

## การพัฒนาเครื่องหมาย ILP จากยีนที่มีความเกี่ยวข้องกับการสังเคราะห์น้ำตาลซูโครสในอ้อย และการศึกษาโครงสร้างพันธุกรรมอ้อย

### Development of ILP markers from genes related to sucrose metabolism in sugarcane and genetic structure of sugarcane

ณัฐภัทร พงศ์ศิริพัฒน์, กิตติพัฒน์ อุโฆษกิจ\*

**Nattapat Pongsiripat, Kittipat Ukoskit\***

ภาควิชาเทคโนโลยีชีวภาพ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ปทุมธานี 12120

Department of Biotechnology, Faculty of Science and Technology, Thammasat University, Pathumthani 12120, Thailand

\*Corresponding author: ku@tu.ac.th

#### บทคัดย่อ

เครื่องหมาย intron length polymorphism (ILP) ได้รับการพัฒนาจากยีนที่เกี่ยวข้องกับการสังเคราะห์น้ำตาลซูโครส โดยใช้ข้อมูลจาก KEGG pathway และผลการศึกษาการทำงานของยีนด้วยไมโครเอเรย์ จากข้อมูลลำดับเบส ESTs ของอ้อย 121,301 ลำดับ (42,377 sugarcane assembled sequences และ 78,924 singleton ESTs) สามารถคัดเลือก ESTs ที่เกี่ยวข้องกับการสังเคราะห์น้ำตาลซูโครสได้ทั้งหมด 91 ลำดับ ทำนายตำแหน่งอินทรอนด้วยเปรียบเทียบลำดับเบส EST ของอ้อยกับลำดับเบสจีโนมของข้าวฟ่าง และออกแบบไพรเมอร์ทั้งหมด 195 คู่ไพรเมอร์จากบริเวณอนุรักษ์ที่คร่อมตำแหน่งอินทรอนอยู่ในการวิเคราะห์โครงสร้างประชากรอ้อย (*Saccharum* spp. L.) 200 ตัวอย่าง พบว่าเครื่องหมาย ILP ตรวจพบแอลลีลได้ 89 แอลลีล มีค่า polymorphism information content (PIC) เฉลี่ย 0.23 ผลการวิเคราะห์หาจำนวนประชากรย่อย ( $K$ ) พบว่ากลุ่มตัวอย่างอ้อยแบ่งออกเป็น 2 ประชากรย่อย ( $K=2$ ) โดยกลุ่มประชากรย่อยที่หนึ่งมีความสัมพันธ์ทางเครือญาติกับอ้อยสายพันธุ์ POJ2878 กลุ่มประชากรย่อยที่สองมีความสัมพันธ์ทางเครือญาติกับอ้อยสายพันธุ์ Trojan และ NCo310 เครื่องหมาย ILP ที่พัฒนาขึ้นในการทดลองนี้สามารถนำไปใช้วิเคราะห์ความสัมพันธ์กับระดับน้ำตาลซูโครสในอ้อยได้ และข้อมูลโครงสร้างพันธุกรรมอ้อยที่ได้จากการทดลองนี้จะเป็นประโยชน์สำหรับแนวทางในการขยายฐานพันธุกรรมอ้อย

#### ABSTRACT

Intron length polymorphism (ILP) markers were developed from genes involving in sucrose metabolism using data from KEGG pathway and a previous study of microarray. Of 121,301 sugarcane ESTs (42,377 sugarcane assembled sequences and 78,924 singleton ESTs), 91 ESTs with predicted function in sucrose metabolism were obtained. Intron positions were determined by aligning the ESTs with genomic sequences of sorghum. The conserved sequences flanking the intron regions were used to design 195 ILP primer pairs. The ILP markers were deployed to assess population structure of 200 sugarcane cultivars (*Saccharum* spp. L.). The polymorphic 89 alleles were detected. The average polymorphism information content (PIC) was 0.23. The number of subpopulations ( $K$ ) of 200 sugarcane cultivars was two ( $K=2$ ). Subpopulation I consisted of most sugarcane varieties sharing the common parent of POJ2878 while most sugarcane varieties in subpopulation II related to the common parent of Trojan and NCo310. The ILP markers developed in this study had a potential for utilizing in association mapping of sucrose content in sugarcane.

The knowledge of population structure obtained in this study should be useful for future breeding programs intended for increasing genetic diversity of sugarcane varieties.

**คำสำคัญ:** อ้อย; เครื่องหมายไอแอลพี; โครงสร้างพันธุกรรม

**Keywords:** sugarcane; ILP markers; genetic structure

## บทนำ

อ้อย (*Saccharum* spp. L.) เป็นพืชใบเลี้ยงเดี่ยว จัดอยู่ในวงศ์หญ้า (Poaceae) ประกอบด้วยชนิด (species) 35-40 ชนิด ในจำนวนนี้ชนิดที่มีความโดดเด่นและเกี่ยวข้องกับสายพันธุ์อ้อยที่ปรากฏอยู่ในปัจจุบัน คือ *S. officinarum* และ *S. spontaneum* เนื่องจากจีโนมของสายพันธุ์อ้อยในปัจจุบันเกิดจากการผสมข้ามชนิดระหว่างอ้อยทั้ง 2 ชนิดนี้ เป็นระยะเวลาเวลานานกว่า 100 ปีแล้ว (Cheavegatti-Gianotto *et al.*, 2011) ในปี พ.ศ. 2557 ประเทศไทยส่งออกน้ำตาลทรายเป็นปริมาณสูงถึง 7,321,575 ตัน คิดเป็นมูลค่า 2.9 พันล้านบาทหรือร้อยละ 1.5 ของมูลค่าการส่งออกสินค้าเกษตรทั้งหมด (Office of the cane and sugar board, online, 2015) การปรับปรุงสายพันธุ์อ้อยไทยโดยมุ่งเน้นที่การเพิ่มระดับน้ำตาลซูโครส จะสามารถลดต้นทุนการผลิต และเพิ่มศักยภาพการแข่งขันให้กับอุตสาหกรรมอ้อยไทยได้เป็นอย่างมาก

การปรับปรุงพันธุ์อ้อยแบบดั้งเดิม (conventional breeding) ใช้เวลาอย่างน้อย 10 ปี เพื่อให้ได้สายพันธุ์อ้อยที่มีลักษณะที่ต้องการ การนำเครื่องหมายดีเอ็นเอมาใช้ในการปรับปรุงพันธุ์สามารถเพิ่มความแม่นยำและความรวดเร็วในการคัดเลือกได้ มีรายงานการใช้เครื่องหมายดีเอ็นเอในการปรับปรุงสายพันธุ์อ้อยให้สามารถต้านทานโรคราสนิม (Rossi *et al.*, 2003) และโรคใบจุดสีเหลือง (Al-Janabi *et al.*, 2007) แต่ยังไม่พบรายงานการใช้เครื่องหมายดีเอ็นเอเพื่อปรับปรุงสายพันธุ์อ้อยให้มีระดับน้ำตาลซูโครสสูงอยู่เลย ดังนั้นการพัฒนาเครื่องหมายดีเอ็นเอโดยมุ่งเน้นไปที่กลุ่มยีนที่เกี่ยวข้องกับการสังเคราะห์น้ำตาลซูโครส จะเป็นจุดเริ่มต้นสำหรับการค้นหาเครื่องหมายดีเอ็นเอสำหรับนำไปใช้ในการปรับปรุงสายพันธุ์อ้อยให้มีระดับน้ำตาลสูงต่อไปในอนาคต

ในปัจจุบันยังไม่มีข้อมูลจีโนมที่สมบูรณ์ของอ้อย ดังนั้นแหล่งข้อมูลสำหรับศึกษายีนของอ้อยที่ดีที่สุดขณะนี้จึงเป็นฐานข้อมูลลำดับเบส expressed sequence tags (ESTs) (Vettore *et al.*, 2003) มีการนำลำดับเบส ESTs ของอ้อยมาใช้เพื่อศึกษาการแสดงออกของยีนในเนื้อเยื่อใบ (McCormick *et al.*, 2008) และลำดับ (Papini-Terzi *et al.*, 2009) ของอ้อย ตลอดจนพัฒนาเครื่องหมายดีเอ็นเอชนิดไมโครแซทเทลไลต์ (EST-SSR) (Oliveira *et al.*, 2009; Ukoskit *et al.*, 2012)

เครื่องหมาย Intron length polymorphism (ILP) คือ เครื่องหมายดีเอ็นเอที่ใช้ตรวจสอบความแตกต่างของความยาวอินทรอนด้วยปฏิกิริยาลูกโซ่พอลิเมอร์เรส (Polymerase Chain Reaction, PCR) โดยใช้ไพรเมอร์ที่จำเพาะกับบริเวณเอ็กซอนที่คร่อมส่วนที่คาดว่าเป็นอินทรอน ซึ่งตำแหน่งของอินทรอนสามารถทำนายได้โดยการเปรียบเทียบระหว่างลำดับเบส ESTs กับลำดับเบสจีโนม (Yang *et al.*, 2007a) ความยาวของอินทรอนมีผลกระทบกับระดับการแสดงออกของยีน และฟีโนไทป์ในข้าวและอะราบิโดพซิส พบว่ากลุ่มยีนที่มีขนาดอินทรอนยาวกว่ามีระดับการแสดงออกของยีนสูงกว่ากลุ่มยีนที่มีขนาดอินทรอนสั้นกว่า (Ren *et al.*, 2006) ในข้าวสาลี พบว่าความยาวของอินทรอนที่สามของยีน Vp-1 (viviparous) มีความสัมพันธ์กับระดับการทนต่อการงอกก่อนการเก็บเกี่ยว (Pre-harvest sprouting, PHS) (Yang *et al.*, 2007b) ด้วยเหตุนี้จึงมีการพัฒนาเครื่องหมาย ILP ในพืชหลายชนิด เช่น ถั่วเหลือง (Shu *et al.*, 2010) ยางพารา (Li *et al.*, 2011) ข้าวโพด (Liu *et al.*, 2012) และสบู่ดำ (Saisug and Ukoskit, 2013)

ระดับความหวานของอ้อยเป็นหนึ่งในลักษณะเป้าหมายที่สำคัญของโครงการปรับปรุงพันธุ์อ้อย ในการศึกษาครั้งนี้จึงมุ่งเน้นไปที่กลุ่มยีนที่เกี่ยวข้องกับการสังเคราะห์แป้งและน้ำตาล โดยการสืบค้นข้อมูลจากฐานข้อมูล KEGG pathway (<http://www.genome.jp/kegg/pathway.html>) และจากผลการศึกษารูปแบบการแสดงออกของยีนในเนื้อเยื่อลำต้นของอ้อยที่มีความหวาน 18 บริกซ์ และ 13 บริกซ์ ด้วยวิธีไมโครแอรเรย์ ของ Papini-Terzi *et al.* (2009) ข้อมูลทั้งสองนี้สามารถนำมาใช้ในการทำนายหน้าที่ของ ESTs ของอ้อยที่มีความเกี่ยวข้องกับการสังเคราะห์แป้งและน้ำตาล และ

นำไปใช้ในการพัฒนาเครื่องหมาย ILP ได้ เครื่องหมาย ILP ที่พัฒนาได้ในการทดลองนี้ได้นำมาใช้ในการวิเคราะห์โครงสร้างประชากรของอ้อยที่เก็บรวบรวมจากแหล่งการปรับปรุงพันธุ์อ้อยทั่วโลก

### อุปกรณ์และวิธีการ

#### ตัวอย่างอ้อยและการสกัดดีเอ็นเอ

สายพันธุ์อ้อยที่ใช้ศึกษามี 200 สายพันธุ์ (Table 1) สายพันธุ์อ้อยเหล่านี้มีความหลากหลายทางฟีโนไทป์ของ

ลักษณะความหวาน และลักษณะองค์ประกอบผลผลิต พันธุ์อ้อยเหล่านี้รวบรวมพันธุ์จากแหล่งในทวีปต่างๆ ของโลก และปลูกที่บริษัท มิตรผลวิจัย พัฒนาอ้อยและน้ำตาล จำกัด อ. ภูเขียว จ. ชัยภูมิ สกัดดีเอ็นเอจากใบอ้อยแต่ละสายพันธุ์ด้วยวิธี cetyltrimethylammonium bromide (CTAB method) ตามวิธีที่ดัดแปลงจากวิธีของ Gawel and Jarret (1991) ตรวจสอบคุณภาพและความเข้มข้นของดีเอ็นเอด้วยวิธีอิเล็กโทรโฟรีซิสในเจลอะกาโรส และการวัดค่าการดูดกลืนแสง ตามลำดับ

**Table 1** List of the 200 sugarcane clones and their country of origin.

Acc. No.	Clones	Country	Acc. No.	Clones	Country	Acc. No.	Clones	Country
1	3-2-023L	Thailand	76	F160	Taiwan	151	Q47	Australia
2	Aiwa	Fiji	77	F164	Taiwan	152	Q66	
3	Apollo	Australia	78	F170	Taiwan	153	Q77	Australia
4	B34104	Barbados	79	F174	Taiwan	154	Q81	Australia
5	B34164	Barbados	80	F175	Taiwan	155	Q82	Australia
6	B41721		81	F176	Taiwan	156	Q83	Australia
7	B4362	Barbados	82	F177	Taiwan	157	Q84	Australia
8	B4715	Barbados	83	F178	Taiwan	158	Q87	Australia
9	B4744	Barbados	84	Fiji105	Fiji	159	Q99	Australia
10	B76718	Barbados	85	Galoa		160	R397	France
11	BL4		86	H47-4991	USA	161	ROC1	Taiwan
12	BO11	India	87	H59-3775	USA	162	ROC2	Taiwan
13	BO14	India	88	Homer	Australia	163	ROC21	Taiwan
14	BO24	India	89	IAC49-131	Brazil	164	ROC23	Taiwan
15	C87-51	Cuba	90	IAC51-205	Brazil	165	ROC24	Taiwan
16	CAC57-11	Philippines	91	IAC52-150	Brazil	166	ROC7	Taiwan
17	CAC57-13	Philippines	92	IAC52-156	Brazil	167	Saipan17	Japan
18	CAC57-23	Philippines	93	IRK67-1		168	Yuangjeing	
19	CAC57-66	Philippines	94	Ja64-19	Cuba	169	SP70-1284	Brazil
20	CB38-22	Brazil	95	K88-92	Thailand	170	TBy20-0214	Thailand
21	CN1	Thailand	96	K93-219	Thailand	171	TBy20-2248	Thailand
22	Co1010	India	97	KK3		172	TBy26-1255	
23	Co1013	India	98	KpK98-40		173	Tuc1149	Argentina
24	Co1148	India	99	Kwt3	China	174	Tuc2645	Argentina
25	Co1287	India	100	L75-2	USA	175	Tuc72-5	Argentina
26	Co245	India	101	LF48-8077	Fiji	176	Tuc74-6	Argentina
27	Co254	India	102	LF63-2705	Fiji	177	UT1	Thailand
28	Co270	India	103	LF70-7143		178	UT3	Thailand
29	Co290	India	104	LF76-2300	Fiji	179	UT5	Thailand
30	Co317	India	105	LF78-960	Fiji	180	Vidar	Australia
31	Co331	India	106	LF79-2964	Fiji	181	Vomo	Fiji
32	Co419	India	107	LF79-640	Fiji	182	Yasawa	Fiji
33	Co449	India	108	LF82-2139	Fiji	183	MP1	
34	Co644	India	109	LF89-1490	Fiji	184	MP2	

Table 1 (continued).

Acc. No.	Clones	Country	Acc. No.	Clones	Country	Acc. No.	Clones	Country
35	Co858	India	110	LF89-2059	Fiji	185	MP3	
36	CoS443	India	111	M124/59	Mauritius	186	MPT02-458	
37	CoS510	India	112	M13/58	Mauritius	187	MPT02-612	
38	CP110	USA	113	M147/158		188	MPT03-166	
39	CP29-211	USA	114	M147/44	Mauritius	189	MPT08-95	
40	CP29-291	USA	115	M336	USA	190	MPT08-191	
41	CP32-224	USA	116	M34/45	Mauritius	191	K84-200	Thailand
42	CP45-150	USA	117	M377/56	Mauritius	192	K92-80	Thailand
43	CP47-193	USA	118	M93/48	Mauritius	193	K95-84	Thailand
44	CP48-103	USA	119	Mackay	Australia	194	K99-72	
45	CP50-11	USA	120	Mana	Fiji	195	Kps01-12	Thailand
46	CP51-13	USA	121	Mossman	Australia	196	Kps01-25	Thailand
47	CP52-148	USA	122	MY5514		197	UT8	Thailand
48	CP52-48	USA	123	N50-211		198	UT91-2-527	Thailand
49	CP52-68	USA	124	N52-211		199	LK92-11	Thailand
50	CP57-603	USA	125	N6	Africa	200	S6	
51	CP61-39	USA	126	N8	Africa			
52	CP63-306	USA	127	NCo293	Africa			
53	CP65-350	USA	128	NCo310	Africa			
54	CP70-32	USA	129	NCo382	Africa			
55	CP72-120	USA	130	NCo387	Africa			
56	CP75-324	USA	131	Phil54-52	Philippines			
57	CP75-327	USA	132	Phil6607	Philippines			
58	CP75-330	USA	133	Pindar	Australia			
59	CP76-340	USA	134	POJ2878	Indonesia			
60	CP77-403	USA	135	PR3067	USA			
61	CP77-418	USA	136	PR69-2110	USA			
62	D158-41		137	PR76-3035	USA			
63	DB42060	Barbados	138	Ps41	Indonesia			
64	DB671760	Barbados	139	PSA14	Philippines			
65	DB7160	Barbados	140	PSA63				
66	Ebene1/37	Mauritius	141	PT43-52	Taiwan			
67	Ehaew	Thailand	142	PT52-227	Taiwan			
68	Eros	Australia	143	Q100	Australia			
69	F140	Taiwan	144	Q107	Australia			
70	F148	Taiwan	145	Q113	Australia			
71	F150	Taiwan	146	Q114	Australia			
72	F151	Taiwan	147	Q115	Australia			
73	F152	Taiwan	148	Q117	Australia			
74	F153	Taiwan	149	Q130	Australia			
75	F156	Taiwan	150	Q146	Australia			

### การพัฒนาเครื่องหมาย ILP

พัฒนาเครื่องหมาย ILP โดยนำข้อมูลลำดับเบส ESTs ของอ้อยจำนวน 121,301 ลำดับ ประกอบด้วย sugarcane assembled sequences 42,377 ลำดับ และ singleton ESTs 78,924 ลำดับ จากฐานข้อมูล Gene Index Project

(ftp://occams.dfc.harvard.edu/pub/bio/tgi/data/Saccharum\_officinarum/) มาดำเนินการคัดเลือกหา ESTs ที่เกี่ยวข้องกับสารสังเคราะห์น้ำตาลซูโครสโดยใช้แหล่งข้อมูล 2 แหล่ง ได้แก่ KEGG pathway database และผลการศึกษาไมโครแอเรย์ การคัดเลือกจาก KEGG

pathway database ใช้การเปรียบเทียบระหว่างลำดับเบส ESTs ของอ้อยกับลำดับกรดอะมิโนของข้าวฟ่างจากฐานข้อมูล Phytozome (<http://phytozome.jgi.doe.gov/pz/portal.html>) ด้วยโปรแกรม blastx ซึ่งลำดับกรดอะมิโนของข้าวฟ่างได้รับการระบุหน้าที่ตามฐานข้อมูล KEGG pathway เรียบร้อยแล้ว ดังนั้นจึงสามารถคัดเลือกลำดับเบส ESTs ที่มีหน้าที่เกี่ยวข้องกับวิถีชีวสังเคราะห์แป้งและน้ำตาล (entry: sbi00500) ได้ สำหรับการคัดเลือกจากผลการศึกษาไมโครแอรีย์ ใช้ลำดับเบส ESTs ที่แสดงออกในเนื้อเยื่อลำต้นของอ้อยที่มีความหวาน 18 บริกซ์ และ 13 บริกซ์ ในรายงานของ Papini-Terzi *et al.* (2009) มาออกแบบไพรเมอร์ นำลำดับเบส ESTs ของอ้อยที่ผ่านการคัดเลือกมาเปรียบเทียบกับลำดับเบสจีโนมของข้าวฟ่างจากฐานข้อมูล Phytozome เพื่อทำนายตำแหน่งของอินทรอน ออกแบบไพรเมอร์ที่จำเพาะกับบริเวณเอ็กซอนที่ครอบคลุมตำแหน่งอินทรอนด้วยโปรแกรม Primer3 (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) โดยกำหนดให้ผลผลิตพีซีอาร์มีขนาดอยู่ในช่วง 150–400 คู่เบส

### การวิเคราะห์จีโนไทป์

วิเคราะห์จีโนไทป์ของอ้อย 200 สายพันธุ์ด้วยเครื่องหมาย ILP ที่พัฒนาขึ้นจำนวน 195 เครื่องหมาย ดำเนินปฏิกิริยาพีซีอาร์ปริมาตร 25  $\mu$ l ประกอบด้วยดีเอ็นเอแม่แบบ 25 ng, 1X PCR buffer, 2.5 mM  $MgCl_2$ , 0.2 mM dNTPs, คู่ไพรเมอร์อย่างละ 0.32  $\mu$ M และ Taq DNA Polymerase 1 unit (AmpOne™  $\alpha$ -Taq premix, General Biotechnology Co., Korea) ใช้สภาวะอุณหภูมิ 94 °C 3 นาที ตามด้วย 35 รอบของ 94 °C 30 วินาที (52–66) °C 1 นาที 72 °C 35 วินาที และสุดท้ายที่อุณหภูมิ 72 °C 5 นาที วิเคราะห์ผลผลิตพีซีอาร์เบื้องต้นด้วยวิธีอิเล็กโทรโฟริซิสในเจลอะกาโรสความเข้มข้น 1.5% และในเจลโพลีอะคริลามิเดความเข้มข้น 6% (denaturing polyacrylamide gel) ย้อมด้วยสารย้อมเงิน (Benbouza *et al.*, 2006) บันทึกแถบดีเอ็นเอผสมสมบูรณ์ โดยให้คะแนนเป็น 1 สำหรับตัวอย่างที่พบแถบดีเอ็นเอที่ตำแหน่งหนึ่งๆ และให้คะแนนเป็น 0 สำหรับตัวอย่างที่ไม่พบแถบดีเอ็นเอที่ตำแหน่งเดียวกันนั้น

### การวิเคราะห์ข้อมูล

คำนวณค่า polymorphism information content (PIC) และ gene diversity (GD) ของเครื่องหมายดีเอ็นเอด้วยโปรแกรม PowerMarker (Liu and Muse, 2004) คำนวณสัมประสิทธิ์ความคล้ายคลึง (similarity coefficient) ด้วยวิธีของ Nei and Li (1979) นำค่าสัมประสิทธิ์ความคล้ายคลึงมาวิเคราะห์การจัดกลุ่มสร้างเป็นเดนโดแกรม (dendrogram) ด้วยวิธี neighbour-joining (NJ) ด้วยโปรแกรม PowerMarker และแสดงผลด้วยโปรแกรม MEGA5 (Tamura *et al.*, 2011) วิเคราะห์การจัดกลุ่มด้วยวิธี Principal Coordinate Analysis (PCoA) และแสดงผลด้วยโปรแกรม NTSYS-PC V.2.0 (Rohlf, 1998) ประเมินจำนวนกลุ่มประชากร ( $K$ ) ที่เหมาะสมในการวิเคราะห์โครงสร้างประชากรจะใช้วิธีวิเคราะห์ความน่าจะเป็นแบบเบย์ (Bayesian) ด้วยโปรแกรม STRUCTURE V.2.2 (Pritchard *et al.*, 2000; Falush *et al.*, 2003) model ที่ใช้เป็นแบบ F-model โดยคำนวณค่าของ  $K$  ตั้งแต่ 1-10 วิเคราะห์ซ้ำจำนวน 10 ครั้ง กำหนดให้ค่า iteration เท่ากับ 100,000 รอบ และ burn-in เท่ากับ 10,000 แล้วใช้ค่า log-likelihood ของ  $K$  ในการหาค่า  $K$  ที่เหมาะสม ซึ่งจำนวนกลุ่มประชากรย่อยที่เหมาะสมจะให้ค่า Delta  $K$  สูงที่สุด (Evanno *et al.*, 2005) โดยกำหนดให้พันธุ์อ้อยที่อยู่ในประชากรเดียวกันค่า  $Q$  มากกว่า 0.75

### ผลการทดลองและวิจารณ์

#### การพัฒนาเครื่องหมาย ILP

จากการสืบค้นลำดับเบส ESTs และระบุหน้าที่การทำงานตามฐานข้อมูล KEGG pathway สามารถค้นหา ESTs ที่เกี่ยวข้องกับวิถีชีวสังเคราะห์แป้งและน้ำตาลได้ 34 ยีน (Table 2) และสามารถคัดเลือก ESTs ของอ้อยจากรายงานของ Papini-Terzi *et al.* (2009) ที่สัมพันธ์กับความหวานได้ 47 ยีน (Table 3) รวมยีนที่เกี่ยวข้องกับความหวานทั้งหมด 81 ยีน จากยีนทั้งหมดนี้สามารถพัฒนาเครื่องหมาย ILP ได้ 195 คู่ไพรเมอร์ Table 4 แสดงเครื่องหมาย ILP ที่มีค่า PIC สูงจำนวน 5 เครื่องหมาย

**Table 2** List of the 34 genes from starch and sucrose metabolism pathway (entry: sbi00500) and the Sugarcane Intron Length (SIL) primer number 101-195 developed within each gene.

SIL no.	Gene	SIL no.	Gene
101-104	1,4-alpha-glucan branching enzyme (chr10)	147-150	glucose-6-phosphate isomerase
105	1,4-beta-D-xylan synthase	151-152	glycogen operon protein
106-110	4-alpha-glucanotransferase (chr02)	153-155	hexokinase (chr09)
111-114	alpha,alpha-trehalase	156	pectinesterase (chr03)
115-119	alpha-1,4-galacturonosyltransferase	157	pectinesterase (chr01)
120	alpha-amylase	158-161	phosphoglucomutase
121	alpha-glucosidase	162-165	starch phosphorylase
122	beta-amylase	166-170	sucrose synthase (chr01)
123-124	beta-D-xylosidase 4	171-173	sucrose synthase (chr10)
125-126	beta-fructofuranosidase (chr03)	174-177	sucrose-phosphate synthase (chr09)
127-128	beta-fructofuranosidase (chr04)	178-180	sucrose-phosphate synthase (chr03)
129-131	beta-glucosidase (chr01)	181-182	trehalose 6-phosphate phosphatase (chr02)
132-135	beta-glucosidase (chr08)	183-184	trehalose 6-phosphate phosphatase (chr07)
136-138	endoglucanase	185	trehalose 6-phosphate synthase_phosphatase (chr2)
139-140	fructokinase (chr03)	186-187	trehalose 6-phosphate synthase_phosphatase (chr7)
141	fructokinase (chr10)	188-190	trehalose 6-phosphate synthase_phosphatase (chr9)
142-146	glucose-1-phosphate adenylyltransferase (chr09)	191-195	UTP--glucose-1-phosphate uridylyltransferase

**Table 3** List of the 47 genes from microarray research (Papini-Terzi *et al.*, 2009) and the Sugarcane Intron Length (SIL) primer number 1-100 developed within each gene.

SIL no.	Gene	SIL no.	Gene
1-3	14-3-3-like protein (chr06)	48-50	Tyrosine specific protein phosphatases
4-6	14-3-3-like protein (chr07)	51-52	Dual specificity protein phosphatases
7	14-3-3-like protein (chr05)	53	PP2C-like (serine/threonine PPM family)
8-13	AAA family ATPase (chr01)	54-56	PP2A/Subunit A (PPM family)
14	Dehydrin	57-60	ScSnRK1-2 (SNF-like kinase)
15-16	Aquaporin 1	61-62	ScPKABA1-3 (SNF-like kinase)
17	Aquaporin 2	63-66	ScOSA PK-7 (SNF-like kinase)
18	Aquaporin 3	67-68	ScCIPK-9 (SNF-like kinase)
19-20	ScC3H95 (zinc finger protein)	69-72	ScCIPK-3 (SNF-like kinase)

**Table 3** (continued).

SIL no.	Gene	SIL no.	Gene
21-23	ScYAB16 (zinc finger protein)	73-74	ScCyclin G-associated kinase-like protein-1
24	ScNAC27 (no apical meristem)	75	ScATN1-2 (protein kinase)
25	ScMYB120 (MYB)	76	ScCDK-6 (cell cycle-related)
26	ScMADS17 (MADS)	77-78	ScCDK-3 (cell cycle-related)
27	LIM (protein-protein interaction)	79	ScCDK-18 (cell cycle-related)
28-29	ScGRAS71 (gibberellin-related)	80-87	ScCKI-3 (casein kinase)
30-31	(Aux-IAA-related)	88	ScCDPK-27 (calcium-dependent)
32-33	ScAUX1134 (Aux-IAA-related)	89-91	HSP70 (Heat shock protein)
34-38	ScARF46 (Auxin-related)	92	LSD1 (zinc finger protein)
39	ScHB2 (homeobox)	93-94	CRT2 Calreticulin 2
40	ScbHLH7 (helix-loop-helix)	95	ABC transporter
41	ScbHLH1 (helix-loop-helix)	96-98	Ca(2+)-ATPase
42-43	HGM (high mobility group protein)	99	Nitrilase
44-45	ScCA2P5 (CCAAT)	100	auxin associated family
46-47	EIN2 (ethylene)		

**Table 4** Primer sequences, EST accession numbers (Gene Index Project), expected size of PCR product, annealing temperature and PIC of the selected polymorphic ILP markers.

SIL no.	Primer sequence (5'→3')	EST acc. no.	Expected size (bp)	Annealing temperature (°C)	PIC
30	TACGAGGACAAGGACGGC GGCTTCCGACCCTTTCATG	TC138898	216	68	0.344
72	GAAGGCAAAGGGGGATACTC TCGCACTTCCACACAACATC	TC115264	180	62	0.370
94	GTTGCTGGGTGGTGTAG AGTGTTCTTGCCATCCTTG	TC113533	216	64	0.375
122	GTTTACACCGACTACATGCG GGATACCTCAGCTCTCCAC	TC112916	237	64	0.325
184	CAAGGTGGAGAACAACAAGT TGAGCACCGACTTCACCA	CA065391	274	62	0.374

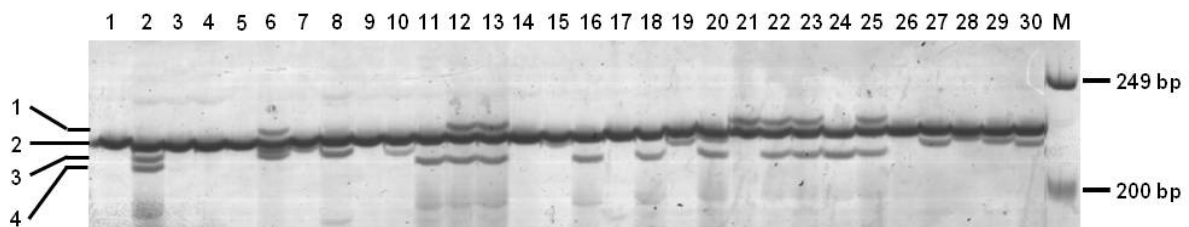
**การวิเคราะห์ความแปรปรวนแอลลีลของเครื่องหมาย ILP**

ผลการทำปฏิกิริยาพีซีอาร์ด้วยไพรเมอร์ ILP จำนวน 195 คู่ พบไพรเมอร์ที่สามารถเพิ่มปริมาณชิ้นดีเอ็นเอได้สำเร็จตรงตามขนาดที่คาดหวังจำนวน 176 คู่

คิดเป็นอัตราความสำเร็จ 90% บ่งชี้ว่าอินทรอนบนจีโนมของอ้อยและข้าวฟ่างมีตำแหน่งและขนาดใกล้เคียงกัน สอดคล้องกับรายงานที่ระบุว่าพืชทั้งสองชนิดมีบรรพบุรุษร่วมกัน (common ancestor) ร่วมกันเมื่อ 8–9 ล้านปีก่อน และจีโนมของพืชทั้งสองชนิดมีลำดับการเรียงตัวของยีนที่

คล้ายกัน (synteny) (Jannoo *et al.*, 2007) ในจำนวนไพรเมอร์ 176 คู่ มีไพรเมอร์ที่ตรวจพบความแตกต่างของขนาดอินทรอนจำนวน 53 คู่ คิดเป็นอัตราโพลิมอร์ฟิซึม 27% ผลการวิเคราะห์ความหลากหลายของแอลลีลพบแถบดีเอ็นเอที่แสดงความแตกต่างระหว่างพันธุ์ 89 แถบ (allele) มีจำนวนแอลลีลผันแปรอยู่ในช่วง 1–5 แอลลีลและค่าเฉลี่ยเท่ากับ 1.68 แอลลีลต่อไพรเมอร์ มีค่า PIC ผันแปรอยู่ในช่วง 0.01–0.375 และค่าเฉลี่ยเท่ากับ 0.23 เมื่อเปรียบเทียบกับค่า PIC เฉลี่ยของเครื่องหมาย EST-SSR ในอ้อยของ Oliveira *et al.* (2009) และ Ukoskit *et al.* (2012) ที่มีค่า PIC เฉลี่ยเท่ากับ 0.73 และ 0.74 ตามลำดับ พบว่าความผันแปรของขนาดอินทรอนในอ้อยมีน้อยกว่าความผันแปรของขนาดลำดับเบสซ้ำสอดคล้องกับการเปรียบเทียบระหว่างเครื่องหมาย ILP

กับ EST-SSR ในพืชชนิดอื่นๆ เช่น ข้าว (Wang *et al.*, 2005) และพืชตระกูลหญ้า (Zeida *et al.*, 2010) ทั้งนี้ อาจเนื่องจากความแปรปรวนของลำดับเบสอันเนื่องมาจากความผันแปรของจำนวนซ้ำที่เกิดจากความผิดพลาดจากการทำงานของเอนไซม์พอลิเมอร์เรสซึ่งมีโอกาสเกิดขึ้นได้มากกว่าความผันแปรของขนาดอินทรอนเกิดจากการแทรกและการขาดหายไปของซันดีเอ็นเอบนโครโมโซมคู่เหมือน (Figure 1) โดยโครโมโซมคู่เหมือนส่วนใหญ่มีขนาดอินทรอนเท่ากัน (แถบดีเอ็นเอที่ 2) เนื่องจากมีความเข้มข้นของแถบดีเอ็นเอมากที่สุด และปรากฏเหมือนกันในทุกสายพันธุ์อ้อยที่ทดสอบ โครโมโซมคู่เหมือนบางแท่งมีการแทรกของซันดีเอ็นเอ (แถบดีเอ็นเอที่ 1) ขณะที่โครโมโซมคู่เหมือนบางแท่งมีการขาดหายของซันดีเอ็นเอ (แถบดีเอ็นเอที่ 3 และ 4)



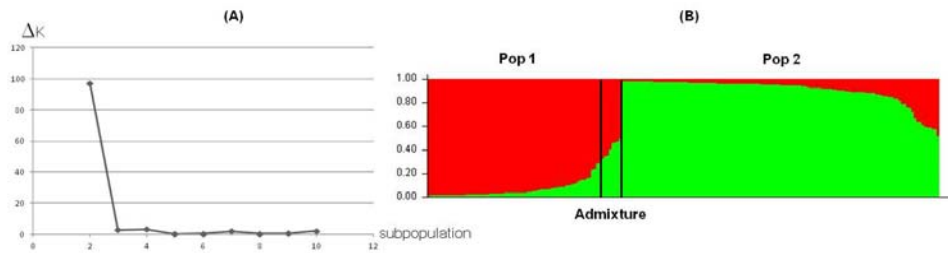
**Figure 1** Polyacrylamide gel electrophoresis of SIL30 markers. Lanes 1–30 indicate sugarcane clone accession number 1 to 30. M is the DNA size marker. The ILP alleles are indicated by numbers 1–4.

### การจัดกลุ่มโครงสร้างประชากรอ้อย

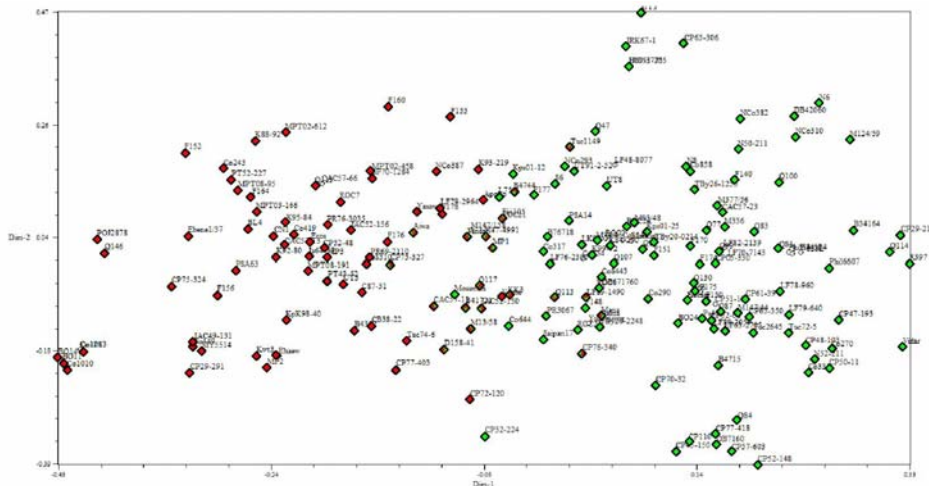
ผลการวิเคราะห์โครงสร้างประชากรอ้อย 200 สายพันธุ์โดยวิธีวิเคราะห์ความน่าจะเป็นแบบเบย์ พบว่าประชากรอ้อยแบ่งออกเป็น 2 กลุ่มประชากรย่อย (Figure 2A) สอดคล้องกับผลการจัดกลุ่มด้วยวิธี PCoA (Figure 3) และผลการสร้างเดนโดรแกรม (Figure 4) โดยมีอ้อยจำนวน 21 สายพันธุ์เป็นพันธุกรรมผสมระหว่างสองประชากร (admixture) เมื่อกำหนดให้พันธุ์มีลักษณะเป็นพันธุกรรมผสมระหว่างสองประชากรมีค่า Q น้อยกว่า 0.75 เมื่อพิจารณาผลการจัดกลุ่มร่วมกับข้อมูลพันธุประวัติของอ้อย (Wongpraneekul, 2008) (Figure 5) พบว่าการเกิดโครงสร้างพันธุกรรมอ้อยมีแนวโน้มที่เกิดจากการมีความสัมพันธ์ทางเครือญาติกันระหว่างสมาชิกภายในประชากรย่อย กลุ่มประชากรย่อยที่หนึ่ง (Figure 2B) มีความสัมพันธ์ทางเครือญาติกับอ้อยสายพันธุ์ POJ2878 (Figure 5A) ส่วนกลุ่มประชากรย่อยที่สอง (Figure 2B) มี

ความสัมพันธ์ผสมระหว่างอ้อยที่มีความสัมพันธ์ทางเครือญาติกับอ้อยสายพันธุ์ Trojan และ NCo310 (Figure 5B, 5C) สอดคล้องกับรายงานของ Heller-Uszynska *et al.* (2011) ที่แสดงการแยกออกเป็นสองกลุ่มระหว่างอ้อยสายพันธุ์ POJ2878 และ Trojan อย่างไรก็ตามเมื่อสืบพันธุประวัติของอ้อย พบว่าอ้อยจากแหล่งในทวีปต่างๆ ของโลกมีฐานพันธุกรรมที่คล้ายกันมาก โดยสายพันธุ์ส่วนใหญ่มีอ้อยพันธุ์ Black Cheribon, Bandjermasin Hitam, Loethers และ Chunnee เป็นบรรพบุรุษร่วมกัน บ่งชี้ถึงการแลกเปลี่ยนสายพันธุ์อ้อยระหว่างนักปรับปรุงพันธุ์อ้อยทั่วโลก ส่งผลให้อ้อยจากโครงการปรับปรุงพันธุ์แต่ละโครงการมีลักษณะไม่แตกต่างกันมากนัก (Raboin *et al.*, 2008) ดังนั้นนักปรับปรุงพันธุ์จึงควรหาแนวทางในการขยายฐานพันธุกรรมอ้อย เช่น การนำเชื้อพันธุกรรมอ้อยชนิดใหม่มาใช้ในโครงการปรับปรุงพันธุ์อ้อย การผสมข้ามระหว่างชนิดโดยใช้แหล่งพันธุกรรมใหม่ เป็นต้น

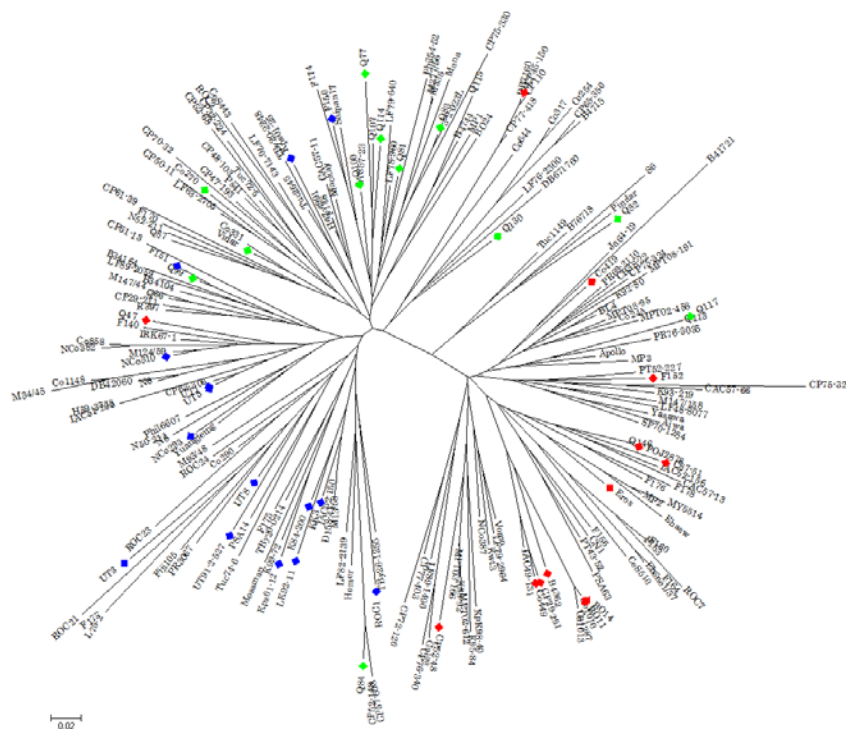




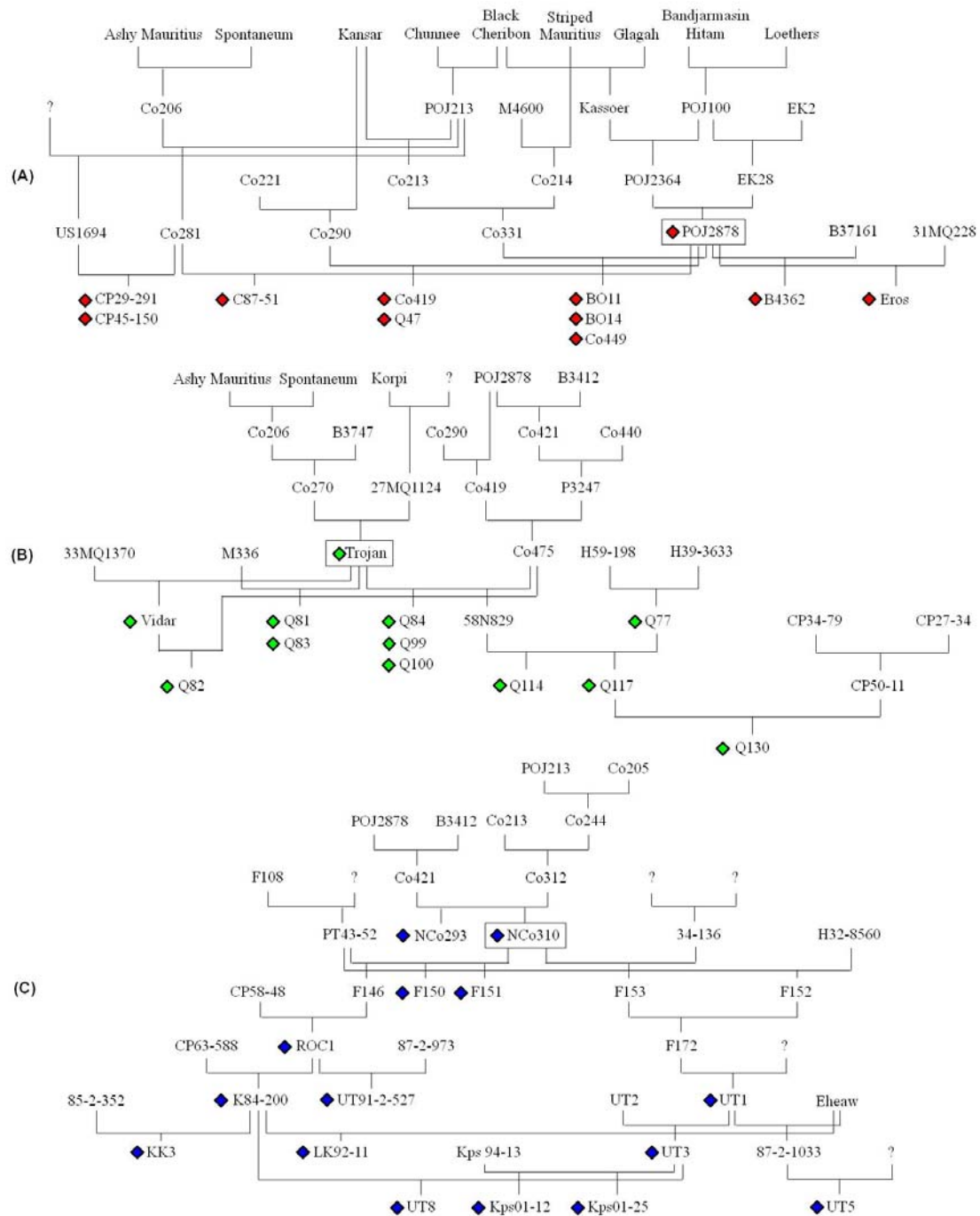
**Figure 2** Population structure of 200 sugarcane varieties. (A)  $\Delta K$  shows a clear peak at subpopulation of 2. (B) The varieties depicted in red (subpopulation 1) and green (subpopulation 2) segments in accordance with Q value.



**Figure 3** Scatter plot of principal coordinate analysis represents the varieties belonging to subpopulation 1 (red cluster), subpopulation 2 (green cluster) and admixture (mixed-color spots).



**Figure 4** Neighbor-joining dendrogram of 200 sugarcane varieties based on 53 polymorphic ILP primer pairs. The red, green and blue spots represent closely related varieties in accordance with genealogy.



**Figure 5** Genealogy of selected varieties with known parentage. (A), (B) and (C) represent varieties sharing common parent of POJ2878, Trojan and NCo310, respectively.

**สรุปผลการวิจัย**

งานวิจัยครั้งนี้สามารถพัฒนาเครื่องหมาย ILP จากกลุ่มยีนที่เกี่ยวข้องกับการสังเคราะห์น้ำตาลซูโครสได้สำเร็จ และสามารถศึกษาโครงสร้างประชากรอ้อยที่รวบรวมพันธุ์จากแหล่งต่างๆ ทั่วโลกได้ โดยกลุ่มประชากรย่อยที่ได้มีความสัมพันธ์กับความสัมพันธ์ทาง

เครือญาติ แต่การแลกเปลี่ยนสายพันธุ์อ้อยระหว่างนักปรับปรุงพันธุ์อ้อยทั่วโลก ส่งผลให้อ้อยจากโครงการปรับปรุงพันธุ์แต่ละโครงการมีลักษณะไม่แตกต่างกันมากนัก ดังนั้นนักปรับปรุงพันธุ์จึงควรขยายฐานพันธุกรรมอ้อย โดยนำเชื้อพันธุกรรมอ้อยชนิดใหม่มาใช้ในโครงการปรับปรุงพันธุ์อ้อยในอนาคต

## กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณทุนสนับสนุนจากศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติ สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ และบริษัทมิตรผลวิจัย พัฒนาอ้อยและน้ำตาล จำกัด

## เอกสารอ้างอิง

- Al-Janabi SM, Parmessur Y, Kross H, Dhayan S, Saumtally S, Ramdoyal K, Autrey LJC, Dookun-Saumtally A (2007) Identification of a major quantitative trait locus (QTL) for yellow spot (*Mycovellosiella koepkei*) disease resistance in sugarcane. *Mol Breed* 19: 1–14.
- Benbouza H, Jacquemin JM, Baudoin T, Mergeai G (2006) Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels. *Biol Agron Soc Env* 10: 77–81.
- Cheavegatti-Gianotto A, Couto de Abreu HM, Arruda P, Filho JCB, Burnquist WE, Creste S, Ciero LD, Ferro JA, Figueira AVDO, Filgueiras TDS, *et al.* (2011) Sugarcane (*Saccharum X officinarum*): A reference study for the regulation of genetically modified cultivars in Brazil. *Trop Plant Biol* 4: 62–89.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611–2620.
- Falush, D, Stephens M, Pritchard JK (2003) Inference of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
- Gawel N, Jarret R (1991) A Modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. *Plant Mol Bio Rep* 9: 262–266.
- Heller-Uszynska K, Uszynski G, Huttner E, Evers M, Carlig J, Caig V, Aitken K, Jackson P, Piperidis G, Cox M, *et al.* (2011) Diversity Arrays Technology effectively reveals DNA polymorphism in a large and complex genome of sugarcane. *Mol Breeding* 28: 37–55.
- Jannoo N, Grivet L, Chantret N, Garsmeur O, Glaszmann JC, Arruda P, D'Hont A (2007) Orthologous comparison in a gene-rich region among grasses reveals stability in the sugarcane polyploid genome. *Plant J* 50: 574–585.
- Li D, Xia Z, Deng Z, Liu X, Dong J, Feng F (2011) Development and characterization of intron-flanking EST-PCR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). *Mol Biotechnol* 51: 148–159.
- Liu H, Lin Y, Chen G, Shen Y, Liu J, Zhang S (2012) Genome-scale identification of resistance gene analogs and the development of their intron length polymorphism markers in maize. *Mol Breeding* 29: 437–447.
- Liu K, Muse SV (2004) PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128–2129.
- McCormick AJ, Cramer MD, Watt DA (2008) Differential expression of genes in the leaves of sugarcane in response to sugar accumulation. *Tropical Plant Biol* 1: 142–158.
- Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* 76: 5269–5273.
- Oliveira KM, Pinto LR, Marconi TG, Mollinari M, Ulian EC, Chabregas M, Falco MC, Burnquist W, Garcia AAF, Souza AP (2009) Characterization of new polymorphic functional markers for sugarcane. *Genome* 52: 191–209.
- Papini-Terzi FS, Rocha FR, Vêncio RZN, Felix JM, Branco DS, Waclawovsky AJ, Bem LED, Lembke CG, Costa MDL, Nishiyama-Jr MY, *et al.* (2009) Sugarcane genes associated with sucrose content. *BMC Genomics* 10: 120.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

- Raboin LM, Pauquet J, Butterfield M, D'Hont A, Glaszmann JC (2008) Analysis of genome-wide linkage disequilibrium in the highly polyploid sugarcane. *Theor Appl Genet* 116: 701–714.
- Ren XY, Vorst O, Fiers MWEJ, Stiekema WJ, Nap JP (2006) In plants, highly expressed genes are the least compact. *Trends Genet* 22: 528–532.
- Rohlf FJ. (1998) NTSYSpc: Numerical Taxonomy and Multivariate Analysis System, version 2.0, Exeter Software, Setauket, New York.
- Rossi M, Araujo PG, Paulet F, Garsmeur O, Dias M, Chen H, VanSluys MA, D'Hont A (2003) Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. *Molec Genet Genomics* 269: 406–419.
- Saisug W, Ukoskit K (2013) Comparative analysis of EST-derived markers for allelic variation in *Jatropha curcas* L. and cross transferability among economically important species of Euphorbiaceae. *Genes Genom* 35: 1–12.
- Shu Y, Li Y, Zhu Y, Zhu Z, Lv D, Bai X, Cai H, Ji W, Guo D (2010) Genome-wide identification of intron fragment insertion mutations and their potential use as SCAR molecular markers in the soybean. *Theor Appl Genet* 121: 1–8.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
- Ukoskit K, Thipmongkolcharoen P, Chatwachirawong P (2012) Novel expressed sequence tag-simple sequence repeats (EST-SSR) markers characterized by new bioinformatic criteria reveal high genetic similarity in sugarcane (*Saccharum* spp.) breeding lines. *Afr J Biotechnol* 11: 1337–1363
- Vettore AL, Silva FRD, Kemper EL, Souza GM, Silva AMD, Ferro MIT, Henrique-Silva F, Giglioti EA, Lemos MVF, Coutinho LL, *et al.* (2003) Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. *Genome Res* 13: 2725–2735.
- Wang X, Zhao X, Zhu J, Wu W (2005) Genome-wide investigation of intron length polymorphisms and their potential as molecular markers in rice (*Oryza sativa* L.). *DNA Res* 12: 417–427.
- Wongpraneekul A (2008) Pedigree analysis of Thai sugarcane germplasm. Master of Science (Plant Breeding) Kasetsart University, Bangkok.
- Yang L, Jin G, Zhao X, Zheng Y, Xu Z, Wu W (2007a) PIP: a database of potential intron polymorphism markers. *Bioinformatics* 23: 2174–2177.
- Yang Y, Zhao XL, Xia LQ, Chen XM, Xia XC, Yu Z, He ZH, Roder M (2007b) Development and validation of a Viviparous-1 STS marker for pre-harvest sprouting tolerance in Chinese wheats. *Theor Appl Genet* 115: 971–980.
- Zeida M, Yu JK, Goldowitz I, Dentond ME, Costichd DE, Jayasuriyaa CT, Sahac M, Elshire R, Benscher D, Breseghellof F, *et al.* (2010) Cross-amplification of EST-derived markers among 16 grass species. *Field Crop Res* 118: 28–35.
- KEGG pathway, <http://www.genome.jp/kegg/pathway.html> (May 2015)
- Office of the cane and sugar board, <http://www.ocsb.go.th/th/home/index.php> (May 2015)
- Phytozome, <http://phytozome.jgi.doe.gov/pz/portal.html> (May 2015)