

## บทปฏิบัติการที่ 8 การใช้บริการค้นหาและวิเคราะห์ข้อมูล ด้านพันธุวิศวกรรมโดยใช้เครือข่ายอินเทอร์เน็ต

### อินเทอร์เน็ต

อินเทอร์เน็ตคือเครือข่ายของเครื่องคอมพิวเตอร์ที่เชื่อมโยงกันทั่วโลก ซึ่งผู้ใช้เครื่องคอมพิวเตอร์สามารถติดต่อและแลกเปลี่ยนข้อมูลกับคอมพิวเตอร์เครื่องอื่นที่อยู่ในเครือข่าย โดยอาศัย software protocol ชื่อ Transmission Control /Internet protocol (TCP/IP) ซึ่งทำหน้าที่เสมือนเครื่องแปลงภาษา ที่จะปรับเปลี่ยนสัญญาณ (ภาษา) จากคอมพิวเตอร์แต่ละเครื่องให้เป็นภาษาเดียวกัน (universal language) เครื่องคอมพิวเตอร์ทุกเครื่องที่อยู่ในเครือข่ายจึงติดต่อสื่อสารกันได้ถึงแม้จะอยู่ห่างไกลกันมากเพียงไรก็ตาม เครือข่ายอินเทอร์เน็ตประดับด้วยโครงสร้างหลัก 3 ประเภท

1. Server program เป็นโปรแกรมที่ทำงานตลอดเวลาบนเครื่องคอมพิวเตอร์แม่น้ำย (host computer) เพื่อรับคำสั่งหรือข้อมูลจากคอมพิวเตอร์เครื่องอื่นๆ (client) ผ่านทางเครือข่ายอินเทอร์เน็ต
2. Client program คือโปรแกรมที่ถูกเรียกโดยใช้เครื่องคอมพิวเตอร์ในเครือข่าย (terminal) เพื่อติดต่อกับคอมพิวเตอร์แม่น้ำย
3. Standard communication protocol คือโปรแกรมที่ทำหน้าที่ป้อนลีกอกลงให้ client ทำการแลกเปลี่ยนข้อมูลกับ server ได้อย่างมีประสิทธิภาพ

### เครื่องมือหรือบริการที่ใช้ติดต่อกันบนเครือข่ายอินเทอร์เน็ต

1. Electronic mail (หรือ e-mail) นอกจากใช้ติดต่อสื่อสารระหว่างกันในรูปแบบของจดหมายแล้ว ยังสามารถใช้คัดลอกข้อมูล และเรียกใช้โปรแกรมของคอมพิวเตอร์อีกเครื่องหนึ่ง (หรือ server) ได้ ถ้ามีคำสั่งเฉพาะที่ถูกต้อง
2. File Transfer Protocol (หรือ FTP) ใช้ถ่ายโอนข้อมูลได้ทุกักษณะ ไม่ว่าจะเป็นภาพ ตัวหนังสือ เสียง หรือแม้แต่ software
3. Usenet คือเวทีหรือที่ชุมนุมของกลุ่มผู้ใช้อินเทอร์เน็ตที่รวมตัวกันบนเครือข่าย อินเทอร์เน็ตเพื่ออภิปรายแสดงความคิดเห็นหรือตอบข้อซักถามในเรื่องที่ก่อสัมสนใจ

อาทิเช่น เรื่องของกีฬา อาหาร ดนตรี หรือเรื่องราวทางวิทยาศาสตร์โดยติดต่อผ่านทาง e-mail

4. Telnet มักนิยมใช้ในกรณีที่ต้องการควบคุมเครื่องคอมพิวเตอร์จากระยะไกล (remote computer) เพื่อวิเคราะห์ข้อมูล คัดลอกข้อมูล หรือเรียกใช้โปรแกรมสำคัญที่มีอยู่เฉพาะในคอมพิวเตอร์นั้นๆ
5. Internet Relay Chat (IRC) มีลักษณะการให้บริการเพื่อแสดงความคิดเห็นเช่นเดียวกับ Usenet ต่างกันที่ทุกคนที่ใช้บริการ IRC จะต้องบันไดทันที ณ เวลาเดียวกัน เมื่อถูกน้ำเสียงร่วมการประชุมจริงๆ
6. Gopher ใช้สืบค้นข้อมูล ระบุที่ตั้งของ website ที่มีข้อมูลที่สนใจ และถ่ายโอนข้อมูลได้โดยใช้ระบบการทำงานแบบง่ายๆ แต่มีประสิทธิภาพสูง สามารถรับส่งข้อมูลทุกประเภท
7. World Wide Web (WWW) เป็นบริการที่ได้รับความนิยมสูงสุดในระบบเครือข่ายอินเตอร์เน็ต เพราะ WWW สามารถเชื่อมต่อเข้ากับแหล่งข้อมูลที่ให้บริการประเภทอื่น เช่น Telnet, Gopher, FTP ได้อย่างไม่มีขีดจำกัด เพียงแต่มีโปรแกรม web browser (หรือ browser) มาทำหน้าที่เชื่อมโยงและนำเสนอข้อมูลใน WWW เท่านั้น ก็สามารถติดต่อกันหาข้อมูลและทำการถ่ายโอนข้อมูลจากแหล่งเก็บข้อมูลต่างๆ ใน WWW (website) โดยตรงหรือจากแหล่งข้อมูลที่ให้บริการประเภทอื่นๆ ซึ่งเชื่อมต่อกันอยู่ โดยไม่ต้องอาศัยการพิมพ์คำสั่งเพื่อใช้งานเหมือนบริการอื่นที่กล่าวมาข้างต้น การใช้งาน WWW จึงเปรียบเสมือนการใช้งานอินเตอร์เน็ตที่รวมบริการแทบทุกประเภทเข้าไว้ด้วยกัน นอกจากนี้ web browser จะแสดงข้อมูลของ WWW ที่มีความสมบูรณ์แบบมาก ผู้ใช้จะมีความรู้สึกเหมือนกำลังอ่านหนังสือสารานุกรมที่มีทั้งตัวอักษร ภาพ และเสียงประกอบ

ปัจจุบันคอมพิวเตอร์และอินเตอร์เน็ตได้รับการพัฒนา จนมีขีดความสามารถในการประมวลผลสูงขึ้นอย่างมาก โดยคอมพิวเตอร์ส่วนบุคคล (PC) มีความสามารถพอกับชุดเบอร์คอมพิวเตอร์เมื่อ 10 กว่าปีก่อน นักวิทยาศาสตร์นิยมใช้คอมพิวเตอร์เพื่อเก็บข้อมูลและประมวลผลมากขึ้น การใช้คอมพิวเตอร์เพื่อสร้างฐานข้อมูลทำให้การเข้าถึงทำได้ง่ายและรวดเร็ว นอกจากคอมพิวเตอร์จะถูกนำมาใช้เพื่อการเก็บข้อมูล วิเคราะห์ข้อมูล และแสดงผลแล้ว ยังสามารถใช้คอมพิวเตอร์เพื่อขอรับรายละเอียดต่างๆ โดยกระบวนการทาง

คณิตศาสตร์ได้ด้วย ดังนั้นจึงเกิดศาสตร์อีกแขนงหนึ่งที่เรียกว่า Computational Biology และถูกพัฒนาจนเป็น Bioinformatics หรือ ชีวสารสนเทศ ในปัจจุบัน

ชีวสารสนเทศเป็นการนำสารสนเทศมาประยุกต์ใช้เพื่อจัดการกับข้อมูลทางชีวภาพในช่วงประมาณ 20 ปีที่ผ่านมา ฐานข้อมูลทางชีวภาพมีการขยายตัวอย่างต่อเนื่อง และรวดเร็วแบบอนุกรมเรขาคณิต ฐานข้อมูลเกี่ยวกับวรรณกรรมหรือเอกสารวิชาการก็มีการเติบโตด้วยอัตราใกล้เคียงกัน ดังนั้นจึงมีการใช้อินเทอร์เน็ตโดยเฉพาะ World Wide Web มาใช้ในการค้นหาข้อมูลเหล่านี้ ในสมัยแรกเริ่มชีวสารสนเทศจัดเป็นศาสตร์ทางด้านชีวภาพ โดยเกี่ยวข้องกับการพัฒนาอัลกอริธึม (algorithms) เพื่อพยายามจะตอบคำถามต่างๆ ในทางชีวภาพ ผู้ที่ทำงานทางด้านชีวสารสนเทศ (Bioinformaticists หรือ Bioinformaticians) จึงเป็นผู้ที่พยายามพัฒนาเครื่องมือ (เช่น โปรแกรมต่างๆ) โดยคนๆ นั้นจำเป็นจะต้องมีความเข้าใจปัญหาทางชีวภาพและในขณะเดียวกันต้องสามารถหาคำตอบแก่คำถามดังกล่าวโดยผ่านกระบวนการทางคณิตศาสตร์ได้ งานวิจัยด้านชีวสารสนเทศจึงมีความเกี่ยวข้องกับความเข้าใจระบบต่างๆ ด้านชีวภาพ และแบบจำลองเชิงคณิตศาสตร์หรือการยกภาพ แล้วสามารถนำหั้งสองส่วนนี้มาสร้างอัลกอริธึมใหม่ๆ เพื่อใช้ในการวิเคราะห์ข้อมูล

ชีวสารสนเทศเป็นวิทยาศาสตร์สาขาหนึ่งที่เกี่ยวข้องกับการใช้ข้อมูลต่างๆ เพื่ออธิบายหรือเพื่อให้เกิดความเข้าใจสิ่งต่างๆ ทางชีวภาพ ชีวสารสนเทศเป็นส่วนหนึ่ง (subset) ของ computational biology ซึ่งเป็นเทคนิคที่ใช้ในการวิเคราะห์แบบจำลองต่างๆ ทางชีวภาพ โดยทำการวิเคราะห์ในเชิงปริมาณ นักวิจัยในสาขาวิชาชีวสารสนเทศมาจากหลาย ๆ แขนง เช่น วิทยาศาสตร์ชีวภาพสาขาต่างๆ คณิตศาสตร์ วิทยาการคอมพิวเตอร์ เป็นต้น การวิจัยด้านชีวสารสนเทศมีหลายส่วน นับตั้งแต่การค้นหาข้อมูล หรือซอฟแวร์บนอินเทอร์เน็ต การใช้ฐานข้อมูลทางชีวภาพที่มีให้บริการสาธารณะ (Public biological database) ไปจนถึงการนำเครือข่ายอินเทอร์เน็ตมาประยุกต์ใช้กับงานทางอณูชีววิทยา

### การนำเครือข่ายอินเทอร์เน็ตมาประยุกต์ใช้กับงานทางอณูชีววิทยา

ในช่วงไม่กี่ปีที่ผ่านมาวิทยาการด้านอณูชีววิทยา (Molecular Biology) มีการพัฒนาและเจริญก้าวหน้าอย่างรวดเร็ว ทำให้มีการศึกษาทดลองและตีพิมพ์เผยแพร่องค์วิจัยใหม่ๆ เป็นจำนวนมาก และมีแนวโน้มที่จะเพิ่มมากขึ้นเป็นทวีคูณในอนาคต เหตุ

นี้อาจก่อให้เกิดปัญหาภัยนักวิทยาศาสตร์หรือผู้สนใจที่ต้องการอ้างอิงหรือนำข้อมูลดังกล่าวมาประยุกต์ใช้ในงานวิจัย เพื่อแก้ไขปัญหาเหล่านี้จึงเริ่มมีการใช้วิทยาการคอมพิวเตอร์มาจัดการรวบรวมข้อมูลด้านอนุชีววิทยาและสาขาวิชาที่เกี่ยวข้องมาจัดเก็บในฐานข้อมูล (Database) อย่างเป็นระบบ เช่น ฐานข้อมูล GenBank (เป็นฐานข้อมูลของ National Center for Biotechnology Information หรือ NCBI; สหรัฐอเมริกา), EMBL (European Molecular Biology Laboratory; อังกฤษ), SWISS-PROT (สวิสเซอร์แลนด์), PIR (สหรัฐอเมริกา), DDBJ (DNA Data Bank of Japan; ญี่ปุ่น) เป็นต้น นอกจากนี้ยังมีการวิจัยและพัฒนาโปรแกรมคอมพิวเตอร์สำหรับใช้ศึกษา ค้นหาและวิเคราะห์ข้อมูลด้านอนุชีววิทยาเหล่านี้มาโดยตลอด เพื่อให้มีศักยภาพสูงเพียงพอที่จะรองรับการวิเคราะห์ข้อมูลทางอนุชีววิทยาได้ ซึ่งบางโปรแกรมสามารถคัดลอกระหว่างคอมพิวเตอร์ผ่านทางเครือข่ายอินเทอร์เน็ต โดยใช้คำสั่ง FTP ผ่านอินเทอร์เน็ตเข้ามายัง FTP server เพื่อโอนหรือคัดลอกไฟล์ข้อมูล (รวมถึง software package) ที่บรรจุอยู่ใน FTP server ไปเก็บไว้ในคอมพิวเตอร์ของตนเองได้ เมื่อได้รับโปรแกรมที่ต้องการแล้วก็สามารถทดลองใช้โปรแกรมนั้น และถ้าพอใจในประสิทธิภาพก็สามารถติดต่อซื้อโปรแกรมนิดเดียวรูปแบบมาใช้วิเคราะห์ในห้องปฏิบัติการได้ (shareware) แต่ก็มีโปรแกรมบางโปรแกรมที่มีประสิทธิภาพสูงและอนุญาตให้ทำการถ่ายโอนมาใช้โดยไม่เสียค่าใช้จ่าย (freeware) เช่น ReadSeq, ClustalW, ClustalX, Phylip, Oliso, GeneDoc, PipMaker (เปรียบเทียบดีเอ็นเอที่มีความยาวถึง 2 เมกะเบส (megabase) มีบริการใน TIGR), TeXshade (ใช้เนื้อหาที่ได้จากการเปรียบเทียบลำดับนิวคลีโอไทด์หรือกรดอะมิโนให้อยู่ในสภาพที่เหมาะสมสำหรับการพิมพ์เผยแพร่ <http://homepages.uni-tuebingen.de/beitz/tse.html>), XGobi [ใช้สำหรับวิเคราะห์แบบหลายมิติ (<http://www.research.att.com/areas/stat/xgobi>)] เป็นต้น

เนื่องจากปัจจุบันนี้วิทยาการด้านเทคโนโลยีสารสนเทศก็มีความเจริญรุ่งเรืองและได้รับความสนใจอย่างกว้างขวาง เครือข่ายอินเทอร์เน็ตจึงเริ่มมีบทบาทสำคัญกับงานทางอนุชีววิทยา เพราะจะเป็นการเปิดโอกาสให้กับนักวิจัยทั่วโลกมีโอกาสแลกเปลี่ยนความรู้ ความคิดเห็นกันผ่านทาง e-mail หรือ usenet ได้อย่างสะดวก รวดเร็วและไม่มีปัญหาอันเนื่องมาจากระยะทางไกลเข้ามามากมาย ด้วยเหตุนี้หน่วยงานที่ทำหน้าที่ดูแลฐานข้อมูลด้านอนุชีววิทยาจึงนำฐานข้อมูลที่มีอยู่แล้วในคอมพิวเตอร์มาบรรจุไว้ในระบบเครือข่ายอินเทอร์เน็ต (เรียกว่า PUBLIC DOMAIN) ซึ่งผู้ที่สนใจเข้ามูลได้ๆในฐานข้อมูลดัง

กล่าวสามารถเรียกมาอ่าน หรือทำการคัดลอก (download) ข้อมูลที่ต้องการมาเก็บไว้ในคอมพิวเตอร์ส่วนตัวเพื่อใช้ในงานวิจัยที่เกี่ยวข้องได้

นอกจากนี้ยังมีหน่วยงานบางแห่งที่ให้บริการวิเคราะห์ข้อมูลผ่านทางเครือข่ายอินเตอร์เน็ต โดยอนุญาตให้ผู้ใช้บริการติดต่อกับคอมพิวเตอร์ของหน่วยงานที่ทำหน้าที่เป็นผู้ให้บริการ และใช้ทรัพยากรที่ถูกบรรจุอยู่ใน public domain นั้น อันได้แก่ ฐานข้อมูล โปรแกรมคอมพิวเตอร์สำหรับวิเคราะห์ข้อมูล หรือข่าวสารงานวิจัยใหม่ๆ เป็นต้น ด้วยตัวอย่างของงานบริการวิเคราะห์ข้อมูลเหล่านี้ ได้แก่ การเปรียบเทียบความเหมือน (homology searching) ของลำดับนิวคลีโอไทด์หรือกรดอะมิโนที่สนใจกับลำดับนิวคลีโอไทด์หรือกรดอะมิโนที่มีอยู่ในฐานข้อมูล การค้นหาลำดับนิวคลีโอไทด์หรือกรดอะมิโนที่สนใจจากฐานข้อมูล (network sequence retrieval) การออกแบบไพรเมอร์สำหรับ PCR และการสร้างโครงร่าง 3 มิติของดีเอ็นเอหรือโปรตีน เป็นต้น

การใช้บริการวิเคราะห์ข้อมูลบนเครือข่ายอินเตอร์เน็ตเป็นบริการที่มีอยู่แล้วบนเครือข่ายอินเตอร์เน็ต โดยเฉพาะอย่างยิ่ง World Wide Web ใน การติดต่อกับ Server ที่ให้บริการสาธารณะเพื่อขอใช้บริการ ในที่นี้จะยกตัวอย่าง Server ที่เป็นที่รู้จักและนิยมใช้กันทั่วไป ในเบื้องต้นควรรู้พื้นฐานของการใช้อินเตอร์เน็ตเพื่อนำมาไปประยุกต์ใช้กับบริการวิเคราะห์ข้อมูลทางอณูชีววิทยาบนเครือข่ายอินเตอร์เน็ต : ผู้ใช้บริการ Web browser เกือบทุกชนิดสามารถใช้คำสั่ง cut, copy หรือ paste ได้เช่นเดียวกับโปรแกรม Text Editor อื่นๆไม่ว่าจะเป็น Notepad, WordPad หรือ Microsoft Word และยังสามารถใช้คำสั่ง cut, copy หรือ paste นี้คัดลอกข้อมูลข้ามไปมาระหว่าง web browser และ Text Editor ได้อีกด้วย

### 1. ฐานข้อมูลทางอณูชีววิทยา (Molecular and Genomic Database)

- ฐานข้อมูลเบื้องต้น (Primary sequence database)

เป็นแหล่งเก็บรวบรวมลำดับนิวคลีโอไทด์หรือลำดับกรดอะมิโนที่ได้มาจากห้องปฏิบัติการวิจัยโดยตรง และยังให้ข้อมูลประกอบอื่นๆ เช่น predicted translated sequence, putative sequence motif, promoter region เป็นต้น ซึ่งโดยส่วนใหญ่而言 สารทางวิทยาศาสตร์มักจะตีพิมพ์เฉพาะบทความที่อ้างถึงลำดับนิวคลีโอไทด์หรือลำดับกรดอะมิโนที่ปรากฏในฐานข้อมูลประเภทนี้แล้วเท่านั้น และมักจะอ้างถึงในรูปของ Accession number ซึ่งจำเพาะกับแต่ละลำดับนิวคลีโอไทด์หรือลำดับกรดอะมิโนนั้นๆ

## ตัวอย่างฐานข้อมูลชนิดนี้

|   |  |
|---|--|
| DDBJ (DNA Data Bank of Japan)                               | <a href="http://www.ddbj.nig.ac.jp">http://www.ddbj.nig.ac.jp</a> ,<br><a href="http://www.nig.ac.jp/home.html">http://www.nig.ac.jp/home.html</a>               |
| EMBL-EBI (EBI : European Bioinformatics Institute, England) | <a href="http://www.ebi.ac.uk/embl/">http://www.ebi.ac.uk/embl/</a> ,<br><a href="http://www.embl-heidelberg.de">http://www.embl-heidelberg.de</a>               |
| GenBank (NCBI)  | <a href="http://www.ncbi.nlm.nih.gov">http://www.ncbi.nlm.nih.gov</a>  |
| SWISS-PROT  | <a href="http://expasy.hcuge.ch/sprot-top.html">http://expasy.hcuge.ch/sprot-top.html</a><br><a href="http://www.expasy.ch/spdbv">http://www.expasy.ch/spdbv</a> |
| PIR (Protein Information Resource)                          | <a href="http://www-nbrf.georgetown.edu/nbrf/scan.html">http://www-nbrf.georgetown.edu/nbrf/scan.html</a>  |
| TIGR  | <a href="http://www.tigr.org">http://www.tigr.org</a>  |
| RASMOL  | <a href="http://www.bernstein-plus-sons.com">http://www.bernstein-plus-sons.com</a>  |

### - ฐานข้อมูลพิเศษ (Specialized database)

เก็บรวบรวมข้อมูลทางอณูชีววิทยาที่มีลักษณะจำเพาะหรือผ่านขั้นตอนการวิเคราะห์ข้อมูลมาบ้างแล้ว ยกตัวอย่างเช่น protein profile, การจำแนกชนิดของโปรตีน (classification of protein), genomic mapping, โครงร่าง 3 มิติของดีเอ็นเอหรือโปรตีน หรือแม้แต่รายละเอียดของเอนไซม์ตัดจำเพาะที่ประยุกต์ใช้ในงานด้านการตัดต่อพันธุกรรม, แหล่งข้อมูลศึกษา Proteomics, ฐานข้อมูลเกี่ยวกับ metabolic pathway เป็นต้น

## ตัวอย่างฐานข้อมูลชนิดนี้

|                         |   |
|-------------------------|---|
| BLOCK (Protein Profile) | <a href="http://blocks.fhcrc.org/">http://blocks.fhcrc.org/</a>   |
| Cn3D                    | <a href="http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3D.html">http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3D.html</a>       |
| GDB (Genome Database)   | <a href="http://www.gdb.org/">http://www.gdb.org/</a>   |
| ISREC                   | <a href="http://hist.isb-sib.ch/cgi-bin/PFSCAN">http://hist.isb-sib.ch/cgi-bin/PFSCAN</a>                                     |
| MMDB                    | <a href="http://www.ncbi.nlm.nih.gov/structure">http://www.ncbi.nlm.nih.gov/structure</a>                                     |
| PDB (Protein Data Bank) | <a href="http://www.pdb.bnl.gov/">http://www.pdb.bnl.gov/</a> , <a href="http://www.rcsb.org/pdb">http://www.rcsb.org/pdb</a> |
| Pfam                    | <a href="http://www.sanger.ac.uk/software/Pfam/">http://www.sanger.ac.uk/software/Pfam/</a>                                   |
| PROSITE                 | <a href="http://expasy.hcuge.ch/sprot/prosite.html">http://expasy.hcuge.ch/sprot/prosite.html</a>                             |
| RasMol                  | <a href="http://www.bernstein-plus-ons.com/software/rasmol/">http://www.bernstein-plus-ons.com/software/rasmol/</a>           |
| ReBASE                  | <a href="http://www.neb.com/rebase">http://www.neb.com/rebase</a>   |
| TREMBL                  | <a href="http://www.embl-heidelberg.de/srs5/">http://www.embl-heidelberg.de/srs5/</a>   |

### - ฐานข้อมูลทาง Proteomics

|        |   |
|--------|---|
| ExPASy | <a href="http://www.expasy.ch/tools">http://www.expasy.ch/tools</a>     |
| PROWL  | <a href="http://prowl.rockefeller.edu">http://prowl.rockefeller.edu</a> |

### - ฐานข้อมูลทาง Metabolic pathway

|      |   |
|------|---|
| KEGG | <a href="http://www.genome.ad.jp/kegg/">http://www.genome.ad.jp/kegg/</a> |
| WIT  | <a href="http://wit.mcs.anl.gov/WIT2/">http://wit.mcs.anl.gov/WIT2/</a>   |

### - ฐานข้อมูลบูรณาการ (Integrated database)

ถ้ามีข้อมูลของสำดับนิวคลีโอไทด์ สำดับกรดอะมิโนหรือส่วนที่เกี่ยวข้องอื่น ไม่ว่าจะเป็น keyword, accession number, field identifier ชื่อผู้ค้นพบ ตำแหน่งบนโครโมโซม ชนิดของสิ่งมีชีวิต หรือแม้แต่ชื่อบทความในวารสารวิทยาศาสตร์ Server ประเภทนี้จะทำการค้นหาในฐานข้อมูลและแสดงผลข้อมูลที่เข้ากัน (match) กับข้อมูลที่ถูกป้อนไว้แต่

แรก (query) บาง Server จะแสดงผลในรูปแบบที่สามารถเชื่อมโยงกับฐานข้อมูลที่เกี่ยวข้องกับ query, website ที่เกี่ยวข้องหรือแม้แต่เอกสารอ้างอิงชนิด MedLine ได้

ตัวอย่าง Server ที่มีบริการชนิดนี้

PubMed (MedLine) <http://www.ncbi.nlm.nih.gov/PubMed>

Entrez (MedLine) <http://www.ncbi.nlm.nih.gov/gquery/gquery.cgi>

## 2. การหาความเหมือนของสายดีเอ็นเอและโปรตีนSequence (Similarity Search Server)

ในการนี้ที่ต้องการทราบว่าลำดับนิวคลีโอไทด์หรือลำดับกรดอะมิโนที่ได้มาจากการปฏิบัติการมีโอกาสที่จะเป็นหรือมีความใกล้เคียงกับยืนหรือโปรตีนชนิดใด จะต้องเรียกใช้บริการ Server ประเภทนี้ Server จะนำเอาลำดับนิวคลีโอไทด์หรือลำดับกรดอะมิโนที่เราป้อนข้อมูลไปเป็นตัวเริ่มต้น (query) และนำไปเปรียบเทียบกับลำดับนิวคลีโอไทด์หรือลำดับกรดอะมิโนในฐานข้อมูล

ตัวอย่าง Server ที่มีบริการชนิดนี้

BLAST <http://www.ncbi.nlm.nih.gov/BLAST>

BLITZ [http://www.ebi.ac.uk/searches/blitz\\_inputhtml](http://www.ebi.ac.uk/searches/blitz_inputhtml)

FASTA <http://www.ddbj.nig.ac.jp/search/fasta-e.html>

MPsrch <http://www.ebi.ac.uk/MPsrch/>

TIGR <http://www.tigr.org>

### **3. การค้นหาชีนบนสายดีเอ็นเอ (Computational Gene Identification Server)**

Server ประเภทนี้จะวิเคราะห์ลำดับนิวคลีโอไทด์แล้วค้นหาบริเวณที่น่าจะสร้างโปรตีน (open reading frame หรือ coding region) ได้ โดยใช้ standard genetic code และ/หรือ alternative genetic code และแสดงผลเป็นรูปภาพ (ที่ประกอบด้วย start codon และ stop codon) และ deduced amino acid sequence

ตัวอย่าง Server ที่มีบริการชนิดนี้

|            |   |
|------------|---|
| ORF finder | <a href="http://www.ncbi.nlm.nih.gov/gorf/gorf.html">http://www.ncbi.nlm.nih.gov/gorf/gorf.html</a>   |
| GeneMark   | <a href="http://opal.biology.gatech.edu/GeneMark/">http://opal.biology.gatech.edu/GeneMark/</a>   |
| GeneScan   | <a href="http://bioweb.pasteur.fr/seqanal/interfaces/genscan-simple.html">http://bioweb.pasteur.fr/seqanal/interfaces/genscan-simple.html</a> |
| Glimmer    | <a href="http://www.tigr.org/software/glimmer/">http://www.tigr.org/software/glimmer/</a>   |
| Procrustes | <a href="http://www.-hto.usc.edu/software/procrustes">http://www.-hto.usc.edu/software/procrustes</a>   |

### **4. การค้นหาตำแหน่งตัดด้วยเอนไซม์เพื่อประโยชน์ในการตัดต่อยน (Restriction Enzyme Mapping Analysis)**

Restriction map หรือตำแหน่งบนชีนดีเอ็นเอที่จำเพาะต่อ restriction enzymes มีความสำคัญเพื่อใช้ในการเลือกตำแหน่งโคลนชีนดีเอ็นเอที่สนใจเป็นไปได้ง่ายขึ้น และยังนำไปประยุกต์ใช้กับเทคนิค RFLP ในการจำแนกสายพันธุ์ของสิ่งมีชีวิตได้อีกด้วย Server ที่สามารถเรียกใช้บริการนี้ได้ คือ

|                |   |
|----------------|---|
| WebCutter      | <a href="http://www.firstmarket.com/cutter/cut2.html">http://www.firstmarket.com/cutter/cut2.html</a>                           |
| WebGene        | <a href="http://www.itb.cnr.it/sun/webgene">http://www.itb.cnr.it/sun/webgene</a>   |
| WWWTacg cutter | <a href="http://www.cmbi.kun.nl/bioinf/tools/tacg/tacg2.form.html">http://www.cmbi.kun.nl/bioinf/tools/tacg/tacg2.form.html</a> |

## 5. การแปลรหัสดีเอ็นเอ (Translate a DNA Sequence into a Protein)

เป็นการแปลงลำดับนิวคลีโอไทด์เป็นลำดับกรดอะมิโน เพื่อใช้กำหนดของโปรตีนที่แสดงออกจากชิ้นดีเอ็นเอที่ทราบลำดับนิวคลีโอไทด์แล้วได้คร่าวๆ โดยเรียกใช้โปรแกรมจาก Server ต่อไปนี้

The Protein Machine <http://www.ebi.ac.uk/emboss/transeq/>

Protein Translation <http://expasy.org/tools/dna.html>

## 6. การเปรียบเทียบความเหมือนของกลุ่มสายดีเอ็นเอ (Sequence Alignment Server)

เป็นการเปรียบเทียบความเหมือนของกลุ่มสายดีเอ็นเอหรือสายโปรตีนที่มีอยู่ เพื่อประโยชน์ในการหา consensus หรือ conserved region ของสายดีเอ็นเอหรือสายโปรตีน

ตัวอย่าง Server ที่มีบริการชนิดนี้

|             |  |
|-------------|--|
| ClustalW1.6 | <a href="http://www.ebi.ac.uk/clustalw/">http://www.ebi.ac.uk/clustalw/</a>  |
| ClustalX    | <a href="ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/">ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/</a><br><a href="http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX/Top.html">http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX/Top.html</a> |
| DIALIGN     | <a href="http://bibiserv.techfak.uni-bielefeld.de/dialign/">http://bibiserv.techfak.uni-bielefeld.de/dialign/</a>  |
| Match-Box   | <a href="http://www.fundp.ac.be/sciences/biologie/bms/">http://www.fundp.ac.be/sciences/biologie/bms/</a><br>matchbox_submit.html  |
| MAP         | <a href="http://genome.cs.mtu.edu/map.html">http://genome.cs.mtu.edu/map.html</a>  |
| MAVID       | <a href="http://baboon.math.berkeley.edu/mavid">http://baboon.math.berkeley.edu/mavid</a>  |
| MSA         | <a href="http://xylian.igh.cnrs.fr/msa/msa.html">http://xylian.igh.cnrs.fr/msa/msa.html</a>  |
| MultAlin    | <a href="http://www.toulouse.inra.fr/multalin.html">http://www.toulouse.inra.fr/multalin.html</a>  |
| T-COFFEE    | <a href="http://www.ch.embnet.org/software/TCoffee.html">http://www.ch.embnet.org/software/TCoffee.html</a>  |
| MUSCA       | <a href="http://cbcdrv.watson.ibm.com/Tmsa.html">http://cbcdrv.watson.ibm.com/Tmsa.html</a>  |

## 7. การหาความสัมพันธ์ทางลำดับวิวัฒนาการ (Phylogenetic Analyses Server)

Server ในกลุ่มนี้จะเก็บรวบรวมข้อมูลเกี่ยวกับความหลากหลายทางชีวภาพ การจัดจำแนกสายพันธุ์สิ่งมีชีวิตและมีโปรแกรมคอมพิวเตอร์จำนวนมากสำหรับใช้ในการวิเคราะห์ข้อมูลเหล่านี้ โดยขึ้นอยู่กับลักษณะของข้อมูลติด วิธีวิเคราะห์ข้อมูล และรูปแบบของผลลัพธ์ โปรแกรมคอมพิวเตอร์ใน Server เหล่านี้จะนำลำดับนิวคลิโอล่าส์หรือลำดับกรดอะมิโนของสิ่งมีชีวิตมาเปรียบเทียบความเหมือน และจะแสดงผลออกมานเป็นความสัมพันธ์ทางลำดับวิวัฒนาการของสิ่งมีชีวิตแต่ละชนิด

|                    |   |
|--------------------|---|
| BIOBOX/Phyliп      | <a href="http://pore.csc.fi/molbio/progs/phyliп/">http://pore.csc.fi/molbio/progs/phyliп/</a>   |
| GeneBee            | <a href="http://www.genebee.msu.su/services/phtree_reduced.html">http://www.genebee.msu.su/services/phtree_reduced.html</a>                         |
| Phyliп program     | <a href="http://bioweb.pasteur.fr/seqanal/interfaces/phyliп-uk.html">http://bioweb.pasteur.fr/seqanal/interfaces/phyliп-uk.html</a>                 |
| Phylogeny Software | <a href="http://evolution.genetics.washington.edu/phyliп/software.serv.html">http://evolution.genetics.washington.edu/phyliп/software.serv.html</a> |
| TreeView           | <a href="http://taxonomy.zoology.gla.ac.uk/rod/treeview.html">http://taxonomy.zoology.gla.ac.uk/rod/treeview.html</a>                               |

## 8. การออกแบบไพรเมอร์ (Primer Design Server)

Server กลุ่มนี้จะเก็บรวบรวมข้อมูลเกี่ยบทั้งหมดที่เกี่ยวข้องกับปฏิกริยาลูกโซ่โพลีเมอเรสหรือ Polymerase Chain Reaction (PCR) เช่น PCR General Theory, PCR Protocols, PCR Enzymes, PCR References, PCR Software เป็นต้น

ตัวอย่างของ Server ในกลุ่มนี้ ได้แก่

|                      |   |
|----------------------|---|
| PCR Jump Station     | <a href="http://www.horizonpress.com/pcr/">http://www.horizonpress.com/pcr/</a>   |
| Bioguide for PCR     | <a href="http://bioinformatics.weizmann.ac.il/mb/bioguide/per/contents.html">http://bioinformatics.weizmann.ac.il/mb/bioguide/per/contents.html</a>   |
| Primer design server | <a href="http://bibiserv.techfak.uni-bielefeld.de/genefisher/">http://bibiserv.techfak.uni-bielefeld.de/genefisher/</a><br><a href="http://genome-www2.stanford.edu/cgi-bin/SGD/web-primer">http://genome-www2.stanford.edu/cgi-bin/SGD/web-primer</a><br><a href="http://www.ebi.ac.uk/dbases/primers/designing.html">http://www.ebi.ac.uk/dbases/primers/designing.html</a> |

|                        |   |
|------------------------|---|
|                        | <a href="http://www.wiliamstone.com/primers/nojavascript/">http://www.wiliamstone.com/primers/nojavascript/</a>                         |
|                        | <a href="http://alces.med.umn.edu/rawprimer.html">http://alces.med.umn.edu/rawprimer.html</a>   |
|                        | <a href="http://doprimer.interactive.de">http://doprimer.interactive.de</a>   |
|                        | <a href="http://www.blocks.fhcrc.org/codehop.html">http://www.blocks.fhcrc.org/codehop.html</a>   |
|                        | <a href="http://ihg.gsf.de/ihg./ExonPrimer.html">http://ihg.gsf.de/ihg./ExonPrimer.html</a>   |
|                        | <a href="http://labtools.stratagene.com">http://labtools.stratagene.com</a>   |
|                        | <a href="http://www.genome.wi.mit.edu/cgi-bin/primer/primer3www.cgi">http://www.genome.wi.mit.edu/cgi-bin/primer/primer3www.cgi</a>     |
| Primer analyses server | <a href="http://www.wiliamstone.com/primers/calculator/">http://www.wiliamstone.com/primers/calculator/</a>                             |
|                        | <a href="http://alces.med.umn.edu/rawtm.html">http://alces.med.umn.edu/rawtm.html</a>   |
|                        | <a href="http://www.med.upenn.edu/%7Ephemmatol/napcore/OligoCalc.html">http://www.med.upenn.edu/%7Ephemmatol/napcore/OligoCalc.html</a> |

## 9.บริการปรับเปลี่ยนข้อมูลดีเอ็นเอหรือโปรตีนให้อยู่ในรูปแบบที่เหมาะสมเพื่อนำไปใช้กับ software (Sequence Manipulation Tools Server)

9.1 เพื่อความเหมาะสมในการนำเสนอสายนิวคลีโอไทด์ที่ได้ไปนำเสนอหรือทำการวิเคราะห์ต่อไป อาจต้องใช้บริการจาก Server ต่อไปนี้ เพื่อกลับทิศ (orientation) ของสายนิวคลีโอไทด์หรือหาสายนิวคลีโอไทด์คู่สม (complementary strand) กับเส้นเดิม

|                           |   |
|---------------------------|---|
| Reverse Complement        | <a href="http://searchlauncher.bcm.tmc.edu/seq-util/Options/revcomp.html">http://searchlauncher.bcm.tmc.edu/seq-util/Options/revcomp.html</a>                     |
| Reverse, Complement       | <a href="http://twod.med.harvard.edu/labgc/adnan/projects/utilities/revcomp.html">http://twod.med.harvard.edu/labgc/adnan/projects/utilities/revcomp.html</a>     |
| or Reverse and Complement | <a href="http://www.bioinformatics.vg/bioinformatics_tools/reversecomplements.html">http://www.bioinformatics.vg/bioinformatics_tools/reversecomplements.html</a> |

### 9.2 Sequence Format Conversion Server

เนื่องจากมี software หรือบริการใน Server หลายชนิดที่ต้องการข้อมูลดิบ (ลำดับนิวคลีโอไทด์หรือลำดับกรดอะมิโน) ที่มีรูปแบบ (format) จำเพาะ รูปแบบเหล่านี้ได้แก่

- FASTA/Pearson

- GenBank
- GCG (Genetic Computer Group)
- EMBL
- NBRF (National Biomedical Research Foundation)
- Phylip (Phylogenetic Inference package)
- MSF (Multiple Sequence format)

ดังนั้นในการค้นหาข้อมูลในแต่ละ Server จะต้องใส่ข้อมูลให้ตรงกับรูปแบบที่ต้องการ ตัวอย่างเช่น

- BLAST, ORF finder ที่ NCBI server ต้องการข้อมูลในรูปแบบ FASTA format
- GeneFisher จะออกแบบไฟร์เมอร์ที่สายพันธุ์ไว้ในรูปแบบ FASTA, GenBank, EMBL format เท่านั้น
- Phylip software ต้องการข้อมูลในรูปแบบ Phylip format  
ดังนั้นก่อนที่จะเรียกใช้ software หรือบริการใน Server ต้องทำการเปลี่ยนแปลงรูปแบบข้อมูลดิบให้เหมาะสมก่อนด้วยโปรแกรม ReadSeq ซึ่งสามารถเรียกใช้ผ่าน Server ดังต่อไปนี้

<http://dot.imgen.bcm.tmc.edu/seq-util/Options/readseq.html>

<http://bimas.dcrt.nih.gov/molbio/readseq/>

<http://searchlauncher.bcm.tmc.edu/seq-util/Options/readseq.html>

การค้นหาหรือวิเคราะห์ข้อมูลบนเครือข่ายอินเตอร์เน็ตนั้น สามารถเข้าไปค้นหาใน Server ของ WWW Search Engine เช่น

|             |   |
|-------------|---|
| Alta Vista  | <a href="http://www.altavista.com/">http://www.altavista.com/</a> |
| Excite      | <a href="http://www.excite.com/">http://www.excite.com/</a>       |
| Infoseek    | <a href="http://www.infoseek.com/">http://www.infoseek.com/</a>   |
| Yahoo Robot | <a href="http://www.yahoo.com/">http://www.yahoo.com/</a>         |
| Google™     | <a href="http://www.google.com/">http://www.google.com/</a>       |
| Lycos Robot | <a href="http://hotbot.lycos.com/">http://hotbot.lycos.com/</a>   |

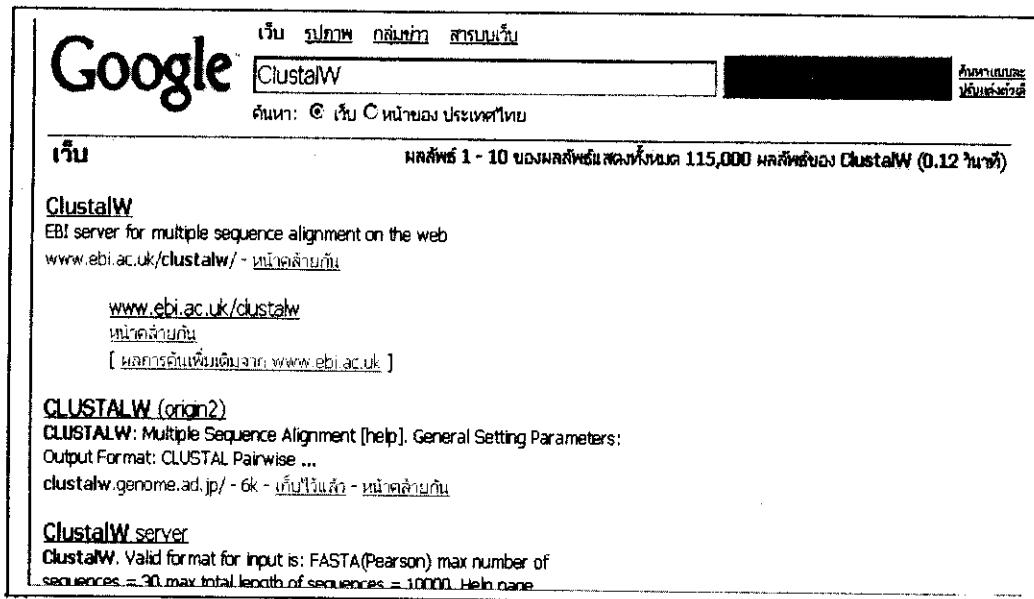
|              |   |
|--------------|---|
| Webcrawler   | <a href="http://web.webcrawler.com/">http://web.webcrawler.com/</a> |
| AOL Anywhere | <a href="http://search.aol.com/">http://search.aol.com/</a>         |

สำหรับการใช้บริการวิเคราะห์ข้อมูลของแต่ละ Server อาจมีลักษณะการใช้ การตั้งค่า parameter รูปแบบข้อมูลดิบที่จำเพาะแตกต่างกันไป ผู้ใช้บริการสามารถทราบรายละเอียดเหล่านี้ได้จาก "HELP" หรือ "More Information" ที่มักปรากฏขณะเรียกใช้บริการของแต่ละ Server อยู่แล้ว และเพื่อเสริมสร้างความเข้าใจในการประยุกต์ใช้เครื่องข่ายอินเตอร์เน็ตกับงานวิจัยทางด้านอณูชีววิทยา จึงขอยกตัวอย่างดังต่อไปนี้

**ตัวอย่างที่ 1** ถ้าต้องการค้นหา Server ที่มีบริการบนเครือข่ายอินเตอร์เน็ตเพื่อวิเคราะห์ข้อมูลลำดับเบสตีเอ็นเอ โดยการใช้ WWW Search Engine ของ Google™

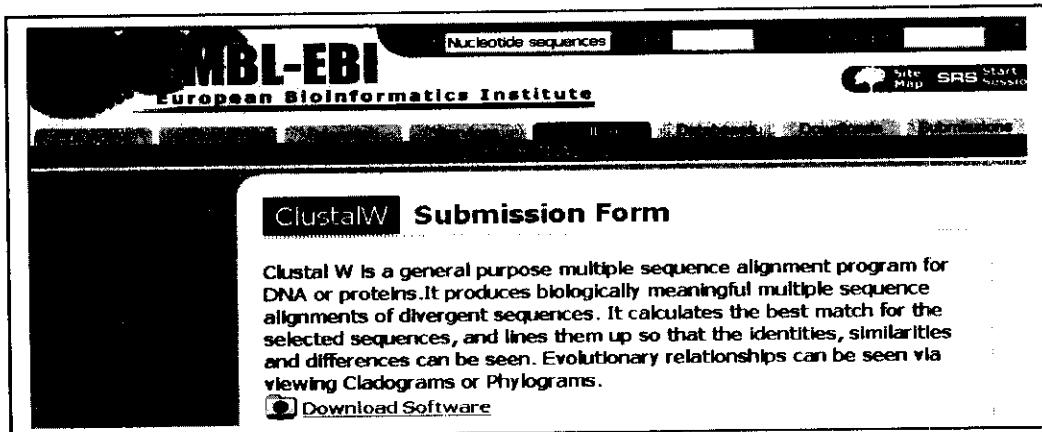
### วิธีทำ

1. เข้า Web Search ค้นหาโปรแกรมเราที่ต้องการ โดยใช้ Google™ ดังรูปที่ 1



รูปที่ 1

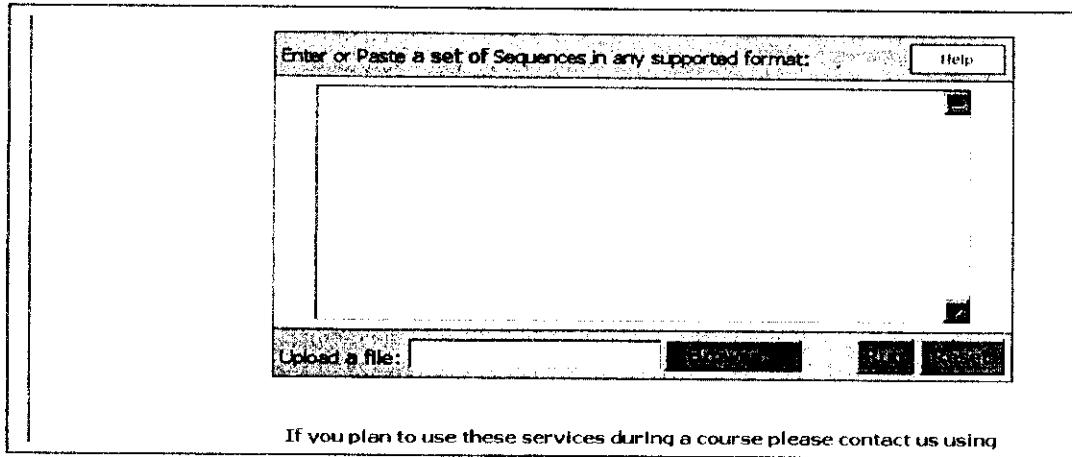
2. เข้าสู่ Web ที่ให้บริการด้านโปรแกรม ClustalW ดังรูปที่ 2, รูปที่ 3 และ รูปที่ 4  
ตามลำดับ



รูปที่ 2

This screenshot shows the "ClustalW Submission Form" interface. It features several rows of input fields. The first row contains five buttons: "Sequence" (highlighted in red), "Interactive", "full", and "single". The second row contains five buttons: "def", "def", "percent", "def", and "def". The third row contains five buttons: "def", "def", "def", "def", and "def". The bottom row contains five buttons: "aln w/numbers", "aligned", "none", "off", and "off". The background of the form is white, and the buttons are rectangular with black outlines.

รูปที่ 3



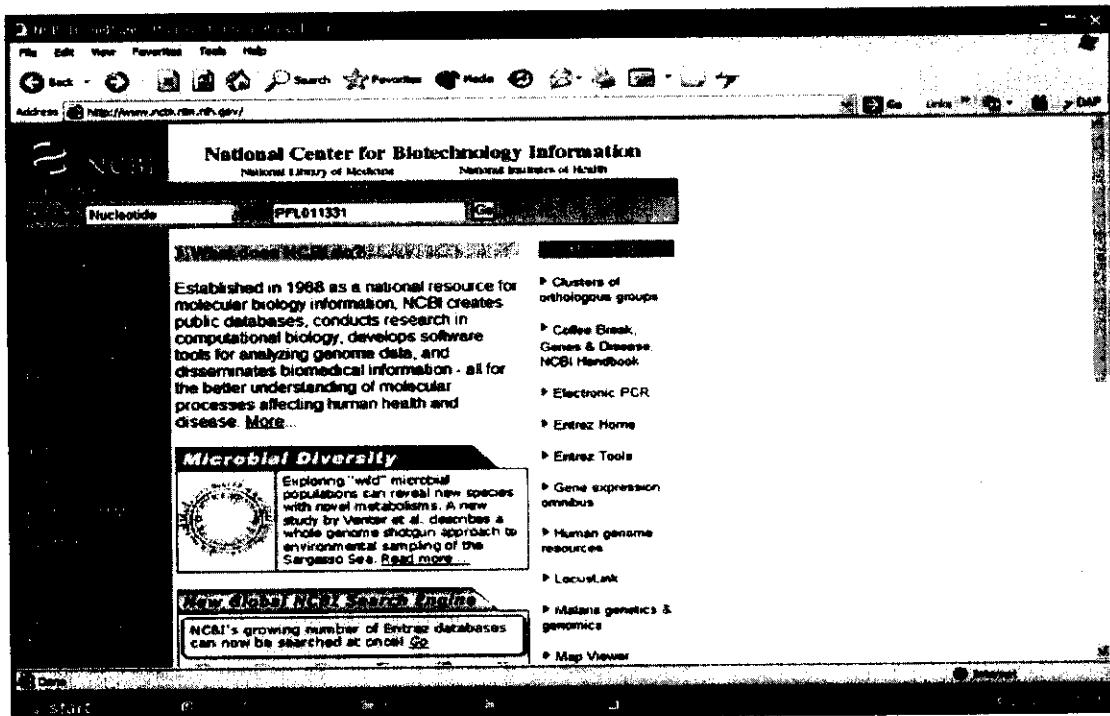
รูปที่ 4

3. ทำการใส่ข้อมูลลำดับเบสดีเอ็นเอที่ต้องการวิเคราะห์ พร้อมกับปรับค่า parameter ตามความเหมาะสม
4. กด run เพื่อทำการประมวลผลข้อมูล

ตัวอย่างที่ 2 ถ้าต้องการออกแบบ primers ที่จำเพาะกับยีน 16S rRNA ของเชื้อ *E. coli* สามารถทำได้ดังนี้

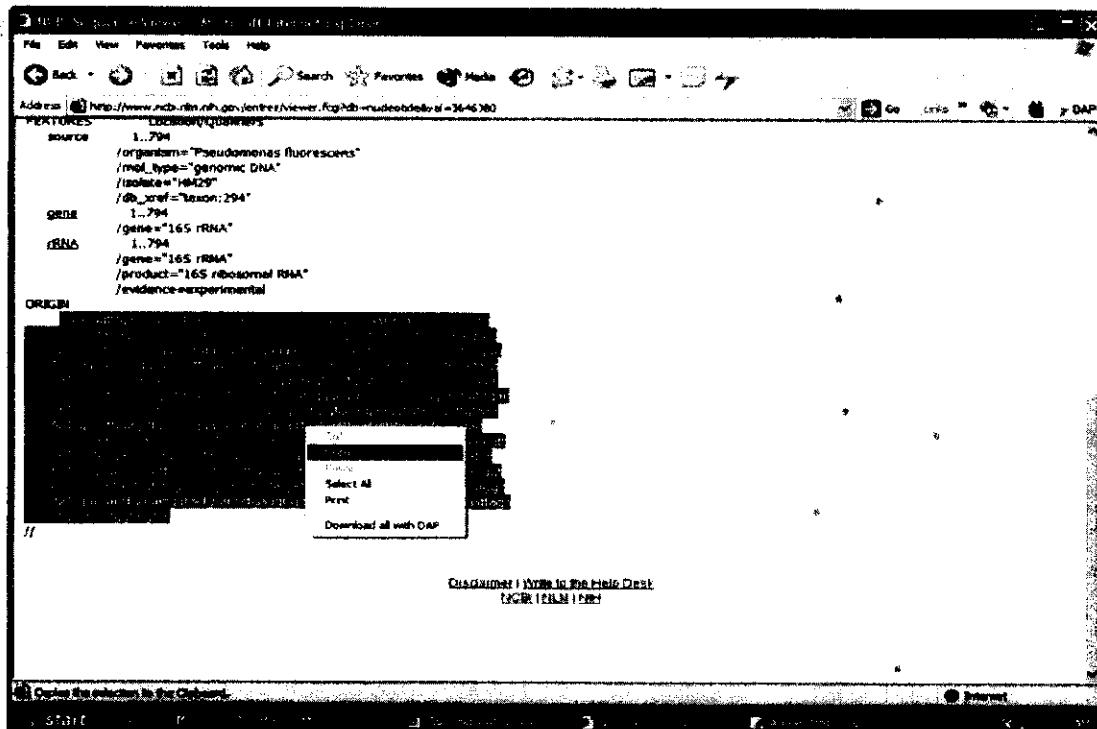
### วิธีทำ

1. เรียกใช้โปรแกรม microsoft word และ web browser
2. ใช้ web browser ติดต่อกับฐานข้อมูล ของ NCBI <http://www.ncbi.nlm.nih.gov> จากนั้นทำการพิมพ์ ชื่อเชื้อและช่วงของลำดับที่เราสนใจลงในช่องที่กำหนด จะดังรูปที่ 1



รูปที่ 1

3. จากนั้นคลิกที่ GO ผลลัพธ์ที่ได้จากการค้นหาจะได้ดังรูปที่ 2 หลังจากนั้นทำการคัดลอกช่วงของลำดับเบสที่ต้องการโดยทำแบบสีดำ (กดแล้วปุ่มซ้ายและลากเมาส์) ที่ลำดับเบส แล้วคลิกขวาเลือกคำสั่ง copy



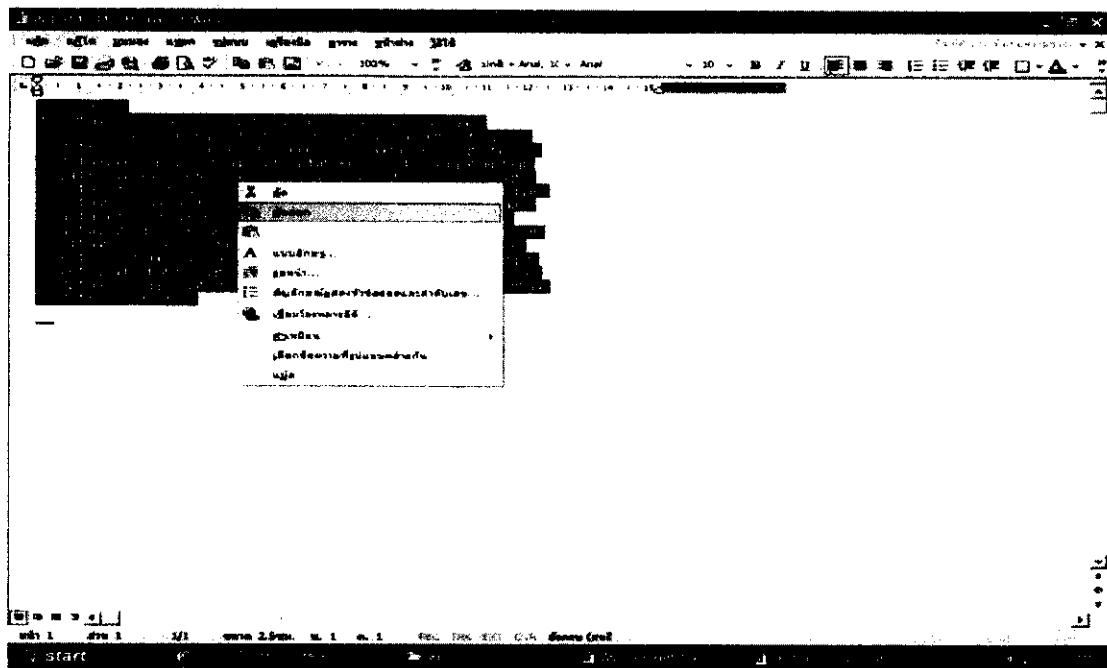
รูปที่ 2

4. เปิดโปรแกรม microsoft word แล้วทำการ edit past จะได้ดังรูปที่ 3 จากนั้นพิมพ์เครื่องหมาย “ > ” ตามด้วยชื่อหรือสัญลักษณ์ของลำดับเบสตั้งกล่าวที่บรรทัดแรก สุด

```
>PFL011331
1 tagetitga tcldgdtca galipeawgc tpeoggcagg otmaekat gaaetcpag
61 opda-yagag aegctygtc otchdngm capopapcgt ptegkamg cttaggeatc
121 tgcttgatg tgppgatc opdcopasas opdcopasas tacxgatc tleactcagg
181 agaaegcagg ppccctggg pccctggg ctacatngg cttaggkag atgatcg
241 tggttgatg atgatcg: aegcpagca tcctgatc gtdpaggg atgatcg
301 acatggatc tppakacgg tccagatc: tccgttggg pccctggg atgatcg
361 aetggatca aegctgatc aegctgatc ctgttggca gaaetttic gtttggaa
421 gacttgg tppaggatc opdcopasas ttaekatc tcttggta cttccopas
481 agaaengt opdcopasat ctgttggatc aegctgatc atcagggg atcagggg
541 atcagggg atcagggg aegctgatc aegctgatc ttaatggta tttggatcc
601 ccggddas cttaggeatc gatcataas cttaggeatc aegctgatc atcagggg
661 gatcttcc tgcttggg! gatcttcc gatcttcc gatcttcc tgcttggg
721 gatcttcc aegctgatc aegctgatc pccctggg ctggggatca ctggggatca
781 taatggta ttc
```

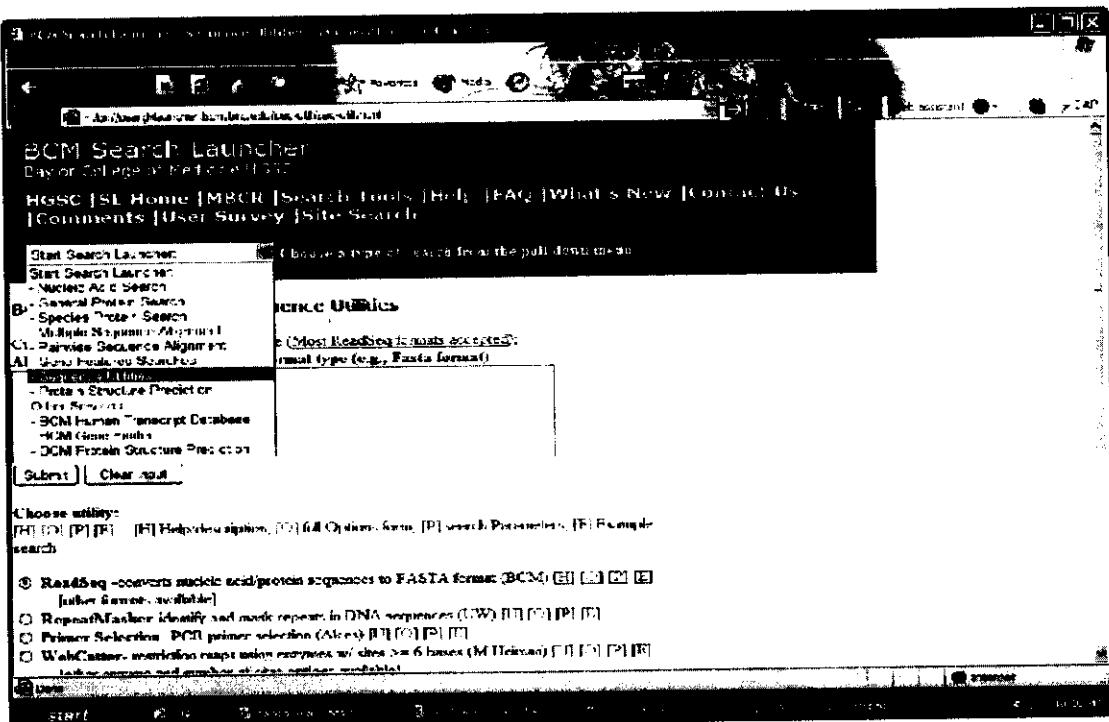
รูปที่ 3

5. ทำการคัดลอกข้อมูลทั้งหมดโดยใช้คำสั่ง edit → select all จากนั้นเลือกคำสั่ง edit → copy ดังแสดงในรูปที่ 4



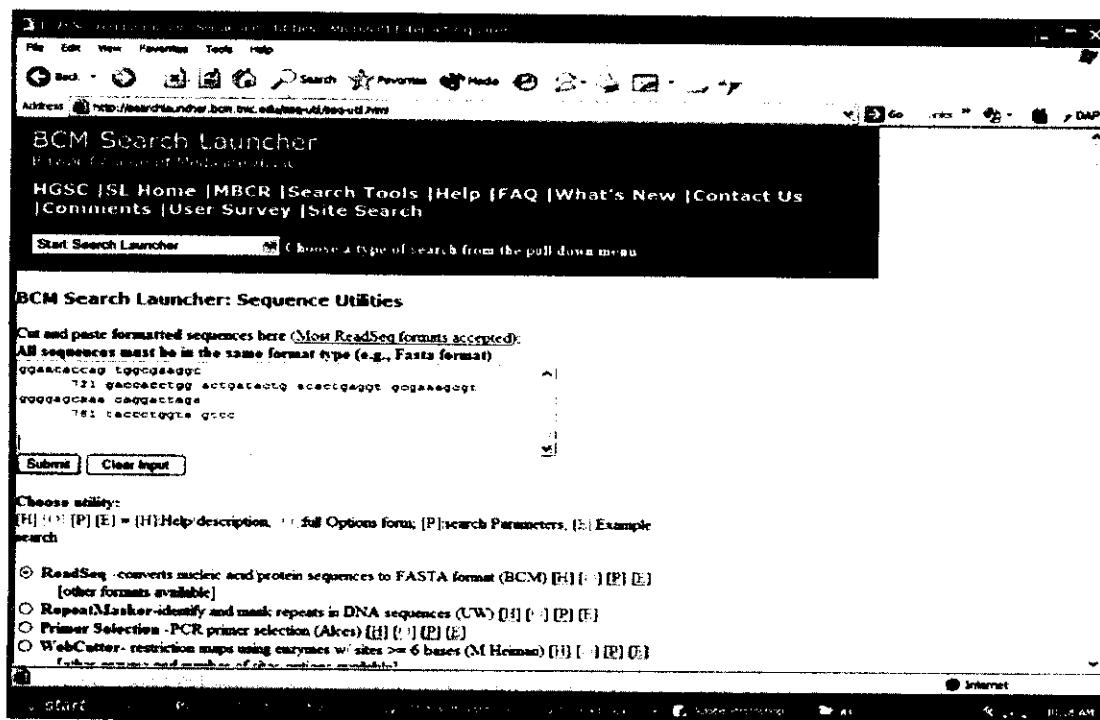
รูปที่ 4

6. เปิด web browser และไปยัง website <http://searchlauncher.bcm.tmc.edu/sequtil/Help/readseq.html> เพื่อทำการเปลี่ยนลำดับเบสให้อยู่ในรูป FASTA โดยใช้บริการ Read Sequence โดยการเลือกที่ sequence utility ดังรูปที่ 5



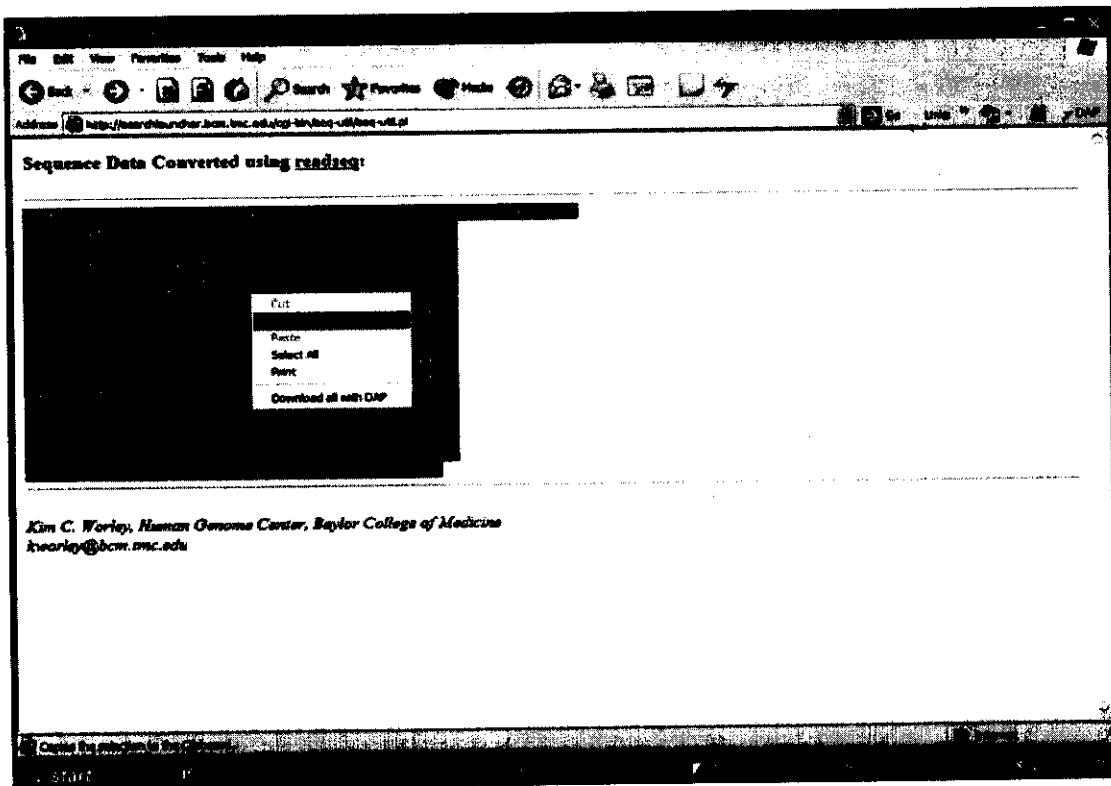
รูปที่ 5

7. จากนั้นใช้คำสั่ง paste ลำดับเบสลงในช่องว่างที่กำหนด และเลือกที่ choose utility เป็น ReadSeq และกดปุ่ม submit ดังรูปที่ 6



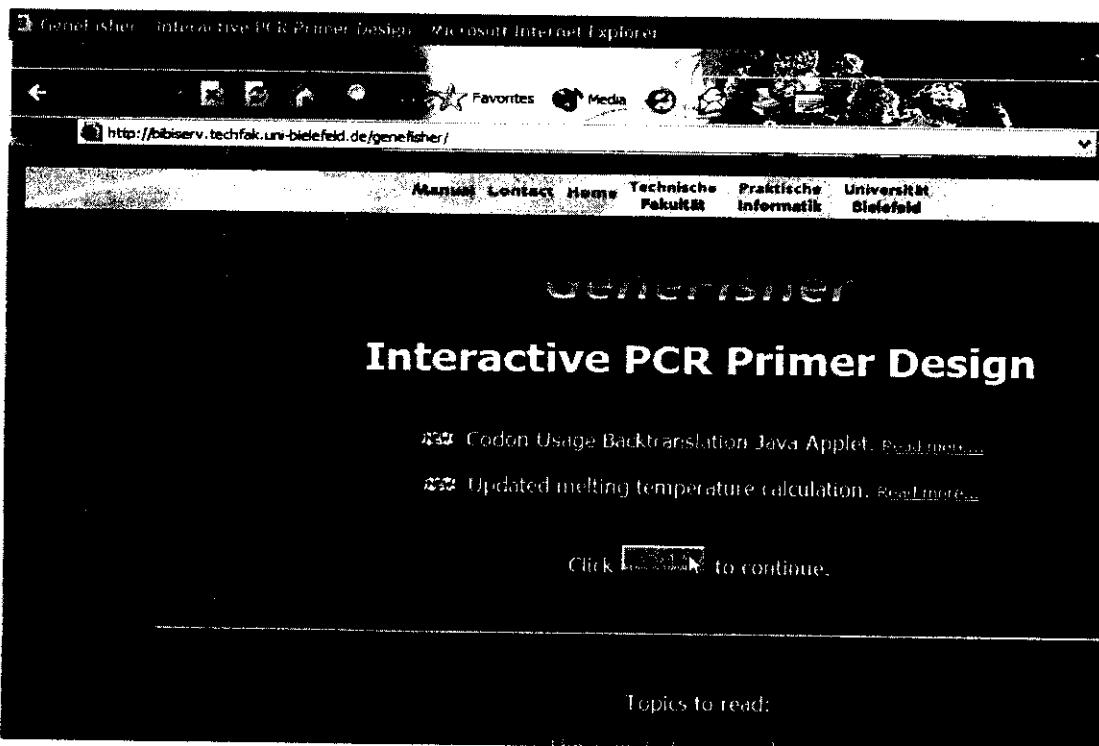
รูปที่ 6

8. ผลลัพธ์ที่ได้จากการเปลี่ยนรูปแบบเป็น FASTA แสดงในรูปที่ 7 หลังจากนั้นทำແກบสี คำที่สำคัญเบส แล้วคลิกขวาเลือกคำสั่ง copy



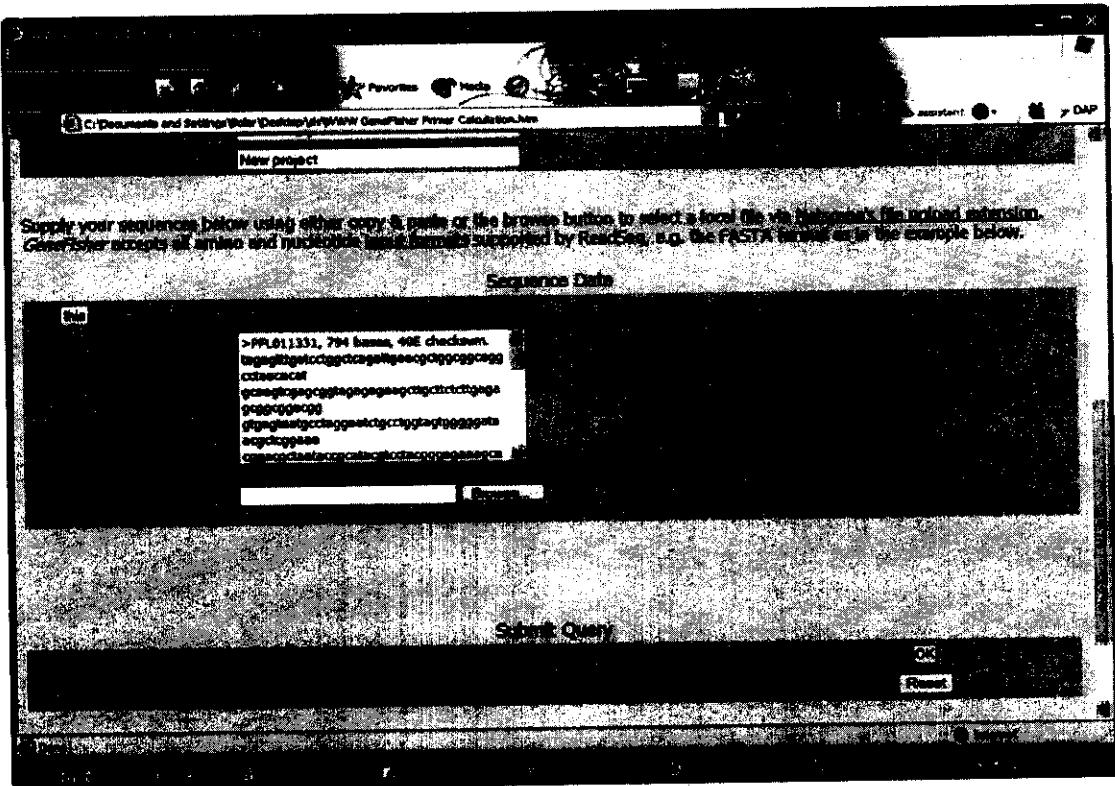
รูปที่ 7

9. เปิดใช้โปรแกรม website <http://bibiserv.techfak.uni-bielefeld.de/genefisher/> เพื่อทำการออกแบบ primers เมื่อเข้าสู่หน้าเมนูหลัก (main menu) ดังแสดงในรูปที่ 8 ให้คลิกที่ start เพื่อเข้าสู่หน้าต่อไป



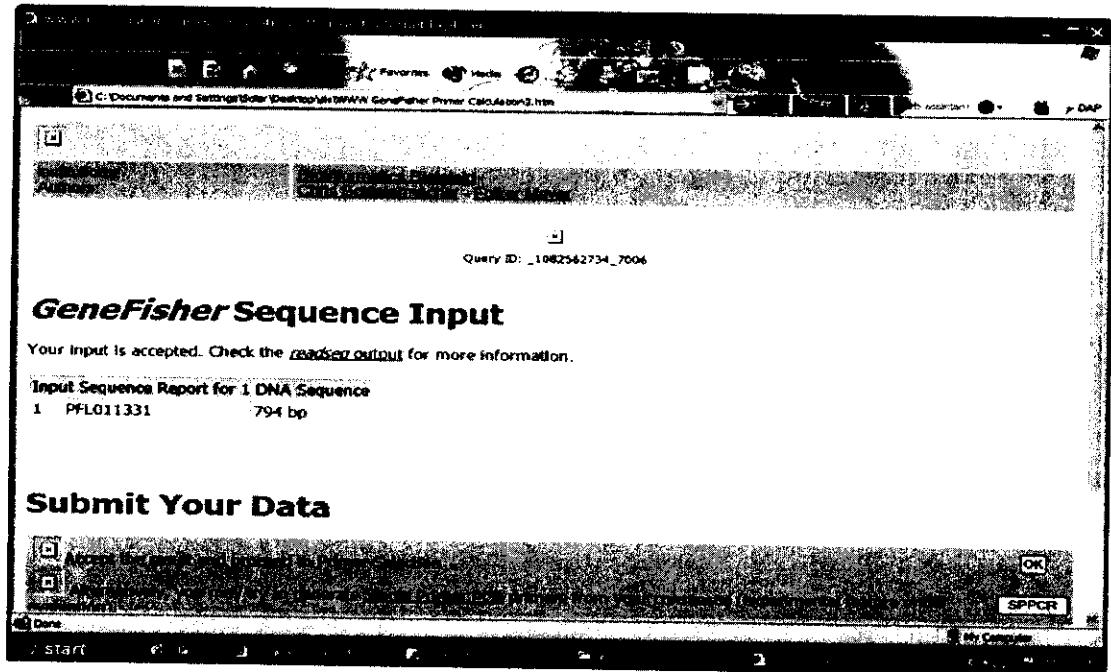
รูปที่ 8

10. จากนั้นใส่ข้อมูลประวัติของ user ได้แก่ email adress และสถานที่ทำงาน ลงในช่อง sequence data แล้วทำการคลิกขวา เลือกคำสั่ง past แล้วคลิกที่ GO เพื่อเข้าสู่หน้าต่อไป ดังรูปที่ 9



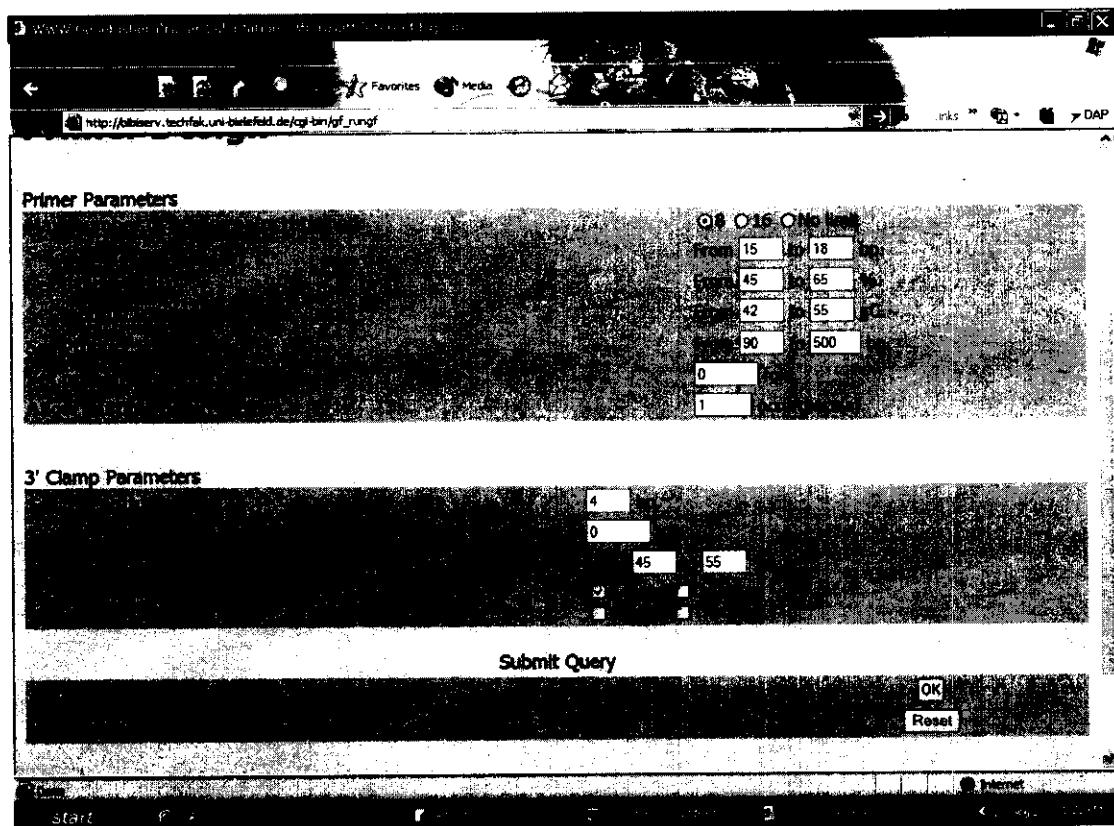
รูปที่ 9

11. หน้ากัดไปจะแสดงถึง sequence input ซึ่งสามารถที่จะดูข้อมูลได้โดยการคลิกที่ sequence out put และในหน้านี้ยังแสดงถึงชื่อของลำดับเบสและลำดับเบสทั้งหมดที่ใส่เข้าไปเพื่อออกรอบน primers ถ้าข้อมูลที่ใส่ถูกต้องแล้ว ให้ทำการคลิกที่ GO เพื่อสู่กระบวนการต่อไป ดังแสดงในรูปที่ 10



รูปที่ 10

12. เมื่อเข้าสู่หน้าต่อไปของการคำนวณ หน้านี้จะแสดงถึงค่า primer parameters ที่เหมาะสม ซึ่งสามารถที่จะตั้งค่าใหม่ได้ ถ้าหากว่าไม่พอใจกับค่าที่โปรแกรมได้ตั้งไว้ และถ้าหากว่าพอใจแล้ว ให้ทำการคลิกที่ OK ดังแสดงในรูปที่ 11 เพื่อให้โปรแกรมทำการคำนวณ และในหน้าตัดไปให้คลิกที่ results เพื่อดูผลที่ได้ โดยผลลัพธ์ที่ได้แสดงในรูปที่ 12



รูปที่ 11

13. ผลที่ได้จากการคำนวณจะแสดงผลออกมาดังรูปที่ 12 โดยสามารถที่จะดูว่า primers คู่ใดที่เหมาะสมที่จะนำไปใช้งาน ซึ่งควรจะพิจารณาดูคู่ที่มีค่า melting temperature  $\langle T_m \rangle$  และขนาดของ PCR products ที่มีความยาวต้องการ

| Primer Calculation Results         |                      |                       |       |            |                      |         |
|------------------------------------|----------------------|-----------------------|-------|------------|----------------------|---------|
| Query ID                           | 1062605544_14768     |                       |       |            |                      |         |
| User email                         | asasd7k@yahoo.com    |                       |       |            |                      |         |
| Project description                | New project          |                       |       |            |                      |         |
| <b>6 best Pairs (of max. 1615)</b> |                      |                       |       |            |                      |         |
| ID                                 | Forward Primer       | Reverse Primer        | Qual. | Prod. Len. | T <sub>m</sub> Diff. | PPProd. |
| 1                                  | GAGCTTAATGCCCTACCCAA | ACGCTTTGCCACCTCA      | 983   | 500        | 0                    | 100     |
| 2                                  | GGGGATATAACCGCTCGGAA | CCATAACCTCTAGGTCTGCA  | 983   | 499        | 2                    | 99      |
| 3                                  | CATAGCTCTAACCGGAGACA | TTCACCGCTTACACAGGAA   | 983   | 500        | 0                    | 99      |
| 4                                  | ATCTCTACGGGAGAAAAGCA | GCATTTTACACGCTCTACACA | 983   | 499        | 0                    | 99      |
| 5                                  | TGTTTACGGTAAATGGGCTC | TCGCGCTTCACTGTCA      | 982   | 498        | 0                    | 99      |
| 6                                  | ATACCGCATAGCTCCCA    | CACAGGAAATTCCACCA     | 982   | 497        | 0                    | 99      |
| 7                                  | CTCTCTACGGGAGAAAAGCA | TTCACCGCTTACACAGGAA   | 981   | 495        | 0                    | 99      |

รูปที่ 12

14. เมื่อทำการคุณภาพ forward primer และ/หรือ reward primer จะเห็นค่าของข้อมูลต่างๆ จากที่ได้จากการคำนวณของโปรแกรม เช่น ค่า Tm , product size ซึ่งข้อมูลเหล่านี้สามารถที่นำมาใช้ในการพิจารณาเลือกคู่ของ primers ที่เหมาะสมได้ เพื่อให้ง่ายต่อการวิเคราะห์ในกระบวนการต่อๆไปของงาน ดังแสดงในรูปที่ 13

WWW.Techfak.Uni-Bielefeld.De - Microsoft Internet Explorer

Favorites Media

[http://bioserv.techfak.uni-bielefeld.de/cgi-bin/fv\\_visualize?pid=108252869\\_19005&item=DAPP=13&NP=39](http://bioserv.techfak.uni-bielefeld.de/cgi-bin/fv_visualize?pid=108252869_19005&item=DAPP=13&NP=39)

Priming Sites

|                                |                       |
|--------------------------------|-----------------------|
| Pair                           | #1                    |
|                                |                       |
| Product Size                   | 481                   |
| T <sub>m</sub> Difference      | 1                     |
| Quality                        | 873                   |
| Pair Data                      |                       |
| Forward Primer Data            | Reverse Primer Data   |
| Sequence: GGC GGTTTG TGA AGTCA | TT CACA ACAC GAG CTGA |
| GC Content: 47                 | 47                    |
| Position: 7                    | 488                   |
| Degeneracy: 0                  | 0                     |
| 3' GC: 50                      | 50                    |
| 3' Degeneracy: 0               | 0                     |
| T <sub>m</sub> : 50.6115       | 51.7714               |
| Quality: 515                   | 515                   |
| Location: 7                    | 488                   |

Done start

จ 13

**ตัวอย่างที่ 3** ถ้าในห้องปฏิบัติการได้โคลนชีนดีเอ็นเอจากไวรัสซึ่งเป็นเชื้อที่สามารถก่อให้เกิดโรคอย่างหนึ่งได้ และจากการทดลองพบว่าชิ้นดีเอ็นเอนี้จะพบจำเพาะในเชื้อไวรัสชนิดนี้เท่านั้น เมื่อนำไปหาลำดับนิวคลีโอไทด์โดยเครื่อง automated sequencer ปรากฏว่าชิ้นดีเอ็นเอจำเพาะนี้มีความยาว 220 bp และมีลำดับนิวคลีโอไทด์ดังแสดงในตารางที่ 1 สมมุติว่าถ้าต้องการทราบโดยคร่าวๆ ว่าชิ้นดีเอ็นเอนี้ควรเป็นส่วนหนึ่งส่วนใดของยีนหรือโกรโมนของไวรัสหรือไม่ และถ้าต้องการออกแบบไพรเมอร์ซึ่งจำเพาะกับชิ้นดีเอ็นเอนี้เพื่อใช้ตรวจสอบการติดเชื้อไวรัสชนิดนี้ในสิ่งส่งตรวจ

### วิธีทำ

1. เรียกใช้โปรแกรม web browser และ microsoft word ที่มีข้อมูลลำดับนิวคลีโอไทด์
2. เปลี่ยนรูปแบบของลำดับนิวคลีโอไทด์เป็น FASTA format โดยใช้โปรแกรมชื่อ ReadSeq ซึ่งมีบริการที่

<http://dot.imgen.bcm.tmc.edu/seq-util/Options/readseq.html>

ตารางที่ 2 แสดงลำดับนิวคลีโอไทด์ที่เป็น FASTA format แล้ว

3. ใช้โปรแกรม web browser ติดต่อกับกลุ่ม Sequence Similarity Server ในที่นี่ใช้โปรแกรม FASTA ที่มีบริการที่ DDBJ server เพราะสามารถฐานข้อมูลของสิ่งมีชีวิตที่สนใจได้ เช่น พีช ไวรัส หรือสัตว์

<http://www.ddbj.nig.ac.jp/search/fasta-e.html>

ตารางที่ 3 แสดงให้เห็นว่าลำดับนิวคลีโอไทด์ที่สนใจมีความเหมือนกับ gag gene ของ human immunodeficiency virus type I มากที่สุด (100% identity) เพราะฉะนั้น จึงสันนิษฐานได้ว่าชิ้นดีเอ็นเอนี้ควรจะเป็นส่วนหนึ่งของ gag gene ของไวรัสโรคเอดส์ แต่เพื่อความมั่นใจควรมีการทดลองชิ้นดีเอ็นเอนี้หรือโปรตีนที่ผลิตจากชิ้นดีเอ็นเอนี้เพิ่มเติมด้วยวิธีทางอนุชีวิทยา เพื่อสนับสนุนสมมุติฐานดังกล่าว

4. ใช้โปรแกรม web browser ติดต่อกับ server ในกลุ่ม Primer design server และป้อนข้อมูลลำดับนิวคลีโอไทด์พร้อมกับกำหนดค่า parameter ที่เหมาะสม

<http://www.williamstone.com/primers/nojavascript/>

5. ได้ผลลัพธ์เป็นคู่ไพรเมอร์ที่เหมาะสมจำนวนมาก ที่มีสมบัติตาม parameter ที่กำหนดไว้แต่แรก ไพรเมอร์ 1 คู่ ที่จะถูกเลือกใช้จะพิจารณาจาก melting temperature,

ความยาวของ PCR product และตำแหน่งที่เพรเมอร์จับน้ำดีอีนเอ ดังแสดงใน  
ตารางที่ 4

ตารางที่ 1

>unknown

```
aaatactggacagctacaaccgtcccttcagacaggatcagaagaacttagatcattatataatacagttagca  
accctctatttgtgtcatcaaaggatagagataaaagacacccaaggaaagcttagacaagatagaggaagag  
caaaaacaaaagtaagaaaaaggcacagcaagcagcagctgacacaggaaacagcagccaggtcagcc  
aaaatta
```

ตารางที่ 2

>unknown, 220 bases, 253B checksum

```
aaatactggacagctacaaccgtcccttcagacaggatcagaagaacttagatcattatataatacagttagca  
accctctatttgtgtcatcaaaggatagagataaaagacacccaaggaaagcttagacaagatagaggaagag  
caaaaacaaaagtaagaaaaaggcacagcaagcagcagctgacacaggaaacagcagccaggtcagcc  
aaaatta
```

ທ່າງການ 3

```

FASTA (3.06 Sept. 1996) function (optimized, DNA matrix) ktop: 6
Join: 47, opt: 32, gap-pen -16/-4, width: 16 reg.-scaled
Scan time: 260.750

The best scores are:
          initn initl opt z-sc E(47173)
L02293|HIV1S1A Human immunodeficiency virus (343) 1100 1100 1100 732.4 7.3e-34
D86068|HIVMCK1 Human immunodeficiency virus (9752) 1073 1073 1073 704.2 9.3e-34
K02013|HIVBRUGG Human immunodeficiency virus (9229) 1073 1073 1073 704.4 9.3e-34
L02296|HIV1S2A Human immunodeficiency virus (343) 1091 1091 1091 726.5 1.5e-33
L02295|HIV1S3A Human immunodeficiency virus (343) 1091 1091 1091 726.5 1.5e-33
L02299|HIV1S7A Human immunodeficiency virus (343) 1091 1091 1091 726.5 1.5e-33
L02294|HIV1S2A Human immunodeficiency virus (343) 1091 1091 1091 726.5 1.5e-33
L02297|HIV1S5A Human immunodeficiency virus (343) 1091 1091 1091 726.5 1.5e-33
L02290|HIV1S6A Human immunodeficiency virus (343) 1091 1091 1091 726.5 1.5e-33
LJ1963|HIVTM475A Human immunodeficiency virus (9795) 1064 1064 1064 698.4 2e-33
K02083|HIVPV22 Human immunodeficiency virus (9770) 1064 1064 1064 698.4 2e-33
X01762|RENTLV Human T-cell leukaemia type (9740) 1064 1064 1064 698.4 2e-33
M15654|HIVB8102 Human immunodeficiency virus (8932) 1064 1064 1064 698.6 2.1e-33
K02012|HIVM3B5 Human immunodeficiency virus (5362) 1064 1064 1064 700.3 2.9e-33
L02303|HIV1S11A Human immunodeficiency virus (343) 1082 1082 1082 720.7 3.3e-33
E11530|HIVF12CG Human immunodeficiency virus (9781) 1055 1055 1055 692.5 4.3e-33
D86069|HIV2132 Human immunodeficiency virus (9754) 1055 1055 1055 692.5 4.3e-33
K03455|HIVHXB2CG Human immunodeficiency virus (9719) 1055 1055 1055 692.5 4.3e-33
L02302|HIV1S10A Human immunodeficiency virus (343) 1073 1073 1073 714.8 7e-33
L02300|HIV1S8A Human immunodeficiency virus (343) 1073 1073 1073 714.8 7e-33

Go to top
>>L02293|HIV1S1A Human immunodeficiency virus type 1 (81 - (343 nt)
initn: 1100 initl: 1100 opt: 1100 z-score: 732.4 expect(): 7.3e-34
100.00% identity in 220 nt overlap

          10      20      30
Query:      A A A T A C T G G G A C A G C T A C A C C G T C C C T T C
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: G C C T G T T A G A A A C A T C A G A G G O C T G T A G A C M A A T C Y G G G A C A G C T A C A C C G T C C C T T C
        40      50      60      70      80      90
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Query: A G A C A G G A T C A G A A G A C T T A G A T C A T T A T A A T A C A G T A G C A A C C C T C T A T T G T G T G C
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: A G A C A G G A T C A G A A G A C T T A G A T C A T T A T A A T A C A G T A G C A A C C C T C T A T T G T G T G C
        100     110     120     130     140     150
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Query: A T C A A G G A T A G A G A T A A A A G A C A C C G A C A G G A C A G C T G A C A C A G G A A C C G C A G G T C A
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: A T C A A G G A T A G A G A T A A A A G A C A C C G A C A G G A C A G C T G A C A C A G G A A C C G C A G G T C A
        160     170     180     190     200     210
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Query: A C A A A A G T A A G A A A A A G G C A C A G C A A G C A G C T G A C A C A G G A A C C G C A G G T C A
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: A C A A A A G T A A G A A A A A G G C A C A G C A A G C A G C T G A C A C A G G A A C C G C A G G T C A
        220     230     240     250     260     270
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
220
Query: G C C A A A A T T A
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: G C C A A A A T T A A C C T T A T A G T G C A G A C A C T C C A G G G C A A A T G G T A C A T C A G G C C A T A T C A C
        280     290     300     310     320     330
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Go to top
>>D86068|HIVMCK1 Human immunodeficiency virus type 1 DNA (9752 nt)
initn: 1073 initl: 1073 opt: 1073 z-score: 704.2 expect(): 9.3e-34
98.63% identity in 220 nt overlap

          10      20      30
Query:      A A A T A C T G G G A C A G C T A C A C C G T C C C T C
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: G C C T G T T A G A A A C A T C A G A G G O C T G T A G A C A A A T C Y G G A C A G C T A C A C C G T C C C T C
        340     350     360     370     380     390
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Query: A G A C A G G A T C A G A A G A C T T A G A T C A T T A T A A T A C A G T A G C A A C C C T C T A T T G T G T G C
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: A G A C A G G A T C A G A A G A C T T A G A T C A T T A T A A T A C A G T A G C A A C C C T C T A T T G T G T G C
        400     410     420     430     440     450
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Query: A T C A A G G A T A G A G A T A A A A G A C A C C G A C A G G A A C C G C A G G T C A
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: A T C A A G G A T A G A G A T A A A A G A C A C C G A C A G G A A C C G C A G G T C A
        460     470     480     490     500     510
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
100     110     120     130     140     150
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Query: A C A A A A G T A A G A A A A A G G C A C A G C A A G C A G C T G A C A C A G G A A C C G C A G G T C A
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: A C A A A A G T A A G A A A A A G G C A C A G C A A G C A G C T G A C A C A G G A A C C G C A G G T C A
        520     530     540     550     560     570
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
220
Query: G C C A A A A T T A
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Subject: G C C A A A A T T A A C C T T A T A G T G C A G A C A C T C C A G G G C A A A T G G T A C A T C A G G C C A T A T C A C
        580     590     600     610     620     630
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

ตารางที่ 4

| ID | Forward primer          | Base number | Tm °C |
|----|-------------------------|-------------|-------|
| 1  | AAATACTGGGACAGCTACAACC  | 1..22       | 57.32 |
| 2  | AATACTGGGACAGCTACAACCG  | 2..23       | 60.42 |
| 3  | ATACTGGGACAGCTACAACCGT  | 3..24       | 59.95 |
| 4  | TACTGGGACAGCTACAACCGTC  | 4..25       | 61.47 |
| 5  | ACAGCTACAACCGTCCCTTCAG  | 11..32      | 62.81 |
| 6  | AGCTACAACCGTCCCTTCAGAC  | 13..34      | 61.92 |
| 7  | CTACAACCGTCCCTTCAGACAG  | 15..36      | 61.08 |
| 8  | GTCCCTTCAGACAGGGATCAGAA | 23..44      | 60.63 |
| 9  | TCCCTTCAGACAGGGATCAGAAG | 24..45      | 60.76 |
| 10 | CCCTTCAGACAGGGATCAGAAGA | 25..46      | 60.76 |

| ID | Reverse primer         | Base number | Tm °C |
|----|------------------------|-------------|-------|
| 1  | TGCTTGCTGTGCCTTTCTTA   | 179..158    | 62.22 |
| 2  | GCTTGCTGTGCCTTTCTTAC   | 178..157    | 60.42 |
| 3  | CTTGCTGTGCCTTTCTTACT   | 177..156    | 57.91 |
| 4  | TTGCTGTGCCTTTCTTACTT   | 176..155    | 58.26 |
| 5  | TGCTGTGCCTTTCTTACTTT   | 175..154    | 58.26 |
| 6  | GCTGTGCCTTTCTTACTTT    | 174..153    | 57.01 |
| 7  | TCTTGTCTAAAGCTTCCTTGGT | 135..114    | 57.32 |
| 8  | CTTGTCTAAAGCTTCCTTGGTG | 134..113    | 58.20 |
| 9  | TTGTCTAAAGCTTCCTTGGTGT | 133..112    | 58.06 |
| 10 | TGTCTAAAGCTTCCTTGGTGT  | 132..111    | 58.51 |

## **แบบฝึกหัดบทปฎิบัติการที่ 8**

1. การค้นหาและวิเคราะห์ข้อมูลด้านพันธุกรรมโดยใช้เครือข่ายอินเตอร์เน็ตมีประโยชน์อย่างไร
2. การค้นหาและวิเคราะห์ข้อมูลโดยใช้เครือข่ายอินเตอร์เน็ต นอกจากใช้ทางด้านพันธุกรรมแล้ว สามารถที่จะนำไปใช้ทางด้านใดได้อีกบ้าง
3. ในอนาคตอันใกล้นี้แนวโน้มการค้นหาและวิเคราะห์ข้อมูลด้านพันธุกรรมโดยใช้เครือข่ายอินเตอร์เน็ต จะมีการเปลี่ยนแปลงไปในทิศทางใด