



ความหลากหลายของลำดับเบสของยีนในวิถีการสังเคราะห์แคโรทีนอยด์ในข้าวไทย จากการวิเคราะห์ลำดับเบสทรานสคริปโตมและจีโนม

Gene Sequence Variations of Carotenoid Biosynthesis Pathway in Thai Rice Varieties (*Oryza sativa* L.) from Transcriptome and Whole Genome Sequencing Analysis

กนกวรรณ จันทร์เพ็ญ¹ วราภรณ์ แสงทอง¹ ช่อทิพา สกกุลสิงหาโรจน์¹ สุภารัตน์ ลิ้นชอุดม¹

อนูรักษ์ โพธิ์เยี่ยม² และแสงทอง พงษ์เจริญกิต^{1*}

Kanokwan Janphen¹ Varaporn Sangtong¹ Chotipa Sakulsingharoj¹ Suparat Leethanatudom¹

Anurug Poeaim² and Saengtong Pongjaroenkit^{1*}

¹สาขาวิชาพันธุศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้

²ภาควิชาชีววิทยา คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹Program of Genetics, Faculty of Science, Maejo University

²Department of Biology, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang

Received : 24 February 2021

Revised : 12 October 2021

Accepted : 19 October 2021

บทคัดย่อ

ลูทีนเป็นสารที่ได้จากวิถีการสังเคราะห์แคโรทีนอยด์ของข้าวไทย พบมากในข้าวที่มีเยื่อหุ้มเมล็ดสีดำ ในการศึกษาครั้งนี้จึงค้นหาความหลากหลายของลำดับเบสของยีนในวิถีการสังเคราะห์แคโรทีนอยด์เปรียบเทียบกับข้าวที่มีเยื่อหุ้มเมล็ดสีขาวที่ได้จากการวิเคราะห์ผลการหาลำดับอาร์เอ็นเอรวมกับการหาลำดับเบสจีโนม ซึ่งเป็นการค้นหาลำดับเบสที่สามารถแยกข้าวที่มีเยื่อหุ้มเมล็ดสีขาวที่ไม่พบสารลูทีนกับข้าวที่มีเยื่อหุ้มเมล็ดสีดำที่ตรวจพบสารลูทีนสูง ผลการค้นหาลำดับเบสที่แยกกลุ่มข้าวที่มีเยื่อหุ้มเมล็ดสีขาวจากสีดำ จำนวน 12 ยีน เป็นการเปลี่ยนแปลงลำดับเบสในส่วนที่เป็นยีนจำนวน 5 ยีน คือ ยีน phytoene desaturase (*PDS*), zeta-carotene isomerase (*ZISO*), lycopene epsilon-cyclase (*LCYe*), cytochrome P450 carotene beta-hydroxylase (*Lut5*, *CYP97A*) และ cytochrome P450 type B (*CYP97B*) โดยพบความหลากหลายของลำดับเบสจำนวนมากทั้งในส่วนที่เป็นยีน บริเวณด้านหน้า และบริเวณด้านหลังยีน ในยีน *ZISO*, *CYP97A* และ *CYP97B* ยีนเหล่านี้จึงน่าจะเป็นยีนสำคัญที่ทำงานเกี่ยวข้องกับการสังเคราะห์สารลูทีนในข้าวไทย ซึ่งความหลากหลายของลำดับเบสดังกล่าวสามารถนำไปใช้ในการพัฒนาเป็นเครื่องหมายดีเอ็นเอเพื่อใช้ในการช่วยคัดเลือกหรือปรับปรุงพันธุ์ข้าวไทยให้มีสารลูทีนสูงต่อไป

คำสำคัญ : ข้าว ; ความหลากหลายของลำดับเบส ; วิถีการสังเคราะห์แคโรทีนอยด์ ; การหาลำดับอาร์เอ็นเอ ; การหาลำดับเบสจีโนม



Abstract

Lutein is a product from carotenoid biosynthetic pathway which is highly found in black rice pericarp. The purpose of this study was to identify sequence variations of genes in carotenoid biosynthetic pathway from Thai rice varieties with white (no lutein) and black (high lutein) pericarp by transcriptome and genome analyses. The sequences of 12 genes in the carotenoid synthesis showing variations between white and black rice were detected. The variations in gene regions were found in 5 genes which were phytoene desaturase (*PDS*), zeta-carotene isomerase (*ZISO*), lycopene epsilon-cyclase (*LCYe*), cytochrome P450 carotene beta-hydroxylase (*Lut5*, *CYP97A*) and cytochrome P450 type B (*CYP97B*). High levels of sequence variations in upstream region, gene region and downstream region were found in *ZISO*, *CYP97A* and *CYP97B* genes. Therefore, these three genes were probably key candidate genes related to lutein biosynthesis. The sequence variations of these genes in the carotenoid biosynthesis pathway will be used to develop DNA markers for selection of high lutein accumulation in rice breeding program.

Keywords : rice ; sequence variation ; carotenoid biosynthesis pathway ; transcriptome sequencing ;
whole genome sequencing



บทนำ

ปัจจุบันเทคโนโลยีการหาลำดับเบสจีโนม (Genome sequencing) ของพืช ได้รับความนิยมน้อยอย่างแพร่หลาย จึงมีการนำผลการวิเคราะห์ลำดับเบสมาประยุกต์ใช้ในการปรับปรุงพันธุ์พืชและการค้นหายีนที่เกี่ยวข้องกับลักษณะทางการเกษตรเพิ่มมากขึ้น เนื่องจากข้าวเป็นพืชอาหารหลักของประชากรโลก และมีการเพาะปลูกเป็นจำนวนมากในหลายทวีป ข้าวจึงเป็นหนึ่งในพืชที่ได้รับความนิยมน้อยในการหาลำดับเบส เพื่อใช้ศึกษาความหลากหลายของลำดับเบสของข้าวแต่ละพันธุ์ เช่น การศึกษาการเปลี่ยนแปลงลำดับเบสของข้าวที่ได้จากการผสมข้ามระหว่างข้าวเอเชีย (WAB56-104) กับข้าวแอฟริกา (CG14) จากการวิเคราะห์ลำดับเบสของข้าวประชากรที่แยกด้วยพีโนไทป์เป็นสองกลุ่ม พบความแตกต่าง single nucleotide polymorphism (SNP) จำนวน 5,152 ตำแหน่ง ระหว่างประชากรข้าวสองกลุ่ม มี 6 ตำแหน่งที่เกี่ยวข้องกับยีนควบคุมลักษณะวันออกดอก การแตกกอ และน้ำหนักข้าวต่อ 1,000 เมล็ดของข้าวเปลือก ซึ่งอาจจะเป็นตำแหน่งสำคัญควบคุมลักษณะดังกล่าว (Badro *et al.*, 2020) เช่นเดียวกับการศึกษาของ Ren *et al.* (2020) ที่ใช้ผลการวิเคราะห์ลำดับเบสจากเทคนิคทรานสคริปโตม (Transcriptome sequencing; RNA-seq) ร่วมกับการหาลำดับเบสจีโนม (Whole Genome Sequencing; WGS) ของข้าวลูกผสมสายพันธุ์ Chuanyou6203 เพื่อวิเคราะห์ลำดับเบสที่มีความสัมพันธ์กับการแสดงออกของยีนที่ส่งผลให้เกิดลักษณะดีเด่นกว่าพันธุ์ของพ่อแม่ (Heterosis) ซึ่งข้อมูลที่ได้สามารถนำไปใช้พัฒนาพันธุ์ข้าวให้ได้ผลผลิตและคุณภาพของเมล็ดสูงต่อไป

จากความต้องการข้าวที่มีคุณค่าทางโภชนาการสูง จึงมีการพัฒนาสายพันธุ์ข้าวให้มีปริมาณสารแคโรทีนอยด์สูงขึ้น เนื่องจากสารในวิถีการสังเคราะห์แคโรทีนอยด์ ได้แก่ แอลฟา-แคโรทีน (alpha-carotene), บีตา-แคโรทีน (beta-carotene), บีตา-คริปโตแซนทิน (beta-cryptoxanthin), ไลโคพีน (lycopene), ลูทีน (lutein) และซีแซนทิน (zeaxanthin) มีความเกี่ยวข้องกับการทำงานของระบบสายตาและการมองเห็น นอกจากนี้ยังเป็นสารต้านอนุมูลอิสระและลดการเกิดมะเร็งบางประเภท จึงมีการศึกษากระบวนการสังเคราะห์และการสะสมของสารดังกล่าวในพืชหลากหลายชนิด (Watkins & Pogson, 2020) จากการศึกษาของ Kongkachuichai & Charoensir (2010) เกี่ยวกับคุณค่าทางโภชนาการของข้าวพื้นเมืองจำนวน 30 สายพันธุ์ ในอำเภอกุดชุม จังหวัดยโสธร เปรียบเทียบระหว่างข้าวขั้วดีและข้าวกล้องที่มีสีของเยื่อหุ้มเมล็ดสีดำ แดง น้ำตาล และขาว พบข้าวที่ให้ปริมาณลูทีนสูงสุดคือข้าวกล้องที่มีเยื่อหุ้มเมล็ดสีดำจำนวน 2 สายพันธุ์ ได้แก่ ข้าวเหนียวก้านน้อยและก้านใหญ่ (132.24 และ 115.15 ไมโครกรัม/100 กรัม) ในข้าวจะพบการสะสมสารฟลาโวนอยด์และแคโรทีนอยด์ที่เยื่อหุ้มเมล็ดและมักไม่พบในเอนโดสเปิร์ม (endosperm) จากการศึกษาปริมาณของสารฟลาโวนอยด์และแคโรทีนอยด์ในข้าวที่มีเยื่อหุ้มเมล็ดสีขาว สีแดง และสีดำ พบว่าข้าวที่มีเยื่อหุ้มเมล็ดสีดำมีปริมาณสารฟลาโวนอยด์และแคโรทีนอยด์สูงกว่าข้าวที่มีเยื่อหุ้มเมล็ดสีขาว และสีแดง (Kim *et al.*, 2010; Pereira-Caro *et al.*, 2013) เช่นเดียวกับการศึกษาการแสดงออกของยีนในวิถีการสังเคราะห์แคโรทีนอยด์จากข้าวไทยพันธุ์ก้านน้อย ทับทิมชุมแพ และปทุมธานี 1 ซึ่งมีเยื่อหุ้มเมล็ดสีดำ สีแดงและสีขาว ตามลำดับ พบว่าข้าวที่มีเยื่อหุ้มเมล็ดสีดำจะมีปริมาณสารลูทีนสูงกว่าข้าวเยื่อหุ้มเมล็ดสีแดงและขาว (Chetry *et al.*, 2019; Janphen *et al.*, 2019)



การศึกษาตำแหน่ง QTLs ที่เกี่ยวข้องกับการสังเคราะห์แคโรทีนอยด์และโฟเลตในข้าว พบตำแหน่ง QTLs ที่สัมพันธ์กับปริมาณโฟเลต จำนวน 3 ตำแหน่ง คือ qQTF-3-1, qQTF-3-2 และ qQTF-3-3 ตั้งอยู่บนโครโมโซมแท่งที่ 3 แต่ไม่พบตำแหน่ง QTL ที่สัมพันธ์กับปริมาณแคโรทีนอยด์ (Ashokkumar *et al.*, 2020) และจากผลการศึกษาการแสดงออกของยีนด้วยเทคนิคทรานสคริปโตมในข้าวที่มีเยื่อหุ้มเมล็ดสีดำ น้ำตาล และขาว จำนวน 6 พันธุ์ พบว่ายีน phytoene synthase (*PSY1*), phytoene desaturase (*PDS*), lycopene beta-cyclase (*LCYb*) และ lycopene epsilon-cyclase (*LCYe*) แสดงออกสูงในข้าวที่มีเยื่อหุ้มเมล็ดสีดำและเป็นการทำงานแบบร่วมกันของยีนอื่น ๆ ในวิถีการสังเคราะห์แคโรทีนอยด์ (Chetry & Chrungoo, 2020; Chetry *et al.*, 2019) เช่นเดียวกับงานวิจัยสร้างข้าวดัดแปลงพันธุกรรมให้มีการสังเคราะห์สาร Provitamin A ซึ่งเป็นสารที่ได้จากวิถีการสังเคราะห์แคโรทีนอยด์ในส่วนเอนโดสเปิร์ม โดยการแทรกยีนจำนวนหลายชุด ได้แก่ ข้าวพันธุ์ Golden Rice 1 จำนวน 2 ชุดยีน คือ ยีน *PSY* จากดอกแดฟโฟดิล และยีน *phytoene desaturase (CRTI)* จากแบคทีเรีย และข้าวพันธุ์ Golden Rice 2 จำนวน 2 ชุดยีน คือ ยีน *PSY* จากข้าวโพด และยีน *CRTI* จากแบคทีเรีย และมีการศึกษาความสมดุลของปริมาณสารตั้งต้นในวิถีการสังเคราะห์แคโรทีนอยด์โดยแทรกยีน *ZmPSY1* จากข้าวโพด, ยีน *PaCRTI* จากแบคทีเรีย และยีน 1-deoxy-D-xylulose 5-phosphate synthase (*AtDXS*) กับยีน *ORANGE (AtOR)* จากอะราบิโดอปซิส เพื่อให้มีการทำงานร่วมกันของยีนมากยิ่งขึ้น พบว่ามีการสร้างสารปีตา-แคโรทีนสูงกว่าข้าวที่ไม่ได้รับการถ่ายยีน (Al-Babili & Beyer, 2005; Bai *et al.*, 2016) การสังเคราะห์สารในวิถีแคโรทีนอยด์พบว่ามีเกี่ยวข้องกับการแสดงออกของยีนจำนวนมาก และยังไม่พบยีนหลักที่ควบคุมการสังเคราะห์ โดยพบว่าเป็นการทำงานร่วมกันของยีนหลายชนิดที่เกี่ยวข้องในวิถีการสังเคราะห์แคโรทีนอยด์ (Stanley & Yuan, 2019)

งานวิจัยนี้มีวัตถุประสงค์ค้นหาลำดับเบสของยีนในวิถีการสังเคราะห์แคโรทีนอยด์ ที่สามารถแยกกลุ่มข้าวไทยที่มีเยื่อหุ้มเมล็ดสีขาวจากกลุ่มที่มีเยื่อหุ้มเมล็ดสีดำ ซึ่งมีปริมาณสารลูทีนต่ำและสูงตามลำดับ ด้วยการวิเคราะห์ข้อมูลลำดับเบสจากเทคนิคทรานสคริปโตม (RNA-seq) ร่วมกับการหาลำดับเบสจีโนม (WGS) ซึ่งเป็นวิธีการที่มีความถูกต้อง แม่นยำสูง เพื่อนำความแตกต่างกันของลำดับเบสที่ได้ของยีนในวิถีการสังเคราะห์แคโรทีนอยด์นั้นมาประยุกต์ใช้ประโยชน์ในการพัฒนาเป็นเครื่องหมายดีเอ็นเอเพื่อใช้ในการปรับปรุงพันธุ์ข้าวไทยให้มีสารแคโรทีนอยด์สูงต่อไปในอนาคต

วิธีดำเนินการวิจัย

พันธุ์ข้าวและการสกัดสารพันธุกรรม

ตัวอย่างพันธุ์ข้าวที่ใช้ในการศึกษาเป็นข้าวที่มีเยื่อหุ้มเมล็ดสีขาว ซึ่งตรวจไม่พบสารลูทีนในเมล็ด จำนวน 2 พันธุ์ คือ ปทุมธานี 1 (PTT1) และ กข-แม่โจ้ 2 (RD-MJU2) จากหน่วยความเป็นเลิศด้านการวิจัยและพัฒนาการปรับปรุงพันธุ์ข้าว มหาวิทยาลัยแม่โจ้ ข้าวที่มีเยื่อหุ้มเมล็ดสีดำซึ่งตรวจพบสารลูทีนสูงในเมล็ด จำนวน 2 พันธุ์ คือ ก้าน้อย (KN) และ ก้าใหญ่ (KY) (Janphen *et al.*, 2019; Kongkachuichai & Charoensir, 2010) จากเขตปฏิรูปที่ดิน สำนักงานการปฏิรูปที่ดินเพื่อการเกษตรกรรม ต.ท่าแมด อ.กุดชุม จ.ยโสธร โดยเก็บตัวอย่างใบอ่อนหลังจากการงอกของเมล็ดนาน 14 วัน มาสกัด Genomic



DNA ด้วยวิธี mCTAB (Li *et al.*, 2013) ละลาย Genomic DNA ที่สกัดได้ด้วย TE buffer, pH 8.0 จากนั้นเก็บตัวอย่างที่อุณหภูมิ -20 องศาเซลเซียส ระหว่างรอส่งตัวอย่างไปวิเคราะห์ลำดับเบส

สำหรับการสกัด Total-RNA จากการศึกษากการแสดงออกของยีนในวิธีการสังเคราะห์แคโรทีนอยด์จากเนื้อเยื่อที่ระยะต่าง ๆ ของข้าวในฐานข้อมูล Expression Atlas (<https://www.ebi.ac.uk/gxa/home>) ในส่วนของ Plant: *Oryza sativa* (*O. sativa*) พบการแสดงออกของยีนสูงในเนื้อเยื่อส่วนใบและช่อดอก จึงได้เลือกใช้ใบธงที่ระยะ 7 วัน ก่อนและหลังออกดอก จากต้นข้าวจำนวน 4 ต้นต่อพันธุ์ ที่ปลูกในโรงเรือนพร้อมกัน มาทำการสกัด Total-RNA ด้วยสารละลาย TRIzol reagent (Invitrogen, USA) แล้วละลายด้วย TE buffer, pH 7.0 กำจัด Genomic DNA ที่ปนเปื้อนด้วยชุดสำเร็จรูป RapidOut DNA Removal Kit (Thermo Scientific, Germany) จากนั้นเก็บตัวอย่างที่อุณหภูมิ -20 องศาเซลเซียส ระหว่างรอส่งตัวอย่างไปวิเคราะห์ลำดับเบส

การวิเคราะห์ผลการหาลำดับอาร์เอ็นเอจากเทคนิคทรานสคริปโตม

Total-RNA ถูกวิเคราะห์ลำดับอาร์เอ็นเอด้วยเทคนิค RNA-seq (NovogeneAIT, Singapore) ด้วยวิธีการอ่านลำดับเบสและวิเคราะห์แบบ *De novo* sequencing ผลจากการหาลำดับเบสที่ได้จะผ่านการวิเคราะห์ข้อมูลตามวิธีการ Trinity นำ read ที่ได้จากการอ่านด้วยเครื่อง Illumina มาสร้าง contig จากนั้นนำ contig ที่มีลำดับเบสตรงกันมาสร้างเป็นสาย transcripts (Grabherr *et al.*, 2011) แล้ววิเคราะห์หน้าที่และการทำงานของสาย transcript ที่ได้จากฐานข้อมูลเป็นลำดับต่อไป ในกระบวนการหาลำดับเบส RNA-seq ได้ทำการ normalization ตามวิธีการของ Dillies *et al.* (2013) ลำดับเบสของ transcripts ที่ได้จะถูกตั้งชื่อในรูปของ Cluster ซึ่งเป็นลำดับเบสจากสาย mRNA ของยีนที่แสดงออกทั้งหมด รวมทั้งวิเคราะห์ความหลากหลายของลำดับเบสในแต่ละ Cluster จากข้าว 4 พันธุ์ คือ พันธุ์ปทุมธานี 1, พันธุ์กข-แม่ใจ 2, พันธุ์ก้าน้อย และพันธุ์กำใหญ่

ในขั้นต้นเลือกเฉพาะ Cluster ที่เกี่ยวข้องกับวิธีการสังเคราะห์แคโรทีนอยด์ (Carotenoid biosynthesis ; PATH : Ko00906) ที่ได้จากการเปรียบเทียบลำดับเบสจากฐานข้อมูล KEGG PATHWAY Database (<https://www.genome.jp/kegg/pathway>) และเลือกเฉพาะยีนที่มีรายงานใน *Oryza sativa japonica* (Japanese rice; Refseq; PATH:osa00906) นำ Cluster ที่ได้จากการเลือกขั้นต้นไปตรวจสอบตำแหน่งของยีนบนโครโมโซมของข้าวพันธุ์ Nipponbare (Ref. sequence) ด้วยการ BLAST ซึ่งเป็นการเปรียบเทียบลำดับเบสของแต่ละ Cluster กับฐานข้อมูล NCBI (<https://blast.ncbi.nlm.nih.gov>) ตามด้วยนำลำดับเบสที่แตกต่างในแต่ละ Cluster ที่ได้จากข้าว 4 พันธุ์ ไปหาตำแหน่งบนโครโมโซมต่อไป จากนั้นเลือกเฉพาะลำดับเบสที่สามารถแยกกลุ่มระหว่างข้าวที่มีเยื่อหุ้มเมล็ดสีขาวและสีดำ

การวิเคราะห์ผลการหาลำดับเบสจีโนม

ส่งตัวอย่าง Genomic DNA ไปอ่านและวิเคราะห์ลำดับเบสด้วยเทคนิค WGS (NovogeneAIT, China) แบบ resequencing โดยผลการหาลำดับเบสและการวิเคราะห์รายงานเป็นตำแหน่งของการเปลี่ยนแปลงลำดับเบสชนิด single nucleotide polymorphism (SNP) หรือ insertion-deletion mutation (InDel) ของข้าวทั้ง 4 พันธุ์ บนโครโมโซมทั้ง 12 แห่ง ในขั้นต้นค้นหาตำแหน่งของยีนในวิธีการสังเคราะห์แคโรทีนอยด์บนโครโมโซมจากฐานข้อมูล Ricebase (<https://ricebase.org/>)

ร่วมกับฐานข้อมูล The Rice Annotation Project (RAP) (<https://rapdb.dna.affrc.go.jp/>) หลังจากได้ตำแหน่งของแต่ละยีนบนโครโมโซมแล้ว นำตำแหน่งที่ได้ไปค้นหาตำแหน่งของ SNP และ InDel จากผลการวิเคราะห์ลำดับจีโนมของข้าวทั้ง 4 พันธุ์ โดยจะค้นหาในส่วนที่เป็นยีน ได้แก่ บริเวณ ปลาย 5' ที่ไม่แปลรหัส (5' untranslated region ; 5' UTR), บริเวณเอกซอน (exon), บริเวณอินทรอน (intron), บริเวณจุดตัด (splice region) และบริเวณ 3' ที่ไม่แปลรหัส (3' untranslated region ; 3' UTR) และบริเวณด้านหน้าและหลังของส่วนที่เป็นยีน คือ บริเวณด้านหน้าของของตำแหน่ง 5' UTR และด้านหลังของของตำแหน่ง 3' UTR ด้านละไม่เกิน 50 กิโลเบส (50,000 เบส) เพื่อรวบรวมตำแหน่งที่มีการเปลี่ยนแปลงลำดับเบสแบบ SNP และ InDel จากนั้นคัดเลือกเฉพาะการเปลี่ยนแปลงลำดับเบสที่สามารถแยกกลุ่มระหว่างข้าวที่มีเยื่อหุ้มเมล็ดสีขาวและสีดำ

การค้นหารูปแบบและตำแหน่งของความหลากหลายของลำดับเบสจากฐานข้อมูล Gramene

นำตำแหน่งการเปลี่ยนแปลงลำดับเบสที่สามารถแยกกลุ่มระหว่างข้าวที่มีเยื่อหุ้มเมล็ดสีขาว (พันธุ์ปทุมธานี 1 และ พันธุ์ข-แม่ใจ 2) กับข้าวที่มีเยื่อหุ้มเมล็ดสีดำ (พันธุ์ก้าน้อย และพันธุ์กำใหญ่) ทั้งแบบ SNP และ InDel ที่พบในส่วนที่เป็นยีนในวิธีการสังเคราะห์แคโรทีนอยด์ทั้ง 14 ยีน ที่ได้จากการวิเคราะห์ข้อมูลทรานสคริปโตมและลำดับเบสของจีโนม โดยใช้ข้อมูลลำดับเบสของข้าวที่มีเยื่อหุ้มเมล็ดสีขาวเป็นหลักและบริเวณที่แสดงความแตกต่างนั้นจะต้องเป็นเบสเดียวกันในข้าวทั้ง 2 พันธุ์ที่มีเยื่อหุ้มเมล็ดสีขาวและมีเบสที่แตกต่างเดียวกันในข้าวที่มีเยื่อหุ้มเมล็ดสีดำ จะเลือกบริเวณดังกล่าวมาทำการมาค้นหาตำแหน่งของการเปลี่ยนแปลงลำดับเบสของ ปลาย 5' ที่ไม่แปลรหัส เอกซอน อินทรอน จุดตัด และปลาย 3' ที่ไม่แปลรหัส ด้วยการนำไปค้นหาจากข้อมูล resequencing ของโครงการ The 3,000 Rice Genomes Project ซึ่งเป็นข้อมูลลำดับเบสจีโนมของข้าวชนิด *japonica*, *indica* และข้าวป่า จำนวน 3,000 สายพันธุ์ ในฐานข้อมูล Gramene (https://ensembl.gramene.org/Oryza_sativa/Info/Index) เพื่อหาผลกระทบที่เกิดขึ้นกับกรดอะมิโนที่เกิดจากการเปลี่ยนแปลงลำดับเบสส่วนที่เป็นยีน เช่น Silent, Missense, Insertion, Deletion และ Pre-mature stop codon (Nonsense)

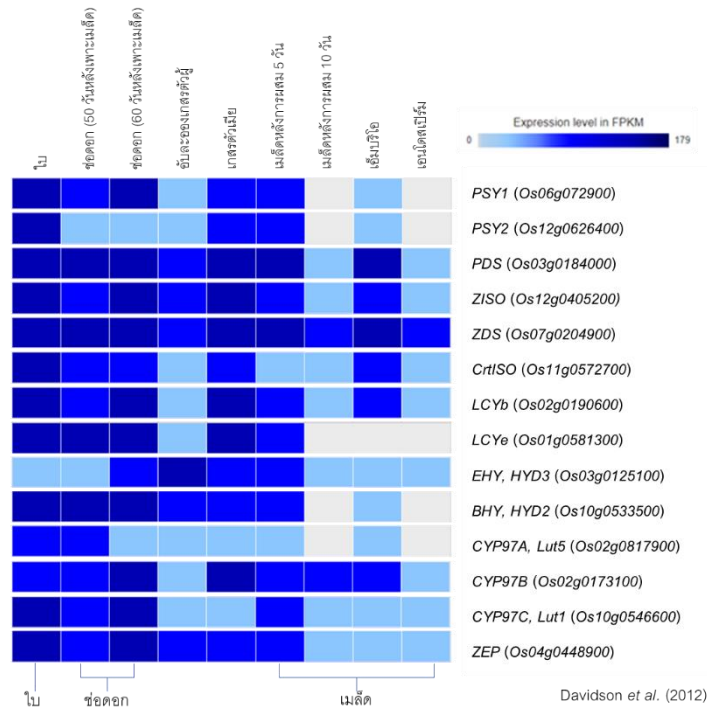
ผลการวิจัย

การค้นหาคความหลากหลายของลำดับเบสของยีนในวิธีการสังเคราะห์แคโรทีนอยด์จากการวิเคราะห์

ทรานสคริปโตม

ในการศึกษา RNA-seq จะต้องใช้ total RNA ปริมาณมากกว่า 3 ไมโครกรัมต่อตัวอย่าง (NovogeneAIT, Singapore) เมื่อบริษัทเลือกเนื้อเยื่อเพื่อสกัด Total-RNA จากการเปรียบเทียบระดับการแสดงออกของยีนในวิธีการสังเคราะห์แคโรทีนอยด์ในเนื้อเยื่อที่บริเวณต่างกัน 9 ชนิด ได้แก่ ใบ (20 วันหลังเพาะเมล็ด), ช่อดอก (50 วันหลังเพาะเมล็ด), ช่อดอก (60 วันหลังเพาะเมล็ด), อับละอองเกสรตัวผู้ (65 วันหลังเพาะเมล็ด), เกสรตัวเมีย (65 วันหลังเพาะเมล็ด), เมล็ดหลังการผสม 5 วัน (65 วันหลังเพาะเมล็ด), เมล็ดหลังการผสม 10 วัน (70 วันหลังเพาะเมล็ด), เอ็มบริโอ (85 วันหลังเพาะเมล็ด) และ เอนโดสเปิร์ม (85 วันหลังเพาะเมล็ด) ของพันธุ์ Nipponbare ที่รายงานในฐานข้อมูล Expression Atlas (<https://www.ebi.ac.uk/gxa/home>) จากการวิเคราะห์ด้วยเทคนิคทรานสคริปโตม (Davidson *et al.*, 2012) พบระดับการ

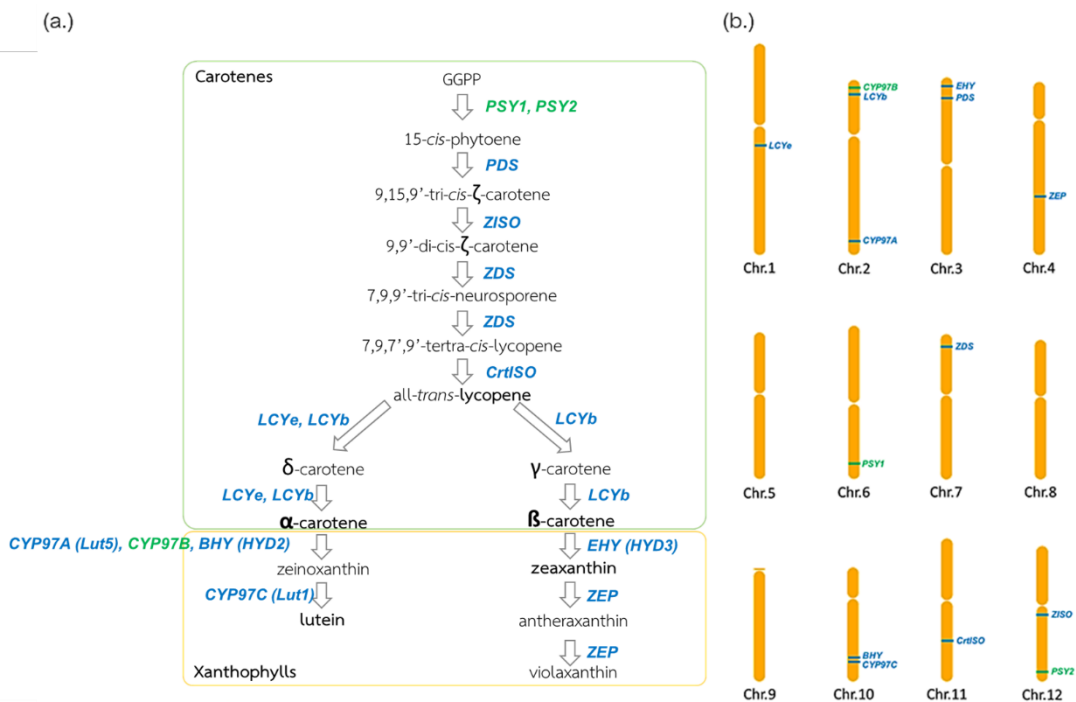
แสดงออกของยีนสูงในส่วนของใบและช่อดอก (ภาพที่ 1) ดังนั้นจึงเลือกเนื้อเยื่อใบธง (ใบที่ติดอยู่กับช่อดอก) ก่อนและหลังวันออกดอก 7 วัน มาสกัด total-RNA พบว่าได้ total-RNA ปริมาณมากและคุณภาพดี เพียงพอต่อการศึกษา RNA-seq ในขั้นตอนนี้ต่อไป



ภาพที่ 1 ผลการเปรียบเทียบระดับการแสดงออกของยีนในวิถีการสังเคราะห์แคโรทีนอยด์จำนวน 14 ยีน จากเนื้อเยื่อที่บริเวณต่างกันของต้นข้าวในฐานข้อมูล Expression Atlas (<https://www.ebi.ac.uk/gxa/home>) ระดับการแสดงออกยีนแสดงด้วยแถบสีที่เป็นค่า FPKM (Fragments Per Kilobase of transcript per Million fragments mapped) สีเทา แสดงค่าต่ำสุด (FPKM = 0) และสีน้ำเงิน แสดงค่าสูงสุด ตามลำดับ

การวิเคราะห์ลำดับเบสด้วยเทคนิค RNA-seq จะเป็นการวิเคราะห์ลำดับเบสของยีนที่ผ่านการถอดรหัส จากผลการวิเคราะห์แต่ละ Cluster ที่ได้จากการนำลำดับเบสไปเปรียบเทียบกับฐานข้อมูล KEGG PATHWAY Database จากข้อมูลจำนวน 15,166 Cluster พบที่เป็นส่วนของ Carotenoid biosynthesis (PATH:Ko00906) ของข้าว *O. sativa japonica* จำนวน 111 Cluster เมื่อตัด Cluster ที่มีลำดับเบสซ้ำกันออกและเลือก Cluster ที่ให้ลำดับเบสยาวที่สุดเป็นตัวแทนของ Cluster ได้ 61 Cluster ซึ่งเป็นรหัสของยีน 24 ยีน เมื่อเลือกเฉพาะยีนของเอนไซม์ที่เกี่ยวข้องกับการสังเคราะห์แคโรทีน (Carotenes) ถึงแซนโทฟิลล์ (Xanthophyll) พบยีนของเอนไซม์ที่เกี่ยวข้องจำนวน 12 เอนไซม์ จำนวน 29 Cluster ได้แก่ เอนไซม์ phytoene synthase (PSY), phytoene desaturase (PDS), zeta-carotene isomerase (ZISO), zeta-carotene desaturase (ZDS), carotene isomerase (CrtISO) ในกระบวนการเปลี่ยน Geranylgeranyl pyrophosphate (GGPP) ไปเป็น lycopene ตามด้วยเอนไซม์

lycopene beta-cyclase (*LCYb*), lycopene epsilon-cyclase (*LCYe*) ในการสังเคราะห์ alpha และ beta carotene จากนั้น alpha carotene จะเปลี่ยนเป็น ลูทีน (lutein) ด้วยเอนไซม์ epsilon-carotene hydroxylase (*BHY, HYD2*), cytochrome P450 carotene beta-hydroxylase (*Lut5, CYP97A*) และ cytochrome P450 carotene epsilon-hydroxylase (*Lut1, CYP97C*) ส่วน beta carotene จะเปลี่ยนเป็น zeaxanthin จนถึง violaxanthin ด้วยเอนไซม์ beta-carotene hydroxylase (*EHY, HYD3*) และ zeaxanthin epoxidase (*ZEP*) (ภาพที่ 2a) โดยแต่ละยีนที่ควบคุมการสร้างเอนไซม์เหล่านี้จะตั้งอยู่บนโครโมโซมของข้าว 9 แห่ง ได้แก่ แห่งที่ 1, 2, 3, 4, 6, 7, 10, 11 และ 12 (ภาพที่ 2b)



ภาพที่ 2 วิธีการสังเคราะห์แคโรทีนอยด์ (Carotenoid biosynthesis) จากฐานข้อมูล KEGG PATHWAY Database (a.) ยีนที่เกี่ยวข้องกับการสังเคราะห์สารแคโรทีนถึงแซนโทฟิล (b.) ตำแหน่งของแต่ละยีนบนโครโมโซมข้าว โดยแถบสีน้ำเงิน คือ ยีนจาก KEGG pathway และแถบสีเขียว คือ ยีนที่ได้จากการสืบค้นเพิ่มเติม

นอกจากยีนที่รายงานว่าเกี่ยวข้องกับการสังเคราะห์สารแคโรทีนและแซนโทฟิล จำนวน 12 ยีนแล้ว จากการสืบค้นข้อมูลยังพบยีน cytochrome P450 type B (*CYP97B*) ที่เกี่ยวข้องกับการสังเคราะห์สารแคโรทีนอยด์ โดยมีความคล้ายคลึงกับลำดับเบสของยีน *CYP97A* (Niu *et al.*, 2020; Villa-Rivera & Ochoa-Alejo, 2020) ซึ่งตั้งอยู่บนโครโมโซมแห่งที่ 2 เช่นเดียวกับยีน *CYP97A* แต่อยู่คนละแขนของโครโมโซม นอกจากนี้ยังพบว่ายีนกลุ่ม *PSY* ในข้าวมี 3 ยีน (isoforms) คือ ยีน *PSY1, PSY2* และ *PSY3* โดยยีน *PSY1* และ *PSY2* เกี่ยวข้องกับการสังเคราะห์แคโรทีนอยด์ตั้งอยู่บนโครโมโซมแห่งที่ 6 และ 12 ตามลำดับ ในขณะที่ยีน *PSY3* นั้นเกี่ยวข้องกับการตอบสนองต่อความเครียดจากปัจจัยทางกายภาพ (abiotic stress) (Dibari *et al.*,

2012; Welsch *et al.*, 2008) ดังนั้นในการศึกษานี้จึงศึกษาทั้งหมด 14 ยีน คือ ยีน *PSY1*, *PSY2*, *PDS*, *ZISO*, *ZDS*, *CrtISO*, *LCYb*, *LCYe*, *BHY*, *CYP97A*, *CYP97B*, *CYP97C*, *EHY* และ *ZEP* (ภาพที่ 2)

ผลการค้นหาลำดับเบสที่สามารถแยกกลุ่มข้าวที่มีเยื่อหุ้มเมล็ดสีขาวกับข้าวที่มีเยื่อหุ้มเมล็ดสีดำ พบว่าบริเวณ SNP หรือ InDel ในบริเวณรหัสของยีนจำนวน 5 ยีน มีการเปลี่ยนแปลงลำดับเบสของ 3 ยีน ที่สามารถแยกกลุ่มข้าวเยื่อหุ้มเมล็ดสีขาวพันธุ์ปทุมธานี 1 และกข-แม่โจ้ 2 ที่ตรวจไม่พบสารลูทีนกับข้าวเยื่อหุ้มเมล็ดสีดำพันธุ์ก้าน้อย และกำใหญ่ ที่มีสารลูทีนสูง คือ ยีน *ZISO* จำนวน 4 ตำแหน่ง, ยีน *CYP97A* จำนวน 16 ตำแหน่ง และยีน *CYP97B* จำนวน 3 ตำแหน่ง (ตารางที่ 1) สำหรับยีน *ZDS* พบ SNP ที่สามารถแยกเฉพาะข้าวพันธุ์ปทุมธานี 1 ออกจากข้าวพันธุ์อื่นได้ เช่นเดียวกับยีน *ZEP* พบ SNP ที่แยกข้าวพันธุ์ก้าน้อยออกจากข้าวพันธุ์อื่นได้ ดังนั้นจากการวิเคราะห์ทรานสคริปโตมพบการเปลี่ยนแปลงลำดับเบสบริเวณรหัสในยีน *ZISO*, *CYP97A* และ *CYP97B* ที่สามารถแยกกลุ่มข้าวที่ไม่มีสารลูทีนและมีสารลูทีนได้

การค้นหาลักษณะหลายของลำดับเบสของยีนในวิถีการสังเคราะห์แคโรทีนอยด์จากการวิเคราะห์ลำดับเบสจีโนม

จากผลการวิเคราะห์ WGS ที่วิเคราะห์ลำดับเบสของยีนที่ตั้งอยู่บนโครโมโซมข้าวทั้งจากส่วนที่เป็นยีนรวมถึงบริเวณด้านหน้ายีนและด้านหลังยีนที่อยู่ห่างจากยีนไม่เกิน 50 กิโลเบส ของยีนในวิถีการสังเคราะห์แคโรทีนอยด์ทั้ง 14 ยีน ผลการค้นหามิพบความหลากหลายของลำดับเบสในยีน *ZDS* และยีน *ZEP* ที่สามารถใช้แยกกลุ่มข้าวที่มีเยื่อหุ้มเมล็ดสีขาวออกจากสีดำ เช่นเดียวกับผลการค้นหาจากการวิเคราะห์ทรานสคริปโตม สำหรับอีก 12 ยีนนั้นพบความหลากหลายของลำดับเบสที่สามารถแยกกลุ่มได้ ได้แก่ ยีน *PSY1* พบ 1 ตำแหน่ง ที่บริเวณหลังยีน, ยีน *PSY2* พบ 2 ตำแหน่ง ที่บริเวณหน้ายีน, ยีน *PDS* พบ 39 ตำแหน่ง ที่บริเวณหน้ายีน 7 ตำแหน่ง ส่วนที่เป็นยีน 2 ตำแหน่ง และบริเวณด้านหลังยีน 30 ตำแหน่ง, ยีน *ZISO* พบ 865 ตำแหน่ง ที่บริเวณหน้ายีน 478 ตำแหน่ง ส่วนที่เป็นยีน 13 ตำแหน่ง และบริเวณหลังยีน 374 ตำแหน่ง, ยีน *CrtISO* พบ 6 ตำแหน่ง ที่บริเวณหน้ายีน, ยีน *LCYb* พบ 60 ตำแหน่ง ที่บริเวณหลังยีน, ยีน *LCYe* พบ 3 ตำแหน่ง ที่บริเวณหน้ายีน 1 ตำแหน่ง ส่วนที่เป็นยีน 1 ตำแหน่ง และบริเวณหลังยีน 1 ตำแหน่ง, ยีน *EHY* พบ 3 ตำแหน่ง ที่บริเวณหน้ายีน 2 ตำแหน่ง และบริเวณหลังยีน 1 ตำแหน่ง, ยีน *BHY* พบ 5 ตำแหน่ง ที่บริเวณหน้ายีน 4 ตำแหน่ง และบริเวณหลังยีน 1 ตำแหน่ง, ยีน *CYP97A* พบ 398 ตำแหน่ง ที่บริเวณหน้ายีน 112 ตำแหน่ง ส่วนที่เป็นยีน 67 ตำแหน่ง และบริเวณหลังยีน 219 ตำแหน่ง, ยีน *CYP97B* พบ 783 ตำแหน่ง ที่บริเวณหน้ายีน 620 ตำแหน่ง ส่วนที่เป็นยีน 20 ตำแหน่ง และบริเวณหลังยีน 139 ตำแหน่ง และยีน *CYP97C* พบ 3 ตำแหน่ง ที่บริเวณหน้ายีน ดังแสดงในตารางที่ 1 ความหลากหลายของลำดับเบสที่สามารถใช้แยกกลุ่มข้าวเยื่อหุ้มเมล็ดสีขาวที่ไม่มีสารลูทีนออกจากเยื่อหุ้มเมล็ดสีดำที่มีสารลูทีนสูง ซึ่งพบจำนวนมากในยีน *ZISO*, *CYP97A* และ *CYP97B* แสดงถึงความหลากหลายของลำดับเบสของยีนดังกล่าวในข้าวไทย

การค้นหาคำแหน่งและรูปแบบความหลากหลายของลำดับเบสจากฐานข้อมูล Gramene

ในการค้นหาคำแหน่งและผลกระทบของความหลากหลายของลำดับเบสจะเลือกเฉพาะการเปลี่ยนแปลงในส่วนที่เป็นยีน เนื่องจากจะมีโอกาสเปลี่ยนแปลงลำดับกรดอะมิโน ส่งผลกับการทำงานของเอนไซม์ แล้วจะส่งผลกับการสังเคราะห์สารลูทีนที่แตกต่างกันในข้าวทั้ง 2 กลุ่ม จึงนำความหลากหลายของลำดับเบสในส่วนที่เป็นยีนวิถีการสังเคราะห์แคโรทีนอยด์



จำนวน 12 ยีน ที่ได้จากการวิเคราะห์ลำดับเบสจากทรานสคริปโตมกับการวิเคราะห์ลำดับเบสจีโนม (ตารางที่ 1) มาค้นหาตำแหน่งของยีนบนโครโมโซมและตำแหน่งการเปลี่ยนแปลงลำดับเบส การเปลี่ยนแปลงของชนิดกรดอะมิโนที่เกิดขึ้นมาจากการเปลี่ยนแปลงนั้นเป็นผลจากการเปรียบเทียบกับลำดับเบสของ *O. sativa* จำนวน 6 พันธุ์ คือ Nipponbare (*O. Sativa* L. ssp *japonica*), IR36 (*O. sativa* L. ssp *indica*), 93-11 (*O. sativa* L. ssp *indica*), Dular (*O. sativa* L. ssp *indica*), Tsipala 421 (*O. sativa* L.), IRIS_313-9516 (*O. sativa* L.) และข้าวป่า (*O. rufipogon*) ในฐานะข้อมูล Gramene พบความหลากหลายของลำดับเบสส่วนที่เป็นยีนวิธีการสังเคราะห์แคโรทีนอยด์ในรูปแบบ SNP หรือ InDel จำนวน 5 ยีน ได้แก่ ยีน *PDS* พบ 2 ตำแหน่ง ที่บริเวณอินทรอน, ยีน *ZISO* พบ 13 ตำแหน่ง ที่บริเวณ 5' UTR เอกซอน อินทรอน และ 3' UTR, ยีน *LCYe* พบ 1 ตำแหน่ง ที่บริเวณอินทรอน, ยีน *CYP97A* พบ 69 ตำแหน่ง ที่บริเวณ 5' UTR เอกซอน อินทรอน จุดตัด และ 3' UTR และยีน *CYP97B* พบ 20 ตำแหน่ง ที่บริเวณเอกซอน อินทรอน จุดตัด และ 3' UTR (ตารางที่ 2)

หากพิจารณาเฉพาะความหลากหลายของลำดับเบสในบริเวณเอกซอนและจุดตัด ที่จะส่งผลกระทบต่อกรดอะมิโน พบการเปลี่ยนแปลงลำดับเบสเหล่านี้ใน 3 ยีน คือ ยีน *ZISO* (*Os12g0405200*) พบ SNP จำนวน 2 ตำแหน่ง ที่ เอกซอน 3 (1 ตำแหน่ง) และ เอกซอน 4 (1 ตำแหน่ง), ยีน *CYP97A* (*Os02g0817900*) พบ SNP จำนวน 23 ตำแหน่ง ที่ เอกซอน 1 (7 ตำแหน่ง) เอกซอน 2 (5 ตำแหน่ง) เอกซอน 4 (1 ตำแหน่ง) เอกซอน 5 (1 ตำแหน่ง) เอกซอน 15 (2 ตำแหน่ง) จุดตัด (3 ตำแหน่ง) และพบ InDel จำนวน 9 ตำแหน่ง ที่ เอกซอน 1 (6 ตำแหน่ง) และจุดตัด (3 ตำแหน่ง) และ ยีน *CYP97B* (*Os02g0173100*) พบ SNP จำนวน 3 ตำแหน่ง ที่ เอกซอน 4 (1 ตำแหน่ง) เอกซอน 7 (1 ตำแหน่ง) และจุดตัด (1 ตำแหน่ง) โดยความหลากหลายของลำดับเบสแบบ SNP และ InDel ในบริเวณยีนที่ส่งผลกระทบต่อเปลี่ยนแปลงลำดับกรดอะมิโน เช่น silent, missense และ nonsense ในแต่ละเอกซอนของยีนแสดงในตารางที่ 2 เช่น ในเอกซอนที่ 5 ของยีน *CYP97A* พบ SNP (T/C) ที่ตำแหน่ง 35,098,135 ทำให้เกิดการเปลี่ยนแปลงลำดับกรดอะมิโนแบบ missense ส่วน SNP (A/T) ที่ตำแหน่ง 35,098,135 ทำให้เกิดการเปลี่ยนแปลงลำดับกรดอะมิโนแบบ nonsense



ตารางที่ 1 การวิเคราะห์การเปลี่ยนแปลงลำดับเบสของยีนในวิธีการสังเคราะห์แคโรทีนอยด์ที่ได้จากเทคนิค RNA-seq และ WGS

ยีน	เอนไซม์; Accessions ID	ความยาว ส่วนที่เป็น ยีน	จำนวนตำแหน่งการเปลี่ยนแปลงลำดับเบส			
			จากเทคนิค RNA-seq	จากเทคนิค WGS		
			ส่วนที่เป็นยีน (ตำแหน่ง)	ด้านหน้ายีน (ตำแหน่ง)	ส่วนที่เป็นยีน (ตำแหน่ง)	ด้านหลังยีน (ตำแหน่ง)
<i>PSY1</i>	Phytoene synthase 1; <i>Os06g072900</i>	3,956 bp	-	-	-	1
<i>PSY2</i>	Phytoene synthase 2; <i>Os12g0626400</i>	3,112 bp	-	2	-	-
<i>PDS</i>	Phytoene desaturase; <i>Os03g0184000</i>	4,416 bp	-	7	2	30
<i>ZISO</i>	15-cis-zeta-carotene isomerase; <i>Os12g0405200</i>	3,307 bp	4	478	13	374
<i>ZDS</i>	zeta-carotene desaturase; <i>Os07g0204900</i>	7,700 bp	-	-	-	-
<i>CrtISO</i>	Carotenoid isomerase; <i>Os11g0572700</i>	4,318 bp	-	6	-	-
<i>LCYb</i>	Lycopene beta cyclase; <i>Os02g0190600</i>	1,918 bp	-	-	-	60
<i>LCYe</i>	Lycopene epsilon cyclase; <i>Os01g0581300</i>	3,633 bp	-	1	1	1
<i>EHY, HYD3</i>	carotene epsilon-ring hydroxylase; <i>Os03g0125100</i>	2,700 bp	-	2	-	1
<i>BHY, HYD2</i>	carotene beta-ring hydroxylase; <i>Os10g0533500</i>	2,026 bp	-	4	-	1
<i>CYP97A, Lut5</i>	beta-ring hydroxylase; cytochrome P450 97A; <i>Os02g0817900</i>	8,098 bp	16	112	67	219
<i>CYP97B</i>	cytochrome P450 97B; <i>Os02g0173100</i>	6,717 bp	3	620	20	139
<i>CYP97C, Lut1</i>	carotenoid epsilon-ring hydroxylase, cytochrome P450 97C; <i>Os10g0546600</i>	4,717 bp	-	3	-	-
<i>ZEP</i>	zeaxanthin epoxidase; <i>Os04g0448900</i>	6,973 bp	-	-	-	-

หมายเหตุ: ความหลากหลายของลำดับเบสที่แสดงจะให้ข้าวที่มีเชื้อหุ้มเมล็ดสีขาว (พันธุ์ปทุมธานี 1 และพันธุ์เกษมใจ 2) เป็นหลักเทียบกับการเปลี่ยนแปลงที่พบในข้าวที่มีเชื้อหุ้มเมล็ดสีดำ (พันธุ์ก้าน้อย และพันธุ์กำใหญ่)



ตารางที่ 2 ตำแหน่งของยีนในวิถีการสังเคราะห์แคโรทีนอยด์ ตำแหน่งและรูปแบบของความหลากหลายของลำดับเบส บริเวณที่ได้รับผลกระทบและลำดับเบสของการเปลี่ยนแปลงที่พบในบริเวณยีนจากการวิเคราะห์ข้อมูลทรานสคริปโตมและลำดับเบสจีโนมจากข้าว 4 พันธุ์

ชื่อยีน	ตำแหน่งยีน	การเปลี่ยนแปลงของลำดับเบส												
		ตำแหน่ง	รูปแบบ	บริเวณที่ได้รับผลกระทบ*	Allele**	ข้อมูล Transcriptome sequencing (RNA-seq)				ข้อมูล Whole Genome Sequencing (WGS)				
						PTT1	RD-MJU2	KN	KY	PTT1	RD-MJU2	KN	KY	
PDS	Chr03:4,410,090..4,414,505 (+ strand)	4,411,054	Insertion	intron	G/GTGCATTGCGCAGC CGCAATTCCTTTCCCC AGTCAGCCACCTTTTCG GTGCAGGT					GTGCATTC	GTGCATTCG	G	G	
		4,412,597	SNP	intron	C/T					T	T	C	C	
ZISO	Chr12:12,204,405..12,207,711 (+ strand)	12,204,484	Deletion	5' UTR	C/A/C					C	C	CA	CA	
		12,204,486	Deletion	5' UTR	CCACCACCACCACCA/ C					C	C	CCACCA	CCACCA	
													CCACCA	CCACCA
													CCA	CCA
		12,205,963	Insertion	intron	A/AAC					AAC	AAC	A	A	
		12,206,218	Insertion	intron	C/CTGTG					CTGTG	CTGTG	C	C	
		12,206,313	Deletion	intron	GAC/G					G	G	GAC	GAC	
		12,206,410	Insertion	intron	A/AT					AT	AT	A	A	
		12,206,881	SNP	exon3 = silent (L223)	T/C		C	C	T	T	C	C	T	T
		12,207,012	SNP	intron	T/A					A	A	T	T	
		12,207,052	Deletion	intron	GTTT/G					G	G	GTTT	GTTT	
		12,207,120	SNP	intron	T/C					C	C	T	T	
		12,207,440	SNP	exon4 = missense (A360Q)	C/A		A	A	C	C	A	A	C	C
12,207,619	SNP	3' UTR	T/A		A	A	T	T	A	A	T	T		
12,207,707	Insertion	downstream gene	A/ATAAAATGCACCAGT TCATCAAACACCTCTCT CTCCTGACCCTTAGAT		AATATAAA	AATATAAA	A	A	?	A	A	A		
LCYe	Chr01:22,535,013..22,538,645 (+ strand)	22,535,599	SNP	intron	A/T					T	T	A	A	
CYP97A, Lut5	Chr02:35,091,194..35,099,291 (- strand)	35,092,097	SNP	3' UTR	G/T					G	G	T	T	
		35,092,112	SNP	3' UTR	C/T					C	C	T	T	
		35,092,234	SNP	3' UTR	C/T					C	C	T	T	
		35,092,952	SNP	3' UTR	T/C		C	C	T	T	C	C	T	T
		35,093,037	SNP	3' UTR	G/T					G	G	T	T	
		35,093,684	SNP	intron	T/A		A	A	T	T	A	A	T	T
		35,093,715	SNP	intron	T/C		T	T	C	C	T	T	C	C
		35,093,967	Deletion	intron	GAA/G/GA		GA/G	GA	GA/G	GA/G	GA	GA	G	G



ชื่อยีน	ตำแหน่งยีน	การเปลี่ยนแปลงของลำดับเบส											
		ตำแหน่ง	รูปแบบ	บริเวณที่ได้รับผลกระทบ*	Allele**	ข้อมูล Transcriptome sequencing (RNA-seq)				ข้อมูล Whole Genome Sequencing (WGS)			
						PTT1	RD-MJU2	KN	KY	PTT1	RD-MJU2	KN	KY
		35,094,024	SNP	intron	T/G	T	T	G	G	T	T	G	G
		35,094,096	SNP	splice region	C/A	A	A	C	C	A	A	C	C
		35,094,201	SNP	exon15 = missense (V623I)	C/T	T	T	C	C	T	T	C	C
		35,094,346	SNP	exon15 = missense (Q574N)	C/A	A	A	C	C	A	A	C	C
		35,094,610	SNP	intron	C/T					T	T	C	C
		35,094,861	SNP	intron	G/A					A	A	G	G
		35,095,097	SNP	intron	A/G					A	A	G	G
		35,095,531	SNP	intron	C/T					C	C	T	T
		35,095,669	SNP	intron	A/G					G	G	A	A
		35,097,198	SNP	splice region	A/G	G	G	A	A	G	G	A	A
		35,097,919	Deletion	intron	TGA/T					TGA	TGA	T	T
		35,097,922	Insertion	intron	A/ACG					A	A	ACG	ACG
		35,097,946	Deletion	intron	CA/C					CA	CA	C	C
		35,098,095	SNP	exon5 = missense (T261S)	T/A					T	T	A	A
		35,098,099	SNP	exon5 = silent (A259)	T/C					T	T	C	C
		35,098,135	SNP	exon5 = nonsense (Y247*)	A/T					A	A	T	T
		35,098,137	SNP	exon5 = missense (Y247D)	A/C					A	A	C	C
		35,098,161	SNP	exon5 = missense (T239A)	T/C					C	C	T	T
		35,098,188	SNP	intron	C/G					C	C	G	G
		35,098,193	SNP	intron	C/G					C	C	G	G
		35,098,197	SNP	intron	C/T					C	C	T	T
		35,098,210	SNP	intron	T/A					T	T	A	A
		35,098,232	SNP	intron	C/A					C	C	A	A
		35,098,234	SNP	intron	G/A					G	G	A	A
		35,098,348	SNP	exon4 = silent (A203)	T/C	T	T	C	C	T	T	C	C
		35,098,374	SNP	intron	A/T					A	A	T	T
		35,098,385	SNP	intron	G/C					G	G	C	C
		35,098,398	SNP	intron	T/C					T	T	C	C
		35,098,399	SNP	intron	A/G					A	A	G	G
		35,098,433	Insertion	intron	T/TG					T	T	TG	TG
		35,098,441	Deletion	splice region	CA/C					CA	CA	C	C
		35,098,520	Deletion	splice region	TG/T					TG	TG	T	T
		35,098,543	SNP	intron	A/T					A	A	T	T
		35,098,550	SNP	intron	G/A					G	G	A	A
		35,098,566	SNP	intron	C/A					C	C	A	A
		35,098,573	Deletion	intron	GGGAGTTGC/G					GGGAGTTG	GGGAGTTG	G	G
										C	C		



ชื่อยีน	ตำแหน่งยีน	การเปลี่ยนแปลงของลำดับเบส											
		ตำแหน่ง	รูปแบบ	บริเวณที่ได้รับผลกระทบ*	Allele**	ข้อมูล Transcriptome sequencing (RNA-seq)				ข้อมูล Whole Genome Sequencing (WGS)			
						PTT1	RD-MJU2	KN	KY	PTT1	RD-MJU2	KN	KY
		35,098,610	SNP	splice region	T/C					T	T	C	C
		35,098,651	SNP	exon2 = silent (Y164)	A/G	A	A	G	G	A	A	G	G
		35,098,657	SNP	exon2 = silent (L162)	A/G	A	A	G	G	A	A	G	G
		35,098,744	SNP	exon2 = silent (G133)	G/A	G	G	A	A	G	G	A	A
		35,098,804	SNP	exon2 = silent (L113)	A/C	A	A	C	C	A	A	C	C
		35,098,811	SNP	exon2 = missense (V111A)	A/G	A	A	G	G	A	A	G	G
		35,098,839	Deletion	intron	AG/A					AG	AG	A	A
		35,098,870	SNP	intron	A/C					A	A	C	C
		35,098,887	SNP	intron	G/C					G	G	C	C
		35,098,896	Deletion	splice region	TGAGAGAGAGAG/TT GAGAGA					TGAGAGA	TGAGAGA	T	T
		35,099,026	Deletion	exon1 = deletion	AGCCGTC/A					AGCCGTC	AGCCGTC	A	A
		35,099,050	Insertion	exon1 = insertion	C/CGTT					C	C	CGTT	CGTT
		35,099,060	SNP	exon1 = silent (S60)	C/G					C	C	G	G
		35,099,063	SNP	exon1 = silent (A59)	G/C					G	G	C	C
		35,099,068	SNP	exon1 = missense (A58S)	C/A					C	C	A	A
		35,099,077	SNP	exon1 = missense (L55I)	G/T					G	G	T	T
		35,099,093	Insertion	exon1 = insertion	A/AC					A	A	AC	AC
		35,099,094	Insertion	exon1 = insertion	G/GAC					G	G	GAC	GAC
		35,099,096	SNP	exon1 = silent (S48)	T/G					T	T	G	G
		35,099,107	Deletion	exon1 = deletion	CAGCAGCAGCA/C					CAGCAGCA GCAGA	CAGCAGCA GCAGA	C	C
		35,099,150	SNP	exon1 = silent (G30)	G/A	G	G	A	A	G	G	A	A
		35,099,173	Insertion	exon1 = insertion	G/GAGC	G	G	GAGC	GAGC				
		35,099,189	SNP	exon1 = silent (Ser17)	G/A	A	A	G	G				
		35,099,274	SNP	5' UTR	A/G					A	A	G	G
		35,099,287	SNP	5' UTR	G/A					G	G	A	A
CYP97B	Chr02:3,995,980..4,002,696 (+ strand)	3,996,654	SNP	splice region	C/T					T	T	C	C
		3,996,743	Deletion	intron	AT/A					A	A	AT	AT
		3,997,137	SNP	intron	G/A					A	A	G	G
		3,997,211	SNP	intron	C/G					G	G	C	C
		3,997,360	SNP	intron	C/T					T	T	C	C
		3,997,758	SNP	intron	T/A					A	A	T	T
		3,998,726	SNP	intron	T/A					A	A	T	T
		3,998,826	SNP	exon4 = missense (N179Q)	T/A	A	A	T	T	A	A	T	T



ชื่อยีน	ตำแหน่งยีน	การเปลี่ยนแปลงของลำดับเบส											
		ตำแหน่ง	รูปแบบ	บริเวณที่ได้รับผลกระทบ*	Allele**	ข้อมูล Transcriptome sequencing (RNA-seq)				ข้อมูล Whole Genome Sequencing (WGS)			
						PTT1	RD-MJU2	KN	KY	PTT1	RD-MJU2	KN	KY
		3,999,100	SNP	intron	C/T					C	A	T	T
		3,999,642	SNP	exon7 = missense (R281H)	G/A					A	A	G	G
		4,000,036	SNP	3' UTR	G/T	G	G	T	T	G	A	T	T
		4,000,560	SNP	intron	G/A					A	A	G	G
		4,000,583	SNP	intron	G/A					G	A	A	A
		4,000,743	SNP	intron	T/A					A	A	T	T
		4,000,869	SNP	intron	T/C					C	A	T	T
		4,000,891	SNP	intron	G/A					A	A	G	G
		4,001,036	SNP	intron	C/T					T	A	C	C
		4,001,826	Deletion	intron	TCGTGACAACATCTCC/ T					T	T	TCGTGAC	TCGTGAC
												AACATCT	AACATCT
												CC	CC
		4,002,179	SNP	intron	T/A					A	A	T	T
		4,002,394	SNP	3' UTR	T/A	A	A	T	T	A	A	T	T

หมายเหตุ: - PTT1 คือข้าวพันธุ์ปทุมธานี 1, RD-MJU2 คือข้าวพันธุ์กข-แม่ใจ 2, KN คือข้าวพันธุ์ก้าน้อย และ KY คือข้าวพันธุ์กำใหญ่

* ผลจากการเปรียบเทียบกับลำดับเบสของ *O. sativa* จำนวน 6 พันธุ์ คือ Nipponbare (*O. sativa* L. ssp *japonica*), IR36 (*O. sativa* L. ssp *indica*), 93-11 (*O. sativa* L. ssp *indica*), Dular (*O. sativa* L. ssp *indica*), Tsipala 421 (*O. sativa* L.), IRIS_313-9516 (*O. sativa* L.) และข้าวป่า (*O. rufipogon*) ในฐานข้อมูล Gramene

** Allele คือ ลำดับเบสที่พบใน Nipponbare (Ref. sequence) / ลำดับเบสที่พบในข้าวพันธุ์ที่นำมาศึกษาที่แตกต่างจาก Nipponbare

วิจารณ์ผลการวิจัย

จากการวิเคราะห์ด้วยเทคนิค WGS พบความหลากหลายของลำดับเบสส่วนที่เป็นยีนมากกว่าการวิเคราะห์ด้วยเทคนิค RNA-seq เนื่องจากการหาลำดับเบสด้วยเทคนิค WGS นั้นสามารถตรวจพบความหลากหลายของลำดับเบสในบริเวณอินทรอน บริเวณด้านหน้าและหลังของส่วนที่เป็นยีนได้ด้วย จึงทำให้พบความหลากหลายของลำดับเบสมากกว่า เช่นเดียวกับการศึกษาในข้าวพันธุ์ Chuanyou6203 ที่เปรียบเทียบการวิเคราะห์ลำดับเบสทั้ง 2 วิธี ผลการวิเคราะห์ที่ได้จากเทคนิค WGS จะพบความหลากหลายของลำดับเบสมากกว่าผลจากเทคนิค RNA-seq (Ren *et al.*, 2020)

ความหลากหลายของลำดับเบสที่พบในส่วนที่เป็นยีน จะทำให้เกิดการเปลี่ยนแปลงลำดับกรดอะมิโน ซึ่งจะส่งผลให้เอนไซม์เหล่านี้ในข้าวเยื่อหุ้มเมล็ดสีขาวและสีดำทำงานได้แตกต่างกัน โดยพบมากในยีน *CYP97A* (*lut5*) เช่น การเพิ่มขึ้นของเบส (Insertion) ในเอกซอน 1 ตำแหน่ง 35,099,173 หรือ SNP ที่จุดตัดในอินทรอน 5 ตำแหน่ง 35,097,198 น่าจะส่งผลให้เอนไซม์ beta-ring hydroxylase ของข้าวเยื่อหุ้มเมล็ดสีขาวและสีดำทำงานได้แตกต่างกัน แล้วกระทบกับการสังเคราะห์ลูทีนในแบบที่เรียดัดแปลงพันธุกรรมลดลง 12-24 เปอร์เซ็นต์ เมื่อเปรียบเทียบกับยีนปกติ (Lv *et al.*, 2012) ซึ่งยีน *CYP97A* และ *CYP97B* มีลำดับเบสคล้ายกันและทำงานในกระบวนการเดียวกัน (Niu *et al.*, 2020; Villa-Rivera & Ochoa-Alejo, 2020) จึงส่งผลให้พบความหลากหลายของลำดับเบสในทั้ง 2 ยีน มีจำนวนมาก ส่วนความหลากหลายของลำดับเบสที่พบในยีน *ZISO* น่าจะมีผลกระทบกับการสังเคราะห์ลูทีนในวิถีการสังเคราะห์แคโรทีนอยด์ เช่น การขาดหายในบริเวณเอกซอนหรืออินทรอน อาจจะมีผลกับการทำงานของเอนไซม์ เหมือนการศึกษาการขาดหายของลำดับเบสของเอกซอนและอินทรอนของยีน *ZISO* ด้วยการกลายพันธุ์ในข้าวโพดและอะราบิดอปซิส พบว่าต้นข้าวโพดและอะราบิดอปซิสที่มียีนกลายพันธุ์จะมีปริมาณแคโรทีนอยด์ ลดลง 20-30 เปอร์เซ็นต์ของต้นที่มียีน *ZISO* ปกติ (Chen *et al.*, 2010) แสดงให้เห็นว่ายีน *ZISO* นั้นมีความสำคัญและหากมีการเปลี่ยนแปลงลำดับเบสอาจจะส่งผลต่อระดับเอนไซม์วิถีการสังเคราะห์แคโรทีนอยด์ (Chen *et al.*, 2010; Wurtzel, 2019)

ความหลากหลายของลำดับเบสที่พบในส่วนด้านหน้ายีน อาจจะมีผลกระทบกับการแสดงออกของ ยีน *PSY2*, *PDS*, *ZISO*, *CrtISO*, *LCYe*, *EHY*, *BHY*, *CYP97A*, *CYP97B* และ *CYP97C* โดยการแสดงออกของยีนสร้างสารลูทีนอาจจะเกี่ยวข้องกับการจับของโปรตีน transcription factor ที่บริเวณโปรโมเตอร์ของยีน สอดคล้องกับการศึกษาของ (Oikawa *et al.*, 2015) ที่ศึกษาโปรโมเตอร์ในข้าวเยื่อหุ้มเมล็ดดำของยีน *B2* พบการเพิ่มขึ้นของลำดับเบสที่เกิดการแทรกของลำดับเบสซ้ำขนาด 11 กิโลเบส เมื่อเปรียบเทียบกับโปรโมเตอร์ของข้าวเยื่อหุ้มเมล็ดสีขาว ส่งผลให้การแสดงออกของยีน *B2* ในข้าวเยื่อหุ้มเมล็ดดำสูงกว่าในข้าวเยื่อหุ้มเมล็ดสีขาว สอดคล้องกับการศึกษา Chetty & Chrungoo (2020) และ Chetty *et al.* (2019) ที่พบการแสดงออกของยีน *PSY1*, *PDS* และ *LCYb* สูงในข้าวเยื่อหุ้มเมล็ดดำที่มีการสะสมของสารในวิถีแคโรทีนอยด์สูง รวมถึงการทำงานร่วมกันของยีน *MADS26* และ *R2R3MYB* ซึ่งเป็นยีนกลุ่ม transcription factor กับยีนในวิถีแคโรทีนอยด์ นอกจากนี้ความหลากหลายของลำดับเบสบริเวณ 5' UTR ของยีน *ZISO* และ *CYP97A* อาจจะมีผลทำให้ระดับการแสดงออก



ของยีนลดลง เช่นเดียวกับการขาดหายไปของบริเวณ 5' UTR ด้วยการแก้ไขจีโนม ในยีน *CrtISO* ของมะเขือเทศ ที่ส่งผลให้ยีนแสดงออกลดลง จนทำให้สีของมะเขือเทศเปลี่ยนแปลงไป (Jayaraj *et al.*, 2021) โดยความหลากหลายของลำดับเบสบริเวณด้านหน้าและด้านหลังยีนสามารถนำไปออกแบบเป็นเครื่องหมายดีเอ็นเอ ที่สามารถแยกความแตกต่างของการสังเคราะห์สารลูทีนของข้าวไทยได้ เช่นเดียวกับส่วนที่เป็นยีน ดังที่มีการศึกษา ยีน *PPR9* ที่เป็นยีนแก้ความเป็นหมันของเรณูในระบบ WA-CMS ที่ออกแบบเครื่องหมายจากการเปลี่ยนแปลงลำดับเบสบริเวณด้านหน้าและด้านหลังยีน 50 กิโลเบส พบว่าเครื่องหมาย RMS-PPR9-1 สามารถแยกข้าวกลุ่มที่มีความสามารถแก้ความเป็นหมันของเรณู (R-line) กับข้าวกลุ่มรักษาความเป็นหมัน (B-line) ได้ (Pranathi *et al.*, 2016)

สรุปผลการวิจัย

การสังเคราะห์สารแคโรทีนอยด์ของข้าวนั้น เป็นการทำงานร่วมกันของหลายยีนในวิถีแคโรทีนอยด์ ซึ่งยังไม่มีรายงานยีนหลักที่ควบคุมการสังเคราะห์ ดังนั้นงานวิจัยนี้จึงเริ่มต้นศึกษาความหลากหลายของลำดับเบสของยีนในวิถีการสังเคราะห์แคโรทีนอยด์จำนวน 14 ยีน จากการวิเคราะห์ข้อมูลลำดับเบสทรานสคริปโตมและลำดับเบสจีโนม พบลำดับเบสที่สามารถใช้แยกกลุ่มข้าวเยื่อหุ้มเมล็ดสีขาวที่ไม่มีสารลูทีนจากข้าวเยื่อหุ้มเมล็ดสีดำที่มีสารลูทีนสูงจำนวน 12 ยีน เป็นการเปลี่ยนแปลงลำดับเบสในส่วนที่เป็นยีนจำนวน 5 ยีน ได้แก่ ยีน *PDS*, *ZISO*, *LCYe*, *CYP97A* และ *CYP97B* โดยพบการเปลี่ยนแปลงลำดับเบสทั้งในส่วนที่เป็นยีน บริเวณด้านหน้าและด้านหลังของยีนจำนวนมากในยีน *ZISO*, *CYP97A* และ *CYP97B* จึงน่าจะเป็นยีนสำคัญในการสังเคราะห์สารลูทีนของพันธุ์ข้าวไทยที่ใช้ในการศึกษานี้ เนื่องจากความหลากหลายของลำดับเบสเหล่านี้่อาจจะส่งผลกระทบต่อลำดับกรดอะมิโนของโปรตีนหรืออาจจะกระทบกับการจับของ transcription factor ซึ่งความหลากหลายของลำดับเบสที่พบในแต่ละยีนนั้น สามารถนำไปพัฒนาเป็นเครื่องหมายดีเอ็นเอใช้ในการคัดเลือกเพื่อการปรับปรุงพันธุ์ข้าวไทยให้มีสารลูทีนสูงต่อไป

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนอุดหนุนวิจัยจากสำนักงานคณะกรรมการส่งเสริมวิทยาศาสตร์ วิจัยและนวัตกรรม (สกสว.) ผ่านมหาวิทยาลัยแม่โจ้ ประจำปีงบประมาณ 2563

เอกสารอ้างอิง

Al-Babili, S., & Beyer, P. (2005). Golden Rice—five years on the road—five years to go? *Trends in plant science*, 10(12), 565-573.



- Ashokkumar, K., Govindaraj, M., Adhimoolam Karthikeyan, V., & Warkentin, T. D. (2020). Genomics-integrated breeding for carotenoids and folates in staple cereal grains to reduce malnutrition. *Frontiers in genetics*, 11.
- Badro, H., Ndjiondjop, M.-N., Furtado, A., & Henry, R. (2020). Sequence variants linked to key traits in interspecific crosses between African and Asian rice. *Plants*, 9(12), 1653.
- Bai, C., Capell, T., Berman, J., Medina, V., Sandmann, G., Christou, P., & Zhu, C. (2016). Bottlenecks in carotenoid biosynthesis and accumulation in rice endosperm are influenced by the precursor-product balance. *Plant Biotechnology Journal*, 14(1), 195-205.
- Chen, Y., Li, F., & Wurtzel, E. T. (2010). Isolation and characterization of the *Z-ISO* gene encoding a missing component of carotenoid biosynthesis in plants. *Plant Physiology*, 153(1), 66-79.
- Chettry, U., & Chrungoo, N. K. (2020). A multifocal approach towards understanding the complexities of carotenoid biosynthesis and accumulation in rice grains. *Briefings in functional genomics*, 19(4), 324-335.
- Chettry, U., Chrungoo, N. K., & Kulkarni, K. (2019). Comparative transcriptomics approach in elucidation of carotenoid biosynthesis regulation in grains of rice (*Oryza sativa* L.). *Scientific reports*, 9(1), 1-12.
- Davidson, R. M., Gowda, M., Moghe, G., Lin, H., Vaillancourt, B., Shiu, S. H., . . . Robin Buell, C. (2012). Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. *The Plant Journal*, 71(3), 492-502.
- Dibari, B., Murat, F., Chosson, A., Gautier, V., Poncet, C., Lecomte, P., . . . Blanco, A. (2012). Deciphering the genomic structure, function and evolution of carotenogenesis related phytoene synthases in grasses. *Bmc Genomics*, 13(1), 1-14.



- Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., . . . Estelle, J. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings in bioinformatics*, 14(6), 671-683.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., . . . Zeng, Q. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), 644-652.
- Janphen, K., Sangtong, V., Sakulsingharoj, C., Leethanatudom, S., Phaisansuthichol, S., & Pongjaroenkit, S. (2019). Expression Analysis of Genes in Carotenoid Biosynthesis Pathway from Thai Rice with Different Pericarp Colors. *Paper presented at the In Proceeding National Genetics Conference (NGC2019) "Genetics for Sustainable Development"*, Burapha University. (In Thai)
- Jayaraj, K. L., Thulasidharan, N., Antony, A., John, M., Augustine, R., Chakravarty, N., ... Thomas, G. (2021). Targeted editing of tomato carotenoid isomerase reveals the role of 5' UTR region in gene expression regulation. *Plant Cell Reports*, 40(4), 621-635.
- Kim, J. K., Lee, S. Y., Chu, S. M., Lim, S. H., Suh, S.-C., Lee, Y.-T., . . . Ha, S.-H. (2010). Variation and correlation analysis of flavonoids and carotenoids in Korean pigmented rice (*Oryza sativa* L.) cultivars. *Journal of agricultural and food chemistry*, 58(24), 12804-12809.
- Kongkachuichai, R., & Charoensir, R. (2010). Nutritional values of land-race rice in the land reform area of Ampur Kudchum, Yasothon Province. *Journal of Nutrition Association of Thailand*, 45(2), 15-32. (in Thai)
- Li, J., Wang, S., Yu, J., Wang, L., & Zhou, S. (2013). A modified CTAB protocol for plant DNA extraction. *Chinese Bulletin of Botany*, 48(1), 72.
- Lv, M.-Z., Chao, D.-Y., Shan, J.-X., Zhu, M.-Z., Shi, M., Gao, J.-P., & Lin, H.-X. (2012). Rice carotenoid β -ring hydroxylase CYP97A4 is involved in lutein biosynthesis. *Plant and Cell Physiology*, 53(6), 987-1002.



Niu, G., Guo, Q., Wang, J., Zhao, S., He, Y., & Liu, L. (2020). Structural basis for plant lutein biosynthesis from α -carotene. *Proceedings of the National Academy of Sciences of the United States of America*, 117(25), 14150-14157.

Oikawa, T., Maeda, H., Oguchi, T., Yamaguchi, T., Tanabe, N., Ebana, K., . . . Izawa, T. (2015). The birth of a black rice gene and its local spread by introgression. *The Plant Cell*, 27(9), 2401-2414.

Pereira-Caro, G., Cros, G., Yokota, T., & Crozier, A. (2013). Phytochemical profiles of black, red, brown, and white rice from the Camargue region of France. *Journal of agricultural and food chemistry*, 61(33), 7976-7986.

Pranathi, K., Viraktamath, B., Neeraja, C., Balachandran, S., Rao, P. K., Revathi, P.,... Naik, S. B. (2016). Development and validation of candidate gene-specific markers for the major fertility restorer genes, *Rf4* and *Rf3* in rice. *Molecular Breeding*, 36(10), 1-14.

Ren, J., Zhang, F., Gao, F., Zeng, L., Lu, X., Zhao, X., ...Dai, M. (2020). Transcriptome and genome sequencing elucidates the molecular basis for the high yield and good quality of the hybrid rice variety Chuanyou6203. *Scientific reports*, 10(1), 1-15.

Stanley, L., & Yuan, Y.-W. (2019). Transcriptional regulation of carotenoid biosynthesis in plants: so many regulators, so little consensus. *Frontiers in plant science*, 10, 1017.

Villa-Rivera, M. G., & Ochoa-Alejo, N. (2020). Chili pepper carotenoids: nutraceutical properties and mechanisms of action. *Molecules*, 25(23), 5573.

Watkins, J. L., & Pogson, B. J. (2020). Prospects for carotenoid biofortification targeting retention and catabolism. *Trends in plant science*, 25(5), 501-512.



Welsch, R., Wüst, F., Bär, C., Al-Babili, S., & Beyer, P. (2008). A third phytoene synthase is devoted to abiotic stress-induced abscisic acid formation in rice and defines functional diversification of phytoene synthase genes. *Plant Physiology*, 147(1), 367-380.

Wurtzel, E. T. (2019). Changing form and function through carotenoids and synthetic biology. *Plant physiology*, 179(3), 830-843.