบถ้วนของโครงสร้างและวิเคราะห์ด้วย Ramachandran plot ซึ่งทำในโปรแกรมออนไลน์ rampage หรือ PROCHECK เพื่อตัดเลือกค่าการคำนวณที่ดีที่สุด แล้วทำการแสดงผลโครงสร้างตติยภูมิของโปรตีนทรานสคริปชั่นแฟคเตอร์ NF-Y ด้วยโปรแกรม PyMOL

ตารางที่ 1 แสดงจำนวนนิวคลีโอไทป์ กรดอะมิโน ค่าน้ำหนักมวลโมเลกุลของโปรตีนและ ค่า isoelectric point (pI) ของโปรตีน NFY จากปาล์มน้ำมัน

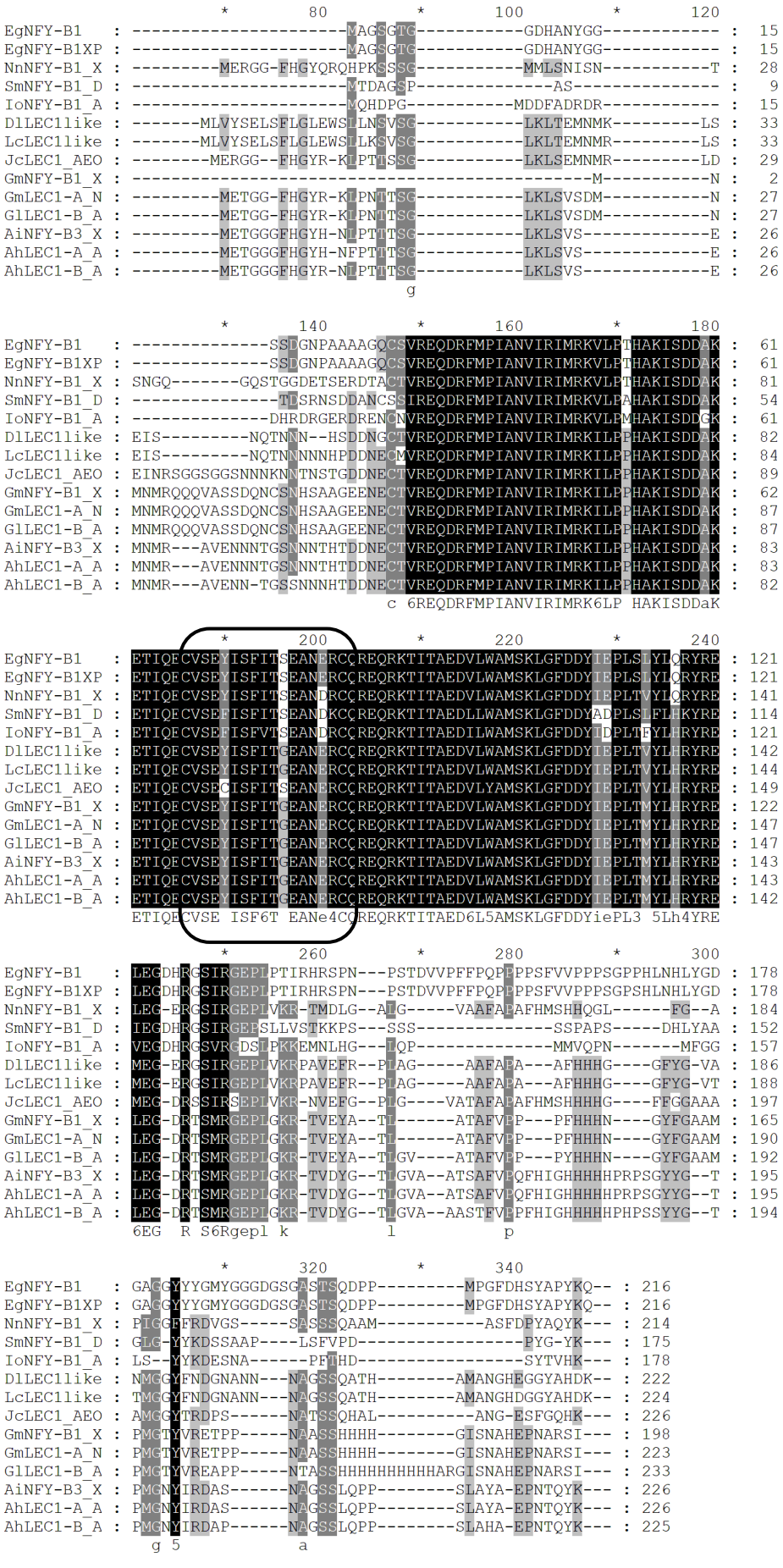
Gene	ORF length	Deduced protein length (aa)	Theoretical protein mol.wt. (Da)	Theoretical protein (pI)
EgNFY-A	705	234	25507.11	8.89
EgNFY-B1	651	216	23549.13	5.45
EgNFY-B3	648	215	22467.85	6.30
EgNFY-C	714	237	25777.47	5.37



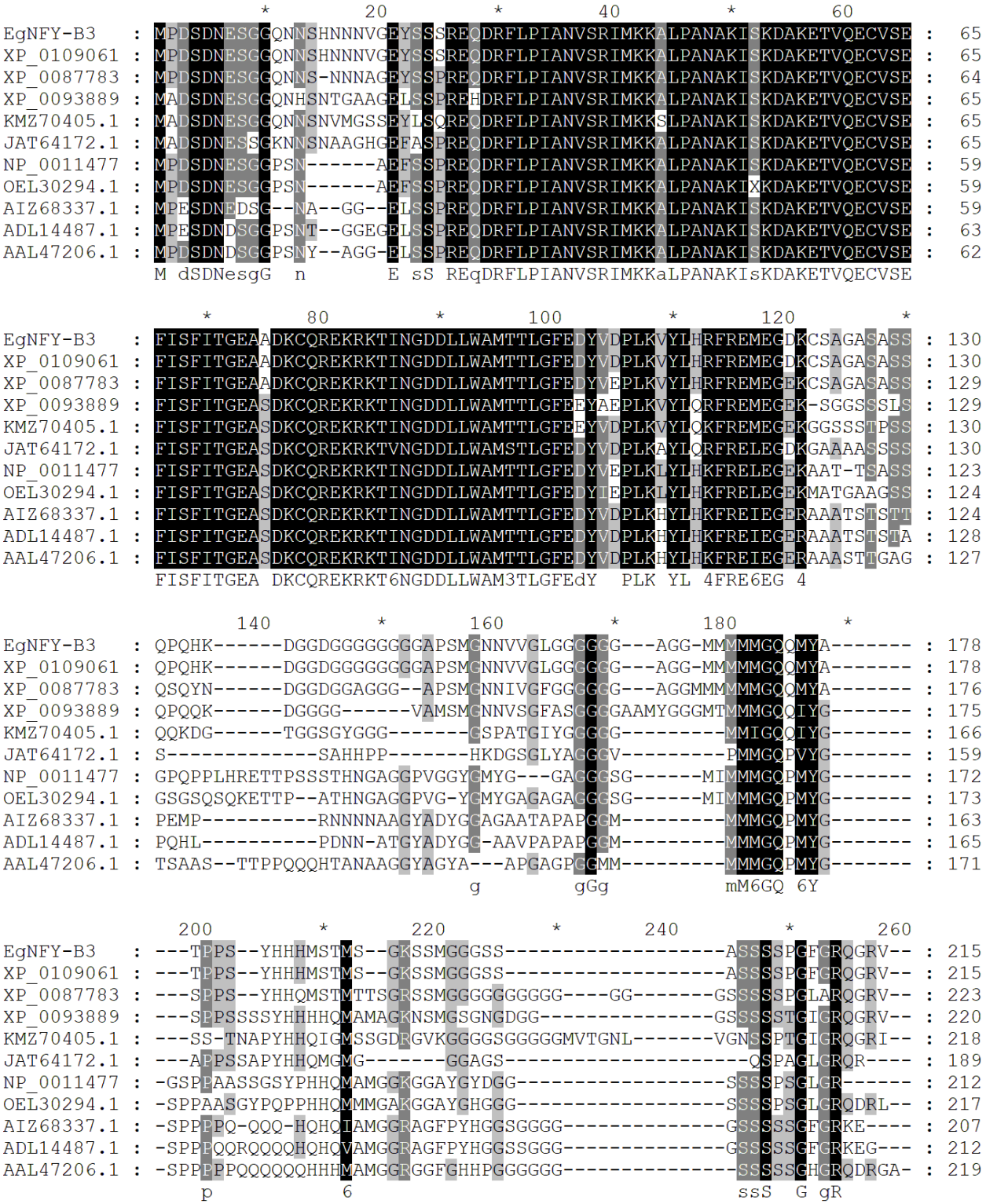
รูปที่1 แสดงการเปรียบเทียบลำดับกรดอะมิโนของของกลุ่มโปรตีน NF-Y จากปาล์มน้ำมัน



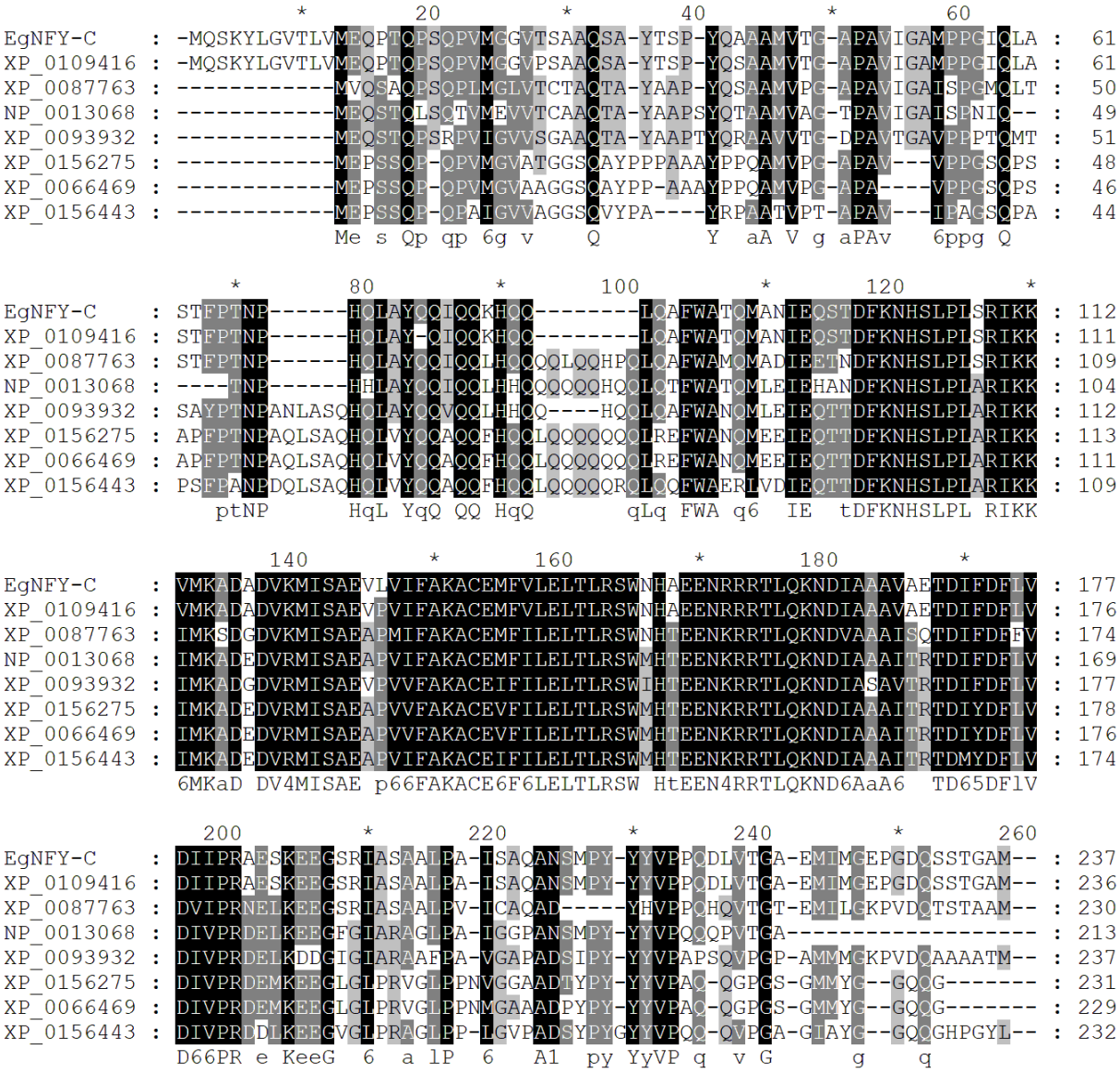
รูปที่2 แสดงการเปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgNF-YA จากปาล์มน้ำมัน เปรียบเทียบกับโปE/nUEMJตีน EgNF-YAที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI



รูปที่ 3 เปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgNF-YB1 จากปาล์มน้ำมันเปรียบเทียบกับโปE/nUEMJตีน EgNF-YB1 ที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI และ ตำแหน่งการทำงานของทรานสคริปชั่นแฟกเตอร์ที่มีการจับจำเพาะกับ promoter บนสายดีเอ็นเอบริเวณ CAAT box ตรงกรดอะมิโนตำแหน่งมีลำดับกรดอะมิโน คือ CVSEyISFITSEANERC ดังในกรอบสี่เหลี่ยมสีดำ



รูปที่4 แสดงการเปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgNF-B3 จากปาล์มน้ำมัน เปรียบเทียบกับโปรตีน EgNF-B3 ที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI



รูปที่ 5 แสดงการเปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgNF-C จากปาล์มน้ำมัน เปรียบเทียบกับโปรตีน EgNF-C ที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI

เมื่อศึกษาบริเวณที่ทำหน้าที่ภายในกลุ่มโปรตีน NF-Y จากปาล์มน้ำมันด้วยโปรแกรม PROSITE พบว่าโปรตีน EgNFY-A พบบริเวณสำคัญที่แสดงสัญลักษณ์ของ NF-YA/HAP2 subunit signature ในตำแหน่งของลำดับกรดอะมิโนที่ 68 – 88 (FVNAKQYhgILRRRqsRAKAE) ซึ่งแสดงลักษณะเฉพาะอนุรักษ์ของโปรตีน NF-YB/HAP2 เป็นตำแหน่งบ่งบอกสมบัติความเป็น CCAAT-binding factor นอกจากยังพบแบบแผนแสดงลักษณะของโปรตีน NF-YA/HAP2 family profile ในตำแหน่งของลำดับกรดอะมิโนที่ 63-123 คือ VEEPVFVNAKQYHGILRRRQSRAKAESENKLIKTRKPYLHESRHLHALKRARGCGGRFLNS โดยตำแหน่งที่ 68-91 (FVNAKQYHGILRRRQSRA KAESEN)จะเป็นตำแหน่งอนุรักษ์

Subunit association domain (SAD) Motifในการจับกับโปรตีน 98-123 (LKRARGCGGRFLNS) จะเป็นตำแหน่งอนุรักษ์ NFYA/HAP2-type ในการจับกับ DNA (DNA bind) โปรตีน EgNFY-B1 พบบริเวณสำคัญที่แสดงสัญลักษณ์ของ NF-YB/HAP3 subunit signature ในตำแหน่งของลำดับกรดอะมิโนที่ 67 – 83 (CVSEyISFITSEANERC) ซึ่งเป็นตำแหน่งการทำงานของทรานสคริปชั่นแฟกเตอร์ที่มีการจับจำเพาะกับ promoter บนสายดีเอ็นเอบริเวณ CAAT box แต่โปรตีน EgNFY-B3 และ EgNFY-C นั้นไม่แสดงตำแหน่งบริเวณสำคัญที่แสดงสัญลักษณ์ใดเลยเมื่อวิเคราะห์ลำดับกรดอะมิโนด้วยโปรแกรม PROSITE (https://prosite.expasy.org/) แต่พบว่าเมื่อเมื่อวิเคราะห์ลำดับกรดอะมิโนของโปรตีน EgNFY-B3ด้วยโปรแกรม Interpro (https://www.ebi.ac.uk/interpro/search/ sequence-search) พบว่าในตำแหน่งของลำดับกรดอะมิโนที่ 28-134 แสดงแบบแผนความเป็น Histone-fold homologous superfamilies ซึ่งโปรตีนกลุ่มนี้จะทำหน้าหน้าควบคุมการแสดงของยีนที่เกี่ยวข้องผ่านการควบคุมการทำงานของโปรตีนฮิสโตนเช่นเดียวกับ EgNFY-C ซึ่งพบในตำแหน่งของลำดับกรดอะมิโนที่ 67-182 ที่แสดงแบบแผนความเป็น Histone-fold homologous superfamilies

3.2 ผลการศึกษาลำดับนิวคลีโอไทด์ของของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ Dofs

เมื่อทำการโคลนและหาลำดับนิวคลีโอไทด์ของยีนกลุ่มทรานสคริปชั่นแฟคเตอร์ Dofs จากเนื้อเยื่อใบอ่อนปาล์มน้ำมัน *E. guineensis* อายุ 2 เดือน โดยใช้เทคนิค RT-PCR เมื่อนำลำดับนิวคลีโอไทด์ไปเปรียบเทียบกับฐานข้อมูล NCBI โดย BLASTX พบว่าโปรตีน Dofs จากปาล์มน้ำมัน 5 รูปแบบคือ EgDof300, EgDof314 EgDof318, EgDof343 และ EgDof347 โดยมีลำดับนิวคลีโอไทด์เส้นสมบูรณ์ (full length) และลำดับกรดอะมิโนที่แตกต่างกันดังแสดงด้านล่าง

>EgDof300

MLSSWDNILPCPTRPVMVDRRWKPNVEIAPNCPRYDSSNTKFCYYNNYSLSQPRYFCK
GCRRYWTKGGSLRNVPVGGGCRKNRRGKSVRLSTVVPPDLSYGRRLSGQIRPDLALD
GMVGNSTGSVQSMDASTTTDGSTIDLALLYAKFLNQHPVVEPGLGVTELPGGIDESFDS
SGRQSDTNLNQQTPAPLPNQYNLVECQRVSEPIAEALLGDGDQQSIGNHDLPVDPSAYE
SAAFPTEGISSFLSSNCPGVPNLASEAANYQGFGSLQLEDSMALHQNLIVNDWSTMDYS
SLETFDWL

>EgDof314

MIQELLGGMVEERKISIPGSLLIDPSSSSSSSSSISSTSLSSASPSPSNTSEQQQNLRCPR
CDSSNTKFCYYNNYNLTQPRHFCKTCRRYWTKGGALRNVPIGGGCRKAKASPAIPSGA
SKSGGTKAKPSPPDLLARSGIIGGLEHELPSSPLFWASPHTSHLMALLRATTVQNPNPNP
PSSSTLNNSVRIKEDGAMIGTHMAADTVGALPSQGLSLDPLGQLPLAPCGSLWRNNNQY
AYQQQPQTQQLQQHSSNMLLGEVPSPAIQELYQRFKSSANYCNDQLQAVMSNVGLFG
FFYFFSYYRYDNNVIDCGY

>EgDof318

MQDFQSIPGVSGRIFGGGGGGGGGGADVRRLRGYPAAAQQQPPVKCPRCDSTNTKFC
YYNNYNLSQPRHFCKACRRYWTKGGVLRNVPVGGGCRKTKRSSSSSSSKTSKSSSTT
GDGDKDSRRPSASRSSSDSSSLTAATTEASASASASASASVSASFPDSTLFASSQTSHP
NLSFDPPPQTVALDTLGHQAPEIFSQVAAGSSFTGLMATAATSPAVLGFNFSDPPPIRVG
PQQSPPKAAATLAEEIRMPGLMDQTAPVDLPPFHGRSGGGGGGLTGLDWPATVDPSLF
DLTSAVDPGAFWNQSHWGDADPSLYLP

>EgDof343

MLSRVLPYPPQRSLMAYHSWKPNMELAPVCPRCDSSNTKFCYYNNYSLTQPRYFCKG
CRRYWTKGGTIRNVPVGGGCRKNRRGKSARSQANHPTSLSNLIYDSHGLLDSDRSGHA
QHDGMRGSSHHSNKSELPPSEGSAIDLALLYTKFVNQCPVLNSDPAVPELPGEIMDESF
SVMGGASTSVMNQMPDQGLNQLGGWEITEARFENQMYPEDSNFSLDPGSSESVTFSS
EGVVTGNDVTASDCWTPQPMFPGFQPLLEDDMFHHQDLVIGDCHENADMFGFLQFTE
MQQLELQSRFKWQHFDSQPKRLRKRCSISRNESVCHSDAPIGSSKTTLLSDHLGL

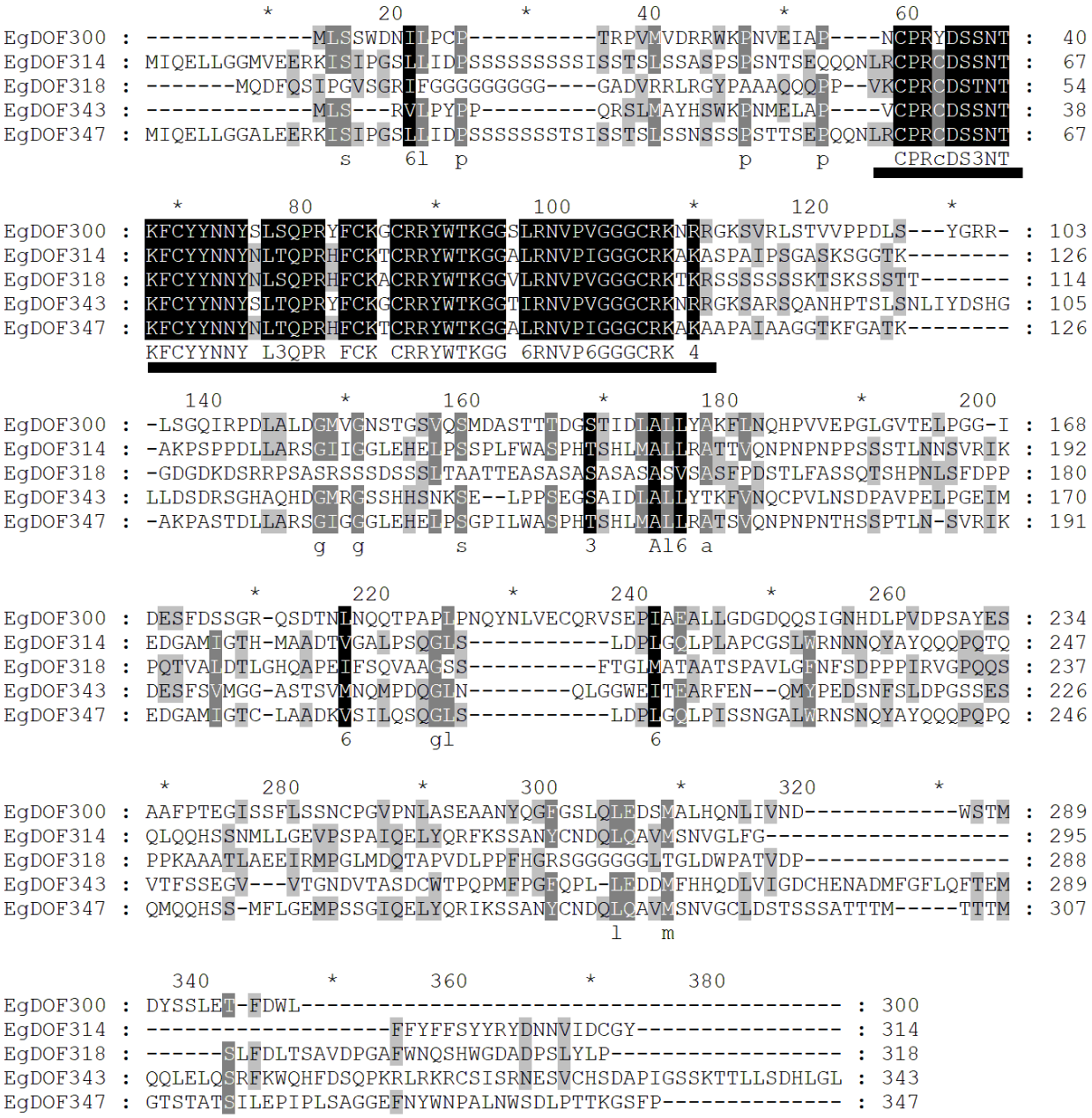
>EgDof347

MIQELLGGALEERKISIPGSLLIDPSSSSSSSTSISSTSLSSNSSSPSTTSEPQQNLRCPRC
DSSNTKFCYYNNY
NLTQPRHFCKTCRRYWTKGGALRNVPIGGGCRKAKAAPAIAAGGTKFGATKAKPASTDL
LARSGIGGGLEHELPSGPILWASPHTSHLMALLRATSVQNPNPNTHSSPTLNSVRIKEDG
AMIGTCLAADKVSILQSQGLSLDPLGQLPISSNGALWRNSNQYAYQQQPQPQQMQQHS
SMFLGEMPSSGIQELYQRIKSSANYCNDQLQAVMSNVGCLDSTSSSATTTMTTTMGTST
ATSILEPIPLSAGGEFNYWNPALNWSDLPTTKGSFP

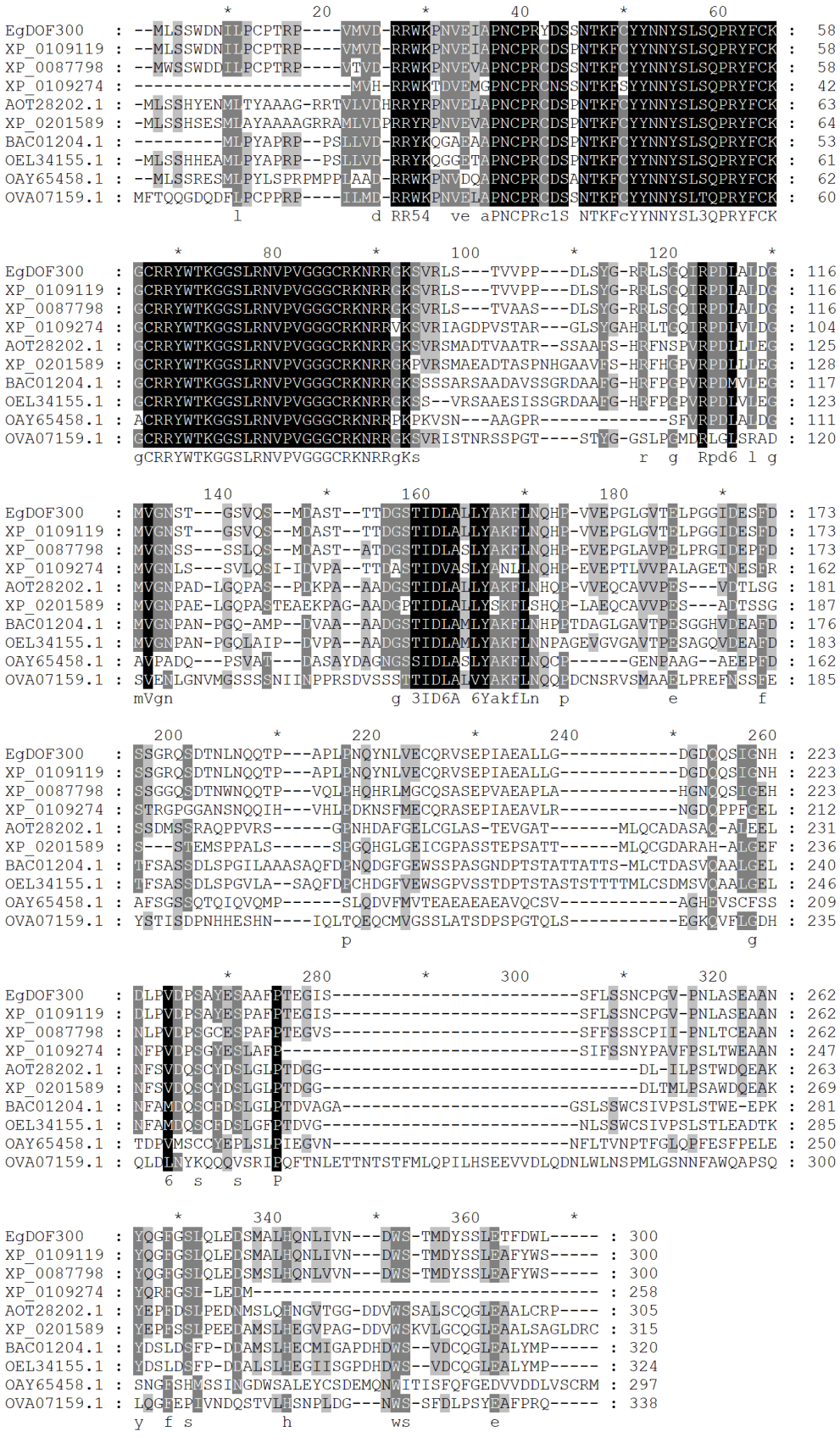
โดยจากข้อมูลลำดับกรดอะมิโนของโปรตีนNF-Y จากปาล์มน้ำมันแต่ละยีนพบว่าโปรตีน คือ EgDof300 (รูปที่ 7), EgDof314 (รูปที่ 8) EgDof318 (รูปที่ 9), EgDof343 (รูปที่ 10) และ EgDof347 (รูปที่ 11) มีลำดับกรดอะมิโนที่แตกต่างกันเมื่อวิเคราะห์เปรียบเทียบลำดับกรดอะมิโนด้วยโปรแกรม ClustalX version 1.81 และโปรแกรม GENEDOC 2.6.001 พบว่าในตำแหน่งที่ กรดอะมิโนที่มีความเหมือนกัน (conserved) จะแสดงผลด้วยตำแหน่งแถบสีดำดังแสดงในรูปที่ 6 เมื่อทำการวิเคราะห์ผลเบื้องต้น ได้แก่คำนวณค่าน้ำหนักมวลโมเลกุลของโปรตีน ค่า isoelectric point (pI) แสดงผลดังในตารางที่ 2 โดยข้อมูลลำดับกรดอะมิโนของโปรตีน Dof จากปาล์มน้ำมันอยู่ในระหว่างการวิเคราะห์ข้อมูลโครงสร้างตติยภูมิของโปรตีนทรานสคริปชั่นแฟคเตอร์ด้วยโปรแกรมในฐานข้อมูล I-TASSER และจะทำการวิเคราะห์หาโมเดลที่เหมาะสมที่สุด โดยจะพิจารณาที่ความสมบูรณ์ครบถ้วนของโครงสร้างและวิเคราะห์ด้วย Ramachandran plot ต่อไป

ตารางที่ 2 แสดงจำนวนนิวคลีโอไทป์ กรดอะมิโน ค่าน้ำหนักมวลโมเลกุลของโปรตีนและ ค่า isoelectric point (pI) ของโปรตีน DOF จากปาล์มน้ำมัน

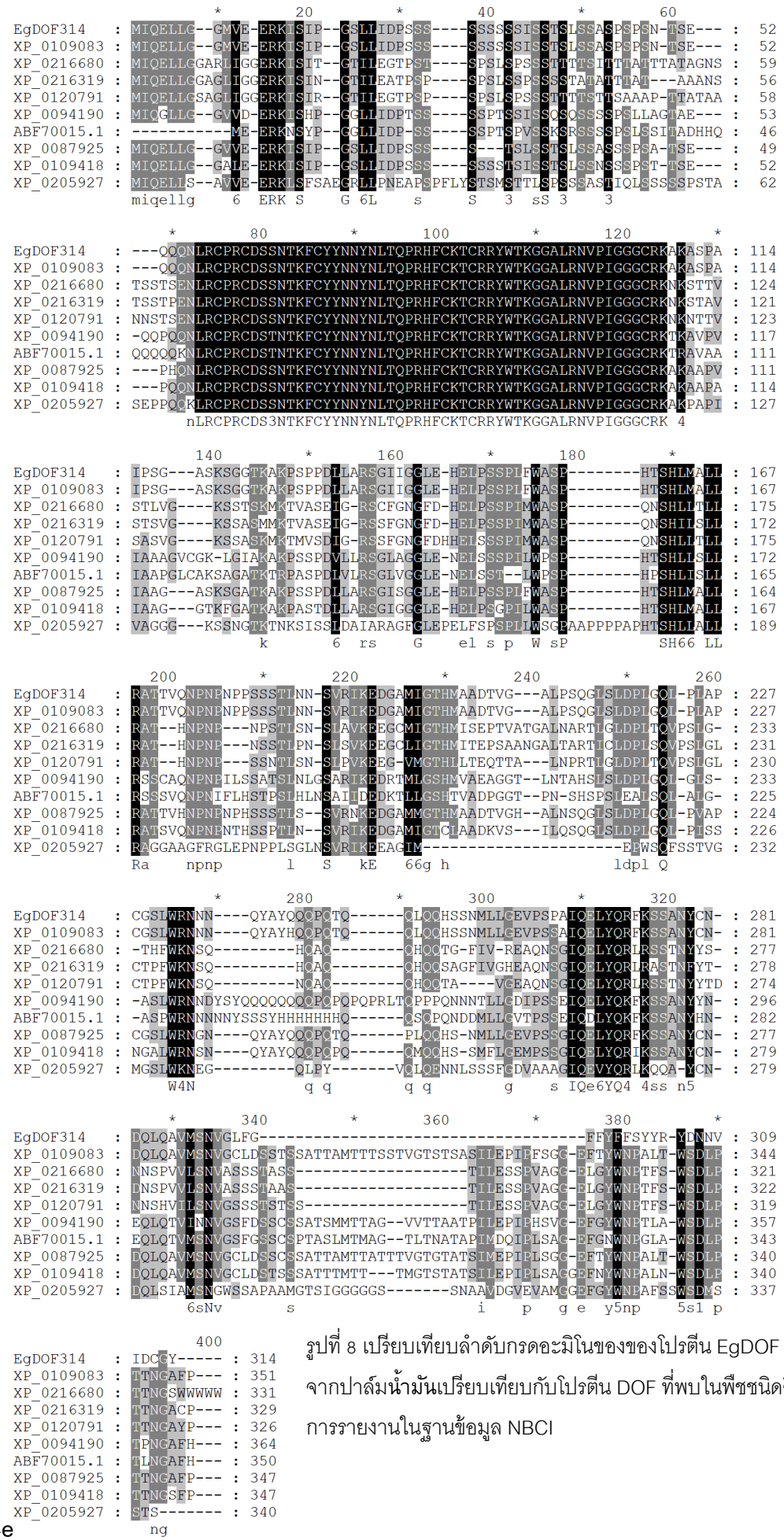
Gene	ORF length	Deduced protein length (aa)	Theoretical protein mol wt. (Da)	Theoretical protein (pI)
EgDof300	903	300	32875.61	4.94
EgDof314	945	314	34068.25	8.93
EgDof318	957	318	32799.16	8.75
EgDof343	1032	343	38387.90	5.99
EgDof347	1044	347	36930.42	8.83



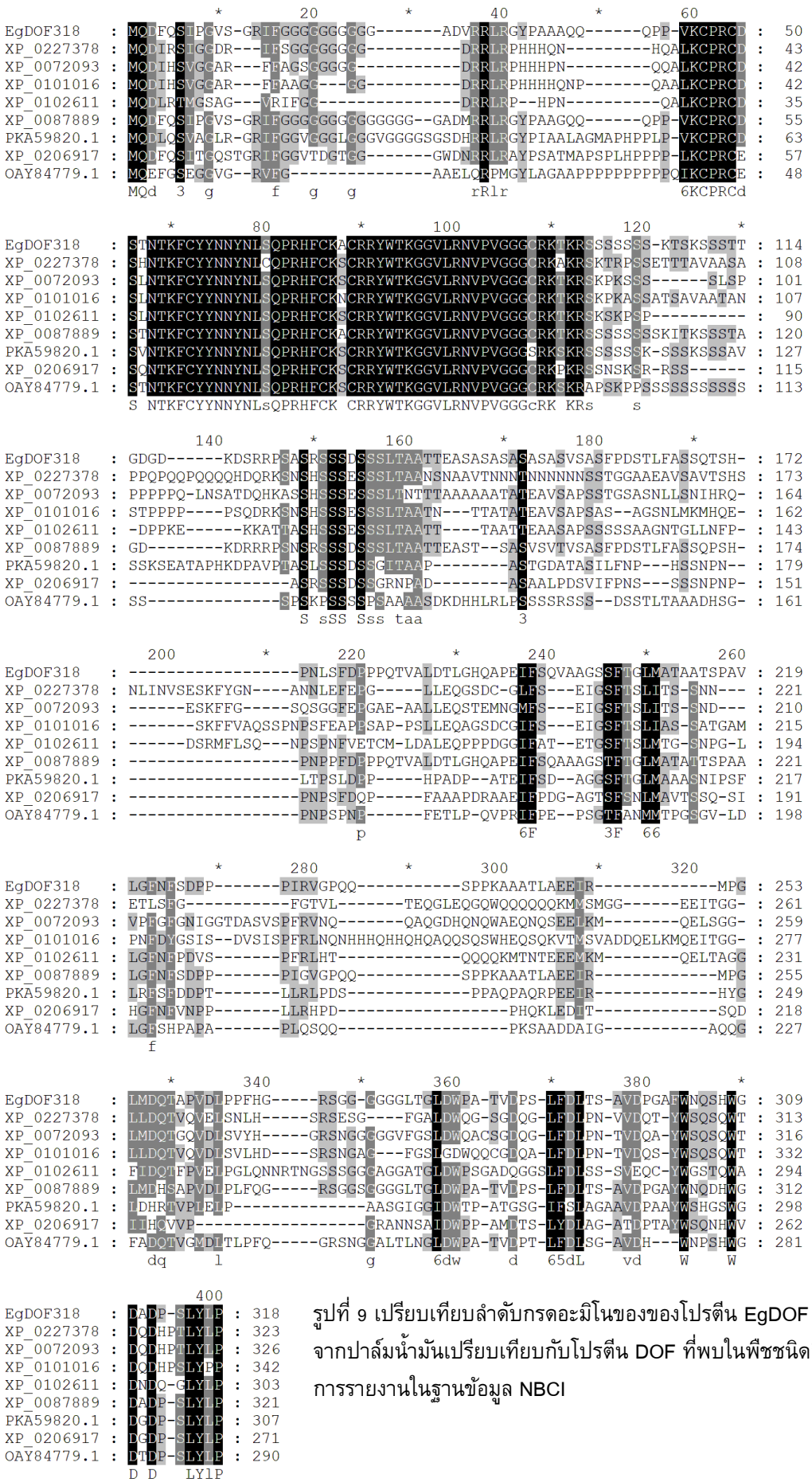
รูปที่ 6 แสดงการเปรียบเทียบลำดับกรดอะมิโนของของกลุ่มโปรตีน DOF จากปาล์มน้ำมันซึ่งแสดงแบบแผนตำแหน่งอนุรักษ์ของ Zinc finger Dof-type signature และ Zinc finger Dof-type profile ดังแสดงตำแหน่งที่เส้นใต้สีดำ



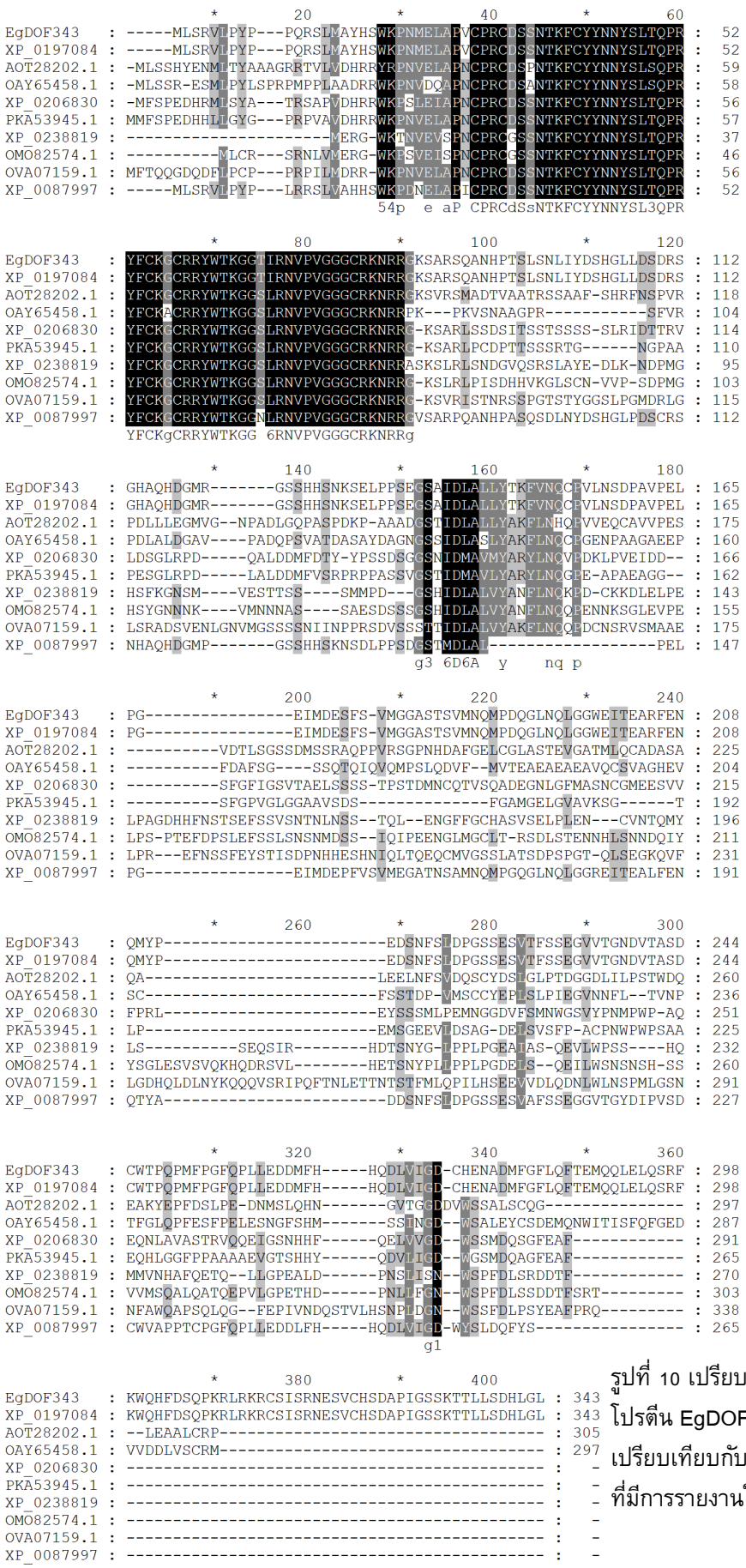
รูปที่ 7 แสดงการเปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgDof300 จากปาล์มน้ำมันเปรียบเทียบกับโปรตีน DOF ที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI



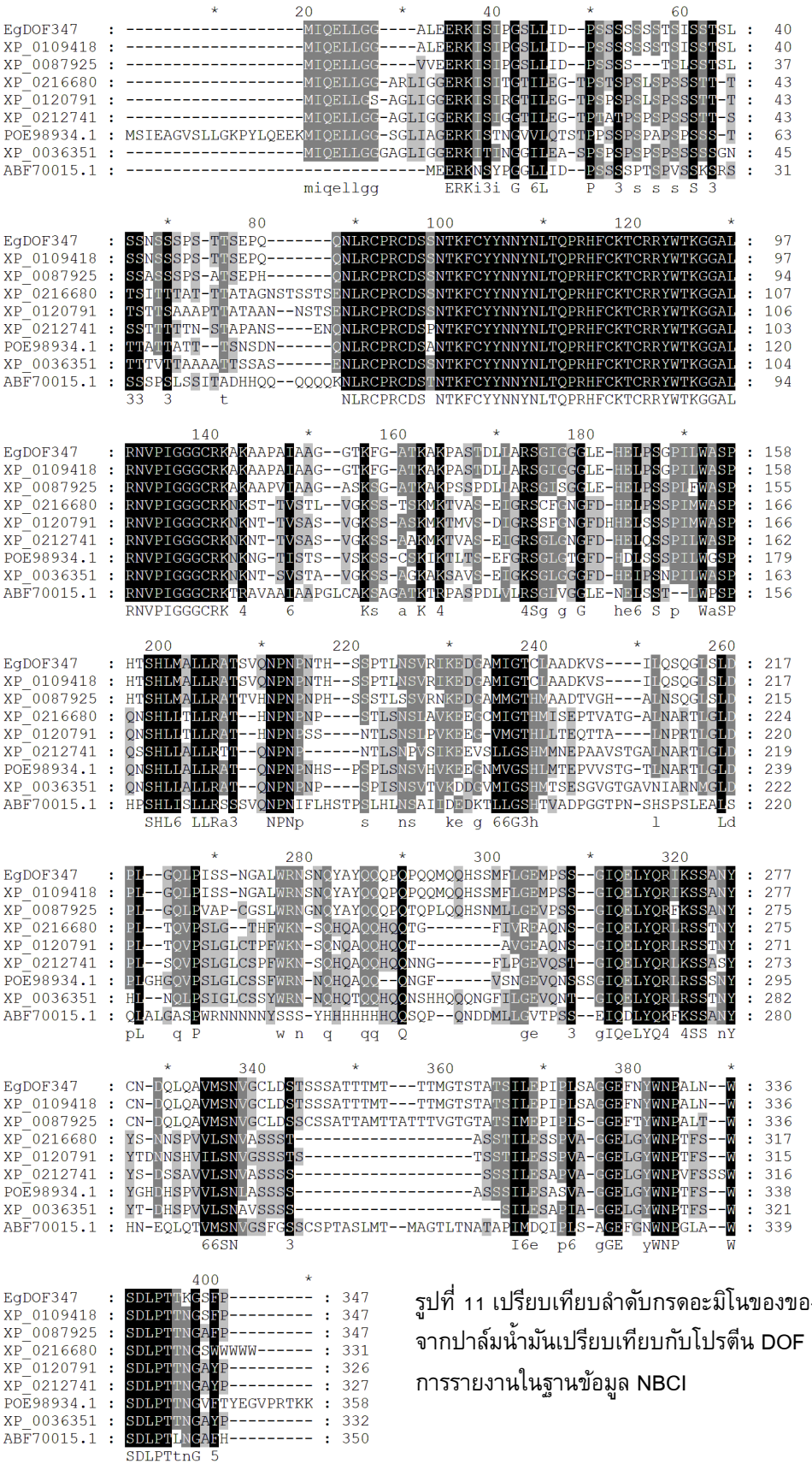
รูปที่ 8 เปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgDOF 314 จากปาล์มน้ำมันเปรียบเทียบกับโปรตีน DOF ที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI



รูปที่ 9 เปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgDOF 318 จากปาล์มน้ำมันเปรียบเทียบกับโปรตีน DOF ที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI



รูปที่ 10 เปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgDOF 343 จากปาล์มน้ำมัน เปรียบเทียบกับโปรตีน DOF ที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI



รูปที่ 11 เปรียบเทียบลำดับกรดอะมิโนของของโปรตีน EgDOF 347 จากปาล์มน้ำมันเปรียบเทียบกับโปรตีน DOF ที่พบในพืชชนิดอื่นที่มีการรายงานในฐานข้อมูล NBCI

เมื่อศึกษาบริเวณที่ทำหน้าที่ภายในกลุ่มโปรตีน EgDof จากปาล์มน้ำมันด้วยโปรแกรม PROSITE พบว่าโปรตีน EgDof300 พบบริเวณสำคัญที่แสดงสัญลักษณ์ของกลุ่มโปรตีน DOF เพียงหนึ่งตำแหน่งคือ ZF_DOF_2 Zinc finger Dof-type profile ตรงตำแหน่งของลำดับกรดอะมิโนที่ 30-84 คือ PNCPRYDSSNTKFCYY NNYSLSQPRYFCKGCRRYWTKGGSLRNVPVGGGCRKNRR เท่านั้นในขณะที่โปรตีน EgDof จากปาล์มน้ำมันอีก 4 ชนิดพบทั้งตำแหน่งอนุรักษ์ของ Zinc finger Dof-type signature และ Zinc finger Dof-type profile โดยEgDof314 พบบริเวณสำคัญที่แสดงสัญลักษณ์ของ ZF_DOF_1 Zinc finger Dof-type signature ในตำแหน่งของลำดับกรดอะมิโนที่ 59-95 (CprCdssntkfCyynnynltqprhfCktCrRyWTkGG) และพบนอกจากยังพบแบบแผนแสดงลักษณะของโปรตีน ZF_DOF_2 Zinc finger Dof-type profile ในตำแหน่งของลำดับกรดอะมิโนที่ 57-111 คือ LRCPRCDSSNTKFCYYNNYNLTQPRHFCKTCRRYWTKGGALRNV-PIGGGCRKAKA โปรตีน EgDof318 พบบริเวณสำคัญที่แสดงสัญลักษณ์ของ ZF_DOF_1 Zinc finger Dof-type signature ในตำแหน่งของลำดับกรดอะมิโนที่ 46-82 (CprCdstntkfCyynnynlsqprhfCkaCrRyWTkGG) และพบแบบแผนแสดงลักษณะของโปรตีน ZF_DOF_2 Zinc finger Dof-type profile ในตำแหน่งของลำดับกรดอะมิโนที่ 44-98 คือ VKCPRCDSTNTKFCYYNNYNLSQPRHFCKACRRYWTKGGVLRNVPVGGGCRKTKR โปรตีน EgDof343 พบบริเวณสำคัญที่แสดงสัญลักษณ์ของ ZF_DOF_1 Zinc finger Dof-type signature ในตำแหน่งของลำดับกรดอะมิโนที่ 30-66 (CprCdssntkfCyynnysltqpryfCkgCrRyWTkGG) และแบบแผนแสดงลักษณะของโปรตีน ZF_DOF_2 Zinc finger Dof-type profile ในตำแหน่งของลำดับกรดอะมิโนที่ 28-82 คือ PVCPRCDSSNTKFCYYNNYSLTQPRYFCKGCRRYWTKGGTIRNVPVGGGCRKNRR และสุดท้ายโปรตีน EgDof347 พบบริเวณสำคัญที่แสดงสัญลักษณ์ของ ZF_DOF_1 Zinc finger Dof-type signature ในตำแหน่งของลำดับกรดอะมิโนที่ 59 – 95 (CprCdssntkfCyynnynltqprhfCktCrRyWTkGG) และแบบแผนแสดงลักษณะของโปรตีน ZF_DOF_2 Zinc finger Dof-type profile ในตำแหน่งของลำดับกรดอะมิโนที่ 55-111 คือ LRCPRCDSSNTKFCYYNNYNLTQPRHFCKTCRRYWTKGGALRNVPIGGGCRKAKA การที่โปรตีน EgDof300ไม่พบ ZF_DOF_1 Zinc finger Dof-type signature อาจจะเป็นส่งให้โปรตีน EgDof300 มีหน้าที่การทำงานที่แตกต่างไปจากกลุ่มโปรตีน EgDof ที่พบจากปาล์มน้ำมัน

3.2 ศึกษาแบบแผนระดับการแสดงออกของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ ด้วยเทคนิค RT-PCR

ศึกษาระดับการแสดงออกยีน *EgNF-YA, EgNF-YB1 , EgNF-YB3* และยีน *EgNF-YC* ในตัวอย่างเนื้อเยื่อส่วนต่างๆ ของต้นปาล์มน้ำมันด้วยวิธี semi-quantitative RT-PCR จาก Total RNA 500 นาโนกรัม ด้วยคู่ไพรเมอร์ที่จำเพาะกับยีน *EgNFY* โดยมียีนมาตรฐาน

(internal control) เทียบคือยีน *18S rDNA* พบว่าการแสดงออกของยีน *EgNF-YA, EgNF-YB1 , EgNF-YB3* และยีน *EgNF-YC* นั้นมีการแสดงออกในทุกชิ้นส่วนที่นำมาศึกษายกเว้น ยีน *EgNF-YA* เท่านั้นที่ไม่พบว่ามีการแสดงออกของยีนในเนื้อเยื่อชั้นกลางของผล (mesocarp), เนื้อเมล็ดในของผล (kernel) อายุ 20 สัปดาห์หลังการติดดอกจากต้นปาล์มที่มีอายุ 7 ปีเลยและมีปริมาณระดับการแสดงออกของยีนในปริมาณน้อยที่สุดเมื่อเทียบกับยีน *EgNF-YB1 , EgNF-YB3* และ *EgNF-YC* ในทุกเนื้อเยื่อที่มีการแสดงออกดังรูปที่ 12 ทั้งนี้พบว่ามีปริมาณการแสดงออกของยีน *EgNFY-B1* ในปริมาณมากที่สุดในเนื้อเยื่อที่สะสมน้ำมันโดยมีค่าการแสดงออกสูงที่สุดในเนื้อเยื่อชั้นกลาง ซึ่งเป็นบริเวณที่มีที่พบน้ำมันมากที่สุด จึงคาดว่า *EgNF-YB1* มีความสัมพันธ์กับปริมาณน้ำมันแสดงผลดังรูปที่ 12 ต่อมาเมื่อศึกษาการแสดงออกของยีน *EgNF-YA, EgNF-YB1 , EgNF-YB3* และยีน *EgNF-YC* จากใบปาล์มน้ำมันของต้นที่มีผลผลิตสูง (n=3) (จำนวนผลของทะลายสูงและมีจำนวนทะลายมาก) (kg/palm) และต้นที่ผลผลิตต่ำ (n=3) (จำนวนผลของทะลายต่ำและมีจำนวนทะลายน้อย) (kg/palm) เปรียบเทียบกับ *18S rDNA* gene เพื่อศึกษาสัมพันธ์ของยีน *EgNFY* กับการสะสมน้ำมันในต้นปาล์มน้ำมันพบว่า *EgNFY-B1* มีค่าการแสดงออกสูงในกลุ่มของใบต้นที่ให้ผลผลิตสูง (High Leave) มากที่สุดเมื่อเทียบกับ *EgNF-YA, EgNF-YB3* และยีน *EgNF-YC* คือมีค่าเท่ากับ 1.01 ± 0.12 ในขณะที่กลุ่มต้นที่ให้ผลผลิตต่ำ (Low Leave) นั้นยีน *EgNFY-B1* มีค่าการแสดงออกต่ำกว่าคือมีค่าเท่ากับ 0.28 ± 0.03 ซึ่งเป็นค่าการแสดงออกที่มีความแตกต่างอย่างมีนัยสำคัญ ดังรูปที่ 12B ซึ่งจากผลการศึกษาที่ได้นี้คาดว่าระดับการแสดงออกของ *EgNFY-B1* นั้นน่าจะมีผลควบคุมปริมาณน้ำมันและการสะสมน้ำมันในต้นปาล์มน้ำมัน

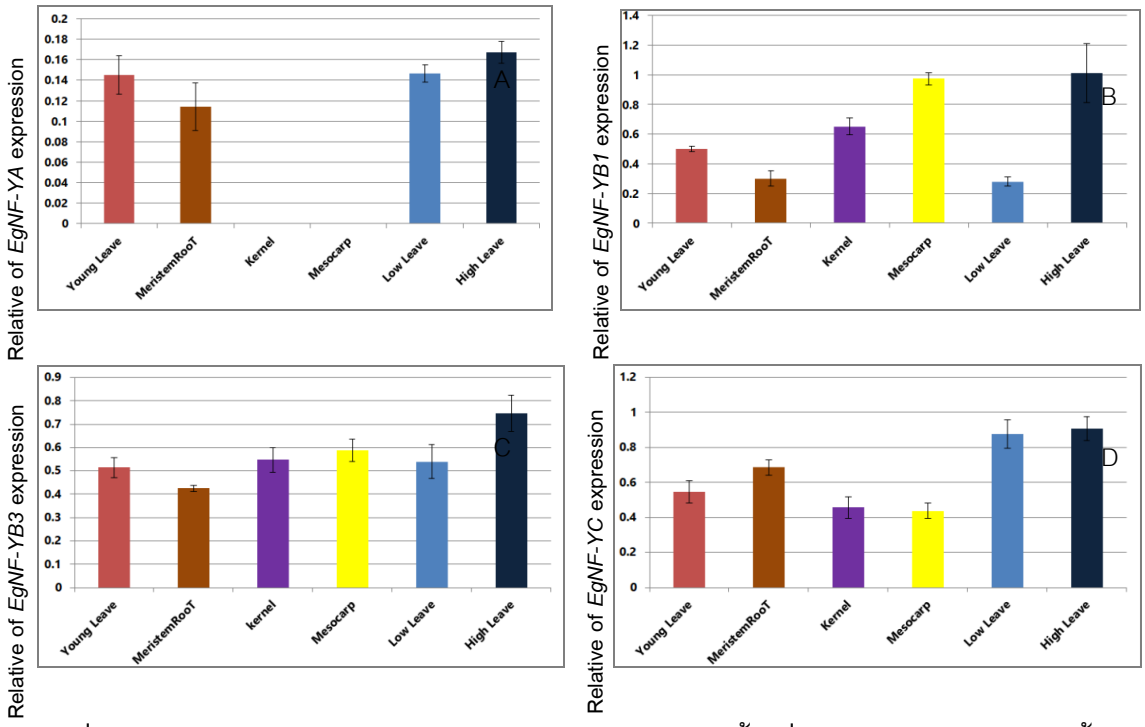
ศึกษาระดับการแสดงออกของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ของยีน *EgDof300, EgDof314 EgDof318, EgDof343* และ *EgDof347* ด้วยวิธี semi-quantitative RT-PCR จาก Total RNA 500 นาโนกรัม ด้วยคู่ไพรเมอร์ที่จำเพาะกับยีน *Eg EgDof* โดยมียีนมาตรฐาน (internal control) เทียบคือยีน *18S rDNA* พบว่ามีเพียงยีน *EgDof314 EgDof318* และ *EgDof343* ที่มีการแสดงออกของยีนในเนื้อเยื่อส่วนชนิด vegetative ของต้นปาล์มดังรูปที่ 13 โดยยีน *EgDof300,* และ *EgDof347* นั้นไม่พบว่ามีการแสดงออกของยีนในเนื้อเยื่อส่วนชนิด vegetative ต่อมาเมื่อศึกษาการแสดงออกของยีนในเนื้อเยื่อชั้นกลางของผล (mesocarp) อายุ 4-20 สัปดาห์หลังการติดดอกจากต้นปาล์มที่มีอายุ 7 ปีพบว่ามีการแสดงออกของยีนเฉพาะยีน *EgDof300 EgDof318*และ *EgDof347* โดยปริมาณระดับการแสดงออกของยีน *EgDof300,*และ *EgDof347* มีปริมาณใกล้เคียงกันในทุกระยะของผลปาล์มน้ำมัน ในขณะที่ *EgDof318* มีการแสดงออกลดลงเมื่ออายุของผลปาล์มน้ำมันผลปาล์มเพิ่มขึ้น ดังรูปที่ 14

เมื่อศึกษาการแสดงออกของยีน *EgDof318* จากใบปาล์มน้ำมันของต้นที่มีผลผลิตสูง (n=3) (จำนวนผลของทะลายสูงและมีจำนวนทะลายมาก) (kg/palm) และต้นที่ผลผลิตต่ำ (n=3) (จำนวนผลของทะลายต่ำและมีจำนวนทะลายน้อย) (kg/palm) เปรียบเทียบกับ *18S rDNA* gene เพื่อศึกษาสัมพันธ์ของยีน *EgDof318* กับการสะสมน้ำมันในต้นปาล์มน้ำมันพบว่า *EgDof318* มีค่าการแสดงออกสูงในกลุ่มของใบต้นที่ให้ผลผลิตต่ำ (Low Leave) มากกว่าที่สุด

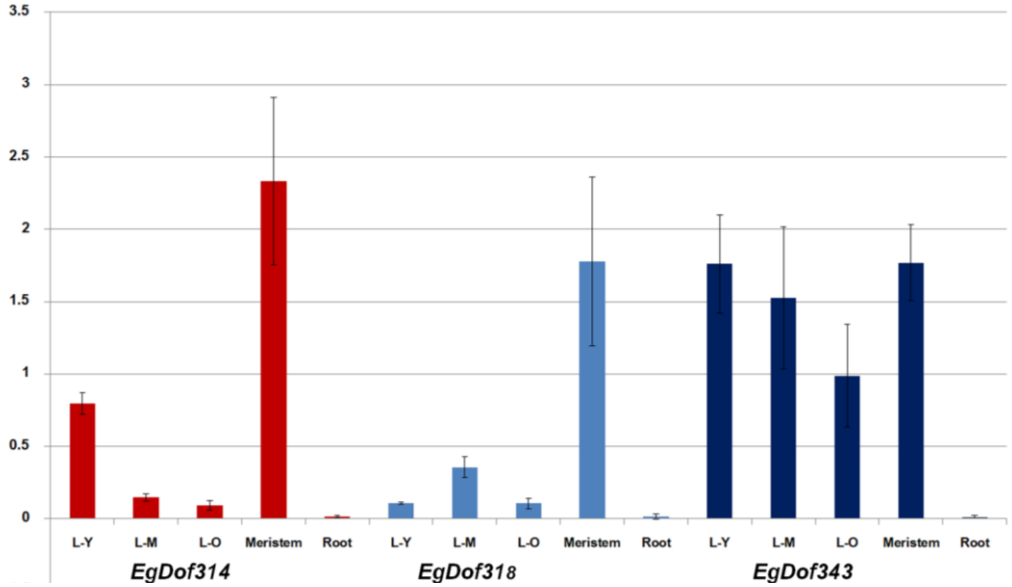
เมื่อเทียบกลุ่มต้นที่ให้ผลผลิตสูง (High Leave) คือมีค่าเท่ากับ 0.95 ± 0.18 และ 0.64 ± 0.06 ตามลำดับซึ่งเป็นค่าการแสดงออกที่มีความแตกต่างอย่างมีนัยสำคัญ ดังรูปที่15 ซึ่งจากผลการศึกษาที่ได้นี้คาดว่าระดับการแสดงออกของ *EgDof318* นั้นน่าจะมีการควบคุมส่งผลต่อปริมาณน้ำมันและการสะสมน้ำมันในต้นปาล์มน้ำมัน

3.3 ผลการทำนายโครงสร้างตติยภูมิของโปรตีน EgNFY และ EgDof จากปาล์มน้ำมัน

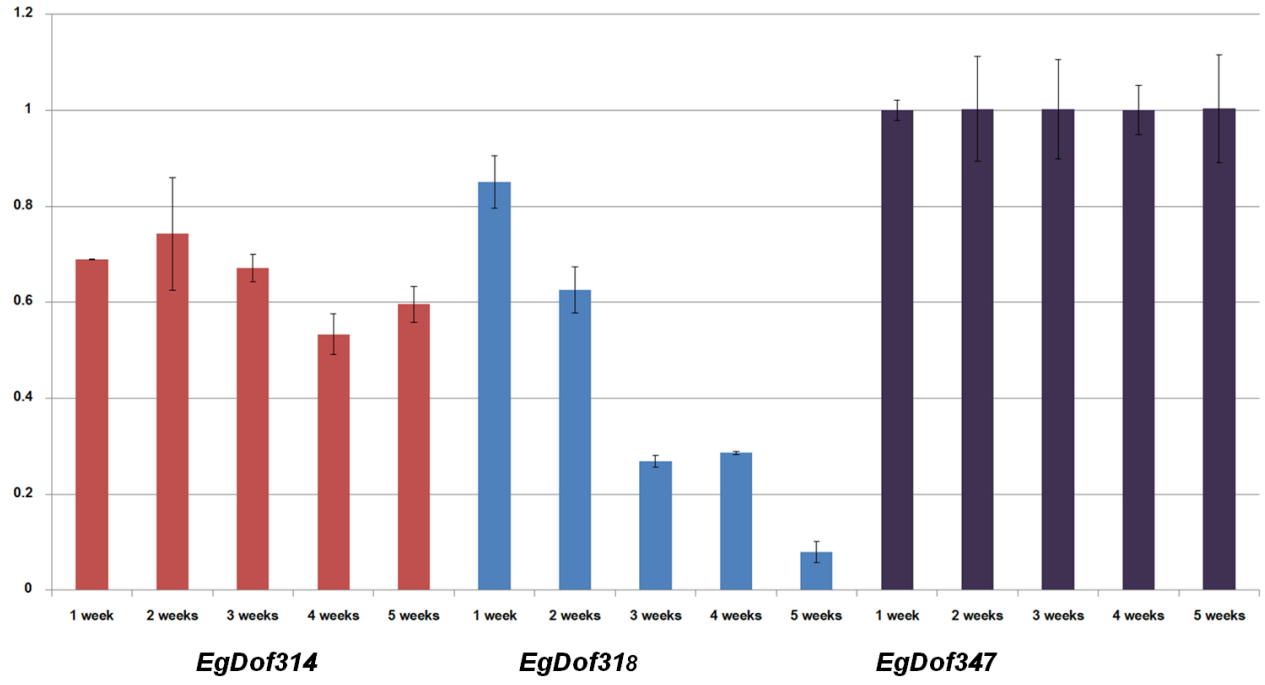
จากการทำนายโครงสร้างตติยภูมิและตำแหน่งการแทรกจับกับ Ligand ต่างๆ ของโปรตีน EgNFY และ EgDof โดยใช้โปรแกรม I-TASSER พบว่าโปรตีน EgNFY-A, EgDof300, EgDof314, EgDof343 และ EgDof347 ไม่มีตำแหน่งการจับกับ DNA ดังแสดงในตารางที่ 3 และ 4 ในขณะที่โปรตีน EgNFY-B3, EgNFY-C และ EgDof318 มีตำแหน่งการจับกับ DNA จำนวน 2 ตำแหน่ง ในขณะที่โปรตีน EgNFY-B1 มีตำแหน่งการจับจับกับ DNA สูงที่สุดคือ จำนวน 2 ตำแหน่ง ดังแสดงในรูปที่ 16 และ 17



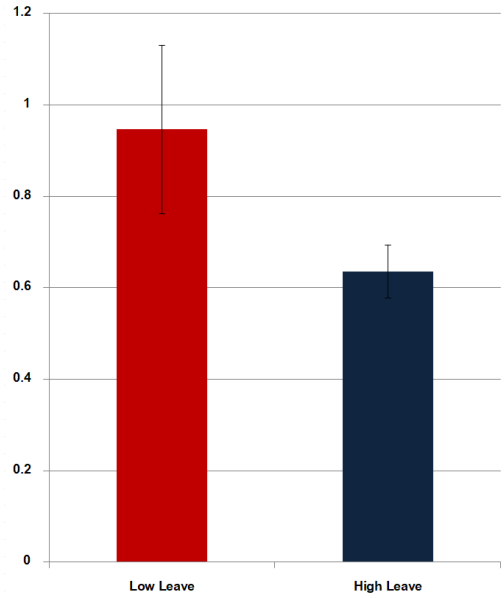
รูปที่ 12 แสดงระดับการแสดงออกของยีน *EgNFY* ในเนื้อเยื่อส่วนต่างๆ ของปาล์มน้ำมัน วิเคราะห์ด้วยวิธี Semi-quantitative RT-PCR โดยใช้โปรแกรม Quantity One เปรียบเทียบกับ 18S rDNA ที่ใช้เป็นตัวควบคุม



รูปที่ 13 แสดงระดับการแสดงออกของยีน *EgDof314* *EgDof318* และ *EgDof343* ในเนื้อเยื่อส่วนต่างๆ ของปาล์มน้ำมันวิเคราะห์ด้วยวิธี Semi-quantitative RT-PCR โดยใช้โปรแกรม Quantity One เปรียบเทียบกับ 18S rDNA ที่ใช้เป็นตัวควบคุม



รูปที่ 14 แสดงระดับการแสดงออกของยีน *EgDof314* *EgDof318* และ *EgDof343* ในเนื้อเยื่อส่วนต่างๆ ของปาล์มน้ำมันวิเคราะห์ด้วยวิธี Semi-quantitative RT-PCR โดยใช้โปรแกรม Quantity One เปรียบเทียบกับ 18S rDNA ที่ใช้เป็นตัวควบคุม



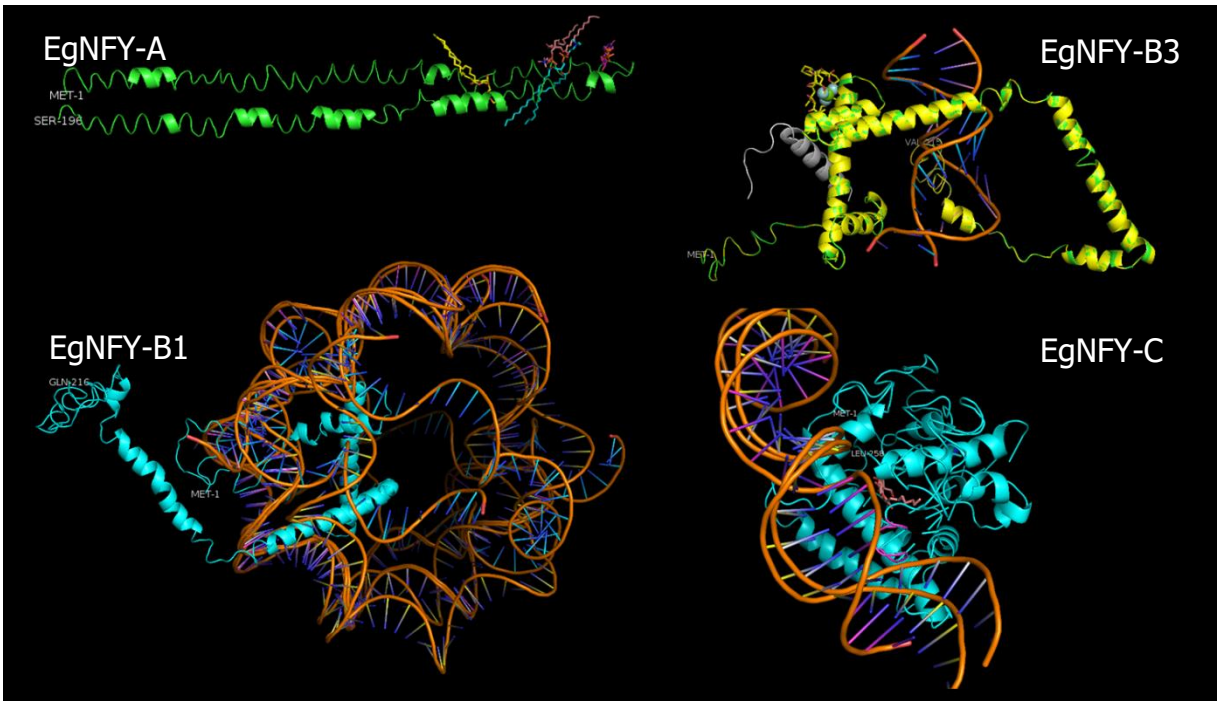
รูปที่ 15 แสดงระดับการแสดงออกของยีน *EgDof31s* ในเนื้อเยื่อใบปาล์มน้ำมันของต้นที่มีผลผลิตสูง (High Leave) และต้นที่ผลผลิตต่ำ(Low Leave) ด้วยวิธี Semi-quantitative RT-PCR และวิเคราะห์ด้วยโปรแกรม Quantity One เปรียบเทียบกับ 18S rDNA ที่ใช้เป็นตัวควบคุม

ตารางที่ 3 แสดงผลทำนายตำแหน่งในการแทรกจับกับ DNA ของโปรตีน EgNFY จากปาล์มน้ำมัน

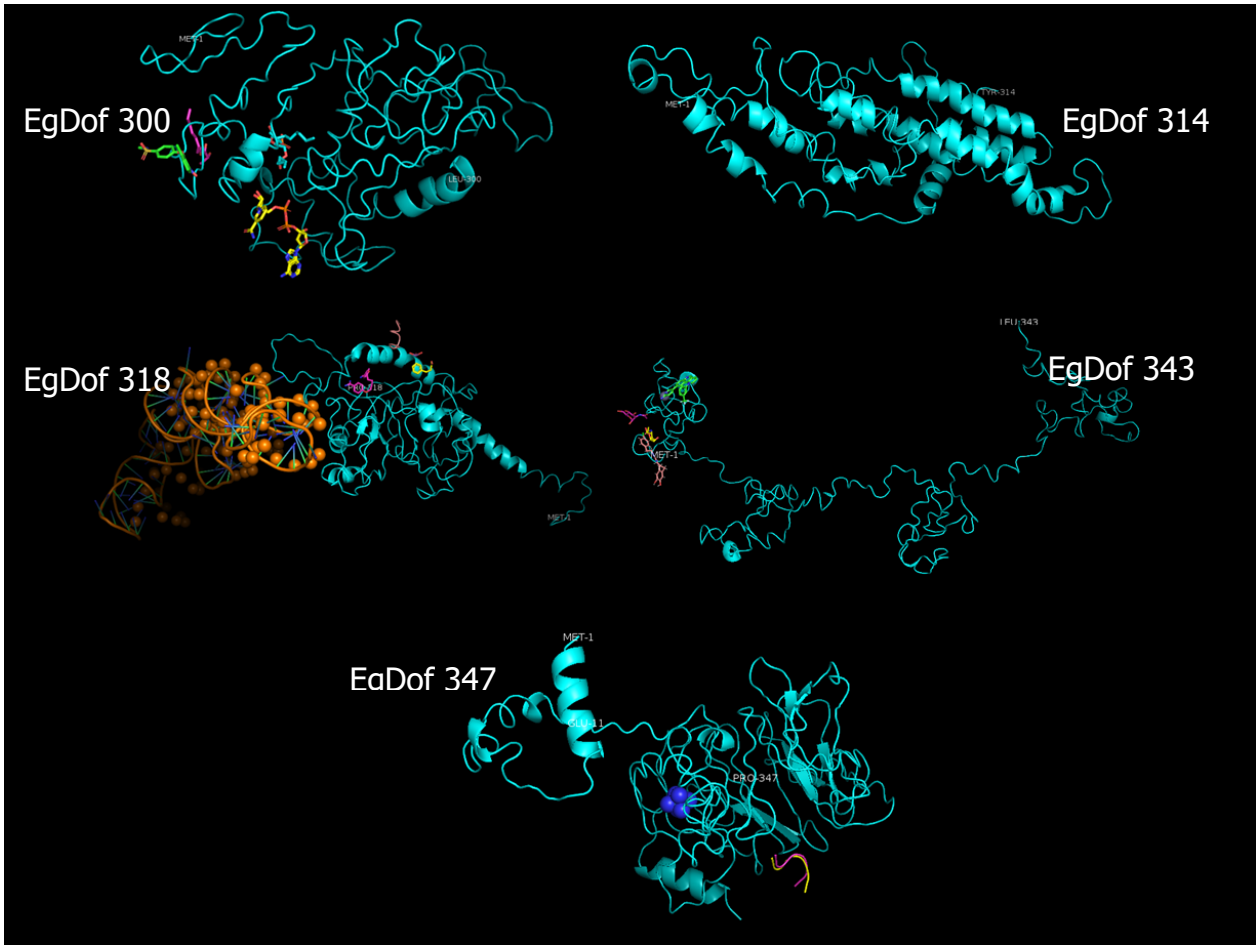
Gene	ORF length	Deduced protein length (aa)	Theoretical protein mol wt. (Da)	Theoretical protein (pI)	Predicted DNA binding Domains
EgNFY-A	705	234	25507.11	8.89	None
EgNFY-B1	651	216	23549.13	5.45	5
EgNFY-B3	648	215	22467.85	6.30	2
EgNFY-C	714	237	25777.47	5.37	2

ตารางที่ 4 แสดงผลทำนายตำแหน่งในการแทรกจับกับ DNAของโปรตีน EgDof จากปาล์มน้ำมัน

Gene	ORF length	Deduced protein length (aa)	Theoretical protein mol wt. (Da)	Theoretical protein (pl)	Predicted DNA binding Domains
EgDof300	903	300	32875.61	4.94	None
EgDof314	945	314	34068.25	8.93	None
EgDof318	957	318	32799.16	8.75	2
EgDof343	1032	343	38387.90	5.99	None
EgDof347	1044	347	36930.42	8.83	None



รูปที่ 16 แสดงการทำนายหน้าที่การทำงานโครงสร้างสามมิติของโปรตีน EgNFY-A EgNFY-B1 EgNFY-B3 และ EgNFY-C โดยโปรแกรม I-TASSER และแสดงผลด้วยโปรแกรม PyMOL แสดงสมบัติของทรานสคริปชั่นแฟกเตอร์และความสามารถในการแทรกจับกับ Ligand ต่างๆ ของโปรตีน

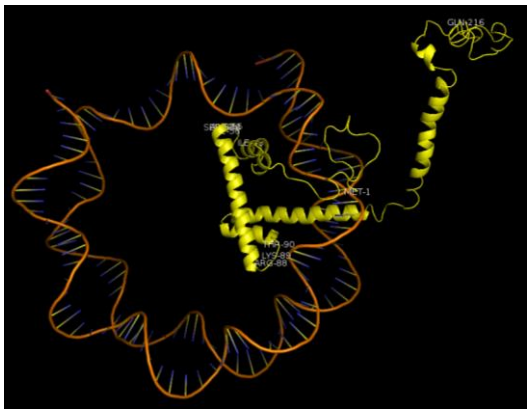


รูปที่ 17 แสดงการทำนายหน้าที่การทำงานโครงสร้างสามมิติของโปรตีน EgDof300, EgDof314 EgDof318, EgDof343 และ EgDof347 โดยโปรแกรม I-TASSER และแสดงผลด้วยโปรแกรม PyMOL แสดงสมบัติของทรานสคริปชั่นแฟกเตอร์และความสามารถในการแทรกจับกับ Ligand ต่างๆ ของโปรตีน

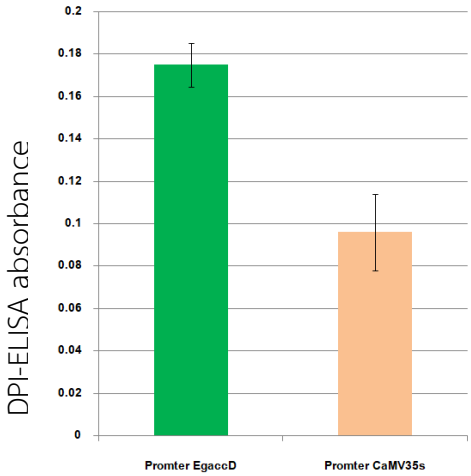
3.4 ศึกษาความจำเพาะของโปรตีน EgNFY-B1 ทรานสคริปชั่นแฟคเตอร์ในการจับกับโปรโมเตอร์ของ *EgaccD* จากปาล์มน้ำมันปาล์มน้ำมันด้วยวิธี DNA-Protein Interaction – ELISA (DPI-ELISA)

เมื่อทำการโคลนยีน *EgNFY-B1* และ *EgNFY-B3* เส้นเต็มเข้าเวคเตอร์ pET-28A เพื่อผลิตโปรตีนลูกผสมดังกล่าวพบว่าไม่สามารถกระตุ้นให้มีการผลิตโปรตีนลูกผสม EgNFY-B1 และ EgNFY-B3 ในแบคทีเรีย *E. coli* สายพันธุ์ BL21 (DE3) ได้เมื่อกระตุ้นด้วย 0.1 mM IPTG ดังนั้นจึงได้ทดสอบผลการผลิตโปรตีนลูกผสม EgNFY-B1 เพียงบางส่วนในตำแหน่งที่ได้จากการวิเคราะห์ข้อมูลโครงสร้างตติยภูมิของ EgNFY-B1 ด้วยโปรแกรมในฐานข้อมูล I-TASSER พบว่ากรดอะมิโนของ EgNFY-B1 ตำแหน่งที่ 43, 55, 56, 57, 58, 88, 89 และ 90 ซึ่งเป็นตำแหน่งที่มีการจับกับสายดีเอ็นเอผลดังรูปที่ 18 ซึ่งเมื่อทำการเข้าโคลนที่มีชิ้นส่วนยีน *EgNFY-B1* ขนาด 183 คู่เบส เวคเตอร์ pET-28A เพื่อผลิตโปรตีนลูกผสม ในแบคทีเรีย *E. coli* สายพันธุ์

BL21 (DE3) ได้เมื่อกระตุ้นด้วย 0.1 mM IPTG นั้นพบว่าสามารถผลิตและทำบริสุทธิ์โปรตีนขนาดประมาณ 7 กิโลดาลตันได้ ดังนั้นจึงได้ทดสอบสมบัติของทรานสคริปชั่นแฟกเตอร์ของโปรตีน EgNFY-B1 ด้วยวิธี DPI-ELISA ซึ่งเป็นวิธีการตรวจสอบการจับกันระหว่างโปรตีนกับดีเอ็นเอภายในหลอดทดลอง (*in vitro*) โดยการวิเคราะห์คุณภาพการจับกันระหว่างโปรตีนลูกผสม EgNFY-B1 ในบริเวณครอบคลุมตำแหน่งอนุรักษ์ที่ 43, 55, 56, 57, 58, 88, 89 และ 90 กับโปรโมเตอร์ EgaccD ซึ่งมีตำแหน่ง CAAT box ทั้งหมด 6 ตำแหน่งและ โดยใช้โปรโมเตอร์ CaMV35s ที่ให้ผลจากการวิเคราะห์โปรโมเตอร์มีจำนวน CAAT box 1 ตำแหน่ง เป็นชุดควบคุม ซึ่งผลการทดสอบพบว่าของทรานสคริปชั่นแฟกเตอร์ของโปรตีน EgNFY-B1 สามารถจับกับโปรโมเตอร์ *EgaccD* ได้ดีกว่าโปรโมเตอร์ CaMV35s ซึ่งผลการทดสอบโดยผลการทำปฏิสัมพันธ์กันสอดคล้องกับจำนวน CAAT box บนโปรโมเตอร์ *EgaccD* ซึ่งเป็นโปรโมเตอร์ที่เกี่ยวข้องกับการผลิตน้ำมันดังแสดงในรูปที่ 19



รูปที่ 18 แสดงการทำนายหน้าที่การทำงานโครงสร้างสามมิติของโปรตีน EgNFY-B1 (สีเหลือง) โดยโปรแกรม I-TASSER แสดงสมบัติของทรานสคริปชั่นแฟกเตอร์ที่สามารถแทรกจับกับแบบจำลองสายดีเอ็นเอได้ในตำแหน่งที่ 43, 55, 56, 57, 58, 88, 89 และ 90



รูปที่ 19 แสดงความจำเพาะของโปรตีน EgNFY-B1 ทรานสคริปชั่นแฟคเตอร์ในการจับกับโปรโมเตอร์ของ *EgaccD*

สรุปผลการทดลอง

1. ได้ทำการโคลนนิวคลีโอไทด์ของยีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y ได้แก่ยีน *EgNF-YA, EgNF-YB1, EgNF-YB3* และยีน *EgNF-YC* และโคลนนิวคลีโอไทด์ของยีนกลุ่มทรานสคริปชั่นแฟคเตอร์ Dofs ได้แก่ยีน *EgDof300, EgDof314 EgDof318, EgDof343* และ *EgDof347* จากเนื้อเยื่อใบและผลอ่อนปาล์มน้ำมัน *E. guineensis* อายุ 2 เดือน
2. ได้ทำการวิเคราะห์ลำดับกรดอะมิโนของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ NF-Y พบว่าโปรตีน EgNFY-A พบแบบแผนแสดงลักษณะของโปรตีน NF-YA/HAP2 family profile ในตำแหน่งของลำดับกรดอะมิโนที่ 63-123 และพบตำแหน่งอนุรักษ์ Subunit association domain (SAD) Motif ในการจับกับโปรตีนหน่วยย่อยอื่นได้ โปรตีน EgNFY-B1 พบบริเวณสำคัญที่แสดงสัญลักษณ์ของ NF-YB/HAP3 subunit signature ในตำแหน่งของลำดับกรดอะมิโนที่ 67 – 83 ซึ่งเป็นตำแหน่งการทำงานของทรานสคริปชั่นแฟกเตอร์ที่มีการจับจำเพาะกับ promoter บนสายดีเอ็นเอบริเวณ CAAT box แต่โปรตีน EgNFY-B3 และ EgNFY-C แต่พบว่าในตำแหน่งของลำดับกรดอะมิโนที่ 28-134 ของโปรตีน EgNFY-B3 แสดงแบบแผนความเป็น Histone-fold homologous superfamilies ซึ่งโปรตีนกลุ่มนี้จะทำหน้าหน้าควบคุมการแสดงของยีนที่เกี่ยวข้องผ่านการควบคุมการทำงานของโปรตีนฮิสโตนเช่นเดียวกับ EgNFY-C ซึ่งพบในตำแหน่งของลำดับกรดอะมิโนที่ 67-182 ที่แสดงแบบแผนความเป็น Histone-fold homologous superfamilies
3. ได้ทำการวิเคราะห์ลำดับกรดอะมิโนของโปรตีนกลุ่มทรานสคริปชั่นแฟคเตอร์ Dofs ได้แก่ยีน EgDof300, EgDof314 EgDof318, EgDof343 และ EgDof347 พบว่าโปรตีน EgDof300 พบบริเวณสำคัญที่แสดงสัญลักษณ์ของกลุ่มโปรตีน DOF เพียงหนึ่งตำแหน่งคือ ZF_DOF_2 Zinc finger Dof-type profile ในขณะที่โปรตีน EgDof314 EgDof318, EgDof343 และ *EgDof347* พบทั้งตำแหน่งอนุรักษ์ของ ZF_DOF_1 หรือ ZF_DOF_2 Zinc finger Dof-type signature และ Zinc finger Dof-type profile
4. เมื่อทำนายความสามารถในการจับกับดีเอ็นเอพบว่า โปรตีน EgNF-YA ไม่จับกับดีเอ็นเอได้ ในขณะที่โปรตีน EgNF-YB1, EgNF-YB3 และ EgNF-YC มีบริเวณของลำดับกรดอะมิโนที่สามารถจับกับดีเอ็นเอได้ โดยโปรตีน EgNF-YB3 และ EgNF-YC มีตำแหน่งที่จับกับดีเอ็นเอได้ 2 ตำแหน่ง ในขณะที่ โปรตีน EgNF-YB1ตำแหน่งที่จับกับดีเอ็นเอได้ 5 ตำแหน่ง
5. เมื่อทำนายความสามารถในการจับกับดีเอ็นเอพบว่า โปรตีน EgDof300, EgDof314 EgDof343 และ EgDof347 ไม่จับกับดีเอ็นเอได้ ในขณะที่มีเพียงโปรตีน EgDof318 ที่จับกับดีเอ็นเอได้โดยตำแหน่งที่จับกับดีเอ็นเอได้ 2 ตำแหน่ง

6. ผลระดับการแสดงออกของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ ได้แก่ *EgNF-YA, EgNF-YB1 , EgNF-YB3* และยีน *EgNF-YC* ในตัวอย่างเนื้อเยื่อส่วนต่าง ๆ ของต้นปาล์มน้ำมัน พบว่าการแสดงออกของยีน *EgNF-YA, EgNF-YB1 , EgNF-YB3* และยีน *EgNF-YC* นั้นมีการแสดงออกในทุกชิ้นส่วนที่นำมาศึกษายกเว้นยีน *EgNF-YA* เท่านั้นที่ไม่พบว่ามีการแสดงออกของยีนในเนื้อเยื่อชั้นกลางของผล และเนื้อเมล็ดในของผลอายุ 20 สัปดาห์หลังการติดดอกจากต้นปาล์มที่มีอายุ 7 ปี นอกจากยีน *EgNFY-B1* ในปริมาณมากที่สุดในเนื้อเยื่อที่สะสมน้ำมันโดยมีค่าการแสดงออกสูงที่สุดในเนื้อเยื่อชั้นกลาง และมีค่าการแสดงออกสูงในกลุ่มของใบต้นที่ให้ผลผลิตสูง (High Leave) มากที่สุดเมื่อเทียบกับ *EgNF-YA, EgNF-YB3* และยีน *EgNF-YC* ผลการศึกษาที่ได้นี้คาดว่าระดับการแสดงออกของ *EgNFY-B1* นั้นน่าจะมีผลควบคุมปริมาณน้ำมันและการสะสมน้ำมันในต้นปาล์มน้ำมัน
7. ผลระดับการแสดงออกของกลุ่มยีนทรานสคริปชั่นแฟคเตอร์ ได้แก่ ยีน *EgDof300, EgDof314 EgDof318, EgDof343* และ *EgDof347* ในตัวอย่างเนื้อเยื่อส่วนต่าง ๆ ของต้นปาล์มน้ำมันพบว่ามีเพียงยีน *EgDof314 EgDof318* และ *EgDof343* ในขณะที่ยีน *EgDof300,*และ *EgDof347* นั้นไม่พบว่ามีการแสดงออกของยีนในเนื้อเยื่อส่วนชนิด vegetative และพบว่ามีการแสดงออกของยีนเฉพาะ *EgDof300 EgDof318,*และ *EgDof343* ในเนื้อเยื่อชั้นกลางของผล สัปดาห์หลังการติดดอกจากต้นปาล์มที่มีอายุ 7 ปี โดยปริมาณระดับการแสดงออกของยีน *EgDof300,*และ *EgDof347* มีปริมาณใกล้เคียงกันในทุกระยะของผลปาล์มน้ำมันในขณะที่ *EgDof318* มีการแสดงออกลดลงเมื่ออายุของผลปาล์มน้ำมันผลปาล์มเพิ่มขึ้นคาดว่า *EgDof318* นอกจากพบว่ายีน *EgDof318* มีค่าการแสดงออกสูงในกลุ่มของใบต้นที่ให้ผลผลิตต่ำมากกว่ากลุ่มของใบต้นที่ให้ผลผลิตสูง
8. พบว่าวิเคราะห์คุณภาพการจับกันระหว่างโปรตีนลูกผสม EgNFY-B1 กับโปรโมเตอร์ของ *EgaccD* จากปาล์มน้ำมันซึ่งมีตำแหน่ง CAAT box ทั้งหมด 6 ตำแหน่งด้วยวิธี DNA-Protein Interaction – ELISA (DPI-ELISA) พบว่าของทรานสคริปชั่นแฟกเตอร์ของโปรตีน EgNFY-B1 สามารถจับกับโปรโมเตอร์ *EgaccD* ได้ดีกว่าโปรโมเตอร์ CaMV35s ซึ่งเป็นชุดควบคุม

ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

.

Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

1. ผลงานตามที่คาดไว้ในสัญญาโครงการ
   เตรียม manuscript เพื่อส่งตีพิมพ์ 2 รายการ
   - Molecular cloning and expression analysis of *EgNFYB1* genes from oil palm (*Elaeis guineensis* Jacq.)
   - Characterize the expression profile of NF-Y and Dofs transcription factor genes involved in fatty acid and lipid biosynthesis (อยู่ในระหว่างการเขียน)
2. การนำผลงานวิจัยไปใช้ประโยชน์
   - มีการนำผลงานวิจัยไปใช้ประโยชน์เชิงสาธารณะ
3. อื่นๆ (เช่น ผลงานตีพิมพ์ในวารสารวิชาการในประเทศ การเสนอผลงานในที่ประชุมวิชาการ หนังสือ การจดสิทธิบัตร)
   - เตรียมยื่นผลงานวิจัยไปการจดสิทธิบัตร 1 ฉบับ

ภาคผนวก

Title: Molecular cloning and expression analysis of *EgNFYB1* genes from oil palm (*Elaeis guineensis* Jacq.)

Authors: Alisa Nakkaew[1,2*],Kanokon Petsean[2] , Amornrat Phongdara[1]

Addresses: [1] Center for Genomic and Bioinformatics Research, Faculty of Science, Prince of Songkla University, Hat-Yai, Songkhla 90110, Thailand

[2] Department of Molecular Biotechnology and Bioinformatics, Faculty of Science, Prince of Songkla University, Hat-Yai, Songkhla 90110, Thailand

*Correspondence to: Dr. Alisa Nakkaew, Email: joy_alisa@yahoo.com/alisa.n@psu.ac.th

ABSTRACT

EgNFYB1 is transcription factor that involved in the oil synthesis in plants. In this study, *EgNFYB1* gene was isolated from mesocarp of oil palm (*Elaeis guineensis* Jacq.), contained 651 bp encoded 216 amino acid. The amino acid sequence analysis of *EgNFYB1* showed highly similarity compared with NFYB1 protein date palm. The secondary structure of *EgNFYB1* was predicted and showed 10 sites of helices and 11 sites of beta strand. Additionally, *EgNFYB1* encodes a HAP3 subunit of the CAAT-binding factor. The binding of *EgNFYB1* protein and *EgaccD* promoter was strongly interaction. *EgNFYB1* gene was highly expressed in mesocarp tissue and high yield leaves. Moreover, the unknown young leaves from different trees were analyzed for their level of expression of *EgNFYB1* considerable variations were observed. These result indicated that the *EgNFYB1* expression may be related to the ability of the palm tree to accumulate oil. Therefore, *EgNFYB1* might be used as a molecular marker during the breeding program for producing high yielding oil palms.

Introduction

Oil palm is the most productive oil crop which plays a key role in meeting current and future demand for vegetable oil. Global demand for plant oils is increasing rapidly and estimated to be doubled by 2030. In order to meet the growing demand, it is necessary to increase plant oil production without expanding plant acreage and improving oil yield. To speed up the process of crop improvement, one approach is by engineering of key

enzymes required for oil biosynthesis (Tester and Langridge 2010; Varshney et al., 2007). Because oil synthesis is a highly coordinated process that involves carbon metabolism through the FA synthesis pathway, the manipulation of a single gene may not be an efficient approach to significantly increase seed oil content (Thelen and Ohlrogge, 2002; Cahoon et al., 2007; Weselake et al., 2009). Plastid pyruvate kinase (PK), pyruvate dehydrogenase (PDH), and acetyl-CoA carboxylase are key enzymes for fatty acid synthesis and are regulated by transcription factors (Baud, et al., 2009; Baud and Lepinec, 2010). Transcription factors (TFs) are important regulators of gene expression that act by binding to plant-specific cis-regulatory elements in the promoter region, thereby activating or repressing the transcriptional rates of their target genes (Wray et al. 2003). The CCAAT box, a common *cis*-acting element in promoters and enhancer regions, exhibits a specific pattern in transcripts specific for several shoot cell types. The CCAAT box core motif is present in many genes that are co-expressed in specific leaf cell types and is flanked by a sequence specific for each leaf cell type, suggesting it has functional significance (Jiao, et al., 2009). The CCAAT-box is a cis-acting element that is often found in the gene promoters of higher eukaryotes and is essential for the transcription process (Mantovani, 1999). In most species, the CCAAT-binding factor, also known as Nuclear Factor Y (NF-Y) or Heme Activator Protein (HAP), is composed of three subunits, NF-YA (HAP2), NF-YB (HAP3), and NFYC (HAP5), and is present in all higher eukaryotes (reviewed by Laloum, et al., 2013). NF-Y members have been identified in several plant species, including model plants. To achieve higher oil content, molecular strategies were employed to identify regulators of plant embryo development, which have been found in several classes of genes involved in this process, such as NFYB1 (LEAFY COTYLEDON 1), L1L (LEAFY COTYLEDON1-LIKE), LEC2 (LEAFY COTYLEDON 2), and WRINKLED (Lotan et al. 1998; Kwong et al. 2003; Lee et al. 2003; Yazawa and Kamada 2007; Mu et al. 2008). The *NFYB1* gene encodes a transcription factor homolog belonging to the CCAAT box-binding factor HAP3 subunit of the CCAAT box-binding factor that binds specifically to DNA fragments containing CCAAT and CAAT sequences (Gusmaroli et al., 2001; Kusnetsov et al., 1999; Lotan et al., 1998; Yazawa and Kamada, 2007).

Previous work, the genes encoding the heteromeric acetyl-CoA carboxylase (ht-ACCase), were found to involve in regulation of the production of fatty acid and enhancing oil production especially *EgaccD*. Which, their promoter was composing of five CCAAT box-binding domains (Nakkaew, et al., 2008, Nakkaew, et al., 2014). This project aims to identify and functional characterize of the NF-Y gene family in oil palm.

That might regulate the expression profile of ht-ACCase and plays the important role in fatty acid and lipid synthesis pathways. Firstly, a subset of NF-Y family genes was cloned and analyzed using bioinformatics tools. Although, expression profiles of all NF-Y family genes were demonstrate. Furthermore, yeast one-hybrid and DNA protein interaction (DPI)-ELISA were estimated DNA binding specificity for defining CCAAT box-binding in oil palm gene promoters. Finally, over-expression of NF-Y family genes in transgenic callus and the expression of a subset of genes involved in fatty acid and lipid biosynthesis were analyzed. The information obtained from this study will provide important knowledge regarding the functions, characterization, and functional regulation of NF-Ys genes that will be the first step to do the metabolic engineering of the key enzymes required for oil biosynthesis and provides a practical approach for the genetic improvement of oil palm.

2. Materials and methods

2.1. Cloning and vector construction

Plant materials and total RNA isolation Mesocarp tissue was excised from *E. guineensis* fruits. Total RNA was extracted from 100 mg of mesocarp tissue using total RNA mini kit (Plant) (Geneaid) according to the manufacturer's instructions. The contaminated genomic DNA was removed by DNaseI treatment. The total RNA obtained was used as a template for cDNA synthesis.

2.2 Isolation of *EgNFYB1* cDNA clones

A 210 bp partial sequence *EgNFYB1* was selected from the EST library. A full-length cDNA was generated by rapid amplification of the cDNA ends (RACE) using the GeneRacer Kit (Invitrogen Corporation, Grand Island, NY, USA). Five micrograms of total RNA from the mesocarp was dephosphorylated, decapped and ligated with the GeneRacer™ RNA Oligo according to the manufacturer's instructions. Amplification of cDNA was carried out with the GeneRacer™ 3′ primer and 3'1NFYB1 primer (Table 1). A second amplification was performed with the Nested primer provided in the kit and a reverse nested primer designed to target the EgNFYB1 nested sequence, 3'2NFYB1 primer (Table 1). The PCR products were separated by electrophoresis on an agarose gel, purified using the QIAquick PCR Purification Kit (Qiagen) and ligated into pGEM-T Easy vector (Promega, Madison, WI, USA) for sequence analysis.

2.3. Sequence analyses and phylogenetic relationships

EgNFYB1 sequences were compared with the database using the BLAST network service (http://www.ncbi.nlm.nih.gov/BLAST/). The potential functional motif of *EgNFYB1* was then predicted by PROSITE using network protein sequence analysis (http://www.expasy.org/tools/). A phylogenetic tree was constructed according to the neighbour-joining method using the PHYLIP package (Choi et al., 2000). One thousand bootstrapped data sets were generated with SEQBOOT. Subsequently, PROTDIST generated a distance matrix, which was used to calculate distance trees, and the Kimura model was used to generate an unrooted minimum distance tree. The distance matrix output file from PROTDIST was then used in the neighbour-joining program to build the tree, and the jumble option was activated to evaluate a different order of the input data. The option of global arrangements was used to incorporate an additional search for the best tree. Finally, the output tree used the Consensus program to make a consensus tree, and the tree file was displayed using the program TREEVIEW (Ver. 1.6.6) with data aligned using the ClustalX 2.0 method (The European Bioinformatics Institute, Saffron Walden, UK) (alignments were shaded), and the conserved regions were identified using GENEDOC.

2.4 Modeling analysis

Modeling of the homology of the 2Dstructures was performed using the ProFunc software packages (http://www.ebi.ac.uk/thornton-srv/databases/profunc) (Laskowski et al., 2005) and the 3D structures of EgNFYB1 was then predicted by I-TASSER network service (http://zhanglab.ccmb.med.umich.edu).

2.5 Preparation of recombinant EgNFYB1 protein

RNA was extracted from mesocarp tissue of *E. guineensis* using Total RNA mini kit (Plant) (Geneaid) according to the manufacturer's instruction for amplified 183 bp of *EgNFYB1* gene which cover the active site of gene.Total RNA from the mesocarp was used for RT-PCR with one-step RT-PCR kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions using NFYB1-BamHIF and NFYB1-XhoIR primers (Table 1). The reaction was started at 50 ^{o}C for 30 min followed by an initial PCR activation step at 95 ^{o}C for 15 min, then 35 cycles at 94 ^{o}C for 1 min, 55^{o}C for 0.5 min and 72 ^{o}C for 1 min. The program was terminated by a 10 min incubation step at 72 ^{o}C. The RT-PCR products were separated by electrophoresis on an agarose gel, purified using the QIAquick PCR Purification Kit (Qiagen) and ligated into pET28a^{-} and transformed into *E. coli* BL21 (DE3). The addition of 0.1 mM IPTG was used to induce the expression of the

recombinant protein, the protein was harvested as described by Favacho (Favacho et al., 2006) and purified using the Ni-NTA Agarose (Qiagen) for DPI-ELISA assay

2.6. Genomic and Preparation of ds-bio DNA probe

Genomic DNA was extracted from mesocarp using a Genomic DNA Mini Kit (Plant) (Geneaid) according to the manufacturer's protocol and was used as a template for probe preparation of *EgaccD* promoter (Nakkaew, et al., 2014) and EgLEA4 promoter (Hualkasin, et al., 2013). The PCR with one 5' biotinylated primer and non biotinylated primer were: accDF and accDR primers (Table 1) for the accD promoter probe; and LEA4F and LEA4R (Table 1) for the LEA4 promoter probe. PCR conditions were as follows: 95 ^{o}C 5 min, 35 cycles of 95 ^{o}C 10 s, 50 ^{o}C 20 s and 72 ^{o}C 1 min; and 72 ^{o}C an additional 5 min. To confirm specific binding of EgNFYB1 transcription factors to DNA probe, CaMV 35S promoter of pBI121 vector was amplified as a control using the CaMVF and CaMVR primers (Table 1) for the CaMV 35s promoter probe. The PCR conditions was the same as described above. The PCR products were separated by electrophoresis on an agarose gel, purified using the QIAquick PCR Purification Kit (Qiagen) for sequence analysis and DPI-ELISA assay. The CCAAT box sequence analysis by PlantPAN program (http://plantpan.mbc.nctu.edu.tw).

2.7 DPI-ELISA

The EgNFYB1 proteins binding to promoters were determined using the DPI-ELISA assays were performed as described (Brand et al., 2010). Promoter biotinylate probes were bound to Streptavidin Coated 96-Well Plates (Thermo Scientific) for 1 h at 37^{o}C. The wells were blocked with 5% (w/v) non-fat dry milk in TBS-T buffer (20 mM Tris-HCl pH7.5, 180 mM NaCl, 0.1% (v/v) Tween20) for 30 min, and then incubated with 100 ng of EgNFYB1 protein for 1 h. After incubation, the wells were washed 3 times with TBS-T buffer, and then were incubated with 1:1000 diluted His-HRP antibody conjugate in PBS-T buffer for 1 h. The wells were then washed twice with PBS-T and PBS buffers, respectively, after incubation. The protein binding was detected by adding the OPD solution (zymed laboratories Inc.), and the reaction was stopped by 2N HCl. The color extinction was measured at 492 nm, using 690 nm as a reference wavelength in the ELISA reader.

2.8. Semi-quantitative RT-PCR

Total RNA was extracted from samples of *E. guineensis* leaves, other vegetative tissues, such as mesocarp , kernel of *E. guineensis* fruits 26 weeks and leaves, cotyledon, stem and root of *E. guineensis* 1-year-old and samples of less than 1-year-

old *E. guineensis* leaves using a Plant total RNA mini kit (Geneaid) according to the manufacturer's protocol. RT-PCR reactions included a total of 500 ng RNA as the template, one-step RT-PCR kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. (Qiagen).The reaction was started at 50 ^{o}C for 30 min followed by an initial PCR activation step at 95 ^{o}C for 15 min, then 30 cycles at 94 ^{o}C for 1 min, 52^{o}C for 0.5 min and 72 ^{o}C for 1 min. The program was terminated by a 10 min incubation step at 72 ^{o}C. RT- PCR was performed to amplify a 133 bp fragment of EgNFYB1 using NFYB1F and NFYB1R primers (Table 1). To confirm tissue-specific expression of NFYB1gene, 540 bp of *E. guineensis* 18S rDNA was amplified as a control using the 18s rDNA-specific primers, 18s rDNAF and 18s rDNAR primers (Table 1). These RT-PCR products were separated by electrophoresis on a 1.5% agarose gel, visualized by ethidium bromide staining and photographed using the Image Analyzer Gel Doc 1000 (Bio-Rad, Hercules, CA, USA). The expression levels were analyzed with the Quantity one program, and statistical analysis was performed using SPSS.

3. Results and discussion

Characterization of EgNFYB1

The oil palm ESTs related to oil synthesis included EgNFYB1.This protein have been reported to be involved in fatty acid pathway. For example, NFYB1 gene act as key regulators to coordinate the expression of fatty acid biosynthetic genes and overexpression of AtNFYB1 gene resulted in a dramatically increased level of FAs in the transgenic Arabidopsis plants (Mu et al., 2008), seed-specific overexpression of maize NFYB1 (ZmNFYB1) resulted in an average 48% increase in seed oil (Shen et al., 2010), Conditional expression of BnNFYB1 increases the seed oil content by 2% (Tan et al., 2011). In this work, we pursued functional studies of the EgNFYB1 in oil palm (*E. guineensis*). The 210-base pair (bp) fragment, a set of gene specific primers was designed to obtain the 3′ ends using a 3′ RACE protocol. The cDNA of the EgNFYB1 gene from oil palm was 613 bp. This sequence had an open reading frame (ORF) of 546 bp ending with the termination codon (TGA). In addition, the EgNFYB1sequence contained a 3′ UTR of 67 bp. The ORF of 181 amino acid residues of the entire sequence was compared with the database using the BLAST network service (NCBI), which revealed an 89% identity with *JcNFYB1* from *Jatropha curcas* and 88% identity with *AhNFYB1* from *Arachis hypogaea* (Fig. 1). The 181 amino acid protein (EgNFYB1) has a predicted molecular weight of 18.93 kDa, with a calculated pI of 7.20. Moreover,

the EgNFYB1amino acid sequences were analysed with PROSITE analysis, which showed significant sequence similarity with the HAP3 subunit signature (C-[VA]-[ST]-E-x-I-S-F-[LIVM]-T-[SGC]-E-A-[SCN]-[DE]-[KRQ]-C. motif; 28-44) involved in DNA-binding (Figure 1) and the predicted 2D structure including the 10 helices, 11 beta-strands (Fig. 2) and predicted 3D structure is shown in Fig. 3. In Arabidopsis, LEAFY COTYLEDON1 (NFYB1) has been identified as key transcription factors involved in the regulation of oil accumulation. NFYB1 encodes a HAP3 subunit of the CCAAT-binding factor (Lotan et al., 1998) and only NFYB1, a member of the HAP3 gene family, has been characterized in detail (Miyoshi et al., 2003).

The multiple alignment and phylogenetic tree was constructed using amino acid sequence with the PHYLIP program using the neighbour-joining method of EgNFYB1 with NFYB1 from plants and bacteria demonstrated that the oil palm *EgNFYB1* amino acid in this work is closest with jatropha *Jc*NFYB1 amino acid (Fig. 4) which in current, the use of Jatropha as a biodiesel feed-stock plant same the oil palm (Fairless, 2007; Sanderson, 2009).

The characteristic of *EgNFYB1* transcription factor binding to promoters by DNA Protein Interaction Enzyme Linked Immunosorbent Assay (DPI-ELISA) (*in vitro*)

The reports by Lotan that The *NFYB1* gene encodes a transcription factor homolog, the CCAAT box-binding factor HAP3 subunit of the CCAAT box-binding factor that bind specifically to DNA fragments containing a CCAAT and CAAT sequence (Gusmaroli et al., 2001; Kusnetsov et al., 1999; Lotan et al., 1998; Yazawa and Kamada, 2007). The CCAAT-box is a cis-acting element that is often found in the gene promoters of higher eukaryotes and essential to transcription process (Mantovani, 1999). We conducted an experiment to determine the characteristic of *EgNFYB1* binding to promoters. The *EgNFYB1* recombinant protein has CCAAT binding site (Laloum et al., 2013) that 61 amino acid and molecular weight of 7 kDa. The *EgNFYB1* recombinant protein expression in pellet and supernatant. Purified supernatant by Ni-NTA Agarose (Qiagen)(Fig. 5). Six CAAT box (CAAT/CCAAT) of biotin carboxylase (*EgaccD*) promoter were identified from oil palm (Table1), the gene of *EgaccD* from *E. guineensis* Jacq. encodes a plastid coded subunit of heteromeric acetyl-CoA carboxylase (AC*Case*) that it is important as it catalyzes the first committed step of fatty acid synthesis (Nakkaew et al., 2008). Two CAAT box (CAAT/CCAAT) of Late-Embryogenesis Abundant group 4 (*EgLEA4*) promoter were identified from oil palm, the *EgLEA*4 genes is a group of genes that have been reported to be involved in stress and hormone responses (Hualkasin et al., 2013) (Table1) and one CAAT box (CAAT/CCAAT) of

cauliflower mosaic virus 35s promoter (CaMV 35s promoter) was identified from pBI 121 vector (positive control), the CaMV 35s promoter is active at a high level in most plant tissues and one of the most widely used (Odell et al., 1985) (Table2).

We tested these interactions by DPI-ELISA. *EgNFYB1* protein bound to the *EgaccD* promoter-probe, the *EgLEA4* promoter-probe and the CaMV35s promoter-probe (Fig. 6). The absorbance of *EgNFYB1* protein - *EgaccD* promoter-probe was 0.175 which more than higher than *EgNFYB1* protein – *EgLEA4* promoter-probe and *EgNFYB1* protein – CaMV35s promoter-probe (Fig. 7). This can be explained by different concentrations of the epitope tagged *EgNFYB1* protein that these differences in binding specificity directly translate to altered residues in the DNA-binding domain. DPI-ELISA analysis shows that *EgNFYB1* protein binding to six CAAT-box of accD promoter is strongly. The absorbance of *EgNFYB1* protein (Don't have promoter probe) was 0.047, *EgNFYB1* protein binding non-specific to residues in well plate and the last well was changes of reaction test that these are negative control.

Suggesting, NFYB1 belongs to the HAP3 or the CCAAT binding factor family, which is highly conserved in promoters of fatty acid synthesis genes are enriched for the consensus NFYB1 binding element that these genes may be direct targets of NFYB1 (Fobert, 2014). In a more extended view, in mammalian cells, the CCAAT-binding factors directly bind to the FAS promoter and play an important role in the regulation of FAS gene expression (Rangan et al., 1996; Roder et al., 1997; Schweizer et al., 2002).

Expression of *EgNFYB1*

Semi-quantitative RT-PCR was performed to evaluate *EgNFYB1* gene expression patterns. Total RNAs extracted from five different tissues (mesocarp , kernel of *E. guineensis* fruits 26 weeks and leaves, cotyledon, stem and root) were amplified using NFYB1F and NFYB1R primer sets and 18S rDNA primers were used as an internal control to ensure. The RT-PCR products size 133 bp and 540 bp, respectively.The RT-PCR for each sample contained a similar amount of total RNA in all the tested samples, the transcripts of this gene are expressed detected in all tissues but at different expression levels in different tissues. *EgNFYB1* showed relatively strong transcription in mesocarp tissue, but relatively weak expression in root tissue (Fig. 8). The high expression of the *EgNFYB1* gene in mesocarp tissue may be associated with the production of palm oil that oil palm can accumulate up to 90% oil in its mesocarp, the highest level in the plant kingdom (Bourgis et al., 2011). This result was similar to the levels of *Eg14-3-3 ω* expression in oil palm. The *Eg14-3-3 ω* expression patterns were strong in the mesocarp that related to the oil accumulation (Nakkaew et al., 2013).To

confirmed the expression of *EgNFYB1* gene that involve in the oil accumulation by using RT-PCR analysis, the total RNA from the six samples of oil palm leaves from individual plants. Samples were divided into two groups are the samples with high production and the samples with low production. The results of RT-PCR demonstrated that the expression level of *EgNFYB1* was high in high-productivity plants and low in low productivity plants (Fig. 9). The RT-PCR analysis indicated the higher level of *EgNFYB1* expression is correlated with the higher productivity and in turn to the oil content. This is in accordance with the levels of *EgaccD* expression in oil palm, RT-PCR analysis in oil palm leaves with high production and oil palm leaves with low production. The expression level of *EgaccD* is correlate to high yield in oil palm (Nakkaew et al., 2008).

Moreover, the total RNA from six unknown young leaves from 2 place and different trees of less than 1-year-old and same size were subjected to gene expression analysis (data not show). The level of expression of *EgNFYB1* considerable variations were observed and indicated that the *EgNFYB1* has transcription at an early age of tree that according with Lotan., et al. reported, *NFYB1* cDNA clone under the control of 35s promoter transformed in plant. The embryo-specific accumulated an oil body compare wild type (Lotan et al., 1998). *NFYB1* increased oil synthesis does appear to be a direct consequence attributed to the natural progression of somatic embryo through the maturation phase (Fobert, 2014). This finding suggests that *EgNFYB1* gene may be related to the ability of the palm tree to accumulate oil. It is interesting to use as a molecular marker during the breeding program and selection of producing high yielding oil palms.

Conclusions

The *EgNFYB1* gene was cloned from EST library. It is highly similar to the *date palm NFYB1* and has the same HAP3 subunit signature which involved in CAAT-binding that *EgNFYB1* protein binding to six CAAT-box of *EgaccD* promoter (gene involved in amount oil seed) is strongly. The *EgNFYB1* expression patterns were strongly expressed in mesocarp tissues and high yield leaves. The finally, the level of expression of *EgNFYB1* considerable variations were the unknown young leaves from different trees. These result indicated that the *EgNFYB1* expression may be related to accumulate oil. Therefore, the *EgNFYB1* may be of benefit to molecular marker during the breeding program in oil palm and other plants.

Acknowledgement

We thank the Thailand Research Fund; TRF (MRG6080358) and Research Assistant (RA) funded by faculty of Science, Prince of Songkla University. Furthermore, this work was financially supported by Prince of Songkla University Fund (SCI580391S) and the Center for Genomics and Bioinformatics Research, the Faculty of Science, Prince of Songkla University.

Legend of Figure

Fig. 1 Multiple amino acid sequence alignments (ClustalX 2.0) of *EgNFYB1* from oil palm (Elaeis guineensis Jacq; compared with the NFYB1 protein from other plants: *Zm*NFYB1 (NP_001105518.1, *Zea mays*), *Os*NFYB1(AAP22065.1, *Oryza sativa* (indica cultivar-group), *Ah*NFYB1-A (ADC33212.1,*Arachis hypogaea*), *Ah*NFYB1-B (ADC33213.1, *Arachis hypogaea*), *Gl*NFYB1-A (ABW71516.1,*Glycine latifolia*), *Gl*NFYB1-B (ABW71517.1,*Glycine latifolia*), *Gm*NFYB1 (NP_001236625.1, *Glycine max*), *Bn*NFYB1 (ACB12186.1, *Brassica napus*, *Bo*NFYB1 (AGK82893.1, *Brassica oleracea*), *Jc*NFYB1 (AEO22132.1, *Jatropha curcas*), *At*NFYB1 (AAC39488.1, *Arabidopsis thaliana*) and *Mt*NFYB1 (XP_003589766.1, *Medicago truncatula*).The putative HAP3-subunit site), which overlap at the NFYB1protein.

Fig. 2 The secondary structure of the putative *EgNFYB1* in oil palm including the 10 helices and 11 beta- strands

Fig. 3 The 3D structure of the putative *EgNFYB1* protein. The image was generated by using the I-TASSER server. Shown are the N- and C- terminal including the 10 helices (red), 11 beta- strands (green) and binding site (yellow)

Fig. 4 The phylogenetic tree constructed with the PHYLIP program using the neighbor-joining method supported for the inferred groups, which were obtained by bootstrap analysis from 1000 replications of the data set using the SEQBOOT in PHYLIP 3.69 package. From *EgNFYB1* of Oil palm (*Elaeis guineensis Jacq* and other living. The accession number and corresponding species are as follows: *Zm*NFYB1 (NP_001105518.1, *Zea mays*), *Os*NFYB1(AAP22065.1, *Oryza sativa* (indica cultivar-group), *Ah*NFYB1-A (ADC33212.1, *Arachis hypogaea*), *Ah*NFYB1-B (ADC33213.1, *Arachis hypogaea*), *Gl*NFYB1-A (ABW71516.1, *Glycine latifolia*), *Gl*NFYB1-B (ABW71517.1, *Glycine latifolia*), *Gm*NFYB1 (NP_001236625.1, *Glycine max*), *BnNFYB1-1 (ADF81044.1, Brassica napus), BnNFYB1-2 (ADF81045.1, Brassica napus*), *Bo*NFYB1 (AGK82893.1, *Brassica oleracea*), *Jc*NFYB1 (AEO22132.1, *Jatropha curcas*), *At*NFYB1

(AAC39488.1, *Arabidopsis thaliana*) and *Eh*CCAAT-binding transcription factor (XP_651413.1, *Entamoeba histolytica*).

Fig. 5 The 12% SDS-PAGE analysis showed the *EgNFYB1* protein expression from *E.coli* BL21(DE3) and *EgNFYB1* protein after protein purification by Ni-NTA Agarose

Fig. 6 The ELISA plate is coated with double stranded biotinylated promoter incubated with *EgNFYB1* recombinant protein. Representative wells of the microtiter plate are shown binding between promoters with *EgNFYB1* proteinLane 1 : The *EgaccD* promoter with *EgNFYB1* recombinant proteinLane 2 : The *EgLEA4* promoter with *EgNFYB1* recombinant protein Lane 3 :The CaMV35s promoter with *EgNFYB1* recombinant protein (Positive control)Lane 4 : The *EgNFYB1* recombinant protein **(**Negative control**)** Lane 5 : reaction test (Negative control)

Fig.7 The binding between promoters with *EgNFYB1* recombinant protein including the *EgNFYB1* protein binding to six CAAT box of *EgaccD* promoter (pink), the *EgNFYB1* protein binding to two CAAT box of *EgLEA4* promoter (green), the *EgNFYB1* protein binding to one CAAT box of CaMV 35s promoter (yellow), the *EgNFYB1* protein binding to well plate (blue) and solution test (Negative control). Each value is the mean of three replicates, and the error bars indicate standard errors at $P < 0.05$.

Fig.8 Semi-quantitative RT-PCR analysis of the *EgNFYB1* gene from various tissues of oil palm. The *EgNFYB1* mRNA levels were analyzed using the Quantity One program with 18S rDNA as the internal control. The gene expression of *EgNFYB1* was strong transcription in mesocarp tissue. Each value is the mean of three replicates and the error bars indicate standard errors at $P < 0.05$.

Fig. 9 Semi-quantitative RT-PCR analysis of *EgNFYB1* expression in six samples of oil palm leaves in high yield (green) and low yield (orange) of oil palm. The *EgNFYB1* mRNA levels were analyzed using the Quantity One program with 18S rDNA as the internal control. The gene expression of *EgNFYB1* was strong transcription in high-productivity plants. Each value is the mean of three replicates and the error bars indicate standard errors at $P < 0.05$.

Table 1: Sequences of the primers used in this study

Primer	Nucleotide sequence (5′----------3′)
GeneRacer™ 3′ 3’1NFYB1 Nested primer 3’2NFYB1	

NFYB1-BamHIF NFYB1-XhoIR accDF accDR LEA4F LEA4R CaMVF CaMVR NFYB1F NFYB1R 18s rDNAF 18s rDNAR	

Table2: The CAAT-box site on promoters analyze by PlanPAN server

Promoter	Number of CCAAT box	Position of CCAAT box
EgaccD	6	68, 269, 285, 599, 649, 689
EgLEA4	2	84, 171
CaMV 35s	1	281

Table 3 The averge of binding binding between promoters with *EgNFYB1* recombinant protein in reaction

Sample	AVERGE of ($O.D_{492}$-$O.D_{690}$)	STDEV ($O.D_{492}$-$O.D_{690}$)
EgaccD	0.175	0.0103
EgLEA4	0.102	0.012
CaMV35s	0.096	0.018
Control	0.047	0.022
Negative	0	0

Table 4: The quantitative of relative of *EgNFYB1* expression level by semi-quantitative RT- PCR in various tissues

Sample	AVERGE of relative of *EgNFYB1* expression level)	STDEV of relative of *EgNFYB1* expression level
Mesocarp	0.972	0.041
Kernel	0.656	0.055
Leafy cotyledon	0.547	0.036
Leave	0.535	0.018
Root	0.235	0.018

Table: 5 The quantitative of relative of *EgNFYB1* expression level by semi-quantitative RT- PCR in oil palm leaves of high yield and low yield

Sample	AVERGE of relative of *EgNFYB1* expression level	STDEV of relative of *EgNFYB1* expression level
High1	0.971	0.194
High2	0.835	0.126
High3	1.224	0.038
Low1	0.293	0.014
Low2	0.247	0.023
Low3	0.301	0.044

Reference

Baud, S., et al., 2009. Role of WRINKLED1 in the transcriptional regulation of glycolytic and fatty acid biosynthetic genes in Arabidopsis. Plant J. 60, 933-47.

Baud, S., et al., 2010. PII is induced by WRINKLED1 and fine-tunes fatty acid composition in seeds of Arabidopsis thaliana. Plant Journal. 64, 291-303.

Baud, S., Lepiniec, L., 2010. Physiological and developmental regulation of seed oil production. Prog Lipid Res. 49, 235-49.

Berger, N., et al., 2011. Transcriptional regulation of Arabidopsis LEAFY COTYLEDON2 involves RLE, a cis-element that regulates trimethylation of histone H3 at lysine-27. Plant Cell. 23, 4065-78.

Billotte, N., et al., 2005. Microsatellite-based high density linkage map in oil palm (Elaeis guineensis Jacq.). Theor Appl Genet. 110, 754-65.

Bourgis, F., et al., 2011. Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning (vol 108, pg 12527, 2011). Proceedings of the National Academy of Sciences of the United States of America. 108, 18186-18186.

Brand, L.H., et al., 2010. DPI-ELISA: a fast and versatile method to specify the binding of plant transcription factors to DNA in vitro. Plant Methods. 6, 25.

Braybrook, S.A., Harada, J.J., 2008. LECs go crazy in embryo development. Trends Plant Sci. 13, 624-30.

Choi, J.H., et al., 2000. PhyloDraw: a phylogenetic tree drawing system. Bioinformatics. 16, 1056-8.

Courchesne, N.M.D., et al., 2009. Enhancement of lipid production using biochemical, genetic and transcription factor engineering approaches. Journal of Biotechnology. 141, 31-41.

Davis, M.S., Solbiati, J., Cronan, J.E., Jr., 2000. Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in Escherichia coli. J Biol Chem. 275, 28593-8.

Fairless, D., 2007. Biofuel: The little shrub that could - maybe. Nature. 449, 652-655.

Favacho, A.R., et al., 2006. Cloning, expression, and purification of recombinant bovine rotavirus hemagglutinin, VP8*, in Escherichia coli. Protein Expr Purif. 46, 196-203.

Fobert, P.R., 2014. Role of transcription factors in storage lipid accumulation in plants. Vol., ed.^eds. The AOCS Lipid Library, pp. 1-7.

Gusmaroli, G., Tonelli, C., Mantovani, R., 2001. Regulation of the CCAAT-Binding NF-Y subunits in Arabidopsis thaliana. Gene. 264, 173-85.

Hualkasin, W., et al., 2013. Molecular cloning and characterization of the late embryogenesis abundant group 4 (EgLEA4) gene from oil palm (Elaeis guineensis Jacq). Songklanakarin J. Sci. Technol. . 35 (3), 275-285.

Kim, K.H., 1997. Regulation of mammalian acetyl-coenzyme A carboxylase. Annu Rev Nutr. 17, 77-99.

Kusnetsov, V., et al., 1999. The assembly of the CAAT-box binding complex at a photosynthesis gene promoter is regulated by light, cytokinin, and the stage of the plastids. J Biol Chem. 274, 36009-14.

Laloum, T., et al., 2013. CCAAT-box binding transcription factors in plants: Y so many? Trends Plant Sci. 18, 157-66.

Laskowski, R.A., Watson, J.D., Thornton, J.M., 2005. ProFunc: a server for predicting protein function from 3D structure. Nucleic Acids Res. 33, W89-93.

Li-Beisson, Y., et al., 2010. Acyl-lipid metabolism. Arabidopsis Book. 8, e0133.

Li, X.Y., et al., 1992. Evolutionary variation of the CCAAT-binding transcription factor NF-Y. Nucleic Acids Res. 20, 1087-91.

Liu, J., et al., 2010. Increasing seed mass and oil content in transgenic Arabidopsis by the overexpression of wri1-like gene from Brassica napus. Plant Physiol Biochem. 48, 9-15.

Lotan, T., et al., 1998. Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. Cell. 93, 1195-205.

Maeo, K., et al., 2009. An AP2-type transcription factor, WRINKLED1, of Arabidopsis thaliana binds to the AW-box sequence conserved among proximal upstream regions of genes involved in fatty acid synthesis. Plant Journal. 60, 476-487.

Mantovani, R., 1999. The molecular biology of the CCAAT-binding factor NF-Y. Gene. 239, 15-27.

Miyoshi, K., et al., 2003. OsHAP3 genes regulate chloroplast biogenesis in rice. Plant J. 36, 532-40.

Mu, J., et al., 2008. LEAFY COTYLEDON1 is a key regulator of fatty acid biosynthesis in Arabidopsis. Plant Physiol. 148, 1042-54.

Murphy, D.J., 2001. The biogenesis and functions of lipid bodies in animals, plants and microorganisms. Prog Lipid Res. 40, 325-438.

Nakkaew, A., et al., 2008. Cloning and expression of a plastid-encoded subunit, beta-carboxyltransferase gene (accD) and a nuclear-encoded subunit, biotin carboxylase of acetyl-CoA carboxylase from oil palm (Elaeis guineensis Jacq.). Plant Science. 175, 497–504.

Nakkaew, A., et al., 2013. Cloning, characterization and overexpression of a 14-3-3 x protein from oil palm (Elaeis guineensis). Plant Breeding.

Odell, J.T., Nagy, F., Chua, N.H., 1985. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature. 313, 810-2.

Peng, F.Y., Weselake, R.J., 2011. Gene coexpression clusters and putative regulatory elements underlying seed storage reserve accumulation in Arabidopsis. BMC Genomics. 12, 286.

Pouvreau, B., et al., 2011. Duplicate maize Wrinkled1 transcription factors activate target genes involved in seed oil biosynthesis. Plant Physiol. 156, 674-86.

Rangan, V., Oskouian, B., Smith, S., 1996. Identification of an inverted CCAAT box motif in the fatty-acid synthase gene as an essential element for modification of transcriptional regulation by cAMP. J Biol Chem 271, 2307–2312.

Roder, K., et al., 1997. Cooperative binding of NF-Y and Sp1 at the DNase I-hypersensitive site, fatty acid synthase insulin-responsive element 1, located at -500 in the rat fatty acid synthase promoter. J Biol Chem 272, 21616–21624.

Sanderson, K., 2009. Wonder weed plans fail to flourish. Nature. 461, 328-329.

Schweizer, M., et al., 2002. Transcription factors acting on the promoter of the rat fatty acid synthase gene. Biochem Soc Trans. 30, 1070-2.

Sendl, A., et al., 1992. Inhibition of cholesterol synthesis in vitro by extracts and isolated compounds prepared from garlic and wild garlic. Atherosclerosis. 94, 79-85.

Shen, B., et al., 2010. Expression of ZmNFYB1 and ZmWRI1 increases seed oil production in maize. Plant Physiol. 153, 980-7.

Singh, R., et al., 2009. Mapping quantitative trait loci (QTLs) for fatty acid composition in an interspecific cross of oil palm. BMC Plant Biol. 9, 114.

Singh, R., et al., 2013. The oil palm SHELL gene controls oil yield and encodes a homologue of SEEDSTICK. Nature. 500, 340-4.

Subrahmanyam, S., Cronan, J.E., Jr., 1998. Overproduction of a functional fatty acid biosynthetic enzyme blocks fatty acid synthesis in Escherichia coli. J Bacteriol. 180, 4596-602.

Suzuki, M., McCarty, D.R., 2008. Functional symmetry of the B3 network controlling seed development. Curr Opin Plant Biol. 11, 548-53.

Tan, H., et al., 2011. Enhanced seed oil production in canola by conditional expression of Brassica napus LEAFY COTYLEDON1 and NFYB1-LIKE in developing seeds. Plant Physiol. 156, 1577-88.

Tranbarger, T.J., et al., 2011. Regulatory mechanisms underlying oil palm fruit mesocarp maturation, ripening, and functional specialization in lipid and carotenoid metabolism. Plant Physiol. 156, 564-84.

Weselake, R.J., et al., 2009. Increasing the flow of carbon into seed oil. Biotechnol Adv. 27, 866-78.

Yazawa, K., Kamada, H., 2007. Identification and characterization of carrot HAP factors that form a complex with the embryo-specific transcription factor C-NFYB1. J Exp Bot. 58, 3819-28.