รายงานการวิจัยฉบับสมบูรณ์

การพัฒนาทัศนจักรอัตโนมัติสำหรับชีวสารสนเทศ

Development of Automatic Machine Vision for Bioinformatics

รศ. ดร. ปิติเขต สู้รักษา (หัวหน้าโครงการ)

นายชัยณรงค์ อมรบัญชรเวช

โครงการวิจัยนี้ได้รับการอุดหนุนโดยเงินรายได้ประจำปี 2553

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

บทคัดย่อ

งานวิจัยนี้เสนอการพัฒนาทัศนจักรสำหรับชีวสารสนเทศ ซึ่งเป็นการศึกษาเบื้องต้นในการนำองค์ความรู้ด้านการประมวลผลภาพร่วมกับระบบทัศนจักร ผลที่ได้จากการพัฒนานี้ได้องค์ความรู้และซอฟท์แวร์ในการวิเคราะห์ดี เอ็น เอ สำหรับประมวลผลภาพที่ได้จากการทดลองสร้างลู่วิ่งดี เอ็น เอ จากเจลอะกาโรส ผลที่ได้จากซอฟท์แวร์และระบบที่พัฒนาขึ้น สามารถใช้ทดแทนโปรแกรมเฉพาะด้านที่มีราคาแพงจากต่างประเทศ

คำสำคัญ: ดี เอ็น เอ, ทัศนจักร, การประมวลผลภาพ

Abstract

This research proposes development of machine vision for bioinformatics, which merges image processing and machine vision. This development yields software for DNA analysis. The input images are obtained from running Agarose gel. The developed software and proposed system can substitute an import and expensive software developed for the same functions.

Keywords: DNA, machine vision, image processing

RCH
QH
324.2
ป615ก

สถหมู่.....
เลขทะเบียน 114490
วัน,เดือน,ปี 20 ส.ค. 2554

b.12290824
i.....

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

งานวิจัยนี้ สำเร็จลงได้ด้วยเงินรายได้ พ.ศ. 2553 ขอขอบคุณคณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังที่ได้สนับสนุนเงินอุดหนุนวิจัยให้กับโครงการ

ขอขอบคุณ ผศ. ดร. สุชาติ ชะนะมา ภาควิชาชีวเคมี จุฬาลงกรณ์มหาวิทยาลัย ที่ให้ความรู้และให้ความช่วยเหลือองค์ความรู้ด้านปฏิบัติการ ดี เอ็น เอ และการวิเคราะห์ภาพดี เอ็น เอ จากเจลอะกาโรส

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

b.....
i.....

คำนำ

โครงงานนี้มุ่งเน้นการสร้างองค์ความรู้ทางด้านศาสตร์ผสานระหว่างชีววิทยากับวิทยาการสารสนเทศ โดยมีเป้าหมายที่จะสร้างต้นแบบแนวคิดทัศนจักรสำหรับ Bioinformatics ขึ้นเองเพื่อประยุกต์กับการวิเคราะห์ข้อมูลและผลการทดลองทางด้านชีววิทยาโมเลกุล และการคำนวณหาความคล้ายคลึงกันระหว่างดีเอ็นเอ เพื่อเป็นการลดการนำเข้าระบบราคาแพงจากต่างประเทศ อีกทั้งเป็นการสร้างองค์ความรู้ขึ้นในประเทศไทยอีกด้วย

ในเฟสของการวิจัยนี้ มุ่งศึกษาเชิงการศึกษาค้นคว้าเพื่อสร้างต้นแบบแนวคิดมากกว่าการผลิตในเชิงพาณิชย์ ผลการวิจัยที่ได้ยังคงต้องการปรับปรุงสู่ต้นแบบพร้อมผลิตอีกครั้งหนึ่ง ดังนั้น การใช้ประโยชน์จากงานวิจัยนี้ควรคำนึงถึงข้อจำกัดนี้ด้วย อย่างไรก็ตาม การศึกษานำร่องนี้น่าจะเป็นประโยชน์ต่อผู้ที่สนใจ

ปิติเขต สู้รักษา (หัวหน้าโครงการวิจัย)
ชัยณรงค์ อมรบัญชรเวช

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

บทที่ **หน้า**

1 บทนำ 1

2 เอกสารและงานวิจัยที่เกี่ยวข้อง 6

3 วิธีดำเนินการวิจัยและผลการวิจัย 18

4 อภิปรายผลการวิจัยและวิจารณ์ 25

5 สรุปและข้อเสนอแนะ 26

บรรณานุกรม 33

ภาคผนวก เทคนิคประมวลผลภาพสำหรับพัฒนาโปรแกรม 34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

รูปที่ **หน้า**

2.1 โครงสร้างโมเลกุลของดีเอ็นเอ.....8

2.2 ส่วนประกอบของดีเอ็นเอและโพลาริตี้ของสายดีเอ็นเอ.....9

2.3 การหาลำดับเบสจากลายภาพพิมพ์ดีเอ็นเอ.....11

2.4 ขั้นตอนในการทดลองโดยวิธี Agarose Gel Electrophoresis.....12

2.5 แสดงวิธี southern blotting 13

2.6 การเปรียบเทียบระหว่าง DNA ของเด็กทารกที่ปกติกับ DNA ของเด็กทารกไม่ปกติ.....15

3.1 แผนภาพขั้นตอนการทำงานของระบบ Bioinformatics 18

3.6 วงจรภายในและวงจรใช้งานจริงของตัวตรวจจับ.....18

3.7 วงจรควบคุมและแสดงผลทาง LCD ของค่าตรวจจับ.....19

4.1 Agarose Gel ที่ได้จากการทดลอง.....25

4.2 Agarose Gel เมื่อผ่านกระบวนการปรับแสงเรียบร้อยแล้ว.....26

4.3 Agarose Gel ก่อนและหลัง การผ่านกระบวนการลดสัญญาณรบกวน.....27

4.4 Agarose Gel เมื่อผ่านการตรวจจับขอบเรียบร้อยแล้ว.....28

4.5 ผลการวิเคราะห์ภาพที่ได้จากภาพ Agarose Gel ในเลนแรก.....29

4.6 โปรแกรม Bio-Images ที่ได้ทำการพัฒนาขึ้น.....29

4.7 ผลการทดลองเปรียบเทียบอัตราส่วนที่พบยีนในเลนจริงเทียบกับการทดลอง.....30

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ **หน้า**

1.1 กิจกรรมวิจัย.....3

4.1 ผลการทดลองยืนยันความถูกต้องของโปรแกรมที่พัฒนาขึ้น 31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ศาสตร์ Bioinformatics เป็นการสนธิระหว่างองค์ความรู้ทางด้าน Information, Computer Science และ Biology เพื่อนำมาประยุกต์ใช้ในงานทางด้านเทคโนโลยีชีวภาพ เช่น ประยุกต์กับงานทางด้านชีวเคมี จุลชีววิทยา เป็นต้น ตัวอย่างงานดังกล่าว เช่น การพิสูจน์เอกลักษณ์บุคคลโดยใช้ DNA Fingerprint การรักษาความปลอดภัย การค้นหายารักษาโรคใหม่ๆ เป็นต้น ซึ่งจะเห็นได้ว่า ระบบ Bioinformatics นั้นถูกนำมาใช้ประโยชน์อย่างแพร่หลายและกว้างขวางมากขึ้น [1-8]

ในประเทศไทยเริ่มมีการนำระบบ Bioinformatics มาใช้งานมากขึ้นโดยเฉพาะงานทางด้านวิทยาศาสตร์ชีวภาพ ได้แก่ ชีวเคมี ที่นำระบบ Bioinformatics มาวิเคราะห์ข้อมูลที่ได้จากผลการทดลอง เพื่อให้ได้ผลลัพธ์ที่ถูกต้องแม่นยำและรวดเร็ว โดยระบบที่เรานำมาใช้กันในประเทศนั้นจะเป็นระบบที่ต้องใช้เครื่องมือสนับสนุนโดยนำเข้าจากต่างประเทศ ซึ่งมีราคาสูงมาก ทำให้เราต้องเสียเงินจำนวนมากให้กับต่างชาติ เพื่อนำระบบ Bioinformatics มาใช้ประโยชน์ หนึ่งในเครื่องมือนั้นคือทัศนจักร หรือ machine vision ซึ่งเป็นการตรวจสอบความถูกต้องหรือวิเคราะห์ข้อมูลด้วยเทคโนโลยีตาหุ่นยนต์

ดังนั้นโครงงานนี้จึงมีเป้าหมายที่จะสร้างต้นแบบแนวคิดทัศนจักรสำหรับ Bioinformatics ขึ้นเองเพื่อประยุกต์กับการวิเคราะห์ข้อมูลและผลการทดลองทางด้านชีววิทยาโมเลกุล และการคำนวณหาความคล้ายคลึงกันระหว่างดีเอ็นเอ เพื่อเป็นการลดการนำเข้าระบบราคาแพงจากต่างประเทศและเป็นการสร้างองค์ความรู้ขึ้นในประเทศไทยอีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์

เพื่อพัฒนาต้นแบบแนวคิดทัศนจักรสำหรับชีวสารสนเทศ

1.3 ขอบเขตของงานวิจัย

งานวิจัยนี้สร้างต้นแบบแนวคิดซึ่งศึกษาในแง่เชิงการประดิษฐ์เป็นส่วนใหญ่ และเนื่องจากงบประมาณที่จำกัดอย่างมาก จึงศึกษาได้เฉพาะโครงสร้างอุปกรณ์ที่มีราคาถูกเท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 1.1 กิจกรรม

ผู้รับผิดชอบ คณะวิจัย

เวลาทำงานวิจัย : 1 ปี

ขั้นตอนการทำงาน	ปีที่1			ปีที่2			ผู้รับผิดชอบ
	เดือน 1 - 4	เดือน 5 - 8	เดือน 9 - 12	เดือน 1 - 4	เดือน 5 - 8	เดือน 9 - 12	
ศึกษาค้นคว้าหาข้อมูลที่จำเป็นในการทำงานวิจัยเกี่ยวกับระบบ Bioinformatics และ Biochemistry และสร้างต้นแบบทัศนจักร							นายชัยณรงค์ อมรบัญชรเวช รศ.ดร.ปิติเขต สู้รักษา ผศ.ดร.สุชาติ ชะนะมา
ทำการเก็บผลการทดลอง ของ Agarose Gel ที่ผ่านกระบวนการ Electrophoresis							นายชัยณรงค์ อมรบัญชรเวช ผศ.ดร.สุชาติ ชะนะมา
ทำการออกแบบและสร้างระบบในส่วน Image Processing							นายชัยณรงค์ อมรบัญชรเวช รศ.ดร.ปิติเขต สู้รักษา
ทำการออกแบบและสร้างระบบในส่วน Pattern Recognition							นายชัยณรงค์ อมรบัญชรเวช
ทำการออกแบบและสร้างระบบในส่วน Numerical							นายชัยณรงค์ อมรบัญชรเวช รศ.ดร.ปิติเขต สู้รักษา
ทำการออกแบบและสร้างระบบในส่วน Data Clustering							นายชัยณรงค์ อมรบัญชรเวช รศ.ดร.ปิติเขต สู้รักษา
ตรวจสอบความถูกต้องและจัดทำรายงานฉบับเต็ม							นายชัยณรงค์ อมรบัญชรเวช รศ.ดร.ปิติเขต สู้รักษา

อุปกรณ์ที่มีอยู่แล้ว

1. Computer 1 เครื่อง
2. กล้องสำหรับเก็บภาพผลการทดลอง 1 เครื่อง
3. ระบบส่องสว่าง 1 ชุด
4. สารเคมีและน้ำยาทางด้านชีววิทยาโมเลกุล
5. หนังสือและเอกสารงานวิจัย

สถานที่ที่ใช้วิจัย

- ห้องปฏิบัติการทางด้านชีววิทยาโมเลกุล (Molecular Biology) จุฬาฯ
- ห้องปฏิบัติการทางด้านวิศวกรรมสารสนเทศ (Information Engineering Lab) สจล.

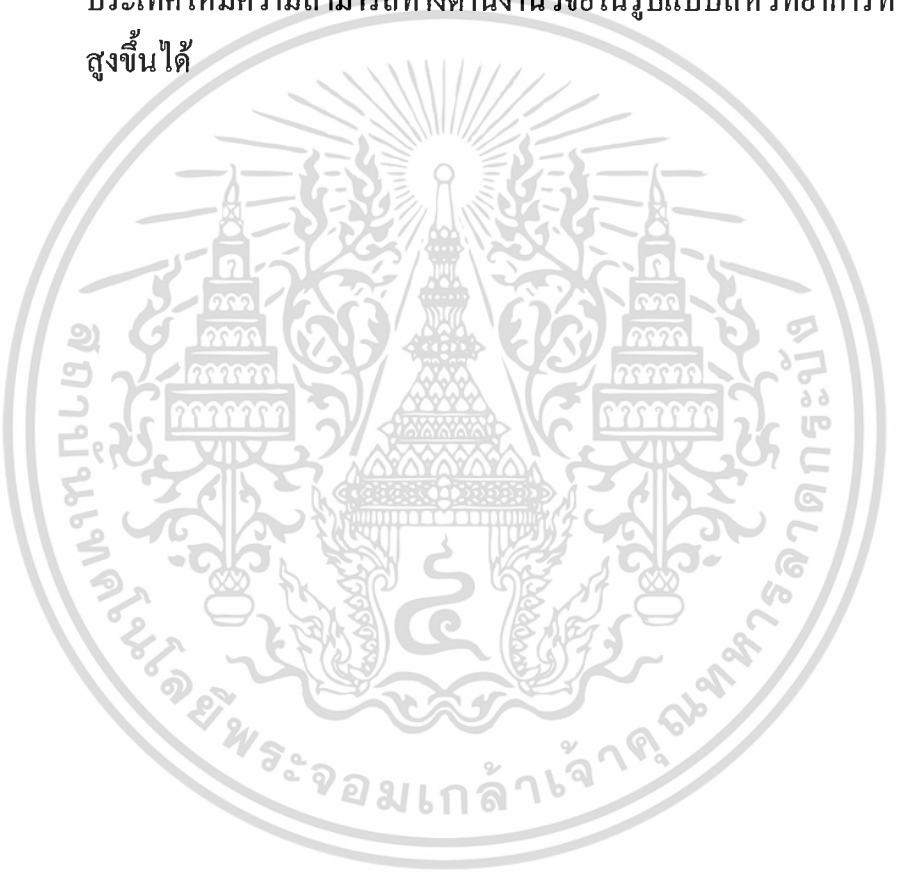
งบประมาณ

รายการประมาณการณ์ค่าใช้จ่าย	งบประมาณ (บาท)
ค่าวัสดุอิเลคทรอนิคส์ สำหรับระบบส่องสว่างและอื่น ๆ	**39,000**
ค่าวัสดุเชิงกลและเคมีสิ้นเปลืองในงานชีววิทยาโมเลกุล	อนุเคราะห์จากห้องปฏิบัติการฯ จุฬาฯ
ค่าวัสดุสำนักงานและวัสดุทางคอมพิวเตอร์และค่าวัสดุใช้สอยอื่นๆ	**20,000**
รวม	**59,000**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลที่คาดว่าจะได้รับ

1. สามารถนำระบบทัศนจักรที่ออกแบบขึ้นมาใช้ทดสอบและเปรียบเทียบคุณสมบัติระหว่างสารชีวโมเลกุล
2. สามารถนำเครื่องต้นแบบแนวคิดที่ได้มาวิเคราะห์ข้อมูลเพื่อนำไปใช้ประโยชน์ทางด้าน Biochemistry และ Biology
3. สามารถนำไปปรับปรุงเพื่อสร้างต้นแบบใหม่ในเฟสถัดไปเพื่อใช้ทดแทนทัศนจักรและโปรแกรมราคาแพงจากต่างประเทศได้
4. ส่งเสริมการสร้างขีดความสามารถและพัฒนาศักยภาพของบุคลกรในประเทศให้มีความสามารถทางด้านงานวิจัยในรูปแบบสหวิทยาการที่สูงขึ้นได้



บทนำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

บทนี้กล่าวถึงทฤษฎีที่เกี่ยวข้องในการประมวลผลภาพเพื่อใช้ในการวิเคราะห์โครงสร้าง DNA และโปรตีน Fragment ของจุลชีพ โดยจะแบ่งเป็นสองส่วนหลักๆ คือ ในส่วนของความรู้เบื้องต้นเกี่ยวกับ DNA หลักการและเทคนิคที่ใช้ใน recombinant DNA technology และในส่วนของการประมวลผลภาพเพื่อใช้ในการวิเคราะห์โครงสร้างของ DNA

2.1 ความรู้เบื้องต้นเกี่ยว DNA

สิ่งมีชีวิตส่วนใหญ่ไม่ว่าจะเป็นต้นไม้ สุนัข คน ล้วนประกอบขึ้นมาจากเซลล์ โดยที่แต่ละเซลล์มีลักษณะพื้นฐานคล้ายๆกัน ก็คือมีเยื่อหุ้มเซลล์ล้อมรอบและมีนิวเคลียสอยู่ตรงกลางคล้ายไข่แดงอยู่กลางไข่ขาว ภายในนิวเคลียสก็จะมีสารพันธุกรรมอยู่ สารพันธุกรรมนี้มีชื่อเรียกว่า โครโมโซม ซึ่งในกรณีของมนุษย์มีอยู่ด้วยกัน 23 คู่ โดยโครโมโซมแต่ละคู่ครึ่งหนึ่งได้รับมาจากแม่ และอีกครึ่งหนึ่งได้รับมาจากพ่อ แต่ละโครโมโซมนั้นประกอบขึ้นสารพันธุกรรมที่เรียกว่า DNA ซึ่งย่อมาจาก Deoxyribonucleic acid มีโครงสร้างเป็นเส้นคู่บิดเข้าหากันเป็นเกลียว สารพันธุกรรมใน DNA ประกอบด้วยการจับตัวกันของสารเคมี 4 ชนิด ที่เรียกว่าเบส คือ A G C T สารเคมี 4 ชนิดนี้จะเรียงตัวกันเป็นลักษณะของการเข้ารหัสโดยเรียงลำดับสลับไปสลับมามีความยาวกว่า 3,000 ล้านต่อเส้น เราเรียกว่ารหัสพันธุกรรมของสิ่งมีชีวิต ซึ่งทำหน้าที่ควบคุมลักษณะของสิ่งมีชีวิต

สิ่งมีชีวิตแต่ละชนิดก็จะมีรหัสพันธุกรรมที่แตกต่างกันออกไป ซึ่งสามารถนำมาใช้ประโยชน์ได้ทั้งในการแพทย์ การสืบสวนสอบสวน เช่น ใช้แทนลายนิ้วมือที่เรียกว่า DNA Fingerprint นอกจากนี้ยังช่วยให้การศึกษาด้านชีววิทยาทำได้ง่ายขึ้นด้วย DNA จะช่วยเสริมข้อมูลให้กับนักชีววิทยาในการหา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความสัมพันธ์ของสิ่งมีชีวิตแต่ละชนิด วิวัฒนาการของมัน หรือแม้กระทั่งถิ่นที่อยู่ของสิ่งมีชีวิตที่สนใจ

2.1.1 โครงสร้างโมเลกุลใน DNA

โมเลกุลดีเอ็นเอ (DNA) หรือ Deoxyribonucleic Acid เป็นสารพันธุกรรมที่กำหนดรหัสพันธุกรรม (Genetic codes) จัดเรียงตามข้อกำหนดจำเพาะของแต่ละบุคคล ทำหน้าที่เก็บสะสมและถ่ายทอดข้อมูลทางพันธุกรรม นักวิทยาศาสตร์ได้ประยุกต์ใช้เทคนิคต่างๆจากความรู้เรื่องคุณสมบัติอันเป็นเอกลักษณ์นี้ นำไปใช้ในการตรวจพิสูจน์เอกลักษณ์บุคคล (Human Identification) นิติวิทยาศาสตร์ (Forensic Science) และโบราณคดี (Molecular Paleontology) เป็นต้น

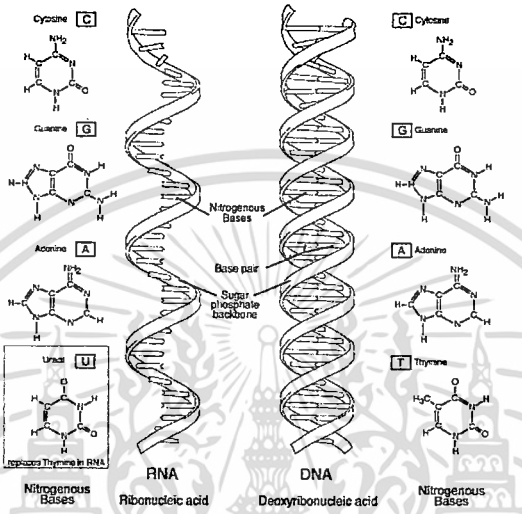
ดีเอ็นเอจะพบในสิ่งมีชีวิตเกือบทุกชนิด ซึ่งพบอยู่ 2 แห่งในเซลล์ที่มีชีวิตเกือบทุกๆเซลล์คือ นิวเคลียร์ (Nuclear DNA) และไมโตคอนเดรียลดีเอ็นเอ (Mitochondrial DNA) นิวเคลียร์ดีเอ็นเอพบในนิวเคลียร์ของเซลล์เท่านั้น โดยเซลล์นิวเคลียร์เป็นศูนย์ควบคุมการทำงานของเซลล์ทั้งหมด นิวเคลียร์ดีเอ็นเอนี้บรรจุลักษณะที่ถูกสืบทอดมาของสิ่งมีชีวิตแต่ละชนิดเช่น สีนัยน์ตา ความสูง ฯลฯ ซึ่งเป็นผลผลิตของดีเอ็นเอที่ได้มาจากพ่อและแม่ ดังนั้นนิวเคลียร์ดีเอ็นเอจึงเป็นครึ่งหนึ่งที่มาจากแม่และครึ่งหนึ่งที่มาจากพ่อ ซึ่งสามารถนำคุณสมบัตินี้มาใช้พิสูจน์ความสัมพันธ์ของพ่อแม่ ลูก และความสัมพันธ์ทางเครือญาติได้

ไมโตคอนเดรียลดีเอ็นเอพบในไมโตคอนเดรียซึ่งอยู่ในตัวเซลล์เท่านั้นโดยไมโตคอลเดรียเป็นแหล่งผลิตพลังงานของเซลล์ ไมโตคอนเดียลดีเอ็นเอจะได้รับการสืบทอดมาจากทางแม่เท่านั้น โดยถูกส่งผ่านจากยายามาสู่แม่ และแม่สู่ลูก จะมีลักษณะเหมือนกับไมโตคอนเดรียลดีเอ็นเอของแม่ทุกประการ ดังนั้นจึงเป็นเครื่องหมายทางลักษณะที่ดีสำหรับการพิสูจน์ว่าเป็นญาติฝ่ายแม่ได้

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงสร้างโมเลกุลของดีเอ็นเอโดยทั่วไปจะมีสายนิวคลีโอไทด์(Nucleotide) สายยาว 2 ยาว ควั่นกันเป็นเกลียวคู่ (Double helix) โดยมีทิศทางสวนกันแบบทิสทางตรงกันข้าม (Antiparallel) โดยนิวคลีโอไทด์ประกอบด้วยน้ำตาลไรโบส (Ribose) ซึ่งประกอบด้วยคาร์บอน 5 อะตอม ฟอสเฟส และ ออร์แกนิคเบส (Organic base) ที่มีไนโตรเจนเป็นส่วนประกอบ ดังแสดงในรูปที่ 2.1

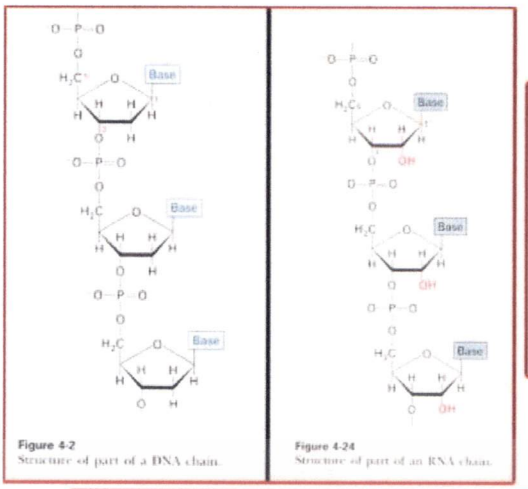


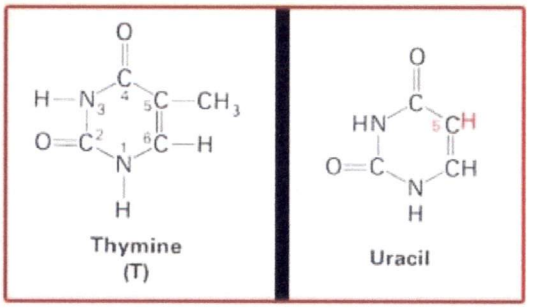
รูปที่ 2.1 โครงสร้างโมเลกุลของดีเอ็นเอ

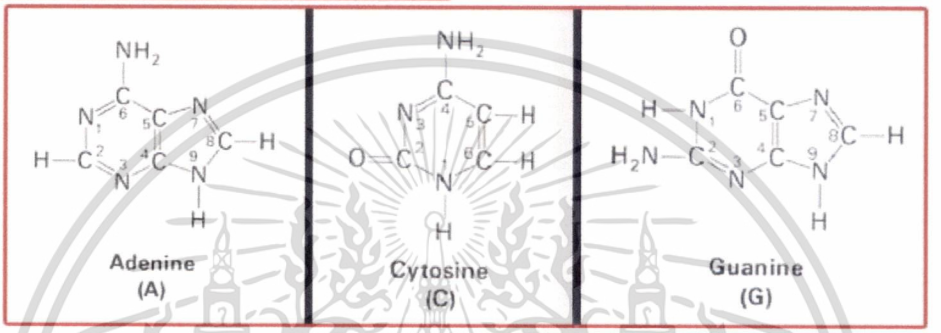
สายดีเอ็นเอ จะมีฟอสเฟตและน้ำตาลเป็นแกนสันหลัง คล้ายราวบันไดที่บิดเป็นเกลียว และมีเบสยื่นออกมาจากแกนสันหลังแต่ละข้างคล้ายขั้นบันได ซึ่งความแตกต่างของแต่ละนิวคลีโอไทด์ขึ้นอยู่กับเบสบางครั้งจึงเรียกว่าเบสแทนนิวคลีโอไทด์ เบสที่เป็นส่วนประกอบของดีเอ็นเอมี 2 ชนิด คือ พัวรีน (Purine) ซึ่งมีโครงสร้างเป็นวงแหวนคู่ และไพริมิดีน (Pyrimidine) ซึ่งมีโครงสร้างเป็นวงแหวนเดี่ยว เบสที่เป็นพัวรีนได้แก่ อะดินีน กับกัวนีน ส่วนเบสที่เป็นไพริมิดีนได้แก่ ไธมีน กับไซโทซีน และเบสจะจับกับคาร์บอนตำแหน่งที่ 1' ของน้ำตาลไรโบส (คาร์บอนในตำแหน่งต่างๆของน้ำตาลใส่เครื่องหมาย ' (Prime) ก็เพื่อความแตกต่างจากตำแหน่งคาร์บอนในเบส) ดังแสดงในรูปที่ 2.2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้







รูปที่ 2.2 ส่วนประกอบของดีเอ็นเอและโพลาริตี้ของสายดีเอ็นเอ

นอกจากนั้นดีเอ็นเอยังถูกสันนิษฐานว่าเป็นหน่วยที่ใช้ในการควบคุมลักษณะนิสัยของบุคคลต่างๆเนื่องจากดีเอ็นเอมีลักษณะที่สำคัญที่บ่งชี้ดังนี้

- ดีเอ็นเอส่วนใหญ่จะปรากฏที่โครโมโซม ส่วนโปรตีนและอาร์เอ็นเอนั้นมักจะพบในไซโตพลาสซึม

- ปริมาณของดีเอ็นเอจะสัมพันธ์กับจำนวนชุดของโครโมโซม คือ ปริมาณของดีเอ็นเอในเซลล์ร่างกาย (diploid) จะเป็น 2 เท่าของปริมาณดีเอ็นเอในหน่วยสืบพันธุ์ (haploid)

- องค์ประกอบของดีเอ็นเอ หรือปริมาณของนิวคลีโอไทด์ จะคงที่ในเซลล์ของทุกส่วนของร่างกาย แต่องค์ประกอบของโปรตีนและอาร์เอ็นเอจะแปรปรวนขึ้นลงในทุกๆเซลล์

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

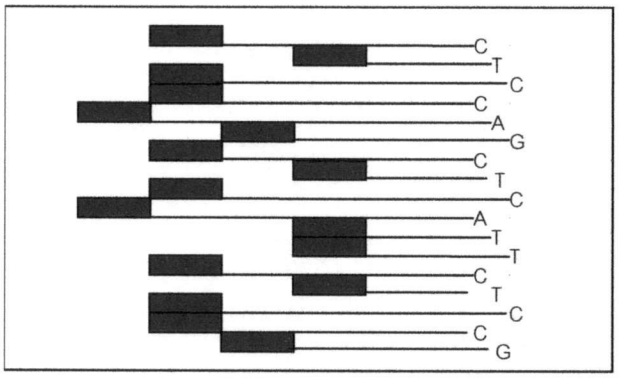
คุณลักษณะ 2 ประการหลังนี้เป็นคุณสมบัติที่หน่วยควบคุมลักษณะจะต้องมี ซึ่งพบคุณสมบัติเช่นนี้ในดีเอ็นเอเท่านั้น แต่จะไม่พบในโปรตีนหรืออาร์เอ็นเอ นอกเหนือจากนั้นยังมีการทดลองที่พิสูจน์ได้ว่าดีเอ็นเอ คือหน่วยควบคุมลักษณะจริงอีกด้วย

2.1.2 ฐานข้อมูลลำดับเบสของดีเอ็นเอ

นักพันธุศาสตร์ได้นำสายดีเอ็นเอไปตรวจวิเคราะห์ และได้ภาพแถบดีเอ็นเอออกมามีลักษณะคล้ายบาร์โคด (Bar-code) ที่สามารถบ่งชี้เอกลักษณ์บุคคลคล้ายกับภาพลายพิมพ์นิ้วมือ ซึ่งได้บัญญัติคำศัพท์ไว้ว่า DNA Fingerprint หรือลายพิมพ์ดีเอ็นเอ ดีเอ็นเอในบุคคลเดียวกันจะมีลายพิมพ์เหมือนกันไม่ว่าจะได้จากเซลล์ส่วนใดของร่างกายเช่น เม็ดเลือดขาว เส้นผม ตัวอสุจิ เป็นต้น ในกรณีฝาแฝดที่เกิดจากการปฏิสนธิจากไข่ฟองเดียวกัน (Monozygotic Twins) จะมีลายพิมพ์ดีเอ็นเอเหมือนกัน ในการนำสายดีเอ็นเอไปใช้ในงานต่างๆ เช่น พิสูจน์ความเป็นพ่อแม่ลูก พิสูจน์ในคดีฆาตกรรม นักพันธุศาสตร์จะนำสายดีเอ็นเอไปตัดด้วยเอนไซม์จำเพาะชนิดที่กำหนดไว้แล้วตามช่วงของสายดีเอ็นเอที่ต้องการพิสูจน์ และทำลายพิมพ์ดีเอ็นเอออกมา จะเห็นว่าในการเก็บข้อมูลเกี่ยวกับดีเอ็นเอหากเก็บเป็นภาพลายพิมพ์ดีเอ็นเอของสิ่งมีชีวิตทุกชนิดบนโลก จะได้ฐานข้อมูลที่มีขนาดใหญ่มาก ดังนั้นการเก็บข้อมูลดีเอ็นเอจึงเก็บในรูปของตัวอักษรเรียงต่อกันเป็นลำดับเบส ซึ่งมีตัวอักษรอยู่ 4 ตัว คือ A G C และ T เป็นชนิดของเบสดังที่กล่าวมาแล้วข้างต้น และตัวอักษร N แทนลำดับเบสที่ไม่รู้จัก ซึ่งการนำลายพิมพ์ดีเอ็นเอที่ได้ไปหาลำดับเบสนั้นมีลักษณะดังรูปที่ 2.3 ซึ่งจะได้ลำดับเบสของดีเอ็นเอคือ CTCCAGCTCATTCTCCG ถึงแม้ว่าฐานข้อมูลจะเก็บอยู่ในรูปของลำดับเบสก็ตาม แต่ก็ยังมีขนาดใหญ่มากอยู่นั่นเอง

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 การหาลำดับเบสจากลายภาพพิมพ์ดีเอ็นเอ

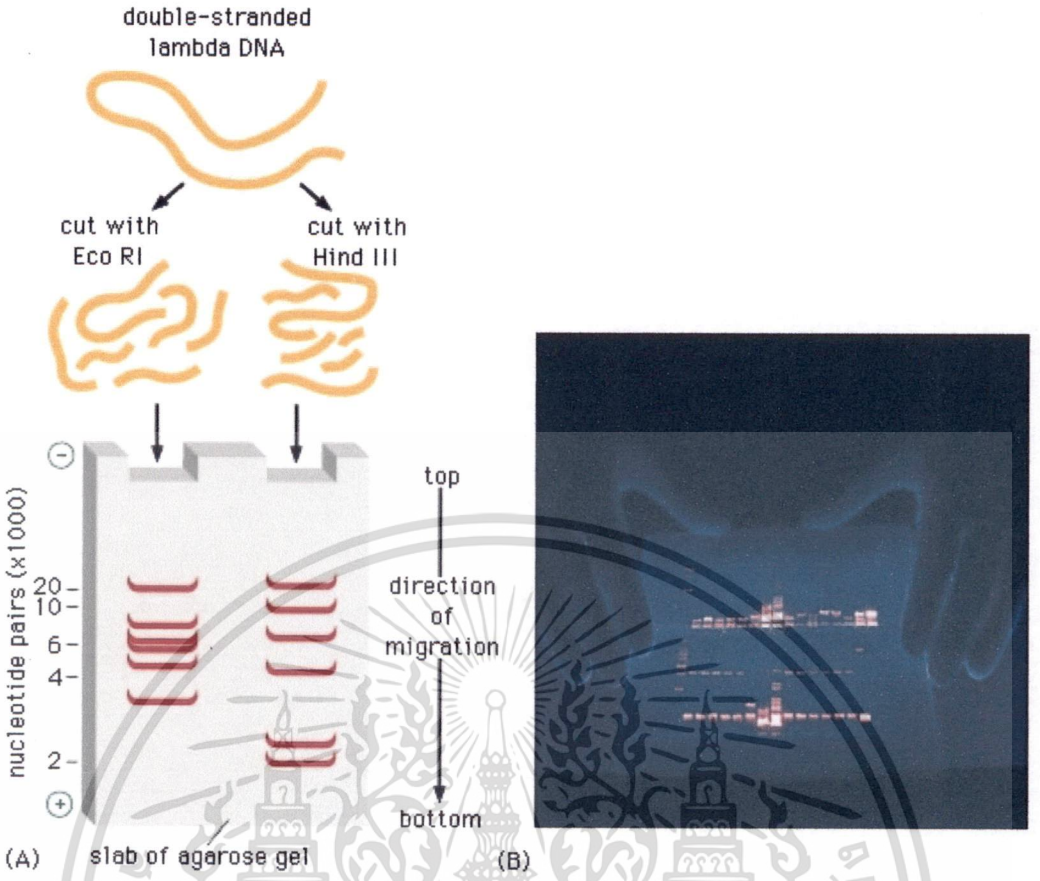
2.2 หลักการและเทคนิคที่ใช้ Recombinant DNA technology

การสร้าง Recombinant DNA และการตัดต่อพันธุกรรมนั้น จะสำเร็จได้ต้องอาศัยหลักการพื้นฐานและเทคนิคหลาย ๆ อย่างมาประกอบกัน ซึ่งมีดังต่อไปนี้

2.2.1 Agarose Gel Electrophoresis

Agarose gel Electrophoresis เป็นเทคนิคการแยกชิ้นดีเอ็นเอที่มีขนาดแตกต่างกันออกจากกันด้วยสนามไฟฟ้า ทั้งนี้เพราะดีเอ็นเอมีประจุลบจากหมู่ฟอสเฟต ดีเอ็นเอที่มีขนาดใหญ่ซึ่งมีมวลมากก็จะมีจำนวนประจุลบมากขึ้นด้วย ดังนั้นเมื่อพิจารณาสัดส่วนของประจุต่อมวลดีเอ็นเอ แล้ว จะเป็นค่าคงที่ agarose gel electrophoresis แยกดีเอ็นเอออกจากกันได้โดยอาศัยแรงเสียดทานการเคลื่อนที่ของดีเอ็นเอ แรงเสียดทานนี้ขึ้นอยู่กับรูปร่างและขนาดของดีเอ็นเอ โดยที่ดีเอ็นเอที่มีรูปร่างอัดแน่นอย่าง supercoiled DNA จะมีแรงเสียดทานน้อย จึงเคลื่อนที่ได้เร็วกว่า circular DNA และ ดี เอ็นเอ รูปร่างเป็นแท่ง ดีเอ็นเอ ที่มีรูปร่างเหมือนกันอย่างชิ้น ดีเอ็นเอ รูปแท่งนั้น ดีเอ็นเอ ที่มีขนาดเล็ก จะเคลื่อนที่ได้เร็วกว่า ดีเอ็นเอที่มีขนาดใหญ่ซึ่งมีแรงเสียดทานการเคลื่อนที่มากกว่า เทคนิคนี้แสดงดังรูปที่ 2.4

เอกสารและงานวิจัยที่เกี่ยวข้อง
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.4 ขั้นตอนในการทดลองโดยวิธี Agarose Gel Electrophoresis

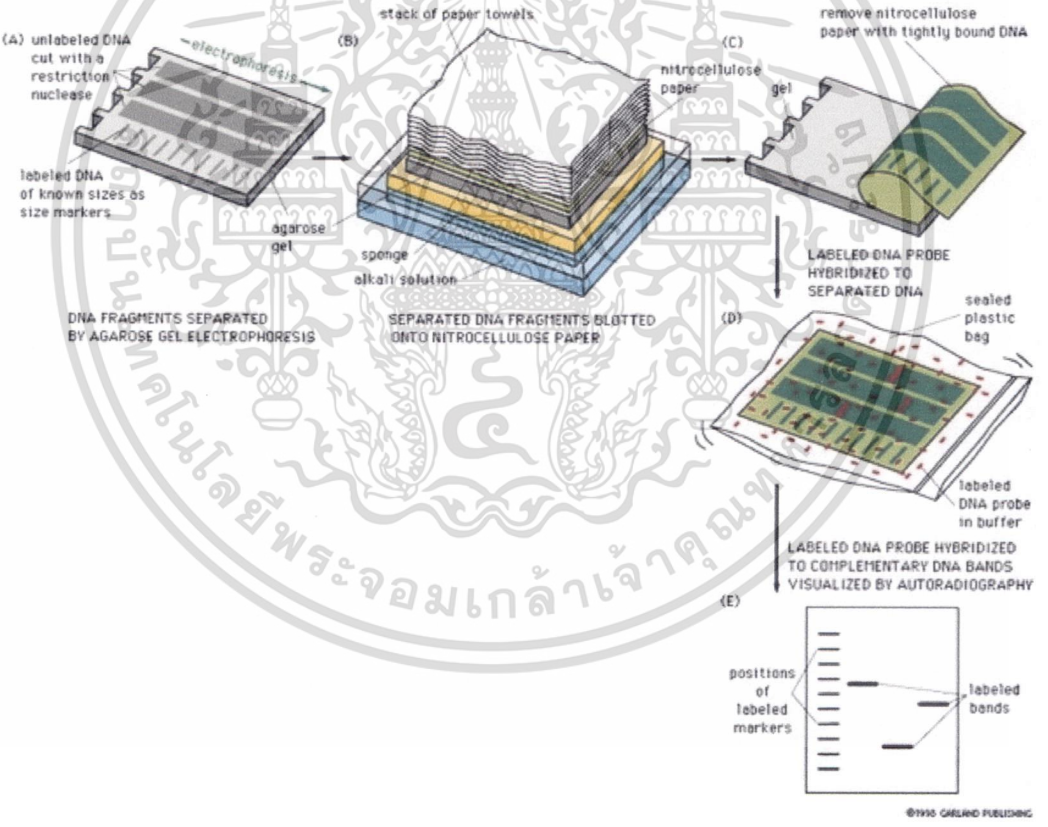
ดังนั้นเมื่อตัดดีเอ็นเอเป็นชิ้น ๆ ด้วย restriction endonuclease จึงสามารถแยกชิ้นดี เอ็น เอ ที่มีขนาดแตกต่างกันได้ด้วย agarose gel electrophoresis กระทำโดยเตรียม agarose gel ที่มีความเข้มข้นที่เหมาะสม หยดสารละลายดีเอ็นเอลงในช่องบน agarose gel แล้วเปิดสนามไฟฟ้าดีเอ็นเอ จะวิ่งจากขั้วลบเข้าหาขั้วบวก เมื่อได้เวลาที่เหมาะสม ปิดสนามไฟฟ้าแล้วย้อมดีเอ็นเอใน agarose gel ด้วยสี acridine เช่น ethidium bromide แล้วนำไปส่องในแสงอัลตราไวโอเลต ethidium bromide ที่ย้อมติดดีเอ็นเออยู่จะเรืองแสงสีส้ม เป็นแถบ ดี เอ็น เอ ซึ่งมองเห็นได้ด้วยตาเปล่า

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2 Southern Blot Hybridization

เมื่อแยกชิ้นดีเอ็นเอที่มีขนาดต่างกันออกจากกันด้วย agarose gel electrophoresis แล้วหากต้องการทำ hybridization เพื่อตรวจหาว่าดีเอ็นเอชิ้นที่มียีนที่ต้องการศึกษาอยู่ในตำแหน่งใดของ electrogram เราจำเป็นที่จะต้องย้ายดีเอ็นเอที่แยกแล้วนั้นออกจาก agaroes gel ไปสู่แผ่น nitrocellulose ซึ่งมีความคงทนมากกว่า ทั้งนี้เพราะว่าระหว่างการ hybridization จะต้องมีการเพิ่มอุณหภูมิเพื่อให้ ดีเอ็นเอแยกสาย ก่อนที่จะจับคู่เบสใหม่กับ DNA probe ความร้อนที่ใช้นี้สามารถทำลาย agaroes gel ให้หลอมละลายหายไปได้ เทคนิคนี้แสดงในรูปที่ 2.5



รูปที่ 2.5 แสดงวิธี southern blotting

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

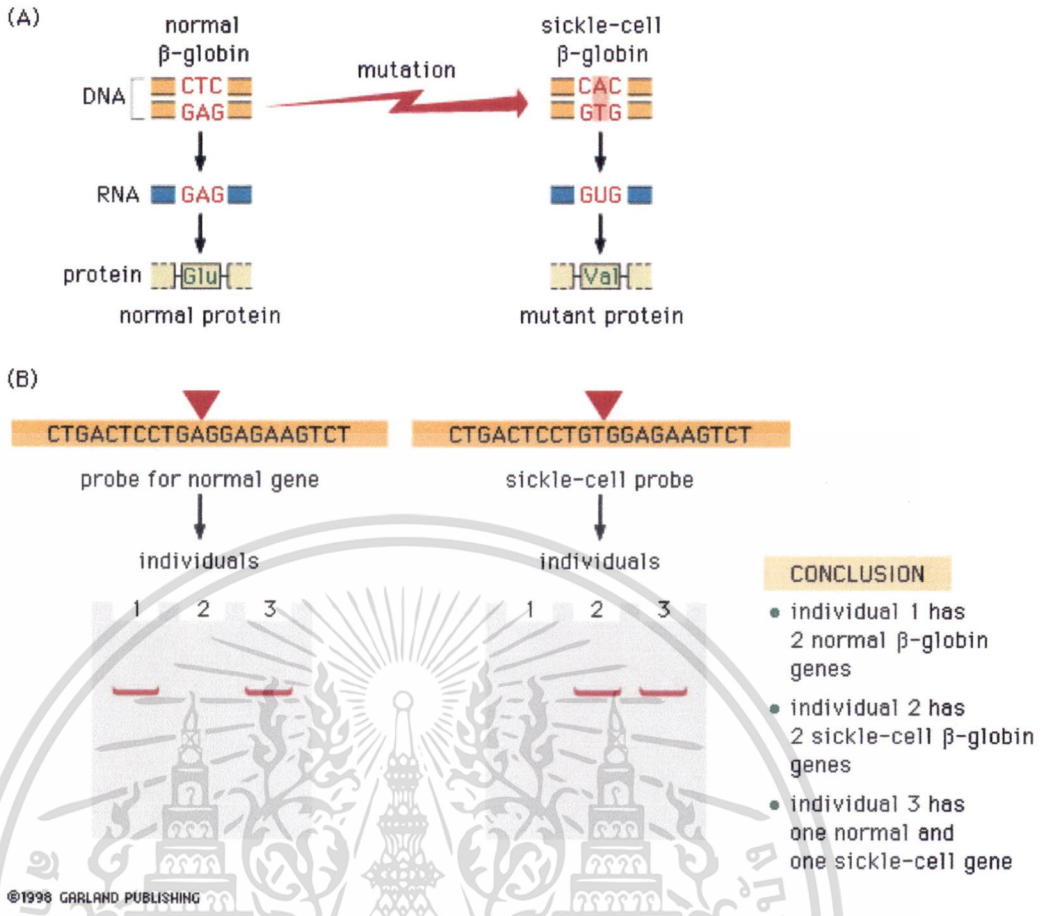
การเคลื่อนย้ายชิ้นดีเอ็นเอ จาก agarose gel ไปสู่แผ่น nitrocellulose ทำโดยการวางแผ่นnitrocellulose ไว้บน agarose gel ที่แช่สารละลายบัฟเฟอร์อยู่ แล้ววางกระดาษซับหลาย ๆ ชั้นบนแผ่น nitrocellulose อีกทีหนึ่ง แผ่นกระดาษซับที่แห้งจะดูดซับสารละลายบัฟเฟอร์ที่อยู่ข้างล่างให้เคลื่อนที่ขึ้น ชะผ่าน agarose gel และแผ่น nitrocellulose สู่กระดาษซับ ในขณะที่เคลื่อนที่ สารละลายบัฟเฟอร์จะชะเอาชิ้นดีเอ็นเอ ให้หลุดออกจาก agarose gel ไปเกาะติดกับแผ่น nitrocellulose ไปด้วย โดยวิธีนี้เราจึงสามารถย้ายชิ้นดีเอ็นเอจาก agarose gel ไปสู่ nitrocellulose ได้ วิธีการนี้เรียกว่า southern blot และถ้านำ DNA probe มาจับกับชิ้นดีเอ็นเอ ที่ย้ายมาอยู่บน nitrocellulose อย่างจำเพาะ ต่ออีกทีหนึ่ง วิธีการทั้งหมดที่ประกอบด้วย agarose gel electrophoresis, southern blot และ hybridization จะเรียกว่า southern blot hybridization

ด้วยกระบวนการเดียวกันกับ southern blot หากกรดนิวคลีอิกที่ย้ายจาก agarose gel สู่ nitrocellulose เป็น อาร์ เอน เอ แล้ว จะเรียกกระบวนการนี้ว่า northern blot ส่วน western blot หมายถึงการย้ายโปรตีนจาก acrilamide gel สู่ nitrocellulose การย้ายโปรตีนไม่ได้ใช้การซับสารละลายบัฟเฟอร์ด้วยกระดาษซับ แต่จะย้ายโปรตีนด้วยสนามไฟฟ้าแทน

Hybridization เป็นเทคนิคที่มีการประยุกต์ใช้อย่างกว้างขวาง ทั้งในด้านงานวิจัย และในทางการแพทย์ ตัวอย่างของการประยุกต์ได้แก่การตรวจหายีน β-globin ที่ผิดปกติของทารกในครรภ์ Sickle cell anemia เกิดจาก base substitution ชนิด transversion ใน β -globin geneส่งผลทำให้มี missense mutation สาย β -globin ตรงที่เคยเป็น Glu ถูกอ่านรหัสผิดไปเป็น Val ดังรูปที่ 2.6

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 การเปรียบเทียบระหว่าง DNA ของเด็กทารกที่ปกติกับ DNA ของเด็กทารกไม่ปกติ

ด้วยเทคนิค hybridization โดยใช้ DNA probe สองชนิดคือ DNA probe สำหรับ β -globin gene ปกติ กับ DNA probe สำหรับ sickle cell gene หากเจาะน้ำคร่ำนำเซลล์จากทารกมาสกัด ดี เอ็น เอ แล้วแยกชิ้น ดี เอ็น เอ ออกจากกันด้วย agarose gel electrophoresis แล้วทำ hybridization ด้วย DNA probe สองชนิดแล้ว หาก ดี เอ็น เอ ของทารกใดให้ผลบวกกับ DNA probe ของ β -globin gene ปกติ ทารกนั้น ไม่เป็นโรค แต่ถ้าหาก ดี เอ็น เอ ของทารกใดให้ผลบวกกับ DNA probe สำหรับ sickle cell gene ทารกนั้นเป็นโรค sickle cell anemiaและหาก ดี เอ็น เอ ของทารกใดให้ผลบวกกับ DNA probe ทั้งสองชนิดแล้ว ทารกนั้นเป็น heterozygote ที่มีทั้ง β-globin geneปกติ และ sickle cell gene

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 การประมวลผลภาพเพื่อใช้ในการวิเคราะห์โครงสร้างของ DNA

การประมวลผลภาพเชิงดิจิตอล (Digital image processing) เป็นการนำคอมพิวเตอร์มาใช้ในการประมวลผลภาพในรูปดิจิตอลฟอร์แมต ซึ่งมีหมวดหมู่ใหญ่ๆ ดังนี้

- Image acquisition เป็นขั้นตอนแรกสุดของการประมวลผลภาพ คือการรับภาพที่ต้องการประมวลผลเข้ามา
- Image enhancement เป็นการปรับปรุงภาพเพื่อให้ภาพมีคุณสมบัติที่เหมาะสมสำหรับการประมวลผลขั้นต่อไป โดยใช้กระบวนการทางคณิตศาสตร์ต่าง ๆ
- Image restoration เป็นกระบวนการปรับปรุงภาพออกมาให้เหมือนภาพตั้งต้น ข้อแตกต่างระหว่าง
- Image restoration กับ Image enhancement คือ Image restoration เป็นการทำให้ภาพเหมือนภาพเดิมมากที่สุด แต่ Image enhancement ไม่สนใจว่าภาพที่ออกมาจะเหมือนภาพเดิมหรือไม่แต่ต้องการให้ภาพมีคุณภาพที่เหมาะสมที่จะนำไปประมวลผลในขั้นตอนถัดไป
- Image compression เป็นเทคนิคในการลดเนื้อที่ในการจัดเก็บรูปภาพ หรือลด Bandwidth ในการส่งข้อมูลภาพผ่านสื่อกลางต่างๆ
- Morphological processing เป็นขบวนการในการประมวลภาพตามรูปร่างของวัตถุในภาพ ส่วนมากเป็นการเปลี่ยนรูปร่างของวัตถุ
- Image segmentation เป็นกระบวนการในการแยกแยะส่วนประกอบของรูปภาพออกเป็นส่วนต่างๆเพื่อประโยชน์ในการวิเคราะห์ภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Image representation and description เป็นการอธิบายรูปภาพหรือแทนข้อมูลภาพด้วยข้อมูลที่สามารถสื่อความหมายได้
- Image recognition เป็นกระบวนการในการจดจำรูปภาพซึ่งเกี่ยวข้องกับความรู้ขั้นสูงในการตีความหมายของรูปภาพ

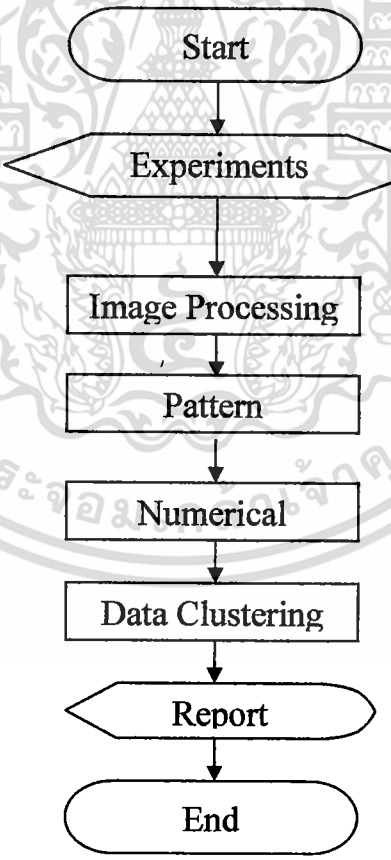
การประมวลผลภาพดิจิตอล หรือ Digital image processing เกี่ยวกับการแปลงข้อมูลภาพให้อยู่ในรูปแบบข้อมูลดิจิตอล (Digital format) ซึ่งสามารถที่จะนำเอาข้อมูลนี้จัดการผ่านกระบวนการต่างๆ ด้วยโปรแกรมคอมพิวเตอร์ได้ ภาพดิจิตอลเป็นภาพที่ประกอบด้วยจุดภาพเล็กๆจำนวนมากเรียกว่า พิกเซล (pixels) โดยใช้ตัวเลขจำนวนหนึ่งแทนค่าของระดับสีหรือระดับความสว่างของแต่ละพิกเซล ซึ่งาสามารถปรับแต่งเพื่อการแสดงผลภาพได้ ในแต่ละพิกเซลจะมีตัวเลขบอกระดับความสว่างหรือระดับสีบรรจุอยู่ ข้อดีของภาพดิจิตอลก็คือสามารถนำมาประมวลผลต่างๆด้วยคอมพิวเตอร์ได้ รายละเอียดของทฤษฎีจะไม่กล่าวในบทนี้ ผู้สนใจสามารถค้นคว้าข้อมูลได้จากตำราการประมวลผลภาพทั่วไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
114490

บทที่ 3
วิธีดำเนินการวิจัยและผลการวิจัย

3.1 แผนงานวิจัย

บทนี้กล่าวถึงศึกษาและทำการพัฒนาระบบวิเคราะห์ภาพถ่าย Agarose Gel จากผลการทดลองของ DNA Fragment ที่ได้ผ่านกระบวนการแยกสารประกอบด้วยไฟฟ้า (Electrophoresis) เพื่อเปรียบเทียบระดับความคล้ายคลึงกันของ DNA Fragment ทำการวิเคราะห์หามวลโมเลกุล และ แสดงผลการคำนวณออกมาในรูปแบบกราฟและแผนภูมิต่างๆ ดังแผนภาพในรูปที่ 3.1



รูปที่ 3.1 แผนภาพขั้นตอนการทำงานของระบบ Bioinformatics

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 ขั้นตอนการดำเนินงาน

สำหรับขั้นตอนในการดำเนินงานวิจัยจะมีทั้งหมด 6 ขั้นตอนดังที่ได้แสดงไว้แล้วในรูปที่ 3.1 ซึ่งจะมีรายละเอียดดังนี้คือ

1. เก็บผลการทดลอง โดยจะทำการเก็บภาพถ่ายผลการทดลอง ซึ่งได้จากกระบวนการ Electrophoresis โดยจะเก็บเป็นไฟล์ภาพเพื่อนำไปประมวลผลต่อไป
2. นำภาพที่ได้ มาผ่านกระบวนการ Image Processing ด้วยระบบทัศนจักร เพื่อให้ได้ภาพที่มีความชัดเจนมากยิ่งขึ้นและ ทำการหาขอบของภาพ เพื่อเตรียมภาพไปใช้สำหรับขั้นตอนต่อไป
3. นำภาพที่ได้มาผ่านกระบวนการ Pattern Recognition เพื่อให้ได้ ข้อมูลที่อยู่ในรูป Scoring Matrix Band Matching เพื่อนำไปประมวลผลต่อไป
4. นำ Scoring Matrix Band Matching ที่ได้มาผ่านกระบวนการ Numerical เพื่อที่จะให้ได้ Dissimilarity Matrix ของ DNA Fragment และ Molecular Weight สำหรับนำไปประมวลผลต่อไป
5. นำ Dissimilarity Matrix ที่ได้มาผ่านกระบวนการ Data Clustering เพื่อเปรียบเทียบความคล้ายคลึงกันของแต่ละ DNA Fragment สำหรับนำไปแสดงผลต่อไป
6. นำข้อมูลที่ได้มาสร้างเป็นแผนภูมิ Phylogenetic เพื่อแสดงผลการเปรียบเทียบความคล้ายคลึงกันของแต่ละ DNA Fragment และแสดงผลของ Molecular Weight ออกมาในรูปกราฟ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 ระเบียบวิจัย

1. เก็บผลการทดลอง

เป้าหมาย	เพื่อนำผลการทดลองของ Agarose Gel ที่ผ่านกระบวนการ Electrophoresis ไปใช้ในกระบวนการวิเคราะห์หาความคล้ายคลึงกันของ DNA Fragment และ Molecular Weight
ขั้นตอนการดำเนินงาน	1. เตรียม Agarose Gel ที่มีความเข้มข้นที่เหมาะสม 2. หยดสารละลายดีเอ็นเอลงในช่องบน Agarose Gel แล้วเปิดสนามไฟฟ้า 3. เมื่อได้เวลาที่เหมาะสม ปิดสนามไฟฟ้าแล้วย้อมดีเอ็นเอใน Agarose gel ด้วยสี Acridine 4. นำไปส่องในแสงอัลตราไวโอเลต Ethidium Bromide ที่ย้อมติดดีเอ็นเออยู่จะเรืองแสงสีส้ม 5. ทำการเก็บภาพถ่ายเพื่อนำไปใช้ในการประมวลผลต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Image Processing

เป้าหมาย	**1. เพื่อนำภาพที่ได้จากผลการทดลองมาปรับปรุงคุณภาพของภาพให้มีความคมชัดและมีรายละเอียดที่จำเป็นมากยิ่งขึ้น** **2. ทำการหาขอบของภาพ เพื่อนำไปใช้ในการสร้าง Matrix ของข้อมูลในขั้นตอนต่อไป**
ขั้นตอนการดำเนินงาน	1. Grayscale Converting เป็นขั้นตอนการเปลี่ยนภาพจากภาพสีเป็นภาพขาวดำ เพื่อให้สามารถนำมาประมวลผลได้ง่าย 2. Brightness and Contrast Adjustment เป็นขั้นตอนการปรับระดับความมืดสว่างและความเข้มแสงให้ภาพมีความคมชัดมากยิ่งขึ้น 3. Noise Reduction เป็นขั้นตอนการกำจัด Noise หรือสัญญาณรบกวนภายในภาพให้ภาพมีความคมชัดยิ่งขึ้น 4. Edge Detection เพื่อหาขอบของภาพสำหรับใช้ในการสร้าง Scoring Matrix Band Matching ในขั้นตอนต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. Pattern Recognition

เป้าหมาย	**เพื่อนำภาพที่ผ่าน Edge Detection มาทำการสร้าง Scoring Matrix Band Matching สำหรับการระบุตำแหน่งของ Band แต่ละ Band บน Agarose Gel ของ DNA Fragment สำหรับนำไปสร้าง Dissimilarity Matrix ต่อไป**
ขั้นตอนการดำเนินงาน	1. นำภาพที่ผ่าน Edge Detection มาแบ่งเป็น Block ตามจำนวน Lane เพื่อทำการประมวลผลในแต่ละ Lane 2. กำหนด Threshold สำหรับพิจารณาการมีอยู่ของ Band จากระดับค่าเฉลี่ยของสีในแต่ละ Block 3. วิเคราะห์ภาพ Front Dye เพื่อกำหนดตำแหน่งสุดท้ายของในแต่ละ Lane 4. ในแต่ละ Block ทำการหาตำแหน่งของ Band ที่อยู่ภายใน Block ว่าอยู่ตำแหน่งใดบ้าง โดยจะทำการประมวลผลแบบบนลงล่างแบบเส้นตรงจากบริเวณตรงกึ่งกลาง Block ซึ่งถ้าหากตำแหน่งใดพบ Pixel ของ Band ให้บันทึกลง Scoring Matrix Band Matching เป็นค่า 1 และตำแหน่งใดที่ไม่มีการพบ Pixel ของ Band ให้บันทึกเป็น 0 5. นำ Scoring Matrix Band Matching ที่ได้ไปใช้ในการประมวลผลต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. Numerical

เป้าหมาย	**เพื่อนำ Scoring Matrix Band Matching มาทำการหาระยะทางในแต่ละ Band เพื่อสร้าง Dissimilarity Matrix สำหรับแต่ละ DNA Fragment เพื่อใช้ในการประมวลผลต่อไป**
ขั้นตอนการดำเนินงาน	1. นำ Scoring Matrix Band Matching มาทำการระบุพิกัดในแต่ละ Column จะแทน Lane แต่ละ Lane ของ DNA Fragment 2. กำหนด Threshold เพื่อแบ่งเปอร์เซ็นต์การยอมรับได้ของตำแหน่ง 3. ทำการหาระยะห่างของบริเวณที่มีค่าเป็น 1 กับจุดอ้างอิงในแต่ละ Column แบบอัตราส่วนเพื่อทำการสร้าง Dissimilarity Matrix ของ DNA Fragment ในแต่ละ Lane

5. Data Clustering

เป้าหมาย	**เพื่อทำการเปรียบเทียบความคล้ายคลึงกันของ DNA Fragment Dissimilarity Matrix ของ DNA Fragment แต่ละอัน**
ขั้นตอนการดำเนินงาน	1. นำข้อมูลจาก Dissimilarity Matrix มาผ่านกระบวนการ Clustering Analysis เพื่อทำการเปรียบเทียบความคล้ายคลึงกันของ DNA Fragment แต่ละอัน 2. นำข้อมูลที่ได้จากกระบวนการ Clustering Analysis มาจัดลำดับความคล้ายคลึงกันของ DNA Fragment

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. Report

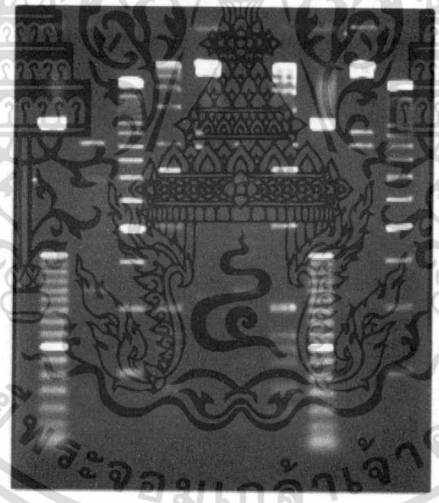
เป้าหมาย	**เพื่อนำผลลัพธ์ที่ได้จากการประมวลผลมาแสดงให้สามารถนำข้อมูลไปใช้ประโยชน์ได้**
ขั้นตอนการดำเนินงาน	1. ทำการแสดงผลภาพที่ผ่านกระประมวลผลในแต่ละขั้นตอน 2. ทำการแสดงผลระดับความคล้ายคลึงกันของ DNA Fragment ในแต่ละ Lane ด้วย แผนภูมิ Phylogenetic 3. ทำการแสดงผล Molecular Weight ด้วยกราฟ และตัวเลข 4. จัดทำรายงานวิจัยฉบับสมบูรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
อภิปรายผลการวิจัยและวิจารณ์

ในบทนี้จะกล่าวถึงการอธิบายผลการทดลองที่ได้จากเมื่อนำภาพ Agarose Gel ไปประมวลผลและวิเคราะห์ในคอมมพิวเตอร์เพื่อให้ได้ค่าที่ออกมาเป็นกราฟ เพื่อที่จะเก็บบันทึกลงในฐานข้อมูลได้

เมื่อทำการทดลองในส่วนของ Agarose Gel Electrophoresis เสร็จสมบูรณ์เรียบร้อยแล้วเราจะได้ภาพ Agarose Gel แสดงดังรูปที่ 4.1 เราจะเห็นว่าภาพ Agarose Gel ที่ได้จากการทดลองนั้นจะมีลักษณะเป็นเลนย่อยๆหลายเลน แล้วในแต่ละเลนจะประกอบไปด้วยแถบสีขาวหลายแถบ ภาพที่ได้นี้สามารถแยกดูว่าเป็นสิ่งมีชีวิดชนิดใดได้ ในขั้นตอนต่อไปเป็นการประมวลผลภาพ Agarose Gel

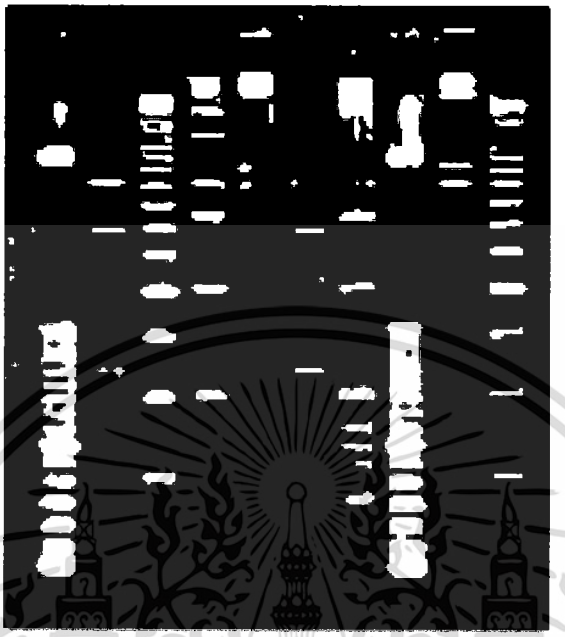


รูปที่ 4.1 Agarose Gel ที่ได้จากการทดลอง

ส่วนแรกในการประมวลผลภาพ Agarose Gel นั้นต้องผ่านขั้นตอนของการปรับแสงของภาพก่อน ในขั้นตอนนี้จะมีการปรับแสงของภาพเพื่อให้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดในภาพ Agarose Gel มีความคมชัดมากยิ่งขึ้น เพื่อที่จะได้แสดงผลค่าออกมาเป็นกราฟได้สำเร็จสมบูรณ์มากที่สุด

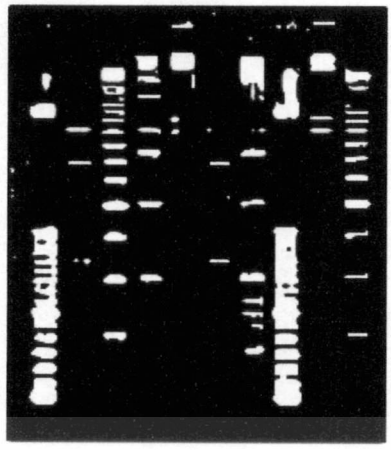


รูปที่ 4.2 Agarose Gel เมื่อผ่านกระบวนการปรับแสงเรียบร้อยแล้ว

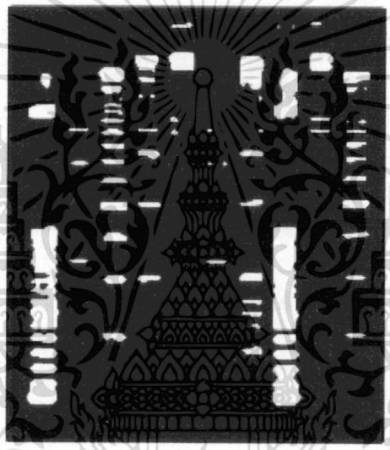
หลังจากผ่านขั้นตอนในการปรับแสงของภาพเรียบร้อยแล้ว ดังแสดงรูปที่ 4.2 จะนำภาพที่ได้มาทำการลดสัญญาณรบกวนของภาพเพื่อให้ได้รายละเอียดของภาพชัดเจนขึ้น

อภิปรายผลการวิจัยและวิจารณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



<ก>



<ข>

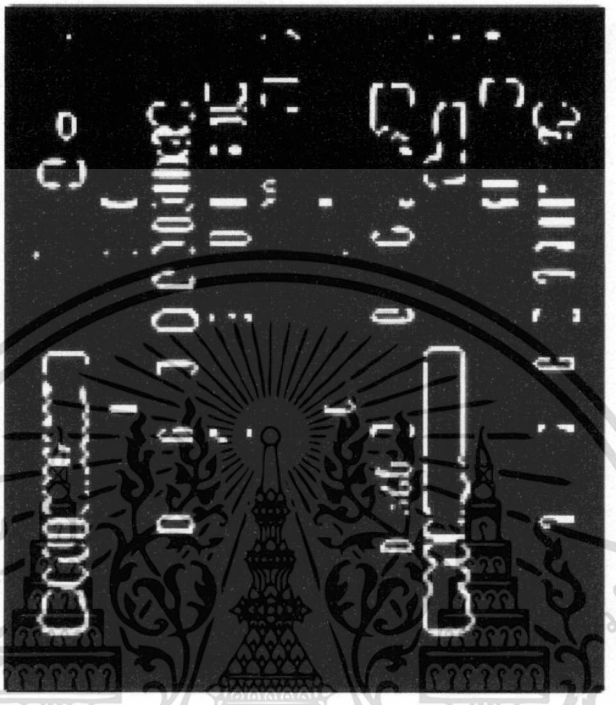
รูปที่ 4.3 Agarose Gel ก่อนและหลัง การผ่านกระบวนการลดสัญญาณรบกวน

จากภาพที่ 4.3 แสดงความแตกต่างก่อนและหลังของภาพ Agarose Gel ที่ผ่านกระบวนการลดสัญญาณรบกวนเรียบร้อยแล้ว ซึ่งในภาพที่ 4.3 <ก> จะเห็นว่าภาพมีสัญญาณรบกวนมากกว่าภาพที่ 4.3 <ข> ที่ผ่านกระบวนการลดค่าสัญญาณรบกวน ซึ่งภาพที่ 4.3 <ข> ได้จากการประมวลผลภาพเพื่อให้ได้ค่าที่สมบูรณ์มากยิ่งขึ้นได้มากกว่าภาพ 4.3< ก>

อภิปรายผลการวิจัยและวิจารณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนต่อไปหลังจากทำการลดสัญญาณรบกวนเรียบร้อยแล้วจะตรวจจับขอบของภาพที่ขาวเพื่อให้เป็นกล่องๆเพื่อง่ายต่อการสร้างกราฟซึ่งผลของการผ่านกระบวนการตรวจจับขอบแล้วจะมีผลดังแสดงรูปที่ 4.4 ดังนี้

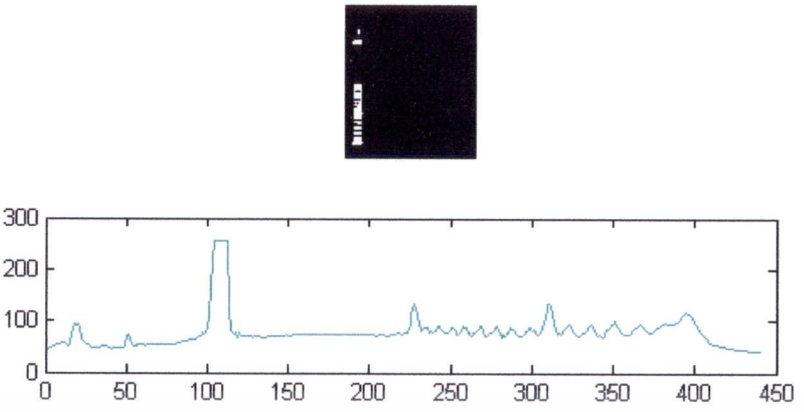


รูปที่ 4.4 Agarose Gel เมื่อผ่านการตรวจจับขอบเรียบร้อยแล้ว

เมื่อตรวจจับขอบของภาพ Agarose Gel เรียบร้อยแล้วขั้นต่อไปเป็นการแยกเลนต่างๆออกจากกันเพื่อให้เราสามารถสร้างกราฟได้ เมื่อเราแยกเลนแต่ละเลนได้แล้วโดยแสดงผลการวิเคราะห์เลนแต่ละเลนเป็นกราฟได้ ดังรูปที่ 4.5

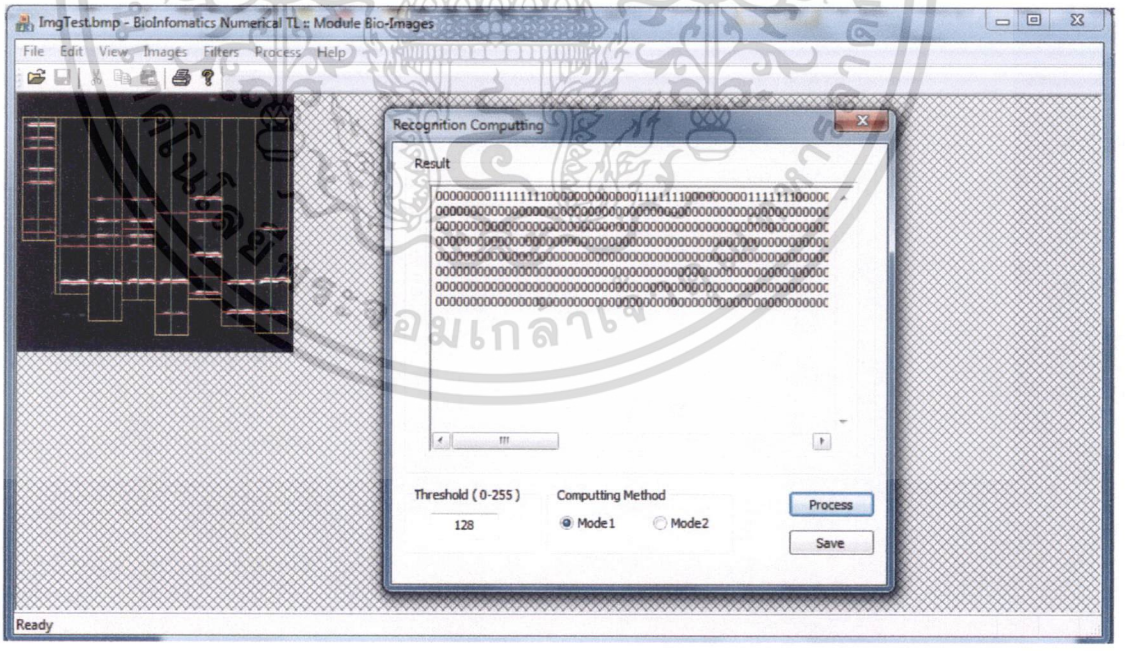
อภิปรายผลการวิจัยและวิจารณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5 ผลการวิเคราะห์ภาพที่ได้จากภาพ Agarose Gel ในเลนแรก

สำหรับการแสดงผลนั้นจะแสดงผลออกมาทางหน้าจอและ Save ออกมาเป็น ไฟล์นามสกุล .txt ทั้งที่เป็น แบบธรรมดา และ แบบที่เป็น Format ของ PHYLIP Module Restdist ซึ่งสามารถนำไปใช้กับ PHYLIP Module นี้ได้เพื่อการประมวลผลหา Distance Matrix ต่อไป ดังแสดงในรูปที่ 4.6 เทคนิคที่ใช้ในการพัฒนาซอฟท์แวร์นี้โปรดดูในภาคผนวก

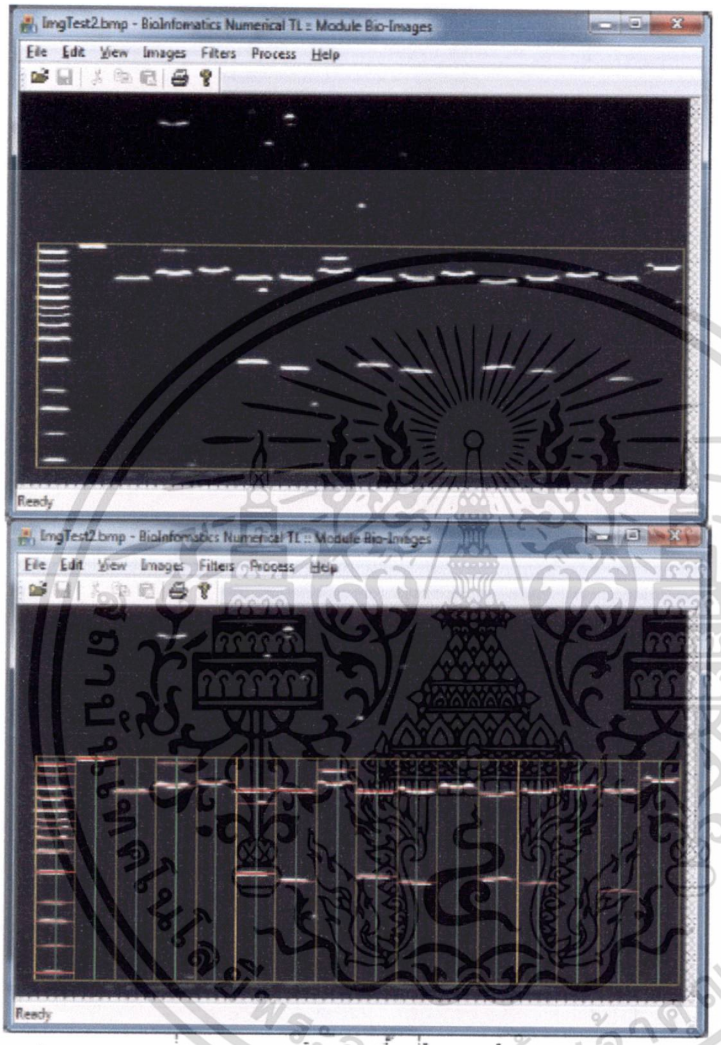


รูปที่4.6 โปรแกรม Bio-Images ที่ได้ทำการพัฒนาขึ้น

อภิปรายผลการวิจัยและวิจารณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองเปรียบเทียบอัตราส่วนที่พบยีนในเลนจริงเทียบกับการทดลองนำเสนอในรูปที่ 4.7 โดยผลความถูกต้องแสดงในตารางที่ 4.1 ซึ่งจะเห็นได้ว่าปรากฏผลสัมฤทธิ์ถูกต้องสมบูรณ์ตามที่ต้องการ



รูปที่4.7 ผลการทดลองเปรียบเทียบอัตราส่วนที่พบยีนในเลนจริงเทียบกับการทดลอง

อภิปรายผลการวิจัยและวิจารณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.1 ผลการทดลองยืนยันความถูกต้องของโปรแกรมที่พัฒนาขึ้น

No. Lane	อัตราส่วนการตรวจพบ Band ในแต่ละ Lane ต่อ Band ที่ปรากฏจริง					ค่าเฉลี่ย (%)
	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3	ครั้งที่ 4	ครั้งที่ 5	
1	14/14	14/14	14/14	14/14	14/14	100
2	1/1	1/1	1/1	1/1	1/1	100
3	1/1	1/1	1/1	1/1	1/1	100
4	2/2	2/2	2/2	2/2	2/2	100
5	1/1	1/1	1/1	1/1	1/1	100
6	3/3	3/3	3/3	3/3	3/3	100
7	2/2	2/2	2/2	2/2	2/2	100
8	2/2	2/2	2/2	2/2	2/2	100
9	2/2	2/2	2/2	2/2	2/2	100
10	2/2	2/2	2/2	2/2	2/2	100
11	1/1	1/1	1/1	1/1	1/1	100
12	2/2	2/2	2/2	2/2	2/2	100
13	2/2	2/2	2/2	2/2	2/2	100
14	1/1	1/1	1/1	1/1	1/1	100
15	2/2	2/2	2/2	2/2	2/2	100
16	1/1	1/1	1/1	1/1	1/1	100

อภิปรายผลการวิจัยและวิจารณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปและข้อเสนอแนะ

จากการอภิปรายผลและวิจารณ์ในบทที่ 4 ได้ข้อสรุปและข้อเสนอแนะดังนี้

1. งานวิจัยในเฟสนี้มีข้อจำกัดในการสร้างอุปกรณ์ทัศนจักรเนื่องจากงบประมาณที่จำกัด กล้องที่ใช้จึงเป็นเกรดทั่วไปไม่ใช้เกรดอุตสาหกรรม ซึ่งอาจเป็นข้อดีสำหรับการลดค่าใช้จ่ายในการศึกษา
2. ผลสัมฤทธิ์ที่ได้ของโปรแกรมคอมพิวเตอร์ในงานวิจัยนี้ ได้ผลตรงกับการวิเคราะห์ที่ถูกตรง สามารถนำไปใช้ในการวิเคราะห์ดี เอ็น เอได้แทนโปรแกรมนำเข้าจากต่างประเทศที่มีราคาแพง
3. การถ่ายรูปภาพของ Agarose Gel เพื่อให้ได้ภาพที่ดีทำได้โดยการควบคุมสภาพแวดล้อมและฉากพื้นหลังของตัวผลการทดลองในแผ่นทดลอง DNAให้มีความสมบูรณ์มากขึ้น กล่าวคือในการถ่ายรูปภาพของ Agarose Gel นั้นควรมีการควบคุมสภาพแวดล้อมและฉากพื้นหลังของภาพให้มีความสมบูรณ์มากขึ้นจะทำให้สามารถลดกระบวนการในการแยกแถบขาวออกจากฉากหลังได้ เนื่องจากไม่จำเป็นต้องทำการ Edge Detection เพราะการทำ Thesholding เพียงพอที่จะแยกส่วนของแถบขาวออกจาก Background ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] I.I. M˘andoiu and A. Zelikovsky, **Bioinformatics Algorithms Techniques and Applications**, John Wiley & Sons, Inc, Canada, 1st, 2008, ISBN 978-0-470-09773-1

[2] I. H. Witten and E. Frank, **Data Mining Practical Machine Learning Tools and Techniques**, Elsevier Inc, USA, 2nd, 2005, ISBN 978-0-12-088407-0

[3] J. Han and M. Kamber, **Data Mining Concepts and Techniques**, Elsevier Inc, USA, 2nd, 2006, ISBN 978-1-55860-901-3

[4] M. Sonka, V. Hlavac and R. Boyle, **Image Processing Analysis and Machine Vision**, Thomson Corporation, USA, 3rd, 2008, ISBN 978-0-495-24428-7

[5] N.A. Campbell, J.B. Reece and Others, **Biology**, Pearson Benjamin Cummings, USA, 8th, 2008, ISBN 978-0-321-53616-7

[6] R.C. Gonzalez, R.E. Woods and S.L. Eddins, **Digital Image Processing Using MATLAB**, Pearson Prentice Hall, USA, 2004, ISBN 0-13-008519-7

[7] A.W.C. Liew, H. Yan and M. Yang, "Pattern Recognition Techniques for the Emerging Field of Bioinformatics", **Science Direct Pattern Recognition**, vol38, Feb. 2005, pp. 2055 – 2073

[8] F. Camicia and Others, "Sequencing, Bioinformatic Characterization and Expression Pattern of a Putative Amino Acid Transporter from the Parasitic Cestode Echinococcus Granulosus", **Science Direct Gene**, vol411, Feb. 2008, pp. 1-9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

เทคนิคประมวลผลภาพสำหรับพัฒนาโปรแกรม

1. การประมวลผลภาพ (Images Processing)

สำหรับการประมวลผลภาพมีวัตถุประสงค์เพื่อเตรียมภาพที่ได้จากผลการทดลองให้มีความคมชัดและสามารถนำไปใช้ในการประมวลผลรู้จำรูปแบบได้แม่นยำยิ่งขึ้น ในการพัฒนา Software นั้น ได้เลือกใช้ Visual C++ และเรียกใช้ MFC Library ในการพัฒนา Software และได้นำ Algorithms ทางการประมวลผลภาพต่างๆมาพัฒนาและใช้งานในงานวิจัยดังนี้

1) Invert Images เป็นการกลับสีของรูปภาพให้มีสีตรงกันข้ามกับสีเดิม โดยมีความสัมพันธ์ดังนี้

$I_i[R,G,B] = M_i[(255-r),(255-g),(255-b)]$ ---------- (1)

เมื่อ I_i คือ pixel ที่ i ของ ภาพชนิด RGB ใดๆที่ถูกทำการ Invert แล้ว

M_i คือ pixel ที่ i ของ ภาพชนิด RGB ต้นแบบ

r ,R คือ ค่าของเฉดสีแดง ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ในช่วง [0,255]

g ,G คือ ค่าของเฉดสีเขียว ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ในช่วง [0,255]

b ,B คือ ค่าของเฉดสีน้ำเงิน ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ในช่วง [0,255]

2) Gray Scale เป็นการแปลงภาพที่เป็นชนิดสีแบบ RGB มาเป็นภาพโทนสีขาวดำโดยมีสมการความสัมพันธ์ดังนี้คือ

$G_i[x] = 0.299r_i + 0.587g_i + 0.114b_i$ ---------- (2)

เมื่อ G_i คือ pixel ที่ i ของ ภาพชนิด Gray Scale ใดๆ ในภาพ Gray Scale

x คือ ค่าของเฉดสีขาวดำ ใน pixel ที่ i ของ ภาพ Gray Scale มีค่าอยู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในช่วง [0,255]

r คือ ค่าของเฉดสีแดง ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ในช่วง [0,255]

g คือ ค่าของเฉดสีเขียว ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ในช่วง [0,255]

b คือ ค่าของเฉดสีน้ำเงิน ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ ในช่วง [0,255]

3) Gamma Correction Transformation เป็นกระบวนการแปลงภาพให้มีความเป็น Linear มากขึ้นเนื่องจากความไม่เป็น Linear ของการแสดงผลทางจอภาพแบบทั่วไปจึงต้องมีการชดเชยค่า Gamma ให้ภาพ โดยมีสมการความสัมพันธ์ดังนี้

$$I[R,G,B]_i = M_i[r^{1/\gamma}, g^{1/\gamma}, b^{1/\gamma}]_i \quad \text{----------(3)}$$

เมื่อ I_i คือ pixel ที่ i ของ ภาพชนิด RGB ใดๆที่ชดเชยค่า Gamma แล้ว

M_i คือ pixel ที่ i ของ ภาพชนิด RGB ต้นแบบ

r ,R คือ ค่าของเฉดสีแดง ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ ในช่วง [0,255]

g ,G คือ ค่าของเฉดสีเขียว ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ ในช่วง [0,255]

b ,B คือ ค่าของเฉดสีน้ำเงิน ใน pixel ที่ i ของ ภาพชนิด RGB มีค่าอยู่ ในช่วง [0,255]

γ คือ ค่า Gamma ที่นำมาปรับแก้ภาพ

4) Low Pass Filter เป็นวิธีการทำภาพให้มัวหรือเบลอขึ้นเพื่อลบ noise ขนาดเล็กๆ โดยจะใช้วิธีการ Convolution ซึ่งมี Mask Coefficient ดังนี้คือ

$$\begin{bmatrix} 1/16 & 1/8 & 1/16 \\ 1/8 & 1/4 & 1/8 \\ 1/16 & 1/8 & 1/16 \end{bmatrix}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) High Pass Filter เป็นการปรับภาพให้มีความคมชัดมากขึ้น แต่ถ้าหากมี noise ก็จะเป็นการขยาย noise ด้วย โดยจะใช้วิธีการ Convolution ซึ่งมี Mask Coefficient ดังนี้คือ

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 9 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

6) Sobel Edge เป็นการหา First Order Derivative ของภาพโดยมีสมการความสัมพันธ์คือ

$$\nabla F = \begin{bmatrix} H_r \\ H_c \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} f(x,y) \\ \frac{\partial}{\partial y} f(x,y) \end{bmatrix} \text{----------(4)}$$

โดย $\frac{\partial}{\partial x} f(x,y)$ หรือ H_r จะเป็นการหาขอบของภาพในแกน x และ $\frac{\partial}{\partial y} f(x,y)$ หรือ H_c จะเป็นการหาขอบของภาพในแกน y ถ้าหาทั้งสองทิศก็จะได้ Magnitude Vector โดยมีสมการดังนี้

$$\nabla f = \sqrt{H_r^2(x,y) + H_c^2(x,y)} \cong |H_r(x,y)| + |H_c(x,y)| \text{ ---------- (5)}$$

และทิศทางของ Vector หาได้จาก

$$\theta = \tan^{-1}\left(\frac{H_c(x,y)}{H_r(x,y)}\right) \text{ ---------- (6)}$$

ซึ่งเราจะทำการทำ Convolution เพื่อหา H_r และ H_c โดยใช้ Sobel Mask Coefficient ดังนี้

Sobel Mask Coefficient $H_r(x,y) = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} or \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$

Sobel Mask Coefficient $H_c(x,y) = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} or \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$

เมื่อได้ค่า H_r หรือ H_c แต่ละ pixel มาแล้วถ้าหากเราใช้เพียงค่า H_r เราก็จะได้ขอบของภาพในแนวแกน x และถ้าใช้เพียงค่า H_c เราก็จะได้ขอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของภาพในแนวแกน y แต่ถ้าหากนำทั้ง 2 ค่ามาบวกกันตามสมการที่ (5) เราก็จะได้ขอบของภาพทั้งในแนวแกน x และ y

7) Prewitt Edge จะเป็นการหาขอบของภาพเช่นเดียวกับวิธีของ Sobel แต่จะแตกต่างกันเพียงค่า Mask Coefficient ซึ่ง Prewitt Edge มีค่า Mask Coefficient ดังนี้คือ

Prewitt Mask Coefficient $H_r(x,y) = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} or \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$

Prewitt Mask Coefficient $H_c(x,y) = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} or \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$

8) Freichen Edeg จะเป็นการหาขอบของภาพเช่นเดียวกับวิธีของ Sobel และ Prewitt แต่จะแตกต่างกันเพียงค่า Mask Coefficient ซึ่ง Freichen Edge มีค่า Mask Coefficient ดังนี้คือ

Freichen Mask Coefficient

$$H_r(x,y) = \begin{bmatrix} -1 & -\sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & \sqrt{2} & 1 \end{bmatrix} or \begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix}$$

Freichen Mask Coefficient

$$H_c(x,y) = \begin{bmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{bmatrix} or \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{bmatrix}$$

9) Laplacian เป็นการหา Second Order Derivative ของภาพโดยมีสมการความสัมพันธ์ดังนี้

$$\nabla^2 F = \begin{bmatrix} \frac{\partial^2}{\partial x^2} f(x,y) \\ \frac{\partial^2}{\partial y^2} f(x,y) \end{bmatrix} \text{---------- (7)}$$

และขนาดของ Magnitude Vector คือ

$$\nabla^2 f = \frac{\partial^2}{\partial x^2} f(x,y) + \frac{\partial^2}{\partial y^2} f(x,y) \text{---------- (8)}$$

โดย Laplacian ใช้ Mask Coefficient ในการทำ Convolution ดังนี้คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\text{Laplacian Mask Coefficient} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} or \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

2. <u>การรู้จำรูปแบบ (Pattern Recognition)</u>

ในส่วนของการรู้จำรูปแบบจะเป็นการแปลงภาพให้เป็นตัวเลขเพื่อนำไปประมวลผลต่อ โดยนำภาพที่มีความคมชัด มาทำการประมวลผลต่อเพื่อให้ได้ Lane Matrix ของภาพทั้งภาพ หรือการทำ Scoring Matrix Band Matching โดยมีขั้นตอนดังนี้คือ

1) กำหนดขอบเขตของพื้นที่ๆจะใช้ในการประมวลผลหรือ Lane Area

เป็นการระบุขอบเขตของ Lane ทั้งหมดที่ต้องการใช้ในการประมวลผล ซึ่งโปรแกรมที่พัฒนาขึ้นมาสามารถกำหนดได้โดยการลากเมาส์ลงไปในบริเวณพื้นผิวของภาพ

2) กำหนดจำนวน Lane ที่มีทั้งหมดที่จะใช้ในการประมวลผล

จะเป็นการระบุจำนวน Lane ที่มีอยู่ใน Lane Area จาก 1) ซึ่งเมื่อกำหนดจำนวน Lane แล้วโปรแกรมก็จะทำการ Mask ขอบเขตของแต่ละ Lane ให้มีความกว้างเท่าๆกัน

3) ระบุขอบเขตสิ้นสุดของแต่ละ Lane และระบุตำแหน่งของเส้นรู้จำของ Lane

สำหรับการประมวลผล โดยการระบุขอบเขตสิ้นสุดของแต่ละ Lane เป็นการกำหนดจุดสิ้นสุดของการประมวลผลในแต่ละ Lane ซึ่งโปรแกรมอนุญาตให้มีตำแหน่งไม่เหมือนกันได้ และสามารถกำหนดตำแหน่งใหม่จากการคลิกเมาส์ในบริเวณขอบเขตความกว้างของ Lane นั้นๆ

สำหรับการระบุตำแหน่งเส้นรู้จำของ Lane จะเป็นเส้นที่ใช้ในการประมวลผลหลัก โดยโปรแกรมจะให้ความสำคัญในการอ่าน pixel ของแต่ละ Band ที่อยู่บนเส้นรู้จำก่อน pixel อื่นๆใน Lane ซึ่งกำหนดได้โดยการคลิกเมาส์เช่นเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4) กำหนด Threshold และ Mode ที่จะใช้ในการประมวลผล

เมื่อได้พื้นที่ในการประมวลผลในแต่ละ Lane แล้ว ต่อมาก็จะเป็นการกำหนดค่าเริ่มต้นต่างๆที่จำเป็นในการประมวลผลนั่นก็คือ Threshold และ Mode โดย Threshold จะใช้ในการแบ่ง pixel ที่อ่านเข้ามาได้เป็น 2 ค่าคือ 0 และ 1 ถ้า pixel ไหนมีค่ามากกว่า Threshold ก็จะให้มีค่าเป็น 1 ส่วน pixel ไหนมีค่าน้อยกว่า Threshold ก็จะให้มีค่าเป็น 0

ส่วน Mode จะมีให้เลือก 2 Mode โดย Mode แรกจะเป็นการอ่านค่าในแต่ละ Lane แล้วทำการแปลงเป็นค่า 0 และ 1 ตามปกติ ส่วน Mode ที่ 2 จะอ่านค่า pixel และแปลงค่า pixel ของ Band เป็น 1 เฉพาะตรงกลาง Band เท่านั้น

5) ประมวลผลและแสดงผลลัพธ์

หลังจากได้กำหนดค่าเริ่มต้นต่างๆจนเสร็จสมบูรณ์แล้วต่อมาก็จะทำการประมวลผลโดยผลลัพธ์ที่ต้องการคือ ชุด Binary Sequence Number ของแต่ละ Lane โดยมีสมการที่เกี่ยวข้องดังนี้คือ

$$f(x) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Im}(i,j)}{N} \quad \text{----------(9)}$$

เมื่อ $f(x)$ คือ ค่าเฉลี่ยของ pixel ณ พื้นที่ใดๆ มีค่าตั้งแต่ 0 - 255

$Im(i,j)$ คือ pixel ของภาพในพิกัด i,j ใดๆ มีค่าตั้งแต่ 0 - 255

m คือความกว้างของพื้นที่ๆใช้ในการหาค่าเฉลี่ย มีหน่วยเป็น pixel

n คือความยาวของพื้นที่ๆใช้ในการหาค่าเฉลี่ย มีหน่วยเป็น pixel

N คือจำนวน pixel ทั้งหมดในพื้นที่ๆถูกนำมาใช้ในการคำนวน

$$g(x) = \begin{cases} 1, if \rightarrow f(x) \geq Threshold \\ 0, if \rightarrow f(x) < Threshold \end{cases} \quad \text{----------(10)}$$

โดยที่ $g(x)$ เป็นค่า Binary Number ของ Band ที่ x โดยหาได้จากการนำ (9) มาหาค่าเฉลี่ยในพื้นที่ของ Band ตามที่ได้กำหนดไว้ในแต่ละ Lane ก่อนการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Process แล้วนำมาเปรียบเทียบกับ *Threshold* ที่ได้กำหนดไว้แล้ว ซึ่งใน Mode1 จะทำการหา Binary Sequence Number ของแต่ละ Lane อย่างตรงไปตรงมาแต่ใน Mode2 จะมีเพียงค่าที่อยู่ตรงกลาง Band เท่านั้นที่จะเป็น 1 โดยผลลัพธ์ที่ได้จะเป็นค่า 0 และ 1 เรียงต่อกันไปโดยใน 1พื้นที่ของ Band ก็จะแทนด้วนค่า 0, 1 เพียงค่าเดียว

นอกจากนี้ถ้าหากแต่ละ Lane มีจำนวน Band ไม่เท่ากันก็จะปรับให้เท่ากัน โดยให้ทุก Lane มีจำนวน Band เท่ากับจำนวน Band ที่มากที่สุดในระบบโดยการเทียบอัตราส่วนดังนี้

$$h(x) = g(x * \frac{U}{V}) \text{ ----------(11)}$$

เมื่อ $h(x)$ เป็นค่าของ Band ใหม่ที่ได้จากการปรับ Band เดิมให้มีจำนวน Band เท่ากับ Lane ที่มี Band ที่มากที่สุด

$g(x)$ เป็นค่าของ Band ของ Lane เก่า

x เป็นค่าตำแหน่งของ Band ของ Lane ใหม่

U เป็นจำนวน Band ของ Lane เก่า

V เป็นจำนวน Band ของ Lane ที่มีจำนวน Band มากที่สุด

หนังสือเป็นสมบัติของท่าน
โปรดช่วยกันรักษา
www.lib.kmitl.ac.th
สำนักหอสมุดกลาง โทร. 0 2739 2221

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้