ภาคผนวก

อาหารเลี้ยงเชื้อ สารละลาย บัฟเฟอร์ เครื่องแก้ว และพลาสติก ที่นำมาใช้ในงานทางพันธุวิศวกรรม ต้องทำให้ปราศจากเชื้อด้วยการใช้ autoclave ที่ความดัน 15 ปอนด์/ตารางนิ้ว ที่อุณหภูมิ 121°C เป็นเวลา 15-20 นาที หรือทำให้ปราศจากเชื้อด้วยการกรองด้วยชุด filter ที่มีแผ่นกรองขนาด 0.22 μm หรือ 0.45 μm ที่ sterile แล้ว และกรองลงสู่หลอด หรือขวด หรือภาชนะที่ sterile แล้วเช่นกัน บัฟเฟอร์หรือสารบางชนิดอาจจะไม่ต้องนำไป sterile ขึ้นกับชนิดของสารละลายนั้นๆ เครื่องแก้วทุกชนิดต้องล้างให้สะอาด และทำให้ปราศจากเชื้อโดยการนำไปอบที่ 180°C เป็นเวลา 2 ชั่วโมง หรือ autoclave อุปกรณ์พลาสติก เช่น microtube, tip, centrifuge tube ถ้าผ่านการฆ่าเชื้อโดยการใช้รังสี (γ-irradiation) ก็สามารถนำมาใช้งานได้เลย แต่ถ้ายังไม่ได้รับการฆ่าเชื้อให้นำไป sterile ด้วยการ autoclave อุปกรณ์พลาสติกควรเลือกแบบใช้แล้วทิ้ง (disposable) ถ้าไม่จำเป็นไม่ควรนำไปล้างแล้วนำกลับมาใช้ใหม่ (reusable) เพราะอาจจะมีการปนเปื้อนของสารเอนไซม์ ดีเอ็นเอ หรืออาร์เอ็นเอ ดังนั้นถ้าจะล้างแล้วนำกลับมาใช้ใหม่ จะต้องทำความสะอาดจนแน่ใจว่าไม่มีการปนเปื้อน โดยอาจจะใช้สารดีเทอร์เจนท์ (detergent) หรือสารเคมี เช่น clorox เพื่อทำลายสารปนเปื้อนนั้นๆ

การทำ siliconized เครื่องแก้ว พลาสติก และใยแก้ว (glasswool) ทำโดยใส่สารละลาย 5% dichlorodimethyl silane ใน chloroform ลงในเครื่องแก้วหรือพลาสติกหรือใยแก้วที่ต้องการทำ siliconized แช่ทิ้งไว้ 5 นาที หลังจากนั้นนำไปล้างน้ำหลายๆ ครั้ง และนำเครื่องแก้วหรือใยแก้วไปอบที่ 180°C เป็นเวลา 2 ชั่วโมง

นอกจากนี้ในการทำปฏิบัติการทางจุลชีววิทยานั้น ผู้ปฏิบัติต้องปฏิบัติตามกฎข้อบังคับและระเบียบต่างๆ อย่างเคร่งครัด ทั้งนี้เพื่อป้องกันการติดเชื้อจากจุลินทรีย์ที่ก่อให้เกิดโรคได้ การป้องกันนี้มิใช่จะเพียงป้องกันเฉพาะตนเองเท่านั้น แต่เป็นการป้องกันผู้อื่นที่อยู่รอบตัวเราด้วย อีกทั้งยังเป็นการป้องกันมิให้การทดลองที่ทำอยู่เกิดการปนเปื้อนจากจุลินทรีย์ที่ไม่ต้องการซึ่งจะทำให้การทดลองเสียไป

1. การเตรียมอาหารเลี้ยงเชื้อ

1.1 LB medium ใช้เลี้ยงแบคทีเรียทั่วๆ ไป

Bacto Tryptone	10	g
Yeast extract	5	g
NaCl	5	g
เติมน้ำกลั่นให้ได้	1,000	ml

ถ้าต้องการเตรียม LB agar ให้เติม Bacto agar ลงไปอีก 1.5% นำไป autoclave ที่ความดัน 15 ปอนด์/ตารางนิ้ว อุณหภูมิ 121°C เป็นเวลา 15-20 นาที ในกรณีที่มีอาหารเลี้ยงเชื้อสำเร็จรูปจากบริษัท ให้เตรียมตาม direction ตามฉลากข้างขวด

2. การเตรียมยาปฏิชีวนะและสารเคมีอื่น ๆ ที่ใช้ผสมใอาหารเลี้ยงเชื้อ

2.1 แอมพิซิลิน

เตรียม stock solution ที่ 10 mg/ml โดยละลาย 0.2 g ในน้ำกลั่น 20 ml จนละลายดี แล้วนำไปกรองด้วยแผ่นกรอง (filter) ขนาด 0.22 μm หรือ 0.45 μm ที่ sterile แล้ว ลงสู่หลอดที่ sterile นำไปเก็บที่ –20°C

2.2 คลอแรมเฟนิคอล

เตรียม stock solution ที่ 35 mg/ml โดยละลาย 0.7 g ในเอธานอล (ethanol) 20 ml จนละลายดี แล้วนำไปกรองด้วยแผ่นกรอง ขนาด 0.22 μm หรือ 0.45 μm ที่ sterile แล้ว ลงสู่หลอดที่ sterile นำไปเก็บที่ –20°C

2.3 Isoprpyl thio-β-D-galactoside (IPTG)

เตรียม stock solution ที่ความเข้มข้น 0.4 M โดยละลาย 0.477 g ในน้ำกลั่น 5 ml จนละลายดี แล้วนำไปกรองด้วยแผ่นกรองขนาด 0.22 μm หรือ 0.45 μm ที่ sterile แล้ว ลงสู่หลอดที่ sterile นำไปเก็บที่ –20°C

3. การเตรียมบัฟเฟอร์และสารละลายอื่น ๆ

3.1 0.1 M $CaCl_2$

ละลาย $CaCl_2.2H_2O$ 147 g ในน้ำกลั่น ปรับปริมาตรให้เป็น 100 ml นำไป sterile ด้วยการ autoclave

3.2 1 M Dithiotreitol (DTT)

ละลาย DTT 1.545 g ในน้ำกลั่น 10 ml ทำให้ sterile ด้วยการกรอง อย่า autoclave สารละลาย DTT หรือสารละลายที่มี DTT ผสมอยู่

3.3 0.5 M EDTA

ชั่ง Disodium ethylenediamine tetraacetate.$2H_2O$ 136.1 g ในน้ำกลั่น 800 ml กวนจนละลายด้วย magnetic stirrer เติมเกล็ด NaOH ลงไปจนกระทั่งได้ pH 8.0 ซึ่งเป็น pH ที่ EDTA ละลายหมดพอดี ปรับปริมาตรให้เป็น 1,000 ml แล้วนำไป autoclave

3.4 Ethidium bromide (10 mg/ml)

ชั่ง ethidium bromide 1 g ละลายในน้ำกลั่น 100 ml กวนจนละลายด้วย magnetic stirrer (ใช้เวลาหลายชั่วโมง) ใส่ขวดหุ้มด้วย aluminium foil หรือเก็บใส่ขวดสีชาเข้มที่ 4°C ต้องสวมถุงมือระหว่างการเตรียม ระวังอย่าหายใจเอาผง ethidium bromide เข้าไปในระหว่างชั่ง เพราะผงของสารนี้ค่อนข้างละเอียดและเบา และสารนี้มีคุณสมบัติเป็น strong mutagen

3.5 1 M HCl

เตรียมปริมาตร 1 Litre โดยการเติม concentrated HCl 86.2 ml ลงในน้ำกลั่น 913.8 ml ผสมให้เข้ากัน ควรทำใน chemical hood เพราะไอกรดของ HCl จะเกิดเป็นควันฟุ้งกระจาย

3.6 Loading dye

สูตรของ loading buffer ที่ใช้กันทั่วไปมีหลายสูตรด้วยกันดังนี้

สูตรที่ 1 : 0.25% Bromphenol blue, 0.25% Xylene cyanol FF และ 40% (w/v) Sucrose ในน้ำกลั่น เก็บที่ 4°C

สูตรที่ 2 : 0.25% Bromphenol blue, 0.25% Xylene cyanol FF และ 15% Ficoll (Type 400, Pharmacia) ในน้ำกลั่น เก็บที่ 4°C

สูตรที่ 3 : 0.25% Bromphenol blue, 0.25% Xylene cyanol FF และ 30% Glycerol ในน้ำกลั่น เก็บที่ 4°C

สูตรที่ 4 : 0.25% Bromphenol blue และ 4% Sucrose ในน้ำกลั่น เก็บที่ 4°C

สูตรที่ 5 : 0.25% Bromphenol blue และ 10% Glycerol ในน้ำกลั่น เก็บที่ 4°C

Loading dye ที่เตรียมตามสูตรเหล่านี้มีความเข้มข้น 6 เท่า ของที่ต้องการใช้ ดังนั้นปริมาณที่ใช้จะเป็น 1/6 ของปริมาตรรวมของสารละลายดีเอ็นเอสุดท้าย เช่น มีสารละลายดีเอ็นเอ 10 μl เติม loading dye 2 μl ซึ่งให้ปริมาตรสุดท้าย 12 μl

3.7 0.1 M $MgCl_2$

ละลาย $MgCl_2.6H_2O$ 2.03 g ในน้ำกลั่น ปรับปริมาตรให้เป็น 100 ml นำไป sterile ด้วยการ autoclave

3.8 1 M NaCl

ละลาย NaCl 29.2 g ในน้ำกลั่น ปรับปริมาตรให้เป็น 100 ml นำไป sterile ด้วยการ autoclave

3.9 10 M NaOH

ละลาย NaOH 40 g ในน้ำกลั่น 80 ml ปรับปริมาตรให้เป็น 100 ml ถ้าต้องการ sterile ก็ทำได้ด้วยการ autoclave การใช้สารละลาย NaOH โดยทั่วๆ ไป เช่น นำไปเจือจางเพื่อใช้ในการเตรียมดีเอ็นเอ ไม่จำเป็นต้อง autoclave แต่อาจละลายหรือเจือจางด้วยน้ำกลั่นที่ sterile ก็เพียงพอ

3.10 สารละลายฟีนอลอิ่มตัว

เตรียมโดยเอาฟีนอลมาทำให้หลอมละลายโดยแช่ใน waterbath ที่ 65°C เติม hydroxy quinoline ซึ่งเป็นสารพวก antioxidant ให้ได้ความเข้มข้น 0.1% เติม 0.1 M Tris-HCl pH 8.0 ปริมาตรเท่าตัว กวนอย่างแรงด้วย magnetic stirrer นาน 10-15 นาที ปล่อยให้แยกชั้นดูด aqueous phase ชั้นบนออก สกัดอีกครั้งด้วย 0.1 M Tris-HCl pH 8.0 จนกระทั่ง pH ของฟีนอล >7.8 (วัดด้วย pH paper) เติม β-mercaptoethanol ให้ได้ 0.2% และเติม 0.1 M Tris-HCl pH 8.0 เพื่อปิดทับผิวของสารละลาย เก็บสารละลายฟีนอลอิ่มตัวที่ได้ในขวดสีชาเข้มหรือขวดแก้วหุ้มด้วย aluminium foil เก็บไว้ที่ 4°C

3.11 Phenol/Chloroform

ผสม melted phenol และ chloroform ปริมาตรเท่าๆ กัน แล้วเติม 0.1 M Tris-HCl pH 7.6 เขย่าแรงๆ ให้ผสมกัน ปล่อยให้แยกชั้น ดูดชั้นบนออกและทิ้งไป สกัดซ้ำด้วย 0.1 M Tris-HCl pH 7.6 อีก 2-3 ครั้ง เติม 0.1 M Tris-HCl pH 7.6 ให้ท่วมผิวสารละลาย เก็บในขวดสีชาเข้มหรือขวดแก้วหุ้มด้วย aluminium foil เก็บไว้ที่ 4°C

3.12 Phenol/Chloroform/Isoamyl alcohol = 25:24:1

ผสม Chloroform และ Isoamyl alcohol ด้วยอัตราส่วน 24:1 นำไปผสมกับสารฟีนอลอิ่มตัวปริมาตรที่เท่ากัน ส่วนใหญ่จะผสมกับฟีนอลอิ่มตัวเมื่อจะนำมาใช้งาน ถ้าต้องการจะเก็บไว้หลังจากผสมกันแล้ว ให้เติม 0.1 M Tris-HCl pH 8.0 เพื่อปิดทับสารละลาย เก็บในขวดสีชาเข้มหรือขวดแก้วหุ้มด้วย aluminium foil เก็บไว้ที่ 4°C ใช้ได้ภายใน 1 เดือน สำหรับส่วนผสมระหว่าง Chloroform และ Isoamyl alcohol เก็บไว้ได้นาน

3.13 3 M Potassium acetate

ชั่ง Potassium acetate 29.46 g ละลายในน้ำกลั่นให้ได้ปริมาตร 60 ml เติม glacial acetic acid 11.5 ml และน้ำกลั่น 28.5 ml จะได้สารละลายที่มี Potassium acetate ความเข้มข้น 3 M และ acetic acid ความเข้มข้น 5 M นำไป sterile ด้วยการ autoclave

3.14 3 M Sodium acetate pH 4.8

ชั่ง Sodium acetate.$3H_2O$ 408.1 g ละลายในน้ำกลั่น 750 ml เติม glacial acetic acid จนกระทั่งได้ pH 4.8 ปรับปริมาตรให้เป็น 1,000 ml และนำไป sterile ด้วยการ autoclave

3.15 10% SDS

ละลาย Sodium dodesyl sulfate 10 g ในน้ำกลั่นที่ sterile แล้ว 100 ml อาจอุ่นเล็กน้อย เพื่อช่วยให้ละลายดีขึ้น เก็บที่อุณหภูมิห้อง

3.16 1 M Tris-HCl pH 8.0

ละลาย Tris base 121.1 g ในน้ำกลั่น 800 ml ปรับ pH โดยการใช้กรด concentrated HCl จนกระทั่งได้ pH 8.0 แล้วปรับปริมาตรให้เป็น 1,000 ml นำไป sterile ด้วยการ autoclave

3.17 TGE buffer (50mM Glucose, 25 mM Tris, 10 mM EDTA, pH 8.0)

ละลาย glucose 9.01 g, Tris base 3.03 g และ EDTA 2.92 g ในน้ำกลั่น 800 ml ปรับ pH โดยการใช้กรด concentrated HCl จนกระทั่งได้ pH 8.0 แล้วปรับปริมาตรเป็น 1,000 ml นำไป sterile ด้วยการ autoclave

3.18 TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0)

ละลาย Tris base 1.21 g และ EDTA 0.29 g ในน้ำกลั่น 800 ml ปรับ pH โดยการใช้กรด concentrated HCl จนกระทั่งได้ pH 8.0 แล้วปรับปริมาตรเป็น 1,000 ml นำไป sterile ด้วยการ autoclave

หมายเหตุ : ในปัจจุบันมีบัฟเฟอร์หรือสารละลายต่างๆ มีขายเป็นน้ำยาสำเร็จรูป (commercial kit) เช่น น้ำยาสำเร็จรูปในการสกัดดีเอ็นเอหรือพลาสมิด (ได้แก่ QIAGEN, Eppendorf) สารละลายฟีนอลอิ่มตัว, loading dye (มาพร้อมกับการซื้อ standard DNA), 10x PCR buffer และ 25 mM $MgCl_2$ (มาพร้อมกับการซื้อ Taq DNA polymerase) น้ำยาเหล่านี้สามารถนำมาใช้งานได้ทันที หรือทำตามขั้นตอนในคู่มือที่มากับน้ำยาสำเร็จรูป

รูปที่ 1 น้ำยาสำเร็จรูปใช้ในการสกัดพลาสมิดดีเอ็นเอ (Eppendor)

4. การเตรียมเอนไซม์

4.1 Lysozyme

เตรียมเป็น stock solution ที่ความเข้มข้น 50 mg/ml ละลายในน้ำกลั่นที่ sterile แบ่งเก็บเป็น aliquots ปริมาตรน้อยๆ (10-20 μl) และปริมาตรมาก (100-1,000 μl) ใน

หลอดที่ sterile เก็บที่ –20°C การ Freeze-thaw หลายๆ ครั้ง จะทำให้เอนไซม์เสียแอคติวิตี ดังนั้นจึงควรหยิบหลอดที่มีเอนไซม์ปริมาณที่พอเหมาะมาใช้งานแต่ละครั้ง หรือจะชั่งเอนไซม์ทุกครั้งที่ใช้งานก็ได้

4.2 Rnase A

ชั่ง Rnase A 100 mg และละลายในสารละลายปริมาตร 10 ml ที่มี 10 mM Tris-HCl pH 7.5 และ 15 mM NaCl ผสมอยู่ ต้มในน้ำเดือด 15 นาที (ระวังอย่าให้ระดับสารละลายเอนไซม์อยู่ต่ำกว่าระดับน้ำเดือด) ปล่อยให้เย็นลงที่อุณหภูมิห้อง แบ่งเป็น aliquots (100-500 μl) เก็บที่ –20°C ที่อุณหภูมินี้เอนไซม์ยังคงทำงานได้ดี แม้จะผ่านการทำ Freeze-thaw หลายๆ ครั้ง

4.3 Proteinase K

เตรียม stock solution 20 mg/ml ในน้ำกลั่นที่ sterile เก็บที่ –20°C และใช้ที่ความเข้มข้น 50 mg/ml ใน reaction buffer ที่ประกอบด้วย 10 mM Tris (pH 7.8), 5 mM EDTA และ 0.5% SDS

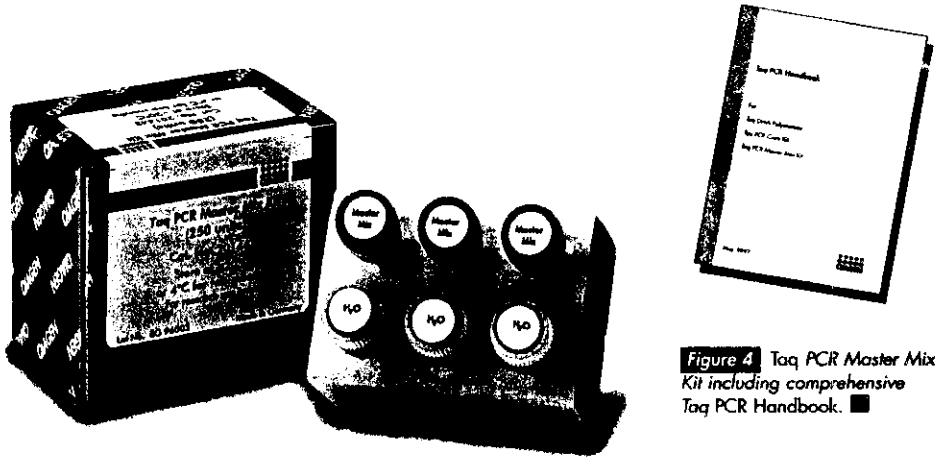


Figure 4 Taq PCR Master Mix Kit including comprehensive Taq PCR Handbook.

รูปที่ 2 น้ำยาสำเร็จรูป Taq PCR Master Kit (QIAGEN)

5. ข้อมูลพื้นฐาน

5.1 หน่วยพื้นฐานในระบบเมตริก

เป็นหน่วยที่ใช้วัดมวลได้แก่ กรัม หน่วยที่วัดความยาวได้แก่ เมตร และหน่วยที่ใช้วัดปริมาตรของไหลได้แก่ ลิตร คำนิยามของหน่วยเหล่านี้คือ

กรัม (gramme, g) เท่ากับ 1/1000 ของมวลของแท่งโลหะเจือระหว่างแพลทินัมกับอิริเดียม (Platinum-iridium) ซึ่งเก็บไว้ที่ International Bureau of Weight and Measures ประเทศฝรั่งเศส

เมตร (metre, m) คือความยาวระหว่างจุดสองจุดบนท่อนโลหะเจือที่ทำขึ้นจากแพลทินัมและอิริเดียม (Platinum-iridium) ซึ่งเก็บไว้ที่ International Bureau of Weight and Measures ประเทศฝรั่งเศส

ลิตร (litre, l) เป็นปริมาตรเท่ากับ 1,000 ลูกบาศก์เซนติเมตร พอดี

ตารางที่ 1 หน่วยย่อยของหน่วยพื้นฐาน คำนำหน้า สัญลักษณ์ และแฟคเตอร์ (เลขคูณ)

คำนำหน้า	สัญลักษณ์	แฟคเตอร์ (เลขคูณ)
Teta	T	10^{12}
Giga	G	10^{9}
Mega	M	10^{6}
kilo	k	10^{3}
hecto*	h	10^{2}
deca*	da	10^{1}
deci*	d	10^{-1}
centi*	c	10^{-2}
milli	m	10^{-3}
micro	μ	10^{-6}
nano	n	10^{-9}
pico	p	10^{-12}
femto	f	10^{-15}
alto	a	10^{-18}

* จะใช้ต่อเมื่อคำอื่นๆ ที่มีอยู่นั้นไม่เหมาะสม

5.2 กรดและด่างเข้มข้น

ตารางที่ 2 แสดง Molarities และ Specific gravities ของสารละลายกรดและด่างเข้มข้น

Acid/base	Molecular weight	%by weight	Molarity (approx.)	1 M Solution (ml/l)	Specific gravity
Acetic acid (glacial)	60.05	99.6	17.4	57.5	1.05
Ammonium hydroxide	35.0	28	14.8	67.6	0.90
Formic acid	46.03	90	23.6	42.4	1.205
		98	25.9	38.5	1.22
Hydrochloric acid	36.46	36	11.6	85.9	1.18
Nitric acid	63.01	70	15.7	63.7	1.42
Perchloric acid	100.46	60	9.2	108.8	1.54
Phosphoric acid	98.00	85	14.7	67.8	1.7
Sulfuric acid	98.07	98	18.3	54.5	1.835

5.3 ความรู้พื้นฐานเกี่ยวกับกรดนิวคลีอิคและโปรตีน

1 bp	=	650 Da
1 MDa ds DNA	=	1.54 kb
dNTP (เฉลี่ย)	=	325 Da
1 A_{260} D/S DNA	=	50 μg/ml
1 A_{260} S/S DNA, S/S RNA	=	40 μg/ml
1 A_{260} Oligonucleotide	=	33 μg/ml
1 amino acid (เฉลี่ย)	=	110 Da

ตารางที่ 3 ขนาดของ Genome ของเซลล์ต่างๆ*

Organism	Base pairs/ Haploid genome	Organism	Base pairs/ Haploid genome
SV40	5,243	*Galius domesticus* (chicken)	1.2×10^8
ϕx174	5,386	*Mus musculus* (mouse)	2.7×10^9
Adenovirus 2	35,937		
Lamda	48,502	*Raltus morvigeticus* (rat)	3.0×10^9
E. coli	4.7×10^6		
Saccharomyces cerevisiae (yeast)	1.5×10^7	*Xenopus laevis* (frog)	3.1×10^9
Arabidopsis thaliana (small plant)	7.0×10^7	*Homo sapiens* (human)	3.3×10^9
		Zea mays (maize)	3.9×10^9
Drosophilia melanogaster (fruit fly)	1.4×10^8	*Nicotiana tabacum* (tabacco)	4.8×10^9

* Genome size หาได้โดยการทำ sequence analysis (viruses), electrophoretic analysis (*E. coli*, *S. cerevisiae*) หรือการใช้วิธีหา DNA content per cell ร่วมกับการทำ hybridization kinetics

5.4 Codons สำหรับ amino acids และ stop codon

ตารางที่ 4 รหัสพันธุกรรม

นิวคลีโอไทด์ตัวที่หนึ่ง	นิวคลีโอไทด์ตัวที่สอง				นิวคลีโอไทด์ตัวที่สาม
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Term	Term[+]	A
	Leu	Ser	Term	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile[+]	Thr	Lys	Arg[+]	A
	Met	Thr	Lys	Arg[+]	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Term = termination ; โคดอนหยุด

[*] นิวคลีโอไทด์ตัวที่หนึ่ง สอง และสาม หมายถึง โคดอนสามเบส (triplet)

[+] ในไมโตคอนเดรียของเซลล์สัตว์ AUA เป็นรหัสโคดอนของ Met, UGA เป็น Trp, AGA และ AGG เป็นโคดอนหยุด

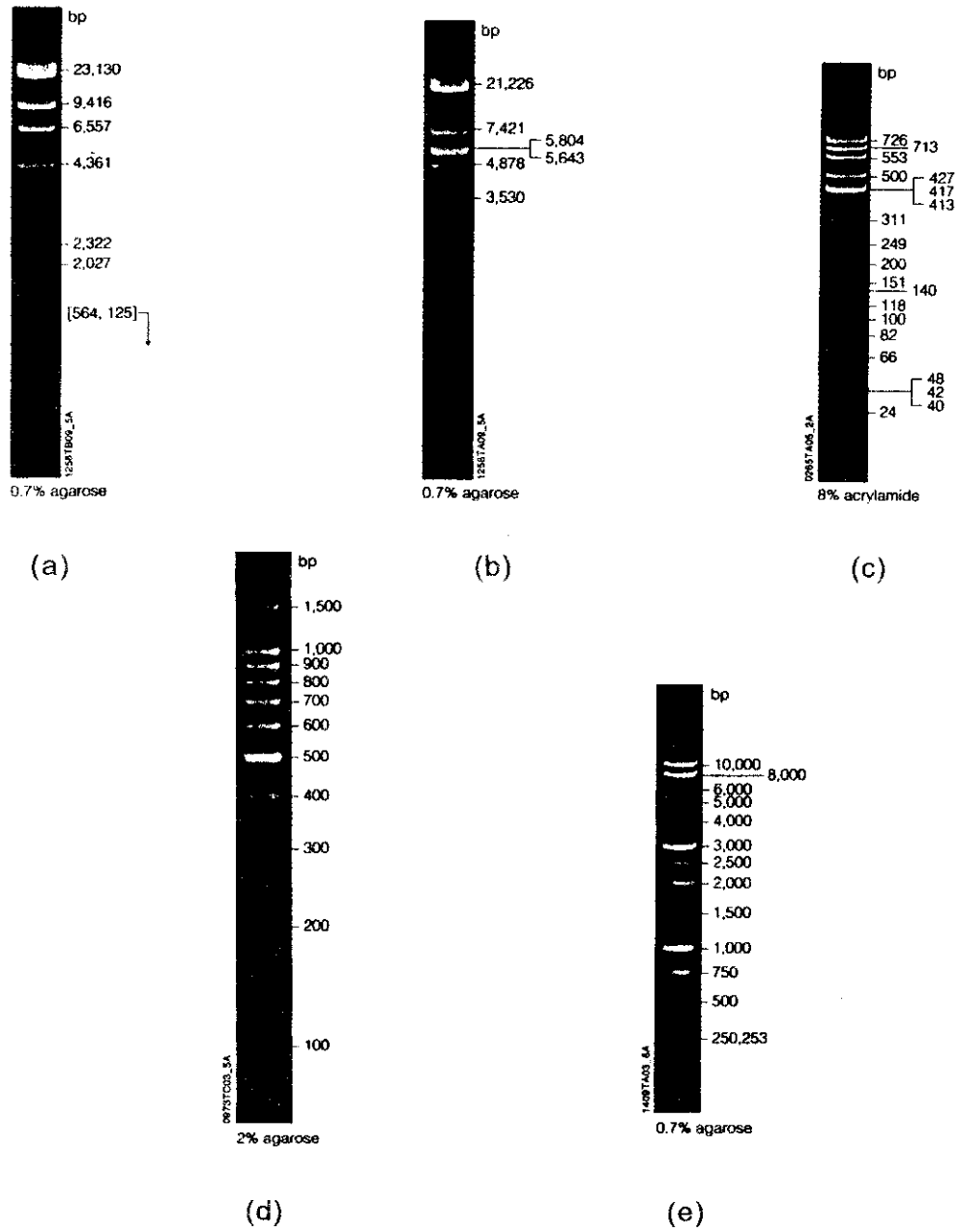
ตารางที่ 5 รหัส (codes) แบบ 1 และ 3 ตัวอักษร สำหรับกรดอะมิโน

กรดอะมิโน	3 ตัวอักษร	1 ตัวอักษร	กรดอะมิโน	3 ตัวอักษร	1 ตัวอักษร
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspatic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamic acid	Glu	E	Serine	Ser	S
Glutamine	Gln	Q	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Lle	I	Valine	Val	V

5.5 Standard, molecular weight marker

ตารางที่ 6 ขนาดของ DNA fragment สายคู่ (double strand) หรือ สายเดี่ยว (single strand) ที่ใช้เป็น size marker แสดงในหน่วยของ bp ยกเว้น λ-*Pst*I แสดงในหน่วยของ kb

λ-*Hin*dIII	λ-*Cla*I	λ-*Pst*I	pBR322/*Hpa*II	pUC/*Sau*IIIAI
23,130	11,850	11.50	622	955
9,416	10,495	5.08	527	585
6,557	4,398	4.75	404	341
4,361	4,198	4.51	309	258
2,322	3,673	2.84	242	141
2,027	2,614	2.55	238	105
564	2,064	2.45	217	78
125	1,915	2.10	201	75
	1,804	2.00	190	46
	1,701	1.70	180	36
	1,112	1.16	160x2	
	973	1.09	147x2	
	657	0.81	123	
	621	0.51	110	
	537	0.47	90	
	354	0.45	76	
		0.34	67	
		0.26, 0.25		
		0.22, 0.2		
		0.16, 0.15		
		0.09		



รูปที่ 3 แสดงขนาดของ DNA fragment ของ standard DNA markers ชนิดต่างๆ บน agarose gel electrophoresis (a) Lamda DNA/*Hin*dIII (b) Lamda DNA/*Eco*RI (c) ϕX174 DNA/*Hin*fI (Promega) (d) 100 bp DNA Ladder (e) 1 kb DNA Ladder

6. การเขียนสัญลักษณ์ Genotype และ Phenotype

Genotype หมายถึง Genetic information ที่มีอยู่ในดีเอ็นเอของเซลล์

Phenotype หมายถึง ลักษณะของเชื้อหรือสิ่งมีชีวิตนั้นๆ ที่แสดงออกมาและสังเกตเห็นได้

ในกรณีของแบคทีเรีย

Genotype : ใช้สัญลักษณ์ตัวอักษรตัวพิมพ์เล็ก 3 ตัวอักษร เป็นชนิดตัวเอน (Italics) สำหรับแต่ละยีน ส่วน specific locus ที่ตามมา จะใช้ตัวอักษรตัวพิมพ์ใหญ่ที่เป็น Italics เช่นกัน ถ้าเป็น specific location ภายในยีนของ locus นั้นๆ ก็จะระบุด้วยตัวเลขที่ไม่ใช่ตัวเอน สำหรับ Genotype ที่เป็น wild type จะเขียนในลักษณะเดียวกัน แต่จะมีเครื่องหมาย superscript

ตัวอย่าง *lac Z* คือ Mutation ที่เกิดที่ *lac Z* gene

lac Z50 คือ Specific allele ที่เกิด Mutation ภายใน *lac Z* gene

lac Z^+ คือ wild type (สามารถสร้างเอนไซม์ β-galactosidase)

Phenotype : ใช้สัญลักษณ์ตัวอักษร 3 ตัว ที่เป็นอักษรตัวตรงเฉพาะตัวอักษรตัวแรกจะเป็นตัวพิมพ์ใหญ่

ตัวอย่าง Lac หมายถึง ความสามารถในการใช้ Lactose

สำหรับสายพันธุ์ที่มีพลาสมิด จะใส่ชื่อ Phage หรือพลาสมิดไว้ในวงเล็บ เช่น

E. coli MC1061(λ) หมายถึง *E. coli* สายพันธุ์ MC1061 ที่มี Prophage DNA อยู่ในโครโมโซม

E. coli BL21DE3plysS(pUC18) หมายถึง *E. coli* สายพันธุ์ BL21DE3plysS ที่มีพลาสมิด pUC18

Character อื่นๆ ที่พบบ่อยในการเขียน Genotype ได้แก่

Δ = Deletion เช่น Δ (lac Z) หมายถึง การ Delete *lac Z* gene

:: = Insertion เครื่องหมายนี้จะเขียนระหว่าง Target และยีนหรือส่วนของดีเอ็นเอที่เข้ามา Insert เช่น *lac Z* :: Tng หมายถึง Tng insert เข้าไปใน *lac Z* gene

หมายเหตุ : Phenotype ที่บ่งบอกลักษณะการดื้อยาของเชื้อ จะใช้ลักษณะที่แตกต่างกันออกไป แต่ที่นิยมใช้กันจะเป็นลักษณะอักษร 2 ตัว ตัวต้นเป็นอักษรพิมพ์ใหญ่ ตามด้วย

อักษรพิมพ์เล็ก และมี superscript เป็น r หรือ s ซึ่งหมายถึงดื้อหรือไวต่อยา เช่น Apr^r, Tc^s หมายถึง ดื้อยาแอมพิซิลิน และไวต่อยาเตตร้าซัยคลิน ตามลำดับ

7. ความปลอดภัยและข้อควรระวังในการทำงานด้านจุลินทรีย์ (Safety and precaution in microbiology works)

คำแนะนำและข้อบังคับบางอย่างในการทำปฏิบัติการ

1. ก่อนเข้าห้องปฏิบัติการทุกครั้งจะต้องศึกษาปฏิบัติการที่จะทำในวันนั้นล่วงหน้าก่อนจากคู่มือศึกษาว่าการทดลองนั้นๆ ทำอย่างไรและมีหลักการอย่างไร การที่จะทำการทดลองให้ได้ผลจะต้องทราบถึงสิ่งที่จะทำก่อนหลังและวัตถุประสงค์ของการทดลอง
2. เมื่อลงมือทำปฏิบัติการ จะต้องสวมเสื้อคลุมทุกครั้ง เก็บของใช้ส่วนตัวทุกชนิดไว้ในโต๊ะ ห้ามวางสิ่งของบนโต๊ะปฏิบัติการ ให้วางได้เฉพาะคู่มือปฏิบัติการและสมุดบันทึกผลการทดลองเท่านั้น
3. ก่อนลงมือทำและหลังจากเสร็จสิ้นการทำปฏิบัติการแล้วทุกครั้งจะต้องเช็ดโต๊ะด้วยฟองน้ำชุบน้ำยาฆ่าเชื้อโรค
4. ล้างมือฟอกสบู่ให้สะอาดทุกครั้งก่อนและหลังทำปฏิบัติการ
5. กระดาษ ก้านไม้ขีดไฟ ฯลฯ ต้องทิ้งลงถังผงทุกครั้ง
6. ห้ามสูบบุหรี่ รับประทานอาหาร ขนม หรือเครื่องดื่มในห้องปฏิบัติการโดยเด็ดขาด
7. ตะเกียงบุนเสน (bunsen burner) ถ้าไม่ใช้ต้องปิดแก๊สทุกครั้ง อย่าเปิดทิ้งไว้เป็นเวลานาน
8. ถ้าไม่ใช้กล้องต้องปิดไฟกล้องจุลทรรศน์ทุกครั้ง
9. ต้องใส่ปิเปตที่ใช้แล้วลงถังซึ่งมีน้ำยาฆ่าเชื้อโรค ต่อจากนั้นจึงนำไปล้างทำความสะอาด ห้ามวางปิเปตที่ใช้แล้วบนโต๊ะ
10. จานเลี้ยงเชื้อที่ใช้แล้วให้แยกใส่ลงในภาชนะที่จัดไว้โดยเฉพาะ ถ้าจุลินทรีย์ที่ใช้เป็นพวกก่อโรค จะต้องนำไปอบฆ่าเชื้อก่อนที่จะนำไปล้างทำความสะอาด
11. ให้แช่สไลด์ย้อมสีที่ใช้แล้วลงในน้ำยาฆ่าเชื้อแล้วจึงนำไปล้างก่อนนำไปใช้ต่อไป

12. หลอดทดลองที่มีจุกสำลี ให้แยกจุกสำลีใส่ลงในภาชนะต่างหาก ส่วนหลอดนั้นถ้ามีจุลินทรีย์ก่อโรคอยู่จะต้องอบฆ่าเชื้อก่อนที่จะนำไปล้าง
13. ถ้ามีอุบัติเหตุเกิดขึ้น เช่น แก้วบาด ไฟลวก ให้รีบรายงานอาจารย์ผู้ควบคุมปฏิบัติการทราบทันที

8. ตัวอย่างบริษัทที่ผลิตเครื่องมือ เอนไซม์ และสารเคมี ในงานทางพันธุวิศวกรรม

เวกเตอร์สำหรับการโคลนยีน (Cloning vectors)	: Amersham International plc, Amersham North America, Boehringer Mannheim Biochemicals (BCL in the UK), CLONTECH Laboratories Inc. (UK distributors, Cambridge BioScience), Gibco-BRL Life Technologies, Invitrogen Corporation (UK distributor, R&D Systems Europe Ltd), Promega, Stratagene
เอนไซม์และสารเคมี (Enzymes and biochemicals reagents	: Amersham International plc, Amersham North America, Boehringer Mannheim Biochemicals (BCL in the UK), CLONTECH Laboratories Inc. (UK distributors, Cambridge BioScience), Gibco-BRL Life Technologies, New England Biolabs (UK distributor, CP Laboratories), Promega, Strtagene
วัสดุวิทยาศาสตร์ต่างๆ (Consumables)	: Costar UK Ltd, Eppendorf, Gilson Inc., Midwest Scientific, P.E. Applied Biosystems, Stratagene, Techne
เครื่องแยกดีเอ็นเอ/โปรตีนด้วยกระแสไฟฟ้า (Gel electrophoresis equipments)	: Appligene, Gibco-BRL Life Technologies, Hoefer Pharmacia Biotech Inc., Bio-Rad Laboratories Ltd, New England Biolabs (UK distributor, CP Laboratories), Stratagene

สังเคราะห์โอลิโกนิวคลีโอไทด์ (Oligonucleotide synthesis)	: Appligene, Genosys Biotechnologies, CLONTECH Laboratories Inc. (UK distributor Cambridge BioScience), Operon, P.E. Applied Biosystms, R&D Systems Europe Ltd, BSU (National Center for Genetic Engineering and Biotechnology, Thailand)
เครื่องเพิ่มปริมาณสารพันธุกรรม (Thermal Cyclers)	: Appligene, Biometra, Ericomp Ltd, Grant Instruments (Cambridge) Ltd, Hybaid, P.E. Applied Biosystems, M.J. Research Inc., Sanyo (UK distributor, Sanyo Gallenkamp plc, Stratagene, Techne
Thermostable DNA Polymerases	: Amersham, Appligene, Boehringer Mannheim (Diagnostics and Biochemicals), CLONTECH, Flowgen Instruments Ltd, Gibco-BRL Life Technological, Hoefer Pharmacia Biotech Inc., Molecular Genetic Resources Inc., NBL Gene Sciences, New England Biolabs (UK distributor, CP Laboratories), P.E. Applied Biosystems, Pharmacia, Promega, Stratagene, Takara Shuzo Co. (UK distributor, Stratech Scientific Ltd; USA distributor, Pan Vera Corp.)
UV transilluminators	: Amersham, Appligene, Fotodyne Inc. (UK distributor, TechGen International Ltd), Hoefer Pharmacia Biotech Inc.

9. อธิบายคำศัพท์

Accession Number ตัวบ่งชี้เฉพาะที่ถูกกำหนดขึ้นเพื่อให้เข้าถึงข้อมูลลำดับเบสดีเอ็นเอหรือโปรตีนหนึ่ง ๆ ในฐานข้อมูล

Adenine เบสชนิดหนึ่งจากทั้งหมด 4 ชนิดที่ประกอบเข้ากันเป็นสายดีเอ็นเอ ตัวย่อของเบสอะดินีนคือ A ส่วนเบสตัวอื่นที่เหลือคือ กวานีน (Guanine=G), ไซโตซีน (Cytosine=C) และ ไธมีน (Thymine=T) เบสอะดีนีนจะจับคู่กับไธมีนเสมอ

Alignment หมายถึงกระบวนการในการเปรียบเทียบลำดับเบส 2 ชุดหรือมากกว่า โดยการตรวจหาลักษณะที่เหมือนกันในลำดับเบสนั้นๆ การเปรียบเทียบลำดับเบสทำได้ 2 ลักษณะคือ การเปรียบเทียบลำดับเบสบางส่วนและการเปรียบเทียบลำดับเบสทั้งหมด โดยทั่วไปมักใช้การเปรียบเทียบเฉพาะที่มากกว่า

Alignment along a tree เป็นการเปรียบเทียบสายลำดับเบสมากกว่า 2 ชุด (multiple alignment) โดยจะเปรียบเทียบกลุ่ม (cluster) ของสายลำดับที่เกี่ยวข้องก่อน ตามด้วยการเปรียบเทียบข้อมูลอื่นของกลุ่ม กลุ่มที่ได้จะเกิดจากสายลำดับที่อยู่ในกลุ่มสาขาเดียวกันในแผนภูมิอนุกรมวิธาน (Hierarchical alignment)

Alignment score ค่าคะแนนที่เครื่องคอมพิวเตอร์คำนวณได้โดยใช้อัลกอริธึ่มในการเปรียบเทียบสายลำดับ ซึ่งมีค่าขึ้นกับจำนวนลำดับเบสหรือกรดอะมิโนที่ตรงกัน (match) จำนวนเบสที่เกิดการแทนที่ (substitution), จำนวนเบสที่เกิดการแทรก (insertion) และจำนวนเบสที่ถูกเอาออก (deletion) เกิดเป็นช่องว่าง (gaps) ค่าคะแนนของลำดับที่ตรงกันกับลำดับที่เกิดการแทนที่ได้มาจากเมทริกซ์ของคะแนน (scoring matrix) เช่น ในกรณีการเปรียบเทียบสายลำดับโปรตีน จะใช้เมทริกซ์บลอสซัม (BLOSUM) หรือเมทริกซ์แพม (PAM) และเลือกใช้การหักคะแนนเมื่อมีการเปิดช่องว่างที่เหมาะสมกับเมทริกซ์ที่ถูกเลือก ค่าคะแนนของการเปรียบเทียบลำดับเบสจะอยู่ในหน่วย log odds มักอยู่ในรูป bit units (log ฐานสอง) ค่าคะแนนสูงแสดงว่าสายลำดับที่นำมาเปรียบเทียบมีความคล้ายกันมาก

Alphabet อักษรที่ใช้เป็นสัญลักษณ์แทนเบสในลำดับดีเอ็นเอมี 4 ตัว คือ A T G และ C และอักษรที่ใช้เป็นสัญลักษณ์แทนกรดอะมิโนในลำดับโปรตีนมี 20 ตัว คือ A, B, C, D, E, F, G, H, I, K, L, M, N, P , Q, R, S, T, U, V, W และ Y

Analogous เป็นคำศัพท์ทางด้านการศึกษาระบบชาติพันธุ์ (phylogenetics) หมายถึงคุณลักษณะที่สืบทอดมาจากบรรพบุรุษที่ไม่มีความเกี่ยวข้องกัน แล้วมีแนวโน้มที่จะเข้าใกล้หรือคล้ายกัน (convergent)

Annotation การทำนายตำแหน่งของยีน (genes) ในจีโนม (genome) ได้แก่ ตำแหน่งของสายลำดับที่มียีนซึ่งบรรจุรหัสสำหรับการสร้างโปรตีน ลำดับกรดอะมิโนในโปรตีนและการเข้าคู่กันได้อย่างมีนัยสำคัญกับโปรตีนอื่นที่ทราบหน้าที่แล้ว รวมทั้งการทำนาย

ตำแหน่งยีนที่ถอดรหัสเป็นอาร์เอ็นเอด้วย การทำนายจะใช้แบบจำลองของยีน เช่น hidden Markov models (HMM) ของ introns และ exons ในยีนที่ถอดรหัสเป็นโปรตีน และแบบจำลองโครงสร้างทุติยภูมิของอาร์เอ็นเอ

Array assay การจัดเรียงลำดับนิวคลีโอไทด์ของ cDNA หรือนิวคลีโอไทด์สายสั้นๆ (oligonucleotide) ไว้บนพื้นผิวของวัสดุรองรับ โดย cDNA หรือ oligonucleotide จะจับเข้ากับดีเอ็นเอเป้าหมาย ซึ่งสามารถตรวจหาลำดับเบสที่เข้าคู่กันได้โดยการวัดการเรืองแสง

Base pairs (bp) ไนโตรเจนเบส 2 ตัวที่จับกันด้วยพันธะที่มีแรงอ่อนๆ (ได้แก่ อะดินีนจับกับไธมีน หรือ กวานีนจับกับไซโตซีน) สายดีเอ็นเอ 2 สายจะจับกันเป็นเกลียวคู่ (double-helix) ด้วยพันธะไฮโดรเจนระหว่างคู่เบสเหล่านี้ จีโนมของมนุษย์ประกอบไปด้วยคู่เบสประมาณ 3 พันล้านคู่ มักเรียก 1 ล้านคู่เบสว่า 1 เมกะเบส (1Mb) และเรียก 1 พันคู่เบสว่า 1 กิโลเบส (1 kb)

Biochips เป็นการจัดเรียงโมเลกุลชีวภาพจำนวนมากลงบนพื้นผิววัสดุขนาดเล็กบางอย่าง (เช่น กระจก) อย่างเป็นระเบียบ โดยเฉพาะนิวคลีโอไทด์สายสั้นๆ โดยถูกจัดเป็นแถวในรูปแบบที่กำหนด มักเรียกว่า DNA microarrays และ Biochips.

Bioinformatics (ชีวสารสนเทศ) เป็นสหวิชาที่รวมเอาศาสตร์ทางชีววิทยา วิทยาศาสตร์คอมพิวเตอร์ คณิตศาสตร์ และสถิติ มาวิเคราะห์ลำดับข้อมูลทางชีววิทยา ส่วนประกอบและการจัดเรียงตัวของยีนและจีโนม รวมถึงทำนายโครงสร้างและหน้าที่ของโมเลกุลขนาดใหญ่

BLAST ย่อมาจาก Basic Local Alignment Search Tool เป็นเทคนิคที่สะดวกรวดเร็วที่ใช้ในการเปรียบเทียบบางส่วนของสายลำดับที่ไม่มีช่องว่าง กับสายลำดับที่สนใจศึกษาเพื่อหาความตรงกัน

Block ส่วนของสายลำดับโปรตีนที่พบว่ารูปแบบ (patterns) ที่พบเหมือนๆ กันเมื่อเปรียบเทียบระหว่างสายลำดับกรดอะมิโนของโปรตีนที่มีความเกี่ยวข้องกันกลุ่มหนึ่ง โดยส่วนดังกล่าวนี้ไม่มีช่องว่าง และประกอบด้วยกรดอะมิโนความยาวประมาณ 3-60 ตำแหน่ง

Browser โปรแกรมที่ใช้เข้าสู่เวบไซต์ต่าง ๆ ใน World Wide Web โดยมีภาษาเอชทีเอ็มแอล (HTML) เป็นตัวช่วยให้ browser แสดงผลของ webpage โดยอิสระจากระบบที่คอมพิวเตอร์ใช้อยู่

cDNA ย่อมาจาก complementary DNA เป็นดีเอ็นเอที่เข้าคู่กับ mRNA (messenger RNA) ซึ่งสามารถนำมาเพิ่มจำนวนและหาลำดับเบสได้ง่าย

cDNA library "ห้องสมุด cDNA" แหล่งรวบรวมลำดับดีเอ็นเอที่สร้างจาก mRNA ซึ่งมีดีเอ็นเอหรือยีนเฉพาะที่มีรหัสสำหรับสร้างโปรตีน (exon) เท่านั้น ไม่นำส่วนที่ไม่มีรหัสโปรตีน (intron) มารวมอยู่ด้วย

Cell หน่วยย่อยที่สุดของสิ่งมีชีวิตที่สามารถดำรงชีวิตได้อย่างอิสระ

Characters and character states ในการศึกษาระบบชาติพันธุ์ หรือวิวัฒนาการ (phylogenetics) คุณสมบัติ (character) หมายถึงลักษณะที่เหมือนกันในสิ่งมีชีวิตต่างชนิดกัน ส่วนสภาพที่แท้จริงของลักษณะหนึ่งๆ ในสิ่งมีชีวิตนั้นจะเรียกว่า สถานสมบัติ (character state) ยกตัวอย่างเช่น คุณสมบัติ "สีขน" อาจมีสถานสมบัติเป็น "สีทอง" "สีแดง" หรือ "สีเหลือง" หรือในระดับอณูชีววิทยา สถานสมบัติอาจหมายถึงเบส 1 ใน 4 ชนิด (A, C, T, G) หรือกรดอะมิโน 1 ใน 20 ชนิดเป็นต้น บางท่านอาจให้ความหมายของคุณสมบัติเหมือนกับสถานะสมบัติ

Chromosome โครงสร้างของดีเอ็นเอและโปรตีนที่พบในนิวเคลียสของเซลล์ โครโมโซมแต่ละอันประกอบด้วยยีนเป็นร้อยเป็นพันยีน ซึ่งสืบทอดมาจากบรรพบุรุษ มนุษย์มีโครโมโซมทั้งหมด 23 คู่ (ในแต่ละคู่ได้มาจากมาจากพ่อและแม่อย่างละอัน) ประกอบด้วยยีนทั้งหมด 50,000-100,000 ยีน (จากการหาลำดับเบสของจีโนมมนุษย์พบว่ามีประมาณ 32,000 ยีน)

Client คอมพิวเตอร์หรือโปรแกรมที่ทำงานบนคอมพิวเตอร์ ซึ่งมีปฏิสัมพันธ์กับคอมพิวเตอร์อื่นที่อยู่ไกลออกไป (server) client แตกต่างจาก user

Clone ID รหัสประจำหรือตัวเลขจำเพาะที่ใช้ในการระบุโคลนของ LifeSeq โดยใน 1 clone ID อาจมี sequence ID ที่เกี่ยวข้องมากกว่า 1 อันก็ได้

Clones กลุ่มเซลล์ที่ได้จากบรรพบุรุษหนึ่งๆ (single ancestor) ซึ่งจะมีลักษณะเหมือนต้นแบบ

Cloning vector โมเลกุลของดีเอ็นเอที่ได้มาจากยีสต์ แบคทีเรีย ไวรัส พลาสมิด คอสมิด (cosmid) หรือ เฟจ (phage) และมีชิ้นส่วนดีเอ็นเอแปลกปลอมแทรกอยู่ สามารถนำเข้าสู่ (transform) เซลล์เจ้าบ้าน (host cells) และเพิ่มจำนวนในเซลล์เจ้าบ้านได้

Clustal alignment เป็นการเปรียบเทียบสายลำดับหลายสายพร้อมกัน (multiple alignment) โดยใช้ชุดโปรแกรม Clustal โดยทำการเปรียบเทียบไล่ไปตามแขนงของแผนภูมิ (alignment along a tree) แผนภูมิจะถูกสร้างขึ้นจากคะแนนที่ได้จากการเปรียบเทียบสายลำดับทีละคู่ (pairwise alignment)

Cluster กลุ่มโคลนที่เกี่ยวข้องกันโดยมีความคล้ายคลึงกันของสายลำดับ (sequence homology) แต่ละ cluster จะมีรหัสประจำ (cluster ID number) เฉพาะเป็นของตัวเองภายใต้เงื่อนไขหนึ่ง

Cluster analysis วิธีการที่ใช้จัดกลุ่มวัตถุที่มีความเกี่ยวข้องสัมพันธ์กันให้อยู่ในกลุ่มเดียวกัน ความสัมพันธ์ขึ้นกับเกณฑ์ที่ใช้ในการพิจารณาความคล้ายคลึงหรือความแตกต่าง ในการวิเคราะห์สายลำดับนั้นจะใช้คะแนนความคล้ายคลึง (similarity score) หรือคะแนนความแตกต่าง (distance score) หรือผลจากการประเมินทางสถิติของคะแนนเหล่านี้

Codon ลำดับเบสของดีเอ็นเอ 3 ตัวที่แปลรหัสได้กรดอะมิโน 1 ตัว

Codon usage การวิเคราะห์ codon ที่มักใช้เฉพาะในยีน หรือสิ่งมีชีวิตชนิดหนึ่ง

Comparative genomics การเปรียบเทียบจำนวน ตำแหน่งและหน้าที่ทางชีวภาพของยีนในจีโนมของสิ่งมีชีวิตต่างๆ จุดประสงค์เพื่อบ่งชี้กลุ่มของยีนที่ทำหน้าที่ทางชีวภาพอันเป็นลักษณะเฉพาะในสิ่งมีชีวิตหนึ่งๆ

Consensus สายลำดับเส้นหนึ่งที่แสดงความผันแปร ณ คอลัมน์ใดๆ ซึ่งได้จากการเปรียบเทียบสายลำดับพร้อมกันหลายเส้น multiple sequence alignment.

Cosmid cloning vector เป็นเวกเตอร์ที่ประดิษฐ์ขึ้น (artificial cloning vector) ประกอบด้วย cos gene ของ lambda phage (ซึ่งใช้ในการ infect *E. coli*) คอสมิดสามารถรับชิ้นส่วนดีเอ็นเอที่มีความยาวถึง 45 กิโลเบสเข้าไปได้ ซึ่งชิ้นส่วนดีเอ็นเอที่แทรกไปในคอสมิดนี้ยาวกว่าที่ใช้กับพลาสมิด

Cytosine เบสชนิดหนึ่งจากทั้งหมด 4 ชนิดที่ประกอบกันเป็นสายดีเอ็นเอ (A T C G) ตัวย่อของไซโตซีน คือ C ส่วนตัวอื่นคือ อะดีนีน กวานีน และไธมีน เบสไซโตซีนจับคู่กับกวานีน

Database แหล่งเก็บรวบรวมข้อมูลบนคอมพิวเตอร์ที่สามารถเรียบเรียง เพิ่ม ลบ และเปลี่ยนแปลงแก้ไขข้อมูลได้ตามวิธีมาตรฐาน

dbEST ข้อมูลสายลำดับรวมทั้งข้อมูลอื่นๆ ของ cDNA ที่ได้จากการแปล mRNA ของสิ่งมีชีวิตชนิดต่างๆ แบบ single-pass ดู ETS.

dbGSS ข้อมูลจีโนมแบบ single-pass, สายลำดับในระดับจีโนมที่มีแต่ exon (exon-trapped genomic sequences) และสายลำดับจาก PCR (Alu PCR sequences) GSS มีความคล้ายคลึงกับ dbEST แต่แตกต่างกันตรงที่สายลำดับของ GSS เป็นจีโนม ส่วนใน dbEST เป็น cDNA ดู GSS.

dbSNP ศูนย์กลางเก็บรวบรวมข้อมูลยีนที่มีลักษณะ polymorphism ซึ่งเกิดการแทนที่ของนิวคลีโอไทด์เพียงหนึ่งตำแหน่ง การเกิด deletion และการเกิด insertion สายสั้นๆ ดู SNP.

dbSTS ข้อมูลสายลำดับเบสและข้อมูลแผนที่บนสายลำดับจีโนมสายสั้นๆ ที่มีตำแหน่งสำคัญ (landmarks) แสดงอยู่ด้วย ดูที่ STS.

Deletion การกลายพันธุ์รูปแบบหนึ่ง โดยเกิดจากการที่ชิ้นส่วนดีเอ็นเอหายไปจากโครโมโซม การหายไปของยีนหรือบางส่วนของยีนซึ่งอาจทำให้เกิดโรคหรือเกิดความผิดปกติต่างๆ ได้

Dendogram แผนภูมิกิ่งที่ใช้แสดงวัตถุที่ถูกนำมาเปรียบเทียบในแนวตั้ง (เช่น สายลำดับหรือยีนในการทำ microarray analysis) วัตถุที่สัมพันธ์กันจะถูกเชื่อมโยงกันด้วยกิ่งที่แตกออกไปและขนานไปทางด้านข้างของวัตถุนั้น

Diploid จำนวนโครโมโซมในเซลล์ทั่วไปยกเว้นเซลล์สืบพันธุ์ ในมนุษย์มีจำนวน diploid เท่ากับ 46

Distance in sequence analysis จำนวนการเปลี่ยนแปลงที่สังเกตได้ในการเปรียบเทียบสายลำดับ 2 เส้น โดยไม่นับรวมช่องว่างที่อยู่ในสายลำดับนั้น

Distance matrix ตารางคะแนนที่ได้จากการเปรียบเทียบสายลำดับเป็นคู่ที่เหมาะสม (optimal pairwise alignment - เรียกอีกอย่างหนึ่งว่า edit distance) สนามเมทริกซ์ (i, j) ประกอบด้วยค่าคะแนนการเปรียบเทียบลำดับเบสที่เหมาะสมของ 2 ตัวอักษรในสายลำดับทั้ง 2 เส้นไปเรื่อยๆ จนถึงตำแหน่ง i และ j และทำการคำนวณจากตำแหน่งใกล้เคียงกันที่อยู่ด้านซ้ายบนโดยใช้สมการ recursive equation (Dynamic Programming Matrix)

Distance measure คือฟังก์ชันที่เชื่อมโยงค่าตัวเลขที่ไม่เป็นลบกับคู่ของสายลำดับเข้าด้วยกัน โดยถือว่าที่ระยะที่สั้นกว่าหมายความว่ามีความคล้ายกันมากกว่าปรกติ distance measures จะสอดคล้องกับกฎทางคณิตศาสตร์ของเมทริกซ์ distance measure เป็นฟังก์ชันที่ใช้ในการให้คะแนนชนิดหนึ่ง

DNA Deoxyribonulceic acid หรือโมเลกุลที่เป็นสาย 2 สายเชื่อมกันด้วยพันธะอ่อนๆ ระหว่างคู่เบสที่แตกต่างกัน 4 ชนิด (A, T, C, G) DNA เป็นแหล่งเก็บข้อมูลพันธุกรรมสำหรับการเจริญเติบโต การพัฒนา และการจำลองแบบ (replication)

DNA chip นิวคลีโอไทด์สายสั้นๆ ที่ถูกพิมพ์อยู่บนผิวของแข็งอย่างเป็นระเบียบ ใช้เพื่อบ่งชี้สายลำดับดีเอ็นเอที่อยู่ติดฉลากด้วยสารเรืองแสง

DNA replication กระบวนการที่เกลียวคู่ของสายดีเอ็นเอคลายออก และเกิดการจำลองตัวเองขึ้นมา ซึ่งเหมือนเดิมทุกประการ

DNA sequencing การหาลำดับเบสที่ถูกต้องในชิ้นส่วนดีเอ็นเอ

Domain name หมายถึงระดับหนึ่งในหลายๆ ระดับของการจัดองค์กรในเครือข่ายอินเตอร์เน็ต มีไว้เพื่อแยกและระบุเครื่องที่ทำหน้าที่เป็นเจ้าบ้านหรือโฮสต์ (host machines) ส่วนชื่อโดเมนในระดับสูงสุด (top level domain name) จะใช้บ่งชี้ชนิดของไซต์หรือประเทศที่โฮสต์นั้นตั้งอยู่

Dot matrix แผนภาพ dot matrix แสดงวิธีการที่ใช้ในการเปรียบเทียบสายลำดับ 2 เส้น สายลำดับเส้นแรกจะถูกเขียนในแนวนอนที่ส่วนบนของกราฟ ส่วนอีกเส้นจะเขียนในแนวตั้งทางซ้ายมือของกราฟ จุดในกราฟแสดงตำแหน่งที่ทั้งสายลำดับ 2 เส้นนั้นมีอักษรเหมือนกัน เส้นทแยงมุมในกราฟแสดงถึงส่วนที่เหมือน หรือคล้ายหรือเข้ากันได้ (alignment) ตารางที่ได้อาจถูกนำไปผ่านการกรองเพื่อหาบริเวณที่มีความคล้ายคลึงกันมากที่สุด โดยกำหนดค่าขั้นต่ำที่ตำแหน่งบนสายลำดับสองเส้นจะต้องเหมือนกันภายในความยาวหรือขอบเขตหน้าต่างที่กำหนด

Double helix การจัดเรียงโครงสร้างดีเอ็นเอที่ดูคล้ายกับบันไดลิงยาวๆ ที่บิดเป็นเกลียวหรือเป็นขด ด้านข้างของ "บันได" ประกอบด้วยโมเลกุลน้ำตาลและฟอสเฟตรวมตัวเป็น backbone ส่วนขั้นบันไดเป็นเบสที่จับกันด้วยแรงพันธะอ่อนๆ ที่เรียกว่าพันธะไฮโดรเจน เบส น้ำตาล และฟอสเฟต รวมตัวกันเป็นนิวคลีโอไทด์ ดังนั้นนิวคลีโอไทด์หลายๆ หน่วยจึงต่อกันเป็นสาย และ 2 สายจับคู่กันเป็นดีเอ็นเอ

Download การส่งถ่ายข้อมูลจากคอมพิวเตอร์โฮสต์หรือเจ้าบ้านที่อยู่ไกลออกไป ไปยังเครื่องคอมพิวเตอร์ที่ใช้งานอยู่ โดยมากจะส่งผ่านทาง FTP.

Duplication การกลายพันธุ์รูปอย่างหนึ่ง เกิดจากการมีชิ้นส่วนดีเอ็นเอที่ซ้ำกันเพิ่มขึ้น 1 ส่วนหรือมากกว่า ซึ่งรวมถึงยีนและโครโมโซมทั้งหมด

e-mail ย่อมาจาก Electronic mail หรือจดหมายอิเล็กทรอนิกส์ หมายถึงข้อความที่สร้างขึ้นบนคอมพิวเตอร์และส่งผ่านอินเทอร์เน็ตไปยังที่ไกลออกไปได้ภายในระยะเวลาไม่กี่วินาที (คำตรงข้าม: snail mail, postal mail)

E-value หมายถึง จำนวนความเหมือน (hits) ที่คาดหวังว่าจะพบด้วยความบังเอิญ ในการสืบค้นฐานข้อมูล และได้ค่าคะแนนหนึ่งๆ หรือดีกว่า ค่า E-value ขึ้นกับขนาดของฐานข้อมูลที่สืบค้น ค่า E-value ยิ่งต่ำคะแนนที่ได้ยิ่งมีความน่าเชื่อถือ ค้นเพิ่มเติมที่ P-value.

EBI ย่อมาจาก European Bioinformatics Institute (เครือข่ายของ EMBL) ข้อมูลเพิ่มเติมติดต่อที่http://www.ebi.ac.uk/.

EMBL ห้องปฏิบัติการทางชีววิทยาโมเลกุลแห่งยุโรป (European Molecular Biology Laboratory) ตั้งอยู่ที่ไฮเดลเบิร์กประเทศเยอรมัน ข้อมูลเพิ่มเติม http://www.embl-heidelberg.de/.

EMBnet เครือข่ายชีววิทยาโมเลกุลแห่งยุโรป (European Molecular Biology network)

Entrez แหล่งข้อมูลออนไลน์ที่จัดทำขึ้นโดยศูนย์ข้อมูลทางเทคโนโลยีชีวภาพแห่งชาติ (NCBI) ซึ่งมีฐานข้อมูลสายลำดับของ GenBank ที่สามารถเชื่อมโยงข้อมูลเหล่านั้นไปยังเอกสารที่ตีพิมพ์ข้อมูลของยีนที่สืบค้นได้

Enzyme โปรตีนที่ทำหน้าที่เร่งปฏิกิริยาชีวเคมี ปกติมักทำให้อัตราเร็วของปฏิกิริยาเพิ่มขึ้น สิ่งมีชีวิตไม่สามารถดำรงอยู่ได้หากไม่มีเอนไซม์

EST Expressed Sequence Tag (ESTs) คือสายลำดับเบสสายสั้นๆ มีความยาวประมาณ 300-500 คู่เบส ที่แปลมาจาก mRNA แล้วได้สายลำดับ cDNA ซึ่งมักมีจำนวนมากๆ EST แสดงให้เห็นการแสดงออกของยีนต่างๆ ณ เวลาหนึ่งๆ ในเนื้อเยื่อที่ศึกษา หรือในช่วงใดช่วงหนึ่งของการเจริญเติบโตและพัฒนาการ EST จะมีฉลาก (tags) ของยีนกำกับ รวมทั้งอาจมีฉลากอื่นๆ ที่ไม่เกี่ยวกับยีนด้วย บันทึกแบบนี้มีข้อมูลประกอบน้อยและมีเฉพาะข้อมูลทางด้าน library และ biosource เท่านั้น สามารถสืบค้น EST ได้จากฐานข้อมูหลายแห่ง เช่น DDBJ/EMBL/GenBank, dbEST และ Unigene หาข้อมูลเพิ่มเติมได้ที่ http://www.ncbi.nlm.nih.gov/dbEST// หรือดูที่ Expressed sequence tag

Exons ส่วนของยีนที่มีรหัสสำหรับการสร้างเป็นโปรตีน แต่ละ exon มีรหัสซึ่งจะถูกแปลเป็นส่วนหนึ่งของโปรตีน ในสิ่งมีชีวิตบางสปีชีส์ (รวมทั้งมนุษย์) exon ต่างๆ ของยีนจะถูก

แยกออกจากกันโดยดีเอ็นเอสายยาวเรียกว่า introns หรือที่บางครั้งเรียกว่า ดีเอ็นเอขยะ (junk DNA) ซึ่งยังไม่ทราบหน้าที่ที่แน่ชัด ดูที่ introns

Expressed sequence tag (EST) ข้อมูลยีนบางส่วนของ cDNA clone ซึ่งมีข้อมูลเพิ่มเติมของสายลำดับ (sequence tag) ของยีนอยู่ด้วย โดยข้อมูลของ EST ได้มาจากการหาสายลำดับแบบครั้งเดียว (single-pass sequencing) เพื่อให้ได้ผลรวดเร็ว ดังนั้นจึงอาจพบอัตราการเกิดความผิดพลาดในสายลำดับนี้สูงมาก บางครั้งสูงถึง 5 %

FAQ แฟ้มข้อมูลคอมพิวเตอร์ที่บันทึกคำถามที่มักถูกถามบ่อยๆ โดยรวบรวมทั้งคำถามและคำตอบเพื่ออำนวยความสะดวกให้แก่ผู้ใช้บริการรายใหม่ของแหล่งข้อมูลนั้นๆ เช่น mailing list หรือ newsgroup

False negative ผลลบที่รายงานไม่ถูกต้องเนื่องจากความผิดพลาดของการทดลอง ถ้าการวัดผลถูกต้อง ข้อมูลที่ได้ต้องเป็นผลบวก

False positive ผลบวกที่รายงานไม่ถูกต้องเนื่องจากความผิดพลาดของการทดลอง ถ้าการวัดผลถูกต้อง ข้อมูลที่ได้ต้องเป็นผลลบ

FASTA โปรแกรมที่ใช้ในการสืบค้นข้อมูลในฐานข้อมูล เพื่อเปรียบเทียบลำดับนิวคลีโอไทด์หรือเปปไทด์ที่มีอยู่กับลำดับเบสหรือนิวคลีโอไทด์ที่มีอยู่ในฐานข้อมูล โปรแกรมนี้พัฒนาขึ้นโดย Lipman และ Pearson โดยอาศัย rapid sequence algorithm

Feature ข้อมูล (annotation) ที่ปรากฏในตำแหน่งที่แน่นอนในสายลำดับหนึ่งๆ

Filtering (window size) ในการเปรียบเทียบสายลำดับทีละคู่โดยใช้วิธี dot matrix การจับคู่แบบสุ่มจะถูกกรองออกโดยใช้เลื่อน "หน้าต่าง" (ซึ่งกำหนดความยาวได้ว่าประกอบด้วยกี่ตัวอักษร) แล้วทำการเปรียบเทียบสายลำดับทั้ง 2 เส้นนั้นทีละหน้าต่าง แทนที่จะทำการเปรียบเทียบที่ละตัวอักษรหรือทีละตำแหน่ง หากตัวอักษรในหน้าต่างตรงกัน (match) จึงทำการจุด (dot) ลงบนกราฟ โดยสามารถกำหนดจำนวนตัวอักษรที่น้อยที่สุดในหน้าต่างที่ตรงกันจึงจะจุด หากน้อยกว่าที่กำหนด โปรแกรมจะไม่จุดให้ During pairwise sequence alignment using the dot matrix method, random matches can be filtered out by using a sliding window to compare the two sequences. Rather than

comparing a single sequence position at a time, a window of adjacent positions in the two sequences is compared and a dot, indicating a match, is generated only if a certain minimal number of matches occur.

Format (file) โปรแกรมต่างๆ ต้องการข้อมูลที่เหมาะสมในรูปแบบที่แตกต่างกัน โดยการใช้คำสำคัญ (keywords) และการจัดลำดับเฉพาะ ลักษณะเฉพาะนี้เรียกว่า file formatt.

Fourier analysis การศึกษาฟังก์ชันของการประมาณการ (approximation) และการเสื่อม (decomposition) โดยใช้ตรีโกณมิติพหุนาม

FTP (File Transfer Protocol) FTP ยินยอมให้บุคคลหนึ่งถ่ายโอนแฟ้มข้อมูลจากเครื่องคอมพิวเตอร์หนึ่งไปยังคอมพิวเตอร์อีกเครื่องหนึ่งโดยผ่านเครือข่าย โดยใช้โปรแกรมลูกข่าย FTP ซึ่งสามารถติดต่อกับเครื่องคอมพิวเตอร์ที่เป็นแม่ข่าย FTP เท่านั้น และแม่ข่ายจะสร้างส่วนข้อมูลเฉพาะของแฟ้มข้อมูลที่มีอยู่เพื่อจำกัดการเข้าถึง โดยที่เครื่องลูกข่ายต้องระบุชื่อผู้ใช้ (username) และรหัสผ่านเพื่อให้สามารถติดต่อแม่ข่ายได้

Functional genomics การประเมินหน้าที่ของยีน โดยการเปรียบเทียบระหว่างจีโนมสามารถตรวจสอบหน้าที่ของยีนใหม่ได้โดยทำให้เกิดการกลายพันธุ์หรือการผ่าเหล่า (mutation) ในยีนนั้น และตรวจหาผลฟีโนไทป์ (phenotype) ที่ผิดปกติไปในสิ่งมีชีวิตนั้น

Gap การเกิดความไม่เหมือนกันในการเปรียบเทียบสายลำดับ 2 เส้นที่เกิดจากการสอดแทรก (insertion) ในสายลำดับเส้นหนึ่ง หรือเกิดจากการสูญหาย (deletion) ในสายลำดับอีกเส้นหนึ่ง

Gap penalty คะแนนที่เป็นตัวเลขที่ได้จากการเปรียบเทียบสายลำดับเบส เป็นคะแนนที่ให้เพื่อเป็น "ค่าปรับ" (penalize) ในกรณีที่มีช่องว่าง (gaps) เกิดขึ้นในการทำการเปรียบเทียบสายลำดับ (sequence alignment) ค่า gap penalties มีผลต่อจำนวนของช่องว่างที่จะเกิดในการเปรียบเทียบสายลำดับซึ่งเป็นผลจากการคำนวณโดยใช้อัลกอริธึม โปรแกรมที่ใช้ในการเปรียบเทียบสายลำดับส่วนใหญ่จะใช้ค่า gap penalties ที่เหมาะสมกับตารางคะแนน scoring matrices

GCG Assembly โปรแกรมที่ใช้ GCG Fragment Assembly System ซึ่งสร้างโดย Genetics Computer Group, Inc. โปรแกรมนี้ใช้ในการประกอบชิ้นส่วนสายลำดับนิวคลีโอไทด์หลายๆ เส้นเข้าด้วยกันเป็น cluster และทำให้เห็นว่าชิ้นส่วนของสายลำดับนิวคลีโอไทด์เหล่านั้นมีส่วนคาบเกี่ยว (overlap) กันอย่างไร

GenBank ฐานข้อมูลสาธารณะเกี่ยวกับสายลำดับดีเอ็นเอ ที่จัดทำโดยศูนย์ข้อมูลทางเทคโนโลยีชีวภาพแห่งชาติ (NCBI) โดยเป็นส่วนหนึ่งของห้องสมุดแพทยศาสตร์แห่งชาติ (National Library of Medicine, NLM)

Gene สายลำดับดีเอ็นเอจำเพาะที่เก็บข้อมูลที่ใช้ในการสร้างโปรตีน (บางยีนอาจมีข้อมูลที่ใช้ในการสร้างโมเลกุลอาร์เอ็นเอ) ประมาณกันว่าจีโนมของมนุษย์มียีนอยู่ประมาณ 50,000-100,000 (จากการหาลำดับเบสของจีโนมมนุษย์พบว่ามีประมาณ 32,000 ยีน)

Gene amplification การเพิ่มจำนวนของสำเนาของชิ้นส่วนดีเอ็นเอจำเพาะ เซลล์เนื้องอกสามารถเพิ่มจำนวนชิ้นส่วนดีเอ็นเอได้เองตามธรรมชาติ ซึ่งสาเหตุอาจเกิดจากสัญญาณเซลล์หรือมีสิ่งแวดล้อมภายนอกเป็นตัวกระตุ้น

Gene expression กระบวนการที่ข้อมูลในยีนgeneถูกเปลี่ยนแปลงไปเป็นโปรตีนหรืออาร์เอ็นเอบางชนิด ในขั้นแรกลำดับดีเอ็นเอจะถูกถอดรหัสเป็นอาร์เอ็นเอ จากนั้นจึงถูกแปลรหัสเป็นโปรตีน

Gene family ยีน 2 ยีนหรือมากกว่าที่เกี่ยวข้องกันโดยมีวิวัฒนาการแบบเบี่ยงเบน (divergent evolution) ออกจากบรรพบุรุษเดียวกัน (common ancestor) โดยอาจเกิดจากการเปลี่ยนแปลงภายในยีนหรือมีการเพิ่มสำเนาซ้ำของยีนขึ้นอีกชุดหนึ่ง (duplicaition)

Gene mapping การตรวจหาตำแหน่งเชิงสัมพัทธ์ (relative position) ของยีนต่างๆ บนโครโมโซม และระยะทางระหว่างยีนเหล่านั้น

Gene pool ยีนทั้งหมดที่อยู่ในสิ่งมีชีวิตสปีชีส์หนึ่งๆ ณ เวลาหนึ่งๆ ทั้งนี้รวมถึงความหลากหลายที่มีอยู่ด้วย

Gene therapy การรักษาโรคโดยการใส่ยีนใหม่เข้าสู่เซลล์ ยีนใหม่นี้อาจถูกใช้แทนยีนเดิมที่ทำหน้าที่บกพร่องไป หรือช่วยในการปกป้องเซลล์ปกติจากผลของยา (เช่น ยาที่ใช้ในการรักษาโรคมะเร็ง ซึ่งอาจเป็นอันตรายต่อเซลล์ปกติ) หรือป้องกันเซลล์จากไวรัส เช่น ไวรัสเอดส์ เป็นต้น

Genetic algorithm อัลกอริธึมที่ใช้ในการสืบค้น ซึ่งได้รับแรงบันดาลใจจากหลักการเกี่ยวกับวิวัฒนาการ โดยคำตอบกลุ่มแรก (ในทางสถิติถือเป็นประชากร) จะถูกนำไปเข้ารหัสก่อน จากนั้นอัลกอริธึมจะทำการสืบค้นประชากรกลุ่มนี้โดยอาศัยค่าที่ตั้งไว้ก่อนสำหรับคำตอบแต่ละอย่าง จากนั้นจึงเลือกคำตอบที่เข้ากันได้ดีที่สุด ในกระบวนการนี้อาจทำให้ได้คำตอบใหม่ๆ โดยเกิดจากการไขว้ (crossover) หรือการกลายพันธุ์ตามที่กำหนดไว้ในคำตอบที่ถูกเข้ารหัส

Genetic code ข้อมูลในยีนที่ใช้กำหนดว่าเซลล์จะสร้างโปรตีนชนิดใดชนิดหนึ่งขึ้นได้อย่างไร รหัสในดีเอ็นเอประกอบด้วยตัวอักษร A, T, G และ C ซึ่งหมายถึงสารเคมีที่เรียกว่านิวคลีโอไทด์ มี 4 ชนิด ได้แก่ adenine, thymine, guanine และ cytosine ตามลำดับ นิวคลีโอไทด์เหล่านี้มาเรียงตัวต่อกันเป็นสายดีเอ็นเอ รหัสยีนแต่ละชนิดคล้ายกับ "คำ" ที่ประกอบด้วยตัวอักษร 3 ตัว เกิดจากการผสมผสานเบสทั้ง 4 ชนิดแตกต่างกันไป และรหัสเหล่านี้เป็นตัวกำหนดชนิดของกรดอะมิโนในเปปไทด์หรือโปรตีน และถูกใช้ในทุกขั้นตอนของการถอดรหัส

Genetic linkage data แผนที่แสดงความเชื่อมโยงด้านพันธุกรรม (genetic linkage map) บนโครโมโซมทำหน้าที่บ่งชี้ตำแหน่งเชิงสัมพัทธ์ (relative location) ของยีนที่รู้จักภายในสายลำดับดีเอ็นเอของโครโมโซม ยีนที่ถูกสืบทอดมาด้วยกันมักจะอยู่ในตำแหน่งใกล้เคียงกันบนโครโมโซมและถือว่ามความเชื่อมโยงกัน (linked)

Genetic map แผนที่ทางกายภาพ (physical map) ของยีนที่ประกอบด้วยตำแหน่งของชิ้นส่วนต่างๆ ของสายลำดับดีเอ็นเอที่ทราบหน้าที่การทำงาน

Genetic mapping เทคนิคที่ใช้บ่งชี้บริเวณต่างๆ บนจีโนมที่อาจจะมียีนที่เกี่ยวกับการก่อโรคอยู่ การทำแผนที่ยีนต้องอาศัยดีเอ็นเอตัวอย่าง ที่ได้มาจากสมาชิกในครอบครัวหลายๆ ครอบครัวที่เป็นโรคที่ต้องการศึกษา

Genetic marker ชิ้นส่วนดีเอ็นเอที่มีตำแหน่งทางกายภาพที่บ่งชี้ได้บนโครโมโซม

Genetic screening การศึกษากลุ่มประชากรเพื่อบ่งชี้กลุ่มย่อยที่มีความเสี่ยงสูงที่จะเป็นโรคหรือแพร่กระจายโรคทางพันธุกรรมได้

Genome ข้อมูลทางพันธุศาสตร์ทั้งหมดที่เก็บไว้ในนิวเคลียสและไมโตคอนเดรีย โดยถูกเก็บอยู่ในกลุ่มโครโมโซมแบบ haploid

Genomic library ห้องสมุดหรือกลุ่มของสายลำดับดีเอ็นเอที่ถูกโคลน (clone) เข้าไปในเวกเตอร์ (vectors) โดยประกอบด้วยจีโนมทั้งหมดของสิ่งมีชีวิตนั้นๆ

Genomics การหาลำดับเบส (sequencing) และการหาลักษณะเฉพาะ (characterization) ของจีโนม รวมทั้งการวิเคราะห์ความสัมพันธ์ระหว่างกิจกรรมของยีนกับหน้าที่ของเซลล์

Genotype ข้อมูลทางพันธุกรรมที่เป็นเอกลักษณ์เฉพาะในสิ่งมีชีวิตหนึ่งๆ

GI ย่อมาจาก GenBank Identifier หมายถึงตัวเลขจำเพาะที่ใช้กำหนดไว้กับสายลำดับโปรตีนและสายลำดับนิวคลีโอไทด์ในฐานข้อมูล GenBank

Gopher ระบบการกระจายเอกสารที่ยินยอมให้สืบค้นและแสดงผลเป็นข้อความ (text-based)

GSS ย่อมาจาก Genome Survey Sequences DDBJ/EMBL/GenBank จะทำงานคล้ายกับ EST แตกต่างกันที่สายลำดับเป็นจีโนมไม่ใช่ cDNA หรือ mRNA, GSS ประกอบด้วยข้อมูลดังนี้: random 'single-pass read' genome survey sequence; single-pass read จากปลายของ cosmid/BAC/YAC (อาจเป็นจำเพาะกับโครโมโซมหรือไม่ก็ได้); exon-trapped genomic sequences; Alu PCR sequences ข้อมูลเพิ่มเติมดู http://www.ncbi.nlm.nih.gov/dbGSS/

Guanine เบสชนิดหนึ่งจากทั้งหมด 4 ชนิดที่ประกอบเข้ากันเป็นสายดีเอ็นเอ (ATGC) ตัวย่อของเบส guanine คือ G ส่วนเบสตัวอื่นคือ adenine, cytosine, และ thymine

Haploid จำนวนโครโมโซมในอสุจิหรือเซลล์ไข่ ซึ่งมีครึ่งหนึ่งของจำนวน diploid

HGMP ย่อมาจาก Human Genome Mapping Project หรือโครงการวิจัยจีโนมมนุษย์ ตั้งอยู่ที่เคมบริดจ์ ประเทศอังกฤษ

High throughput screening เทคนิคในการวิเคราะห์แบบอัตโนมัติเพื่อสืบค้นหากิจกรรม (activity) ที่ต้องการในสารประกอบหลายชนิดจำนวนมากๆ

Highly conserved sequences สายลำดับดีเอ็นเอที่มีความเหมือนกันมากในสิ่งมีชีวิตที่แตกต่างกันหลายชนิด นักวิทยาศาสตร์เชื่อว่าลำดับดีเอ็นเอที่เหมือนกันในสิ่งมีชีวิตต่างชนิดกันนี้ คือส่วนของยีนที่ทำหน้าที่พื้นฐานที่สำคัญบางอย่างในการดำรงชีวิตในสิ่งมีชีวิตรูปแบบต่างๆ และโครงสร้างของยีนนี้ยังคงได้รับการอนุรักษ์ในสายวิวัฒนาการ โดยเกิดการเปลี่ยนแปลงmutationsเพียงเล็กน้อยเท่านั้น

Homolog ส่วนประกอบที่เหมือนกันในสิ่งมีชีวิต 2 ชนิด (เช่น ยีนที่มีลำดับเบสเหมือนกันมาก) ซึ่งทำให้เชื่อได้ว่าน่าจะได้รับการถ่ายทอดมาจากบรรพบุรุษร่วมของสิ่งมีชีวิต 2 ชนิดในสายวิวัฒนาการ

Homologous ในวิชาว่าด้วยการศึกษา phylogenetics นั้น homologous หมายถึงลักษณะเฉพาะที่เหมือนกันของสิ่งมีชีวิตที่ต่างกัน โดยเป็นการสืบทอดพันธุกรรมมาจากลักษณะเดียวกันในบรรพบุรุษ ส่วนในทางอณูชีววิทยามักแปลความหมายง่ายๆ ว่าเหมือนกัน โดยไม่ได้คำนึงถึงความสัมพันธ์ด้านพันธุกรรม

Homologous recombination การแลกเปลี่ยนชิ้นส่วนดีเอ็นเอDNA ในระหว่างการสร้างเซลล์ไข่และอสุจิ การเกิด recombination จะทำให้โครโมโซมมีการสับเปลี่ยนยีนหรือข้อมูลพันธุกรรม

Homoplasy ความเหมือนที่เกิดขึ้นอย่างเป็นอิสระต่อกัน และไม่ได้บ่งชี้ว่ามาจากต้นกำเนิดเดียวกันใน phylogenetic

Horizontal transfer การส่งถ่ายข้อมูลทางพันธุกรรมระหว่างสิ่งมีชีวิต 2 สปีชีส์ที่ต่างกัน ซึ่งตามปกติจะไม่มีการส่งถ่ายข้อมูลพันธุกรรมเดิม ดีเอ็นเอที่ถูกส่งผ่านมาจะอยู่ในจีโนม

ของเซลล์ผู้รับ และสามารถตรวจหาได้จากประวัติใหม่ๆ ด้าน phylogenetics รวมทั้ง codon content ซึ่งเปรียบเทียบกับจีโนมส่วนที่เหลือ

Host คอมพิวเตอร์ใดๆ ที่อยู่บนเครือข่ายอินเตอร์เน็ตซึ่งสามารถระบุที่อยู่ได้โดยตรงผ่าน IP address หรือเจ้าบ้าน

HSPs ย่อมาจาก High-Scoring Segment Pairs หรือชิ้นส่วนของสายลำดับ 2 เส้นที่มีความยาวเท่ากันเหมือนกัน เมื่อนำมาเปรียบเทียบเฉพาะบริเวณจะมีค่าคะแนนการเปรียบเทียบ alignment score เท่ากับหรือมากกว่าระดับต่ำสุด (threshold) ที่ตั้งไว้ (cutoff)

HT ฐานข้อมูลเชิงกายวิภาคของยีนที่แสดงออก (Expressed Gene Anatomy Database, EGAD) ของ TIGR ประกอบด้วยชุดของสายลำดับนิวคลีโอไทด์ที่ไม่ซ้ำกัน และเป็นตัวแทนของ mRNA ที่สมบูรณ์ (mature transcript-HTs) HTs ได้รับการระบุชื่อและหน้าที่ (annotated) โดยมีข้อมูลเกี่ยวกับหน้าที่ในระดับเซลล์และข้อมูลอื่นๆ ที่เกี่ยวข้อง Expressed Gene Anatomy Database (EGAD) contains a non-redundant set of nucleotide sequences that represent mature transcripts (HTs).The HTs are annotated with cellular roles and links to related information.

HTGS (HTG) ย่อมาจาก High-throughput genomes sequences (HTG คือ HTGS Division ใน DDBJ/EMBL/GenBank) ศูนย์วิจัยหาข้อมูลจีโนมหลายๆ แห่งทั่วโลกเริ่มทำการค้นหาลำดับเบสของมนุษย์ และยูคารีโอทชั้นสูงในระดับใหญ่ขึ้น เชื่อกันว่าฐานข้อมูลนี้จะช่วยในการรวบรวมข้อมูลลำดับเบส ที่ยังไม่เสร็จสมบูรณ์จากหน่วยงานต่างๆ เข้าด้วยกัน ซึ่งส่วนใหญ่แล้วจะเป็นลำดับเบส ที่มีจำนวนช่องว่างที่สำคัญเป็นจำนวนมาก มีความถูกต้องแม่นยำต่ำ และยังไม่ทราบหน้าที่ สายลำดับเหล่านี้ไม่สามารถผ่านมาตรฐานขั้นสูงสำหรับข้อมูลที่ DDBJ/EMBL/GenBank ตั้งไว้ ข้อมูลเพิ่มเติมดู http://www.ncbi.nlm.nih.gov/HTGS/

HTML ย่อมาจาก Hyper-text Markup Language เป็นภาษามาตรฐานและใช้ตัวอักษรเพื่อกำหนดformatของเอกสาร World Wide Web ไฟล์ HTML จะถูกแปลและแสดงผลผ่านโปรแกรมเว็บบราวเซอร์ (browsers)

Human artificial chromosome พาหะหรือเวกเตอร์ที่ใช้สำหรับส่งผ่านหรือสำหรับการแสดงออกของชิ้นส่วนดีเอ็นเอของมนุษย์ HACs ถูกสร้างขึ้นมาและทำหน้าที่คล้ายกับโครโมโซมของมนุษย์

Human Genome project โครงการวิจัยระดับนานาชาติเพื่อสืบหาตำแหน่งยีนแต่ละยีนของมนุษย์ให้ได้เป็นจีโนมที่สมบูรณ์และมีรายละเอียดที่สมบูรณ์

Hybridization การที่สายลำดับดีเอ็นเอหรืออาร์เอ็นเอสายเดี่ยวสั้นๆ เข้าคู่กับลำดับดีเอ็นเออีกเส้นหนึ่ง (complememtary DNA or RNA) ได้ Hybridization ที่ใช้ในการวิเคราะห์กรดนิวคลีอิคจะต้องมี probe ที่เข้าคู่กับลำดับเบสเป้าหมายในดีเอ็นเอตัวอย่างได้

Hybridization array technologies เทคโนโลยีและวิธีการสร้าง DNA chip หรือ biochip DNA หรือ biochips และเกี่ยวข้องกับ microarray assays

Hyperlink รูปภาพหรือตัวอักษรที่อยู่ในเอกสาร World Wide Web ซึ่งสามารถใช้เม้าส์เลือกและคลิกเพื่อนำพาผู้ใช้ไปยังส่วนอื่นของหน้าเว็บเพจเดียวกันหรือหน้าเว็บเพจอื่นก็ได้ โดยไม่ขึ้นกับสถานที่ที่ใช้งาน

Hypertext ตัวอักษรในเว็บเพจที่มีสีแตกต่างออกไปจากตัวอักษรทั่วไป หรือถูกขีดเส้นใต้จะทำหน้าที่เป็น hyperlink คือสามารถใช้เพื่อเชื่อมโยงไปยังจุดอื่นๆ ได้

Indel ตัวย่อของคำว่า "Insertion and deletion" ใช้ประยุกต์เข้ากับการเปรียบเทียบสายลำดับสายเส้นที่มีความยาวของบริเวณที่เปรียบเทียบแตกต่างกัน multiple alignment และไม่สามารถระบุได้ว่าความยาวลำดับเบสที่แตกต่างกันนั้นเกิดจาก insertions หรือ deletions

Information content (of a scoring matrix) การแสดงระดับการอนุรักษ์ของส่วนหนึ่งในสายลำดับ (sequence conservation) ในคอลัมน์ของเมทริกซ์คะแนน (scoring matrix) ที่แสดงการเปรียบเทียบสายลำดับที่เกี่ยวข้องกัน ซึ่งรวมถึงจำนวนคำถามที่ถูกถามเพื่อใช้ในการจับคู่ตำแหน่งที่เหมาะสมในคอลัมน์ของสายลำดับที่นำมาทดสอบ จำนวนที่เป็นไปได้ที่มากที่สุดสำหรับเบสคือ 2 และ 4.32 สำหรับโปรตีน (ลอการิธึมฐาน 2 ของจำนวนตัวอักษรที่เป็นไปได้ในสายลำดับ)

***in situ* hybridization** การจับคู่เบสของดีเอ็นเอกับโครโมโซมที่อยู่ช่วงเมตาเฟสบนสไลด์ของกล้องจุลทรรศน์

Insertion ลักษณะความผิดปกติของโครโมโซมแบบหนึ่ง เกิดจากการที่มีลำดับดีเอ็นเอแทรกเข้ามาในยีน และไปรบกวนลักษณะโครงสร้างและการทำหน้าที่ของยีนนั้น

Internet ระบบเครือข่ายเชื่อมโยงคอมพิวเตอร์ ที่ใช้สำหรับส่งแฟ้มข้อมูลหรือข้อความระหว่างกันได้ จัดเป็นเครือข่ายที่ซ้อนอยู่บนเครือข่ายคอมพิวเตอร์อีกชั้นหนึ่ง

Intranet เป็นการนำเอาเทคโนโลยีและโปรโตคอลทางอินเตอร์เน็ตมาใช้ในการสร้างเครือข่ายภายใน โดยไม่ได้ติดต่อกับอินเตอร์เน็ตและได้รับการป้องกันโดย firewall

Introns ลำดับดีเอ็นเอส่วนที่ไม่มีรหัสสำหรับแปลเป็นโปรตีน ลำดับดีเอ็นเอส่วนอื่นๆ ที่ไม่มีรหัสสำหรับโปรตีนคือ ลำดับเบสควบคุม (control sequences) และบริเวณที่คั่นระหว่าง 2 ยีน (intergenic region) ซึ่งยังไม่ทราบหน้าที่ดูเพิ่มเติมที่ Exons

Java ภาษาที่ใช้ในการเขียนโปรแกรม พัฒนาขึ้นโดยบริษัท Sun Microsystems ที่ยินยอมให้โปรแกรมขนาดเล็ก (applets) ทำงานบนเครื่องคอมพิวเตอร์ใดๆ ก็ได้ Java applets จะเริ่มทำงานเมื่อกดเลือกบริเวณที่เป็น hyperlink บนเว็บเพจ

LAN Local Area Network เครือข่ายคอมพิวเตอร์ที่ครอบคลุมพื้นที่ขนาดเล็ก ใช้ติดต่อสื่อสารกันในพื้นที่จำกัด เช่น ภายในสำนักงาน ปีกใดปีกหนึ่งของตึก หรือระหว่างกลุ่มอาคารที่อยู่ใกล้เคียงกัน

Library แหล่งเก็บรวบรวมยีนที่แสดงออกในตัวอย่างเนื้อเยื่อจำเพาะ และรายละเอียดเพิ่มเติมของยีนเหล่านั้น

Ligand โมเลกุลที่ถูกโมเลกุลอื่นมาจับ มักจับกันด้วยพันธะที่ไม่ใช่พันธะโควาเลนท์ (non-covalent bonds)

Linear discriminant analysis การวิเคราะห์ที่มีการแบ่งแยกชุดข้อมูลออกเป็น 2 ชุด โดยอาศัยใช้เส้นตรงที่อยู่บนกราฟในตำแหน่งที่ดีที่สุดเป็นตัวแบ่ง

Linkage แนวโน้มของยีนหรือเครื่องหมายบอกตำแหน่ง (makers) ต่าง ๆ ที่ถูกสืบทอดมาด้วยกัน Linkage ของยีน 2 ยีนเครื่องบ่งชี้ว่ายีนเหล่านั้นอยู่ใกล้ซึ่งกันและกันในจีโนม

Linux ระบบปฏิบัติการ UNIX รูปแบบหนึ่งที่หาใช้ได้ฟรีแต่มีความแข็งแกร่งในเชิงพาณิชย์ และเหมาะกับนักชีวสารสนเทศมือใหม่ที่มีงบประมาณน้อย สามารถติดตั้งและใช้งานควบคู่กับโปรแกรมวินโดวส์บนคอมพิวเตอร์เครื่องเดียวกันได้

Local alignment เทคนิคการเปรียบเทียบสายลำดับ alignment ที่มักใช้ในกรณีที่สายลำดับเส้นแรก (s) สั้นกว่าสายลำดับเส้นที่ 2 (t) และพบว่าบางส่วนบนสายลำดับ t เหมือนหรือคล้ายกับสายลำดับ s มาก โดยเราไม่คิดคะแนนของการขาดหายไป (deletion) ของยีนส่วนที่อยู่ด้านหน้าและหลังของ t ได้ ในทางชีววิทยา คำว่า local alignment อาจหมายถึงการเปรียบเทียบความเหมือนเฉพาะที่ (local similarity - Approximately Matching)

Local similarity เทคนิคการเปรียบเทียบสายลำดับ alignment แบบหนึ่ง ที่ใช้สำหรับค้นหาเฉพาะส่วนของสายลำดับ 2 เส้นที่มีความคล้ายคลึงกันมากที่สุด ในทางชีววิทยามักหมายถึง local alignment

Locus เป็นการระบุตำแหน่งของยีนบนโครโมโซมรูปแบบหนึ่ง พหูพจน์ของ locus คือ loci

Marker เครื่องหมายซึ่งใช้ระบุบริเวณจำเพาะหนึ่งๆ ในจีโนม marker ทุกอันบน transcript map ของมนุษย์เป็น STSs

Maximum parsimony จำนวนขั้นตอนการเกิดวิวัฒนาการที่น้อยที่สุดที่สามารถทำให้เกิดความผันแปรในกลุ่มของสายลำดับที่กำลังทำการศึกษา คำนวณได้จากการเปรียบเทียบจำนวนขั้นตอนในแผนภูมิชาติพันธุ์ (phylogenetic trees) ที่เป็นไปได้ทั้งหมด

Minimum spanning tree หากกำหนดให้มีวัตถุชุดหนึ่งที่สัมพันธ์กัน อัลกอริธึมซึ่งถูกจำแนกออกจากกันโดยอาศัยค่าคะแนนความเหมือนกันหรือแตกต่างกัน วัตถุที่มีความคล้ายกันมากที่สุดจะอยู่ติดกัน ณ ตำแหน่งกิ่งนอกสุดของแผนภูมิ และวัตถุที่มีความเหมือนกันน้อยกว่าจะเรียงตามลำดับไปตามกิ่งด้านในของแผนภูมิ ความยาวของกิ่งในแผนภูมิต้นไม้คำนวณได้จาก heighbor-joining algorithm ซึ่งใช้สำหรับสร้างแผนภูมิชาติ

พันธุ์ของสายลำดับเบสจาก distance matrixdistance matrix ผลรวมของความยาวกิ่งที่ได้ระหว่างวัตถุแต่ละคู่จะมีค่าใกล้เคียงกับค่าที่ได้จากโครงสร้างการจำแนก (classification scheme)

Mitochondrial DNA สารพันธุกรรมของไมโตคอนเดรีย ซึ่งเป็นออร์กาเนลล์ที่สร้างพลังงานให้แก่เซลล์

Molecular clock สมมติฐานว่าด้วยการแทนที่ (substitutions) ของนิวคลีโอไทด์หรือกรดอะมิโนที่เกิดขึ้นด้วยอัตราคงที่ ในช่วงเวลาของการวิวัฒนาการ (fixed rat over evolutionary time) คล้ายกับเข็มนาฬิกาที่เคลื่อนที่อย่างช้าๆ หากกำหนดข้อมูลสำหรับการสอบเทียบ (calibration data) ร่วมกับค่าคงที่ของ molecular clock ให้ จะสามารถคำนวณหาเวลาที่ผ่านไปหลังจากโมเลกุล 2 โมเลกุลเบี่ยงเบนหรือแยกออกจากกันได้ โดยอาศัยปริมาณความแตกต่างกันของสายลำดับคู่ดังกล่าว

Molecular modeling การทำนายรูปทรง 3 มิติของโมเลกุลโดยการประเมินข้อมูลสายลำดับของรูปร่างโมเลกุลที่ทราบรูปทรงอยู่ก่อนแล้ว

Molecular evolution ปฏิกิริยาชีวเคมีพื้นฐานที่เกี่ยวข้องกับการเจริญของเซลล์และการแบ่งเซลล์เหมือนกันในเซลล์ทุกสปีชีส์ ฐานข้อมูลของวิถีชีวเคมี (biochemical pathway) ของสปีชีส์ที่สูญพันธุ์แล้วอาจถูกนำมาเปรียบเทียบกับสายลำดับดีเอ็นเอของสปีชีส์ที่ยังคงดำรงเผ่าพันธุ์อยู่ หรือในการตัดสินว่ามนุษย์มีความสัมพันธ์ทางชีวเคมีกับลิงกอริลลาและลิงชิมแปนซีหรือไม่ สามารถทำได้โดยการเปรียบเทียบสายลำดับดีเอ็นเอ

Monte Carlo A method that samples possible solutions to a complex problem as a way to estimate a more general solution.

mRNA (messenger RNA) เป็นยีนที่แสดงลักษณะทางพันธุกรรม ซึ่งจะถูกแปลรหัสเป็นโปรตีนต่อไป

MSA alignment เป็นการเปรียบเทียบที่คล้ายกับการเปรีบเทียบสายลำดับพร้อมกันหลายเส้นอย่างเหมาะสม (optimal multiple alignment) คำนวณโดยใช้ชุดโปรแกรม MSA ซึ่งอาศัยวิธีการคำนวณแบบ carrillo-lipman bound

Multiple alignment การเปรียบเทียบสายลำดับของโมเลกุลที่ได้จากสิ่งมีชีวิตพร้อมกัน 3 เส้นหรือมากกว่าโดยจัดเรียงเป็นรูปสี่เหลี่ยมผืนผ้า ซึ่งแต่ละแถวจะประกอบด้วยสายลำดับ 1 เส้นที่อาจมีช่องว่างแทรกอยู่ และมีการเน้นตำแหน่งที่เหมือนกัน (similarity) หรือการอนุรักษ์ (conservsation) การเปรียบเทียบสายลำดับพร้อมกับหลายเส้นอย่างเหมาะสมจะต้องให้ค่าคะแนนที่เหมาะสม หมายถึงระดับความเหมือนกันที่มากที่สุด หรือมีการสูญเสียต่ำที่สุด (Multiple sequence alignment)

Mutation การเปลี่ยนแปลงทางโครงสร้างของสายลำดับดีเอ็นเอ เกิดเป็นกลายพันธุ์หรือผ่าเหล่า

Mutation data matrix ตารางคะแนน scoring matrix ที่ได้จากรวบรวมผลการสังเกตุการกลายพันธุ์เฉพาะจุด (point mutation) ในการเปรียบเทียบสายลำดับ อาจรวมถึง Dayhoff PAM matrix ที่ให้ค่าคะแนนเป็น log odds scores

Mutation studies ใน Sequin ชุดของสายลำดับเบสของยีนเดียวกันในสิ่งมีชีวิตสปีชีส์เดียวกัน หรืออาจจะในสิ่งมีชีวิตตัวเดียวกัน ที่แยกได้จากการเหนี่ยวนำให้เกิดการกลายพันธุ์หลายๆ แบบ และถูกนำไปวิเคราะห์หาลำดับเบส

NCBI ย่อมาจาก National Center for Biotechnology Information ตั้งอยู่ ณ เมืองวอชิงตัน ดี ซี ประเทศสหรัฐอเมริกา

Neighbor-joining method การนำวัตถุที่คล้ายกับมาอยู่ในกลุ่มเดียวกัน (cluster) โดยเลือกจากกลุ่มวัตถุที่มีความเกี่ยวข้องกัน (เช่น ยีนที่มีสายลำดับคล้ายคลึงกัน)) เพื่อสร้างแผนภูมิต้นไม้ซึ่งแต่ละกิ่งจะสะท้อนความแตกต่างระหว่างวัตถุ (ในที่นี้คือยีนหรือสายลำดับต่างๆ ที่นำมาศึกษา)

NIH สถาบันสุขภาพแห่งชาติ (National Institute of Health) ของประเทศสหรัฐอเมริกา

Non-coding DNA สายดีเอ็นเอที่ไม่มีข้อมูลที่จำเป็นสำหรับการสร้างโปรตีน สาย non-coding นี้เป็นเสมือนภาพสะท้อนในกระจกเงาของสายดีเอ็นเอที่สามารถแปลรหัสเป็นโปรตีนได้ (coding strand) หรือที่เรียกว่า antisense strand

Nonsense mutation การกลายพันธุ์รูปแบบหนึ่งซึ่งมีแทนที่เบส (substitution) ในดีเอ็นเอเพียง 1 ตำแหน่ง เป็นผลทำให้เกิดรหัสหยุด (stop codon)

Normal distribution การกระจายตัวที่พบในข้อมูลหลายๆ แบบ เช่น น้ำหนักและขนาดร่างกาย และคะแนนสอบ ลักษณะการกระจายเป็นรูประฆังคว่ำ (bell-shaped curve) ซึ่งสามารถอธิบายได้โดยใช้ค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐานของค่าเฉลี่ย (standard deviation of mean) ค่าคะแนนที่ได้จากการเปรียบเทียบสายลำดับบางส่วน (local sequence alignment) โดยเป็นสายลำดับที่ได้จากการสุ่ม หรือเป็นสายลำดับที่ไม่เกี่ยวข้องกัน จะไม่เป็นไปตามการกระจายแบบนี้ แต่จะกระจายแบบ extream value distribution ซึ่ง ณ ตำแหน่งที่มีค่าคะแนนสูงจะมีหางยาวแทน

Normalized library แหล่งเก็บรวบรวมลำดับ cDNA (cDNA library) ซึ่งได้จากการกำจัดเอาสายลำดับที่มีการแสดงออกสูงๆ ออกไป เพื่อให้ได้สัดส่วนของ messenger RNAs ที่มีจำนวนน้อยให้มากขึ้น ดังนั้น Normalized library จึงไม่ได้แสดงภาพรวมของการแสดงออกของยีนในเนื้อเยื่ออย่างถูกต้องแม่นยำ

Northern blot เทคนิคที่ใช้ในการบ่งชี้และหาตำแหน่งของ mRNA ในลำดับเบสที่สามารถเข้าคู่กับชิ้นส่วนดีเอ็นเอที่ใช้ในการตรวจสอบที่เรียกว่า probe ได้

Nucleotide ส่วนประกอบทางโครงสร้างหรือองค์ประกอบพื้นฐาน (building blocks) ของดีเอ็นเอและอาร์เอ็นเอ นิวคลีโอไทด์ประกอบด้วยเบสหนึ่งตัว (1 ใน 4 ชนิดต่อไปนี้ ได้แก่ adenine, thymine, cytosine และ guanine) ต่ออยู่กับโมเลกุลน้ำตาลและฟอสเฟตอย่างละ 1 โมเลกุล

Nucleus ศูนย์กลางของโครงสร้างของเซลล์เป็นที่อยู่ของโครโมโซม

Object Management Group (OMG) องค์กรที่ไม่แสวงหากำไร ถูกจัดตั้งขึ้นเพื่อประชาสัมพันธ์ข้อมูลข่าวสารเกี่ยวกับโปรแกรมประเภท component-based software โดยการแนะนำ standardize object software OMG ได้กำหนดแนวทางในระดับอุตสาหกรรมและแจกแจงรายละเอียดเกี่ยวกับข้อกำหนดต่างๆ ของ object management เพื่อให้ได้กรอบโครงสร้างที่เหมาะสมในการพัฒนาโปรแกรมประยุกต์ ภายใน OMG ประกอบไปด้วยกลุ่ม

วิจัยวิทยาศาสตร์ชีวภาพ กลุ่มสมาคมบริษัทยา สถาบันการศึกษา ผู้ค้า software และ hardware ซึ่งทำงานร่วมกันเพื่อปรับปรุงประสิทธิภาพในการติดต่อสื่อสารและการทำงานร่วมกันระหว่างแหล่งข้อมูลทางคอมพิวเตอร์ที่เกี่ยวข้องกับการวิจัยทางชีวภาพ

Oligonucleotide array assay DNA chip ได้จากการพิมพ์หรือสังเคราะห์ oligonucleotide probes ต่างๆ ลงบนวัสดุรองรับ (solid support) Oligonucleotide array ถูกออกแบบให้สามารถตรวจหาตำแหน่งที่เกิดการกลายพันธุ์ พื้นที่เป้าหมาย (target region) หาตำแหน่งของยีนที่มีการแสดงออก (gene expression) และบ่งชี้หน้าที่ของยีน (gene function) ได้

Oligonucleotide probe ลำดับดีเอ็นเอสายสั้นๆ ที่ถูกสังเคราะห์ขึ้นโดยเฉพาะ เพื่อให้เข้าคู่ (hybridization) กับดีเอ็นเอเป้าหมายได้

Oncogene ยีนที่มีความสามารถทำให้เซลล์ปกติให้กลายเป็นเซลล์มะเร็ง

Pairwise alignment คือการเปรียบเทียบสายลำดับทีละคู่ โดยสายลำดับทั้ง 2 เส้นจะถูกเติมช่องว่างแทรกเข้าไปเพื่อให้สายลำดับทั้งสองมีความยาวเท่ากัน แล้วจึงแสดงค่าความเหมือนกันที่สูงที่สุด โดยเทียบทีละตัวอักษร ส่วน optimal pairwise alignment คือการเปรียบเทียบสายลำดับที่มีคะแนนเหมาะสมที่สุด (คือ มีความเหมือนกันที่สุดหรือมีค่าสูญเสียต่ำที่สุด)

P-value คล้ายกับ E-value แต่ P-value คือความน่าจะเป็นของโอกาสที่จะเกิดผลขึ้นจากคะแนนที่มีอยู่หรือมากกว่า ซึ่งตรงข้ามกับจำนวนผลที่คาดไว้ ค่าสูงสุดของ P-value คือ 1.0 ในขณะที่ E-value คือจำนวนสายลำดับในฐานข้อมูลที่มากที่สุดที่ถูกสืบค้นมาได้ หาก P-value มีค่าน้อย ๆ (คือมีนัยสำคัญ) ค่า P และ E จะมีค่าเท่ากัน ดังนั้นชุดโปรแกรมต่าง ๆ จึงเลือกใช้ค่าใดก็ได้ ตัวอย่างเช่น NCBI BLAST 2.0, FASTA และ HMM จะรายงานผลซึ่งแสดงE-value ในขณะที่ WU-BLAST 2.0 จะรายงานผลโดยแสดง P-value ดูเพิ่มเติมที่ E-value

Pearson correlation coefficent การวัดความสัมพันธ์ระหว่างตัวแปร 2 ตัว ซึ่งส่งผลถึงระดับความสัมพันธ์ของตัวแปรทั้งสอง ตัวอย่างเช่น ค่าสัมประสิทธิ์ที่ใช้วัดความเหมือนกันของแสดงออกของยีน (gene expression) ในการทดลองทำ microarray

Peptide กรดอะมิโนตั้งแต่ 2 ตัวขึ้นไปต่อกันด้วยพันธะเปปไทด์

Percent identity ร้อยละของคอลัมน์ที่ได้จากการเปรียบเทียบ alignment สายลำดับเบส 2 เส้น ซึ่งประกอบด้วยกรดอะมิโนที่เหมือนกัน คอลัมน์ของการเปรียบเทียบสายลำดับที่เป็นช่องว่าง (gaps) จะไม่ถูกนำมาคำนวณด้วย

Percent similarity ร้อยละของคอลัมน์ที่ได้จากการเปรียบเทียบ alignment สายลำดับเบส 2 เส้น ซึ่งประกอบด้วยกรดอะมิโนที่เหมือนกัน หรือกรดอะมิโนถูกแทนที่บ่อยๆ ในสายลำดับโปรตีนที่เกี่ยวข้องกัน (conservative substitutions) การแทนที่แบบนี้อาจพบในเมทริกซ์การแทนที่ (substitution matrix) ของกรดอะมิโน เช่น เมทริกซ์ PAM ของ Dayhoff และเมทริกซ์ BLOSUM ของ Henikoff เป็นต้น แต่คอลัมน์ของการเปรียบเทียบสายลำดับที่เป็นช่องว่างจะไม่ถูกนำมาคำนวณด้วย

Phenotype ลักษณะเฉพาะ (traits) ของสิ่งมีชีวิตที่สังเกตได้ ตัวอย่างเช่น สีผม น้ำหนัก หรือการมีหรือไม่มีโรคต่างๆ ลักษณะดังกล่าวนี้ไม่จำเป็นต้องเกี่ยวข้องกับพันธุกรรมก็ได้

Phrap โปรแกรมที่ถูกพัฒนาโดย Phil Green จากมหาวิทยาลัยวอชิงตัน ย่อมาจาก PHil's Revised Assembly Program เป็นเครื่องมือที่ใช้ประกอบชิ้นส่วนดีเอ็นเอที่ได้จากเทคนิค shotgun เข้าด้วยกัน

Physical map แผนที่โครโมโซมของสิ่งมีชีวิตสปีชีส์หนึ่งๆ โดยแสดงตำแหน่งทางกายภาพของยีนหรือ markers บนแต่ละโครโมโซม แผนที่กายภาพมีความสำคัญในการค้นหายีนที่ก่อให้เกิดโรคโดยใช้วิธี positional cloning และการหาลำดับเบสบนดีเอ็นเอ (DNA sequencing)

PHYLIP โปรแกรมที่สร้างโดย J. Felsenstein ใช้สำหรับศึกษา Phylogeneticity

Phylogenetic studies ในโปรแกรม Sequin นั้น ตัวอย่างที่นำมาศึกษาเป็นชุดของสายลำดับของยีนเดียวกันที่มาจากสิ่งมีชีวิตต่างๆ จากหลายสปีชีส์ โดยมีสมมติฐานว่า สิ่งมีชีวิตแต่ละหน่วยนั้นไม่สามารถผสมพันธุ์กันได้ ในกรณีนี้ Sequin ไม่ยินยอมให้ระบุชื่อสิ่งมีชีวิตนั้นได้ แต่จะให้ใส่รหัสใน definition line แทน สำหรับโปรแกรม Sequin นั้น ชุดลำดับเบสที่มียีนเหมือนกันแต่อยู่ในสิ่งมีชีวิตคนละตัว มีสมมติฐานว่าสิ่งมีชีวิตแต่ละตัวไม่สามารถผสมพันธุ์กันได้ sequin ไม่ยินยอมให้ระบุชื่อสิ่งมีชีวิต แต่จะให้รหัสใน definition line แทน โดยทำหน้าที่เป็นตัวควบคุมการกำหนดรหัสพันธุกรรม (genetic codes)

Phylogenetic trees เป็นแผนภูมิการสืบทอดลักษณะต่างๆ จากบรรพบุรุษ โดยแผนภูมินี้สร้างจากข้อมูลที่ได้จากการเปรียบเทียบสายลำดับกรดอะมิโนในโปรตีนตัวอย่างจากสิ่งมีชีวิตสปีชีส์ต่างๆ เช่น cytochrome C เป็นต้น

PIR ย่อมาจาก Protein Identification Resource International ผู้ให้บริการฐานข้อมูลโปรตีน

Plasmid ดีเอ็นเอที่อยู่นอกโครโมโซม (Extrachromosomal DNA)

Platform หมายถึงระบบปฏิบัติการบนคอมพิวเตอร์ที่โปรแกรมต่างๆ จำเป็นต้องใช้เพื่อให้สามารถทำงานได้ (เช่น UNIX หรือ Windows 95) โดยทั่วไปมักหมายถึงชนิดของคอมพิวเตอร์ที่ใช้งาน เช่น Macintosh หรือ PC

Poisson distribution ใช้ในการทำนายความเป็นไปได้ในการเกิดเหตุการณ์ที่มีความถี่ในการเกิดน้อยในช่วงระยะเวลาหนึ่งๆ หรือเมื่อมีจำนวนครั้งในการทดลองมากๆ ในการวิเคราะห์สายลำดับนั้น การกระจายแบบนี้ถูกนำมาใช้ในการคำนวณโอกาสที่จะมีเบสในสายลำดับที่ไม่เกี่ยวข้องกันจับคู่กันได้ และให้คะแนนในการเปรียบเทียบแบบ local alignment สูง

Polymerase chain reaction (PCR) เทคนิคในการสร้างชิ้นส่วนดีเอ็นเอจำนวนมากได้อย่างรวดเร็วและประหยัด บางครั้งอาจเรียกว่า "การทำสำเนาระดับโมเลกุล (molecular photocopying)" เทคนิค PCR นี้เป็นประโยชน์มากในการค้นคว้าวิจัยทางด้านชีววิทยาแพทยศาสตร์ โดยเฉพาะอย่างยิ่งด้านพันธุศาสตร์

Polymorphic Marker หมายถึง marker ที่แสดงความหลากหลายของประชากร ซึ่งช่วยให้สามารถติดตามการสืบทอดลักษระทางพันธุกรรมได้ marker ต้องมีลักษณะเป็น polymorphic เพื่อที่จะถูกใช้วางตำแหน่งด้วยวิธีการทางพันธุศาสตร์ ถึงแม้ว่า marker ใดๆ จะสามารถถูกวางไว้บนแผนที่กายภาพ (physical map) ได้ก็ตาม

Polymorphism การที่มี alleles ของ locus หนึ่งๆ ในประชากรที่กำหนดให้ตั้งแต่ 2 alleles ขึ้นไป โดยความถี่ในการปรากฏจะสูงกว่าความถี่ของการเกิดการกลายพันธุ์

Population studies ในโปรแกรม Sequin นั้น ประชากรหมายถึงชุดของสายลำดับของยีนเดียวกันที่มาจากสิ่งมีชีวิตต่างๆ ในสปีชีส์เดียวกัน โดยมีสมมติฐานว่า สิ่งมีชีวิตแต่ละหน่วยนั้นสามารถผสมพันธุ์ระหว่างกันได้ Sequin จะยินยอมให้ตั้งชื่อสิ่งมีชีวิตนั้นได้ โดยใช้ข้อมูลที่ช่วยในการจำแนก เช่น สายพันธุ์ clone หรือ isolate ซึ่งข้อมูลที่เหล่านี้จะช่วยให้โปรแกรมสามารถทำงานได้ถูกต้องยิ่งขึ้น

Position-specific scoring matrix (PSSM) แสดงค่าความผันแปรที่พบในคอลัมน์ต่างๆ ของการเปรียบเทียบ alignment สายลำดับหลายเส้นที่มีความสัมพันธ์กัน โดยคอลัมน์ของเมทริกซ์แต่ละคอลัมน์จะมีลักษณะสอดคล้องกับคอลัมน์ถัดไป และในแต่ละแถวจะสอดคล้องกับตัวอักษรหนึ่งๆ ในสายลำดับ (ตัวอักษรดังกล่าวอาจเป็นเบส 1 ใน 4 ชนิดในลำดับดีเอ็นเอ หรือกรดอะมิโนใดๆ ใน 20 ชนิดในลำดับโปรตีน) ค่าเมทริกซ์ที่ได้จะอยู่ในรูป log odds scores ซึ่งได้จากจำนวนนับ (counts) ของตัวอักษรที่ในสายลำดับ หารด้วยค่าคาดหวังของจำนวนนับซึ่งขึ้นกับส่วนประกอบในสายลำดับ แล้วจึงเปลี่ยนอัตราส่วนที่ได้ไปเป็นค่า log ตารางคะแนนหรือเมทริกซ์นี้จะถูกนำไปใช้ประเมินหรือเคลื่อนที่ไปตามสายลำดับเพื่อหาบริเวณที่คล้ายกันโดยหา log odds scores ที่เข้าคู่กัน และหาค่าคะแนนที่สูงที่สุด กรณีนี้จะไม่มีช่องว่าง (gaps) ในการทดสอบ บางครั้งอาจเรียกว่า weight matrix หรือ scoring matrix

Positional cloning เทคนิคในการบ่งชี้ยีนที่เกี่ยวข้องกับโรคใดโรคหนึ่ง ประกอบไปด้วยการทำแผนที่ยีน (genetic mapping) การทำแผนที่เชิงกายภาพ (physical mapping) และการหาลำดับเบส (sequencing)

Posterior (Bayesian analysis) ความน่าจะเป็นแบบมีเงื่อนไขซึ่งขึ้นกับความรู้ที่มีมาก่อนหน้านี้และความสัมพันธ์ที่ประเมินขึ้นใหม่จากตัวแปรต่างๆ โดยใช้กฎของ Bayes (Bayes' rule)

Primer ลำดับนิวคลีโอไทด์สายสั้นๆ ซึ่งใช้ในเทคนิค PCR (polymerase chain reaction)

Prior (Bayesian analysis) ค่าการกระจายที่คาดหวังของตัวแปรที่ขึ้นอยู่กับข้อมูลที่มีมาก่อนหน้านี้

Profile ตารางคะแนนของบริเวณที่ถูกอนุรักษ์ไว้จากการเปรียบเทียบสายลำดับหลายเส้น multiple alignmentที่สอดคล้องกัน โดยยอมให้มีช่องว่าง (Gaps) ปรากฏในการเปรียบเทียบด้วย โดยแต่ละคอลัมน์ของการเปรียบเทียบโปรตีนนั้น โปรไฟล์จะกำหนดคะแนนให้แก่กรดอะมิโนทั้ง 20 ชนิด (กรดอะมิโนแต่ละชนิดมีคะแนนเฉพาะตัว) และกำหนดคะแนนให้แก่การเปิดช่องว่างที่เรียกว่า gap penalties ซึ่งอาจเกิดจากการแทรกสอด (insertions) ในบริเวณที่อยู่ติดกับคอลัมน์ หรืออาจเกิดจากมีการหายไป (deletion) ภายในคอลัมน์ อาจเรียกโปรไฟล์แบบนี้ว่า "Position specific scoring matrices" (PSSMs) โปรไฟล์ที่ไม่มีการแทรกสอดและการหายไปจะเรียกว่า "weight matrices"

Profile alignment หมายถึง การปรับเปลี่ยนกระบวนการเปรียบเทียบ alignment สายลำดับอย่างง่ายให้เหมาะกับสายลำดับของความถี่ของตัวอักษร แทนที่จะเป็นสายลำดับของตัวอักษร

Profile hidden Markov model แบบจำลอง hidden Markov model ของบริเวณที่ถูกอนุรักษ์ไว้ (Conserved region) ซึ่งได้จากการเปรียบเทียบสายลำดับพร้อมกันหลายเส้น multiple sequence alignment โดยรวมถึงช่องว่าง (gaps) ด้วย แบบจำลองนี้สามารถนำไปใช้ในการค้นหาสายลำดับใหม่เพื่อหาความคล้ายคลึงกับสายลำดับที่ได้รับการเปรียบเทียบแล้ว

Promoter องค์ประกอบหนึ่งของยีนที่มีข้อมูลในการเปิดหรือปิดการทำงานของยีน กระบวนการสร้างอาร์เอ็นเอจากดีเอ็นเอ หรือ transcription จะเริ่มขึ้นที่ promoter

Protein โมเลกุลที่ประกอบด้วยกรดอะมิโนต่อกันด้วยพันธะเปปไทด์ ซึ่งลำดับของกรดอะมิโนถูกเรียงตามข้อมูลในยีนบนสายลำดับดีเอ็นเอ โปรตีนมีหน้าที่หลากหลาย ซึ่งรวมถึงการเป็นตัวเร่งปฏิกิริยาหรือเอนไซม์ เป็นโครงสร้างของเซลล์ หรือโมเลกุลที่ทำหน้าที่ถ่ายทอดสัญญาณ

Proteome ภาพรวมของกลุ่มโปรตีนอย่างสมบูรณ์ ที่แสดงออกในเนื้อเยื่อ เซลล์ หรือระบบชีวภาพ ณ เวลาหนึ่งๆ การศึกษา proteome เริ่มโดยการทำนายยีน (gene prediction) และการระบุหน้าที่ของยีน (annotation method) แต่ในที่สุดจะต้องทำการทบทวน โดยใช้ข้อมูลที่ได้รับจากสายลำดับเบสที่มียีนที่มีแสดงออก

Proteomics การวิเคราะห์การแสดงออกของโปรตีนในเนื้อเยื่อที่ปกติและเนื้อเยื่อที่ผิดปกติหรือเป็นโรค อย่างเป็นระบบ

Pseudogene สายลำดับดีเอ็นเอที่คล้ายกับยีนปกติมาก แตกต่างกันเพียงเล็กน้อยเท่านั้น จึงทำให้ไม่สามารถแสดงออกได้ ยีนดังกล่าวนี้อาจเคยทำงานได้ แต่เมื่อเวลาผ่านไปยีนนี้อาจมีการเปลี่ยนแปลงหรือกลายพันธุ์ไป จนไม่สามารถสร้างโปรตีนproteinได้

PubEST ตัวย่อของสายลำดับเบสที่ได้จากฐานข้อมูลสาธารณะ เช่น โครงการ WashU-Merck EST หรือสถาบันวิจัย Banting

Radiation Hybrid (RH) Mapping วิธีการทำ physical mapping วิธีหนึ่งที่ใช้ความถี่ของการแตกพันธะในสายดีเอ็นเอที่เกิดจากการเหนี่ยวนำโดยรังสีเอ็กซ์ ซึ่งสามารถระบุระยะทางระหว่าง marker gene ได้

Radiation Hybrid (RH) Panel A set of DNA samples prepared from a collection of radiation hybrids. Each radiation hybrid is a clonal population of cells derived by fusion of lethally X-irradiated donor cells with rodent cells.

RASMOL ชุดโปรแกรมที่ R. Sayle คิดค้นขึ้นเพื่อแสดงโครงสร้างของโปรตีน

Rational drug design วิธีการสร้างสารประกอบอย่างเป็นระบบ โดยการวิเคราะห์โครงสร้าง หน้าที่การทำงาน และปฏิสัมพันธ์

Recombinant DNA technology เทคนิคต่างๆ ที่ที่นักวิทยาศาสตร์สาขาอณูชีววิทยาใช้ในการปรับแต่งโมเลกุลของดีเอ็นเอเพื่อศึกษาการแสดงออกของยีน

Regular expressions โปรแกรมที่ทำการคำนวณเพื่อใช้หาความแปรผันที่พบในกลุ่มสายลำดับที่ใกล้เคียงกัน ซึ่งรวมถึงโอกาสที่จะมีการเปลี่ยนแปลงลำดับเบสเพียงตำแหน่งเดียว การแทรกสอด การเกิดซ้ำ ฯลฯ ตัวอย่างเช่น โปรแกรมที่ใช้หาความผันแปรต่างๆ ที่พบในโดเมนของโปรตีนต่างๆ ในฐานข้อมูล PROSITE

Relational database การจัดการข้อมูลให้อยู่ในรูปตาราง โดยที่แต่ละคอลัมน์จะแสดงฟีลด์ (Field) หรือชนิดของข้อมูลที่เก็บไว้ในบันทึก (Record) แต่ละแถวของตารางแสดงส่วนประกอบที่เกี่ยวข้องในบันทึกแต่ละบันทึกนั้น ฐานข้อมูลหนึ่งๆ สามารถมีได้หลายตาราง และใช้ภาษาคำถาม (Queru language) เพื่อเข้าสู่ข้อมูล

Repressor โปรตีนที่ทำหน้าที่ระงับการทำงานของยีน

Restriction enzymes เอนไซม์ที่มีความจำเพาะกับลำดับเบสเฉพาะในสายคู่ดีเอ็นเอและจะตัดดีเอ็นเอ ณ ตำแหน่งนั้นเท่านั้น มักเรียก Restriction enzyme ว่ากรรไกรระดับโมเลกุล (Molecular scissors)

Restriction fragment length polymorphisms (RFLP) ความแปรผันทางพันธุกรรม ณ ตำแหน่งที่ restriction enzyme ไปตัดชิ้นส่วนดีเอ็นเอ และพบว่าความแปรผันดังกล่าวมีผลต่อขนาดของชิ้นส่วนดีเอ็นเอที่ได้ด้วย ลำดับเบสชนิดนี้สามารถใช้เป็นเครื่องหมายใน physical maps และ Linkage maps ได้ RFLP อ่านออกเสียงว่า "rif"

Ribonucleic acid (RNA) มีองค์ประกอบทางเคมีใกล้เคียงกับสายดีเอ็นเอสายเดี่ยว แต่ในอาร์เอ็นเอนั้น รหัสพันธุกรรมจะมีนิวคลีโอไทด์ U (uracil) แทน T (Thymine) อาร์เอ็นเอทำหน้าที่ขนส่งข้อมูลทางพันธุกรรมของดีเอ็นเอ (genetic code) ไปยังไซโตพลาสมของเซลล์เพื่อสร้างเป็นโปรตีน

Ribosome องค์ประกอบภายในเซลล์ เป็นออร์กาเนลล์ (organelles) ที่ทำหน้าที่สร้างโปรตีน

Score ค่าคะแนนของการเปรียบเทียบสายลำดับซึ่งคำนวณจาก scoring function (Cost of an Alignment, Alignment Score)

Scoring along a tree คือวิธีที่ใช้ในการคำนวณคะแนนของการเปรียบเทียบสายลำดับพร้อมกันหลายสาย (multiple alignment) ซึ่งได้มาจากผลรวมของการคำนวณค่าคะแนนทั้งหมดของสายลำดับที่เกี่ยวข้องเท่านั้น (คือสายลำดับที่มีความใกล้เคียงกันในสายวิวัฒนาการ)

Scoring function ทำหน้าที่เปลี่ยนแนวคิดเชิงนามธรรมให้เป็นค่าตัวเลข Scoring function ถูกนำมาใช้ในการประเมิน (คิดคะแนน) ขั้นตอนการแก้ไขที่เกิดขึ้นเพียง 1 ขั้นตอน (หรือการรวมลักษณะที่เกี่ยวข้องอย่างง่ายๆ) การประเมินการเปรียบเทียบสายลำดับพร้อมกันหลายสายmultiple alignmentเป็นคอลัมน์ การเปรียบเทียบสายลำดับทีละคู่แบบทั้งสาย (whole pairwise alignments) การเปรียบเทียบสายลำดับพร้อมกันหลายแบบทั้งสาย (whole multiple alignment) โดยทั่วไปคะแนนการเปรียบเทียบสายลำดับ 2 สาย 2 อันคือ s และ t เป็นผลรวมของคะแนนจากการแก้ไข สำหรับการเปรียบเทียบสายลำดับพร้อมกันหลายสายนั้น มีวิธีการให้คะแนนได้หลายแบบ เช่น sum-of-pairs score หรือ scoring along a tree เป็นต้น ปกติ Cost Function และ Weight Function มักถูกนำมาใช้เหมือนๆ กัน โดยอาศัยแนวคิดเหมือนกันคือ "ค่ายิ่งน้อยยิ่งดี" การวัดความคล้ายคลึง (similarity measure) และการวัดความห่างไกล (distance measure) เป็นรูปแบบของ scoring function

Selectivity (in database similarity searches) ความสามารถของวิธีที่ใช้ในการสืบค้นเพื่อจัดโปรตีนให้อยู่ในกลุ่ม (family) โดยไม่ทำให้เกิดผลบวกแบบไม่ถูกต้องในการจัดกลุ่มของโปรตีนในกลุ่มอื่น

Sensitivity (in database similarity searches) ความสามารถของวิธีที่ใช้ในการสืบค้นเพื่อจัดโปรตีนให้อยู่ในกลุ่ม (family) ให้ได้มากที่สุดเท่าที่จะเป็นไปได้ โดยรวมโปรตีนที่มีความคล้ายคลึงอย่างจำกัดเข้าไปด้วย

Sequence alphabet สัญลักษณ์หรือชุดของตัวอักษรซึ่งนำไปเรียงต่อกันเป็นสายลำดับ เช่น ตัวอักษรในสายดีเอ็นเอ (A,T,G,C) หรือตัวอักษรที่ใช้แทนกรดอะมิโนในโปรตีน

Sequence Tagged Site (STS) ตำแหน่งในจีโนม ซึ่งมีข้อมูลสายลำดับจำเพาะ สามารถตรวจพบได้โดยปฏิกิริยา PCR

Server คอมพิวเตอร์ที่ทำงานตามคำสั่งที่ได้รับจากคอมพิวเตอร์ลูกข่าย (client) ที่อยู่ห่างไกลออกไป

Signal transduction กระบวนการทางชีววิทยาที่เกิดขึ้นเมื่อมีสัญญาณจากภายนอกไปทำให้เกิดสัญญาณขึ้นภายในเซลล์ ทำให้เกิดการตอบสนองทางสรีรวิทยาของเซลล์

Significance ผลลัพธ์ที่มีนัยสำคัญหมายถึงผลลัพธ์ที่ไม่ได้เกิดขึ้นด้วยความบังเอิญ ดังนั้นจึงเป็นผลที่น่าจะเกิดขึ้นตามความเป็นจริง ระดับของนัยสำคัญเป็นสิ่งที่แสดงว่าผลลัพธ์นั้นเป็นผลมาจากความบังเอิญมาน้อยเพียงใด โดยแสดงในรูปของความน่าจะเป็น ในการวิเคราะห์สายลำดับนั้น ความมีนัยสำคัญของคะแนนที่ได้จากการเปรียบเทียบสายลำดับalignment scoreอาจคำนวณในรูปของโอกาสที่คะแนนนั้นจะพบได้ในระหว่างสายลำดับที่ไม่มีความเกี่ยวข้องกันหรือสายลำดับแบบสุ่ม

Similarity score (sequence alignment) ผลรวมของตำแหน่งที่เหมือนกัน (identical matches) และตำแหน่งที่มีการแทนที่แบบอนุรักษ์ (conservative substitutions) ซึ่งมีคะแนนสูง ที่ได้จากการเปรียบเทียบสายลำดับ หารด้วยจำนวนตัวอักษรทั้งหมดในสายลำดับที่นำมาเปรียบเทียบ โดยทั่วไปในกรณีนี้มักจะไม่ให้มีช่องว่าง (gaps) ปรากฏ

Similarity measure เป็นฟังก์ชันที่แสดงความสัมพันธ์ระหว่างค่าตัวเลขกับคู่ของสายลำดับ โดยมีแนวความคิดคือ ค่ายิ่งสูงแสดงว่าสายลำดับคู่นั้นยิ่งคล้ายกัน Similarity measure เป็นฟังก์ชันในการให้คะแนนแบบหนึ่ง scoring function(Similarity Function, Similarity Score)

Single-linkage cluster analysis การวิเคราะห์วัตถุที่มีความเกี่ยวข้องสัมพันธ์กัน เช่น โปรตีนที่คล้ายกันแต่อยู่ในจีโนมต่างกัน เพื่อแจกแจงความสัมพันธ์ทั้งแบบใกล้เคียงและห่างไกลกัน แล้วแสดงเป็นแผนภูมิต้นไม้ (tree หรือ dendogram) วิธีนี้นำเอาวัตถุที่ใกล้ชิดกัน ซึ่งได้จาก Neighbor-joining algorithm มาแสดงอยู่บนสาขาด้านนอกของแผนภูมิ ส่วนวัตถุที่ห่างไกลกันมากกว่าจะถูกเติมลงในสาขาระดับล่างๆ ของแผนภูมิ นอกจากนี้วิธี

นี้ยังใช้เพื่อพยากรณ์ความสัมพันธ์ในการสืบทอดลักษณะ (phylogenetics relationships) ด้วย distance method

Site คอลัมน์หนึ่งๆ ที่ได้จากการเปรียบเทียบ alignment กรดอะมิโนหรือนิวคลีโอไทด์หน่วยต่างๆ ที่อยู่บนตำแหน่ง (site) เดียวกันถือว่าเหมือนกัน (homologous)

SNP ย่อมาจาก Single nucleotide polymorphism เป็น polymorphism ของนิวคลีโอไทด์ 1 ตำแหน่งในสายลำดับดีเอ็นเอ

Somatic cells เซลล์ทั้งหมดที่เป็นส่วนประกอบของร่างกายในสิ่งมีชีวิต ยกเว้นเซลล์สืบพันธุ์ (reproductive cells)

Southern blot เทคนิคที่ใช้ในการหาตำแหน่งของดีเอ็นเอที่สามารถจับคู่กันได้ด้วยเบสคู่สม (complementary) กับชิ้นส่วนดีเอ็นเออีกสายหนึ่งที่เรียกว่า probe (ตัวติดตาม)

Specificity (in database similarity searches) ความสามารถของวิธีที่ใช้ในการสืบค้นเพื่อจัดโปรตีนให้อยู่ในกลุ่ม (family) ให้ได้มากที่สุดเท่าที่จะเป็นไปได้ โดยรวมโปรตีน proteinที่มีความคล้ายคลึงอย่างจำกัดเข้าไปด้วย

SRS (Sequence Retrieval System) โปรแกรมที่ใช้ในการค้นหาข้อมูลในฐานข้อมูลชีวภาพ สร้างโดย T. Etzold (EMBL)

SRS WWW Sequence Retrieval System ที่ทำงานภายใต้สภาพแวดล้อมแบบ WWW

Structurally verified alignment ปรกติเป็นการเปรียบเทียบสายลำดับโปรตีนหลายเส้นพร้อมกัน (multiple alignment) โดยได้รับการยืนยันด้วยข้อมูลโครงสร้าง โดยเฉพาะอย่างยิ่งข้อมูลเกี่ยวกับตำแหน่งของกรดอะมิโนช่องวาง 3 มิติ (Structurally Verified Multiple Alignment)

Structure base drug design การวิจัยทางเภสัชวิทยาที่อาศัยข้อมูลโครงสร้างตติยภูมิของตำแหน่งเป้าหมายของยา

STS ย่อมาจาก Sequences Tagged Site STSs เป็นสายลำดับที่มีหน้าที่เฉพาะตัว คือ ใช้ในการจำแนกคู่ของ primer ที่ใช้ในการทดลอง PCR ที่ทำให้ได้ mapping reagents ที่สามารถระบุตำแหน่งใดตำแหน่งหนึ่งบนจีโนมได้ GenBank มีแผนก STS เพื่อสนับสนุนการเปรียบเทียบ STSs กับสายลำดับของแผนกอื่น เพื่อเชื่อมโยงตำแหน่งต่างๆ บนสายลำดับที่เป็นปริศนากับยีนที่ทราบข้อมูลแล้ว รายละเอียดเพิ่มเติมดูที่ http://www.ncbi.nlm.nih.gov/dbSTS

STS-Content Mapping วิธีจัดทำแผนที่กายภาพ (physical mapping) โดยใช้ STSs ที่ปรากฏอยู่ในโคลนที่มีขนาดใหญ่ (เช่น YACs) เพื่อเชื่อมโยงไปยังตำแหน่งต่างๆ ที่เรียงลำดับและสัมพันธ์กัน

Substitution การแทนที่นิวคลีโอไทด์หนึ่งๆ ในสายลำดับดีเอ็นเอด้วยนิวคลีไทด์อีกชนิดหนึ่ง หรือการแทนที่กรดอะมิโนหนึ่งๆ ในโปรตีนด้วยกรดอะมิโนอีกชนิดหนึ่ง

Suicide gene เป็นกลยุทธ์ที่ทำให้เซลล์มะเร็งมีความอ่อนแอต่อการรักษาด้วยเคมีบำบัด (chemotherapy) มาขึ้น วิธีหนึ่งที่ใช้ก็คือ การเชื่อมโยงยีนต่างๆ ที่แสดงออกในเซลล์มะเร็งกัลป์ยีนอื่นๆ ที่ทำหน้าที่ในการผลิตเอนไซม์ซึ่งไม่พบในสัตว์เลี้ยงลูกด้วยนม ทำให้สามารถเปลี่ยนสารซึ่งไม่มีอันตรายให้กลายเป็นสารที่เป็นพิษต่อเซลล์มะเร็งได้

Sum-of-pairs score เป็นวิธีที่ใช้คำนวณหาคะแนนของการเปรียบเทียบสายลำดับแบบหลายเส้นพร้อมกัน (multiple alignment) โดยคะแนนที่ได้จากแต่ละคู่ของการเปรียบเทียบจะถูกเติมเข้าไปเรื่อยๆ เพื่อให้ได้คะแนนสุดท้าย

Synteny การปรากฏอยู่ของชุดของยีนที่มีความเหมือนกัน (homologous gene) โดยมีการจัดลำดับแบบเดียวกันในจีโนม 2 จีโนม

Telnet โปรโตคอลอินเตอร์เน็ตหรือโปรแกรม ที่ช่วยให้ผู้ใช้สามารถเชื่อมต่อกับคอมพิวเตอร์ที่อยู่ห่างไกลออกไปได้

THC ย่อมาจาก Tentative Human Consensus sequences ซึ่งสร้างขึ้นจากการรวบรวม ESTs เข้าด้วยกันให้เป็น transcripts เสมือน THC จัดทำโดย TIGR และสามารถสืบค้นได้ที่ NCBI

Threading การทำนายโครงสร้างโปรตีนโดยการเปรียบเทียบสายลำดับโปรตีนที่ไม่ทราบโครงสร้างกับโปรตีนที่ทราบโรงสร้าง 3 มิติแล้ว เพื่อพิจารณาว่าสายลำดับกรดอะมิโนนั้นสอดคล้องกับโครงสร้างหรือไม่ ทั้งในแง่ของตำแหน่งที่อยู่และคุณสมบัติทางเคมี

Thymine เบสชนิดหนึ่งในจำนวน 4 ชนิด (ATGC) ซึ่งเป็นองค์ประกอบของดีเอ็นเอ สัญญลักษณ์แทน thymine คือ T ส่วนเบสชนิดอื่นๆ ได้แก่ adenine (A), guanine (G) และ cytosine (C) thymine จะจับคู่กับ adenine เสมอ

Training Set กลุ่มของสายลำดับนิวคลีโอไทด์หรือกรดอะมิโนที่เชื่อถือได้ ซึ่งได้จากการเปรียบเทียบสายลำดับแบบหลายเส้นพร้อมกัน (multiple sequence alignment) แล้วสร้างแบบจำลอง HMM (HMM profile) ขึ้นมาA collection of trusted sequences (amino acid or nucleotide) from which a multiple sequence alignment, profile or profile-HMM is built.

Transgenic สิ่งมีชีวิตที่สร้างขึ้นจากการทดลอง โดยการใส่ดีเอ็นเอภายนอกเข้าไปในเซลล์เริ่มต้น (germ line) ปกติมักทำโดยการฉีดดีเอ็นเอแปลกปลอมเข้าไปในนิวเคลียสของเซลล์ตัวอ่อน

Transgenic knockout สัตว์ที่เป็นแบบจำลอง ซึ่งยีนที่จำเพาะ 1 หรือ 2 ชุดถูกตัดออกไป (deleted) หรือถูกยับยั้งไม่ให้แสดงออก (inactivated)

Translocation การที่ชิ้นส่วนดีเอ็นเอขนาดใหญ่หลุดหรือขาดออกจากโครโมโซมหนึ่งๆ แล้วไปติดอยู่กับอีกโครโมโซมหนึ่ง

TrEMBL ย่อมาจาก Translated EMBL เป็นฐานข้อมูล SRS SRS ที่รวบรวมจากฐานข้อมูลดีเอ็นเอของ EMBL ท่านสามารถไป TrEMBL ได้โดยคลิกที่นี่Translated, a-based compilation of the EMBL DNA data library.

Trisomy การที่มีโครโมโซมใดโครโมโซมหนึ่งจำนวน 3 ชุด ซึ่งปกติจะมีเพียง 2 ชุดเท่านั้น Possessing three copies of a particular

Tumor suppressor gene ยีนที่ทำหน้าที่ในการปกป้อง โดยปกติจะจำกัดการเจริญเติบโตของเนื้องอกหรือมะเร็ง หากยีนดังกล่าวนี้กลายพันธุ์ไปจะทำให้ความสามารถในการจำกัดการเจริญเติบโตของเซลล์มะเร็งลดลง Tumor suppressor gene ที่รู้จักกันดีได้แก่ BRCA1 และ p53A

Uncertainty จากทฤษฎีสารสนเทศ (information theory) หมายถึงค่า log ของค่าเฉลี่ยของจำนวนตัวเลือกที่ใช้เพื่อการจำแนก ดูเพิ่มเติมที่ Information content

UniGene database ฐานข้อมูลสาธารณะซึ่งดูแลโดย NCBI โดยรวบรวมสายลำดับของ GenBank ที่เป็นผลผลิตจากการถอดรหัสยีนต่างๆ (transcript products)

Unified Modeling Language (UML) มาตรฐานที่สนับสนุนโดย Object Management Group ซึ่งกำหนดรูปแบบเครื่องหมายที่ใช้ในการออกแบบชนิดปรับตามวัตถุ (object-oriented design)

Unique singleton singleton ที่ไม่พบในฐานข้อมูลสาธารณะ

Uracil เบสชนิดหนึ่งในจำนวน 4 ชนิด (A, U, G, C) ซึ่งเป็นองค์ประกอบของอาร์เอ็นเอ สัญญลักษณ์แทน uracil คือ U uracil คือเบสที่แทนที่ thymine ในดีเอ็นเอ uracil สามารถจับคู่กับ adenine เช่นเดียวกับ thymine

URL ย่อมาจาก Uniformed Resource Locator ใช้ร่วมกับ web browsersเพื่อกำหนด site ที่จะเข้าไปใช้งาน (FTP, Gopher, or Web) รวมทั้งถือว่าเป็นที่อยู่ของเว็บไซต์ด้วย

Vector ตัวกลางที่ทำหน้าที่ในการนำพายีนที่ถูกดัดแปลงหรือยีนแปลกปลอม เช่น ไวรัส หรือดีเอ็นเอขนาดเล็กที่เรียกว่าพลาสมิด (plasmid) ในการรักษาด้วยยีน (gene therapy) นั้น พาหะหรือเวกเตอร์นี้จะปลดปล่อยยีนที่กำหนดไปยังเซลล์เป้าหมาย

Western blot เทคนิคที่ใช้ในการหาตำแหน่งของโปรตีนที่สามารถจับกับแอนติบอดีอย่างจำเพาะเจาะจงได้

World Wide Web ระบบที่ใช้ในการเผยแพร่เอกสาร ซึ่งสามารถใช้กับเอกสารหลายชนิดที่ไม่ใช่มีแค่ตัวอักษร

WPI ย่อมาจาก Wisconsin Package Interface, เป็นอินเตอร์เฟสแบบกราฟิกสำหรับชุดโปรแกรม GCG

X-ray crystallography เทคนิคที่ใช้ในการหาโครงสร้าง 3 มิติของโปรตีน

Yeast artificial chromosome (YAC) ดีเอ็นเอขนาดใหญ่ที่มีสายลำดับส่วนหนึ่งที่จำเป็นสำหรับทำหน้าที่เป็นโครโมโซมของยีสต์ โดยอยู่ร่วมกับชิ้นส่วนดีเอ็นเอแปลกปลอม (เช่น ดีเอ็นเอของคน) ซึ่งมีอาจมีขนาดใหญ่ถึง 1 Mb

พิมพ์ที่... สำนักพิมพ์มหาวิทยาลัยรามคำแหง
Ramkhamhaeng University Press.