

Chương VI

DỰ ÁN BỘ GEN NGƯỜI - ĐỈNH CAO CỦA DI TRUYỀN HỌC HIỆN ĐẠI

6.1. Ý NGHĨA Y SINH HỌC .

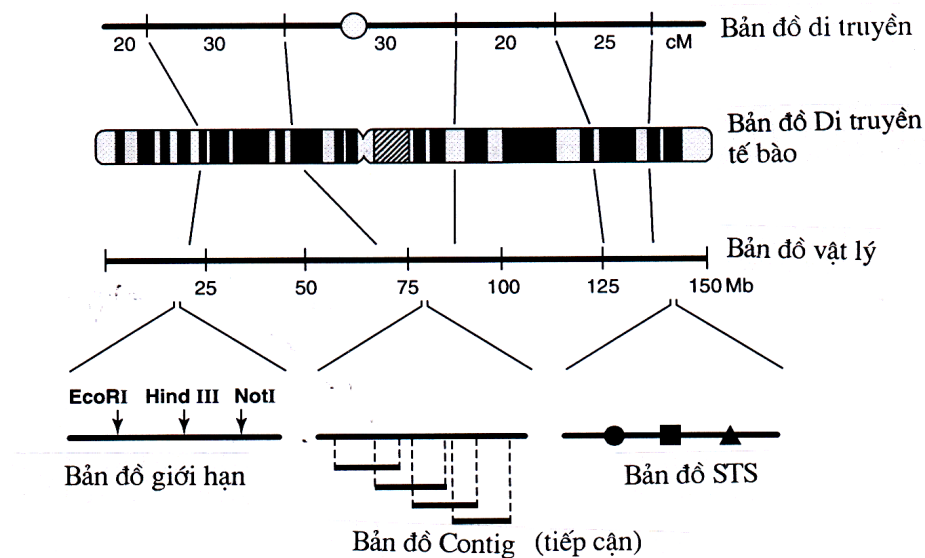
Những thông tin về việc xác định các trình tự bộ gen người và các tổ chức khác sẽ làm thay đổi lớn trong sinh học và y học . Chẳng hạn như với việc đối chiếu với bộ gen người , đang bùng nổ những thông tin mới về nguồn gốc sự sống của chúng ta , về các gen gây bệnh hay việc chẩn đoán và những khả năng tiến tới trị liệu . Những tiến bộ đạt được trong các lĩnh vực như hệ gen , Proteomics (hệ protein) , tin sinh học (bioinformatics) , hệ gen các chế phẩm thuốc (pharmacogenomics) cũng tăng lên rất nhanh .

Trong chương này chúng tôi giới thiệu một cách vắn tắt những phát hiện chính của dự án bộ gen người (The Human genome Project) (HGP) và những mối liên quan của chúng với sinh , y học .

6.2. DỰ ÁN BỘ GEN NGƯỜI CÓ NHIỀU MỤC ĐÍCH .

Dự án bộ gen người bắt đầu từ năm 1990 , đó là một nỗ lực quốc tế mà mục tiêu chính là xác định trình tự toàn bộ bộ gen người và bộ gen của một vài tổ chức khác đã có các cơ sở về nghiên cứu di truyền (như *Escherichia Coli* , nấm men (*Sacharomyces cerevisiae*) , ruồi dấm (*Drosophila melanogaster*) , giun đũa (*Caenorhabditis elegans*) và một loại chuột nhà thường gặp (*Mus Musculus*) .

Cho tới nay những mục tiêu chính của dự án đã được hoàn tất . Tại hoa kỳ , trung tâm quốc gia nghiên cứu bộ gen người (the National Center for Human Genome Research) (NCHGR) đã được thành lập năm 1989, khởi đầu được hướng dẫn bởi **Jame D. Watson** và sau này là **Francis Collins** . NCHGR giữ vai trò hàng đầu trong việc định hướng các nỗ lực quốc gia đối với dự án bộ gen người . Năm 1997 cơ quan này trở thành viện nghiên cứu quốc gia bộ gen người (the National Human Genome Research Institute —NHGRI) , trong đó có sự cộng tác quốc tế của các nhóm từ Hoa kỳ , Anh , Nhật bản , Pháp , Đức và Trung quốc và nó trở thành một tập đoàn quốc tế xác định trình tự bộ gen người (The International Human Genome Sequencing Consortium) —(IHGSC) . Lúc đầu cơ quan này chỉ đặt ra một số mục tiêu ngắn hạn như tạo ra bản đồ bộ gen người với các dấu chuẩn (marker) 2-5 xăng ti morgan (cM) riêng rẽ rồi xây dựng nên một bản đồ vật lý của tất cả 24 NST (22 nhiễm sắc thể thường và 2 NST giới tính là X và Y) với các dấu chuẩn có khoảng 100.000 cặp bazơ (bp) . (Hình .1) Tóm tắt những điểm khác biệt giữa **bản đồ di truyền** , **bản đồ di truyền tế bào** và **bản đồ vật lý của một NST** . Những mục tiêu này và những mục tiêu ban đầu đã thu được thành quả vượt trội vào những năm giữa 1990 . Năm 1998 một mục tiêu mới của HGP quốc gia đã được thông báo . Đó là việc mong muốn hoàn tất toàn bộ trình tự vào cuối 2003 hoặc sớm hơn . Những mục tiêu đặc biệt khác có liên quan tới công nghệ xác định trình tự , so sánh các hệ gen , tin sinh học , các vấn đề liên quan tới luân lý và các vấn đề khác .



Hình 6.1 Các phương pháp chính để xác định và phân lập các gen bình thường và gen gây bệnh. Với bản đồ di truyền, vị trí của một vài dấu chuẩn di truyền (marker) giả thuyết đã được chỉ rõ cùng với các khoảng cách di truyền tính bằng xăng ti morgan. Vòng tròn chỉ vị trí của tâm động (centromere). Với bản đồ di truyền tế bào những vị trí gần chính xác của các dấu chuẩn đã được chỉ rõ cùng với các khoảng cách vật lý tương đối tính bằng cặp megabazơ. Những ví dụ về bản đồ giới hạn, điểm tiếp cận (contig) và bản đồ STS cũng đều được chỉ rõ.

Mùa thu 1998, đã có khoảng 6% trình tự bộ gen người đã được hoàn tất và sẽ đặt ra cơ sở cho những công việc tương lai. Một tiến triển nữa là có một nhóm thứ 2, đó là một công ty tư nhân **Celera Genomics** đứng đầu là **Craig Venter** thông báo sẽ nhận trách nhiệm xác định trình tự bộ gen người. Năm 1995 Venter và đồng nghiệp đã xuất bản công trình về trình tự toàn bộ bộ gen của *Haemophilus influenzae* và *Mycoplasma genitalium*, đó là các mẫu đầu tiên đã được xác định trình tự gen. Một yếu tố quan trọng mang lại thành lợi cho những người làm công việc này là sử dụng một loại súng để bắn tức là *sonicating DNA*, xác định trình tự các đoạn và nối các trình tự trên cơ sở gộp lên nhau. Đối với việc so sánh thì có rất nhiều cách tiếp cận được sử dụng ở những thời điểm khác nhau khi nghiên cứu các gen bình thường và các gen gây bệnh (được ghi trong bảng 6.1)

Bảng 6.1 Những phương pháp chính dùng để xác định và phân lập các gen bình thường và gen gây bệnh.

Quy trình	Ghi chú
Phát hiện các bệnh liên quan tới di truyền tế bào đặc biệt.	Như trường hợp đứt nhỏ ở băng Xp 21.2 có liên quan đến gen gây bệnh đau cơ Deuchenne.
Những nghiên cứu tổng quát về NST.	Xác định nội giống của các gia đình lớn. Các gen trội dễ nhận dạng hơn các gen lặn.
Dùng đầu dò để xác định các locus dấu chuẩn	Đầu dò có thể xác định được hàng ngàn STS, RFLP, SNP □ khắp trên NST. Nó thăm dò cả 2 phía và phác hoạ rõ ràng.
Lập bản đồ lai bức xạ (radiation hybrid mapping)	Hiện nay là phương pháp nhanh nhất định vị một gen hay một đoạn DNA để phân miền NST người và xây dựng bản đồ vật lý.
Sử dụng các tế bào lai sinh dưỡng của người hay loài gặm nhấm.	Cho phép chuyển một gen tới một NST đặc hiệu nhưng không tới các phân miền.

Huỳnh quang trong lai situ.	Cho phép định vị một gen của một băng NST .
Dùng điện di gel xung điện (PFGE) để tách chiết các đoạn DNA lớn.	Cho phép phân lập các đoạn DNA dài thu nhận từ việc sử dụng <i>Endonucleaza</i> giới hạn để cắt rất hạn chế DNA.
NST đi dạo	Phân dòng các đoạn lặp DNA ; đó là quy trình trong phòng thí nghiệm thường chỉ khoảng 100-200 kb.
NST nhảy	Cắt DNA thành những mảnh tương đối lớn và quay vòng , nó có thể được loại đi nhanh chóng hơn và với các đoạn DNA dài hơn so với NST đi dạo .
Phân dòng qua YAC, BAC, Cosmid , Phage, và plasmid .	Cho phép phân lập các đoạn có chiều dài thay đổi .
Phát hiện sự biểu hiện của mRNA trong các mô bởi cách thâm Northern khi dùng một hay nhiều đoạn của gen như là một đầu dò .	mRNA có thể biểu hiện trong các mô bị tác động .
PCR	Sử dụng để khuếch đại các đoạn gen và các ứng dụng khác
Xác định trình tự DNA.	Thiết lập bản đồ vật lý giải pháp cao nhất . Xác định khung đọc mở .Tạo thuận lợi cho các phương tiện để có thể xác định được hàng triệu cặp bazơ mỗi ngày .
Cơ sở dữ liệu	Số sánh DNA và các trình tự protein thu nhận từ gen chưa biết với trình tự đã biết trong dữ liệu có thể làm dễ dàng việc định loại .

Ghi chú: STS : *Sequence tagged site* .

RFLP : *Restriction fragment linked polymorphism* .

SNP : *Single nucleotide polymorphism* .

YAC : *Yeast artificial chromosome* .

BAC : *Bacterial artificial chromosome* .

PCR : *polymerase chain Reaction* .

6.3. TRÌNH TỰ PHÁC THẢO CỦA BỘ GEN NGƯỜI ĐƯỢC THÔNG BÁO THÁNG 6 NĂM 2000

Tháng 6 -2000 , người đứng đầu IHGSC và công ty tư nhân Celera Genomics thông báo đã hoàn tất bản thảo trình tự bộ gen người với khoảng 90 % tổng số . Những phát hiện chính của 2 nhóm được xuất bản tách biệt ở 2 tờ báo vào tháng 1-2001 , IHGSC ở báo *Nature* và Celera thì ở *Science* . Bản thảo được tập đoàn này xuất bản là sản phẩm nghiên cứu ít nhất là 10 năm của 20 trung tâm tại 6 quốc gia . Xuất bản phẩm của Celera và các trung tâm hợp tác khác là sản phẩm của một công trình nghiên cứu với thời gian ít nhất là 3 năm hoặc ngắn hơn đôi chút . Còn một phần dữ liệu là của IHGSC . *Thành tựu của sự kết hợp này đã thu hút sự chú ý của công chúng và nó đã cung cấp cho chúng ta một thư viện về sự sống , một bảng liệt kê các thời kỳ của sự sống , còn Holy Grail thì cung cấp cho chúng ta những phát hiện về di truyền con người .*

6.4. HAI NHÓM SỬ DỤNG CÁC PHƯƠNG PHÁP KHÁC NHAU .

Dưới đây là tóm tắt những phát hiện lớn được ghi trong 2 bản thảo và những ghi chú về các ứng dụng của chúng . Những sự khác biệt không được nêu ra ở đây bởi vì những phần thống nhất chiếm tỷ lệ quá lớn . Tuy nhiên cũng cần phải tóm tắt những phương pháp khác nhau mà 2 nhóm đã sử dụng . Về cơ bản thì IHGSC trước hết dùng bản đồ sau đó rồi mới đến trình tự . Sở dĩ như vậy một phần là do việc xác định trình tự là một quá trình tiến hành chậm chạp khi mà dự án chung đã bắt đầu . Vì thế chiến lược của của tập đoàn là phải vượt thời gian bởi đã đạt được những tiến bộ trong việc xác định trình tự và các kỹ thuật khác . Cách tiếp cận tổng thể là đối chiếu trong khi xác định trình tự các mảnh nhỏ (shotgun) có thứ bậc (hierarchical shotgun sequencing) , bao hàm việc phá vỡ ra từng

mảnh toàn bộ bộ gen thành những đoạn xấp xỉ 100-200 kb và cài chúng vào NST nhân tạo của vi khuẩn (BAC). Những BAC này sau đó được định vị trên NST người bằng cách tìm các trình tự dấu chuẩn như là các vị trí bắt đầu các trình tự (sequence —tagged sites -STS), những vị trí này thì đã được xác định. Các STS đều ngắn (thường nhỏ hơn 500 bp), đó là locus hệ gen độc quyền cho các thử nghiệm PCR. Các dòng BAC sau đó được làm vỡ thành những mảnh nhỏ (shotgunning). Sau đó mỗi mảnh nhỏ đều được xác định trình tự và sử dụng những giải thuật máy tính để nhận dạng các thông tin trình tự thích ứng từ các đoạn gộp nhau thành những mảnh có các trình tự đầy đủ.

Celera thì sử dụng phương pháp Whole genome shotgun (bản toàn bộ bộ gen). Các đoạn shotgun được gắn với nhau bởi những giải thuật thành những bộ khung lớn và dùng STS để sửa chữa các vị trí trên khung này trong bộ gen. Một khung bao gồm một dãy các contig nằm ở phía bên phải nhưng không cần thiết phải nối vào các trình tự liên tiếp.

Contig là các trình tự kề nhau của DNA tạo bởi sự ghép lại các đoạn trình tự gộp nhau của một NST tự nhiên hay một BAC. Có thể tìm kiếm được máy xác định trình tự với số lượng lớn (high-throughput sequenator), các chương trình máy tính mạnh (powerful computer programs), yếu tố cạnh tranh và các yếu tố khác đã được tính toán để làm nhanh các tiến bộ

6.5. VIỆC XÁC ĐỊNH TRÌNH TỰ BỘ GEN NGƯỜI ĐEM LẠI ÍCH LỢI CHO CÁC PHÁT HIỆN MỚI.

Chúng ta hãy trích dẫn một vài phát hiện chính trong dự án bộ gen người ở bảng dưới đây:

Bảng 6.2. Những phát hiện chính trong bản thảo thô về bộ gen người.

- *Trên 90 % Bộ gen đã được xác định trình tự, còn những chỗ trống lớn và nhỏ vẫn phải tiếp tục bổ sung làm đầy.
- *Xác định được số gen mã cho protein là 30.000 đến 40.000.
- *Chỉ có 1,1 đến 1,5 % bộ gen là mã cho các protein.
- *Có những thay đổi lớn về hình ảnh NST người (tức là số gen theo từng Mb, mật độ SNP, lượng GC, số yếu tố có thể được vận chuyển và các đảo CpG, tốc độ tái tổ hợp).
- *Các gen của người thực hiện nhiều công việc hơn là gen của giun đũa hay ruồi dấm.
- *Hệ protein (proteome) người phức tạp hơn so với các động vật có xương sống khác.
- *Các trình tự lặp chiếm khoảng 50% bộ gen.
- *Khoảng 100 vùng mã hoá đã được sao chép và được loại đi bằng những hạt vận chuyển có cơ sở là RNA.
- *Khoảng 200 gen có thể truyền từ vi khuẩn sang.
- *Trên 3 triệu SNP đã được xác định.

6.6. HẦU HẾT BỘ GEN NGƯỜI ĐƯỢC XÁC ĐỊNH TRÌNH TỰ.

Trên 90 % bộ gen người đã được xác định trình tự vào tháng 7 -2000, bỏ xa bộ gen lớn nhất đã được xác định trình tự với kích cỡ khoảng 3,2 gigabazo (Gb). Dưới bộ gen người thì bộ gen của ruồi dấm là lớn nhất (khoảng 180 Mb). Vẫn còn những chỗ trống lớn và nhỏ và chất lượng của một số dữ liệu về trình tự sẽ phải xử lý lại vì một số phát hiện có lẽ là chưa được thật sự chính xác.

6.7. ĐƯỢC XÁC ĐỊNH ĐƯỢC RẰNG BỘ GEN NGƯỜI MÃ CHO KHOẢNG 30.000 -40.000 PROTEIN.

Điều ngạc nhiên nhất là so với những kết quả thu được từ trước thì rõ ràng là số gen mã cho protein là thấp, chỉ nằm giữa khoảng 30.000 đến 40.000. Theo những số liệu mới thì con số này cao hơn gấp khoảng 2 lần đối với giun đũa (19.099) và gấp 3 lần đối với ruồi dấm (13.061). Những hình ảnh này gợi ý rằng phải có những giải thích khác chính xác hơn về số gen của người vì người có tổ chức phức tạp hơn so với 2 tổ chức đơn giản hơn.

6.8. CHỈ CÓ 1,1 ĐẾN 1,5 % BỘ GEN NGƯỜI DÙNG ĐỂ MÃ CHO PROTEIN.

Phân tích các số liệu thấy rằng chỉ có 1,1 -1,5 % bộ gen là các exon. Khoảng 24 % là các intron và khoảng 75% các trình tự nằm trung gian (gen trung gian —intergenic). So sánh các số liệu trên giun đũa và ruồi dấm cho thấy kích thước của exon ở cả 3 mẫu tương đối hằng định (đối với người là 145 bp). Tuy nhiên, kích thước của intron người có thể biến đổi nhiều hơn (khoảng trên 3300bp). Kết quả cho thấy có sự thay đổi lớn trong kích thước của gen.

6.9. CẤU TRÚC CỦA CÁC NHIỄM SẮC THỂ NGƯỜI THAY ĐỔI RẤT LỚN.

Có nhiều đặc tính khác nhau giữa các NST của người, chẳng hạn như số megabazo, mật độ của các đa hình nucleotit đơn (SNP), lượng GC, số các yếu tố vận chuyển và các đảo CpG và tốc độ tái tổ hợp. Ta hãy lấy một ví dụ: NST số 19 chứa số gen nhiều nhất (23 gen/ megabazo), trong khi đó NST số 13 và nhiễm sắc thể Y thì chỉ thừa thớt (5 gen/ megabazo). Tới thời điểm này vẫn chưa giải thích được những sự thay đổi này.

6.10. CÁC GEN CỦA NGƯỜI HOẠT ĐỘNG NHIỀU HƠN CÁC GEN CỦA CÁC TỔ CHỨC ĐƠN GIẢN HƠN.

Sự nổi thay thế là khá phổ biến ở người, chiếm ít nhất là 35% các gen của chúng. Các số liệu chỉ ra rằng con số trung bình của các bản sao khác biệt /gen của NST số 22 và 19 là 2,6 và 3,2. Những đặc tính này là cao hơn giun tròn, nó chỉ có 12,2% gen là có nổi thay thế và chỉ có 1,34 thay đổi nổi /gen.

6.11. HỆ PROTEIN (PROTEOME) NGƯỜI PHỨC TẠP HƠN ĐỘNG VẬT CÓ XƯƠNG SỐNG.

Có tương đối ít các Domain protein mới xuất hiện ở các động vật có xương sống. Tuy nhiên, ở người số cấu trúc Domain phân biệt (khoảng 1800) của protein gấp 1,8 lần giun đũa và ruồi dấm. Khoảng 90 họ protein đặc biệt của động vật có xương sống có nhiều trong hệ miễn dịch và hệ thần kinh.

Những kết quả trong 2 bản thảo rất giàu thông tin về các họ và lớp protein. Ví dụ trong bảng 54-3 đã liệt kê các lớp chính của protein được mã bởi các gen của người. Và lớp lớn nhất là lớp “chưa rõ —unknown”. Việc xác định các protein chưa rõ này là một tiêu điểm lớn của nhiều phòng thí nghiệm.

Bảng 6.3. Những lớp chính của protein được mã bởi các gen của người.

Lớp protein	Số %
Chưa rõ	12.809 (41%)
Enzym axit nucleic	2.308 (7,5%)
Các yếu tố phiên mã	1.850 (6%)
Các Receptor	1.534 (5%)
Hydrolaza	1.227 (4,0%)
Các phân tử điều hoà chọn lọc (tức là protein G, các chất điều hoà chu kỳ tế bào).	988 (3,2%)
Các tiền gen gây ung thư	902 (2,8%)
Các pprotein cấu trúc bộ khung tế bào	876 (2,8%)
Kinaza	868 (2,8%)

6.12. CÁC TRÌNH TỰ LẶP CHIẾM TRÊN 50% BỘ GEN NGƯỜI.

Các trình tự lặp lại tính được ít nhất là chiếm một nửa bộ gen. Chúng rơi vào 5 lớp sau đây:

1. Các lặp rải rác (transposon derived repeats) (interspersed repeats).
2. Xử lý các gen giả.
3. Lặp các trình tự đơn.
4. Sao chép từng đoạn, tạo nên các bản sao 10-300kb từ một vùng của bộ gen thành một vùng khác.
5. Các khối trình tự lặp nối tiếp nhau có ở đoạn giữa (centromere) và đoạn cuối (telomere) và các vùng khác của NST.

Những thông tin chính về trình tự của hầu hết các lớp trên có giá trị trong việc tìm hiểu kiến trúc và sự phát triển của bộ gen người. Có 2 điểm cần quan tâm là nguyên tố nhôm được coi là thành viên nổi bật nhất (chiếm khoảng 10% tổng số bộ gen) của các yếu tố rải rác ngắn (short interspersed elements —SINE), có thể có các vùng giàu GC. Sự trùng đôi các đoạn ở người thấy phổ biến hơn giun đũa và ruồi dấm. Điều đó có thể là những cấu trúc này có liên quan tới sự tráo exon và làm tăng sự biến đổi protein ở người.

6.13. NHỮNG PHÁT HIỆN KHÁC ĐÁNG QUAN TÂM.

Ba điểm chính trong bảng 54-2 đã tóm tắt các vấn đề này. Đại để gồm các nội dung sau: Có khoảng 100 vùng mã hoá được xác định là có sự sao chép và được loại đi bởi các hạt chuyển cơ sở là RNA (retrotransposons). Điều đó có thể là một số gen có thể chấp nhận những vai trò theo tiến trình của thời gian. Một phát hiện đáng ngạc nhiên là trên 200 gen có thể có nguồn gốc từ vi khuẩn truyền tới. Không thấy các gen này ở các tế bào nhân chuẩn mà lại không phải là động vật có vú. Trên 3 triệu SNP đã được xác định. Điều này chứng minh về sự vô giá của các khía cạnh bản đồ gen.

Cũng cần phải nhấn mạnh rằng những phát hiện được liệt kê ở đây chỉ là một số nhỏ trong nội dung có trong các bản thảo . Ban đọc có thể bàn luận các vấn đề một cách đầy đủ trên các bản gốc của các báo cáo (phần đối chiếu) .

6.14. LẬP KẾ HOẠCH CHO NHỮNG CÔNG VIỆC TIẾP THEO VỀ BỘ GEN NGƯỜI VÀ CÁC TỔ CHỨC KHÁC .

IHGSC đã chỉ rõ rằng sẽ xác định đầy đủ các trình tự vào năm 2003 . Nhiệm vụ này có liên quan tới việc lấp đầy chỗ trống và xác định các gen mới cũng như vị trí và chức năng của chúng . Sẽ phải xác định các vùng điều hoà và các trình tự của các bộ gen lớn khác (như chuột nhà *Rattus norvegicus* , một loại chuột của Na uy ; *Danio rerio* —một loại cá vằn ; *Fugu rubripes* —một loại cá nóc hổ (the tiger puffer fish) và một hay nhiều động vật linh trưởng) . Thực vậy , bản thảo bộ gen cá nóc hổ đã được xuất bản năm 2002 . Các SNP sẽ được xác định bổ sung ; một Catalog đầy đủ của các phương án hy vọng sẽ có giá trị trong việc lập bản đồ gen với các chủng phức tạp và cho các ứng dụng khác . Cùng với các vấn đề trên , các cơ sở dữ liệu sẽ được thêm dần các thông tin mới và những cơ sở dữ liệu mới sẽ được thiết lập để phục vụ cho các mục đích đặc biệt . Sẽ tiến hành những nghiên cứu về hệ gen chức năng (functional genomics) (tức là nghiên cứu bộ gen để xác định chức năng của tất cả các gen và các sản phẩm của chúng) .

6.15. NHỮNG LIÊN QUAN TỚI HỆ PROTEIN , CÔNG NGHỆ SINH HỌC VÀ TIN SINH HỌC .

Nhiều lĩnh vực sẽ bị ảnh hưởng bởi những hiểu biết về bộ gen người . Sau đây chỉ nêu một vài ví dụ ngắn gọn :

Proteomics (Hệ protein) theo nghĩa rộng là nghiên cứu tất cả các protein đã được mã trong một cơ thể (tức là Proteome), bao gồm cấu trúc , chức năng , sự tương tác , và những thay đổi của chúng . Còn theo nghĩa hẹp tức là sự định loại và sự nghiên cứu các protein gắn với các hoạt động của tế bào , nhưng không cần thiết phải là toàn bộ Proteome . Đối với con người thì nhiều protein cá thể sẽ được xác định cùng các đặc tính của chúng . Sự tương tác và hàm lượng của chúng sẽ được xác định ở cả 2 trạng thái : Sinh lý và bệnh lý . Các thông tin thu được sẽ được đưa vào các dữ liệu cơ sở thích hợp . Các kỹ thuật như điện di 2 chiều , khối phổ và các kháng thể sẽ là trung tâm để mở rộng lĩnh vực đang phát triển nhanh chóng này . Trên hết là hệ protein sẽ tạo thuận lợi lớn cho chúng ta hiểu cơ bản về protein cũng như nuôi dưỡng Công nghệ sinh học khi các protein mới được dùng để chẩn đoán , trị liệu và các ứng dụng khác cũng như các biện pháp nhằm phát triển sản xuất kinh tế . Cũng đòi hỏi các chuyên gia trong lĩnh vực Tin sinh học , bởi vì nó sẽ số hoá rất nhanh cho công việc quản lý và sử dụng một số khổng lồ các dữ liệu từ các nghiên cứu về hệ gen và hệ protein .

6.16. NHỮNG LIÊN QUAN TỚI Y HỌC .

Trên thực tế là mỗi lĩnh vực của y học đều chịu tác động bởi các thông tin về bộ gen người , đặc biệt việc theo dõi các gen gây bệnh sẽ được thuận lợi rất nhiều . Như đã đề cập ở trên , bản đồ SNP sẽ giúp ích rất nhiều trong việc xác định các gen liên quan tới các bệnh phức tạp . Đầu dò cho bất kỳ một gen nào cũng sẽ được áp dụng nếu cần thiết , trước tiên là việc cải tiến các test chẩn đoán các bệnh có liên quan tới gen và các gen liên quan trực tiếp đến căn nguyên của các bệnh đặc biệt . Trong lĩnh vực Dược phẩm (pharmacogenomics) cũng đang được mở rộng và có khả năng trong tương lai thuốc sẽ được bào chế để **điều tiết những biến đổi trong các enzym** và các **protein khác có liên quan tới hoạt động và chuyển hoá của thuốc trong các cơ thể** . Nghiên cứu về gen còn liên quan tới các hành vi để có thể hiểu biết sâu sa hơn nữa về các căn nguyên và khả năng điều trị các bệnh tâm thần .

Nhiều vấn đề thuộc luân lý , sự riêng tư và việc sử dụng thông tin hệ gen với mục đích thương mại cũng phải được đề cập .

Một điểm nữa cũng quan trọng là mang những lợi ích về y học và kinh tế cho những người thuộc Thế giới thứ ba trong các lĩnh vực dịch vụ y tế , chẩn đoán và điều trị bệnh .

GIẢI THÍCH MỘT SỐ TỪ CHUYÊN MÔN .

A

activator protein Một protein phản ứng với vị trí điều hoà trên phân tử DNA và kích lên sự biểu hiện gen .

adenosine một nucleotit có chứa riboza hoặc deoxyriboza, phot phat và adenin .

adenovirus một virut thông thường bao gồm một bộ gen DNA và một capsid icosahedral ; sử dụng như một vec tơ trong gen trị liệu .

agarose một chất liệu gelatin dùng trong kỹ thuật điện di .

AIDS một bệnh do virut có đặc tính là làm mất các tế bào Lympho —T của hệ miễn dịch và thường đi kèm với các bệnh do vi sinh vật gây nên các nhiễm trùng cơ hội gây tử vong .

amino acid một hợp chất hoá học có chứa ít nhất một nhóm amin và một nhóm axit hữu cơ ; nó là các đơn vị cấu thành của protein .

anticodon một trình tự gồm 3 bazơ của phân tử RNA bổ cứu mã trên phân tử mRNA .

antisense molecule phân tử mRNA phản ứng và làm trung hoà phân tử mRNA dùng để tổng hợp một protein đặc biệt ; phân tử antisense có tác dụng làm ngừng sự sản xuất protein .

autosome 44 nhiễm sắc thể của bộ gen người không liên quan tới việc xác định giới tính (nhiễm sắc thể thường) .

B

Bacillus thuringiensis vi khuẩn hình que mà các tinh thể độc tố của nó tác động như một chất diệt côn trùng dùng để chống lại một số loài thuộc chân khớp ; nguồn gen để cảm ứng kháng lại loài chân khớp trong thực vật .

bacteriophage một virut sao chép bên trong vi khuẩn hoặc hợp nhất với các chất liệu di truyền của nó với NST của vi khuẩn ; sử dụng như một vec tơ trong công nghệ DNA .

beta-galactosidase một enzym tiêu hoá được mã bởi một operon có trong tế bào vi khuẩn .

biolistic một xy lanh có một ống bằng ny lon mang các hạt hình cầu dùng để bắn gen vào các tế bào thực vật .

biotechnology nguyên lý sinh học trong đó các quá trình dựa trên cơ sở Hoá sinh như công nghệ DNA được dùng cho các mục đích thực tiễn .

bovine growth hormone (BGH) một hormon được sản xuất bằng công nghệ di truyền ; BGH kích thích tăng sản xuất sữa của trâu bò sữa ; cũng được gọi là somatotropin .

C

“**cape**” một nucleotit chứa 7-metyl guanozin được thêm vào đầu dẫn của phân tử mRNA trong khi tổng hợp phân tử này .

cDNA phân tử DNA được tổng hợp bởi enzym phiên mã ngược (transcriptaza) trên RNA ; trình tự bazơ ni tơ trong DNA bổ cứu với trình tự bazơ trong RNA .

cDNA library một khối tế bào được đưa vào các plasmid của chúng một bộ gen đặc biệt , bằng cách đó có thể thu nhận gen tương đối dễ dàng .

centiMorgan (cM) một đơn vị của di truyền học phân tử tương đương với một triệu cặp bazơ trong một phân tử DNA .

chimera một plasmid hay một vec tơ khác có chứa DNA không thường thấy trong plasmid .

chimeric plasmid một plasmid đã được công nghệ hoá để có mang một hay nhiều gen ngoại lai .

Chromosome Nhiễm sắc thể .

chromosome walking một phương pháp xác định trình tự bazơ trong đó một NST được phân tích bằng cách trải dài ra rồi chọn lấy một mẫu ở một thời điểm và xác định trình tự bazơ của mẫu đó .

clone nhóm của các tổ chức , hay tế bào , phân tử hay các đối tượng khác - tất cả đều đi từ một cá thể đơn ; đồng nghĩa với tập đoàn hay khuẩn lạc .

codon một trình tự gồm 3 bazơ trên phân tử mRNA, hoạt động cùng với anticodon trên phân tử tRNA để định rõ nơi một amino axit được đặt vào trong một protein .

complementary thành phần đối nghịch của một cặp bazơ đã được nhận dạng (bổ trợ , bổ sung) ; đặc biệt là Adenin bổ cứu với thymin , Xyrosin bổ cứu với Guanin .

Conjugation quá trình tái tổ hợp trong đó một vi khuẩn sống nhận được các đoạn DNA từ một vi khuẩn sống khác và biểu lộ protein được mã bởi DNA nhận được .

contig map bản đồ NST trong đó các đoạn DNA nối với nhau tạo nên một phân tử DNA kề nhau (kế tiếp nhau) .

cosmid một đoạn thẳng DNA mà phần đuôi của nó có thể đan vào nhau tạo nên một vòng ; cosmid có thể sử dụng như một vec tơ trong công nghệ sinh học .

cytogenetic map bản đồ NST trong đó các gen hoạt tính đáp ứng với một thuốc nhuộm hóa học và biểu lộ như các băng trên NST .

D

dideoxynucleotide một nucleotit thiếu 2 nguyên tử oxy trong phần cacbohydrat .

DNA axit deoxyribonucleic —một hợp chất hữu cơ của gen.

DNA fingerprinting một kỹ thuật tương hợp DNA trong đó các trình tự nhất định trong DNA của một mẫu tế bào đượ xác định và so sánh với trình tự bazơ DNA của một mẫu tế bào thứ hai để xem chúng đã **đồng nhất ở mức thiết lập tin tưởng chưa** .

DNA ligase một enzym nối các đoạn DNA với nhau bằng liên kết giữa nhóm phot phat của một nucleotit với phân tử deoxyriboza của một nucleotit tiếp theo .

DNA probe một phân tử DNA tổng hợp chuỗi đơn , nhỏ gắn với chuỗi DNA bổ cứu đích trong một hỗn hợp các chất sinh học .

double helix phân tử chuỗi kép bị nứt ra để hình thành một xoắn ; dạng DNA trong NST .

Drosophila melanogaster tên khoa học của một loại ruồi đượ sử dụng trong công nghệ sinh học .

E

EcoRI một enzym giới hạn từ một chủng R của vi khuẩn Escherichia Coli.

electrophoresis một kỹ thuật Hoá sinh trong đó các phân tử đượ dịch chuyển theo kích thước và các đặc tính hoá học của chúng trong Trường điện từ .

endonuclease một enzym tế bào phân cắt phân tử DNA .

enzyme một protein xúc tác một phản ứng hoá học của sự chuyển hoá mà nó vẫn còn giữ nguyên sau phản ứng .

escherichia Coli một chủng vi khuẩn thường thấy trong đường tiêu hoá của người đượ dùng trong công nghệ DNA như một tổ chức nhận trình tự gen .

eucaryote một tổ chức mà các tế bào của nó có nhân chuẩn (như nấm , protozoa , thực vật , động vật) .

eucaryotic thuộc về một tổ chức phức tạp mà các tế bào của nó có nhân và các cơ quan tử , nhân lên bằng gián phân (mitose) và có các đặc tính phân biệt với các ế bào procaryote đơn giản hơn . Các tế bào thực vật , động vật và người là nhân chuẩn

exon phần mRNA đặc hiệu cho một trình tự của một amino axit ; exon vẫn đượ giữ lại trong phân tử mRNA cuối cùng và biểu hiện trình tự amino axit của protein .

F

forensic làm theo luật pháp .

G

gene một đoạn DNA cung cấp thông tin hoá học cho việc tổng hợp protein trong tế bào ; tập hợp các gen và xen vào các trình tự DNA hợp thành một NST .

gene bank nơi chứa các kho tế bào của cơ thể để sau này có thể tiến hành phân tích di truyền .

gene library một dãy các tế bào hoặc các cơ thể đơn bào thường chứa các gen từ các tế bào ngoại lai ; các gen lấy từ các tế bào của thư viện gen đượ dùng trong các thí nghiệm công nghệ DNA .

gene linkage map bản đồ NST trong đó các gen hoạt động đượ định vị bằng cách định vị kết hợp chặt chẽ với các gen dấu chuẩn .

gene probe phân tử DNA hoặc RNA kết hợp với đoạn DNA hoặc RNA bổ cứu khi đưa vào một khối lớn các phân tử axit nucleic .

genetic code trình tự bazơ ni tơ trong phân tử DNA chuyên hoá cho một trình tự của một amino axit trong một protein .

genome toàn bộ các thành phần DNA nhân của tất cả các gen của một tế bào hay một cơ thể đơn bào

genomic library một tập hợp của các tế bào đã hợp nhất trong các plasmid của nó tất cả các gen của một tế bào khác . Bằng cách này , việc thu nhận gen tương đối dễ dàng .

glycoprotein một protein có chứa một hoặc nhiều phân tử cacbohydrat gắn với một hoặc nhiều amino axit trong một protein .

glycosylation quá trình sinh hoá học trong đó các phân tử cacbohydrat đượ gắn vào amino axit của một protein ở những điểm khác nhau .

Golgi apparatus một dãy các màng dát mỏng trong tế bào nhân chuẩn cùng với những thứ để xử lý protein trước khi sự tổng hợp hoàn tất ; cũng còn gọi là thể Golgi .

gp 120 một glycoprotein định vị ở lớp vỏ của virus HIV ; một chất cần thiết để gắn HIV vào tế bào chủ trong quá trình sao chép ; có thể tổng hợp được bằng công nghệ DNA để sử dụng trong việc sản xuất vắc xin .

H

helix giống như một cuộn dây nối với máy điện thoại (xoắn) , dạng DNA trong NST .

hemophilia B một dạng ưa chảy máu trong đó bệnh nhân không tạo được cục máu đông do thiếu một yếu tố cần thiết cho sự đông máu , yếu tố đông máu tổng hợp có thể làm nhẹ bớt triệu chứng này .

hepatitis B một bệnh viêm gan do virus truyền qua máu ; vắc xin hepatitis B đã được sản xuất bằng công nghệ DNA .

human genome chứa khoảng 100.000 gen trong một tế bào người .

“human mouse” chuột được thay đổi về giải phẫu để có chứa các tế bào chức năng của hệ miễn dịch .

Huntington’s diseases một bệnh di truyền đặc trưng bởi sự hư hỏng tiến triển hệ thần kinh cùng với sự đập phá .

hypercholesterolemia một bệnh di truyền đi kèm với việc tăng mức cholesterol trong máu vì vắng mặt receptor cholesterol .

I

ice —minus bacterium một vi khuẩn biến đổi gen tạo ra các protein cảm ứng sự hình thành các tinh thể đá ở nhiệt độ thấp hơn bình thường .

insulin hormon tuyến tụy bao gồm 51 amino axit , 2 chuỗi nối với nhau . Insulin làm thuận lợi cho việc chuyển glucoza vào máu . Insulin cũng được sản xuất bằng vi khuẩn biến đổi gen .

interferon các protein được sản xuất bởi các tế bào của người trong đáp ứng với virus . Interferon kích thích sự tổng hợp các chất nhằm chống lại sự thâm nhập của virus vào trong tế bào ; nó cũng được tổng hợp bằng các tế bào vi khuẩn biến đổi gen .

intron phần xen vào phân tử mRNA đầu tiên , sau đó nó được loại đi trước khi tạo thành phân tử mRNA cuối cùng .

K

kilobase một đơn vị gồm 1000 bazơ ni tơ trong một phân tử DNA hay RNA .

“knockout mouse” chuột biến đổi di truyền để làm thiếu các gen cho cả cơ quan hay cả hệ thống cơ quan .

L

lactoferrin một protein có trong sữa có khả năng gắn các phân tử sắt ; lactoferrin được sản xuất bằng các tế bào động vật chuyển gen .

M

mammalian Động vật có xương sống , máu nóng có tuyến vú , có lông tóc , chẳng hạn như loài gặm nhấm , động vật nuôi , động vật bậc cao và người .

mega YAC một loại nấm men lớn ít gấp có NST nhân tạo .

messenger RNA (mRNA) phân tử RNA được tổng hợp với các bazơ ni tơ mã bộ cứu với DNA ; mRNA mang các thông điệp tới tế bào chất của tế bào để tổng hợp protein .

microinjection một kỹ thuật Hoá sinh trong đó một xy lanh nhỏ được sử dụng để thâm nhập vào trong tế bào và đẩy các đoạn DNA vào nhân của tế bào .

mus musculus một loại chuột thường dùng trong công nghệ DNA .

mutation thay đổi đặc tính của một tổ chức do thay đổi các gen trong tổ chức đó .

N

nanometer Đơn vị đo lường tương đương với một phần tỷ met .

nucleic acid một hợp chất hữu cơ cấu thành từ các nucleotit nối với nhau bởi các liên kết phot phot .

nucleotide một đơn vị cấu thành của axit nucleic .

O

oligonucleotide các đoạn tổng hợp nhỏ của RNA hoặc DNA ; một dạng của gen tổng hợp ; cũng được gọi là oligo ; cũng là các đoạn nhỏ DNA hoạt động như các phân tử antisense .

oncogenes gen gây khối u hay ung thư .
“**oncomouse**” chuột biến đổi gen có khả năng bị ung thư vú cao .
operator một dãy các bazơ ni tơ trong DNA của một operon ở đó protein ức chế hay hoạt hoá có thể điều chỉnh sự biểu hiện của gen .
operon phức hợp các gen cấu trúc và gen điều hoà trong NST .
origin of replication một trình tự bazơ ni tơ trên phân tử DNA làm tín hiệu cho điểm bắt đầu của sự sao chép của phân tử DNA .

P

peptide một protein nhỏ .
phage viết tắt của bacteriophage ; một virus sao chép trong vi khuẩn hoặc hợp nhất axit nucleic của nó với bộ gen vi khuẩn ; được dùng như một vec tơ .
“**pharm**” **animal** động vật chuyển gen để sản xuất ra một sản phẩm thuốc xác định .
phosphate group một phân tử gồm phot pho , oxy , và nguyên tử hydro đi từ axit phot pho ric nó có mặt cả trong RNA và DNA .
phosphodiester bond liên kết hoá học hình thành giữa 5' phot phat tự do của một nucleotit và deoxyriboza của nucleotit thứ hai .
physical map bản đồ NST trong đó vị trí của gen hoạt động được xác định và đã biết số bazơ giữa các gen hoạt hoá .
plasmid một vòng đóng chặt của DNA chứa khoảng trên mười gen và các bản sao nhân lên trong tế bào chất của vi khuẩn ; nó có thể là một vec tơ của gen trong công nghệ DNA .
poly —A tail một trình tự của nucleotit chứa adenin ở cuối phân tử mRNA không rõ về chức năng nhưng có thể được dùng trong công nghệ DNA để xác định phân tử mRNA .
polygalacturonate một enzym có trong thực vật có thể tiêu hoá được pectin .
polymer một phân tử hoá học bao gồm các đơn vị lặp đi lặp lại của một chất đặc biệt .
polymerase chain reaction một quá trình sinh hoá trong đó một đoạn DNA được nhân lên hàng triệu lần nhờ sự hoạt động của enzym polymeraza .
primer DNA một đoạn DNA để khởi đầu tổng hợp DNA , chẳng hạn như trong phản ứng chuỗi PCR .
procaryote một tổ chức mà các tế bào của nó có nhân rải rác (chẳng hạn như vi khuẩn) .
prokaryotic thuộc về một tổ chức đơn giản mà các tế bào của nó thiếu nhân hoặc các bào quan , nhân lên bằng cách sinh sản đơn giản , nó có các đặc điểm khác biệt hẳn so với các tế bào nhân chuẩn . Các tế bào vi khuẩn là nhân rải rác .
promoter site một trình tự của nucleotit khởi đầu cho sự phiên mã của một mã di truyền trong DNA thành mRNA .
protease một enzym tiêu hoá protein ; ở vi khuẩn , proteaza thường tiêu hoá các protein ngoại lai .
pseudomonas syringae một protein ở đầu một vi khuẩn gây cảm ứng hình thành các tinh thể đá một cách tự nhiên . vi khuẩn có thể được biến đổi gen để cảm ứng hình thành các tinh thể ở nhiệt độ thấp hơn .
purine một bazơ ni tơ . Ví dụ như adenin và guanin .

Q

Q-beta replicase một enzym xúc tác tổng hợp RNA sử dụng RNA như một cái khuôn ; dùng để khuếch đại tín hiệu trong PCR .

R

RAC Ủy ban cố vấn tái tổ hợp DNA ; một nhóm chuyên viên trong viện y học quốc gia Hoa kỳ NIH . Nhóm này có quyền phê chuẩn hoặc bãi miễn các dự kiến thực nghiệm có liên qua tới gen trị liệu .
radioactive có thể phát ra bức xạ năng lượng có thể phát hiện được .
recognition sequence một trình tự bazơ trên DNA được nhận dạng hoá học bằng một enzym giới hạn .
regulatory site một trình tự bazơ trên phân tử DNA ở đó sự biểu hiện gen có thể được kiểm soát bằng phản ứng với các protein kiểm chế hoặc hoạt hoá .
repressor protein một protein phản ứng với vị trí điều hoà trên phân tử DNA và kiểm chế sự biểu hiện của gen bằng cách ức chế sự phiên mã .
resolution giải pháp .
restriction enzyme một enzym có thể cắt DNA ở một vị trí giới hạn .
restriction fragment length polymorphism (RFLP) đa hình độ dài các đoạn giới hạn .
retrovirus một virus thâm nhập vào các tế bào chủ và sử dụng enzym phiên mã ngược để mã cho phân tử DNA bổ cứu cho RNA retrovirus ; nó được dùng như một vec tơ DNA .

reverse transcriptase một enzym sử dụng các trình tự bazơ trong phân tử DNA như một mô hình để tổng hợp một phân tử DNA bổ cứu .

RFLP (restriction fragment length polymorphism) một đoạn DNA phân lập bởi các enzym giới hạn và có chiều dài rất thay đổi ; RFLP được dùng trong chẩn đoán và in dấu DNA nhưng chưa rõ chức năng của chúng .

ribose phân tử 5 cac bon và là một trong các thành phần của RNA ; cũng tương tự như deoxyriboza chỉ khác là có thêm vào một nguyên tử oxy .

ribosome binding site một dãy các bazơ ni tơ trong DNA của một operon ở đó mã bazơ trong mRNA là đặc hiệu cho phép mRNA gắn vào riboxom .

RNA polymerase một enzym có chức năng trong phiên mã và tổng hợp phân tử RNA với các bazơ bổ cứu với nó trong DNA .

S

sequence map bản đồ NST trong đó bản chất và vị trí của mỗi bazơ ni tơ đã rõ **sequence tag site** (STS) một đoạn DNA được gắn vào một vị trí chọn lọc trên phân tử DNA để đánh dấu vị trí ưu tiên để phân tích .

signal peptide một protein nhỏ tạo thuận lợi cho việc vận chuyển một protein đã được sản xuất ra khỏi tế bào để đi vào môi Trường ngoài tế bào .

Southern blotting một kỹ thuật trong phòng thí nghiệm trong đó các phân tử đã được tách chiết bằng điện di sau đó thấm khỏi gel để định loại .

“sticky end” một chuỗi đơn nucleotit nối rộng ở phần đuôi của một đoạn DNA (đầu dính).

structural genes một nhóm gen trong operon mã cho một protein và sự biểu hiện của nó được kiểm soát bởi một số gen điều hoà .

subunits các phân tử vi khuẩn hoặc các đoạn virus có thể sử dụng làm vac xin ; nó được sản xuất bằng công nghệ DNA .

T

tag polymerase một enzym chịu nhiệt đi từ vi khuẩn *Thermus aquaticus* ; enzym này là một DNA polymeraza được sử dụng trong PCR .

telomere một đoạn DNA ở đầu NST của nấm men để bảo vệ NST khỏi bị phân giải bởi enzym nucleaza .

termination site một dãy các nucleotit làm tín hiệu kết thúc sự phiên mã từ DNA thành mRNA .

thymidine một nucleotit có chứa deoxyriboza , phot phat và thymin .

thymidine kinase (TK) một enzym có chức năng trong tổng hợp DNA bằng cách giải phóng thymin từ DNA sử dụng trước để tổng hợp DNA mới ; nó được sử dụng như một dấu chuẩn để xác định chắc chắn cho việc cài gen .

tissue plasminogen activator (TPA) một protein tác động như một proteaza và phân giải các cục máu đông . TPA cũng có thể sản xuất bằng công nghệ DNA .

transcription một quá trình mà enzym tổng hợp phân tử RNA khi sử dụng một chuỗi DNA như một cái khuôn ; các bazơ RNA bổ cứu với các bazơ DNA .

transduction quá trình tái tổ hợp gen trong đó virus của vi khuẩn nhận được các các đoạn DNA của vi khuẩn khi sao chép virus và chuyển những đoạn DNA này vào một vi khuẩn sống khác ở đó DNA được biểu hiện .

transfer RNA (tRNA) một phân tử RNA gắn vào một amino axit đặc hiệu và vận chuyển chúng tới riboxom , ở đó xảy ra sự tổng hợp protein ; có nhiều tRNA cho một amino axit .

transformation quá trình tái tổ hợp gen trong đó vi khuẩn nhận được các đoạn DNA từ môi Trường bên ngoài và biểu hiện các protein được mã bởi các gen có trong các đoạn đó .

translation một quá trình sinh hoá có liên quan tới các enzym, riboxom, và các thành phần khác trong đó phân tử mRNA cung cấp một bazơ ni tơ mã cho sự thay thế của các amino axit trong sự tổng hợp protein .

tumor-infiltrating lymphocyte (TIL) các tế bào lympho của hệ miễn dịch tấn công và phân huỷ các tế bào khối u .

tumor necrosis factor (TNF) một protein được tạo ra bởi đại thực bào của người nó kích thích sự phá vỡ các tế bào khối u ; TNF có thể được tạo ra bằng công nghệ DNA .

U

ultramicroscopic dưới khả năng quan sát của hiển vi quang học .

V

variable —number tandem repeat (xem VNTR) = đoạn lặp dài .

vector thực chất của việc vận chuyển DNA ngoại lai vào trong một tế bào của một tổ chức, ở đó DNA được biểu hiện .; một plasmid , cosmid hay một virus có thể như là một vec tơ .

vector incompetence khi một tổ chức không có khả năng tự thiết lập trong vật chủ (bất lực) .

virus một đoạn axit nucleic được bao quanh bởi một lớp protein bao ngoài , trong một số Trường hợp có cả một lớp vỏ ; có khả năng sao chép trong các tế bào vật chủ và gây bệnh trong một số Trường hợp ; mặt khác , axit nucleic của virus có thể hợp nhất với NST của tế bào ; virus thường được sử dụng như các vec tơ .

VNTR (variable —number tandem repeat) một đoạn DNA có chứa một số đặc biệt các trình tự bazơ ni tơ lặp đi lặp lại . VNTR được dùng trong kỹ thuật in vân tay DNA .(Đoạn lặp dài)

Y

yeast artificial chromosome (YAC) một đoạn DNA từ một tế bào nấm men và được dùng như phân tử cơ sở để gắn vào các đoạn DNA khác .