

บทคัดย่อ

การพัฒนาเครื่องหมายไมโครแซทเทลไลท์ด้วยวิธีมาตรฐานเป็นวิธีที่ค่อนข้างยุ่งยาก ใช้ระยะเวลานานและเสียค่าใช้จ่ายสูง ในปัจจุบันได้มีการค้นพบเครื่องหมายที่เป็นลำดับเบสซ้ำ (simple sequence repeat :SSR) ภายในลำดับเบสที่มีการแสดงออก (expressed sequence tags :ESTs) ซึ่งการพัฒนาเครื่องหมายไมโครแซทเทลไลท์แบบนี้เป็นวิธีที่ง่ายและรวดเร็วโดยการใช้ข้อมูลจากเหมืองข้อมูล EST ในงานวิจัยนี้ข้อมูล EST ปาล์มน้ำมันได้จากฐานข้อมูลสาธารณะ PalmGenes และศูนย์พันธุวิศวกรรมและเทคโนโลยีชีวภาพแห่งชาติ จำนวน 5,610 และ 75,948 ลำดับเบส ตามลำดับ นำข้อมูลลำดับเบสทั้งหมดทำการจัดกลุ่ม (cluster) และรวบรวมข้อมูล (assemble) ได้ 47,256 กลุ่ม จากนั้นค้นหา SSR โดยกำหนดชุดซ้ำของ mono- , di- , tri- , tetra- , penta- และ hexa nucleotide ให้มีจำนวนซ้ำ 10, 5, 5, 3, 3 และ 3 ตามลำดับ พบว่ามี EST ที่ภายในมี SSR จำนวน 1,570 EST-SSRs จากนั้นวิเคราะห์หน้าที่ พบว่า 120 ลำดับเบส เป็นยีนที่ทราบหน้าที่การทำงาน 5 ลำดับเบส เป็นโปรตีนสมมุติฐาน และ 1,445 ลำดับเบสเป็นกลุ่มที่ไม่พบความคล้ายคลึงกับอะไรเลยในฐานข้อมูล จากนั้นนำ EST-SSRs ที่มีจำนวนซ้ำที่สูงจำนวน 374 EST-SSRs ไปออกแบบไพรเมอร์ พบว่า 302 คู่ไพรเมอร์สามารถทำปฏิกิริยาพีซีอาร์สำเร็จ จากนั้นการประเมินโพลิมอร์ฟิซึมของเครื่องหมาย EST-SSR โดยใช้ปาล์มน้ำมัน 57 ตัวอย่าง พบว่ามี 284 คู่ไพรเมอร์ ที่แสดงโพลิมอร์ฟิซึม และสามารถสร้างแถบดีเอ็นเอเฉลี่ย 2.68 แอลลิลต่อคู่ไพรเมอร์ มีค่า PIC และ D เฉลี่ยเท่ากับ 0.34 และ 0.48 ตามลำดับ ในขณะที่เครื่องหมาย EST-SSRs สามารถเพิ่มปริมาณพีซีอาร์ได้ใน *E. oleifera* จำนวน 254 คู่ไพรเมอร์และ *Cocos nucifera* จำนวน 195 คู่ไพรเมอร์ เมื่อวิเคราะห์แนวโน้มการกระจายตัวของเครื่องหมาย EST-SSR พบว่า 65 เครื่องหมายมีการกระจายตัวเป็นไปตามกฎของเมนเดลในลูก F₁ จำนวน 10 ตัวและพ่อแม่ จากนั้นเปรียบเทียบระหว่างเครื่องหมาย EST-SSRs และ genomic SSRs (gSSR) พบว่า ค่า PIC ของเครื่องหมาย EST-SSRs มีค่าที่ต่ำกว่า จากนั้นทำการวิเคราะห์ Linkage disequilibrium ระหว่างเครื่องหมาย EST-SSRs และ gSSRs ที่สร้างแผนที่ได้ในงานวิจัยอื่น พบว่าไม่มีเครื่องหมาย SSR ใดที่แสดง linkage disequilibrium ผลการทดลองนี้ชี้ให้เห็นว่าการพัฒนาเครื่องหมายที่สัมพันธ์กับลักษณะที่ศึกษาโดยการวิเคราะห์ association mapping ควรใช้เทคนิคการวิเคราะห์ Linkage disequilibrium ภายในยีนที่เกี่ยวข้องกับลักษณะที่ศึกษา โดยใช้แหล่งพันธุกรรมที่หลากหลายของปาล์มน้ำมัน

Abstract

The conventional method for development of microsatellite makers have been time consuming and expensive. Currently, the discovery of simple sequence repeat (SSR) marker in expressed sequence tags (ESTs) provided the opportunity to develop microsatellites in a simple and direct way, by data mining of EST database. In the present study 5,610 and 75,948 oil palm ESTs generated by the PalmGenes and National center of genetic engineering and biotechnology, respectively, were investigated for identification and development of simple sequence repeat markers derived from expressed sequence tags (EST-SSR). All sequence data were clustered and assembled into 47,256 clusters. Based on the minimum repeat unit defined as 10, 5, 5, 3, 3 and 3 for mono-, di-, tri-, tetra-, penta- and hexa nucleotides, respectively, 1,570 EST containing SSRs were identified. Homology searching revealed that EST-SSRs represented 120, 5 and 1,445 clusters of known gene products, hypothetical proteins and unknown proteins, respectively. The 374 EST-SSRs based on high repeat unit were selected for primer designed. Three hundred and two primer pairs successfully amplified and 284 of them showed polymorphism among 57 oil palm accessions with an average allele number of 2.68 per locus, average Polymorphism Information Content of 0.34 and average Discrimination Power of 0.48. Two hundred and fifty-four and 195 primer pairs successfully cross-amplified in *E. oleifera* and *Cocos nucifera* (coconut), respectively. Sixty-five markers putatively fitted Mendelian segregation evaluated in 10 F1 randomly selected progenies and their two parents. Comparing EST-SSRs and genomic SSRs (gSSR) for their informativeness indicated the lower Polymorphism Information Content in EST-SSRs. Linkage disequilibrium between pairs of linked EST-SSR and gSSR markers mapped in the other study was investigated. None of linked SSR markers showed linkage disequilibrium suggesting that the large number of markers may be used for whole genome scan association mapping in diverse germplasm of highly heterozygous oil palm.